

# O Impacto dos Fatores Familiares Sobre a Defasagem Idade-Série de Crianças no Brasil\*

Danielle Carusi Machado<sup>†</sup>, Gustavo Gonzaga<sup>‡</sup>

**Sumário:** 1. Introdução; 2. Resenha da literatura; 3. Breves considerações sobre a transmissão ; 4. Estratégia econométrica: a escolha dos instrumentos; 5. Resultados; 6. Considerações finais; A. Base de dados e conceitos utilizados.

**Palavras-chave:** Escolaridade; atraso educacional; defasagem idade-série.

**Códigos JEL:** I21; J13; J18; J24.

Analisamos o efeito da renda e da educação dos pais sobre a probabilidade de as crianças terem defasagem idade-série usando a PNAD 1996. Com a adoção de hipóteses sobre os vínculos existentes entre gerações, controlamos para a existência de fatores não observados que afetam a formação da renda dos pais e as decisões referentes à escolaridade das crianças (viés de simultaneidade) ou que são passados de uma geração à outra (viés de hereditariedade). Usamos três instrumentos: oferta educacional dos pais; fatores familiares entre as gerações de pais e avós; e a mudança educacional de 1971.

*We study the impacts of family income and parental education on the probability of children's schooling delay using the 1996 PNAD. With the adoption of some hypotheses about the links between generations of children, parents and grandparents, we control for the existence of non observable factors that simultaneously affect the income formation of parents and decisions concerning children schooling (simultaneity bias) or that are transmitted from one generation to another (hereditary bias). We work with three instruments: number of schools by the time parents were children; family factor changes across generations (parents/grandparents); and the 1971 change of the educational system.*

\*Este artigo é uma versão revista de uma parte da tese de doutorado de Machado (2005). Gostaríamos de agradecer Thierry Magnac e aos membros da banca, Sérgio Firpo, André Portela, Eduardo Rios Neto, André Urani e aos participantes de seminários/workshops no Departamento de Economia da PUC-Rio, no IPEA e no Encontro Nacional da Sociedade Brasileira de Econometria. Gostaríamos também de agradecer as sugestões e comentários de um parecerista anônimo. Os erros remanescentes são de inteira responsabilidade dos autores.

<sup>†</sup>E-mail: daniellecarusi@vm.uff.br Faculdade de Economia. Universidade Federal Fluminense – UFF. End: Rua Tiradentes, 17 - Ingá. 24.210-510 Niterói, RJ. Tel : 00 55 21 26299724

<sup>‡</sup>E-mail: gonzaga@econ.puc-rio.br Departamento de Economia. Pontifícia Universidade Católica do Rio de Janeiro – PUC/Rio. End: Rua Marquês de São Vicente, 225, - Gávea. 22453-900 Rio de Janeiro, RJ. Tel: 00 55 21 35271078



## 1. INTRODUÇÃO

Apesar do aumento da inserção da maioria das crianças brasileiras na escola, muitas não progredem ao longo do sistema educacional de forma contínua e adequada (Gomes-Neto e Hanushek, 1994, Duryea, 1998, Leon e Menezes-Filho, 2002, Barros e Mendonça, 1998, Ribeiro, 1991, Fletcher, 1985). São comuns os seguintes problemas: (i) entrada tardia na escola – crianças com 7 e 8 anos de idade que ainda não iniciaram o ciclo fundamental; (ii) repetências e (iii) abandono da escola.<sup>1</sup>

O progresso educacional da criança pode ser influenciado por características relacionadas à escola e ao *background familiar* (educação e renda familiar), bem como à complementaridade entre esses fatores. O ambiente familiar e as características dos pais têm impacto nas decisões relacionadas às crianças e jovens. Jovens que vivem em famílias cujo “*background*” é melhor (renda mais alta e nível educacional superior) têm melhores condições de permanecer por um período mais longo na escola. Nestes casos, o custo de permanência nesta etapa de estudante não seria restritivo.<sup>2</sup> Além disso, o status sócio-econômico e educacional está usualmente positivamente correlacionado entre as gerações.

No Brasil, há uma vasta literatura empírica sobre as relações existentes entre escolaridade das crianças e aspectos familiares como instrução dos pais e nível de renda. Uma parte aborda a mobilidade em termos educacionais (Barros e Lam, 1993, Ferreira e Veloso, 2003, Marteleto, 2004). Outra vertente estuda a relação entre desigualdade de oportunidades e desigualdade de renda (Bourguignon et alii, 2003, Menezes-Filho, 2001). Já uma terceira linha enfatiza a relação entre trabalho infantil, pobreza e escolaridade (Emerson e Souza, 2003, 2002, 2006, Kassouf, 2001, Barros et alii, 1996, Menezes-Filho e alii, 2000).

Neste artigo, analisamos a influência dos fatores familiares na escolaridade de uma amostra de crianças que ainda não finalizaram o ciclo educacional com base nos dados da Pesquisa Nacional por Amostra de Domicílios de 1996 (PNAD). Não abordamos o nível de escolaridade atingido, mas uma medida que capta aspectos ao longo do processo escolar: a defasagem idade-série.<sup>3</sup>

O objetivo é mostrar o efeito da renda familiar e do nível educacional dos pais sobre a probabilidade de as crianças terem defasagem idade-série. Com a adoção de algumas hipóteses sobre os vínculos existentes entre as gerações de crianças, pais e avós, controlamos para a existência de fatores não observados que afetam simultaneamente a formação da renda dos pais e as decisões referentes à escolaridade das crianças (viés de simultaneidade) ou que são passados de uma geração à outra da mesma família (viés de hereditariedade). Usando o procedimento de variáveis instrumentais (IV) (Angrist e Krueger, 2001), selecionamos três grupos de instrumentos: a oferta educacional da segunda geração (os pais das crianças),<sup>4</sup> os fatores familiares entre as gerações<sup>5</sup> (explorando as informações da terceira e da segunda geração, ou seja, pais e avós, conforme sugerido em Maurin, 2002, Cogneau e Maurin, 2001, Couralet, 2002) e a mudança no sistema educacional obrigatório de 1971, seguindo Black et alii, 2005, Chevalier, 2004, Chevalier et alii, 2005.

Mostramos que a renda familiar per capita e o nível educacional dos pais têm efeito negativo na probabilidade da criança ter defasagem idade-série. O efeito da renda estimado usando variáveis instrumentais é mais forte que os das regressões padrões, ocorrendo o inverso para o efeito dos níveis

<sup>1</sup> Leon e Menezes-Filho (2002) mostram que a repetição é um dos principais fatores que explica a evasão. A proporção de alunos que interrompe os estudos após experiências de repetência é maior que daqueles que saem da escola sem terem repetido. Ver Bonelli e Veiga (2004) para uma descrição da evasão escolar de jovens no Brasil.

<sup>2</sup> Haveman e Wolfe (1995) fazem uma resenha sobre o impacto da família nas crianças (escolaridade, renda, etc).

<sup>3</sup> Segundo Horowitz e Souza (2004), como o processo de acumulação de capital humano das crianças não está completo, a taxa de progressão pode ser uma boa medida para captar o aprendizado. A literatura educacional estabelece que quanto maior o atraso educacional menor o nível de escolaridade atingido.

<sup>4</sup> Alguns autores também seguem essa linha, como Emerson e Souza (2006) e Currie e Moretti (2003).

<sup>5</sup> No questionário da PNAD/IBGE de 1996, há um suplemento de mobilidade social que pergunta a todos moradores chefes e cônjuges com mais de 15 anos informações sobre a ocupação do pai durante sua adolescência (quando tinha 15 anos de idade). Existem perguntas sobre a escolaridade de seus pais (mãe e pai).

educacionais. Numa estimação padrão, via probit ou mínimos quadrados ordinários, o viés da escolaridade dos pais é para cima enquanto o da renda familiar é para baixo. Esse resultado sugere que aspectos familiares que passam de uma geração para outra são importantes na transmissão de conhecimentos de pais para filhos e que o processo de formação da renda familiar e de escolarização das crianças estão fortemente interligados. Logo, a implementação de políticas que visam à ampliação da escolaridade também perpassa pelo entendimento dos mecanismos de transmissão da educação entre as gerações.

O artigo está organizado da seguinte forma: na próxima seção, apresentamos uma breve resenha sobre os principais artigos empíricos produzidos nesta área, focando particularmente nas estratégias econométricas adotadas para separar os efeitos renda familiar e educação dos pais; na seção 3, descrevemos o modelo que fundamenta a existência dos vínculos intergeracionais na transmissão de capital humano e os principais canais de endogeneidade; na seção 4, apresentamos a estratégia econométrica; na seção 5, os resultados; e, na seção 6, as considerações finais.

## 2. RESENHA DA LITERATURA

Vários artigos examinam diferentes fenômenos relacionados à escolaridade do público infantil e suas características familiares no Brasil (Psacharopoulos e Arriagada, 1989, Duryea, 1998, Saha, 2004, Barros e Lam, 1993). A literatura de demanda educacional também está fortemente atrelada às análises de trabalho infantil.<sup>6</sup> A principal questão é investigar as razões pelas quais os pais preferem que as crianças trabalhem ao invés de estudarem (Psacharopoulos e Arriagada, 1989, Kassouf, 2001, Barros et alii, 1996, Menezes-Filho e alii, 2000).

Não somente para o Brasil, mas também para outros países em desenvolvimento, existem vários fatos estilizados mostrando que crianças de famílias pobres cujos pais têm menor grau de instrução passam por maiores dificuldades ao longo da infância e quando adultas (Haveman e Wolfe, 1995). Os acréscimos na renda familiar poderiam, por um lado, ter um papel não desprezível no acúmulo de capital humano se usados na obtenção de bens facilitadores do aprendizado escolar (compra de livros, cadernos, etc) e se diminuíssem a oferta de trabalho infantil (efeito renda).

Por outro lado, acréscimos na renda dos pais poderiam não ser suficientes para que as crianças de famílias mais pobres alcançassem o nível educacional dos filhos dos mais ricos.<sup>7</sup> As crianças de famílias mais ricas estão usualmente inseridas em um contexto sócio-econômico e cultural favorável ao acúmulo de capital humano, que muitas vezes lhes proporcionam acesso a melhores escolas próximas ao local de moradia e contato com pessoas instruídas no ambiente familiar. Essas facilidades podem ser transmitidas de pais para filhos.

Com relação aos efeitos do grau de escolaridade dos pais sobre o aprendizado de seus filhos, podem ser diretos ou indiretos. Segundo Currie e Moretti (2003), existem quatro potenciais canais dessa influência: (i) como mães/pais mais escolarizados têm rendimentos do trabalho mais altos, podem adquirir mais bens para o aprendizado de seus filhos, conforme já discutido acima; (ii) uma mãe/pai mais escolarizada/o geralmente tem como parceiro/a alguém com nível de escolaridade similar, o que potencializa o efeito da escolaridade dos pais; (iii) pais ou mães mais educados tendem a ter um padrão de comportamento em que nutrem mais expectativas em relação à escolaridade dos filhos, (Marteletto, 2004); e (iv) pais mais educados tendem a possuir menos filhos, (Hanushek, 1992). Além desses quatro aspectos, destaca-se que pais mais instruídos têm mais contato com o cenário social e cultural, tornando-os potencialmente mais aptos para transmitir conhecimentos.<sup>8</sup>

<sup>6</sup>Ver, por exemplo, Emerson e Souza (2006), Kassouf (2001), Barros et alii (1996). Na literatura internacional, ver Basu (1999) e Basu e Tzannatos (2003).

<sup>7</sup>Vários artigos discutem a relação entre renda e escolaridade: Blau (1999), Shea (2000), Goux e Maurin (2005), Wolfe (1982).

<sup>8</sup>Esses pais são mais produtivos no processo de aprendizado, o chamado efeito "criação", conforme Chevalier et alii (2005). Ver também Black et alii (2005).



Se, de fato, a transmissão de conhecimentos de uma geração para outra decorre mais de fatores não observados, como uma habilidade ou tradição familiar, a relação entre a escolaridade da criança e dos pais estaria sendo mais explicada por aspectos não controlados pelo formulador de políticas, não refletindo um efeito causal. Para Chevalier et alii (2005), esse canal de transmissão intergeracional seria natural, não sendo influenciado pelo contexto social no qual a criança está inserida. Filhos de pais mais hábeis seriam também mais hábeis.

O principal problema apontado acima é a existência de correlação entre as variáveis de renda familiar e de educação dos pais com fatores não observáveis que afetam a escolaridade da criança. Conforme já descrito na introdução, existem dois tipos de viés: o de simultaneidade, pois fatores não observados podem afetar simultaneamente a formação da renda dos pais e as decisões referentes à escolaridade das suas crianças; e o de hereditariedade, que se refere a características específicas da família que são passadas de uma geração à outra, como por exemplo, o hábito de leitura.

Na literatura internacional, duas estratégias econométricas são usualmente implementadas para lidar com esse problema:<sup>9</sup> (i) encontrar fontes de variações exógenas da renda familiar e do nível educacional dos pais, ou pelo menos, variações exógenas em alguns componentes que não sejam correlacionados com as características não observadas; e (ii) eliminar o efeito das características não observadas (adotando a hipótese de que são fixas entre as gerações) através de diferenças ao longo do tempo e entre membros da mesma família.

No primeiro grupo, face à quase inexistência de experimentos aleatórios,<sup>10</sup> os artigos inserem-se na literatura de avaliação de políticas. Os instrumentos utilizados são, em geral, fundamentados em experimentos institucionais, reformas que afetam de forma exógena o grau de instrução dos pais e a renda familiar.

Carvalho (2000) adota essa estratégia para o caso brasileiro.<sup>11</sup> Analisa a escolaridade das crianças no meio rural, utilizando a reforma da seguridade social dos trabalhadores rurais em 1991 como um experimento institucional para estudar o efeito renda puro sobre a oferta de trabalho dos homens e sobre a participação no mercado de trabalho e na escola das crianças. Os resultados mostram que as meninas de 12 a 14 anos que residem com beneficiários do programa aumentaram sua participação na escola.

Com relação à escolaridade dos pais, os experimentos naturais usados referem-se a modificações no sistema educacional, correlacionados com o grau de instrução, mas não correlacionados com fatores não observados (Black et alii, 2005, Chevalier, 2004, Chevalier et alii, 2005).

Quando não se dispõe de um experimento institucional, deve-se encontrar um componente de variação da renda que não afete a progressão educacional da criança.

Shea (2000), usando dados longitudinais, isola os determinantes observados da renda do pai usando as variações da renda provocadas pelo seu status ocupacional (se sindicalizado ou se operário em uma indústria) e pela perda de emprego devido ao fechamento da empresa. Adota a hipótese de que essas mudanças na renda não estão correlacionadas com a habilidade dos pais, mas sim com os *rents* econômicos gerados no setor industrial e sindicalizado e com a “sorte/azar” da empresa, na qual o pai trabalha,

<sup>9</sup>Uma parte desta literatura incorpora à análise indicadores capazes de refletir aspectos não observáveis, como genéticos, comparando, por exemplo, o efeito da renda dos pais em filhos adotados e biológicos. O impacto da renda dos pais sobre a educação dos filhos é positivo e significativo para os filhos biológicos e não significativo para os adotados. Esses artigos estão sujeitos a críticas quanto ao tamanho da amostra, conforme Shea (2000). Para o caso brasileiro, ver Verona (2004).

<sup>10</sup>Um exemplo desse tipo de experimento é quando uma família é premiada com renda adicional através de uma loteria. Nesse caso, a variação da renda produzida é completamente aleatória e não está correlacionada com fatores não observáveis dos pais e das crianças. Esse argumento pode ser questionável se pais que jogam na loteria são sistematicamente diferentes em algumas características de pais que não jogam.

<sup>11</sup>Duflo (2003) faz uma análise similar na área de saúde. Ela utiliza a expansão não esperada do programa de aposentadorias para os negros da África do Sul nos anos 90 para examinar o impacto dos recursos adicionais dos avós na saúde das crianças. Compara os resultados das crianças que vivem com seus avós com as que não vivem. Suas estimativas sugerem que a disponibilidade monetária adicional tem um impacto positivo no peso e na altura das crianças.

continuar operando. Seu principal resultado é que para pais com pouca instrução, a variação exógena da renda é muito importante para o investimento em capital humano infantil.

Maurin (2002) e Couralet (2002) usam como instrumentos características educacionais e ocupacionais dos pais e dos avós. O primeiro, a partir de hipóteses sobre a transmissão de habilidades educacionais de pais para filhos, encontra que as diferenças de renda entre os pais afetam as transições das crianças ao longo da escola elementar e que a subestimação do efeito da renda nas estimações padrões é ocasionada pela existência do viés de simultaneidade. O segundo artigo foca na probabilidade da criança trabalhar, considerando as decisões de oferta de trabalho dos outros membros da família e as decisões referentes à escolaridade.

Existem artigos que abordam o problema de endogeneidade da educação dos pais e da renda familiar conjuntamente. Dumas e Lambert (2005) analisam a probabilidade da criança ir para a escola e o nível de escolaridade atingido quando adulta. Focam no efeito da educação dos pais e da riqueza familiar. Usam um grupo de instrumentos que caracterizam a infra-estrutura da área de residência dos pais quando eles tinham 10 anos de idade e as condições de saúde de seus pais. Chevalier et alii (2005) usam estratégia econométrica similar a Shea (2000) para identificar o efeito da renda do pai (status sindical do pai) e no caso da escolaridade, adota procedimento similar a Black et alii (2005) (reforma institucional do sistema educacional). Emerson e Souza (2006) investigam se o fato de um adulto ter trabalhado quando criança prejudica a sua aquisição de rendimentos. Utilizam um conjunto de instrumentos que refletem a oferta educacional na infância, sendo, portanto, correlacionados com as decisões escolares das crianças, mas não com os seus ganhos monetários quando adultas.

O segundo grupo de estratégia econométrica citado acima considera que uma das fontes de correlação da renda familiar e do nível educacional dos pais com os fatores não observados presentes na equação de aprendizado da criança é um efeito fixo da família. Esses fatores familiares não observados, fixos ao longo do tempo, podem ser eliminados através da comparação de componentes da mesma dinastia familiar em diferentes períodos, portanto sujeitos a diferentes valores da renda familiar. Neste caso, o resultado escolar  $educ$  da criança  $i$  da família  $f$  poderia ser representado por uma função linear, conforme a eq. (1) abaixo:

$$educ_{if} = Z_{if}\beta + u_f + v_{if} \quad (1)$$

Onde:

- $Z_{if}$  são características exógenas que afetam a escolaridade da criança  $i$  da família  $f$ .
- $u_f$  representa um componente familiar do erro (não observado)
- $v_{if}$  corresponde ao erro aleatório individual da criança  $i$  da família  $f$

Ou seja, os fatores não observados são decompostos em uma parte fixa ( $u_f$ ) e específica à família (como habilidade, por exemplo) e em uma parte particular à criança e aleatória ( $v_{if}$ ). Quando tomamos a diferença entre o nível educacional atingido por pais e avós, eliminamos este efeito fixo da família. Cogneau e Maurin (2001) adotam essa estratégia para identificar o impacto dos recursos de ambos, pai e mãe, e do nível educacional das mães nas decisões escolares das crianças. Usam como instrumento os resultados passados da família – diferenças entre os níveis de educação e de status ocupacional entre pais e avós – que são considerados não correlacionados com os choques transitórios da renda e com os fatores não observados específicos da família.<sup>12</sup>

No Brasil, com relação à renda familiar, alguns trabalhos relatam o problema de endogeneidade, devido à possibilidade de existência de variáveis omitidas (como habilidade) ou erros de medida. Em

<sup>12</sup>Ver também Blau (1999) que adota a seguinte hipótese de identificação: fatores que simultaneamente determinam a renda dos pais e a educação dos filhos não variam temporalmente e dentro das dinastias. São estimados três modelos alternativos de efeitos fixos que controlam por características não observadas: (i) comuns às mães e suas respectivas irmãs, (ii) comuns às mães e (iii) comuns ao domicílio.



alguns artigos, a instrução dos pais é diretamente usada como *proxy* do capital sócio-econômico familiar, estratégia não adequada se tivermos interesse também no impacto puro da educação dos pais e não somente no efeito da renda familiar.

Um exemplo recente que aborda a questão da endogeneidade da renda é Vasconcellos (2005). Essa autora analisa a relação entre a frequência escolar e a renda familiar de crianças entre 7 e 14 anos de idade. Utiliza como instrumento da renda familiar as variações na distribuição de renda no período de 1981 a 1999. Mostra que o efeito renda somente é significativo para o grupo mais pobre, enquanto para os demais não há relação causal entre a renda e a frequência escolar das crianças.

Courealet (2002) aplica a mesma metodologia de Cogneau e Maurin (2001) para estudar a probabilidade das crianças trabalharem no Brasil e em outros dois países. Está preocupado com a endogeneidade da renda da família sobre as decisões da criança referentes a ir para a escola ou trabalhar e sobre as decisões dos outros membros da família. Utiliza uma série de instrumentos, como níveis educacionais dos avós e as diferenças em relação à geração de pais, dentre outros. Esse artigo, apesar de usar uma abordagem econométrica semelhante à nossa, não investiga o impacto da renda familiar e do nível de educação dos pais sobre a defasagem idade-série.

No tocante ao nível educacional dos pais, o efeito da habilidade individual deles sobre o aprendizado das crianças reflete-se na escolaridade destes. Pais mais hábeis contribuem para o desempenho educacional de seus filhos porque são mais escolarizados e auferem mais renda. A variável omitida na nossa equação de aprendizado da criança está relacionada às habilidades herdadas da família. Características intrínsecas a uma dinastia são passadas de uma geração para outra, influenciando tanto a escolaridade dos pais quanto diretamente a escolaridade das crianças.

### 3. BREVES CONSIDERAÇÕES SOBRE A TRANSMISSÃO

#### 3.1. Um modelo empírico<sup>13</sup>

Para ilustrar a dinâmica do acúmulo de capital humano e os vínculos existentes entre as três (filhos, pais e avós) de uma mesma dinastia, descrevemos brevemente o modelo empírico de Maurin (2002) com algumas alterações.<sup>14</sup> A apresentação desse modelo é importante para demonstrar os canais através dos quais a endogeneidade da renda familiar e da escolaridade dos pais se expressa na equação de atraso educacional das crianças. Por meio dele, justificaremos a escolha de um dos instrumentos utilizados na abordagem empírica (diferença de escolaridade entre pais e avós) para controlar para efeitos hereditários não observados entre as gerações.

Dentro de cada dinastia (família), a dinâmica de acúmulo de capital humano é determinada da seguinte forma (para crianças e adultos):

- A escolaridade das crianças da geração  $t + 1$  é produzida a partir dos insumos comprados pelos pais, pelo nível de esforço nos estudos e por outros aspectos herdados da família observados, como o nível educacional dos pais, e não observados, como hábitos familiares de estudo e leitura. A produção educacional é, portanto, função de aspectos escolares, familiares e individuais.<sup>15</sup>

<sup>13</sup>Esta seção se baseia em Maurin (2002) e Cogneau e Maurin (2001). Ressaltamos que nesses artigos, a escolaridade dos pais está somente presente na equação que determina a renda familiar, portanto afetaria a probabilidade de as crianças estarem atrasadas em termos educacionais apenas indiretamente.

<sup>14</sup>Incluimos o impacto direto do grau de instrução dos pais sobre a escolaridade das crianças e enfatizamos a probabilidade delas estarem ou não com defasagem idade-série em um determinado período do ciclo de vida.

<sup>15</sup>Esses aspectos afetam os custos e benefícios da criança prosseguir nos estudos bem como o seu esforço de aprendizado. Para uma criança mais velha, por exemplo, o custo de oportunidade pode ser mais elevado pois as alternativas no mercado de trabalho são maiores.

- Os pais, crianças da geração  $t$  adultos em  $t + 1$ , decidem o volume de recursos alocados para a educação das crianças e para o consumo, bem como outros fatores que afetam a capacidade de aprendizado de seus filhos. Essas escolhas são feitas com base na maximização da função de utilidade da família – na qual um dos argumentos é a educação dos filhos – considerando a restrição orçamentária e a função de produção educacional das crianças.
- Seguindo a mesma dinâmica, os pais acumularam capital humano graças às decisões de seus pais (os avós das crianças), pertencentes à geração de crianças em  $t - 1$  – e de outros fatores, individuais ou escolares. Os gastos em educação dos pais são função da renda dos avós.

Em termos analíticos, podemos escrever as equações (para cada dinastia  $f$ ) que determinam esse processo de aquisição de escolaridade, conforme as equações (2) a (6) abaixo:

$$educ_{1t+1} = \alpha.Y_{2t+1} + \beta.educ_{1t} + \lambda X_{1t+1} + u_{t+1} \quad (2)$$

$$Y_{2t+1} = a.educ_{1t} + \theta.X_{2t+1} + v_{t+1} \quad (3)$$

$$educ_{1t} = \alpha.Y_{2t} + \beta.educ_{1t-1} + \lambda.X_{1t} + u_t \quad (4)$$

$$Y_{2t} = a.educ_{1t-1} + \theta.X_{2t} + v_t \quad (5)$$

$$educ_{1t-1} = \alpha.Y_{2t-1} + \beta.educ_{1t-2} + \lambda X_{1t-1} + u_{t-1} \quad (6)$$

Onde  $educ_{1t+1}$ ,  $educ_{1t}$  e  $educ_{1t-1}$  são as medidas de anos de estudo completo das três gerações de uma mesma dinastia: crianças, pais e avós, respectivamente. O subscrito 1 representa as crianças, enquanto o subscrito 2, os adultos. O aprendizado das crianças de cada geração, descrito nas equações 2, 4 e 6, depende da renda, do volume de aprendizado de seus respectivos pais (definido quando eram crianças), de outras características agrupadas em  $X$  e de fatores não observados presentes no erro  $u$ .

Os fatores aleatórios que afetam a escolaridade das gerações de crianças, pais e avós estão presentes nos termos  $u_{t+1}$ ,  $u_t$  e  $u_{t-1}$ , respectivamente, e podem ser agrupados conforme abaixo:

- Específicos a cada indivíduo, como, por exemplo, habilidade ou gosto pelos estudos. Indivíduos idênticos em todos os aspectos podem ter desempenho educacional diferente simplesmente porque um gosta mais de estudar do que o outro. Um indivíduo também pode ter uma habilidade inata para aprender melhor e mais rápido que outro.
- Específicos à dinastia, pois as escolhas com relação à escolarização são diversas entre as famílias não somente devido às diferenças observadas mas também a fatores aleatórios, que podem relacionar-se à tradição ou preferências familiares (hábito de leitura, por exemplo).
- Específicos à geração, pois ao longo das décadas, as preferências por escolarização e a oferta educacional variam. Existem choques na oferta educacional que são específicos a uma geração, como, por exemplo, o aumento do volume de recursos para construção de escolas de nível médio ou a valorização da escolaridade no mercado de trabalho.

As equações (3) e (5) representam os níveis de renda dos pais e dos avós, respectivamente. A capacidade de obtenção de rendimentos depende do aprendizado adquirido e de outras qualidades não relacionadas à escolaridade (presentes nos termos  $v_{t+1}$  e  $v_t$ ). Esses fatores também podem ser individuais (talento para uma tarefa cuja remuneração é alta), familiares (habilidade em produzir algo, que



passa de pai para filho) ou específicos de uma geração (preferência por lazer aumenta de uma geração para outra).

A estrutura da matriz de variância e covariância dos erros das equações (2) a (6) define, portanto, os vínculos entre as gerações de crianças, pais e avós, e as hipóteses sobre essa estrutura serão importantes para a abordagem empírica.<sup>16</sup>

Tendo explicitado a dinâmica do aprendizado, voltamos ao fenômeno de interesse que é a defasagem idade-série num determinado período  $\tau$  da vida da criança da geração  $t + 1$ .

De acordo com a definição de defasagem idade-série, a criança é considerada atrasada em termos educacionais se não tem o total de anos de estudo completos compatível com a sua idade no início de cada ano letivo. Conforme a legislação educacional brasileira, uma criança que completa sete anos de idade antes do início do ano letivo (1º. de março) deve entrar na 1ª. série do ensino fundamental. Se prosseguir ao longo do sistema continuamente, ou seja, sem repetir ou abandonar o curso, ela não terá defasagem idade-série.

Para as crianças com sete anos de idade exatos no início do ano letivo, estão atrasadas as que não freqüentam a 1ª. série do ensino fundamental.<sup>17</sup> Para as que têm mais de sete anos de idade, podemos definir a defasagem idade-série num determinado período  $\tau$  de sua vida conforme a eq. (7) abaixo:<sup>18</sup>

$$\begin{cases} A_{1,t+1,\tau} = 1 & \text{se } \frac{educ_{1t+1,\tau}}{Idade_{1t+1,\tau-7}} < 1 \\ A_{1,t+1,\tau} = 0 & \text{se } \frac{educ_{1t+1,\tau}}{Idade_{1t+1,\tau-7}} \geq 1 \end{cases} \quad (7)$$

Onde  $educ_{1,t+1,\tau}$  e  $Idade_{1t+1,\tau}$  são: total de anos de estudos e idade completos da criança no início do ano letivo do período  $\tau$ , respectivamente.

Para as crianças que já saíram da escola, a eq. (2) descreve diretamente o total de anos de estudo completos neste período  $\tau$ . Em termos analíticos:  $educ_{1t+1,\tau} = educ_{1,t+1}$ . Para as crianças que ainda freqüentam a escola, a escolaridade final não foi atingida. Ou seja,  $educ_{1,t+1,\tau} < educ_{1,t+1}$ .

Como o total de anos de estudos completos da criança nesse período  $\tau$  é uma parte da sua desejada, os mesmos fatores que influenciam a determinação da escolaridade desejada também afetam o total de anos de estudos acumulados até o início do ano letivo deste período. Assim, a da criança no período  $\tau$  ter defasagem idade-série é também função dos aspectos que influenciam o nível educacional desejado, conforme a eq. (8) abaixo:

$$P(A_{1t+1,\tau} = 1) = P(\alpha.Y_{2t+1} + \beta.educ_{1t} + \lambda_a.X_{a,1t+1} + u_{t+1} < 0) \quad (8)$$

O denominador da equação (7), formado pela idade da criança no dia primeiro de março e pela idade de entrada legal no sistema educacional, foi incorporado a  $X$ . Logo, a equação (8) diferencia-se da equação (2) por descrever a probabilidade da criança no período  $\tau$  ter defasagem idade-série e pelos termos  $\lambda_a$  e  $X_a$ , pois dentre as variáveis explicativas estão a diferença entre a idade completa da criança e a idade legal de entrada na escola, bem como a sua maturidade no ano, medida em meses, no dia 1º de março.

O interesse é identificar o impacto dos fatores familiares [ $\alpha$  e  $\beta$  na eq. (8), acima] sobre a probabilidade da criança da geração  $t + 1$  no período  $\tau$  estar ou não atrasada em termos educacionais e a direção

<sup>16</sup>Na próxima sub-seção aprofundaremos mais essa questão.

<sup>17</sup>A pesquisa utilizada (PNAD de 1996) foi realizada no mês de setembro. Desta forma, o atraso educacional, medido pela defasagem idade-série pode estar viesado para baixo. Crianças que ainda não repetiram o ano ou que não saíram da escola até a data de realização da pesquisa não podem ser computadas com atraso educacional. No final do ano, a defasagem idade-série é, portanto, potencialmente mais alta do que no mês de setembro. Como este problema de medida na variável dependente é decorrente da base de dados que utilizamos, ao analisar os resultados, devemos sempre ter em vista que o resultado do atraso educacional é calculado para um mês específico do ano, antes do término do período letivo.

<sup>18</sup>Seguimos a sugestão de Horowitz e Souza (2004) para a construção desse indicador. Contudo, a diferença com relação a esses autores é que usamos o total de anos de estudos completos da criança ao invés do grau de instrução freqüentado.

do viés provocado pela existência de fatores omitidos familiares que influenciam a determinação da renda e da escolaridade.

### 3.2. Os vínculos “não observados” entre as gerações

A tabela 1 abaixo apresenta a estrutura da matriz de variância e covariância das eqs (2) a (6). As correlações (expressas pelas letras da tabela 1) não nulas entre os erros seriam consequência da existência de fatores não observados que afetam simultaneamente a formação da renda e as decisões referentes à escolaridade (*simultaneidade*) ou que são passados de uma geração à outra da mesma dinastia (*hereditariedade*).<sup>19</sup>

**Tabela 1** – Estrutura da matriz de variância e covariância

	$u_{t+1}$	$u_t$	$u_{t-1}$	$v_{t+1}$	$v_t$
$u_{t+1}$	-	(a)	(a)	(c)	(d)
$u_t$	-	-	(a)	(b)	(c)
$u_{t-1}$	-	-	-	(b)	(b)
$v_{t+1}$	-	-	-	-	(b)
$v_t$	-	-	-	-	-

Obs: para simplificar não incluímos as variâncias na diagonal.

As equações (2), (4) e (6) mostram que as decisões de escolaridade de uma geração são tomadas durante a infância. Quando as gerações de pais e avós tornam-se adultas, supõe-se que não mais frequentam a escola. Assim, conforme esse modelo, as correlações que podem existir entre os erros dessas equações não se associam à *simultaneidade*. Entretanto, nessas equações podem constar aspectos herdados como hábito de leitura e preferências. Por exemplo, um indivíduo pode ter um melhor aprendizado que outro com características observáveis idênticas porque sua família tinha o costume de incentivar a leitura e o estudo em casa. Esse fator não é observado, mas está presente nos termos de erro  $u_{t+1}$ ,  $u_t$  e  $u_{t-1}$ . Se essas características passam de pais para filhos, ou de avós para netos, há provavelmente correlação entre os erros das eqs. (2), (4) e (6), ou seja, (a) não é nulo.

Somente para ilustrar, se a educação da criança da geração  $t + 1$  pode ser escrita em termos analíticos como uma função linear da escolaridade dos seus pais, crianças da geração  $t$  ( $educ_{1t}$ ),<sup>20</sup> de outros fatores familiares (ou dinásticos)  $u_f$  e individuais ( $u_{t+1}^i$ ), temos que:

$$educ_{1t+1} = \beta \cdot educ_{1t} + u_f + u_{t+1}^i \quad (9)$$

Onde  $E(educ_{1t}, u_f) \neq 0$  e  $E(educ_{1t}, u_{t+1}^i) = 0$ . Ou seja, esse fator familiar explica as de escolaridade entre crianças, pais e avós de diferentes dinastias. Se a equação (9) acima for estimada utilizando-se mínimos quadrados ordinários (MQO), como não temos informação sobre esse termo familiar, o coeficiente que mede o efeito da escolaridade dos pais seria viesado. Teríamos que:

$$\begin{aligned} p \lim \hat{\beta}_{MQO} &= E(educ_{1t} \cdot educ_{1t})^{-1} E(educ_{1t} \cdot educ_{1t+1}) = \\ &= \beta + E(educ_{1t} \cdot educ_{1t})^{-1} E(educ'_{1t} u_f) \end{aligned} \quad (10)$$

O viés do parâmetro  $\beta$  dependerá da persistência da influência deste termo familiar de uma geração para outra.

<sup>19</sup>Segundo Cogneau e Maurin (2001) e Maurin (2002).

<sup>20</sup>Excluímos a renda familiar per capita e as características observáveis da criança apenas para simplificar a discussão.



Outra forma desse componente hereditário influir as variações de escolaridade entre indivíduos de diferentes dinastias seria através do efeito sobre a aquisição de renda. Dependendo do tipo de ocupação, aspectos herdados da família podem ser valorizados no mercado de trabalho. Por exemplo, se um pai é professor, hábitos de leitura transmitidos pelos seus pais podem se tornar instrumentos essenciais para seu sucesso profissional, influenciando sua posição no mercado de trabalho e, conseqüentemente, sua aferição de rendimentos. Neste caso, a correlação (b) não seria nula. Se esse componente familiar herdado que influencia a aquisição de rendimentos é similar ao que afeta o acúmulo de capital humano dos membros de uma mesma dinastia, as correlações (a) e (b) seriam iguais.

Além dessas correlações potenciais citadas acima, a endogeneidade pode ter origem na presença de erros de medida e de simultaneidade. De acordo com Angrist e Krueger (2001), erros de medida surgem por diferentes razões: primeiro, devido à qualidade do dado reportado e, segundo, devido à incapacidade de se obter a informação tal como especificada na teoria. No tocante a este último aspecto, como o acúmulo de conhecimentos é feito ao longo do ciclo de vida do indivíduo, a renda relevante é a permanente.

Há também uma possibilidade de viés produzido pela simultaneidade das decisões de oferta de trabalho e escolaridade de diferentes membros da família. Se a renda incluída na análise é a corrente, pode-se gerar um viés na estimativa do seu impacto sobre a escolaridade, pois essa renda afeta as decisões referentes à alocação do tempo e, conseqüentemente, ao aprendizado da criança. Nesse caso, existe um componente não observado no erro da equação de aprendizado da criança que afeta simultaneamente a equação de formação da renda familiar (presente no erro desta equação).

Conforme a literatura de capital humano e de retorno à educação (Card, 1999) nas equações de rendimentos podem existir fatores omitidos que determinam alguns aspectos referentes à decisão do indivíduo de se escolarizar. Um exemplo é a habilidade ou talento individual. Se essa característica é valorizada no mercado de trabalho, uma pessoa mais educada ganha mais não somente porque tem mais instrução mas também porque detém mais habilidade.

Essas características individuais, por sua vez, também poderiam afetar o aprendizado dos filhos à medida que os pais se tornassem mais produtivos na transmissão de conhecimentos. Assim, a correlação (c) da tabela 1 não seria nula, pois existem fatores que simultaneamente estão afetando a aquisição de rendimentos por parte dos pais e o acúmulo de escolaridade por parte de seus filhos. Por exemplo, comparando crianças idênticas em todos os aspectos com exceção da habilidade dos pais, as que têm pais mais talentosos terão mais chances de ter um melhor desempenho educacional não somente porque o nível de renda familiar aumenta (se habilidade é valorada no mercado de trabalho e se renda tem um impacto na escolaridade), mas também porque os pais podem ganhar mais trabalhando menos e, portanto, podem alocar uma parte maior do seu tempo para contribuir para a aprendizagem das crianças. Por sua vez, a correlação (d) da tabela 1 também poderia ser não nula caso os avós contribuíssem diretamente para o aprendizado de seus netos.

Em todos os casos listados acima, a escolaridade dos pais (avós) e a renda familiar estão correlacionadas com o termo de erro da equação de atraso educacional da criança. Nosso problema, portanto, é: como estimar os parâmetros  $\alpha$  e  $\beta$  da eq. (8), dada a estrutura dinâmica de vínculos não observados existentes entre a renda (equação 2) e a escolaridade dos pais (equação 3).

Segundo Maurin (2002) e Cogneau e Maurin (2001), a estratégia econométrica depende das hipóteses feitas sobre a estrutura de variância e covariância dos erros das eqs. (2) a (6), ou seja, nossa tabela 1. Esses autores, usando a informação de renda e educação de três gerações sucessivas da França e de Madagascar, respectivamente, estimam o impacto da renda sobre a escolaridade das crianças usando diferentes procedimentos para testar a existência de viés de hereditariedade ou de simultaneidade.

Um dos resultados apresentados por Cogneau e Maurin (2001) e Maurin (2002) é que as diferenças entre o grau de instrução dos pais e dos avós podem ser um instrumento potencial para o nível de renda da família, mesmo na presença de viés de hereditariedade e simultaneidade. Este será um dos instrumentos utilizados para estimarmos o impacto da renda familiar e da escolaridade dos pais na probabilidade das crianças terem defasagem idade-série.

#### 4. ESTRATÉGIA ECONOMETRICA: A ESCOLHA DOS INSTRUMENTOS

A análise empírica adotada para investigar o efeito dos fatores familiares sobre a probabilidade da criança ter defasagem idade-série enfatizará a existência de variáveis omitidas entre as gerações. O principal problema é que queremos estimar a eq. (8), considerando as possíveis correlações existentes entre os termos de erros das eqs. (8), (3) e (4), transcritas novamente abaixo:

$$P(A_{1t+1,\tau} = 1) = P(\alpha.Y_{2t+1} + \beta.educ_{1t} + \lambda_a.X_{a,1t+1} + u_{t+1} < 0)$$

$$Y_{2,t+1} = a.educ_{1t} + \theta.X_{2t+1} + v_{t+1}$$

$$educ_{1t} = \alpha.Y_{2t} + \beta.educ_{1t-1} + \lambda.X_{1t} + u_t$$

Adotamos também as seguintes hipóteses com relação ao termo de erro da eq. (8):

$$u_{t+1} = u_{t+1}^i + u_f \quad (11)$$

$$E(u_f.u_j^i) = 0 \text{ onde } j = t+1, t, t-1, t-2, \text{ etc e } E(u_j^i.u_{j'}^i) = 0 \text{ se } j \neq j' \quad (12)$$

Ou seja, o termo de erro é escrito como uma função linear de fatores familiares fixos no tempo ( $u_f$ ) e outros aspectos individuais ( $u_{t+1}^i$ ). De acordo com a eq. (12), os fatores familiares não são correlacionados com os termos individuais e os termos individuais não são serialmente correlacionados.

Face aos problemas de endogeneidade, podemos estimar os parâmetros  $\alpha$  e  $\beta$  da eq. (8) de forma consistente usando a técnica de variáveis instrumentais. Faremos esse procedimento e iremos comparar os resultados gerados usando o modelo probabilidade linear e probit, identificando, desta forma, a direção do viés, tanto da renda familiar quanto da instrução do pai e da mãe.

Abaixo, listamos os instrumentos selecionados para estimar o efeito da renda familiar e da escolaridade dos pais sobre a probabilidade da criança ter defasagem idade-série:

1. Um instrumento institucional – a introdução da lei de diretrizes e bases da educação de 1971 que ampliou a escolaridade fundamental obrigatória de 4 para 8 anos.
2. A média do total de escolas de cada estado quando os pais estavam em idade escolar (7 a 14 anos de idade).<sup>21</sup>
3. A diferença entre a escolaridade dos pais e dos avós.<sup>22</sup>

Explicamos essas escolhas detalhadamente nas sub-seções seguintes.

<sup>21</sup>Construímos também a razão entre o total de professores pelo total de escolas. Como os resultados foram similares, optamos por não apresentá-los. Usamos as séries históricas do livro Estatísticas do Século XX, FIBGE (2003) e como para alguns períodos não foi possível separar escolas primárias de secundárias, optamos por utilizar o total. Através da parte de migração da PNAD, não temos o local de moradia dos pais na exata idade escolar. Consideramos o estado de nascimento como a medida mais razoável.

<sup>22</sup>Como não foi possível fazer o cálculo dos anos de estudos completos dos avós, comparamos a classificação do grau de instrução dos avós com os anos de estudos completos dos pais.



#### 4.1. A lei 5.692 de 1971

No início da década de 60, o sistema educacional brasileiro era dividido em duas etapas: o ensino primário e o médio. O primário era composto por, no mínimo, quatro anos, sendo factível sua extensão para seis anos. Todas as crianças a partir dos sete anos de idade deveriam se matricular na escola primária.

A educação média era destinada à formação dos adolescentes. Esta fase era constituída por dois ciclos: o ginásial e o colegial. A inscrição na primeira série do ciclo ginásial somente era feita após a conclusão do curso primário e a criança deveria ter ou fazer onze anos de idade ao longo do ano letivo. Para a matrícula na 1ª série do ciclo colegial, o pré-requisito era a conclusão do ciclo anterior. O ginásial tinha quatro séries anuais e o colegial pelo menos três séries.

A partir da reforma do ensino de 1971, incorporou-se à escola primária básica de quatro anos a antiga escola secundária de 1º ciclo (ginásial), ampliando, assim, de quatro para oito anos a escolaridade obrigatória. Desta forma, a reforma tornou obrigatória a presença na escola para todas as crianças entre 7 e 14 anos. Criou-se o ensino de 1º grau composto de oito anos letivos (da 1ª. a 8ª série) e o ensino de 2º. grau, formado por três ou quatro séries anuais, que corresponderiam ao ensino médio colegial.<sup>23</sup>

Parte das variações na escolaridade dos pais pode ser explicada por essas alterações no sistema educacional. A modificação institucional ocorrida em 1971 pode ser considerada exógena e correlacionada com a escolaridade dos pais, constituindo, portanto, um instrumento potencial.

Como a obrigatoriedade é definida pela faixa etária, ou seja, crianças de 7 a 14 anos de idade deveriam estar na escola, os pais afetados por essa reforma são os que tinham menos de 14 anos de idade em 1972, quando a lei entrou em exercício. Considerando a possibilidade de um ano de atraso educacional e as diferenças nos meses de nascimento ao longo do ano, optamos por criar uma variável igual a 1 para todos os pais e mães nascidos a partir de 1957. Quando crianças, por lei, eles deveriam permanecer na escola até completar os 14 anos de idade.

#### 4.2. A oferta de escolas

Para cada pai e mãe, imputamos a média do total de escolas conforme a unidade da federação em que nasceram e os anos em que estavam em idade escolar (7 a 14 anos). Ou seja, para cada pai e mãe, teríamos oito variáveis de oferta educacional, das quais fizemos uma média. Essas variações na oferta educacional por estado não estão relacionadas aos hábitos familiares (como talento para leitura) que influenciaram a aquisição de escolaridade por parte dos pais e de seus filhos.

Pais que nasceram em locais onde existiam melhores condições de escolarização tinham menores custos em frequentar a escola. Além disso, ao inserirmos as variáveis de oferta educacional do local de nascimento dos pais, captamos condições de vida que influíram apenas a sua infância e não a de seus filhos. Esse fenômeno é ainda mais forte se migraram do local de nascimento, pois nesse caso o conjunto de variáveis de oferta educacional afetou apenas a escolaridade dos pais. O impacto de mais ou menos escolas no local de nascimento dos pais influenciaria o aprendizado das crianças apenas através do efeito que teria sobre o nível educacional dos pais.

Por outro lado, se os pais não migraram do local de nascimento, o conjunto de indicadores da oferta educacional poderia estar impactando diretamente na escolaridade das crianças. Se o pai nasceu e continuou vivendo numa região com boas escolas, ou seja, não migrou, é provável que seu filho também usufrua facilidades similares. O mesmo ocorre se migrou para estudar e continuou vivendo no mesmo local em que estudou. Nestes dois casos, o conjunto de variáveis que utilizamos como instrumento

<sup>23</sup>É importante destacar que a medida de escolaridade dos pais usada ao longo deste artigo refere-se aos anos de estudos completos. Questões relacionadas à qualidade da educação após a implantação da lei 5.692 de 1971 não são abordadas. Existe uma enorme discussão na área de educação sobre a massificação educacional provocada por essa lei, sobretudo pela não adequabilidade do sistema educacional da 5ª a 8ª série ao novo público atingido. Ver Nunes (1996) e Gusso (1993).

para o nível educacional dos pais também estaria tendo um efeito direto na escolaridade das crianças, tornando nossa restrição de exclusão não válida.

Argumentamos, contudo, que o regime escolar das décadas de 60 e 70, época em que a maior parte dos pais da amostra de crianças estava na escola, também é muito diferente do regime escolar freqüentado por seus filhos. Como houve enorme transformação do sistema educacional a partir dos anos 80, tanto em termos quantitativos quanto qualitativos, não acreditamos que os vínculos existentes entre as instituições escolares passadas e as freqüentadas pelas crianças de nossa amostra sejam importantes para determinar o seu aprendizado, principalmente no tocante ao acúmulo de defasagem idade-série. Destacamos também que enquanto na época dos pais, a existência ou não de escolas no entorno constituía uma enorme barreira à educação fundamental, a partir dos anos 80, outros fatores são mais determinantes para a decisão de se escolarizar e, principalmente, para não ocorrer o acúmulo de defasagem idade-série. Com base nesta argumentação, usamos a oferta educacional do local de nascimento dos pais como instrumento para seus anos de estudos completos.

### 4.3. Diferença de escolaridade entre as gerações

O terceiro conjunto de instrumentos que utilizamos refere-se às diferenças entre o grau de instrução do avô/avó do lado materno (paterno) e a mãe (pai).<sup>24</sup> Esses instrumentos são construídos com base na idéia de “primeiras diferenças”. Como forma de eliminar o efeito fixo hereditário da dinastia que afeta a escolaridade das crianças e dos pais, tiramos a diferença de escolaridade entre duas gerações antecedentes à geração de crianças (de pais e avós). Usamos esses indicadores como instrumentos da renda familiar e do grau de instrução na equação de escolaridade da criança.

Na figura 1, podemos visualizar a proporção de domicílios pela nossa escala de pontos, segundo os quatros indicadores de diferenças educacionais (“educação mãe-educação avô materno”, “educação mãe-educação avó materna”, “educação pai-educação avô paterno”, “educação pai-educação avó paterna”). A maior parte dos pais e das mães ascenderam de 1 a 2 pontos na escala educacional e a proporção de pais que não mudaram de posição educacional em relação a seus pais é sempre superior a 18% (ponto zero no eixo x).

Pais que ascenderam na escala educacional em relação a seus pais conseguiram completar mais anos de estudo.<sup>25</sup> Pais que subiram na escala educacional em relação a seus pais também são os que possuem renda familiar mais alta.

A diferença entre o grau de instrução dos pais e dos avós afeta a probabilidade da criança ter defasagem idade-série apenas através da educação dos pais e do nível de renda familiar. O argumento defendido para justificar essa hipótese é que os fatores não observados que afetaram a escolaridade dos pais e a formação da renda familiar e que também influenciam a probabilidade da criança ter defasagem idade-série são principalmente familiares e transmitidos de geração para geração (*efeitos fixos*). Como esses fatores também influenciam a escolaridade dos avós, são eliminados quando calculamos as diferenças (o termo  $u_f$  da eq. (9)).

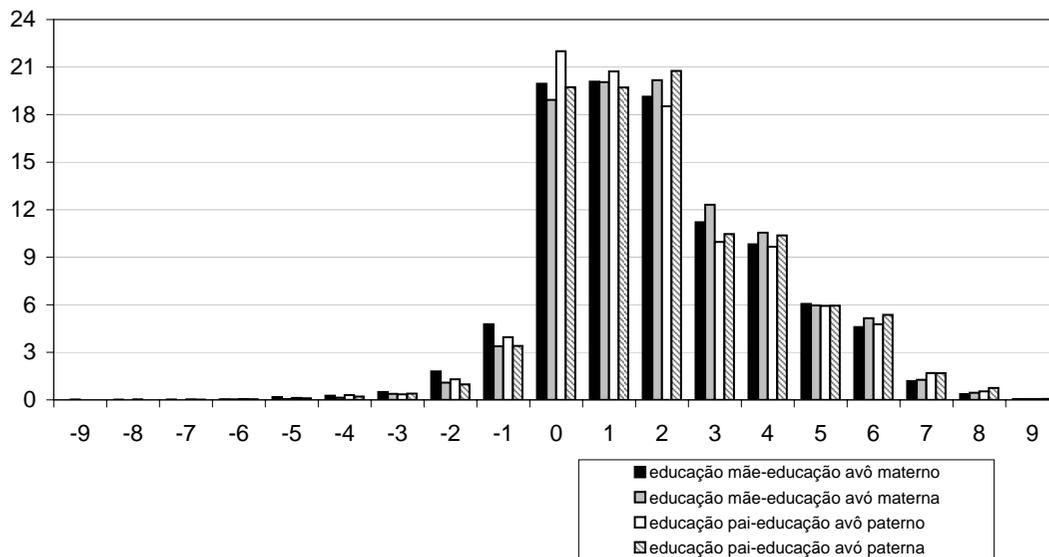
Vale destacar que essa diferença também poderia indicar que alguns pais são mais hábeis do que outros para se escolarizar, distanciando-se mais da escala educacional dos seus pais (terceira geração). Ou seja, o desempenho educacional de um indivíduo é igualmente influenciado pelas suas habilidades inatas para aprender. Indivíduos mais hábeis no aprendizado seriam mais eficientes em ensinar e transmitir conhecimentos e, portanto, mais produtivos em auxiliar seus filhos nos estudos. Entretanto, esse

<sup>24</sup>Uma alternativa apontada na literatura (Couralet, 2002, Maurin, 2002) é usar o nível educacional dos avós como instrumento para a renda familiar per capita. Neste caso, supõe-se que as transmissões de pais para filhos “esgotariam” o viés de hereditariedade. Alguns autores (Dumas e Lambert, 2005, Dahl e Lochner, 2005) criticam esse argumento tendo em vista que podem persistir fatores não observáveis relacionados com a formação educacional da geração de avós que afetam igualmente a geração de crianças e jovens.

<sup>25</sup>A correlação entre as 4 variáveis de diferenças educacionais e o nível educacional dos pais ou a renda familiar per capita é positiva e significativa a 1%.



**Figura 1** – Proporção de domicílios segundo as diferenças entre o grau de instrução de pais e avós



Fonte: Pnad/IBGE, 1996

efeito da habilidade dos pais sobre o aprendizado dos filhos está sendo captado diretamente pelo nível de escolaridade dos pais, definido em um período anterior.

## 5. RESULTADOS

Nesta seção, a eq. (8), transcrita abaixo, é estimada através dos modelos de probabilidade linear e probit em dois estágios usando nosso conjunto de instrumentos, conforme descrito na seção 4.

$$P(A_{1t+1,\tau} = 1) = P(\alpha.Y_{2t+1} + \beta.educ_{1t} + \lambda_a.X_{a,1t+1} + u_{t+1} < 0)$$

A critério de comparação, estimamos um modelo de probabilidade linear usando mínimos quadrados ordinários e um modelo *probit* por máxima verossimilhança (estimação padrão). Nesses dois últimos casos, iremos comparar o viés provocado nos efeitos renda e educação dos pais quando não consideramos a possibilidade de existência de endogeneidade.

Nossa variável dependente é um indicador igual a um se a criança tem defasagem idade-série sendo regredida no logaritmo da renda familiar per capita e no total de anos de estudos do pai e da mãe. Também incluímos outros controles na nossa estimação, referentes às características da família, da criança e do local de moradia.

Abaixo, discutimos os principais resultados referentes às estimações dos parâmetros  $\alpha$  e  $\beta$ , que medem os efeitos da renda familiar e da educação dos pais, respectivamente, sobre a defasagem idade-série.

### 5.1. Efeito da renda familiar per capita e da escolaridade dos pais

Da coluna (1) a (2) da tabela 2, na próxima página, mostramos os resultados da estimação dos parâmetros  $\alpha$  e  $\beta$  da eq. (8) que descreve a probabilidade da criança ter defasagem idade-série, usando o modelo de probabilidade linear em dois estágios. Nas colunas (3) e (4) reportamos os resultados do modelo de probabilidade linear usando mínimos quadrados ordinários. Nas colunas (5) a (8), a estimação *probit* sem instrumentos e, nas colunas (9) a (12), a estimação *probit* em dois estágios.

Nas três primeiras linhas da tabela 2, verificamos a forte correlação negativa entre a probabilidade da criança ter defasagem idade-série e o nível de renda familiar per capita e o grau de instrução dos pais, com exceção do coeficiente dos anos de estudo do pai na estimação do *probit* em dois estágios que é positivo mas não significativo (ver coluna 9 da tabela 2). Os sinais dos parâmetros  $\alpha$  e  $\beta$  estimados de forma significativa são sempre negativos. Crianças de pais mais instruídos e que vivem em famílias com renda per capita mais elevada têm menos chances de acumular defasagem idade-série.

O efeito<sup>26</sup> do nível educacional da mãe é mais intenso que o do pai. A diferença na probabilidade das crianças terem defasagem idade-série causada pelo aumento de um ano a mais de escolaridade da mãe é de 0,008 (coluna (1) da tabela 2), segundo o modelo de probabilidade linear com instrumentos. O impacto da escolaridade do pai na probabilidade da criança ter defasagem idade-série é também negativo, contudo perde a significância quando consideramos a possibilidade da endogeneidade (ver coluna (1) da tabela 2, linha dois).<sup>27</sup>

Em todas as especificações, o impacto da renda familiar<sup>28</sup> na redução da probabilidade da criança ter defasagem idade-série é sempre significativo. Quando utilizamos o método de variáveis instrumentais em dois estágios (veja colunas 1 e 9 da tabela 2), esse resultado torna-se mais forte.

Usando o método de variáveis instrumentais, o efeito direto da renda familiar per capita sobre a probabilidade das crianças terem defasagem idade-série é quase 6 vezes maior do que o encontrado nas estimações padrões (sem usar instrumentos). Por outro lado, o efeito de um ano a mais de estudo do pai é reduzido significativamente (de 0,01 para 0,001, comparando os modelos de probabilidade linear, sem e com instrumentos, respectivamente). Fenômeno similar ocorre para as mães (de 0,02 para 0,008).

<sup>26</sup>Os efeitos marginais foram calculados para uma população de referência. Veja nota ao final da tabela 2.

<sup>27</sup>Quando as estimações são feitas pelo procedimento *probit* em dois estágios, a significância das variáveis de educação dos pais cai. Priorizamos, portanto, a análise do modelo de probabilidade linear. O importante é notar que o efeito é similar: perda de significância e redução em valor absoluto.

<sup>28</sup>Quando falamos do impacto da renda, subentende-se que se refere ao logaritmo da renda familiar per capita incluindo a renda dos pais e dos demais membros da família que moravam no domicílio no momento de realização da pesquisa. O suplemento de mobilidade social, de onde foram retiradas as informações sobre os avós das crianças, não continha dados sobre o nível de renda. A renda dos avós somente é computada no total da renda familiar se eles viviam no domicílio.



Tabela 2 – Resultados das estimações da probabilidade da criança ter defasagem idade-série

Variável dependente:	IV - probabilidade linear			OLS - probabilidade linear			Probit			IV - Probit		
	Coef (1)	ep (2)	Coef (3)	ep (4)	Coef (5)	ep (6)	Ef marg (7)	ep (8)	Coef (9)	ep (10)	Ef marg (11)	ep (12)
Ter defasagem idade-série	-0,008 <sup>**</sup>	0,004	-0,018 <sup>*</sup>	0,001	-0,07 <sup>*</sup>	0,004	-0,019 <sup>*</sup>	0,001	-0,015	0,014	-0,005	0,004
anos de estudos completos da mãe	-0,001	0,004	-0,012 <sup>*</sup>	0,001	-0,049 <sup>*</sup>	0,004	-0,013 <sup>*</sup>	0,001	0,006	0,014	0,002	0,005
anos de estudos completos do pai	-0,203 <sup>*</sup>	0,056	-0,033 <sup>*</sup>	0,004	-0,134 <sup>*</sup>	0,013	-0,036 <sup>*</sup>	0,004	-1,008 <sup>*</sup>	0,194	-0,33 <sup>*</sup>	0,071
In da renda familiar per capita												
idade em 10. de março	0,076 <sup>*</sup>	0,003	0,067 <sup>*</sup>	0,001	0,257 <sup>*</sup>	0,006	0,069 <sup>*</sup>	0,002	0,309 <sup>*</sup>	0,012	0,101 <sup>*</sup>	0,007
parte inteira da idade	-0,125 <sup>*</sup>	0,011	-0,126 <sup>*</sup>	0,01	-0,505 <sup>*</sup>	0,041	-0,135 <sup>*</sup>	0,012	-0,507 <sup>*</sup>	0,042	-0,166 <sup>*</sup>	0,014
maturidade	-0,088 <sup>*</sup>	0,007	-0,084 <sup>*</sup>	0,006	-0,33 <sup>*</sup>	0,024	-0,101 <sup>*</sup>	0,008	-0,336 <sup>*</sup>	0,025	-0,12 <sup>*</sup>	0,009
ser menina	-0,021	0,016	-0,063 <sup>*</sup>	0,008	-0,207 <sup>*</sup>	0,028	-0,06 <sup>*</sup>	0,008	0,002	0,053	0,001	0,017
ser de cor branca ou amarela												
domicílio localizado na:												
área metropolitana	0,038 <sup>*</sup>	0,012	0,011	0,007	0,047 <sup>***</sup>	0,028	0,012 <sup>***</sup>	0,007	0,171 <sup>*</sup>	0,035	0,053 <sup>*</sup>	0,011
região nordeste	0	0,02	0,05 <sup>*</sup>	0,01	0,161 <sup>*</sup>	0,036	0,046 <sup>*</sup>	0,011	-0,123 <sup>***</sup>	0,067	-0,039 <sup>***</sup>	0,02
região sudeste	-0,073 <sup>*</sup>	0,012	-0,089 <sup>*</sup>	0,01	-0,315 <sup>*</sup>	0,036	-0,096 <sup>*</sup>	0,012	-0,235 <sup>*</sup>	0,038	-0,082 <sup>*</sup>	0,013
região sul	-0,082 <sup>*</sup>	0,013	-0,094 <sup>*</sup>	0,011	-0,341 <sup>*</sup>	0,043	-0,077 <sup>*</sup>	0,009	-0,274 <sup>*</sup>	0,045	-0,081 <sup>*</sup>	0,011
total de irmãos/irmãs mais novos	-0,005	0,015	0,039 <sup>*</sup>	0,003	0,13 <sup>*</sup>	0,012	0,035 <sup>*</sup>	0,004	-0,097 <sup>***</sup>	0,051	-0,032 <sup>***</sup>	0,017
total de irmãos/irmãs mais velhos	0,021 <sup>*</sup>	0,007	0,035 <sup>*</sup>	0,004	0,124 <sup>*</sup>	0,016	0,033 <sup>*</sup>	0,004	0,057 <sup>*</sup>	0,023	0,019 <sup>*</sup>	0,007
total de irmãs mais velhas	-0,015 <sup>**</sup>	0,007	-0,012 <sup>**</sup>	0,006	-0,039 <sup>***</sup>	0,024	-0,01 <sup>***</sup>	0,006	-0,062 <sup>*</sup>	0,024	-0,02 <sup>*</sup>	0,008
diferença entre a idade do pai e da mãe	0,002 <sup>***</sup>	0,001	0,001 <sup>***</sup>	0,001	0,006 <sup>**</sup>	0,003	0,002 <sup>**</sup>	0,001	0,008 <sup>*</sup>	0,003	0,003 <sup>*</sup>	0,001
constante	0,733 <sup>*</sup>	0,205	0,105 <sup>*</sup>	0,023	-1,425 <sup>*</sup>	0,088			1,694 <sup>**</sup>	0,709		
Obs:	20.332		20.332		Obs:		20.332		Obs:		20.332	
F(15,12831)	515,9		802,15		Wald chi2(15)		3.749,80		Wald chi2(15)		4503,1	
Prob > F	0,000		0,000		Prob > chi2		0,000		Prob > chi2		0,000	
R <sup>2</sup>	0,222		0,334									

Fonte: Pnad/IBGE 1996. Obs: (1) os instrumentos são: diferença entre educação da mãe e de seus pais e a educação do pai e de seus pais, variável categórica igual a 1 para pais nascidos depois de 1957, média do total de escolas no período escolar e no estado de nascimento do pai e da mãe; (2) o efeito marginal foi calculado para a seguinte população de referência: meninas, brancas ou amarelas, moradoras da região sudeste e de área metropolitana, para as demais variáveis incluímos a média; (3) \* p-valor < = 1%, \*\* p-valor entre 1% e 5%, \*\*\* p-valor > 5% e < = 10%. (4) Os erros padrões estimados para os modelos lineares são robustos, sendo corrigidos para potenciais problemas decorrentes de heterocedasticidade. (5) ep - erro padrão. (6) Teste de sobreidentificação: Hansen J estatística Chi2(5) = 25,802 para o modelo de probabilidade linear.

Encontramos que a renda familiar é um fator muito importante para explicar aspectos referentes à escolaridade das crianças. De acordo com as colunas (1) e (9) da tabela 2, quando a renda familiar per capita aumenta em 1%, a probabilidade da criança ter defasagem idade-série é reduzida em 0,20 e 0,33, usando os modelos de probabilidade linear e probit, com instrumentos, respectivamente. Esse impacto é significativamente maior que o efeito estimado por mínimos quadrados ordinários ou pelo *probit*.

Outro resultado é que o efeito da escolaridade dos pais nas estimações padrões é superior ao estimado em dois estágios. Uma possível explicação para esse resultado é a existência de fatores familiares omitidos que afetaram a escolaridade dos pais e que também influenciam a probabilidade da criança ter defasagem idade-série. Pais mais educados já detêm uma “bagagem” herdada dos seus antecedentes que influencia seus filhos, independentemente do número de anos de estudo que adquiram.

Com relação aos resultados das outras características incluídas na nossa regressão, observamos que não variam entre nossas especificações e estão em conformidade com a literatura de educação infantil.

Meninas têm menor probabilidade de terem defasagem idade-série comparativamente aos meninos. Usualmente, crianças do sexo masculino ajudam seus pais no trabalho ao contrário das meninas, que em geral cooperam mais intensamente em atividades no domicílio (cuidar de irmãos mais novos, afazeres domésticos, etc) não computadas entre as atividades produtivas. O custo de oportunidade dos meninos também é superior, pois a demanda da sociedade para que trabalhem mais cedo é maior.

Crianças brancas ou amarelas, em geral, têm menor probabilidade de terem defasagem idade-série comparativamente às demais. Estas pertencem aos grupos sociais com as piores condições econômicas, dificultando sua progressão ao longo do ciclo escolar: entram tarde na escola ou têm dificuldades em continuar os estudos. É interessante notar que esse efeito é reduzido quando controlamos por aspectos familiares não observados. Esse resultado sugere que obstáculos ao progresso educacional de crianças não brancas e amarelas podem refletir dificuldades já herdadas das gerações passadas.

As variáveis indicadoras das regiões onde as crianças moram refletem o efeito das condições estruturais oferecidas em termos de oferta educacional ou de serviços públicos. O Sul é onde as crianças possuem maiores oportunidades de instrução, ao contrário das regiões Norte/Nordeste e Centro-Oeste. Por outro lado, nas áreas metropolitanas, a probabilidade da criança acumular defasagem idade-série é maior que nas não metropolitanas. Apesar da oferta educacional possivelmente ser melhor nestas localidades, podem existir outros fatores, como os associados à atratividade do mercado de trabalho, incentivando a saída da escola e a não dedicação aos estudos.

Com relação à composição da família, o número de irmãs ou irmãos mais velhos tem o efeito de aumentar a probabilidade da criança ter defasagem escolar. Quanto maior o número de irmãos, maior a divisão do tempo e da renda dos pais, diminuindo também o total desses insumos na produção educacional.

Comparando crianças que têm a mesma quantidade de irmãos/irmãs mais velhos, aquelas com mais irmãs com idade superior possuem uma menor probabilidade de terem defasagem idade-série. Essas irmãs mais velhas podem contribuir para o serviço doméstico e o cuidado dos irmãos. Neste caso, as crianças mais novas são beneficiadas, pois as mais velhas assumem o papel do “adulto” na ausência dos pais.

No tocante à idade da criança, quanto mais alta a faixa etária, maior a probabilidade de ocorrer defasagem idade-série. Esse efeito parece natural tendo em vista que nossa variável dependente está captando não somente o atraso causado pela entrada tardia na escola mas também por outros fenômenos, como repetência e evasão escolar. À medida que a criança cresce e evolui ao longo do sistema escolar, maiores são as chances de acumular defasagem idade-série.

Destacamos que para crianças da mesma faixa etária, as que têm mais maturidade possuem uma menor probabilidade de terem defasagem idade-série. Como pode ser visto na sexta linha da tabela 2, o coeficiente da variável maturidade é negativo em todas as três especificações. Comparando duas crianças com 8 anos de idade completa em 1º. de março, a que fez aniversário no início do ano tem uma probabilidade mais alta de estar atrasada do que outra que fez aniversário no final do ano anterior. Isto pode estar ocorrendo devido às regras de inserção no sistema educacional. A regra usualmente aplicada



para a criança ingressar na 1ª. série do ensino fundamental é ter 7 anos completos em 1º. de março, ou fazer aniversário no 1º. semestre.

## 5.2. Algumas observações sobre o primeiro estágio

Como pode ser visto no final da tabela 3 que apresenta o primeiro estágio do modelo de probabilidade linear, o conjunto dos instrumentos possui poder explicativo (teste F) em todas as três equações: grau de instrução do pai e da mãe e logaritmo da renda familiar per capita.<sup>29</sup>

O resultado da variável que indica que a mãe nasceu depois de 1957 teve o sinal esperado na escolaridade da mãe. O coeficiente estimado é positivo, ou seja, mães que nasceram depois de 1957 têm maior grau de instrução do que as demais. Já para os pais, a variação da escolaridade não é explicada de forma significativa por essa diferença no ano de nascimento. Isto ocorre possivelmente porque a maioria dos pais na nossa amostra de crianças nasceu antes de 1957.

Os coeficientes estimados para a média do total de escolas no período em que mães e pais estavam em idade escolar foram positivos. Ou seja, pais e mães que nasceram em locais com maior oferta escolar, adquiriram mais anos de estudos do que outros pais e mães, nascidos em lugares com infra-estrutura escolar menos adequada.

Os coeficientes estimados das diferenças entre o grau de escolaridade da segunda (pais) e da terceira geração (avós e avôs das crianças) também têm os sinais positivos esperados. Pais e mães que mais subiram na escala educacional em relação aos seus pais têm grau de instrução mais elevado e uma renda familiar mais alta. Esse efeito é mais forte quando o grau de instrução dos pais e das mães da segunda geração é comparado com a escolaridade das mães da terceira geração (as avós).

Uma possível extensão aos resultados apresentados acima é investigar se os resultados modificam se considerarmos o impacto direto da escolaridade dos avós sobre a escolaridade das crianças. Como pode ser visto na tabela 4, o fato de avós e avôs terem baixa escolaridade não influencia de forma significativa a probabilidade da criança ter defasagem idade-série. Todos os coeficientes dos avós e avôs foram estimados de forma imprecisa.

No entanto, o coeficiente estimado para a escolaridade da mãe aumenta, em termos absolutos, enquanto da renda familiar per capita diminui em relação à estimação em dois estágios anterior, em que não incluímos entre os regressores variáveis explicativas da instrução dos avós (compare as colunas 1 das tabelas 2 e 4). A escolaridade dos pais permanece tendo um efeito não significativo sobre a probabilidade da criança ter defasagem idade-série. Apesar disso, o mesmo resultado da comparação com as estimações padrões permanece: a renda familiar nas estimações padrões é viesada para baixo e o inverso ocorre para a escolaridade dos pais.

<sup>29</sup>Fizemos o teste F para alguns grupos de instrumentos separadamente e o poder explicativo permanece como pode ser visto nos valores reportados ao final da tabela 3. O poder explicativo é maior para o conjunto de instrumentos relacionados às diferenças educacionais entre as gerações de pais e avós.

Tabela 3 – Estimativas do primeiro estágio

	Anos de estudo da mãe		Anos de estudo do pai		Log da renda familiar per capita	
	Coef (1)	ep (2)	Coef (3)	ep (4)	Coef (5)	ep (6)
idade em 1o. de março	-0,03*	0,008	-0,017**	0,009	0,038*	0,003
parte inteira da idade	-0,03	0,06	0,053	0,062	-0,002	0,025
maturidade	-0,022	0,035	-0,044	0,036	-0,025***	0,014
ser menina	1,036*	0,04	1,119*	0,042	0,365*	0,017
ser de cor branca ou amarela						
domicílio localizado na:						
área metropolitana	0,535*	0,038	0,664*	0,04	0,228*	0,016
região nordeste	-0,372*	0,062	-0,408*	0,064	-0,36*	0,026
região sudeste	-0,51*	0,066	-0,298*	0,068	-0,011	0,027
região sul	-0,504*	0,07	-0,368*	0,073	-0,015	0,029
total de irmãos/irmãs mais novos	-0,25*	0,017	-0,226*	0,017	-0,275*	0,007
total de irmãos/irmãs mais velhos	-0,333	0,024	-0,274*	0,025	-0,137*	0,01
total de irmãs mais velhas	0,064**	0,035	0,066**	0,036	-0,002	0,014
diferença entre a idade do pai e da mãe	-0,04	0,004	-0,035*	0,005	-0,007*	0,002
educação mãe-educação avô materno (a)	0,632	0,014	-0,002	0,015	0,031	0,006
educação mãe-educação avó materna (b)	0,923	0,015	0,262*	0,015	0,085*	0,006
educação pai-educação avô paterno (c)	0,054	0,015	0,575*	0,016	0,043*	0,006
educação pai-educação avó paterna (d)	0,254*	0,015	1,036*	0,016	0,089*	0,006
média de escolas mãe (e)	0,008**	0,005	-0,008**	0,005	0,008*	0,002
média escolas pai (f)	-0,001	0,005	0,011**	0,005	-0,001	0,002
pai nasceu depois de 1957 (g)	-0,079	0,048	-0,026	0,049	-0,102*	0,02
mãe nasceu depois de 1957 (h)	0,133*	0,046	0,269*	0,047	-0,057*	0,019
constante	2,99	0,133	2,396	0,138	4,195	0,055
Obs:	20.332		20.332		20.332	
Partial R2	0,632		0,636		0,137	
F(8,20311)	4.353,59		4.426,20		401,62	
H0: a = b = c = d = 0 F(4, 20311)	6.724,09		6.219,76		648,8	
H0: e = f = 0 F(2, 20311)	2,54		2,13		8,05	
H0: g = h = 0 F(2, 20311)	3,7		15,79		25,67	

Fonte: Pnad/IBGE 1996.

Observações: (1) os instrumentos são: diferença entre educação da mãe e de seus pais e a educação do pai e de seus pais; variável categórica igual a 1 para pais nascidos depois de 1957, média do total de escolas no período escolar e no estado de nascimento do pai e da mãe; (2) \* p-valor < = 1%, \*\* p-valor entre 1% e 5%, \*\*\* p-valor > 5% e < = 10%. (3) Os erros padrões estimados são robustos, sendo corrigidos para potenciais problemas decorrentes de heterocedasticidade. ep - erro padrão.



## 6. CONSIDERAÇÕES FINAIS

No tocante à estimação das variáveis potencialmente endógenas, encontramos que todas as três reduzem a probabilidade da criança ter defasagem idade-série. O efeito da renda familiar foi bem superior ao encontrado nas estimações padrões enquanto o impacto da educação da mãe e do pai foi reduzido, sobretudo desse último.

A renda, quando exógena, não capta aspectos permanentes do nível sócio econômico da família e existem fatores omitidos que simultaneamente afetam sua determinação e o aprendizado das crianças. Como na nossa lista de instrumentos, utilizamos a diferença entre o grau de instrução dos pais em relação aos avós, estaríamos captando mudanças permanentes na escala social e educacional, portanto, traduzindo uma mudança mais de longo prazo que impacta na vida das crianças. Destaca-se que como a escolaridade dos pais foi determinada num período anterior ao processo de aprendizado das crianças, também estaríamos reduzindo o viés causado por simultaneidade das decisões.

O efeito do nível educacional dos pais quando calculado de forma exógena poderia estar sendo influenciado por algum viés de hereditariedade. Ou seja, seu efeito nas estimações padrões estaria sendo super-estimado pois captaria o impacto de variáveis omitidas, herdadas da família e que influenciam o acúmulo de capital humano das crianças.

Em termos de política, um programa de transferências de renda parece ter um forte impacto na redução da defasagem idade-série mesmo quando controlamos por fatores familiares omitidos passados de uma geração para outra. Vale destacar também que neste artigo focamos apenas em um lado do processo de acúmulo de capital humano (a demanda educacional por parte das famílias), não abordando a oferta educacional oferecida à primeira geração. Uma extensão a esta pesquisa seria incorporar à análise aspectos relacionados à escola identificando as complementaridades entre fatores familiares e escolares, de forma a potencializar qualquer tipo de ação política que priorizasse o público infantil em idade escolar.

Com relação às características gerais das crianças e da família, nossos resultados estão em conformidade com a literatura existente. O público infantil mais vulnerável à ocorrência de defasagem idade-série é formado por: meninos, não brancos e amarelos, moradores das regiões onde há maiores dificuldades de oferta educacional, como o Nordeste, pertencentes às famílias maiores, com pais que têm menor nível de instrução e com menor renda familiar per capita.

As matrizes de mobilidade mostraram que pais/avós de nível educacional mais baixo possuem / netos mais vulneráveis ao atraso educacional. O mesmo ocorre quando investigamos as desagregações pelos décimos da renda familiar: as crianças pertencentes às famílias mais pobres tinham mais chances que as outras de acumularem defasagem idade-série.

**Tabela 4** – Impacto da escolaridade dos avôs e avós - probabilidade linear com instrumentos

Variável dependente: Ter defasagem idade-série	Coefficiente (1)	ep (2)
anos de estudos completos da mãe	-0,011*	0,003
anos de estudos completos do pai	-0,003	0,004
ln da renda familiar per capita	-0,168*	0,049
idade em 1o. de março		
parte inteira da idade	0,074*	0,003
maturidade	-0,125*	0,011
ser menina	-0,087*	0,007
ser de cor branca ou amarela	-0,03**	0,013
domicílio localizado na:		
área metropolitana	0,032*	0,011
região nordeste	0,01	0,017
região sudeste	-0,077*	0,012
região sul	-0,085*	0,012
total de irmãos/irmãs mais novos	0,004	0,013
total de irmãos/irmãs mais velhos	0,024*	0,006
total de irmãs mais velhas	-0,014**	0,007
diferença entre a idade do pai e da mãe	0,002***	0,001
avô do lado materno sem escolaridade	0,013	0,01
avó do lado materno sem escolaridade	-0,002	0,011
avô do lado paterno sem escolaridade	-0,014	0,01
avó do lado paterno sem escolaridade	-0,006	0,01
constante	0,61*	0,184
Obs:		20.332
F(19, 12831)		483,53
Prob > F		0
R2		0,263

Fonte: Pnad/IBGE 1996.

Observações: (1) os instrumentos são: diferença entre educação da mãe e de seus pais e a educação do pai e de seus pais, variável categórica igual a 1 para pais nascidos depois de 1957, média do total de escolas no período escolar e no estado de nascimento do pai e da mãe; (2) \* p-valor  $\leq 1\%$ , \*\* p-valor entre 1% e 5%, \*\*\* p-valor  $> 5\%$  e  $\leq 10\%$ . (3) Os erros padrões estimados são robustos, sendo corrigidos para potenciais problemas decorrentes de heterocedasticidade.

(2) ep - erro padrão.



## Referências Bibliográficas

- Albernaz, A., Ferreira, F., & Franco, C. (2002). Qualidade e equidade no ensino fundamental Brasileiro. *Pesquisa e Planejamento Econômico – IPEA*, 32(3).
- Angrist, J. D. & Krueger, A. B. (2001). Instrumental variables and the search for identification: From supply and demand to natural experiments. *Journal of Economic Perspectives*, 15(4):69–85. <http://ideas.repec.org/a/aea/jecper/v15y2001i4p69-85.html>.
- Barros, R. & Lam, D. (1993). Desigualdade de renda, desigualdade em educação e escolaridade das crianças no Brasil. *Pesquisa e Planejamento Econômico – IPEA*, 23(2).
- Barros, R. P. d. & Mendonça, R. (1998). Conseqüências da repetência sobre o desempenho educacional. Série Estudos – Projeto de Educação Básica para o Nordeste 7, AMinistério da Educação e do Desporto, Brasília.
- Barros, R. P. d., Mendonça, R., Santos, D. D. d., & Quintaes, G. (2001). Determinantes do desempenho educacional no Brasil. Texto para discussão 834, IPEA, Rio de Janeiro. [http://www.ipea.gov.br/pub/td/td\\_2001/td\\_0834.pdf](http://www.ipea.gov.br/pub/td/td_2001/td_0834.pdf).
- Barros, R. P. d., Mendonça, R., & Velazco, T. (1996). A pobreza é a principal causa do trabalho infantil no Brasil urbano? In IPEA, editor, *Economia brasileira em perspectiva*, volume 2. IPEA.
- Basu, K. (1999). Child labor: Cause, consequence and cure, with remarks on international labor standards. *Pesquisa e Planejamento Econômico – IPEA*, 37.
- Basu, K. & Tzannatos, Z. (2003). The global child labor problem: What do we know and what can we do? *World Bank Economic Review*, 17(2):147–173. <http://ideas.repec.org/a/oup/wbecrv/v17y2003i2p147-173.html>.
- Black, S. E., Devereux, P. J., & Salvanes, K. G. (2005). Why the apple doesn't fall far: Understanding intergenerational transmission of human capital. *American Economic Review*, 95(1):437–449. <http://ideas.repec.org/a/aea/aecrev/v95y2005i1p437-449.html>.
- Blau, D. M. (1999). The effect of income on child development. *The Review of Economics and Statistics*, 81(2):261–276. <http://ideas.repec.org/a/tpr/restat/v81y1999i2p261-276.html>.
- Bonelli, R. & Veiga, A. (2004). Determinantes de exclusão educacional em cinco estados do Brasil. In Pena, M. V. J. & Santos, M. M. R. d., editors, *Vulnerabilidade entre crianças e jovens: pobreza, exclusão e risco social e cinco estados brasileiros*. Banco Mundial, Brasília, 1a edition.
- Bourguignon, F., Ferreira, F. H. G., & Menendez, M. (2003). Inequality of outcomes and inequality of opportunities in Brazil. Textos para discussão 478, Departament of Economics PUC-Rio (Brazil). <http://ideas.repec.org/p/rio/texdis/478.html>.
- Card, D. (1999). The causal effect of education on earnings. In Ashenfelter, O. & Card, D., editors, *Handbook of Labor Economics*, volume 3 of *Handbook of Labor Economics*, chapter 30, pages 1801–1863. Elsevier. <http://ideas.repec.org/h/eee/labchp/3-30.html>.
- Carvalho, I. E. (2000). *Income Effects on Child Labor and School Enrollment in Brazil*. Tese de Doutorado, Department of Economics of Massachusetts – Institute of Technology.
- Cascio, E. & Lewis, E. (2005). Schooling and the afqt: evidence from school entry laws. IZA Discussion Paper Series 1481, Institute for the Study of Labor (IZA).

- Chevalier, A. (2004). Parental education and child's education: A natural experiment. IZA Discussion Papers 1153, Institute for the Study of Labor (IZA). <http://ideas.repec.org/p/iza/izadps/dp1153.html>.
- Chevalier, A., Harmon, C., O'Sullivan, V., & Walker, I. (2005). The impact of parental income and education on the schooling of their children. IZA Discussion Papers 1496, Institute for the Study of Labor (IZA). <http://ideas.repec.org/p/iza/izadps/dp1496.html>.
- Cogneau, D. & Maurin, E. (2001). Parental income and school attendance in a low-income country : A semi-parametric analysis. Papers 2001-08, Institut National de la Statistique et des Etudes Economiques-. <http://ideas.repec.org/p/fth/inseep/2001-08.html>.
- Couralet, P. E. (2002). *Une analyse économique du travail des enfants*. Tese de Doutorado, École des Hautes Études en Sciences Sociales, Paris.
- Currie, J. & Moretti, E. (2003). Mother's education and the intergenerational transmission of human capital: Evidence from college openings. *The Quarterly Journal of Economics*, 118(4):1495–1532. <http://ideas.repec.org/a/tpr/qjecon/v118y2003i4p1495-1532.html>.
- Dahl, G. B. & Lochner, L. (2005). The impact of family income on child achievement. NBER Working Papers 11279, National Bureau of Economic Research, Inc. <http://ideas.repec.org/p/nbr/nberwo/11279.html>.
- De Vreyer, P., Lambert, S., & Magnac, T. (1999). Educating children: a look at household behavior in côte d'ivoire. Document de Recherche EPEE 99-13, Centre d'Étude des Politiques Économiques de l'Université d'Evry.
- Dubois, P. & alli (2002). Effets sur l'inscription et la performance à l'école d'un programme de transferts conditionnels au mexique. Cahier de Recherche 3, Institut National de la Recherche Agronomique (INRA), Toulouse.
- Duflo, E. (2003). Grandmothers and granddaughters: Old-age pensions and intrahousehold allocation in south africa. *World Bank Economic Review*, 17(1):1–25. <http://ideas.repec.org/a/oup/wbecrv/v17y2003i1p1-25.html>.
- Dumas, C. & Lambert, S. (2005). Children education in senegal : how does family background influence achievement. Research Unit Working Papers 0503, Laboratoire d'Economie Appliquee, INRA. <http://ideas.repec.org/p/lea/leawpi/0503.html>.
- Duryea, S. (1998). Children's advancement through school in brazil: the role of transitory shocks to household income. RES Working Papers 376, Inter-American Development Bank, Research Department. <http://ideas.repec.org/p/idb/wpaper/376.html>.
- Emerson, P. M. & Souza, A. P. (2002). Birth order, child labor and school attendance in Brazil. Working Papers 0212, Department of Economics, Vanderbilt University. <http://ideas.repec.org/p/van/wpaper/0212.html>.
- Emerson, P. M. & Souza, A. P. (2003). Is there a child labor trap? intergenerational persistence of child labor in Brazil. *Economic Development and Cultural Change*, 51(2):375–98. <http://ideas.repec.org/a/ucp/ecdecc/v51y2003i2p375-98.html>.
- Emerson, P. M. & Souza, A. P. (2006). Is child labor harmful? the impact of working earlier in life on adult earnings. Anais do congresso, Latin America Econometric Society Meeting (LAMES 2006), Cidade do México.



- Ferreira, S. & Veloso, F. A. (2003). Mobilidade intergeracional de educação no Brasil. *Pesquisa e Planejamento Econômico – IPEA*, 33(3).
- Filmer, D. & Pritchett, L. (1998). The effect of household wealth on educational attainment: demographic and health survey evidence. Policy Research Working Paper Series 1980, The World Bank. <http://ideas.repec.org/p/wbk/wbrwps/1980.html>.
- Fletcher, P. R. (1985). A repetência no ensino de 1o. grau: um problema negligenciado da educação Brasileira. *Revista Brasileira de Administração da Educação*, 3(1).
- Gomes-Neto, J. B. & Hanushek, E. A. (1994). Causes and consequences of grade repetition: Evidence from Brazil. *Economic Development and Cultural Change*, 43(1):117–48. <http://ideas.repec.org/a/ucp/ecdecc/v43y1994i1p117-48.html>.
- Goux, D. & Maurin, E. (2005). The effect of overcrowded housing on children's performance at school. *Journal of Public Economics*, 89(5–6):797–819. <http://ideas.repec.org/a/eee/pubeco/v89y2005i5-6p797-819.html>.
- Gusso, D. A. (1993). A educação básica. In Albuquerque, R. C., editor, *O Brasil social: realidades, desafios, opções*, number 139 in IPEA. IPEA.
- Hanushek, E. A. (1992). The trade-off between child quantity and quality. *Journal of Political Economy*, 100(1):84–117. <http://ideas.repec.org/a/ucp/jpolec/v100y1992i1p84-117.html>.
- Hasenbalg, C. (2003). Estatísticas do século xx: Educação. In IBGE, editor, *Estatísticas do século XX*. CDDI/IBGE.
- Haveman, R. & Wolfe, B. (1995). The determinants of children's attainments: A review of methods and findings. *Journal of Economic Literature*, 33(4):1829–1878. <http://ideas.repec.org/a/aea/jecolit/v33y1995i4p1829-1878.html>.
- Horowitz, A. W. & Souza, A. P. (2004). The dispersion of intra-household human capital across children: A measurement strategy and evidence. Working Papers 0408, Department of Economics, Vanderbilt University. <http://ideas.repec.org/p/van/wpaper/0408.html>.
- Jacoby, H. G. (1994). Borrowing constraints and progress through school: Evidence from Peru. *The Review of Economics and Statistics*, 76(1):151–60. <http://ideas.repec.org/a/tpr/restat/v76y1994i1p151-60.html>.
- Kassouf, A. L. (2001). Trabalho infantil. In Lisboa, M. d. B. & Menezes-Filho, N. A., editors, *Microeconomia e sociedade no Brasil*. Contra Capa.
- Leon, F. L. & Menezes-Filho, N. A. (2002). Reprovação, avanço e evasão escolar no Brasil. *Pesquisa e Planejamento Econômico – IPEA*, 32(3).
- Machado, D. C. (2005). *Escolaridade das Crianças no Brasil: três ensaios sobre a defasagem idade-série*. Tese de Doutorado, Departamento de Economia da PUC-RJ.
- Marteletto, L. J. (2004). Desigualdade intergeracional de oportunidades educacionais: uma análise da matrícula e escolaridade das crianças brasileiras. Texto para discussão 242, UFMG/CEDEPLAR, Belo Horizonte.
- Maurin, E. (2002). The impact of parental income on early schooling transitions: A re-examination using data over three generations. *Journal of Public Economics*, 85(3):301–332. <http://ideas.repec.org/a/eee/pubeco/v85y2002i3p301-332.html>.

- Mayer, S. (1997). *What money can't buy: family income and children's life chances*. Harvard University Press.
- Menezes-Filho, N. A. (2001). Educação e desigualdade. In Lisboa, M. d. B. & Menezes-Filho, N. A., editors, *Microeconomia e sociedade no Brasil*. Contra Capa.
- Menezes-Filho, N. A. (2002). Equações de rendimentos: questões metodológicas. In Corseuil, C. H. & alli, editors, *Estrutura Salarial: aspectos conceituais e novos resultados para o Brasil*. IPEA/Ministério do Trabalho.
- Menezes-Filho, N. A. & alli (2000). Explaining the time allocation decisions of adolescents in latin américa and the caribbean. mimeo.
- Nunes, L. C. P. (1996). Ensino de primeiro grau e qualidade no Brasil hoje. *Economia brasileira em perspectiva*, 2.
- Psacharopoulos, G. & Arriagada, A. M. (1989). The determinants of early age human capital formation: Evidence from Brazil. *Economic Development and Cultural Change*, 37(4):683-708. <http://ideas.repec.org/a/ucp/ecdecc/v37y1989i4p683-708.html>.
- Ribeiro, S. C. (1991). A pedagogia da repetência. *Estudos Avançados*, 5(12).
- Rios-Neto, E. L. G., César, C. C., & Riani, J. L. R. (2002). Estratificação educacional e progressão escolar por série no Brasil. *Pesquisa e Planejamento Econômico*, 32(3).
- Saha, R. (2004). Educational grade gaps of boys and girls in Brazil. mimeo.
- Shea, J. (2000). Does parents' money matter? *Journal of Public Economics*, 77(2):155-184. <http://ideas.repec.org/a/eee/pubeco/v77y2000i2p155-184.html>.
- Vasconcellos, L. (2005). A relação entre frequência escolar e renda familiar no Brasil – 1981 a 1999. *Pesquisa e Planejamento Econômico*, 35(2).
- Verona, A. P. d. A. (2004). A relação entre fecundidade e educação dos filhos: um experimento natural usando gêmeos. Dissertação de mestrado, CEDEPLAR/FEA-UFMG.
- Wolfe, J. R. (1982). The impact of family resources on childhood iq. *The Journal of Human Resources*, 17(2).



## A. BASE DE DADOS E CONCEITOS UTILIZADOS

### A.1. Base de dados

A fonte de dados é a Pesquisa Nacional por Amostra de Domicílios de 1996 (PNAD/IBGE) que contém informações sobre as condições gerais do domicílio e da família, sobre as características gerais e educacionais de todos os moradores e um levantamento sobre mobilidade social, realizado para as pessoas com mais de 15 anos (pessoa de referência ou cônjuge). Quem respondia esta parte do questionário fornecia o nível de instrução da mãe e do pai<sup>30</sup> e a ocupação exercida por este último quando o entrevistado tinha 15 anos de idade. No caso do pai ter falecido antes de a pessoa ter atingido essa idade, era reportada a última ocupação exercida. Restringimos nossa amostra às crianças que eram filhos de 7 a 14 anos de idade,<sup>31</sup> moradoras da área urbana,<sup>32</sup> que tinham pai e mãe presentes no domicílio (20.332 crianças). Os pais respondem ao questionário de mobilidade social. Portanto, para cada criança, obtemos também dados sobre os pais de seus pais, ou seja, seus avós paternos e maternos. Para cada avô, temos informações sobre a sua ocupação, e para ambos, avôs e avós, indicadores educacionais. Na tabela 5 apresentamos as estatísticas descritivas da amostra pesquisada.

### A.2. Conceitos

Criamos uma variável categórica (*defasagem idade-série*) igual a um para todas as crianças da nossa amostra que: (i) freqüentavam a escola, mas não tinham os anos de estudos completos compatíveis com a sua idade e (ii) não freqüentavam a escola e não tinham os anos de estudos apropriados para a sua idade.<sup>33</sup> No total da amostra selecionada 35,5% das crianças têm defasagem idade-série.

São consideradas as seguintes variáveis explanatórias da probabilidade de atraso educacional:

- Características individuais das crianças e jovens:
  - (a) Variável binária para sexo, com valor 1 para as meninas (*ser menina*);
  - (b) Variável categórica indicando que a criança se autodeclara de cor branca ou amarela, tem valor nulo quando a cor declarada é preta, parda, indígena ou ignorada (*ser de cor branca ou amarela*);
  - (c) Idade da criança; como a data de referência da pesquisa é setembro,<sup>34</sup> normalizamos a idade para o dia 1º. de março de 1996. O cálculo foi feito da seguinte forma: (i) para as crianças que nasceram antes da data de referência da pesquisa, usamos a idade reportada na PNAD somada aos meses e dias até 1º. de março de 1996; (ii) para as que nasceram depois, diminuimos da idade reportada na PNAD meses e dias até 1º. de março de 1996. Separamos a idade da criança em duas parcelas: a parte não inteira da idade da criança no dia 1º. de março de 1996 (*maturidade*) e a parte inteira da idade da criança no dia 1º. de março de 1996 (*idade completa*). Com essas duas variáveis, introduzimos na análise de progresso no sistema educacional uma variável pouco utilizada na literatura empírica sobre educação no Brasil que é o conceito de maturidade da criança no ano,<sup>35</sup> construída a partir do dia e do mês de nascimento. No nosso país, o dia e mês do nascimento da

<sup>30</sup> A mãe (pai) é a mulher (homem) responsável pela criação da pessoa, mesmo sem ser a genitora (o genitor).

<sup>31</sup> Anos completos em 1º. de março de 1996.

<sup>32</sup> Optamos por apresentar apenas os resultados da área urbana, pois os erros de mensuração da renda no meio rural são potencialmente maiores.

<sup>33</sup> Uma criança que segue um padrão de escolaridade normal entra na escola com 7 anos de idade e termina a 8ª. série do primeiro grau com 15 anos. Seguimos este padrão para definir a defasagem idade-série.

<sup>34</sup> Para a normalização consideramos o dia 25 de setembro de 1996.

<sup>35</sup> Alguns autores, como Maurin (2002), Cogneau e Maurin (2001) e Goux e Maurin (2005) usam definições similares.

criança afetam a decisão dos pais de inserir o filho na escola, pois há uma lei obrigando crianças com 7 anos de idade a ingressarem no sistema escolar e aí permanecerem até os 14 anos de idade. Usualmente, a regra adotada pelas escolas públicas é a criança fazer 7 anos no primeiro semestre do ano, o que a torna apta a ingressar na escola. Se ela faz 7 anos após o primeiro semestre, os pais somente são obrigados a inserir a criança na escola no ano seguinte, por isso escolhemos a normalização para o dia 1º. de março.<sup>36</sup>

- Características do domicílio *de moradia*
  - (a) Variável categórica indicando que o domicílio situa-se na área metropolitana (*área metropolitana*);
  - (b) Variáveis categóricas indicando a região na qual o domicílio está localizado (Norte<sup>37</sup> e Centro Oeste, Nordeste, Sudeste e Sul).
- Características da família:
  - (a) Logaritmo da renda total familiar per capita (*log da renda familiar per capita*), a renda total familiar per capita foi construída a partir das informações de todos seus componentes e de todas as fontes de renda. A renda dos avós somente entra no cálculo da renda total se eles ainda morarem no domicílio, tendo em vista que não temos essa informação através do questionário de mobilidade social, apenas do seu nível educacional e do status ocupacional dos avós.
  - (b) Anos completos de estudo da mãe e do pai;
  - (c) Total de irmãos e irmãs mais novos e mais velhos.
  - (d) Diferença, em módulo, entre as idades das mães e dos pais;
  - (e) Variáveis categóricas indicando o grau de instrução dos avós e dos avôs (3 variáveis categóricas para cada um dos 4 avós).

### A.3. Algumas estatísticas descritivas

Como pode ser visto na tabela 5, 35,5% das crianças da amostra selecionada possuíam defasagem idade-série. A distribuição das crianças da amostra por sexo era praticamente igual, 49,7% eram meninas. Em termos de cor, mais da metade era branca ou amarela (61,3%). A idade média das crianças normalizada para o dia 1º. de março era de 10,53, a idade máxima era de 14 e a idade mínima de 7 anos.

Com relação à localização dos domicílios onde as crianças moravam, observamos que a maioria não estava nas áreas metropolitanas (66%) e quase 50% ficavam na região Sudeste.

A média da renda familiar per capita, em setembro de 1996, estava em torno de 125 reais, variando de zero a 6 mil reais.

Em média, as crianças tinham um irmão ou uma irmã, mas velha ou mais nova. A média dos irmãos mais velhos era mais alta, pois a amostra foi formada por crianças entre 7 e 14 anos de idade.

A média de escolaridade dos pais, tanto das mães quanto dos pais, estava em 6 anos de estudos, ou seja, menos que o ensino fundamental completo. Há uma grande disparidade entre os pais com relação aos anos de estudo completos – o desvio padrão da média de escolaridade ficou em 4,4 para as mães e 4,6 para os pais.

Os indicadores de escolaridade dos avós e das avôs mostram que a maioria não estudou e que isto é mais grave para as avôs. Cerca de 45% das avôs do lado paterno e do lado materno não estudaram. Estas porcentagens para os avôs foram de 37% e 39% para o lado materno e paterno, respectivamente.

<sup>36</sup>Também fizemos as estimações normalizando a idade para o dia 1 de julho de 1996, os resultados não modificaram.

<sup>37</sup>A Pnad de 1996 não investigou a área rural da região norte, mas como nos restringimos à área urbana, isso não afeta nossos resultados.

**Tabela 5** – Estatísticas descritivas da amostra

Variáveis	Média	dp	Mínimo	Máximo
ter defasagem idade série	0,355	0,478	0	1
ser menina	0,497	0,500	0	1
ser de cor branca ou amarela	0,613	0,487	0	1
idade em 1o. de março				
parte inteira da idade	10,538	2,301	7	14
maturidade	0,504	0,288	0	0,997
domicílio localizado na:				
área metropolitana	0,336	0,472	0	1
região norte ou centro-oeste	0,122	0,328	0	1
região nordeste	0,222	0,415	0	1
região sudeste	0,488	0,500	0	1
região sul	0,168	0,374	0	1
ln da renda familiar per capita	4,830	1,265	0	8,719
diferença entre a idade do pai e da mãe	4,951	4,467	0	35
total de irmãs e irmãos mais novos	1,021	1,123	0	9
total de irmãs e irmãos mais velhos	0,975	1,065	0	9
total de irmãs mais velhas	0,464	0,706	0	6
anos de estudos completos da mãe	6,166	4,402	0	17
anos de estudos completos do pai	6,170	4,614	0	17
grau de instrução do avô do lado paterno				
zero ano de estudo	0,390	0,488	0	1
1 a 3 anos de estudo	0,306	0,461	0	1
4 ou mais anos de estudo	0,305	0,460	0	1
grau de instrução da avó do lado materno				
zero ano de estudo	0,445	0,497	0	1
1 a 3 anos de estudo	0,265	0,441	0	1
4 ou mais anos de estudo	0,291	0,454	0	1
grau de instrução da avó do lado paterno				
zero ano de estudo	0,456	0,498	0	1
1 a 3 anos de estudo	0,265	0,441	0	1
4 ou mais anos de estudo	0,280	0,449	0	1
grau de instrução do avô do lado materno				
zero ano de estudo	0,377	0,485	0	1
1 a 3 anos de estudo	0,303	0,459	0	1
4 ou mais anos de estudo	0,320	0,467	0	1

Fonte: Pnad/1996 (IBGE). Obs: (1) A amostra é constituída de 20.332 crianças que quando expandida corresponde a 9.598.694 crianças. (2) dp = desvio padrão