

IMPACTO DOS FATORES FAMILIARES, ESCOLARES E COMUNITÁRIOS NA QUANTIDADE E QUALIDADE DO ENSINO NO ESTADO DE MINAS GERAIS

Juliana de Lucena Ruas Riani^{*}
Eduardo Luiz Gonçalves Rios-Neto^{**}

Palavras chaves: educação, modelos hierárquicos, recursos escolares.

Resumo

Na análise do desempenho escolar existem vários fatores familiares, escolares e comunitários que afetam a educação. No Brasil, a maioria dos estudos concentra suas análises apenas no aspecto intradomiciliar não levando em consideração questões de oferta da educação. Essa linha de pensamento é importante por determinar os condicionantes intradomiciliares da educação. Porém, pouco pode-se dizer em relação a elaboração de políticas públicas que possam melhorar a educação em seus aspectos quantitativos e qualitativos. Neste sentido, a incorporação de variáveis de caráter da oferta do sistema de ensino é importante.

Desta forma, o objetivo deste artigo foi investigar como que os fatores de background familiar, comunitários e escolares determinam a quantidade e qualidade do ensino fundamental em Minas Gerais, recorrendo aos modelos logísticos hierárquicos de dois níveis. O primeiro nível é o indivíduo e o segundo, os municípios mineiros. A quantidade do ensino foi captada pela probabilidade de freqüentar a escola e a qualidade pela probabilidade de freqüentar a escola na idade adequada à série que se cursa.

As variáveis explicativas foram agrupadas em três esferas. A primeira foi as variáveis de background familiar, como educação dos pais, posição ocupacional do pai e número de filhos no domicílio. A segunda esfera incluiu as variáveis relacionadas à condição de oferta escolar como tamanho das salas de aula, educação e salário dos professores, média de horas-aula, porcentagem de escolas com bibliotecas, quadras, laboratórios de informática e ciências. Na terceira estão as variáveis relacionadas aos recursos dos municípios, como anos médios de estudo da população adulta.

A conclusão a que se chegou é que os fatores familiares são possuem importante impacto sobre as duas transições, porém, os fatores escolares também, abrindo possibilidades para a elaboração de políticas públicas que beneficiem a melhoria dos insumos escolares, ou seja, melhor qualificação dos professores e infra-estrutura das escolas.

^{*} Doutoranda em Demografia pelo CEDEPLAR/UFMG.

^{**} Professor titular do curso de Demografia do CEDEPLAR/UFMG.

1) Introdução

O sistema educacional brasileiro tem evoluído significativamente na última década, observando-se melhora nos seus indicadores, principalmente em relação aos anos médios de estudo e matrícula. Apesar desta melhora, o Brasil ainda apresenta índices baixos quando comparado a outros países, além de uma grande desigualdade educacional. De acordo com dados do Ministério da Educação e Cultura (MEC), o tempo médio de escolaridade do Brasil é de 6,2 anos, enquanto que no Paraguai é de 8 anos, Argentina 9 anos e Europa, 16 anos.

O baixo desempenho 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. 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, desta forma, um ciclo vicioso.

Desta forma, estudos que tentam compreender a dinâmica e os fatores associados ao processo educacional são importantes para entender a baixa escolaridade do Brasil e para auxiliar nas políticas educacionais.

Na análise do desempenho escolar, existem três conjuntos de fatores que podem influenciar a educação, seja através da mudança na demanda escolar ou na oferta educacional. O primeiro conjunto incorpora os fatores relacionados à família que alteram a demanda por educação, como por exemplo a estrutura familiar e o nível sócio-econômico. O segundo agrupa aqueles associados à escola como infra-estrutura, professores 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.

Os diferentes fatores que afetam a educação têm sido foco de análise de vários estudos, porém, a maioria concentra suas análises principalmente nos fatores micros, ou seja, nos fatores familiares e atributos individuais. Esta análise é importante para o entendimento do desempenho do ensino, principalmente em termos da desigualdade educacional, pois os estudos demonstram que pais pobres e com menor educação geralmente transferem baixa escolaridade para os filhos. Porém, é também importante analisar como os fatores de condições de oferta do ensino público e os fatores comunitários exercem influência na escolaridade, para desta forma conduzir políticas públicas educacionais mais eficazes.

Algumas tentativas neste ponto já têm sido feitas no Brasil, porém, a maioria dos estudos utiliza a estimação de Mínimos Quadrados Ordinários (MQO), colocando as variáveis

agregadas no mesmo nível das individuais. Esse método assume que os erros são independentes e com variância constante e na ausência desses pressupostos as estimativas podem ser enviesadas e não eficientes¹.

Numa análise agregada do Brasil ou dos estados é bem provável que estes pressupostos não se apliquem. Se considerarmos que a relação entre as características dos estudantes e seus desempenhos escolares pode variar através dos municípios devido as diferentes disponibilidades e qualidades de recursos educacionais, bem como à diferença de estoque de capital e riqueza municipal.

Para lidar com a quebra dos pressupostos dos modelos MQO tem-se recorrido aos modelos hierárquicos. Esses modelos, além de possibilitar a obtenção de melhores estimativas, permitem estudar a interação das variáveis independentes entre os níveis.

O objetivo deste trabalho foi investigar como os fatores de *background* familiar, comunitários e escolares determinam a quantidade e qualidade do ensino fundamental em Minas Gerais, recorrendo-se a modelos logísticos hierárquicos de dois níveis. O primeiro nível contemplou o indivíduo e o segundo nível o município. Com a modelagem hierárquica também foi possível analisar a partição da variância em componentes, ou seja, determinar a importância dos diferentes níveis na variabilidade dos dados.

A quantidade do ensino foi captada pela probabilidade do indivíduo de 7 a 14 anos está matriculado na escola e a qualidade pela probabilidade do aluno de 7 a 14 anos está cursando a série adequada a sua idade. As variáveis explicativas foram agrupadas em três grupos. O primeiro composto por variáveis de *background* familiar, como educação dos pais, *status* ocupacional do pai e número de filhos no domicílio. O segundo grupo incluem as variáveis relacionadas à condição de oferta escolar como tamanho das salas de aula, educação e salário dos professores, média de horas-aula, porcentagem de escolas com bibliotecas, quadras, laboratórios de informática e ciências. No terceiro grupo estão as variáveis relacionadas aos recursos dos municípios, como escolaridade média da população adulta.

Esse trabalho foi dividido em mais três seções, além dessa introdutória. Na próxima seção, são vistos alguns trabalhos que analisam os determinantes educacionais, os quais auxiliaram na escolha das variáveis independentes utilizadas no modelo econométrico. Na terceira seção, são apresentados a base de dados e os aspectos metodológicos do trabalho. Na quarta parte são comentados os resultados do modelo hierárquico. Por fim, na última são inferidas algumas conclusões a respeito dos determinantes familiares, escolares e

¹ Um estimador é não enviesado quando possui distribuição amostral com média igual ao parâmetro que será estimado e é eficiente quando possui a menor dispersão possível.

comunitários na probabilidade de freqüentar a escola e na probabilidade de freqüentar a escola na idade adequada à série do ensino fundamental nos municípios de Minas Gerais.

2) Determinantes do Desempenho Escolar

Como já foi supracitado, existem três conjuntos de fatores (familiares, escolares e comunitários) que afetam o desempenho educacional do indivíduo. Esses três conjuntos afetam a educação de maneira distinta. O primeiro altera a demanda intradomiciliar por educação enquanto que os outros dois alteram o lado da oferta educacional. A maioria dos estudos que tentam entender os fatores associados ao desempenho dos indicadores educacionais enfoca principalmente o primeiro conjunto de fatores.

Dentro deste contexto, Silva e Hasenbalg (2001) destacam 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.

Os autores estudam o efeito destas dimensões nas transições escolares entre séries para o Brasil com dados da Pesquisa Nacional de Amostra por Domicílio - 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 extensão familiar. Como esperado, seus resultados mostram uma relação positiva para as variáveis de capital econômico e cultural e negativa para as variáveis de capital social familiar. O padrão do efeito das variáveis explicativas é crescente até a 4ª série do ensino fundamental, onde 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, eles encontraram 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. Um fato importante ...

Outros estudos como os de Marteleto (2001), Psacharopoulos e Arriagada (1989) e Mare (1979 e 1981), dão 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 sócio-econômicas. 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 enfocam os fatores familiares nas análises dos determinantes do desempenho escolar. Poucos estudos brasileiros têm focado os outros dois conjuntos de fatores mencionados acima – escolares e comunitários. Um dos primeiros estudos que tenta incorporar variáveis relacionadas à oferta da educação na análise dos determinantes da escolaridade no Brasil é o de Barros *et alii* (2001). Eles consideraram cinco fatores determinantes do desempenho escolar, quais sejam: recursos familiares, características pessoais e geográficas, qualidade e disponibilidade da oferta do ensino, recursos comunitários e custo de oportunidade do estudo.

Para os recursos familiares, os autores utilizaram as variáveis clássicas, renda per capita da família e o nível de escolaridade dos pais. 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. Para apurar a disponibilidade, eles empregaram a razão entre o número de escolas do ensino médio e a população de 11 a 25 anos e a duração média do percurso casa-escola.

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 nível educacionais maiores podem auxiliar melhor as tarefas escolares das crianças, diminuindo o custo da educação e em comunidades mais ricas, menor nível educacional pode reduzir as chances de emprego e melhores salários.

Os autores utilizam dados da PNAD de 1996 e da Pesquisa de Padrão de Vida (PPV) 1996/97 para explorar suas conjecturas e encontram que os coeficientes das variáveis comunitárias, qualidade da infra-estrutura e de disponibilidade de serviços educacionais foram positivos e significativos. No caso da educação média dos professores, os resultados não foram muito claro, pois considerando a educação média dos professores do ensino fundamental o impacto foi positivo, porém, no caso dos professores do ensino médio, o impacto foi negativo. Porém, as variáveis de *background* familiar apresentaram maior impacto.

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 essas variáveis realmente afetam a educação e não podem ser postas a margem dos estudos educacionais. Uma ressalva que pode ser feita a esse trabalho, é que os autores utilizam as variáveis relacionadas à escola e comunidade no mesmo nível das variáveis intradomiciliares e individuais, estimando os determinantes educacionais através do modelo MQO.

Conforme Lee & Bryk (1989), dados que possuem natureza hierárquica (como dados do Censo Demográfico e PNAD) e que são tratados da forma convencional (MQO) podem produzir resultados imprecisos. Os principais motivos que levariam a esta imprecisão seriam o vício de agregação, independência dos erros e heterogeneidade da regressão.

O primeiro motivo pode ocorrer quando variáveis de níveis de agregação diferentes possuem efeitos distintos sobre a variável resposta. Para exemplificar, pode-se considerar que o efeito do nível sócio-econômico do indivíduo sobre o seu desempenho escolar é diferente do efeito do nível sócio-econômico do município. Os erros podem não ser independentes e com variância não constante quando o meio influencia os resultados educacionais dos indivíduos de forma diferente. Em outras palavras, os resultados educacionais dos estudantes de um mesmo município serão provavelmente influenciados de maneira semelhante pela estrutura e disponibilidade de recursos educacionais do município. Como esses possuem estrutura e recursos educacionais diferentes, os indivíduos de municípios diferentes serão influenciados de maneira distinta.

Quando se ignoram os problemas descritos acima, o efeito da variável agregada pode ser subestimado tendendo a apresentar-se não significativo. Provavelmente muitos estudos, dentre eles o de Barros *et alii* (2001), encontraram um efeito menor das variáveis escolares sobre a educação, em consequência desses problemas.

Uma metodologia que trabalha com a quebra dos pressupostos das regressões de MQO é os modelos hierárquicos. Rios-Neto, César e Riani (2002) utilizou essa metodologia para estudar os determinantes das probabilidades de progressão da 1ª e 5ª séries utilizando variáveis de caráter familiar e de oferta do sistema de ensino, considerando dois níveis de hierarquia. No nível 1, o indivíduo, as variáveis independentes utilizadas foram de *background* familiar (educação da mãe e *dummies* 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 nível 2, que contemplou o ano, UF, situação do domicílio e tipo de área (metropolitana e não-metropolitana), utilizou-se como variáveis independentes, os anos médios de estudo e salário médio dos professores do 1º grau e a razão professor do 1º grau e população em idade escolar. Estas variáveis tentaram captar qualidade e disponibilidade da educação ofertada. Os autores utilizaram dados das PNADs de 1981 a 1999 e encontraram um importante impacto da escolaridade média do professor na probabilidade de progredir na 1ª e 5ª séries.

Pelo que foi exposto, percebe-se que existem poucos trabalhos no Brasil que analisam o impacto das variáveis escolares na educação, e eles apontam para a importância das variáveis relacionadas à oferta educacional. Já na literatura internacional há controvérsia sobre o efeito das variáveis de insumos escolares na educação.

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.

3) Metodologia

3.1) Base de Dados

Foram utilizados os dados do Censo Demográfico de 2000, fornecidos pelo IBGE e o Censo Escolar de 2000, obtido pelo INEP. Do Censo Demográfico foram retirados os dados referentes às características pessoais e familiares do indivíduo. Nesta base de dados, as informações são obtidas no nível do indivíduo, e não há quesitos referentes às características das escolas onde o indivíduo estuda. Os quesitos sobre educação só fornecem informações referentes à série e grau que o indivíduo frequenta ou frequentou.

O Censo Escolar tem informação no nível da escola, sendo possível obter informação sobre as características da escola, tais como os níveis e modalidades de ensino, a caracterização física da escola, número de professores e sua escolaridade, matrícula, repetência, etc. Porém, as informações para os alunos são bem restritas, só possuindo dados referente ao sexo e data de nascimento² para os indivíduos matriculados.

Desta forma, para se ter à informação sobre as características individuais e familiares dos indivíduos e as características da escola foi feita uma junção nas duas bases, utilizando-se como variável de ligação o código do município (de acordo com a codificação do IBGE), que é comum as duas bases.

Contudo, ainda fica uma lacuna, pois no Censo Demográfico não é possível identificar a escola onde o indivíduo estuda e, por outro lado, no Censo Escolar não se tem como identificar os indivíduos que estudam em uma determinada escola, portanto, as variáveis referentes à escola serão agrupadas a nível municipal, como por exemplo, a porcentagem de escolas com quadras de esporte no município, porcentagem de professores com curso superior no município, etc³.

3.2) O modelo Hierárquico

Para estudar os determinantes da probabilidade de estar matriculado na escola e de freqüentar a escola na idade adequada a série foi utilizado o modelo logístico hierárquico. Os modelos logísticos, hierárquicos ou não, modela a probabilidade de um evento ocorrer, sendo sua equação dada por:

$$\ln\left(\frac{\theta_i}{1-\theta_i}\right) = \beta_0 + \beta_{ki} X_{ki} \quad (1)$$

Onde: θ_i é a probabilidade do evento ocorrer e X_{ki} são as variáveis independentes.

Quando se incorpora a modelagem hierárquica no modelo logístico, assume-se a existência de um conjunto de dados hierárquicos⁴, onde a variável dependente é medida no nível mais baixo e as variáveis explicativas em todos os níveis existentes (HOX, JJ, 1995). Nesse trabalho o nível mais baixo contemplou o indivíduo – nível 1 – e o nível mais alto, o

² Mesmo assim, estas duas informações não se cruzam, ou seja, sabe-se o montante de alunos do sexo masculino e feminino em determinada série, mas não se sabe a idade deles e vice-e-versa.

³ As variáveis dependentes serão listadas mais a frente.

⁴ Os dados utilizados nesse trabalho foram retirados do questionário mais amplo do Censo Demográfico que é aplicado à uma amostra. Esta amostra possui uma estrutura hierárquica de coleta porque primeiro é selecionado uma unidade geográfica de análise, conhecida como distrito censitário. Num segundo momento, são escolhidos os domicílios, de maneira aleatória dentro do distrito censitário, e todos os indivíduos destes domicílios são entrevistados.

município – nível 2. A modelagem hierárquica mais simples considera apenas o intercepto como tendo efeito aleatório. Nesse caso, a equação do primeiro nível, fica descrita da seguinte forma:

$$\ln\left(\frac{\theta_{ij}}{1-\theta_{ij}}\right) = \beta_{0j} + \beta_{kij} X_{iji} + e_{ij} \quad (2)$$

E a modelagem do nível 2 se dá da seguinte forma:

$$\beta_{0j} = \gamma_{00} + u_{0j} \quad (3)$$

$$\beta_{kj} = \gamma_{k0} \quad (4)$$

Onde: i indexa o nível 1, nesse caso o indivíduo; j as unidades de nível 2 que são os municípios; o termo β_{0j} é coeficiente do intercepto que contém o termo aleatório u_{0j} , β_{kj} são os coeficientes das k variáveis incluídas no modelo e que só possuem efeito fixo.

Além de considerar o intercepto com efeito aleatório, pode-se também incorporar variáveis explicativas no nível 2 que determinam o coeficiente aleatório, que será chamada por Z_j . Desta forma, no nível 1 tem-se:

$$\ln\left(\frac{\theta_{ij}}{1-\theta_{ij}}\right) = \beta_{0j} + \beta_{kj} X_{kj} \quad (5)$$

E no nível 2:

$$\beta_{0j} = \gamma_{00} + \gamma_{01} Z_j + u_{0j} \quad (6)$$

$$\beta_{k1j} = \gamma_{k10} \quad (7)$$

Para modelar a probabilidade dos indivíduos de 7 a 14 anos estar matriculado na escola foram utilizadas as seguintes categorias para a variável resposta: “1”, caso o indivíduo nessa faixa etária esteja freqüentando a escola e, “0”, caso contrário. Para modelar a probabilidade de estar cursando a série adequada a sua idade considerou-se apenas os indivíduos de 7 a 14 anos que freqüentam a escola e as categorias para a variável resposta foram: “1”, para o aluno que esteja cursando a série adequada a sua idade e, “0”, caso contrário. Ou seja, a pessoa que tem 10 anos e está na 2ª série terá valor 0, e quem está na 4ª série terá valor 1.

Com relação as variáveis explicativas, estas foram escolhidas de acordo com a literatura sobre os determinantes educacionais vista na seção anterior. Dessa forma, no nível 1 foram utilizadas variáveis de caráter familiar, como educação dos pais, número de filhos totais no domicílio, variáveis categóricas identificando a posição de ocupação do pai e se o

domicílio é chefiado por mulher, sexo e idade. Essas são variáveis normalmente utilizadas nos trabalhos sobre determinantes educacionais e já se comprovou possuírem impactos no desempenho educacional do indivíduo.

A grande questão é com relação à modelagem do nível 2, pois como já foi visto, existem poucos estudos que utilizam variáveis de escola numa modelagem hierárquica. As covariáveis incluídas nesse nível tiveram a finalidade de captar as condições de oferta escolar do município e características municipais que possam influenciar o desempenho escolar dos indivíduos. Portanto, os fatores escolares foram divididos em recursos humanos escolares, quais sejam: tamanho médio da sala de aula, porcentagem de professores com nível superior, salário médio dos professores e média de horas aula diárias; e infra-estrutura da escola cujas variáveis são: porcentagem de escolas no município que possuem bibliotecas, quadras esportivas, laboratório de ciência e laboratório de informática. Todas essas variáveis tiveram a finalidade de captar o efeito da qualidade da escola. Para os fatores comunitários foi incluída a educação média da população adulta, considerada como o capital social do município. Tanto as variáveis de caráter escolar quanto comunitário são agregadas por município.

Cabe ressaltar, que as variáveis independentes do nível 1 foram expressas em termos do desvio em relação à média total da amostra. No caso do modelo logístico hierárquico tal procedimento é vantajoso à medida que facilita a estimativa das probabilidades de frequência à escola e de cursar a série adequada à idade, já que podem ser obtidas a partir apenas do intercepto, facilitando o exercício de simulação de variações em algumas covariáveis que foi feito na seção seguinte. Para as variáveis de nível 2, considerou-se seus valores métricos.

Nas tabelas 1, encontram-se as estatísticas descritivas para as duas amostras. Com relação ao nível 1, pode-se destacar a baixa média de domicílios chefiados por mulher, a homogeneidade com relação a distribuição dos sexos e o baixo percentual de pessoas freqüentando a escola na idade correta.

Para as variáveis de nível 2, destaca-se que mais da metade dos professores do ensino fundamental não possuem curso superior e o baixo percentual de escolas com laboratórios de ciências e informática.

Tabela 1: Descrição das variáveis de nível 1 e 2 para as diferentes amostras

Amostra nível 1: filhos de 7 a 14 anos	N	MÉDIA	D. PADRÃO	MÍNIMO	MÁXIMO
Idade	296515	10,53	2,28	7,00	14,00
Dummy sexo; 0=homem e 1=mulher	296515	0,49	0,50	0,00	1,00
Frequenta escola	296515	0,96	0,20	0,00	1,00
Educação da mãe	296515	5,08	3,90	0,00	17,00
Domicílio chefiado por mulher	296515	0,03	0,18	0,00	1,00
Educação do pai	296515	4,70	3,89	0,00	17,00
Pai empregado com carteira	296515	0,30	0,46	0,00	1,00
Pai empregado no setor público	296515	0,04	0,19	0,00	1,00
Pai empregador	296515	0,03	0,18	0,00	1,00
Pai conta própria	296515	0,26	0,44	0,00	1,00
Pai fora da ocupação	296515	0,18	0,38	0,00	1,00
Pai empregado sem carteira ou doméstico	296515	0,20	0,40	0,00	1,00
Total de filhos no domicílio	296515	3,42	1,80	1,00	16,00
Amostra nível 1: filhos de 7 a 14 anos que frequentam a escola					
Idade	283564	10,49	2,26	7,00	14,00
Dummy sexo; 0=homem e 1=mulher	283564	0,49	0,50	0,00	1,00
Frequenta escola na idade certa	283564	0,50	0,50	0,00	1,00
Educação da mãe	283564	5,19	3,91	0,00	17,00
Domicílio chefiado por mulher	283564	0,03	0,18	0,00	1,00
Educação do pai	283564	4,81	3,91	0,00	17,00
Pai empregado com carteira	283564	0,30	0,46	0,00	1,00
Pai empregado no setor público	283564	0,04	0,19	0,00	1,00
Pai empregador	283564	0,03	0,18	0,00	1,00
Pai conta própria	283564	0,26	0,44	0,00	1,00
Pai fora da ocupação	283564	0,18	0,38	0,00	1,00
Pai empregado sem carteira ou doméstico	283564	0,19	0,39	0,00	1,00
Total de filhos no domicílio	283564	3,38	1,77	1,00	16,00
Amostra nível 2: municípios					
Recursos Humanos Escolares					
Média de alunos por turma	853	31,38	5,40	18,47	63,56
Porcentagem de professores com curso superior	853	45,02	21,19	0,00	94,70
Média de horas-aulas diárias	853	4,39	0,10	4,00	5,40
Salário médio dos professores	853	439,76	142,49	151,00	1488,79
Capital Social					
Educação média da população adulta	853	4,50	1,01	1,87	8,21
Infra-estrutura da escola					
Proporção de escolas com biblioteca	853	0,37	0,26	0,00	1,00
Proporção de escolas com quadra de esportes	853	0,23	0,24	0,00	1,00
Proporção de escolas com laboratório de informática	853	0,08	0,11	0,00	0,90
Proporção de escolas com laboratório de ciências	853	0,09	0,14	0,00	1,00

4) Análise dos resultados

4.1) Frequência à Escola

Na tabela 2 são apresentados os resultados das regressões para a modelagem da probabilidade de frequentar a escola. Foram estimados três modelos de regressão. O modelo 1 considerou o intercepto como tendo efeito aleatório (equações 2 a 4), enquanto que no modelo 2 e 3 foram incluídas covariáveis de nível 2 para determinar os

coeficientes aleatórios do intercepto (equações 5 a 7), sendo que no último modelo a educação média da população adulta não entrou por possuir multicolinearidade com a porcentagem de professores com nível superior, o que pode enviesar os resultados.

A análise do efeito aleatório de modelo 1 (parte inferior da tabela), mostra que é aceitável a hipótese do intercepto ser considerado com tendo efeito aleatório, já que ele foi significativo. Em outras palavras, é aceitável que os municípios apresentem médias diferentes para a frequência à escola. Nesse modelo mais simples, o componente da variância estimado para o intercepto (0,3865) é denominado variância não-condicional. À medida que são incluídas variáveis no nível 2, a variância passa a ser condicional e pode-se calcular a proporção da variância que é explicada pelas variáveis de nível 2⁵. Desta forma, percebe-se que no segundo modelo, as variáveis de nível 2 explicam 53% do componente aleatório do intercepto, enquanto que no modelo 3, 36%.

Partindo para a análise dos efeitos fixos, percebe-se que dentre as variáveis de nível 1, apenas a *dummy* domicílio chefiado por mulher não apresentou-se estatisticamente significativa, enquanto que todas as demais mostraram-se significativas e com sinal esperado. As variáveis relacionadas aos atributos individuais, idade e sexo, mostram que a idade possui uma relação negativa com a frequência à escola e que as mulheres possuem uma probabilidade maior de frequentar a escola. Com relação à educação dos pais, fica comprovada a sua importância na frequência à escola, sendo que a educação materna possui impacto um pouco maior, corroborando com o trabalho de Barros e Lam, 1996. Para as variáveis categóricas de posição de ocupação do pai, já que a categoria omitida é pai empregado sem carteira assinada ou doméstico, é de se esperar o sinal positivos de todas as demais posições, ressaltando que aquelas que possuem maior *status* socioeconômico apresentam maior impacto. Por fim, o número de filhos no domicílio apresenta efeito negativo na frequência à escola, confirmando a existência do *trade off* entre quantidade e qualidade dos filhos em Minas Gerais.

Analisando o segundo modelo, que considera todas as variáveis do nível 2, percebe-se que os recursos humanos escolares não se mostraram importantes para probabilidade dos filhos de 7 a 14 anos frequentarem a escola, já que quase todos não foram estatisticamente significativos, com exceção da porcentagem de professores com

⁵ Para calcular a porcentagem da variância explicada, utiliza-se a seguinte fórmula:

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

curso superior, que apesar de ter apresentado significância seu sinal foi contrário ao esperado.

Com relação à infra-estrutura da escola, os resultados do modelo 2 mostraram que a porcentagem de escolas com biblioteca, quadras e laboratórios de ciências possuem impacto positivo e significativo⁶ para a frequência à escola.

A educação média da população adulta apresentou-se significativa e com sinal positivo, demonstrando que um melhor estoque de capital humano pode aumentar os ganhos de escala da educação influenciando de maneira positiva a probabilidade de frequentar a escola.

Mesmo com a educação média da população adulta sendo estatisticamente significativa, ela não foi considerada no modelo 3 por apresentar multicolinearidade com as demais variáveis. Os resultados desse modelo diferem bastante do anterior, principalmente com relação as variáveis de recursos humanos escolares. O salário médio do professor e o número médio de alunos por turma passaram a apresentar significância e a porcentagem de professores com curso superior apresentou o sinal esperado. As variáveis de infra-estrutura da escola se mostraram bastantes importantes para a frequência escolar, indicando que uma política de melhoria da infraestrutura das escolas poderá resultar em ganhos escolares.

Com o objetivo de melhor visualizar o impacto das variáveis escolares na probabilidade de frequentar a escola, foi feito um exercício de simulação considerando variações na média de educação materna e paterna e em algumas variáveis escolares selecionadas, levando em consideração os coeficientes do modelo 3 (gráficos 1 a 5).

⁶ Significativo ao nível de 0,05.

Tabela 2: Regressão para quem frequenta ou não a escola

Efeito Fixo	Modelo 1	Valor-p	Modelo 2	Valor-p	Modelo 3	Valor-p
Intercepto	3,3528	0,000	1,4597	0,084	2,9126	0,002
Recursos Humanos Escolares						
Média de alunos por turma			-0,0028	0,466	-0,0108	0,010
Porcentagem de professores com curso superior			-0,0080	0,000	0,0021	0,094
Média de horas-aulas diárias			0,0526	0,784	0,0310	0,885
Salário médio dos professores			0,0001	0,456	0,0004	0,030
Capital Social						
Educação média da população adulta			0,4143	0,000		
Infra-estrutura da escola						
Porcentagem de escolas com biblioteca			0,2180	0,051	0,4943	0,000
Porcentagem de escolas com quadra de esportes			0,2619	0,029	0,4299	0,002
Porcentagem de escolas com laboratório de informática			0,0329	0,878	0,5748	0,015
Porcentagem de escolas com laboratório de ciências			0,4632	0,017	0,7155	0,001
Educação da Mãe						
idade	0,1376	0,000	0,1398	0,000	0,1388	0,000
sexo; 0=homem e 1=mulher	-0,1830	0,000	-0,1828	0,000	-0,1830	0,000
Domicílio chefiado por mulher	0,1115	0,000	0,1115	0,000	0,1116	0,000
Educação do pai	-0,0657	0,237	-0,0696	0,212	-0,0681	0,221
Pai empregado com carteira	0,1055	0,000	0,1092	0,000	0,1074	0,000
Pai empregado no setor público	0,2911	0,000	0,2922	0,000	0,2909	0,000
Pai empregador	0,5196	0,000	0,5090	0,000	0,5114	0,000
Pai empregador	0,5840	0,000	0,6256	0,000	0,6126	0,000
Pai conta própria	0,1992	0,000	0,1960	0,000	0,1982	0,000
Pai fora da ocupação	0,0920	0,001	0,0933	0,001	0,0938	0,001
Total de filhos no domicílio	-0,0764	0,000	-0,0747	0,000	-0,0750	0,000
EFEITO ALEATÓRIO						
intercepto	0,3865	0,0000	0,1825	0,000	0,2489	0,000
Porcentagem da variância explicada			52,77		35,60	

Gráfico 1: Impacto da educação dos pais e média de alunos por turma na probabilidade de frequentar à escola.

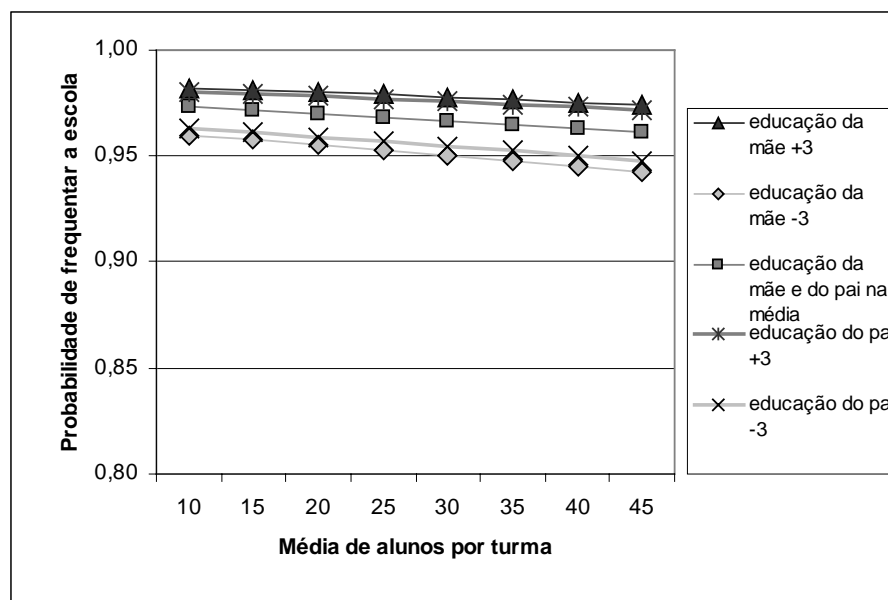


Gráfico 2: Impacto da educação dos pais e da porcentagem de escolas com bibliotecas na probabilidade de freqüentar a escola.

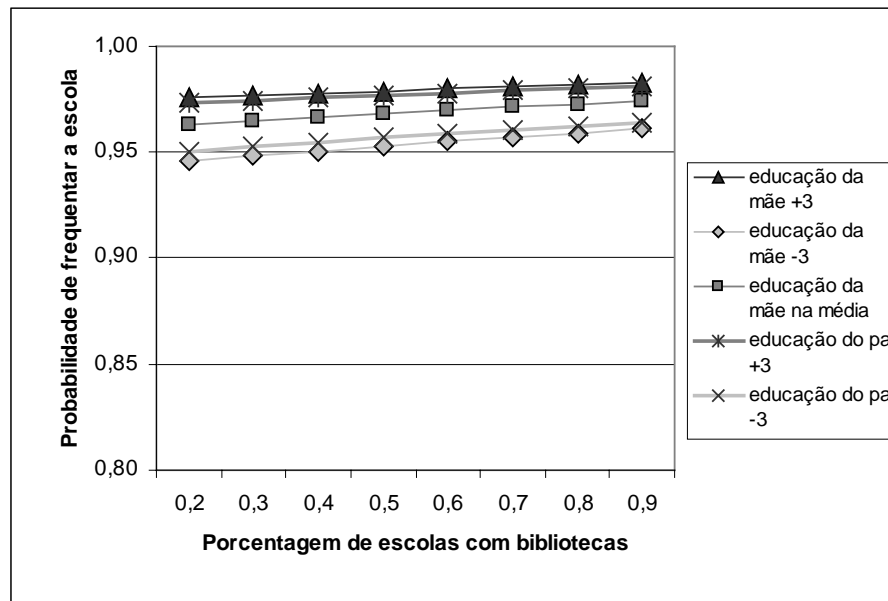


Gráfico 3: Impacto da educação dos pais e da porcentagem de escolas com quadras na probabilidade de freqüentar a escola.

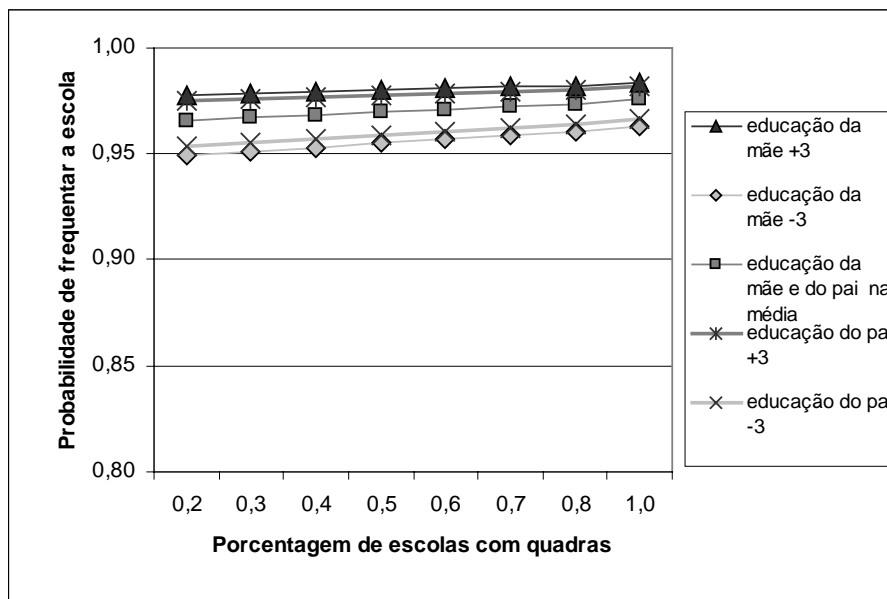


Gráfico 4: Impacto da educação dos pais e da porcentagem de escolas com laboratório de informática na probabilidade de freqüentar a escola.

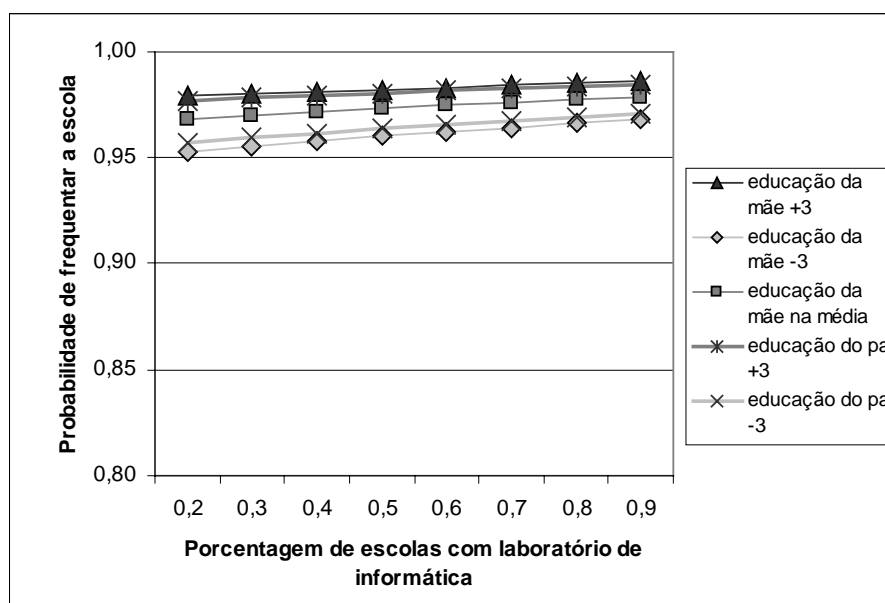
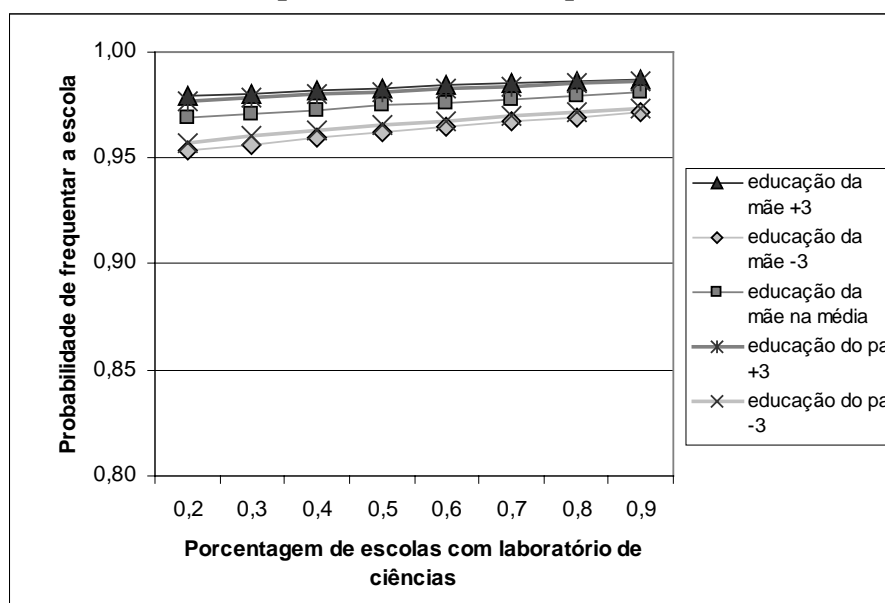


Gráfico 5: Impacto da educação dos pais e da porcentagem de escolas com laboratório de ciências na probabilidade de freqüentar a escola.



4.2) Freqüência à escola na idade adequada

Na tabela 3 encontram-se os resultados das regressões para as pessoas que freqüentam escola na idade correta, percebe-se que nesse caso também é aceitável a hipótese de considerar o intercepto com efeito aleatório. Quando se acrescenta as covariáveis de nível 2 (modelo 2 e 3) a porcentagem da variância explicada do intercepto é respectivamente de 53% e 40%.

Analisando o efeito fixo das variáveis de nível 1, verifica-se que a educação dos pais e as variáveis categóricas de ocupação do pai, apresentarem coeficientes significativos e com o sinal esperado, porém, seus impactos foram menores nesse caso do que para a frequência à escola. Já para a chefia feminina e o total de filhos do domicílio ocorre o inverso.

No modelo 2, dentre os recursos humanos escolares, a média de alunos por turma e a porcentagem de professores com curso superior apresentaram significativos e com o sinal esperado. Já para a infra-estrutura da escola apenas a porcentagem de escolas com quadras se apresentou estatisticamente significativo e com o sinal esperado, enquanto que a porcentagem de escolas com laboratório de informática e ciências tiveram resultados contraditórios, ou seja, impacto negativo na probabilidade de frequentar a escola na idade adequada.

No modelo 3, a média de alunos por turma, a porcentagem de professores com curso superior, a porcentagem de escolas com bibliotecas e quadras passam a ter maior impacto na frequência à escola na idade correta.

Nos gráficos 6 a 10 são apresentadas as simulações do impacto da variação da educação dos pais e de algumas variáveis relacionadas à escola na probabilidade de frequentar a escola na idade adequada.

Tabela 3: Regressão para quem freqüenta a escola na idade adequada

	Modelo 1	Valor-p	Modelo 2	Valor-p	Modelo 3	Valor-p
Intercepto	-0,1394	0,000	-0,5317	0,339	0,3431	0,574
Recursos Humanos Escolares						
Média de alunos por turma			-0,0074	0,004	-0,0123	0,000
Porcentagem de professores com curso superior			0,0037	0,000	0,0100	0,000
Média de horas-aulas diárias			-0,1433	0,255	-0,1413	0,311
Salário médio dos professores			-0,0002	0,020	-0,0001	0,529
Capital Social						
Educação média da população adulta			0,2679	0,000		
Infra-estrutura da escola						
Porcentagem de escolas com biblioteca			-0,0037	0,958	0,1727	0,023
Porcentagem de escolas com quadra de esportes			0,1672	0,024	0,2686	0,001
Porcentagem de escolas com laboratório de informática			-0,4574	0,001	-0,1373	0,341
Porcentagem de escolas com laboratório de ciências			-0,2069	0,074	-0,0815	0,522
Educação da Mãe						
idade	0,0745	0,000	0,0744	0,000	0,0745	0,000
sexo; 0=homem e 1=mulher	-0,1860	0,000	-0,1861	0,000	-0,1861	0,000
Domicílio chefiado por mulher	0,3002	0,000	0,3003	0,000	0,3003	0,000
Educação do pai	-0,1757	0,000	-0,1757	0,000	-0,1757	0,000
Pai empregado com carteira	0,0492	0,000	0,0492	0,000	0,0492	0,000
Pai empregado no setor público	0,1096	0,000	0,1095	0,000	0,1095	0,000
Pai empregador	0,1322	0,000	0,1319	0,000	0,1320	0,000
Pai empregador	0,2017	0,000	0,2019	0,000	0,2019	0,000
Pai conta própria	0,1543	0,000	0,1541	0,000	0,1542	0,000
Pai fora da ocupação	0,0365	0,008	0,0363	0,009	0,0364	0,009
Total de filhos no domicílio	-0,1200	0,000	-0,1201	0,000	-0,1201	0,000
EFEITO ALEATÓRIO						
intercepto	0,2244	0,000	0,1052	0,000	0,1350	0,000
Porcentagem da variância explicada			53,10		39,83	

Gráfico 6: Impacto da educação dos pais e da média de alunos por turma na probabilidade de freqüentar a escola na idade adequada.

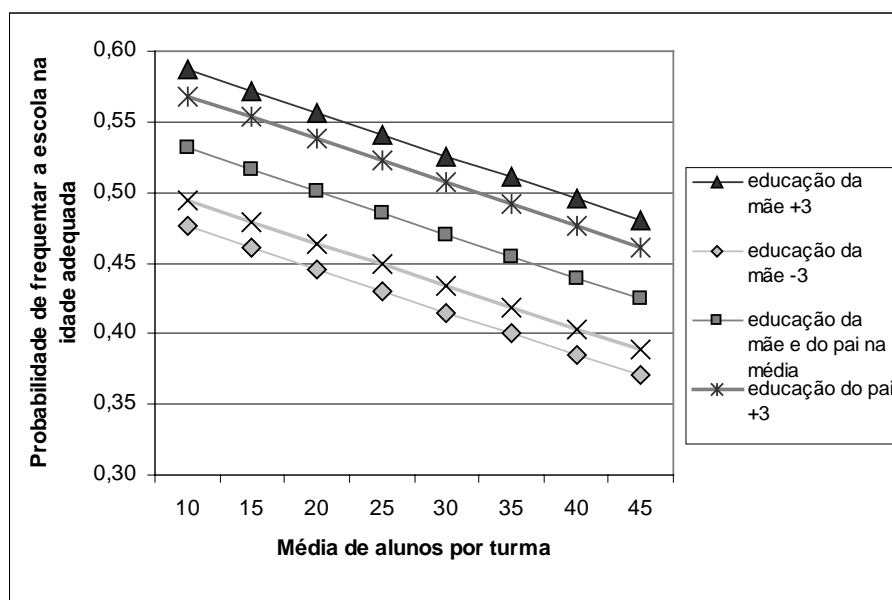


Gráfico 7: Impacto da educação dos pais e da média de alunos por turma na probabilidade de freqüentar a escola na idade adequada.

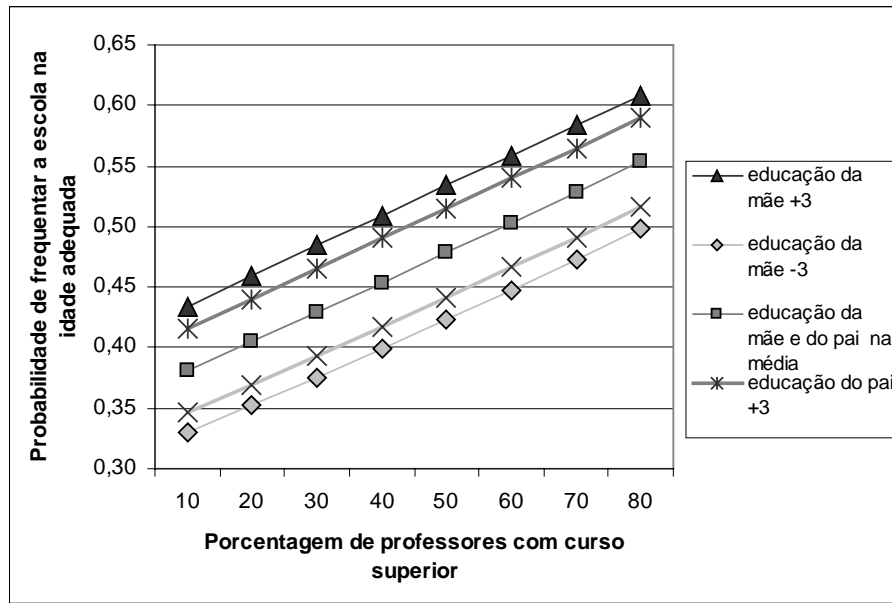


Gráfico 8: Impacto da educação dos pais e da porcentagem de professores com curso superior na probabilidade de freqüentar a escola na idade adequada.

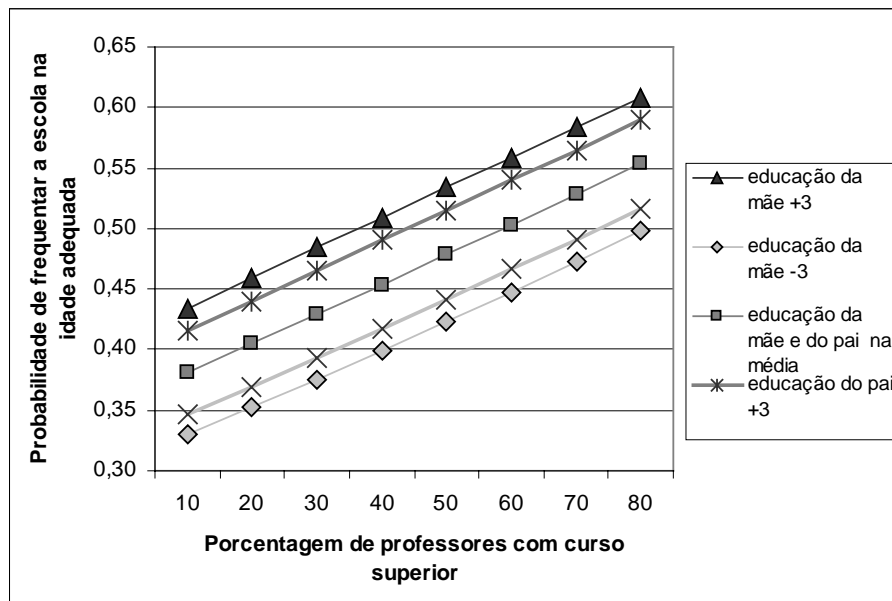


Gráfico 9: Impacto da educação dos pais e da porcentagem de escolas com bibliotecas na probabilidade de frequentar a escola na idade adequada.

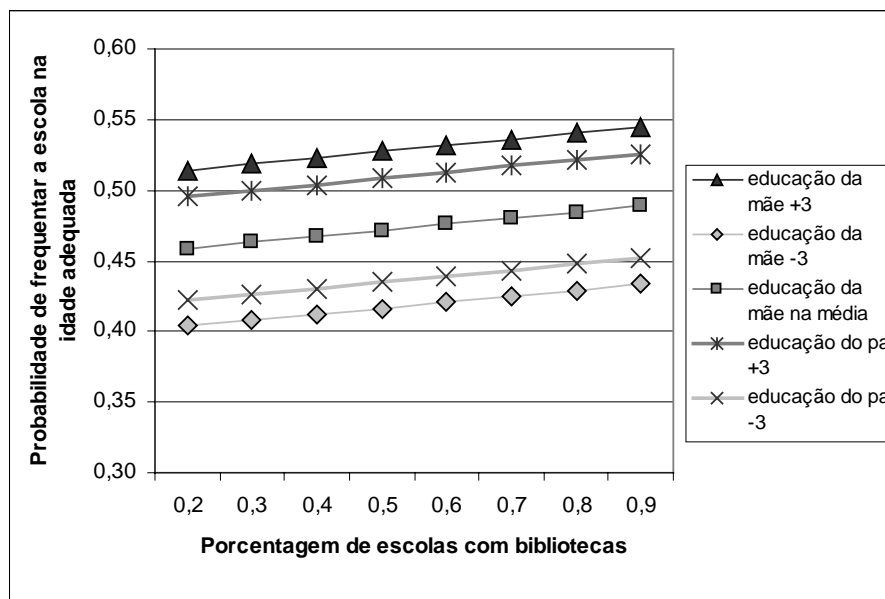
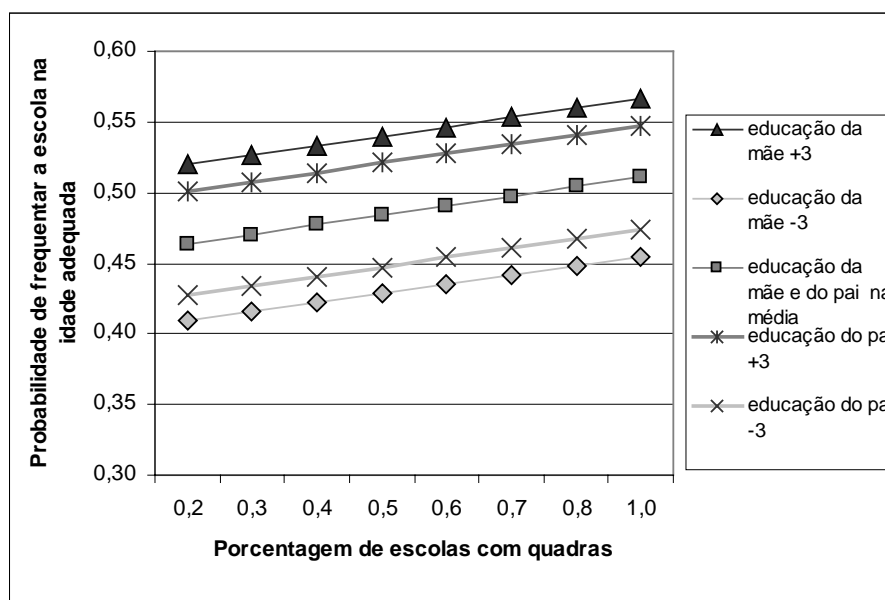


Gráfico 10: Impacto da educação dos pais e da porcentagem de escolas com quadras na probabilidade de frequentar a escola na idade adequada.



5) Conclusão

O presente trabalho objetivou analisar como os fatores familiares, escolares e comunitários afetam a probabilidade dos indivíduos de 7 a 14 anos frequentar a escola e frequentar a escola na idade adequada a série nos municípios de Minas Gerais.

O conjunto de fatores familiares considerados foi: anos de estudo do pai e da mãe, posição ocupacional do pai, chefia feminina e número de filhos no domicílio. Os fatores escolares que entraram na análise foram: porcentagem de professores com curso superior, salário médio dos professores, número médio de alunos por turma, média de horas-aula diária, porcentagem de escolas dentro do município com bibliotecas, quadras de esporte, laboratório de informática e de ciências. Para os fatores comunitários utilizou os anos médios de estudo da população adulta.

Toda a análise foi feita utilizando o modelo logístico hierárquico, que além de permitir trabalhar conjuntamente com variáveis relacionadas à família do indivíduo, aos fatores escolares e aos comunitários, trabalha com a quebra de pressupostos dos modelos MQO que advém da estrutura hierárquica dos dados.

Os resultados demonstraram que os fatores familiares são importantes determinantes das duas probabilidades estudadas. Sendo que a educação dos pais e a posição na ocupação do pai têm maior impacto na probabilidade de freqüentar a escola. Para a probabilidade de freqüentar a escola na idade correta seus impactos diminuem, enquanto que a chefia feminina e o número de filhos passam a ter maior importância.

Com relação aos fatores relacionados à escola, eles também se apresentaram importantes para a determinação da quantidade e qualidade educacional. As variáveis relacionadas à infra-estrutura da escola tiveram maior impacto na probabilidade de freqüentar a escola (quantidade), enquanto que para a probabilidade de freqüentar a escola na idade correta (qualidade) a porcentagem de professores com curso superior apresentou maior impacto.

A educação média do município (fator comunitário) mostrou-se importante para ambas as probabilidades, sugerindo que o estoque de capital humano traz ganhos de escala para a educação. Entretanto, como ela possui correlação com outras variáveis que entraram no modelo, principalmente a porcentagem de professores com curso superior, também foi feita uma análise sem considerá-la.

Esses resultados sugerem que políticas voltadas para a melhoria dos insumos escolares, como melhor qualificação dos professores e infra-estrutura das escolas, são importantes para o desenvolvimento educacional, principalmente se focalizada em regiões onde a estrutura escolar e os indicadores educacionais são mais precários.

6) Bibliografia

- BARROS, R. P. e LAM, D. Income and educational inequality and children's schooling attainment. In Nancy Birdsall e Richard Sabot (edt.). **Opportunity forgone: education in Brazil**. 1996 Inter-American development Bank. The Johns Hopkins University, Washington.
- BIDWELL, E. e KASARDA, D. (1980). "Conceptualizing and Measuring the effects of school and schooling", **Journal of Education**. Pp. 401-430.
- BUCHMANN C. & HANNUM, E. (2001). "Education and stratification in developing countries: a review of theories and research", **Annual review of sociology**. 27:77-102.
- CARD, D. & KRUEGER, A. B. (1996) "School resources and student outcomes: an overview of the literature and new evidence from North and South Carolina." **Journal of Economic perspectives**. Volume 10, number 4, pp.31-50.
- GREENE, W. (1993). **Econometric analysis**. Macmillan, 2nd ed.,
- HANUSHEK, E. A. (1998). **The evidence on class size. Occasional**. Paper Number 98-1, W. Allen Wallis Institute os political Economy. University of Rochester. February.
- HOX, J. J. (1995). **Applied Multilevel Analysis**. Chap. 2 and 4. TT-Publikaties, Amsterdam. Instituto Nacional de Estudos e Pesquisas Educacionais – Inep. **Geografia da Educação Brasileira**. Brasília: Inep, 2000.
- KRUEGER, A. B. (1999). "Experimental estimates of education production functions". **The Quartely Journal of Economics**. May. 1999. pp. 497-532.
- LAZEAR, E. P. (1999). "Education Production". Working paper 7349, **National Bureau of Economic Research**.
- LEE, J. & BARRO. R. J. (1997). Schooling quality in a cross section of countries. Working paper 6198, **National Bureau of Economic Research**.
- MARTELETO, L. (2001). "Children's schooling in Brazil: do number and composition of siblings matter?". Paper prepared for presentation at the **General Conference of the International Union for the Scientific Study of Population**, Salvador, Brazil, August.
- MELLO E SOUZA, A. e SILVA, N. V. Family background, Quality of Education and Public and Private Schools: Effects on school Transitions. In Nancy Birdsall e Richard Sabot (edt.). **Opportunity forgone: education in Brazil**. 1996 Inter-American development Bank.
- RIOS-NETO, E. L. G., CÉSAR, C. C. e RIANI, J.L.R. (2002) "**Estratificação educacional e progressão escolar por série no Brasil**". Artigo apresentado no encontro Determinantes do Sucesso Educacional. IPEA, 27-28 maio.