

A RELAÇÃO ENTRE FREQUÊNCIA ESCOLAR E RENDA FAMILIAR NO BRASIL — 1981-1999*

Lígia Vasconcellos
Do Banco Itaú

Este artigo explora as variações na distribuição de renda no período de 1981 a 1999 para determinar, por meio do método de variáveis instrumentais, se há relação causal entre a renda familiar e a frequência escolar de crianças em idade escolar compulsória no Brasil. Considerando-se um efeito-renda homogêneo, não se acha uma relação causal significativa. Quando, porém, se permite que o efeito varie entre grupos de renda, definidos por quartis de renda, encontra-se um efeito-renda significativo da renda familiar sobre a frequência escolar do grupo de renda mais pobre.

1 INTRODUÇÃO

A educação é, hoje, uma das maiores preocupações em termos de política pública no Brasil. Embora o nível de gastos com educação, em relação ao PIB, possa ser comparado com o dos países desenvolvidos, e o aumento das taxas de frequência escolar tenha acelerado na última década, os indicadores educacionais estão bem distantes do desejável.

A Tabela 1 compara alguns indicadores educacionais entre países latino-americanos. O Brasil apresenta o pior resultado em relação a taxa de alfabetização e anos completos de escola,¹ ambos indicadores relativos à população adulta. A taxa média bruta de matrícula,² que pondera as taxas de matrícula dos três níveis de ensino — fundamental, médio e superior —, mostra um quadro favorável, sugerindo que a próxima geração terá um nível educacional melhor. Essa medida bruta, porém, esconde um forte desequilíbrio entre a idade da população frequentando cada nível de ensino e a idade esperada para cada nível. Esse problema está relacionado às altas taxas de repetência — por exemplo, a taxa bruta de matrícula do ensino fundamental era de 130,5% em 1999.

Foi implementada uma série de políticas voltadas para a educação escolar na década de 1990, representando uma mudança de foco das políticas públicas do ensino superior para os níveis escolar e pré-escolar. Entre as novas políticas³

* Pesquisa realizada para Tese de Doutorado na Faculdade de Economia, Administração e Contabilidade (FEA) da USP. A autora agradece à Capes pelo apoio financeiro. Agradece, também, a Jorge Arbache, Richard Dickens, Reynaldo Fernandes, Ana Lúcia Kassouf, Marco Manacorda, Naércio Menezes-Filho, Paulo Picchetti e Adriana Schor por comentários e sugestões sobre versões anteriores desta pesquisa.

1. Esses são os indicadores educacionais usados para medir o Índice de Desenvolvimento Humano (IDH), publicado pelas Nações Unidas.
2. Razão entre a população estudantil em determinado nível de ensino, independentemente da idade, e o total da população com idade esperada para o mesmo nível.
3. Fundo de Manutenção e Desenvolvimento do Ensino Fundamental e de Valorização do Magistério (Fundef) e Lei de Diretrizes e Bases da Educação Nacional (LDB), ambos de 1996.

TABELA 1
AMÉRICA LATINA: INDICADORES EDUCACIONAIS

País	Taxa de alfabetização ^a	Taxa bruta de matrícula ^b	Anos de escolaridade ^c
	2000	1999	2000
Argentina	96,8	83	8,49
Chile	95,8	78	7,89
Costa Rica	95,6	67	6,00
México	91,4	71	6,73
Cuba	96,7	76	7,78
Colômbia	91,7	73	5,01
Venezuela	92,6	65	5,60
Brasil	85,2	80	4,56

Fontes: PNUD (2000) e Banco Mundial (2002).

^a Taxa de alfabetização (15 anos de idade ou mais).

^b Taxa combinada, incluindo ensinos fundamental, médio e superior.

^c População adulta.

estão a vinculação de parte das receitas públicas de todas as esferas de governo para gastos educacionais em geral e especificamente para o ensino fundamental e um nível mínimo de gasto por aluno, garantido por transferência federal caso o governo local não tenha recursos suficientes. Além disso, novas regras foram implementadas, introduzindo, por exemplo, um ano letivo mais longo e a definição de uma qualificação mínima para os professores. Essas políticas estão certamente relacionadas com a melhora observada na frequência escolar na década de 1990. Mais de 95% das crianças entre 9 e 12 anos estavam na escola em 1999. Além disso, a média de escolaridade dos jovens de 17 anos em 1999 era de 6,9 anos, resultado bem melhor que a média da população adulta vista na Tabela 1, mas ainda assim menor que o esperado para esta idade — 10 anos de estudo.

Os resultados educacionais, no entanto, não são explicados apenas por políticas públicas, mas também por características individuais e familiares, como o *background* familiar e as habilidades inatas. Este artigo foca sua atenção na renda familiar. Embora haja muitos estudos investigando o impacto da renda familiar sobre a educação, não é claro se essa relação é de causalidade, pois a renda pode estar captando o efeito de outras características familiares.

Haveman e Wolfe (1995) resumem uma série de artigos analisando várias medidas de resultado escolar para os Estados Unidos; a renda familiar é quase sempre considerada como variável explicativa e, na maioria dos estudos, tem efeito positivo e significativo. Os autores mostram que estudos incluindo maior

número de características familiares como controles apresentam efeito-renda menor, o que confirma a hipótese de que a renda se correlaciona com características familiares. Mesmo controlando-se por essas características, no entanto, características não-observadas ou erros de medida na renda podem ainda afetar o efeito estimado da renda, como será discutido.

A relação potencial entre renda familiar e resultados educacionais é um bom indicador de oportunidades iguais e mobilidade social entre gerações. Barros (1997) resume vários artigos para o Brasil e mostra que, depois de se controlar por educação, a diferença da renda reduz-se em 35% a 50% — um poder explicativo alto e que reforça a importância da educação para a distribuição de renda e mobilidade intergeracional. Além disso, o retorno da educação no Brasil é alto, especialmente no nível superior. Relatório do Banco Mundial de 2002 estima o retorno dos níveis médio e superior no Brasil em 19%.

A teoria do capital humano postula que o investimento em educação é baseado em seu retorno em termos de salário futuro, comparado ao custo da educação. Com mercados perfeitos, o problema de maximização individual gera o nível privado ótimo de educação. Custos e retornos heterogêneos, em função de diferentes habilidades, preferências familiares ou ainda diferenças nos mercados locais podem gerar diferentes decisões de investimento entre famílias. A decisão educacional pode também ser afetada por restrições de crédito, caso em que o nível de investimento possível pode ser menor que a escolha ótima sem restrição. Na presença de restrição de crédito, o investimento em educação depende, basicamente, da renda familiar e do subsídio do governo.

Mesmo se a escola for pública, caso da maioria no Brasil — 91% dos alunos do ensino fundamental estavam em escolas públicas em 2001 —, há custos indiretos, como transporte e de oportunidade, que podem explicar a importância da renda familiar. Muitas crianças trabalham para complementar a renda dos pais, apesar de haver proibição legal do trabalho infantil até os 14 anos de idade,⁴ a mesma idade considerada compulsória para frequentar a escola. Mais de 10% das crianças entre 10 e 13 anos de idade estavam trabalhando em 1999; esse percentual aumenta para 30% no caso de jovens entre 14 e 17 anos no mesmo ano.

Se as famílias têm diferentes níveis de restrição de crédito, o impacto da renda sobre a educação também pode diferir entre elas. É, portanto, possível que a restrição de crédito importe apenas a partir de determinado nível de renda. Essa possibilidade de efeito heterogêneo da renda entre famílias é investigada. Não se vai distinguir entre variações de renda transitória e permanente. A teoria do capital humano, em geral, considera a renda permanente. A restrição orçamentária familiar, no entanto, também é afetada por choques transitórios

4. Desde 1998, a legislação estabelece um mínimo de 16 anos como regra geral e outra de 14 anos para a situação de aprendiz.

na renda, e a decisão de frequência escolar pode ser alterada a qualquer momento, ou pelo menos a cada novo início de ano letivo. Choques transitórios, portanto, também impactam na frequência escolar.

Kassouf (2000), Barros *et alii* (2001) e Carvalho (2001) analisam os determinantes da frequência escolar no Brasil usando dados da Pesquisa Nacional por Amostra de Domicílios (PNAD) do Instituto Brasileiro de Geografia e Estatística (IBGE), a mesma que será aqui utilizada. Kassouf analisa a frequência escolar para uma amostra do ano de 1995 de crianças de 7 a 14 anos. Controlando por características familiares, a autora encontra efeito sempre positivo e significativo para a renda salarial do pai; a significância da renda da mãe depende da inclusão ou não da escolaridade dos pais; e a renda não-salarial nunca é significativa. Barros *et alii* consideram anos completos de escola para jovens de 11 a 25 anos no ano de 1996. Os autores incluem como variáveis explicativas adicionais às características familiares variáveis de comunidade e da escola; também estimam que a renda familiar é significativa. Ambos os artigos estimam que a escolaridade dos pais é a variável mais importante afetando a educação dos filhos.

A existência de correlação entre renda familiar e frequência escolar não necessariamente implica causalidade. A renda familiar pode estar captando o efeito de outras variáveis que também afetam a educação, como *background* familiar. Carvalho (2001) tenta isolar o efeito-renda de não-observáveis. Sua amostra inclui jovens na faixa de 10 a 14 anos no período 1989-1993. O autor, para isolar o efeito da renda, usa a reforma da previdência implementada no início de 1990, que afetou fortemente e de forma exógena a renda de trabalhadores rurais. Ele não acha efeito causal para a maioria das subamostras que estuda: todas as amostras de meninos, meninas entre 10 e 11 anos de idade e meninas advindas de famílias em que o benefício é recebido por um homem. Na presença de efeitos heterogêneos, sua estimativa mede um efeito médio local — Local Average Treatment Effect (LATE) —,⁵ isto é, as famílias rurais compostas por pessoa com direito ao benefício previdenciário e criança têm maior peso na ponderação do efeito-renda, pois são as famílias afetadas pela reforma. Essa parte da população é mais pobre que a média, portanto, a não-significância da renda em suas estimativas é uma surpresa. Uma possível explicação está na fonte de variação da renda utilizada, apenas a advinda de um programa social e não do salário. Haveman e Wolfe (1995) comentam que a fonte de renda parece importante para determinar o seu efeito e Kassouf (2000) não encontrou nenhum efeito para a renda não-salarial.

5. Angrist e Imbens (1995) demonstram esse resultado.

Este artigo também pretende isolar o efeito da renda familiar sobre a frequência escolar. A estratégia para isolar esse efeito está em explorar as variações na renda familiar causadas por variações na distribuição de renda, que provavelmente não estão correlacionadas com as variáveis não-observadas, como preferências e habilidades familiares. Acemoglu e Pischke (2001) usam a mesma estratégia de identificação para estimar o efeito da renda dos pais sobre a matrícula no ensino superior nos Estados Unidos. Esse instrumento deve prover uma estimativa mais geral do efeito-renda do que a estimativa de Carvalho (2001), pois a variação da renda usada aqui não está limitada a um grupo específico de famílias; além disso, considera-se um período de tempo mais longo.

A frequência escolar é o indicador menos afetado pela dinâmica dentro da escola, comparada, por exemplo, com anos completos de escolaridade, que é afetado pela política de reprovação. Dessa forma, espera-se que o fato de aspectos relacionados diretamente à qualidade do ensino não serem considerados não afete a análise aqui apresentada.

A existência ou não de causalidade entre renda e frequência escolar é informação relevante para decisões de política pública. Desde 1995 há experimentos de transferência de renda com condicionalidades para famílias pobres com crianças em idade escolar, e desde 2001 programas federais⁶ incorporaram esses programas; há também programas similares no sentido de evitar o trabalho infantil. Esses programas criam uma relação de dependência entre frequência escolar e renda, pois a frequência é compulsória para o recebimento do benefício. Se existisse uma relação causal independentemente do benefício, não haveria necessidade do custo relacionado à fiscalização da frequência.

Este artigo estima o efeito da renda familiar sobre a frequência escolar para crianças em idade escolar compulsória no período 1981-1999 no Brasil. Investiga-se a possibilidade de efeitos-renda heterogêneos no tempo e entre grupos de renda, determinados entre quartis de renda.

2 DESCRIÇÃO DOS DADOS

Os dados são das PNADs do IBGE. O período de análise vai de 1981 a 1999. Não há pesquisa em ano de Censo Demográfico — 1991 no período considerado — e não houve pesquisa em 1994. As PNADs incluem todo o país, com exceção da área rural da região Norte, que representa 2,3%⁷ da população.

6. Programa Nacional de Renda Mínima vinculada à educação — “bolsa-escola” — iniciado em 2001 e incorporado ao programa “bolsa-família” em 2003.

7. Censo Demográfico de 2000, do IBGE.

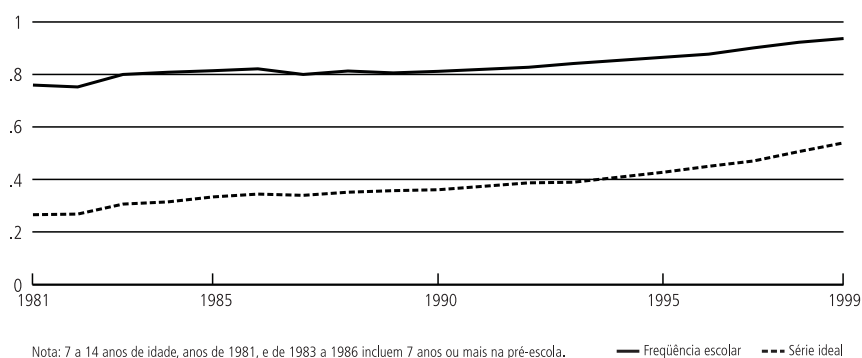
A faixa etária está restrita à idade de escola compulsória — 7 a 14 anos. Seria necessário um modelo mais completo para incluir a dinâmica relacionada à decisão de oferta de trabalho dentro da família e sua relação com a frequência escolar e a renda da família, mas a escolha dessa faixa etária reduz o problema de não se levar em conta essa dinâmica, pois o custo de oportunidade de não se trabalhar é menor.

A frequência escolar refere-se à semana de referência da entrevista, em setembro. As crianças são consideradas na escola se respondem estar freqüentando o ensino regular no período da entrevista — não é considerada a frequência à pré-escola, mas o ensino de jovens e adultos (antigo supletivo) sim.

O Gráfico 1 mostra a evolução da frequência escolar no tempo. Há um aumento constante no tempo, acelerado nos últimos anos do período. Nas pesquisas de 1981 e de 1983 a 1986, as PNADs realocaram para a 1ª série do ensino fundamental as crianças de 7 anos de idade ou mais que informaram estar na pré-escola, e não é possível distingui-las das crianças efetivamente freqüentando a 1ª série. O aumento de frequência observado nesses anos é provavelmente apenas relacionado a esse erro de medida. Como essa era uma regra geral,⁸ para toda a amostra no período, espera-se que seu efeito seja capturado pelas *dummies* de ano e, portanto, não afete as estimativas do efeito da renda.

Uma medida de atraso escolar que também é apresentada no gráfico é a proporção de crianças na série esperada (ideal) de acordo com a idade. Essa

GRÁFICO 1
PROPORÇÃO DE FREQUÊNCIA ESCOLAR E DE SÉRIE IDEAL NO TEMPO



8. A regra da PNAD entre 1987 e 1989 ainda aloca crianças de 10 ou mais anos de idade à 1ª série caso respondam estar na pré-escola; após 1989, crianças de 12 anos ou mais são realocadas. Como o percentual de crianças na pré-escola nessas idades é baixo, a realocação nesses períodos não é tão importante.

medida reflete diferentes aspectos da educação: idade de entrada, reprovação e evasão, e apesar de o indicador melhorar no tempo, o atraso ainda é muito alto.

A PNAD pesquisa todos os tipos de renda dos membros do domicílio: renda de todos os trabalhos, pensões, ganhos de juros, entre outros. A renda mensal informada se refere à semana de referência da pesquisa em setembro. A informação da renda diz respeito aos indivíduos de 10 ou mais anos de idade.

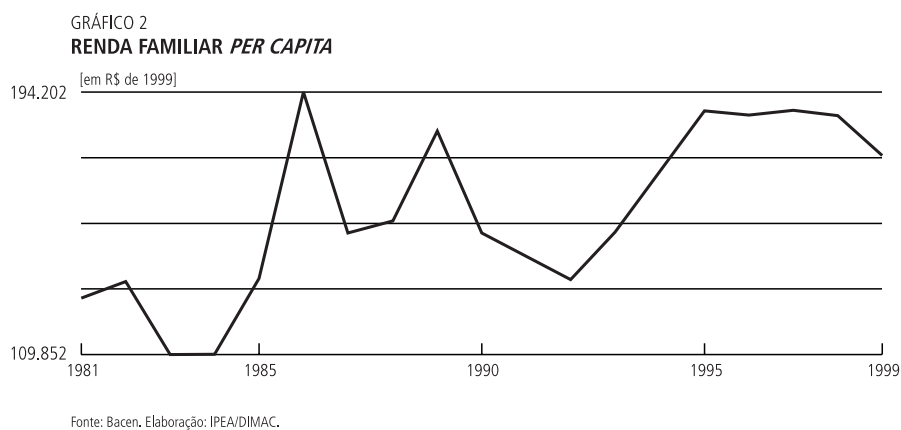
A unidade da pesquisa é o domicílio, mas a análise será feita com a unidade familiar. Consideram-se como parte da família o chefe, seu cônjuge, filhos, outros parentes e agregados (indivíduos que não sejam parentes nem empregados domésticos e que não paguem aluguel) morando no domicílio. Utiliza-se a renda *per capita*: renda familiar, incluindo todas as fontes de renda, dividida pelo número de pessoas na família. A renda da criança não é considerada, em função de a decisão de frequência estar provavelmente vinculada à decisão de trabalho. A renda de irmãos é considerada.

Há um *trade-off* na escolha da definição de renda. Com a inclusão de mais rendas, além da renda dos pais, enfatiza-se o aspecto de riqueza e esta é uma medida melhor para ser relacionada às decisões de educação. Por outro lado, por exemplo, a renda de irmãos pode ser uma fonte de endogeneidade, pois sua decisão de trabalho e educação pode estar relacionada à possibilidade de o outro irmão ir à escola.

A renda nominal de cada ano foi atualizada a valores de setembro de 1999 de acordo com o Índice Nacional de Preços ao Consumidor (INPC) do IBGE. Como a análise inclui um período de quase hiperinflação e cinco mudanças de moeda, mesmo a correção mais acurada dos valores nominais deve ser vista com cautela. Mesmo dentro do mesmo ano e mês, pode haver ruído na comparação de valores; por exemplo, as negociações de salário por classes sindicais ocorrem uma vez por ano e em diferentes meses, dependendo da classe e da localidade. O erro de medida na informação de renda pode, portanto, ser importante.

O Gráfico 2 mostra a trajetória da renda familiar *per capita* no tempo. Essa trajetória envolve as mudanças na oferta de trabalho dos membros da família e a mudança na composição da família, que inclui um número decrescente de filhos e crescente de famílias com um só pai presente. Não houve aumento importante na renda total familiar comparando-se os anos de 1981 e 1999, mas o gráfico mostra que aumentou a renda familiar *per capita* no tempo.

Algumas características familiares e demográficas são usadas como variáveis de controle na análise. As características individuais são gênero e raça. Não há informação de raça para os anos de 1981 e 1983, e para os anos de 1984 e 1985 a informação está disponível apenas para parte da amostra. Como raça é



uma variável explicativa importante, foi criada a categoria “sem informação” para essas observações, de forma a permitir sua inclusão na análise econométrica. As categorias para raça são branco, pardo, negro e oriental. As características familiares são educação dos pais, número de irmãos vivendo no domicílio, e se é uma família com um só pai vivendo no domicílio. A escolaridade dos pais é definida como o número de anos completos de educação de cada um: de 0 a 16 anos. Para essa última variável também foi criada a categoria “sem informação” — a falta de informação deve-se ao fato de na maioria dos casos um dos pais não morar no domicílio. Variáveis adicionais de controle são: *dummies* regionais, área urbana ou rural, *dummies* de ano, e a posição na distribuição de renda, medida por grupos definidos pelos quartis de renda, criados dentro dos anos e das regiões.

Como o interesse do artigo está centrado na relação da renda familiar com a frequência escolar dos filhos, indivíduos na faixa etária definida mas que tinham outra condição na família foram desconsiderados. Famílias sem informação de renda (ou com renda 0) também foram excluídas. A Tabela 2 resume a seleção da amostra.

A Tabela 3 mostra a renda para alguns percentis depois de feita a seleção da amostra descrita. Após a primeira seleção, uma segunda seleção foi feita, desconsiderando-se as observações com renda familiar nos extremos da distribuição de renda, 5% de cada ponta, para diminuir problemas de erro de medida existentes especialmente nos extremos da distribuição. A amostra final contém 959.264 crianças de 7 a 14 anos.⁹ O menor grupo dentro de ano,

9. Ver Apêndice para distribuição da amostra por anos e regiões.

TABELA 2
CARACTERÍSTICAS DAS OBSERVAÇÕES
 [em %]

Condição na família	Renda		
Filho	93,45	Renda > 0	96,83
Outro parente	5,68	= 0 ou sem informação	3,17
Chefe/cônjuge	0,19		
Outros	0,68		

Nota: 7 a 14 anos de idade.

TABELA 3
RENDA PER CAPITA POR PERCENTIS
 [em R\$ de 1999]

Ano	p1	p5	Mediana	p95	p99
1981	8,45	16,19	72,41	451,84	1.101,95
1985	7,38	15,51	69,49	490,83	1.194,51
1990	5,83	13,34	71,48	556,03	1.334,47
1995	8,59	18,91	96,70	706,34	1.611,76
1999	8,33	18,57	91,66	658,33	1.500,00

região e grupo de renda tem 394 observações (região Norte, 1981), portanto, não há problema de tamanho de amostra para o estimador de dados agrupados que será utilizado.

3 ESTRATÉGIA EMPÍRICA

Para determinar se há, de fato, uma relação causal entre renda familiar e frequência escolar será explorada a variação da distribuição de renda no tempo e entre regiões. A distribuição de renda é medida pela diferença das rendas médias dos grupos definidos pelos quartis de renda, criados dentro de cada região e ano.

São apresentadas várias especificações, sendo possível acompanhar a mudança da estimativa do efeito da renda devido à inclusão de variáveis de controle adicionais. Os termos de erros são corrigidos para possíveis efeitos de *cluster*, como sugerido por Moulton (1990), sendo os *clusters* definidos por grupos de ano-região-quartil de renda. O fato de haver irmãos entre as observações não está sendo controlado, eles são tratados como observações independentes.

São apresentadas estimativas por mínimos quadrados ordinários (MQO) e por *probit*.¹⁰ Essas estimativas têm provavelmente dois problemas importantes: correlação da renda com o termo de erro e erro de medida na variável renda familiar, os dois problemas levando a estimativas com viés.

O viés de variável omitida ocorre quando há variáveis não-controladas que estão correlacionadas tanto com a frequência escolar como com a renda familiar. Portanto, a inclusão de variáveis explicativas observadas é necessária, mas o problema não é resolvido se fatores como habilidades ou preferências familiares relacionadas à educação restam no termo de erro. Como a decisão sobre frequência para a faixa etária considerada é dos pais, importam suas habilidades e preferências e essas características são muito provavelmente correlacionadas com a renda familiar. Se essas são as variáveis omitidas, é esperado um viés para cima nas estimativas por MQO, pois em geral habilidades e preferências são ambas correlacionadas positivamente com renda e frequência. O problema de erro de medida leva a um viés de atenuação.

O uso do método de variáveis instrumentais deve resolver esses problemas de variável omitida (as hipóteses de identificação são explicitadas em seguida) e de erro de medida. É usada a estimação de mínimos quadrados em dois estágios (2SLS).¹¹ O primeiro estágio estima a renda a partir do instrumento. Variáveis de controle são incluídas para ajudar a isolar o efeito da renda, diminuindo o risco de correlação entre o instrumento e o termo de erro.

O instrumento é a variação da distribuição de renda no tempo e entre regiões. Como região de residência e ano podem ter impacto direto sobre a frequência escolar, pois capturam o efeito de variáveis agregadas, devem ser incluídos na regressão do segundo estágio. Da mesma forma como a posição na distribuição (grupos definidos pelos quartis de renda) pode estar correlacionada com habilidades (que incluem habilidades adquiridas da família), essa posição também deve entrar como controle. Portanto, as restrições de exclusão se referem somente à variação da distribuição de renda que ocorre de forma diferenciada entre regiões e tempo e, para que o instrumento seja válido, as variáveis não-observadas não podem variar de forma diferenciada no tempo entre os grupos de renda (definidos pelos quartis) e regiões.

A hipótese de identificação, portanto, é a de que a trajetória das taxas de frequência escolar entre crianças com pais de diferentes grupos de renda na mesma região, após controle por variáveis observáveis, deveria ser a mesma, não

10. Como o método de variáveis instrumentais usado posteriormente é linear, os resultados por MQO são analisados com maior atenção, mas as estimativas por *probit* apresentam resultados similares.

11. Como mostram Davidson e MacKinnon (1993), nem sempre é possível obter estimativas consistentes com funções não-lineares em métodos em dois estágios, daí a escolha de um modelo linear.

fosse a diferença na variação da renda de cada família. A identificação é possível em função da existência de variação na renda condicional em ano, região, posição na distribuição, e características familiares observáveis.

Entender a renda como resultado da valorização das características familiares no mercado pode esclarecer melhor a hipótese de identificação: essas características afetam tanto a renda como a decisão de educação, mas seu preço no mercado de trabalho está apenas correlacionado com a renda. Dessa forma, diferentes valorizações das mesmas características familiares no tempo e entre regiões geram a variação exógena necessária para o instrumento.

Em outros termos:

$$edu_{iqr} = a + b \ln w_{iqr} + \sum c_j \cdot F_{jiqr} + u_{iqr} \quad (1)$$

$$\ln w_{iqr} = D_{iqr} + D_q + D_r + D_t + \sum d_j \cdot F_{jiqr} + v_{iqr} \quad (2)$$

Onde:

edu = frequência escolar;

$\ln w$ = logaritmo da renda familiar;

F_j = característica familiar observada j ;

a, b, c_j, d_j = coeficientes;

u, v = termos de erro;

D = *dummies* ou interações de *dummies*;

i = indivíduo;

t = ano;

q = grupo de renda; e

r = região.

As hipóteses de identificação são:¹²

a) Condição de posto

$$\left[E(\ln w | q, r, t) - E(\ln w | q) - E(\ln w | r) - E(\ln w | t) \right]^2 \neq 0$$

12. Essa notação é encontrada em Blundell, Duncan e Meghir (1998).

A renda familiar varia de forma diferente entre grupos, pois, mesmo depois de controlar por quartil, região e tempo, ainda existe variação da renda. A variação exógena usada como instrumento é representada pelo conjunto de interações de quartil-região-ano, condicional no nível dessas variáveis e nas características familiares observadas.

b) Restrições de exclusão

$$E(u_{iqr} | \text{quartil, região, ano}) = D_q + D_r + D_t$$

Diferenças não esclarecidas na média da frequência escolar entre grupos são explicadas por efeitos fixos (região e quartil) e por um efeito-tempo aditivo. Portanto, as interações entre essas *dummies* não têm impacto sobre a frequência, a não ser através da renda, e portanto podem ser usadas como instrumento.

A estratégia de identificação, portanto, depende de as variáveis não-observáveis não variarem de forma diferenciada entre grupos. Como os “suspeitos” prováveis são habilidades e preferências, essa parece ser uma hipótese adequada. Poderíamos ser mais estritos em relação às restrições de exclusão e incluir também como controles as interações de segunda ordem de tempo, região e quartil de renda. Essa possibilidade foi testada, e não sobra variação na renda para a identificação, pois esses controles adicionais eliminam muito da variação dos dados. De qualquer forma, como veremos, sua inclusão não altera a estimativa do efeito-renda e, portanto, não parece haver problema de correlação entre essas interações e a renda.

Esse estimador de variáveis instrumentais equivale a um estimador de dados agrupados, em que temos 340 grupos (17 anos x 4 quartis de renda x 5 regiões). Se não houvesse controles adicionais além das *dummies* de quartil, tempo e região, os dois estimadores, de variáveis instrumentais e de mínimos quadrados generalizados sobre uma amostra de médias seriam idênticos, caso contrário, como é o caso aqui, eles serão diferentes em função da existência de variação das variáveis de controle dentro dos grupos. Será usada a estimação sobre os dados individuais.

Para verificar se vale a equação do primeiro estágio, isto é, se o instrumento excluído é correlacionado com a variável endógena, é calculado um teste-F para o instrumento e o R^2 parcial para a regressão de primeiro estágio, como sugerido em Bound, Jaeger e Baker (1995). O teste de superidentificação de Hansen, que é robusto para heteroscedasticidade, testa a exogeneidade do instrumento. Esse teste foi escolhido em função da suposição de que há algum grau de dependência entre as observações dentro de cada grupo.

Os resultados das especificações são apresentados para a amostra total e também para três subperíodos (1981-1986, 1987-1992 e 1993-1999) para investigar a possibilidade de uma variação no tempo do efeito-renda. Cada um dos subperíodos escolhidos inclui uma variação importante da distribuição de renda, o que garante as propriedades do instrumento. O primeiro período inclui o Plano Cruzado, primeiro plano de estabilização em 1985, que causou um período curto de recuperação econômica e alguma redistribuição de renda; o segundo é um período de quase hiperinflação, em geral relacionada à perda de renda real pelos mais pobres; e o terceiro período inclui o Plano Real, o último e bem-sucedido plano de estabilização, que também diminuiu as taxas de pobreza.

Vale notar que a interpretação das estimativas com variáveis instrumentais deve ser feita com cuidado na presença de efeitos-renda heterogêneos e de efeitos de grupo. Na presença de efeito heterogêneo, a estimativa com variáveis instrumentais equivale a uma média ponderada dos efeitos individuais (LATE). Como mostram Angrist e Imbens (1995), os pesos estão relacionados à variação relativa da renda familiar induzida pelo instrumento, isto é, indivíduos com suas rendas familiares mais afetadas pela variação na distribuição de renda terão maior peso. Como o instrumento inclui variações entre regiões e quartis ao longo de vários anos, a estimativa tende a ser representativa para o grupo inteiro. Se há um efeito da renda do grupo (aqui definido pelas variáveis ano, região e grupo de renda) adicional ao efeito da renda familiar, o tipo de estimador usado aqui, por equivaler ao estimador sobre dados agrupados, acaba somando ambos os efeitos-renda, individual e de grupo. Nesse caso, o efeito-renda estimado incluiria o do grupo.

4 ESTATÍSTICA DESCRITIVA

As próximas estatísticas¹³ estabelecem a relação entre frequência escolar e renda familiar. Além disso, existe evidência da variação diferenciada da renda entre quartis e regiões ao longo do tempo, hipótese de identificação do instrumento escolhido.

A Tabela 4 confirma o aumento de frequência escolar no período 1981-1999 já visto no Gráfico 1, aumento especialmente forte no final da década de 1990. Esse aumento deu-se de forma diferenciada entre grupos de renda e de regiões: ele é maior para os grupos de renda mais pobres; a região Nordeste quase alcançou os indicadores de frequência das outras regiões na década de 1990. O grupo mais pobre na região Nordeste passou de uma frequência de 63,8% no período 1981-1986 para 75,5% em 1993-1999. No mesmo período,

13. Todas as estatísticas descritivas foram calculadas utilizando-se os pesos amostrais das PNADs.

TABELA 4
PROBABILIDADE MÉDIA DE FREQUÊNCIA ESCOLAR POR REGIÃO E QUARTIL

Região/quartil	1º	2º	3º	4º	Total
1981-1986					
Sudeste	0,748	0,809	0,874	0,948	0,848
Norte	0,815	0,860	0,887	0,925	0,872
Nordeste	0,638	0,686	0,760	0,882	0,731
Sul	0,710	0,778	0,847	0,919	0,804
Centro-Oeste	0,689	0,770	0,847	0,931	0,805
Total	0,701	0,762	0,833	0,923	0,802
1987-1992					
Sudeste	0,769	0,838	0,904	0,954	0,871
Norte	0,799	0,850	0,894	0,943	0,870
Nordeste	0,624	0,690	0,772	0,885	0,733
Sul	0,751	0,831	0,894	0,956	0,849
Centro-Oeste	0,734	0,815	0,885	0,946	0,844
Total	0,715	0,788	0,861	0,934	0,823
1993-1999					
Sudeste	0,880	0,921	0,951	0,976	0,935
Norte	0,833	0,891	0,929	0,959	0,901
Nordeste	0,755	0,818	0,874	0,942	0,840
Sul	0,877	0,927	0,957	0,982	0,933
Centro-Oeste	0,851	0,898	0,937	0,969	0,913
Total	0,833	0,886	0,927	0,966	0,902

Nota: Crianças de 7 a 14 anos; dados de 1981 e de 1983 a 1986 incluem crianças na pré-escola.

o grupo de renda mais rico do Sudeste aumentou a frequência de 94,8% para 97,6%. É importante notar que a diminuição da frequência entre os dois primeiros períodos presente para alguns grupos é provavelmente efeito da regra da PNAD de incluir pré-escola no primeiro período e não representa uma piora de frequência.

A Tabela 5 resume a evolução da renda familiar por região e quartil. Há uma diminuição de renda para o grupo mais pobre entre o primeiro e o segundo

TABELA 5
RENDA MENSAL FAMILIAR PER CAPITA MÉDIA POR REGIÃO E QUARTIL
 [em R\$ de 1999]

Região/quartil	1º	2º	3º	4º	Total
1981-1986					
Sudeste	34,92	69,11	119,21	257,97	124,42
Norte	36,65	67,73	111,70	240,49	114,37
Nordeste	23,75	39,66	64,23	160,56	66,10
Sul	36,46	71,10	118,80	252,29	110,32
Centro-Oeste	32,77	60,67	101,59	239,57	105,18
Total	31,25	59,02	100,77	228,12	101,96
1987-1992					
Sudeste	33,83	71,68	128,38	293,83	138,54
Norte	32,18	64,13	112,59	267,51	116,94
Nordeste	20,69	36,56	63,10	180,76	69,65
Sul	35,94	76,18	131,64	298,24	125,94
Centro-Oeste	30,81	61,34	108,31	273,08	117,18
Total	29,44	59,85	107,03	260,63	112,71
1993-1999					
Sudeste	43,93	91,38	159,53	355,57	170,26
Norte	34,15	65,54	115,31	297,08	124,91
Nordeste	26,41	46,14	78,82	228,30	87,69
Sul	46,09	96,64	167,86	368,43	163,62
Centro-Oeste	37,95	75,04	131,26	326,18	140,17
Total	37,35	75,16	132,75	318,06	139,41

Nota: Crianças de 7 a 14 anos; renda familiar sem a renda da própria criança, 5% de cada ponta da distribuição de renda foram aparados.

períodos. A falta de crescimento durante a “década perdida” afetou de forma desigual a população. A diferença na variação entre os grupos de renda pode ser parcialmente explicada pela diferente capacidade dos indivíduos de defender sua renda real da inflação. Os diversos planos de estabilização tiveram diferentes impactos distributivos, em função de suas regras, por exemplo, sobre ajuste de preços e oferta de crédito, e também em função do cenário econômico gerado, de crescimento ou recessão. O Plano Real em 1994 levou a um aumento de

renda, que foi maior para os mais pobres. Parte desse ganho foi perdida ao longo da década de 1990, mas apenas o Norte urbano pobre (os dois grupos de renda mais pobres) chegou ao terceiro período com uma média de renda menor do que a média do início dos anos 1980.

Vemos nas tabelas apresentadas que, embora a frequência escolar tenha crescido ao longo de todo o período de análise e a renda tenha passado por ciclos de crescimento e queda, os maiores aumentos de frequência estão relacionados aos períodos de crescimento de renda.

Para melhor observar a relação entre a variação da frequência escolar e a da renda, o procedimento de diferenças em diferenças é utilizado, comparando-se a variação no tempo da desigualdade na escola com a desigualdade de renda por região. O Gráfico 3 (conjunto de gráficos) mostra que há uma relação entre as variações no tempo da diferença de frequência escolar e na desigualdade de renda (lembramos que a hipótese de identificação diz que não há outro efeito da desigualdade de renda sobre a frequência que não seja o efeito via nível de renda).

A desigualdade de renda é medida pelo logaritmo da razão das rendas dos grupos definidos pelos 4º e 1º quartis. A diferença de frequência escolar refere-se à diferença das probabilidades médias desses mesmos grupos.

A evolução da diferença da frequência escolar por região está no Gráfico 3. As regiões Sudeste e Centro-Oeste apresentam uma diminuição relativamente constante dessa diferença no tempo, enquanto a diminuição nas outras regiões começa apenas na década de 1990. As mais altas diferenças de frequência são observadas no primeiro período nas regiões Nordeste e Centro-Oeste.

O Gráfico 4 (conjunto de gráficos) mostra a evolução da desigualdade de renda no tempo por região. Como esperado, pela Tabela 5, há um aumento geral na desigualdade na década de 1980, que pára na década seguinte. A região Sudeste mostra o menor aumento, enquanto Norte e Nordeste, apesar de começarem o período com desigualdades menores, mostram o maior aumento. Nos anos 1990 a desigualdade fica praticamente inalterada, com alguma melhora nos últimos anos.

A análise da variação das desigualdades de renda e frequência escolar sugere que a estabilização e a distribuição de renda relativamente estável da década de 1990 são relacionadas com o decréscimo mais acentuado na desigualdade de frequência escolar. Isso aponta para uma relação positiva entre renda e escola, relação a ser estimada na próxima seção.

Os Gráficos 3 e 4 restringem-se à análise da variação diferencial da evolução da renda no tempo para o 1º e o 4º quartil. Gráficos similares de diferenças em

GRÁFICO 3
DIFERENÇA DE PROBABILIDADE DE FREQUÊNCIA ESCOLAR ENTRE O 4º E O 1º GRUPO DE RENDA

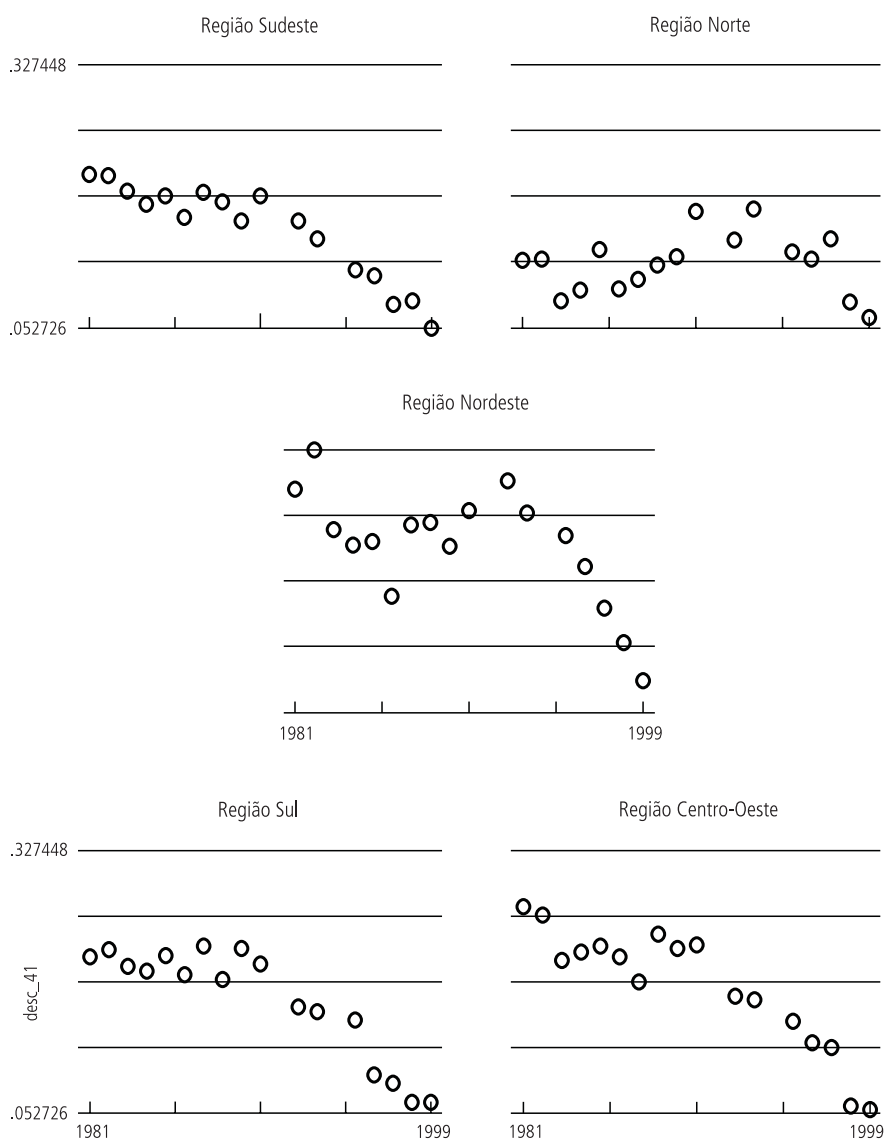
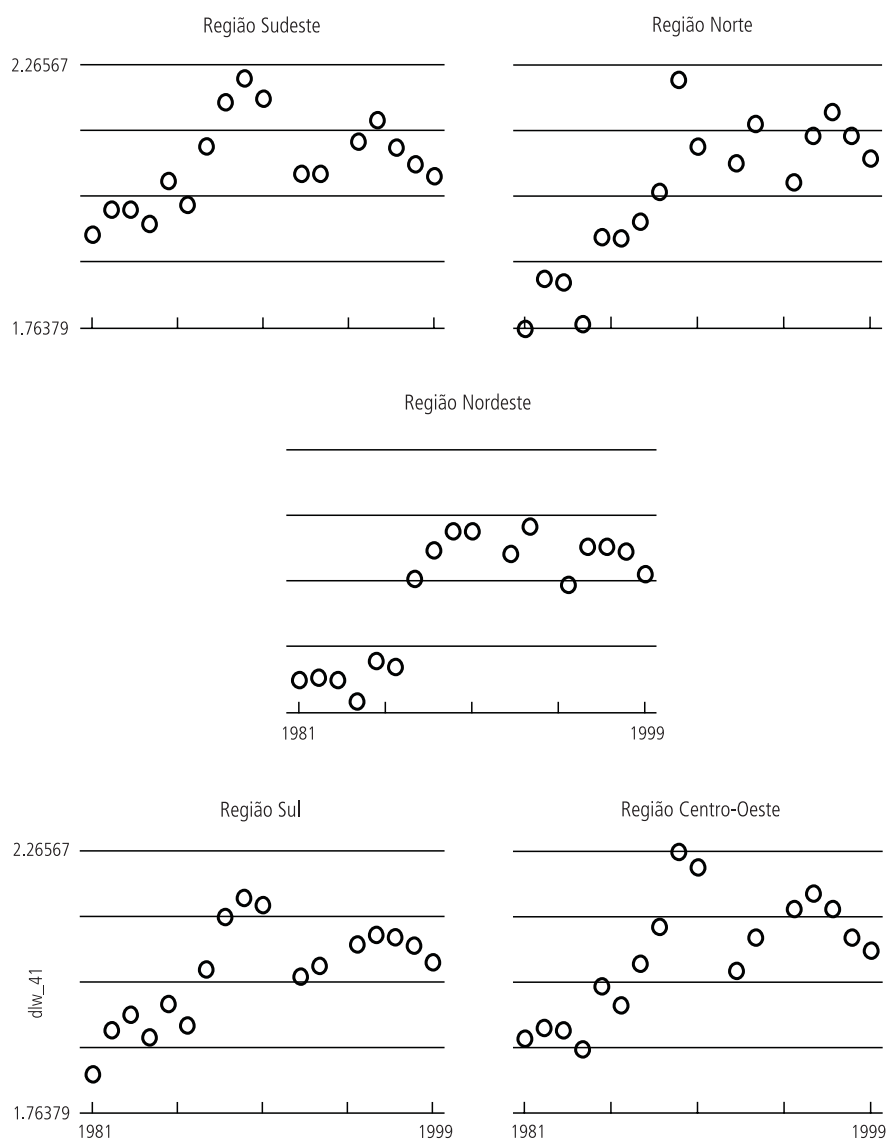


GRÁFICO 4

DIFERENÇA DE RENDA (ln) ENTRE O 4º E O 1º GRUPO DE RENDA ENTRE ANOS E REGIÕES



diferenças foram construídos relacionando as médias desses grupos de renda à mediana (se renda) e à média (se probabilidade de frequência) da amostra total. A evolução dessas desigualdades mostra padrão similar ao apresentado nos Gráficos 3 e 4. Destacam-se os fatos de o aumento de desigualdade das regiões Norte e Nordeste referir-se apenas aos grupos mais ricos e da diminuição de desigualdade escolar das regiões Sudeste e Centro-Oeste aplicar-se apenas aos grupos mais ricos.

Vale lembrar que o Gráfico 4 (conjunto de gráficos) descreve apenas parte da variação de renda usada como instrumento, pois a variação entre todos os quartis de renda é considerada, e não apenas entre o 4º e o 1º quartil.

Outra forma de se comprovar se há variação suficiente de renda é realizada por meio de uma regressão da desigualdade de renda em função de tendências por região, controladas por *dummies* de ano e região. Procurou-se verificar se as tendências por região são diferentes. A Tabela 6 mostra os coeficientes das tendências por região, em que a variável dependente é a desigualdade de renda entre o 4º e o 1º quartil e a tendência é calculada como a interação das *dummies* de região com uma variável categórica para os anos. Comparando-se o Sudeste com as outras regiões, observam-se tendências diferentes, tendo o Norte e o Nordeste evolução similar, assim como o Sul e o Centro-Oeste. A igualdade de todos os coeficientes foi rejeitada por um teste-F.

TABELA 6
TENDÊNCIA POR REGIÃO DA DESIGUALDADE DE RENDA

Região	Coefficiente	Erro-padrão	Teste-F
Sudeste	0,005	0,002	$F(4,60) = 12,84$
Norte	0,018	0,002	Valor-p = 0,00
Nordeste	0,015	0,002	
Sul	0,009	0,002	
Centro-Oeste	0,011	0,002	

Nota: Desigualdade = log da razão da renda entre o 4º e o 1º quartil; tendência = variável categórica por região: 1 a 17 para os anos 1981 a 1999, 0 se outra região; teste-F: se os coeficientes das tendências são iguais; estão incluídas *dummies* de ano e região.

5 RESULTADOS

A Tabela 7 resume as estimativas do efeito-renda para diferentes especificações e métodos de estimação. A variável dependente é uma variável binária para frequência escolar. A definição de renda familiar é o logaritmo da renda familiar *per capita* (sem a renda da própria criança, quando houver). A primeira regressão não inclui controles, dessa forma estima-se o efeito “bruto” da renda familiar,

TABELA 7
EFEITO DA RENDA FAMILIAR (LN) SOBRE FREQUÊNCIA ESCOLAR

Especificação	MQO				Probit ^a
	Total	1981-1986	1987-1992	1993-1999	Total
1	0,100 (0,004)	0,111 (0,006)	0,103 (0,005)	0,068 (0,006)	0,104 (0,004)
2	0,067 (0,003)	0,089 (0,005)	0,070 (0,004)	0,043 (0,003)	0,084 (0,003)
3	0,043 (0,003)	0,058 (0,005)	0,041 (0,004)	0,030 (0,003)	0,059 (0,003)
4	0,020 (0,003)	0,024 (0,004)	0,016 (0,004)	0,014 (0,003)	0,027 (0,002)
5	0,022 (0,002)	0,027 (0,003)	0,017 (0,003)	0,014 (0,002)	0,026 (0,002)
IV — 4 ^b	-0,027 (0,029)	-0,055 (0,056)	-0,003 (0,031)	0,003 (0,062)	-
Teste-F ^c	173,89 [0,0]	141,80 [0,0]	124,18 [0,0]	97,75 [0,0]	
R ² parcial ^d	0,070	0,033	0,044	0,036	
Hansen ^e	219,835 [0,768]	36,570 [0,708]	24,780 [0,531]	43,106 [0,635]	
Número de observações	959.264	441.380	236.235	281.649	

Nota: Erro-padrão entre parênteses; valor-*p* entre colchetes. É usado estimador de variância robusto à existência de *cluster*, onde o *cluster* é definido por ano-região-quartil.

^a Os efeitos marginais do *probit* foram calculados para os valores médios dos regressores.

^b A estimativa de variáveis instrumentais para a amostra total não inclui o ano de 1981 (em função de problemas computacionais relacionados ao tamanho da amostra).

^c Significância conjunta dos instrumentos excluídos no primeiro estágio do 2SLS.

^d Relativo à regressão do primeiro estágio.

^e Teste de superidentificação.

isto é, o efeito da renda e de suas correlacionadas. As regressões seguintes adicionam novos controles. A regressão 2 adiciona as *dummies* de ano, região e quartil. As regressões seguintes incluem também características familiares e individuais. Como escolaridade dos pais e renda familiar são muito correlacionadas — cerca de 50% de correlação —, a escolaridade dos pais entra apenas na regressão 4, e mostra como sua inclusão é importante para corrigir o efeito-renda estimado. A especificação 5 adiciona o segundo nível de interações para ano, região e quartil de renda. Todas as variáveis de controle são incluídas como

conjunto de *dummies*, à exceção da renda, permitindo, assim, efeitos diferenciados entre os níveis dessas variáveis.

Os conjuntos de controles são os seguintes:

1. nenhum controle;
2. ano, região, quartil de renda;
3. gênero, área urbana, idade, número de filhos, raça, presença de apenas um dos pais, ano, região e quartil;
4. escolaridade do pai, escolaridade da mãe, gênero, área urbana, idade, número de filhos, raça, presença de apenas um dos pais, ano, região e quartil; e
5. escolaridade do pai, escolaridade da mãe, gênero, área urbana, idade, número de filhos, raça, presença de apenas um dos pais, ano, região, quartil, interações de segunda ordem de ano, região e quartil.

A estimativa por MQO do efeito da renda diminui no tempo e com a inclusão de controles adicionais. A inclusão de educação dos pais diminui a magnitude da estimativa à metade. As estimativas por *probit* mostram o mesmo padrão. Essa diminuição de magnitude entre as especificações 1 e 4, com a inclusão de mais controles, revela que realmente havia problema de variável omitida. Por outro lado, não há mudança significativa das estimativas com a inclusão das interações de segunda ordem entre quartil, ano e região na especificação 5, o que mostra que essas interações não têm impacto direto sobre a frequência escolar. Esse fato é importante, pois reforça a possibilidade de inclusão dessa variação como parte do instrumento excluído.

O efeito da renda se mantém significativo para todas as estimativas por MQO, apesar do aumento do erro-padrão (relativamente à magnitude da estimativa) com a inclusão da educação dos pais. As estimativas caem de 0,024 no primeiro período para 0,014 no último (especificação 4). Também ocorre uma diminuição da magnitude das estimativas das outras variáveis relacionadas às características familiares, assim como uma diminuição do poder explicativo (R^2) das estimativas no tempo.

Sendo a especificação 4 a preferida para a análise, os comentários que seguem sobre as variáveis de controle referem-se a esta especificação (estimativas no Apêndice). A educação dos pais tem efeito positivo e significativo sobre frequência escolar. A magnitude dos efeitos de primário completo ou mais anos de escolaridade é a mesma quando comparada a nenhuma educação dos pais. Além disso, o efeito da escolaridade da mãe é maior que o da escolaridade do pai. Ser do sexo feminino e morar em área urbana têm efeito positivo sobre a frequência escolar. Idade apresenta um efeito côncavo, mas o efeito é positivo

para todas as idades se comparadas aos 7 anos de idade.¹⁴ Ser branco ou oriental leva a uma probabilidade maior de estar na escola. A existência de mais de 4 filhos na família afeta negativamente a probabilidade de estar na escola (ter mais de 11 filhos não tem efeito significativo, provavelmente em função das poucas observações com essa característica). É interessante notar que ter até 4 filhos gera um efeito pequeno, mas positivo e significativo em relação a apenas um filho, sugerindo algum ganho de escala em mandar mais de um filho para a escola. A probabilidade de estar na escola é maior na região Sudeste, e é maior para os primeiros anos da década de 1980 (capturando a presença de crianças na pré-escola na amostra) e final da década de 1990. As *dummies* para os grupos de renda continuam significativas para os dois grupos de renda mais ricos, relativamente ao grupo mais pobre omitido mesmo após a introdução de todos os controles de características familiares.

Os próximos comentários referem-se ao método de 2SLS aplicado à especificação 4. O teste-F de significância dos instrumentos excluídos na regressão do primeiro estágio e o R^2 parcial confirmam a existência de correlação entre a variável endógena, a renda familiar e o instrumento excluído. O teste de superidentificação sugere que não há correlação entre o instrumento e o termo de erro.

As estimativas por variáveis instrumentais não mostram nenhuma relação entre renda familiar e frequência escolar. O fato de esse método de estimação também resolver o problema de viés de atenuação apenas reforça seu resultado não-significativo. Esse resultado, à primeira vista, sugere que a correlação freqüentemente encontrada entre renda familiar e frequência escolar representa antes o efeito de variáveis não-observadas, como habilidades e preferências familiares.

É preciso, no entanto, levar em conta que a estimativa mede o efeito médio individual. Na presença de efeito heterogêneo entre indivíduos, é possível que alguns grupos apresentem efeito significativo da renda. Vimos que o efeito da renda diminuiu ao longo do tempo; é possível que ele também varie entre grupos de renda, o que veremos a seguir.

Para verificar essa possibilidade, criaram-se quatro variáveis de renda por meio da interação da renda com as *dummies* de grupos de renda, e essas interações foram usadas na estimativa por MQO. Como mostrado na Tabela 8, o efeito-renda é realmente diferente entre grupos de renda. Ele nunca é significativo para o grupo de renda mais alta, e permanece significativo no período mais recente apenas para o grupo mais pobre; o efeito-renda, inclusive, aumenta no último período para esse grupo mais pobre.¹⁵ As estimativas do efeito-renda

14. Idade no mês de junho.

15. A inclusão dos 5% mais pobres da distribuição de renda, que foram aparados, iria provavelmente reforçar a significância do efeito-renda para as famílias pobres.

TABELA 8
EFEITO DA RENDA FAMILIAR (LN) POR GRUPOS DE RENDA

	MQO				Probit total	IV 1993-1999
	Total	1981-1986	1987-1992	1993-1999		
1° quartil	0,049 (0,006)	0,037 (0,008)	0,036 (0,009)	0,047 (0,012)	0,042	0,238 (0,109)
2° quartil	0,043 (0,008)	0,045 (0,013)	0,035 (0,010)	0,025 (0,015)	0,034	0,133 (0,090)
3° quartil	0,022 (0,007)	0,032 (0,011)	0,022 (0,008)	0,006 (0,011)	0,024	0,107 (0,087)
4° quartil	-0,002 (0,004)	0,010 (0,006)	-0,004 (0,004)	-0,006 (0,005)	0,011	0,105 (0,114)
Número de observações	959.264	441.380	236.235	281.649		

Nota: Estimativas baseadas na especificação 4; erro-padrão entre parênteses; MQO — 4 variáveis-renda criadas por quartil; *probit* — 1 variável-renda, efeitos marginais calculados na média de cada quartil.

por *probit*, calculadas sobre as médias de renda de cada grupo de renda, confirmam o maior efeito para os mais pobres.

Essas estimativas, porém, ainda podem sofrer do problema de viés de variável omitida. A última coluna da tabela mostra as estimativas por variáveis instrumentais para o último período, 1993 a 1999, e mostra o mesmo padrão de resultado, isto é, o efeito-renda é significativo apenas para o grupo mais pobre. Portanto, mesmo controlando-se por variáveis omitidas, a renda continua importante para explicar a frequência escolar para o grupo de renda definido pelo menor quartil de renda.

É importante, no entanto, observar que o instrumento foi criado originalmente para incluir a variação da renda entre os grupos definidos pelos quartis de renda. Estimando-se quatro efeitos-renda, definidos justamente pelos quartis de renda, utiliza-se apenas a variação da renda familiar entre regiões e anos.¹⁶ A maior magnitude das estimativas por variável instrumental pode ser consequência dessa menor variação no instrumento.

Resumindo, a estimativa por MQO para o quartil mais pobre é de 0,047, muito maior do que a estimativa média, de 0,014, e a estimativa por variáveis

16. Uma alternativa para manter a variação entre quartis seria criar grupos definidos por outras medidas de renda, de forma a ainda haver variação de quartis de renda dentro dos grupos. O melhor candidato seria a escolaridade dos pais, mas essa medida é muito correlacionada com a renda e o resultado seria o mesmo.

instrumentais para esse mesmo grupo de renda é de 0,238, sendo a estimativa média não-significativa. Essas estimativas, portanto, corroboram a hipótese de que o efeito da renda familiar sobre a frequência escolar é heterogêneo entre grupos de renda e é significativo para o grupo de renda mais pobre.

6 COMENTÁRIOS

Este trabalho explorou a variação da distribuição de renda no período 1981-1999 para determinar se há relação causal entre a renda familiar e a frequência escolar para crianças em idade de escola obrigatória no Brasil. Não foi encontrado efeito causal quando se supôs um efeito-renda homogêneo, mas se estimou um efeito-renda significativo para o grupo definido pelo quartil de renda mais pobre, quando se permitiu que o efeito da renda variasse entre os grupos de renda definidos pelos quartis de renda.

Políticas como o Programa Bolsa-Escola, hoje incluído no Bolsa-Família, podem ser, portanto, efetivas em aumentar a frequência escolar se focarem os grupos de renda mais pobres. A compulsoriedade de frequência escolar para o recebimento da transferência de renda garante a melhora de escolaridade mesmo sem uma relação causal entre renda e frequência, mas a existência dessa relação causal sugere que, mesmo sem essa compulsoriedade (ou fiscalização de seu cumprimento), a frequência escolar aumentaria com o aumento de renda trazido pela transferência.

Fica para uma pesquisa futura a determinação da magnitude da transferência de renda necessária para aumentar em pelo menos um ano a frequência escolar. Outra pesquisa interessante seria refazer as estimativas aqui apresentadas para a faixa etária fora da idade de escola obrigatória, pois o custo de oportunidade de se estar na escola aumenta com a idade, e, pelo menos para o ensino superior, existem custos diretos relevantes relacionados ao estudo.

APÊNDICE

TABELA A.1
NÚMERO DE OBSERVAÇÕES POR REGIÃO E ANO

Ano	Região					Total
	Sudeste	Norte	Nordeste	Sul	Centro-Oeste	
1981	21.138	1.576	16.173	8.758	3.705	51.349
1982	21.154	1.627	16.181	8.563	3.869	51.394
1983	21.653	1.706	16.387	8.416	3.970	52.133
1984	21.766	1.800	16.939	8.463	3.931	52.900
1985	22.922	2.031	17.098	8.562	4.163	54.777
1986	23.513	2.072	17.830	8.634	4.332	56.380
1987	23.938	2.204	18.100	8.654	4.226	57.122
1988	24.721	2.210	18.379	8.924	4.388	58.622
1989	25.571	2.331	18.911	9.328	4.528	60.671
1990	24.511	2.569	18.945	8.994	4.511	59.529
1992	24.951	2.702	17.023	8.703	4.556	57.935
1993	25.123	2.866	17.529	9.000	4.672	59.190
1995	24.610	2.858	18.475	8.809	4.686	59.437
1996	24.082	2.867	17.775	8.739	4.611	58.075
1997	23.461	2.793	17.593	8.768	4.729	57.344
1998	23.212	2.817	17.119	8.539	4.566	56.254
1999	23.039	2.773	17.327	8.510	4.505	56.154
Total	399.365	39.802	297.785	148.364	73.948	959.264

Fonte: PNADs de 1981 a 1999.

Nota: Crianças de 7 a 14 anos de idade, de famílias com renda positiva; 5% de cada ponta da distribuição de renda foram cortados da amostra.

TABELA A.2
REGRESSÃO — FREQUÊNCIA ESCOLAR

MQO com erros robustos (especificação 4)					
		R^2			
					0,149
					340
					959.264
	Coefficiente	Erro-padrão		Coefficiente	Erro-padrão
LN(renda)	0,019	0,003	Número de irmãos_2	0,005	0,001
Escolaridade do pai_1	0,047	0,004	Número de irmãos_3	0,009	0,002
Escolaridade do pai_2	0,052	0,003	Número de irmãos_4	0,002	0,002
Escolaridade do pai_3	0,062	0,003	Número de irmãos_5	-0,013	0,003
Escolaridade do pai_4	0,076	0,002	Número de irmãos_6	-0,030	0,003
Escolaridade do pai_5	0,065	0,004	Número de irmãos_7	-0,037	0,004
Escolaridade do pai_6	0,071	0,004	Número de irmãos_8	-0,056	-0,005
Escolaridade do pai_7	0,066	0,003	Número de irmãos_9	-0,074	0,009
Escolaridade do pai_8	0,070	0,003	Número de irmãos_10	-0,091	0,012
Escolaridade do pai_9	0,073	0,004	Número de irmãos_11	-0,074	0,023
Escolaridade do pai_10	0,068	0,003	Número de irmãos_12	-0,058	0,045
Escolaridade do pai_11	0,062	0,003	Número de irmãos_13	-0,135	0,074
Escolaridade do pai_12	0,063	0,006	Número de irmãos_14	0,015	0,190
Escolaridade do pai_13	0,064	0,004	Número de irmãos_15	-0,575	-0,024
Escolaridade do pai_14	0,063	0,004	Número de irmãos_16	-0,092	-0,007
Escolaridade do pai_15	0,058	0,003	Número de irmãos_22	0,185	0,011
Escolaridade do pai_16	0,058	0,004	Negro	-0,026	0,003
Escolaridade do pai_sem	0,082	0,016	Oriental	0,006	0,007
Escolaridade da mãe_1	0,070	0,004	Pardo	-0,008	0,002
Escolaridade da mãe_2	0,083	0,004	Sem raça	-0,024	0,007
Escolaridade da mãe_3	0,095	0,004	Um pai	-0,058	0,016
Escolaridade da mãe_4	0,120	0,003	Região Norte	-0,004	0,004
Escolaridade da mãe_5	0,115	0,004	Região Nordeste	-0,025	0,005

(continua)

(continuação)

MQO com erros robustos (especificação 4)					
			R^2	0,149	
			Número de <i>clusters</i>	340	
			Número de observações	959.264	
	Coefficiente	Erro-padrão		Coefficiente	Erro-padrão
Escolaridade da mãe_6	0,122	0,004	Região Sul	-0,010	0,004
Escolaridade da mãe_7	0,122	0,004	Região Centro-Oeste	-0,003	0,004
Escolaridade da mãe_8	0,125	0,004	Ano 1982	-0,023	0,016
Escolaridade da mãe_9	0,129	0,005	Ano 1983	0,038	0,010
Escolaridade da mãe_10	0,127	0,005	Ano 1984	0,041	0,010
Escolaridade da mãe_11	0,130	0,004	Ano 1985	0,037	0,010
Escolaridade da mãe_12	0,129	0,005	Ano 1986	0,018	0,014
Escolaridade da mãe_13	0,129	0,005	Ano 1987	0,001	0,013
Escolaridade da mãe_14	0,130	0,004	Ano 1988	0,009	0,013
Escolaridade da mãe_15	0,123	0,004	Ano 1989	-0,002	0,013
Escolaridade da mãe_16	0,119	0,005	Ano 1990	0,004	0,013
Escolaridade da mãe_sem	0,035	0,017	Ano 1992	0,013	0,014
Homem	-0,018	0,002	Ano 1993	0,025	0,013
Urbano	0,085	0,003	Ano 1995	0,035	0,013
Idade_8	0,109	0,003	Ano 1996	0,048	0,013
Idade_9	0,148	0,005	Ano 1997	0,069	0,014
Idade_10	0,166	0,007	Ano 1998	0,083	0,016
Idade_11	0,156	0,007	Ano 1999	0,096	0,019
Idade_12	0,128	0,008	2º quartil	0,008	0,006
Idade_13	0,079	0,009	3º quartil	0,025	0,006
Idade_14	0,013	0,010	4º quartil	0,035	0,008
			Constante	0,459	0,018

ABSTRACT

This paper exploits the changes in the income distribution in the period 1981 to 1999 to determine, using IV-estimates, if there is a causal relationship between family income and school attendance for children in compulsory school age in Brazil. No significant causal link was found when a homogeneous income effect across groups was assumed, but a significant income effect emerged for the poorest income group when the effect of family income on school attendance was allowed to vary between income quartile groups.

BIBLIOGRAFIA

- ACEMOGLU, D., PISCHKE, J. Changes in the wage structure, family income, and children's education. *European Economic Review*, v. 45, 2001.
- ANGRIST, J., IMBENS, G. Two-stage least squares estimation of average causal effects in models with variable treatment intensity. *Journal of the American Statistical Association*, v. 90, n. 430, June 1995.
- BANCO MUNDIAL. *Closing the gap in education and technology*, Cap. 5. Latin American and Caribbean Studies, 2002. Acessível em: <<http://www.worldbank.org>>.
- BARROS, R. *Os determinantes da desigualdade no Brasil*. IPE/USP, 1997 (Seminário, 22/97).
- BARROS, R. *et alii. Determinantes do desempenho educacional no Brasil*. IPEA, out. 2001 (Texto para Discussão, 834).
- BIRDSALL, N., SABOT, R. (eds.). *Opportunity foregone: education in Brasil*. Inter-American Bank of Development, 1996.
- BLUNDELL, R., DUNCAN, A., MEGHIR, C. Estimating labor supply responses using tax reforms. *Econometrica*, v. 66, n. 4, July 1998.
- BOUND, J., JAEGER, A., BAKER, R. Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *Journal of the American Statistical Association*, v. 90, n. 430, June 1995.
- BRASIL. *Constituição de 1988. Emenda Constitucional 14, de 1996*. Fundo de Manutenção e Desenvolvimento do Ensino Fundamental e de Valorização do Magistério (Fundef). Acessível em: <<http://prolei.cibec.inep.gov.br>>.
- . *Lei 9.394, dezembro de 1996. Lei de Diretrizes e Bases*. Acessível em: <<http://prolei.cibec.inep.gov.br>>.
- . *Lei 10.219, abril de 2001. Programa Nacional de Renda Mínima Vinculada à Educação Bolsa-Escola*. Acessível em: <<http://prolei.cibec.inep.gov.br>>.
- CARD, D. The causal effect of education on earnings. In: ASHENFELTER, O., CARD, D. (eds.). *Handbook of Labor Economics*. Elsevier, Amsterdam: North-Holland, v. 3, Cap. 30, 1999.
- . Estimating the return to schooling: progress on some persistent econometric problems. *Econometrica*, v. 69, n. 5, Sep. 2001.
- CARVALHO, I. *Household income as a determinant of child labor and school enrolment in Brasil: evidence from a social security reform*. Northeast Universities Development Consortium Conference (NEUDC). Boston: Boston University/Institute for Economic Development, 2001. Acessível em: <<http://www.bu.edu>>.

- DAVIDSON, R., MACKINNON, J. *Estimation and inference in econometrics*. Oxford University Press, 1993.
- HAVEMAN, R., WOLFE, B. The determinants of children's attainments: a review of methods and findings. *Journal of Economic Literature*, v. 33, Dec. 1995.
- IBGE. *Censo Demográfico de 2000*. 2002. Acessível em: <<http://www.ibge.gov.br>>.
- INEP/MEC. Instituto Nacional de Estudos e Pesquisas Educacionais/Ministério da Educação. *Dados estatísticos: números da educação no Brasil*. 2001. Acessível em: <<http://www.ibge.gov.br>>.
- . *Informe estatístico da educação básica: evolução recente das estatísticas da educação básica no Brasil*. Brasília, 1998. Acessível em: <<http://www.ibge.gov.br>>.
- KASSOUF, A. Trabalho infantil: escolaridade x emprego. *XXVIII Encontro Nacional de Economia (Anpec)*, 2000. Campinas: Unicamp, 2000. Acessível em: <<http://www.anpec.org.br>>.
- MOULTON, B. An illustration of a pitfall in estimating the effects of aggregate variables on micro units. *The Review of Economics and Statistics*, v. 72, n. 2, 1990.
- PNUD. *Informe sobre Desarrollo Humano*, 2000. Acessível em: <<http://www.undp.org>>.
- VASCONCELLOS, L. *Frequência e atraso escolar e sua relação com a renda familiar no Brasil*. São Paulo: FEA/USP, 2003 (Tese de Doutorado).

(Originais recebidos em julho de 2004. Revistos em maio de 2005.)

