

## 2

# O impacto dos fatores familiares sobre a defasagem idade-série de crianças no Brasil

### 2.1

#### Introdução

Apesar da inserção da maioria das crianças na escola, muitas não progridem ao longo do sistema educacional de forma contínua e adequada (*Gomes-Neto e Hanushek, 1994; Dureya, 1998 e Leon e Menezes-Filho, 2002; Barros e Mendonça 1998*). 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>2</sup> Crianças que, de acordo com a sua idade, deveriam frequentar uma determinada série escolar estão em uma inferior, acumulando o que denominamos defasagem idade-série ou atraso educacional. Não alcançam os níveis de escolaridade compatíveis com sua faixa etária, permanecendo mais tempo que o estipulado para uma etapa, sobretudo nas últimas séries do ciclo escolar (*Ribeiro, 1991; Fletcher, 1985*).

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 status sócio-econômico e educacional está usualmente positivamente correlacionado entre as gerações. O ambiente familiar e as características dos pais têm impacto no comportamento e 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 para passarem mais tempo na escola. Nestes casos, o custo de permanência nesta etapa de estudante não seria restritivo e as chances de

---

<sup>2</sup> *Leon e Menezes-Filho (2002)* mostram que um dos fatores capazes de explicar a evasão é a repetição. A proporção de alunos que interrompe os estudos após experiências de repetência é maior que daqueles que saem da escola sem terem repetido. *Bonelli e Veiga (2004)* fazem uma descrição da evasão escolar de jovens no Brasil.

tornarem-se pobres seriam usualmente ínfimas.<sup>3</sup> Um dos grandes desafios aos formuladores de política educacional atualmente é fornecer às crianças pobres as condições para entrarem na idade adequada na escola e progredirem ao longo do sistema educacional sem interrupções, assimilando conhecimentos com qualidade.

No Brasil, há uma vasta literatura empírica sobre as relações existentes entre escolaridade das crianças e os 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*). Existe outra vertente que estuda a relação entre desigualdade de oportunidades e desigualdade de renda, por exemplo *Bourguignon, Ferreira e Menéndez (2003) e Menezes-Filho (2001)*. E uma terceira linha que enfatiza a relação entre trabalho infantil, pobreza e escolaridade (*Emerson e Souza, 2000, 2002, 2004; Kassouf, 2001; Barros, Mendonça e Velazco, 1996, Menezes-Filho et alii, 2000, dentre outros*).

Neste capítulo, analisamos a influência dos fatores familiares na escolaridade de uma amostra de crianças que ainda não finalizaram o ciclo educacional usando os dados da Pesquisa Nacional por Amostra de Domicílios de 1996 (PNAD). Não abordamos o nível de escolaridade atingido, mas uma medida que capta aspectos quantitativos ao longo do processo escolar: a defasagem idade-série.<sup>4</sup> Nosso objetivo é mostrar o efeito da renda familiar e do nível educacional dos pais sobre a probabilidade de as crianças terem defasagem idade-série usando o procedimento de variáveis instrumentais (IV) (*Angrist e Krueger, 2001*). Como a implementação de políticas que visam à ampliação da escolaridade também perpassa pelo entendimento dos mecanismos de transmissão da educação entre as gerações, iremos investigar esses efeitos acima considerando esses aspectos.

---

<sup>3</sup> *Haveman e Wolfe (1995)* fazem uma resenha sobre o impacto da família nas crianças (escolaridade, renda, etc)

<sup>4</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 tem como estabelecida o fato de que quanto maior o atraso educacional menor o nível de escolaridade atingido. Logo, uma política que vise combater as desigualdades educacionais deveria também abordar as desigualdades no progresso ao longo do sistema escolar. Apesar da defasagem idade-série também refletir aspectos relacionados à qualidade da escola, essas relações podem ser confusas tendo em vista que no caso brasileiro, escolas de baixa qualidade e com poucos recursos podem adotar a promoção automática. Por outro lado, em escolas de melhor qualidade, as regras de promoção podem ser mais rígidas, sendo o atraso da criança não necessariamente condicionado à falta de recursos familiares ou escolares.

Selecionamos três grupos de instrumentos: a oferta educacional da segunda geração (os pais das crianças),<sup>5</sup> os fatores familiares entre as gerações<sup>6</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 al (2003), Chevalier (2004), Chevalier et al. (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 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.

Na próxima seção, apresentamos uma resenha sobre os principais artigos empíricos produzidos nesta área, focando particularmente nas estratégias econométricas adotadas para separar o efeito renda e educação dos pais; na seção 2.3, descrevemos o modelo que fundamenta a existência dos vínculos intergeracionais na transmissão de capital humano e os principais canais de endogeneidade; na seção 2.4, descrevemos nossos dados, os conceitos e fazemos uma análise descritiva das correlações existentes entre as três gerações (crianças, pais e avós); na seção 2.5, apresentamos a estratégia econométrica; e na seção 2.6, os resultados. Na seção 2.7, traçamos nossas considerações finais.

---

<sup>5</sup> Alguns autores também seguem essa linha, como *Emerson e Souza (2004)* e *Currie e Moretti (2003)*.

<sup>6</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).

## 2.2 Resenha da literatura

Vários artigos brasileiros examinam diferentes fenômenos relacionados à escolaridade do público infantil e suas características familiares (*Psacharopoulos e Arriagada, 1989; Dureya, 1998; Saha, 2004; Barros e Lam, 1993, etc.*). No tocante ao grau de instrução dos pais, acham que a sua influência sobre a participação das crianças na escola é crucial e que o efeito da escolaridade materna é mais forte que o paterno, sobretudo nas meninas. Alguns estudos, por sua vez, estão preocupados em mostrar a importância de um fator face a outros aspectos que também afetam a escolaridade da criança (*Rios-Neto, César e Riani, 2002*), ou, em identificar os grupos com maiores dificuldades no processo de escolarização.

A literatura de demanda educacional também está fortemente atrelada às análises de trabalho infantil. Muitas crianças não freqüentam a escola ou não podem se dedicar integralmente ao aprendizado porque no mesmo período estão exercendo alguma atividade econômica seja no mercado de trabalho ou no domicílio.<sup>7</sup> A principal questão é investigar as razões pelas quais os pais preferem as crianças trabalhando ao invés de se dedicando aos estudos (*Psacharopoulos e Arraigada, 1989; Kassouf, 2001, Barros, Mendonça e Velazco, 1996; Menezes-Filho et alii, 2000*).

Assim, 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*). Essas evidências justificam a adoção generalizada de programas de apoio financeiro às famílias situadas na escala inferior da distribuição de renda.

Contudo, não é claro que os resultados das crianças realmente melhoram devido aos benefícios monetários recebidos ou ao fato dos pais adquirirem mais escolaridade. Uma das preocupações básicas da literatura de demanda educacional

---

<sup>7</sup> *Emerson e Souza (2004); Kassouf (2001); Barros, Mendonça e Velazco (1996)*, etc. Na literatura internacional, ver *Basu (1999)* e *Basu e Tzannatos (2003)*..

é investigar o real impacto de fatores familiares, como renda e escolaridade dos pais sobre os resultados em termos de aprendizado.<sup>8</sup>

Os acréscimos na renda familiar poderiam 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). Neste caso, ter mais recursos afetaria positivamente o processo de aprendizagem das crianças. Entretanto, crianças de famílias mais ricas estão usualmente inseridas em um contexto sócio-econômico e cultural favorável ao acúmulo de capital humano (melhores escolas próximas ao local de moradia, contato com pessoas instruídas no ambiente familiar, etc). Essas facilidades podem ser transmitidas de pais para filhos. Independentemente dos adicionais de renda em um determinado momento, as crianças de famílias mais pobres poderiam continuar em uma situação educacional inferior aos filhos dos mais ricos.<sup>9</sup>

No tocante ao grau de escolaridade dos pais, seu efeito pode ser direto ou indireto sobre o aprendizado de seus filhos. As escolaridades da mãe ou do pai, segundo *Currie e Moretti (2003)*, influenciariam o aprendizado da criança por quatro potenciais canais:

(i) Como mães/pais mais escolarizados têm rendimentos do trabalho mais altos, podem adquirir mais bens para o aprendizado de seus filhos;

(ii) Uma mãe/pai mais escolarizada/o geralmente tem como parceiro/a alguém com nível de escolaridade similar. Homens mais educados usualmente se casam com mulheres mais educadas e vice-versa, ou seja, a educação afeta a escolha dos parceiros. O parceiro/a com grau de instrução mais alto contribui positivamente para o volume de rendimentos na família, logo, mais recursos podem ser gastos na escolaridade da criança;

(iii) Pais ou mães mais educados tendem a ter um padrão de comportamento em que nutrem mais expectativas em relação à escolaridade dos filhos, incentivando-os mais no processo de aprendizagem, seguindo a literatura sociológica (*Marteletto, 2004*);

---

<sup>8</sup> A literatura utiliza diferentes medidas de escolaridade da criança, como testes de proficiência, número de repetências, evasão e frequência escolar, etc. Ao longo desta seção, apesar de diferentes óticas do aprendizado, trataremos todas igualmente.

<sup>9</sup> Vários artigos discutem a relação entre renda e escolaridade: *Blau (1999)*, *Shea (2000)*, *Goux e Maurin (2005)*, *Wolfe (1982)*.

(iv) Pais mais educados possuem menos filhos, preocupando-se mais com a qualidade da criança, em termos de bem-estar e capital humano. De acordo com *Hanushek (1992)*, há um *trade-off* entre qualidade e quantidade de filhos.

Além desses quatro aspectos acima, como pais mais instruídos têm mais contato com o cenário social e cultural, sua capacidade e facilidade de transmissão de conhecimentos são potencialmente maiores.<sup>10</sup>

As externalidades associadas à escolaridade de uma geração, ou seja, o retorno social da educação, podem ser amplas, impactando positivamente no aprendizado dos seus descendentes. Na literatura, este argumento prioriza o *spillover* da expansão educacional, com conseqüências de longo prazo estendidas a toda sociedade, *Black et al. (2003)*.

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 seria espúria. O vínculo entre pais e filhos seria explicado por aspectos não controlados pelo formulador de políticas e não refletiria 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, também seriam mais hábeis, conforme essa abordagem.

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.

Na literatura internacional, duas estratégias econométricas são usualmente implementadas para tratar esse problema:<sup>11</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

---

<sup>10</sup> Esses pais são mais produtivos no processo de aprendizado, o chamado efeito “criação”, conforme *Chevalier et alii (2005)*.

<sup>11</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 é positivo e significativo para os 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)*.

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>12</sup> os artigos inserem-se na literatura de avaliação de políticas. Os instrumentos utilizados são fundamentados em experimentos institucionais, reformas que afetam de forma exógena o grau de instrução dos pais e a renda familiar.

*Carvalho* (2000)<sup>13</sup> adota essa estratégia para o caso brasileiro. Analisa 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 na oferta de trabalho dos homens e sobre a participação no mercado de trabalho e na escola das crianças. O resultado mostra 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 às modificações no sistema educacional, correlacionados com o grau de instrução, mas não correlacionados com fatores não observados (*Black et al., 2003; Chevalier, 2004; Chevalier et al., 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. Vários são os artigos internacionais que procuram identificar o efeito puro da renda e da escolaridade dos pais sobre os diferentes aspectos escolares da criança a partir da seleção de instrumentos específicos.

No quadro 1 da página 30 apresentamos um resumo de alguns artigos que selecionaram instrumentos para estimar o impacto da renda ou do nível educacional dos pais sobre diferentes óticas da escolaridade (proficiência, evasão escolar, progresso educacional, etc.) ou da qualidade de vida das crianças (se

---

<sup>12</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>13</sup> *Duflo* (2000) 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.



trabalham ou não, rendimento quando adultas, se tiveram filho na adolescência, etc).

Os seis primeiros artigos enfatizam a estimação do efeito renda puro. Abordaremos cinco deles e discutiremos *Blau (1999)* mais à frente.

*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” da empresa, na qual o pai trabalha, continuar operando. Seu principal resultado é que para pais com pouca instrução a renda é muito importante para o investimento em capital humano infantil.

Dois artigos, *Maurin (2002)* e *Couralet (2002)* usam como instrumentos características educacionais e ocupacionais dos pais e dos avós. O primeiro está particularmente interessado em testar a existência de fatores hereditários nas transições escolares. 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. Mostra 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.

*Mayer (1997)* e *Dahl e Lochner (2005)* utilizam variações de benefícios recebidos pela família para identificar o efeito da renda sobre a proficiência e sobre diferentes características das crianças, respectivamente. As diferenças nos benefícios recebidos não estariam correlacionadas com fatores não observados que afetam a família.

Os artigos seguintes abordam o problema de endogeneidade da educação dos pais e da renda familiar. *Dumas e Lambert (2005)* analisam a probabilidade da criança ir para a escola e o nível de escolaridade atingido quando adultas. Focam no efeito da educação dos pais e da riqueza familiar. Usam um grupo de instrumentos que caracterizam a área de residência dos pais quando eles tinham



10 anos de idade (presença de infra-estrutura) e as condições de saúde de seus pais. *Chevalier et al* (2005) usam a 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 al (2003)*, a reforma institucional do sistema educacional. *Emerson e Souza (2004)* investigam se o fato de um adulto ter trabalhado quando criança prejudica a sua aquisição de rendimentos. Utiliza um conjunto de instrumentos (oferta educacional na infância) correlacionados com as decisões escolares das crianças, mas não relacionado aos 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:

$$(1) \text{educ}_{if} = Z_{if} \beta + u_f + v_{if}$$

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 (habilidade, etc) e uma parte particular à criança e aleatória ( $v_{if}$ ). Quando tomamos a diferença, controlamos por este efeito fixo da família. *Cogneau e Maurin (2001)* adotam essa estratégia para identificar o impacto dos recursos dos pais 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.

Nesta linha, *Blau* (1999) 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. Estimam 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.

No Brasil, com relação à renda familiar, alguns trabalhos<sup>14</sup> relatam o problema de endogeneidade, devido à possibilidade de existência de variáveis omitidas (como habilidade) ou erros de medida. Em 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.

*Couralet* (2002) também aplica a mesma metodologia de *Cogneau e Maurin* (2001) para estudar a probabilidade das crianças trabalharem ou não 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 com a 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 problema da endogeneidade na determinação do aprendizado da criança deve ser abordado de forma diferente dos

---

<sup>14</sup> Um trabalho já discutido acima é *Carvalho* (2000).

estudos sobre retornos à escolaridade do próprio indivíduo no mercado de trabalho. Nestes, a equação de rendimento do trabalho é influenciada diretamente pela habilidade individual, valorizada no mercado de trabalho à medida que também afeta a capacidade produtiva. Teríamos um problema de endogeneidade causado por variável omitida. A habilidade omitida da equação de rendimento do trabalho afetaria tanto o total de anos de estudos quanto o rendimento auferido na atividade econômica (*Emerson e Souza, 2004; Menezes-Filho, 2002*).

No nosso caso, o efeito da habilidade individual dos pais sobre o aprendizado das crianças se reflete 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 e que são passadas de uma geração para outra, influenciando tanto a escolaridade dos pais quanto diretamente a escolaridade das crianças.

Quadro 1: Resumo dos artigos que utilizam a técnica de variáveis instrumentais para investigar características das crianças			
Artigo	Variável dependente	Instrumentos	Justificativa
Shea (2000)	Anos de estudo	- Pai ser sindicalizado e estar na indústria - Perda involuntária do emprego	Variações na renda não correlacionadas com habilidade não observada herdada dos pais. Pais sindicalizados não são mais hábeis que pais não sindicalizados com as mesmas características observáveis.
Mayer (1997)	Anos de estudo, renda quando adultas, ter filho na adolescência, ser mãe solteira, evadir da escola	- Crescimento da renda dos pais depois do resultado da criança ser efetivado. - Variações estaduais do benefício para famílias monoparentais com crianças	- Crescimento da renda futura não é antecipado, ou, é antecipado mas as respostas das pessoas às mudanças nas rendas futuras são aleatórias com respeito às características não observadas que determinam os resultados das crianças. - As diferenças interestaduais entre as crianças de famílias monoparentais e com ambos os pais são apenas afetadas pelas diferenças no nível dos benefícios.
Maurin (2002)	Transições escolares	- Educação dos pais e status sócio econômico dos avós.	- Componente permanente da renda e testar a existência de determinantes hereditários.
Biau (1999)	Desenvolvimento cognitivo da criança	- Três alternativas de modelo de efeitos fixos (efeito fixo da família): - Mãe tem uma irmã, crianças da mesma mãe, criança em dois pontos no tempo	- Efeito fixo é considerado a única fonte de correlação entre os regressores e o termo de erro.
Couralet (2002)	Probabilidade da criança trabalhar	- Anos de estudos dos pais e renda familiar dos avós (proxy de categoria sócio profissional). - Educação dos avós. - Diferença dos anos de estudos entre pais e avós.	- As preferências dos pais não são transmitidas de uma geração a outra e a qualidade da educação de duas gerações sucessivas de um mesmo domicílio não são correlacionadas. - Supõe que as transmissões de pais para filhos esgotam o viés de hereditariedade. - Se o resíduo da equação de escolarização é a soma de uma componente própria a uma geração (não correlacionada com os resíduos dos períodos anteriores) e uma componente dinâmica (esse efeito é expurgado pela diferença).
Dahl e Lochner (2005)	Proficiência de matemática e leitura das crianças	- Mudanças não lineares no Earned Income tax credit (EITC) e na estrutura nacional de ganhos ao longo do tempo.	- Mudanças não correlacionadas com choques idiossincráticos na família - O aumento do retorno à educação ao longo das duas últimas décadas. Renda familiar cresceu mais para as famílias com nível educacional mais alto.
Dumas e Lambert (2005)	Probabilidade de ir à escola e nível de escolaridade atingido	- <i>Dummies</i> indicando a presença de escolas primárias, secundárias na vizinhança quando o pai tinha 10 anos de idade, se os pais viviam na zona rural ou urbana, existência de centro de saúde na comunidade. - Para cada um dos pais, sabemos se seus pais estavam vivos quando tinham 10 anos de idade. Se sofriam algum problema grave de saúde que o impedisse de trabalhar. - Ranking dos pais em relação a seus irmãos. Duas variáveis: não tem irmã mais nova e não tem irmão mais novo.	- Famílias mais ricas têm mais acesso aos serviços de saúde, logo tem uma probabilidade maior de ter uma melhor saúde e de viver mais. A probabilidade de morrer ou de ficar doente antes da criança ter 10 anos de idade não está correlacionada com a riqueza. - Saúde não está correlacionada com preferências e habilidade, sendo explicada principalmente pela idade e por choques - O único rank exógeno é o filho mais velho. O mais novo não é exógeno pois os pais decidiram não ter mais filhos.
Chevalier et al (2005)	Continuar na escola após terminar escolaridade obrigatória Proficiência	- Mudança da escolaridade obrigatória: aumento da idade de saída. - Status sindical do pai	- Mudança legal gera uma descontinuidade nos anos de estudos do pai independente da sua habilidade. - Variação da renda provocada por participar ou não de um sindicato é exógena à habilidade do pai que afeta a escolaridade do filho.
Cogneau e Maurin (2001)	Matricular a criança na idade correta (6 anos) e enviar a criança para a escola antes dos 9 anos.	- Diferença entre o nível educacional dos pais e dos avós. - Diferença entre o setor de atividade dos pais e dos avós (agrícola vs. não agrícola)	- Resultado passado da família que não está correlacionado com choques transitórios presentes no termo de erro. - Tirando a primeira diferença desses resultados retiramos as características permanentes não observadas que afetam simultaneamente as decisões da família e seus recursos
Emerson e Souza (2004)	Efeito do trabalho infantil na renda	- Número de escolas primárias, secundárias e universidades para cada estado de 1933 a 1976. - Número de professores por escola em cada estado e cada ano. - Reforma do sistema educacional	- As variáveis institucionais que selecionam como instrumentos estão correlacionadas com as decisões escolares e de trabalho das crianças, mas não correlacionadas com a variação não explicada dos ganhos dos adultos.
Black et al. (2003)	Nível educacional das crianças		- Variações temporais e geográficas independem da habilidade dos pais

## 2.3

### Breves considerações sobre a transmissão intergeracional

#### 2.3.1

##### Um modelo empírico<sup>15</sup>

Para ilustrar a dinâmica do acúmulo de capital humano e os vínculos existentes entre as três gerações (filhos, pais e avós) de uma mesma dinastia, descrevemos brevemente o modelo empírico de *Maurin* (2002) com algumas alterações. Incluímos 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.

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 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 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 função de aspectos escolares, familiares e individuais.<sup>16</sup>
- Os pais, crianças da geração  $t$ , adultos em  $t+1$ , decidem o volume de recursos alocados para educação das crianças e para o consumo, bem como outros fatores que afetam a capacidade de aprendizado de seus filhos (entre eles, seu grau de instrução). Essas escolhas são feitas com base na maximização da

<sup>15</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>16</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.

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 eq. (2) a eq. (6) abaixo:

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

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

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

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

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

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 um representa as crianças, enquanto o subscrito dois, os adultos. O aprendizado das crianças de cada geração, descritos 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  $u$ .

Para a geração de crianças em  $t+1$ , a renda, o nível de escolaridade de seus pais ( $Y_{2t+1}$  e  $educ_{1t}$ , respectivamente),  $X_{1t+1}$  e outros fatores agrupados em  $u_{t+1}$  determinam o total de anos de estudos.

Fatores aleatórios que afetam a escolaridade da geração 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, 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. Esse componente de aleatoriedade é individual.
- Específicos à dinastia, pois cada uma tem uma forma peculiar de cuidar de seus filhos. 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). É claro que, dependendo da família, alguns desses aspectos podem ser mais ou menos fortes de geração para geração.
- Específicos à geração, pois ao longo das décadas, as preferências por escolarização e a oferta educacional variam, afetando de forma significativa decisões de uma geração. Existem choques na oferta educacional que são específicos a uma geração, por exemplo, o aumento do volume de recursos para construção de escolas de nível médio. O aumento da valorização da escolaridade no mercado de trabalho também cria mais incentivos para completar um maior número de anos de estudos.

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 aleatórios que afetam o nível de rendimentos dos pais e avós 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 eq. (2) a eq. (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. Na próxima sub-seção aprofundaremos mais essa questão.



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. Para as que têm mais de sete anos de idade, podemos definir a defasagem idade-série num determinado período  $\tau$  de sua vida conforme eq. (7) abaixo:<sup>17</sup>

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

Onde:

$educ_{1,t+1,\tau}$  e  $Idade_{1,t+1,\tau}$  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:  $educ_{1,t+1,\tau} = educ_{1,t+1}$ . 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 completo da criança nesse período  $\tau$  é uma parte da sua escolaridade 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.

<sup>17</sup> Seguimos a sugestão de *Horowitz e Souza (2004)* para 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.

Assim, a probabilidade da criança no período  $\tau$  ter defasagem idade série é também função dos aspectos que influenciam o nível educacional desejado, conforme eq. (8) abaixo:

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

O denominador da eq. (7), formado pela idade da criança no dia primeiro de março e pela idade de entrada legal no sistema educacional, foi incorporado a  $X$ .

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

### 2.3.2

#### 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), pág. 32. 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>18</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.

<sup>18</sup> Segundo Cogneau e Maurin (2001) e Maurin (2002).

As eqs. (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 familiares 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>19</sup>, de outros fatores familiares (ou dinásticos)  $u_f$  e individuais ( $u_{t+1}^i$ ), temos que:

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

Onde  $E(educ_{1t}, u_f) \neq 0$  e  $E(educ_{1t}, u_{t+1}^i) = 0$ . Ou seja, esse fator familiar explica as diferenças 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 o limite em probabilidade do coeficiente estimado seria igual a eq. (10) abaixo:

(10)

$$p \lim \hat{\beta}_{MQO} = E(educ_{1t} \cdot educ_{1t})^{-1} E(educ_{1t} \cdot educ_{1t+1}) = \beta + E(educ_{1t} \cdot educ_{1t})^{-1} E(educ_{1t} \cdot u_f)$$

---

<sup>19</sup> Excluimos a renda familiar per capita e as características observáveis da criança apenas para simplificar a discussão.

O viés do parâmetro  $\beta$  dependerá da persistência da influência deste termo familiar de uma geração para outra. Por exemplo, uma criança que vive numa família com hábito de leitura e que dá um enorme valor à instrução, teria maiores chances e mais apoio em prosseguir nos estudos do que outra criança numa situação distinta. Mesmo que as duas crianças tenham pais com grau de escolaridade similar, há a possibilidade do aprendizado de ambas se distinguirem face à existência desse componente familiar hereditário.

Em termos analíticos, esse exemplo acima implica a existência de correlação positiva entre  $educ_{1t}$  e  $u_f$ . Conseqüentemente, a direção do viés do coeficiente estimado pelo método de mínimos quadrados ordinários (MQO), segundo eq.(10) da página anterior, seria para cima.

Por outro lado, nada impede que tenhamos uma situação contrária. Por exemplo, uma família que não nutre expectativas quanto aos benefícios da escolaridade para os filhos. A taxa de desconto intertemporal dos pais poderia ser muito elevada e haveria um forte incentivo para que as crianças saíssem cedo da escola e ingressassem no mercado de trabalho. A escolaridade dos pais estaria negativamente associada com esse aspecto familiar, logo, a direção do viés nesse caso seria para baixo. Em suma, a direção do viés do coeficiente da escolaridade dos pais depende do sinal da correlação com esse termo familiar não observado.

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, (b) não seria nulo. 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 letras (a) e (b) seriam iguais.

Além dessas correlações potenciais acima citadas, 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, o 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 mais 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 letra (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 aspectos com exceção da habilidade dos pais, as que têm pais mais talentosos terão mais chances de terem 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 alocarem uma parte maior do seu tempo para contribuir para a aprendizagem das crianças. Por sua vez, a letra (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 é: como estimar os parâmetros  $\alpha$  e  $\beta$  da eq. (8), página 35, 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 depende das hipóteses feitas sobre a estrutura de variância e covariância dos erros das eqs. de (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. *Couralet (2002)*, por exemplo, usa técnica similar para testar a simultaneidade nas decisões das crianças irem para a escola e trabalhar, considerando também a oferta de trabalho de outros membros da família.

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 pode ser um instrumento potencial para o nível de renda da família, mesmo na presença de viés de hereditariedade e simultaneidade. Utilizaremos esse resultado para estimarmos o impacto da renda familiar e da escolaridade dos pais na probabilidade das crianças terem defasagem idade-série.

## 2.4

### Base de dados e conceitos utilizados

#### 2.4.1

##### 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>20</sup> e a ocupação exercida por este último quando o entrevistado tinha 15 anos de idade. No caso do pai ter

---

<sup>20</sup> A mãe (pai) é a mulher (homem) responsável pela criação da pessoa, mesmo sem ser a genitora (o genitor).

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>21</sup> moradoras da área urbana,<sup>22</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 2 abaixo apresentamos as estatísticas descritivas da amostra pesquisada.

Tabela 2: 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: elaboração própria a partir da PNAD/IBGE, 1996.

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

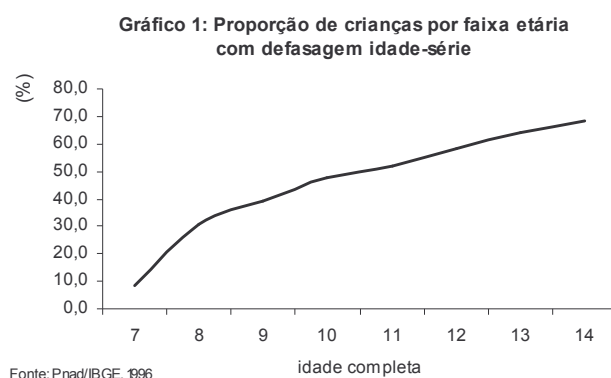
<sup>21</sup> Anos completos em 1º de março de 1996.

<sup>22</sup> Optamos por apresentar apenas os resultados da área urbana, pois os erros de mensuração da renda no meio rural são potencialmente maiores.



## 2.4.2 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>23</sup> No total da amostra selecionada 46,5% das crianças têm defasagem idade-série. Como pode ser visto no gráfico 1, a proporção de crianças de cada faixa etária com defasagem idade-série aumenta continuamente com a idade. Quase 69% das crianças com 14 anos de idade possuem defasamento escolar.



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>24</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

<sup>23</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>24</sup> Para a normalização consideramos o dia 25 de setembro de 1996.

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>25</sup> construída a partir do dia e do mês de nascimento. No nosso país, o dia e mês do nascimento da 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>26</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>27</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 cômputo 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.

<sup>25</sup> Alguns autores, como Maurin (2002), Cogneau e Maurin (2001) e Goux e Maurin (2005) usam definições similares.

<sup>26</sup> Também fizemos as estimações normalizando a idade para o dia 1 de julho de 1996, os resultados não modificaram.

<sup>27</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.

(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, ver Apêndice 7.1.3, pág. 130).

### 2.4.3

#### **Análise descritiva dos dados: matrizes do atraso educacional e da origem familiar**

Nesta sub-seção, faremos uma análise descritiva dos dados priorizando as relações existentes entre as três gerações: crianças, pais e avós. O objetivo é mostrar para quais grupos a herança educacional e o status econômico da família são fatores cruciais para o aprendizado. Ou seja, identificar grupos de crianças onde a vulnerabilidade familiar (baixo grau de escolaridade dos avós e dos pais e baixa renda familiar) é traduzida em maiores chances de ocorrência do atraso educacional.

Construímos “matrizes de mobilidade” que relacionam a origem da criança<sup>28</sup> (em duas perspectivas: educacional e de renda familiar) com sua situação em termos de defasagem escolar. Faremos os seguintes cortes para a origem familiar: sete níveis educacionais<sup>29</sup> (para a mãe, o pai, o avô e a avó do lado paterno e o avô e a avó do lado materno) e dez classes de renda familiar per capita, de acordo com os décimos da sua distribuição.<sup>30</sup>

Cada coeficiente da matriz representa a razão das razões de chances (*odds ratio*), calculada a partir de regressões logísticas cuja variável dependente é se a criança tem ou não defasagem idade-série.

A razão de chances ( $r$ ) é definida como a razão entre a probabilidade do evento  $j$  ocorrer ( $p_j = \frac{n_j}{n}$ ) e a probabilidade do evento não ocorrer ( $1 - p_j$ ). Ou

<sup>28</sup> Dumas e Lambert (2005) fazem análise relativamente similar usando dados de crianças no Senegal.

<sup>29</sup> Sem escolaridade, de 1 a 3 anos de estudos completos, 4 anos de estudos completos, 5 a 7 anos de estudos completos, 8 anos de estudos completos, 9 a 11 anos de estudos completos e 12 ou mais anos de estudos completos.

<sup>30</sup> Para os valores médios da renda familiar per capita por décimos da distribuição ver tabela 23, no Apêndice 7.1.1, pág. 125.

seja, a probabilidade da criança ter defasagem idade-série sobre a probabilidade dela não ter defasagem idade-série ( $r = \frac{n_j}{n - n_j}$ ).

A razão das razões de chances, por sua vez, mede a relação entre as chances do evento ocorrer ou não entre as origens 1 e 2 (classe de renda familiar, nível educacional dos pais e dos avós). Definindo  $r1$  e  $r2$  como a razão de chances da criança de origem familiar 1 ou 2, respectivamente, temos que a razão das razões de chances é  $r1/r2$ .

A primeira célula da tabela 3 significa que a situação na qual uma criança de uma família do primeiro décimo estar atrasada e do segundo décimo não estar atrasada é 1,005 vez mais provável que o inverso: a criança do primeiro décimo não estar atrasada e do segundo décimo estar atrasada.

Tabela 3: Razão das razões de chances de ter ou não defasagem idade-série segundo os décimos da distribuição da renda familiar per capita

	1	2	3	4	5	6	7	8	9
2	1,005 0,049								
3	1,284 0,038	1,277 0,038							
4	1,755 0,028	1,746 0,028	1,367 0,035						
5	2,318 0,021	2,306 0,021	1,806 0,027	1,321 0,036					
6	2,946 0,017	2,931 0,017	2,295 0,022	1,679 0,029	1,271 0,039				
7	3,793 0,014	3,773 0,014	2,955 0,017	2,162 0,023	1,636 0,031	1,287 0,041			
8	4,311 0,012	4,289 0,012	3,358 0,016	2,457 0,021	1,860 0,029	1,463 0,037	1,137 0,049		
9	8,141 0,007	8,098 0,007	6,341 0,009	4,639 0,012	3,511 0,017	2,763 0,022	2,146 0,028	1,888 0,033	
10	12,979 0,005	12,911 0,005	10,110 0,006	7,396 0,009	5,598 0,012	4,405 0,015	3,422 0,020	3,010 0,023	1,594 0,045

Fonte: PNAD/96

Nota: em cada célula, na primeira linha a razão das razões de chances e na segunda linha os erros padrões.

Na primeira coluna, à medida que as crianças se diferenciam mais em termos de renda familiar, a razão das razões de chances aumenta. Comparando uma criança que está no primeiro décimo com uma que está no último décimo, essa razão aumenta para 12,979. Ou seja, há 12,979 mais chances de observamos uma criança de família pobre atrasada e uma criança de família rica não atrasada do que a situação inversa. Essa razão também se reduz à medida que se caminha para o lado direito da tabela 3, pois as diferenças em termos de renda familiar são menores.

O mesmo exercício acima é feito para cada nível educacional do pai, da mãe e dos avós do lado paterno e materno.

Na tabela 4 abaixo, a origem familiar é caracterizada pelo grau de instrução da mãe. Temos uma criança de mãe analfabeta atrasada em termos educacionais e uma de mãe com mais de 12 anos de estudo não atrasada tem 21,811 vezes mais chances de ocorrer que o oposto. A situação em que uma criança cuja mãe tem de 9 a 11 anos de estudo estar atrasada e uma criança cuja mãe tem mais de 12 anos de estudo não estar atrasada acontece 2,037 mais vezes que o inverso. As diferenças maiores encontram-se entre as mães que são analfabetas e que nem finalizaram o ensino primário.

Tabela 4: Razão das razões de chances de ter ou não defasagem idade-série segundo o nível educacional da mãe

	sem estudo	1 a 3 anos de estudo	4 anos de estudo	5 a 7 anos de estudo	8 anos de estudo	9 a 11 anos de estudo
<b>1 a 3 anos de estudo</b>	1,726 0,025					
<b>4 anos de estudo</b>	3,105 0,013	1,799 0,021				
<b>5 a 7 anos de estudo</b>	4,136 0,011	2,397 0,017	1,332 0,029			
<b>8 anos de estudo</b>	5,560 0,009	3,222 0,014	1,791 0,025	1,344 0,035		
<b>9 a 11 anos de estudo</b>	10,709 0,004	6,205 0,007	3,449 0,012	2,589 0,017	1,926 0,026	
<b>12 ou + anos de estudo</b>	21,811 0,003	12,639 0,006	7,025 0,010	5,273 0,014	3,923 0,019	2,037 0,037

Fonte: PNAD/96

Nota: em cada célula, na primeira linha a razão das razões de chances e na segunda linha os erros padrões.

Para os pais (tabela 5, na página seguinte), as razões entre os diversos níveis educacionais são praticamente iguais às encontradas na tabela 4 referentes às mães, seguindo direções similares.

Fizemos essas tabelas de mobilidade considerando o grau de instrução dos avós das crianças, tanto do lado materno quanto do paterno (tabelas 24 a 27 do Apêndice 7.1.1, págs. 125 e 126). As razões sempre são superiores a um, com exceção das células onde comparamos o avô e a avó de lado paterno de 4 anos de escolaridade com os que tinham entre 5 e 7 anos de escolaridade (coeficientes próximos de um). O resultado “crianças de familiares mais educados têm mais

chances de não estarem atrasadas do que crianças de familiares menos instruídos” permanece.

Tabela 5: Razão das razões de chances de ter ou não defasagem idade-série segundo o nível educacional do pai

	sem estudo	1 a 3 anos de estudo	4 anos de estudo	5 a 7 anos de estudo	8 anos de estudo	9 a 11 anos de estudo
<b>1 a 3 anos de estudo</b>	1,747 0,023					
<b>4 anos de estudo</b>	3,078 0,013	1,762 0,021				
<b>5 a 7 anos de estudo</b>	3,852 0,012	2,205 0,019	1,252 0,032			
<b>8 anos de estudo</b>	5,012 0,009	2,869 0,016	1,628 0,027	1,301 0,037		
<b>9 a 11 anos de estudo</b>	8,146 0,006	4,663 0,009	2,647 0,016	2,115 0,022	1,625 0,030	
<b>12 ou + anos de estudo</b>	24,293 0,003	13,906 0,005	7,893 0,009	6,306 0,012	4,847 0,016	2,982 0,025

Fonte: PNAD/96

Nota: em cada célula, na primeira linha a razão das razões de chances e na segunda linha os erros padrões.

Os gráficos 13 e 14, no Apêndice 7.1.1, pág. 127, mostram a razão das razões de chances quando confrontamos o grupo familiar sem instrução com outros grupos de escolaridade. Os valores são praticamente similares nos dois primeiros grupos educacionais tanto para o lado paterno quanto materno (sem escolaridade versus 1 a 3 anos de estudo, sem escolaridade e 4 anos de estudo).

A situação em que uma criança cujo pai/mãe não têm instrução estar atrasada e outra criança cujo pai/mãe têm de 1 a 3 ou 4 anos de estudos não estar atrasada é mais provável do que a inversa. Este fato também ocorre na mesma proporção para crianças e avós/avôs.

A diferença entre as gerações somente é revelada quando focamos os grupos educacionais mais elevados. Enquanto para os pais e para as mães as razões sobem de 5 a 21, para os avós e avôs esses valores giram em torno de 3 a 10 quando comparamos os sem escolaridade com os grupos educacionais de mais de 8 anos de estudos.<sup>31</sup> As chances de ocorrer a seguinte situação: uma criança com avós de baixo nível de instrução estar atrasada e outra com avós de mais de 8 anos de instrução não estar atrasada são maiores do que o inverso. Essas chances são

<sup>31</sup> Fizemos gráficos similares para todos os grupos educacionais e encontramos resultados similares.

bem mais elevadas se considerássemos não o nível educacional dos avós, mas a origem familiar da segunda geração – de pais e mães.

Como a geração de avós como um todo possuía um baixo nível educacional, devido principalmente à escassez de oferta escolar e às preferências pela inserção no mercado de trabalho, sua influência sobre o atraso educacional das crianças parece ser inferior ao da geração de pais.

Por sua vez, pais com nível de escolaridade mais baixo possuem maiores dificuldades na aquisição de rendimentos que avós e avôs, tendo em vista as mudanças estruturais ocorridas no mercado de trabalho. Crianças cujos pais têm pouca escolaridade, conseqüentemente com menores rendimentos familiares, estão mais vulneráveis ao atraso educacional. Resta saber se ao compararmos crianças semelhantes em diversos aspectos e que se diferenciam somente conforme a renda familiar ou ao nível educacional dos pais esse efeito permanece.

## 2.5

### Estratégia econométrica: 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) abaixo, considerando as possíveis correlações existentes entre os termos de erros das eqs. (3), (4) e (8), transcritas novamente abaixo:

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

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

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

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

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

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



Ou seja, o termo de erro é escrito com uma função linear de fatores familiares fixos no tempo ( $u_f$ ) e outros aspectos individuais ( $u_{t+1}^i$ ). De acordo com 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 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. Média do total de escolas de cada estado quando os pais estavam em idade escolar (7 a 14 anos de idade).<sup>32</sup>
3. A diferença entre a escolaridade dos pais e dos avós.<sup>33</sup>

Explicamos essas escolhas detalhadamente nas sub-seções seguintes.

### 2.5.1 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.

<sup>32</sup> Construimos também a razão entre o total de professores pelo total de escolas, os resultados foram similares, logo optamos por não apresentá-los. Usamos as séries históricas do livro *Estatísticas do Século XX, FIBGE (2003)*. 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>33</sup> Como não foi possível fazer o cálculo dos anos de estudos completos dos avós, construimos a tabela 29, que compara a classificação do grau de instrução dos avós com os anos de estudos completos dos pais, conforme pode ser visto no Apêndice 7.1.3, pág. 131.

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 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. 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>34</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, 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 pais e mães nascidos a partir de 1957. Quando crianças, por lei, deveriam permanecer na escola até completar os 14 anos de idade.

---

<sup>34</sup> É importante destacar que a medida de escolaridade dos pais usada ao longo deste capítulo 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)*.

## 2.5.2 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).<sup>35</sup> Ou seja, para cada um, 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 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 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,

---

<sup>35</sup> No gráfico 15 do Apêndice 7.1.1, pág. 128, podemos visualizar a tendência de crescimento do total de escolas e professores do ensino de 1ª a 8ª. série de 1932 até o ano de 1975. Grande parte dessa expansão ocorreu ao longo das décadas de 60 e início de 70, beneficiando os indivíduos nascidos no final dos anos 50 e início dos anos 60. O gráfico 16, por sua vez, mostra que os estados que tinham um baixo número de escolas possuíam uma baixa média de escolaridade, reforçando o uso da variável “total de escolas” para explicar variações da escolaridade dos pais.

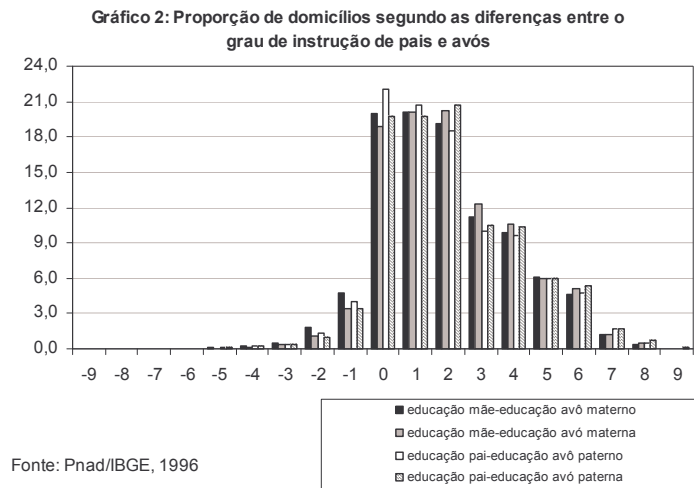
também é muito diferente do regime escolar freqüentado por seus filhos. Como houve enorme transformação do sistema educacional a partir dos anos 80, tanto em termos quantitativos quanto qualitativos, não acreditamos que os vínculos existentes entre as instituições escolares passadas e as freqüentadas 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.

### **2.5.3 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>36</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.

---

<sup>36</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 à geração de crianças e jovens.

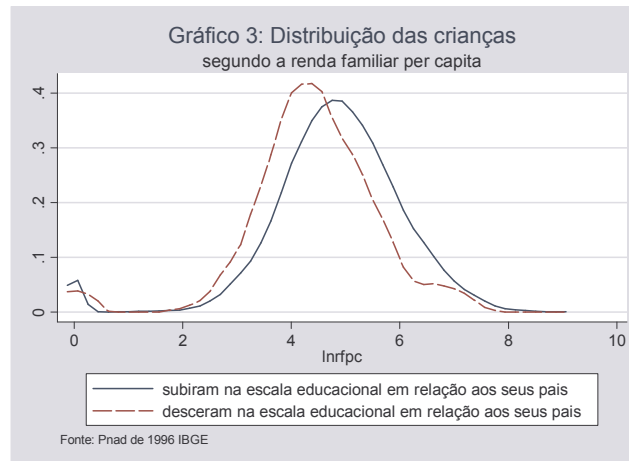


No gráfico 2 acima podemos visualizar a proporção de domicílios pela nossa escala de pontos, segundo os quatros indicadores de diferenças educacionais (“*educação mãe-educação avô materno*”, “*educação mãe-educação avô materna*”, “*educação pai-educação avô paterno*”, “*educação pai-educação avô paterna*”). 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>37</sup> Pais que subiram na escala educacional em relação a seus pais também são os que possuem renda familiar mais alta, como pode ser visto no gráfico 3 da distribuição das crianças por renda familiar per capita.

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) apresentada na sub-seção 2.3.2).

<sup>37</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%.



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 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.

## 2.6 Resultados

Nesta seção, a eq. (8), transcrita abaixo, é estimada através de um modelo de probabilidade linear em dois estágios usando nosso conjunto de instrumentos, conforme descrito na seção anterior 2.5.

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

A critério de comparação, estimamos um modelo de probabilidade linear usando mínimos quadrados ordinários e um modelo *probit* por máxima verossimilhança (estimação padrão). Nesses três 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.

### 2.6.1

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

Da coluna (1) a (2) da tabela 6, 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 usando mínimos quadrados ordinários e da (5) a (8), a estimação *probit* sem instrumentos.

Nas três primeiras linhas de todas as colunas, 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. Os sinais dos parâmetros estimados  $\alpha$  e  $\beta$  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>38</sup> do nível educacional da mãe é mais intenso que o do pai, sendo sempre significativo. 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 6). 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 6, linha dois).

---

<sup>38</sup> Chamamos a atenção que os efeitos marginais foram calculados para uma população de referência. Veja nota ao final da tabela 6.



Em todas as especificações, o impacto da renda familiar<sup>39</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 coluna 1 da tabela 6), 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). Fenômeno similar ocorre para as mães (de 0,02 para 0,008).

Tabela 6: Resultados das estimações da probabilidade da criança ter defasagem idade-série

Variável dependente:	IV - probabilidade linear		OLS - probabilidade linear		Probit			
	Coefficiente (1)	dp (2)	Coefficiente (3)	dp (4)	Coefficiente (5)	dp (6)	Efeito marginal (7)	dp (8)
<b>Ter defasagem idade-série</b>								
anos de estudos completos da mãe	-0,008 **	0,004	-0,018 *	0,001	-0,070 *	0,004	-0,019 *	0,001
anos de estudos completos do pai	-0,001	0,004	-0,012 *	0,001	-0,049 *	0,004	-0,013 *	0,001
ln da renda familiar per capita	-0,203 *	0,056	-0,033 *	0,004	-0,134 *	0,013	-0,036 *	0,004
idade em 1o. de março								
parte inteira da idade	0,076 *	0,003	0,067 *	0,001	0,257 *	0,006	0,069 *	0,002
maturidade	-0,125 *	0,011	-0,126 *	0,010	-0,505 *	0,041	-0,135 *	0,012
ser menina	-0,088 *	0,007	-0,084 *	0,006	-0,330 *	0,024	-0,101 *	0,008
ser de cor branca ou amarela	-0,021	0,016	-0,063 *	0,008	-0,207 *	0,028	-0,060 *	0,008
domicílio localizado na:								
área metropolitana	0,038 *	0,012	0,011	0,007	0,047 ***	0,028	0,012 ***	0,007
região nordeste	0,000	0,020	0,050 *	0,010	0,161 *	0,036	0,046 *	0,011
região sudeste	-0,073 *	0,012	-0,089 *	0,010	-0,315 *	0,036	-0,096 *	0,012
região sul	-0,082 *	0,013	-0,094 *	0,011	-0,341 *	0,043	-0,077 *	0,009
total de irmãs/irmãs mais novos	-0,005	0,015	0,039 *	0,003	0,130 *	0,012	0,035 *	0,004
total de irmãs/irmãs mais velhos	0,021 *	0,007	0,035 *	0,004	0,124 *	0,016	0,033 *	0,004
total de irmãs mais velhas	-0,015 **	0,007	-0,012 **	0,006	-0,039 ***	0,024	-0,010 ***	0,006
diferença entre a idade do pai e da mãe	0,002 ***	0,001	0,001 ***	0,001	0,006 **	0,003	0,002 **	0,001
constante	0,733 *	0,205	0,105 *	0,023	-1,425 *	0,088		
Obs:	20.332		20.332		Obs:		20.332	
F(15, 12831)	515,90		802,15		Wald chi2(15)		3.749,83	
Prob > F	0,000		0,000		Prob > chi2		0,000	
R <sup>2</sup>	0,222		0,334		Pseudo R2		0,308	

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%.

Encontramos que a renda familiar é um fator muito importante para explicar aspectos referentes à escolaridade das crianças. De acordo com a coluna (1) da tabela 6, quando dobramos a renda familiar per capita, a probabilidade da criança

<sup>39</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.

ter defasagem idade-série é reduzida em 0,20 pontos. Esse impacto é significativamente maior que o efeito estimado por mínimos quadrados ordinários ou por máxima verossimilhança (*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, independente do número de anos de estudo que adquiriram.

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 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 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 acumularem 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 6, 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.

### 2.6.2

#### Algumas observações sobre o primeiro estágio

Como pode ser visto no final da tabela 7 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. Quando testamos separadamente cada grupo de instrumentos, encontramos resultado similar: têm poder explicativo sobre as variáveis de interesse.

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 avôs sobre a escolaridade das crianças. Como pode ser visto na tabela 8 na página 60, o fato de avós e avôs terem baixa escolaridade não influencia de forma significativa a probabilidade da criança ter defasagem idade-série. Todos coeficientes dos avôs e avós foram estimados de forma imprecisa.

Tabela 7: Estimções do primeiro estágio do modelo de probabilidade linear

	Anos de estudo da mãe		Anos de estudo do pai		Log da renda familiar per capita	
	Coefficiente (1)	dp (2)	Coefficiente (3)	dp (4)	Coefficiente (5)	dp (6)
idade em 1o. de março						
parte inteira da idade	-0,030 *	0,008	-0,017 **	0,009	0,038 *	0,003
maturidade	-0,030	0,060	0,053	0,062	-0,002	0,025
ser menina	-0,022	0,035	-0,044	0,036	-0,025 ***	0,014
ser de cor branca ou amarela	1,036 *	0,040	1,119 *	0,042	0,365 *	0,017
domicílio localizado na:						
área metropolitana	0,535 *	0,038	0,664 *	0,040	0,228 *	0,016
região nordeste	-0,372 *	0,062	-0,408 *	0,064	-0,360 *	0,026
região sudeste	-0,510 *	0,066	-0,298 *	0,068	-0,011	0,027
região sul	-0,504 *	0,070	-0,368 *	0,073	-0,015	0,029
total de irmãos/irmãs mais novos	-0,250 *	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 *	0,010
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	-0,040 *	0,004	-0,035 *	0,005	-0,007 *	0,002
educação mãe-educação avô materno	0,632 *	0,014	-0,002	0,015	0,031 *	0,006
educação mãe-educação avó materna	0,923 *	0,015	0,262 *	0,015	0,085 *	0,006
educação pai-educação avô paterno	0,054 *	0,015	0,575 *	0,016	0,043 *	0,006
educação pai-educação avó paterna	0,254 *	0,015	1,036 *	0,016	0,089 *	0,006
média de escolas mãe	0,008 ***	0,005	-0,008 ***	0,005	0,008 *	0,002
média escolas pai	-0,001	0,005	0,011 **	0,005	-0,001	0,002
pai nasceu depois de 1957	-0,079	0,048	-0,026	0,049	-0,102 *	0,020
mãe nasceu depois de 1957	0,133 *	0,046	0,269 *	0,047	-0,057 *	0,019
constante	2,990 *	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		401,620	

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%.

No entanto, o coeficiente estimado para a escolaridade da mãe aumentou enquanto da renda familiar per capita diminuiu 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 6 e 8). 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 estimções padrões permanece: a renda familiar nas estimções padrões é viesada para baixo e o inverso ocorre para a escolaridade dos pais.

Tabela 8: Impacto da escolaridade dos avôs e avós na probabilidade das crianças terem defasagem idade-série - probabilidade linear com instrumentos

<i>Variável dependente:</i>	<b>Coeficiente</b>	<b>dp</b>
<b>Ter defasagem idade-série</b>	<b>(1)</b>	<b>(2)</b>
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	-0,030 **	0,013
domicílio localizado na:		
área metropolitana	0,032 *	0,011
região nordeste	0,010	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	0,010
avó do lado materno sem escolaridade	-0,002	0,011
avô do lado paterno sem escolaridade	-0,014	0,010
avó do lado paterno sem escolaridade	-0,006	0,010
constante	0,610 *	0,184
<hr/>		
Obs:		20.332
F(19, 12831)		483,53
Prob > F		0,000
R <sup>2</sup>		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 <= 1%, \*\* p-valor entre 1% e 5%, \*\*\* p-valor > 5% e <=10%.

## 2.7

### Considerações finais

No tocante à estimação das variáveis potencialmente endógenas, encontramos que todas 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 estimções 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 capítulo focamos apenas em um lado do processo de acúmulo de capital humano (a demanda educacional por parte das famílias), não abordamos 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 filhos/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.