Juliana de Lucena Ruas Riani

Determinantes Do Resultado Educacional No Brasil: Família, Perfil Escolar Dos Municípios E Dividendo Demográfico Numa Abordagem Hierárquica E Espacial

Belo Horizonte, MG
UFMG/CEDEPLAR
2005

Juliana de Lucena Ruas Riani

Determinantes Do Resultado Educacional No Brasil: Família, Perfil Escolar Dos Municípios E Dividendo Demográfico Numa Abordagem Hierárquica E Espacial

Tese apresentada ao curso de doutorado do Centro de Desenvolvimento e Planejamento Regional da Faculdade de Ciências Econômicas da Universidade Federal de Minas Gerais, como requisito parcial à obtenção do Título de Doutor em Demografia.

Orientador: Prof. Eduardo Luis Gonçalves Rios-Neto.
Co-orientador: Profª. Cibele Comini César
Profª. Sueli Moro

Belo Horizonte, MG
Centro de Desenvolvimento e Planejamento Regional
Faculdade de Ciências Econômicas – UFMG
2005

FOLHA DE APROVAÇÃO

Ao Flávio, Flavinho e Teteu.
E a todos aqueles que, de alguma
forma, buscam na educação, um Brasil melhor.

AGRADECIMENTOS

No decorrer do meu doutorado muitas foram as pessoas que contribuíram de maneira direta ou indireta para chegar até aqui.

Como não poderia deixar de ser, sou imensamente grata ao meu orientador, Eduardo Rios-Neto, não apenas pelas suas idéias, que foram essenciais para a construção dessa tese, mas também pela oportunidade de trabalhar na pesquisa com o INEP, de onde brotou grande parte do conteúdo dessa tese. Porém, falar dos seus *insights* é "chover no molhado", pois aqueles que já tiveram oportunidade de trabalhar com ele, logo percebem tal qualidade. Contudo, ao meu ver, a sua maior qualidade é o que podemos chamar de "generosidade acadêmica". Essa generosidade se traduz no seu desprendimento em compartilhar suas idéias, dando oportunidades a todos que com ele trabalham. Se não fosse por essa qualidade, não teria sido co-organizadora do livro Introdução à Demografia da Educação, não teria apresentado um trabalho na IUSSP e tantas outras atividades profissionais.

Sou imensamente grata também as minhas co-orientadoras, Cibele Comini e Sueli Moro, cujas sugestões foram fundamentais para o resultado final da tese. Somado a isso, o carinho e amizade das duas foram essenciais.

E aos demais professores do CEDEPLAR, que mesmo não tendo participação direta, contribuíram para a minha formação como demógrafa. Em especial ao José Alberto de Carvalho pelos conselhos, sempre sábios.

Gostaria de agradecer também aos funcionários da biblioteca pela revisão bibliografia.

Não poderia deixar de citar o Instituto Nacional de Estudos e Pesquisas Educacionais – INEP – órgão do Ministério da Educação - MEC, cujo convênio com o CEDEPLAR, proporcionou grande material para elaboração desse estudo. Essa tese é, sem dúvida, um dos frutos dessa parceria.

Agradeço também ao pessoal da sala 820 - Vânia, André, Vicente, Fabiana, Carlos e Luciana - que além de me ajudarem diretamente em algumas dúvidas de redação e formatação, agüentaram meu período de stress.

Nesse período de doutorado uma das boas coisas foram as amizades construídas, mesmo que em algum tempo, devido aos caminhos diferentes, os contatos não sejam constantes. Dessa forma, gostaria de agradecer a todos os meus amigos do CEDEPLAR (para não correr o risco de esquecer alguém, não vou citar nomes) que proporcionaram

agradáveis “conversas de corredor”. Com certeza, essas conversas tornam o dia mais agradável. Um agradecimento especial vai para Andréa, pela solicitude e paciência em fazer uma revisão final do texto da tese, e para Flávia Chein, pela ajuda no SpaceStat.

Também sou imensamente grata ao meu marido, Flávio, que em muitos finais de semana “segurou as pontas” com os meninos, para que eu pudesse escrever aquele capítulo, estimar aquela regressão .... Somado a isso, o seu incentivo e compreensão foram de suma importância.

Agradeço também ao meu *background* familiar, ou seja, aos meus pais, Amarílio e Reivane, que mesmo sem saber o que é Função de Produção Educacional, sempre preocuparam em mostrar a importância da educação para seus filhos.

Por fim, gostaria de agradecer aos meus filhos, Flavinho e Teteu, pelos momentos de imensa alegria que me proporcionam.

RESUMO

O objetivo central dessa tese é investigar os determinantes do resultado educacional nos níveis de ensino fundamental e médio, combinando também a tradição da demografia econômica, ou seja, incluindo os efeitos demográficos. A análise é feita utilizando duas metodologias diferentes: hierárquica e hierárquica-espacial. Essa última constitui-se em uma contribuição metodológica importante dessa tese.

A análise dos determinantes educacionais é baseada em três questões chaves. A primeira, diz respeito ao impacto do dividendo demográfico. Busca-se avaliar até que ponto as profundas mudanças demográficas ocorridas no Brasil, com 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".

A segunda questão refere-se a importância dos fatores da rede escolar dos municípios, principalmente aqueles relacionados com a qualidade dos recursos humanos e infra-estrutura. Essa análise é importante por examinar quais fatores do perfil escolar dos municípios podem diminuir a importância dos fatores de *background* familiar no sentido de diminuir a estratificação educacional, ou seja, diminuir a relação entre a trajetória escolar do indivíduo e a sua origem social.

Por fim, a terceira questão busca investigar como o contexto espacial mais amplo interfere no resultado educacional. Essa análise é importante por dois motivos. Primeiro, porque quando se consideram os efeitos espaciais – autocorrelação e heterogeneidade espacial – a estimativa dos parâmetros relacionados as variáveis contextuais são mais confiáveis. Segundo, porque será possível analisar até que ponto políticas públicas educacionais são mais eficazes quando focalizadas apenas em determinado município ou quando levam em conta um espaço geográfico mais abrangente, ou seja, os municípios adjacentes.

Os resultados demonstram que apesar do forte impacto da educação materna nos indicadores educacionais, uma melhoria no perfil escolar dos municípios pode minimizar seus efeitos. Porém, essa melhoria possui maior eficácia quando abrange um contexto espacial mais amplo.

Somado a isso, a diminuição do tamanho relativo de crianças e jovens se constitui em um fator positivo, tanto para os indicadores de acesso quanto para os de eficiência e rendimento.

ABSTRACT

The central aim of this dissertation is to investigate the determinants of educational outcomes at elementary and high school levels, using the economic demographic tradition. The analysis is done via two different methodologies: hierarchical and spatial-hierarchical. The last one represents an important methodological contribution to this dissertation.

The analysis of the educational determinants is based on tree main points: The first one deal with the demographic dividend impact. It attempts to evaluate how relevant demographic changes in Brazil – which had as a consequence the decline in the young dependency ratio – could bring a demographic bonus to the Brazilian education, for opening the "window of opportunities".

The second point is related with the importance of several factors present at local schools, mainly those connected to infra-structure and humans resources quality. This analysis is important for examining which local school structure aspects may reduce the family background factors relevance, allowing for an educational stratification attenuation - that is - thinning the relationship between the individual's school path and your social origin.

Finally, the third point tries to detect how a larger spatial context interferes in the educational output. This analysis is important for two reasons: First, because when spatial effects –autocorrelation and spatial heterogeneity – are considered, the parameters estimates associated with contextual variables are more trustworthy. Second, because it will be possible to analyze if educational public policies are more effective when centralized solely in a specific municipality or when taking into account a greater geographical area (i.e. those including adjacent municipalities).

The results show that, despite the strong impact of mother's education over educational indicators, an improvement in the local schools' quality can minimize these effects. However, this enhancement is more successful when a greater spatial context is considered.

Adding to these points, smaller sizes of children and youth cohorts turn out to be a positive aspect for both school indicators: access and efficiency.

SUMÁRIO

INTRODUÇÃO 1
1) INDICADORES EDUCACIONAIS E SUA EVOLUÇÃO NAS ÚLTIMAS DÉCADAS 10
1.1) Taxa de Analfabetismo 10
1.2) Taxa de Atendimento Escolar 16
1.3) Taxa de Escolarização Bruta 20
1.4) Taxa de Escolarização Líquida 24
1.5) Taxa de Distorção idade/série 28
1.6) Probabilidade de Progressão Por Série (PPS) 31
1.7) Anos Médios de Estudo ou Escolaridade Média 35
1.8) Conclusão 40
2) MARCO TEÓRICO E PRINCIPAIS ESTUDOS SOBRE OS DETERMINANTES DA EDUCAÇÃO 41
2.1) Função de Produção Educacional 41
2.2) Alguns Estudos Sobre os Determinantes da Educação 44
2.3) Conclusão 53
3) ESTRATÉGIA EMPÍRICA, BASE DE DADOS E TRATAMENTO DAS VARIÁVEIS 54
3.1) Especificação da Função de Produção Educacional 54
3.2) Como medir resultado educacional? 58
3.3) Variáveis Independentes 60
3.4) Base de Dados e amostra 67
3.5) Conclusão 72
4) ARCABOUÇO METODOLÓGICO 74
4.1) O Modelo Logístico Hierárquico 74
4.2) Análise espacial 80
4.2.1) Análise Exploratória Espacial 80
4.2.2) Econometria Espacial 82
4.3) Modelo hierárquico-espacial 84
4.4) Conclusão 86
5) ANÁLISE HIERÁRQUICA DOS DETERMINANTES EDUCACIONAIS 88
5.1) Probabilidade de freqüentar a escola 88
5.2) Probabilidade de freqüentar a escola na idade adequada 99
5.3) Probabilidade de Progressão por Série (PPS) 112
5.4) Conclusão 120
6) INCLUSÃO DO CONTEXTO ESPACIAL NA ANÁLISE 122
6.1) Resultados da Análise Exploratória Espacial 123
6.3) Regressão Espacial 143
6.4) Conclusão 151
7) ANÁLISE HIERÁRQUICO-ESPACIAL DOS DETERMINANTES EDUCACIONAIS 154
7.1) Probabilidade de freqüentar a escola 154
7.2) Probabilidade de freqüentar a escola na idade adequada 165
7.3) Probabilidade de Progressão por Série (PPS) 174
7.4) Conclusão 183
CONCLUSÃO 186
BIBLIOGRAFIA 196
ANEXOS 204

LISTA DE TABELAS

TABELA 1. 1: Taxa de distorção idade/série por sexo, cor e nível de ensino – Brasil, 2000. 30
TABELA 1. 2: Probabilidade de progressão por série por grupo etários – Brasil, 2000. 33
TABELA 3. 1: Estatísticas descritivas das variáveis do nível 1 para as diferentes amostras do Ensino Fundamental 69
TABELA 3. 2: Estatísticas descritivas das variáveis do nível 2 para as diferentes amostras do Ensino Fundamental 70
TABELA 3. 3: Estatísticas descritivas das variáveis do nível 1 para as diferentes amostras do Ensino Médio 71
TABELA 3. 4: Estatísticas descritivas das variáveis do nível 2 para as diferentes amostras do Ensino Médio 72
TABELA 5. 1: Resultado da regressão para a probabilidade de freqüentar a escola – indivíduos de 7 a 14 anos de idade 90
TABELA 5. 2: Resultado da regressão para a probabilidade de freqüentar a escola – indivíduos de 15 a 17 anos de idade 94
TABELA 5. 3: Resultado da regressão para a probabilidade de freqüentar a escola – indivíduos de 15 a 17 anos de idade que já concluíram o fundamental 98
TABELA 5. 4: Variações na probabilidade de freqüentar a escola segundo variações nas variáveis de *background* familiar para diferentes amostras 99
TABELA 5. 5: Resultado da regressão para a probabilidade de freqüentar a escola na idade correta – indivíduos de 7 a 14 anos de idade que freqüentam a escola 102
TABELA 5. 6: Resultado da regressão para a probabilidade de freqüentar a escola na idade correta – indivíduos de 15 a 17 anos de idade que freqüentam a escola ... 103
TABELA 5. 7: Variações na probabilidade de freqüentar a escola na idade adequada segundo variações nas variáveis de *background* familiar e rede escolar – ensino fundamental e médio 107
TABELA 5. 8: Resultado da regressão para a probabilidade de progressão na 1ª série 114
TABELA 5. 9: Resultado da regressão para a probabilidade de progressão na 5ª série 115
TABELA 5. 10: Resultado da regressão para a probabilidade de progressão na 1ª série do ensino médio 115
TABELA 6. 1: Estatística I de Moran para as variáveis do Ensino Fundamental 124
TABELA 6. 2: Estatística I de Moran para as variáveis do Ensino Médio 125
TABELA 7. 1: Resultado dos diferentes modelos para a probabilidade de freqüentar a escola – indivíduos de 7 a 14 anos. 159
TABELA 7. 2: Resultado dos diferentes modelos para a probabilidade de freqüentar a escola – indivíduos de 15 a 17 anos. 160
TABELA 7. 3: Resultado dos diferentes modelos para a probabilidade de freqüentar a escola – indivíduos de 15 a 17 anos que já concluíram o ensino fundamental. .. 161
TABELA 7. 4: Resultado dos diferentes modelos para o intercepto da probabilidade de freqüentar a escola na idade correta – ensino fundamental. 167
TABELA 7. 5: Resultado dos diferentes modelos para o coeficiente da educação materna da probabilidade de freqüentar a escola na idade correta – ensino fundamental. 169
TABELA 7. 6: Resultado dos diferentes modelos para o intercepto da probabilidade de freqüentar a escola na idade correta – ensino médio. 172

TABELA 7. 7: Resultado dos diferentes modelos para o coeficiente da educação materna da probabilidade de freqüentar a escola na idade correta – ensino médio. 174
TABELA 7. 8: Resultado dos diferentes modelos para o intercepto da probabilidade de progressão na 1ª série. 175
TABELA 7. 9: Resultado dos diferentes modelos para o coeficiente da educação materna da probabilidade de progressão na 1ª série. 177
TABELA 7. 10: Resultado dos diferentes modelos para o intercepto da probabilidade de progressão na 5ª série. 180
TABELA 7. 11: Resultado dos diferentes modelos para o intercepto da probabilidade de progressão na 1ª série do ensino médio. 182
TABELA A1. 1: Resultado da regressão para a probabilidade de freqüentar a escola – indivíduos de 7 a 14 anos de idade - diferentes modelos 204
TABELA A1. 2: Resultado da regressão para a probabilidade de freqüentar a escola – indivíduos de 15 a 17 anos de idade - diferentes modelos 205
TABELA A1. 3: Resultado da regressão para a probabilidade de freqüentar a escola – indivíduos de 15 a 17 anos de idade que já concluíram o ensino fundamental - diferentes modelos. 206
TABELA A1. 4: Resultado da regressão para a probabilidade de freqüentar a escola – indivíduos de 15 a 17 anos de idade - diferentes modelos. 207
TABELA A1. 5: Resultado da regressão para a probabilidade de freqüentar a escola na idade correta – indivíduos de 7 a 14 anos de idade - diferentes modelos..... 208
TABELA A1. 6: Resultado da regressão para a probabilidade de freqüentar a escola na idade correta – indivíduos de 15 a 17 anos de idade - diferentes modelos..... 209
TABELA A1. 7: Resultado da regressão para a probabilidade de progressão na 1ª série – diferentes modelos. 210
TABELA A1. 8: Resultado da regressão para a probabilidade de progressão na 5ª série – diferentes modelos. 211
TABELA A1. 9: Resultado da regressão para a probabilidade de progressão na 1ª série do ensino médio – diferentes modelos. 211
TABELA A2. 1: Matriz de correlação para as variáveis do ensino fundamental 212
TABELA A2. 2: Matriz de correlação para as variáveis do ensino médio. 213
TABELA A3. 1: Resultado da regressão para a probabilidade de freqüentar a escola – indivíduos de 15 a 17 anos de idade. 214
TABELA A3. 2: Resultado da regressão para a probabilidade de freqüentar a escola – indivíduos de 15 a 17 anos de idade que já concluíram o ensino médio. 214
TABELA A3. 3: Resultado da regressão para a probabilidade de freqüentar a escola na idade correta– indivíduos de 15 a 17 anos de idade. 215
TABELA A3. 4: Resultado da regressão para a Probabilidade de Progressão na 1ª série do ensino médio. 215

LISTA DE GRÁFICOS

GRÁFICO 1. 1: Taxa de analfabetismo na faixa de 15 anos ou mais – Brasil, 1950/2000 11
GRÁFICO 1. 2: Taxa de analfabetismo da população de 15 anos ou mais para diferentes países – 2000 12
GRÁFICO 1. 3: Taxa de analfabetismo da população de 15 anos ou mais – Estados 2000 13
GRÁFICO 1. 4: Taxa de analfabetismo por grupo etário – Brasil, 2000. 14

GRÁFICO 1. 5: Taxa de analfabetismo para a população de 15 anos ou mais por sexo e situação de domicílio – Brasil, 2000. 15
GRÁFICO 1. 6: Taxa de analfabetismo para a população de 15 anos ou mais segundo grupos de cor – Brasil, 2000 15
GRÁFICO 1. 7: Taxa de atendimento segundo grupos de idade – Brasil, 1970/2000... 17
GRÁFICO 1. 8: Taxa de atendimento por grupos de idade e sexo – Brasil, 1970 a 2000. 17
GRÁFICO 1. 9: Taxa de atendimento segundo grupos de idade e situação de domicílio – Brasil, 1970/2000 18
GRÁFICO 1. 10: Taxa de atendimento para a faixa etária de 7 a 14 anos para Brasil e grandes regiões 1970 a 2000 19
GRÁFICO 1. 11: Taxa de atendimento para a faixa etária de 15 a 17 anos para Brasil e grandes regiões 1970 a 2000 20
GRÁFICO 1. 12: Taxa de escolarização bruta do ensino fundamental – Brasil e regiões, 1980 a 2000 21
GRÁFICO 1. 13: Taxa de escolarização bruta do ensino fundamental – Brasil e estados, 2000 22
GRÁFICO 1. 14: Taxa de escolarização bruta do ensino médio – Brasil e regiões, 1980 a 2000 23
GRÁFICO 1. 15: Taxa de escolarização bruta do ensino médio – Brasil e estados, 2000 23
GRÁFICO 1. 16: Taxa de escolarização líquida do ensino fundamental – Brasil e Regiões, 1980 a 2000. 25
GRÁFICO 1. 17: Taxa de escolarização líquida do ensino médio – Brasil e Regiões, 1980 a 2000. 25
GRÁFICO 1. 18: Taxa de escolarização líquida do ensino fundamental – Brasil e estados, 2000 26
GRÁFICO 1. 19: Taxa de escolarização líquida do ensino médio – Brasil e estados, 2000 27
GRÁFICO 1. 20: Taxa de Escolarização Líquida do Ensino Fundamental por cor – Brasil, 1980 a 2000 27
GRÁFICO 1. 21: Taxa de Escolarização Líquida do Ensino Médio por cor – Brasil, 1980 a 2000 28
GRÁFICO 1. 22: Taxa de Distorção idade/série do Ensino Fundamental – Brasil e regiões, 1980 a 2000. 30
GRÁFICO 1. 23: Taxa de Distorção idade/série do Ensino Médio – Brasil e regiões, 1980 a 2000. 31
GRÁFICO 1. 24: Taxa de progressão nas quatro primeiras séries do ensino fundamental – Brasil, 2000. 34
GRÁFICO 1. 25: Taxa de progressão nas quatro últimas séries do ensino fundamental – Brasil, 2000. 34
GRÁFICO 1. 26: Taxa de progressão do ensino médio – Brasil, 2000. 35
GRÁFICO 1. 27: Anos médios de estudo da população de 15 ou mais anos de idade – Brasil, 1980 a 2000. 37
GRÁFICO 1. 28: Anos médios de estudo da população de 7 a 25 anos de idade segundo educação do chefe da família – Brasil, 1980 a 2002. 38
GRÁFICO 1. 29: Anos médios de estudo da população de 7 a 25 anos de idade segundo quintil de renda familiar *per capita* – Brasil, 1981 a 2002. 39
GRÁFICO 1. 30: Anos médios de estudo da população de 7 a 25 anos de idade segundo a cor do chefe da família – Brasil, 1981 a 2002. 39

GRÁFICO 5. 1: Impacto da pressão demográfica e da restrição de oferta na probabilidade de freqüentar a escola para os indivíduos de 7 a 14 anos, considerando os municípios com menor e maior valores. 92
GRÁFICO 5. 2: Impacto da pressão demográfica e da restrição de oferta na probabilidade de freqüentar a escola para os indivíduos de 7 a 14 anos 92
GRÁFICO 5. 3: Impacto da pressão demográfica e da restrição de oferta na probabilidade de freqüentar a escola para os indivíduos de 15 a 17 anos, considerando os municípios com menor e maior valores. 96
GRÁFICO 5. 4: Impacto da pressão demográfica e da restrição de oferta na probabilidade de freqüentar a escola para os indivíduos de 15 a 17 anos 96
GRÁFICO 5. 5: Impacto da pressão demográfica e da restrição de oferta na probabilidade de freqüentar a escola para os indivíduos de 15 a 17 anos que já concluíram o fundamental, considerando os municípios com menor e maior valores 99
GRÁFICO 5. 6: Impacto da variação dos fatores de estrutura escolar do município na probabilidade de freqüentar a escola na idade adequada - Ensino Fundamental . 105
GRÁFICO 5. 7: Impacto da variação dos fatores de estrutura escolar do município na probabilidade de freqüentar a escola na idade adequada - Ensino Médio 106
GRÁFICO 5. 8: Impacto da variação da educação da mãe e do porcentual de professores com curso superior na probabilidade de freqüentar a escola na idade correta - ensino fundamental 109
GRÁFICO 5. 9: Impacto da variação da educação da mãe e da média de alunos por turma na probabilidade de freqüentar a escola na idade correta - ensino fundamental 110
GRÁFICO 5. 10: Impacto da variação da educação da mãe e do fator de infra-estrutura na probabilidade de freqüentar a escola na idade correta - ensino fundamental .. 111
GRÁFICO 5. 11: Impacto da variação da educação da mãe e da qualidade da rede escolar do município na probabilidade de freqüentar a escola na idade correta - ensino fundamental 111
GRÁFICO 5. 12: Impacto da variação dos fatores de rede escolar do município na probabilidade de freqüentar a escola na idade adequada – Probabilidade de Progressão na 1ª série 116
GRÁFICO 5. 13: Impacto da variação dos fatores de rede escolar do município na probabilidade de freqüentar a escola na idade adequada – Probabilidade de Progressão na 5ª série 117
GRÁFICO 5. 14: Impacto da variação dos fatores de rede escolar do município na probabilidade de freqüentar a escola na idade adequada – Probabilidade de Progressão na 1ª série do ensino médio 117
GRÁFICO 5. 15: Impacto da variação da educação da mãe e da proporção de matrículas em estabelecimentos com ciclos na Probabilidade de Progressão na 1ª série 118
GRÁFICO 5. 16: Impacto da variação da educação da mãe e do porcentual de professores com curso superior na Probabilidade de Progressão na 1ª série 119
GRÁFICO 5. 17: Impacto da variação da educação da mãe e do fator de infra-estrutura na Probabilidade de Progressão na 1ª série 119
GRÁFICO 5. 18: Impacto da variação da educação da mãe e do fator de infra-estrutura na Probabilidade de Progressão na 1ª série 120
GRÁFICO 7. 1: 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 os indivíduos de 7 a 14 anos. 162
GRÁFICO 7. 2: 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 os indivíduos de 7 a 14 anos 162
GRÁFICO 7. 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 os indivíduos de 15 a 17 anos. 163
GRÁFICO 7. 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 diferentes modelos para os indivíduos de 15 a 17 anos. 163
GRÁFICO 7. 5: 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 diferentes modelos para os indivíduos de 15 a 17 anos que já concluíram o ensino fundamental. 164
GRÁFICO 7. 6: 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 diferentes modelos para os indivíduos de 15 a 17 anos que já concluíram o ensino fundamental. 164
GRÁFICO 7. 7: Variação na probabilidade de cursar a escola na idade adequada decorrente do aumento de um desvio padrão em relação à média de variáveis selecionadas – simulações segundo diferentes modelos para os indivíduos de 7 a 14 anos. 168
GRÁFICO 7. 8: Impacto da variação da educação da mãe e da qualidade da rede escolar do município na probabilidade de freqüentar a escola na idade correta (Modelo hierárquico-espacial com o resíduo *bayesiano*) - ensino fundamental 170
GRÁFICO 7. 9: Impacto da variação da educação da mãe e da qualidade da rede escolar do município na probabilidade de freqüentar a escola na idade correta - ensino fundamental (Modelo hierárquico) 170
GRÁFICO 7. 10: Variação na probabilidade de cursar a escola na idade adequada decorrente do aumento de um desvio padrão em relação à média de variáveis selecionadas – simulações segundo diferentes modelos para os indivíduos de 15 a 17 anos. 173
GRÁFICO 7. 11: Variação na probabilidade de progressão na 1ª série decorrente do aumento de um desvio padrão em relação à média de variáveis selecionadas – simulações segundo diferentes modelos. 176
GRÁFICO 7. 12: Impacto da variação da educação da mãe e da qualidade da rede escolar do município na probabilidade de freqüentar a escola na probabilidade de progressão da 1ª série (Modelo hierárquico-espacial com resíduo *bayesiano*) ... 178
GRÁFICO 7. 13: Impacto da variação da educação da mãe e do fator de infra-estrutura na Probabilidade de Progressão na 1ª série (Modelo hierárquico) 178
GRÁFICO 7. 14: Variação na probabilidade de progressão na 5ª série decorrente do aumento de um desvio padrão em relação à média de variáveis selecionadas - diferentes modelos. 181
GRÁFICO 7. 15: Variação na probabilidade de progressão na 1ª série do ensino médio decorrente do aumento de um desvio padrão em relação à média de variáveis selecionadas - diferentes modelos. 183

GRÁFICO A4. 1: Variação na probabilidade de cursar a escola na idade adequada decorrente do aumento de um desvio padrão em relação à média – simulações segundo diferentes modelos para os indivíduos de 7 a 14 anos. 216
GRÁFICO A4. 2: Variação na probabilidade de cursar a escola na idade adequada decorrente do aumento de um desvio padrão em relação à média – simulações segundo diferentes modelos para os indivíduos de 15 a 17 anos. 216
GRÁFICO A4. 3: Variação na probabilidade de progressão na 1ª série decorrente do aumento de um desvio padrão em relação à média – simulações segundo diferentes modelos..... 217
GRÁFICO A4. 4: Variação na probabilidade de progressão na 5ª série decorrente do aumento de um desvio padrão em relação à média – simulações segundo diferentes modelos..... 217
GRÁFICO A4. 5: Variação na probabilidade de progressão na 1ª série do ensino médio decorrente do aumento de um desvio padrão em relação à média – simulações segundo diferentes modelos. 218

LISTA DE MAPAS

MAPA 6. 1: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a taxa de atendimento do ensino fundamental 127
MAPA 6. 2: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para o tamanho relativo da coorte de 7 a 14 anos 127
MAPA 6. 3: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a relação professor do nível fundamental/população de 7 a 14 anos..... 128
MAPA 6. 4: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a taxa de atendimento do ensino médio 129
MAPA 6. 5: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a taxa de escolarização líquida do ensino médio 129
MAPA 6. 6: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para o tamanho relativo da coorte de 15 a 17 anos 130
MAPA 6. 7: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a relação professor do nível médio/população de 15 a 17 anos..... 130
MAPA 6. 8: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a taxa de eficiência do ensino fundamental 131
MAPA 6. 9: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a probabilidade de passar de zero para um ano de estudo (e0) 131
MAPA 6. 10: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a probabilidade de passar de quatro para cinco anos de estudo (e_4)..... 132
MAPA 6. 11: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a média de horas-aula diária no ensino fundamental 132
MAPA 6. 12: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a média de alunos por turma no ensino fundamental..... 133
MAPA 6. 13: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a porcentagem de professores com curso superior do ensino fundamental 133
MAPA 6. 14: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para o fator de infra-estrutura do ensino fundamental..... 134
MAPA 6. 15: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para o percentual de matrículas em escolas com ciclos do ensino fundamental 135
MAPA 6. 16: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a proporção de matrículas públicas do ensino fundamental..... 135

MAPA 6. 17: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a proporção de matrículas municipais dentre as públicas do ensino fundamental .. 135
MAPA 6. 18: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a proporção de estabelecimentos com até 25 matrículas no ensino fundamental ... 136
MAPA 6. 19: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a proporção de estabelecimentos com 25 a 50 matrículas no ensino fundamental . 136
MAPA 6. 20: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a proporção de estabelecimentos com 50 a 200 matrículas no ensino fundamental 136
MAPA 6. 21: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a proporção de estabelecimentos com 201 a 1000 matrículas no ensino fundamental 137
MAPA 6. 22: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a proporção de estabelecimentos com mais de 1000 matrículas no ensino fundamental 137
MAPA 6. 23: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a educação média da população adulta 137
MAPA 6. 24: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a taxa de eficiência do ensino médio 139
MAPA 6. 25: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a probabilidade de progressão na 1ª serie do ensino médio(e8)..... 139
MAPA 6. 26: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a media de horas-aula diária do ensino médio 139
MAPA 6. 27: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a media de alunos por turma do ensino médio..... 140
MAPA 6. 28: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a porcentagem de professores com curso superior do ensino médio 140
MAPA 6. 29: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para o fator de infra-estrutura do ensino médio 140
MAPA 6. 30: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a proporção de matrículas públicas no ensino médio..... 141
MAPA 6. 31: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a proporção de matrículas municipais dentre as públicas no ensino médio..... 141
MAPA 6. 32: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a proporção de estabelecimentos com menos de 25 matrículas no ensino médio... 141
MAPA 6. 33: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a proporção de estabelecimentos com 25 a 50 matrículas no ensino médio 142
MAPA 6. 34: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a proporção de estabelecimentos com 50 a 200 matrículas no ensino médio 142
MAPA 6. 35: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a proporção de estabelecimentos com 201 a 1000 matrículas no ensino médio 142
MAPA 6. 36: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a proporção de estabelecimentos com mais de 1000 matrículas no ensino médio.. 143

INTRODUÇÃO

O sistema educacional brasileiro tem evoluído significativamente nas últimas décadas, observando-se melhora nos seus indicadores, principalmente em relação ao acesso no ensino fundamental. Esse nível de ensino já se encontra praticamente universalizado, já que em 2000, 94,5% das crianças de 7 a 14 anos freqüentavam a escola, sendo que 90% delas estavam matriculadas no fundamental (IBGE, 2000).

Apesar da melhora geral dos indicadores, alguns problemas ainda persistem, como a baixa cobertura do ensino médio e a alta distorção entre a idade e a série adequada cursada. De acordo com os dados do Censo Demográfico de 2000, fornecido pelo Instituto Brasileiro de Geografia e Estatística – IBGE – do total dos jovens de 15 a 17 anos, somente 77,7% freqüentavam escola em 2000, e apenas 34,4% cursavam o ensino médio.

Com relação à taxa de distorção idade/série, os dados revelam que 54% das matrículas do ensino fundamental em 2000 eram de alunos com idade superior a considerada adequada à série. No nível médio, o quadro é ainda pior, pois 66% dos alunos matriculados nesse nível encontravam-se com idade superior a considerada adequada.

Outro fato marcante do sistema educacional brasileiro é a grande desigualdade existente nas regiões brasileiras e em grupos específicos da população. De modo geral, as regiões Norte e Nordeste são as que apresentam os piores indicadores educacionais, as regiões Sul e Sudeste, os melhores, enquanto que o Centro-Oeste está próximo da média brasileira. Para os diferentes estratos da população, percebe-se que indivíduos negros, moradores da área rural e pertencentes a famílias com menor nível sócio-econômico possuem pior performance educacional.

Com o acesso do ensino fundamental caminhando para sua universalização, essa desigualdade vem diminuindo nesse nível, pois os avanços se dão nos grupos sociais mais desprovidos de atendimento educacional. No ensino médio, observa-se um movimento inverso, ou seja, ampliação dos diferenciais entre as regiões e grupos específicos da população, conforme foi detectado com os dados dos Censos Demográficos de 1980 a 2000. Esse fato reflete a baixa cobertura desse nível de ensino, de tal forma que sua expansão atinge primeiramente as regiões mais desenvolvidas, assim como os estratos populacionais mais privilegiados.

Os baixos níveis de cobertura do ensino médio e a retenção dos alunos nas séries resultam em uma baixa escolaridade média para o Brasil. Os dados do IBGE revelam ainda que apesar dos anos médios de estudo da população brasileira acima de 15 anos de idade ter aumentado ao longo das últimas décadas, passando de 4,0 em 1980 para 6,2 em 2000, ainda se constitui em um nível bastante baixo, já que essa população não possui ao menos o ensino fundamental completo.

Esse baixo nível educacional traz para o Brasil vários problemas econômicos e sociais. No lado econômico, o baixo nível educacional prejudica o crescimento da economia e aumenta o atraso tecnológico do país em relação às outras economias (BECKER,1993). No lado social, o baixo nível educacional é considerado um dos grandes causadores da pobreza e da má distribuição de renda do Brasil. Os trabalhadores não-qualificados possuem um baixo rendimento em relação aos qualificados, com isso, investem pouco em educação nos filhos, que também terão um baixo rendimento, alimentando, portanto, um ciclo vicioso (CAMARGO & BARROS, 1991).

Desta forma, estudos que tentem compreender a dinâmica e os fatores associados ao processo educacional são importantes para entender a baixa escolaridade no Brasil e para auxiliar nas políticas educacionais.

Quando se pretende analisar os fatores que afetam a educação, objetiva-se analisar como os diversos insumos do processo educacional podem afetar os resultados educacionais de um indivíduo. Um dos principais arcabouços teóricos sobre a associação entre os insumos e o resultado educacional é dada pela Função de Produção Educacional (FPE). Ela é importante por guiar na escolha dos principais determinantes educacionais e foi utilizada como ponto de partida para essa tese.

De acordo com a abordagem da FPE, os insumos educacionais podem ser divididos em três conjuntos de fatores. O primeiro conjunto incorpora os fatores relacionados à família que alteram a demanda por educação, como, por exemplo, os recursos, a estrutura e o nível sócio-econômico familiar. O segundo agrupa aqueles associados à escola como infra-estrutura, nível educacional dos professores, tamanho de classes e organização escolar. E por fim, o terceiro conjunto engloba fatores relacionados à comunidade como estoque de capital e recursos, os quais, juntamente com o anterior, alteram a oferta educacional.

No Brasil, a maioria dos estudos sobre os determinantes educacionais tem enfocado principalmente o papel da família na educação dos filhos. Esses estudos

demonstram que os fatores de *background* familiar, principalmente a educação dos pais, possuem grande impacto no resultado educacional. Esse tipo de análise é importante para o entendimento da educação, principalmente em termos da estratificação educacional, uma vez que demonstram que pais pobres e com nível educacional mais baixo geralmente transferem baixa escolaridade para os filhos.

Por outro lado, os estudos que incluem os fatores relacionados à escola e ao ambiente comunitário, além dos fatores familiares, são mais recentes no Brasil. A maioria desses trabalhos demonstra que, apesar do importante impacto da família, os fatores relacionados à escola e à comunidade também possuem impacto no resultado educacional. Essa análise é também importante por auxiliar na formulação de políticas públicas educacionais mais eficazes.

No contexto da literatura internacional, o debate sobre o impacto dos diversos insumos no resultado educacional já está bem desenvolvido, mas longe de se chegar a um consenso. Um dos primeiro estudos a abordar essa questão foi o Relatório de Coleman (COLEMAN *et alli*, 1966), que concluiu que fatores de *background* familiar são mais importantes que fatores escolares na determinação do desempenho educacional da criança.

Esse trabalho deu subsídio a vários outros que estenderam a análise para outros indicadores educacionais. Esses estudos tanto confirmaram os achados de Coleman como discordaram. Numa revisão literária sobre os principais trabalhos relacionados aos determinantes educacionais, BUCHMANN & HANNUM (2001) concluem que o impacto da qualidade da escola e dos professores é maior em países em desenvolvimento do que em países desenvolvidos. Tal fato pode estar refletindo a maior variabilidade dos recursos escolares nos países em desenvolvimento.

No campo dos estudos educacionais há ainda outra tradição que vem da literatura demográfica econômica e que está associada com o dividendo demográfico. Um dos principais trabalhos nessa área é o de COALE & HOOVER (1956) que associa o declínio da taxa de fecundidade, com conseqüente diminuição da proporção de crianças e jovens, ao declínio do número de matrículas, levando a uma poupança governamental. Porém, se nem toda criança está matriculada, o dividendo demográfico pode levar a uma maior cobertura escolar. Estudos posteriores, principalmente o de SCHULTZ (1987), demonstram que os gastos do governo podem não diminuir na mesma proporção da diminuição da coorte em idade escolar, podendo ocorrer uma

melhora na qualidade do ensino. Cabe ressaltar, que a maioria dos estudos sobre o dividendo demográfico, não incluem variáveis relacionadas ao *background* familiar.

Dentro desse contexto, essa tese busca investigar os determinantes do resultado educacional, combinando também a tradição da demografia econômica, ou seja, incluindo os efeitos demográficos na análise. Será utilizada a abordagem da FPE como "pano de fundo" para as investigações. Essa abordagem será importante por permitir organizar de forma mais clara a escolha das variáveis relevantes para o processo educacional. Entretanto, é importante ressaltar que algumas modificações foram feitas para justá-la a estrutura e disponibilidade dos dados.

O primeiro ajuste foi a utilização de modelo logístico hierárquico de dois níveis na estimação da FPE, onde o primeiro nível considera o indivíduo e o segundo, o município. Esses modelos incorporam a estrutura hierárquica ou de agrupamento da população analisada. A vantagem de estimar a FPE utilizando os modelos hierárquicos, ao invés dos modelos clássicos, é que nesse último os resultados podem ser enviesados, devido a quebra de seus pressupostos, que são erros independentes e identicamente distribuídos (i.i.d.).

Numa análise agregada do Brasil é bem provável que estes pressupostos não se apliquem, uma vez que a relação entre as características dos estudantes e seus resultados escolares pode variar através dos municípios. Isso pode acontecer devido às diferentes disponibilidades e qualidades de recursos educacionais, bem como à diferença de estoque de capital e riqueza dos municípios. Dessa forma, ao lidar com a quebra dos pressupostos dos modelos clássicos, a modelagem hierárquica proporciona uma melhor obtenção das estimativas para os parâmetros.

Outra vantagem dos modelos hierárquicos, que será importante nessa tese, é que eles permitam estudar a interação entre as variáveis nos diferentes níveis. É possível, portanto, analisar como as variáveis independentes medidas no nível 2 modificam o efeito das variáveis independentes medidas no nível 1.

Devido à indisponibilidade de alguns dados, a segunda adaptação da FPE foi agregar os fatores escolares para a média municipal. Em outras palavras pode-se dizer que os insumos escolares são representados pelos insumos que caracterizam o perfil escolar do município.

Dessa forma, os insumos educacionais compreenderam os seguintes conjuntos de fatores. Fatores de *background* familiar, tais como educação materna, categoria de ocupação do chefe da família e estrutura familiar (chefia feminina e famílias

conviventes), que foram mensurados no primeiro nível do modelo hierárquico, ou seja, ao nível do indivíduo. Nesse nível também foram incluídas variáveis relacionadas às características individuais, como sexo, idade, cor e situação de domicílio, que tiveram a principal finalidade de controle nas regressões.

O outro conjunto de insumos reflete a qualidade e disponibilidade da oferta escolar dos municípios, ou seja, está relacionado com o perfil escolar dos municípios. Dentro desse conjunto se encontram variáveis relacionadas com: i) qualidade dos recursos humanos e da infra-estrutura das escolas; ii) restrição de oferta; iii) distribuição das matrículas segundo tamanho dos estabelecimentos; iiii) distribuição das matrículas segundo dependência administrativa e iiiii) organização do ensino em ciclos[1]. Esses fatores foram medidos ao nível do município (nível 2) e constituem um dos principais focos dessa tese.

Os municípios brasileiros possuem um perfil escolar bastante heterogêneo, com alguns possuindo a maioria dos estabelecimentos com alta qualidade da infra-estrutura e dos recursos humanos escolares, além de grande oferta em todos os níveis de ensino. No outro extremo, há municípios que possuem uma rede escolar bastante precária, que, além de ter uma grande restrição de oferta - alguns nem sequer oferecem a segunda metade do ensino fundamental e o ensino médio – possuem baixa qualidade no que se refere à infra-estrutura e recursos humanos escolares. Dessa forma, é importante saber se a rede escolar do município faz diferença para o resultado educacional de suas crianças.

O terceiro conjunto dos insumos educacionais está relacionado com a literatura demográfica econômica associada ao dividendo demográfico. Assim, 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. Esse fator é o outro foco dessa tese e também foi medido ao nível municipal.

Nesse nível também se encontram algumas variáveis de controle para as regressões, como a educação média da população adulta e o tamanho da população.

Por fim, a terceira adaptação da FPE foi tentar incluir os efeitos espaciais na análise. Quando se considera uma análise agregada da educação, dois efeitos espaciais devem ser levados em conta: a autocorrelação e a heterogeneidade espacial. A

[1] Forma de organização que distribui os anos de escolaridade em grupos.

dependência espacial ocorre quando as variáveis observadas em determinado lugar dependem dos valores observados nos locais vizinhos e a heterogeneidade, quando médias e variâncias dos erros não são constantes no espaço.

Em outras palavras pode-se dizer que as variáveis agregadas de um município podem estar sendo influenciadas pelas variáveis dos municípios vizinhos, levando a uma autocorrelação espacial. Por outro lado, podem existir unidades espaciais com comportamento diferenciado com relação aos seus vizinhos, sugerindo heterogeneidade espacial.

A econometria espacial incorpora os dois efeitos descritos acima, com modelos de regressões, quais sejam: autoregressivo de defasagem espacial e erro espacial. 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 será chamado de modelo hierárquico-espacial e trata-se de 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.

Somado a isso, a combinação das duas metodologias possibilita trabalhar com a quebra de alguns pressupostos do modelo hierárquico como a independência entre os efeitos das áreas de nível dois, nesse caso entre os municípios. Considerando que o Brasil possui uma forte heterogeneidade social e econômica, a qual se traduz numa grande heterogeneidade espacial, pode-se esperar que as áreas mais próximas sejam mais similares que as áreas mais distantes, ocorrendo, portanto, correlação espacial no nível macro.

Dessa forma, a vantagem de combinar as duas metodologias é que a modelagem hierárquica permitirá analisar conjuntamente as variáveis relacionadas aos indivíduos e as contextuais (municipais) em níveis diferentes de agregação, diminuindo o problema relacionado à falácia ecológica, além de possibilitar analisar como as variáveis interagem entre os diferentes níveis. Por sua vez, o modelo espacial permitirá 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.

Tendo em mente essas modificações na FPE, nessa tese, a análise dos determinantes educacionais será baseada em três questões chaves. O primeiro "fio condutor" será a análise do impacto do dividendo demográfico nos indicadores educacionais. Nessa análise busca-se avaliar até que ponto as profundas mudanças demográficas ocorridas no Brasil, com a significativa queda da Taxa de Fecundidade Total, e, por conseguinte, diminuição da razão de dependência dos jovens, podem trazer algum bônus demográfico para educação brasileira, no sentido de abrir uma "janela de oportunidades".

A segunda questão chave será com relação à importância dos fatores da rede escolar dos municípios, principalmente aqueles relacionados com a qualidade dos recursos humanos e infra-estrutura. Essa análise é importante por examinar quais fatores do perfil escolar dos municípios podem diminuir a importância dos fatores de *background* familiar no sentido de diminuir a estratificação educacional, ou seja, diminuir a relação entre a trajetória escolar do indivíduo e a sua origem social.

Nessa análise, a utilização dos modelos hierárquicos é importante por permitir analisar a interação entre os fatores de *background* familiar e as variáveis da rede escolar dos municípios. Pode analisar, por exemplo, como uma melhora na infra-estrutura e qualidade dos serviços educacionais ofertados em um município pode minimizar a associação entre a educação materna e o resultado educacional.

Por fim, o terceiro "fio condutor" dessa tese será investigar como o contexto espacial mais amplo interfere no resultado educacional. Essa análise é importante por dois motivos. Primeiro, porque quando se consideram os efeitos espaciais – autocorrelação e heterogeneidade espacial – a estimativa dos parâmetros relacionados as variáveis contextuais são mais confiáveis. Segundo, porque será possível analisar até que ponto políticas públicas educacionais são mais eficazes quando focalizadas apenas em determinado município ou quando levam em conta um espaço geográfico mais abrangente, ou seja, os municípios adjacentes.

Em síntese, o objetivo central dessa tese será investigar os determinantes do resultado educacional dos níveis de ensino fundamental e médio, focalizando as três questões descritas acima. A análise é feita utilizando duas metodologias diferentes: hierárquica e hierárquica-espacial. Essa última constitui em uma contribuição metodológica importante dessa tese.

Considerou-se como resultado educacional duas diferentes dimensões da educação. A primeira, diz respeito ao acesso da criança no sistema de ensino. Enquanto

que a segunda está relacionada com a eficiência e rendimento do aluno na escola, ou seja, repetência e evasão escolar.

O acesso das crianças ao sistema de ensino foi captado através da probabilidade de freqüentar a escola para os indivíduos de 7 a 14 anos e 15 a 17 anos. Essas faixas etárias são as correspondentes à idade adequada de cursar respectivamente o ensino fundamental e médio. Por sua vez, a eficiência e rendimento do sistema escolar estão relacionados com a repetência, entrada tardia no sistema de ensino e evasão. Para captar essa dimensão utilizou-se a probabilidade de cursar a escola na idade certa para as duas faixas etárias citadas acima e a Probabilidade de Progressão por Série – PPS.

A abordagem diferenciada do ensino em seus níveis médio e fundamental é importante, uma vez que estes apresentam características bastante distintas. O ensino fundamental já está quase universalizado e possui uma ampla oferta. O ensino médio, por sua vez, apresenta uma taxa de cobertura bem abaixo do fundamental. Além disto, existe uma grande restrição na oferta, uma vez que vários municípios não oferecem esse tipo de ensino e, entre os que oferecem, alguns não fazem frente à massa de alunos que está entrando no ensino médio, devido à diminuição da repetência e à universalização do ensino fundamental.

Essa tese está estruturada em sete capítulos. No primeiro capítulo é feita uma análise dos indicadores educacionais e sua evolução nas últimas décadas. Através dessa análise foi possível verificar a melhora em praticamente todos os indicadores, principalmente os relacionados ao acesso no sistema de ensino, embora o acesso dos jovens de 15 a 17 anos ainda encontra-se aquém do desejado. Por outro lado, constatou-se que os indicadores de rendimento e eficiência do sistema de ensino apesar de terem melhorado ao longo dos anos, ainda se constitui em um dos principais problemas da educação brasileira.

No segundo capítulo percorre-se a abordagem da Função de Produção Educacional (FPE) que, como já mencionado, serviu de "pano de fundo" para a análise dos determinantes educacionais. Nesse capítulo também foi feita uma revisão dos principais trabalhos que analisam os determinantes educacionais, tanto no Brasil como em outros países. O terceiro capítulo contém a estratégia empírica adotada nessa tese, com a especificação dos modelos utilizados, tratamento das variáveis dependentes e independentes e base de dados. Utilizou-se os dados do Censo Demográfico e do Censo Escolar de 2000, o primeiro fornecido pelo IBGE e o segundo pelo Instituto Nacional de Estudos e Pesquisas Educacionais – INEP. Foi feita uma combinação dessas duas bases

para obter as informações a nível individual e a nível municipal. O capítulo quatro refere-se ao arcabouço metodológico, onde se descreve os modelos hierárquico e espacial, além de discutir a metodologia de construção do modelo hierárquico-espacial. No capítulo seguinte, são apresentados os resultados do modelo hierárquico, ou seja, da primeira especificação da FPE, que não leva em consideração os efeitos espaciais na análise. No sexto capítulo objetiva-se incluir o espaço na análise, verificando a existência da autocorrelação e heterogeneidade espacial. Para tanto foram utilizadas as técnicas de Análise Exploratória Espacial dos Dados (ESDA) e a econometria espacial. Através da ESDA foi possível descrever a distribuição espacial das variáveis, identificando padrões de associação e heterogeneidade espacial. Com a regressão espacial foi possível identificar o verdadeiro processo de autocorrelação espacial das variáveis. No capítulo 7 são apresentados os resultados do modelo hierárquico-espacial. Seus resultados foram comparados com os do modelo hierárquico não-espacial. Finalmente, na última parte da tese encontra-se a conclusão, com a síntese dos principais resultados dessa tese.

1) INDICADORES EDUCACIONAIS E SUA EVOLUÇÃO NAS ÚLTIMAS DÉCADAS

Nesse capítulo pretende-se fazer uma análise descritiva de alguns indicadores educacionais. Os indicadores considerados foram: taxa de analfabetismo, taxa de atendimento, taxa de escolarização bruta e líquida, taxa de distorção idade/série, anos médios de estudo e Probabilidade de Progressão por Serie (PPS), calculados através dos Censos Demográficos e Pesquisa Nacional de Amostra por Domicílio (PNAD) mais recentes. Esse conjunto de indicadores possibilita identificar alguns aspectos e problemas do sistema de ensino brasileiro, bem como a sua evolução nas últimas décadas, dando subsídios para as análises posteriores e interpretação dos resultados dessa tese.

1.1) Taxa de Analfabetismo

A taxa de analfabetismo é bastante utilizada por órgãos internacionais como um indicador que mede os níveis de desenvolvimento socioeconômico dos países. Representa o quociente entre a população analfabeta e a população total de um mesmo grupo etário. Geralmente, calcula-se essa taxa para as pessoas acima de 10 ou 15 anos de idade Considera-se analfabeto aquele indivíduo que é incapaz de ler e escrever pelo menos um bilhete simples na sua língua de origem. A fórmula abaixo mostra como calcular a taxa de analfabetismo:

$$TA = \frac{P_{ana}}{P} \times 100 \qquad (1.1)$$

Sendo:

TA é a taxa de analfabetismo;

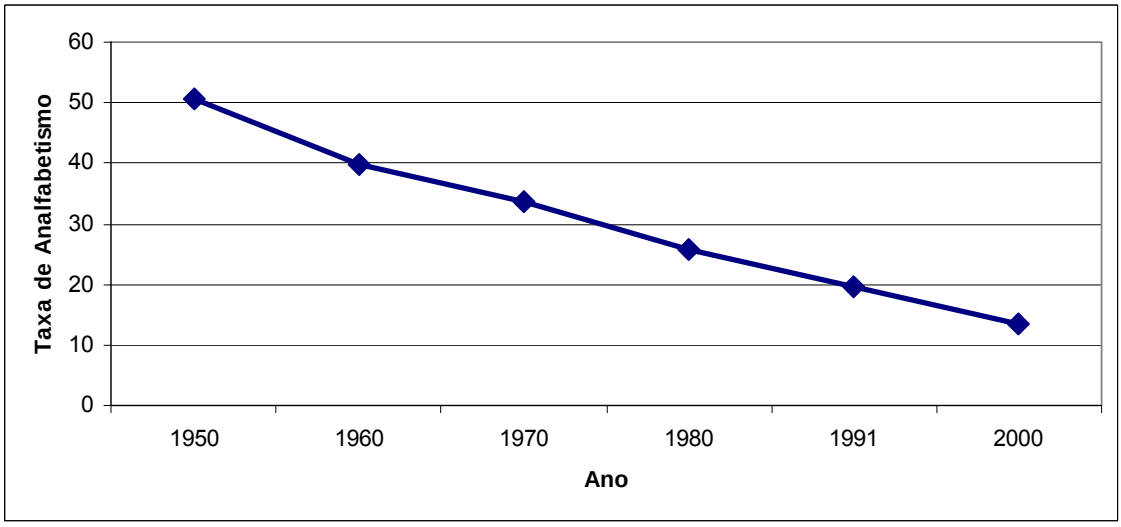
P_{ana} é a população analfabeta de um determinado grupo etário;

P é a população total nesse mesmo grupo etário.

Como se sabe, a taxa de analfabetismo no Brasil apresentou uma significativa redução nas últimas décadas. O GRAF.1.1 mostra a evolução da taxa de analfabetismo para a população brasileira de 15 anos ou mais no período de 1950 a 2000. Percebe-se

que esta taxa diminuiu muito na segunda metade do século vinte: de 50% da população, nos anos 50, para aproximadamente 14% ao final do período.

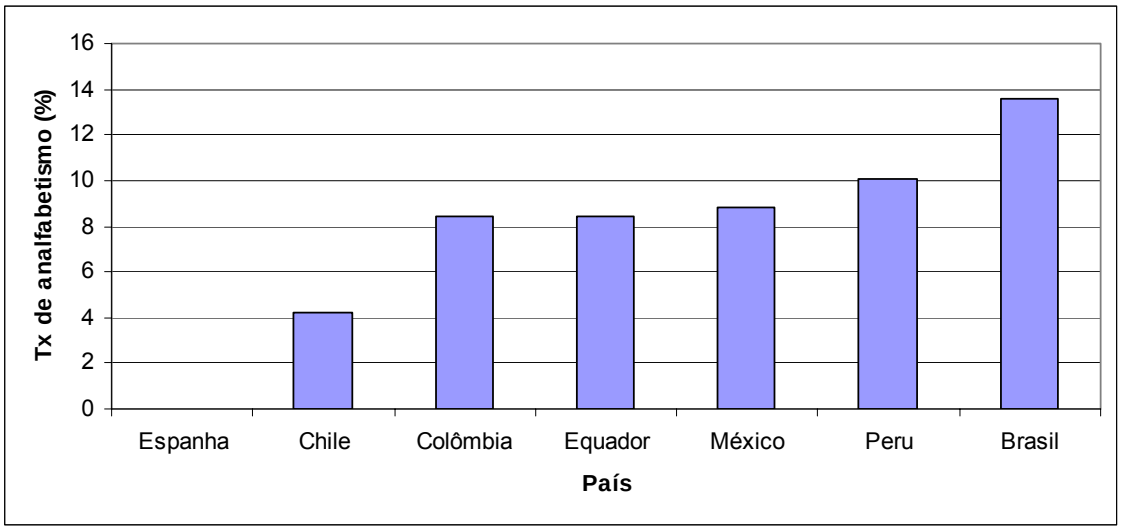
GRÁFICO 1. 1: Taxa de analfabetismo na faixa de 15 anos ou mais – Brasil, 1950/2000



Fonte: IBGE, Censos Demográficos.

Apesar dessa marcante queda, a taxa de analfabetismo brasileira ainda apresenta um nível muito elevado quando comparada com outros países, inclusive países com nível de renda semelhante ao nosso. O GRAF. 1.2 compara os dados brasileiros com os de países selecionados, alguns com renda *per capita* inferior à nossa. Como se pode notar, a Espanha apresenta uma taxa de analfabetismo desprezível.

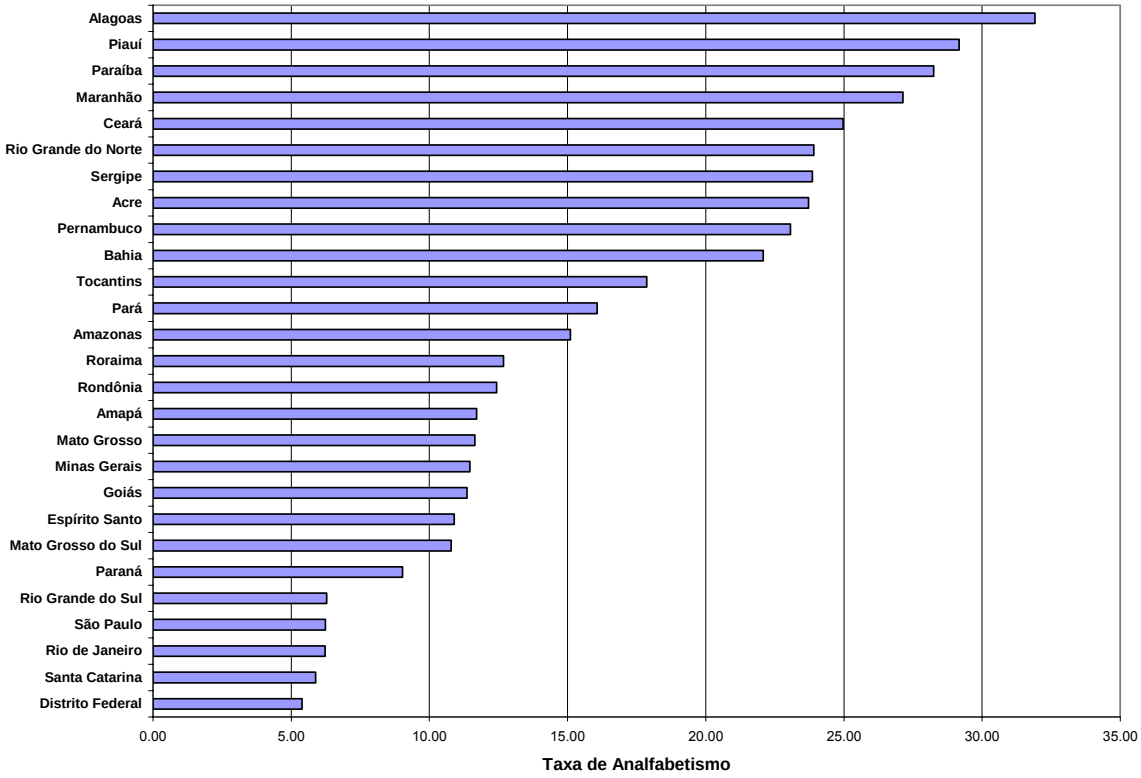
GRÁFICO 1. 2: Taxa de analfabetismo da população de 15 anos ou mais para diferentes países – 2000



Fonte: PNUD e Unesco.

Além de exibir elevadas taxas de analfabetismo, o Brasil é um país extremamente heterogêneo. O GRAF. 1.3 mostra a grande desigualdade regional nessas taxas, apresentando os dados desagregados por estado. As taxas variam de 5,7%, no Distrito Federal, a 33,4% em Alagoas. De maneira geral, os estados das regiões Nordeste e Norte são os que apresentam as maiores taxas, o contrário ocorrendo nas regiões Sul e Sudeste. Grosso modo, de acordo com o GRAF. 1.3, os estados brasileiros podem ser classificados em três grupos distintos quanto ao nível dessa taxa: os de melhores índices, do Distrito Federal até o Rio Grande do Sul; os intermediários, do Paraná até Tocantins; e os que apresentam as maiores taxas, da Bahia em diante. Essa situação evidencia a necessidade de implementação de políticas públicas específicas voltadas para as áreas mais atrasadas em termos educacionais.

GRÁFICO 1. 3: Taxa de analfabetismo da população de 15 anos ou mais – Estados 2000



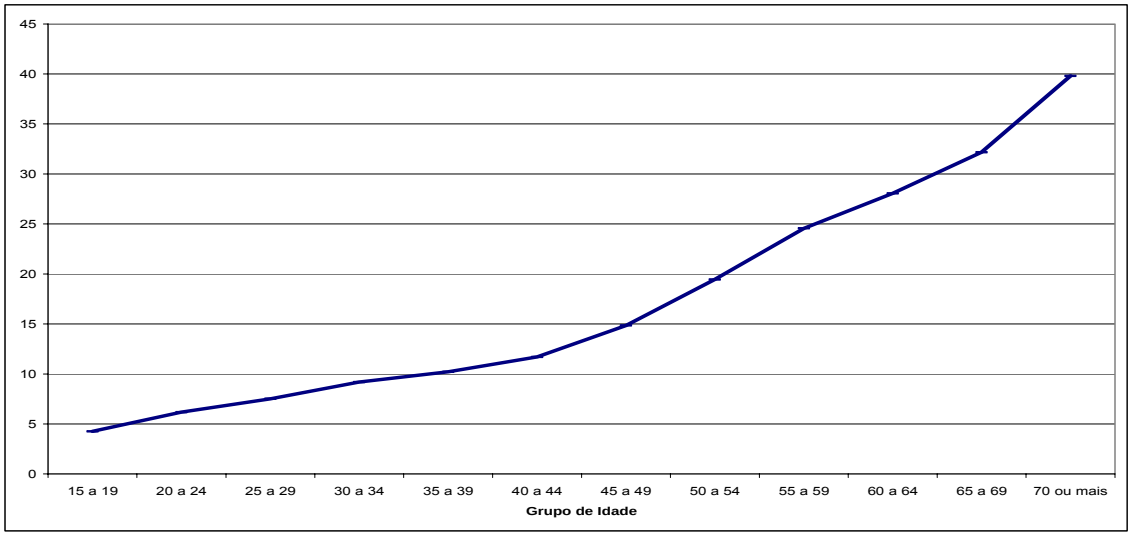
Fonte: Censo Demográfico de 2000.

Quando a taxa de analfabetismo é calculada para toda a população acima de uma certa idade, como é o caso dos gráficos acima, referentes a pessoas com 15 anos ou mais de idade, ela engloba toda a história passada do sistema de educação, ou seja, é uma medida de estoque. Desta forma, ela não seria apropriada para avaliar os recentes avanços na educação, uma vez que políticas educacionais geralmente são voltadas para a população em idade escolar, com prioridade para o sistema de ensino básico. Avanços educacionais recentes são, portanto, diluídos pelas deficiências do passado.

Para captar o avanço de políticas voltadas para o ensino básico e auxiliar na formulação de políticas de alfabetização de adultos, é conveniente analisar a taxa de analfabetismo por faixa etária. Desta forma, também se estaria tirando o efeito composição da estrutura etária que está embutido nesta taxa. Como o maior número de analfabetos concentra-se nas idades mais avançadas, uma população mais envelhecida pode ter uma taxa de analfabetismo mais alta que outra mais jovem, mesmo que elas tenham as mesmas taxas por grupos etários.

No GRAF. 1.4 encontram-se as taxas de analfabetismo para diferentes grupos de idade em 2000. Fica bastante evidenciado o diferencial entre as gerações. A maior taxa nas populações mais idosas é resultado da deficiência do sistema de ensino em anos passados. Os níveis próximos de zero para os grupos etários mais jovens captam o avanço na educação que vem ocorrendo de maneira gradativa e constante nas últimas décadas. Este quadro evidencia a importância de políticas de alfabetização de adultos para garantir um nível mínimo de dignidade e cidadania para a população mais idosa.

GRÁFICO 1. 4: Taxa de analfabetismo por grupo etário – Brasil, 2000.

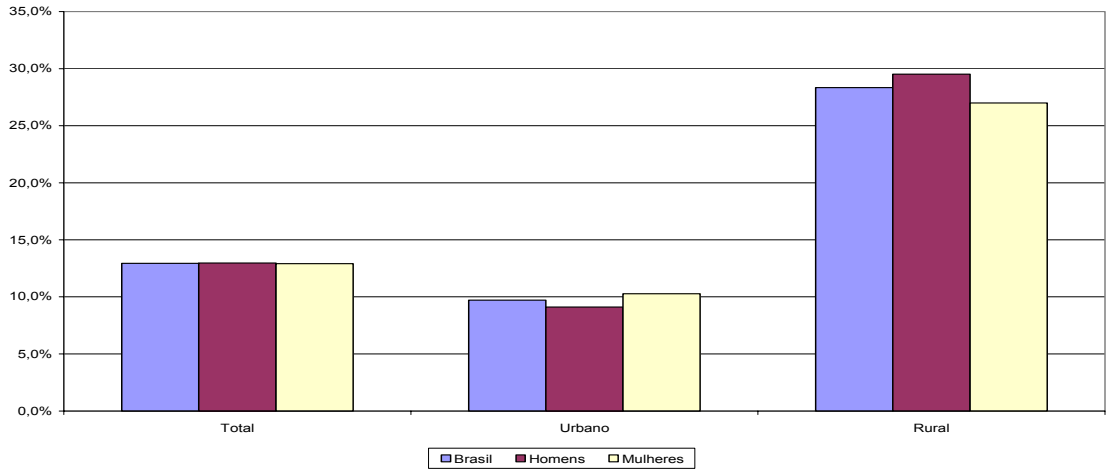


Fonte: Censo Demográfico de 2000.

Além das diferenças observadas entre estados e grupos etários, as taxas de analfabetismo também diferem bastante entre outros grupos específicos da população, conforme pode ser observado nos GRAFs. 1.5 e 1.6. No primeiro gráfico, as taxas são calculadas segundo sexo e situação de domicílio e percebe-se que há uma grande diferença entre os setores rural e urbano, sendo este último bem mais avançado em termos educacionais que o primeiro. Exemplificando, essa taxa em 2000 era por volta de 9,7% nas cidades, contra 28,3% no setor rural. Com relação ao sexo, verifica-se nenhuma diferença no total e pequenas diferenças por situação de domicílio: no setor rural, as mulheres apresentavam uma menor taxa de analfabetismo, o contrário sendo observado no setor urbano. Cabe ressaltar, que essas diferenças entre os sexos podem estar sendo minimizadas ou aumentadas dependendo da diferente estrutura etária entre

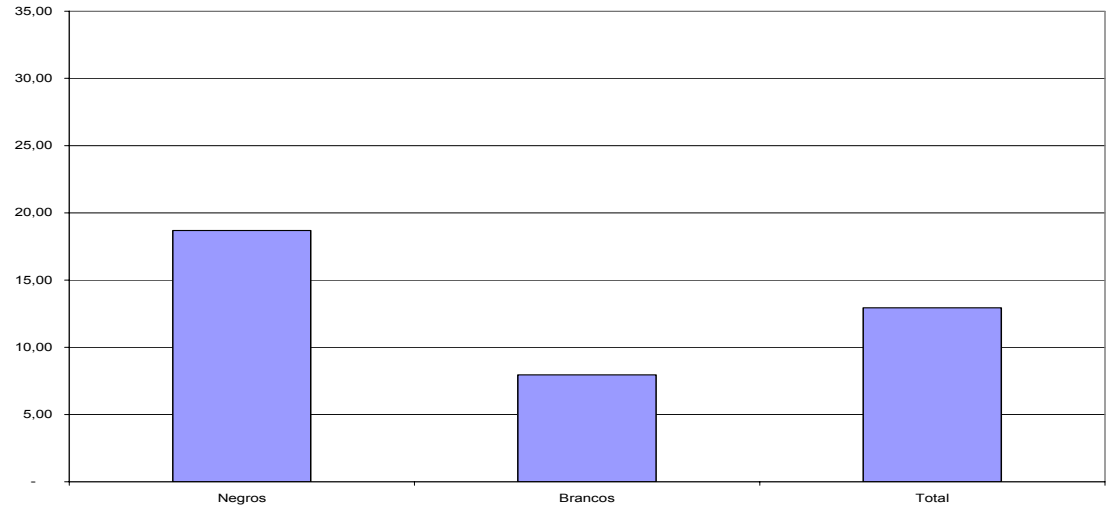
homens e mulheres, pois, como já mencionado, populações mais velhas possuem maiores índices de analfabetismo.

GRÁFICO 1. 5: Taxa de analfabetismo para a população de 15 anos ou mais por sexo e situação de domicílio – Brasil, 2000.



Fonte: Censo Demográfico de 2000.

GRÁFICO 1. 6: Taxa de analfabetismo para a população de 15 anos ou mais segundo grupos de cor – Brasil, 2000



Fonte: Censo Demográfico de 2000.

Quando se analisa as taxas de analfabetismo segundo grupos de cor (GRAF. 1.6), percebe-se que as diferenças entre brancos/amarelos e negros (pretos e pardos) são bem acentuadas. Os negros possuem mais que o dobro da taxa de analfabetismo que os brancos e amarelos, sendo de 18,69% e 7,95% respectivamente. Cabe ressaltar, que essa

desigualdade educacional entre negros e brancos pode estar refletindo o diferencial de rendimentos entre esses grupos.

1.2) Taxa de Atendimento Escolar

A taxa de atendimento escolar ou taxa de freqüência escolar por grupo etário capta a proporção da população em uma determinada faixa etária que freqüenta escola, podendo avaliar a capacidade do sistema de ensino de manter as crianças e adolescentes nas escolas. Em geral, considera-se a faixa etária adequada para se cursar determinado grau, ou seja, 4 a 6 anos para o ensino infantil, 7 a 14 para o ensino fundamental, 15 a 17 para o ensino médio e 18 a 22-24 anos para o ensino superior. A fórmula que expressa esse indicador é mostrada a seguir:

$$TAE_i = \frac{MAT_{(i)}}{P_i} \times 100 \qquad (1.2)$$

Sendo:

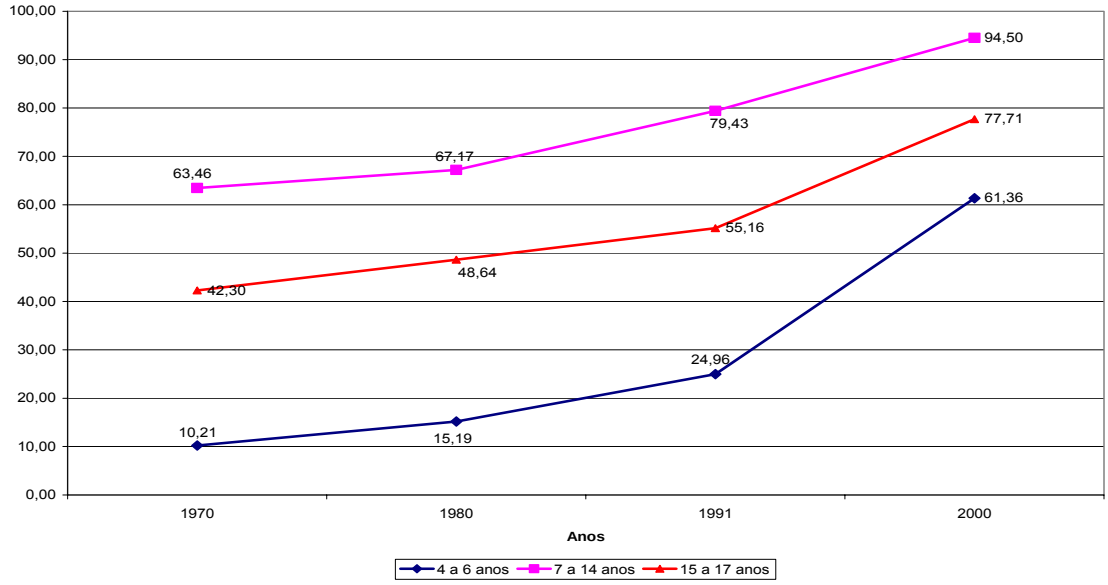
TAE_i é a taxa de atendimento escolar para determinada faixa etária;

MAT_i é a matrícula em todos os níveis de ensino na faixa etária selecionada;

P_i é a população na mesma faixa etária.

O GRAF. 1.7 traz as taxas de atendimento para as faixas etárias correspondentes ao ensino infantil, fundamental e médio no período de 1970 a 2000. Percebe-se que nas três décadas, ocorreu um aumento no atendimento escolar em todas as faixas etárias analisadas, sendo que o maior ritmo de crescimento para as crianças de 7 a 14 anos se deu com maior intensidade a partir da década de 80, enquanto que nos demais grupos etários na década de 90. Para as pessoas de 7 a 14 anos, pode-se dizer que já há uma quase universalização do acesso no ensino fundamental, uma vez que em 2000, 94,5% das pessoas nessa faixa freqüentavam escola. Parte disso pode ser explicado pelas recentes políticas educacionais de retenção de crianças na escola e ampliação da oferta de escolas e vagas.

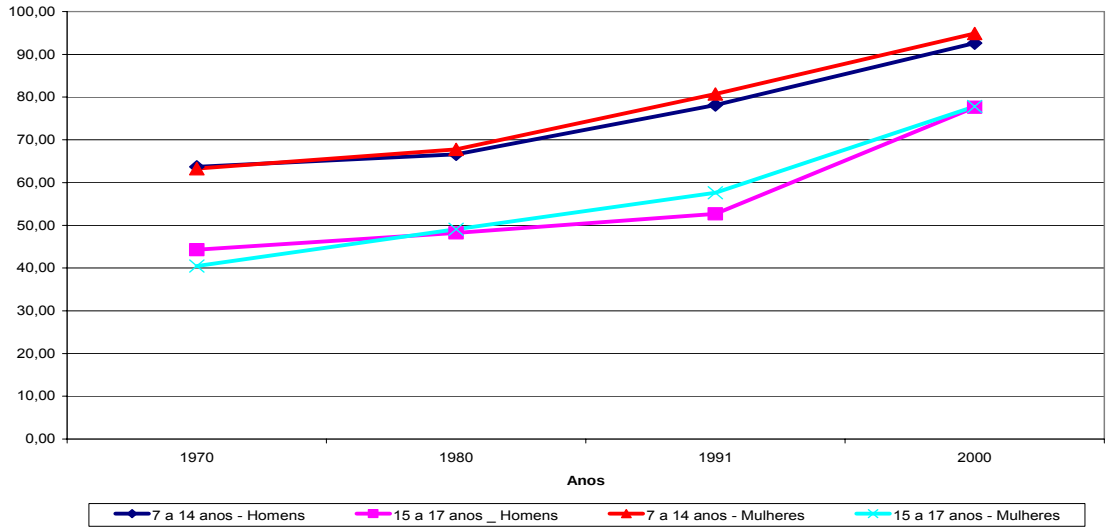
GRÁFICO 1. 7: Taxa de atendimento segundo grupos de idade – Brasil, 1970/2000



Fonte: Censos Demográficos de 1970, 1980, 1991 e 2000.

Quando se considera a evolução da taxa de atendimento por sexo, GRAF. 1.8, verifica-se que em 1970 a taxa de atendimento masculina, nas duas faixas etárias, era um pouco superior a feminina, fato que se inverte nos Censos de 1980 e 1991. Porém, os dados mais recentes de 2000, mostram que praticamente não existe diferencial entre os sexos.

GRÁFICO 1. 8: Taxa de atendimento por grupos de idade e sexo – Brasil, 1970 a 2000.

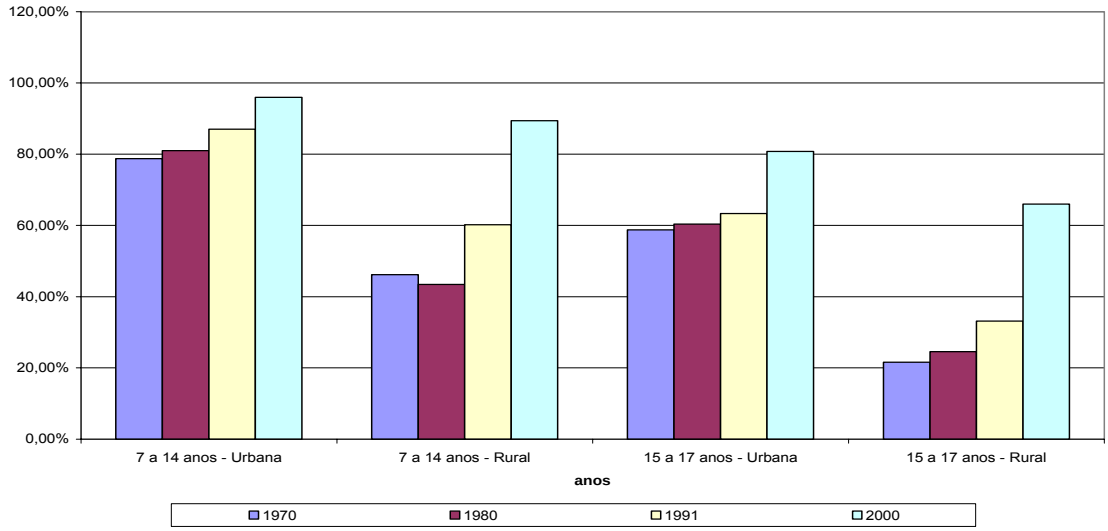


Fonte: Censos Demográficos de 1970, 1980, 1991 e 2000.

Um exame dessas taxas por situação de domicílio, GRAF. 1.9, mostra que apesar de ainda persistirem diferenças entre o urbano e o rural, a taxa desse último apresentou nos anos 90 maior crescimento para as duas faixas etárias analisadas. Tal fato, minimiza as desigualdades entre os setores.

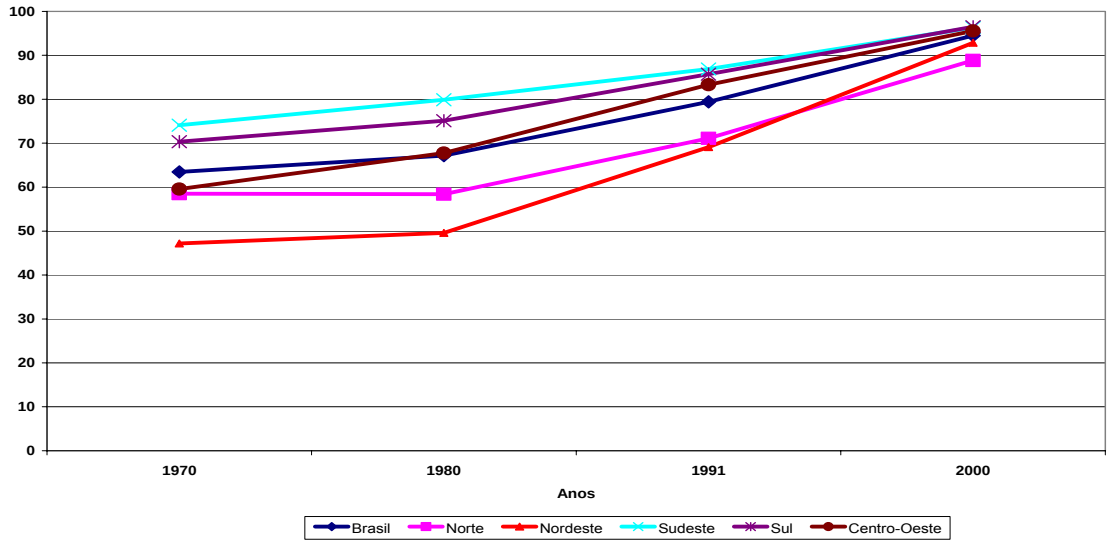
Da mesma forma, quando se analisa a taxa de atendimento da faixa etária de 7 a 14 anos por Grandes Regiões (GRAF. 1.10), observa-se que as diferenças regionais também vêm diminuindo. Apesar de todas as regiões apresentarem crescimento desta taxa no período analisado, a região Nordeste foi a que teve maior aumento. A diminuição das diferenças regionais e entre os setores rural e urbano pode ser atribuída, em parte, ao reflexo de políticas focalizadas para corrigi-las, dentre elas, a criação de um novo modelo de financiamento da educação – o Fundo de Manutenção e Desenvolvimento de Ensino Fundamental e de Valorização do Magistério (FUNDEF).

GRÁFICO 1. 9: Taxa de atendimento segundo grupos de idade e situação de domicílio – Brasil, 1970/2000



Fonte: Censos Demográficos de 1970, 1980, 1991 e 2000.

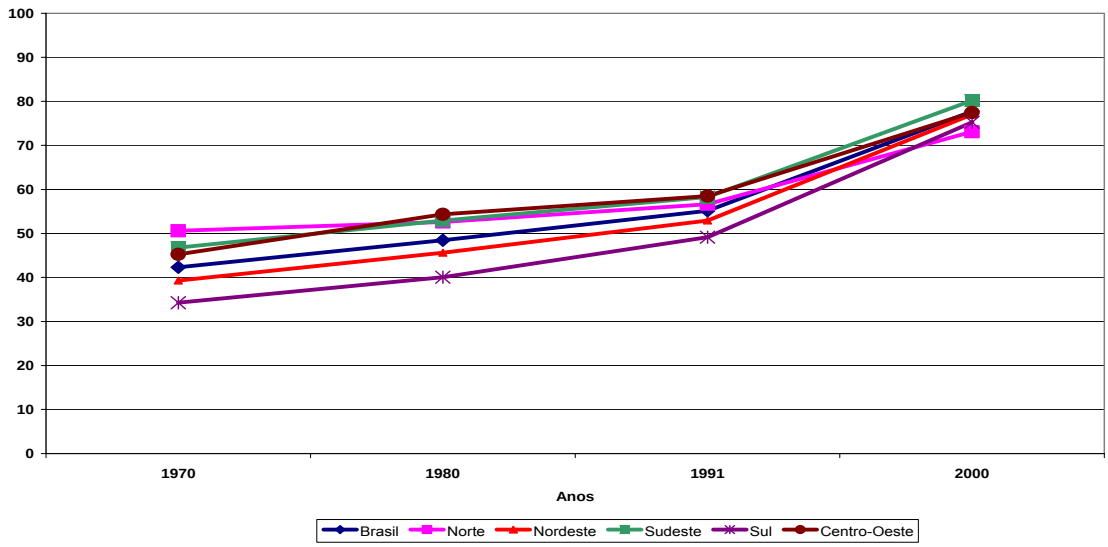
GRÁFICO 1. 10: Taxa de atendimento para a faixa etária de 7 a 14 anos para Brasil e grandes regiões 1970 a 2000



Fonte: Censos Demográficos de 1970, 1980, 1991 e 2000.

A análise da evolução histórica da taxa de atendimento para a faixa etária de 15 a 17 anos revela um quadro ambíguo (GRAF. 1.11). A região Sul possui taxa menor que as regiões Norte e Nordeste em quase todo o período analisado. Fato contraditório, pois é sabido que o sistema educacional dessas duas últimas regiões é mais precário que o da primeira. Uma explicação pode ser a elevada repetência da região Nordeste e Norte, o que acarreta um alto percentual de pessoas ainda cursando o Ensino Fundamental. Associado a isso, como o Ensino Médio ainda não tem sido foco de políticas de retenção, há uma maior evasão das pessoas nesta faixa de idade que já concluíram o Ensino Fundamental. Dessa forma, como na região Sul a repetência é menor, os jovens de 15 a 17 anos que já concluíram o Ensino Fundamental evadem, o que causa um menor taxa de atendimento para esta faixa etária.

GRÁFICO 1. 11: Taxa de atendimento para a faixa etária de 15 a 17 anos para Brasil e grandes regiões 1970 a 2000



Fonte: Censos Demográficos de 1970, 1980, 1991 e 2000.

1.3) Taxa de Escolarização Bruta

Este indicador é dado pela razão entre as matrículas em um determinado nível de ensino e a população em idade adequada para cursar tal nível. Ele possibilita avaliar o volume de matrículas nesse nível em função da demanda potencial na faixa etária adequada.

A expressão que calcula esse indicador é dada pela fórmula a seguir:

$$TEB = \frac{MAT_{(j)}}{P_i} \times 100 \qquad (1.3)$$

Sendo:

TEB é a taxa de escolarização bruta;

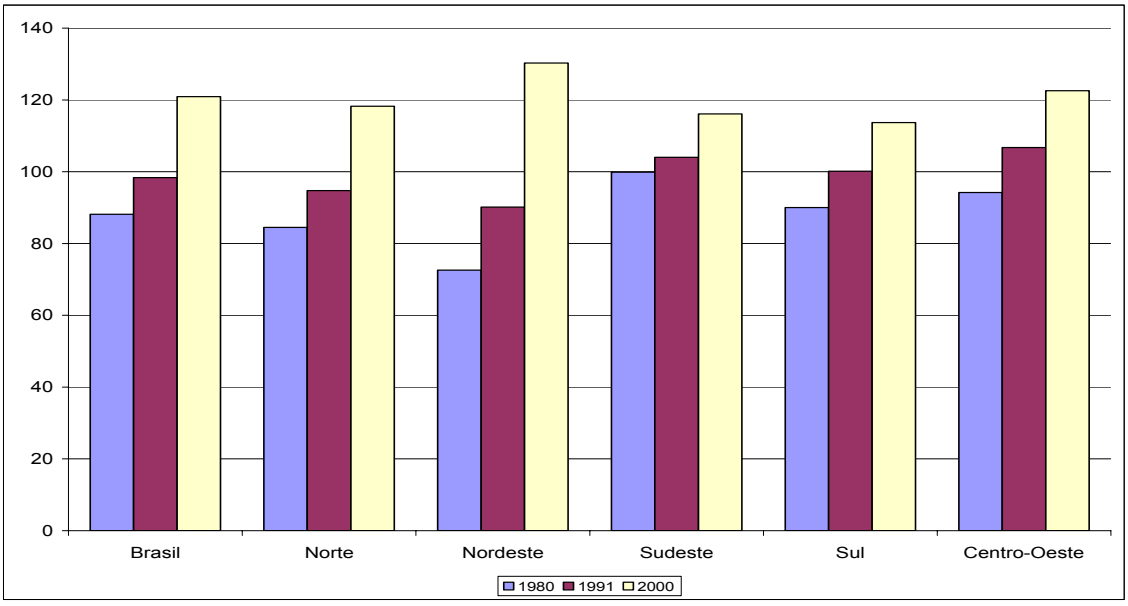
MAT_j é a matrícula total em um determinado nível de ensino; e

P_i é a população na faixa etária adequada a esse nível de ensino.

Como o numerador desta taxa é a matrícula total independente da idade, ela pode ser inflada devido ao grande número de alunos que se encontram fora da idade adequada de cursar determinado nível de ensino. A grande proporção de pessoas fora da faixa apropriada ocorre devido à entrada tardia na escola ou à repetência.

Durante as duas últimas décadas, a taxa de escolarização bruta do ensino fundamental apresentou crescimento para todas regiões, ultrapassando 100% em todos os casos, de tal forma que o Brasil alcançou em 2000 uma taxa de 120,91% (GRAF. 1.12). Esse número indica que existia nesse ano quase 21% de pessoas cursando o ensino fundamental a mais que o tamanho da população de 7 a 14 anos.

GRÁFICO 1. 12: Taxa de escolarização bruta do ensino fundamental – Brasil e regiões, 1980 a 2000

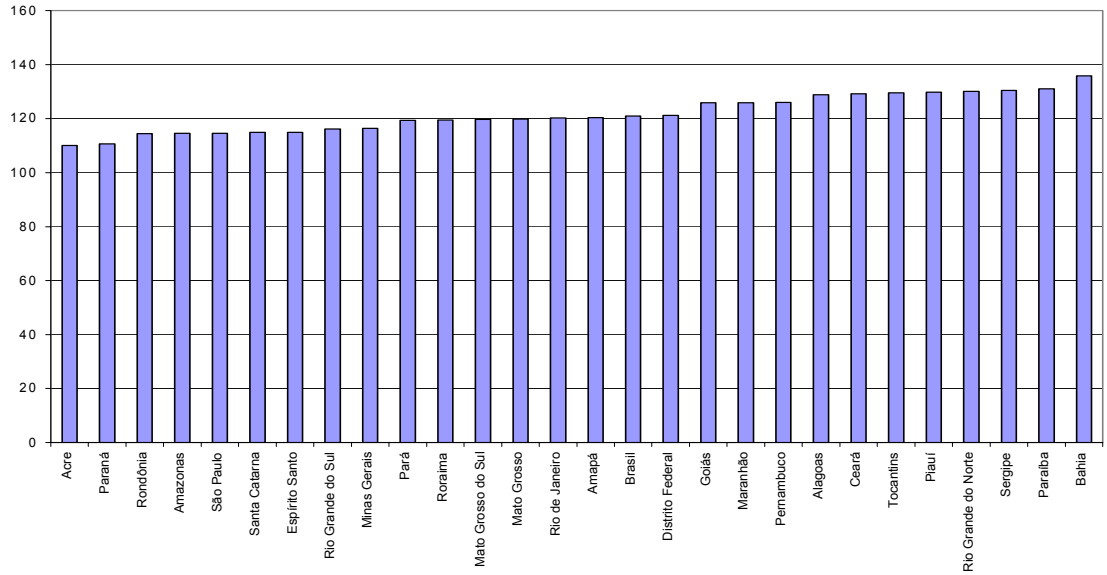


Fonte: Censos Demográficos de 1980, 1991 e 2000.

Percebe-se, portanto, que este indicador é um pouco ambíguo. Ele pode aumentar tanto por aspectos positivos como por negativos. Um aumento do atendimento escolar eleva a taxa de escolarização bruta, sendo este um fator positivo, já que significa que há mais crianças sendo atendidas pelo sistema de ensino. Em contrapartida, uma maior repetência pode aumentar essa taxa, o que caracteriza um fato negativo, que é a maior retenção de estudantes em determinada série.

A análise da taxa de escolarização bruta do ensino fundamental para os estados em 2000 mostra que em todos os estados essa taxa ultrapassa 100%, evidenciando um quadro de ineficiência do sistema de ensino, com alto grau de repetência escolar e/ou entrada tardia no sistema de ensino. Observa-se também que a maioria dos estados que apresentam valores maiores que a média são os que possuem nível socioeconômico mais baixo, com exceção do Distrito Federal, o que indica que os aspectos negativos que trazem ambigüidade ao indicador ainda são os preponderantes.

GRÁFICO 1. 13: Taxa de escolarização bruta do ensino fundamental – Brasil e estados, 2000

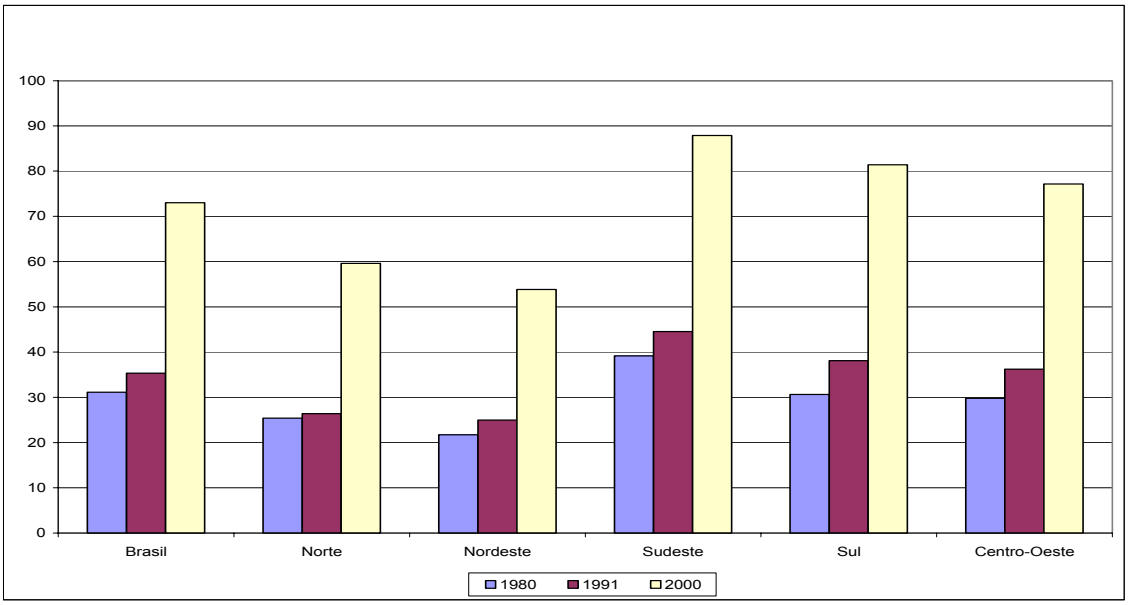


Fonte: Censo Demográfico de 2000.

A taxa de escolarização bruta do ensino médio, apesar de se encontrar em patamares bem inferiores que do ensino fundamental, apresentou um acentuado crescimento durante os anos de 80 e 90, passando a média nacional de 31,11% em 1980 para 73,02 em 2000. Esse crescimento expressivo foi observado em todas as regiões, sendo mais acentuado na região Norte, principalmente durante a década de 80 (GRAF. 1.14).

O fato do ensino médio apresentar taxa de escolarização bruta menor que o fundamental é devido, em parte, ao grande contingente de pessoas de 15 a 17 anos que ainda estão cursando o Ensino Fundamental e, em parte, à maior evasão nas idades mais velhas.

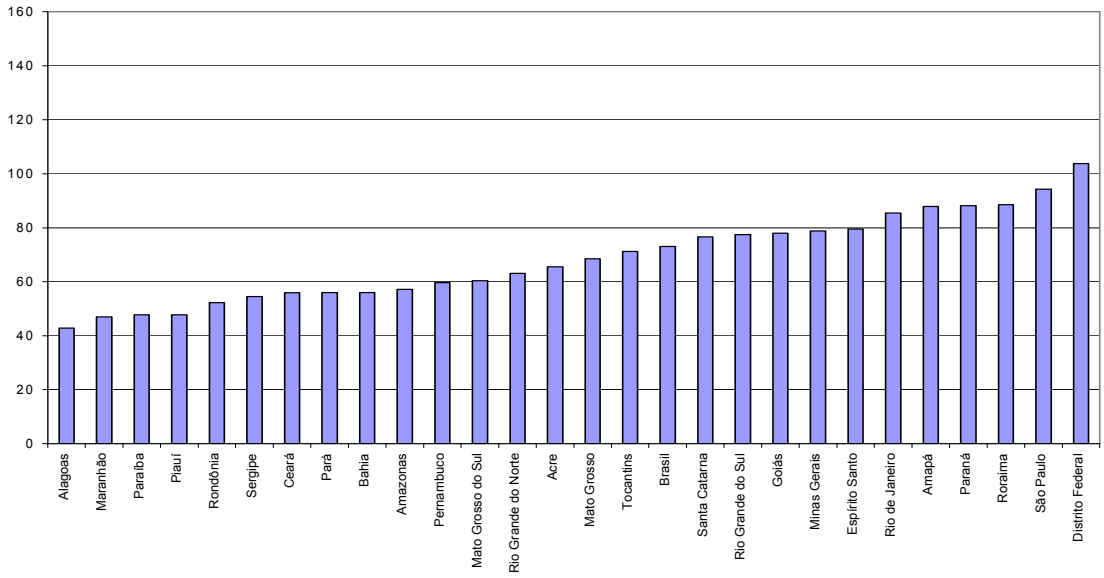
GRÁFICO 1. 14: Taxa de escolarização bruta do ensino médio – Brasil e regiões, 1980 a 2000



Fonte: Censos Demográficos de 1980, 1991 e 2000.

O GRAF. 1.15 mostra que a ambigüidade do indicador nesse nível de ensino tende para os aspectos positivos. Os estados que apresentam os maiores valores são, de forma geral, aqueles com os melhores indicadores socioeconômicos.

GRÁFICO 1. 15: Taxa de escolarização bruta do ensino médio – Brasil e estados, 2000



Fonte: Censo Demográfico de 2000.

1.4) Taxa de Escolarização Líquida

A ambigüidade observada no indicador anterior não está presente na taxa de escolarização líquida. Este indicador corresponde à razão entre as matrículas das pessoas em idade adequada para estar cursando um determinado nível e a população total na mesma idade, ou seja, indica a porcentagem da população na faixa etária que está matriculada no nível de ensino adequado. Como este indicador não capta os estudantes que estão atrasados e fora de seu nível adequado de ensino, ele é mais apropriado para avaliar a eficiência do sistema de ensino do que o anterior, já que um crescimento dessa taxa só ocorre, basicamente, por fatores positivos. Note-se que um estudante de 10 anos de idade na 1ª série do ensino fundamental tem idade certa para estar no nível em que se encontra, mas não está na série adequada.

A expressão que calcula esse indicador é dada pela fórmula (1.4):

$$TEL = \frac{MAT_{(ij)}}{P_i} \times 100 \qquad (1.4)$$

Sendo:

TEL é a taxa de escolarização líquida;

MAT_{ij} é a matrícula na faixa etária adequada a um determinado nível de ensino; e

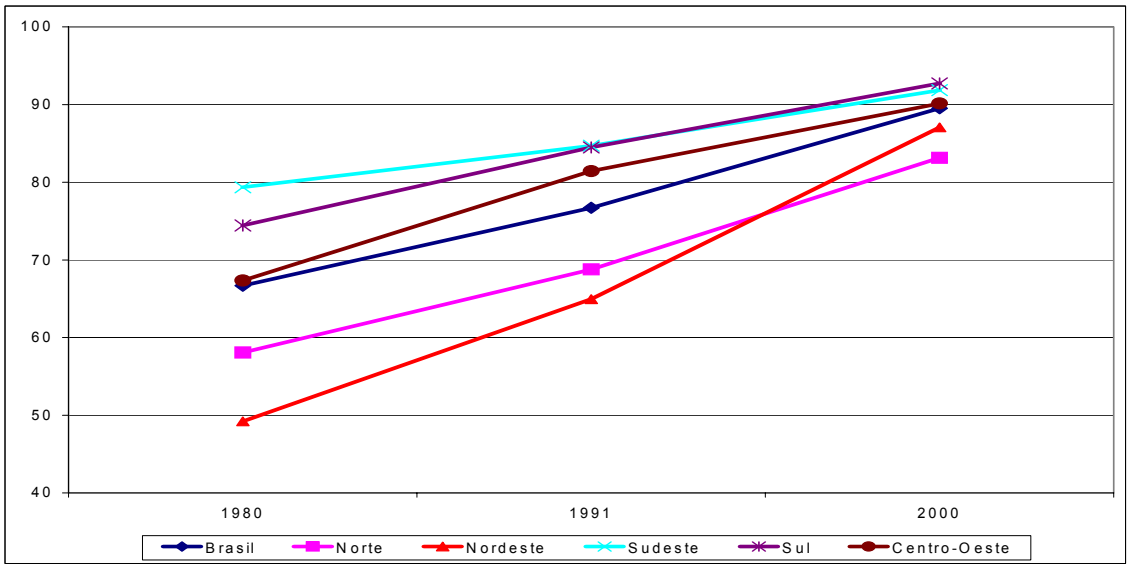
P_i é a população na mesma faixa etária.

Quando analisa-se a evolução dessa taxa entre os Censos de 1980 e 2000 tanto para ensino fundamental quanto para o médio (GRAF. 1.16 e 1.17), observa-se essas taxas também apresentaram crescimento nesse período, sendo bem mais significativo para o ensino médio, com variação de 15,02% entre 1980 e 1991 e de 119,60% entre 1991 e 2000. Apesar desse expressivo aumento, o ensino médio ainda apresenta um nível bem abaixo do desejado: apenas 34,39% dos estudantes brasileiros e 15 a 17 anos freqüentam este nível.

Com relação às diferenças regionais, cabe destacar que no ensino fundamental, apesar de ainda existirem, a tendência é de diminuição. Por outro lado, no ensino médio observa-se um aumento das desigualdades entre as regiões, formando três segmentos bem diferenciados, Sul e Sudeste com taxas bem semelhantes e acima da brasileira,

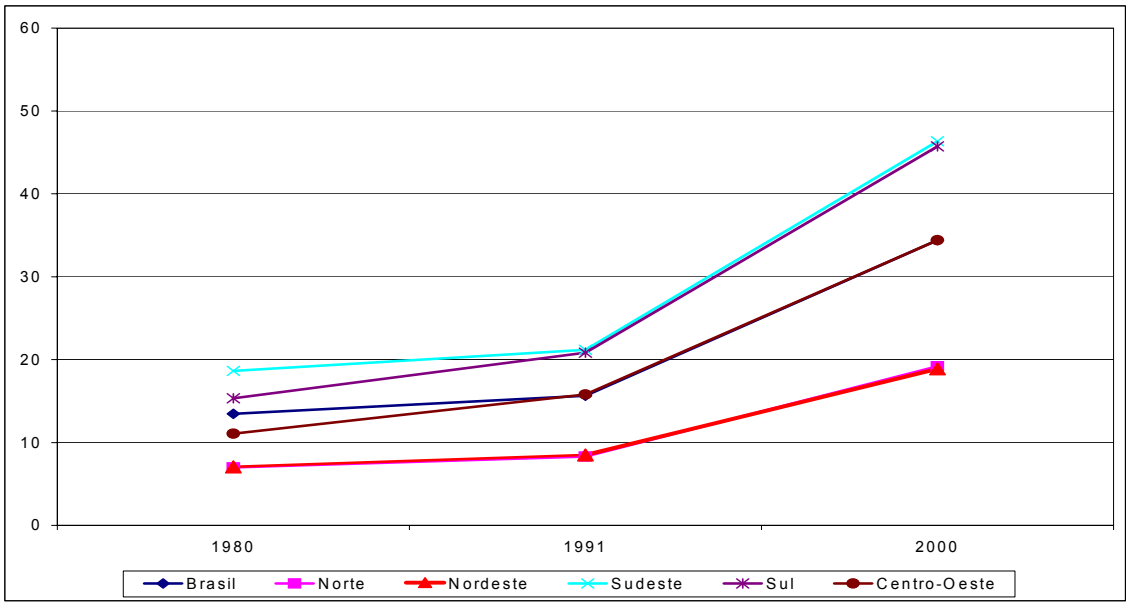
Norte e Nordeste com taxas próximas e abaixo a do Brasil e o Centro-Oeste representando a média nacional, ou seja, taxa quase igual a do Brasil. Cabe ressaltar que no ensino fundamental o Norte e Nordeste foram as regiões que apresentaram maior crescimento, enquanto que no médio foram o Sul e Sudeste.

GRÁFICO 1. 16: Taxa de escolarização líquida do ensino fundamental – Brasil e Regiões, 1980 a 2000.



Fonte: Censos Demográficos de 1980, 1991 e 2000.

GRÁFICO 1. 17: Taxa de escolarização líquida do ensino médio – Brasil e Regiões, 1980 a 2000.



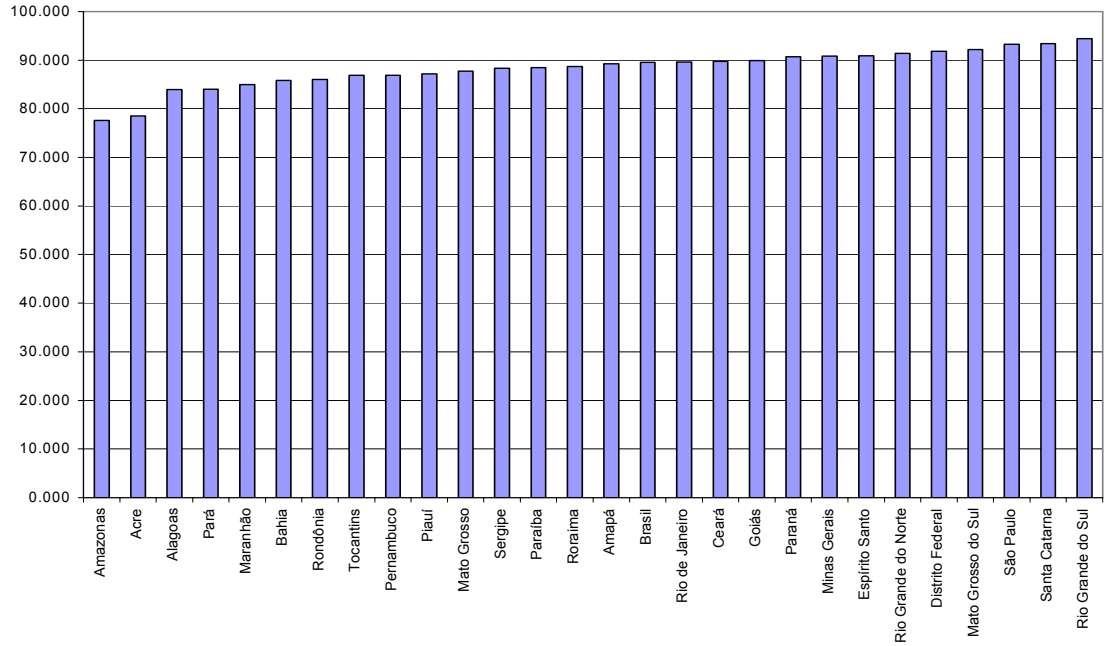
Fonte: Censos Demográficos de 1980, 1991 e 2000.

O GRAF. 1.18 mostra os resultados da taxa de escolarização líquida do ensino fundamental por estado. Observe-se que no caso da escolarização líquida a ambigüidade existente na escolarização bruta não é mais verificada. Os estados que reconhecidamente detêm os melhores indicadores socioeconômicos estão entre os que têm as maiores cifras. O contrário é verdadeiro no outro extremo. Chama a atenção o desempenho de dois estados do Nordeste, Rio Grande do Norte e Ceará, que apresentam valores acima da média nacional. Esses estados mostraram um grande avanço recente em vários aspectos sociais e esse indicador está captando parte dessas melhorias.

Apesar de haver pequenas diferenças estaduais nas taxas de escolarização líquida do Ensino Fundamental, com a maioria dos estados do Sul, Sudeste e Centro-Oeste apresentando maiores taxas que os estados do Norte e Nordeste, pode-se dizer que este nível de ensino está praticamente universalizado, posto que 90% dos estudantes brasileiros de 7 a 14 anos hoje estão freqüentando o Ensino Fundamental.

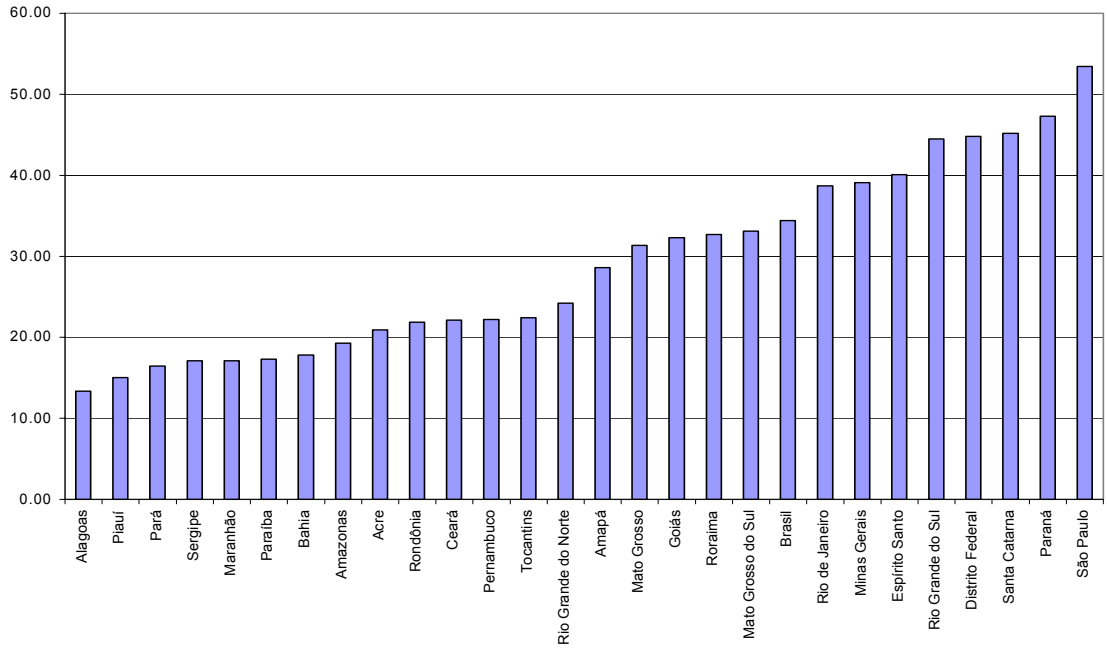
No caso do ensino médio (GRAF. 1.19), as disparidades entre os estados também são bem mais acentuadas, com Alagoas apresentando o pior desempenho (13,34%) e São Paulo, o melhor (53,41%).

GRÁFICO 1. 18: Taxa de escolarização líquida do ensino fundamental – Brasil e estados, 2000



Fonte: Censo Demográfico de 2000.

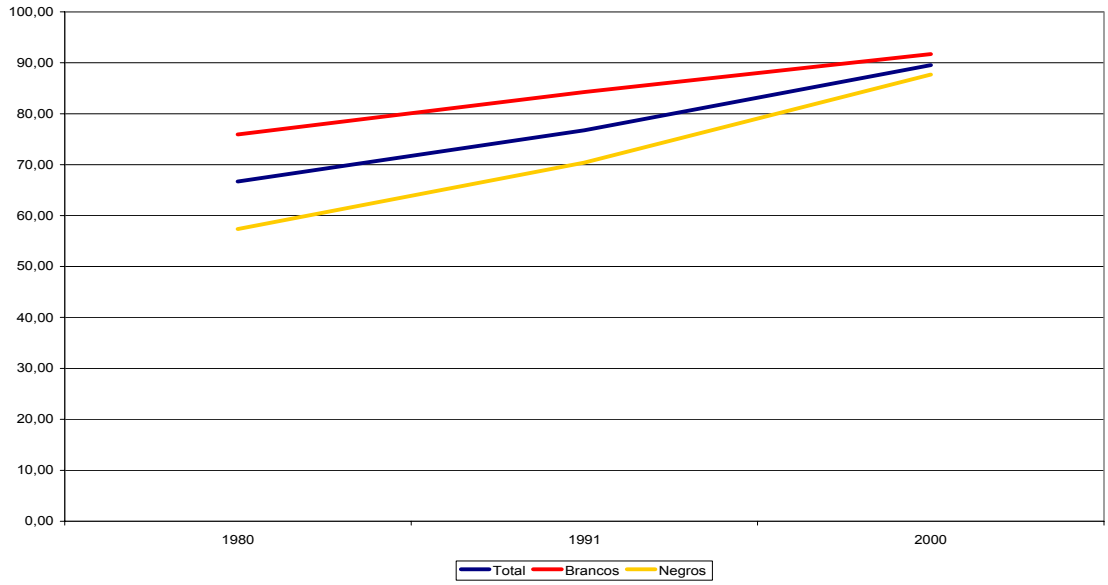
GRÁFICO 1. 19: Taxa de escolarização líquida do ensino médio – Brasil e estados, 2000



Fonte: Censo Demográfico de 2000.

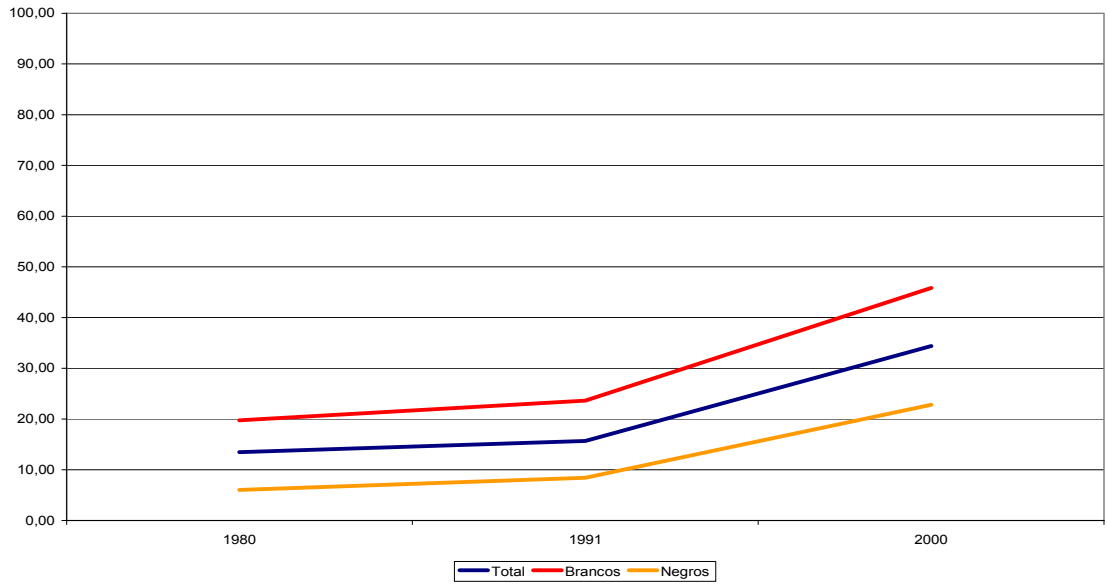
Quando se analisa o diferencial por cor (GRAFs. 1.20 e 1.21) verifica-se que para o ensino fundamental a desigualdade entre as raças apresentou diminuição entre os período analisado, já no nível médio ocorre um pequeno aumento da diferença entre brancos e negros.

GRÁFICO 1. 20: Taxa de Escolarização Líquida do Ensino Fundamental por cor – Brasil, 1980 a 2000



Fonte: Censos Demográficos de 1980, 1991 e 2000.

GRÁFICO 1. 21: Taxa de Escolarização Líquida do Ensino Médio por cor – Brasil, 1980 a 2000



Fonte: Censos Demográficos de 1980, 1991 e 2000.

1.5) Taxa de Distorção idade/série

A taxa de distorção idade/série corresponde ao total de matrículas de pessoas que estão cursando determinada série em idade superior à considerada ideal sobre o total de matrículas na série em questão. Esse índice é importante por determinar problemas relacionados, principalmente, com a alta repetência em determinada série, que é um dos principais problemas no sistema de ensino brasileiro, com graves conseqüências para os níveis de escolaridade da população.

A expressão que calcula esse indicador é dada pela fórmula a seguir:

$$TDIS_{js} = \frac{MAT_{jsi_sup}}{MAT_{js}} \times 100 \qquad (1.5)$$

Sendo:

$TDIS_{js}$ é taxa de distorção idade/série da série s do nível de ensino j;

MAT_{jsi_sup} é o número de matrículas de pessoas com idade superior à idade adequada de estar cursando uma determinada série s do nível de ensino j; e

MAT_{js} é o número total de matrículas na série s do nível de ensino j.

Pode-se calcular também o complemento da taxa de distorção idade/série, que é a taxa de eficiência. Esse indicador nada mais é que um menos a taxa de distorção

idade/série, ou seja, ele representa a proporção de matrículas de pessoas que estão em idade adequada de cursar determinada série sobre a matrícula total da série.

A metodologia de cálculo desta taxa utilizando dados do IBGE é diferente daquela proposta pelo INEP, baseada nos dados do Censo Escolar. Os Censos Demográficos e PNADs trazem informações sobre a idade e a data de nascimento da pessoa. Além disso, a data de referência dessas bases é no segundo semestre do ano. Esses dois aspectos permitem determinar com clareza quais estudantes têm idade adequada para estar cursando uma determinada série e quais não têm. Por outro lado, o Censo Escolar não conta com a informação sobre a idade do indivíduo, nem sobre o mês de nascimento do estudante. A informação presente nessa base é somente o ano de nascimento da pessoa. Além disso, a data de referência da base é maio, o que impossibilita o uso da mesma metodologia aplicada na base do IBGE. O indicador distorção idade/série, quando calculado pelo Censo Escolar, inclui um maior grupo de estudantes com idade adequada.

A taxa de distorção idade/série do ensino fundamental apresentou queda entre os anos investigados, passando de 77,97% em 1980 para 54% em 2000. Apesar dessa queda, o quadro ainda é preocupante, pois mais da metade dos alunos do ensino fundamental estão fora da idade adequada. Com relação ao ensino médio, o panorama é ainda pior, pois em 2000 havia 66,36% de alunos matriculados nesse nível de ensino que se encontravam fora da idade adequada.

No que toca a questão das disparidades regionais, percebe-se através dos GRAFs. 1.22 e 1.23 que nos dois níveis de ensino, ao contrário do que ocorreu nos demais indicadores, as desigualdades regionais estão se acentuando, com o Norte e Nordeste apresentando pior performance.

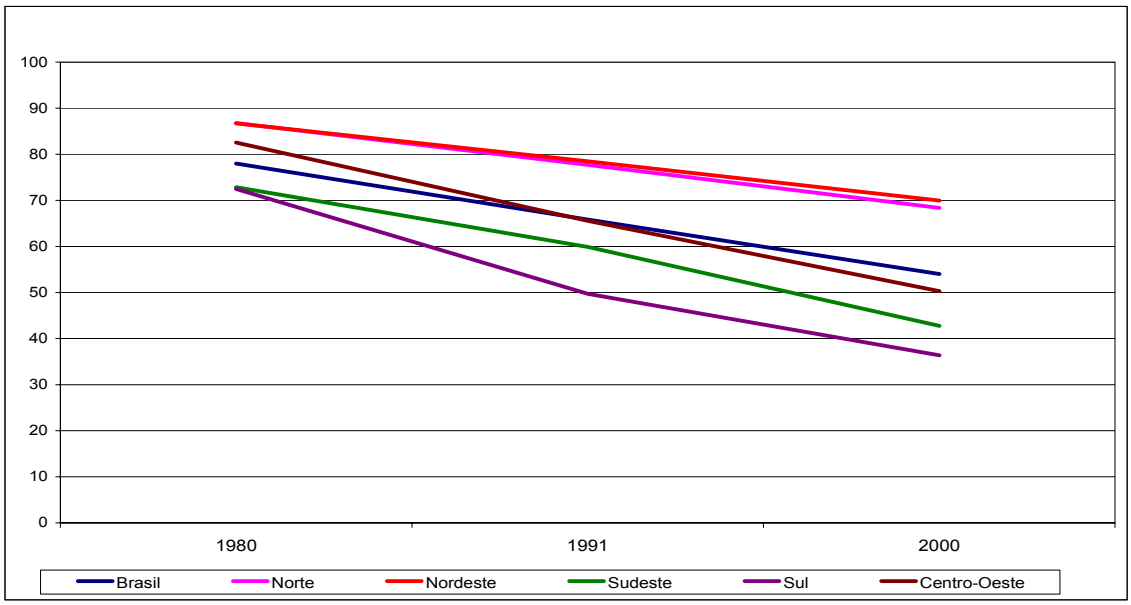
Além das disparidades regionais, percebe-se também as diferenças desta taxa em relação ao sexo, em menor grau, e a cor, mais acentuada. Tanto no ensino fundamental quanto no médio os homens possuem maior distorção entre a idade adequada e a série cursada do que as mulheres, e os negros maior que os brancos.

TABELA 1. 1: Taxa de distorção idade/série por sexo, cor e nível de ensino – Brasil, 2000.

Categoria	Ensino Fundamental	Ensino Médio
Total	54,03	66,36
Mulheres	50,79	64,13
Homens	57,15	69,05
Brancos	42,67	57,05
Negros	64,69	76,77

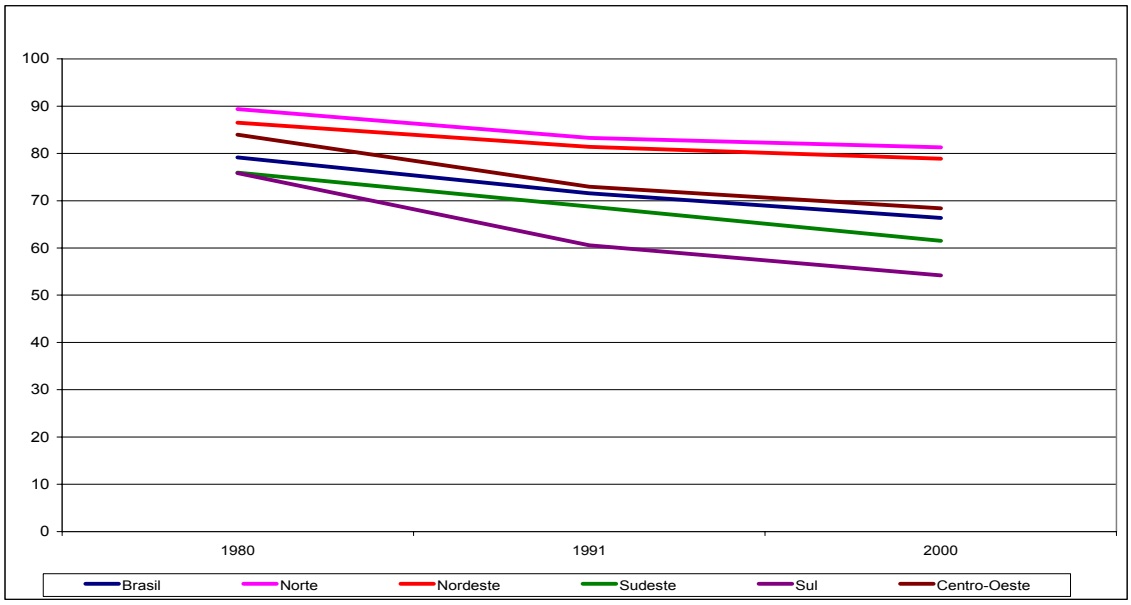
Fonte: Censo Demográfico de 2000.

GRÁFICO 1. 22: Taxa de Distorção idade/série do Ensino Fundamental – Brasil e regiões, 1980 a 2000.



Fonte: Censos Demográficos de 1980, 1991 e 2000.

GRÁFICO 1. 23: Taxa de Distorção idade/série do Ensino Médio – Brasil e regiões, 1980 a 2000.



Fonte: Censos Demográficos de 1980, 1991 e 2000.

1.6) Probabilidade de Progressão Por Série (PPS)

A probabilidade de Progressão por Série foi um método desenvolvido no âmbito do projeto INEP/CEDEPLAR por Eduardo Rios-Neto. Ele foi inspirado no método demográfico da Razão de Progressão por Parturição aplicado para a estimativa da fecundidade, desenvolvido originalmente por Norman Ryder e Louis Henry.

A PPS é útil por permitir analisar as diversas transições do sistema de ensino, o padrão de evasão definitiva por série escolar, gerar o cálculo dos anos médios de estudo além de permitir o desenho de cenários para a projeção dos anos médios de estudo.

Segundo Rios-Neto, esse método é mais adequado quando aplicado a coortes, e para o caso de se desejar estimar os anos completos de estudo concluído, ele deve ser aplicado à coortes mais velhas, ou seja, que por hipótese já deveriam ter concluído seus estudos. Porém, esse pressuposto pode ser relaxado, principalmente no caso da analise das transições entre as séries dos níveis de ensino inferiores. O seu cálculo se dá da seguinte forma:

$$e_k = \frac{P_{k+1}}{P_k} \qquad (1.6)$$

Sendo:

e_k = probabilidade de progressão da série k para a série k+1;

P_{k+1} = pessoas na coorte que concluíram pelo menos a série k+1;

P_k = pessoas na coorte que concluíram pelo menos a série k.

Pode-se dizer, portanto, que e_0 é a probabilidade de progressão de zero para um ano de estudo, e e_{k-1} é a probabilidade de conclusão de k-1 anos de estudo para k anos de estudo, entre os que completaram k-1 anos de estudo.

Com o conjunto de probabilidades de progressão por série pode-se chegar a proporção da coorte com pelo menos *k* anos de estudo. Essa proporção é o produto das PPS até k – 1 anos de estudo, denominada por $e_{0,k}$. Dessa forma tem-se que:

$$e_{0,k} = \prod_{j=0}^{k-1} ej \qquad (1.7)$$

Sendo:

$e_{0,k}$ é a proporção que possui pelo menos *k* anos de estudo.

A proporção com pelo menos *k* anos de estudo permite calcular o número médio de anos de estudo da coorte que é o somatório do conjunto da população com pelo menos *k* anos de estudo.

Os gráficos e tabela abaixo foram retirados do artigo de Rios-Neto (2004) que serviram de base para a análise das PPSs do ensino fundamental e médio. Eles foram construídos com base nos dados do Censo Demográfico de 2000 para os grupos etários qüinqüenais considerando a população de 20 a 64 anos.

TABELA 1. 2: Probabilidade de progressão por série por grupo etários – Brasil, 2000.

PPS	Grupos de Idade								
	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64
	Ano de entrada no ensino sistema de ensino*								
	1983-87	1978-82	1973-77	1968-72	1963-67	1958-62	1953-57	1948-52	1943-47
e0	0,9516	0,9402	0,9249	0,9164	0,8997	0,8680	0,8197	0,7635	0,7202
e1	0,9712	0,9677	0,9604	0,9567	0,9517	0,9411	0,9266	0,9116	0,9000
e2	0,9604	0,9553	0,9498	0,9433	0,9371	0,9221	0,9033	0,8838	0,8711
e3	0,9410	0,9325	0,9261	0,9212	0,9138	0,8959	0,8722	0,8462	0,8244
e4	0,8845	0,8497	0,8243	0,7923	0,7302	0,6778	0,6139	0,5541	0,4985
e5	0,8886	0,8728	0,8759	0,8855	0,8994	0,8978	0,8990	0,8996	0,9014
e6	0,9127	0,9064	0,9085	0,9195	0,9349	0,9434	0,9461	0,9461	0,9468
e7	0,8959	0,8953	0,8956	0,9050	0,9090	0,9256	0,9345	0,9379	0,9375
e8	0,8078	0,7823	0,7586	0,7625	0,7554	0,7521	0,7478	0,7265	0,7016
e9	0,8783	0,9113	0,9229	0,9349	0,9471	0,9583	0,9657	0,9702	0,9709
e10	0,8425	0,8944	0,9066	0,9158	0,9229	0,9335	0,9426	0,9466	0,9468
e11	0,2701	0,3168	0,3458	0,3840	0,4244	0,4669	0,4858	0,4738	0,4461
e12	0,7184	0,8720	0,8992	0,9126	0,9271	0,9429	0,9573	0,9657	0,9698
e13	0,6263	0,8465	0,8792	0,8928	0,9031	0,9175	0,9366	0,9482	0,9507
e14	0,5545	0,8041	0,8376	0,8472	0,8598	0,8728	0,8919	0,8988	0,8964
e15	0,3288	0,4415	0,4173	0,4076	0,4065	0,4073	0,4190	0,4355	0,4409
e16	0,2116	0,2676	0,2940	0,2998	0,2952	0,3124	0,3142	0,2922	0,2704

Fonte: Rios-Neto, E., 2004.

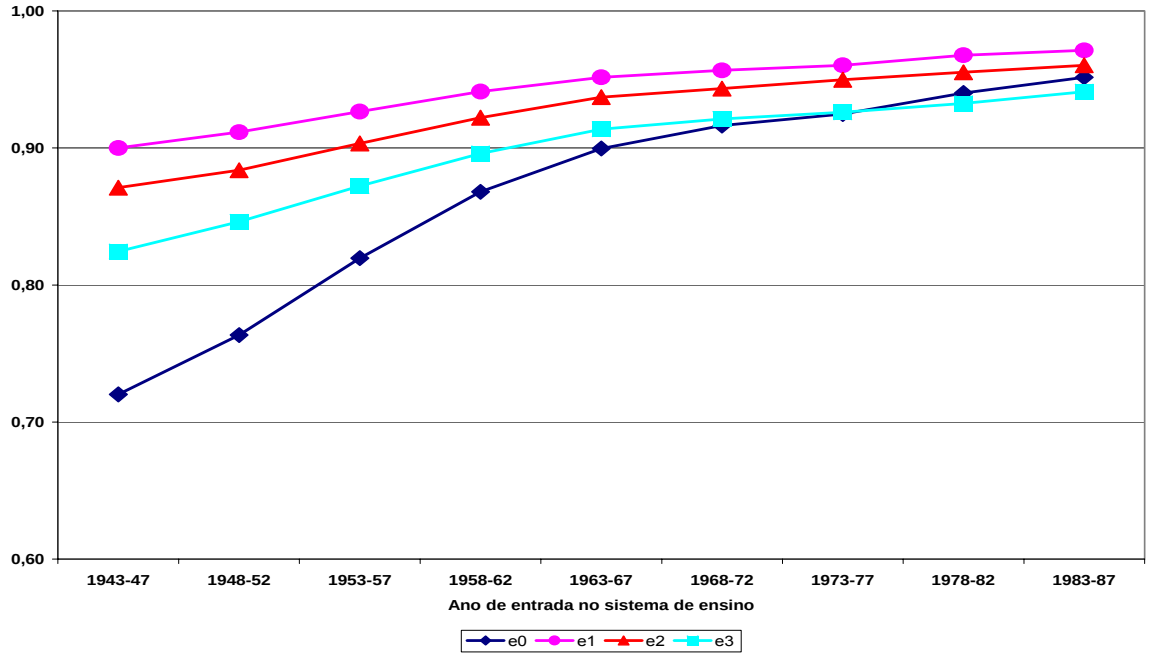
* Considera-se como ano de entrada no sistema de ensino o período em que cada coorte tinha 7 anos de idade, pois esta seria a idade adequada de se ingressar no sistema escolar.

Percebe-se no GRAF. 1.24 que para as coortes mais velhas a probabilidade de passar de zero para um ano de estudo era o grande gargalo das primeiras quatro séries do ensino fundamental. Essa progressão sofreu uma significativa melhora, alcançando patamares semelhantes às três restantes. Essa melhora é resultado tanto da maior cobertura educacional quanto da diminuição da repetência nessa série.

Com relação à progressão das quatro últimas séries do ensino fundamental (GRAF. 1.25), a maior retenção é na probabilidade de passar de quatro para cinco anos de estudo, mas vale ressaltar que ela também apresentou significativa melhora durante o período, e o inverso ocorrendo para as demais progressões, que apresentaram pequena queda.

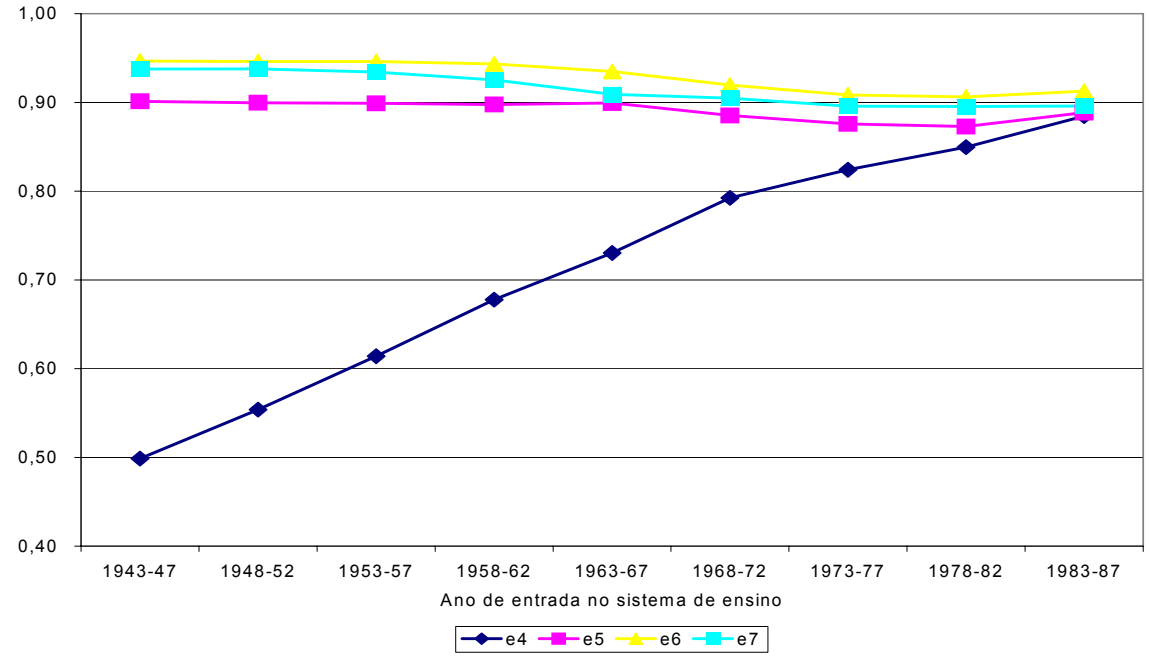
As transições do ensino médio, GRAF. 1.26, mostram uma menor probabilidade de passar de oito para nove anos de estudo, ressalta-se que ela vem apresentando crescimento enquanto as demais apresentam tendência de queda.

GRÁFICO 1. 24: Taxa de progressão nas quatro primeiras séries do ensino fundamental – Brasil, 2000.



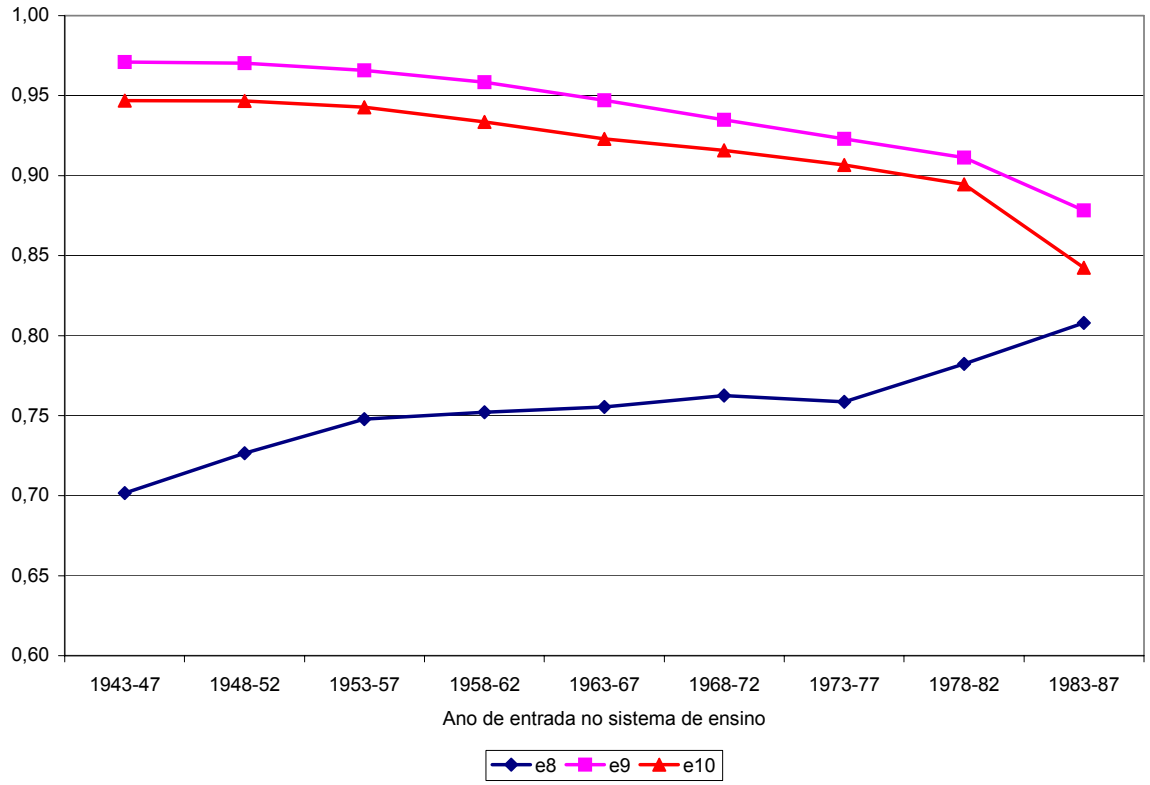
Fonte: Rios-Neto, E., 2004.

GRÁFICO 1. 25: Taxa de progressão nas quatro últimas séries do ensino fundamental – Brasil, 2000.



Fonte: Rios-Neto, E., 2004.

GRÁFICO 1. 26: Taxa de progressão do ensino médio – Brasil, 2000.



Fonte: Rios-Neto, E., 2004.

1.7) Anos Médios de Estudo ou Escolaridade Média

A escolaridade média corresponde à média dos anos de estudos concluídos por uma determinada população. Vale ressaltar que não se consideram os anos que a pessoa passou na escola, mas apenas aqueles em que ela teve aprovação. A formula abaixo exemplifica como são feitos os cálculos.

$$\text{Anos de estudos} = \frac{\left(0 \times P_0 + 1 \times P_1 + \ldots + 17 \times P_{17}\right)}{\text{P}} \qquad (1.8)$$

Sendo:

P_0 é a população com zero ano de estudo;

P_1 é a população com um ano de estudo;

P_{17} é a população com 17 anos de estudo; e

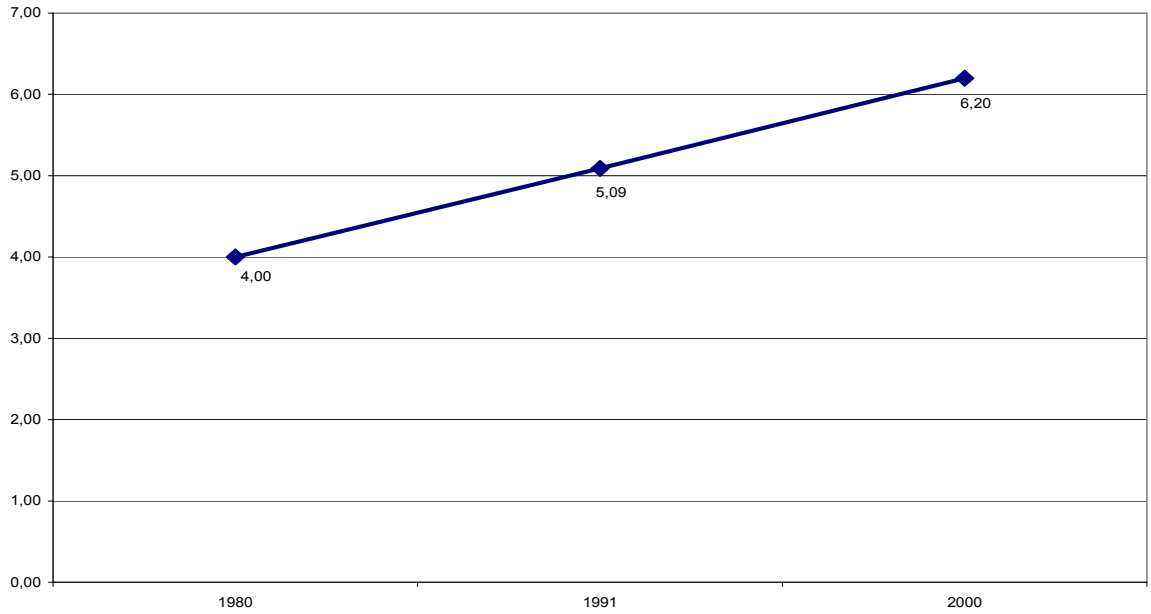
P é a população total.

As pessoas com zero ano de estudos são aqueles que nunca estudaram e os que, se freqüentaram a escola, não conseguiram ser aprovados nem na 1ª série do Ensino Fundamental. Os indivíduos com um ano de estudo são aqueles que já foram aprovados na 1ª série do Ensino Fundamental, mas ainda não foram aprovados na série seguinte, e assim sucessivamente.

A escolaridade média é um importante indicador educacional porque na sua estimativa estão embutidas as taxas de rendimento escolar – aprovação, reprovação e evasão –, bem como o grau de atendimento do sistema de ensino. Portanto, altos níveis de atendimento escolar e taxas de aprovação tendem a elevar a escolaridade média, uma vez que há maior número de pessoas dentro da escola e que estão progredindo para séries mais avançadas. Por outro lado, taxas de evasão e reprovação maiores tendem a diminuir a escolaridade média. Desta forma, esta é uma boa medida síntese das taxas de rendimento escolar e do nível de atendimento do sistema de ensino.

Através da análise da evolução dos anos médios de estudo, verifica-se que o Brasil apresenta um aumento de mais de 2 anos de estudo, conseqüência principalmente da maior cobertura e fluxo entre as séries do ensino fundamental, conforme visto acima, alcançando em 2000 uma escolaridade média de 6,20 anos de estudo. Apesar da melhora esse ainda é um patamar bastante baixo se considerarmos que a população de 15 anos ou mais não possui, em média, nem o ensino fundamental completo.

GRÁFICO 1. 27: Anos médios de estudo da população de 15 ou mais anos de idade – Brasil, 1980 a 2000.



Fonte: Censos Demográficos de 1980, 1991 e 2000.

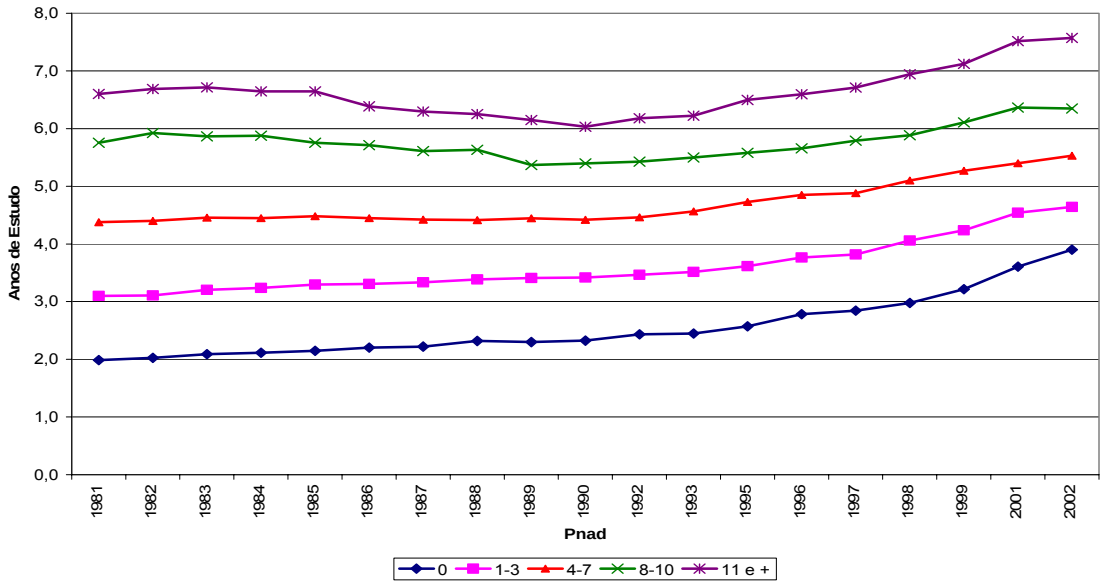
Usando dados da Pesquisa Nacional de Amostra por Domicílio (PNAD) de 1981 a 2002, analisou-se a escolaridade média da população de 7 a 25 anos segundo grupo de anos de estudo e cor do chefe da família e quintil de renda familiar *per capita*, conforme gráficos abaixo. Com esse exame é possível verificar a desigualdade de oportunidade educacional entre os diferentes grupos da população.

Crianças que pertencem a famílias cujo chefe tem no mínimo o ensino médio completo (11 anos e mais de estudo) possuem escolaridade média bem maior que crianças cujo chefe do domicílio tem zero ano de estudo. Tal fato demonstra a existência de um ciclo vicioso, em que pais com baixa escolaridade transmitem pouca escolaridade aos seus filhos e assim sucessivamente. Cabe ressaltar que a escolaridade média dos três menores grupos de anos de estudo do chefe da família – as três últimas curvas do GRAF. 1.28 – apresentou um crescimento maior do que a dos outros dois grupos. O que indica que a diferença entre eles diminuiu nos últimos anos, ou seja, parece estar havendo uma homogeneização, mesmo que ainda fraca, nos níveis escolares de diferentes camadas sociais brasileiras. Por outro lado, o exame dos anos médios de estudo por quintil de renda familiar *per capita*, GRAF. 1.29, mostra que a diferença da escolaridade entre os extratos de renda praticamente não diminuiu, com

famílias mais ricas (pertencentes ao quinto quintil) possuindo quase que o dobro de escolaridade das famílias mais pobres (pertencentes ao primeiro quintil).

Quando se analisa a escolaridade média por cor do chefe da família (GRAF. 1.30), nota-se que crianças cujas famílias são chefiadas por brancos ou amarelos possuem maior escolaridade do que aquelas nascidas em famílias com chefes pretos ou pardos[2]. Esta diferença, como já mencionado, pode estar refletindo as características socioeconômicas das famílias, já que chefes pretos e pardos são os que possuem menor nível educacional e renda.

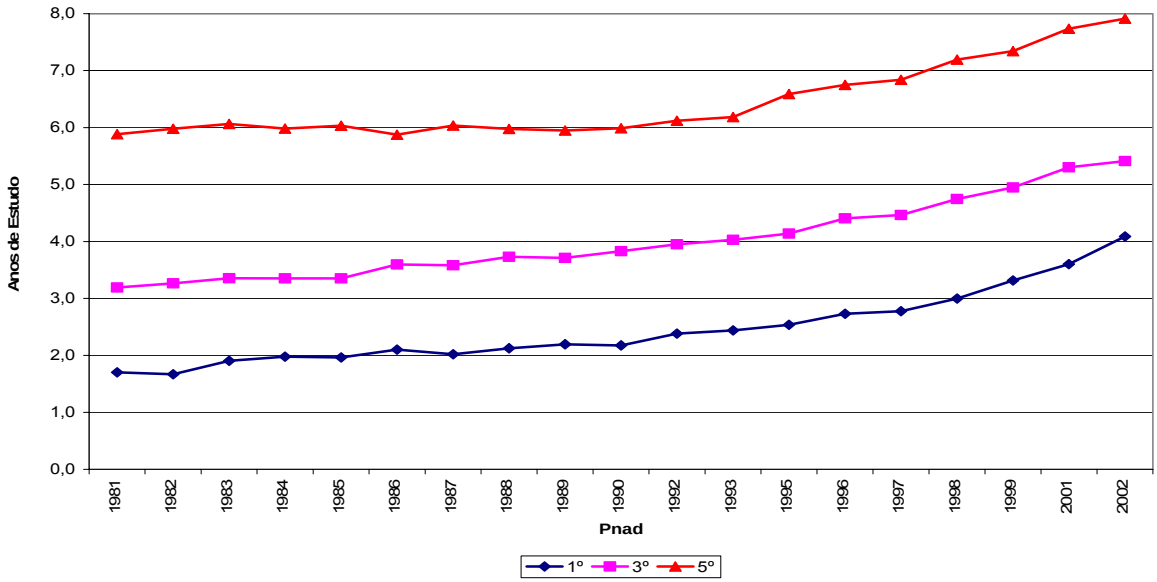
GRÁFICO 1. 28: Anos médios de estudo da população de 7 a 25 anos de idade segundo educação do chefe da família – Brasil, 1980 a 2002.



Fonte: Pesquisa Nacional de Amostra por Domicílio

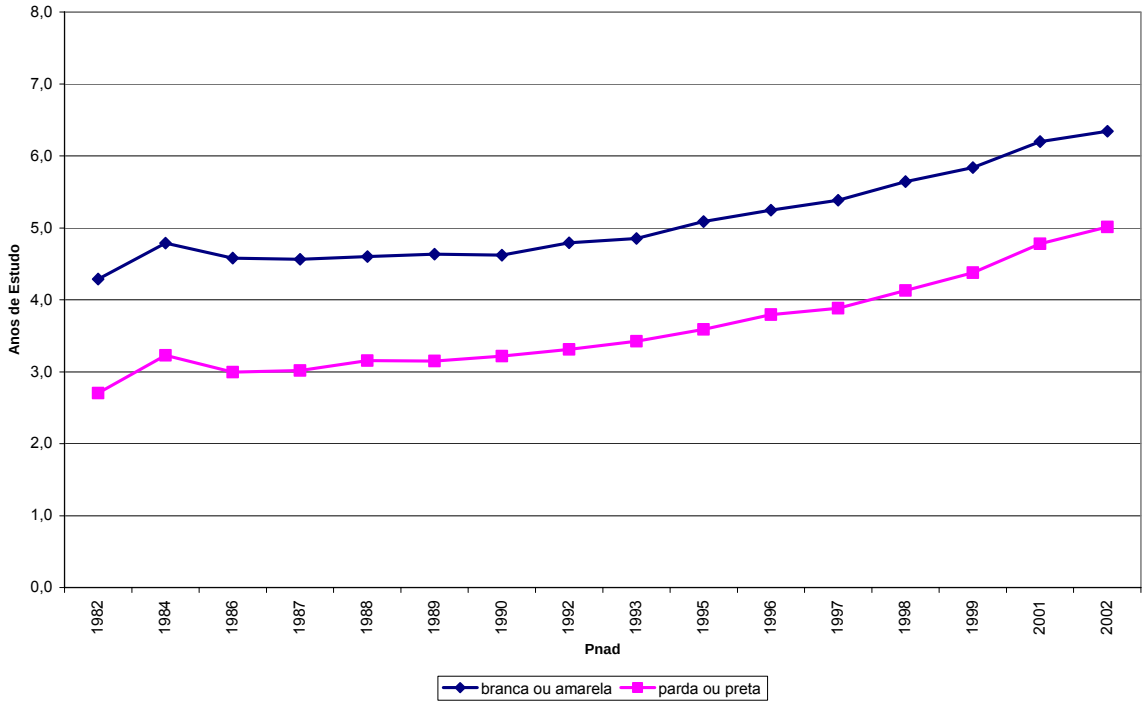
[2] Nos anos de 1981, 1983 e 1985 a PNAD não incluiu o quesito sobre cor ou raça, impossibilitando a desagregação da escolaridade média nesses anos.

GRÁFICO 1. 29: Anos médios de estudo da população de 7 a 25 anos de idade segundo quintil de renda familiar *per capita* – Brasil, 1981 a 2002.



Fonte: Pesquisa Nacional de Amostra por Domicílio

GRÁFICO 1. 30: Anos médios de estudo da população de 7 a 25 anos de idade segundo a cor do chefe da família – Brasil, 1981 a 2002.



Fonte: Pesquisa Nacional de Amostra por Domicílio

1.8) Conclusão

Nesse capítulo foi feita uma análise descritiva dos principais indicadores educacionais no Brasil, analisando sua evolução entre as últimas décadas, bem como as diferenças regionais, por sexo, cor, situação de residência e outras características sócio-econômicas.

Com base nessa descrição, observou-se que o Brasil apresentou melhora em quase todos os indicadores analisados entre os anos de 1980 a 2000. Pode-se dizer que no caso do ensino fundamental os problemas relacionados com a cobertura estão praticamente solucionados, com a maioria das pessoas de 7 a 14 anos freqüentando a escola, inclusive com uma diminuição no que se refere às desigualdades regionais, de cor e situação de domicílio.

No caso do ensino médio, apesar de ter ocorrido uma melhora, os níveis de cobertura ainda se apresentam bastante baixos, e com aumento nas diferenças regionais e por outras características que ainda persistem.

Esse quadro de diminuição das desigualdades no ensino fundamental e aumento no médio é esperado, já que o primeiro está caminhando para a universalização e a tendência é de expansão para as camadas mais desprovidas de atendimento educacional. Já no ensino médio, como a cobertura é bem baixa, sua expansão ocorre primeiro para as classes com melhores condições sócio-econômicas.

Também ficou bastante evidenciado que a distorção idade/série, conseqüência seja de uma alta repetência ou entrada tardia no sistema de ensino, é um dos problemas mais graves do sistema educacional brasileiro, sendo um dos principais enclaves para o aumento nos anos médios de estudo.

2) MARCO TEÓRICO E PRINCIPAIS ESTUDOS SOBRE OS DETERMINANTES DA EDUCAÇÃO

Como já mencionado, o objetivo dessa tese é investigar os principais determinantes do resultado educacional das crianças e jovens. Para tal fim, pretende-se nesse capítulo percorrer uma das principais abordagens do estudo dos determinantes escolares, que é modelada com base na "Função de Produção Educacional". Tal abordagem servirá como ponto de partida, porém, devido a indisponibilidade de alguns dados, foi necessário adaptá-la, conforme será visto no Capítulo 3. Em seguida, são revistos alguns estudos que analisam os determinantes da educação, principalmente para o caso brasileiro.

2.1) Função de Produção Educacional

Quando se pretende analisar os fatores que afetam a educação, objetiva-se analisar como os diversos insumos do processo educacional podem afetar os resultados educacionais de um indivíduo. A relação entre os insumos e o resultado educacional, seja esse desempenho acadêmico, anos de estudo ou qualquer outra variável de resultado educacional, é dada pela Função de Produção Educacional. A FPE é um ramo da literatura que faz uma analogia entre o processo de aquisição do conhecimento humano e o processo produtivo, o que guia a escolha das variáveis e a interpretação coerente dos seus efeitos (TODD & WOLPIN, 2003). A Função de Produção Educacional pode ser especificada da seguinte forma[3]:

$$A_t = f(F_t, S_t, O_t, \varepsilon_t) \qquad (2.1)$$

Sendo:

A_t é o resultado educacional;

F_t é o vetor das características e *background* da família do estudante acumulativos no tempo t;

S_t é o vetor de insumos escolares e dos professores acumulativo no tempo t;

O_t é o vetor de outros insumos relevantes, tal como fatores comunitários e da turma, acumulativos no tempo t; e

[3] Para maiores detalhes da Função de Produção Educacional ver: HANUSHEK, E. A. (2002), HANUSHEK, GOMES-NETO & HARBISON (1996), LEE & BARRO (1992), LAZEAR, E. P. (1999), TODD, P. E. & WOLPIN, K. (2003) e KRUEGER, A. B. (1999).

ε_t é o termo de erro aleatório que reflete os fatores não mensurados que contribuem para o resultado educacional no tempo t.

Segundo HANUSHEK (2002) uma das questões chaves da Função de Produção Educacional é como medir o resultado educacional (A_t). A principal justificativa para os estudos educacionais é seu efeito no mercado de trabalho, ou seja, no rendimento do indivíduo. O ideal, portanto, seria analisar a associação dos insumos escolares e o rendimento subseqüente do indivíduo. Entretanto, tal analise é inviabilizada pela dificuldade de obtenção de dados longitudinais. Dessa forma, a maioria dos estudos tem contextualizado o problema em dois estágios. Primeiramente, é feita uma associação entre os insumos escolares e os resultados educacionais, mensurados pelos testes de desempenho, anos médios de estudo ou outros resultados intermediários como taxa de repetência ou evasão. E, num segundo momento, esses resultados educacionais são relacionados com o sucesso no mercado de trabalho. Apesar de importante analisar esses dois estágios, nessa tese, só será investigado o primeiro, ou seja, como os insumos escolares afetam o resultado educacional.

Com relação aos insumos familiares, F_t, a maioria dos estudos concorda com o papel importante das variáveis de *background* familiar na educação. Quando elas não são incorporados, os coeficientes dos outros insumos, principalmente os escolares, podem ser enviesados, devido à correlação entre eles (HANUSHEK, 2002 e KRUEGER, 1999).

Apesar da importância dos insumos familiares, principalmente para a melhor especificação da Função de Produção Educacional, Hanushek argumenta que mudanças nas características da família pode ser de difícil manipulação por parte de políticas do Estado, tanto por serem pouco sensíveis no curto prazo, quanto por serem caras para mudanças de curto prazo. Por exemplo, é sabido que a educação da mãe é importante determinante do resultado educacional da criança, mas a melhor política de curto prazo não é, provavelmente, enviar toda a mãe de volta para a escola. Por outro lado, para políticas de longo prazo é importante considerar o impacto dos insumos familiares, principalmente nos países em desenvolvimento. Para auxiliar na direção de políticas de longo prazo é importante conhecer a natureza exata e o impacto causal dos fatores familiares. O problema tem sido a falta de uma convincente identificação estratégica que distingue o verdadeiro aspecto estrutural da família da variedade de *proxy* correlacionadas.

Para o caso dos insumos escolares, a questão de quais fatores considerar é difícil e importante, principalmente porque eles são freqüentemente o objeto de políticas governamentais para melhoria do resultado educacional. As medidas mais freqüentemente utilizadas nos diversos estudos são: i) recursos reais das salas de aula, que são medidos através do nível educacional e experiência do professor e tamanho de classe ou taxa professor/aluno; ii) recursos financeiros agregados, que são gasto por aluno e salário dos professores; iii) medidas de outros recursos da escola, como característica especifica do professor, insumos administrativos (salário ou qualidade dos administradores) e instalações, tais como: laboratórios e tamanho e presença de bibliotecas (HANUSHEK, 2002).

Com relação aos outros insumos relevantes que fazem parte da Função de Produção Educacional, O_t, pode-se dividi-los em fatores comunitários e da turma. É importante considerar os fatores comunitários porque os estados possuem políticas diferentes, e os gastos educacionais, leis trabalhistas e outras regulamentações variam entre eles. Dessa forma, análises sobre a educação que considerem vários estados ou distritos devem incluir alguma medida da política educacional que diferencie a viabilidade ou custos das escolas ou o custo de oportunidade dos diferentes mercados de trabalho locais para não tornar os parâmetros enviesados. A direção do viés dependerá da correlação da variável omitida com as que foram incluídas no modelo.

As características da turma, tal como o quão preparado os alunos da turma estão para o currículo escolar ou sua motivação e perspectiva geral, podem influenciar o desempenho educacional do indivíduo, porém o principal problema é com relação à mensuração dessas características. A dificuldade se dá principalmente em separar o efeito da turma do efeito individual. Segundo Hanushek, estimativas da Função de Produção Educacional que consideram o efeito da turma podem sobreestimar seu efeito, devido a fatores individuais mal mensurados ou omitidos que são comuns a turma.

O termo de erro aleatório da equação, $\varepsilon_{t,}$ merece especial atenção. Ele reflete os fatores que não foram incluídos na análise, mas que afetam o resultado educacional do indivíduo. Se os fatores não incluídos não são correlacionados com os que entraram na análise, não há problema, pois os parâmetros estimados não serão enviesados. Por outro lado, se eles são correlacionados os coeficientes da análise serão enviesados.

HANUSHEK, GOMES-NETO & HARBISON (1996) colocam que o risco dos parâmetros estimados pela Função de Produção Educacional serem enviesados é substancial. Primeiro, porque o processo educacional é acumulativo, dessa forma a situação passada da escola tende a ser correlacionada com as características familiares e os insumos escolares contemporâneos. Segundo, porque alguns fatores são difíceis de mensurar, tal como habilidade natural, motivação e aspiração do estudante, que tendem a ser correlacionados com as características familiares observadas.

Na maioria das vezes esses problemas são difíceis de serem contornados devido à limitação dos dados disponíveis. Eles estão disponíveis apenas para um ponto do tempo, não acompanhando o aluno durante a sua permanência na escola. Somado a isso, muitas pesquisas possuem dados limitados com relação à qualidade de alguns insumos escolares, como por exemplo, alguma ação qualitativa diferente do professor, além da impossibilidade de mensurar a capacidade genética. Tais dificuldades podem levar a problemas de especificação do modelo como foi visto acima.

2.2) Alguns Estudos Sobre os Determinantes da Educação[4]

No Brasil, a maioria dos estudos que tenta entender os determinantes do desempenho dos indicadores educacionais enfoca principalmente aqueles ligados à família do indivíduo, ou seja, o termo Ft da equação (2.1).

Dentro deste contexto, para SILVA E HASENBALG (2001) existem três dimensões importantes com relação à estrutura familiar que afeta o desempenho do aluno na escola. A primeira seria relacionada aos recursos econômicos disponíveis para os gastos educacionais dos filhos, denominada de capital econômico. A segunda dimensão seria relacionada aos recursos educacionais ou capital cultural da família, que pode propiciar um ambiente mais adequado ao aprendizado. A última dimensão de estrutura familiar que afeta o desempenho do indivíduo no sistema de ensino seria em relação à estrutura dos arranjos familiares, que podem facilitar ou prejudicar a ação dos indivíduos dentro da estrutura social.

[4] O uso da terminologia "determinantes educacionais" é controverso por estabelecer um aspecto causal dos fatores associados ao desempenho escolar. Porém, como SOARES (2002) coloca "para fazer sugestões políticas é necessário a interpretação causal de pelo menos algumas das diversas correlações entre as variáveis educacionais, sociais e familiares e o desempenho educacional". Dessa maneira, como este trabalho pretende de alguma forma colaborar com políticas educacionais, esse termo se torna pertinente.

SILVA E HASENBALG (2001) estudam o efeito destas dimensões nas transições escolares entre séries para o Brasil com dados da PNAD de 1999, utilizando a renda *per capita* para dimensionar o capital econômico, os anos de estudo do chefe da família para determinar o capital cultural, e, para apurar o capital social familiar, os autores utilizam a chefia feminina, o número de filhos no domicílio e a extensão familiar. Como esperado, seus resultados mostram uma relação positiva para as variáveis de capital econômico e cultural e negativa para as variáveis que medem o capital social familiar. O padrão do efeito das variáveis explicativas é crescente até a 4ª série do ensino fundamental, quando começa a declinar, mostrando uma maior influência das variáveis de *background* familiar até o meio do ensino fundamental.

No intuito de estudar a diferença educacional entre São Paulo e a região Nordeste, BARROS E LAM (1996) utilizaram a educação dos pais e a renda do chefe do domicílio como determinantes da demanda intradomiciliar por educação, utilizando dados da PNAD para 1982. Seus resultados mostraram um importante papel da educação dos pais, principalmente das mães, na determinação dos anos médios de estudo de crianças de 14 anos de idade, sendo que o efeito observado foi maior para as crianças nordestinas.

Alguns estudos, como o de MARTELETO (2001), dão um importante enfoque ao número de irmãos presentes no domicílio. A hipótese é de que um maior número de irmãos, principalmente em idade escolar, implicaria em uma diluição dos recursos financeiros e recursos educacionais destinados aos filhos (diminuição do tempo, energia e atenção que os pais oferecem a cada filho). Todos os estudos demonstraram uma relação inversa entre número de irmãos no domicílio e o desempenho escolar, tanto nos anos médios de estudo quanto na taxa de matrícula. O efeito persistiu mesmo quando se controlou para variáveis socioeconômicas. Porém, a variável número de irmãos no domicílio pode estar sendo sobreestimada nos modelos, uma vez que pode haver um viés de simultaneidade entre ela e a escolaridade, ou seja, "pais podem escolher a combinação de baixa fecundidade e alta escolaridade dos seus filhos. Isso só seria verdade, entretanto, quando se acredita que os pais são racionais em suas decisões de fecundidade e expectativa educacional para seus filhos, mesmo em países em desenvolvimento" (MARTELETO, 2001, pp13).

Os estudos mencionados até aqui concentram suas análises nos fatores familiares e atributos individuais como determinantes do resultado escolar. Fatores esses que se comprovaram ter impactos importantes na educação dos indivíduos no Brasil.

Por outro lado, do ponto de vista dos fatores comunitários e da qualidade e disponibilidade dos serviços educacionais, os termos S_t e O_t da equação da Função de Produção Educacional, no Brasil, só recentemente eles têm sido analisados.

Um dos primeiros estudos que consideram variáveis escolares na determinação do desempenho escolar é o de HANUSHEK, GOMES-NETO & HARBISON (1996), que utilizou a base de dados longitudinal do projeto EDURURAL, realizado no Nordeste do Brasil entre os anos de 1981 e 1985. Essa base permitiu que eles estudassem diversos componentes escolares, que foram divididos em três grupos, quais sejam: i) abastecimento de água e energia elétrica, instalações sanitárias, mobiliário para estudantes e professores, denominados de "hardware"; ii) livros textos, guias para professores, recursos audiovisuais, *notebooks* e matérias de escritório, denominados de "software"; e, iii) variáveis relacionadas à qualidade dos professores, como escolaridade, experiência e salário dos professores, taxa aluno/professor, resultado dos testes de desempenho aplicados nos professores e se o professor participou de programas específicos de treinamento.

Seus resultados demonstram que para o Nordeste rural do Brasil, de uma maneira geral, a qualidade física da escola e recursos para as aulas, os componentes de "hardware" e "software", são positivamente relacionados com o desempenho do estudante. Por outro lado, as variáveis relacionadas com os professores não se mostraram importantes para o desempenho dos estudantes.

BARROS *et alii* (2001) também incorporou variáveis de caráter comunitário e de disponibilidade e qualidade dos serviços educacionais na análise dos determinantes da escolaridade, medida pelo número de séries completadas pelo indivíduo. Eles consideraram cinco fatores determinantes do desempenho escolar, quais sejam: recursos familiares, características pessoais e geográficas, qualidade e disponibilidade da oferta do ensino, ambiente comunitários e custo de oportunidade do tempo.

Para os recursos familiares, os autores utilizaram as variáveis clássicas, renda *per capita* da família e o nível de escolaridade do pai e da mãe separadamente. Segundo os autores a associação do desempenho educacional com a escolaridade dos pais tem

sido mais estreita do que com renda familiar *per capita*. Tal associação pode ser atribuída ao fato que o processo de escolaridade é acumulativo ao longo do tempo, sendo mais influenciada por variações na renda permanente do que na corrente. Como a escolaridade dos pais reflete a renda permanente e a renda *per capita* a corrente, o primeiro terá maior influência na escolaridade dos filhos.

Os autores colocam ainda mais dois motivos que podem justificar a importância da educação dos pais. Primeiro, o de que a escolaridade dos pais também pode ter um importante impacto na redução do custo da educação dos filhos, devido à diminuição da dificuldade e custo de aprendizagem. O segundo motivo está relacionados ao fato de que para muitos pais a educação não é apenas um bem de investimento, que eleva a produtividade e renda futura, mas um bem de consumo, que tem um fim em si mesmo. Dessa forma, pais mais escolarizados vêem maior utilidade na escolaridade dos filhos.

Com relação aos fatores de qualidade e disponibilidade dos serviços escolares da comunidade, BARROS *et alii* utilizaram a qualidade dos professores e a qualidade da infra-estrutura das escolas para captar o lado da qualidade escolar. O primeiro foi medido em termos da escolaridade média dos professores que reside no município e o segundo foi medido ao nível da comunidade, representada pelo setor censitário. Os autores optaram por utilizar um índice sintético de qualidade da infra-estrutura educacional disponível que inclui a disponibilidade de equipamentos para uso dos alunos (computadores, livros, vídeo e TV), duração da jornada letiva, incidência do curso noturno e distância casa-escola.

Para apurar a disponibilidade, eles empregaram a razão entre o número de escolas que fornecem cursos da 5ª a 8ª séries ou o ensino médio e a população de 11 a 25 anos e a duração média do percurso casa-escola. Esses dois indicadores foram agregados ao nível do município.

Para captar os recursos comunitários foram utilizadas a escolaridade média da população adulta e a renda *per capita* da comunidade. Os autores argumentam que a inclusão destas variáveis é importante, pois adultos com maior nível educacional podem auxiliar melhor as tarefas escolares das crianças, diminuindo o custo da educação e que em comunidades mais ricas, menor nível educacional pode reduzir as chances de emprego e melhores salários.

Por último, outro insumo utilizado pelos autores foi o custo de oportunidade do tempo, que é o tempo utilizado pelos indivíduos na escola e que deixa de ser aproveitado em oportunidades eventuais no mercado de trabalho. Dessa forma, quanto mais atrativo for o mercado de trabalho para o jovem na idade escolar, maior a probabilidade dele largar os estudo para trabalhar. As medidas de custo de oportunidade utilizadas pelos autores envolvem o salário médio recebido pelos trabalhadores com as mesmas características individuais e que vivem na mesma comunidade e a probabilidade que esses trabalhadores estejam empregados.

Os autores utilizam dados da PNAD de 1996 e da Pesquisa de Padrão de Vida (PPV) 1996/97 para explorar suas conjecturas[5]. Seus resultados confirmam que a escolaridade dos pais, em particular a da mãe, é o mais importante fator para determinar o desempenho educacional dos jovens. Um ano a mais de escolaridade dos pais possui maior efeito nos anos médios de estudo dos filhos que três anos a mais de escolaridade dos professores. Os autores ressaltam que apesar do impacto da educação dos pais ser maior que dos professores, do ponto de vista de políticas públicas pode ser mais barato elevar a escolaridade dos professores que dos pais, uma vez que o contingente de professores é menor.

Um resultado particularmente interessante encontrado pelos autores é o de que a qualidade da infra-estrutura da escola possui impacto tão ou mais importante que a escolaridade dos professores. A sua importância é maior na segunda etapa do ensino fundamental, mas perde importância no ensino médio.

Com relação à disponibilidade de serviços educacionais, o resultado para a razão entre o número de escolas e a população em idade escolar foi o esperado, ou seja, positivo e significativo, o inverso ocorrendo para a duração média do percurso casa-escola, cujo coeficiente apresentou sinal positivo, não esperado, contudo, não foi significativo.

Para o restante dos insumos, variáveis comunitárias e custo de oportunidade do tempo, os resultados foram os esperados. Comunidades mais ricas e com maior nível educacional dos seus adultos, afetam positivamente a escolaridade de seus membros, sendo que a educação média dos adultos mostrou-se mais importante. Por outro lado,

[5] Suas análises foram feitas apenas para as regiões Nordeste e Sudeste.

quanto mais atraente o mercado de trabalho local em relação a media nacional, pior o desempenho educacional dos membros dessa comunidade.

O trabalho de BARROS *et alli* (2001) é importante por trazer à discussão dos determinantes de educação, a importância das variáveis de caráter comunitário e de oferta dos serviços educacionais, mostrando que eles afetam a educação. Porém, eles utilizam o método de Mínimos Quadrados Ordinários (MQO), os quais consideram as variáveis agregadas no mesmo nível de análise das variáveis individuais. A estrutura da base de dado utilizada pelos autores, pesquisas domiciliares, possui natureza hierárquica e quando tratados de forma convencional podem acarretar um viés sistemático de subestimação da variável agregada, devido à quebra de pressupostos do modelo MQO. Para contornar esse problema, a literatura tem recorrido aos modelos hierárquicos, que serão melhor especificados no Capítulo 4 dessa tese.

No Brasil, os estudos clássicos que utilizam os modelos hierárquicos são o referente ao estudo da escola eficaz. Alguns trabalhos nessa linha são os de FLETCHER (1998), ALBERNAZ *et alii*, (2002), CÉSAR, C. C. & SOARES, J. F. (2001), FERRÃO, M. E. *et alii* (2001) e BARBOSA e FERNANDES (2001) que, utilizando a terminologia dos modelos hierárquicos, consideram o individuo no primeiro nível de análise e as escolas, no segundo. Esses estudos utilizam a base de dados do Sistema Nacional de Avaliação da Educação Básica (SAEB). Essa base possui informações sobre testes padronizados de desempenho, além de outras informações a respeito do nível socioeconômico do aluno e algumas características da escola, professores e diretores. Cabe ressaltar, que o plano amostral do SAEB é representativo apenas para as Unidades da Federação.

Os estudos da escola eficaz mostram que existe uma fonte de variabilidade que são as escolas, ou seja, alunos com características semelhantes que freqüentam escolas diferentes, têm desempenhos distintos. Eles atribuem grande parte dessa variabilidade ao nível socioeconômico médio de cada escola, mas mesmo controlando por essa variável ainda resta parte dessa variância que pode ser devida a fatores propriamente escolares.

No intuito de explicar quais outros fatores escolares contribuem para a eficácia da escola, ALBERNAZ *et alii* (2002) considerou na análise, além do nível socioeconômico da escola, outros fatores relacionados à infra-estrutura da escola (salas de aula arejadas e nível de ruído das salas de aulas), aos recursos financeiros e as

práticas escolares (incidência de faltas de aula por parte do aluno e nível de escolaridade do professor). Todas esses fatores mostraram-se importante no desempenho do aluno, mostrando que escolas melhores levam a melhores resultados.

Utilizando também os modelos hierárquicos, mas em uma outra linha de análise, onde as variáveis dependentes analisadas são as probabilidades de progressão da 1ª e 5ª séries, RIOS-NETO, CÉSAR E RIANI (2002) investiga o impacto de variáveis de caráter familiar e de oferta do sistema de ensino, utilizando dados das PNADs de 1981 a 1999. Diferentemente dos trabalhos anteriores, o primeiro nível do modelo hierárquico contemplou o indivíduo e o segundo nível, o ano, UF, situação do domicílio e tipo de área (metropolitana e não-metropolitana).

No primeiro nível, as variáveis independentes utilizadas foram de *background* familiar (educação da mãe e variáveis indicadores de ocupação e posição na ocupação do pai). Como esperado, a educação da mãe apresentou grande efeito sobre a probabilidade de progressão por série. Para o segundo nível, utilizou-se como variáveis independentes, os anos médios de estudo e salário médio dos professores do ensino fundamental e razão professor do ensino fundamental e população em idade escolar. Estas variáveis tentaram captar qualidade e disponibilidade da educação ofertada. Os resultados encontrados pelos autores mostram que a escolaridade média do professor é importante para analisar o comportamento individual e possui um importante efeito substituição com a educação materna.

Um resultado importante e inovador desse trabalho é o *trade off* entre o nível médio educacional dos professores e a educação materna, existente principalmente na probabilidade de passar na 1ª série. Em outras palavras, o aumento da média de escolaridade do professor, medida *proxy* da qualidade educacional, tem um impacto positivo maior para filhos de mães com menor nível educacional.

No contexto da literatura internacional, o debate sobre o impacto dos diversos insumos no resultado educacional já está bem desenvolvido, mas longe de se chegar a um consenso. Um dos primeiro estudos a abordar essa questão foi o Relatório de Coleman (COLEMAN *et alli*, 1966), que concluiu que fatores de *background* familiar são mais importantes que fatores escolares na determinação do desempenho educacional da criança.

Esse trabalho deu subsidio a vários outros que estenderam a análise para outros indicadores educacionais. Esses estudos tanto confirmaram os achados de Coleman como discordaram. Um exemplo desse desacordo acadêmico é o recente debate entre HANUSHEK (1998 e 2002) e KRUEGER (1998 e 2000), que gira em torno principalmente do impacto do tamanho de turma, que tem sido alvo de muitas discussões nos países industrializados, por se tratar de um importante instrumento de política educacional. O tamanho de turma é considerado por muitos autores (SCHULTZ, 1987; JONES, 1990) um indicador de qualidade do ensino, pois considera-se que em classes pequenas os alunos têm maior atenção, tendo um extra-monitoramento.

No trabalho de 1998, HANUSHEK investiga o impacto do tamanho da sala de aula e da relação professor/aluno[6] na performance da educação por dois caminhos. Primeiro, ele seleciona alguns trabalhos publicados desde 1994 que analisam a relação professor/aluno na Função de Produção Escolar e conclui que em apenas 15% destes estudos esta variável é significativa e positiva. Alguns autores criticam este método, como KRUEGER (2003), argüindo que a amostra de estimativas utilizadas por Hanushek é enviesada para se tirar essa conclusão e a simples agregação dos estudos não é um processo correto para se tirar conclusões a respeito dos efeitos do tamanho das classes no desempenho do aluno.

Em um outro caminho, Hanushek utiliza dados do experimento americano do estado de Tennessee denominado Student/Teacher Achievement Ratio - STAR[7] - na qual conclui que classes menores têm impacto positivo apenas no jardim de infância, mas não tem efeito nos anos escolares subseqüentes.

KRUEGER (1999) examina o experimento STAR utilizando a Função de Produção Educacional, em que são incluídas variáveis de controle para os estudantes e professores, como sexo e cor dos estudantes e nível de escolaridade dos professores (diferentemente de trabalho de Hanushek visto acima que faz uma análise descritiva dos dados). Os resultados encontrados por Krueger sugerem que o principal benefício no atendimento em classes menores ocorre no primeiro ano de escola, depois o impacto é menor, porém ainda significativo.

[6] Tamanho de classe e relação professor/aluno não são iguais, na segunda pode haver professores que não dão aula, apenas trabalham na parte administrativa da escola.
[7] STAR é um estudo longitudinal onde os estudantes permanecem no mesmo tipo de classe (classes pequenas e médias) durante quatro anos.

Tanto Krueger quanto Hanushek concordam que a incerteza, na literatura, sobre o efeito do tamanho de classe (ou outro insumo escolar) é derivada da imperfeição da modelagem da FPE utilizada. KRUEGER (1999) coloca "grande parte da incerteza na literatura é devido ao fato que a especificação apropriada – incluindo a forma funcional, nível de agregação, variáveis de controle relevante e identificação – da Função de Produção Educacional é incerta".

Para LAZEAR (1999) a discordância entre o impacto da variável tamanho de classe nos estudos empíricos é devida em parte à relação que existe entre ela e estudantes desatentos ou de difícil aprendizagem. Classes pequenas podem representar que crianças precisam de mais ajuda, ou seja, o número ótimo de professores declina em classes com melhores estudantes, e desta forma, as classes menores vão ter estudantes mais desatentos, com pior rendimento.

Para LEE & BARRO (1997) quando a análise dos determinantes escolares é feita dentro de um mesmo país não fica claro como interpretar a associação positiva entre o resultado educacional e os insumos escolares devido a endogeneidade dos insumos escolhidos. A mobilidade dos residentes em busca de melhores escolas torna esse problema mais sério dentro de países que através deles. Dessa forma, os autores utilizam dados agregados de 58 países para analisar a relação entre os insumos familiares e escolares e o produto escolar - captado pelo desempenho do aluno, taxa de repetência e taxa de evasão – utilizando a Função de Produção Educacional.

Os insumos familiares e escolares foram agregados para países. Dessa forma, para captar os fatores familiares eles utilizam o produto *per capita* real, como *proxy* da renda da família, e os anos médios de escolaridade primária dos adultos de 25 anos e mais, como *proxy* da educação dos pais. Os recursos escolares foram captados pela taxa aluno/professor na educação, gasto público educacional real por estudante, salário real dos professores e comprimento do ano escolar, todos para o nível primário e agregados para país. Os resultados de LEE & BARRO indicam que os fatores familiares têm uma grande importância nas variáveis de resultado escolar, mas os recursos escolares também são positivamente relacionados à performance do estudante, principalmente a taxa aluno/professor.

Numa revisão literária sobre os principais trabalhos relacionados aos determinantes educacionais, BUCHMANN & HANNUM (2001) concluem que o impacto da qualidade da escola e dos professores é maior em países em

desenvolvimento do que em países desenvolvidos. Conclusão semelhante é vista em HEYNEMEN E LOXLEY (1983), que estudam uma amostra de 29 países, encontrando que as variáveis escolares explicam duas ou três vezes mais a variação nos testes de desempenho nos países em desenvolvimento do que em países desenvolvidos. Esse resultado pode ser devido a maior variabilidade dos recursos escolares nos países em desenvolvimento.

2.3) Conclusão

Nesse capítulo foi feita primeiramente uma revisão da Função de Produção Educacional, que servirá como ponto de partida para as análises dessa tese. Através dessa revisão percebe-se que ao se adaptar o arcabouço teórico da Função de Produção Educacional aos dados disponíveis surgem vários problemas, relacionados principalmente com a impossibilidade de mensurar alguns insumos do processo educacional tal como a capacidade genética - e a falta de informação completa sobre os insumos, devido a carência de estudos longitudinais. Tais problemas podem trazer algum viés nas estimativas da Função de Produção Educacional, contudo alguns praticamente não têm solução.

Cabe destacar, que o exame da FPE é importante por permitir organizar de forma mais clara a escolha das variáveis relevantes ao processo educacional, mesmo que sua especificação seja modificada, como é o caso dessa tese.

Com base na revisão dos principais trabalhos sobre os determinantes educacionais, verifica-se que tanto as características individuais e familiares quanto às características contextuais, ou seja, da escola ou da comunidade, afetam o resultado educacional. No Brasil, tem-se estudado com mais freqüência o impacto da família, e os resultados comprovam que realmente ela é importante para a educação das crianças e jovens. Porém, os estudos que consideram além da família, outros fatores, como a qualidade e disponibilidade do ensino e ambiente comunitário, demonstram que seus efeitos também são importantes.

3) ESTRATÉGIA EMPÍRICA, BASE DE DADOS E TRATAMENTO DAS VARIÁVEIS

Busca-se nesse capítulo expor ao leitor a estratégica empírica adotada nessa tese, bem como a descrição e mensuração das variáveis utilizadas e base de dados.

3.1) Especificação da Função de Produção Educacional

Para analisar o impacto de alguns insumos educacionais no resultado educacional utiliza-se como ponto de partida o arcabouço teórico da Função de Produção Educacional (FPE) visto na seção anterior, dada pela equação (2.1). Porém, foram feitas algumas modificações para ajustar à estrutura e disponibilidade dos dados.

O primeiro ajuste foi a utilização de modelos hierárquicos[8] de dois níveis na estimação da FPE, onde o primeiro nível considera o indivíduo e o segundo, o município. Esses modelos incorporam a estrutura hierárquica ou de agrupamento da população analisada.

O modelos hierárquicos são modelos de coeficientes aleatórios, apropriados para análise de variáveis aleatórias cujas distribuições sobre a população não são independentes e idênticas, mas sim agrupadas de forma não aleatória em sub-grupos relevantes (ALBERNAZ *et alii*, 2004).

A vantagem de estimar a FPE utilizando os modelos hierárquicos, ao invés dos modelos clássicos, é que nesse último os resultados podem ser enviesados, devido a quebra de seus pressupostos, que são erros independentes e identicamente distribuídos (i.i.d.).

No caso da análise dessa tese, os erros podem não ser i.i.d. quando se considera que há uma dependência entre o resultado educacional dos estudantes que são atendidos dentro de um mesmo município e que a relação entre as características dos estudantes e seu resultado educacional varia através dos municípios.

No Brasil, os trabalhos clássicos que utilizam os modelos hierárquicos para a análise escolar consideram no primeiro nível o aluno e no segundo, a escola. Esses estudos utilizam a base de dados do Sistema de Avaliação da Educação Básica - SAEB - e estão voltados ao estudo dos fatores associados ao desempenho escolar dos alunos.

[8] Os modelos hierárquicos serão melhor especificados no capítulo referente ao arcabouço metodológico.

Apesar de muito importante esse foco, já que os resultados do SAEB indicam que a qualidade do ensino é baixa, o interesse dessa tese é complementar esses estudos, analisando os demais problemas do ensino brasileiro que são a baixa cobertura do ensino médio, a distorção idade/série - conseqüência de uma alta repetência e/ou entrada tardia na escola - e a evasão.

Diferente dos trabalhos citados acima, nessa tese, a unidade de nível 2 são os municípios. Dessa forma, os fatores de insumos escolares e de professores, o termo S_t da equação (2.1), foram agrupados por município. Essa agregação pode incorrer em alguns problemas, que serão abordados a seguir.

Um exemplo é o debate entre Hanushek e Kruger sobre o efeito do tamanho de classe, já mencionado no Capítulo 2, que se estende também para a questão da agregação dos dados escolares. O primeiro considera dados ao nível da escola enquanto que o segundo utiliza dados escolares agregados.

A principal crítica de HANUSHEK (1996) aos trabalhos que utilizam dados agregados é que eles não consideram medidas relacionadas à família. Porém, esse não é o caso desse trabalho, já que no primeiro nível se encontram as variáveis de *background* familiar.

Outro problema colocado por ele, é que a política educacional dos estados pode não ser a mesma. Portanto, se estados variam em muitos aspectos da política educacional e operacionalização escolar, quando se ignoram esses efeitos, as estimativas dos coeficientes ligados as escolas poderão estar sujeitas a problemas de viés. A direção do viés irá depender do grau de correlação da política estadual e dos recursos escolares. Se positiva, o efeito dos recursos escolares agregados por estado tendem a ser positivo.

Essa crítica de Hanushek não se aplica apenas as análises que utilizam dados agregados, mas também aquelas que analisam a FPE utilizando os fatores escolares ao nível da escola em múltiplos estados. Ele coloca que apenas trabalhos que analisam os fatores escolares dentre de um mesmo estado, onde a política educacional é constante para todas as escolas, irão produzir estatísticas não viesadas.

CARD & KRUEGER (1996) colocam que um dos caminhos para resolver o problema da agregação dos dados da escola é analisar estudantes que foram expostos a recursos educacionais drasticamente diferentes. Eles argumentam ainda, que estudos que utilizam dados no nível da escola podem incorrer no problema da endogeneidade

dos recursos escolares dentro da escola. A endogeneidade ocorre, por exemplo, quando crianças com baixa performance escolar podem ser designadas para classes menores, o que pode levar para estimativas com viés descendente. Por outro lado, crianças altamente motivadas podem ser atraídas para escolas com maiores recursos por aluno, levando para estimativas com viés ascendente. Nesse caso, o uso de medidas escolares agregadas tende a minimizar o problema da endogeneidade dos recursos escolares dentro da escola. Somado a isso, a agregação dos dados pode reduzir ou eliminar erros aleatórios de medidas escolares quando há importantes erros de mensuração, sendo benéfica porque poderá levar para a média algum problema de mensuração.

Todos os problemas que podem surgir da agregação dos dados escolares para a média municipal podem não existir se as escolas dentro do município são homogêneas. Nesse caso, a agregação, através da média, representaria a escola típica do município. Porém, esse não é o caso das escolas brasileiras, onde há uma grande diferença das escolas dependendo da dependência administrativa (federal, municipal, estadual e particular), e mesmo entre escolas da mesma dependência administrativa.

Tal fato seria problema se o propósito fosse fazer um estudo da escola eficaz, o que não é o caso. Aqui, os insumos escolares são representados pelos insumos que caracterizam a rede escolar do município, dessa forma, não se pretende saber quais fatores da escola influenciam o resultado educacional. O objetivo passa a ser a rede de ensino eficaz, assim busca-se saber como o perfil da oferta escolar dos municípios determina o resultado educacional do estudante.

O perfil escolar dos municípios é bastante diferente, com alguns possuindo a maioria dos estabelecimentos com boa infra-estrutura e recursos humanos, além de grande oferta em todos os níveis de ensino, e por outro lado, outros com situação bastante precária, com poucas escolas, e alguns municípios nem sequer oferece a segunda metade do ensino fundamental e o ensino médio. Dessa forma, a pergunta que se coloca é: o que faz a rede escolar do município ser boa, no sentido de garantir um maior acesso e eficiência da suas crianças e jovens?

Nesse trabalho, além do modelo hierárquico de dois níveis, foi feita uma outra tentativa de melhorar a especificação da FPE. Essa nova especificação leva em conta os efeitos espaciais advindos do fato de se trabalhar com os municípios brasileiros. Esses efeitos são a autocorrelação e heterogeneidade espacial. Em outras palavras pode-se dizer que as variáveis agregadas de um município podem estar correlacionadas com as

dos municípios vizinhos, ou seja, podem existir aglomerados regionais para o resultado educacional (autocorrelação espacial). Da mesma forma, podem existir unidades espaciais com comportamento diferenciado com relação aos seus vizinhos, os chamados *outliers espaciais,* os quais sugerem heterogeneidade espacial.

Os modelos hierárquicos levam em conta a heterogeneidade espacial, ao passo que considera que há um efeito aleatório ligado a cada unidade de nível 2, nesse caso os municípios. Porem, eles não consideram 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. Esse pressuposto pode não ser realista, quando se considera que o Brasil possui uma forte heterogeneidade social e econômica, a qual se traduz numa grande heterogeneidade espacial, pode-se esperar que as áreas mais próximas sejam mais similares que as áreas mais distantes, ocorrendo, portanto, correlação espacial no nível macro. De fato, numa análise espacial exploratória feita no Capítulo 6, foi detectada alta correlação espacial nas covariáveis. Quando não se considera esses efeitos espaciais, os resultados obtidos para os coeficientes das variáveis de nível 2 podem ser enviesados e não eficientes.

Somado a isso, quando se está discutindo a rede escolar do município faz sentido levar em conta a correlação espacial entre as variáveis dependentes e independentes utilizadas na FPE, já que existe uma certa mobilidade dos indivíduos em busca de municípios que ofereçam determinado nível de ensino, ou que ofereçam um sistema escolar melhor, seja no aspecto da infra-estrutura ou da qualidade dos recursos humanos escolares. A externalidade espacial pode surgir também quando essa mobilidade ocorre nos próprios insumos escolares, como, por exemplo, professores mais bem qualificados atenderem vários municípios. Ou ainda, determinada atividade econômica pode beneficiar um conjunto de municípios próximos, tornando-os mais ricos e mais propícios a investir na educação, equipando melhor suas escolas.

Dessa forma, a segunda especificação da FPE foi feita incorporando a autocorrelação e a heterogeneidade espacial, características dos modelos espaciais, no modelo hierárquico, que possui uma natureza não espacial, mas que permite considerar diferentes níveis de agregação em uma mesma análise. Tal modelo é chamado de modelo hierárquico-espacial e será melhor explicado no capítulo que contém o arcabouço metodológico.

Vale a pena ressaltar, que os dados considerados na análise não são longitudinais, o que impossibilita uma análise do valor adicionado, já que o processo educacional é acumulativo no tempo. Assim, se estará supondo que o termo de erro aleatório dos dois níveis inclui a história passada dos insumos, sendo que o termo de erro do primeiro nível inclui os insumos familiares passados e o do segundo nível, os escolares e comunitários passados. Como visto no capítulo 2, esse pressuposto pode tornar a estimação dos coeficientes enviesados, se os insumos passados são correlacionados com os contemporâneos.

3.2) Como medir resultado educacional?

Como visto no primeiro capítulo, o acesso ao ensino fundamental no Brasil não se configura mais como um problema, posto que no ano de 2000, 94,50% das crianças de 7 a 14 anos estavam cursando esse nível de ensino. Já para o ensino médio, sua cobertura está bem aquém do desejável, com 77,71% das crianças de 15 a 17 anos cursando a escola e apenas 34,39% delas cursando o nível médio. Ao mesmo tempo foi detectado que um dos maiores problemas do ensino brasileiro são os elevados índices de distorção idade/série. Através da Probabilidade de Progressão por Série (PPS) percebeu-se também uma baixa progressão em algumas séries, conseqüência da evasão e/ou repetência. Os níveis desses três fatores – acesso, repetência e evasão – resultam na escolaridade média da população.

Dessa forma, nesse trabalho o resultado educacional englobou duas diferentes dimensões da educação. A primeira, diz respeito ao acesso da criança ao sistema de ensino que foi captada pela freqüência à escola.

Como as variáveis dependentes são medidas ao nível do aluno, a freqüência a escola foi dimensionada através de variáveis categóricas que tentaram modelar as probabilidades de freqüentar a escola tanto no ensino fundamental quanto no médio. Para modelar essa probabilidade foi utilizada as seguintes categorias para a variável resposta: “1”, caso o indivíduo esteja freqüentando a escola e, “0”, caso contrário. Como se pretende analisar o ensino fundamental e médio, foram utilizadas duas amostras com diferentes cortes etários, quais sejam: filhos de 7 a14 anos e de 15 a 17 anos. Como já foi mencionado, essas faixas etárias são as consideradas adequadas de cursar o ensino fundamental e médio, respectivamente.

Optou-se também por analisar a freqüência a escola da amostra referente a população de 15 a 17 anos que já concluiu o ensino fundamental. Já que através da análise da taxa de atendimento (GRAF. 1.11), pode-se perceber que no ensino médio ela apresenta caráter ambíguo, ou seja, a região Sul, que sabidamente apresenta melhor performance educacional, possui taxa de atendimento para a faixa etária de 15 a 17 anos menor que a região Norte, que possui pior performance.

A segunda dimensão está relacionada com o rendimento e eficiência do aluno na escola, ou seja, com a repetência e evasão escolar. Com já foi visto, um dos indicadores que expressa o grau de repetência, e também a entrada tardia no sistema de ensino, em uma comunidade é a taxa de distorção idade/série ou o seu complemento, a taxa de eficiência. No nível micro, ou seja, do indivíduo, optou-se por captar a taxa de eficiência através da probabilidade do aluno estar matriculado na série adequada a sua idade. Essa probabilidade foi modelada utilizando as seguintes categorias como variável resposta: "1", para o aluno que esteja cursando a série adequada a sua idade e, "0", caso contrário. Nesse caso, a amostra para o ensino fundamental foi composta pelos filhos de 7 a 14 anos que estejam freqüentando a escola, e para o médio apenas os filhos de 15 a 17 anos que estejam cursando a escola.

Por outro lado, tanto a repetência quanto a evasão escolar foram captadas pela Probabilidade de Progressão por Série (PPS). Para analisar o ensino fundamental, considerou-se a probabilidade de passar de zero para um ano de estudo, denominada por e_0, e a probabilidade de passar de quatro para cinco anos de estudo, e_4. Optou-se por estudar apenas estas duas transições do fundamental, pois de acordo com o que foi visto no Capítulo 1, essas PPSs são as mais importantes, constituindo os principais gargalos do ensino fundamental.

A primeira PPS foi modelada utilizando como variável reposta as seguintes categorias: "1" para os indivíduos que possuem um ou mais anos de estudo e "0" caso contrário. Nesse caso, a amostra foi composta pelos filhos de 10 anos de idade. A escolha dessa idade se deve ao fato de que, no Brasil, os alunos demoram em média pouco mais de dois anos para serem aprovados na 1ª série.

Para modelar e_4, foi utilizada como amostra apenas os indivíduos de 14 anos que possuem pelo menos quatro anos de estudo (concluiu a quarta série do ensino fundamental), e a variável resposta adotou os seguintes valores: "1" para os indivíduos que possuem cinco ou mais anos de estudo e "0" para os indivíduos que possuem

apenas quatro anos de estudo. Da mesma forma, considerou uma defasagem de dois anos na idade de conclusão da quinta série.

Para o ensino médio, foi estudada a probabilidade de passar de 8 para 9 anos de estudo (e_8)[9], que foi modelada considerando os filhos 18 anos de idade que possuem pelo menos oito anos de estudo e a variável resposta assumirá os seguintes valores: "1"para os indivíduos que tiverem nove ou mais anos de estudo e "0" caso contrário.

O uso das PPSs também tem a vantagem de permitir saber o efeito das variáveis independentes em cada etapa do processo educacional. Pode-se testar, por exemplo, se os fatores de *background* familiar são mais importantes no começo da vida escolar que nas transições mais avançadas ou, por outro lado, se o inverso ocorre para os fatores relacionados à oferta escolar do município. Quando a variável dependente é os anos médios de estudo, os coeficientes das variáveis independentes correspondem à média dos efeitos em todas as transições, não captando o impacto diferenciado nas diversas transições escolares que resultam nos anos médios de estudo.

Resumindo, as variáveis dependentes utilizadas nessa tese tentaram analisar duas diferentes dimensões da educação brasileira. A primeira refere-se ao acesso das crianças e jovens ao sistema de ensino, que foi captada pela probabilidade de cursar a escola. A segunda dimensão está relacionada com a eficiência e rendimento do aluno. Duas medidas foram utilizadas para avaliar essa ultima dimensão. A primeira foi a probabilidade dos alunos estarem cursando a escola na idade correta, que está captando variações tanto na repetência quanto na entrada tardia ao sistema de ensino. A segunda medida é a probabilidade de progressão por série, que buscou captar variações na repetência e evasão.

3.3) Variáveis Independentes

As variáveis independentes foram escolhidas de acordo com o arcabouço teórico revisado no capítulo anterior. No primeiro nível do modelo hierárquico se encontram aquelas relacionadas às características dos indivíduos e ao seu ambiente familiar.

[9] A escolha dessa PPS, também se deve ao fato dessa transição ser a mais importante do ensino médio, conforme visto no Capítulo 1.

CARACTERÍSTICAS DOS INDIVÍDUOS

As variáveis que representam as características individuais têm a principal função de controle na regressão. Nesse conjunto foi considerada a idade, o sexo, a cor e a localização do indivíduo.

No caso da idade, ela foi medida de forma contínua (anos completos de idade) e sua inclusão é importante, pois vários estudos demonstram que existe uma relação negativa entre a idade e o acesso à escola e a repetência, ou seja, quanto mais velha a criança maior a probabilidade de não freqüentar a escola e de freqüentar séries abaixo da adequada para a sua idade.

O sexo também foi incluído para controlar as possíveis diferenças em relação ao gênero, portanto, foi incluída uma variável categórica sendo "1", para os homens e "0", para as mulheres. Espera-se que o sinal dessa variável seja negativo, ou seja, que os homens tenham uma pior performance em relação às mulheres, já que a maioria dos estudos mostra que as mulheres jovens apresentam resultados educacionais superiores aos dos homens.

Outro fator de controle muito importante é a cor do indivíduo. A importância dessa variável para o caso brasileiro é analisada em vários estudos de estratificação social (FERNADES, 2001; SILVA e HASENBALG, 2001) que encontram um forte efeito da cor sobre as transições escolares. Os resultados mostram que indivíduos negros (cor preta ou parda) possuem desvantagem educacional em relação aos brancos, mesmo depois de se controlar por variáveis sócio-econômicas. Tal fato pode estar refletindo fatores discriminatórios ou culturais. Dessa forma, a variável cor foi considerada nas regressões de forma categórica, sendo "1" para brancos e amarelos; e "0" para pretos e pardos (ou negros).

A situação de residência do indivíduo, morador da zona rural ou urbana, foi outra variável controle, codificada da seguinte forma: "1", caso o indivíduo residisse na área urbana; e "0", caso contrário. Essa variável tentou captar os diferenciais urbano/rural no desempenho do sistema escolar.

BACKGROUND FAMILIAR

Como visto na revisão bibliográfica do capítulo 2, o conjunto de variáveis que representa o ambiente familiar é extremamente importante para o resultado educacional dos indivíduos. Dessa forma, as variáveis escolhidas são as normalmente utilizadas nos trabalhos sobre determinantes educacionais e que já se comprovou possuírem impactos no desempenho educacional do indivíduo. São elas:

1) Escolaridade da mãe. Foi medida em termos dos anos de estudo completo da mãe, que varia de 0 a 16 anos de estudo. De acordo com SILVA e HASENBALG (2001) ela representa o "capital cultural" familiar, sendo que a hipótese é de que crianças pertencentes a famílias com baixo capital cultural (nesse caso mães com baixa escolaridade), apresentam pior desempenho na escola, com maiores chances de evasão e de repetência. Alguns estudos utilizam a educação do chefe ou dos pais separadamente, porém, a escolha da escolaridade da mãe se deveu ao fato dela ser mais importante na determinação do resultado educacional dos filhos que a do pai, como mostra vários trabalhos, dentre eles o de LAM e BARROS (1995) e RIOS-NETO, CESAR e RIANI (2001).

2) Categoria de ocupação do chefe. Foram construídas categorias sócio-ocupacionais para o chefe da família, que levou em conta 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. A categoria de ocupação do chefe é indicadora tanto do "capital econômico" quanto do "capital cultural" da família, já que expressa tanto o nível de rendimento de cada ocupação quanto o de instrução. Pode-se supor também que há uma correlação dessas variáveis com a educação paterna, já que a maioria das famílias são chefiadas por homens.

3) Chefia feminina. Ela foi incluída como variável dicotômica, sendo "1" para os domicílios chefiados por mulher e "0", caso contrário. Segundo SILVA e HASENBALG (2001) essa variável está relacionada com o "capital social" e reflete o

fato da família ser quebrada ou não, já que a maioria das famílias chefiadas por mulheres é monoparental.

4) Família convivente. Também foi medida de forma dicotômica, assumindo os seguintes valores: "1" para famílias que convivem com outras no mesmo domicílio e "0", caso contrário. Pode-se dizer que essa variável também reflete o "capital social", no sentido de que está relacionada com os aspectos da estrutura familiar, tome-se como exemplo, o caso de filhos de mães solteiras que moram com os avós. Ela também foi incluída como forma de controlar o fato de outras pessoas adultas pertencentes a outra família, mas que convivem no mesmo domicílio, interferir na educação do indivíduo, seja por possuir maior nível educacional ou maior renda.

Cabe ressaltar, que algumas variáveis de *background* familiar, reconhecidamente importantes para o desempenho escolar do indivíduo, como é o caso da renda familiar *per capita* e do número de filhos na família, foram omitidas na análise. Optou-se por essa omissão por elas estarem sujeitas a possíveis vieses de simultaneidade, o que requer uma discussão teórica mais aprofundada. Somado a isso, a educação materna e a categoria de ocupação do chefe podem possuir uma forte correlação com essas variáveis.

Com relação as variáveis de nível 2, elas refletem a disponibilidade e qualidade da oferta escolar do município, além de aspectos demográficos e ambiente comunitário.

PERFIL ESCOLAR DO MUNICÍPIO

O efeito das variáveis relacionadas à rede ou perfil escolar do município no resultado educacional do indivíduo é um dos principais focos dessa tese. Os fatores da rede escolar do município estão relacionados à qualidade, disponibilidade e distribuição dos serviços educacionais.

1) Qualidade dos serviços educacionais. Para medir a qualidade dos serviços educacionais foi considerado dois conjuntos de insumos, um ligado à qualidade dos recursos humanos e outro à qualidade da infra-estrutura das escolas. Com relação ao primeiro, foram utilizadas três variáveis, quais sejam: média de horas-aula diária, média de alunos por turma e percentual de docentes com curso superior. No calculo da média municipal das duas primeiras variáveis, levou-se em consideração o número de matrículas em cada escola. Tal procedimento foi necessário, porque o INEP

disponibiliza esses indicadores por escola, como as matrículas não são distribuídas uniformemente, podendo haver um número maior de alunos matrículas em escolas com melhores recursos humanos, ou o inverso, a média municipal foi feita utilizando o número de matrículas como peso. A fórmula de cálculo dessas variáveis para cada escola é dada pelas seguintes expressões:

$$NATU_{ks} = \frac{M_{ks}}{T_{ks}} \quad (3.1)$$

Sendo:

$NATU_{ks}$ é o número médio de alunos por turma, no nível de ensino k, na série ou grupo de séries *s*;

M_{ks} é o número de matrículas em cada escola no nível de ensino k, na série ou grupo de séries *s*;

T_{ks} é número de turmas em cada escola no nível de ensino k, na série ou grupo de séries *s*.

$$NHAD_{ks} = \frac{\sum_{h=0}^{H} h_{ks} \times M_{ksh}}{M_{ks}} \quad (3.2)$$

Sendo:

$NHAD_{ks}$ é o número médio de horas-aula diária no nível de ensino k, na série ou grupo de séries *s*;

h_{ks} é o número de horas-aula diária no nível de ensino k, na série ou grupo de séries *s*;

M_{kh} é a matrícula no nível de ensino k, na série ou grupo de séries *s*, com *h* horas-aula diária;

M_k é o número de matrículas no nível de ensino k, na série ou grupo de séries *s*;

$$PD_k^{ES} = \frac{D_k^{ES}}{D_k} \times 100 \quad (3.3)$$

Sendo:

PD^{ES}_k é o percentual de docentes com formação superior no nível de ensino k;

D^{ES}_k é o número de docentes com formação superior, atuando no nível de ensino k;

D_k é o número total de docentes atuando no nível de ensino k.

Para a qualidade da infra-estrutura das escolas foi feita uma análise fatorial por componentes principais dos seguintes indicadores: percentual de alunos no município beneficiados por estabelecimentos com quadras de esporte, bibliotecas, laboratório de informática e laboratório de ciências. Esse índice foi chamado de fator de infra-estrutura.

2) Restrição de oferta. Com esse indicador é possível averiguar até que ponto a disponibilidade de serviços educacionais em cada município interfere no resultado educacional dos indivíduos, principalmente em relação ao acesso á escola. Para medir a restrição de oferta 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.

3) Distribuição das matrículas segundo o tamanho do estabelecimento. 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 Mistério da Educação e Cultura (MEC) é a nucleação das escolas, que procura aglomerar as escolas pequenas em um único espaço físico, de forma a tornar mais efetiva a atenção do governo sobre elas, facilitando assim o investimento em infra-estrutura. Dessa forma, foram criadas três variáveis que reflitam a disposição dos alunos por tamanho de estabelecimentos, quais sejam: proporção de estabelecimentos com 50 a 200 matrículas, proporção de estabelecimentos com 201 a 1000 matrículas e proporção de estabelecimentos com mais de 1000 matrículas.

4) Distribuição segundo dependência administrativa. Mesmo controlando o perfil da rede de ensino dos municípios pela qualidade da infra-estrutura e dos recursos escolares e demais fatores, pode haver diferenças nos resultados dos alunos devido a diferenças pedagógicas ou de financiamento dependendo da categoria administrativa da escola, principalmente entre a escola pública e privada. Mesmo dentre as escolas públicas pode haver diferença entre as federais, estaduais e municipais. Dessa forma, foram incluídas variáveis que discriminem a distribuição das matrículas do município por dependência administrativa, quais sejam: porcentagem de matrículas públicas no

município e porcentagem de matrículas municipais dentre as públicas. Optou-se por discriminar apenas as escolas municipais dentre as públicas porque as federais são um número muito pequeno de escolas, principalmente no caso do ensino fundamental.

5) Existência de ciclos. Foi medida em termos do número de matrículas em escolas com ciclo sobre o total de matrículas do município. Os ciclos são uma forma de organização do ensino que distribui os anos de escolaridade ou séries em grupos, só existindo para o nível fundamental. Essa variável serviu principalmente como forma de controle, ressaltando que a análise do seu efeito é importante no sentido de saber até que ponto a distorção idade/série está sendo afetada por tal política.

CARACTERISTICAS DEMOGRÁFICAS

O Brasil tem passado por mudanças demográficas profundas, com significativa queda da Taxa de Fecundidade Total, alcançando, em alguns municípios, níveis abaixo do de reposição, o que resulta na diminuição do peso relativo da população em idade escolar e até absoluto em alguns casos. Na literatura de demografia econômica há um clássico debate entre COALE & HOOVER (1958) e SCHULTZ (1987) sobre os efeitos de um aumento ou queda da fecundidade na educação. De acordo com o primeiro, uma menor fecundidade leva a uma diminuição da taxa de dependência de jovens, reduzindo o número de pessoas matriculadas na escola. Tal redução provoca um aumento da poupança do governo relativo a diminuição nos gastos educacionais. Já para Schultz, os gastos do governo podem não diminuir na mesma proporção da diminuição da coorte em idade escolar, podendo ocorrer uma melhora na qualidade do ensino.

Dessa forma, pode-se esperar que o impacto demográfico tenha influência nos resultados educacionais, no sentido de abrir uma janela de oportunidades para a educação brasileira. Assim, foi incluída nesse estudo uma variável que represente a pressão demográfica e que possa captar possíveis dividendos demográficos associados com o tamanho relativo da população em idade escolar adequada para cursar determinado nível de ensino.

Também foram incluídas variáveis categóricas que identificam o tamanho do município como forma de captar a economia de escala do setor de serviços educacionais em municípios que possuem maior população. Essa variável também pode refletir o grau de urbanização do município.

AMBIENTE COMUNITÁRIO

Como visto na revisão bibliográfica, o artigo de BARROS *et alli* (2001) propõe utilizar a escolaridade média da população adulta e a renda *per capita* da comunidade como medidas do ambiente comunitário. Porém, nesse trabalho optou-se por utilizar apenas a escolaridade média da população adulta do município, que foi medida em termos da média aritmética da escolaridade dos indivíduos com mais de 20 anos de idade. A renda *per capita* não foi utilizada por possuir alta correlação com a primeira variável e por estar sujeita a possíveis vieses de simultaneidade. Tal omissão não deve trazer grandes problemas, já que no próprio trabalho de BARROS *et alli* a renda *per capita* da comunidade mostrou não possuir importância, o contrário ocorrendo com a escolaridade media da população adulta.

Cabe ressaltar que as variáveis independentes do segundo nível listadas acima, não necessariamente foram utilizadas em todas as estimações de regressão. Dependendo da dimensão educacional analisada (acesso e eficiência ou rendimento) alguns fatores não têm importância, não sendo, portanto, considerados nos modelos de regressão.

3.4) Base de Dados e amostra

Para poder analisar o impacto de fatores relacionados ao indivíduo e a rede escolar do município, bem como alguns fatores demográficos e municipais, foram utilizados os dados do Censo Demográfico de 2000, fornecidos pelo IBGE e do Censo Escolar de 2000, obtido junto ao INEP. Na primeira base, as informações são obtidas no nível do indivíduo, e não há quesitos referentes às características escolares. Os quesitos sobre educação só fornecem informações referentes à série e grau que o indivíduo freqüenta ou freqüentou.

Os dados do Censo Escolar são desagregados ao nível de escola, sendo possível obter informação sobre suas características, tais como estrutura física, níveis e modalidades de ensino, quantidade de professores e sua escolaridade, número de matrículas, repetência, etc, também é possível identificar o município onde a escola se localiza. Infelizmente, nessa base de dados não existem informações a respeito das características do aluno e da sua família.

Assim, para se ter a informação sobre as características individuais e familiares dos indivíduos e as características da rede escolar do município, foi feita uma junção nas duas bases, utilizando como variável de ligação o código do município (de acordo com a codificação do IBGE), comum às duas bases.

Dessa forma, do Censo Demográfico foram retirados os dados referentes às características pessoais, familiares, residência do indivíduo e características municipais, tais como o tamanho da população em idade escolar e total e a escolaridade média da população adulta. Do Censo Escolar obteve-se informações referentes à rede escolar do município e do número de professores.

Nas TAB. 3.1 e 3.2 encontra-se o número de casos bem como as estatísticas descritivas das variáveis utilizadas para as diferentes amostras relativas à análise do ensino fundamental. No nível 1, a amostra contempla os filhos de famílias cuja mãe está presente, enquanto que no nível 2 a amostra considera os municípios brasileiros no ano de 2000, que somam 5.507. No caso das PPS, o menor número da amostra no nível 2 é devido ao fato de que em alguns municípios não foi amostrado nenhuma criança de 10 ou 14 anos de idade. Percebe-se que em todos os níveis as amostras são relativamente grandes, o que diminui possíveis problemas relacionados a multicolinearidade.

Através da TAB. 3.1, percebe-se que 95% das crianças de 7 a 14 anos da amostra freqüentam a escola, mas apenas 42%, dentre as que cursam a escola, estão na séria adequada a sua idade. A probabilidade de passar de zero para um ano de estudo, para os filhos de 10 anos de idade, é de 87%, e para os filhos de 14 anos que já concluíram a quarta série concluir a quinta é de 84%. Com relação as variáveis independentes do nível 1, pode-se destacar a baixa média de domicílios chefiados por mulher, a homogeneidade com relação a distribuição dos sexos e cor, a concentração dos chefes ocupados na categoria de nível manual e o baixo percentual de pessoas morando em áreas rurais.

Para as variáveis do segundo nível (TAB. 3.2), ressalta-se o baixo percentual de professores com curso superior, além de uma grande variação entre os municípios, com alguns não possuindo nenhum professor com curso superior e outros com todos. Também foi verificada grande variação no número médio de alunos por turma. Com relação à qualidade da infra-estrutura da rede escolar, percebe-se uma baixa média de

estabelecimentos com laboratórios de informática e ciências. Apesar do número diferente da amostra, a média das variáveis foi a mesma.

TABELA 3. 1: Estatísticas descritivas das variáveis do nível 1 para as diferentes amostras do Ensino Fundamental

NOME DA VARIÁVEL	N	MÉDIA	D. PADRÃO	MÍNIMO	MÁXIMO
Probabilidade de frequentar escola - filhos de 7 a 14 anos					
Frequenta escola	2972471	0,95	0,22	0	1
Idade	2972471	10,54	2,28	7	14
Dummy situação de domicílio; 1=urbano e 0=rural	2972471	0,71	0,45	0	1
Dummy cor; 1=branco e amarelo e 0=preto e pardo	2972471	0,50	0,50	0	1
Dummy sexo; 1=homem e 0=mulher	2972471	0,51	0,50	0	1
Educação da mãe	2972471	5,15	4,06	0	16
Domicílio chefiado por mulher	2972471	0,18	0,38	0	1
Dummy família convivente	2972471	0,04	0,18	0	1
Categoria de ocupação do chefe nível superior	2972471	0,07	0,25	0	1
Categoria de ocupação do chefe nível médio	2972471	0,16	0,37	0	1
Categoria de ocupação do chefe nível manual	2972471	0,49	0,50	0	1
Chefe não ocupado	2972471	0,28	0,45	0	1
Probabilidade de frequentar escola na idade adequada - filhos de 7 a 14 anos que frequentam escola					
Frequenta escola na idade adequada	2814140	0,42	0,49	0	1
Idade	2814140	10,53	2,26	7	14
Dummy situação de domicílio; 1=urbano e 0=rural	2814140	0,73	0,45	0	1
Dummy cor; 1=branco e amarelo e 0=preto e pardo	2814140	0,51	0,50	0	1
Dummy sexo; 1=homem e 0=mulher	2814140	0,51	0,50	0	1
Educação da mãe	2814140	5,30	4,06	0	16
Domicílio chefiado por mulher	2814140	0,18	0,38	0	1
Dummy família convivente	2814140	0,03	0,18	0	1
Categoria de ocupação do chefe - nível superior	2814140	0,07	0,26	0	1
Categoria de ocupação do chefe - nível médio	2814140	0,16	0,37	0	1
Categoria de ocupação do chefe - nível manual	2814140	0,49	0,50	0	1
Chefe não ocupado	2814140	0,27	0,45	0	1
Probabilidade de passar na 1ª série - filhos de 10 anos					
Aprovado na 1ª série	365561	0,87	0,33	0	1
Dummy situação de domicílio; 1=urbano e 0=rural	365561	0,71	0,45	0	1
Dummy cor; 1=branco e amarelo e 0=preto e pardo	365561	0,50	0,50	0	1
Dummy sexo; 1=homem e 0=mulher	365561	0,52	0,50	0	1
Educação da mãe	365561	5,18	4,06	0	16
Domicílio chefiado por mulher	365561	0,17	0,38	0	1
Dummy família convivente	365561	0,03	0,18	0	1
Categoria de ocupação do chefe - nível superior	365561	0,07	0,25	0	1
Categoria de ocupação do chefe - nível médio	365561	0,17	0,37	0	1
Categoria de ocupação do chefe - nível manual	365561	0,49	0,50	0	1
Chefe não ocupado	365561	0,28	0,45	0	1
Probabilidade de passar na 5ª série - filhos de 14 anos com no mínimo 4 anos de estudo					
Aprovado na 5ª série	293097	0,84	0,37	0	1
Dummy situação de domicílio; 1=urbano e 0=rural	293097	0,79	0,41	0	1
Dummy cor; 1=branco e amarelo e 0=preto e pardo	293097	0,56	0,50	0	1
Dummy sexo; 1=homem e 0=mulher	293097	0,49	0,50	0	1
Educação da mãe	293097	5,65	4,08	0	16
Domicílio chefiado por mulher	293097	0,20	0,40	0	1
Dummy família convivente	293097	0,02	0,14	0	1
Categoria de ocupação do chefe - nível superior	293097	0,08	0,28	0	1
Categoria de ocupação do chefe - nível médio	293097	0,16	0,36	0	1
Categoria de ocupação do chefe - nível manual	293097	0,49	0,50	0	1
Chefe não ocupado	293097	0,27	0,45	0	1

TABELA 3. 2: Estatísticas descritivas das variáveis do nível 2 para as diferentes amostras do Ensino Fundamental

NOME DA VARIÁVEL	Probabilidade de frequentar escola e frequentar na idade adequada		Probabilidade de passar na 1ª série (e_0)		Probabilidade de passar na 5ª série (e_4)	
Número da amostra	5507		5506		5505	
	MÉDIA	D. PADRÃO	MÉDIA	D. PADRÃO	MÉDIA	D. PADRÃO
1. PERFIL DAS ESCOLAS MUNICIPAIS						
1.1 Recursos Humanos Escolares						
Média de horas-aulas diárias	4.27	0.30	4.27	0.31	4.27	0.30
Média de alunos por turma	32.32	9.26	32.32	9.26	32.32	9.26
Porcentagem de professores com curso superior	0.37	0.28	0.37	0.28	0.37	0.28
1.2 Infra-estrutura Escolar						
Fator de infra-estrutura	0.00	1.00	0.00	1.00	0.00	1.00
Proporção de matrículas beneficiadas em escolas com bibliotecas	0.51	0.32	0.51	0.32	0.51	0.32
Proporção de matrículas beneficiadas em escolas com quadras	0.41	0.33	0.41	0.33	0.41	0.33
Proporção de matrículas beneficiadas em escolas com laboratório de informática	0.14	0.21	0.14	0.21	0.14	0.21
Proporção de matrículas beneficiadas em escolas com laboratório de ciências	0.17	0.24	0.17	0.24	0.17	0.24
1.3 Restrição de Oferta						
Razão professor por população em idade escolar	0.06	0.02	0.06	0.02	0.06	0.02
1.4 Nucleação						
Proporção de estabelecimentos com 50 a 200 matrículas	0.27	0.19	0.27	0.19	0.27	0.19
Proporção de estabelecimentos com 201 a 1000 matrículas	0.30	0.25	0.30	0.25	0.30	0.25
Proporção de estabelecimentos com mais de 1000 matrículas	0.02	0.05	0.02	0.05	0.02	0.05
1.5 Distribuição das Matrículas Segundo Dependência Administrativa						
Proporção de matrículas públicas	0.97	0.04	0.97	0.04	0.97	0.04
Proporção de matrículas municipais dentre as públicas	0.54	0.25	0.54	0.25	0.54	0.25
1.6 Existência de Ciclos						
Percentual de matrículas em escolas com ciclos	0.30	0.37	0.30	0.37	0.30	0.37
2. CARACTERÍSTICAS DEMOGRÁFICAS						
2.1 Pressão Demográfica						
Tamanho relativo da coorte de 7 a 14 anos	0.17	0.03	0.17	0.03	0.17	0.03
2.2 Dummies para Tamanho da População do Município						
População de 0 a 4999	0.24	0.43	0.24	0.43	0.24	0.43
População de 5000 a 9999	0.24	0.43	0.24	0.43	0.24	0.43
População de 10000 a 19999	0.25	0.43	0.25	0.43	0.25	0.43
População de 20000 a 49999	0.18	0.38	0.18	0.38	0.18	0.38
População de50000 a 99999	0.05	0.23	0.05	0.23	0.05	0.23
População de100000 a 199999	0.02	0.14	0.02	0.14	0.02	0.14
Acima de 200000	0.02	0.14	0.02	0.14	0.02	0.14
3 AMBIENTE COMUNITÁRIO						
Educação média da população adulta	4.37	1.28	4.37	1.28	4.37	1.28

Nas TABs. 3.3 e 3.4 encontram-se as estatísticas descritivas dos respectivos níveis 1 e 2 das amostras utilizadas na análise do ensino médio. O menor número de casos no nível 2 é conseqüência, principalmente, do fato de que vários municípios não possuem estabelecimentos que ofereçam esse nível de ensino.

A análise descritiva das variáveis de nível 1, revelam que 81% dos filhos de 15 a 17 anos freqüentam a escola, e dos que já concluíram o ensino fundamental, apenas 48% continuam na escola. Observa-se também uma baixa incidência dos filhos nessa faixa etária que estão na série adequada. Diferentemente do caso anterior, nas amostras relativas ao ensino médio, há uma incidência um pouco maior de filhos de cor branca ou amarela.

Com relação as variáveis de nível 2, ressalta-se que, de modo geral, no ensino médio há uma maior predominância nos municípios de estabelecimento com melhor

recursos humanos e infra-estrutura do que no fundamental, além de uma maior incidência de escolas maiores.

TABELA 3. 3: Estatísticas descritivas das variáveis do nível 1 para as diferentes amostras do Ensino Médio

NOME DA VARIÁVEL	N	MÉDIA	D. PADRÃO	MÍNIMO	MÁXIMO
Probabilidade de frequentar escola - filhos de 15 a 17 anos					
Frequenta escola	1035760	0,81	0,39	0	1
Idade	1035760	15,98	0,82	15	17
Dummy situação de domicílio; 1=urbano e 0=rural	1035760	0,74	0,44	0	1
Dummy cor; 1=branco e amarelo e 0=preto e pardo	1035760	0,51	0,50	0	1
Dummy sexo; 1=homem e 0=mulher	1035760	0,54	0,50	0	1
Educação da mãe	1035760	4,92	4,13	0	16
Domicílio chefiado por mulher	1035760	0,22	0,41	0	1
Dummy família convivente	1035760	0,02	0,12	0	1
Categoria de ocupação do chefe nível superior	1035760	0,07	0,26	0	1
Categoria de ocupação do chefe nível médio	1035760	0,16	0,37	0	1
Categoria de ocupação do chefe nível manual	1035760	0,45	0,50	0	1
Chefe não ocupado	1035760	0,31	0,46	0	1
Probabilidade de frequentar escola - filhos de 15 a 17 anos que já concluíram o ensino fundamental					
Frequenta escola	419983	0,48	0,5	0	1
Idade	419983	16,16	0,79	15	17
Dummy situação de domicílio; 1=urbano e 0=rural	419983	0,87	0,34	0	1
Dummy cor; 1=branco e amarelo e 0=preto e pardo	419983	0,68	0,47	0	1
Dummy sexo; 1=homem e 0=mulher	419983	0,46	0,50	0	1
Educação da mãe	419983	7,09	4,28	0	16
Domicílio chefiado por mulher	419983	0,20	0,40	0	1
Dummy família convivente	419983	0,01	0,12	0	1
Categoria de ocupação do chefe - nível superior	419983	0,14	0,34	0	1
Categoria de ocupação do chefe - nível médio	419983	0,17	0,37	0	1
Categoria de ocupação do chefe - nível manual	419983	0,44	0,50	0	1
Chefe não ocupado	419983	0,25	0,43	0	1
Probabilidade de frequentar escola na idade correta - filhos de 15 a 17 anos que frequentam escola					
Frequenta escola na idade adequada	841420	0,24	0,43	0	1
Idade	841420	15,93	0,82	15	17
Dummy situação de domicílio; 1=urbano e 0=rural	841420	0,77	0,42	0	1
Dummy cor; 1=branco e amarelo e 0=preto e pardo	841420	0,53	0,50	0	1
Dummy sexo; 1=homem e 0=mulher	841420	0,52	0,50	0	1
Educação da mãe	841420	5,38	4,22	0	16
Domicílio chefiado por mulher	841420	0,21	0,41	0	1
Dummy família convivente	841420	0,02	0,12	0	1
Categoria de ocupação do chefe - nível superior	841420	0,09	0,28	0	1
Categoria de ocupação do chefe - nível médio	841420	0,16	0,37	0	1
Categoria de ocupação do chefe - nível manual	841420	0,45	0,50	0	1
Chefe não ocupado	841420	0,30	0,46	0	1
Probabilidade de passar na 1ª série do ensino médio (e_8) - filhos de 18 anos					
Aprovado na 1ª série do ensino médio	179789	0,79	0,41	0	1
Dummy situação de domicílio; 1=urbano e 0=rural	179789	0,87	0,33	0	1
Dummy cor; 1=branco e amarelo e 0=preto e pardo	179789	0,64	0,48	0	1
Dummy sexo; 1=homem e 0=mulher	179789	0,49	0,50	0	1
Educação da mãe	179789	6,54	4,26	0	16
Domicílio chefiado por mulher	179789	0,23	0,42	0	1
Dummy família convivente	179789	0,01	0,11	0	1
Categoria de ocupação do chefe - nível superior	179789	0,12	0,33	0	1
Categoria de ocupação do chefe - nível médio	179789	0,15	0,36	0	1
Categoria de ocupação do chefe - nível manual	179789	0,44	0,50	0	1
Chefe não ocupado	179789	0,29	0,45	0	1

TABELA 3. 4: Estatísticas descritivas das variáveis do nível 2 para as diferentes amostras do Ensino Médio

NOME DA VARIÁVEL	Probabilidade de frequentar escola e frequentar na idade adequada		Probabilidade de frequenar a escola para ensino médio condicional		Probabilidade de passar na 5ª série (e_4)	
Número da amostra	5227		5218		5188	
	MÉDIA	D. PADRÃO	MÉDIA	D. PADRÃO	MÉDIA	D. PADRÃO
1. PERFIL DAS ESCOLAS MUNICIPAIS						
1.1 Recursos Humanos Escolares						
Média de horas-aulas diárias	4.16	0.38	4.16	0.38	4.16	0.38
Média de alunos por turma	35.03	6.89	35.04	6.87	35.08	6.84
Porcentagem de professores com curso superior	0.77	0.29	0.77	0.29	0.77	0.29
1.2 Infra-estrutura Escolar						
Fator de infra-estrutura						
Proporção de matrículas beneficiadas em escolas com bibliotecas	0.76	0.40	0.76	0.40	0.76	0.40
Proporção de matrículas beneficiadas em escolas com quadras	0.61	0.45	0.61	0.45	0.61	0.45
Proporção de matrículas beneficiadas em escolas com laboratório de informática	0.33	0.43	0.33	0.43	0.33	0.43
Proporção de matrículas beneficiadas em escolas com laboratório de ciências	0.35	0.44	0.35	0.44	0.35	0.44
1.3 Restrição de Oferta						
Razão professor por população em idade escolar	0.04	0.03	0.04	0.03	0.04	0.03
1.4 Nucleação						
Proporção de estabelecimentos com 50 a 200 matrículas	0.41	0.42	0.41	0.42	0.41	0.41
Proporção de estabelecimentos com 201 a 1000 matrículas	0.50	0.42	0.50	0.42	0.51	0.42
Proporção de estabelecimentos com mais de 1000 matrículas	0.03	0.12	0.03	0.12	0.03	0.12
1.5 Distribuição das Matrículas Segundo Dependência Administrativa						
Proporção de matrículas públicas	0.95	0.12	0.95	0.12	0.95	0.12
Proporção de matrículas municipais dentre as públicas	0.09	0.27	0.09	0.27	0.09	0.27
2. CARACTERÍSTICAS DEMOGRÁFICAS						
2.1 Pressão Demográfica						
Tamanho relativo da coorte de 15 a 17 anos	0.07	0.01	0.07	0.01	0.07	0.01
2.2 Dummies para Tamanho da População do Município						
População de 0 a 4999	0.22	0.41	0.22	0.41	0.21	0.41
População de 5000 a 9999	0.24	0.43	0.24	0.43	0.24	0.43
População de 10000 a 19999	0.26	0.44	0.26	0.44	0.26	0.44
População de 20000 a 49999	0.18	0.39	0.18	0.39	0.19	0.39
População de50000 a 99999	0.06	0.23	0.06	0.23	0.06	0.23
População de100000 a 199999	0.02	0.15	0.02	0.15	0.02	0.15
Acima de 200000	0.02	0.14	0.02	0.14	0.02	0.14
3 AMBIENTE COMUNITÁRIO						
Educação média da população adulta	4.43	1.25	4.44	1.25	4.44	1.25

3.5) Conclusão

Nesse capítulo foi feita a especificação da Função de Produção Educacional utilizada nessa tese. Serão consideradas duas formas de especificação. Primeiramente, a FPE será estimada aplicando o modelo hierárquico de dois níveis, onde o primeiro nível considera o aluno e o segundo, o município. Os resultados dessa especificação são discutidos no Capítulo 5. Num segundo momento, será utilizado o modelo hierárquico-espacial que considera a autocorrelação e heterogeneidade espacial nas variáveis pertencentes ao segundo nível. As estimações desse modelo será o tema do Capítulo 7.

Também foram definidas as variáveis dependentes, que englobam duas diferentes dimensões da educação brasileira, quais sejam: o acesso das crianças à escola e o rendimento e/ou a eficiência do aluno na escola. A primeira dimensão está

relacionada com a quantidade de educação demandada e a segunda com a qualidade educacional.

Com relação as variáveis independentes, no primeiro nível se encontram aquelas relacionadas às características individuais e *background* familiar, e no segundo nível, as variáveis consideradas estão relacionadas à disponibilidade e qualidade da oferta escolar do município, aspectos demográficos e ambiente comunitário.

4) ARCABOUÇO METODOLÓGICO

Nesse capítulo são discutidos os dois modelos estatísticos que serviram de base para a especificação da Função de Produção Educacional. Trata-se do modelo logístico hierárquico e do modelo de regressão espacial. Discute-se também a metodologia de construção do modelo hierárquico-espacial, que é uma tentativa de inserção da autocorrelação espacial na estimação do modelo hierárquico.

4.1) O Modelo Logístico Hierárquico

Para uma melhor compreensão dos modelos hierárquicos, ou multiníveis como também são chamados, primeiramente será feita uma explanação do modelo hierárquico linear para em seguida especificar o modelo logístico hierárquico.

Os modelos hierárquicos incorporam a estrutura hierárquica dos dados, assumindo que a variável dependente é medida no menor nível de agregação e as variáveis independentes em todos os outros níveis (HOX, 1995). Essa modelagem traz inúmeras vantagens em relação aos modelos lineares clássicos estimados por Mínimos Quadrados Ordinários (MQO) uma vez que permite incorporar a quebra de um pressuposto importante do modelo linear, que é o de termos de erro independentes e identicamente distribuídos (i.i.d), decorrente da dependência que os indivíduos tem dentro de uma mesma unidade (escola, município, Estado, hospital, etc). Somado a isso, esses modelos resolvem o problema da unidade de análise, quando se tem variáveis independentes medidas em níveis distintos de agregação.

Os modelos hierárquicos além de proporcionarem a obtenção de melhores estimativas para os parâmetros, também possibilitam testes de hipóteses do efeito entre os níveis, ou seja, permite analisar como as variáveis independentes medidas no nível 2 modificam o efeito das variáveis independentes medidas no nível 1. Permite também a partição da variância e covariância em componentes entre os níveis, explicando a variabilidade dos dados nos diversos níveis (BRYK, A. S. & RAUDENBUSH, S. W., 2002).

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 análise hierárquica é conveniente considerar primeiro o modelo mais simples, sem variáveis explicativas, e o intercepto tendo efeito aleatório, ou seja, considera que o valor médio esperado da variável dependente varia entre as unidades de nível 2. Esse modelo é chamado de componente da variância completamente não-condicional, ou modelo hierárquico nulo, e é útil por permitir estimar a variabilidade associada a cada nível, o que possibilita identificar a contribuição de cada nível para a variação total. 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 conforme abaixo:

$$y_{ij} = \beta_{0j} + e_{ij} \qquad (4.1)$$

E a do nível 2 como:

$$\beta_{0j} = \gamma_{00} + u_{0j} \qquad (4.2)$$

Substituindo (4.2) em (4.1), tem-se a equação final da seguinte forma:

$$y_{ij} = \gamma_{00} + u_{oj} + e_{ij} \qquad (4.3)$$

Em que:

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

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

β_{0j} é o resultado médio para a j*th* unidade, ou seja, $\beta_{0j} = \mu_{yj}$;

u_{oj} é o efeito aleatório associado ao nível 2, ou seja, aplicada a todos os indivíduos na j*th* unidade;

e_{ij} é o efeito aleatório associado ao nível 1, ou seja, é a quantidade aleatória para o i*th* indivíduo na j*th* unidade e

γ_{00} é a média do resultado para todos os indivíduos.

Assume-se que u_{0j} e e_{ij} são independentes e seguem distribuição normal com média zero e variâncias constantes σ^2_{uo} e σ^2_{e}, respectivamente. O termo γ_{00} é o parâmetro da parte fixa do modelo.

Os parâmetros σ^2_e e σ^2_{uo} representam respectivamente a variabilidade dentro do grupo e entre os grupos, e são referidos como os parâmetros aleatórios do modelo. Dessa forma, a variância de y_{ij} é resultado dessas duas variâncias e é dada por VAR (y_{ij}) = $\sigma^2_{uo} + \sigma^2_e$.

Com os valores de σ^2_{uo} e σ^2_e desse modelo pode-se calcular uma medida denominada correlação intra-classe, que mede a proporção da variância devida ao nível 2.

$$\rho = \frac{\sigma^2_{u0}}{\sigma^2_e + \sigma^2_{uo}} \quad (4.4)$$

O valor dessa estatística varia entre 0 e 1. Um caso extremo de coeficiente de correlação igual a zero, significaria que as unidades de nível 2 seriam homogêneas. O extremo oposto significaria que toda a variabilidade de y_{ij} é devida às diferenças entre as unidades do nível 2.

O próximo passo é acrescentar ao modelo descrito na equação (4.1) as variáveis independentes medidas no nível 1, ou seja, aquelas relacionadas ao indivíduo, e que possam explicar parte da variabilidade associada ao nível 1. Para facilitar as anotações das equações será considerada apenas uma variável independente, mas o modelo pode ser expandido para k variáveis independentes. Assim, a equação no primeiro nível é dada por:

$$y_{ij} = \beta_{0j} + \beta_{1j} X_{1ij} + e_{ij} \qquad (4.5)$$

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

$$\beta_{0j} = \gamma_{00} + u_{0j} \qquad (4.6)$$

$$\beta_{1j} = \gamma_{10} \qquad (4.7)$$

Substituindo (4.6) e (4.7) em (4.5) teremos a equação:

$$y_{ij} = \gamma_{00} + \gamma_{10} X_{1ij} + u_{oj} + e_{ij} \qquad (4.8)$$

Sendo:

γ_{00} e γ_{10} são os parâmetros da parte fixa do modelo, comum a todos os indivíduos.

Quando se considera apenas o intercepto com efeito aleatório, assume-se que a inclinação é a mesma para todas as unidades de nível 2, mudando apenas o intercepto de cada unidade j, que é dado por ($\gamma_{00} + u_{0j}$). Entretanto, pode-se também ampliar o modelo considerando-se que a inclinação difere aleatoriamente de j para j. Em outras palavras pode-se dizer que a inclinação não é a mesma para cada unidade de nível 2. A equação de nível 1 continua igual a (4.5), porém, o modelo de nível 2 é definido conforme o conjunto de equações abaixo:

$$\beta_{0j} = \gamma_{00} + u_{0j} \qquad (4.9)$$

$$\beta_{1j} = \gamma_{10} + u_{1j} \qquad (4.10)$$

Substituindo as equações do segundo nível na do primeiro, tem-se:

$$y_{ij} = \gamma_{00} + \gamma_{10} X_{1ij} + u_{1j} X_{1ij} + u_{0j} + e_{ij} \qquad (4.11)$$

Note que esse modelo possui três componentes aleatórios, quais sejam: u_{oj}, que é o efeito aleatório da unidade j na média; $u_{1j}X_{1ij}$, em que u_{1j} é o efeito aleatório na inclinação e e_{ij} que é o termo de erro do nível 1.

Supõe-se que os termos residuais (u_{oj}, u_{1j} e e_{ij}) possuem distribuição normal com média zero e variâncias constantes especificadas por σ^2_{u0}, σ^2_{u1} e σ^2_{e}, respectivamente. Os resíduos do nível 2 são assumidos independentes dos resíduos do nível 1, mas correlacionados entre si, com covariância dada por σ_{u01}.

Pose-se expandir o modelo e considerar mais de uma variável independente e todas possuindo um termo aleatório. Esse seria um modelo mais complexo e de difícil estimação, pois são muitos os parâmetros desconhecidos, o que resulta num processo complexo e lento. Por essa razão, na maioria das vezes, considera-se apenas o intercepto possuindo efeito aleatório. Porém, de acordo com o arcabouço teórico, pode-se assumir outros coeficientes como aleatório. Essa hipótese pode ser testada através de teste disponível no pacote estatístico HLM.

Os modelos vistos até o presente podem ser chamados de modelos de variância não-condicional, pois não possuem nenhuma variável independente no nível superior.

Esses modelos, no entanto, podem ser expandidos considerando-se que um conjunto de variáveis independentes medidas no nível 2 explicam a variação aleatória do intercepto e/ou da inclinação nesse nível. Dessa forma, a especificação da equação de nível 1 continua sendo descrita conforme a equação (4.5). O conjunto de equações que descrevem o nível 2, por sua vez, torna-se mais complexo, pois existem covariáveis do nível agregado envolvidas na análise, o que resulta no seguinte conjunto de equações do nível 2:

$$\beta_{0j} = \gamma_{00} + \gamma_{01} W_j + u_{0j} \qquad (4.12)$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11} W_j + u_{1j} \qquad (4.13)$$

Sendo:

W_j o vetor de variáveis independentes medidas no nível 2;

β_{0j} e β_{1j} os coeficientes do intercepto e da variável independente do nível 1, assumidos aleatórios;

u_{0j} e u_{1j} os efeitos aleatórios do nível 2.

Substituindo esses três conjuntos de equação em (4.5), tem-se:

$$y_{ij} = \gamma_{00} + \gamma_{01} W_j + \gamma_{10} X_{1ij} + \gamma_{11} W_j X_{1ij} + u_{1j} X_{1ij} + u_{0j} + e_{ij} \qquad (4.14)$$

A variância do segundo nível é agora residual ou condicional, pois representa a variabilidade de β_{00} e β_{1j} depois de controlada pelas variáveis independentes do segundo nível. Se essas variáveis conseguissem captar toda a variabilidade do intercepto e da inclinação existente no nível superior, não sobraria nenhuma parte aleatória, porém, dificilmente consegue-se incluir variáveis que expliquem toda a aleatoriedade.

Os modelos descritos acima são hierárquicos lineares, e segundo BRYK & RAUDENBUSH (2002), são adequados para dados amostrados em conglomerados quando: i) o resultado esperado de cada nível pode ser representado como função linear dos coeficientes de regressão; e ii) o efeito aleatório de cada nível pode razoavelmente ser assumido com distribuição normal.

As variáveis respostas utilizadas nesse trabalho são binárias, assumindo valores 0 ou 1. Nesse caso, o uso de um modelo hierárquico linear não é apropriado, pois o termo de erro do nível 1 não é normalmente distribuído e não possui variância

homogênea. Segundo BRYK & RAUDENBUSH (2002), isso ocorre porque dado o valor predito, o erro do nível 1 só pode assumir um de dois valores ($e_{ij} = 1 - \beta_0 - \beta_1 X_{kij}$, quando $y_i = 1$; e $e_{ij} = - \beta_0 - \beta_1 X_{kij}$, quando $y_i = 0$) e a variância dependerá do valor predito (considerando o valor predito em termos da probabilidade do evento ocorrer, tem-se que valores próximos de 0 e 1, possuem menores variâncias e próximos de 0,25, maiores)[10].

Outro problema é que no modelo linear não há restrição para o valor predito, o qual pode assumir qualquer valor real. Entretanto, o valor predito para a variável binária, se visto como probabilidade do evento ocorrer, só poderá assumir valores entre zero e um.

A forma mais comum de tratar variáveis binárias num modelo hierárquico é por meio do modelo *logit* hierárquico, em que o valor predito é transformado no logaritmo da chance de sucesso, conhecido como *log-odds*. A equação final do modelo *logit* hierárquico é dada por:

$$\eta_{ij} = \ln\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right) = \gamma_{00} + \gamma_{01}W_j + \gamma_{10}X_{1ij} + \gamma_{11}W_j X_{1ij} + \gamma_{20}X_{2ij} + u_{1j}X_{1ij} + u_{0j} + e_{ij} \quad (4.15)$$

Sendo:

η_{ij} o log da chance (*log-odds*) de sucesso para o indivíduo e

π_{ij} a probabilidade de sucesso para o indivíduo i.

O valor predito do *log-odds* pode ser convertido para o valor predito da probabilidade, conforme equação abaixo:

$$\pi_{ij} = \frac{1}{1 + \exp\{-\eta_{ij}\}} \quad (4.16)$$

Se a probabilidade de sucesso, π_{ik}, é 0,5, o *odds* do sucesso será 1,00 e o *log-odds* igual a zero. Quando a probabilidade de sucesso é menor que 0,5, o *odds* é menor que 1 e o *log-odds* é negativo. Para probabilidades de sucesso maior que 0,5, o *odds* é maior que 1 e o *log-odds* é positivo (BRYK & RAUDENBUSH, 2002).

[10] Para maiores detalhes sobre regressão logística ver ALLISON, P. D., 1999, e para o caso específico do modelo hierárquico ver BRYK & RAUDENBUSH, 2002.

4.2) Análise espacial

A análise espacial - estudos quantitativos dos fenômenos localizados no espaço – lida com duas formas de efeitos relacionados ao espaço: 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, é a instabilidade estrutural nas médias e variância das variáveis no espaço. A econometria espacial trabalha com estes dois efeitos espaciais, incorporando-os nos modelos econométricos.

Apesar de outros modelos, como o hierárquico, tratarem da heterogeneidade, de acordo com ANSELIN (1999), é importante estudá-la em um contexto espacial porque a heterogeneidade normalmente ocorre juntamente com a autocorrelação espacial. Somado a isso, a heterogeneidade e a autocorrelação espacial podem mostrar efeitos confundidos.

Os tipos de análise espacial a serem empregados nesse estudo são a ESDA (Exploratory Spatial Data Analysis – Análise Exploratória Espacial) e a Econometria Espacial.

A ESDA é uma esfera da análise exploratória que permite descrever e visualizar as distribuições espaciais, detectando padrões de associação espacial e formas de heterogeneidade espacial nas variáveis. A econometria espacial trabalha com estes dois efeitos espaciais, incorporando-os nos modelos econométricos.

4.2.1) Análise Exploratória Espacial

As técnicas de ESDA podem ser divididas em medidas de autocorrelação espacial global e indicadores de associação espacial local, conhecidos como LISA.

As estatísticas globais de autocorrelação espacial mais conhecidas são o I de Moran e o C de Geary. Elas são semelhantes às estatísticas clássicas de autocorrelação não espacial, sendo a principal diferença a inclusão de uma matriz de pesos espaciais que representa a interação entre as unidades georeferenciadas. .

No caso do presente trabalho, a matriz de pesos espacial utilizada foi a de contigüidade binária de fronteiras comuns[11]. Quando se utiliza essa matriz de peso, a estatística global apropriada é a estatística I de Moran, cuja fórmula é dada por:

$$I = \frac{N}{S_0} \frac{\sum_i \sum_j W_{ij}(x_i - \mu)(x_j - \mu)}{\sum (x_i - \mu)^2} \qquad (4.17)$$

Onde:

N é o número de observações;

Wij é o elemento da Matriz W correspondente ao par i e j;

xi e xj são os valores das variáveis nos locais i e j;

S_0 é a soma de todos os pesos ($S_0 = \sum_i \sum i w_{ij}$) e para matrizes padronizadas S_0 = N;

μ é a média da variável x.

A inferência do teste I é feita com base nos valores padronizados de z, cujo cálculo é z = (I – E(I) / SD(I)), onde: E(I) é a média teórica da estatística I e SD(I) é o seu desvio padrão teórico. A média teórica, E(I)= –1/(N-1), enquanto que o desvio padrão teórico depende das hipóteses aleatórias. Assume-se que a estatística I é assintoticamente normalmente distribuída. Para o cálculo da variância de I podem ser feitas três pressuposições sobre a sua distribuição de probabilidade , quais sejam: i) valores seguem distribuição normal; ii) uniformidade, ou seja, assume-se que cada valor observado é igualmente provável de ocorrer em qualquer localização; iii) permutação, semelhante à anterior porém, com média e distribuição geradas através de permutações, que foi a utilizada nessa tese.

[11] A matriz de pesos de contigüidade binária de fronteiras comuns é positiva e simétrica com N x N elementos. Os elementos w_{ij}, correspondem a cada par de observações ij. Os elementos não-zero da matriz wij são os vizinhos e os elementos iguais a zero são os não vizinhos. Por convenção a diagonal é zero, ou seja, a localização i não é vizinha dela mesma.
Formalizando tem-se:

$$W = \sum_{j=1,..,n} w_{ij}$$

Onde: $w_{ij} = 1$, quando i e j são vizinhos;
$w_{ij} = 0$, quando i e j não são vizinhos;
$w_{ii} = 0$, por convenção.

Na interpretação do I de Moran, valores significativos e positivos indicam autocorrelação espacial positiva e valores negativos e significativos indicam autocorrelação espacial negativa.

Entretanto, medidas de autocorrelação espacial global resumem em um único valor a existência de autocorrelação espacial, podendo encobrir ou mesmo levar a conclusões enganosas no caso da existência de forte heterogeneidade espacial. Somado a isso, essas medidas não permitem detectar padrão local de autocorrelação espacial. Para este fim, são utilizadas as medidas de autocorrelação espacial local conhecidas como LISA – Local Indicators of Spacial Association. Elas são importantes porque além de permitirem identificar a heterogeneidade espacial, também indicam autocorrelacão local mesmo quando as medidas globais não apontam.

As principais medidas LISA são o *Moran Scatterplot* e o I de Moran local. O primeiro é a representação gráfica do I de Moran, que pode ser interpretado através de uma regressão linear de y em Wy, onde W é a matriz de peso espacial. A inclinação dessa regressão, ρ, é um indicador da extensão da autocorrelação espacial entre cada observação e a média dos seus vizinhos. Seu valor corresponde à estatística global do I de Moran e sua representação gráfica ao *Moran Scatterplot.* No gráfico do *Moran Scatterplot*, os 1º e 3º quadrantes representam a autocorrelação espacial positiva, ou seja, observações com altos (baixos) valores são rodeados por altos (baixos) valores para a média dos seus vizinhos. Já os 2º e 4º quadrantes indicam autocorrelação espacial negativa, valores altos (baixos) rodeados por valores baixos (altos) para a média dos vizinhos. O 1º e 3º quadrantes sugerem, portanto, a presença de autocorrelação espacial enquanto que os outros dois, são indicativos de heterogeneidade espacial. A outra estatística local, I de Moran Local, é semelhante a global, porém calculadas para cada área i.

4.2.2) Econometria Espacial

Os principais modelos de regressão espacial são o autoregressivo de defasagem espacial e o de erro espacial. O primeiro modelo é o processo autoregressivo da variável dependente. Sua fórmula é dada por:

$$y = \rho Wy + X\beta + \varepsilon \qquad (4.18)$$

Sendo:

ρ é o parâmetro autoregressivo espacial;

W é a matriz de peso;

X é a matriz de variáveis independentes;

β é a matriz de coeficientes da regressão e

ε é o vetor do termo de erro aleatório.

A equação (4.18) 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:

$$y = (I - \rho W)^{-1} X\beta + (I - \rho W)^{-1} \varepsilon \qquad (4.19)$$

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 ε de todas as localizações no sistema, não apenas ao erro em *i* (ANSELIN, 2003). Desta forma, a equação (4.19) é importante por mostrar que o efeito espacial opera através das covariáveis X e do termo de erro não observado, ou seja, de fatores não observados que não entraram no modelo.

O modelo de erro espacial indica que há correlação espacial em variáveis que não foram incluídas no modelo. Esse modelo pode ser definido por dois processos, o autoregressivo espacial (SAR) e o de média móvel espacial (SMA). O primeiro processo indica um processo global nos erros e sua fórmula estrutural é dada por:

$$y = X\beta + \lambda W\varepsilon + u \qquad (4.20)$$

Sendo:

λ é o coeficiente autoregressivo;

ε é o termo de erro que possui dependência espacial e

u é o termo de erro aleatório.

A forma reduzida é dada por:

$$y = X\beta + (I - \lambda W)^{-1} u \qquad (4.21)$$

O modelo de média móvel espacial do erro pressupõe que os efeitos espaciais das variáveis omitidas se limitam aos vizinhos mais próximos. Tanto a forma estrutural quanto a reduzida é dada por:

$$y = X\beta + \lambda Wu + u \qquad (4.22)$$

Além dos modelos de defasagem e erro espacial existem modelos mais complexos como o SARMA (S*patial Autoregressive Moving Average Process).* Este modelo considera que há um processo autoregressivo na variável dependente, ou seja, efeito global, e um de média móvel no erro, efeito local[12].

Através dos testes de diagnósticos disponíveis nos programas computacionais é possível detectar se a dependência espacial é melhor capturada pelo modelo de defasagem espacial, erro espacial ou SARMA. Nessa tese, utilizou-se o teste do Multiplicador de Lagrange (LM) para fazer tal diagnóstico[13]. Esse teste, além de permitir a distinção entre os modelos de erro espacial e defasagem espacial, possuem uma classe de testes robustos para outros tipos de má especificação local.

Os modelos espaciais podem ser estimados por três métodos: máxima verossimilhança, variáveis instrumentais (ou mínimos quadrados de dois estágios) e métodos generalizado dos momentos[14] (GMM).

Importante ressaltar, que quando se ignora a autocorrelação espacial e heterogeneidade espacial dos dados, os parâmetros estimados nas regressões podem ser enviesados e ineficientes, se o processo em curso for de defasagem espacial, e apenas ineficientes, se o modelo detectado for de erro espacial.

4.3) Modelo hierárquico-espacial

Como já mencionado, essa tese busca desenvolver uma metodologia que combina os modelos hierárquico e espacial. A vantagem de combinar essas duas metodologias é que a primeira permite analisar conjuntamente as variáveis relacionadas aos indivíduos e as contextuais em níveis diferentes de agregação, diminuindo o problema relacionado à falácia ecológica, além de possibilitar analisar como as

[12] Para maiores detalhes dos modelos de ordens maiores ver: ANSELIN & BERA (1998).
[13] Para maiores detalhes dos testes ver: ANSELIN et alii (1996) e ANSELIN (1988b).
[14] Para maiores detalhes dos métodos de estimação ver ANSELIN (1999).

variáveis interagem entre os diferentes níveis. Já o modelo espacial leva em conta a possível autocorrelação espacial nas variáveis dependentes e independentes e/ou nos termos de erro no segundo nível do modelo hierárquico. Acredita-se que a introdução do caráter espacial das variáveis no modelo hierárquico possibilitará obter estimativas mais consistentes e eficientes dos parâmetros nas regressões uma vez que a presença de autocorrelação espacial nas variáveis do modelo hierárquico incorre na quebra de um pressuposto importante desse modelo que é o de erros i.i.d. no nível mais alto (nível 2). De fato, como será mostrado no Capitulo 6, as variáveis analisadas apresentaram forte autocorrelação e heterogeneidade espacial.

Este campo de estudo está começando a ser desenvolvido, com a utilização de métodos *baeysianos*, que utilizam o programa computacional WinBUGS juntamente com MLwiN. Porém, a proposta aqui é desenvolver uma técnica utilizando os programas HLM e SpaceStat. Tentativa semelhante foi desenvolvida no trabalho de MORENOFF (2003) para o estudo do peso da criança ao nascer. Desta forma, esse trabalho será o ponto de partida para a metodologia da presente análise.

MORENOFF (2003) parte da equação estrutural do modelo de defasagem espacial - equação (4.18), a qual pressupõe que os efeitos espaciais operam através de um processo auto-regressivo de defasagem espacial da variável dependente. Porém, o modelo assim representado possui uma variável endógena, WY, que só pode ser estimada através do método de máxima verossimilhança, variáveis instrumentais ou método dos momentos. A alternativa encontrada por MORENOFF para introduzir o efeito espacial na análise hierárquica foi a inclusão de defasagens espaciais das variáveis independentes[15] no modelo hierárquico, ou seja, a inclusão de WX.

Como pode ser observado na equação (4.19), que é a forma reduzida do modelo de defasagem espacial, o processo espacial opera através das covariáveis X e dos termos de erro, em outras palavras, existe uma dependência espacial nas variáveis observadas e não-observadas. A estratégia adotada por MORENOFF permite corrigir a dependência espacial apenas das variáveis observadas

A estratégia adotada nessa tese consiste em três passos. Em primeiro lugar são obtidos os resíduos do segundo nível do modelo hierárquico não-condicional, ou seja,

[15] A defasagem espacial de determinada variável na localidade j é a sua média nos seus vizinhos. Os vizinhos são determinados pela matriz de peso espacial.

quando não se inclui variáveis de nível 2 na análise, apenas as de nível 1. Esses resíduos são os termos u_{0j} e u_{kj} das equações (4.9) e (4.10), respectivamente.

Em seguida são feitos testes econométricos espaciais nesses resíduos. Através desses testes é possível perceber se o efeito espacial opera na forma de defasagem ou de erro espacial. Os resultados desses testes mostraram que na maioria dos casos o processo espacial ocorre mais intensamente na forma de defasagem espacial.

Na terceira etapa é estimado o modelo hierárquico-espacial, com duas especificações diferentes. A primeira consiste em acrescentar a defasagem dos resíduos descritos acima no segundo nível do modelo hierárquico, ou seja, nas equações (4.12) e (4.13). Dessa forma, se está corrigindo a dependência espacial tanto das variáveis observadas quanto das não-observadas.

A segunda especificação consiste em incluir no nível 2 da regressão hierárquica a 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, apenas das observadas. Porém, ela é útil por permitir observar qual covariável é mais espacialmente correlacionada com as varáveis independentes.

4.4) Conclusão

Nesse capítulo foram apresentados os modelos estatísticos utilizados nessa tese, que são os modelos hierárquico, espacial e hierárquico-espacial. O primeiro modelo permite analisar os dados no nível micro (indivíduos) e macro (municípios), ou seja, as variáveis individuais e contextuais são analisadas em níveis diferentes. Dessa forma, é possível calcular a variação observada nos dois níveis de análise. Em outras palavras, é possível saber se a diferença no resultado educacional do indivíduo em relação à média é decorrente das características do próprio indivíduo ou das características municipais, ou de ambas. Outra vantagem relevante dos modelos hierárquicos, para essa tese, é possibilitar o estudo de interações entre as variáveis do nível micro e macro.

A abordagem espacial, será útil na medida em que permite diagnosticar a existência de dependência espacial nos dados agregados, o que possibilitará incluir na análise dos determinantes educacionais não apenas os fatores contextuais internos (do

próprio município), mas também os fatores contextuais mais amplos, ou seja, das localidades adjacentes.

O modelo hierárquico-espacial combina as duas metodologias anteriores. A vantagem desse modelo é permitir analisar a influências dos fatores de *background* familiar e contextuais (variáveis municipais) em seus respectivos níveis, além de incluir as "externalidades espaciais", isto é, a influência das variáveis contextuais dos municípios vizinhos.

5) ANÁLISE HIERÁRQUICA DOS DETERMINANTES EDUCACIONAIS

Nesse capítulo é feita a análise dos resultados do modelo hierárquico para os indicadores educacionais investigados nessa tese. Para facilitar a análise, os resultados serão comentados separando as variáveis dependentes em seções distintas. Dessa forma, serão primeiramente abordados os resultados para a probabilidade de freqüentar a escola, ou seja, os determinantes do acesso ao sistema de ensino. Em seguida, comenta-se os resultados para a probabilidade de freqüentar a escola na idade adequada, e por último, as progressões. Essas duas últimas dizem respeito ao rendimento e eficiência do ensino. As duas dimensões do ensino – acesso e eficiência – possuem dinâmicas diferentes, portanto, os fatores mais importantes que determinam cada uma delas são de natureza distintas.

5.1) Probabilidade de freqüentar a escola

Para o acesso das crianças e jovens ao sistema de ensino é provável que alterações no nível da demanda e oferta escolar sejam os fatores que mais o influenciam. No nível micro, as variáveis de *background* familiar exercem influencia 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. Tal fato é relevante nos dois níveis de ensino, no fundamental por estar praticamente universalizado e pressões demográficas podem ter uma maior influencia; no ensino médio, por possuir grande restrição de oferta.

Dessa forma, para essa dimensão foram estimados 5 modelos, conforme pode ser visto nas TAB. 5.1 – ensino fundamental - e TAB. 5.2 – ensino médio, todos com o pressuposto de que apenas o intercepto possui efeito aleatório. 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 foram incluídas 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)[16].

A razão para incluir separadamente as variáveis de nível dois mais relevantes é a possibilidade de verificar em 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:

$$\%\ \text{da variância explicada} = \frac{\hat{\sigma}_{00(não-condicional)} - \hat{\sigma}_{00(condicional)}}{\hat{\sigma}_{00(não-condicional)}} \qquad (5.1)$$

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 (TAB. 5.1). Através da análise do efeito aleatório do modelo 1 (parte inferior da tabela) percebe-se que é aceitável a hipótese do intercepto ser considerado como tendo efeito aleatório, já que ele foi significativo. Em outras palavras, é aceitável que os municípios apresentem valores distintos para a probabilidade de freqüentar 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 (39,70%), enquanto que 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 uma relação negativa com a freqüência escolar, e que mulheres, brancos e amarelos e moradores da área urbana tem maior probabilidade de freqüentar a escola.

[16] Na TAB. A1.1 do Anexo 1 encontram-se modelos de regressão com todas as variáveis independentes consideradas nesse estudo, exceto o tamanho relativo da coorte que possui forte correlação com a maioria das demais variáveis. Através deles percebeu-se que as variáveis de nucleação e educação média da população adulta possuíam significativo impacto no ensino fundamental. A opção de incluir apenas as variáveis de nucleação, se deve ao fato da educação média da população adulta possuir forte correlação com o tamanho relativo da coorte (-0,720). O modelo que inclui essa última variável (além das do modelo 5) também se encontra nessa tabela do anexo.

TABELA 5. 1: Resultado da regressão para a probabilidade de freqüentar a escola – indivíduos de 7 a 14 anos de idade

Variáveis Independentes	Modelo 1		Modelo 2		Modelo 3		Modelo 4		Modelo 5	
Efeito Fixo	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig
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 a14					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 1000							1.636	0.000	1.111	0.000
Proporção de estabelecimentos com mais de 1000 matrículas							0.723	0.000	0.924	0.000
População até 4999									-0.010	0.868
População entre 5000 e 9999									-0.028	0.648
População entre 10000 e 19999									-0.051	0.399
População entre 20000 e 49999									-0.103	0.087
População entre 50000 e 99999									-0.003	0.959
População entre 100000 e 199999									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 nível superior	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 nível médio	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
Efeito Aleatório										
Coeficiente	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	

Com relação aos fatores de *background* familiar, observa-se que o nível educacional da mãe é mais importante que as demais variáveis, o que corrobora com os resultados dos estudos vistos anteriormente. Para melhor ilustrar o impacto desses fatores, a TAB. 5.4 contém 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 o aumento de 1,97% na probabilidade das crianças de 7 a 14 anos freqüentarem a escola, enquanto que o efeito das demais variáveis é menor. Cabe ressaltar, que o pequeno aumento dessa probabilidade, nesse nível de ensino, se deve ao fato dele 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 etárias provocam um aumento na probabilidade de freqüentar a escola numa magnitude próxima ao verificado na simulação da variação na educação materna. Por outro lado, a restrição de oferta, embora apresente coeficiente significativo, tem impacto menor. Os GRAFs. 5.1 e 5.2 ilustram o impacto de variações na probabilidade de freqüentar a escola para os indivíduos de 7 a 14 anos decorrente de variações no tamanho relativo da coorte e restrição de oferta. A partir deles percebe-se mais claramente o maior impacto da razão de dependência de jovens, evidenciando que o acesso das crianças de 7 a 14 anos a escola é beneficiado por um maior dividendo demográfico, proveniente de uma redução da fecundidade.

GRÁFICO 5. 1: Impacto da pressão demográfica e da restrição de oferta na probabilidade de freqüentar a escola para os indivíduos de 7 a 14 anos, considerando os municípios com menor e maior valores.

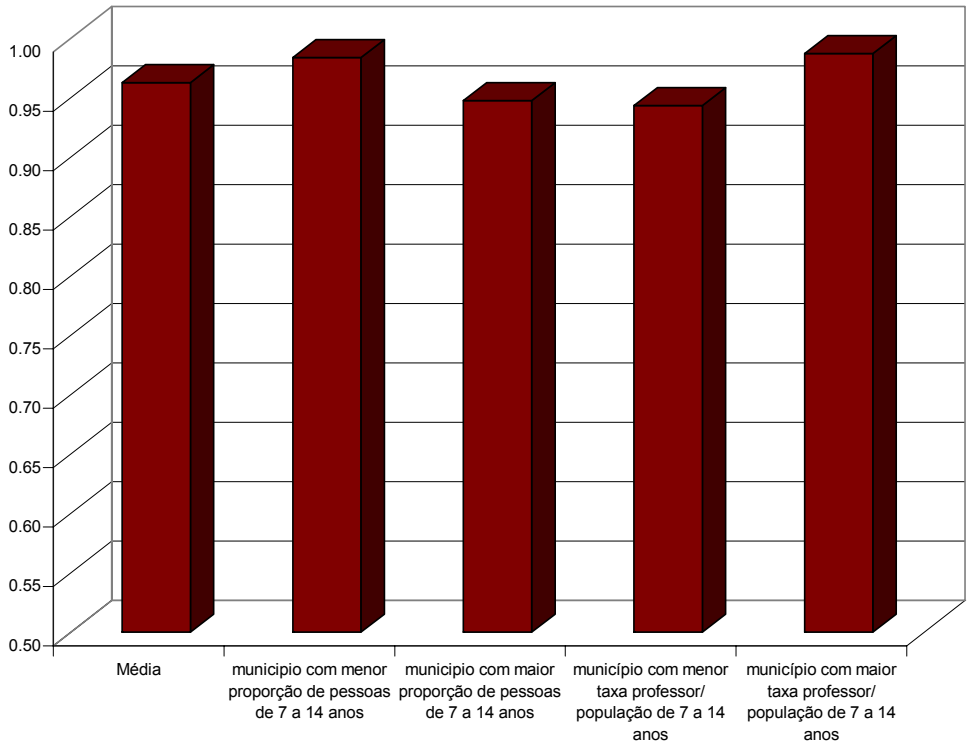
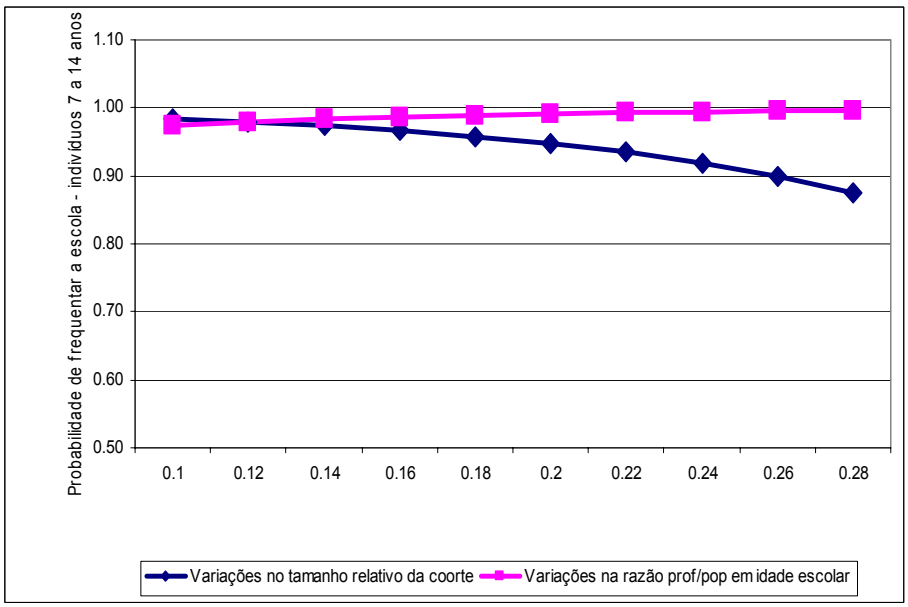


GRÁFICO 5. 2: Impacto da pressão demográfica e da restrição de oferta na probabilidade de freqüentar a escola para os indivíduos de 7 a 14 anos



Os resultados para os indivíduos de 15 a 17 anos se encontram na TAB. 5.2. Através deles percebe-ser que para esse nível de ensino também é aceitável a hipótese do intercepto ser considerado com efeito aleatório. Quando se acrescenta 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 as variáveis de nível 1, os resultados são muito semelhantes aos encontrados para os indivíduos de 7 a 14 anos, as principais diferenças são: i) impacto negativo da idade, sexo e famílias chefiadas por mulher mais forte; ii) as variáveis de posição da ocupação do chefe da família possuem maior impacto, sendo o efeito positivo maior para as com maior *status* socioeconômico e negativo para chefe fora do mercado de trabalho; iii) família convivente teve impacto positivo, inverso do ocorrido no caso anterior.

Através da TAB. 5.4 observa-se que no ensino médio a variação decorrente das simulações[17] de mudanças nas variáveis de nível 1 é bem maior que no fundamental, principalmente a educação materna. Tal fato é atribuído a 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.

[17] Para o ensino médio as simulações foram feitas considerando os coeficientes do modelo 6.

TABELA 5. 2: Resultado da regressão para a probabilidade de freqüentar a escola – indivíduos de 15 a 17 anos de idade

Variáveis Independentes	Modelo 1		Modelo 2		Modelo 3		Modelo 4		Modelo 5		Modelo 6	
Efeito Fixo	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig
Intercepto	1.440	0.000	1.912	0.000	1.141	0.000	1.518	0.000	1.000	0.000	1.113	0.000
Tamanho relativo da coorte de 15 a 17 anos			-7.081	0.000					9.421	0.000	9.241	0.000
Razão professor/pop. 15 a 17					8.115	0.000			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 1000							-0.098	0.081	0.081	0.119		
Proporção de estabelecimentos com mais de 1000 matrículas							0.637	0.000	0.416	0.000		
População até 4999									-0.693	0.000	-0.788	0.000
População entre 5000 e 9999									-0.725	0.000	-0.781	0.000
População entre 10000 e 19999									-0.665	0.000	-0.711	0.000
População entre 20000 e 49999									-0.609	0.000	-0.635	0.000
População entre 50000 e 99999									-0.406	0.000	-0.413	0.000
População entre 100000 e 199999									-0.191	0.003	-0.197	0.003
Idade	-0.441	0.000	-0.441	0.000	-0.441	0.000	-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	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	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	-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	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	-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	0.069	0.003	0.069	0.003	0.069	0.003
Categoria de ocupação do chefe nível superior	0.479	0.000	0.481	0.000	0.483	0.000	0.479	0.000	0.485	0.000	0.485	0.000
Categoria de ocupação do chefe nível médio	0.141	0.000	0.141	0.000	0.141	0.000	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	-0.048	0.000	-0.047	0.000	-0.047	0.000
Efeito Aleatório												
Coeficiente	0.294	0.000	0.290	0.000	0.262	0.000	0.282	0.000	0.220	0.000	0.223	0.000
% da variância explicada			1.43		10.88		3.93		25.03		24.08	

Partindo para a análise das variáveis municipais, percebe-se que, para esse nível de ensino, a restrição de oferta tem impacto positivo importante. Quando se considera o município com maior número de professores do ensino médio por população de 15 a 17 anos, a probabilidade de freqüentar a escola chega a quase 100%, conforme pode ser visto no GRAF. 5.3. Destaca-se também o sinal positivo encontrado para o tamanho relativo da coorte em idade escolar, resultado não esperado. Porém, quando considera apenas essa variável de nível 2 na análise (modelo 2) seu sinal torna-se negativo. Através da TAB. A1.4 do Anexo 1, 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. Como vai ser visto mais à frente, o coeficiente do tamanho relativo da coorte nas regressões para a freqüência das pessoas de 15 a 17 anos que já concluíram o ensino fundamental é positivo, mesmo quando se considera apenas o tamanho relativo da coorte. O que corrobora essa hipótese já que a cobertura dos jovens que já concluíram o ensino básico está em um nível bem mais baixo.

GRÁFICO 5. 3: Impacto da pressão demográfica e da restrição de oferta na probabilidade de freqüentar a escola para os indivíduos de 15 a 17 anos, considerando os municípios com menor e maior valores.

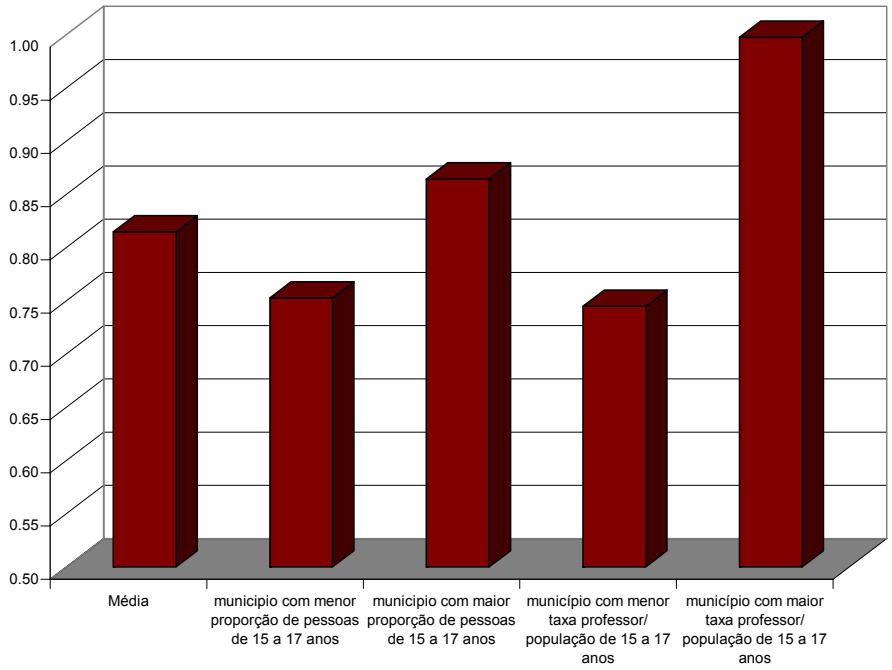
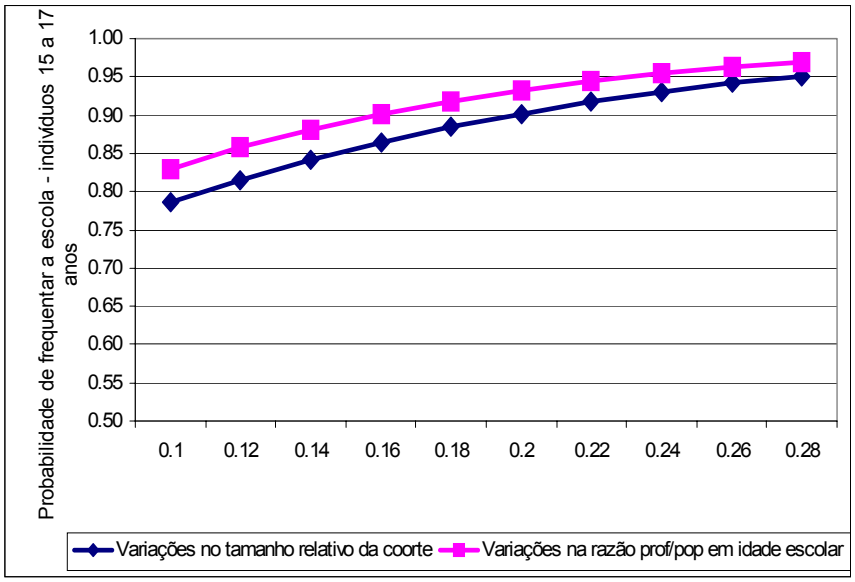


GRÁFICO 5. 4: Impacto da pressão demográfica e da restrição de oferta na probabilidade de freqüentar a escola para os indivíduos de 15 a 17 anos



Como já mencionado, devido à ambigüidade da taxa de atendimento das pessoas de 15 a 17 anos, visto no Capítulo 1, foi também analisada a probabilidade de freqüentar a escola das pessoas de 15 a 17 anos que já concluíram o ensino fundamental, na tentativa de captar mudanças na taxa de escolarização líquida do ensino médio e ver até que ponto os resultados são muito diferentes do anterior. Quando essa probabilidade é

analisada (TAB 5.3), observa-se que os resultados para as variáveis de *background* familiar e características do indivíduo são muito parecidos com o da freqüência dos jovens de 15 a 17 anos não-condicional, não cabendo, portanto, tecer maiores comentários. Já para as variáveis contextuais, percebe-se uma diminuição do impacto da restrição de oferta e um pequeno aumento do efeito da pressão demográfica, GRAF. 5.5, ressaltando que essa última continua com o sinal não esperado, ou seja, positivo.

TABELA 5. 3: Resultado da regressão para a probabilidade de freqüentar a escola – indivíduos de 15 a 17 anos de idade que já concluíram o fundamental

Variáveis Independentes	Modelo 1		Modelo 2		Modelo 3		Modelo 4		Modelo 5		Modelo 6	
	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig
Efeito Fixo												
Intercepto	2.306	0.000	1.934	0.000	2.156	0.000	2.492	0.000	1.363	0.000	1.377	0.000
Tamanho relativo da coorte de 15 a 17 anos			5.760	0.000					18.249	0.000	18.258	0.000
Razão professor/pop. 15 a 17					3.669	0.000			8.185	0.000	8.152	0.000
Proporção de estabelecimentos com 50 a 200 matrículas							-0.233	0.013	-0.059	0.514		
Proporção de estabelecimentos com 201 a 1000							-0.246	0.008	-0.019	0.832		
Proporção de estabelecimentos com mais de 1000 matrículas							0.576	0.000	0.250	0.059		
População até 4999									-0.700	0.000	-0.761	0.000
População entre 5000 e 9999									-0.745	0.000	-0.787	0.000
População entre 10000 e 19999									-0.626	0.000	-0.663	0.000
População entre 20000 e 49999									-0.508	0.000	-0.524	0.000
População entre 50000 e 99999									-0.288	0.000	-0.292	0.000
População entre 100000 e 199999									-0.164	0.046	-0.168	0.042
Idade	-0.391	0.000	-0.391	0.000	-0.391	0.000	-0.391	0.000	-0.391	0.000	-0.391	0.000
Situação de domicílio; 1=urbano e 0=rural	0.689	0.000	0.689	0.000	0.688	0.000	0.687	0.000	0.681	0.000	0.681	0.000
Cor; 1=branco e amarelo e 0=preto e pardo	0.184	0.000	0.184	0.000	0.184	0.000	0.184	0.000	0.185	0.000	0.185	0.000
Sexo; 1=homem e 0=mulher	-0.327	0.000	-0.327	0.000	-0.327	0.000	-0.327	0.000	-0.327	0.000	-0.327	0.000
Educação da mãe	0.121	0.000	0.121	0.000	0.121	0.000	0.121	0.000	0.121	0.000	0.121	0.000
Domicílio chefiado por mulher	-0.361	0.000	-0.361	0.000	-0.361	0.000	-0.361	0.000	-0.362	0.000	-0.362	0.000
Família convivente	0.080	0.100	0.081	0.099	0.080	0.102	0.080	0.101	0.081	0.100	0.081	0.099
Categoria de ocupação do chefe nível superior	0.454	0.000	0.453	0.000	0.455	0.000	0.455	0.000	0.458	0.000	0.458	0.000
Categoria de ocupação do chefe nível médio	0.187	0.000	0.187	0.000	0.187	0.000	0.187	0.000	0.187	0.000	0.187	0.000
Chefe não ocupado	-0.011	0.430	-0.011	0.408	-0.010	0.463	-0.010	0.438	-0.010	0.456	-0.010	0.456
Efeito Aleatório												
Coeficiente	0.445	0.000	0.447	0.000	0.431	0.000	0.423	0.000	0.344	0.000	0.346	0.000
% da variância explicada			-0.50		3.02		4.94		22.66		22.31	

GRÁFICO 5. 5: Impacto da pressão demográfica e da restrição de oferta na probabilidade de freqüentar a escola para os indivíduos de 15 a 17 anos que já concluíram o fundamental, considerando os municípios com menor e maior valores.

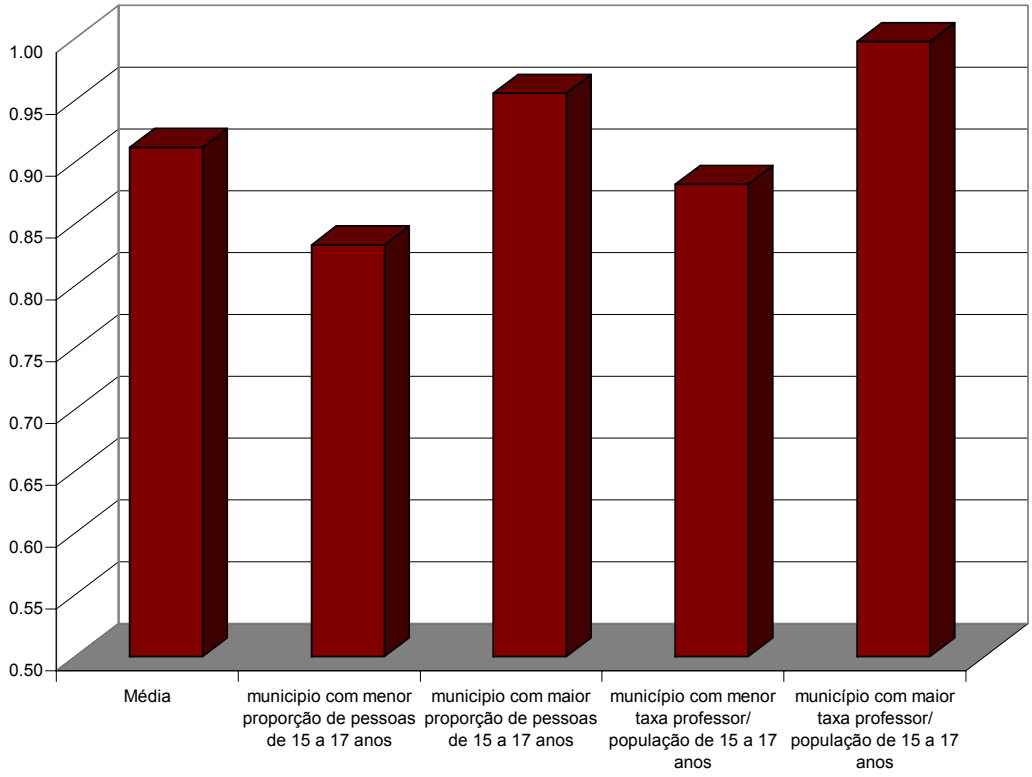


TABELA 5. 4: Variações na probabilidade de freqüentar a escola segundo variações nas variáveis de *background* familiar para diferentes amostras*

Variáveis Independentes	Indivíduos de 7 a 14 anos		Indivíduos de 15 a 17 anos		Indivíduos de 15 a 17 anos que já concluíam o ensino fundamental	
	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

Nota : * Utilizou-se os coeficientes do modelo 5 para a amostra de 7 a 14 anos e o modelo 6 para as demais.

5.2) Probabilidade de freqüentar a escola na idade adequada

Nessa dimensão do ensino - eficiência escolar – além dos fatores municipais que afetam a demanda e oferta, ou seja, o dividendo demográfico e a restrição de oferta, os demais, relacionados às características da rede de ensino do município, principalmente qualidade, também podem ser relevantes.

Dessa forma, a análise do impacto dos fatores contextuais nessa dimensão será pautada em duas direções. Primeiro, será focalizado o efeito dos fatores de demanda e oferta, e em seguida o impacto dos demais fatores relacionados à rede escolar do município, principalmente os referente à qualidade dos recursos humanos e infra-estrutura escolar.

No nível micro, pode-se supor também, que para a eficiência escolar, os fatores relacionados ao ambiente familiar possuem maior importância que para o acesso a escola, e que seus efeitos não são os mesmos em todos os municípios. Assim, os modelos de regressões para essa dimensão consideraram que tanto o intercepto quanto a inclinação não são os mesmo para os municípios. Para definir qual coeficiente das variáveis independentes do primeiro nível possui efeito aleatório, utilizou como base o artigo de RIOS-NETO, CÉSAR E RIANI (2002), que testa a educação materna e paterna como tendo efeitos aleatórios e seus resultados apontam para uma variância aleatória do coeficiente ligado a educação da mãe. A vantagem de considerar que há variação no coeficiente da educação materna é que possibilita analisar a sua interação com as variáveis independentes do nível 2 na probabilidade de cursar a escola na idade adequada. Assim, além de avaliar a importância da educação materna, será possível verificar até que ponto seu efeito pode ser diminuído pela melhora na rede escolar dos municípios.

Foram estimados quatro modelos de regressão. O Modelo 1 é o não-condicional, ou seja, não possui covariável de nível 2. O modelo 2 inclui as covariáveis relacionadas ao dividendo e restrição de oferta, que são respectivamente o tamanho relativo da coorte e a razão professor/população em idade escolar, além das *dummies* para o tamanho da população. O modelo seguinte, inclui as demais covariáveis de nível 2. Optou-se também por estimar mais um modelo que exclui o tamanho relativo da coorte - Modelo 4. A exclusão dessa variável se deve basicamente a dois motivos. Primeiro, pela sua alta correlação com as variáveis relacionadas aos fatores de qualidade da rede escolar dos municípios, conforme pode ser visto nas TABs. A2.1 a A2.2 do Anexo 2. Segundo porque ela também está associada com o nível socioeconômico do município, dessa forma, além de captar o verdadeiro efeito de variações na razão de dependência de jovens, também pode captar diferenças socioeconômicas dos municípios. Dessa forma, no modelo 4, a restrição de oferta passa a ser a síntese da oferta e demanda, já que no seu denominador contém a população em idade escolar.

Os resultados das regressões para o ensino fundamental e médio encontram-se nas TABs. 5.5 e 5.6. A análise do efeito aleatório nos dois níveis de ensino mostra que é aceitável a hipótese de considerar o intercepto e a educação materna como tendo efeitos aleatórios, ressaltando que para essa última o componente da variância é pequeno. Comparando os dois níveis de ensino, percebe-se que o efeito aleatório é maior para o ensino médio que o fundamental, tanto para o intercepto quanto para a educação materna. Quando todas as variáveis de nível 2 são incluídas, a variabilidade do intercepto é explicada em mais de 80% nos dois níveis de ensino. Para a variabilidade do coeficiente da educação materna, a inclusão das variáveis de nível 2 explicam mais no nível médio (76,67%) que no fundamental (48,28%). Nota-se também que a exclusão do tamanho relativo da coorte em idade escolar não diminui a porcentagem da variância explicada na educação materna do fundamental.

Após a análise do efeito aleatório, considera-se o resultado do efeito fixo. Com relação as variáveis relacionadas às características e ambiente familiar do indivíduo, percebe-se que seus resultados são semelhantes aos encontrados para a probabilidade de freqüentar a escola, destacando o forte impacto da educação materna. Comparando os efeitos diferenciados no ensino fundamental e médio, observa-se que o impacto tanto dos atributos individuais (exceto idade) quanto do *background* familiar é bem maior para o ensino médio.

TABELA 5. 5: Resultado da regressão para a probabilidade de freqüentar a escola na idade correta – indivíduos de 7 a 14 anos de idade que freqüentam a escola

Variáveis Independentes	Modelo 1		Modelo 2		Modelo 3		Modelo 4	
Efeito Fixo	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig
Intercepto	-0.604	0.000	3.775	0.000	-3.130	0.000	-4.520	0.000
Tamanho relativo da coorte de 7 a 14 anos			-22.272	0.000	-5.354	0.000		
Razão professor/pop. 7 a14			-7.386	0.000	-2.556	0.000	-0.800	0.029
Proporção de matrículas em escolas com ciclo					0.188	0.000	0.179	0.000
Proporção de matrículas públicas					1.361	0.000	1.301	0.000
Proporção de matrículas municipais dentre as públicas					-0.189	0.000	-0.173	0.000
Proporção de estabelecimentos com 50 a 200 matrículas					-0.115	0.000	-0.139	0.000
Proporção de estabelecimentos com 201 a 1000					-0.047	0.084	-0.062	0.027
Proporção de estabelecimentos com mais de 1000 matrículas					0.203	0.047	0.164	0.118
Educação média da população adulta					0.219	0.000	0.261	0.000
População até 4999			0.016	0.751	0.343	0.000	0.392	0.000
População entre 5000 e 9999			-0.001	0.982	0.341	0.000	0.381	0.000
População entre 10000 e 19999			-0.095	0.052	0.298	0.000	0.338	0.000
População entre 20000 e 49999			-0.084	0.089	0.244	0.000	0.269	0.000
População entre 50000 e 99999			-0.008	0.888	0.202	0.000	0.218	0.000
População entre 100000 e 199999			0.046	0.475	0.123	0.005	0.135	0.003
Horas-aula diárias					0.277	0.000	0.314	0.000
Alunos por turma					-0.006	0.000	-0.007	0.000
Proporção professores com curso superior					0.401	0.000	0.563	0.000
Fator de infra-estrutura					0.117	0.000	0.135	0.000
Idade	-0.231	0.000	-0.231	0.000	-0.231	0.000	-0.231	0.000
Situação de domicílio; 1=urbano e 0=rural	0.313	0.000	0.313	0.000	0.313	0.000	0.313	0.000
Cor; 1=branco e amarelo e 0=preto e pardo	0.181	0.000	0.181	0.000	0.181	0.000	0.181	0.000
Sexo; 1=homem e 0=mulher	-0.226	0.000	-0.226	0.000	-0.226	0.000	-0.226	0.000
Educação da mãe	0.091	0.000	-0.095	0.000	0.190	0.000	0.210	0.000
Tamanho relativo da coorte de 7 a 14 anos			1.039	0.000	0.076	0.138		
Razão professor/pop. 7 a14			0.394	0.000	0.314	0.000	0.292	0.000
Proporção de matrículas em escolas com ciclo					0.005	0.035	0.006	0.030
Proporção de matrículas públicas					-0.166	0.000	-0.166	0.000
Proporção de matrículas municipais dentre as públicas					0.014	0.000	0.014	0.001
Proporção de estabelecimentos com 50 a 200 matrículas					-0.003	0.562	-0.002	0.634
Proporção de estabelecimentos com 201 a 1000					0.006	0.217	0.006	0.193
Proporção de estabelecimentos com mais de 1000 matrículas					0.031	0.038	0.032	0.034
Educação média da população adulta					-0.013	0.000	-0.014	0.000
População até 4999			-0.036	0.000	-0.040	0.000	-0.041	0.000
População entre 5000 e 9999			-0.027	0.000	-0.033	0.000	-0.033	0.000
População entre 10000 e 19999			-0.018	0.000	-0.027	0.000	-0.027	0.000
População entre 20000 e 49999			-0.012	0.012	-0.020	0.000	-0.020	0.000
População entre 50000 e 99999			-0.010	0.058	-0.012	0.016	-0.012	0.014
População entre 100000 e 199999			-0.006	0.302	-0.005	0.318	-0.005	0.303
Horas-aula diárias					0.024	0.000	0.024	0.000
Alunos por turma					0.001	0.000	0.001	0.000
Proporção professores com curso superior					-0.060	0.000	-0.062	0.000
Fator de infra-estrutura					-0.007	0.000	-0.008	0.000
Domicílio chefiado por mulher	-0.139	0.000	-0.139	0.000	-0.139	0.000	-0.139	0.000
Família convivente	0.009	0.240	0.009	0.246	0.008	0.261	0.008	0.260
Categoria de ocupação do chefe nível superior	0.127	0.000	0.126	0.000	0.128	0.000	0.128	0.000
Categoria de ocupação do chefe nível médio	0.088	0.000	0.089	0.000	0.089	0.000	0.089	0.000
Chefe não ocupado	-0.063	0.000	-0.062	0.000	-0.062	0.000	-0.062	0.000
Efeito Aleatório								
Coeficiente	0.580	0.000	0.226	0.000	0.100	0.000	0.107	0.000
Educação da mãe	0.003	0.000	0.002	0.000	0.001	0.000	0.001	0.000
% da variância explicada - intercepto			60.98		82.79		81.63	
% da variância explicada - educação materna			29.50		48.28		48.28	

TABELA 5. 6: Resultado da regressão para a probabilidade de freqüentar a escola na idade correta – indivíduos de 15 a 17 anos de idade que freqüentam a escola

Variáveis Independentes	Modelo 1		Modelo 2		Modelo 3		Modelo 4	
Efeito Fixo	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig
Intercepto	-1.889	0.000	1.861	0.000	-5.000	0.000	-6.634	0.000
Tamanho relativo da coorte de 15 a 17 anos			-62.633	0.000	-17.548	0.000		
Razão professor/pop. 15 a17			18.595	0.000	4.976	0.000	6.539	0.000
Proporção de estabelecimentos com 50 a 200 matrículas					0.144	0.033	0.164	0.018
Proporção de estabelecimentos com 201 a 1000					0.311	0.000	0.315	0.000
Proporção de estabelecimentos com mais de 1000 matrículas					0.274	0.009	0.208	0.051
Proporção de matrículas públicas					0.558	0.000	0.659	0.000
Proporção de matrículas municipais dentre as públicas					-0.295	0.000	-0.327	0.000
Educação média da população adulta					0.559	0.000	0.620	0.000
População até 4999			-0.430	0.000	1.257	0.000	1.316	0.000
População entre 5000 e 9999			-0.215	0.009	1.096	0.000	1.172	0.000
População entre 10000 e 19999			-0.354	0.000	0.896	0.000	0.978	0.000
População entre 20000 e 49999			-0.356	0.000	0.702	0.000	0.763	0.000
População entre 50000 e 99999			-0.154	0.088	0.467	0.000	0.502	0.000
População entre 100000 e 199999			0.014	0.896	0.332	0.000	0.358	0.000
Fator de infra-estrutura					0.288	0.000	0.304	0.000
Horas-aula diárias					0.000	0.993	-0.002	0.948
Alunos por turma					-0.008	0.000	-0.010	0.000
Proporção professores com curso superior					0.271	0.000	0.334	0.000
Idade	-0.098	0.000	-0.098	0.000	-0.098	0.000	-0.098	0.000
Situação de domicílio; 1=urbano e 0=rural	0.521	0.000	0.523	0.000	0.525	0.000	0.525	0.000
Cor; 1=branco e amarelo e 0=preto e pardo	0.396	0.000	0.396	0.000	0.396	0.000	0.396	0.000
Sexo; 1=homem e 0=mulher	-0.546	0.000	-0.546	0.000	-0.547	0.000	-0.547	0.000
Educação da mãe	0.173	0.000	-0.067	0.000	0.201	0.000	0.330	0.000
Tamanho relativo da coorte de 15 a 17 anos			3.854	0.000	1.396	0.000		
Razão professor/pop. 15 a17			-0.583	0.000	-0.061	0.437	-0.171	0.029
Proporção de estabelecimentos com 50 a 200 matrículas					-0.018	0.057	-0.019	0.045
Proporção de estabelecimentos com 201 a 1000					-0.028	0.003	-0.028	0.003
Proporção de estabelecimentos com mais de 1000 matrículas					-0.027	0.043	-0.022	0.099
Proporção de matrículas públicas					-0.058	0.000	-0.065	0.000
Proporção de matrículas municipais dentre as públicas					0.022	0.003	0.024	0.001
Educação média da população adulta					-0.028	0.000	-0.033	0.000
População até 4999			0.000	0.968	-0.070	0.000	-0.075	0.000
População entre 5000 e 9999			-0.002	0.670	-0.057	0.000	-0.063	0.000
População entre 10000 e 19999			0.010	0.064	-0.045	0.000	-0.051	0.000
População entre 20000 e 49999			0.013	0.019	-0.036	0.000	-0.041	0.000
População entre 50000 e 99999			-0.004	0.459	-0.032	0.000	-0.035	0.000
População entre 100000 e 199999			-0.009	0.173	-0.023	0.000	-0.026	0.000
Fator de infra-estrutura					-0.015	0.000	-0.016	0.000
Horas-aula diárias					0.028	0.000	0.028	0.000
Alunos por turma					0.001	0.000	0.001	0.000
Proporção professores com curso superior					-0.027	0.000	-0.032	0.000
Domicílio chefiado por mulher	-0.319	0.000	-0.319	0.000	-0.319	0.000	-0.319	0.000
Família convivente	-0.057	0.017	-0.057	0.017	-0.057	0.017	-0.057	0.017
Categoria de ocupação do chefe nível superior	0.337	0.000	0.336	0.000	0.337	0.000	0.336	0.000
Categoria de ocupação do chefe nível médio	0.273	0.000	0.272	0.000	0.272	0.000	0.272	0.000
Chefe não ocupado	0.019	0.008	0.019	0.009	0.019	0.011	0.018	0.012
Efeito Aleatório								
Coeficiente	1.615	0.000	0.634	0.000	0.263	0.000	0.278	0.000
Educação da mãe	0.005	0.000	0.002	0.000	0.001	0.000	0.001	0.000
% da variância explicada - intercepto			60.77		83.71		82.79	
% da variância explicada - educação materna			55.11		76.67		74.44	

Para as variáveis de nível 2, duas análises são feitas. Primeiro, referente ao impacto direto no intercepto e segundo ao *trade-off* entre educação da mãe e as variáveis de nível 2. Na primeira análise detecta-se o efeito das variáveis contextuais na probabilidade média de freqüentar a escola na idade correta. Na segunda, é possível perceber até que ponto o efeito da educação materna pode ser minimizado pelos fatores municipais.

Focalizando primeiramente o impacto direto dos fatores de demanda e oferta (Modelo 2), percebe-se o forte impacto negativo do tamanho relativo da coorte em idade escolar, nos dois níveis de ensino, com maior intensidade para o médio. Considerando os coeficientes desse modelo, a simulação do aumento de um desvio padrão em relação á média dessa variável, diminui a probabilidade de freqüentar a escola na idade certa em 37% no fundamental e 43,7% no médio. Para a razão professor/população em idade escolar, o resultado para o ensino fundamental não foi o esperado, ou seja, foi negativo, já para o nível médio o sinal foi positivo.

Com relação ao impacto do tamanho relativo da coorte em idade escolar no coeficiente da educação materna, percebe-se que ele é bem menor, o que pode ser um indicativo de que essa variável apesar de estar associada ao nível socioeconômico do município, também está captando variações na razão de dependência, já que teoricamente não seria de se esperar que a razão de dependência tivesse efeito na interação com a educação materna.

Apesar do importante impacto da variável de pressão demográfica, a análise do restante dos fatores de rede escolar dos municípios é feita com o Modelo 4, que a exclui. Como já mencionado, essa opção se deve ao fato da alta correlação do tamanho relativo da coorte com praticamente todas as demais variáveis.

O impacto direto dos fatores da rede escolar do município mostrou-se importante na probabilidade de cursar a escola na série adequada, para os dois níveis de ensino. Para melhor visualizar o efeito desses fatores, os GRAFs. 5.6 e 5.7 mostram o valor predito segundo simulações de mudanças nas variáveis de rede escolar, utilizando os coeficientes do Modelo 4. Considerou-se a variação da média mais e menos o desvio padrão de cada variável.

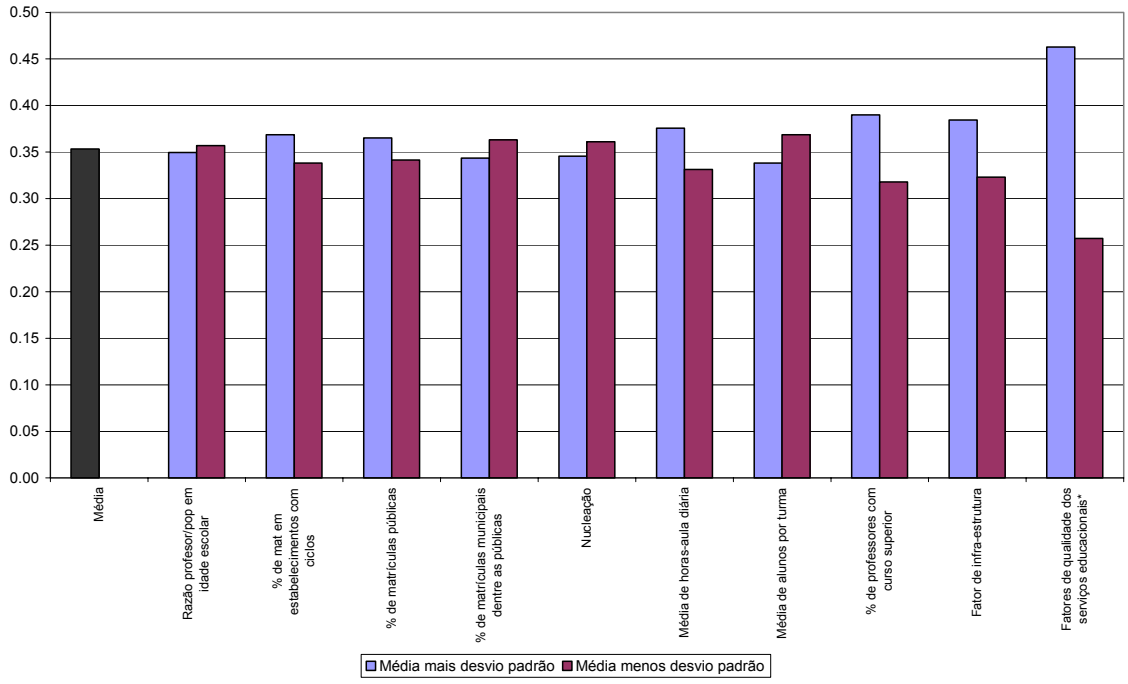
Para o ensino fundamental, percebe-se que as variáveis relacionadas à qualidade dos serviços educacionais são os fatores com maior impacto. Dentre elas, destaca-se o grande efeito da proporção de professores com curso superior – variável de qualidade dos recursos humanos – e o fator de infra-estrutura – medida *proxy* da qualidade da infra-estrutura média das escolas do município. Com relação às outras variáveis da rede escolar do município, os principais resultados no ensino fundamental são: i) pequeno impacto das variáveis de nucleação; ii) efeito negativo para a porcentagem de matrículas municipais dentre as públicas, podendo inferir que a municipalização do ensino pode

não ser acompanhada de uma boa eficiência do ensino e iii) coeficiente negativo e significativo apenas a 3% para a razão professor/população de 7 a 14 anos.

No caso do ensino médio (GRAF. 5.7), dentre os fatores relacionados ao perfil das escolas do município, destaca-se o significativo impacto do fator de infra-estrutura. Verifica-se também que as variáveis de nucleação e restrição de oferta são mais importantes nesse nível de ensino que no fundamental.

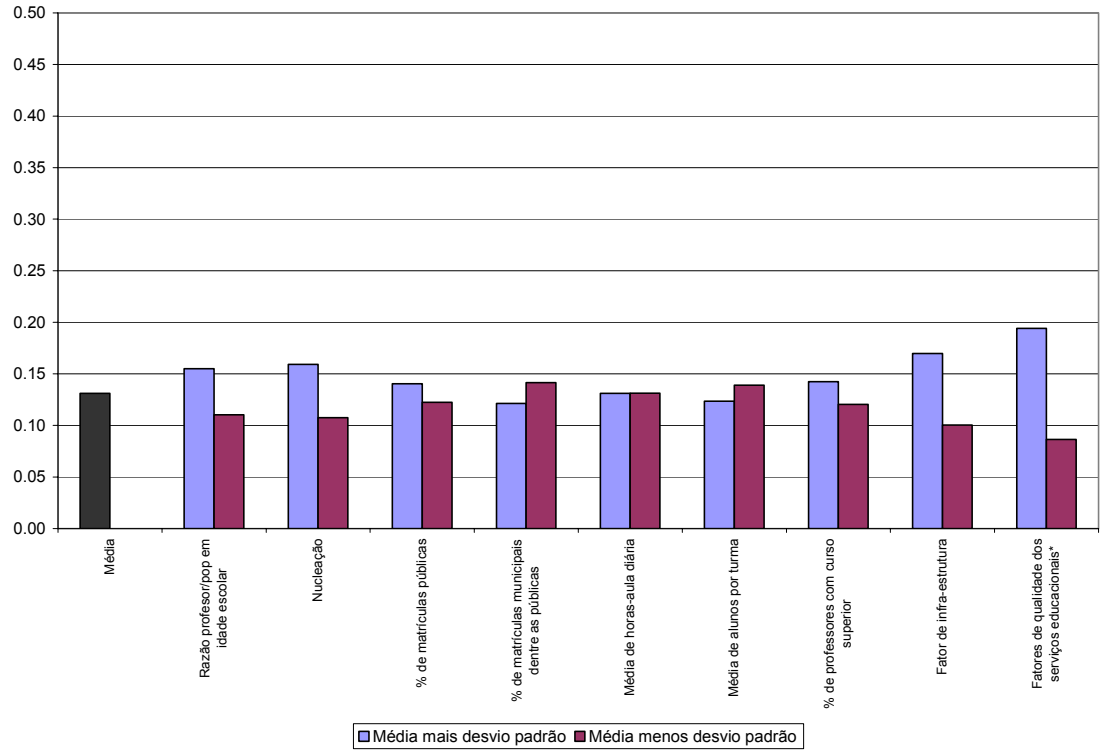
Cabe destacar, que a melhora simultânea dos fatores de qualidade da rede escolar do município provoca um significativo aumento na freqüência na idade correta nos dois níveis de ensino, como pode ser visto nos gráficos abaixo. Tal fato evidencia a importância de políticas públicas que estimulem a melhoria do perfil escolar dos municípios.

GRÁFICO 5. 6: Impacto da variação dos fatores de estrutura escolar do município na probabilidade de freqüentar a escola na idade adequada - Ensino Fundamental



*Nota: Na simulação da melhoria nos fatores de qualidade dos serviços educacionais, para média de alunos por turma, considerou a média menos o desvio padrão; e para a piora, a média mais o desvio padrão.

GRÁFICO 5. 7: Impacto da variação dos fatores de estrutura escolar do município na probabilidade de freqüentar a escola na idade adequada - Ensino Médio



*Nota: Na simulação da melhoria nos fatores de qualidade dos serviços educacionais, para média de alunos por turma, considerou a média menos o desvio padrão; e para a piora, a média mais o desvio padrão.

De acordo com os resultados das regressões vistos acima, percebe-se que tanto os fatores de *background* familiar quanto os relacionados à estrutura escolar dos municípios, são importantes na determinação da probabilidade de cursar a escola na idade correta, para os dois níveis de ensino. Para melhor comparar os efeitos de ambos fatores, a TAB. 5.7 traz as variações nessa probabilidade decorrente de variações desses fatores (considerou como variação a média mais e menos o desvio padrão de cada variável). Verifica-se que a educação materna realmente é o mais importante determinante da probabilidade de cursar a idade adequada, mas cabe ressaltar que os fatores de rede escolar dos municípios apresentam impacto maior que as demais variáveis de ambiente familiar e bastante significativo quando considera uma melhora simultânea nos fatores de qualidade dos serviços educacionais. Tal resultado traz implicações importantes para a política educacional, pois demonstra que de certa forma o contexto pode minimizar o papel da família sobre a eficiência do aluno. Vale a pena ressaltar, que da mesma forma que na seção anterior, o ensino médio apresenta maior variação que o fundamental por estar em um nível menor, já que apenas 24% dos

estudantes de 15 a 17 anos estavam na série adequada a sua idade, ao passo que para os alunos de 7 a 14 anos esse percentual é de 42%.

TABELA 5. 7: Variações na probabilidade de freqüentar a escola na idade adequada segundo variações nas variáveis de *background* familiar e rede escolar – ensino fundamental e médio

Variáveis Independentes	Ensino Fundamental		Ensino Médio	
Background Familiar	Média mais desvio padrão	Média menos desvio padrão	Média mais desvio padrão	Média menos desvio padrão
Educação da mãe	24,25	-21,85	81,30	-48,10
Domicílio chefiado por mulher	-3,39	3,44	-10,82	11,91
Família convivente	0,10	-0,10	-0,59	0,59
Categoria de ocupação do chefe nível superior	2,16	-2,14	8,47	-7,90
Categoria de ocupação do chefe nível médio	2,13	-2,11	9,07	-8,42
Chefe não ocupado	-1,80	1,81	0,74	-0,73
Rede Escolar				
Razão profesor/pop. em idade escolar	-1,03	1,04	18,31	-15,85
% de estabelecimentoa com ciclo	4,33	-4,25	-	-
% de matrículas públicas	3,39	-3,34	7,08	-6,68
% de matrículas municipais dentre as públicas	-2,78	2,81	-7,42	7,92
Nucleação	-2,16	2,18	21,34	-18,07
Média de horas-aula diária	6,39	-6,21	-0,06	0,06
Média de alunos por turma	-4,25	4,33	-5,75	6,04
% de professores com curso superior	10,42	-9,95	8,71	-8,11
Fator de infra-estrutura da rede escolar	8,88	-8,54	29,51	-23,60
Fatores de qualidade dos serviços educacionais*	31,06	-27,22	48,10	-34,15

*Nota: Na simulação da melhoria nos fatores de qualidade dos serviços educacionais, para média de alunos por turma, considerou a média menos o desvio padrão; e para a piora, a média mais o desvio padrão.

Passando para a outra análise, referente ao efeito das variáveis agregadas no intercepto da educação materna, observa-se que nos dois níveis de ensino ocorre o *trade off* entre essa variável e as relacionadas ao perfil escolar do município.

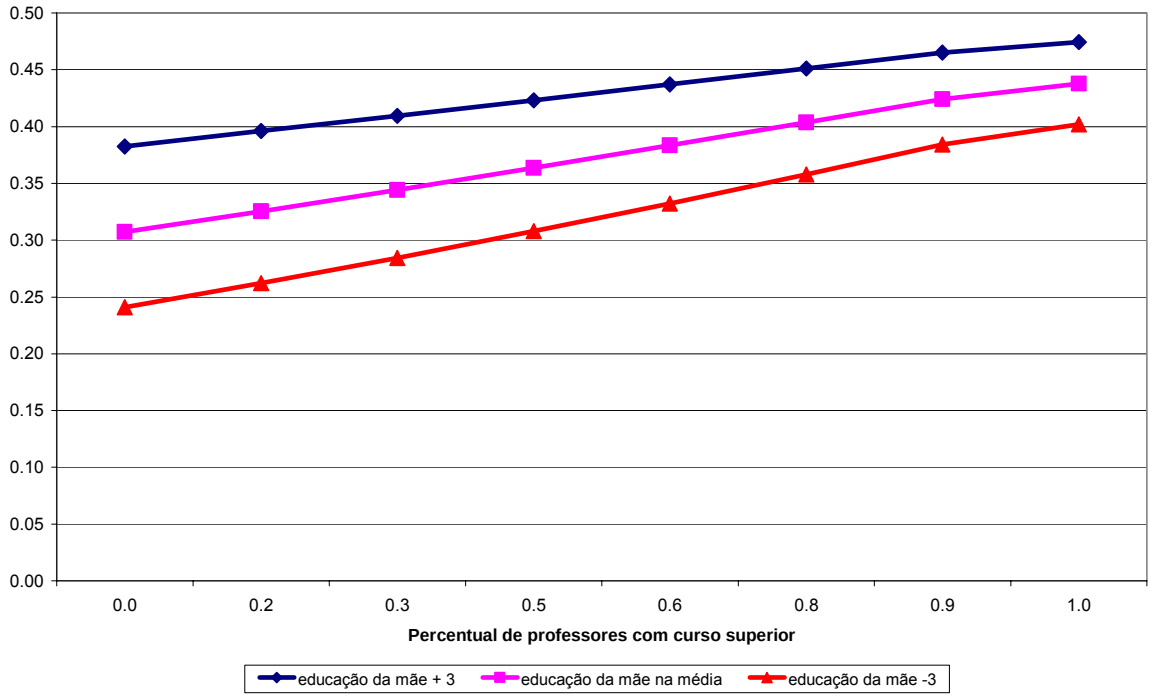
No ensino fundamental o efeito substituição, mencionado acima, ocorre com a porcentagem de professores com curso superior, alunos por turma e fator de infra-estrutura, ou seja, com a maioria dos fatores de qualidade escolar do município. Os coeficientes dessas variáveis são negativos e significativos no coeficiente da educação materna (para média de alunos por turma o coeficiente é positivo), em outras palavras pode-se dizer que esses fatores são substitutos da educação materna. Assim, filhos de mães com baixa escolaridade são mais beneficiados pela melhora da qualidade da rede de ensino dos municípios que filhos de mães mais escolarizadas, o que sugere que a

melhoria dos fatores referentes ao perfil escolar do município tem um importante papel na diminuição das desigualdades educacionais entre os alunos de origem socioeconômica diferentes.

Com o objetivo de melhor visualizar o *trade-off* mencionado acima, são apresentados exercícios de simulação que mostra o impacto de uma variação simultânea da educação da mãe e dos fatores da rede escolar do município mais relevantes, no que tange ao efeito substituição (GRAFs. 5.8 a 5.11).

No GRAF. 5.8 a visualização do efeito substituição entre a educação da mãe e a porcentagem de professores com curso superior é clara. A linha do meio representa a educação da mãe na média, a superior quando a educação da mãe é três anos a mais que a média e a inferior, três anos a menos. Considerando o início da curva, onde a porcentagem de professores com curso superior no município é igual a zero, o aumento da educação materna provoca um grande aumento na probabilidade de freqüentar a escola na idade adequada para as crianças de 7 a 14 anos, na ordem de 24,5%. Por outro lado, uma diminuição reduz a probabilidade em 21,6%. À medida que o percentual de professores com curso superior aumenta, o impacto da educação materna diminui. Dessa forma, quando o percentual de professores com curso superior é de 100%, um aumento de três anos em relação a media na educação materna aumenta a probabilidade em 8,35% e uma diminuição reduz em 8,20% essa probabilidade. Pode-se dizer, portanto, que o impacto da educação da mãe na probabilidade de cursar a série na idade adequada no ensino fundamental é diminuído à medida que aumenta o porcentual de professores com curso superior no município.

GRÁFICO 5. 8: Impacto da variação da educação da mãe e do porcentual de professores com curso superior na probabilidade de freqüentar a escola na idade correta - ensino fundamental



Da mesma forma, através dos GRAFs. 5.9 e 5.10 percebe-se o efeito substituição da educação materna e o número médio de alunos por turma e o fator de infra-estrutura do município, porém de maneira menos acentuada. Variações na educação materna possuem menor efeito em turmas menores, à medida que o número de alunos por turma aumenta, o efeito da educação da mãe é maior. No caso do fator de infra-estrutura, quanto maior a qualidade da infra-estrutura dos municípios, com aumento na porcentagem de matrículas beneficiadas em escolas com bibliotecas, quadras de esportes, laboratórios de informática e ciências, menor o efeito da educação materna sobre a probabilidade do filho.

O GRAF. 5.11 mostra a variação simultânea nos três fatores de qualidade da rede de ensino (fator de infra-estrutura, porcentagem de professores com curso superior e média de alunos por turma). A escala do eixo das abscissas representa a melhoria dos três fatores, correspondente as escalas dos três gráficos anteriores. Assim, a escala 1 representa um município com nenhum professor com curso superior, média de alunos por turma de 80 e fator de infra-estrutura escolar igual a –1,3. No outro extremo, a escala 8 representa todos os professores com curso superior, média de alunos por turma igual a 10 e fator de infra-estrutura de 3,6.

Observa-se que uma variação simultânea nos três fatores de qualidade escolar possui um grande efeito substituição com a educação materna. Em uma rede de ensino com alta qualidade de recursos humanos e de infra-estrutura, ou seja, salas pequenas, todos os professores com ensino superior e todas as escolas com quadras, bibliotecas e laboratórios de ciências e informática, a probabilidade do aluno freqüentar a escola na idade correta não difere quanto ao nível educacional da mãe. Nesse caso, tem-se uma situação onde as oportunidades educacionais não dependem da origem social do aluno, ou seja, uma sociedade com alta mobilidade social.

GRÁFICO 5. 9: Impacto da variação da educação da mãe e da média de alunos por turma na probabilidade de freqüentar a escola na idade correta - ensino fundamental

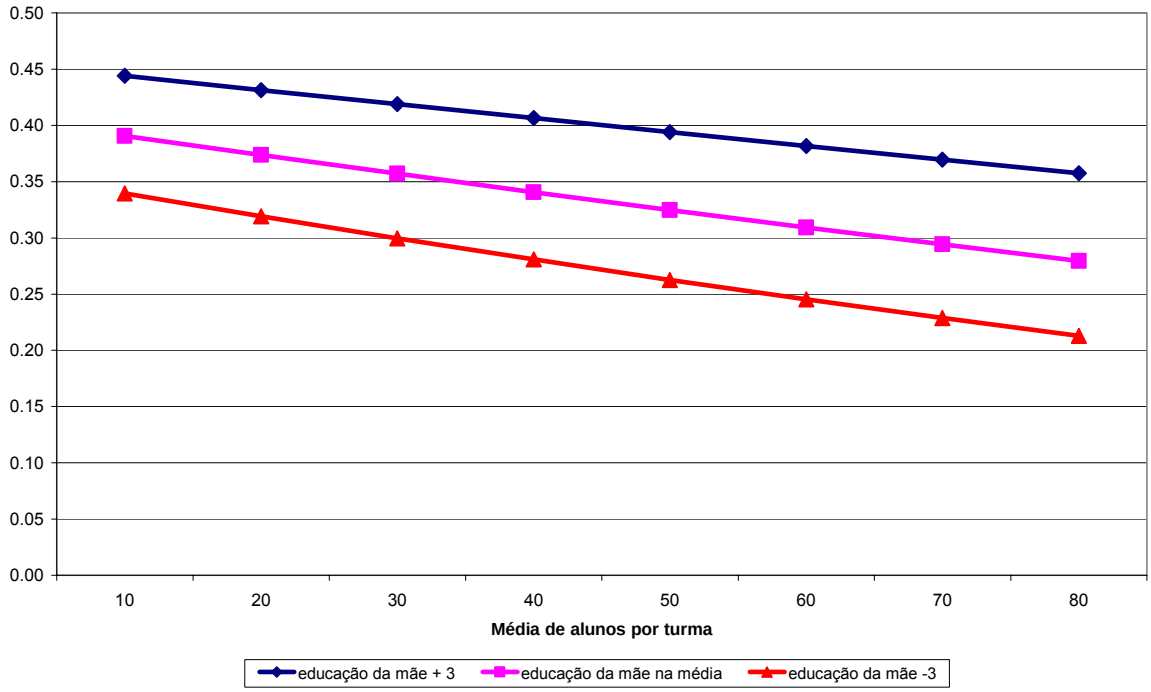


GRÁFICO 5. 10: Impacto da variação da educação da mãe e do fator de infra-estrutura na probabilidade de freqüentar a escola na idade correta - ensino fundamental

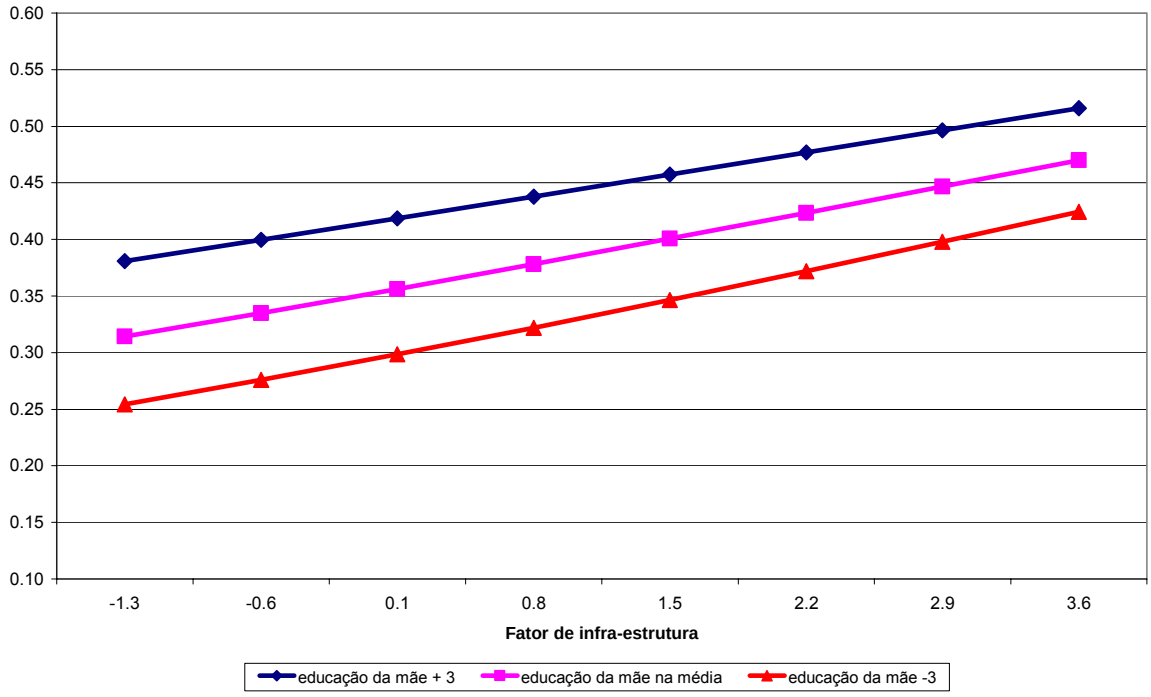
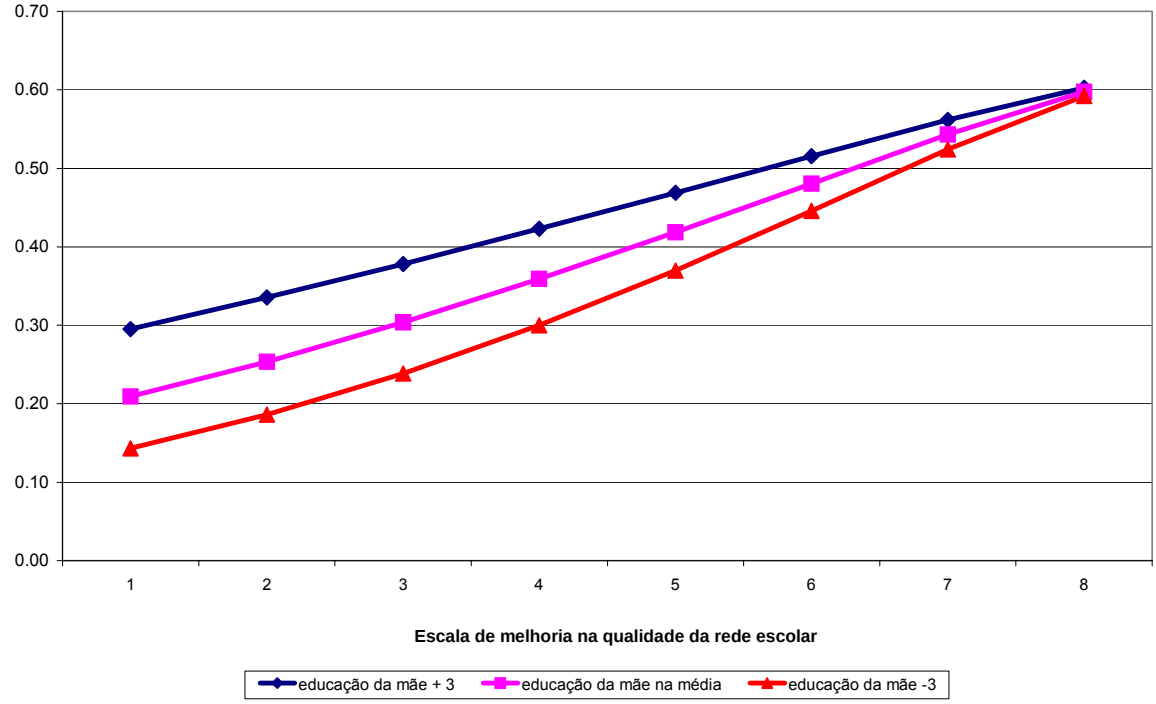


GRÁFICO 5. 11: Impacto da variação da educação da mãe e da qualidade da rede escolar do município na probabilidade de freqüentar a escola na idade correta - ensino fundamental



Para o ensino médio, o *trade off* entre as varáveis relacionadas à rede escolar do município e a educação materna também ocorre e para as mesmas variáveis do caso anterior, porém, de forma bem mais sutil, não sendo possível visualizar graficamente esse efeito. Porém, quando considera municípios com 0% de professores com curso superior, um aumento de educação materna de 3 anos em relação a média aumenta a probabilidade dos alunos de 15 a 17 anos cursar a escola na idade correta em 66,53%. No outro extremo, considerando um valor de 100% para a porcentagem de professores com curso superior, um mesmo aumento na educação materna provoca um aumento menor, de 50,39%. O mesmo ocorre para o número de alunos por turma e para o fator de infra-estrutura, ou seja, o impacto da educação materna é minimizado pela melhora desse fatores .

Com base no que foi exposto, pode-se dizer que a melhora do perfil escolar dos municípios possui dois efeitos. Primeiro, aumentar a probabilidade de freqüentar a escola na idade correta para os dois níveis de ensino, e, segundo, diminuir a importância dos fatores relacionados ao ambiente familiar do aluno (no caso, educação da mãe), diminuindo a desigualdade intergeracional.

5.3) Probabilidade de Progressão por Série (PPS)

Nessa seção são estimadas as regressões para as PPSs. Considerou-se os mesmos modelos do caso anterior, porém, os testes para a aleatoriedade do efeito da educação da mãe, mostrou que eles só são aleatórios para a probabilidade de passar de zero para um ano de estudo (e_0), dessa forma, para as outras PPSs foi considerado apenas o intercepto como tendo efeito aleatório.

Começando a análise pelo efeito aleatório, os resultados das TAB. 5.8 a 5.10 mostram que o componente da variância no modelo não-condicional é bastante alto para e_0, e diminui nas transições mais elevadas. Quando se incorpora variáveis contextuais nos modelos, o poder de explicação delas é maior para e_4, explica mais de 84% da variabilidade do intercepto.

Passando para os efeitos fixos, e analisando primeiramente as variáveis de *background* familiar, percebe-se que no modelo 1 os coeficientes para a educação da mãe e as *dummies* de posição de ocupação do chefe (exceto chefe fora da ocupação)

aumentam um pouco de e_0 para e_4 e caiem novamente, em e_8. Corroborando com o trabalho de SILVA E HASENBALG (2001), que encontraram um efeito crescente das variáveis explicativas até a 4ª série do ensino fundamental, quando então começam a declinar. Por outro lado, as variáveis categóricas "domicílios chefiados por mulheres" e "família convivente" não apresentam esse padrão. A primeira apresenta um aumento ao longo das transições, porém bem mais significativo de e_0 para e_4, enquanto que a última perde significância em e_4 e e_8.

Com relação ao impacto direto das variáveis agregadas, a sua análise também será feita considerando os dois enfoques da seção anterior, ou seja, o efeito dos fatores de demanda e oferta (Modelo 2) e o efeito dos demais fatores da rede escolar do município (Modelo 4). Verifica-se que para as progressões das três séries analisadas, o efeito do tamanho relativo da coorte em idade escolar também é bastante forte, evidenciando que o bônus demográfico também é importante para a melhora do fluxo dos alunos entre as séries.

Partindo para a análise das demais variáveis relacionadas ao perfil médio do ensino dos municípios (Modelo 4), percebe-se que seus efeitos também são crescentes de e_0 para e_4 e decrescentes para e_8, tornando-se não significativos. Apenas a variável de restrição de oferta – razão professor/população em idade escolar – apresenta comportamento diferente, pois seu efeito é significativo e forte apenas para a progressão do ensino médio (e_8), apontando que para a transição do ensino médio a disponibilidade de oferta é importante.

Um resultado interessante, é que a proporção de matrículas em escolas com ciclo não tem efeito para e_0, já que não foi estatisticamente significativo, já para e_4 ela foi significativa e com sinal positivo.

TABELA 5. 8: Resultado da regressão para a probabilidade de progressão na 1ª série (e_0)

Variáveis Independentes	Modelo 1		Modelo 2		Modelo 3		Modelo 4	
Efeito Fixo	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig
Intercepto	2.1485	0.0000	7.086	0.000	0.867	0.084	-1.122	0.018
Tamanho relativo da coorte de 7 a 14 anos			-26.068	0.000	-8.186	0.000		
Razão professor/pop. 7 a14			-3.095	0.001	-1.518	0.112	1.414	0.126
Proporção de matrículas em escolas com ciclo					-0.025	0.555	-0.024	0.575
Proporção de matrículas públicas					1.354	0.000	1.157	0.002
Proporção de matrículas municipais dentre as públicas					-0.242	0.000	-0.203	0.002
Proporção de estabelecimentos com 50 a 200 matrículas					0.131	0.079	0.107	0.155
Proporção de estabelecimentos com 201 a 1000					0.005	0.944	0.001	0.986
Proporção de estabelecimentos com mais de 1000 matrículas					-0.091	0.722	-0.114	0.659
Educação média da população adulta					0.294	0.000	0.353	0.000
População até 4999			-0.176	0.037	0.307	0.003	0.378	0.000
População entre 5000 e 9999			-0.126	0.118	0.347	0.000	0.405	0.000
População entre 10000 e 19999			-0.195	0.014	0.326	0.001	0.390	0.000
População entre 20000 e 49999			-0.230	0.004	0.203	0.019	0.242	0.006
População entre 50000 e 99999			-0.142	0.102	0.111	0.190	0.139	0.105
População entre 100000 e 199999			-0.022	0.831	0.088	0.341	0.109	0.249
Horas-aula diárias					0.089	0.116	0.133	0.020
Alunos por turma					-0.013	0.000	-0.014	0.000
Proporção professores com curso superior					0.291	0.001	0.556	0.000
Fator de infra-estrutura					0.138	0.000	0.169	0.000
Situação de domicílio; 1=urbano e 0=rural	0.5276	0.0000	0.523	0.000	0.520	0.000	0.520	0.000
Cor; 1=branco e amarelo e 0=preto e pardo	0.1539	0.0000	0.157	0.000	0.159	0.000	0.159	0.000
Sexo; 1=homem e 0=mulher	-0.4247	0.0000	-0.425	0.000	-0.426	0.000	-0.426	0.000
Educação da mãe	0.1579	0.0000	-0.180	0.000	0.492	0.000	0.549	0.000
Tamanho relativo da coorte de 7 a 14 anos			1.834	0.000	0.244	0.187		
Razão professor/pop. 7 a14			0.602	0.004	0.169	0.454	0.068	0.751
Proporção de matrículas em escolas com ciclo					-0.062	0.000	-0.063	0.000
Proporção de matrículas públicas					-0.273	0.000	-0.262	0.001
Proporção de matrículas municipais dentre as públicas					0.019	0.191	0.018	0.199
Proporção de estabelecimentos com 50 a 200 matrículas					0.013	0.490	0.012	0.511
Proporção de estabelecimentos com 201 a 1000					-0.018	0.314	-0.021	0.244
Proporção de estabelecimentos com mais de 1000 matrículas					-0.045	0.384	-0.049	0.336
Educação média da população adulta					-0.019	0.000	-0.020	0.000
População até 4999			-0.054	0.000	-0.066	0.001	-0.067	0.001
População entre 5000 e 9999			-0.027	0.031	-0.042	0.018	-0.043	0.014
População entre 10000 e 19999			-0.012	0.286	-0.034	0.034	-0.035	0.028
População entre 20000 e 49999			-0.010	0.369	-0.032	0.033	-0.032	0.030
População entre 50000 e 99999			-0.015	0.204	-0.030	0.026	-0.031	0.023
População entre 100000 e 199999			-0.023	0.109	-0.025	0.079	-0.025	0.071
Horas-aula diárias					0.005	0.673	0.003	0.834
Alunos por turma					0.000	0.619	0.000	0.505
Proporção professores com curso superior					-0.035	0.070	-0.040	0.028
Fator de infra-estrutura					-0.010	0.071	-0.011	0.048
Domicílio chefiado por mulher	-0.0854	0.0000	-0.086	0.000	-0.085	0.000	-0.084	0.000
Família convivente	-0.0838	0.0090	-0.089	0.006	-0.090	0.006	-0.089	0.006
Categoria de ocupação do chefe nível superior	0.1412	0.0000	0.169	0.000	0.189	0.000	0.187	0.000
Categoria de ocupação do chefe nível médio	0.0364	0.0320	0.047	0.006	0.049	0.004	0.048	0.005
Chefe não ocupado	-0.1552	0.0000	-0.149	0.000	-0.146	0.000	-0.147	0.000
Efeito Aleatório								
Coeficiente	1.01348	0.000	0.52259	0.000	0.382	0.000	0.39849	0.000
Educação da mãe	0.00817	0.000	0.00643	0.000	0.005	0.000	0.00513	0.000
% da variância explicada - intercepto			48.44		62.27		60.68	
% da variância explicada - educação materna			21.30		37.09		37.21	

TABELA 5. 9: Resultado da regressão para a probabilidade de progressão na 5ª série (e_4)

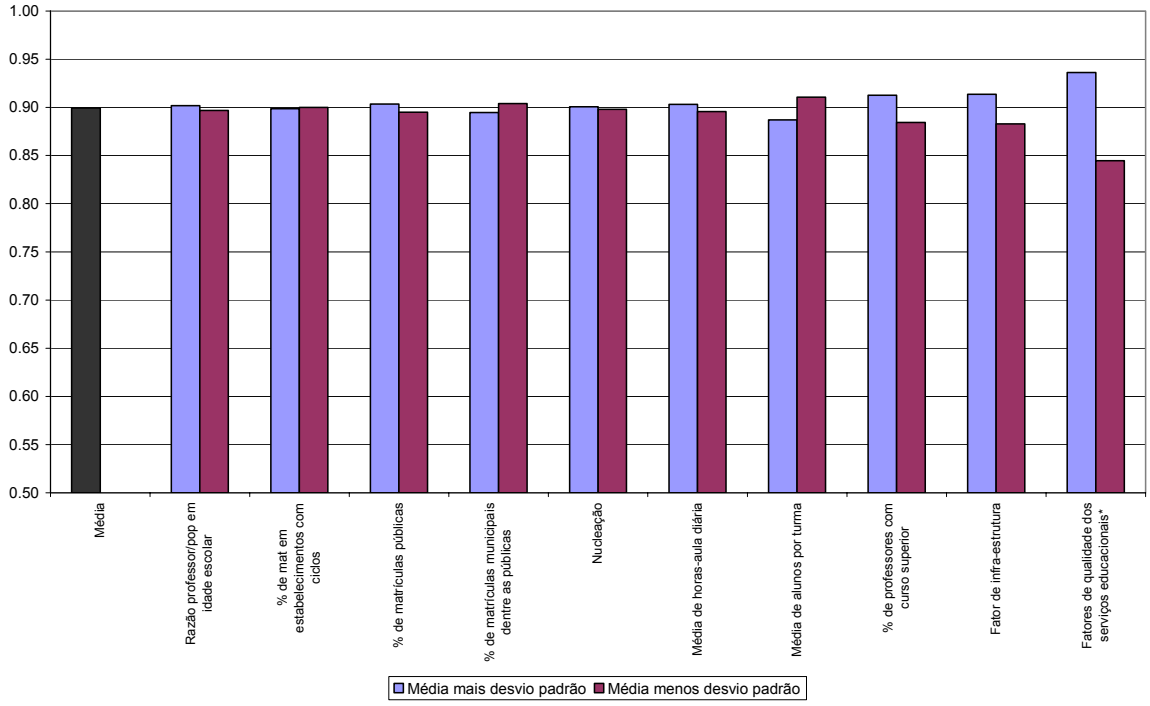
Variáveis Independentes	Modelo 1		Modelo 2		Modelo 3		Modelo 4	
Efeito Fixo	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig
Intercepto	1.666	0.000	7.080	0.000	-1.479	0.000	-2.423	0.000
Tamanho relativo da coorte de 7 a 14 anos			-28.009	0.000	-3.966	0.000		
Razão professor/pop. 7 a14			-6.509	0.000	-0.155	0.833	1.191	0.091
Proporção de matrículas em escolas com ciclo					0.161	0.000	0.160	0.000
Proporção de matrículas públicas					1.673	0.000	1.567	0.000
Proporção de matrículas municipais dentre as públicas					-0.321	0.000	-0.303	0.000
Proporção de estabelecimentos com 50 a 200 matrículas					0.136	0.017	0.118	0.039
Proporção de estabelecimentos com 201 a 1000					0.201	0.001	0.194	0.001
Proporção de estabelecimentos com mais de 1000 matrículas					-0.226	0.214	-0.235	0.198
Educação média da população adulta					0.282	0.000	0.311	0.000
População até 4999			-0.016	0.823	0.421	0.000	0.457	0.000
População entre 5000 e 9999			-0.100	0.140	0.359	0.000	0.390	0.000
População entre 10000 e 19999			-0.204	0.003	0.326	0.000	0.357	0.000
População entre 20000 e 49999			-0.182	0.007	0.253	0.000	0.273	0.000
População entre 50000 e 99999			-0.033	0.651	0.221	0.000	0.235	0.000
População entre 100000 e 199999			-0.027	0.749	0.061	0.303	0.071	0.231
Horas-aula diárias					0.142	0.001	0.164	0.000
Alunos por turma					-0.004	0.003	-0.005	0.000
Proporção professores com curso superior					0.617	0.000	0.735	0.000
Fator de infra-estrutura					0.246	0.000	0.261	0.000
Situação de domicílio; 1=urbano e 0=rural	0.528	0.000	0.527	0.000	0.527	0.000	0.526	0.000
Cor; 1=branco e amarelo e 0=preto e pardo	0.312	0.000	0.313	0.000	0.316	0.000	0.316	0.000
Sexo; 1=homem e 0=mulher	-0.515	0.000	-0.518	0.000	-0.522	0.000	-0.522	0.000
Educação da mãe	0.181	0.000	0.182	0.000	0.183	0.000	0.183	0.000
Domicílio chefiado por mulher	-0.251	0.000	-0.252	0.000	-0.254	0.000	-0.254	0.000
Família convivente	-0.039	0.327	-0.040	0.319	-0.043	0.293	-0.043	0.291
Categoria de ocupação do chefe nível superior	0.651	0.000	0.672	0.000	0.695	0.000	0.695	0.000
Categoria de ocupação do chefe nível médio	0.180	0.000	0.186	0.000	0.186	0.000	0.186	0.000
Chefe não ocupado	-0.026	0.044	-0.021	0.101	-0.019	0.151	-0.019	0.143
Efeito Aleatório								
coeficiente	0.915	0.000	0.36961	0.000	0.142	0.000	0.144	0.000
% da variância explicada - intercepto			59.62		84.48		84.25	

TABELA 5. 10: Resultado da regressão para a probabilidade de progressão na 1ª série do ensino médio (e_8)

Variáveis Independentes	Modelo 1		Modelo 2		Modelo 3		Modelo 4	
Efeito Fixo	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig
Intercepto	1.216	0.000	2.129	0.000	-0.197	0.394	-0.599	0.002
Tamanho relativo da coorte de 15 a 17 anos			-16.933	0.000	-4.509	0.003		
Razão professor/pop. 15 a17			11.106	0.000	5.786	0.000	6.223	0.000
Proporção de estabelecimentos com 50 a 200 matrículas					0.073	0.320	0.077	0.295
Proporção de estabelecimentos com 201 a 1000					0.121	0.097	0.122	0.094
Proporção de estabelecimentos com mais de 1000 matrículas					0.077	0.461	0.059	0.571
Proporção de matrículas públicas					-0.036	0.696	-0.011	0.904
Proporção de matrículas municipais dentre as públicas					-0.018	0.713	-0.027	0.573
Educação média da população adulta					0.176	0.000	0.190	0.000
População até 4999			-0.448	0.000	0.150	0.018	0.163	0.010
População entre 5000 e 9999			-0.285	0.000	0.164	0.002	0.179	0.001
População entre 10000 e 19999			-0.251	0.000	0.150	0.002	0.167	0.001
População entre 20000 e 49999			-0.238	0.000	0.087	0.048	0.100	0.024
População entre 50000 e 99999			-0.172	0.000	0.028	0.516	0.035	0.422
População entre 100000 e 199999			-0.076	0.130	0.036	0.437	0.042	0.366
Fator de infra-estrutura					0.091	0.000	0.096	0.000
Horas-aula diárias					0.081	0.004	0.080	0.004
Alunos por turma					0.003	0.162	0.002	0.278
Proporção professores com curso superior					-0.009	0.853	0.009	0.848
Situação de domicílio; 1=urbano e 0=rural	0.474	0.000	0.470	0.000	0.469	0.000	0.469	0.000
Cor; 1=branco e amarelo e 0=preto e pardo	0.295	0.000	0.296	0.000	0.295	0.000	0.295	0.000
Sexo; 1=homem e 0=mulher	-0.572	0.000	-0.575	0.000	-0.577	0.000	-0.577	0.000
Educação da mãe	0.124	0.000	0.125	0.000	0.125	0.000	0.125	0.000
mãe chefe	-0.262	0.000	-0.263	0.000	-0.264	0.000	-0.264	0.000
Família convivente	-0.039	0.475	-0.043	0.441	-0.043	0.437	-0.043	0.437
Categoria de ocupação do chefe nível superior	0.557	0.000	0.572	0.000	0.577	0.000	0.577	0.000
Categoria de ocupação do chefe nível médio	0.209	0.000	0.211	0.000	0.211	0.000	0.211	0.000
Chefe não ocupado	0.050	0.001	0.053	0.001	0.054	0.000	0.054	0.000
Efeito Aleatório								
Coeficiente	0.260	0.000	0.10962	0.000	0.082	0.000	0.082	0.000
% da variância explicada - intercepto			57.80		68.51		68.30	

Os gráficos abaixo ilustram o efeito direto dos fatores de rede escolar do município nas diferentes probabilidades, através de simulações de variações nesses fatores. Percebe-se o maior impacto da porcentagem de professores com curso superior, média de alunos por turma e do fator de infra-estrutura para e_0 e e_4. No caso do e_8, além desse último, a razão professor/população em idade escolar também se mostra bastante importante, um aumento dela tem maior efeito que a melhora simultânea dos fatores de qualidade dos serviços educacionais.

GRÁFICO 5. 12: Impacto da variação dos fatores de rede escolar do município na probabilidade de freqüentar a escola na idade adequada – Probabilidade de Progressão na 1ª série (e_0)



*Nota: Na simulação da melhoria nos fatores de qualidade dos serviços educacionais, para média de alunos por turma, considerou a média menos o desvio padrão; e para a piora, a média mais o desvio padrão.

GRÁFICO 5. 13: Impacto da variação dos fatores de rede escolar do município na probabilidade de freqüentar a escola na idade adequada – Probabilidade de Progressão na 5ª série (e_4)

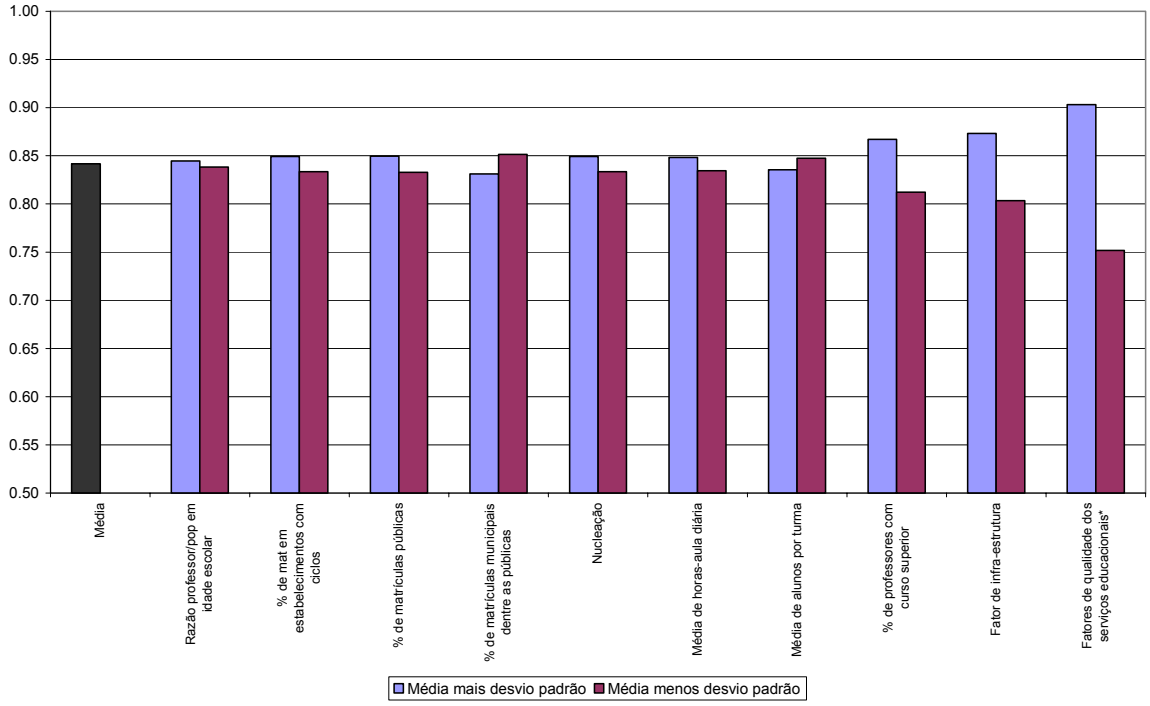
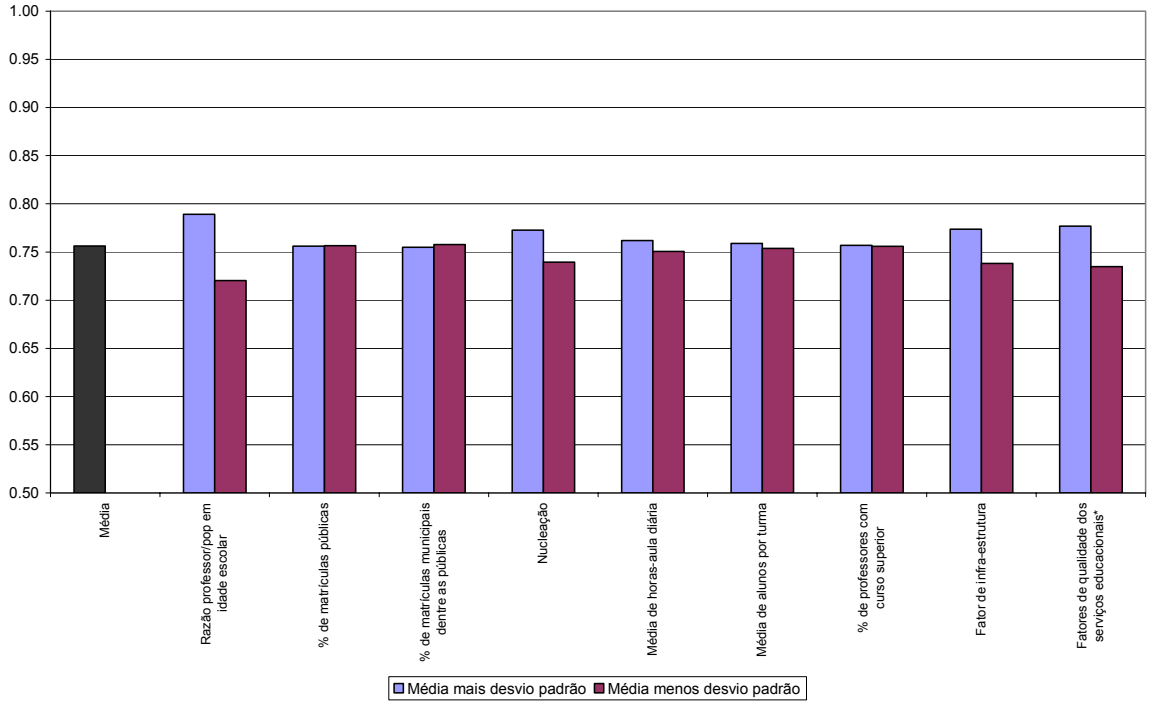


GRÁFICO 5. 14: Impacto da variação dos fatores de rede escolar do município na probabilidade de freqüentar a escola na idade adequada – Probabilidade de Progressão na 1ª série do ensino médio (e_8)



*Nota: Na simulação da melhoria nos fatores de qualidade dos serviços educacionais, para média de alunos por turma, considerou a média menos o desvio padrão; e para a piora, a média mais o desvio padrão.

Analisando o *trade-off* entre a educação materna e alguns fatores de rede escolar do município na primeira progressão, observa-se que ele é mais expressivo para a proporção de matrículas em estabelecimentos com ciclos, proporção de professores com curso superior e o fator de infra-estrutura. À medida que esses fatores aumentam, a distância vertical entre as curvas de probabilidades, para os diferentes níveis de escolaridade materna, diminui.

A variação simultânea desses três fatores (GRAF. 5.18), mostra a importância da melhoria do perfil escolar dos municípios, já que provoca uma "democratização" das transições entre as séries, ou seja, a origem socioeconômica do aluno passa a não ter efeito.

GRÁFICO 5. 15: Impacto da variação da educação da mãe e da proporção de matrículas em estabelecimentos com ciclos na Probabilidade de Progressão na 1ª série (e_0)

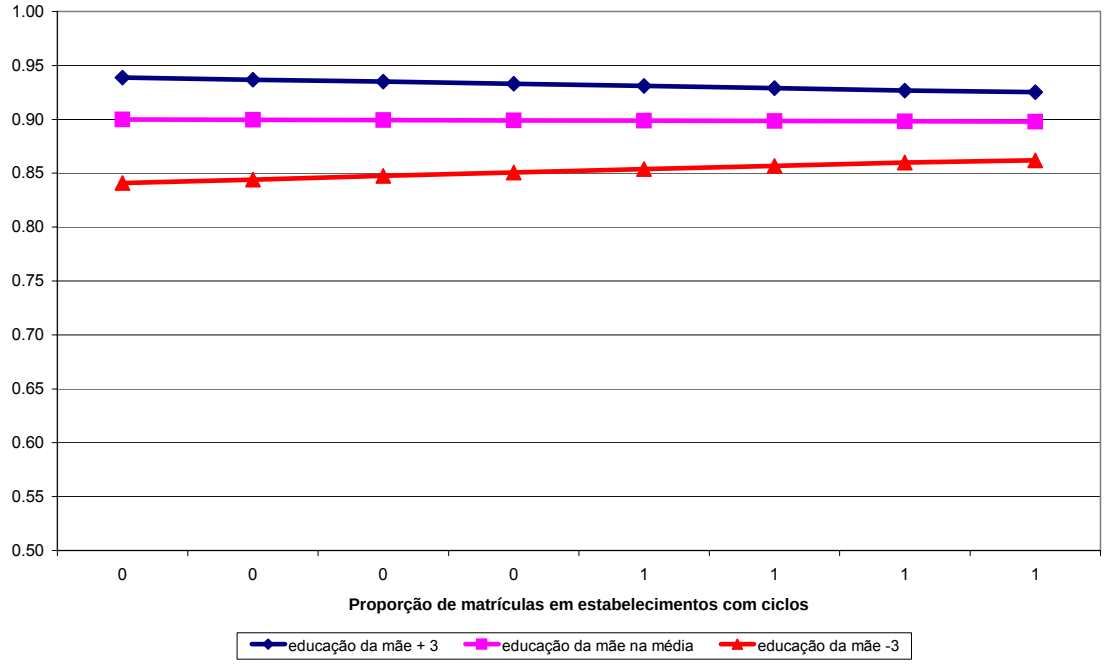


GRÁFICO 5. 16: Impacto da variação da educação da mãe e do porcentual de professores com curso superior na Probabilidade de Progressão na 1ª série (e_0)

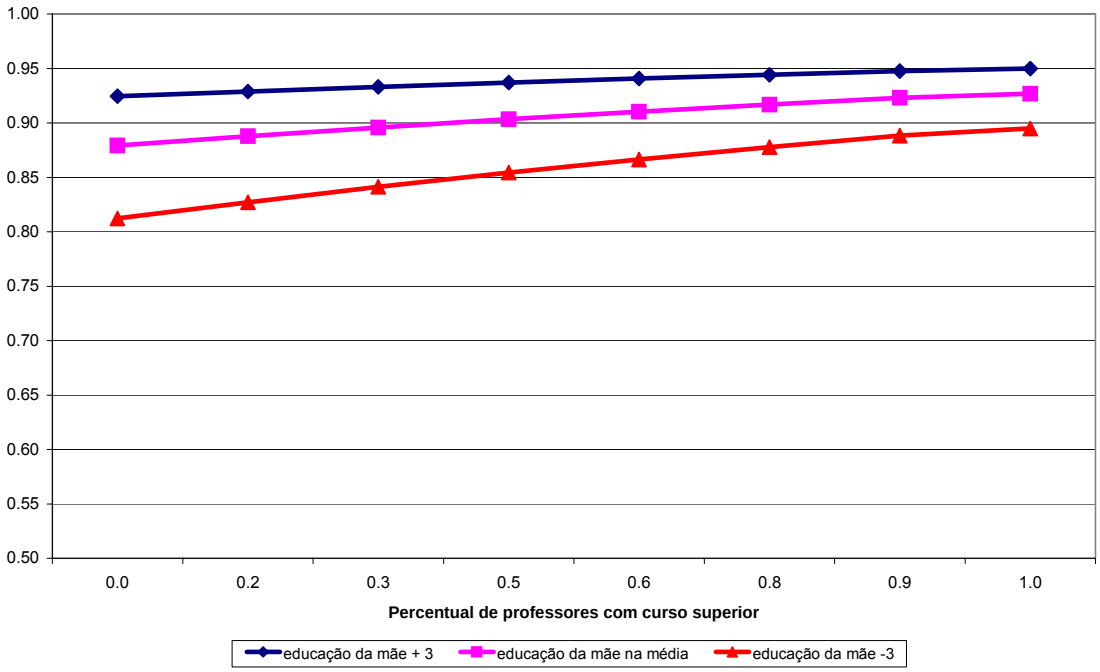


GRÁFICO 5. 17: Impacto da variação da educação da mãe e do fator de infra-estrutura na Probabilidade de Progressão na 1ª série (e_0)

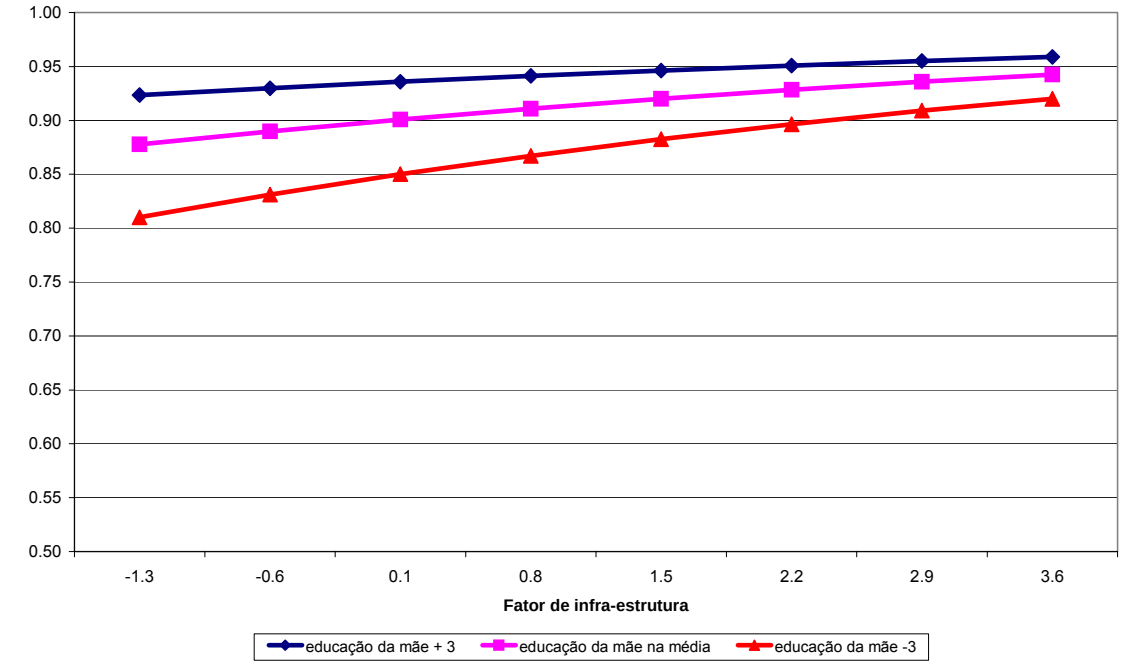
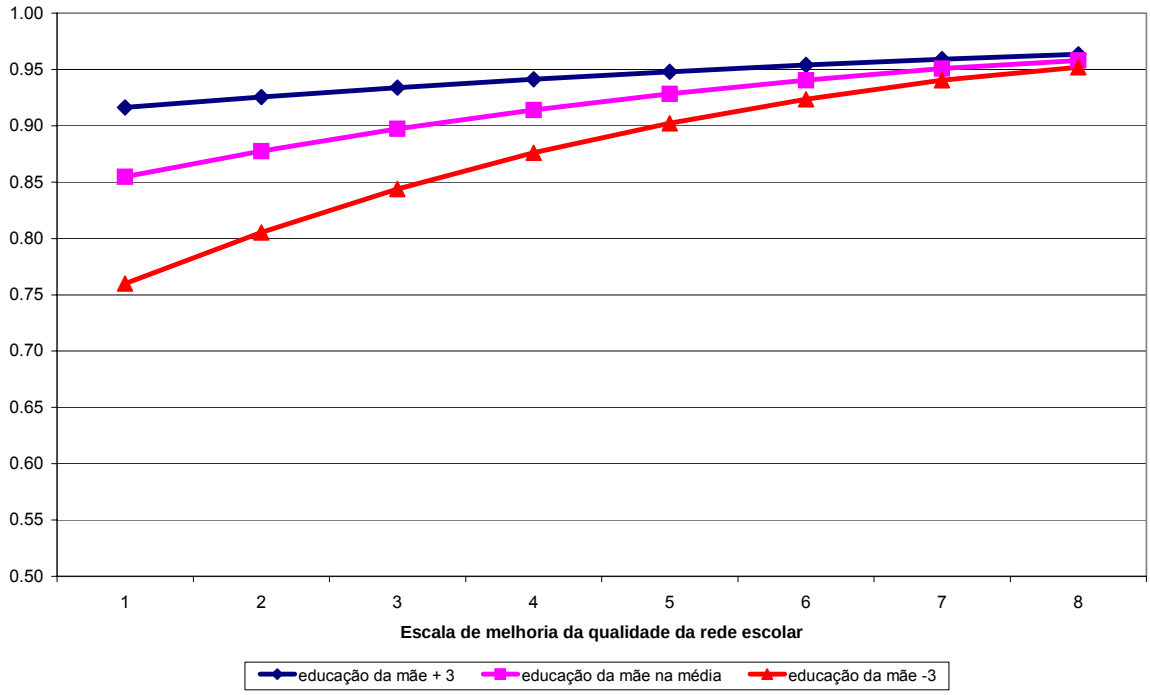


GRÁFICO 5. 18: Impacto da variação da educação da mãe e do fator de infra-estrutura na Probabilidade de Progressão na 1ª série (e_0)



5.4) Conclusão

Esse capítulo teve o objetivo de estudar os determinantes dos indicadores relacionados ao acesso e eficiência do sistema de ensino brasileiro, utilizando a modelagem hierárquica. A leitura dos resultados enfocou três pontos. Primeiro, a importância do dividendo demográfico nesses indicadores, ou seja, procurou investigar até que ponto a menor razão de dependência de jovens, conseqüência de uma menor taxa de fecundidade, pode aumentar as oportunidades educacionais brasileiras. O segundo ponto refere-se ao estudo da importância dos fatores da rede escolar dos municípios, principalmente aqueles relacionados com a qualidade dos recursos humanos e infra-estrutura. Por fim, procurou-se também determinar quais fatores relacionados ao perfil escolar dos municípios podem minimizar a importância dos fatores de *background* familiar na educação dos filhos. Esse último ponto é importante, já que os resultados obtidos para todos os indicadores demonstraram ser a educação materna o fator mais relevante. Tal resultado comprova a alta estratificação educacional existente no Brasil, no sentido de que a trajetória escolar do aluno está bastante relacionada com a sua origem social.

Com relação ao impacto do dividendo demográfico, 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 de idade. Já para a freqüência 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. Para a outra dimensão educacional – eficiência e rendimento – também se verificou o impacto favorável de uma menor pressão demográfica. Destaca-se também, que a disponibilidade de oferta mostrou-se importante em todos os indicadores do ensino médio, sinalizando a necessidade de políticas públicas que aumentam a oferta desse nível de ensino nos municípios brasileiros.

Como já mencionado, os resultados comprovam a grande importância da educação materna nos indicadores educacionais analisados nessa tese, evidenciando a desigualdade de oportunidades educacionais no Brasil. Em termos de políticas públicas é importante saber quais fatores podem minimizar a importância da origem social do aluno, a fim de tornar o sistema mais igualitário. Nesse sentido, os resultados mostraram dois pontos importantes. Primeiro, os fatores relacionados à rede escolar dos municípios, principalmente os de qualidade dos recursos humanos e infra-estrutura dos serviços educacionais, aumentam a probabilidade média de freqüentar a escola na idade correta e de progredir entre as séries, ou seja, melhora os indicadores de eficiência do sistema de ensino.

Segundo, existe um efeito substituição entre a educação materna e os fatores de qualidade da rede escolar dos municípios, de tal forma que uma melhora simultânea nesses fatores, torna as oportunidades educacionais igualitárias para os alunos, ou seja, o sucesso educacional, entendido aqui como a probabilidade de freqüentar a escola na idade correta e a probabilidade de progressão por série, passa a não depender da origem social do aluno.

6) INCLUSÃO DO CONTEXTO ESPACIAL NA ANÁLISE

O objetivo desse capítulo é inserir o contexto espacial na análise dos indicadores educacionais. Na análise anterior foram abordados os determinantes internos dos municípios nos indicadores educacionais. Entretanto, nessa análise, as características dos municípios foram tratadas como sendo independentes das características dos municípios vizinhos. Isso significa que foram ignoradas quaisquer dependências, seja do ambiente social, da restrição de oferta ou do perfil da rede escolar que poderia haver entre municípios vizinhos.

A autocorrelação espacial dos indicadores de acesso e eficiência, conseqüência da associação espacial tanto das variáveis observadas quanto das não-observadas (erro) pode tornar os coeficientes do modelo hierárquico enviesados e não-eficientes.

Dessa forma, fica evidenciada a importância de se verificar até que ponto os indicadores educacionais e seus determinantes estão sendo influenciados pelo contexto espacial mais amplo, que ultrapassa os limites dos municípios e tentar introduzir essa associação espacial nos modelos hierárquicos a fim de obter melhores estimativas dos parâmetros. Esse será o assunto do próximo capítulo, ressaltando-se que essa abordagem ainda não foi considerada nas análises dos determinantes educacionais. Pretende-se nesse trabalho fazer uma primeira tentativa de inserção da influência espacial nos estudos educacionais.

O diagnóstico para detectar os padrões de interação espacial será feito por meio de técnicas de ESDA aplicadas às variáveis independentes do nível 2 dos modelos hierárquicos visto na seção anterior, bem como à taxa de atendimento, taxa de eficiência e PPSs calculadas para os municípios. Essas três taxas representam as variáveis dependentes do modelo hierárquico, porém, como a análise espacial requer que as medidas sejam agregadas em unidades geográficas, elas são medidas em termos de taxas municipais ao invés de dicotômicas, como no modelo anterior. Por fim, foi feita uma análise de regressão espacial para detectar o verdadeiro processo espacial dos dados.

6.1) Resultados da Análise Exploratória Espacial

Inicialmente, foi calculada a estatística I de Moran com o objetivo de detectar autocorrelação espacial global nas variáveis. Os resultados para o ensino fundamental encontram-se na TAB. 6.1. Através deles foi possível perceber a presença de autocorrelação espacial positiva e significativa em todas as variáveis, sendo bastante alta para a taxa de eficiência, percentual de docentes com curso superior no município e tamanho relativo da coorte de 7 a 14 anos.

Na TAB. 6.2 encontram-se os resultados do I de Moran para o ensino médio. Da mesma forma que no caso anterior, detectou-se autocorrelação espacial positiva e significativa em todas as variáveis, porém, com valores em geral menores que no ensino fundamental. A taxa de escolarização líquida, a taxa de eficiência, o tamanho relativo da coorte de 15 a 17 anos e a escolaridade média da população adulta foram as variáveis que apresentaram os maiores valores para a estatística de autocorrelação espacial global.

TABELA 6. 1: Estatística I de Moran para as variáveis do Ensino Fundamental

Variáveis	I de Moran	Prob
INDICADORES EDUCACIONAIS		
Taxa de eficiência	0,890	0,000
Taxa de atendimento	0,579	0,000
Probabilidade de passar na 1ª séire (e_0)	0,570	0,000
Probabilidade de passar na 5ª séire (e_4)	0,654	0,000
1. PERFIL DAS ESCOLAS MUNICIPAIS		
1.1 Recursos Humanos Escolares		
Média de horas-aula diárias	0,770	0,000
Média de alunos por turma	0,420	0,000
Porcentagem de professores com curso superior	0,822	0,000
1.2 Infra-estrutura Escolar		
Fator de infra-estrutura	0,704	0,000
1.3 Restrição de Oferta		
Razão professor por população em idade escolar	0,428	0,000
1.4 Nucleação		
Proporção de estabelecimentos com 50 a 200 matrículas	0,206	0,000
Proporção de estabelecimentos com 201 a 1000 matrículas	0,530	0,000
Proporção de estabelecimentos com mais de 1000 matrículas	0,215	0,000
1.5 Distribuição das Matrículas Segundo Dependência Administrativa		
Proporção de matrículas públicas	0,173	0,000
Proporção de matrículas municipais dentre as públicas	0,509	0,000
1.6 Existência de Ciclos		
Percentual de matrículas em escolas com ciclos	0,671	0,000
2. CARACTERÍSTICAS DEMOGRÁFICAS		
2.1 Pressão Demográfica		
Tamanho relativo da coorte de 7 a 14 anos	0,811	0,000
2.2 Variáveis Categóricas para Tamanho da População do Município		
População de 0 a 4999	0,161	0,000
População de 5000 a 9999	0,030	0,000
População de 10000 a 19999	0,037	0,000
População de 20000 a 49999	0,053	0,000
População de50000 a 99999	0,014	0,000
População de100000 a 199999	0,052	0,000
Acima de 200000		
3 AMBIENTE COMUNITÁRIO		
Educação média da população adulta	0,721	0,000

TABELA 6. 2: Estatística I de Moran para as variáveis do Ensino Médio

Variáveis	I de Moran	Prob
INDICADORES EDUCACIONAIS		
Taxa de eficiência	0,751	0,000
Taxa de atendimento	0,392	0,000
Taxa de escolarização líquida	0,789	0,000
Probabilidade de progressão na 1ª série do ensino médio (e_8)	0,210	0,000
1. PERFIL DAS ESCOLAS MUNICIPAIS		
1.1 Recursos Humanos Escolares		
Média de horas-aula diárias	0,182	0,000
Média de alunos por turma	0,194	0,000
Porcentagem de professores com curso superior	0,431	0,000
1.2 Infra-estrutura Escolar		
Fator de infra-estrutura	0,492	0,000
1.3 Restrição de Oferta		
Razão professor por população em idade escolar	0,439	0,000
1.4 Nucleação		
Proporção de estabelecimentos com 50 a 200 matrículas	0,077	0,000
Proporção de estabelecimentos com 201 a 1000 matrículas	0,085	0,000
Proporção de estabelecimentos com mais de 1000 matrículas	0,100	0,000
1.5 Distribuição das Matrículas Segundo Dependência Administrativa		
Proporção de matrículas públicas	0,150	0,000
Proporção de matrículas municipais dentre as públicas	0,309	0,000
2. CARACTERÍSTICAS DEMOGRÁFICAS		
2.1 Pressão Demográfica		
Tamanho relativo da coorte de 15 a 17 anos	0,652	0,000
2.2 Variáveis Categóricas para Tamanho da População do Município		
População de 0 a 4999	0,161	0,000
População de 5000 a 9999	0,030	0,000
População de 10000 a 19999	0,037	0,000
População de 20000 a 49999	0,053	0,000
População de50000 a 99999	0,014	0,030
População de100000 a 199999	0,052	0,000
Acima de 200000	0,193	0
3 AMBIENTE COMUNITÁRIO		
Educação média da população adulta	0,721	0,000

Entretanto, como mencionado no Capítulo 4, seção 4.2.1, a estatística I de Moran é uma estatística global que não permite determinar possíveis *clusters* e *outliers* espaciais entre os municípios. As medidas locais de autocorrelação espacial são mais apropriadas para esse tipo de inferência. Nesse trabalho foram utilizadas duas medidas locais de autocorrelação espacial, o *Moran Scatterplot* e a estatística I de Moran local.

Os Mapas 6.1 a 6.36 representam graficamente essas duas medidas para as variáveis do estudo, sendo que a estatística Moran local (mapas a direita) inclui os valores municipais do *Moran Scatterplot* (mapas a esquerda) que foram significativos aos níveis de 1% e 5%.

Primeiramente, foi feita uma análise da distribuição geográfica da taxa de atendimento do ensino fundamental e médio e dos seus principais determinantes municipais, o tamanho relativo da coorte em idade escolar e a razão

professor/população em idade escolar, conforme detectado no modelo 5, da análise hierárquica.

Os Mapas 6.1 a 6.3 representam graficamente os valores do *Moran Scatterplot* e do Moran local para as variáveis do ensino fundamental. Pode-se observar pelos mapas do *Moran Scatterplot* a existência de *clusters* para as três variáveis analisadas, ou seja, municípios com altos (baixos) valores circundados por municípios com altos (baixos) valores, comprovando a existência de autocorrelação espacial positiva para essas variáveis.

Grosso modo, o Brasil pode praticamente ser dividido em dois: as regiões Sul, Sudeste e a parte sul do Centro-Oeste apresentando valores altos de taxa de atendimento do fundamental e professor/população em idade escolar (Alto-Alto) e baixos para o tamanho relativo da coorte de 7 a 14 anos (Baixo-Baixo); as regiões Norte, Nordeste e a parte norte do Centro-Oeste apresentando um quadro oposto. No que concerne à razão professor/população em idade escolar essa divisão do território Brasileiro não se dá de maneira tão evidente, com São Paulo, Paraná e Santa Catarina apresentando baixos valores. Interessante destacar também a presença de alguns *outliers* em todas as regiões – valores baixos rodeados por valores altos, ou vice-versa.

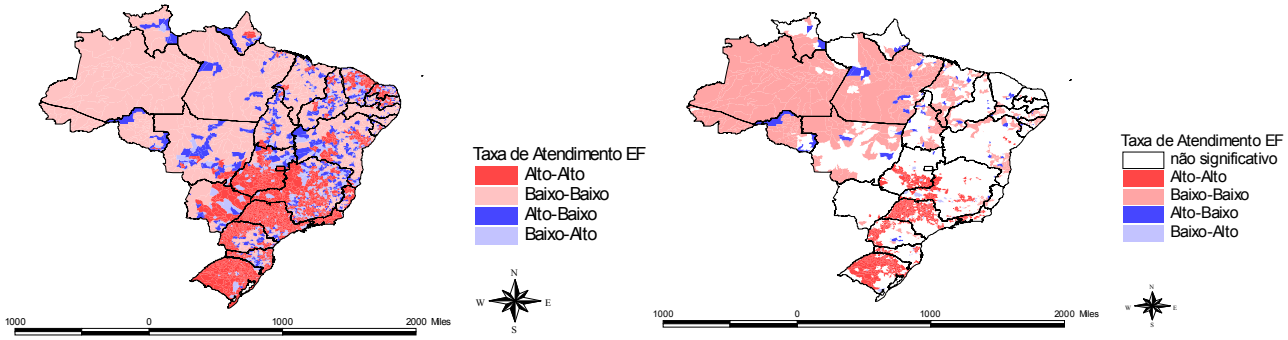
Como pode ser percebido pelos mapas do *Moran Scatterplot* os valores das variáveis para municípios pertencentes a um mesmo estado são próximos e correlacionados, ou seja, valores baixos rodeados por valores baixos ou vice-versa. Tal resultado sugere que os fatores contextuais dos estados exercem forte influência nas variáveis analisadas. Porém, quando se analisam apenas os valores significativos - I de Moran Local – percebe-se que a hegemonia dos estados desaparece, ou seja, surgem alguns aglomerados de municípios dentro de um mesmo estado, destacando-se positivamente ou negativamente, principalmente nos estados do Nordeste, Sudeste e Sul. Cabe ressaltar, que os resultados das significâncias são sempre relativos aos vizinhos, ou seja, o mesmo valor pode ser considerado alto num lugar e baixo em outro.

A investigação visual dessas três taxas confirma a existência de autocorrelação espacial, além de sugerir que a taxa de atendimento tem uma associação espacial positiva com a razão professor/população em idade escolar e negativa com o tamanho relativo da coorte em idade escolar. No caso da restrição de oferta, pode-se inferir que uma maior disponibilidade de oferta escolar dos municípios vizinhos possui uma

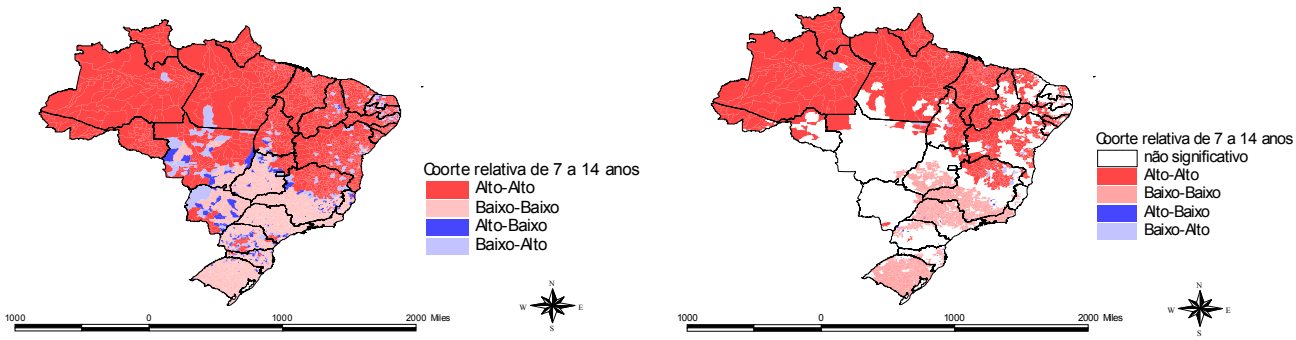
externalidade espacial positiva com os indicadores de acesso de determinado município, no sentido de atender parte dos seus alunos.

Por outro lado, o tamanho relativo da coorte de 7 a 14 anos tem uma autocorrelação espacial negativa com a taxa de atendimento. Esse resultado é esperado já que essa variável está muito associada com as condições socioeconômicas dos municípios. Pode-se pensar também que determinado município rodeado por vizinhos com maior proporção de pessoas em idade escolar possa incorrer em maior prejuízo no atendimento de seus residentes se parte da população vizinha for atendida por seu sistema de ensino.

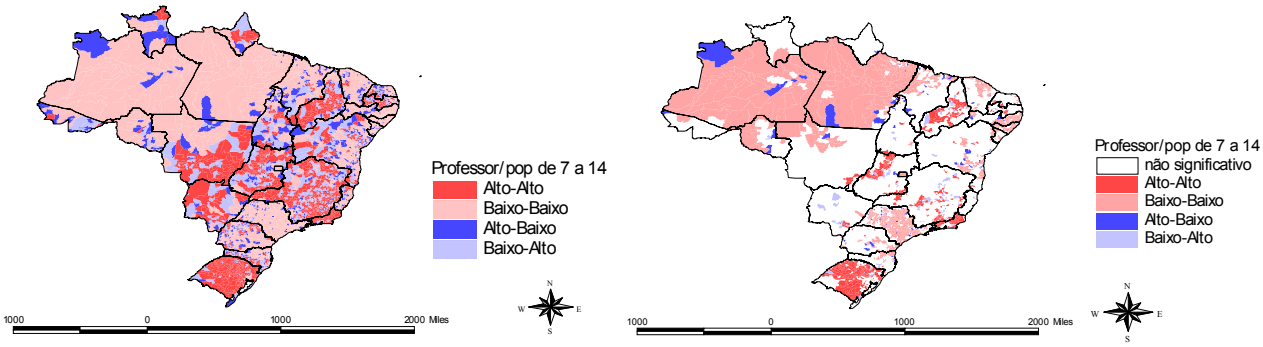
MAPA 6. 1: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a taxa de atendimento do ensino fundamental



MAPA 6. 2: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para o tamanho relativo da coorte de 7 a 14 anos



MAPA 6. 3: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a relação professor do nível fundamental/população de 7 a 14 anos



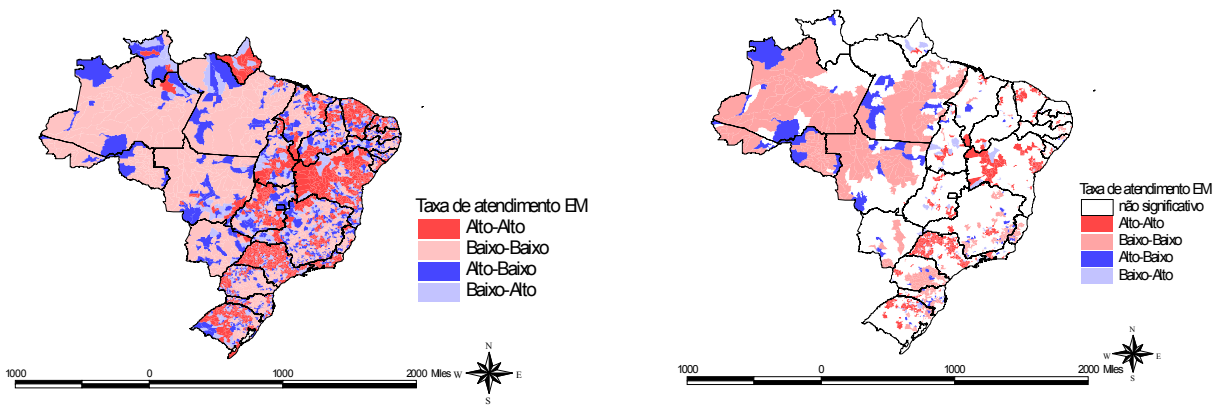
Com relação ao ensino médio, a inspeção visual da taxa de atendimento confirma a ambigüidade desse indicador já comentada no Capítulo 1. O padrão espacial encontrado no MAPA 6.4 mostra que os valores altos-altos são predominantes nas regiões Nordeste e Sudeste. Como já comentado, o fato desse valor ser mais alto para a região Nordeste que a Sul, pode ser atribuído a maior repetência na primeira região, onde muitos jovens ainda cursam a escola porque não terminaram nem o ensino fundamental, ao passo que no Sul, como a repetência é menor e a evasão nessa faixa etária ainda é alta, os jovens terminam o fundamental e evadem. Dessa forma, optou-se por fazer a análise espacial para a taxa de escolarização líquida do ensino médio, que representa a forma agregada da variável dependente do modelo hierárquico do capítulo anterior para a freqüência à escola da amostra referente à população de 15 a 17 anos que já concluiu o ensino fundamental, que foi medida de forma dicotômica no capítulo anterior.

Analisando-se, portanto, conjuntamente os mapas referentes à taxa de escolarização líquida, a relação professor/população em idade escolar[18] e o tamanho relativo da coorte de 15 a 17 anos, MAPAS 6.5 a 6.7, percebe-se que, da mesma forma que no fundamental, as variáveis estão aglomeradas e com o mesmo padrão espacial, ou seja, predominam valores alto-alto nas regiões Sudeste e Sul e baixo-baixo nas demais regiões, nas duas primeiras variáveis, o contrário acontecendo para a última variável. Os mapas também sugerem que a taxa de escolarização líquida possui uma autocorrelação

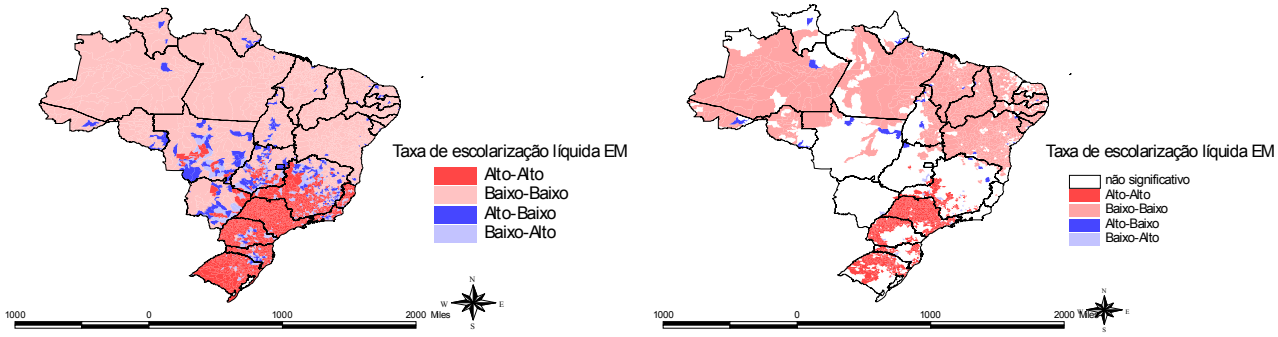
[18] Em 2000 havia 278 municípios sem escolas que oferecessem o nível médio. Como a análise espacial requer valores para todas as unidades georeferenciadas, optou-se por colocar zero para a relação professor/população em idade escolar nesses municípios, já que realmente não existem professores lecionando no nível de ensino médio.

espacial positiva com a relação professor/população em idade escolar, em outras palavras, a freqüência dos jovens de 15 a 17 anos no ensino médio em determinado município pode estar sendo influenciada não apenas pela oferta desse nível de ensino no próprio município, mas também pela dos municípios adjacentes. Para o tamanho relativo da coorte, parece existir autocorrelação espacial negativa com a taxa de escolarização líquida. Como já comentado, uma maior razão de dependência de jovens está associada ao nível de desenvolvimento das localidades, o que pode explicar tal associação negativa entre o tamanho relativo da coorte e a taxa de escolarização líquida do ensino médio.

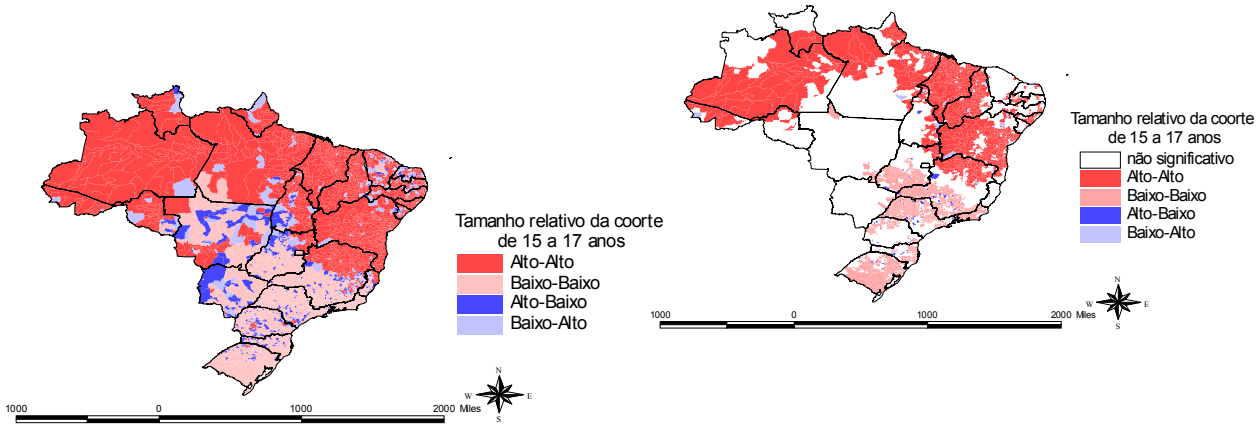
MAPA 6. 4: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a taxa de atendimento do ensino médio



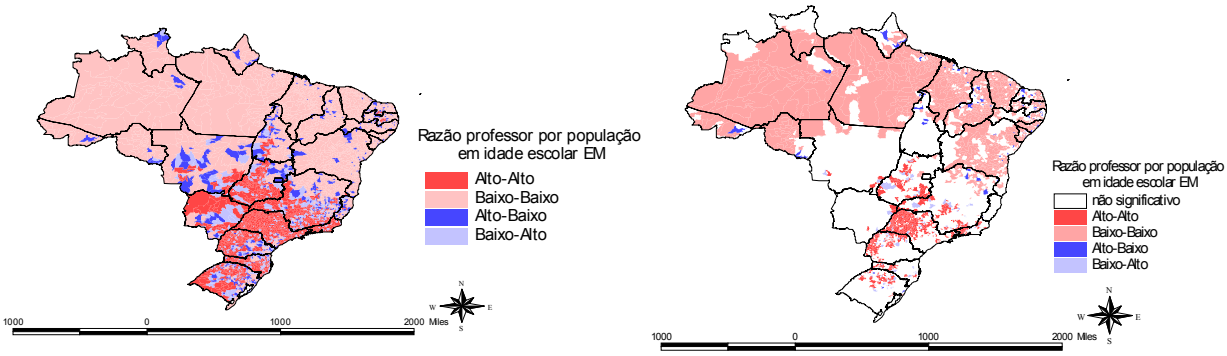
MAPA 6. 5: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a taxa de escolarização líquida do ensino médio



MAPA 6. 6: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para o tamanho relativo da coorte de 15 a 17 anos



MAPA 6. 7: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a relação professor do nível médio/população de 15 a 17 anos

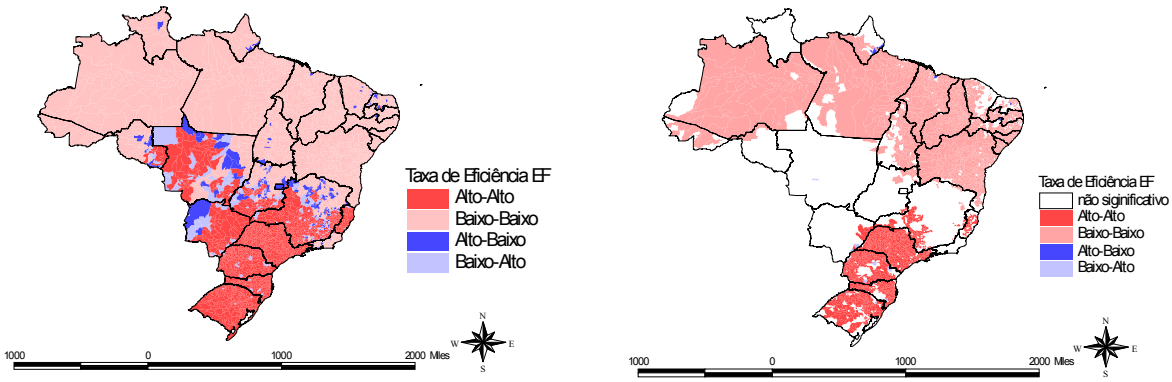


A análise seguinte contempla os diagnósticos para a taxa de eficiência, para as PPSs e para as demais variáveis independentes do nível dois do modelo hierárquico da seção anterior. Para facilitar a exposição, a análise é feita separadamente por nível de ensino, iniciando-se pelo fundamental.

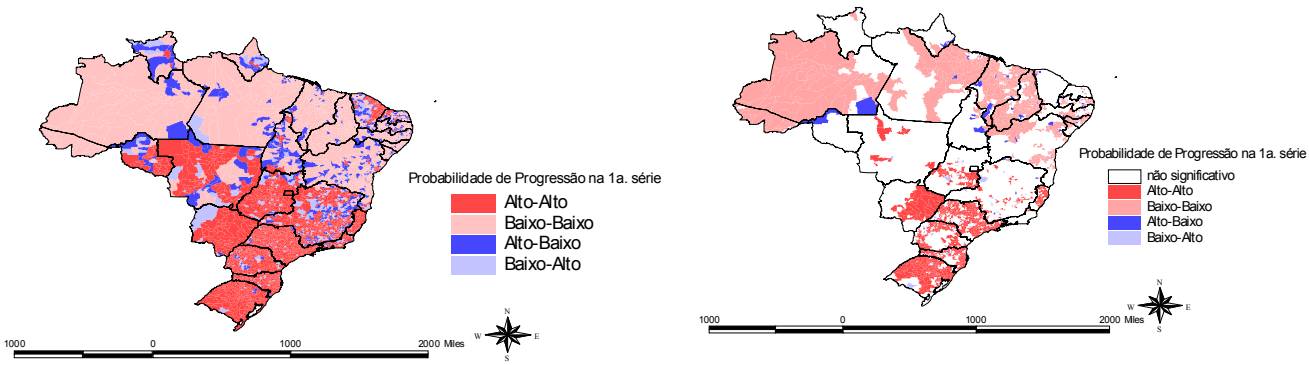
Através da inspeção visual do *Moran Scatterplot* dos MAPAS 6.8 a 6.10 percebe-se que as três taxas de resultado educacional (taxa de eficiência do fundamental, probabilidade de progressão na 1º série - e_0 - e probabilidade de progressão na 5º série - e_4), de um modo geral, possuem o mesmo padrão espacial observado para a taxa de atendimento. Interessante notar, o caso da taxa de eficiência do fundamental, que apresenta poucos *outliers* nas regiões Norte, Nordeste e Sul, ou seja, valores baixos (altos) rodeados por altos (baixos).

Quando se analisa apenas o valor significativo – I de Moran Local -, observa-se que a hegemonia dos estados permanece forte para a taxa de eficiência, ou seja, quase a totalidade dos municípios de um mesmo estado apresentam valores semelhantes para essa taxa ou valores não significativos. Dessa forma, as manchas vermelhas escuras (Alto-Alto), vermelhas claras (Baixo-Baixo) ou brancas (valores não significativos) cobrem quase a totalidade dos estados. Para as PPSs já existem aglomerados de municípios com valores significativos se destacando dentro de um mesmo estado, ou seja, as manchas cobrem apenas uma parte do estado, o que pode ser um indício da existência de algum fator contextual de diferenciação em relação ao restante do estado.

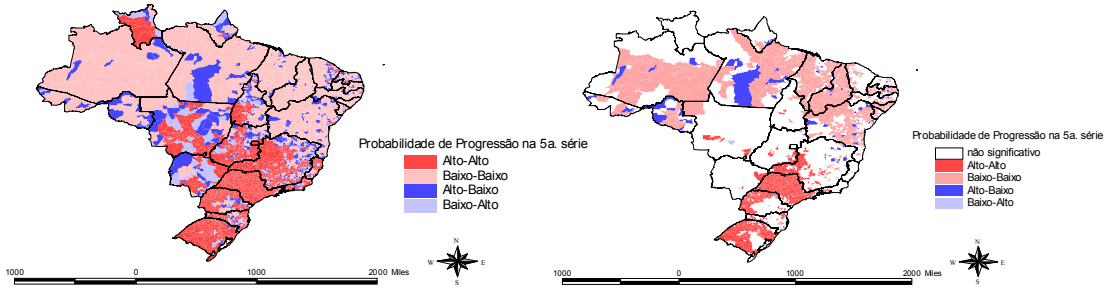
MAPA 6. 8: ***Moran Scatterplot*** **(esquerda) e I de Moran local (direita) para a taxa de eficiência do ensino fundamental**



MAPA 6. 9: ***Moran Scatterplot*** **(esquerda) e I de Moran local (direita) para a probabilidade de passar de zero para um ano de estudo (e0)**

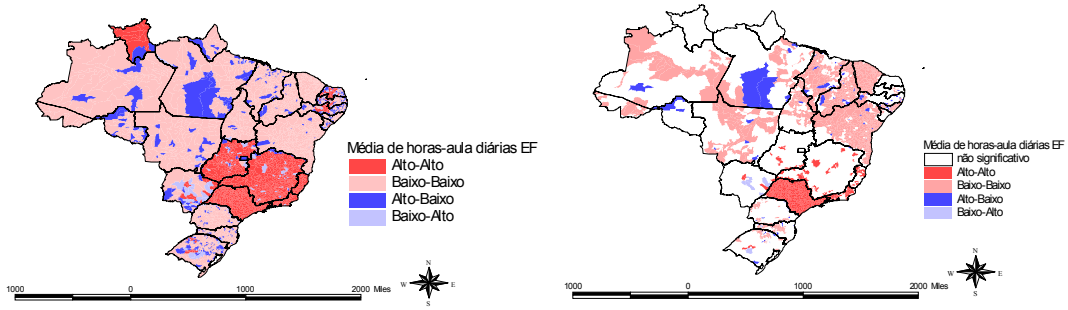


MAPA 6. 10: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a probabilidade de passar de quatro para cinco anos de estudo (e_4)

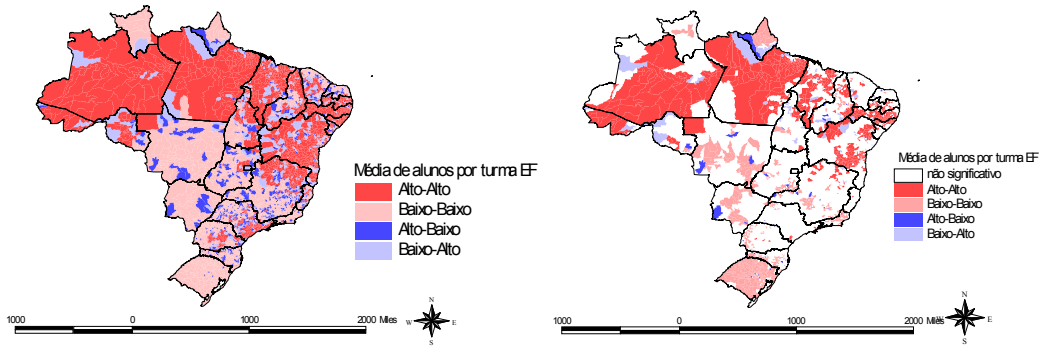


Nos MAPAS 6.11 a 6.13 estão representados o *Moran Scatterplot* e o I de Moran local para as variáveis de recursos humanos escolares do nível fundamental. Observa-se que a média de horas-aula diária apresenta um padrão espacial um pouco diferenciado, com os estados de Roraima e Goiás apresentando valores altos rodeados por valores altos (manchas vermelhas), além dos estados da região Sudeste. Já a média de alunos por turma e a porcentagem de professores com curso superior apresentam o padrão semelhante ao observado nas variáveis analisadas anteriormente. Os resultados também sugerem que a porcentagem de professores com curso superior possui uma autocorrelação espacial positiva com a taxa de eficiência, e_0 e e_4, enquanto que a média de alunos por turma apresenta autocorrelação espacial negativa com essas variáveis.

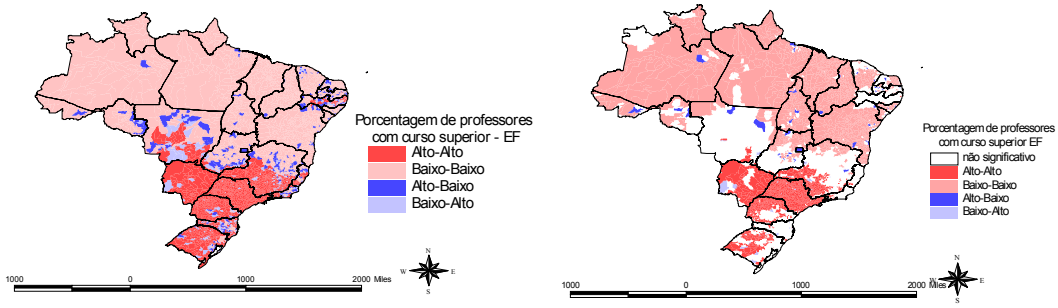
MAPA 6. 11: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a média de horas-aula diária no ensino fundamental



MAPA 6. 12: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a média de alunos por turma no ensino fundamental

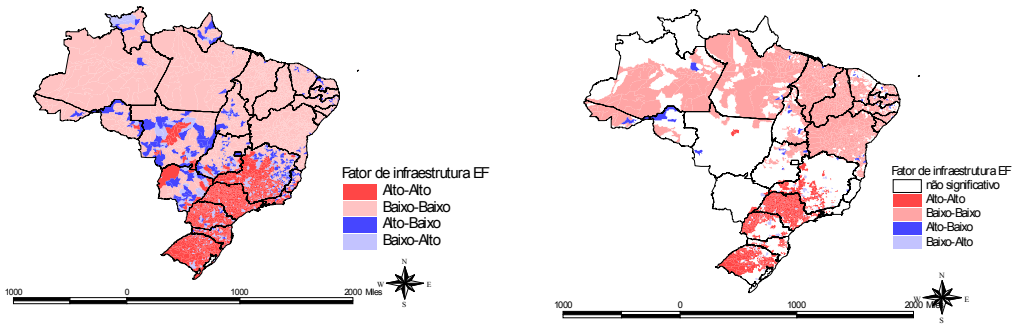


MAPA 6. 13: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a porcentagem de professores com curso superior do ensino fundamental



Com relação ao fator de infra-estrutura da rede escolar do fundamental, este também se apresenta aglomerado com maiores valores nas regiões Sul e Sudeste. Dessa forma, os mapas também sugerem, que há uma autocorrelação espacial positiva entre o fator de infra-estrutura e a taxa de eficiência, probabilidade de passar na 1ª série (e_0) e probabilidade de passar na 5ª série (e_4).

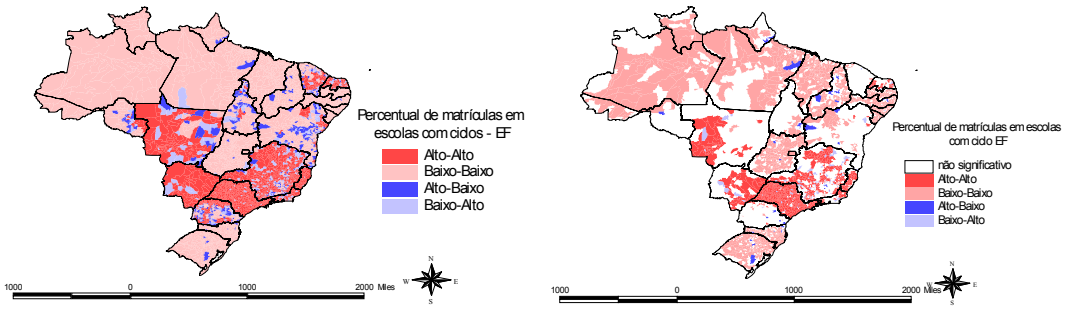
MAPA 6. 14: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para o fator de infra-estrutura do ensino fundamental



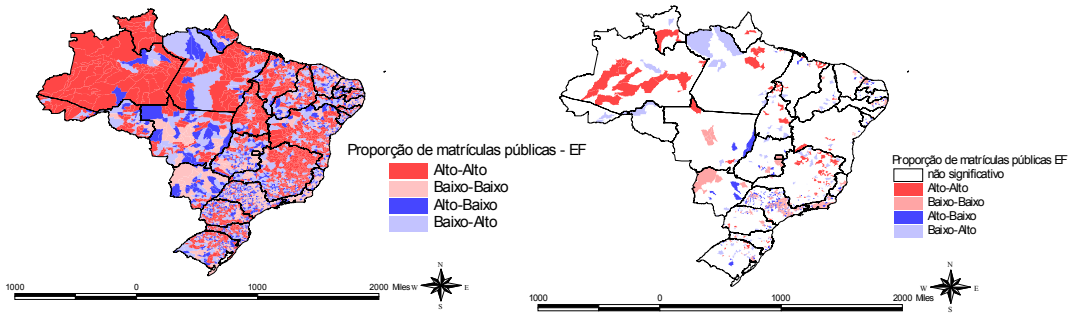
Com relação às demais variáveis, os MAPAS 6.15 a 6.23 mostram que apesar de existirem aglomerações, com valores baixo-baixo e alto-alto, estas se dão de forma menos acentuada que nas variáveis analisadas anteriormente. Destaca-se o percentual de matrículas em escolas públicas que é espacialmente heterogêneo.

Ainda com relação aos MAPAS 6.15 a 6.23, os principais resultados encontrados são: i) as regiões Sudeste e Centro-Oeste apresentam maior proporção de matrículas em escolas com ciclos; ii) a proporção de matrículas municipais entre as públicas concentra-se nas regiões Nordeste e parte da Norte; iii) as escolas de pequeno porte, ou seja, com até 50 matrículas prevalecem nas regiões Norte e Nordeste e as de grande porte, com mais de 1000 matrículas, são bastante escassas nos municípios brasileiros e iiii) a educação média da população adulta é maior para os municípios da região Sudeste, Sul e Centro-Oeste.

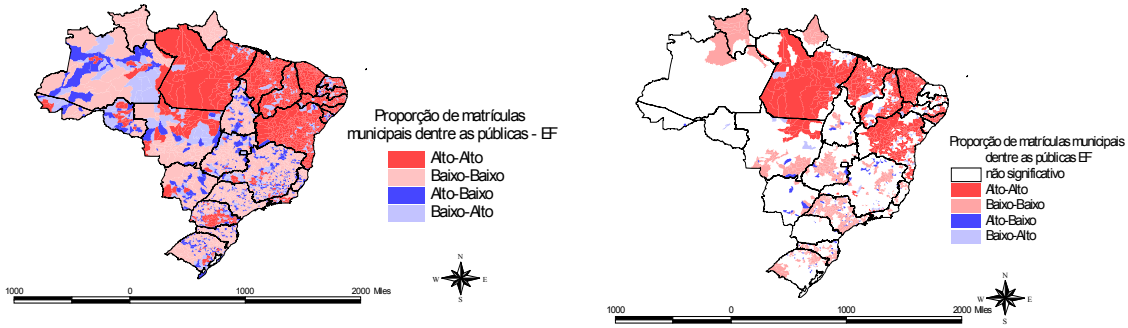
MAPA 6. 15: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para o percentual de matrículas em escolas com ciclos do ensino fundamental



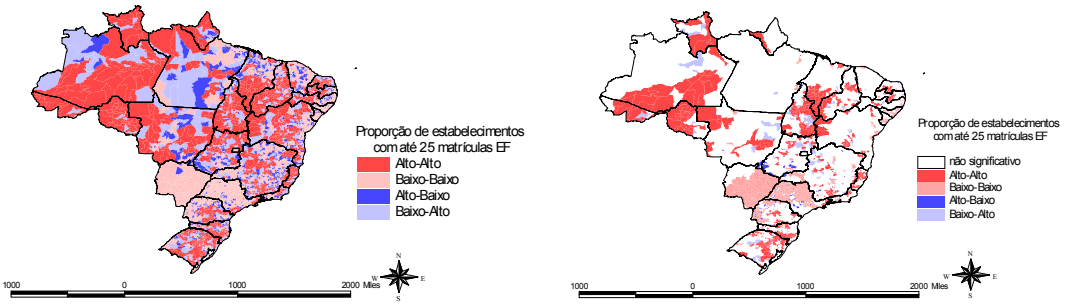
MAPA 6. 16: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a proporção de matrículas públicas do ensino fundamental



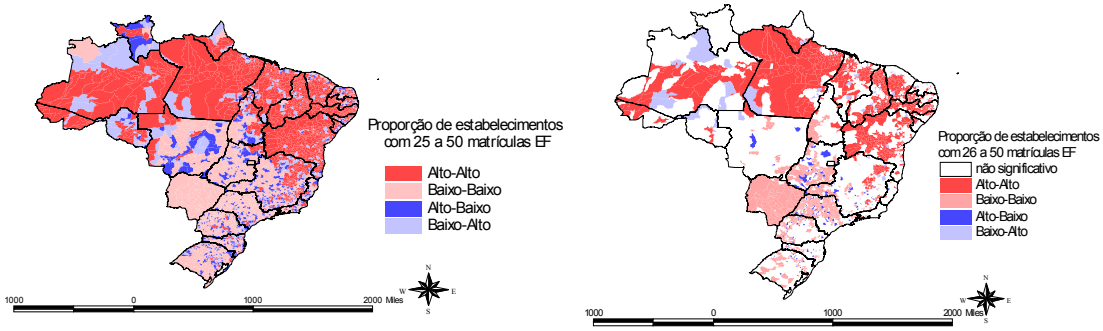
MAPA 6. 17: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a proporção de matrículas municipais dentre as públicas do ensino fundamental



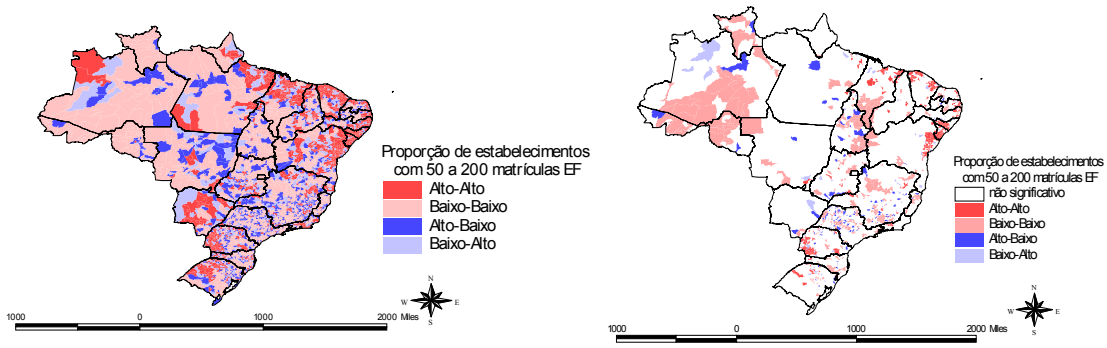
MAPA 6. 18: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a proporção de estabelecimentos com até 25 matrículas no ensino fundamental



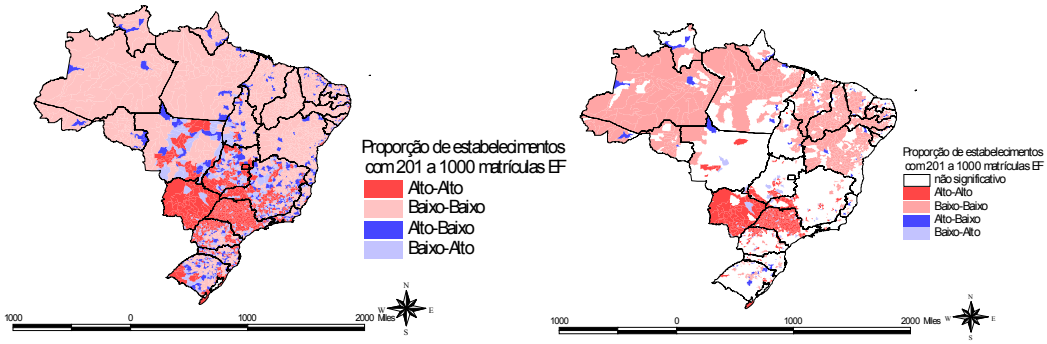
MAPA 6. 19: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a proporção de estabelecimentos com 25 a 50 matrículas no ensino fundamental



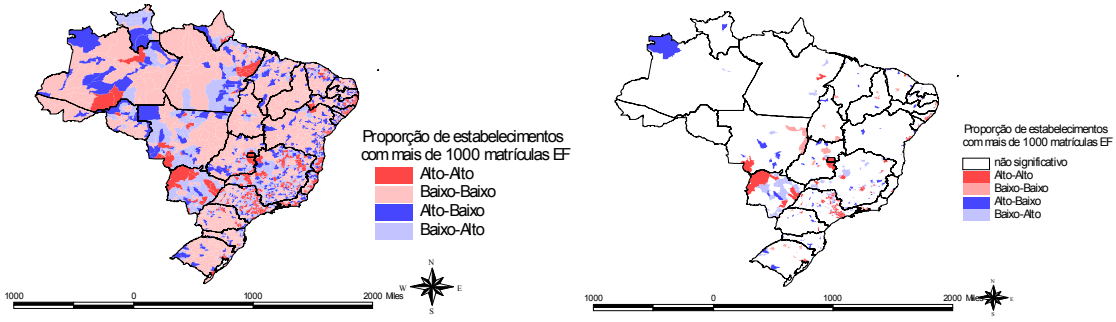
MAPA 6. 20: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a proporção de estabelecimentos com 50 a 200 matrículas no ensino fundamental



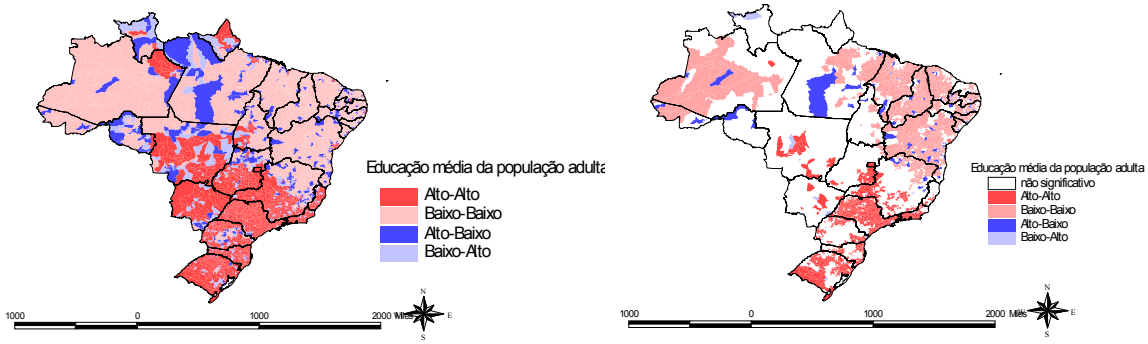
MAPA 6. 21: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a proporção de estabelecimentos com 201 a 1000 matrículas no ensino fundamental



MAPA 6. 22: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a proporção de estabelecimentos com mais de 1000 matrículas no ensino fundamental



MAPA 6. 23: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a educação média da população adulta



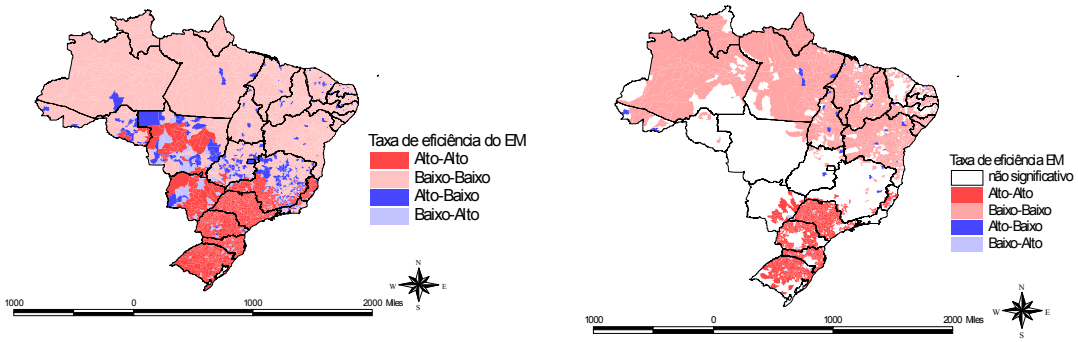
Através da análise da taxa de eficiência do ensino médio e da probabilidade de progressão de oito para nove anos de estudo (e_8) - MAPAS 6.24 e 6.25, percebe-se que apenas a primeira possui a distribuição padrão verificada nas taxas anteriores, enquanto que em e_8 os valores alto-alto concentram-se nos municípios do Sudeste, Nordeste e Centro-Oeste.

Os mapas restantes referem-se às variáveis independentes do nível dois do modelo hierárquico para o nível médio. Da mesma forma que na relação professor/população em idade escolar, nos municípios onde não havia estabelecimentos oferecendo esse nível de ensino em 2000 foi imputado o valor zero para as variáveis relacionadas à rede escolar do município. Tal procedimento foi necessário já que a análise espacial requer valores para todas as unidades georeferenciadas.

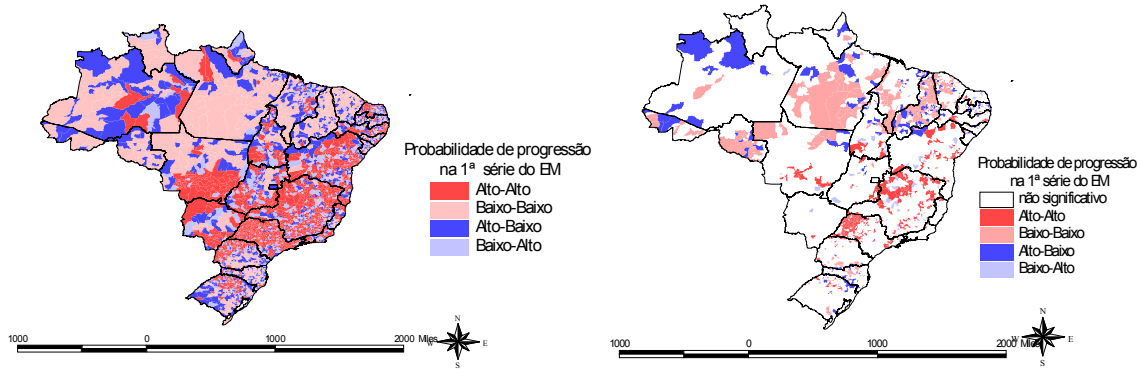
O principal resultado evidenciado pelos mapas é que a porcentagem de professores com curso superior e principalmente o fator de infra-estrutura da rede escolar são as variáveis que mais se assemelham com o padrão de aglomeração espacial da taxa de eficiência. Destaca-se também que quando são considerados os valores significativos do I de Moran Local a hegemonia dos estado quase não existe e surgem aglomerados de municípios destacando-se dentro dos estados. Ou em outras palavras pode-se dizer que não há manchas com valores semelhantes cobrindo a totalidade do estado, mas cobrindo apenas alguns municípios dentro do estado, sugerindo que há algum fator contextual ou de política do município que o diferencie do restante do estado.

Destaca-se também: i) a porcentagem de escolas públicas, a porcentagem de escolas municipais entre as públicas e as variáveis de nucleação são mais homogêneas que as demais; ii) os estados da região Nordeste são os que possuem a maior proporção de matrículas municipais dentre as públicas e iii) os estabelecimentos de menor porte concentram-se nas regiões Norte e Nordeste, enquanto os de maior porte (acima de 200 matrículas) são mais freqüentes no Mato Grosso do Sul e São Paulo..

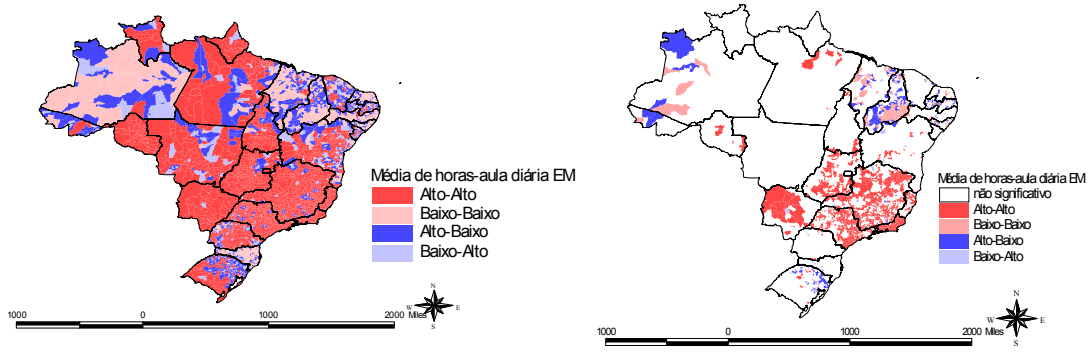
MAPA 6. 24: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a taxa de eficiência do ensino médio



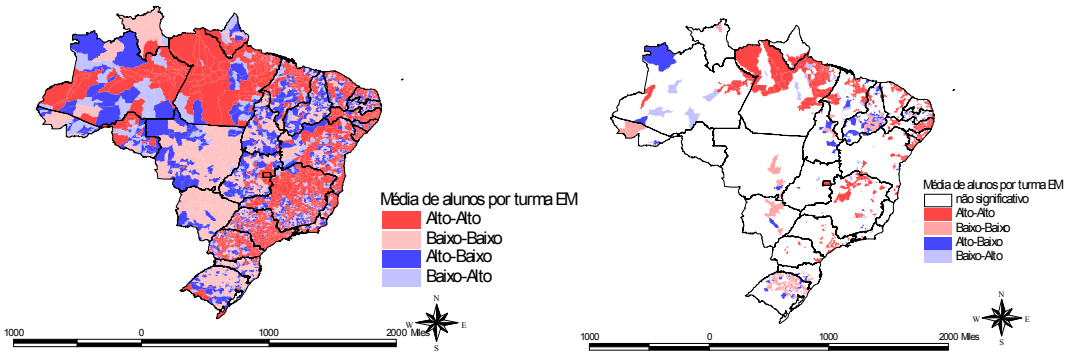
MAPA 6. 25: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a probabilidade de progressão na 1ª serie do ensino médio(e8)



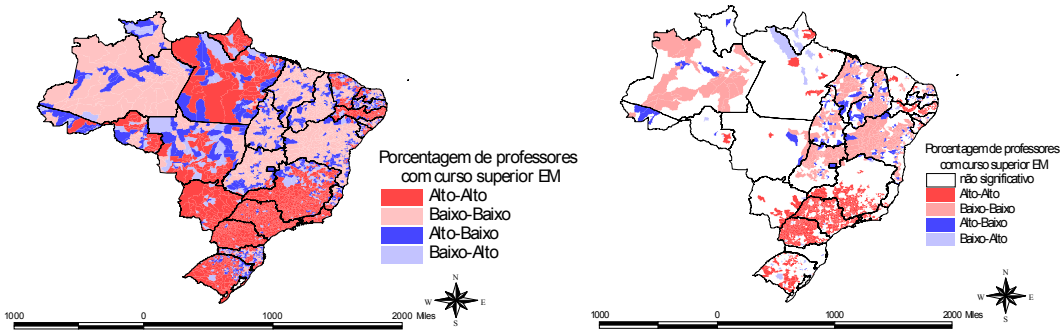
MAPA 6. 26: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a media de horas-aula diária do ensino médio



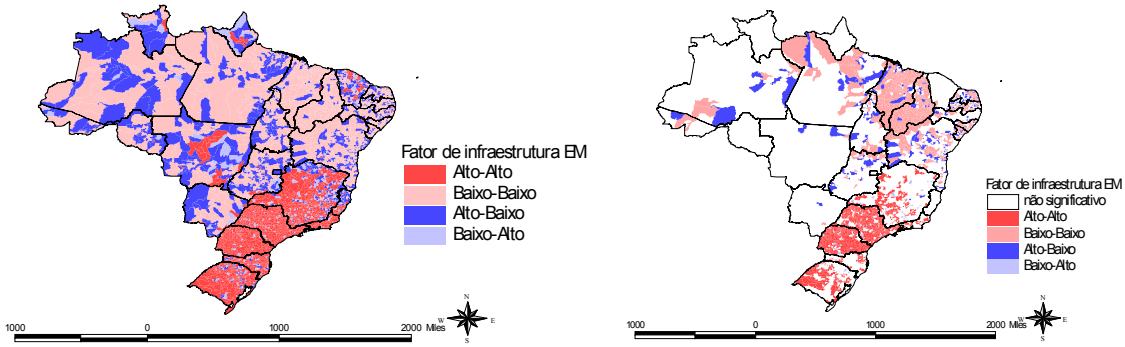
MAPA 6. 27: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a media de alunos por turma do ensino médio



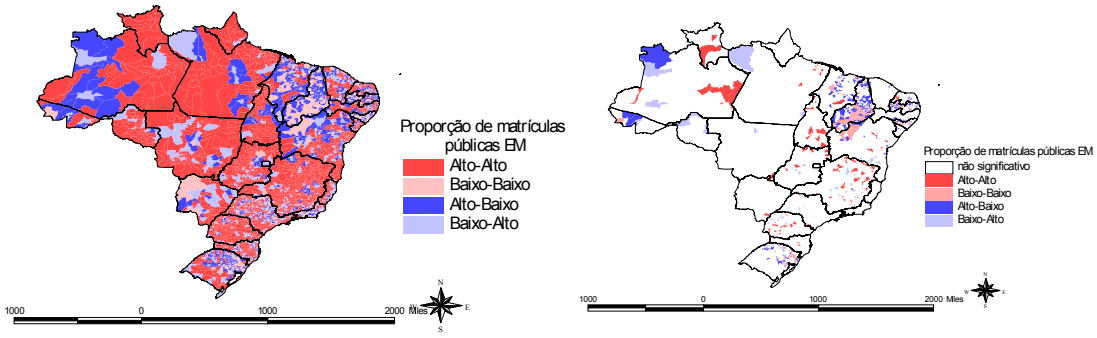
MAPA 6. 28: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a porcentagem de professores com curso superior do ensino médio



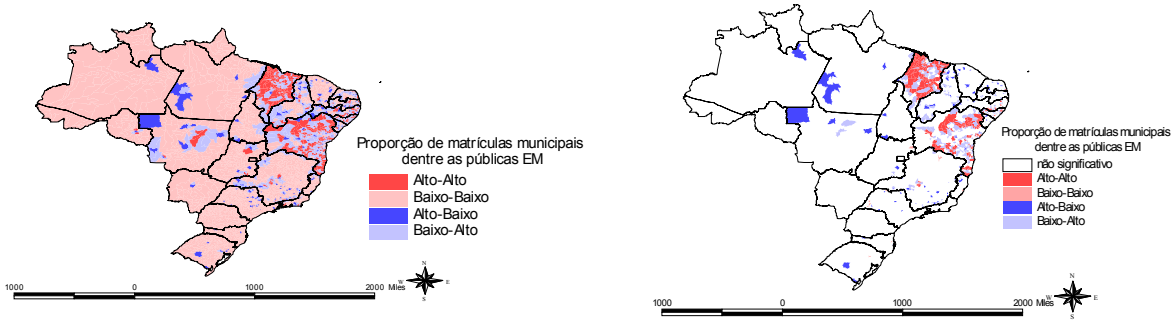
MAPA 6. 29: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para o fator de infra-estrutura do ensino médio



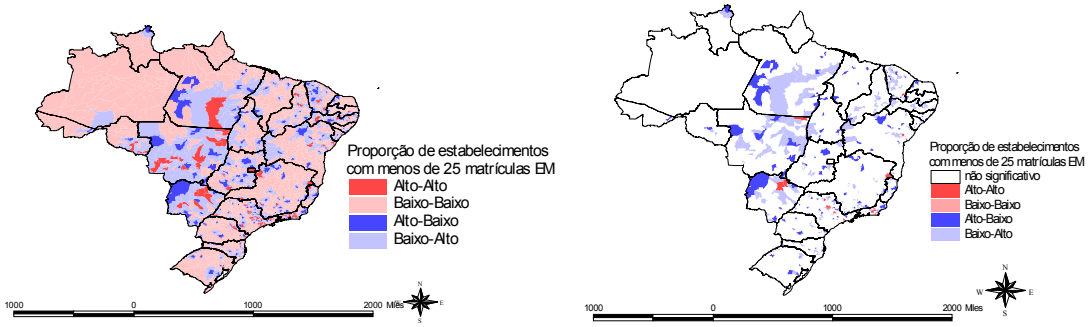
MAPA 6. 30: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a proporção de matrículas públicas no ensino médio



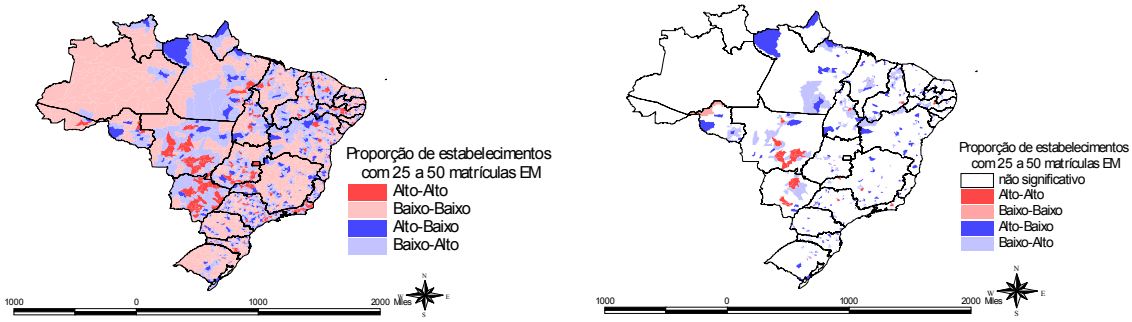
MAPA 6. 31: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a proporção de matrículas municipais dentre as públicas no ensino médio



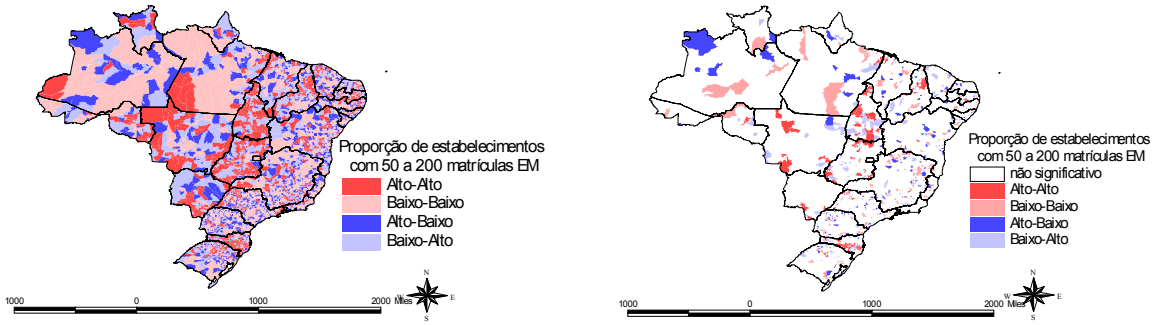
MAPA 6. 32: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a proporção de estabelecimentos com menos de 25 matrículas no ensino médio



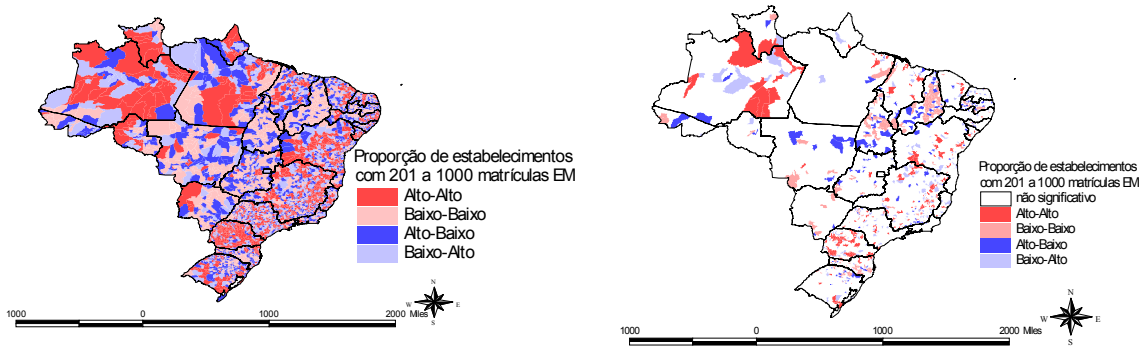
MAPA 6. 33: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a proporção de estabelecimentos com 25 a 50 matrículas no ensino médio



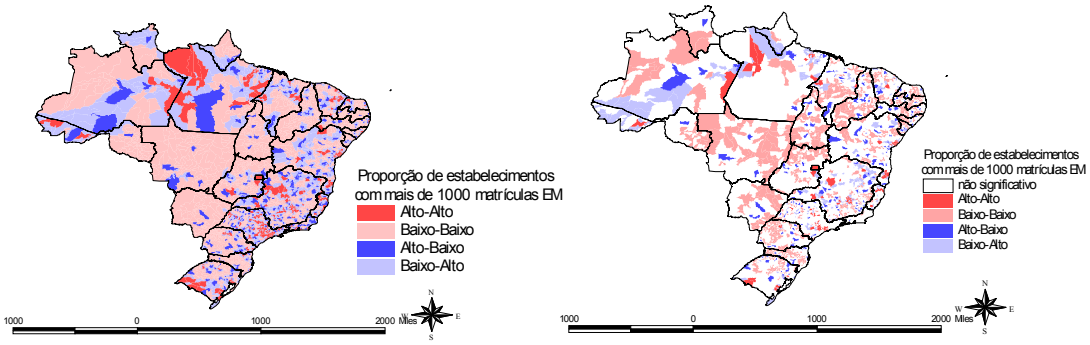
MAPA 6. 34: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a proporção de estabelecimentos com 50 a 200 matrículas no ensino médio



MAPA 6. 35: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a proporção de estabelecimentos com 201 a 1000 matrículas no ensino médio



MAPA 6. 36: *Moran Scatterplot* (esquerda) e I de Moran local (direita) para a proporção de estabelecimentos com mais de 1000 matrículas no ensino médio



6.3) Regressão Espacial

As representações gráficas do *Moran Scatterplot* e do I Moran Local na seção anterior, revelaram a existência de autocorrelação espacial nos indicadores de acesso e eficiência; cogitou-se a hipótese de que essa dependência espacial poderia ser devida a externalidades positivas e negativas dos seus preditores, ou seja, à influência das variáveis independentes dos municípios vizinhos. Porém, uma hipótese alternativa para essa autocorrelação espacial positiva é que ela existe apenas porque seus preditores são também espacialmente aglomerados. Se essa hipótese for verdadeira, quando os indicadores de acesso e eficiência forem regredidos controlando-se pelas variáveis independentes, a autocorrelação espacial deixará de ser significativa. Por outro lado, caso a autocorrelação espacial persista isso significa que realmente existe uma influência das covariáveis dos municípios vizinhos (MORENOFF, 2003, p. 1004-1005).

Para testar essas conjecturas, foi feito um exercício econométrico espacial, para detectar o processo espacial em curso. Esse exercício foi feito com os resíduos do nível 2 do modelo hierárquico quando são consideradas apenas as variáveis de características do indivíduo e de *background* familiar, ou seja, os resíduos do nível 2 do Modelo 1 da análise hierárquica feita no Capítulo 5[19]. Apenas para relembrar, esse modelo considera a existência de um termo de erro no intercepto e no coeficiente da educação materna (para alguns casos) associado ao nível 2. Nesse modelo mais simples não considerou a existência de nenhuma covariável de nível 2 para explicar essa aleatoriedade. Em

[19] O HLM gera dois tipos de resíduo: *Bayesiano* e MQO. A análise foi feita para os dois resíduos.

seguida, estimou-se uma regressão linear através do método dos MQO no programa *SpaceStat*, tendo como variáveis dependentes esses resíduos e como independentes as variáveis do nível 2 do modelo hierárquico. Esse procedimento teve o objetivo de detectar o tipo de processo espacial presente nos dados, se autoregressivo de defasagem espacial ou de erro espacial. Por fim, estimou-se um modelo espacial, considerando que o processo espacial é de defasagem espacial[20]. Cabe ressaltar que na maioria dos casos, foram detectados os dois tipos de processo espacial, com maior intensidade para o de defasagem espacial. Porém, para facilitar a análise, considerou-se a existência apenas do processo de defasagem espacial.

Os resultados das regressões para os resíduos do Modelo 1 da análise hierárquica para a probabilidade de freqüentar a escola das duas faixas etárias encontram-se nas TABs. 6.3 a 6.5[21]. Para a amostra das crianças de 7 a 14 anos, o modelo de regressão incluiu as mesmas variáveis de nível 2 do Modelo 5 da análise hierárquica do capitulo anterior, e para a faixa etária de 15 a 17 anos, as mesmas do Modelo 6.

Os testes de diagnóstico para autocorrelação espacial realizados nas regressões lineares clássicas revelam a sua existência, sendo mais forte para o processo de defasagem espacial. Já que o teste de multiplicador de lagrange robusto é significativo tanto para o modelo de erro espacial quanto para o de defasagem espacial, porém o seu valor é maior para o segundo caso. A estimação do modelo de defasagem espacial confirma que o coeficiente do termo WY, ou seja, o parâmetro autorregressivo espacial (ρ), é altamente significativo em todos os casos.

Considerando primeiramente a probabilidade dos indivíduos de 7 a 14 anos freqüentarem a escola (TAB. 6.3), o coeficiente do termo de defasagem espacial (WY) foi igual a 0,61 para o resíduo *bayesiano*, muito próximo ao do resíduo MQO, que foi de 0,614. Esses coeficientes significam que a taxa de variação na freqüência escolar média das crianças de 7 a 14 anos em determinado município está associada a mudanças nas variáveis observadas (incluídas no modelo) e das não-observadas nos municípios

[20] O teste de diagnóstico para a normalidade dos resíduos mostrou que eles não são normais, dessa forma utiliou-se o processo de mínimo quadrado de dois estágios (2SLS), para estimar o modelo. Nessa abordagem, as variáveis X defasadas (WX) servem como instrumento para o termo de defasagem de Y (WY), para remover a correlação dessa última com o termo de erro (ANSELIN, 1995).

[21] No caso do ensino médio, para obter os resíduos do nível 2 para todos os municípios, as regressões hierárquicas foram feitas imputando o valor zero nas variáveis relacionadas aos fatores escolares para os municípios que não possuem estabelecimentos que ofertem esse nível de ensino (278 municípios). Os resultados não diferem muito do anterior, conforme pode ser visto nas TABs. A3.1 a A3.4 do Anexo 3. Nas regressões lineares e espaciais desse capítulo foram acrescentada uma variável *dummy* para identificar esses municípios.

vizinhos. Eles também indicam a magnitude do multiplicador espacial, ou seja, o coeficiente do termo de defasagem espacial vezes o da variável independente X, representa o efeito na freqüência escolar proveniente da variação de uma unidade da variável X dos seus vizinhos. Importante notar, que quando se considera no modelo o termo de defasagem espacial, as demais variáveis têm seu efeito diminuído, principalmente o tamanho relativo da coorte.

Com relação à freqüência escolar dos indivíduos de 15 a 17 anos (TAB. 6.4) e dos indivíduos dessa mesma faixa etária que já concluíram o ensino fundamental (TAB. 6.5), a mesma análise pode ser feita, destacando-se que o termo de defasagem espacial possui um efeito maior no primeiro caso. Para o último caso, a inclusão da defasagem espacial não diminui muito o efeito das outras variáveis. Finalmente, observa-se também que o teste de diagnóstico aponta para a existência de erro espacial mesmo após a introdução do termo de defasagem espacial, com exceção para o resíduo *bayesiano* da freqüência escolar condicional dos jovens de 15 a 17 anos.

TABELA 6.3: Estimativas de regressão para os resíduos de nível 2 para a probabilidade de freqüentar a escola – indivíduos de 7 a 14 anos (resíduos do modelo hierárquico que utiliza apenas covariáveis de nível 1)

Variáveis	Modelo MQO				Modelo de defasagem espacial - 2SLS			
	Resíduo Bayesiano		Resíduo MQO		Resíduo Bayesiano		Resíduo MQO	
	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig
Constante	0.968	0.000	1.185	0.000	0.146	0.113	0.120	0.310
Termo de defasagem espacial (Wy)	**-**	**-**	**-**	**-**	**0.610**	**0.000**	**0.614**	**0.000**
Tamanho relativo da coorte de 7 a 14 anos	-10.088	0.000	-12.596	0.000	-3.117	0.000	-3.969	0.000
Razão professor/pop. 7 a14	8.198	0.000	9.339	0.000	6.472	0.000	7.412	0.000
Proporção de estabelecimentos com 50 a 200 matrículas	0.497	0.000	0.640	0.000	0.310	0.000	0.418	0.000
Proporção de estabelecimentos com 201 a 1000	0.840	0.000	0.999	0.000	0.512	0.000	0.624	0.000
Proporção de estabelecimentos com mais de 1000 matrículas	0.857	0.000	0.858	0.000	0.350	0.011	0.373	0.030
População até 4999	-0.144	0.013	0.109	0.131	-0.337	0.000	-0.168	0.010
População entre 5000 e 9999	-0.105	0.070	0.042	0.558	-0.254	0.000	-0.167	0.009
População entre 10000 e 19999	-0.141	0.015	-0.019	0.794	-0.254	0.000	-0.185	0.003
População entre 20000 e 49999	-0.186	0.001	-0.073	0.306	-0.261	0.000	-0.195	0.002
População entre 50000 e 99999	-0.063	0.307	0.004	0.954	-0.095	0.072	-0.063	0.351
População entre 100000 e 199999	-0.001	0.989	0.024	0.784	-0.057	0.355	-0.046	0.558
R2	0.422		0.416		0.537		0.524	
R2-adj	0.421		0.415					
Diagnóstico para dependência espacial	Valor	Sig	Valor	Sig	Valor	Sig	Valor	Sig
Multiplicador de langrange (erro espacial)	1449.438	0.000	1163.689	0.000	28.041	0.000	24.910	0.000
Multiplicador de langrange Robusto (erro espacial)	29.229	0.000	8.467	0.004				
Multiplicador de langrange (defasagem espacial)	1663.476	0.000	1399.098	0.000				
Multiplicador de langrange Robusto (defasagem espacial)	243.267	0.000	243.877	0.000				

TABELA 6.4: Estimativas de regressão para os resíduos de nível 2 para a probabilidade de freqüentar a escola – indivíduos de 15 a 17 anos (resíduos do modelo hierárquico que utiliza apenas covariáveis de nível 1)

Variáveis	Modelo MQO				Modelo de defasagem espacial - 2SLS			
	Resíduo Bayesiano		Resíduo MQO		Resíduo Bayesiano		Resíduo MQO	
	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig
Constante	-0.896	0.000	-1.298	0.000	-0.294	0.000	-0.542	0.000
Termo de defasagem espacial (Wy)					**0.896**	**0.000**	**0.876**	**0.000**
Tamanho relativo da coorte de 15 a 17 anos	15.275	0.000	20.328	0.000	4.845	0.000	7.632	0.000
Razão professor/pop. 15 a 17	2.859	0.000	4.892	0.000	2.173	0.000	3.692	0.000
População até 4999	-0.214	0.000	-0.165	0.003	-0.086	0.007	-0.050	0.280
População entre 5000 e 9999	-0.275	0.000	-0.275	0.000	-0.116	0.000	-0.115	0.012
População entre 10000 e 19999	-0.241	0.000	-0.233	0.000	-0.118	0.000	-0.109	0.017
População entre 20000 e 49999	-0.235	0.000	-0.233	0.000	-0.131	0.000	-0.128	0.006
População entre 50000 e 99999	-0.166	0.000	-0.165	0.009	-0.058	0.098	-0.049	0.337
População entre 100000 e 199999	-0.097	0.071	-0.093	0.213	-0.039	0.352	-0.029	0.633
Dummy para município sem escola de EM	0.134	0.000	0.194	0.000	0.057	0.011	0.088	0.007
R2	0.101		0.095		0.502		0.433	
R2-adj	0.100		0.094					
Diagnóstico para dependência espacial								
Multiplicador de langrange (erro espacial)	2963.600	0.000	2199.480	0.000	142.818	0.000	102.151	0.000
Multiplicador de langrange Robusto (erro espacial)	3.445	0.063	9.617	0.002				
Multiplicador de langrange (defasagem espacial)	3263.700	0.000	2446.120	0.000				
Multiplicador de langrange Robusto (defasagem espacial)	303.544	0.000	256.258	0.000				

TABELA 6.5: Estimativas de regressão para os resíduos de nível 2 para a probabilidade de freqüentar a escola – indivíduos de 15 a 17 anos que já concluíram o Ensino Fundamental (resíduos do modelo hierárquico que utiliza apenas covariáveis de nível 1)

Variáveis	Modelo MQO				Modelo de defasagem espacial - 2SLS			
	Resíduo Bayesiano		Resíduo MQO		Resíduo Bayesiano		Resíduo MQO	
	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig
Constante	-0.009	0.890	-0.441	0.006	0.019	0.760	-0.294	0.056
Termo de defasagem espacial (Wy)					**0.420**	**0.000**	**0.637**	**0.000**
Tamanho relativo da coorte de 15 a 17 anos	6.938	0.000	13.138	0.000	5.504	0.000	8.568	0.000
Razão professor/pop. 15 a 17	3.028	0.000	4.682	0.000	2.746	0.000	4.097	0.000
População até 4999	-0.618	0.000	-0.495	0.000	-0.543	0.000	-0.400	0.000
População entre 5000 e 9999	-0.647	0.000	-0.574	0.000	-0.556	0.000	-0.428	0.000
População entre 10000 e 19999	-0.600	0.000	-0.529	0.000	-0.525	0.000	-0.420	0.000
População entre 20000 e 49999	-0.515	0.000	-0.410	0.000	-0.454	0.000	-0.330	0.002
População entre 50000 e 99999	-0.320	0.000	-0.254	0.036	-0.262	0.000	-0.170	0.144
População entre 100000 e 199999	-0.188	0.002	-0.159	0.267	-0.164	0.004	-0.114	0.406
Dummy para município sem escola de EM	0.031	0.342	-0.277	0.000	-0.002	0.957	-0.340	0.000
R2	0.088		0.032		0.158		0.1491	
R2-adj	0.087		0.031					
Diagnóstico para dependência espacial								
Multiplicador de langrange (erro espacial)	1174.945	0.000	387.273	0.000	0.055	0.815	13.576	0.000
Multiplicador de langrange Robusto (erro espacial)	38.190	0.000	0.014	0.907				
Multiplicador de langrange (defasagem espacial)	1150.543	0.000	396.051	0.000				
Multiplicador de langrange Robusto (defasagem espacial)	13.787	0.000	8.792	0.003				

No caso da freqüência na idade correta (TABs 6.6 a 6.9) as análises foram feitas tanto para o resíduo do intercepto quanto para o da educação materna, já que nos modelos hierárquicos dessa dimensão educacional o coeficiente da educação materna era aleatório. As variáveis contextuais consideradas nos modelos de regressão foram as mesmas do Modelo 4 da análise hierárquica. Da mesma forma que no caso anterior, foi

detectado autocorrelação espacial em todos os modelos, sendo o termo de defasagem espacial positivo e significativo.

É importante ressaltar, que a inclusão do termo de defasagem espacial diminui o efeito das variáveis relacionadas à rede escolar dos municípios, com maior intensidade para o ensino médio, onde apenas o fator de infra-estrutura escolar e a razão professor/população em idade escolar continuam significativos.

Considerando o efeito espacial no coeficiente da educação materna, o termo de defasagem espacial foi significativo e alto para ambos os níveis de ensino. Como nos demais casos, as variáveis contextuais perdem importância quando o termo de defasagem espacial é incluído na regressão. No modelo para o ensino fundamental, apenas a porcentagem de professores com curso superior e proporção de matrículas públicas continuaram significativas entre as variáveis relacionadas aos fatores educacionais dos municípios. No modelo para o ensino médio, continuaram significativo o fator de infra-estrutura e a razão professor/população em idade escolar.

TABELA 6.6: Estimativas de regressão para os resíduos de nível 2 para o intercepto da probabilidade de freqüentar a escola na idade correta – Ensino Fundamental (resíduos do modelo hierárquico que utiliza apenas covariáveis de nível 1).

Variáveis	Modelo MQO				Modelo de defasagem espacial - 2SLS			
	Resíduo Bayesiano		Resíduo MQO		Resíduo Bayesiano		Resíduo MQO	
	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig
Constante	-3.784	0.000	-4.161	0.000	-1.295	0.000	-1.513	0.000
Termo de defasagem espacial (Wy)					**0.528**	**0.000**	**0.531**	**0.000**
Razão professor/pop. 7 a 14	-0.617	0.078	-0.649	0.087	-0.211	0.468	-0.209	0.511
Proporção de matrículas em escolas com ciclo	0.160	0.000	0.168	0.000	0.078	0.000	0.081	0.000
Proporção de matrículas públicas	1.214	0.000	1.428	0.000	0.201	0.134	0.342	0.020
Proporção de matrículas municipais dentre as públicas	-0.165	0.000	-0.169	0.000	-0.125	0.000	-0.129	0.000
Proporção de estabelecimentos com 50 a 200 matrículas	-0.133	0.000	-0.135	0.000	-0.078	0.001	-0.080	0.002
Proporção de estabelecimentos com 201 a 1000	-0.071	0.009	-0.065	0.027	-0.036	0.109	-0.029	0.248
Proporção de estabelecimentos com mais de 1000 matrículas	0.188	0.079	0.149	0.200	0.042	0.636	0.011	0.914
Educação média da população adulta	0.262	0.000	0.274	0.000	0.152	0.000	0.160	0.000
População até 4999	0.376	0.000	0.390	0.000	0.081	0.041	0.083	0.056
População entre 5000 e 9999	0.371	0.000	0.387	0.000	0.099	0.011	0.103	0.015
População entre 10000 e 19999	0.326	0.000	0.349	0.000	0.083	0.025	0.093	0.022
População entre 20000 e 49999	0.256	0.000	0.279	0.000	0.067	0.058	0.080	0.041
População entre 50000 e 99999	0.209	0.000	0.226	0.000	0.089	0.013	0.098	0.012
População entre 100000 e 199999	0.134	0.006	0.137	0.009	0.038	0.344	0.038	0.389
Horas-aula diárias	0.296	0.000	0.319	0.000	0.112	0.000	0.123	0.000
Alunos por turma	-0.006	0.000	-0.007	0.000	-0.004	0.000	-0.004	0.000
Proporção professores com curso superior	0.554	0.000	0.583	0.000	0.225	0.000	0.234	0.000
Fator de infra-estrutura	0.123	0.000	0.130	0.000	0.044	0.000	0.047	0.000
R2	0.772		0.761		0.820		0.810	
R2-adj	0.771		0.760					
Diagnóstico para dependência espacial	Valor	Sig	Valor	Sig	Valor	Sig	Valor	Sig
Multiplicador de langrange (erro espacial)	1491.215	0.000	1374.024	0.000	0.145	0.703	1.052	0.305
Multiplicador de langrange Robusto (erro espacial)	130.628	0.000	101.749	0.000				
Multiplicador de langrange (defasagem espacial)	1871.237	0.000	1755.008	0.000				
Multiplicador de langrange Robusto (defasagem espacial)	510.651	0.000	482.733	0.000				

TABELA 6.7: Estimativas de regressão para os resíduos de nível 2 para o coeficiente da educação materna da probabilidade de freqüentar a escola na idade correta – Ensino Fundamental (resíduos do modelo hierárquico que utiliza apenas covariáveis de nível 1).

Variáveis	Modelo MQO				Modelo de defasagem espacial - 2SLS			
	Resíduo Bayesiano		Resíduo MQO		Resíduo Bayesiano		Resíduo MQO	
	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig
Constante	0.152	0.000	0.146	0.000	0.081	0.000	0.112	0.001
Termo de defasagem espacial (Wy)					**0.741**	**0.000**	**0.844**	**0.000**
Razão professor/pop. 7 a 14	0.122	0.000	0.227	0.001	0.011	0.689	0.040	0.534
Proporção de matrículas em escolas com ciclo	0.000	0.902	0.008	0.013	-0.002	0.059	-0.004	0.243
Proporção de matrículas públicas	-0.116	0.000	-0.179	0.000	-0.046	0.000	-0.083	0.004
Proporção de matrículas municipais dentre as públicas	0.009	0.000	0.020	0.000	-0.003	0.098	-0.005	0.293
Proporção de estabelecimentos com 50 a 200 matrículas	0.001	0.535	-0.007	0.179	0.000	0.985	-0.008	0.126
Proporção de estabelecimentos com 201 a 1000	0.003	0.232	0.014	0.007	0.000	0.948	0.009	0.064
Proporção de estabelecimentos com mais de 1000 matrículas	0.016	0.075	0.025	0.215	0.010	0.233	0.008	0.684
Educação média da população adulta	-0.012	0.000	-0.015	0.000	-0.006	0.000	-0.007	0.000
População até 4999	-0.031	0.000	-0.039	0.000	-0.008	0.023	-0.006	0.478
População entre 5000 e 9999	-0.030	0.000	-0.033	0.000	-0.010	0.004	-0.005	0.569
População entre 10000 e 19999	-0.026	0.000	-0.028	0.001	-0.010	0.004	-0.003	0.683
População entre 20000 e 49999	-0.021	0.000	-0.021	0.010	-0.007	0.023	-0.001	0.908
População entre 50000 e 99999	-0.014	0.000	-0.013	0.115	-0.006	0.050	-0.002	0.775
População entre 100000 e 199999	-0.007	0.077	-0.006	0.489	-0.003	0.485	0.000	0.961
Horas-aula diárias	0.007	0.000	0.024	0.000	0.000	0.996	0.001	0.836
Alunos por turma	0.000	0.000	0.000	0.000	0.000	0.054	0.000	0.665
Proporção professores com curso superior	-0.043	0.000	-0.064	0.000	-0.009	0.001	-0.013	0.046
Fator de infra-estrutura	-0.006	0.000	-0.007	0.000	0.000	0.682	0.002	0.286
R2	0.489		0.226		0.621		0.398	
R2-adj	0.487		0.224					
Diagnóstico para dependência espacial	Valor	Sig	Valor	Sig	Valor	Sig	Valor	Sig
Multiplicador de langrange (erro espacial)	703.582	0.000	330.357	0.000	179.490	0.000	143.375	0.000
Multiplicador de langrange Robusto (erro espacial)	11.314	0.001	42.854	0.000				
Multiplicador de langrange (defasagem espacial)	1031.783	0.000	451.730	0.000				
Multiplicador de langrange Robusto (defasagem espacial)	339.516	0.000	164.226	0.000				

TABELA 6.8: Estimativas de regressão para os resíduos de nível 2 para o intercepto da probabilidade de freqüentar a escola na idade correta – Ensino Médio (resíduos do modelo hierárquico que utiliza apenas covariáveis de nível 1).

Variáveis	Modelo MQO				Modelo de defasagem espacial - 2SLS			
	Resíduo Bayesiano		Resíduo MQO		Resíduo Bayesiano		Resíduo MQO	
	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig
Constante	-4.361	0.000	-5.449	0.000	-1.705	0.000	-2.287	0.000
Termo de defasagem espacial (Wy)					**0.587**	**0.000**	**0.585**	**0.000**
Razão professor/pop. 15 a 17	4.672	0.000	5.746	0.000	2.613	0.000	3.495	0.000
Proporção de estabelecimentos com 50 a 200 matrículas	0.097	0.100	0.189	0.011	-0.046	0.341	0.007	0.907
Proporção de estabelecimentos com 201 a 1000	0.225	0.000	0.377	0.000	0.030	0.537	0.131	0.041
Proporção de estabelecimentos com mais de 1000 matrículas	0.075	0.470	0.210	0.106	0.085	0.309	0.206	0.061
Proporção de matrículas públicas	0.378	0.000	0.490	0.000	0.029	0.644	0.051	0.529
Proporção de matrículas municipais dentre as públicas	-0.135	0.000	-0.202	0.000	-0.074	0.012	-0.129	0.001
Educação média da população adulta	0.605	0.000	0.721	0.000	0.293	0.000	0.356	0.000
População até 4999	1.256	0.000	1.518	0.000	0.366	0.000	0.501	0.000
População entre 5000 e 9999	1.115	0.000	1.351	0.000	0.327	0.000	0.451	0.000
População entre 10000 e 19999	0.927	0.000	1.159	0.000	0.233	0.000	0.360	0.000
População entre 20000 e 49999	0.723	0.000	0.944	0.000	0.175	0.002	0.310	0.000
População entre 50000 e 99999	0.473	0.000	0.645	0.000	0.154	0.009	0.273	0.000
População entre 100000 e 199999	0.345	0.000	0.431	0.000	0.077	0.259	0.125	0.161
Horas-aula diárias	0.027	0.272	0.013	0.676	0.002	0.931	-0.012	0.641
Alunos por turma	-0.006	0.000	-0.006	0.004	0.000	0.997	0.002	0.178
Proporção professores com curso superior	0.242	0.000	0.272	0.000	0.039	0.186	0.020	0.597
Dummy para município sem escola de EM	0.468	0.002	0.501	0.008	-0.030	0.805	-0.059	0.713
Fator de infra-estrutura	0.276	0.000	0.308	0.000	0.109	0.000	0.119	0.000
R2	0.739		0.713		0.813		0.785	
R2-adj	0.738		0.712					
Diagnóstico para dependência espacial	Valor	Sig	Valor	Sig	Valor	Sig	Valor	Sig
Multiplicador de langrange (erro espacial)	1236.525	0.000	950.954	0.000	24.615	0.000	35.257	0.000
Multiplicador de langrange Robusto (erro espacial)	47.813	0.000	14.997	0.000				
Multiplicador de langrange (defasagem espacial)	1976.160	0.000	1608.943	0.000				
Multiplicador de langrange Robusto (defasagem espacial)	787.449	0.000	672.986	0.000				

TABELA 6.9: Estimativas de regressão para os resíduos de nível 2 para o coeficiente da educação materna da probabilidade de freqüentar a escola na idade correta – Ensino Médio (resíduos do modelo hierárquico que utiliza apenas covariáveis de nível 1).

Variáveis	Modelo MQO				Modelo de defasagem espacial - 2SLS			
	Resíduo Bayesiano		Resíduo MQO		Resíduo Bayesiano		Resíduo MQO	
	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig
Constante	0.200	0.000	0.191	0.000	0.075	0.000	0.086	0.016
Termo de defasagem espacial (Wy)					**0.629**	**0.000**	**0.652**	**0.000**
Razão professor/pop. 15 a 17	-0.213	0.000	-0.459	0.000	-0.115	0.000	-0.382	0.000
Proporção de estabelecimentos com 50 a 200 matrículas	-0.006	0.038	-0.013	0.272	0.001	0.607	-0.010	0.413
Proporção de estabelecimentos com 201 a 1000	-0.013	0.000	-0.020	0.106	-0.003	0.226	-0.016	0.206
Proporção de estabelecimentos com mais de 1000 matrículas	-0.005	0.374	-0.022	0.280	-0.006	0.159	-0.037	0.084
Proporção de matrículas públicas	-0.020	0.000	-0.058	0.000	-0.002	0.564	-0.033	0.038
Proporção de matrículas municipais dentre as públicas	0.006	0.000	0.013	0.064	0.003	0.080	0.006	0.381
Educação média da população adulta	-0.029	0.000	-0.037	0.000	-0.013	0.000	-0.015	0.000
População até 4999	-0.065	0.000	-0.060	0.000	-0.021	0.000	-0.011	0.490
População entre 5000 e 9999	-0.058	0.000	-0.064	0.000	-0.019	0.000	-0.020	0.182
População entre 10000 e 19999	-0.049	0.000	-0.056	0.000	-0.015	0.000	-0.017	0.254
População entre 20000 e 49999	-0.039	0.000	-0.047	0.001	-0.013	0.000	-0.015	0.294
População entre 50000 e 99999	-0.029	0.000	-0.038	0.008	-0.014	0.000	-0.019	0.210
População entre 100000 e 199999	-0.023	0.000	-0.028	0.100	-0.010	0.003	-0.015	0.391
Horas-aula diárias	0.002	0.067	0.026	0.000	0.001	0.166	0.013	0.008
Alunos por turma	0.000	0.000	0.001	0.021	0.000	0.941	0.000	0.705
Proporção professores com curso superior	-0.012	0.000	-0.012	0.095	-0.001	0.376	0.006	0.384
Dummy para município sem escola de EM	-0.011	0.126	0.027	0.371	0.005	0.386	-0.008	0.797
Fator de infra-estrutura	-0.013	0.000	-0.016	0.000	-0.005	0.000	-0.006	0.030
R2	0.725		0.203		0.809		0.2752	
R2-adj	0.724		0.200					
Diagnóstico para dependência espacial	Valor	Sig	Valor	Sig	Valor	Sig	Valor	Sig
Multiplicador de langrange (erro espacial)	1009.109	0.000	3.201	0.074	77.581	0.000	93.907	0.000
Multiplicador de langrange Robusto (erro espacial)	6.904	0.009	65.782	0.000				
Multiplicador de langrange (defasagem espacial)	1874.159	0.000	20.912	0.000				
Multiplicador de langrange Robusto (defasagem espacial)	871.955	0.000	83.493	0.000				

Com relação às PPSs (TABs. 6.10 a 6.13), os resultados dos testes para autocorrelação espacial e as regressões espaciais não diferem muito das análises anteriores, não cabendo portanto tecer maiores comentários. Ressalta-se apenas que o coeficiente do termo de defasagem espacial para o resíduo do intercepto foi maior para a probabilidade de progressão na 1ª série (e_0).

TABELA 6.10: Estimativas de regressão para os resíduos de nível 2 para o intercepto da probabilidade de progressão na 1ª série do ensino fundamental (resíduos do modelo hierárquico que utiliza apenas covariáveis de nível 1).

Variáveis	Modelo MQO				Modelo de defasagem espacial - 2SLS			
	Resíduo Bayesiano		Resíduo MQO		Resíduo Bayesiano		Resíduo MQO	
	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig
Constante	-2.011	0.000	-3.257	0.000	-0.594	0.058	-1.310	0.008
Termo de defasagem espacial (Wy)					**0.475**	**0.000**	**0.568**	**0.000**
Razão professor/pop. 7 a 14	0.780	0.180	0.568	0.528	0.817	0.129	0.744	0.386
Proporção de matrículas em escolas com ciclo	0.050	0.076	-0.086	0.046	0.013	0.610	-0.065	0.118
Proporção de matrículas públicas	0.449	0.090	1.273	0.002	-0.189	0.447	0.285	0.473
Proporção de matrículas municipais dentre as públicas	-0.197	0.000	-0.324	0.000	-0.135	0.000	-0.185	0.002
Proporção de estabelecimentos com 50 a 200 matrículas	-0.024	0.615	0.060	0.411	0.005	0.909	0.080	0.255
Proporção de estabelecimentos com 201 a 1000	-0.049	0.286	-0.043	0.547	-0.016	0.702	0.016	0.813
Proporção de estabelecimentos com mais de 1000 matrículas	0.193	0.289	-0.095	0.736	0.053	0.753	-0.121	0.651
Educação média da população adulta	0.302	0.000	0.377	0.000	0.190	0.000	0.213	0.000
População até 4999	0.188	0.017	0.654	0.000	-0.078	0.299	0.164	0.176
População entre 5000 e 9999	0.253	0.001	0.655	0.000	0.014	0.846	0.221	0.059
População entre 10000 e 19999	0.209	0.005	0.565	0.000	0.001	0.991	0.179	0.112
População entre 20000 e 49999	0.076	0.287	0.376	0.001	-0.077	0.251	0.073	0.498
População entre 50000 e 99999	0.045	0.540	0.234	0.038	-0.035	0.608	0.061	0.570
População entre 100000 e 199999	0.074	0.373	0.151	0.236	0.005	0.948	0.041	0.733
Horas-aula diárias	0.094	0.011	0.122	0.033	0.046	0.177	0.078	0.154
Alunos por turma	-0.012	0.000	-0.016	0.000	-0.008	0.000	-0.009	0.000
Proporção professores com curso superior	0.354	0.000	0.554	0.000	0.115	0.025	0.102	0.222
Fator de infra-estrutura	0.080	0.000	0.161	0.000	0.024	0.081	0.052	0.020
R2	0.507		0.434		0.559		0.500	
R2-adj	0.505		0.432					
Diagnóstico para dependência espacial	Valor	Sig	Valor	Sig	Valor	Sig	Valor	Sig
Multiplicador de langrange (erro espacial)	651.050	0.000	268.643	0.000	9.636	0.002	56.310	0.000
Multiplicador de langrange Robusto (erro espacial)	6.240	0.012	14.259	0.000				
Multiplicador de langrange (defasagem espacial)	800.623	0.000	410.312	0.000				
Multiplicador de langrange Robusto (defasagem espacial)	155.812	0.000	155.927	0.000				

TABELA 6.11: Estimativas de regressão para os resíduos de nível 2 para o coeficiente da educação materna da probabilidade de progressão na 1ª série do ensino fundamental (resíduos do modelo hierárquico que utiliza apenas covariáveis de nível 1).

Variáveis	Modelo MQO				Modelo de defasagem espacial - 2SLS			
	Resíduo Bayesiano		Resíduo MQO		Resíduo Bayesiano		Resíduo MQO	
	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig
Constante	0.142	0.000	0.539	0.000	0.050	0.018	0.315	0.045
Termo de defasagem espacial (Wy)					**0.550**	**0.000**	**0.522**	**0.000**
Razão professor/pop. 7 a 14	-0.057	0.124	-0.741	0.004	-0.032	0.371	-0.739	0.004
Proporção de matrículas em escolas com ciclo	-0.012	0.000	-0.053	0.000	-0.006	0.002	-0.033	0.014
Proporção de matrículas públicas	-0.059	0.000	-0.218	0.060	-0.015	0.374	-0.095	0.441
Proporção de matrículas municipais dentre as públicas	0.010	0.000	-0.004	0.807	0.005	0.034	-0.013	0.488
Proporção de estabelecimentos com 50 a 200 matrículas	0.003	0.359	-0.020	0.338	-0.002	0.472	-0.037	0.086
Proporção de estabelecimentos com 201 a 1000	0.002	0.457	-0.040	0.046	0.001	0.802	-0.041	0.046
Proporção de estabelecimentos com mais de 1000 matrículas	-0.036	0.002	-0.060	0.451	-0.018	0.106	-0.025	0.762
Educação média da população adulta	-0.015	0.000	-0.038	0.000	-0.009	0.000	-0.027	0.000
População até 4999	-0.018	0.000	-0.054	0.120	-0.004	0.377	-0.032	0.375
População entre 5000 e 9999	-0.020	0.000	-0.047	0.164	-0.007	0.126	-0.024	0.483
População entre 10000 e 19999	-0.017	0.000	-0.046	0.156	-0.007	0.156	-0.027	0.413
População entre 20000 e 49999	-0.011	0.014	-0.041	0.195	-0.004	0.399	-0.026	0.411
População entre 50000 e 99999	-0.011	0.020	-0.036	0.254	-0.006	0.152	-0.026	0.426
População entre 100000 e 199999	-0.015	0.004	-0.025	0.480	-0.010	0.063	-0.013	0.724
Horas-aula diárias	-0.002	0.381	0.005	0.776	0.001	0.573	0.012	0.481
Alunos por turma	0.000	0.000	-0.001	0.235	0.000	0.009	-0.001	0.148
Proporção professores com curso superior	-0.016	0.000	-0.034	0.142	-0.005	0.161	-0.021	0.385
Fator de infra-estrutura	-0.004	0.000	0.006	0.304	-0.001	0.499	0.012	0.072
R2	0.374		0.049		0.441			
R2-adj	0.372		0.046					
Diagnóstico para dependência espacial	Valor	Sig	Valor	Sig	Valor	Sig	Valor	Sig
Multiplicador de langrange (erro espacial)	158.344	0.000	0.2984	0.5849	78.492	0.000	13.882	0.000
Multiplicador de langrange Robusto (erro espacial)	24.908	0.000	0.7727	0.3794				
Multiplicador de langrange (defasagem espacial)	257.279	0.000	0.4528	0.5010				
Multiplicador de langrange Robusto (defasagem espacial)	123.844	0.000	0.9271	0.3356				

TABELA 6. 12: Estimativas de regressão para os resíduos de nível 2 para o intercepto da probabilidade de progressão na 4ª série do ensino fundamental (resíduos do modelo hierárquico que utiliza apenas covariáveis de nível 1).

Variáveis	Modelo MQO				Modelo de defasagem espacial - 2SLS			
	Resíduo Bayesiano		Resíduo MQO		Resíduo Bayesiano		Resíduo MQO	
	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig
Constante	-2.305	0.000	-4.318	0.000	-0.763	0.007	-2.122	0.000
Termo de defasagem espacial (Wy)					**0.382**	**0.000**	**0.411**	**0.000**
Razão professor/pop. 7 a 14	1.018	0.036	0.800	0.296	1.211	0.010	1.142	0.126
Proporção de matrículas em escolas com ciclo	0.086	0.000	0.103	0.005	0.063	0.005	0.082	0.022
Proporção de matrículas públicas	0.421	0.057	1.690	0.000	-0.288	0.190	0.627	0.074
Proporção de matrículas municipais dentre as públicas	-0.231	0.000	-0.330	0.000	-0.190	0.000	-0.255	0.000
Proporção de estabelecimentos com 50 a 200 matrículas	0.072	0.071	0.217	0.001	0.087	0.023	0.230	0.000
Proporção de estabelecimentos com 201 a 1000	0.032	0.411	0.129	0.033	0.028	0.450	0.122	0.039
Proporção de estabelecimentos com mais de 1000 matrículas	0.076	0.617	-0.254	0.287	-0.062	0.671	-0.319	0.171
Educação média da população adulta	0.269	0.000	0.379	0.000	0.189	0.000	0.263	0.000
População até 4999	0.239	0.000	0.651	0.000	0.030	0.650	0.301	0.004
População entre 5000 e 9999	0.233	0.000	0.555	0.000	0.046	0.471	0.244	0.017
População entre 10000 e 19999	0.171	0.006	0.488	0.000	0.015	0.803	0.219	0.025
População entre 20000 e 49999	0.083	0.165	0.381	0.000	-0.037	0.523	0.164	0.079
População entre 50000 e 99999	0.102	0.094	0.308	0.001	0.035	0.551	0.180	0.054
População entre 100000 e 199999	0.037	0.593	0.096	0.376	-0.022	0.737	0.003	0.980
Horas-aula diárias	0.099	0.001	0.117	0.016	0.025	0.411	0.029	0.548
Alunos por turma	-0.002	0.020	-0.006	0.000	-0.001	0.176	-0.004	0.003
Proporção professores com curso superior	0.531	0.000	0.811	0.000	0.280	0.000	0.396	0.000
Fator de infra-estrutura	0.131	0.000	0.197	0.000	0.072	0.000	0.103	0.000
R2	0.613		0.568		0.638		0.594	
R2-adj	0.612		0.567					
Diagnóstico para dependência espacial	Valor	Sig	Valor	Sig	Valor	Sig	Valor	Sig
Multiplicador de langrange (erro espacial)	208.296	0.000	104.583	0.000	14.604	0.000	30.382	0.000
Multiplicador de langrange Robusto (erro espacial)	0.320	0.571	12.440	0.000				
Multiplicador de langrange (defasagem espacial)	329.047	0.000	205.100	0.000				
Multiplicador de langrange Robusto (defasagem espacial)	121.071	0.000	112.957	0.000				

TABELA 6. 13: Estimativas de regressão para os resíduos de nível 2 para o intercepto da probabilidade de progressão na 1ª série do ensino médio (resíduos do modelo hierárquico que utiliza apenas covariáveis de nível 1).

Variáveis	Modelo MQO				Modelo de defasagem espacial - 2SLS			
	Resíduo Bayesiano		Resíduo MQO		Resíduo Bayesiano		Resíduo MQO	
	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig
Constante	-0.378	0.000	-1.262	0.000	-0.156	0.044	-0.459	0.114
Termo de defasagem espacial (Wy)					**0.334**	**0.000**	**0.424**	**0.000**
Educação média da população adulta	0.075	0.000	0.203	0.000	0.053	0.000	0.129	0.000
População até 4999	-0.169	0.000	0.207	0.074	-0.219	0.000	-0.027	0.825
População entre 5000 e 9999	-0.139	0.000	0.244	0.030	-0.176	0.000	0.054	0.637
População entre 10000 e 19999	-0.149	0.000	0.190	0.085	-0.179	0.000	0.025	0.827
População entre 20000 e 49999	-0.177	0.000	0.146	0.175	-0.195	0.000	0.018	0.869
População entre 50000 e 99999	-0.170	0.000	0.071	0.534	-0.172	0.000	0.005	0.965
População entre 100000 e 199999	-0.079	0.031	0.052	0.695	-0.091	0.012	0.003	0.983
Razão professor/pop. 15 a 17	1.486	0.000	5.253	0.000	1.104	0.000	3.914	0.000
Proporção de estabelecimentos com 50 a 200 matrículas	0.004	0.873	-0.063	0.498	0.004	0.867	-0.103	0.272
Proporção de estabelecimentos com 201 a 1000	-0.014	0.596	-0.004	0.967	-0.012	0.630	-0.044	0.642
Proporção de estabelecimentos com mais de 1000 matrículas	-0.057	0.206	-0.026	0.873	-0.047	0.289	-0.018	0.911
Proporção de matrículas públicas	-0.053	0.111	-0.095	0.429	-0.070	0.032	-0.139	0.243
Proporção de matrículas municipais dentre as públicas	0.025	0.108	-0.066	0.248	0.022	0.156	-0.076	0.180
Fator de infra-estrutura	0.034	0.000	0.095	0.000	0.023	0.000	0.057	0.003
Horas-aula diárias	0.045	0.000	0.025	0.514	0.036	0.001	0.000	1.000
Alunos por turma	0.000	0.641	0.002		0.000	0.784	-0.001	0.709
Proporção professores com curso superior	0.003	0.844	-0.043	0.444	-0.013	0.413	-0.101	0.074
Dummy para município sem escola de EM	0.202	0.002	-0.284	0.234	0.111	0.091	-0.516	0.032
R2	0.238		0.155		0.266		0.1903	
R2-adj	0.235		0.153					
Diagnóstico para dependência espacial	Valor	Sig	Valor	Sig	Valor	Sig	Valor	Sig
Multiplicador de langrange (erro espacial)	144.044	0.000	36.408	0.000	8.272	0.004	21.018	0.000
Multiplicador de langrange Robusto (erro espacial)	0.721	0.396	6.623	0.010				
Multiplicador de langrange (defasagem espacial)	159.717	0.000	49.294	0.000				
Multiplicador de langrange Robusto (defasagem espacial)	16.394	0.000	19.509	0.000				

6.4) Conclusão

Esse capítulo teve como objetivo discutir a existência da autocorrelação e heterogeneidade espacial dos dados utilizados nessa tese. O diagnóstico para detectar os

padrões de interação espacial foi feito utilizando as técnicas de Análise Exploratória Espacial dos Dados (ESDA) e econometria espacial.

As ferramentas da ESDA utilizadas foram o I de Moran, o *Moran Scatterplot* e o I de Moran Local. Através da estatística I de Moran foi possível perceber que todas as variáveis tratadas no modelo hierárquico do capítulo anterior possuem autocorrelação espacial global, sendo mais forte para o ensino fundamental que para o médio.

A confecção dos mapas do *Moran Scatterplot* foi importante por revelar evidências de que a maioria das variáveis é aglomerada espacialmente, ou seja, municípios próximos possuem níveis similares, evidenciando uma associação espacial positiva. O Brasil apresenta um padrão espacial bem definido para quase todas as variáveis, com as regiões Norte e Nordeste e os estados ao norte do Centro-Oeste apresentando piores indicadores, enquanto que o Sul, Sudeste e o sul do Centro-Oeste os melhores.

Através do *Moran Scatterplot* verificou-se também que os valores das variáveis para os municípios dentro de um mesmo estado são muito próximos e correlacionados, ou seja, as manchas com valores alto-alto ou baixo-baixo cobrem a totalidade do estado. Tal fato sugere uma hegemonia do estado seja através das suas políticas ou outros fatores contextuais. Entretanto, quando considera apenas os valores significativos do I de Moran local verifica-se o aparecimento de *clusters* de municípios, ou seja, as manchas com valores semelhantes cobrem apenas alguns municípios dentro do estado. Dessa forma, pode-se dizer que nesses municípios há algum fator municipal que o diferencie do restante do estado. Porém, há algumas exceções como os estados de São Paulo, Maranhão, Alagoas e Piauí que na maioria das variáveis não apresentam aglomerados de municípios.

Somado a isso, os padrões dos mapas sugerem que a taxa de atendimento do ensino fundamental possui uma autocorrelação espacial positiva com a relação professor/população em idade escolar e negativa com o tamanho relativo da coorte em idade escolar. Para o ensino médio, essa mesma associação foi verificada para a taxa de escolarização líquida, já que a taxa de atendimento desse nível de ensino apresenta caráter ambíguo.

Através da investigação visual, detectou-se também que no ensino fundamental a taxa de eficiência, e_0 e_4 possuem uma possível autocorrelação espacial positiva com a

porcentagem de professores com curso superior, razão professor/população em idade escolar e o fator de infra-estrutura e negativa com o tamanho relativo da coorte e média de alunos por turma.

Para o ensino médio observou-se uma autocorrelação espacial positiva apenas entre a taxa de eficiência e a porcentagem de professores com curso superior, razão professor/população em idade escolar e o fator de infra-estrutura da rede escolar e negativa para o tamanho relativo da coorte.

Dessa forma, pode-se dizer que os indicadores de acesso (taxa de atendimento do nível fundamental e a taxa de escolarização líquida do médio) são influenciados positivamente pela disponibilidade de oferta dos municípios vizinhos e negativamente pelo tamanho relativo da coorte dos municípios vizinhos.

Por outro lado, os indicadores de eficiência e progressão (apenas no fundamental) podem estar sendo influenciados positivamente pela restrição de oferta e pela qualidade da rede escolar dos municípios vizinhos e negativamente pelo tamanho da coorte em idade escolar dos seus vizinhos.

Para averiguar se a autocorrelação espacial dos indicadores educacionais continua, mesmo quando são regredidos com as variáveis contextuais, estimou-se um modelo de defasagem espacial. Seus resultados mostraram que o termo de defasagem espacial da variável dependente (WY) é alto e significativo para todos os casos.

Pode-se concluir, portanto, que os indicadores educacionais analisados nessa tese, possuem uma autocorrelação espacial com as variáveis observadas e não-observadas dos municípios vizinhos. Quando não se corrige essa dependência espacial, os coeficientes estimados do modelo hierárquico podem ser inconsistentes e ineficientes. Dessa forma, fica comprovada a importância da construção de um modelo hierárquico-espacial que amenize esses problemas. A construção e estimação desse modelo será o tema do próximo capítulo.

7) ANÁLISE HIERÁRQUICO-ESPACIAL DOS DETERMINANTES EDUCACIONAIS

Nesse capítulo são comparados os resultados do modelo hierárquico estimado no Capítulo 5 com as estimativas hierárquico-espaciais, que buscam controlar para os efeitos espaciais nos dados[22]. Como constatado no capítulo anterior, o estudo dos resíduos do nível 2 do modelo hierárquico apresenta autocorrelação espacial, principalmente na forma de defasagens espaciais das variáveis observadas e não-observadas. A presença de um processo espacial nos resíduos faz com que as estimativas dos coeficientes não sejam confiáveis, podendo ser enviesadas e/ou ineficientes.

Assim, nesse capítulo, tentou-se incorporar o efeito espacial no modelo hierárquico de duas formas diferentes. Na primeira, incluiu-se as defasagens espaciais dos resíduos, tanto *bayesianos* quanto MQO, do nível 2 do modelo hierárquico analisado no Capítulo 5, enquanto que na segunda forma incluiu-se as defasagens espaciais das variáveis independentes utilizadas nos modelos. Vale lembrar que a defasagem espacial de uma variável é a média, dessa variável, nos seus vizinhos.

Quando se acrescenta a defasagem espacial dos resíduos controla-se para a autocorrelação espacial existente tanto nas variáveis observadas quanto nas não-observadas. Já quando se considera apenas as defasagens espaciais das variáveis independentes, não se está controlando para a dependência espacial das variáveis não-observadas, porém, é útil na medida em que evidencia qual variável independente é mais correlacionada espacialmente com os indicadores educacionais, ou seja, as variáveis dependentes.

7.1) Probabilidade de freqüentar a escola

Os resultados do modelo hierárquico-espacial para a freqüência escolar dos indivíduos de 7 a 14 anos e de 15 a 17 anos encontram-se nas TABs. 7.1 a 7.3. Para facilitar as comparações, os resultados do modelo hierárquico do Capítulo 5 também se

[22] Como já mencionado no Capítulo 4, pode-se estimar um modelo hierárquico-espacial com os programas MLwiN e WinBUGS, que possibilita incluir a matriz de pesos espaciais. Tal procedimento torna as estimativas mais consistentes e eficientes (para exemplo dessa aplicação ver LANGFORD & LEYLAND, 1999; BORGONI & BILLARI, 2003). Cabe ressaltar, que na análise proposta nessa tese, os parâmetros podem ser menos eficientes. Porém, tal fato pode ser minimizado devido ao grande número de casos no nível 2 do modelo hierárquico.

encontram nas tabelas. Visto que nesse capítulo, 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, percebe-se pela TAB. 7.1 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*, uma diminuição de um desvio padrão em relação à média, do tamanho relativo da coorte, provoca uma variação percentual na probabilidade das crianças de 7 a 14 anos freqüentarem a escola de 0,31%, passando de 0,961 para 0,964. Quando não se incluiu o termo espacial na análise, a variação foi de 1,13%. Com relação a variável de restrição de oferta, a diminuição do seu efeito é bem menor, de forma que ela passa a ser mais importante 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 um aumento percentual de 0,46% na probabilidade de freqüentar a escola, alcançando 0,966. Vale lembrar, que essa pequena variação na probabilidade se deve ao fato da freqüência das crianças de 7 a 14 anos já estar praticamente no seu nível máximo, ou seja, 100%.

No que diz respeito aos indivíduos de 15 a 17 anos, percebe-se pela TAB. 7.2 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, que é o sinal esperado. Por outro lado, no modelo que inclui apenas as defasagens das variáveis independentes, o sinal continua positivo, apesar do 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, e que esses são autocorrelacionados espacialmente, já que a defasagem espacial do resíduo capta os efeitos espaciais das variáveis observadas e não-observadas.

Entre as variáveis independentes defasadas, observa-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 a probabilidade de freqüência escolar dessa faixa etária no município local em 2,32%.

Com relação à freqüência escolar dos indivíduos de 15 a 17 anos que já concluíram o ensino fundamental, ou seja, um grupo mais seletivo, os resultados foram diferentes dos anteriores. Observa-se pela TAB. 7.3, que ao contrário do que ocorreu na freqüência não condicional dessa faixa etária, o sinal do tamanho relativo da coorte foi positivo para todos os casos. Outra diferença encontrada, é que a restrição de oferta não se transforma no fator mais importante, como nas análises anteriores. Somado a isso, no modelo com a defasagem das variáveis independentes, a restrição de oferta dos municípios vizinhos não foi significativa, enquanto que a defasagem espacial do tamanho relativo da coorte mostrou um efeito expressivo.

Com base nos resultados para a probabilidade de freqüência escolar percebe-se que quando se tenta controlar a autocorrelação espacial nos modelo hierárquicos, diminuindo os problemas de viés e eficiência das estimativas, o efeito das variáveis de nível 2 diminui. E, por outro lado, observa-se 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 freqüência escolar. O efeito direto seria o impacto na probabilidade de cursar a escola decorrente de mudanças nas variáveis contextuais no próprio município. Já o efeito indireto seria o impacto decorrente de mudanças das variáveis contextuais nos municípios vizinhos. O primeiro efeito é captado pelos coeficientes das covariáveis do nível dois nas duas especificações do modelo hierárquico-espacial (com defasagem dos resíduos e com defasagem das covariáveis) e o segundo é captado 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 como visto no capítulo anterior essas variáveis são altamente correlacionadas espacialmente, de tal forma que podem variar conjuntamente.

Os GRAFs. 7.1 a 7.6 ajudam a compreender melhor os efeitos diretos e indiretos. O GRAF. 7.1 mostra a mudança em pontos percentuais na probabilidade de

cursar a escola para os indivíduos de 7 a 14 anos decorrente de uma 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 uma diminuição na pressão demográfica no próprio município, sendo que na primeira as simulações foram feitas com base nos coeficientes do modelo hierárquico do capítulo 5[23], que não separa os efeitos diretos e indiretos. Na segunda e terceira colunas as simulações foram calculadas utilizando-se os 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 variáveis 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 encontrado nos dois modelos hierárquico-espaciais. A quarta coluna mede o aumento percentual quando ocorre uma diminuição no valor dessa variável nos municípios vizinhos, que foi considerado como o efeito indireto do dividendo demográfico. Por fim, a última coluna é a soma da terceira e quarta colunas, ou seja, é a soma do efeito direto e indireto do tamanho relativo da coorte em idade escolar.

No GRAF. 7.2 esse mesmo exercício foi realizado para a razão professor/população 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 GRAFs. 7.3 a 7.6 são semelhantes aos dois primeiros, porém medem os impactos na probabilidade dos indivíduos de 15 a 17 anos cursarem a escola independentemente da sua escolaridade (GRAF. 7.3 e 7.4) e para os que já completaram o ensino fundamental (GRAF. 7.5 e 7.6), respectivamente. A variação negativa nas probabilidades, ou seja, a diminuição em pontos percentuais nas probabilidades observada nos GRAFs 7.3 e 7.5 ocorre pelo fato do sinal do tamanho relativo da coorte não ser o esperado.

Através dos gráficos percebe-se que quando se tenta controlar a autocorrelação espacial dos dados, realmente ocorre a diminuição do impacto do tamanho relativo da

[23] Para a freqüência escolar dos indivíduos de 7 a 14 anos, utilizou o Modelo 5 do Capítulo 5; para os indivíduos de 15 a 17 anos o Modelo 6.

coorte e da razão professor/população em idade escolar. Porém, quando se soma o efeito indireto dessas variáveis, ou seja, o efeito via seus vizinhos, observa-se que o aumento ou diminuição são próximos dos encontrado no modelo hierárquico não-espacial.

Esse resultado pode suscitar uma importante questão. Apesar da 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 o efeito direto e indireto das variáveis contextuais, enquanto que o modelo hierárquico-espacial separa esses dois efeitos.

TABELA 7. 1: Resultado dos diferentes modelos para a probabilidade de freqüentar a escola – indivíduos de 7 a 14 anos.

Variáveis Independentes	Modeo Hierarquico (Modelo 5)		Modelo com a defasagem do resíduo bayesiano		Modelo com a defasagem do resíduo MQO		Modelo com a defasagem das variáveis independentes	
Efeito Fixo	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig
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 a14	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 1000	1.111	0.000	0.629	0.000	0.619	0.000	0.846	0.000
Proporção de estabelecimentos com mais de 1000 matrículas	0.924	0.000	0.210	0.175	0.317	0.040	0.718	0.000
População até 4999	-0.010	0.868	-0.286	0.000	-0.329	0.000	-0.283	0.000
População entre 5000 e 9999	-0.028	0.648	-0.236	0.000	-0.268	0.000	-0.258	0.000
População entre 10000 e 19999	-0.051	0.399	-0.217	0.000	-0.249	0.000	-0.236	0.000
População entre 20000 e 49999	-0.103	0.087	-0.218	0.000	-0.250	0.000	-0.252	0.000
População entre 50000 e 99999	-0.003	0.959	-0.055	0.302	-0.083	0.116	-0.093	0.148
População entre 100000 e 199999	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 a14							5.040	0.000
Defasagem espacial Proporção de estabelecimentos com 50 a 200 matrículas							0.696	0.000
Defasagem espacial Proporção de estabelecimentos com 201 a 1000							0.477	0.000
Defasagem espacial Proporção de estabelecimentos com mais de 1000 matrículas							-0.129	0.733
Defasagem espacial População até 4999							0.519	0.000
Defasagem espacial População entre 5000 e 9999							0.124	0.269
Defasagem espacial População entre 10000 e 19999							0.106	0.337
Defasagem espacial População entre 20000 e 49999							0.068	0.537
Defasagem espacial População entre 50000 e 99999							0.031	0.796
Defasagem espacial População entre 100000 e 199999							0.241	0.089
Efeito Aleatório								
Coeficiente	0.287	0.000	0.192	0.000	0.190	0.000	0.268	0.000

TABELA 7. 2: Resultado dos diferentes modelos para a probabilidade de freqüentar a escola – indivíduos de 15 a 17 anos.

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	
Efeito Fixo	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig
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	10.285	0.000	9.602	0.000	9.277	0.000	9.284	0.000
População até 4999	-0.788	0.000	-0.624	0.000	-0.658	0.000	-0.714	0.000
População entre 5000 e 9999	-0.781	0.000	-0.589	0.000	-0.622	0.000	-0.696	0.000
População entre 10000 e 19999	-0.711	0.000	-0.557	0.000	-0.584	0.000	-0.625	0.000
População entre 20000 e 49999	-0.635	0.000	-0.503	0.000	-0.525	0.000	-0.548	0.000
População entre 50000 e 99999	-0.413	0.000	-0.284	0.000	-0.301	0.000	-0.328	0.000
População entre 100000 e 199999	-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 espacial 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 4999							-0.166	0.062
Defasagem espacial População entre 5000 e 9999							-0.578	0.000
Defasagem espacial População entre 10000 e 19999							-0.513	0.000
Defasagem espacial População entre 20000 e 49999							-0.404	0.000
Defasagem espacial População entre 50000 e 99999							-0.370	0.001
Defasagem espacial População entre 100000 e 199999							-0.049	0.695
Efeito Aleatório								
Coeficiente	0.223	0.000	0.125	0.000	0.126	0.000	0.213	0.000

TABELA 7. 3: Resultado dos diferentes modelos para a probabilidade de freqüentar a escola – indivíduos de 15 a 17 anos que já concluíram o ensino fundamental.

Variáveis Independentes	Modelo Hierarquico (Modelo 6)		Modelo com a defasagem do resíduo bayesiano		Modelo com a defasagem do resíduo MQO		Modelo com a defasagem das variáveis independentes	
Efeito Fixo	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig
Intercepto	1.377	0.000	1.545	0.000	1.616	0.000	0.737	0.001
Tamanho relativo da coorte de 15 a 17 anos	18.258	0.000	13.574	0.000	12.772	0.000	9.604	0.000
Razão professor/pop. 15 a 17	8.152	0.000	6.755	0.000	6.904	0.000	8.355	0.000
População até 4999	-0.761	0.000	-0.532	0.000	-0.618	0.000	-0.719	0.000
População entre 5000 e 9999	-0.787	0.000	-0.522	0.000	-0.618	0.000	-0.737	0.000
População entre 10000 e 19999	-0.663	0.000	-0.436	0.000	-0.524	0.000	-0.621	0.000
População entre 20000 e 49999	-0.524	0.000	-0.350	0.000	-0.425	0.000	-0.484	0.000
População entre 50000 e 99999	-0.292	0.000	-0.131	0.025	-0.196	0.001	-0.257	0.001
População entre 100000 e 199999	-0.168	0.042	-0.102	0.132	-0.121	0.076	-0.155	0.057
Defasagem espacial do resíduo			1.063	0.000	0.609	0.000		
Defasagem espacial Tamanho relativo da coorte de 15 a 17 anos							18.124	0.000
Defasagem espacial Razão professor/pop. 15 a 17							1.233	0.266
Defasagem espacial População até 4999							0.225	0.074
Defasagem espacial População entre 5000 e 9999							-0.363	0.004
Defasagem espacial População entre 10000 e 19999							-0.210	0.092
Defasagem espacial População entre 20000 e 49999							0.049	0.700
Defasagem espacial População entre 50000 e 99999							0.070	0.645
Defasagem espacial População entre 100000 e 199999							0.115	0.518
Efeito Aleatório								
Coeficiente	0.346	0.000	0.221	0.000	0.226	0.000	0.332	0.000

GRÁFICO 7. 1: 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 os indivíduos de 7 a 14 anos.

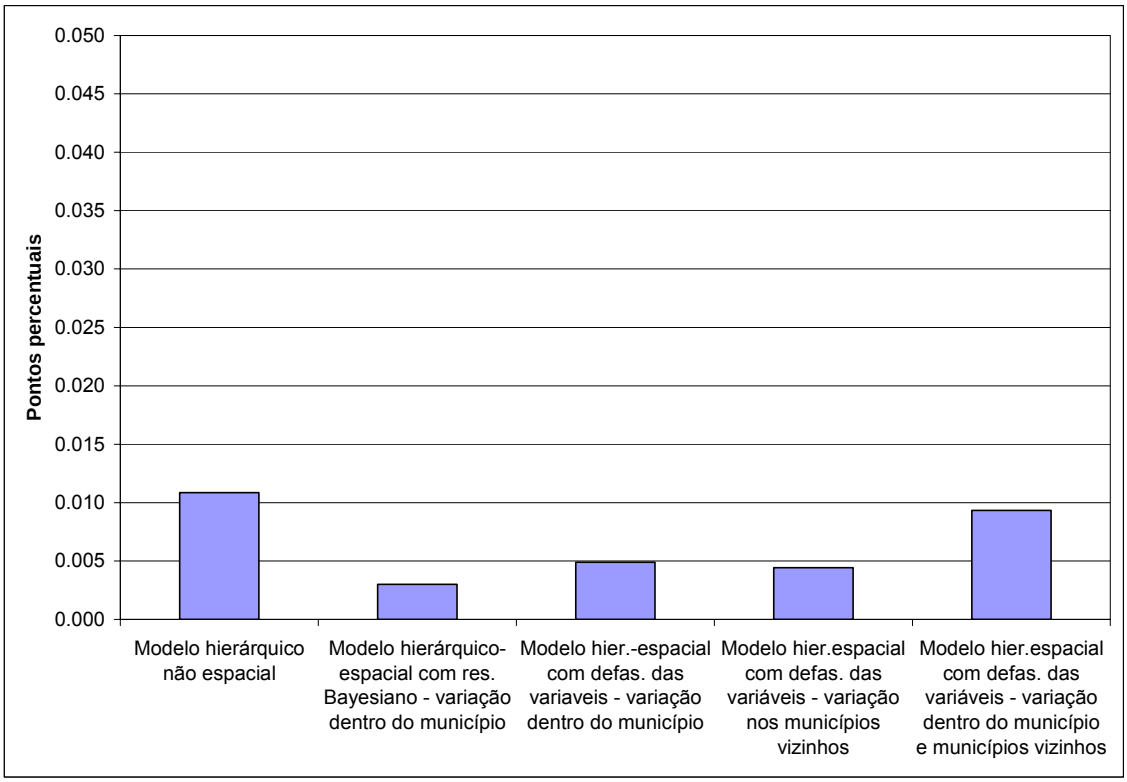


GRÁFICO 7. 2: Variação na probabilidade de cursar a escola decorrente do aumento de um desvio padrão em relação à média da razão professor/população em idade escolar – simulações segundo os diferentes modelos para os indivíduos de 7 a 14 anos

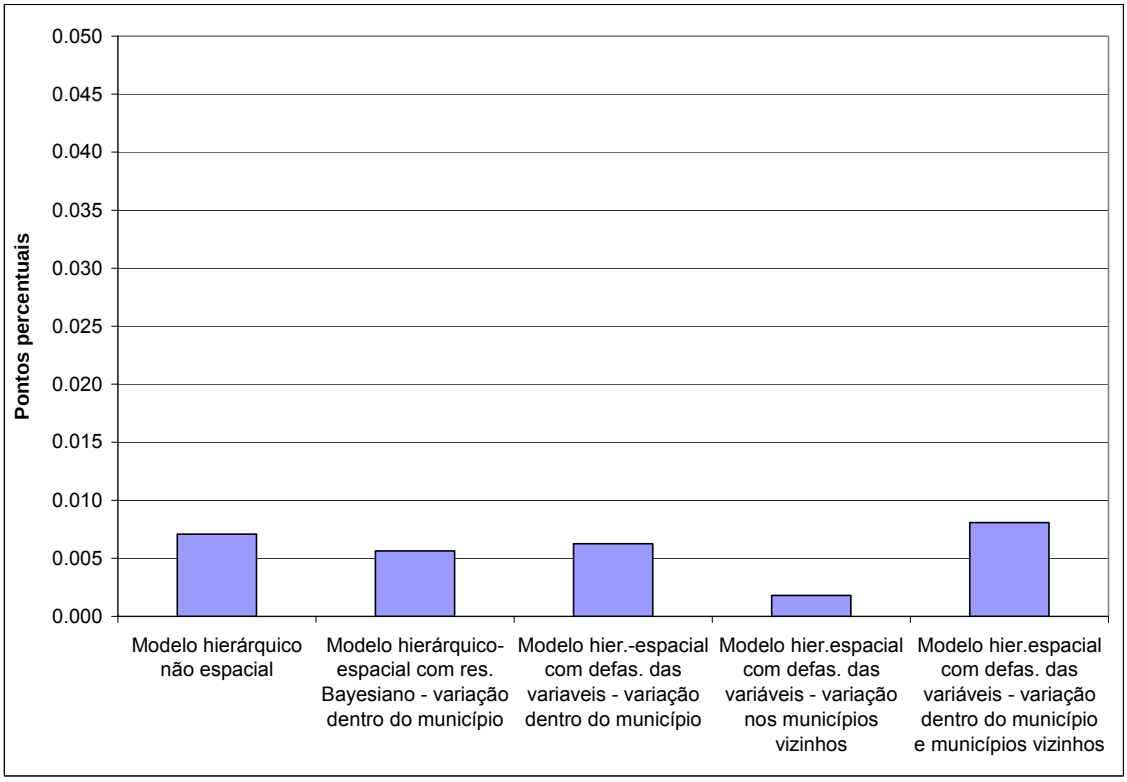


GRÁFICO 7. 3: Variação na probabilidade de cursar a escola decorrente da diminuição de um desvio padrão em relação à média do tamanho relativo da coorte em idade escolar – simulações segundo os diferentes modelos para os indivíduos de 15 a 17 anos.

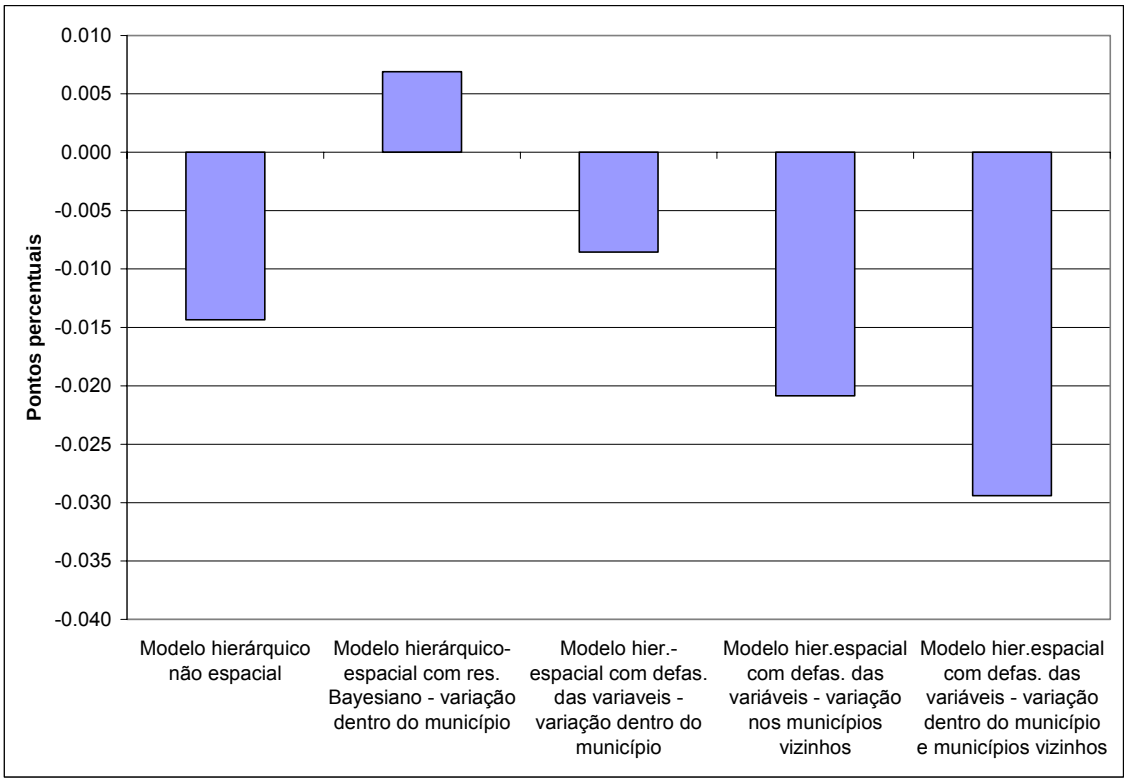


GRÁFICO 7. 4: Variação na probabilidade de cursar a escola decorrente do aumento de um desvio padrão em relação à média da razão professor/população em idade escolar – simulações segundo diferentes modelos para os indivíduos de 15 a 17 anos.

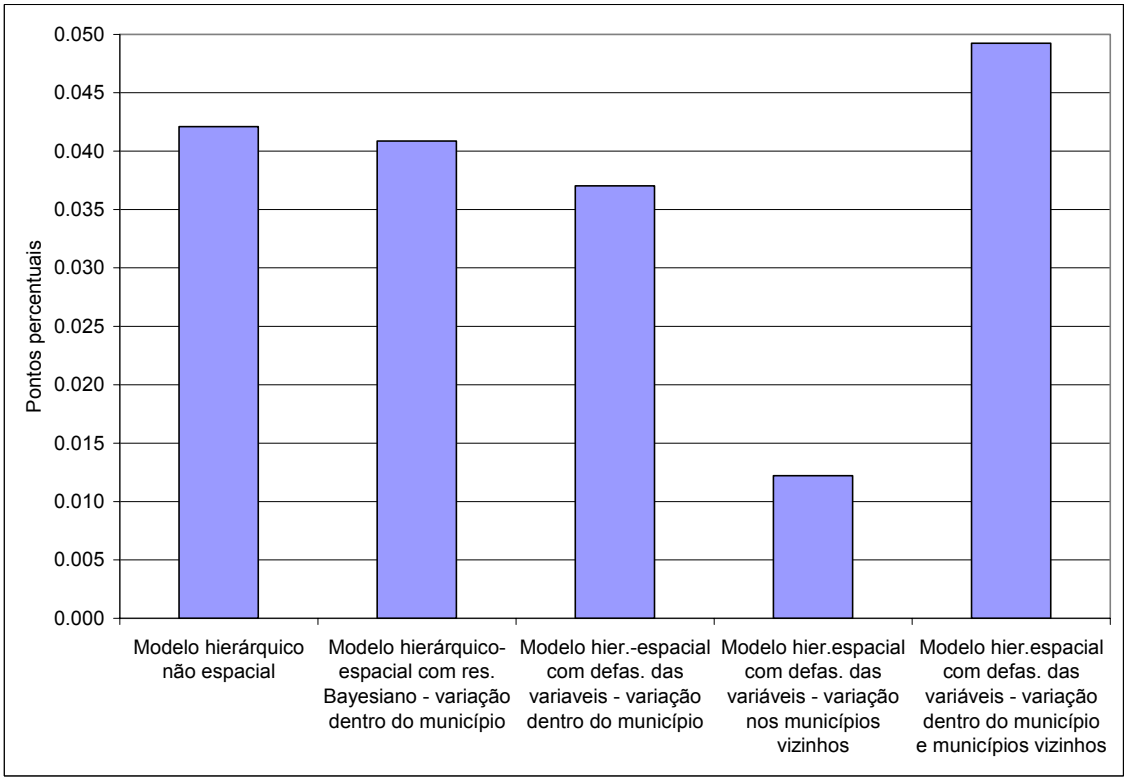


GRÁFICO 7. 5: Variação na probabilidade de cursar a escola decorrente da diminuição de um desvio padrão em relação à média do tamanho relativo da coorte em idade escolar – simulações segundo diferentes modelos para os indivíduos de 15 a 17 anos que já concluíram o ensino fundamental.

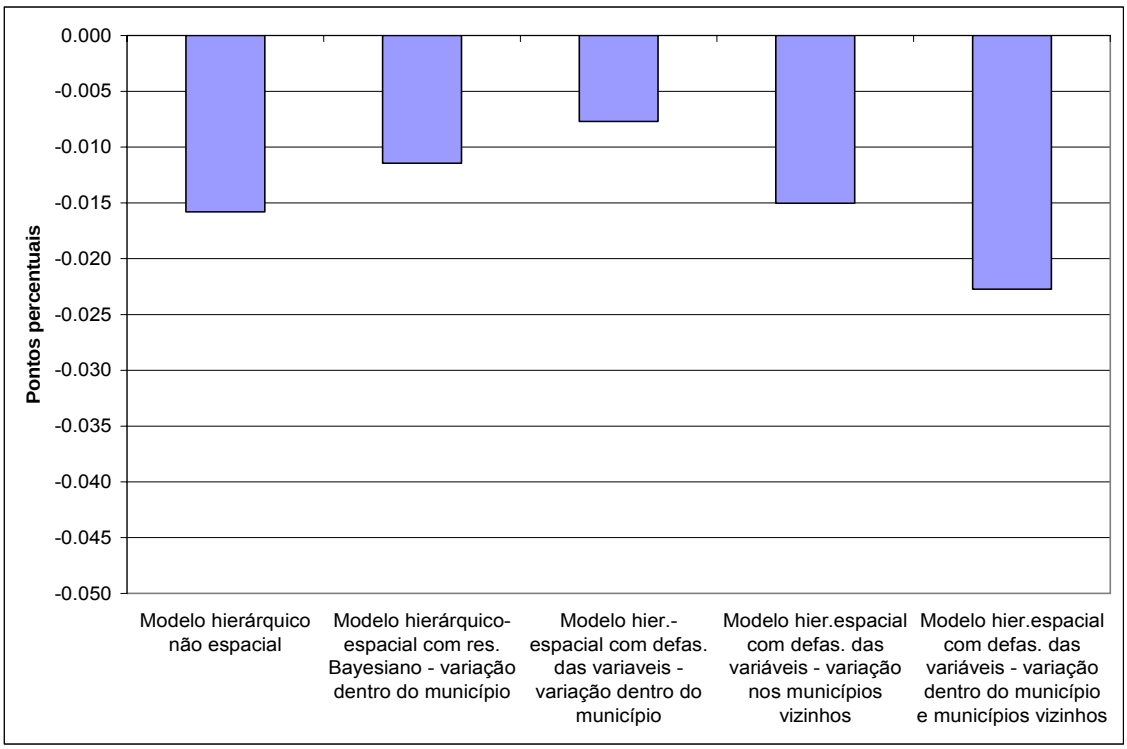
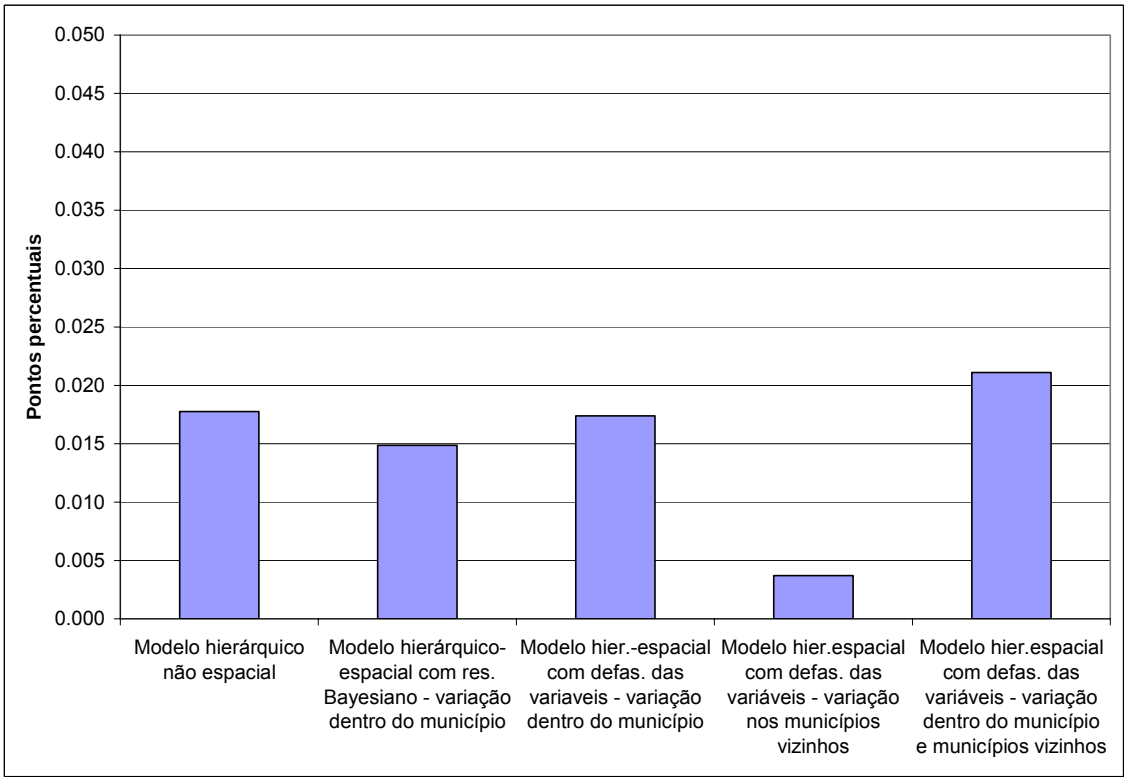


GRÁFICO 7. 6: Variação na probabilidade de cursar a escola decorrente do aumento de um desvio padrão em relação à média da razão professor/população em idade escolar – simulações segundo diferentes modelos para os indivíduos de 15 a 17 anos que já concluíram o ensino fundamental.



7.2) Probabilidade de freqüentar a escola na idade adequada

Para essa dimensão da educação, a análise hierárquica-espacial foi realizada para os coeficientes do intercepto e da educação materna, já que esse último também foi considerado como tendo efeito aleatório na análise do modelo hierárquico do Capítulo 5. Apenas para relembrar, as estimações hierárquicas para a probabilidade de freqüência escolar na idade correta e para as PPSs foram feitas assumindo-se que o intercepto e o coeficiente da educação materna possuíam efeitos aleatórios e que havia um conjunto de covariáveis do nível 2 que explicava parte dessa aleatoriedade. No caso do modelo hierárquico-espacial, devido ao grande número de variáveis envolvidas na análise, principalmente no modelo com a defasagem espacial das variáveis independentes, não foram incluídas variáveis de nível 2 no coeficiente da educação materna, quando a ênfase na estimativa das regressões era o intercepto, apesar de considerá-lo aleatório. Do mesmo modo, nas regressões cujo interesse era o impacto das variáveis contextuais na educação materna, considerou-se o intercepto aleatório, mas não foram incluídas variáveis no nível 2.

Considerando-se primeiramente o impacto no intercepto dos fatores da rede escolar dos municípios na probabilidade das crianças de 7 a 14 anos freqüentarem a escola na idade adequada (TAB. 7.4), percebe-se que nos modelos em que se corrige para a autocorrelação espacial, a maioria dos coeficientes continua significativa estatisticamente, porém, seus efeitos são atenuados. Comparando-se os dois modelos hierárquico-espaciais, observa-se que a principal diferença é que no modelo que inclui as defasagens espaciais das variáveis independentes[24], o coeficiente da porcentagem de professores com curso superior passa a ser não significativo, enquanto que a queda do coeficiente da média de horas-aula diárias é menor.

Para melhor visualização, o GRAF. 7.7 mostra a diminuição ou aumento nessa probabilidade decorrente do aumento de um desvio padrão em relação à média de alguns fatores de perfil da rede escolar, considerando os diferentes modelos. Com o objetivo de não poluir o gráfico foram selecionados apenas os fatores com maior impacto, no GRAF. A4.1 do Anexo 4 encontra-se o efeito de todos os fatores.

[24] Não foram incluídas as variáveis defasadas espacialmente das *dummies* de tamanho da população. Essa exclusão se deve ao fato do modelo não ter convergido quando as considerava. Tal procedimento não deve trazer grandes problemas, já que a estatística I de Moran aponta baixa autocorrelação espacial nessas variáveis (ver TAB. 6.1).

Através desse gráfico observa-se melhor os efeitos diretos e indiretos das variáveis contextuais. A barra azul escura refere-se ao efeito encontrado no modelo hierárquico visto no Capítulo 5, o qual não separa esses dois efeitos. Já as barras rosa e amarela mostram o efeito direto, ou seja, a variação na probabilidade decorrente de mudanças nos fatores contextuais do próprio município. No caso da barra rosa o efeito direto foi medido considerando-se o modelo hierárquico-espacial com a defasagem dos resíduos *bayesianos* e a amarela com as defasagens espaciais das variáveis de nível 2. Como já mencionado, a diferença do efeito direto encontrada utilizando os diferentes modelos deve-se ao fato das variáveis serem correlacionadas com outros fatores que não foram incluídos na análise e desses apresentarem autocorrelação espacial. Entretanto, as diferenças são pequenas indicando que o efeito espacial das variáveis omitidas estão influenciando pouco os coeficientes estimados.

O efeito indireto, ou seja, uma mudança na probabilidade em função de mudanças nas variáveis de nível 2 dos municípios vizinhos, é mostrado pela barra azul clara. A barra vermelha mostra a soma dos efeitos diretos e indiretos. Percebe-se que a incorporação dos efeitos espaciais diminui o efeito das variáveis de nível 2 quando se compara com o modelo hierárquico, porém, essa perda de certa forma é recuperada pelo efeito via seus vizinhos (indireto), de modo que a soma dos dois efeitos é em muitos casos próxima do modelo hierárquico ou maior.

Esses resultados fornecem indícios de que uma grande parte do efeito dos fatores de rede escolar dos municípios, obtidos através do modelo hierárquico não-espacial, está associada ao contexto espacial mais amplo no qual o município está inserido. Em outras palavras, pode-se dizer que políticas de melhoria do perfil escolar dos municípios têm maior impacto quando não são focalizadas apenas em um município, mas quando abrangem uma dimensão espacial mais ampla.

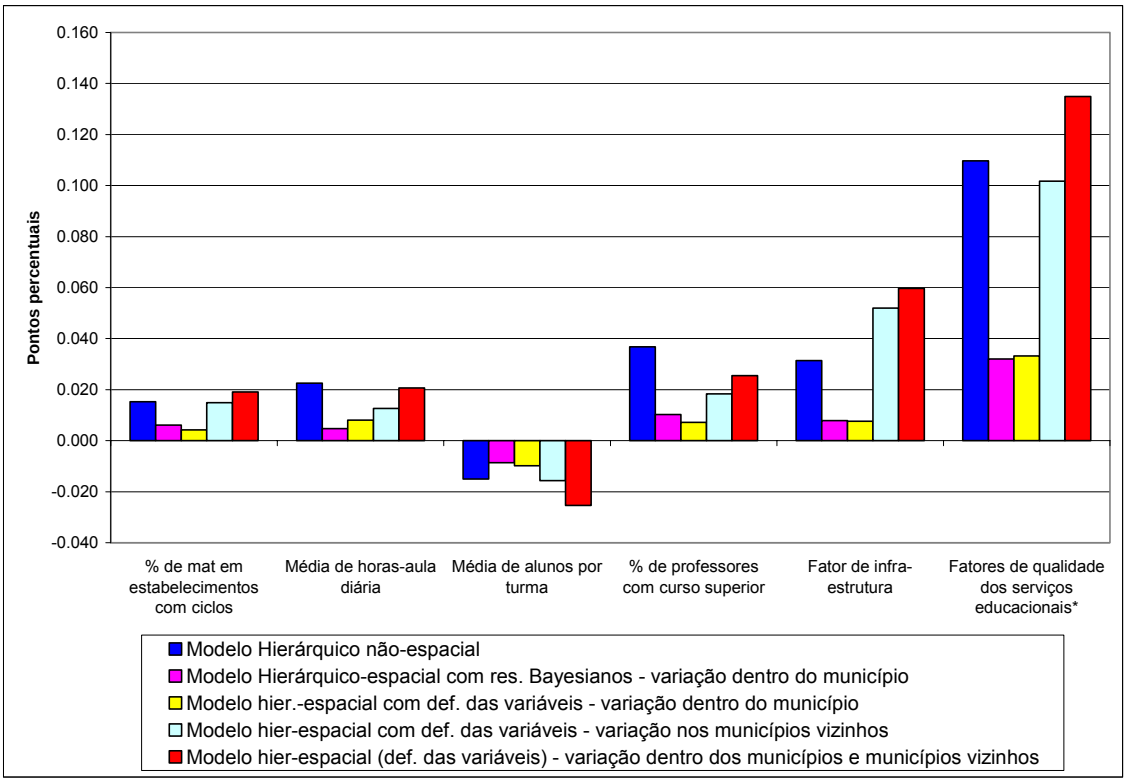
Dois fatos também merecem ser destacados. Em primeiro lugar, o expressivo impacto da variação do fator de infra-estrutura e da variação simultânea dos fatores de qualidade dos serviços educacionais dos municípios vizinhos na probabilidade dos indivíduos de 7 a 14 anos cursarem a escola na idade correta no município local. O efeito indireto dessas duas variações é bem maior que o efeito direto, conforme pode ser observado no GRAF. 7.7. Uma melhora simultânea nos fatores de qualidade escolar de um desvio padrão em relação à média nos municípios vizinhos aumenta essa

probabilidade em quase 14 pontos percentuais, enquanto que a melhora no próprio município provoca um aumento próximo de 3 pontos percentuais. Em segundo lugar, a média de alunos por turma foi a que apresentou a menor diminuição do efeito, quando comparada ao modelo hierárquico não-espacial. Tal fato pode ser importante, já que o MAPA 6.12 mostra que para os valores significativos do I de Moran local, há manchas de aglomerados de municípios se destacando em praticamente todos os estados, enquanto que na maioria das demais variáveis, as manchas de valores significativos cobrem praticamente todo o estado ou, ao contrário, todo o estado não possui valor significativo. Dessa forma, pode-se cogitar a hipótese de que a média de horas-aula diárias seja uma variável importante de diferenciação entre a política municipal e a estadual.

TABELA 7. 4: Resultado dos diferentes modelos para o intercepto da probabilidade de freqüentar a escola na idade correta – ensino fundamental.

Variáveis Independentes	Modelo Hierarquico (Modelo 4)		Modelo com a defasagem do resíduo bayesiano		Modelo com a defasagem do resíduo MQO		Modelo com a defasagem das variáveis independentes	
Efeito Fixo	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig
Intercepto	-4.520	0.000	-1.356	0.000	-1.358	0.000	-4.123	0.000
Razão professor/pop. 7 a 14	-0.800	0.029	-0.306	0.307	-0.291	0.331	-0.604	0.122
Proporção de matrículas em escolas com ciclo	0.179	0.000	0.072	0.000	0.073	0.000	0.050	0.018
Proporção de matrículas públicas	1.301	0.000	0.031	0.803	0.023	0.852	0.474	0.002
Proporção de matrículas municipais dentre as públicas	-0.173	0.000	-0.119	0.000	-0.123	0.000	-0.123	0.000
Proporção de estabelecimentos com 50 a 200 matrículas	-0.139	0.000	-0.062	0.010	-0.065	0.007	-0.064	0.030
Proporção de estabelecimentos com 201 a 1000	-0.062	0.027	-0.024	0.291	-0.023	0.308	0.004	0.906
Proporção de estabelecimentos com mais de 1000 matrículas	0.164	0.118	-0.039	0.642	-0.018	0.834	0.186	0.067
Educação média da população adulta	0.261	0.000	0.127	0.000	0.129	0.000	0.219	0.000
População até 4999	0.392	0.000	0.030	0.399	0.035	0.322	0.119	0.006
População entre 5000 e 9999	0.381	0.000	0.045	0.184	0.050	0.147	0.134	0.002
População entre 10000 e 19999	0.338	0.000	0.037	0.257	0.039	0.231	0.117	0.003
População entre 20000 e 49999	0.269	0.000	0.034	0.275	0.035	0.264	0.090	0.016
População entre 50000 e 99999	0.218	0.000	0.066	0.031	0.066	0.031	0.093	0.012
População entre 100000 e 199999	0.135	0.003	0.015	0.658	0.018	0.606	0.028	0.494
Horas-aula diárias	0.314	0.000	0.067	0.001	0.069	0.000	0.112	0.000
Alunos por turma	-0.007	0.000	-0.004	0.000	-0.004	0.000	-0.005	0.000
Proporção professores com curso superior	0.563	0.000	0.159	0.000	0.158	0.000	0.111	0.014
Fator de infra-estrutura	0.135	0.000	0.034	0.000	0.034	0.000	0.033	0.000
Defasagem espacial do resíduo			0.673	0.000	0.639	0.000		
Defasagem espacial Razão professor/pop. 7 a14							-2.487	0.000
Defasagem espacial da educação média da população adulta							-0.024	0.067
Def. espacial da proporção de matrículas em escolas com ciclo							0.207	0.000
Def. espacial da proporção de matrículas públicas							1.551	0.000
Def. espacial proporção de matrículas municipais dentre as públicas							-0.133	0.003
Defasagem espacial do fator de infra-estrutura							0.256	0.000
Def. espacial Proporção de estabelecimentos com 50 a 200 matrículas							-0.246	0.000
Def. espacial Proporção de estabelecimentos com 201 a 1000							-0.278	0.000
Def. espacial Proporção de estabelecimentos com mais de 1000 matrículas							-0.214	0.258
Defasagem espacial da média de horas-aula diárias							0.203	0.000
Defasagem espacial de alunos por turma							-0.011	0.000
Def. espacial da proporção de professores com curso superior							0.304	0.000
Efeito Aleatório								
Coeficiente	0.107	0.000	0.062	0.000	0.062	0.000	0.087	0.000

GRÁFICO 7. 7: Variação na probabilidade de cursar a escola na idade adequada decorrente do aumento de um desvio padrão em relação à média de variáveis selecionadas – simulações segundo diferentes modelos para os indivíduos de 7 a 14 anos.



*Nota: Na simulação da melhoria nos fatores de qualidade dos serviços educacionais, para média de alunos por turma, considerou a diminuição de um desvio padrão em relação à média.

Quanto ao efeito das variáveis agregadas no coeficiente da educação materna, observa-se pela TAB. 7.5, que o *trade off*[25] que existia anteriormente entre essa variável e aquelas relacionadas ao perfil escolar dos municípios, praticamente desaparece, restando apenas um fraco efeito substituição com a proporção de professores com curso superior. Desta forma, pode-se dizer que a melhoria na qualidade da rede escolar dos municípios minimiza pouco a dependência que existe entre o sucesso educacional do aluno e a sua origem socioeconômica, quando o espaço é considerado na análise.

Para melhor ilustrar, o GRAF. 7.8 mostra o impacto de uma variação simultânea na porcentagem de professores com curso superior, na média de alunos por turma e no fator de infra-estrutura – que são os que apresentam efeito substituição no modelo hierárquico não-espacial - na freqüência em idade correta das crianças de 7 a 14 anos,

[25] O *trade off* refere-se ao efeito substituição das variáveis agregadas no coeficiente da educação materna, ou seja, filhos de mães com baixa escolaridade são mais beneficiados pela melhora da qualidade da rede de ensino municipal que filhos de mães mais escolarizadas.

considerando os coeficientes do modelo hierárquico-espacial com resíduos *bayesianos*. Comparando esse gráfico com o GRAF. 5.11 – que foi reproduzido para facilitar a leitura (GRAF. 7.9) - percebe-se mais claramente a diminuição do efeito substituição desses fatores com a educação da mãe.

TABELA 7. 5: Resultado dos diferentes modelos para o coeficiente da educação materna da probabilidade de freqüentar a escola na idade correta – ensino fundamental.

Variáveis Independentes	Modelo Hierarquico (Modelo 4)		Modelo com a defasagem do resíduo bayesiano		Modelo com a defasagem do resíduo MQO		Modelo com a defasagem das variáveis independentes	
Efeito Fixo	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig
Intercepto	0.210	0.000	0.133	0.000	0.171	0.000	-0.031	0.528
Razão professor/pop. 7 a 14	0.292	0.000	0.069	0.199	0.107	0.046	0.118	0.076
Proporção de matrículas em escolas com ciclo	0.006	0.030	0.001	0.704	-0.001	0.622	-0.010	0.004
Proporção de matrículas públicas	-0.166	0.000	-0.066	0.001	-0.090	0.000	-0.083	0.000
Proporção de matrículas municipais dentre as públicas	0.014	0.001	0.000	0.966	0.000	0.933	-0.005	0.268
Proporção de estabelecimentos com 50 a 200 matrículas	-0.002	0.634	-0.006	0.175	-0.005	0.291	-0.003	0.600
Proporção de estabelecimentos com 201 a 1000	0.006	0.193	0.002	0.698	0.001	0.817	0.002	0.639
Proporção de estabelecimentos com mais de 1000 matrículas	0.032	0.034	0.024	0.065	0.022	0.088	0.015	0.328
Educação média da população adulta	-0.014	0.000	-0.006	0.000	-0.006	0.000	-0.007	0.000
População até 4999	-0.041	0.000	-0.011	0.037	-0.016	0.002	-0.012	0.050
População entre 5000 e 9999	-0.033	0.000	-0.007	0.151	-0.012	0.014	-0.008	0.152
População entre 10000 e 19999	-0.027	0.000	-0.004	0.327	-0.008	0.066	-0.006	0.285
População entre 20000 e 49999	-0.020	0.000	-0.002	0.599	-0.005	0.217	-0.003	0.558
População entre 50000 e 99999	-0.012	0.014	-0.002	0.547	-0.004	0.350	-0.003	0.577
População entre 100000 e 199999	-0.005	0.303	0.000	0.918	0.000	0.988	0.000	0.921
Horas-aula diárias	0.024	0.000	0.011	0.001	0.009	0.003	-0.006	0.198
Alunos por turma	0.001	0.000	0.000	0.039	0.000	0.052	0.000	0.036
Proporção professores com curso superior	-0.062	0.000	-0.014	0.002	-0.023	0.000	-0.019	0.009
Fator de infra-estrutura	-0.008	0.000	0.001	0.470	-0.001	0.488	0.000	0.999
Defasagem espacial do resíduo			1.107	0.000	0.607	0.000		
Defasagem espacial Razão professor/pop. 7 a14							0.620	0.000
Defasagem espacial da educação média da população adulta							0.000	0.915
Def. espacial da proporção de matrículas em escolas com ciclo							0.021	0.000
Def. espacial da proporção de matrículas públicas							-0.084	0.024
Def. espacial proporção de matrículas municipais dentre as públicas							0.060	0.000
Defasagem espacial do fator de infra-estrutura							-0.010	0.000
Def. espacial Proporção de estabelecimentos com 50 a 200 matrículas							0.018	0.034
Def. espacial Proporção de estabelecimentos com 201 a 1000							0.007	0.378
Def. espacial Proporção de estabelecimentos com mais de 1000 matrículas							0.076	0.006
Defasagem espacial da média de horas-aula diárias							0.055	0.000
Defasagem espacial de alunos por turma							0.001	0.000
Def. espacial da proporção de professores com curso superior							-0.059	0.000
Efeito Aleatório								
Educaçã da mãe	0.001	0.000	0.001	0.000	0.001	0.000	0.001	0.000

GRÁFICO 7. 8: Impacto da variação da educação da mãe e da qualidade da rede escolar do município na probabilidade de freqüentar a escola na idade correta (Modelo hierárquico-espacial com o resíduo *bayesiano*) - ensino fundamental

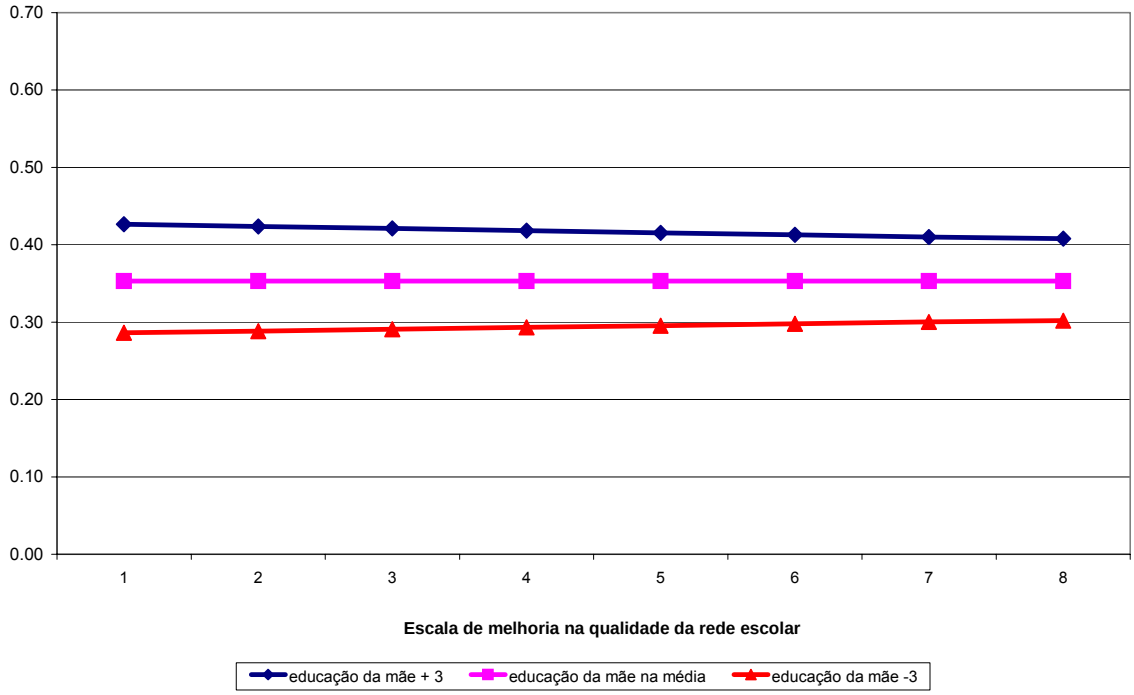
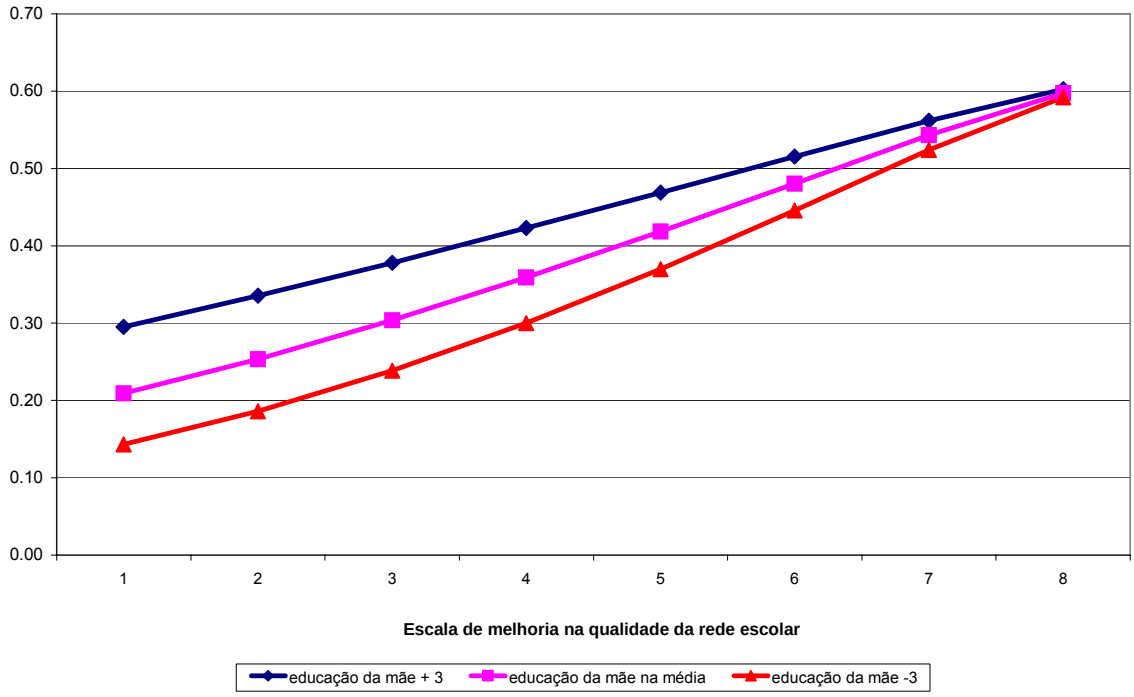


GRÁFICO 7. 9: Impacto da variação da educação da mãe e da qualidade da rede escolar do município na probabilidade de freqüentar a escola na idade correta - ensino fundamental (Modelo hierárquico)



Os resultados do modelo hierárquico-espacial para a probabilidade dos jovens de 15 a 17 anos freqüentarem a escola na idade correta para os coeficientes do intercepto e da educação materna, encontram-se respectivamente nas TABs. 7.6 e 7.7.

Analisando-se primeiramente o resultado para o intercepto, percebe-se que, da mesma forma que ocorreu no ensino fundamental, em todos os modelos hierárquico-espaciais, as variáveis relacionadas à rede escolar do município perdem importância. Dentre essas variáveis, apenas a porcentagem de escolas municipais dentre as públicas, o fator de infra-estrutura e a razão professor/população em idade escolar, continuam significativas estatisticamente, embora com coeficientes menores. A razão professor/população em idade escolar é o fator que possui maior impacto direto, ou seja, uma variação na restrição de oferta escolar dentro do próprio município é o fator que produz maior impacto nessa probabilidade, o que reafirma a sua importância para o ensino médio. Por outro lado, observa-se que mudanças nos fatores de perfil escolar e da razão professor/população em idade escolar dos municípios vizinhos possuem maior impacto que uma variação local, confirmando a importância de políticas educacionais mais abrangentes, não apenas locais.

O GRAF. 7.10 mostra visualmente os efeitos diretos e indiretos do aumento de um desvio padrão em relação à média de alguns fatores selecionados, considerando-se os diferentes modelos. Através desse gráfico observa-se que média de horas-aula diárias, número médio de alunos por turma e porcentagem de professores com curso superior praticamente não possuem efeito direto, mas verifica-se a existência do efeito indireto.

Já a razão professor/população em idade escolar, a nucleação, o fator de infra-estrutura e uma variação simultânea nos fatores de qualidade escolar apresentam efeitos direto e indireto importantes. No caso da nucleação, esse resultado é contrário ao encontrado para o ensino fundamental. Pode-se dizer, portanto, que um maior número de estabelecimentos de nível médio de grande porte possui importante impacto, tanto direto quanto indireto, na probabilidade dos indivíduos de 15 a 17 anos cursarem a escola na idade correta, e que a soma dos dois efeitos é próxima ao captado no modelo hierárquico não-espacial.

É importante destacar que, em uma análise comparativa entre os dois níveis de ensino, observa-se que o impacto dos fatores contextuais é bem maior para o ensino

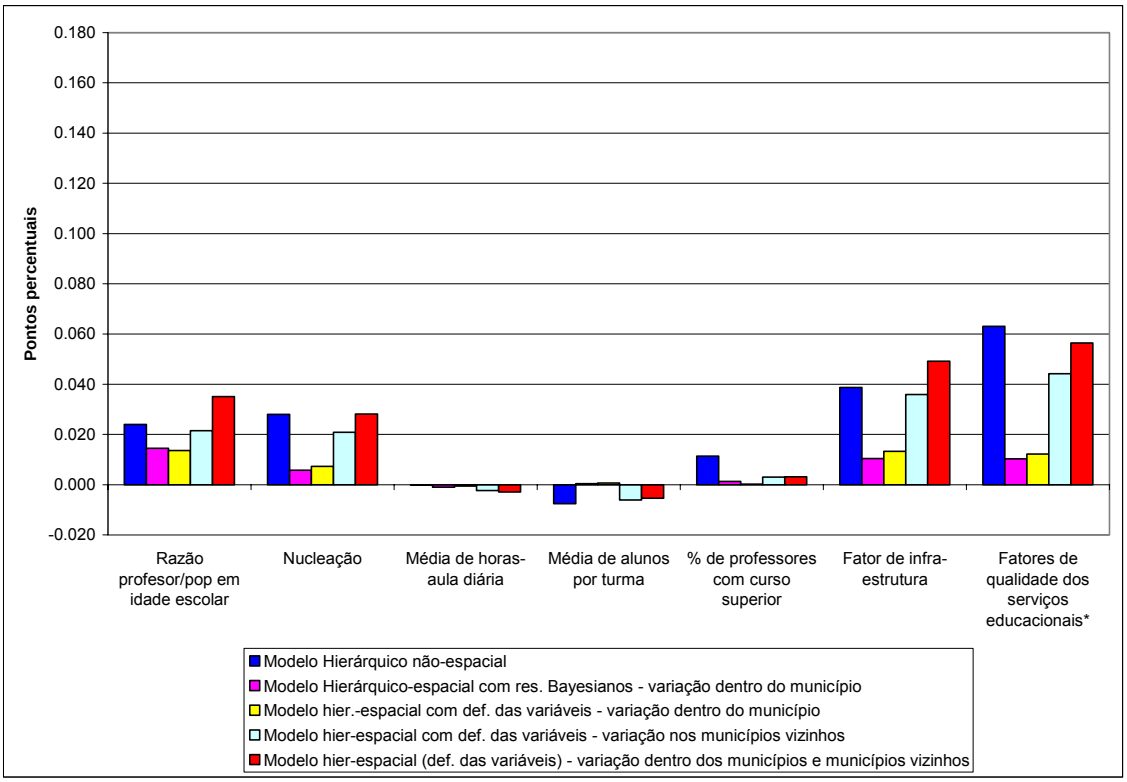
fundamental que para o médio, com exceção da razão professor/população em idade escolar e da nucleação.

Partindo para a análise das regressões para o coeficiente da educação materna (TAB. 7.7), verifica-se que nos modelos hierárquico-espaciais ocorre a total perda de significância das variáveis contextuais, apenas média de horas-aula diárias continuou significativa estatisticamente, porém sem efeito substituição. Interessante notar, que no modelo que inclui a defasagem espacial das variáveis independentes, média de alunos por turma e porcentagem de professores com curso superior defasadas espacialmente possuem efeito substituição com a educação materna.

TABELA 7. 6: Resultado dos diferentes modelos para o intercepto da probabilidade de freqüentar a escola na idade correta – ensino médio.

Variáveis Independentes	Modelo Hierarquico (Modelo 4)		Modelo com a defasagem do resíduo bayesiano		Modelo com a defasagem do resíduo MQO		Modelo com a defasagem das variáveis independentes	
Efeito Fixo	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig
Intercepto	-6.634	0.000	-3.323	0.000	-3.804	0.000	-5.894	0.000
Razão professor/pop. 15 a17	6.539	0.000	3.912	0.000	4.561	0.000	3.638	0.000
Proporção de estabelecimentos com 50 a 200 matrículas	0.164	0.018	-0.009	0.865	0.028	0.593	-0.001	0.992
Proporção de estabelecimentos com 201 a 1000	0.315	0.000	0.073	0.153	0.126	0.017	0.101	0.083
Proporção de estabelecimentos com mais de 1000 matrículas	0.208	0.051	0.176	0.021	0.195	0.015	0.145	0.105
Proporção de matrículas públicas	0.659	0.000	0.014	0.829	0.052	0.450	0.161	0.036
Proporção de matrículas municipais dentre as públicas	-0.327	0.000	-0.176	0.000	-0.187	0.000	-0.164	0.000
Educação média da população adulta	0.620	0.000	0.244	0.000	0.302	0.000	0.363	0.000
População até 4999	1.316	0.000	0.311	0.000	0.491	0.000	0.419	0.000
População entre 5000 e 9999	1.172	0.000	0.275	0.000	0.443	0.000	0.381	0.000
População entre 10000 e 19999	0.978	0.000	0.206	0.000	0.353	0.000	0.289	0.000
População entre 20000 e 49999	0.763	0.000	0.159	0.000	0.278	0.000	0.208	0.000
População entre 50000 e 99999	0.502	0.000	0.139	0.001	0.204	0.000	0.143	0.007
População entre 100000 e 199999	0.358	0.000	0.050	0.278	0.106	0.031	0.091	0.119
Fator de infra-estrutura	0.304	0.000	0.085	0.000	0.116	0.000	0.106	0.000
Horas-aula diárias	-0.002	0.948	-0.022	0.266	-0.006	0.772	-0.011	0.652
Alunos por turma	-0.010	0.000	0.001	0.690	-0.001	0.555	0.001	0.617
Proporção professores com curso superior	0.334	0.000	0.038	0.242	0.063	0.055	0.006	0.897
Defasagem espacial do resíduo			0.665	0.000	0.477	0.000		
Defasagem espacial da educação média da população adulta							0.142	0.000
Defasagem espacial População até 4999							0.859	0.000
Defasagem espacial População entre 5000 e 9999							1.023	0.000
Defasagem espacial População entre 10000 e 19999							0.955	0.000
Defasagem espacial População entre 20000 e 49999							0.695	0.000
Defasagem espacial População entre 50000 e 99999							0.488	0.000
Defasagem espacial População entre 100000 e 199999							0.475	0.000
Defasagem espacial Razão professor/pop. 15 a 17							8.423	0.000
Def. espacial Proporção de estabelecimentos com 50 a 200 matrículas							0.362	0.003
Def. espacial Proporção de estabelecimentos com 201 a 1000							0.446	0.000
Def. espacial Proporção de estabelecimentos com mais de 1000 matrículas							0.097	0.608
Def. espacial da proporção de matrículas públicas							0.084	0.531
Defasagem espacial da média de horas-aula diárias							-0.040	0.224
Defasagem espacial de alunos por turma							-0.009	0.001
Def. espacial da proporção de professores com curso superior							0.107	0.106
Defasagem espacial do fator de infra-estrutura							0.362	0.000
Def. espacial proporção de matrículas municipais dentre as públicas							-0.089	0.215
Efeito Aleatório								
Coeficiente	0.278	0.000	0.107	0.000	0.124	0.000	0.176	0.000

GRÁFICO 7. 10: Variação na probabilidade de cursar a escola na idade adequada decorrente do aumento de um desvio padrão em relação à média de variáveis selecionadas – simulações segundo diferentes modelos para os indivíduos de 15 a 17 anos.



*Nota: Na simulação da melhoria nos fatores de qualidade dos serviços educacionais, para média de alunos por turma, considerou a diminuição de um desvio padrão em relação à média.

TABELA 7. 7: Resultado dos diferentes modelos para o coeficiente da educação materna da probabilidade de freqüentar a escola na idade correta – ensino médio.

Variáveis Independentes	Modelo Hierarquico (Modelo 4)		Modelo com a defasagem do resíduo bayesiano		Modelo com a defasagem do resíduo MQO		Modelo com a defasagem das variáveis independentes	
Efeito Fixo	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig
Intercepto	0.330	0.000	0.063	0.006	0.078	0.001	0.055	0.045
Razão professor/pop. 15 a17	-0.171	0.029	0.080	0.281	0.097	0.198	0.031	0.690
Proporção de estabelecimentos com 50 a 200 matrículas	-0.019	0.045	-0.011	0.204	-0.012	0.173	-0.009	0.290
Proporção de estabelecimentos com 201 a 1000	-0.028	0.003	-0.013	0.133	-0.014	0.105	-0.012	0.179
Proporção de estabelecimentos com mais de 1000 matrículas	-0.022	0.099	-0.013	0.269	-0.011	0.378	-0.015	0.207
Proporção de matrículas públicas	-0.065	0.000	-0.010	0.338	-0.018	0.096	-0.006	0.557
Proporção de matrículas municipais dentre as públicas	0.024	0.001	-0.006	0.331	-0.006	0.309	-0.012	0.060
Educação média da população adulta	-0.033	0.000	0.000	0.787	-0.002	0.211	0.004	0.029
População até 4999	-0.075	0.000	-0.001	0.904	-0.008	0.289	0.003	0.687
População entre 5000 e 9999	-0.063	0.000	0.002	0.775	-0.004	0.494	0.004	0.493
População entre 10000 e 19999	-0.051	0.000	0.003	0.584	-0.002	0.647	0.005	0.391
População entre 20000 e 49999	-0.041	0.000	0.002	0.705	-0.002	0.615	0.003	0.564
População entre 50000 e 99999	-0.035	0.000	-0.008	0.074	-0.009	0.033	-0.007	0.114
População entre 100000 e 199999	-0.026	0.000	-0.005	0.209	-0.008	0.094	-0.005	0.258
Fator de infra-estrutura	-0.016	0.000	0.001	0.643	-0.001	0.422	0.001	0.597
Horas-aula diárias	0.028	0.000	0.029	0.000	0.029	0.000	0.018	0.000
Alunos por turma	0.001	0.000	0.000	0.084	0.001	0.022	0.000	0.437
Proporção professores com curso superior	-0.032	0.000	-0.008	0.163	-0.010	0.086	0.009	0.206
Defasagem espacial do resíduo			0.279	0.000	0.006	0.137		
Defasagem espacial da educação média da população adulta							-0.012	0.000
Defasagem espacial População até 4999							0.021	0.108
Defasagem espacial População entre 5000 e 9999							0.014	0.229
Defasagem espacial População entre 10000 e 19999							0.021	0.047
Defasagem espacial População entre 20000 e 49999							0.022	0.026
Defasagem espacial População entre 50000 e 99999							0.018	0.111
Defasagem espacial População entre 100000 e 199999							0.024	0.058
Defasagem espacial Razão professor/pop. 15 a 17							0.135	0.300
Def. espacial Proporção de estabelecimentos com 50 a 200 matrículas							-0.017	0.296
Def. espacial Proporção de estabelecimentos com 201 a 1000							-0.004	0.809
Def. espacial Proporção de estabelecimentos com mais de 1000 matrículas							0.004	0.887
Def. espacial da proporção de matrículas públicas							-0.066	0.001
Defasagem espacial da média de horas-aula diárias							0.029	0.000
Defasagem espacial de alunos por turma							0.001	0.020
Def. espacial da proporção de professores com curso superior							-0.042	0.000
Defasagem espacial do fator de infra-estrutura							-0.004	0.133
Def. espacial proporção de matrículas municipais dentre as públicas							0.009	0.343
Efeito Aleatório								
Coeficiente	0.001	0.000	0.003	0.000	0.004	0.000	0.004	0.000

7.3) Probabilidade de Progressão por Série (PPS)

Para a probabilidade de progressão na 1ª série (e_0), a regressão hierárquico-espacial foi estimada tanto para o coeficiente do intercepto quanto para o da educação materna, adotando-se o procedimento anteriormente descrito, ou seja, quando a análise é para o intercepto não são incluídas variáveis de nível 2 para o coeficiente da educação materna – apesar de considerá-lo aleatório. Por outro lado, quando a análise está voltada para o efeito substituição entre a educação materna e alguns fatores do perfil escolar dos municípios, não são incluídas covariáveis de nível 2 no coeficiente do intercepto.

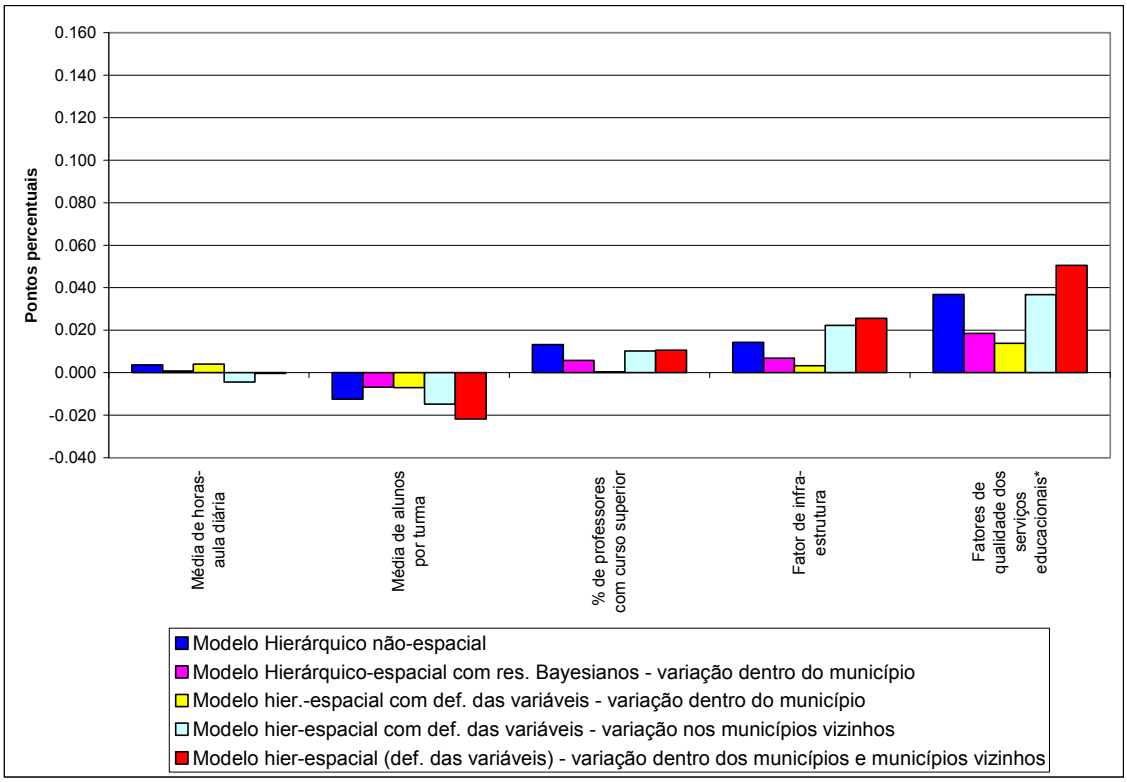
Analisando-se primeiramente os resultados para o intercepto (TAB. 7.8), observa-se que nos modelos em que se inclui a defasagem espacial dos resíduos, os fatores de rede escolar do município, que eram significativos no modelo hierárquico não-espacial, continuam significativos, apesar dos seus efeitos terem diminuído.

Conforme pode ser visto no GRAF. 7.11, no caso de e_0 o efeito direto não diminui muito quando comparado ao modelo hierárquico não-espacial. Parte da perda é compensada pelo efeito indireto desses fatores e a soma dos dois efeitos é próxima do encontrado no modelo hierárquico não-espacial. Entre os fatores de perfil da rede escolar do município, a média de alunos por turma, a porcentagem de professores com curso superior e o fator de infra-estrutura são as variáveis que apresentam maior efeito direto, sendo que esse último apresenta um impacto importante via seus vizinhos. Devido ao fato das variáveis de qualidade dos serviços educacionais serem importantes, uma variação simultânea desses fatores possui forte impacto. Uma melhoria nesses fatores de um desvio padrão em relação á média aumenta em cerca de 5 pontos percentuais essa probabilidade, sendo que quase 4 pontos percentuais são referentes ao efeito indireto.

TABELA 7. 8: Resultado dos diferentes modelos para o intercepto da probabilidade de progressão na 1ª série (e_0).

Variáveis Independentes	Modelo Hierárquico (Modelo 4)		Modelo com a defasagem do resíduo bayesiano		Modelo com a defasagem do resíduo MQO		Modelo com a defasagem das variáveis independentes	
Efeito Fixo	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig
Intercepto	-1.122	0.018	0.717	0.081	-0.950	0.029	2.836	0.001
Razão professor/pop. 7 a 14	1.414	0.126	1.000	0.214	1.493	0.080	0.532	0.593
Proporção de matrículas em escolas com ciclo	-0.024	0.575	-0.023	0.532	0.064	0.102	-0.023	0.672
Proporção de matrículas públicas	1.157	0.002	0.480	0.134	1.145	0.001	0.920	0.015
Proporção de matrículas municipais dentre as públicas	-0.203	0.002	-0.157	0.005	-0.213	0.001	-0.126	0.074
Proporção de estabelecimentos com 50 a 200 matrículas	0.107	0.155	0.106	0.107	0.075	0.280	0.098	0.203
Proporção de estabelecimentos com 201 a 1000	0.001	0.986	0.019	0.778	0.040	0.577	0.022	0.791
Proporção de estabelecimentos com mais de 1000 matrículas	-0.114	0.659	-0.208	0.356	0.040	0.870	0.230	0.385
Educação média da população adulta	0.353	0.000	0.234	0.000	0.340	0.000	0.371	0.000
População até 4999	0.378	0.000	0.038	0.662	0.343	0.000	0.095	0.338
População entre 5000 e 9999	0.405	0.000	0.088	0.281	0.346	0.000	0.134	0.151
População entre 10000 e 19999	0.390	0.000	0.097	0.210	0.339	0.000	0.145	0.100
População entre 20000 e 49999	0.242	0.006	0.030	0.675	0.200	0.012	0.055	0.504
População entre 50000 e 99999	0.139	0.105	0.038	0.590	0.126	0.101	0.042	0.597
População entre 100000 e 199999	0.109	0.249	0.050	0.512	0.123	0.147	0.052	0.548
Horas-aula diárias	0.133	0.020	0.028	0.572	0.103	0.051	0.153	0.045
Alunos por turma	-0.014	0.000	-0.008	0.000	-0.012	0.000	-0.008	0.000
Proporção professores com curso superior	0.556	0.000	0.241	0.001	0.488	0.000	0.016	0.891
Fator de infra-estrutura	0.169	0.000	0.080	0.000	0.151	0.000	0.039	0.129
Defasagem espacial do resíduo			0.706	0.000	0.168	0.000		
Defasagem espacial Razão professor/pop. 7 a14							-1.020	0.478
Defasagem espacial da educação média da população adulta							-0.134	0.000
Def. espacial da proporção de matrículas em escolas com ciclo							0.175	0.015
Def. espacial da proporção de matrículas públicas							-1.329	0.031
Def. espacial proporção de matrículas municipais dentre as públicas							-0.331	0.003
Defasagem espacial do fator de infra-estrutura							0.332	0.000
Def. espacial Proporção de estabelecimentos com 50 a 200 matrículas							-0.221	0.095
Def. espacial Proporção de estabelecimentos com 201 a 1000							-0.188	0.149
Def. espacial Proporção de estabelecimentos com mais de 1000 matrículas							-1.295	0.007
Defasagem espacial da média de horas-aula diárias							-0.184	0.070
Defasagem espacial de alunos por turma							-0.025	0.000
Def. espacial da proporção de professores com curso superior							0.472	0.003
Efeito Aleatório								
Coeficiente	0.398	0.000	0.335	0.000	0.390	0.000	0.396	0.000

GRÁFICO 7. 11: Variação na probabilidade de progressão na 1ª série decorrente do aumento de um desvio padrão em relação à média de variáveis selecionadas – simulações segundo diferentes modelos.



*Nota: Na simulação da melhoria nos fatores de qualidade dos serviços educacionais, para média de alunos por turma, considerou a diminuição de um desvio padrão em relação à média.

Com relação ao resultado da regressão hierárquico-espacial no coeficiente da educação materna (TAB. 7.9), observa-se, no modelo que inclui a defasagem dos resíduos, a permanência do *trade off* entre a educação materna e a porcentagem de professores com curso superior, porcentagem de matrículas em escolas com ciclos e fator de infra-estrutura (GRAF. 7.12), mesmo que em menor escala, conforme pode ser visto na comparação com o gráfico do modelo hierárquico do Capítulo 5 que é replicado aqui (GRAF. 7.13).

TABELA 7. 9: Resultado dos diferentes modelos para o coeficiente da educação materna da probabilidade de progressão na 1ª série (e_0).

Variáveis Independentes	Modelo Hierarquico (Modelo 4)		Modelo com a defasagem do resíduo bayesiano		Modelo com a defasagem do resíduo MQO		Modelo com a defasagem das variáveis independentes	
Efeito Fixo	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig
Intercepto	0.549	0.000	0.589	0.000	0.781	0.000	0.423	0.009
Razão professor/pop. 7 a 14	0.068	0.751	0.077	0.698	0.004	0.985	0.093	0.693
Proporção de matrículas em escolas com ciclo	-0.063	0.000	-0.037	0.000	-0.053	0.000	-0.020	0.091
Proporção de matrículas públicas	-0.262	0.001	-0.259	0.000	-0.358	0.000	-0.218	0.004
Proporção de matrículas municipais dentre as públicas	0.018	0.199	0.031	0.017	0.037	0.005	0.018	0.247
Proporção de estabelecimentos com 50 a 200 matrículas	0.012	0.511	-0.006	0.711	0.003	0.849	-0.028	0.143
Proporção de estabelecimentos com 201 a 1000	-0.021	0.244	-0.005	0.742	-0.015	0.379	-0.012	0.525
Proporção de estabelecimentos com mais de 1000 matrículas	-0.049	0.336	0.024	0.615	-0.036	0.453	-0.043	0.415
Educação média da população adulta	-0.020	0.000	-0.039	0.000	-0.045	0.000	-0.042	0.000
População até 4999	-0.067	0.001	-0.055	0.003	-0.083	0.000	-0.036	0.060
População entre 5000 e 9999	-0.043	0.014	-0.036	0.024	-0.062	0.000	-0.022	0.202
População entre 10000 e 19999	-0.035	0.028	-0.036	0.013	-0.059	0.000	-0.022	0.151
População entre 20000 e 49999	-0.032	0.030	-0.033	0.014	-0.049	0.001	-0.020	0.151
População entre 50000 e 99999	-0.031	0.023	-0.033	0.008	-0.044	0.001	-0.025	0.046
População entre 100000 e 199999	-0.025	0.071	-0.025	0.047	-0.035	0.008	-0.020	0.134
Horas-aula diárias	0.003	0.834	0.005	0.694	-0.006	0.614	0.006	0.733
Alunos por turma	0.000	0.505	0.001	0.003	0.001	0.000	0.001	0.019
Proporção professores com curso superior	-0.040	0.028	-0.063	0.000	-0.081	0.000	-0.030	0.252
Fator de infra-estrutura	-0.011	0.048	-0.014	0.009	-0.022	0.000	-0.003	0.659
Defasagem espacial do resíduo			1.066	0.000	0.003	0.497		
Defasagem espacial Razão professor/pop. 7 a14							0.137	0.672
Defasagem espacial da educação média da população adulta							0.007	0.350
Def. espacial da proporção de matrículas em escolas com ciclo							-0.061	0.000
Def. espacial da proporção de matrículas públicas							-0.128	0.304
Def. espacial proporção de matrículas municipais dentre as públicas							0.060	0.011
Defasagem espacial do fator de infra-estrutura							-0.037	0.000
Def. espacial Proporção de estabelecimentos com 50 a 200 matrículas							0.111	0.000
Def. espacial Proporção de estabelecimentos com 201 a 1000							0.020	0.466
Def. espacial Proporção de estabelecimentos com mais de 1000 matrículas							0.079	0.381
Defasagem espacial da média de horas-aula diárias							0.024	0.280
Defasagem espacial de alunos por turma							0.002	0.002
Def. espacial da proporção de professores com curso superior							-0.029	0.412
Efeito Aleatório								
Educação da mãe	0.005	0.000	0.009	0.000	0.008	0.000	0.008	0.000

GRÁFICO 7. 12: Impacto da variação da educação da mãe e da qualidade da rede escolar do município na probabilidade de freqüentar a escola na probabilidade de progressão da 1ª série (Modelo hierárquico-espacial com resíduo *bayesiano*)

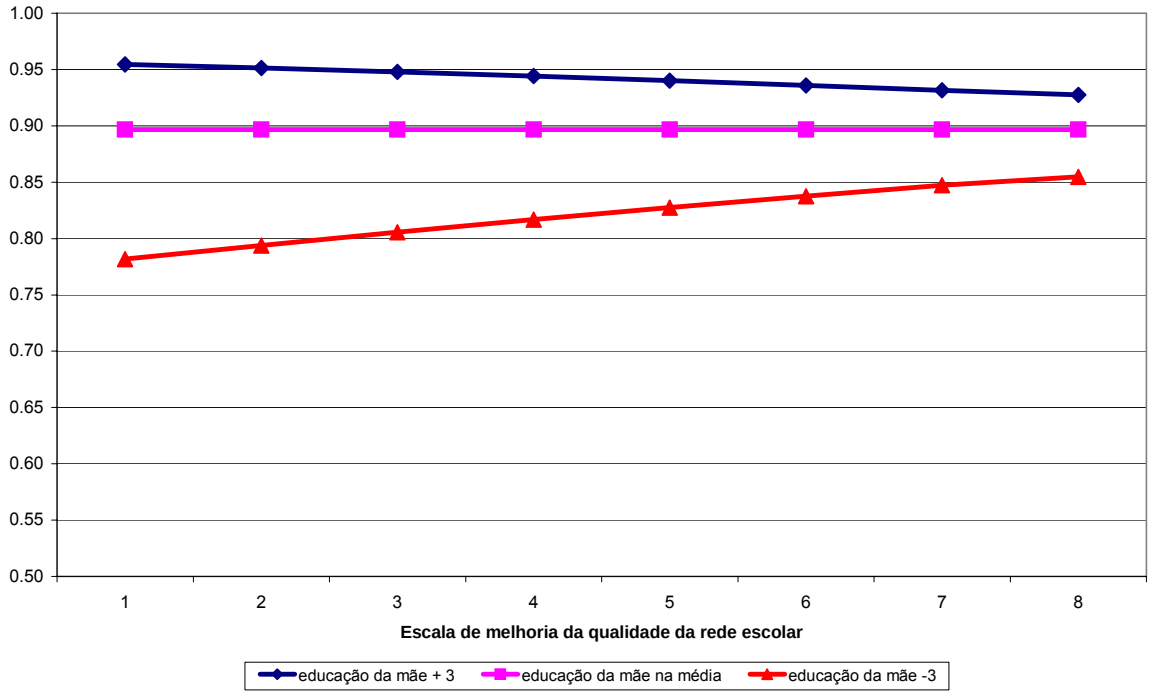
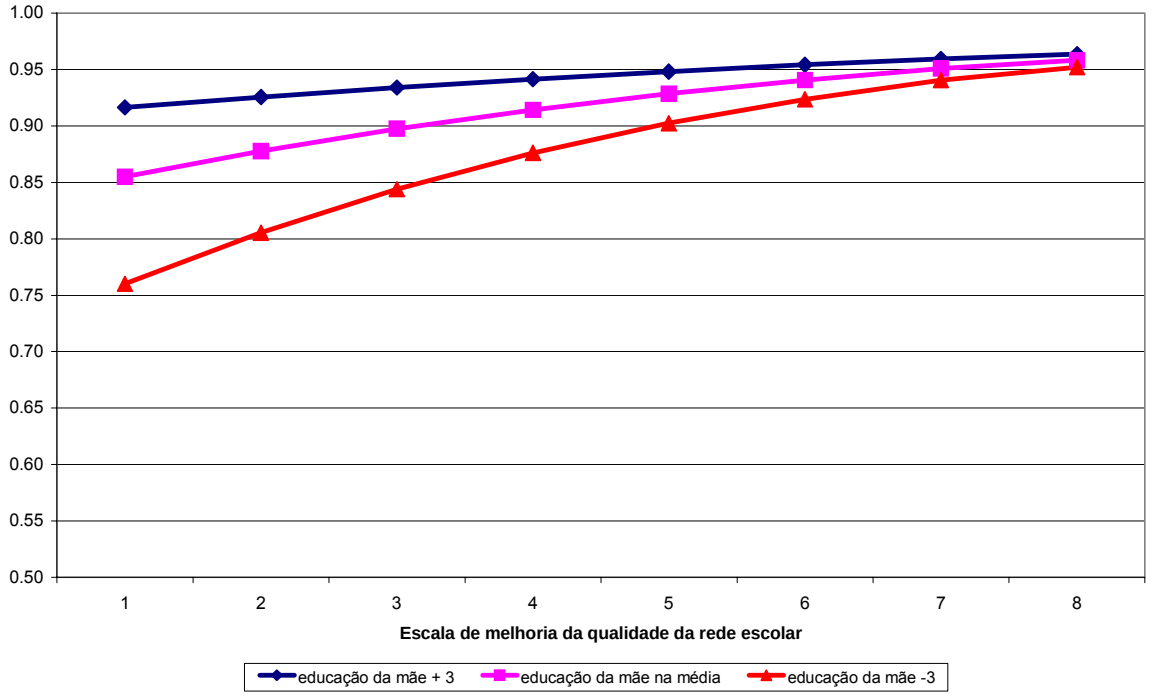


GRÁFICO 7. 13: Impacto da variação da educação da mãe e do fator de infra-estrutura na Probabilidade de Progressão na 1ª série (Modelo hierárquico)



Com relação às outras duas PPSs, probabilidade de progressão na 5ª série e na 1ª série do ensino médio, ou seja, e_4 e e_8, percebe-se pelas TABs. 7.10 e 7.11 que a inclusão dos efeitos espaciais provocam uma diminuição menor no impacto das variáveis de rede escolar do município, principalmente para e_8.

Nos GRAFs. 7.14 e 7.15 encontram-se os efeitos diretos e indiretos para e_4 e e_8, respectivamente. Para e_4, observa-se que os fatores de rede escolar do município que possuem maiores efeitos diretos e indiretos são a porcentagem de professores com curso superior e o fator de infra-estrutura, sendo o efeito indireto maior, o que reafirma a importância de políticas públicas voltadas para um contexto espacial mais amplo.

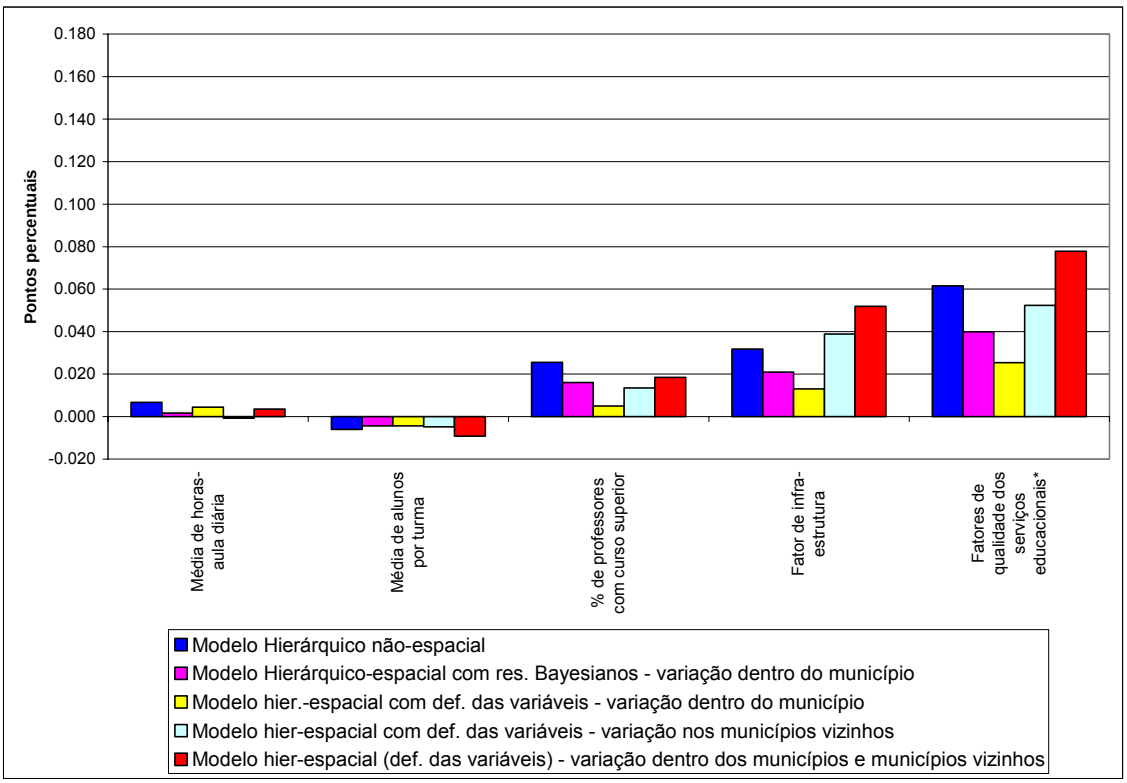
Para a probabilidade do ensino médio, e_8, a restrição de oferta continua sendo o fator mais importante, com expressivo impacto direto e indireto. Interessante notar, que o efeito direto da nucleação é o segundo mais importante, porém a defasagem espacial dessa variável possui efeito negativo.

Comparando as três probabilidades de progressão, observa-se que de modo geral, os efeitos diretos e indiretos dos fatores associados ao perfil escolar dos municípios aumentam de e_0 para e_4 e diminuem de e_4 para e_8.

TABELA 7. 10: Resultado dos diferentes modelos para o intercepto da probabilidade de progressão na 5ª série (e_4).

Variáveis Independentes	Modelo Hierarquico (Modelo 4)		Modelo com a defasagem do resíduo bayesiano		Modelo com a defasagem do resíduo MQO		Modelo com a defasagem das variáveis independentes	
Efeito Fixo	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig
Intercepto	-2.423	0.000	-0.337	0.327	-0.402	0.237	-1.480	0.013
Razão professor/pop. 7 a 14	1.191	0.091	1.482	0.030	1.550	0.023	2.220	0.006
Proporção de matrículas em escolas com ciclo	0.160	0.000	0.124	0.000	0.129	0.000	0.082	0.044
Proporção de matrículas públicas	1.567	0.000	0.564	0.027	0.554	0.029	0.765	0.007
Proporção de matrículas municipais dentre as públicas	-0.303	0.000	-0.239	0.000	-0.239	0.000	-0.235	0.000
Proporção de estabelecimentos com 50 a 200 matrículas	0.118	0.039	0.138	0.013	0.129	0.020	0.127	0.038
Proporção de estabelecimentos com 201 a 1000	0.194	0.001	0.132	0.016	0.133	0.016	0.138	0.034
Proporção de estabelecimentos com mais de 1000 matrículas	-0.235	0.198	-0.535	0.002	-0.390	0.023	-0.160	0.407
Educação média da população adulta	0.311	0.000	0.228	0.000	0.226	0.000	0.337	0.000
População até 4999	0.457	0.000	0.195	0.004	0.162	0.017	0.214	0.003
População entre 5000 e 9999	0.390	0.000	0.159	0.011	0.128	0.040	0.170	0.010
População entre 10000 e 19999	0.357	0.000	0.160	0.007	0.129	0.028	0.162	0.009
População entre 20000 e 49999	0.273	0.000	0.118	0.031	0.088	0.108	0.114	0.047
População entre 50000 e 99999	0.235	0.000	0.146	0.005	0.125	0.016	0.127	0.018
População entre 100000 e 199999	0.071	0.231	0.004	0.947	-0.004	0.948	-0.007	0.903
Horas-aula diárias	0.164	0.000	0.042	0.297	0.064	0.112	0.109	0.063
Alunos por turma	-0.005	0.000	-0.003	0.007	-0.003	0.011	-0.003	0.012
Proporção professores com curso superior	0.735	0.000	0.448	0.000	0.409	0.000	0.137	0.129
Fator de infra-estrutura	0.261	0.000	0.166	0.000	0.158	0.000	0.102	0.000
Defasagem espacial do resíduo			0.469	0.000	0.357	0.000		
Defasagem espacial Razão professor/pop. 7 a14							-3.409	0.003
Defasagem espacial da educação média da população adulta							-0.135	0.000
Def. espacial da proporção de matrículas em escolas com ciclo							0.096	0.078
Def. espacial da proporção de matrículas públicas							1.319	0.004
Def. espacial proporção de matrículas municipais dentre as públicas							-0.187	0.027
Defasagem espacial do fator de infra-estrutura							0.383	0.000
Def. espacial Proporção de estabelecimentos com 50 a 200 matrículas							-0.062	0.548
Def. espacial Proporção de estabelecimentos com 201 a 1000							0.008	0.932
Def. espacial Proporção de estabelecimentos com mais de 1000 matrículas							-0.161	0.640
Defasagem espacial da média de horas-aula diárias							-0.022	0.776
Defasagem espacial de alunos por turma							-0.006	0.015
Def. espacial da proporção de professores com curso superior							0.406	0.001
Efeito Aleatório								
Coeficiente	0.144	0.000	0.119	0.000	0.117	0.000	0.122	0.000

GRÁFICO 7. 14: Variação na probabilidade de progressão na 5ª série decorrente do aumento de um desvio padrão em relação à média de variáveis selecionadas - diferentes modelos.

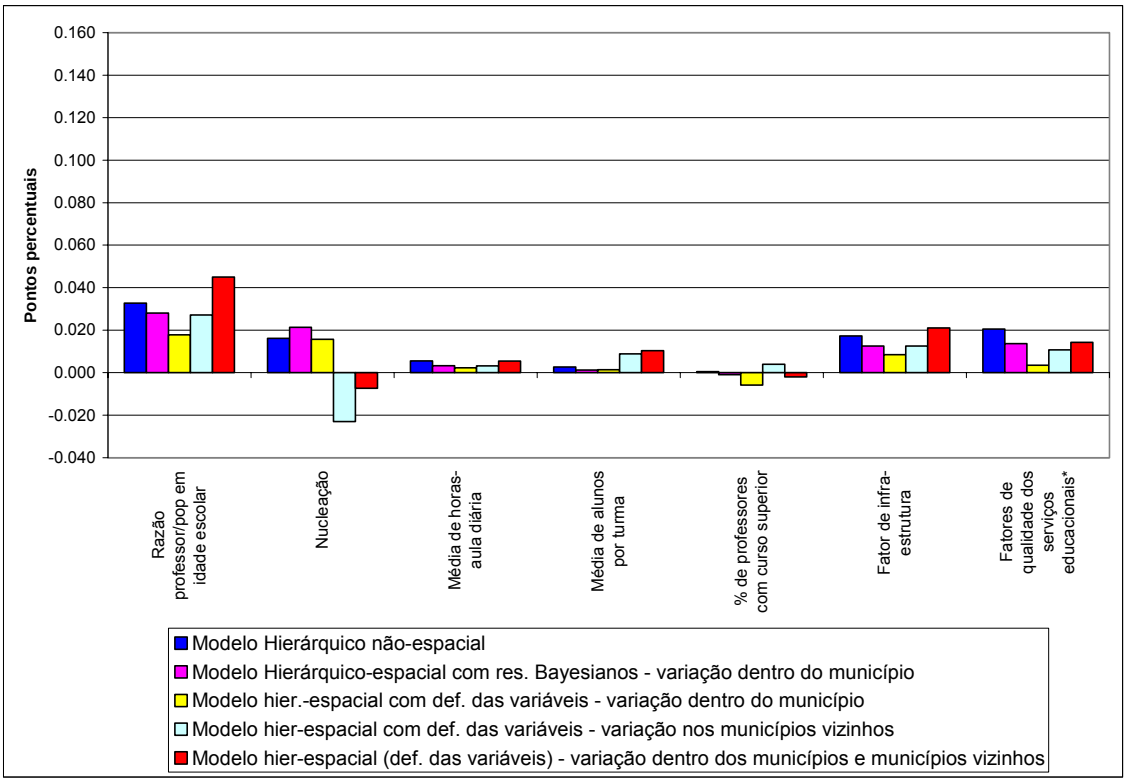


*Nota: Na simulação da melhoria nos fatores de qualidade dos serviços educacionais, para média de alunos por turma, considerou a diminuição de um desvio padrão em relação à média..

TABELA 7. 11: Resultado dos diferentes modelos para o intercepto da probabilidade de progressão na 1ª série do ensino médio (e_8).

Variáveis Independentes	Modelo Hierarquico (Modelo 4)		Modelo com a defasagem do resíduo bayesiano		Modelo com a defasagem do resíduo MQO		Modelo com a defasagem das variáveis independentes	
Efeito Fixo	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig
Intercepto	-0.599	0.002	-0.057	0.759	-0.087	0.641	-0.235	0.323
Razão professor/pop. 15 a17	6.223	0.000	5.309	0.000	5.204	0.000	3.346	0.000
Proporção de estabelecimentos com 50 a 200 matrículas	0.077	0.295	0.109	0.131	0.105	0.143	0.067	0.365
Proporção de estabelecimentos com 201 a 1000	0.122	0.094	0.148	0.037	0.144	0.043	0.108	0.138
Proporção de estabelecimentos com mais de 1000 matrículas	0.059	0.571	0.110	0.276	0.115	0.252	0.124	0.234
Proporção de matrículas públicas	-0.011	0.904	-0.119	0.192	-0.090	0.321	-0.077	0.426
Proporção de matrículas municipais dentre as públicas	-0.027	0.573	-0.037	0.442	-0.035	0.460	-0.009	0.858
Educação média da população adulta	0.190	0.000	0.148	0.000	0.150	0.000	0.181	0.000
População até 4999	0.163	0.010	0.064	0.300	0.026	0.669	-0.036	0.617
População entre 5000 e 9999	0.179	0.001	0.109	0.031	0.072	0.159	0.002	0.976
População entre 10000 e 19999	0.167	0.001	0.112	0.015	0.076	0.100	0.010	0.859
População entre 20000 e 49999	0.100	0.024	0.064	0.126	0.030	0.476	-0.026	0.589
População entre 50000 e 99999	0.035	0.422	0.028	0.488	-0.001	0.977	-0.039	0.381
População entre 100000 e 199999	0.042	0.366	0.017	0.699	0.013	0.771	-0.020	0.656
Fator de infra-estrutura	0.096	0.000	0.070	0.000	0.069	0.000	0.047	0.002
Horas-aula diárias	0.080	0.004	0.048	0.074	0.054	0.046	0.033	0.293
Alunos por turma	0.002	0.278	0.001	0.615	0.002	0.268	0.001	0.585
Proporção professores com curso superior	0.009	0.848	-0.018	0.689	-0.031	0.495	-0.110	0.045
Defasagem espacial do resíduo			0.650	0.000	0.269	0.000		
Defasagem espacial da educação média da população adulta							-0.063	0.010
Defasagem espacial População até 4999							0.320	0.007
Defasagem espacial População entre 5000 e 9999							0.148	0.168
Defasagem espacial População entre 10000 e 19999							0.128	0.207
Defasagem espacial População entre 20000 e 49999							0.069	0.461
Defasagem espacial População entre 50000 e 99999							-0.020	0.847
Defasagem espacial População entre 100000 e 199999							0.047	0.699
Defasagem espacial Razão professor/pop. 15 a 17							7.770	0.000
Def. espacial Proporção de estabelecimentos com 50 a 200 matrículas							-0.251	0.080
Def. espacial Proporção de estabelecimentos com 201 a 1000							-0.300	0.036
Def. espacial Proporção de estabelecimentos com mais de 1000 matrículas							-0.242	0.264
Def. espacial da proporção de matrículas públicas							-0.109	0.497
Defasagem espacial da média de horas-aula diárias							0.036	0.365
Defasagem espacial de alunos por turma							0.009	0.007
Def. espacial da proporção de professores com curso superior							0.095	0.239
Defasagem espacial do fator de infra-estrutura							0.093	0.000
Def. espacial proporção de matrículas municipais dentre as públicas							0.021	0.804
Efeito Aleatório								
Coeficiente	0.082	0.000	0.066	0.000	0.066	0.000	0.074	0.000

GRÁFICO 7. 15: Variação na probabilidade de progressão na 1ª série do ensino médio decorrente do aumento de um desvio padrão em relação à média de variáveis selecionadas - diferentes modelos.



*Nota: Na simulação da melhoria nos fatores de qualidade dos serviços educacionais, para média de alunos por turma, considerou a diminuição de um desvio padrão em relação à média..

7.4) Conclusão

Esse capítulo estimou o modelo hierárquico-espacial, que busca incluir os efeitos espaciais na análise hierárquica dos determinantes educacionais. Os resultados foram comparados aos do modelo hierárquico não-espacial estimado no Capítulo 5.

Com relação a um dos pontos enfocados nessa tese – importância do dividendo demográfico – os resultados para o acesso das crianças ao sistema de ensino, demonstram que quando se considera o efeito espacial, o impacto do dividendo demográfico cai significativamente, tornando a restrição de oferta o fator contextual mais importante para as duas faixas etárias, principalmente para os indivíduos de 15 a 17 anos.

A mudança do efeito do tamanho relativo da coorte nos modelos hierárquico-espaciais foi maior que para a restrição de oferta, ocorrendo inclusive mudança de sinal para o acesso das crianças de 15 a 17 anos na escola, independente da sua escolaridade.

A análise do impacto dos fatores de rede escolar dos municípios nos indicadores de eficiência e rendimento, de um modo geral demonstrou que quando se tentou corrigir a autocorrelação espacial dos dados, ocorreu uma diminuição do impacto direto de todas as variáveis contextuais nos indicadores, com menor intensidade para as PPSs. Por outro lado, percebe-se que as defasagens espaciais dessas variáveis possuem efeitos significativos

Esses resultados sugerem que as variáveis contextuais possuem dois efeitos. Um seria o efeito direto que diz respeito a mudanças nas probabilidades em virtude de variações desses fatores dentro do próprio município. O outro seria o efeito indireto decorrente de variações nas probabilidades quando ocorre mudança nos fatores contextuais dos municípios vizinhos. Os modelos hierárquico-espaciais conseguem captar esses dois efeitos. O primeiro efeito é captado pelos coeficientes das covariáveis do nível dois nas duas especificações do modelo hierárquico-espacial (com defasagem dos resíduos e com defasagem das covariáveis) e o segundo é captado pelos coeficientes das defasagens espaciais das variáveis do modelo hierárquico-espacial que as inclui. A soma dos dois efeitos é o efeito final de mudanças das variáveis contextuais.

Para tentar visualizar esses dois efeitos, foram construídos gráficos com simulações de mudanças nas probabilidades decorrente de variações dos fatores contextuais para os diferentes modelos. Através desse exercício foi possível perceber que a soma dos efeitos diretos e indiretos calculados a partir dos coeficientes do modelo hierárquico-espacial com a defasagem das variáveis de nível 2 é em muitos casos próxima do encontrado no modelo hierárquico não-espacial.

Assim, pode-se dizer que apesar do modelo hierárquico não-espacial possuir o problema de viés nas estimativas, já que não corrige o problema da autocorrelação espacial dos dados, seus resultados podem ser interpretados como a soma dos efeitos diretos e indiretos das variáveis contextuais, ou seja, podem estar captando o efeito local e o efeito dos municípios vizinhos.

Somado a isso, pode-se inferir que políticas públicas de melhoria de qualidade do perfil escolar dos municípios são mais eficazes quando levam em conta um espaço geográfico mais abrangente, ou seja, quando não se limitam a um determinado município, mas envolvem todos os municípios adjacentes.

Com relação ao *trade-off* entre a educação materna e os fatores de rede escolar dos municípios, observa-se uma redução significativa para o caso da probabilidade de freqüentar a escola na idade correta para os indivíduos de 7 a 14 anos e a total perda para essa probabilidade dos indivíduos de 15 a 17 anos. Por outro lado, da mesma forma que ocorreu na análise do intercepto, as variáveis defasadas espacialmente passam a apresentar esse efeito substituição, principalmente para o primeiro caso, o que reafirma a importância de políticas educacionais mais abrangentes geograficamente. Interessante notar, que para a probabilidade de passar na 1ª série (e_0) o efeito substituição existente entre a educação materna e alguns fatores de rede escolar que ocorria no modelo hierárquico contínua significativo.

Outro resultado interessante encontrado nesse capítulo foi o sinal negativo para a defasagem espacial da variável educação média da população adulta em praticamente todos os casos (exceto para a probabilidade de freqüentar a escola na idade correta para os indivíduos de 15 a 17 anos). Essa relação negativa entre a educação média da população adulta dos municípios vizinhos e os indicadores educacionais pode ser devida ao fato dessa variável estar associada ao nível de desenvolvimento socioeconômico dos municípios. Assim, pode-se dizer que ser vizinho de municípios mais desenvolvidos constitui-se em um fator negativo para os indicadores educacionais de determinado município.

Conclui-se também, que a restrição de oferta é um fator importante para as políticas voltadas para a população alvo do ensino médio, seja para o acesso escolar, seja para a freqüência na idade correta ou para a progressão desse nível de ensino.

CONCLUSÃO

Para diagnosticar os principais problemas da educação no Brasil, foi feita, primeiramente, uma análise descritiva da trajetória dos principais indicadores educacionais nas últimas décadas. Tais indicadores foram calculados a partir das pesquisas domiciliares do IBGE, ou seja, dos Censos Demográficos e Pesquisa Nacional de Amostra por Domicílio – PNAD. Os indicadores analisados foram: taxa de analfabetismo, taxa de atendimento, taxa de escolarização bruta e líquida, taxa de distorção idade/série, Probabilidade de Progressão por Série (PPS) e anos médios de estudo.

Através dessa análise foi possível perceber que ocorreu uma melhora significativa dos indicadores nas três últimas décadas, com significativa redução da taxa de analfabetismo, aumento da cobertura e do fluxo dos estudantes entre as séries, o que resultou em um aumento dos anos médios de estudo entre esse período.

Os resultados revelaram que o acesso ao ensino fundamental já está praticamente universalizado, com 94,50% das crianças de 7 a 14 anos freqüentando a escola e 89,53% deles freqüentando o ensino fundamental. Devido a essa quase universalização do ensino fundamental, observa-se que as desigualdades entre os estados e entre diferentes grupos da população, como brancos e negros, diminuíram significativamente. Por outro lado, a participação dos jovens de 15 a 17 anos no sistema de ensino, apesar de ter melhorado nas últimas décadas, ainda está aquém do desejável. A taxa de atendimento nesse grupo etário é de 77,71%, sendo que apenas 34,39% desses jovens freqüentam o nível de ensino adequado a sua faixa etária, ou seja, o nível médio. Somado a isso, as desigualdades regionais e entre grupos específicos da população apresentaram tendência de ampliação.

Outro fato marcante do sistema de ensino brasileiro revelado por essa análise descritiva foi a alta distorção entre a idade e a série adequada cursada. No ensino fundamental havia, em 2000, 54,03% de matrículas de pessoas que estavam fora da série adequada a sua idade, no nível médio esse percentual era de 66,36%.

Pode-se dizer, portanto, que apesar da melhora dos anos médios de estudo da população brasileira, seus níveis ainda estão abaixo do desejável, já que, em 2000, a população acima de 15 anos de idade não possuía em média nem o ensino fundamental

completo. Tal escolaridade é conseqüência principalmente do baixo acesso dos jovens de 15 a 17 anos na escola e da elevada distorção idade/série, causada principalmente pela alta repetência.

Através da análise descritiva percebeu-se também que os grupos mais desprivilegiados são os com piores indicadores educacionais. Tal fato demonstra a forte estratificação educacional existente no Brasil, fato comprovado por vários estudos (FERNADES, 2001; SILVA e HASENBALG, 2001). Em outras palavras, pode-se dizer, que o resultado educacional de uma criança está fortemente associada com o seu ambiente familiar, principalmente com o nível educacional dos pais. Segundo CAMARGO E BARROS (1991), pais com menor nível educacional podem não perceber mais claramente a importância da educação e o seu efeito benéfico nos rendimentos futuros, ou ainda, por serem mais pobres, são mais desencorajados a investir na educação dos seus filhos, devido ao elevado custo de oportunidade da retirada dos seus filhos do mercado de trabalho.

A associação positiva entre os fatores de *background* familiar e o resultado educacional da criança alimenta o ciclo vicioso da pobreza. Trabalhadores com menor nível educacional possuem taxas de rendimento menores. Dessa forma, por serem pobres e devido a limitação de crédito da economia brasileira, investem menos na educação de seus filhos, perpetuando dessa forma a desigualdade social.

Surge, portanto, algumas questões importantes. Como políticas públicas podem tornar a educação menos estratificada, ou seja, diminuir a dependência do resultado educacional da criança com o seu *background* familiar? Seria a estrutura escolar dos municípios um dos fatores a ser utilizado para minimizar tal dependência?

Somado a isso, questões de natureza estrutural, que fogem do domínio de políticas públicas de curto prazo, podem trazer benefícios para a melhora dos indicadores educacionais nos grupos menos privilegiados da população. Dentre tais questões estaria aquela relacionada ao aumento geral da escolaridade materna. A melhora dos níveis de escolaridade, mesmo que lenta, ocorreu também nas camadas mais pobres, o que por si só gera benefícios para as gerações mais novas.

Outro fator está relacionado com a mudança da composição etária da população brasileira. O Brasil vem apresentando constantes quedas nas taxas de fecundidade, inclusive nas camadas mais pobres. Essa queda resulta em uma menor razão de

dependência de jovens o que pode se constituir em uma "Janela de Oportunidades" para políticas que têm como população alvo esses grupos populacionais, conforme coloca CRAVALHO & WONG (1995).

Dessa forma, essa tese buscar responder algumas dessas questões tendo como um dos objetivos centrais a investigação dos determinantes do resultado educacional no Brasil tendo como "pano de fundo" a Função de Produção Educacional - FPE – associada à literatura demográfica econômica do dividendo demográfico.

Com base na análise descritiva dos indicadores educacionais, optou-se em considerar como resultado educacional, duas dimensões diferentes da educação, quais sejam: acesso ao sistema de ensino e eficiência e rendimento na escola. A primeira dimensão foi captada através da probabilidade de freqüentar a escola para os indivíduos de 7 a 14 anos e de 15 a 17 anos. Essas são as faixas etárias adequadas de cursar respectivamente o ensino fundamental e médio.

A segunda dimensão foi captada pela probabilidade do aluno estar matriculado na série adequada a sua idade, considerou-se as duas faixas etárias anteriormente mencionadas e a Probabilidade de Progressão por Série (PPS). Devido ao elevado número de progressões no sistema de ensino, preferiu-se analisar apenas as transições mais importantes para cada nível de ensino. Conforme evidenciado pela a análise descritiva, os principais gargalos das duas primeiras etapas do sistema de ensino brasileiro são as probabilidades de passar na 1º série, na 5º série e na 1º série do ensino médio.

Para analisar os determinantes dessas duas dimensões utilizou-se a FPE como ponto de partida. Ela foi importante por permitir organizar de forma mais clara a escolha das variáveis independentes relevantes ao processo educacional.

Foram consideradas duas formas de especificação da FPE. Num primeiro momento, a FPE foi estimada aplicando o modelo hierárquico de dois níveis, onde o primeiro nível considera o aluno e o segundo, o município. Num segundo momento, foi utilizado o modelo hierárquico-espacial. Esse modelo agrega a modelagem espacial com a hierárquica, permitindo trabalhar com a possível correlação espacial das variáveis dependentes e independentes e/ou dos erros no segundo nível do modelo hierárquico, o que possibilita obter melhores estimativas dos parâmetros das regressões.

Os insumos educacionais considerados na análise foram agrupados da seguinte forma. No primeiro nível foram considerados os insumos educacionais relacionados aos fatores de *background* familiar, quais sejam: educação materna, categoria de ocupação do chefe, chefia feminina e a presença de famílias conviventes. Nesse nível também foram incluídas variáveis relacionados aos atributos individuais, como idade, sexo, cor e localização de residência.

No segundo nível do modelo hierárquico, encontram-se as variáveis relacionadas ao perfil da rede escolar do município e ao dividendo demográfico.

Para o perfil escolar dos municípios foram consideradas variáveis relacionadas com a qualidade e disponibilidade da oferta escolar. Dessa forma, foram considerados variáveis referentes à: qualidade dos recursos humanos (média de alunos por turma, média de horas-aula diárias e porcentagem de professores com curso superior); qualidade da infra-estrutura escolar; restrição de oferta (razão professor/população em idade escolar); nucleação (distribuição das matrículas segundo tamanho dos estabelecimentos); distribuição das matrículas segundo dependência administrativa e organização do ensino em ciclos. O dividendo demográfico foi captado pelo tamanho relativo da coorte em idade escolar.

A base de dados utilizada para esse estudo foi o Censo Demográfico de 2000 e o Censo Escolar do mesmo ano. O resultado da primeira especificação da FPE encontra-se no Capítulo 5 e foi pautada em dois pontos específicos.

1) Importância do dividendo demográfico.

Os resultados demonstraram que uma menor pressão demográfica, ou seja, uma menor proporção de crianças e jovens, provoca importante impacto em todas as dimensões analisadas. No caso do acesso escolar das crianças de 7 a 14 anos o dividendo demográfico se constitui no fator contextual mais importante. Menores coortes nessa faixa etária provocam um aumento na probabilidade de freqüentar a escola. Esse resultado vai de encontro às preposições do estudo de COALE & 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. Devido à contradição desse resultado, buscou-se verificar qual fator que torna o sinal do coeficiente do tamanho relativo da coorte positivo, já que quando se estima a regressão apenas com essa variável, o sinal é positivo. Com essa verificação pode perceber que o

sinal do tamanho relativo da coorte passa a ser positivo quando se incorpora a razão professor/população em idade escolar na análise, ou seja, a restrição de oferta. Surge, portanto, uma possível explicação para essa contradição. Como a restrição de oferta é o fator contextual mais importante para a freqüência escolar dos indivíduos de 15 a 17 anos, e dado o baixo atendimento dessa faixa etária, um aumento da população jovem pode acarretar uma maior demanda por educação, quando se controla pela oferta escolar.

Para a segunda dimensão da educação – rendimento e eficiência – o bônus demográfico também possui importante impacto favorável para os dois níveis de ensino.

Esses resultados indicam que mudança na estrutura etária brasileira com diminuição da razão de dependência de crianças e jovens, decorrente da diminuição da Taxa de Fecundidade Total que vem ocorrendo no Brasil nas últimas décadas, pode se constituir em uma "Janela de Oportunidades" no sentido de melhorar a distorção entre a idade e a série adequada cursada, principalmente do ensino médio, e os fluxos entre as séries, tal como já argumentado por CARVALHO & WONG (1995).

2) Importância dos fatores do perfil escolar dos municípios.

A relação positiva entre os fatores de *background* familiar e o resultado educacional, como já foi mencionado, é bastante conhecida e estudada na literatura brasileira. Os resultados dessa tese confirmaram essa associação positiva, comprovando também a maior importância do nível educacional materno para os indicadores de acesso e eficiência dos filhos.

Dada a desigualdade de oportunidades existente no sistema educacional brasileiro, buscou-se determinar quais fatores relacionados à rede escolar dos municípios podem diminuir essa estratificação social. Tal questão é importante para a eficácia de políticas públicas educacionais, no sentido de torná-las mais igualitárias.

Para os indicadores de acesso, a análise dos fatores de rede escolar dos municípios levou em conta basicamente apenas o impacto da restrição de oferta, que se constituiu o principal fator para o maior acesso escolar dos jovens de 15 a 17 anos. Dessa forma, ficou evidenciada a necessidade de políticas públicas que aumentem a disponibilidade de oferta do nível de ensino médio nos municípios brasileiros.

Com relação aos indicadores de eficiência e rendimento, a análise englobou todos os fatores de rede escolar estudados nessa tese e foram feitas duas análises. A primeira, refere-se ao efeito das variáveis contextuais na probabilidade média de freqüentar a escola na idade correta, ou seja, ao impacto no intercepto. A segunda análise diz respeito ao *trade-off* entre educação materna e as variáveis de nível 2, o que possibilita averiguar até que ponto o efeito da escolaridade da mãe pode ser minimizado pelos fatores municipais.

Na primeira análise pode-se observar que para a probabilidade dos indivíduos de 7 a 14 anos freqüentarem a escola na idade correta os fatores relacionados à qualidade dos serviços educacionais, principalmente a proporção de professores com curso superior e o fator de infra-estrutura, são os que possuem maior impacto na probabilidade média. Já para a probabilidade dos indivíduos de 15 a 17 anos freqüentarem a escola na idade correta percebe-se o maior impacto do fator de infra-estrutura e da restrição de oferta.

Com relação as PPS, verificou-se o maior impacto da porcentagem de professores com curso superior, média de alunos por turma e do fator de infra-estrutura nas probabilidade de passar de zero para um ano de estudo (e_0), quatro para cinco anos de estudo (e_4) e oito para nove anos de estudo (e_8). A razão professor/população em idade escolar também se mostrou importante para essa última progressão.

Dessa forma, pode-se dizer que, de modo geral, uma melhoria nos fatores de qualidade da rede escolar do município provoca uma melhora nos indicadores de eficiência e rendimento em ambos níveis de ensino. Destaca-se, também, que um aumento na oferta do ensino médio possui importante efeito nos indicadores desse nível de ensino. Tal fato evidencia a importância de políticas públicas que estimulem a melhoria do perfil escolar dos municípios, já que demonstra que de certa forma o contexto pode minimizar o papel da família sobre a eficiência do aluno.

Na segunda análise, referente ao *trade off* entre a educação materna e os fatores relacionados à rede escolar dos municípios, os resultados evidenciam que os fatores de qualidade escolar possuem um efeito substituição com a educação materna, principalmente para a probabilidade dos indivíduos de 7 a 14 anos freqüentarem a escola e para a probabilidade de progressão na 1ª série. Tal fato significa que filhos de mães com baixa escolaridade são mais beneficiados pela melhora da qualidade da rede

de ensino do município do que filhos de mães mais escolarizadas. Pode-se inferir, portanto, que a melhoria dos fatores referentes ao perfil escolar do município tem um importante papel na diminuição das desigualdades educacionais entre os alunos de origem socioeconômica diferentes.

Em síntese, os resultados da análise hierárquica corroboram com a hipótese de que a melhora da rede escolar dos municípios se constitui um importante fator de diminuição da estratificação educacional, primeiro pelo seu efeito direto no aumento médio das probabilidades estudadas, segundo por diminuir a importância dos fatores relacionados ao ambiente familiar do aluno (no caso, escolaridade da mãe), diminuindo a desigualdade intergeracional.

A segunda especificação da FPE levou em consideração a autocorrelação e heterogeneidade espacial das variáveis analisadas, através da construção de um modelo hierárquico-espacial. Esse modelo buscou incorporar o processo espacial no modelo hierárquico, o que resultou em duas importantes contribuições para a análise dos determinantes educacionais.

A primeira contribuição diz respeito à melhora na estimação dos parâmetros das regressões, já que incorpora a dependência espacial dos dados. Quando os dados apresentam autocorrelação e heterogeneidade espacial, e o processo espacial não é incorporado na análise, as estimativas dos parâmetros nas regressões podem ser enviesados e/ou ineficientes. Somado a isso, a presença de dependência espacial nas variáveis do modelo hierárquico incorre na quebra de um pressuposto importante desse modelo que é o de erros i.i.d. no nível mais alto (nível 2).

Através da análise espacial feita no Capítulo 6, comprovou-se a existência de forte correlação espacial em praticamente todas as variáveis analisadas, ressaltando a importância da construção do modelo hierárquico-espacial. A dependência espacial ocorre quando as variáveis observadas em determinado lugar dependem dos valores observados nos locais vizinhos. Com a análise do Capítulo 6 também foi possível detectar que o processo espacial ocorre principalmente na forma de defasagem espacial. Esse modelo considera que o efeito espacial opera através das covariáveis e do termo de erro, ou seja, dos fatores observados e não observados.

O efeito espacial foi incorporado no modelo hierárquico de duas formas diferentes. Em uma delas incluiu-se as defasagens espaciais dos resíduos do nível 2 do

modelo hierárquico analisado no Capítulo 5. Na outra forma, incluiu-se nos modelos de regressões as defasagens espaciais das variáveis independentes utilizadas na análise. No primeiro caso, se está corrigindo a dependência espacial tanto das variáveis observadas quanto das não-observadas. Já no segundo caso, é corrigido apenas o efeito espacial das variáveis observadas.

A primeira forma do modelo hierárquico-espacial possibilita determinar o efeito nas variáveis dependentes decorrente de mudanças nos fatores contextuais dentro do próprio município, sem o problema relacionado ao viés das estimativas. Esse efeito foi chamado de direto. Por outro lado, a segunda especificação possibilita avaliar o impacto nos indicadores decorrente de mudanças nas variáveis contextuais no próprio município e nos seus vizinhos, esse último impacto foi chamado de efeito indireto. Essa seria, portanto, a segunda contribuição do modelo hierárquico-espacial, ou seja, a possibilidade de separar o efeito direto – relativo a variações nos determinantes dentro do próprio município – do efeito indireto – relativo a mudanças nos fatores contextuais nos municípios vizinhos.

Dessa forma, foi possível avaliar a importância do dividendo demográfico e dos fatores de rede escolar dos municípios levando em consideração a autocorrelação e heterogeneidade espacial, o que possibilitou obter melhores estimativas para os parâmetros. Somado a isso, essa metodologia ajuda a entender melhor como o processo educacional se dá no espaço, já que consegue captar o efeito de mudanças dos determinantes educacionais do próprio município e o efeito via seus vizinhos.

Analisando primeiramente a importância do dividendo demográfico, observa-se que quando se separa os efeitos diretos e indiretos, o primeiro cai significativamente, tornando a restrição de oferta o fator contextual mais importante para o acesso à escola. Destaca-se também, que para a freqüência escolar dos indivíduos 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, que é o sinal 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.

Com relação 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 uma importante mudança nas probabilidades das crianças e jovens freqüentarem a escola.

Partindo para o outro ponto enfocado nessa tese – importância dos fatores de rede escolar dos municípios – observa-se que a inclusão do espaço na análise também diminui seus impactos diretos nos indicadores de eficiência e rendimento, com menor proporção para a probabilidade de progressão na 1ª série do fundamental. Por outro lado, percebe-se que as defasagens espaciais dessas variáveis possuem efeitos significativos, ou seja, mudanças nos fatores contextuais nos municípios vizinhos possuem importantes impactos nos indicadores de determinando município.

Pode-se dizer, portanto, que a soma dos efeitos diretos e indiretos resulta no impacto final nos indicadores educacionais decorrente de mudanças nas variáveis contextuais, já que esses são altamente correlacionados, de tal forma que podem variar conjuntamente.

Esse resultado pode suscitar duas importantes questões. A primeira questão refere-se a maior eficácia das políticas públicas educacionais. Pode-se dizer que políticas de melhoria do perfil escolar dos municípios têm maior impacto quando não são focalizadas apenas no próprio município, mas quando abrangem uma dimensão espacial mais ampla, ou seja, os municípios adjacentes.

A segunda questão diz respeito à interpretação dos resultados do modelo hierárquico não espacial. Com base nos resultados, pode-se inferir que apesar das estimativas do modelo hierárquico serem sujeitas a problema de viés e ineficiência, devido a autocorrelação espacial das variáveis, sua análise pode ser relevante se considera-se que esse modelo está captando o efeito direto e indireto das variáveis contextuais, enquanto que o modelo hierárquico-espacial separa esses dois efeitos.

Todavia, cabe ressaltar que a interpretação tem que levar em conta que o impacto dos fatores contextuais nos indicadores educacionais está associada ao contexto espacial mais amplo no qual o município está inserido.

Essa tese buscou averiguar os determinantes do acesso e eficiência do sistema escolar do ensino fundamental e médio no Brasil, focalizando, principalmente, no impacto da rede escolar dos municípios e no dividendo demográfico. Porém, a busca do entendimento dos determinantes e o auxílio que tal entendimento pode fornecer para

melhoria de políticas públicas nesta área está longe de se esgotar. Algumas lacunas e possíveis desdobramentos restaram dessa tese.

Um possível desdobramento seria um exame, ou estudo de caso, dos *outliers* de municípios encontrados no Capítulo 6, cujos valores das variáveis analisadas se destacaram positivamente ou negativamente em relação ao valor médio dos seus vizinhos. Esse exame pode possibilitar uma maior compreensão espacial dos fatores de diferenciação de políticas educacionais.

Outro possível desdobramento, ou melhor, preenchimento de uma lacuna, seria expandir esse estudo para o ensino superior. Os estudos que enfocam esse nível de ensino são mais escassos e sua análise seria bastante útil para a compreensão da baixa escolaridade brasileira, já que de acordo com RIOS-NETO (2004), a probabilidade de progressão na 1ª série do superior em 2000 foi de 31,7% para coorte de 25 a 29 anos, se constituindo no principal gargalo do sistema de ensino brasileiro. Somado a isso, a análise evolutiva dessa probabilidade mostra que ela vem diminuindo ao longo dos anos.

Por fim, a busca da melhor especificação da FPE não se esgota. Um modelo ideal seria aquele que considera três níveis de análise. Sendo o primeiro nível o indivíduo, o segundo a escola e o último nível o município. A dificuldade desse exame consiste na obtenção de dados que possa obter esses três níveis de análise. Nessa linha de aperfeiçoamento dos modelos, é necessário também tentar incluir o modelo de erro na análise, já que os modelos de defasagens espaciais realizados no Capítulo 6, mostraram que o processo espacial também ocorre na forma de erro espacial, mesmo que em menor escala.

BIBLIOGRAFIA

ALBERNAZ, A., FERREIRA, F. H. G., FRANCO, C. Qualidade e equidade no ensino fundamental brasileiro. **Pesquisa e Planejamento Econômico**, Rio de Janeiro, v.32, n.3, p.453-476, dez. 2002.

ALLISON, P. D. **Logistic regression using the SAS system**: theory and application. Cary, NC: SAS Institute, 1999. 288p.

ANGRIST, J., LAVY, V. Using maimonides' rule to estimate the effect of class size on scholastic achievement. **Quarterly Journal of Economics**, v.114, n.2, p.533-75, May 1999.

ANSELEN, L. Under the hood: issues in the specification and interpretation of spatial regression models. **Agricultural Economics**, v.27, n.3, p.247-267, 2002.

ANSELIN, L. *et al*. Simple diagnostic test for spatial dependence. **Regional Science and Urban Economics**, v.26, n.1, p.77-104, Feb. 1996.

ANSELIN, L. Lagrange multiplier test diagnostics for spatial dependence and spatial heterogeneity. **Geographical Analysis**, v.20, p.1-17, 1988.

ANSELIN, L. Local indicators of spatial association – LISA. **Geographical Analysis**, v.27, n.2, p.93-114, Apr. 1995.

ANSELIN, L. **SpaceStat tutorial**: a workbook for using SpaceStat in the analysis of spatial data. Santa Barbara: NCGIA, 1992.

ANSELIN, L. **Spatial econometrics**. Dallas: University of Texas, School of Social Science, 1999. 284p.

ANSELIN, L. Spatial externalities, spatial multipliers, and spatial econometrics. **International Regional Science Review**, v.26, n.2, p.153-166, 2003.

ANSELIN, L., BERA, A. Spatial dependence in linear regression models with an introduction to spatial econometrics. In: ULLAH, A., GILES, D. E. (Eds.) **Handbook of applied economic statistics**. New York: Marcel Dekker, 1998. p.237-289.

ARMOR, D. J. School and family effects on black and white achievment: a reexamination of the USOE data. In: MOYNIHAN, D. P., MOYNIHAN, D. P. (Eds.). **On equality of educational opportunity**: papers deriving from the Harvard University Faculty Seminar on the Coleman Report. New York: Vintage Books, 1972. p.168-229.

BARBOSA, M. E. F, FERNANDES, C. A escola brasileira faz diferença? uma investigação dos efeitos da escola na proficiência em matemática dos alunos da 4ª série.

In: FRANCO. C. (Org.) **Avaliação, ciclos e promoção na educação**. Porto Alegre: ArtMed, 2001. p.155-172.

BARBOSA, M. L. O. **Efeitos da qualidade da escola sobre o desempenho dos alunos**. Washington, DC.: Latin American Studies Association, 2001. (Mimeogr.) Disponível em: <http://136.142.158.105/Lasa2001/DeOliveiraBarbosaMaria.pdf>.

BARROS, R. P., LAM, D. Income and educational inequality and children's schooling attainment. In: BIRDSALL, N., SABOT, R. H (Eds.) **Opportunity forgone**: education in Brazil. Washington DC.: Inter-American development Bank. 1996. p.337-367

BARROS, R. P., MENDONÇA, R. S. P., SHOPE, J. A. **Regional disparities in education within Brazil**: the role of quality of education. Brasilia: IPEA, 1993. 9p. (Texto para discussão, 311)

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.

BECKER, G. S. **Human capital**: a theoretical and empirical analysis with special reference to education. 3.ed. Chicago: The University of Chicago, 1993. 390p.

BEHRMAN, J. R. Is child schooling a poor proxy for child quality?. **Demography**, v.24, n.3, p.341-359, Aug. 1987.

BEHRMAN, J. R., BIRDSALL, N. The quality of schooling: quality alone is misleading. **American Economic Review**, v.73, n.5. p.928-946, .Dec. 1983.

BEHRMAN, J. R., BIRDSALL, N. The quality of schooling: reply. **The American Economic Review**, v.75, n.5, p.1202-1205, Dec. 1985.

BIDWELL, E., KASARDA, D. Conceptualizing and measuring the effects of school and schooling. **Journal of Education**, v.88, n.4, p.401-430, 1980.

BIRDSALL, N., BRUNS, B. SABOT, R. H. Education in Brazil: planning a bad hand badly. In: BIRDSALL, N., SABOT, R. H (Eds.) **Opportunity forgone**: education in Brazil. Washington, DC.: Inter-American development Bank.1996. p.7-47.

BLACK. S. E. Do better school matter? parental valuation of elementary education. **The Quarterly Journal of economics**, v.114, n.2, p.577-599, May 1999.

BORGONI, R., BILLARI, F. C. Bayesian spatial analysis of demographic survey data: an application to contraceptive use at first sexual intercourse. **Demographic Research**, v.8, n.3, 2003. Disponível em: <www.demographic-research.org>

BRYK, A. S., RAUDENBUSH, S. W. **Hierarchical linear models**: applications and data analysis methods. 2.ed. Newbury Park, California: Sage, 2002. 485p.

BUCHMANN C., HANNUM, E. Education and stratification in developing countries: a review of theories and research. **Annual review of sociology**, v.27, p.77-102, Jan. 2001.

CAMARGO, J., BARROS, R. As causas da pobreza no Brasil: Porca Miséria!. In: IPEA. **Perspectivas da economia brasileira**: 2002. Brasília, D.F.: IPEA, 1991. Cap.27. p.525-544.

CARD, D., ALAN, B. School resources and student outcomes: an overview of the literature and new evidence from North and South Carolina. **Journal of Economic Perspectives**, v.10, n.4, p.31-50, 1996.

CARD, D., KRUEGER, A. B. School resources and student outcomes: an overview of the literature and new evidence from North and South Carolina. **Journal of Economic perspectives**, v.10, n.4, p.31-50, 1996.

CARVALHO, J. A. M., WONG, L. A. (1995). **Window of opportunity**: some demographic and socioeconomic implication of rapid fertility decline in Brazil. Belo Horizonte: UFMG/CEDEPLAR, 1995. 36p. (Texto para discussão, 91)

CASTRO, M. H. G., DAVANZO, A. M. Q. **Situação da educação básica no Brasil.** Brasília: INEP. 1999. 134 p.

CÉSAR, C. C., SOARES, J. F. Desigualdades acadêmicas induzidas pelo contexto escolar. **Revista Brasileira de Estudos Populacionais,** v.18, n.1/2, jan./dez, p.97-110, 2001.

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. 389p.

COLEMAN, J. S. *et al*. **Equality of educational opportunity**. Washington, DC: US Department of Health, Education & Welfare, 1966. 548p.

COLEMAN, J. S. Social capital in the creation of human capital. In: LESSER, E. L. (Ed) **Knowledge and social capital**: foundation and application. Boston: Butterworth Hewinemann, 2000. Cap.2. p.17-41.

EATON, P. J. The quality of schooling: comment. **The American Economic Review**, v.75, n.5, p.1195-1201, Dec. 1985.

FERNADES, D. C. **Raça, origem socioeconômica e desigualdade educacional no Brasil, uma análise longitudinal**. 2001. Trabalho apresentado na Reunião Anual da Associação Nacional de Pós-graduação e pesquisa em Ciências Sociais (ANPOCS). Outubro. 2001.

FERRÃO, M. E. *et al*. O SAEB – Sistema Nacional de Avaliação da Educação Básica: objetivos, características e contribuições na investigação da escola eficaz. **Revista Brasileira de Estudos Populacionais,** v.18, n.1/2, p.111-130, jan./dez, 2001.

FERREIRA, F. H. G., BARROS, R. P. Education and income distribution in urban Brazil, 1976-1996. **Cepal Review**, n.71, p.41-61, Aug. 2000. Disponível em: <http://www.iets.inf.br/acervo/Artigos/Education%20and%20income%20distribution%20in%20urban%20Brazil,%201976-1996.pdf>

FLETCHER, P. R. **À procura do ensino eficaz**. Rio de Janeiro, RJ: MEC, 1998.

FULLER, B., CLARK, P. Raising school effects while ignoring culture? local conditions and the influence of classroom tools, rule and pedagogy. **Review of educational research**, v.64, n.1, p.119-157, 1994.

GOLDSTEIN, H. **Multilevel statistical models**. 2.ed. Londres: Edward Arnold, 1995.

GOLDSTEIN, H. Nonlinear multilevel models, with an application to discrete response data. **Biometrika**, v.78, n.1, p.45-51, 1991.

GOMES-NETO, J. B., HANUSHEK, E. A. The causes and effects of grade repetition. In: BIRDSALL, N., SABOT, R. H (Eds.) **Opportunity forgone**: education in Brazil. Washington, DC.: Inter-American development Bank.1996. p.425-460.

GREENE, W. **Econometric analysis**. 2.ed. Englewood Cliffs, N.J.: Prentice Hall, 1993. 791p.

GRUSKY, D. B., DIPRETE, T. A. Recent trends in the process of stratification. **Demography,** v.27, n.4, p.617-637, Nov. 1990

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. Evidence, politics, and the class size debate. In: MISHEL, L., ROTHSTEIN, R. (Eds). **The class size debate**. Washington, DC: Economic policy Institute, 2002. p.37-63.

HANUSHEK, E. A. Measuring investment in education. **Journal of Economics Perspectives**, v.10, n.4, p.9-30, 1996.

HANUSHEK, E. A. Publicly provided education. In: AUERBACH, A. J., FELDSTEIN, M. (Eds.) **Handbook of public economics**. Amsterdam: Elsevier, 2002. v.4, p. 2045-2141.

HANUSHEK, E. A. **The evidence on class size**. Rochester, N.Y.: University of Rochester, W. Allen Wallis Institute of political Economy, 1998. (Occasional Paper, 98-1)

HEALY, M. J. R. Multilevel data and their analysis. In: LEYLAND, A. H., GOLDSTEIN, H. (Eds.). **Multilevel modeling of health statistics**. New York: Wiley, 2001. p.1-12.

HEYNEMAN, S. P., LOXLEY, W. A. The effect of primary school quality on academic achievement across twenty-nine high- and low – income countries. **American Journal Sociology**, v.88, n.6, p.1162-1194, May 1983.

HOX, J. J. **Applied multilevel analysis**. Amsterdam: TT-Publicities, 1995. Caps. 2, 4.

HOX, J. J. Multilevel modeling: when and why. In: BALDERJAHN, I., MATHAR R., SCHADER, M. (Eds). **Classification, data analysis, and data highways**. New York: Springer Verlag, 1998. p.147-154.

IBGE. **Censo demográfico de 1991**: documentação dos microdados da amostra. Rio de Janeiro: IBGE, 1996.

IBGE. **Censo Demográfico de 2000**: documentação dos microdados da amostra. 2.ed. Rio de Janeiro: IBGE, 2003. Disponível em CD-ROM.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS – INEP. **Geografia da educação brasileira**. Brasília: INEP.2000. 144p.

JONES, G. J. **Population dynamics and education and health planning**. Geneva: International Labour Office, 1990. 88p. (Wold Employment Programme. Paper, 8)

KLEIN, R., RIBEIRO, S. C. O censo educacional e o modelo de fluxo: o problema da repetência. **Revista Brasileira de Estatística**, v.52, n.197, p.5-45, 1991.

KLEIN, R., RIBEIRO, S. C. **Produção e utilização de indicadores educacionais**. Rio de Janeiro: LNCC/CNPQ, 1995. p.1-35.

KRUEGER, A. B. **An economist's view of class size research.** [Maryland]**:** Milkmen Institute, 2000. (Research paper)

KRUEGER, A. B. Economic consideration and class size. **The Economic Journal**, v.113, n.485, p.34-63, Feb. 2003.

KRUEGER, A. B. Experimental estimates of education production functions. **The Quarterly Journal of Economics,** v.114, n.2, p.497-532, May. 1999

KRUEGER, A. B. Reassessing the view that American schools are broken. **Economic Research bank of New York**, v.4, n.1, p.29-46, Mar. 1998.

LAM, D., MARTELETO, L. **A dinâmica da escolaridade das crianças brasileiras durante a transição demográfica**: aumento no tamanho da coorte versus diminuição no tamanho da família. Belo Horizonte: CEDEPLAR/UFMG, 2004. (Texto para discussão, 243)

LANGFORD, I. H. *et al.* Multilevel, modeling of the geographical distribution of diseases. **Journal of the Royal Statistical Society,** Series C (Applied Statistics), v.48, n.2, p.253-268, 1999.

LAZEAR, E. P. **Educational production**. Cambridge: National Bureau of Economic Research, 1999. 48p. (Working paper, 7349)

LEE, J., BARRO. R. J. **Schooling quality in a cross section of countries**. Cambridge: National Bureau of Economic Research, 1997. 33, [9] p. (Working paper, 6198)

LEE, V. E., BRIK, A. S. A multilevel model of the social distribution of high school achievement. **Sociology of Education**, v.62, n.3, p.172-192, July 1979.

MACEDO, G. A. Fatores associados ao rendimento escolar de alunos da 5ª série (2000) – uma abordagem do valor adicionado. In: ENCONTRO NACIONAL DE ESTUDOS POPULACIONAIS, 14, 2004, Caxambu. **Anais**. Campinas: ABEP, 2004. Disponível em CD-ROM.

MARE, R. D., CHEN, M. D. Further evidence on sibship size and educational stratification. **American Sociological Review**, v.51, n.3, p.403-412, June 1986.

MARE, ROBERT D. Change and stability in educational stratification. **American Sociological Review**, v.46, n.1, p.72-87, 1981.

MARE, ROBERT D. Social background composition and educational growth. **Demography**, v.16, n.1, p.55-71, Feb. 1979.

MARSHALL, G., SWIFT, A. ROBERTS, S. **Against the odds?** social class and social justice in industrial societies. Oxford: Clarendon, 1997. 278p.

MARTELETO, L. Children's schooling in Brazil: do number and composition of siblings matter? In: INTERNATIONAL UNION FOR THE SCIENTIFIC STUDY OF POPULATION GENERAL CONFERENCE, 24, 2001, Salvador, BA. **XXIV IUSSP general conference, Salvador** - Brazil / 18-24 August 2001. Paris: IUSSP, 2001. Disponível em CD-ROM.

MARTELETO, L. Desigualdade regional e intergeracional de oportunidades: a matrícula e a escolaridade de crianças e jovens no Brasil. In: ENCONTRO DA ASSOCIAÇÃO BRASILEIRA DE ESTUDOS POPULACIONAIS, 13, 2002, Ouro Preto-MG. **Violências, o estado e a qualidade de vida da população brasileira**: anais. Belo Horizonte: ABEP, 2002. Disponível em CD-ROM.

MELLO E SOUZA, A., SILVA, N. V. Family background, quality of education and public and private schools: effects on school transitions. In: BIRDSALL, N., SABOT, R. H (Eds.) **Opportunity forgone**: education in Brazil. Washington, DC.: Inter-American development Bank.1996. p.367-384.

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.

MORO, S., CHEIN, F., MACHADO, A. Self-employment in Brazil and its determinants: a spatial analysis. Belo Horizonte. CEDEPLAR/UFMG, 2003. (Texto para discussão, 204)

PARK, Y.B., ROSS, D.R; SABOT, R.H. Educational expansion and the inequality of pay in Brazil and Korea. In: BIRDSALL, N., SABOT, R. H (Eds.) **Opportunity forgone**: education in Brazil. Washington, DC.: Inter-American development Bank.1996. p.267-287.

RIANI, J. L. R. Impactos da estrutura etária em indicadores de educação no Brasil, 1991. **Revista Brasileira de Estudos Populacionais**, v.18, n1/2, p.15-34, jan./dez, 2002

RICE, N. Binomial regression. In: LEYLAND, A. H., GOLDSTEIN, H. (Eds.) **Multilevel modeling of health statistics**. New York: Wiley, 2001. p.27-43.

RIOS-NETO, E. L. G. **Análise da progressão de matrícula e população**: o método probabilidade de progressão por série. Versão preliminar apresentada no Workshop de Demografia da Educação. Salvador, Bahia Associação Brasileira de Estudos Populacionais, ABEP, Junho 2001.

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.

RIOS-NETO, E. L. G., RIANI, J. L. R. (Orgs) **Introdução à demografia da educação**. Campinas: ABEP, 2004. 207p.

RIOS-NETO, E. L. G., RIANI, J. L. R., CÉSAR, C. C. **Mother's or teacher's education?** educational stratification and grade progression in Brazil. Belo Horizonte: UFMG/CEDEPLAR, 2003. 26p. (Testo para discussão, 193)

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.

SHAVIT, Y., BLOSSFELD, H. P. **Persistent inequality**: changing educational attainment in thirteen countries. Boulder, CO: Westview, 1993.

SILVA, N. D. V., KASSOUF, A. L. O trabalho e a escolaridade dos brasileiros jovens. In: ENCONTRO DA ASSOCIAÇÃO BRASILEIRA DE ESTUDOS POPULACIONAIS, 13, 2002, Ouro Preto-MG. **Violências, o estado e a qualidade de vida da população brasileira**: anais. Belo Horizonte: ABEP, 2002. Disponível em CD-ROM.

SILVA, N. V. **Atualização da escala socioeconômica de ocupação para 1980**. Laboratório Nacional de Computação Científica, CNPq, 1985. 17p. (Mimeogr.)

SILVA, N. V. **Posição social das ocupações**. Rio de Janeiro: IBGE, Centro de Informática, 1973. (Mimeogr.)

SILVA, N. V. **Posição social nas ocupações**. Rio de Janeiro: IBGE, 1974. 20p. (Mimeogr.)

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

SOARES, J. F., CÉSAR, C. C., MANBRINI, J. Determinantes do desempenho dos alunos do ensino básico brasileiro: evidências do SAEB de 1997. In: FRANCO, C. (Org) **Avaliação, ciclos e promoção na educação.** Porto Alegre: ArtMed, 2001. p.121-154.

SOARES, S. Os fatores que determinam o sucesso educacional. **Pesquisa e Planejamento Econômico**, Rio de Janeiro, v.32, n.3, p.385-394, Dez. 2002.

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.

WILLMS, J. D., SOMERS, M. A. **Schooling outcomes in Latin America**. report prepared for UNESCO, ago. 1999. p.1-43.

WILLMS. J. D. Monitoring school performance for "standards-based reform". **Evaluation and Research in Education**, v.14, n.3&4, p.237-253, 2000.

ANEXOS

ANEXO 1: Modelos hierárquicos complementares

TABELA A1. 1: Resultado da regressão para a probabilidade de freqüentar a escola – indivíduos de 7 a 14 anos de idade - diferentes modelos

Variáveis Independentes	Modelo 7		Modelo 8	
Efeito Fixo	Coeficiente	Sig	Coeficiente	Sig
Intercepto	0.625	0.051	1.5373	0.0000
Tamanho relativo da coorte de 7 a 14 anos			-4.8723	0.0000
Razão professor/pop. 7 a14	10.712	0.000	9.8804	0.0000
Proporção de matrículas em escolas com ciclo	-0.030	0.295		
Proporção de matrículas públicas	-0.080	0.754		
Proporção de matrículas municipais dentre as públicas	0.114	0.007		
Proporção de estabelecimentos com 50 a 200 matrículas	0.554	0.000	0.6277	0.0000
Proporção de estabelecimentos com 201 a 1000	0.650	0.000	0.7328	0.0000
Proporção de estabelecimentos com mais de 1000 matrículas	0.553	0.002	0.6430	0.0000
Educação média da população adulta	0.256	0.000	0.2593	0.0000
População até 4999	0.446	0.000	0.4964	0.0000
População entre 5000 e 9999	0.392	0.000	0.4429	0.0000
População entre 10000 e 19999	0.346	0.000	0.3863	0.0000
População entre 20000 e 49999	0.219	0.001	0.2658	0.0000
População entre 50000 e 99999	0.195	0.003	0.2254	0.0000
População entre 100000 e 199999	0.141	0.048	0.1483	0.0350
Horas-aula diárias	0.078	0.037		
Alunos por turma	-0.007	0.000		
Proporção professores com curso superior	0.089	0.102		
Fator de infra-estrutura	0.091	0.000		
Idade	-0.024	0.000	-0.024	0.000
Situação de domicílio; 1=urbano e 0=rural	0.611	0.000	0.611	0.000
Cor; 1=branco e amarelo e 0=preto e pardo	0.139	0.000	0.139	0.000
Sexo; 1=homem e 0=mulher	-0.199	0.000	-0.199	0.000
Educação da mãe	0.179	0.000	0.179	0.000
Domicílio chefiado por mulher	-0.228	0.000	-0.229	0.000
Família convivente	-0.217	0.000	-0.217	0.000
Categoria de ocupação do chefe nível superior	0.315	0.000	0.315	0.000
Categoria de ocupação do chefe nível médio	0.014	0.090	0.015	0.084
Chefe não ocupado	-0.187	0.000	-0.187	0.000
Efeito Aleatório				
coeficiente	0.258	0.000	0.257	0.000
% da variância explicada	55.61		55.77	

TABELA A1. 2: Resultado da regressão para a probabilidade de freqüentar a escola – indivíduos de 15 a 17 anos de idade - diferentes modelos

Variáveis Independentes	Modelo 7		Modelo 8		Modelo 9	
Efeito Fixo	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig
Intercepto	1.201	0.000	-0.216	0.097	-0.235	0.053
Tamanho relativo da coorte de 15 a 17 anos			16.334	0.000	16.637	0.000
Razão professor/pop. 15 a17	6.666	0.000	6.941	0.000	6.761	0.000
Proporção de estabelecimentos com 50 a 200 matrículas	-0.056	0.276	-0.019	0.712		
Proporção de estabelecimentos com 201 a 1000	0.024	0.638	0.028	0.586		
Proporção de estabelecimentos com mais de 1000 matrículas	0.349	0.000	0.240	0.004		
Proporção de matrículas públicas	-0.276	0.000				
Proporção de matrículas municipais dentre as públicas	0.209	0.000				
Educação média da população adulta	0.165	0.000	0.144	0.000	0.151	0.000
População até 4999	-0.212	0.000	-0.351	0.000	-0.385	0.000
População entre 5000 e 9999	-0.288	0.000	-0.413	0.000	-0.430	0.000
População entre 10000 e 19999	-0.276	0.000	-0.367	0.000	-0.380	0.000
População entre 20000 e 49999	-0.256	0.000	-0.350	0.000	-0.350	0.000
População entre 50000 e 99999	-0.154	0.004	-0.238	0.000	-0.232	0.000
População entre 100000 e 199999	-0.070	0.253	-0.108	0.086	-0.106	0.091
Fator de infra-estrutura	-0.096	0.000				
Horas-aula diárias	-0.099	0.000				
Alunos por turma	0.010	0.000				
Proporção professores com curso superior	-0.248	0.000				
Idade	-0.441	0.000	-0.441	0.000	-0.441	0.000
Situação de domicílio; 1=urbano e 0=rural	0.467	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	-0.444	0.000	-0.444	0.000	-0.444	0.000
Educação da mãe	0.159	0.000	0.159	0.000	0.159	0.000
Domicílio chefiado por mulher	-0.362	0.000	-0.362	0.000	-0.362	0.000
Família convivente	0.069	0.003	0.488	0.000	0.488	0.000
Categoria de ocupação do chefe nível superior	0.488	0.000	0.141	0.000	0.141	0.000
Categoria de ocupação do chefe nível médio	0.141	0.000	-0.047	0.000	-0.047	0.000
Chefe não ocupado	-0.047	0.000	0.069	0.003	0.069	0.003
Efeito Aleatório						
coeficiente	0.19666	0.000	0.209	0.000	0.210	0.000
% da variância explicada	33.06		28.82		28.56	

TABELA A1. 3: Resultado da regressão para a probabilidade de freqüentar a escola – indivíduos de 15 a 17 anos de idade que já concluíram o ensino fundamental - diferentes modelos.

Variáveis Independentes	Modelo 7		Modelo 8		Modelo 9	
Efeito Fixo	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig
Intercepto	2.103	0.000	0.720	0.001	0.662	0.001
Tamanho relativo da coorte de 15 a 17 anos			21.956	0.000	22.235	0.000
Razão professor/pop. 15 a17	6.808	0.000	6.720	0.000	6.631	0.000
Proporção de estabelecimentos com 50 a 200 matrículas	-0.120	0.177	-0.063	0.483		
Proporção de estabelecimentos com 201 a 1000	-0.055	0.532	-0.042	0.632		
Proporção de estabelecimentos com mais de 1000 matrículas	0.296	0.027	0.176	0.184		
Proporção de matrículas públicas	-0.478	0.000				
Proporção de matrículas municipais dentre as públicas	0.450	0.000				
Educação média da população adulta	0.106	0.000	0.073	0.000	0.077	0.000
População até 4999	-0.255	0.003	-0.527	0.000	-0.560	0.000
População entre 5000 e 9999	-0.344	0.000	-0.589	0.000	-0.614	0.000
População entre 10000 e 19999	-0.294	0.000	-0.483	0.000	-0.504	0.000
População entre 20000 e 49999	-0.223	0.002	-0.387	0.000	-0.389	0.000
População entre 50000 e 99999	-0.073	0.298	-0.209	0.004	-0.205	0.005
População entre 100000 e 199999	-0.060	0.450	-0.122	0.137	-0.122	0.139
Fator de infra-estrutura	-0.166	0.000				
Horas-aula diárias	-0.060	0.069				
Alunos por turma	0.019	0.000				
Proporção professores com curso superior	-0.301	0.000				
Idade	-0.391	0.000	-0.391	0.000	-0.391	0.000
Situação de domicílio; 1=urbano e 0=rural	0.678	0.000	0.679	0.000	0.680	0.000
Cor; 1=branco e amarelo e 0=preto e pardo	0.185	0.000	0.185	0.000	0.185	0.000
Sexo; 1=homem e 0=mulher	-0.327	0.000	-0.327	0.000	-0.327	0.000
Educação da mãe	0.121	0.000	0.121	0.000	0.121	0.000
Domicílio chefiado por mulher	-0.362	0.000	-0.362	0.000	-0.362	0.000
Família convivente	0.081	0.099	0.459	0.000	0.459	0.000
Categoria de ocupação do chefe nível superior	0.459	0.000	0.187	0.000	0.187	0.000
Categoria de ocupação do chefe nível médio	0.188	0.000	-0.010	0.474	-0.010	0.474
Chefe não ocupado	-0.009	0.479	0.081	0.100	0.081	0.099
Efeito Aleatório						
coeficiente	0.3116	0.000	0.337	0.000	0.338	0.000
% da variância explicada	29.94		24.18		24.04	

TABELA A1. 4: Resultado da regressão para a probabilidade de freqüentar a escola – indivíduos de 15 a 17 anos de idade - diferentes modelos.

Variáveis Independentes	Modelo coorte		Modelo coorte e restrição		Modelo coorte e nucleação	
Efeito Fixo	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig
Intercepto	2.427	0.000	0.626	0.000	1.969	0.000
Tamanho relativo da coorte de 15 a 17 anos	-4.282	0.000	6.795	0.000	-6.887	0.000
Razão professor/pop. 15 a 17			9.806	0.000		
Proporção de estabelecimentos com 50 a 200 matrículas					-0.128	0.023
Proporção de estabelecimentos com 201 a 1000					-0.090	0.109
Proporção de estabelecimentos com mais de 1000 matrículas					0.638	0.000
População até 4999	-0.660	0.000				
População entre 5000 e 9999	-0.817	0.000				
População entre 10000 e 19999	-0.807	0.000				
População entre 20000 e 49999	-0.710	0.000				
População entre 50000 e 99999	-0.445	0.000				
População entre 100000 e 199999	-0.215	0.002				
Idade	-0.441	0.000	-0.441	0.000	-0.441	0.000
Situação de domicílio; 1=urbano e 0=rural	0.468	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	-0.444	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	0.482	0.000	0.481	0.000
Categoria de ocupação do chefe nível superior	0.483	0.000	0.141	0.000	0.141	0.000
Categoria de ocupação do chefe nível médio	0.141	0.000	-0.047	0.000	-0.048	0.000
Chefe não ocupado	-0.047	0.000	0.069	0.003	0.069	0.003
Efeito Aleatório						
Coeficiente	0.253	0.000	0.259	0.000	0.278	0.000
% da variância explicada	13.71		11.79		5.25	

TABELA A1. 5: Resultado da regressão para a probabilidade de freqüentar a escola na idade correta – indivíduos de 7 a 14 anos de idade - diferentes modelos.

Variáveis Independentes	Modelo restrição de oferta		Modelo ciclo		Modelo recursos humanos		Modelo fator de infraestrutura		Modelo dependência administrativa		Modelo nucleação	
Efeito Fixo	Coef	Sig	Coef	Sig	Coef	Sig	Coef	Sig	Coef	Sig	Coef	Sig
Intercepto	-0.835	0.000	-0.891	0.000	-3.280	0.000	-0.605	0.000	2.991	0.000	-1.105	0.000
Razão professor/pop. 7 a14	3.638	0.000										
Proporção de matrículas em escolas com ciclo			0.959	0.000								
Proporção de matrículas públicas									-2.796	0.000		
Proporção de matrículas municipais dentre as públicas									-1.622	0.000		
Proporção de estabelecimentos com 50 a 200 matrículas											-0.051	0.292
Proporção de estabelecimentos com 201 a 1000											1.639	0.000
Proporção de estabelecimentos com mais de 1000 matrículas											1.629	0.000
Educação média da população adulta												
População até 4999												
População entre 5000 e 9999												
População entre 10000 e 19999												
População entre 20000 e 49999												
População entre 50000 e 99999												
População entre 100000 e 199999												
Horas-aula diárias					0.631	0.000						
Alunos por turma					-0.018	0.000						
Proporção professores com curso superior					1.554	0.000						
Fator de infra-estrutura							0.593	0.000				
Idade	-0.231	0.000	-0.231	0.000	-0.231	0.000	-0.231	0.000	-0.231	0.000	-0.231	0.000
Situação de domicílio; 1=urbano e 0=rural	0.313	0.000	0.313	0.000	0.313	0.000	0.313	0.000	0.314	0.000	0.313	0.000
Cor; 1=branco e amarelo e 0=preto e pardo	0.181	0.000	0.181	0.000	0.181	0.000	0.181	0.000	0.181	0.000	0.181	0.000
Sexo; 1=homem e 0=mulher	-0.226	0.000	-0.226	0.000	-0.226	0.000	-0.226	0.000	-0.226	0.000	-0.226	0.000
Educação da mãe	0.109	0.000	0.098	0.000	0.025	0.045	0.091	0.000	0.064	0.000	0.117	0.000
Razão professor/pop. 7 a14	-0.291	0.000										
Proporção de matrículas em escolas com ciclo			-0.025	0.000								
Proporção de matrículas públicas									-0.017	0.328		
Proporção de matrículas municipais dentre as públicas									0.078	0.000		
Proporção de estabelecimentos com 50 a 200 matrículas											-0.013	0.012
Proporção de estabelecimentos com 201 a 1000											-0.076	0.000
Proporção de estabelecimentos com mais de 1000 matrículas											0.018	0.244
Educação média da população adulta												
População até 4999												
População entre 5000 e 9999												
População entre 10000 e 19999												
População entre 20000 e 49999												
População entre 50000 e 99999												
População entre 100000 e 199999												
Horas-aula diárias					0.014	0.000						
Alunos por turma					0.001	0.000						
Proporção professores com curso superior					-0.099	0.000						
Fator de infra-estrutura							-0.028	0.000				
Domicílio chefiado por mulher	-0.139	0.000	-0.139	0.000	-0.139	0.000	-0.139	0.000	-0.139	0.000	-0.139	0.000
Família convivente	0.009	0.242	0.009	0.241	0.008	0.257	0.009	0.250	0.009	0.240	0.009	0.240
Categoria de ocupação do chefe nível superior	0.127	0.000	0.127	0.000	0.128	0.000	0.127	0.000	0.126	0.000	0.127	0.000
Categoria de ocupação do chefe nível médio	0.088	0.000	0.088	0.000	0.089	0.000	0.089	0.000	0.088	0.000	0.088	0.000
Chefe não ocupado	-0.063	0.000	-0.063	0.000	-0.062	0.000	-0.062	0.000	-0.063	0.000	-0.062	0.000
Efeito Aleatório												
Coeficiente	0.577	0.000	0.454	0.000	0.171	0.000	0.224	0.000	0.381	0.000	0.391	0.000
Educação da mãe	0.003	0.000	0.003	0.000	0.002	0.000	0.002	0.000	0.002	0.000	0.002	0.000
% da variância explicada - intercepto	0.63		21.71		70.58		61.42		34.32		32.59	
% da variância explicada - educação materna	0.38		2.30		37.93		31.03		13.79		13.03	

TABELA A1. 6: Resultado da regressão para a probabilidade de freqüentar a escola na idade correta – indivíduos de 15 a 17 anos de idade - diferentes modelos.

Variáveis Independentes	Modelo restrição de oferta		Modelo recursos humanos		Modelo infraestrutura		Modelo depedência administrativa		Modelo nucleação	
Efeito Fixo	Coef	Sig	Coef	Sig	Coef	Sig	Coef	Sig	Coef	Sig
Intercepto	-3.107	0.000	-4.375	0.000	-1.915	0.000	-2.375	0.000	-2.219	0.000
Razão professor/pop. 15 a17	31.807	0.000								
Proporção de estabelecimentos com 50 a 200 matrículas									0.195	0.129
Proporção de estabelecimentos com 201 a 1000									0.416	0.002
Proporção de estabelecimentos com mais de 1000 matrículas									1.084	0.000
Proporção de matrículas públicas							0.696	0.000		
Proporção de matrículas municipais dentre as públicas							-1.987	0.000		
Educação média da população adulta										
População até 4999										
População entre 5000 e 9999										
População entre 10000 e 19999										
População entre 20000 e 49999										
População entre 50000 e 99999										
População entre 100000 e 199999										
Fator de infra-estrutura					0.866	0.000				
Horas-aula diárias			0.584	0.000						
Alunos por turma			-0.043	0.000						
Proporção professores com curso superior			2.044	0.000						
Idade	-0.098	0.000	-0.098	0.000	-0.098	0.000	-0.098	0.000	-0.098	0.000
Situação de domicílio; 1=urbano e 0=rural	0.522	0.000	0.522	0.000	0.521	0.000	0.521	0.000	0.521	0.000
Cor; 1=branco e amarelo e 0=preto e pardo	0.396	0.000	0.395	0.000	0.396	0.000	0.396	0.000	0.396	0.000
Sexo; 1=homem e 0=mulher	-0.546	0.000	-0.546	0.000	-0.546	0.000	-0.546	0.000	-0.546	0.000
Educação da mãe	0.228	0.000	0.172	0.000	0.175	0.000	0.222	0.000	0.207	0.000
Razão professor/pop. 15 a17	-1.432	0.000								
Proporção de estabelecimentos com 50 a 200 matrículas									-0.029	0.007
Proporção de estabelecimentos com 201 a 1000									-0.041	0.000
Proporção de estabelecimentos com mais de 1000 matrículas									-0.064	0.000
Proporção de matrículas públicas							-0.063	0.000		
Proporção de matrículas municipais dentre as públicas							0.112	0.000		
Educação média da população adulta										
População até 4999										
População entre 5000 e 9999										
População entre 10000 e 19999										
População entre 20000 e 49999										
População entre 50000 e 99999										
População entre 100000 e 199999										
Fator de infra-estrutura					-0.043	0.000				
Horas-aula diárias			0.000	0.986						
Alunos por turma			0.003	0.000						
Proporção professores com curso superior			-0.128	0.000						
Domicílio chefiado por mulher	-0.319	0.000	-0.319	0.000	-0.319	0.000	-0.319	0.000	-0.319	0.000
Família convivente	-0.057	0.017	-0.057	0.017	-0.056	0.017	-0.056	0.018	-0.057	0.017
Categoria de ocupação do chefe nível superior	0.337	0.000	0.336	0.000	0.338	0.000	0.337	0.000	0.337	0.000
Categoria de ocupação do chefe nível médio	0.273	0.000	0.271	0.000	0.273	0.000	0.272	0.000	0.273	0.000
Chefe não ocupado	0.019	0.009	0.019	0.009	0.019	0.008	0.019	0.008	0.019	0.008
Efeito Aleatório										
Coeficiente	0.914	0.000	1.077	0.000	0.873	0.000	1.373	0.000	1.605	0.000
Educação da mãe	0.003	0.000	0.003	0.000	0.003	0.000	0.004	0.000	0.004	0.000
% da variância explicada - intercepto	43.40		33.31		45.92		15.02		0.65	
% da variância explicada - educação materna	32.89		29.56		41.33		14.44		0.89	

TABELA A1. 7: Resultado da regressão para a probabilidade de progressão na 1ª série – diferentes modelos.

Variáveis independentes	Modelo restrição deoferta		Modelo ciclo		Modelo recursos humanos		Modelo infraestrutura		Modelo dependência administrativa		Modelo nucleação	
Efeito Fixo	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig
Intercepto	1.578	0.000	1.887	0.000	0.542	0.014	2.189	0.000	7.327	0.000	1.477	0.000
Razão professor/pop. 7 a14	9.179	0.000										
Proporção de matrículas em escolas com ciclo			0.885	0.000								
Proporção de matrículas públicas									-4.289	0.000		
Proporção de matrículas municipais dentre as públicas									-1.871	0.000		
Proporção de estabelecimentos com 50 a 200 matrículas											0.123	0.147
Proporção de estabelecimentos com 201 a 1000											2.135	0.000
Proporção de estabelecimentos com mais de 1000 matrículas											1.357	0.000
Educação média da população adulta												
População até 4999												
População entre 5000 e 9999												
População entre 10000 e 19999												
População entre 20000 e 49999												
População entre 50000 e 99999												
População entre 100000 e 199999												
Horas-aula diárias					0.450	0.000						
Alunos por turma					-0.029	0.000						
Proporção professores com curso superior					1.856	0.000						
Fator de infra-estrutura							0.731	0.000				
Situação de domicílio; 1=urbano e 0=rural	0.528	0.000	0.528	0.000	0.524	0.000	0.523	0.000	0.525	0.000	0.519	0.000
Cor; 1=branco e amarelo e 0=preto e pardo	0.154	0.000	0.156	0.000	0.158	0.000	0.156	0.000	0.155	0.000	0.154	0.000
Sexo; 1=homem e 0=mulher	-0.425	0.000	-0.425	0.000	-0.426	0.000	-0.425	0.000	-0.425	0.000	-0.425	0.000
Educação da mãe	0.191	0.000	0.196	0.000	0.402	0.000	0.160	0.000	0.022	0.624	0.205	0.000
Razão professor/pop. 7 a14	-0.531	0.005										
Proporção de matrículas em escolas com ciclo			-0.121	0.000								
Proporção de matrículas públicas									0.051	0.298		
Proporção de matrículas municipais dentre as públicas									0.160	0.000		
Proporção de estabelecimentos com 50 a 200 matrículas											0.036	0.042
Proporção de estabelecimentos com 201 a 1000											-0.171	0.000
Proporção de estabelecimentos com mais de 1000 matrículas											-0.104	0.017
Educação média da população adulta												
População até 4999												
População entre 5000 e 9999												
População entre 10000 e 19999												
População entre 20000 e 49999												
População entre 50000 e 99999												
População entre 100000 e 199999												
Horas-aula diárias					-0.057	0.000						
Alunos por turma					0.001	0.000						
Proporção professores com curso superior					-0.126	0.000						
Fator de infra-estrutura							-0.054	0.000				
Domicílio chefiado por mulher	-0.085	0.000	-0.085	0.000	-0.085	0.000	-0.084	0.000	-0.084	0.000	-0.085	0.000
Família convivente	-0.084	0.009	-0.084	0.009	-0.088	0.007	-0.087	0.007	-0.086	0.008	-0.085	0.009
Categoria de ocupação do chefe nível superior	0.142	0.000	0.146	0.000	0.173	0.000	0.172	0.000	0.155	0.000	0.153	0.000
Categoria de ocupação do chefe nível médio	0.036	0.032	0.040	0.018	0.046	0.007	0.043	0.012	0.039	0.020	0.041	0.015
Chefe não ocupado	-0.155	0.000	-0.154	0.000	-0.149	0.000	-0.151	0.000	-0.154	0.000	-0.153	0.000
Efeito Aleatório												
Coeficiente	0.999	0.000	0.908	0.000	0.493	0.000	0.554	0.000	0.711	0.000	0.756	0.000
Educação da mãe	0.008	0.000	0.006	0.000	0.006	0.000	0.006	0.000	0.007	0.000	0.007	0.000
% da variância explicada - intercepto	1.48		10.44		51.36		45.32		29.82		25.37	
% da variância explicada - educação materna	-0.24		27.29		25.21		27.78		16.77		18.97	

TABELA A1. 8: Resultado da regressão para a probabilidade de progressão na 5ª série – diferentes modelos.

Variáveis independentes	Modelo restrição de oferta		Modelo ciclo		Modelo recursos humanos		Modelo infraestrutura		Modelo dependência		Modelo nucleação	
Efeito Fixo	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig
Intercepto	1.224	0.000	1.321	0.000	-0.948	0.000	1.673	0.000	6.921	0.000	0.898	0.000
Razão professor/pop. 7 a14	7.105	0.000										
Proporção de matrículas em escolas com ciclo			1.140	0.000								
Proporção de matrículas públicas									-4.237	0.000		
Proporção de matrículas municipais dentre as públicas									-2.134	0.000		
Proporção de estabelecimentos com 50 a 200 matrículas											0.071	0.339
Proporção de estabelecimentos com 201 a 1000											2.455	0.000
Proporção de estabelecimentos com mais de 1000 matrículas											1.437	0.000
Educação média da população adulta												
População até 4999												
População entre 5000 e 9999												
População entre 10000 e 19999												
População entre 20000 e 49999												
População entre 50000 e 99999												
População entre 100000 e 199999												
Horas-aula diárias					0.609	0.000						
Alunos por turma					-0.025	0.000						
Proporção professores com curso superior					2.221	0.000						
Fator de infra-estrutura							0.840	0.000				
Situação de domicílio; 1=urbano e 0=rural	0.528	0.000	0.528	0.000	0.530	0.000	0.527	0.000	0.526	0.000	0.520	0.000
Cor; 1=branco e amarelo e 0=preto e pardo	0.312	0.000	0.314	0.000	0.316	0.000	0.314	0.000	0.313	0.000	0.314	0.000
Sexo; 1=homem e 0=mulher	-0.515	0.000	-0.515	0.000	-0.519	0.000	-0.518	0.000	-0.516	0.000	-0.516	0.000
Educação da mãe	0.181	0.000	0.181	0.000	0.182	0.000	0.182	0.000	0.182	0.000	0.181	0.000
Domicílio chefiado por mulher	-0.251	0.000	-0.251	0.000	-0.253	0.000	-0.252	0.000	-0.251	0.000	-0.251	0.000
Família convivente	-0.039	0.328	-0.039	0.329	-0.042	0.298	-0.042	0.302	-0.042	0.304	-0.040	0.324
Categoria de ocupação do chefe nível superior	0.650	0.000	0.650	0.000	0.677	0.000	0.672	0.000	0.666	0.000	0.661	0.000
Categoria de ocupação do chefe nível médio	0.180	0.000	0.182	0.000	0.185	0.000	0.182	0.000	0.182	0.000	0.183	0.000
Chefe não ocupado	-0.026	0.044	-0.026	0.050	-0.021	0.104	-0.022	0.086	-0.025	0.058	-0.024	0.064
Efeito Aleatório												
Coeficiente	0.908	0.000	0.742	0.000	0.253	0.000	0.265	0.000	0.541	0.000	0.560	0.000
% da variância explicada - intercepto	0.85		19.00		72.39		71.03		40.87		38.81	

TABELA A1. 9: Resultado da regressão para a probabilidade de progressão na 1ª série do ensino médio – diferentes modelos.

Variáveis Independentes	Modelo restrição de oferta		Modelo recursos humanos		Modelo infraestrutura		Modelo dependência		Modelo nucleação	
Efeito Fixo	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig	Coeficiente	Sig
Intercepto	0.633	0.000	-0.436	0.003	1.154	0.000	1.698	0.000	1.276	0.000
Razão professor/pop. 15 a 17	15.165	0.000								
Proporção de matrículas públicas							-0.467	0.000		
Proporção de matrículas municipais dentre as públicas							-0.607	0.000		
Proporção de estabelecimentos com 50 a 200 matrículas									-0.124	0.148
Proporção de estabelecimentos com 201 a 1000									-0.061	0.473
Proporção de estabelecimentos com mais de 1000 matrículas									0.275	0.022
Educação média da população adulta										
População até 4999										
População entre 5000 e 9999										
População entre 10000 e 19999										
População entre 20000 e 49999										
População entre 50000 e 99999										
População entre 100000 e 199999										
Horas-aula diárias			0.354	0.000						
Alunos por turma			-0.010	0.000						
Proporção professores com curso superior			0.650	0.000						
Fator de infra-estrutura					0.310	0.000				
Situação de domicílio; 1=urbano e 0=rural	0.476	0.000	0.476	0.000	0.475	0.000	0.473	0.000	0.473	0.000
Cor; 1=branco e amarelo e 0=preto e pardo	0.296	0.000	0.296	0.000	0.295	0.000	0.295	0.000	0.296	0.000
Sexo; 1=homem e 0=mulher	-0.574	0.000	-0.573	0.000	-0.574	0.000	-0.572	0.000	-0.572	0.000
Educação da mãe	0.124	0.000	0.124	0.000	0.124	0.000	0.124	0.000	0.124	0.000
Domicílio chefiado por mulher	-0.263	0.000	-0.263	0.000	-0.263	0.000	-0.262	0.000	-0.262	0.000
Família convivente	-0.042	0.447	-0.041	0.455	-0.041	0.460	-0.039	0.475	-0.039	0.476
Categoria de ocupação do chefe nível superior	0.564	0.000	0.561	0.000	0.563	0.000	0.559	0.000	0.558	0.000
Categoria de ocupação do chefe nível médio	0.210	0.000	0.209	0.000	0.209	0.000	0.209	0.000	0.209	0.000
Chefe não ocupado	0.051	0.001	0.051	0.001	0.052	0.001	0.050	0.001	0.050	0.001
Efeito Aleatório										
Coeficiente	0.148	0.000	0.192	0.000	0.169	0.000	0.232	0.000	0.256	0.000
% da variância explicada - intercepto	42.95		26.00		34.95		10.69		1.56	

ANEXO 2: Matriz de correlação das variáveis do segundo nível do modelo hierárquico.

TABELA A2. 1: Matriz de correlação para as variáveis do ensino fundamental

Variáveis												
Tamanho relativo da coorte de 7 a 14 anos	1.000	-0.291	-0.666	-0.452	0.449	-0.708	-0.140	0.045	0.126	0.077	-0.061	-0.092
Razão professor/pop. 7 a 14	-0.291	1.000	0.021	-0.067	-0.371	0.015	0.336	0.050	-0.135	-0.153	-0.110	-0.086
Fator de infra-estrutura	-0.666	0.021	1.000	0.501	-0.353	0.768	0.054	-0.038	-0.111	-0.026	0.083	0.099
Horas-aula diárias	-0.452	-0.067	0.501	1.000	-0.112	0.559	0.059	-0.022	-0.085	-0.014	0.029	0.069
Alunos por turma	0.449	-0.371	-0.353	-0.112	1.000	-0.324	-0.178	-0.013	0.116	0.078	0.009	0.000
Proporção professores com curso superior	-0.708	0.015	0.768	0.559	-0.324	1.000	0.045	-0.050	-0.107	-0.024	0.100	0.119
População até 4999	-0.140	0.336	0.054	0.059	-0.178	0.045	1.000	-0.315	-0.325	-0.259	-0.136	-0.083
População entre 5000 e 9999	0.045	0.050	-0.038	-0.022	-0.013	-0.050	-0.315	1.000	-0.323	-0.258	-0.135	-0.082
População entre 10000 e 19999	0.126	-0.135	-0.111	-0.085	0.116	-0.107	-0.325	-0.323	1.000	-0.266	-0.139	-0.085
População entre 20000 e 49999	0.077	-0.153	-0.026	-0.014	0.078	-0.024	-0.259	-0.258	-0.266	1.000	-0.111	-0.068
População entre 50000 e 99999	-0.061	-0.110	0.083	0.029	0.009	0.100	-0.136	-0.135	-0.139	-0.111	1.000	-0.035
População entre 100000 e 199999	-0.092	-0.086	0.099	0.069	0.000	0.119	-0.083	-0.082	-0.085	-0.068	-0.035	1.000
Educação média da população adulta	-0.720	0.075	0.725	0.461	-0.434	0.724	-0.038	-0.120	-0.137	0.020	0.203	0.224
Proporção de estabelecimentos com 50 a 200 matrículas	0.108	0.027	-0.136	-0.196	-0.109	-0.094	-0.113	-0.024	0.086	0.088	0.002	-0.038
Proporção de estabelecimentos com 201 a 1000	-0.429	-0.153	0.543	0.478	-0.243	0.579	0.090	-0.041	-0.137	-0.042	0.086	0.124
Proporção de estabelecimentos com mais de 1000 matrículas	-0.089	-0.193	0.173	0.149	0.037	0.168	-0.155	-0.102	-0.042	0.057	0.134	0.198
Proporção de matrículas em escolas com ciclo	-0.292	-0.078	0.381	0.552	-0.082	0.445	-0.008	-0.012	-0.042	0.003	0.040	0.060
Proporção de matrículas públicas	0.266	0.081	-0.257	-0.185	0.095	-0.298	0.302	0.225	0.035	-0.222	-0.304	-0.251
Proporção de matrículas municipais dentre as públicas	0.420	-0.057	-0.465	-0.421	0.286	-0.458	-0.083	-0.008	0.103	0.053	-0.033	-0.057

TABELA A2. 2: Matriz de correlação para as variáveis do ensino médio.

Variáveis												
Tamanho relativo da coorte de 15 a 17 anos	1.000	-0.567	-0.028	0.033	-0.003	-0.612	-0.134	0.043	0.087	0.075	-0.034	-0.064
Razão professor/pop. 15 a 17	-0.567	1.000	0.155	-0.161	-0.036	0.537	0.385	-0.051	-0.217	-0.142	0.004	0.033
Proporção de estabelecimentos com 50 a 200 matrículas	-0.028	0.155	1.000	-0.890	-0.103	-0.117	0.481	-0.079	-0.224	-0.113	-0.033	-0.041
Proporção de estabelecimentos com 201 a 1000	0.033	-0.161	-0.890	1.000	-0.189	0.045	-0.410	0.152	0.254	0.009	-0.048	-0.008
Proporção de estabelecimentos com mais de 1000 matrículas	-0.003	-0.036	-0.103	-0.189	1.000	0.220	-0.152	-0.162	-0.137	0.232	0.258	0.144
Educação média da população adulta	-0.612	0.537	-0.117	0.045	0.220	1.000	-0.038	-0.120	-0.137	0.020	0.203	0.224
População até 4999	-0.134	0.385	0.481	-0.410	-0.152	-0.038	1.000	-0.315	-0.325	-0.259	-0.136	-0.083
População entre 5000 e 9999	0.043	-0.051	-0.079	0.152	-0.162	-0.120	-0.315	1.000	-0.323	-0.258	-0.135	-0.082
População entre 10000 e 19999	0.087	-0.217	-0.224	0.254	-0.137	-0.137	-0.325	-0.323	1.000	-0.266	-0.139	-0.085
População entre 20000 e 49999	0.075	-0.142	-0.113	0.009	0.232	0.020	-0.259	-0.258	-0.266	1.000	-0.111	-0.068
População entre 50000 e 99999	-0.034	0.004	-0.033	-0.048	0.258	0.203	-0.136	-0.135	-0.139	-0.111	1.000	-0.035
População entre 100000 e 199999	-0.064	0.033	-0.041	-0.008	0.144	0.224	-0.083	-0.082	-0.085	-0.068	-0.035	1.000
Fator de infra-estrutura	-0.467	0.384	-0.176	0.135	0.164	0.595	-0.080	-0.046	-0.056	0.051	0.128	0.092
Horas-aula diárias	-0.213	0.225	-0.086	0.041	0.108	0.296	-0.055	-0.015	-0.048	0.039	0.065	0.066
Alunos por turma	0.246	-0.360	-0.342	0.320	0.209	-0.087	-0.396	-0.071	0.108	0.218	0.143	0.093
Proporção professores com curso superior	-0.384	0.269	-0.103	0.085	0.086	0.409	-0.007	-0.014	-0.076	0.012	0.076	0.067
Proporção de matrículas públicas	-0.040	-0.018	-0.037	0.081	-0.064	-0.118	0.156	0.141	-0.011	-0.137	-0.131	-0.097
Proporção de matrículas municipais dentre as públicas	0.278	-0.174	0.044	-0.029	-0.060	-0.335	-0.062	0.007	0.084	0.028	-0.057	-0.040

ANEXO 3: Resultado dos modelos hierárquicos que imputam o valor zero para os municípios que não possuem estabelecimentos de ensino médio

TABELA A3. 1: Resultado da regressão para a probabilidade de freqüentar a escola – indivíduos de 15 a 17 anos de idade.

Variáveis Independentes	Modelo 1	
Efeito Fixo	Coeficiente	Sig
Intercepto	1.433	0.000
Idade	-0.441	0.000
Situação de domicílio; 1=urbano e 0=rural	0.467	0.000
Cor; 1=branco e amarelo e 0=preto e pardo	0.162	0.000
Sexo; 1=homem e 0=mulher	-0.446	0.000
Educação da mãe	0.158	0.000
Domicílio chefiado por mulher	-0.360	0.000
Família convivente	0.071	0.003
Categoria de ocupação do chefe nível superior	0.480	0.000
Categoria de ocupação do chefe nível médio	0.139	0.000
Chefe não ocupado	-0.047	0.000
Efeito Aleatório		
coeficiente	0.291	0.000

TABELA A3. 2: Resultado da regressão para a probabilidade de freqüentar a escola – indivíduos de 15 a 17 anos de idade que já concluíram o ensino médio.

Variáveis Independentes	Modelo 1	
Efeito Fixo	Coeficiente	Sig
Intercepto	2.294	0.000
Idade	-0.390	0.000
Situação de domicílio; 1=urbano e 0=rural	0.686	0.000
Cor; 1=branco e amarelo e 0=preto e pardo	0.184	0.000
Sexo; 1=homem e 0=mulher	-0.325	0.000
Educação da mãe	0.121	0.000
Domicílio chefiado por mulher	-0.359	0.000
Família convivente	0.078	0.108
Categoria de ocupação do chefe nível superior	0.456	0.000
Categoria de ocupação do chefe nível médio	0.184	0.000
Chefe não ocupado	-0.011	0.427
Efeito Aleatório		
coeficiente	0.451	0.000

TABELA A3. 3: Resultado da regressão para a probabilidade de freqüentar a escola na idade correta– indivíduos de 15 a 17 anos de idade.

Variáveis Independentes	Modelo 1	
Efeito Fixo	Coeficiente	Sig
Intercepto	-1.935	0.000
Idade	-0.098	0.000
situação urbana	0.522	0.000
branco	0.396	0.000
homem	-0.547	0.000
educação da mãe	0.174	0.000
mãe chefe	-0.318	0.000
posição na ocupação superior	0.338	0.000
posição na ocupação média	0.274	0.000
não ocupado	0.019	0.010
família convivente	-0.057	0.017
Efeito Aleatório		
coeficiente	1.682	0.000
Educação da mãe	0.005	0.000

TABELA A3. 4: Resultado da regressão para a Probabilidade de Progressão na 1ª série do ensino médio.

Variáveis Independentes	Modelo 1	
Efeito Fixo	Coeficiente	Sig
Intercepto	1.202	0.000
Situação de domicílio; 1=urbano e 0=rural	0.474	0.000
Cor; 1=branco e amarelo e 0=preto e pardo	0.293	0.000
Sexo; 1=homem e 0=mulher	-0.571	0.000
Educação da mãe	0.124	0.000
Domicílio chefiado por mulher	-0.260	0.000
Família convivente	-0.036	0.509
Categoria de ocupação do chefe nível superior	0.557	0.000
Categoria de ocupação do chefe nível médio	0.209	0.000
Chefe não ocupado	0.048	0.001
Efeito Aleatório		
coeficiente	0.265	0.000

ANEXO 4: Gráficos com simulações de variações em todos os fatores do nível 2

GRÁFICO A4. 1: Variação na probabilidade de cursar a escola na idade adequada decorrente do aumento de um desvio padrão em relação à média – simulações segundo diferentes modelos para os indivíduos de 7 a 14 anos.

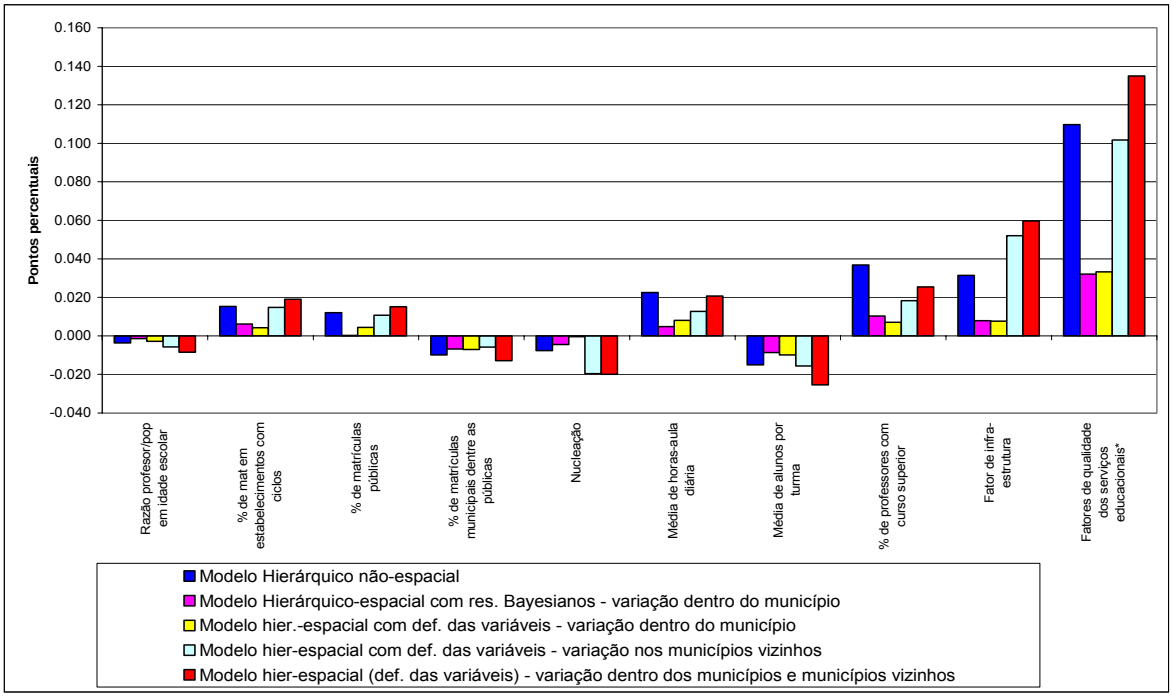


GRÁFICO A4. 2: Variação na probabilidade de cursar a escola na idade adequada decorrente do aumento de um desvio padrão em relação à média – simulações segundo diferentes modelos para os indivíduos de 15 a 17 anos.

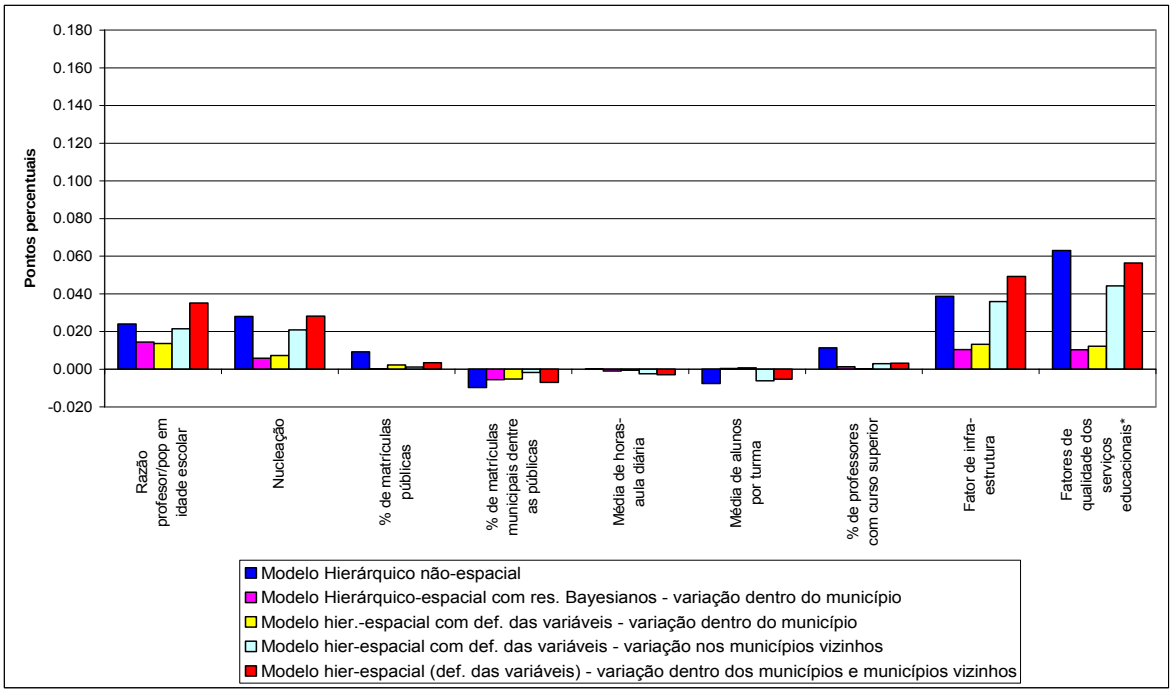


GRÁFICO A4. 3: Variação na probabilidade de progressão na 1ª série decorrente do aumento de um desvio padrão em relação à média – simulações segundo diferentes modelos.

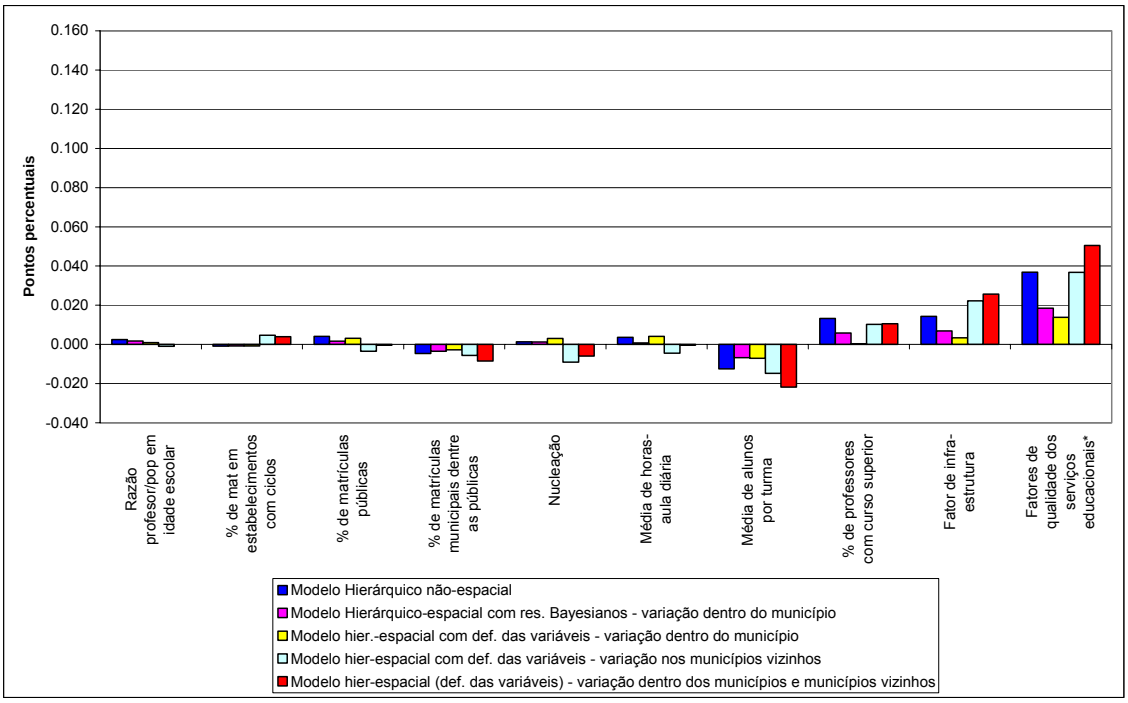


GRÁFICO A4. 4: Variação na probabilidade de progressão na 5ª série decorrente do aumento de um desvio padrão em relação à média – simulações segundo diferentes modelos.

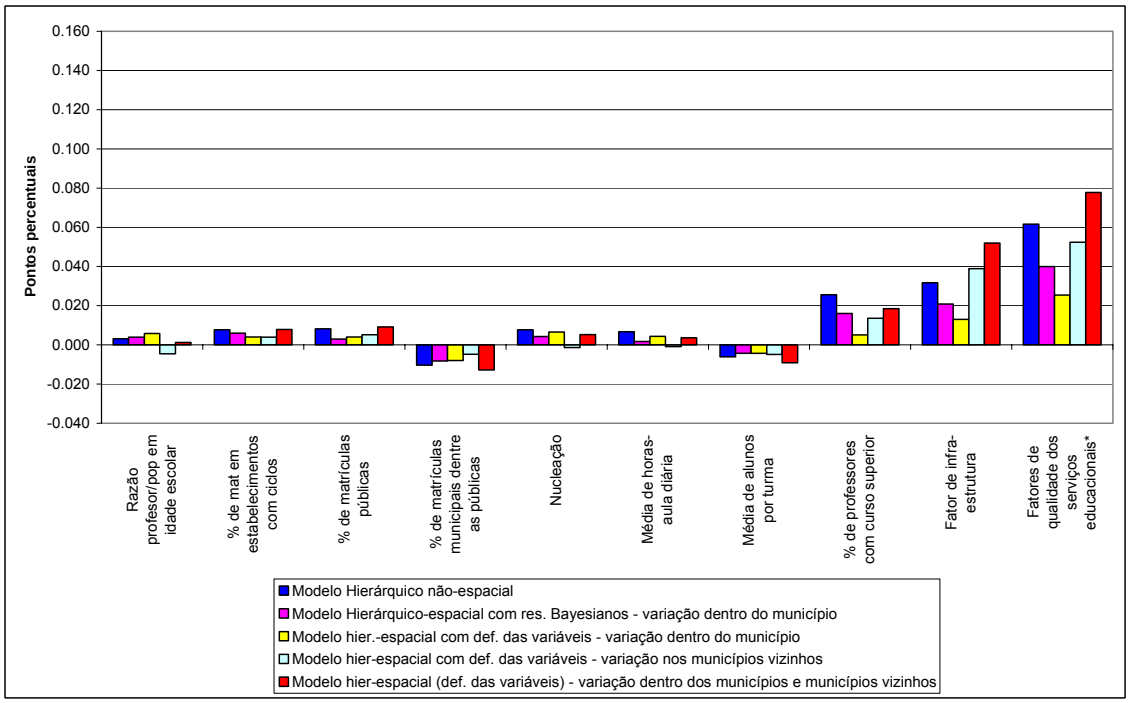


GRÁFICO A4. 5: Variação na probabilidade de progressão na 1ª série do ensino médio decorrente do aumento de um desvio padrão em relação à média – simulações segundo diferentes modelos.