

"A FEA e a USP respeitam os direitos autorais deste trabalho. Nós acreditamos que a melhor proteção contra o uso ilegítimo deste texto é a publicação online. Além de preservar o conteúdo motiva-nos oferecer à sociedade o conhecimento produzido no âmbito da universidade pública e dar publicidade ao esforço do pesquisador. Entretanto, caso não seja do interesse do autor manter o documento online, pedimos compreensão em relação à iniciativa e o contato pelo e-mail bibfea@usp.br para que possamos tomar as providências cabíveis (remoção da tese ou dissertação da BDTD)."

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E CONTABILIDADE
DEPARTAMENTO DE ECONOMIA

LÍGIA MARIA DE VASCONCELLOS

**FREQUÊNCIA E ATRASO ESCOLAR E SUA RELAÇÃO COM A RENDA
FAMILIAR NO BRASIL**

SÃO PAULO

2003

Reitor da Universidade de São Paulo

Prof. Dr. Adolpho José Melfi

Diretora da Faculdade de Economia, Administração e Contabilidade

Profa. Dra. Maria Tereza Leme Fleury

Chefe de Departamento de Economia

Elizabeth Maria Mercier Querido Farina

LÍGIA MARIA DE VASCONCELLOS

**FREQUÊNCIA E ATRASO ESCOLAR E SUA RELAÇÃO COM A RENDA
FAMILIAR NO BRASIL**

Tese apresentada ao Programa de Pós-graduação em Economia da FEA-USP como requisito para obtenção de título de doutor em Economia.

Orientador: Prof. Reynaldo Fernandes

SÃO PAULO

2003

FICHA CATALOGRÁFICA

Vasconcellos, Lígia Maria de
Frequência e atraso escolar e sua relação com a renda
familiar no Brasil / Lígia Maria de Vasconcellos. -- São Paulo :
FEA/USP, 2003.
128 p.

Tese - Doutorado
Bibliografia

1. Educação – Aspectos econômicos 2. Renda familiar
- Brasil I. Faculdade de Economia, Administração e
Contabilidade da USP.

CDD – 370

Este trabalho não teria sido possível sem o apoio financeiro da CAPES - Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Ministério da Educação e da FIPE - Fundação Instituto de Pesquisas Econômicas.

Agradeço a minha família;

a Adriana Schor, amiga e interlocutora, pela leitura atenta e pelos comentários que acompanharam este trabalho, desde um primeiro esboço até sua versão final;

ao professor Naércio Aquino Menezes Filho, por suas sugestões e comentários por ocasião do exame de qualificação e na fase final do trabalho;

a Marislei Nishijima, amiga e interlocutora, pela leitura minuciosa de textos relacionados a esta pesquisa;

ao professor Paulo Picchetti, por suas sugestões por ocasião do exame de qualificação;

aos co-orientadores Richard Dickens e Marco Manacorda, pelo incentivo e atenção durante minha estada no Queen Mary, University of London;

agradeço, sobretudo, a meu orientador Reynaldo Fernandes pelas sugestões e comentários que estimularam esta pesquisa e pela autonomia que me foi dada durante todo o trabalho.

Sumário

Introdução.....	pág. 01
Capítulo 1. Indicadores educacionais.....	pág. 04
1.1. Educação no Brasil e na América Latina.....	pág. 04
1.2. Frequência e atraso escolar.....	pág. 11
Capítulo 2. Revisão da literatura.....	pág. 21
2.1. Trabalhos teóricos.....	pág. 21
2.2. Trabalhos empíricos.....	pág. 25
2.2.1. Os determinantes dos resultados educacionais.....	pág. 27
2.2.2. Renda familiar e resultados educacionais.....	pág. 33
Capítulo 3. Fatores determinantes da educação no Brasil.	pág. 39
3.1. Políticas públicas.....	pág. 40
3.2. Características familiares e individuais.....	pág. 46
3.3. Estimativas.....	pág. 52
Capítulo 4. Análise de decomposição no tempo.....	pág. 62
4.1. Decomposição para a amostra total.....	pág. 69
4.2. Decomposição por quartil de renda.....	pág. 74
Capítulo 5. Frequência e atraso escolar e sua relação com a renda familiar.....	pág. 79
5.1. Estratégia empírica.....	pág. 81
5.1.1. Apresentação do instrumento.....	pág. 81
5.1.2. Especificações e questões econométricas.....	pág. 86
5.1.3. Efeitos de grupo e efeitos heterogêneos.....	pág. 89
5.2. Evolução da renda e da educação.....	pág. 90
5.3. Estimativas.....	pág. 96
Comentários finais.....	pág.105

Anexos

A1. Descrição dos dados da PNAD.....	pág.107
A2. Definição de variáveis e da amostra.....	pág.110
A3. Outros indicadores educacionais.....	pág.115
A4. Outras estatísticas da evolução da renda e da educação.....	pág.121
 Bibliografia.....	 pág.125

RESUMO

Analisam-se a freqüência e o atraso escolar no Brasil e seus principais determinantes no período de 1981 a 1999. Em especial, utilizando-se a variação da distribuição da renda no tempo e entre regiões, verifica-se se há relação causal entre renda familiar e os indicadores educacionais: estimou-se que não há efeito direto da renda familiar sobre a freqüência escolar.

ABSTRACT

The evolution of school attendance and delay on school and their determinants are analysed for Brazil in the period from 1981 to 1999. Specifically the change in income distribution across regions and years is exploit to verify the existence of a causal relationship between family income and the educational indicators: no direct effect of income on attendance was found.

INTRODUÇÃO

O objetivo deste trabalho é determinar a existência de relação causal entre renda familiar e educação dos filhos no Brasil.

Esta análise é especialmente relevante para o Brasil, em função de seus ainda baixos indicadores de educação no começo deste século e da importância que têm tomado programas de renda mínima vinculada à educação como forma de manter as crianças na escola. Os resultados deste trabalho podem gerar informações relevantes para o melhor desenho deste tipo de programa.

A renda familiar, contudo, não é o único fator determinante da educação, portanto, para melhor situar sua importância, será analisado o papel de outros importantes determinantes da educação para depois concentrar a atenção no papel da renda.

A análise será feita a partir dos dados da Pesquisa Nacional por Amostra de Domicílios - PNAD¹, considerando o período de 1981 a 1999.

Dada a grande mudança nos indicadores educacionais neste período, principalmente no final da década de 90, será apresentada a trajetória destes indicadores e de seus determinantes no período e analisada a variação ocorrida na importância destes determinantes no tempo.

Não há dúvida sobre a existência de correlação entre renda familiar e educação dos filhos, porém não é claro se esta relação é direta ou se apenas capta o efeito de outras variáveis sobre a decisão de educação, como o nível educacional dos pais ou suas preferências em relação à formação dos filhos. A existência de

¹ Pesquisa realizada pelo Instituto Brasileiro de Geografia e Estatística - IBGE.

efeito causal reforça o uso de políticas de renda mínima que visem afetar a educação; contudo, a não-existência de relação causal não impede a eficácia de tais políticas, mas indica a necessidade delas criarem o vínculo entre o recebimento do benefício e a educação.

A estratégia deste trabalho será explorar as variações ocorridas na distribuição de renda familiar no período considerado para determinar a existência de efeito causal da renda familiar sobre a educação.

O primeiro capítulo apresenta alguns indicadores educacionais para o Brasil e países da América Latina e, em especial, os dois indicadores para o Brasil que serão examinados com maior atenção no decorrer do trabalho, frequência e atraso escolar. Veremos que os indicadores do Brasil estão em muitos casos aquém de seus pares latino-americanos, por outro lado, houve grande melhora do desempenho escolar nos últimos anos, principalmente no que tange à frequência escolar de crianças em idade de escola obrigatória.

O segundo capítulo faz uma breve revisão da literatura, apresentando trabalhos teóricos e empíricos que buscam entender quais são os principais fatores determinantes da educação. Estes fatores podem ser resumidos em dois grandes blocos, os determinados por políticas públicas e as características das famílias.

O terceiro capítulo apresenta, então, para o Brasil, as principais políticas públicas ligadas à educação do período e as características familiares e individuais que a literatura considera mais relevantes para explicar a educação. Apresenta também as primeiras estimativas para os indicadores educacionais, frequência e atraso escolar. A frequência escolar refere-se a crianças que freqüentam ensino fundamental ou médio, e é definido como atraso escolar o caso em que a criança ou jovem não está na

série esperada de acordo com sua idade, independentemente de estar cursando a escola.

O capítulo seguinte compara as mesmas estimativas, agora calculadas para o início e o final do período, de forma a verificar como a mudança no tempo das características familiares observadas e do impacto por elas gerado contribuíram para a variação no período das médias de frequência e atraso escolar.

O quinto capítulo apresenta a análise específica do efeito da renda familiar sobre estes dois indicadores educacionais. É explicada a estratégia empírica utilizada e apresentados e discutidos os resultados estimados.

Por fim, os principais resultados encontrados são resumidos e comentados.

capítulo 1: INDICADORES EDUCACIONAIS

1.1. Educação no Brasil e na América Latina

A educação no Brasil, apesar de uma melhora expressiva nos últimos anos, ainda apresenta resultados aquém dos desejados, como sugere a comparação com países com grau de desenvolvimento semelhante, que têm, portanto, as mesmas condições econômicas para atingir um determinado nível de resultados educacionais.

Em geral são comparados os níveis educacionais das populações adultas dos países, pois de alguma forma estes níveis refletem a qualificação dos trabalhadores e sua capacidade de gerar renda, para si próprios e para o crescimento do país. Não menos importantes são os indicadores do desempenho educacional da população na faixa etária escolar.

A tabela 1 compara alguns indicadores do Brasil com países da América Latina. É apresentada a ordenação dos países de acordo com seu índice de desenvolvimento humano - IDH, publicado pelo Programa das Nações Unidas para o Desenvolvimento - PNUD. O índice é construído combinando medidas de renda, saúde e educação. O Brasil era o 73^o país em uma lista de 150 países no ano de 2000. O índice educacional utilizado é uma combinação da taxa de alfabetização de adultos e da taxa bruta de escolarização, ambas incluídas na tabela 1 e abaixo comentadas.

A taxa de alfabetização de adultos no Brasil está aquém do valor alcançado por seus pares, 85% contra médias bem acima de 90%. A taxa bruta de escolarização é medida como a razão entre o número de alunos em um dado nível escolar, independentemente da idade, e o número de crianças na faixa etária correspondente a este nível escolar. A taxa de escolarização apresentada combina as taxas dos três níveis de ensino (fundamental, médio, e

superior) e está entre as melhores dos países selecionados, porém, como veremos adiante, ela esconde as grandes distorções entre idade e série escolar cursada nos diversos níveis de ensino.

Tabela 1. Indicadores educacionais

País	Ordenação IDH ¹	Taxa de alfabetização ²	Taxa bruta de escolarização ³	Anos de escolaridade ⁴
ano	2000	2000	1999	2000
Argentina	34	96,8	83	8,49
Chile	38	95,8	78	7,89
Uruguay	40	97,7	79	7,25
Costa Rica	43	95,6	67	6,00
México	54	91,4	71	6,73
Cuba	55	96,7	76	7,78
Colômbia	68	91,7	73	5,01
Venezuela	69	92,6	65	5,60
Brasil	73	85,2	80	4,56
Peru	82	89,9	80	7,33
Bolívia	114	85,5	70	5,54

Nota. 1: ordenação do IDH entre 150 países; 2: taxa de alfabetização de adultos (15 anos de idade ou mais); 3: taxa bruta combinada de escolarização, incluindo primário, secundário e terciário; 4: população adulta.

Fontes. 1,2,3: Informe sobre Desarrollo Humano - PNUD, 2000; 4: Closing the gap in education and technology - Banco Mundial, 2002.

Os anos de escolaridade média no Brasil estão também aquém dos países selecionados; a população adulta teria, em média, não mais que o antigo primário completo. Por outro lado, os jovens de 14 a 17 anos de idade na década de 90 já possuíam, em média, o primário completo¹. Jovens de 17 anos em 1999 tinham uma média de 6,8 anos completos de educação, uma grande melhora se comparado à média da população adulta. Não se pode esquecer, contudo, que estes mesmos jovens já poderiam ter 10 anos completos de estudo caso estivessem cursando a série escolar de acordo com sua idade.

FILMER e PRICHETT (1998) apresentam os resultados educacionais de frequência e evasão para jovens entre 15 e 19 anos de diversos

¹ Ver anexo A3, tabela A7, para informações mais detalhadas.

países, separando os jovens em 3 grupos de renda, os 40% mais pobres, os 40% seguintes, e os 20% mais ricos. Os autores agrupam os países de acordo com os resultados encontrados. Por exemplo, o sul da Ásia apresenta a situação ideal de alto índice de entrada na escola e baixa evasão, já a África Central apresenta o pior desempenho, com baixo índice de entrada e alta evasão. A América Latina tem também um padrão comum de resultados entre seus países: existe praticamente universalização do ensino se considerada a entrada na escola, isto é, alta taxa de frequência no 1º ano escolar, porém a evasão também é alta, gerando baixo nível de escolaridade. O grupo dos 40% mais pobres é praticamente o único responsável pelos índices de evasão.

Apesar do padrão comum latino-americano, podemos ver algumas diferenças entre os países analisados, como mostra a tabela 2. O Brasil apresenta a menor proporção de jovens de 15 a 19 anos que completaram pelo menos a 5ª série, principalmente se comparados os grupos mais pobres de cada país. Em relação à série completada pelo jovem mediano de cada grupo de renda, o resultado não é muito diferente entre os países, todos com uma diferença muito grande entre mais pobres e mais ricos: a mediana dos mais pobres equivale basicamente ao nível primário, e a dos mais ricos equivale a completar o ensino fundamental.

Tabela 2. Indicadores educacionais - jovens de 15 a 19 anos

País	Ano	Proporção completou 5ª série		Série completada - mediana	
		+ pobres	+ ricos	+ pobres	+ ricos
Colômbia	1995	63,0	94,2	5	9
Brasil	1996	45,7	89,5	4	8
Peru	1996	74,6	96,4	5	9
Bolívia	1993-94	70,5	93,7	6	10

Fonte. FILMER e PRICHETT (1998).

Apesar dos autores concluírem que o grande problema é a evasão, na verdade grande parte do problema, pelo menos no caso

do Brasil, é a alta repetência. Parte dos jovens de 15 a 19 anos ainda estão na escola, e portanto a diferença entre as proporções de jovens que completaram duas séries consecutivas não representa apenas a proporção de evadidos entre estas séries, pois inclui o resultado da repetência.

Voltando a atenção para o Brasil, abaixo são apresentados e comentados mais alguns indicadores.

A taxa de atendimento, número de alunos na escola em determinada faixa etária em relação à população na mesma faixa etária, é apresentada na tabela 3 e comprova que praticamente se atingiu a universalização do ensino para a faixa etária de 7 a 14 anos no final da década de 90. Pode-se constatar também a melhora expressiva da taxa de atendimento para a faixa de 15 a 17 anos de idade no período curto de 1994 para 1999. Estas faixas etárias equivalem, respectivamente, às idades esperadas para os alunos do ensino fundamental e do ensino médio.

Apesar da freqüência escolar, medida pela taxa de atendimento, apresentar bons resultados, o atraso escolar é ainda um grave problema na educação brasileira, como atestam as taxas de escolarização apresentadas na tabela 3. A taxa bruta de escolarização é medida relacionando a população de alunos, independentemente de sua idade, com a população da faixa etária esperada para o nível de ensino. Ela ultrapassa 100% para o ensino fundamental, mostrando a existência de alunos fora da faixa etária adequada para este nível de ensino; em 1999, a taxa é surpreendentemente alta, 130.5%. A taxa de escolarização líquida, que considera para seu cálculo apenas os alunos na idade esperada, apresenta valores mais razoáveis, 95.4% em 1999. A mesma distorção ocorre nas taxas do ensino médio, onde a maioria dos alunos está acima da faixa etária esperada. A taxa bruta é muito alta, pois inclui todos os alunos, e a taxa líquida é muito baixa, apenas 32.6% em 1999. Outra vez, este dado deve ser analisado com cautela, pois na verdade grande parte destes

adolescentes não está fora da escola, mas sim atrasada ainda no nível anterior, como atestam as taxas de atendimento.

Tabela 3. Taxas de Atendimento e de Escolarização por Nível de Ensino

Ano	Taxa de Atendimento		Taxa de Escolarização			
			Ensino Fundamental		Ensino Médio	
	7 a 14 anos	15 a 17 anos	Bruta	Líquida	Bruta	Líquida
1994	92.7	68.7	110.2	87.5	47.6	20.8
1999	97.0	84.5	130.5	95.4	74.8	32.6

Fonte: Ministério da Educação e Desporto - MEC/Instituto Nacional de Estudos e Pesquisas Educacionais - INEP (2001).

O Censo escolar de 1996 registrou que mais de 46% dos alunos frequentando o ensino fundamental têm idade superior à faixa etária correspondente a cada série. Este fato não é apenas resultado das altas taxas de repetência, mas também da entrada tardia na escola. As classes de alfabetização, que se configuram em um nível intermediário entre educação infantil e ensino fundamental e se concentram na região Nordeste, contribuem para este resultado, pois grande parte das crianças aí matriculada tem mais de 7 anos de idade. As classes de alfabetização não são mais incentivadas, mas em 1997 ainda abrigavam 25% das crianças matriculas entre pré-escola e classes de alfabetização.

Em decorrência do atraso no ensino fundamental, o ensino médio atende majoritariamente jovens acima da faixa etária prevista (acima de 17 anos). A análise deste capítulo, no entanto, concentrar-se-á na faixa etária correspondente à idade esperada de se cursar os ensinos fundamental e médio, 7 a 17 anos de idade, e portanto não estará considerando este contingente de jovens acima de 17 anos ainda na escola.

O ensino médio apresentou o maior nível de crescimento de matrículas na década de 90, mas o nível de matrícula ainda é

baixo relativamente à população na faixa etária correspondente. O custo de oportunidade de não trabalhar deve ser fator importante, principalmente levando-se em conta que os alunos concluem o 1º grau com idade superior ao esperado. Além disso, ainda existe deficiência na oferta de vagas: 15% dos municípios brasileiros não ofereciam o 2º grau completo até 1998, e apenas 2% dos alunos matriculados neste nível de ensino estavam localizados em áreas rurais no ano de 2001.

Os ensinos fundamental e médio são em sua maioria atendidos por escolas públicas - 91% e 87% em 2001, respectivamente. Desta forma, a análise dos dados educacionais refere-se praticamente aos resultados obtidos pela escola pública. É, contudo, interessante observar a grande diferença de atraso escolar entre estas dependências administrativas. As taxas de distorção idade-série, diferentemente das taxas de escolarização, incluem apenas os alunos matriculados, e portanto medem apenas a atraso de entrada e repetência dos alunos, não considerando jovens que nunca entraram ou que evadiram. A evasão irá antes melhorar este indicador, pois é provável que os alunos remanescentes apresentem melhores resultados educacionais que os que evadiram.

A tabela 4 apresenta as taxas de distorção idade-série e a idade mediana por dependência administrativa. 14.5% dos alunos da escola pública já estavam atrasados na 1ª série (em função de entrada tardia ou repetência), e 51.1% deles estarão atrasados no último ano do 2º grau. Estes números poderiam ser ainda piores, não fosse a evasão nos últimos anos de ensino, como comprova a melhora na taxa de distorção da 8ª série para o 3º colegial. Nas escolas privadas apenas 2.1% estão atrasados na 1ª série, e 12.9%, na última série. Nos dois casos, público e privado, as transições que levam a maior aumento de atraso são da 1ª para a 2ª série, e da 4ª para a 5ª série. Enquanto o aluno mediano da escola privada cursa cada série com a idade ideal, o aluno mediano da escola pública começa com 1 ano de atraso e termina com 2 anos de atraso em relação à idade ideal.

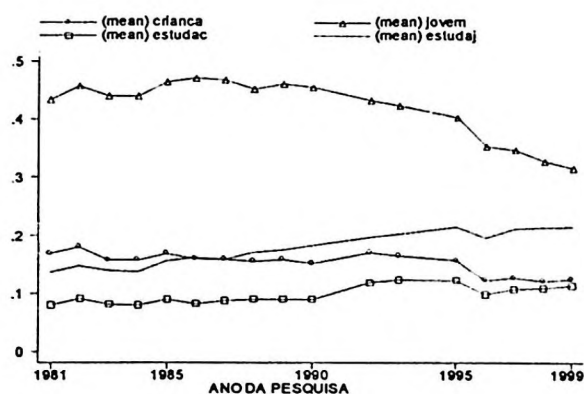
Tabela 4. Taxa de Distorção Idade-Série e Idade Mediana por série

Indicador	ano 2001	Séries		
		1 ^a	8 ^a	3 ^o (2 ^o grau)
Taxa distorção Idade-Série				
Total		13.9	50.8	45.5
Pública		14.5	54.3	51.1
Privada		2.1	9.6	12.9
Idade Mediana				
Total		7	15	19
Pública		8	16	19
Privada		7	14	17

Fonte. MEC/INEP (2001).

Um aspecto peculiar da educação no Brasil refere-se ao alto número de jovens que trabalham e estudam. Até 1998, jovens a partir de 14 anos podiam entrar no mercado de trabalho formal (atualmente a partir de 16 anos de idade); se por um lado a evasão a partir desta idade realmente aumenta, muitos continuam os estudos concomitantemente ao trabalho.

Gráfico 1. Trabalho infantil: proporção de crianças e jovens que trabalham e, destes, proporção que estuda



Nota. criança: trabalho na faixa etária de 10 a 13 anos; jovem: trabalho na faixa etária de 14 a 17 anos; estudac_i (i = c - criança ou j - jovem): proporção que estuda entre os que trabalham.

Fonte. a partir dos dados das PNADs de 1981 a 1999.

A proporção de crianças de 10 a 13 anos de idade e de jovens de 14 a 17 anos trabalhando é mostrada no gráfico 1. Mostra-se também a proporção destas crianças e jovens que, além de trabalhar, estuda. Praticamente todas as crianças que trabalham também estudam no final da década de 90; para os jovens, há uma tendência de queda de trabalho e de aumento de estudo. O trabalho junto com os estudos está refletido no alto número de alunos de ensino médio que freqüentam turnos noturnos, 51% em 2001.

A próxima seção apresenta a trajetória dos dois indicadores educacionais que serão utilizados no restante do trabalho: probabilidade de estar freqüentando escola e probabilidade de se estar na série ideal (série esperada de acordo com a idade). Como veremos, apesar dos resultados educacionais ainda estarem aquém do desejado, eles melhoraram muito no decorrer das décadas de 80 e de 90.

1.2. Freqüência e atraso escolar

Os indicadores de freqüência e de estar na série ideal são os indicadores escolhidos para serem investigados no restante do trabalho.

As informações foram calculadas a partir dos dados da Pesquisa Nacional por Amostra de Domicílios - PNAD² realizada pelo Instituto Brasileiro de Geografia e Estatística - IBGE, e referem-se ao período de 1981 a 1999. Esta pesquisa foi escolhida em função de sua abrangência temporal e geográfica, incluindo praticamente todo o país, com exceção apenas da zona rural da região Norte.

A freqüência escolar refere-se à freqüência em ensino fundamental ou médio, inclusive curso supletivo. Estar na série ideal acrescenta à informação de freqüência escolar se a criança entrou na escola com a idade correta e se continuou freqüentando

² A descrição dos dados da PNAD encontra-se no anexo A1.

sem repetir nenhum ano escolar, portanto refere-se à frequência da série esperada de acordo com a idade.

Os indicadores serão apresentados por coorte e idade, acompanhando crianças nascidas em um mesmo ano, e portanto com expectativa de entrar na escola no mesmo ano. O efeito de políticas educacionais, realizadas em dado ano e relacionadas a determinada série escolar, estará disseminado por várias coortes e idades.

A idade de entrada obrigatória na escola é 7 anos. Considera-se neste trabalho que crianças com 7 anos de idade completos até junho de cada ano deveriam entrar na escola naquele ano. As coortes, portanto, foram construídas incluindo nascidos de julho de determinado ano a junho do ano seguinte. No decorrer de todo o trabalho a idade utilizada refere-se ao mês de junho de cada ano.

A população etária que será investigada é a equivalente ao esperado para os ensinos fundamental e médio, de 7 a 17 anos. Como já há muitas crianças de 6 anos no ensino fundamental, sua informação também será apresentada. Desta forma apresenta-se informação a partir dos nascidos em 1964 (17 anos em 1981) até os nascidos em 1993 (6 anos em 1999).

As tabelas apresentam a coorte de nascimento nas linhas e idade nas colunas; as informações de um ano específico da pesquisa podem ser lidas em cada diagonal. Há duas formas, portanto, de analisar a evolução dos indicadores educacionais, acompanhando-se os dados a cada idade para uma dada geração (cada linha representa uma geração), e acompanhando-se a evolução dos indicadores para uma mesma idade entre gerações (cada coluna representa uma idade).

A tabela 5 apresenta a abertura por coorte e idade da proporção das crianças que frequentam escola³. Já existem jovens a partir de 15 anos que concluíram o ensino médio, e eles serão

³ O anexo A3 traz as tabelas das outras classificações, abertas por coorte e idade, que complementam a proporção que frequenta escola, basicamente 'nunca ter frequentado' e 'ter evadido'.

considerados junto aos que freqüentam, pois o fizeram até concluir.

Tabela 5. Proporção que freqüenta a escola (ou já concluiu)

coorte	idade											
	6	7	8	9	10	11	12	13	14	15	16	17
1964												43.540
1965											50.509	41.810
1966										57.523	47.877	43.101
1967									65.851	56.169	48.587	42.105
1968								73.909	64.796	57.913	49.723	42.701
1969							79.118	72.218	65.308	56.426	49.163	42.070
1970						82.584	78.834	74.727	66.660	57.798	49.684	42.113
1971				84.415	83.025	81.610	75.695	67.468	57.631	50.336	45.029	
1972			81.319	84.358	85.785	81.808	75.773	67.247	57.351	51.861	45.445	
1973		77.061	81.634	87.827	85.697	82.004	74.688	66.271	60.358	53.198	46.553	
1974		63.400	76.437	86.819	87.455	86.102	82.152	75.920	68.725	60.344	53.000	
1975	22.551	60.973	83.354	87.545	88.210	86.495	83.253	77.745	70.221	62.698		50.234
1976	18.031	73.991	84.191	87.547	88.467	87.249	83.094	77.277	71.248		57.279	52.053
1977	29.756	76.774	84.795	88.285	89.227	87.630	82.780	78.143		65.378	58.502	
1978	29.719	78.678	87.706	86.446	89.922	87.864	84.428		74.501	68.127		56.278
1979	29.051	80.451	82.349	87.381	89.874	88.006		80.900	76.605		63.725	59.257
1980	30.990	69.694	83.076	86.950	89.737		86.364	83.992		72.897	67.505	62.925
1981	21.307	71.502	81.172	87.172		89.350	90.511		80.338	75.860	71.514	66.583
1982	21.073	68.817	81.608		89.351	89.900		86.646	82.741	78.595	75.853	70.348
1983	21.934	69.060		87.327	90.551		90.739	88.728	85.768	81.468	77.283	
1984	20.192		82.249	88.572		92.099	92.063	91.022	88.284	84.232		
1985		70.506	81.939		91.606	92.851	94.378	92.989	90.095			
1986	22.886	70.689		89.973	92.343	94.390	94.769	93.367				
1987	23.377		85.297	90.794	94.062	95.843	95.997					
1988		75.442	85.767	92.630	95.975	97.312						
1989	25.902	76.935	89.407	95.397	97.029							
1990	26.935	78.761	91.694	95.822								
1991	26.014	82.774	94.034									
1992	30.018	85.164										
1993	35.033											

Nota. os percentuais em negrito incluem crianças freqüentando a pré-escola (por critério da PNAD não é possível distingui-las das crianças na 1ª série)⁴; as diagonais em branco referem-se aos anos em que não houve PNAD, 1991 e 1994.

O aumento da proporção de jovens na escola entre gerações ocorre em todas as idades. Nota-se também que o aumento de freqüência é pequeno durante a década de 80, sendo o aumento expressivo na década de 90. A partir da coorte de 1985, mais de 90% das crianças de 10 a 14 anos estão na escola.

A proporção de crianças entrando com 6 e 7 anos aumenta muito na última década, porém ainda é significativa a entrada até 9 anos de idade (medida pelo aumento de freqüência dos 7 aos 9 anos

da mesma geração). Este atraso na entrada, mesmo que diminuindo, ainda é importante: comparando-se a frequência com 7 e 9 anos de idade, calcula-se que 17% das crianças estão entrando na escola com 8 e 9 anos de idade, proporção que se mantém desde a coorte de 1980 até a de 1990.

A evasão, que pode ser medida pela diminuição de frequência a cada idade, apresenta tendência de ocorrer em idade mais avançada, após a idade de escola obrigatória de 14 anos.

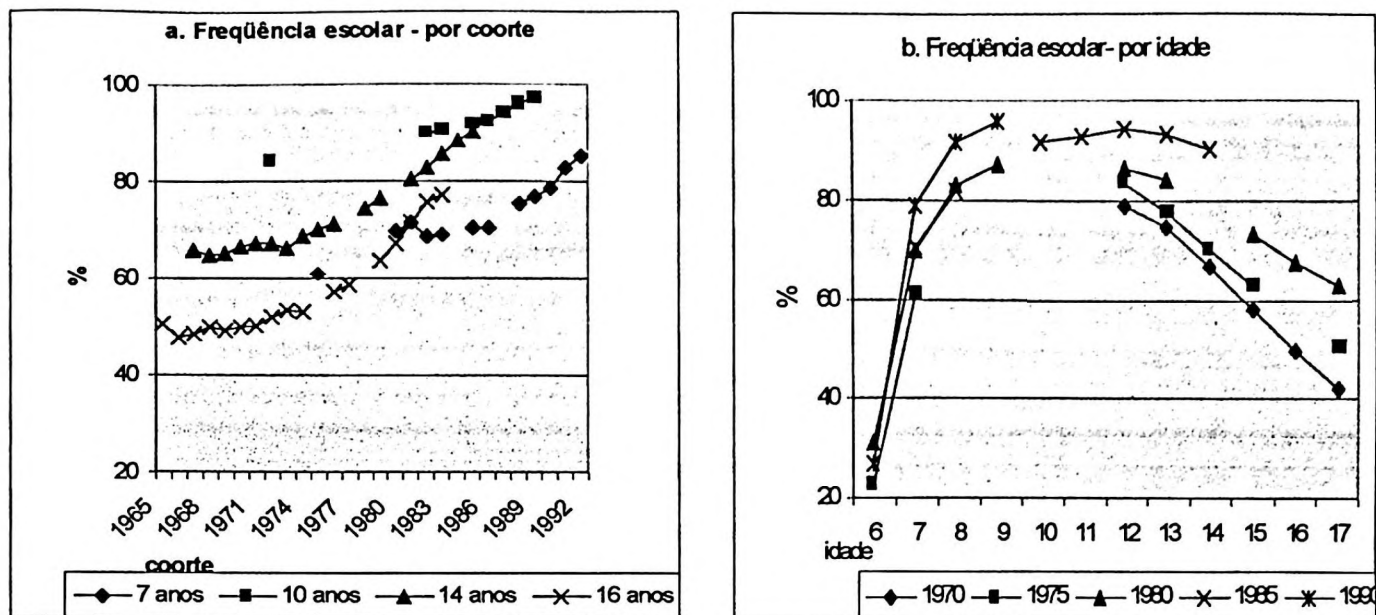
Os dados, portanto, mostram que grande parte das crianças frequenta a escola, principalmente até a idade de escola obrigatória, como confirmam os gráficos 2a e 2b. Eles reproduzem a informação da tabela 1: o primeiro mostra a evolução da frequência entre coortes para as idades 7, 10, 14 e 16 anos; e o segundo gráfico mostra a evolução a cada idade para as coortes de 1970, 1975, 1980, 1985 e 1990.

O gráfico 2a mostra a aceleração do aumento de frequência no final do período analisado, pois as curvas de frequência apresentam maior inclinação para as coortes mais jovens para todas as idades. As idades com menor frequência são a idade de entrada, 7 anos, e de escola não-obrigatória, 16 anos. O grupo de 10 anos de idade apresenta a maior proporção de crianças frequentando em relação a todas as coortes.

O gráfico 2b por idade confirma que a maior proporção de crianças na escola ocorre nas idades entre 9 e 12 anos. A melhora de frequência é contínua entre coortes para todas as idades (curvas mais altas para coortes mais recentes), mas é interessante observar que da coorte de 1980 para a de 1985 a melhora ocorreu apenas para as crianças com mais idade, enquanto a coorte de 1990 já apresenta grande melhora na idade de entrada em relação às coortes anteriores. Apesar do pico de frequência ser aos 12 anos de idade, a coorte de 1985 já mostra uma acentuada queda na evasão para as idades de 13 e 14 anos.

⁴ Explicação dos critérios da PNAD no anexo A1.

Gráficos 2. Freqüência escolar



Nota. os pontos faltantes das curvas referem-se aos anos em que não houve PNAD, 1991 e 1994, e aos anos em que não foi possível separar as crianças em pré-escola das crianças na 1ª série do ensino fundamental (dados em negrito na tabela 5).

A freqüência à escola proporciona uma série de vantagens à criança, não necessariamente associadas ao aprendizado. A merenda escolar é uma delas, principalmente para crianças oriundas de famílias pobres. A escola também promove desde cedo a socialização da criança, seu contato com outras crianças e professores é a primeira experiência para a vida em comunidade e para a cidadania. Além disso, estando na escola, a criança não está nas ruas, podendo estar mais protegida da criminalidade.

Estes fatos por si só já são suficientes para comprovar a importância da escola. O objetivo da escola, contudo, vai além do assistencialismo e da socialização, por isso é interessante analisar também medidas de desempenho escolar. Este trabalho não pretende analisar o sistema educacional em si, mas é possível olhar medidas quantitativas que tragam alguma informação além da freqüência escolar.

Já vimos na seção anterior que o atraso escolar é grande no Brasil, ele é em parte função da entrada tardia na escola, em parte função da evasão, e, dentro da escola, o atraso ocorre devido à repetência. O atraso escolar advindo da repetência não só onera o sistema escolar em termos de custos, pois os alunos ficarão mais anos na escola para adquirir o mesmo número de anos de estudo completos, como pode incentivar a evasão de alunos desestimulados e relativamente velhos para a série cursada.

A segunda medida investigada é uma medida de atraso escolar. A série ideal é definida como a série em que a criança deveria estar caso entrasse na escola na idade obrigatória, 7 anos, e não repetisse nenhum ano. Estar na série ideal é portanto resultado de várias ações: ter em primeiro lugar entrado na escola, tê-lo feito aos 7 ou menos anos de idade, ter continuado freqüentando a escola, e não ter repetido nenhuma série. No caso dos que entraram adiantados, considerarei também a série ideal de acordo com a entrada aos 7 anos de idade, supondo portanto de menor importância a repetência no caso de quem estava adiantado. Estão também incluídos na série ideal os que já concluíram a escola.

Mesmo que a taxa de repetência fique constante no tempo, a diminuição do atraso na idade de entrada melhoraria todos os indicadores para as idades futuras de uma dada geração. A evasão contribui para a piora do indicador de série ideal. Para os mais velhos fica difícil distinguir qual o fator mais importante para a diminuição de crianças na série ideal, a repetência ou a evasão.

A tabela 6 apresenta as proporções por coorte e idade de crianças na série ideal. Há uma melhora da proporção na série ideal entre coortes em todas as idades e a melhora é contínua no tempo e acelera-se na década de 90. Contudo apenas 22% dos jovens de 17 anos estavam na série ideal em 1999 e cerca de 15% das crianças de 7 anos ainda não haviam entrado na escola no mesmo ano.

Tabela 6. Proporção na série ideal

coorte	idade										
	7	8	9	10	11	12	13	14	15	16	17
1964											8.624
1965										8.682	8.561
1966									10.086	8.269	9.346
1967								11.643	9.950	8.997	8.897
1968							13.745	11.368	10.534	9.723	9.029
1969						15.769	13.672	12.460	9.675	8.876	8.699
1970					18.919	16.098	14.522	12.716	11.263	10.089	9.769
1971				23.775	19.965	17.678	15.719	13.858	12.303	10.846	9.494
1972			29.123	24.080	22.982	18.159	15.961	14.125	12.679	10.126	10.958
1973		35.677	29.529	27.640	23.048	18.417	16.334	13.959	12.160	11.321	9.872
1974	63.400	36.590	32.407	27.373	22.666	19.116	15.869	15.928	12.350	10.026	
1975	60.973	39.340	33.191	29.292	23.807	20.080	17.911	16.195	12.785		11.568
1976	73.991	40.758	35.163	30.276	26.967	21.045	19.057	16.677		12.404	11.448
1977	76.774	45.753	37.739	30.835	27.568	22.395	18.810		15.697	13.271	
1978	78.678	47.230	37.382	33.078	28.328	22.645		18.759	16.354		13.858
1979	80.451	47.521	38.876	34.598	29.200		21.692	18.085		15.882	16.049
1980	69.694	49.672	40.651	34.506		25.721	23.068		19.410	18.090	16.448
1981	71.502	49.627	40.800		33.791	28.384		23.133	21.590	20.015	19.970
1982	68.817	51.033		38.586	33.741		26.993	25.288	22.869	23.270	22.210
1983	69.060		44.896	39.164		31.950	30.229	27.387	25.440	24.424	
1984		54.831	46.652		37.377	34.353	30.372	30.015	28.316		
1985	70.506	53.540		42.645	40.452	36.312	33.781	32.842			
1986	70.689		49.430	44.850	41.543	36.616	36.353				
1987		58.741	52.437	46.580	45.645	42.471					
1988	75.442	60.607	54.820	52.725	49.907						
1989	76.935	63.921	60.414	55.980							
1990	78.761	67.649	62.175								
1991	82.774	72.724									
1992	85.164										

Nota. os percentuais em negrito incluem crianças em pré-escola; as diagonais em branco referem-se aos anos em que não houve PNAD, 1991 e 1994.

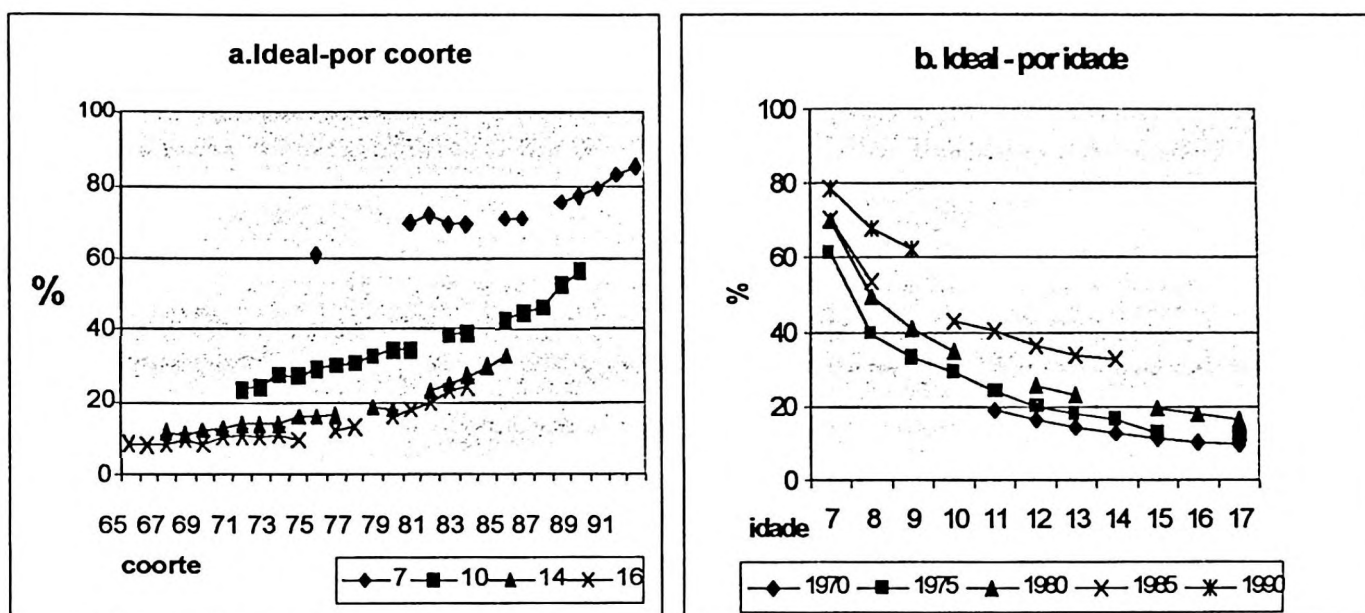
O gráfico 3a mostra a tendência de melhora na proporção na série ideal entre coortes para as idades apresentadas. É surpreendente o aumento de crianças na série ideal aos 10 anos de idade, principalmente nos últimos 2 anos analisados, 1998 e 1999.

Como se pode melhor observar no gráfico 3b, da mesma forma que para a frequência, a melhora de estar na série ideal da coorte de 1980 para a de 1985 é maior para as crianças mais velhas. Já os nascidos em 1990 entram mais cedo na escola e sua taxa de repetência na 1ª série do ensino fundamental diminui: cai a inclinação da curva entre coortes na passagem dos 7 para os 8 anos de idade.

A medida 'ideal' esconde quão atrasadas estão as crianças em relação à série ideal. Se a maioria estivesse atrasada apenas um ano, talvez isto não fosse uma questão tão relevante. Condicional

em ainda estar freqüentando, grande parte do atraso é de mais de um ano, o que mostra a gravidade desta situação. Mais de 50% dos jovens de 16 e 17 anos cursando a escola estão mais de 1 ano atrasados. Como os outros indicadores, este indicador de atraso também melhora no tempo: por exemplo, aos 10 anos de idade, 45% da coorte de 1971 estava mais de 1 ano atrasado, mas apenas 20% estavam atrasados na mesma idade para a coorte de 1989⁵.

Gráficos 3. Estar na série ideal



Nota. os pontos faltantes das curvas referem-se aos anos em que não houve PNAD, 1991 e 1994, e aos anos em que não foi possível separar as crianças em pré-escola das crianças na 1ª série do ensino fundamental (dados em negrito na tabela 6).

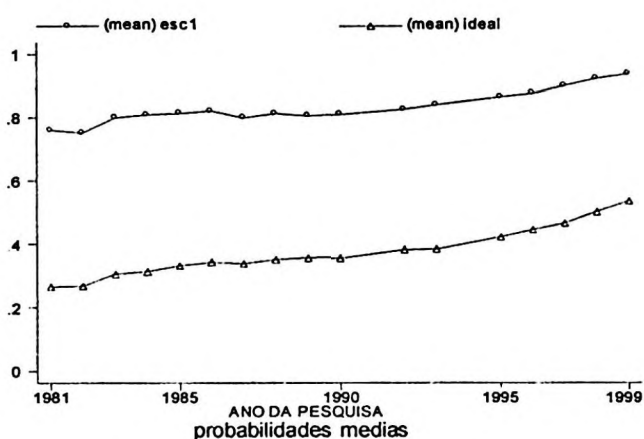
Os dois indicadores acima apresentados, freqüência escolar e estar na série ideal, serão os indicadores usados nos capítulos que se seguem para as análises dos determinantes da educação e do efeito da renda familiar. Antes de analisar os determinantes dos indicadores educacionais, contudo, vale a pena apresentá-los também abertos por ano e por região de residência.

⁵ Ver anexo A3, tabela A6.

Abaixo são apresentadas as proporções de frequência escolar e série ideal no tempo apenas para as crianças de 7 a 14 anos de idade, pois esta faixa etária é o foco de atenção no restante do trabalho.

O gráfico 4 mostra a média dos indicadores educacionais no tempo. Confirma-se a melhora dos indicadores já vista na abertura por coortes.

Gráfico 4. Proporção de frequência e série ideal para crianças de 7 a 14 anos



Nota. Os anos de 1981 e 1983 a 1986 incluem crianças na pré-escola.

É interessante também observar os indicadores para as crianças de 7 a 14 anos entre regiões. Como mostra a tabela 7, a região NE, que apresenta a menor proporção de crianças na escola, é também a região que apresenta o maior crescimento da frequência escolar no tempo, chegando a valores bem próximos das outras regiões em 1999. A tabela 8 apresenta a mesma abertura para estar na série ideal, mostrando também a pior performance da região Nordeste. As regiões S e SE apresentam o melhor resultado durante todo o período, e praticamente mantêm a mesma diferença em relação às regiões N e NE.

Tabela 7. Proporção de frequência escolar entre regiões e anos

ano/região	N	NE	SE	S	CO	média
1981	0.833	0.656	0.829	0.785	0.767	0.760
1982	0.825	0.635	0.831	0.787	0.759	0.753
1983	0.884	0.727	0.850	0.800	0.816	0.799
1984	0.882	0.740	0.858	0.808	0.817	0.808
1985	0.887	0.745	0.862	0.817	0.822	0.814
1986	0.878	0.762	0.864	0.823	0.829	0.821
1987	0.859	0.719	0.852	0.813	0.837	0.800
1988	0.870	0.722	0.865	0.843	0.843	0.813
1989	0.852	0.703	0.869	0.844	0.826	0.806
1990	0.853	0.710	0.873	0.857	0.836	0.812
1992	0.859	0.716	0.892	0.877	0.856	0.828
1993	0.861	0.739	0.905	0.890	0.859	0.842
1995	0.881	0.775	0.922	0.907	0.885	0.866
1996	0.884	0.794	0.925	0.929	0.905	0.878
1997	0.880	0.834	0.943	0.942	0.914	0.901
1998	0.915	0.877	0.949	0.951	0.936	0.923
1999	0.933	0.903	0.955	0.961	0.944	0.937
Média	0.875	0.752	0.886	0.862	0.854	0.835

Nota. 7 a 14 anos de idade.

Tabela 8. Proporção na série ideal entre regiões e anos

ano/região	N	NE	SE	S	CO	média
1981	0.217	0.155	0.338	0.340	0.237	0.266
1982	0.222	0.143	0.345	0.357	0.247	0.268
1983	0.256	0.182	0.386	0.387	0.287	0.306
1984	0.273	0.188	0.393	0.400	0.311	0.315
1985	0.293	0.196	0.418	0.429	0.311	0.333
1986	0.315	0.205	0.427	0.444	0.340	0.344
1987	0.288	0.185	0.430	0.461	0.331	0.339
1988	0.286	0.189	0.445	0.481	0.343	0.352
1989	0.301	0.187	0.451	0.502	0.351	0.357
1990	0.307	0.189	0.457	0.507	0.368	0.361
1992	0.318	0.201	0.489	0.548	0.378	0.387
1993	0.308	0.210	0.488	0.560	0.397	0.390
1995	0.351	0.239	0.540	0.591	0.427	0.428
1996	0.357	0.251	0.570	0.615	0.466	0.450
1997	0.348	0.263	0.598	0.628	0.491	0.470
1998	0.395	0.304	0.627	0.662	0.528	0.506
1999	0.434	0.340	0.658	0.701	0.558	0.540
Média	0.320	0.215	0.478	0.508	0.382	0.381

Nota. 7 a 14 anos de idade.

No próximo capítulo será feita a revisão da literatura que tenta explicar os resultados educacionais.

capítulo 2: REVISÃO DA LITERATURA

2.1. Trabalhos teóricos

Os primeiros modelos que analisam as decisões relativas à educação, como BECKER (1967), concentram-se nas decisões das famílias. A família é vista como uma unidade de produção que maximiza sua utilidade; a unidade familiar decide a oferta de trabalho de seus membros e como usarão seus recursos. A utilidade é função do consumo atual da família e do consumo futuro dos filhos, portanto considera-se certo altruísmo dos pais. O investimento em capital humano dos filhos trará maior renda para a geração seguinte. Se existe mercado de crédito, ou melhor, se os pais podem tomar emprestado para investir em seus filhos usando a renda futura dos filhos como colateral, a decisão de investimento em capital humano dos filhos não necessariamente afeta o consumo atual, e portanto não precisa depender da renda dos pais. Alguns modelos tomam como dadas as características e decisões das famílias e maximizam apenas a utilidade do próprio indivíduo em função de sua educação. A educação em geral é considerada como bem de investimento puro, desprezando seu aspecto como bem de consumo.

GRILICHES (1979) apresenta um modelo de maximização de riqueza individual, baseado na decisão sobre educação, para discutir alguns problemas econométricos relacionados à estimação da taxa de retorno da educação. Estes problemas estão basicamente relacionados à omissão da habilidade como variável de controle numa estimativa relacionando renda do trabalho e educação. Ele discute as possíveis direções de um viés gerado na estimação do retorno da educação, caso a habilidade ou a endogeneidade da educação não sejam consideradas.

CARD (2001) estende a análise de GRILICHES (1979), apresentando um modelo de capital humano que especifica explicitamente formas funcionais para os benefício e custo marginais do investimento em capital humano. Desta forma, ele insere explicitamente no modelo a heterogeneidade individual, que levaria a diferentes decisões relativas à educação. A heterogeneidade dos indivíduos pode estar relacionada a diferentes fatores: habilidades inatas da criança ou da família, que podem afetar tanto o custo (de aprendizado) como o benefício da educação; custos da educação, como a distância escola-casa; contatos sociais da família que podem ser importantes para a inserção da criança no mercado de trabalho, entre outros.

Nos dois casos os autores supõem mercados de capitais perfeitos, sendo possível portanto tomar emprestado no presente contra pagamento futuro. É difícil, porém, acreditar em mercado perfeito de capitais, principalmente quando o empréstimo seria feito contra o capital humano futuro (e incerto) da criança. Se não houver mercado de crédito, os recursos familiares serão a fonte de recurso para o investimento em educação. A estes recursos somar-se-iam gastos do governo, que podem aliviar o custo para as famílias. De qualquer forma, famílias de menor renda podem ter uma restrição orçamentária que as impeça de educar seus filhos, mesmo que a educação seja vantajosa em termos de benefícios futuros. Neste caso, portanto, importa a renda familiar para a educação dos filhos.

BECKER e TOMES (1986), interessados na desigualdade entre gerações, apresentam um modelo de decisão da família que considera a hipótese de inexistência de mercado de crédito. A utilidade da família inclui seu consumo atual e o consumo futuro dos filhos, e sua decisão refere-se à alocação de recursos para atingir seu objetivo de maximização de utilidade. No caso com mercado de capitais perfeito, o capital humano dos filhos depende apenas do gasto de seus pais com sua educação e de sua dotação inicial. Como dotação os autores entendem habilidades inatas e

características familiares que não dependem de investimento para serem transmitidas. A maximização da utilidade ocorre quando custo e receita marginais se interceptam: o custo equivale à taxa de juros de mercado e é constante para todos, e a receita (ou taxa de retorno do investimento em capital humano) é função do nível de gastos em educação feito pelos pais e da dotação inicial de seus filhos. O nível de gastos com educação e a dotação inicial afetam positivamente a acumulação de capital humano.

No caso de não haver acesso ao mercado de capitais, a decisão de gastos com educação dos pais passa a depender também da renda dos pais e de seu grau de altruísmo. Isto ocorre porque, dada a restrição de crédito, para investir em seus filhos os pais terão que alterar seu próprio consumo presente. O custo do investimento deixa de ser constante e passa a ser crescente em relação ao nível de gastos dos pais, pois a taxa de desconto (subjéctiva) dos pais aumenta quanto mais eles deixam de consumir em favor de seus filhos. Além disso, o efeito da dotação inicial dos filhos passa a ter efeito ambíguo sobre o investimento em capital humano.

ACEMOGLU e PISCHKE (2001) baseiam-se no modelo de BECKER e TOMES (1986) para chegar a uma forma funcional simples para estimação do efeito da renda dos pais sobre decisões de estudo de seus filhos. Consideram que não existe mercado de crédito. A renda familiar importa para as famílias que tiverem renda abaixo de determinado nível, nível a partir do qual o investimento óptimo em capital humano dos filhos passa a afetar o próprio consumo da família.

Consideram dois períodos: no primeiro, a renda familiar é usada para consumir e investir na educação dos filhos, e no segundo, o filho consome de acordo com a renda gerada pelo capital humano adquirido. Seja a utilidade da família dada por:

$$\ln c_t + b \ln c_{t+1} \quad (1)$$

investimento em educação. O investimento em educação trará maior utilidade se

$$A_i \leq \ln[y_i ((w_h - w)/w_h)],$$

como se vê, para famílias ricas, a decisão depende apenas do custo da educação, medido por A , e de seu retorno, medido pelo ganho de renda em função da educação, $w_h - w$. Para famílias pobres, a decisão depende adicionalmente da renda familiar, y_i .

É possível, portanto, que estimativas da probabilidade de frequência escolar apresentem diferentes coeficientes para a renda familiar, dependendo do nível de renda. O coeficiente seria significativo para os mais pobres, e não significativo para mais ricos. Na próxima seção comentarei os resultados empíricos encontrados por ACEMOGLU e PISCHKE (2001).

A literatura de uma forma geral considera a renda permanente para a análise das decisões relacionadas à educação. É possível, porém, que choques transitórios na renda também afetem estas decisões. Choques negativos na renda dos pais podem obrigar o filho a trabalhar, levando a um menor tempo de dedicação à escola e eventualmente à evasão. A decisão de anos de estudo, portanto, mesmo que seja a princípio baseada na renda permanente, é reavaliada a cada momento, ou pelo menos a cada início de ano letivo.

2.2. Trabalhos empíricos

A análise empírica dos resultados educacionais das crianças pode ser considerada como parte de uma literatura mais ampla, que procura determinar o grau de mobilidade intergeracional. Os trabalhos de mobilidade comparam em geral renda ou ocupação de pais e filhos. Um dos fatores mais importantes para gerar oportunidades iguais no mercado de trabalho, e portanto capaz de aumentar a mobilidade social entre gerações, é a educação. Desta forma, a educação é uma variável-chave para proporcionar

mobilidade de renda entre gerações. Em vez de comparar diretamente a renda de pais e filhos, mesmo porque estes dados não estão tão freqüentemente disponíveis como dados de seção transversal, pode-se analisar a relação entre características dos pais e o resultado educacional dos filhos. Este indicador educacional estará sinalizando o rendimento futuro do filho e portanto o grau de mobilidade possível.

HAVEMAN e WOLFE (1995) resumem alguns trabalhos sobre mobilidade social: estudos utilizando dados de painel para os EUA encontram correlação significativa entre renda do pai e renda do filho, o que justifica o interesse de outros estudos em determinar por que via esta correlação se dá, estendendo o número de variáveis sociais e familiares para explicar o resultado do filho.

BLANDEN *et alli* (2001) mostram que a mobilidade intergeracional diminuiu na Grã-Bretanha, comparando-se nascidos em 1958 com nascidos em 1970. Aumenta a elasticidade da renda dos pais para explicar renda dos filhos. Esta mudança de elasticidade é explicada em parte pelo fato de o aumento de resultado educacional no tempo ter sido maior para os mais ricos. Os autores comentam que este resultado difere de resultados encontrados para os EUA, em que a mobilidade parece aumentar no tempo, principalmente em função de uma relação mais fraca entre educação e renda dos pais.

SHEA (2000) analisa o efeito da renda dos pais sobre a renda dos filhos para os EUA. Confirmando a existência de uma correlação positiva entre renda do filho e renda do pai, o autor procura separar o efeito da renda de variáveis correlacionadas. Seu resultado geral é que não existe efeito da renda, porém para uma subamostra de pais com baixa educação, importaria a renda do pai para a renda futura do filho.

Outro bloco da literatura, em vez de olhar diretamente a relação entre renda dos pais e renda dos filhos, concentra-se ou

na decisão de educação ou na relação da educação com retorno futuro no mercado de trabalho. As pesquisas mais relacionadas com este trabalho analisam quais fatores familiares, individuais e de política pública influenciam o resultado escolar. Há também uma extensa pesquisa relativa à relação entre educação e renda posterior no mercado de trabalho: a estimação do retorno da educação é feita a partir de uma equação de Mincer¹ (basicamente, logaritmo da renda em função da educação e experiência no trabalho), que representa o resultado de maximização do investimento em capital humano. Este conjunto de trabalhos, portanto, analisa duas etapas relacionadas à mobilidade social: a primeira observa a aquisição de educação, fator relevante para gerar oportunidades iguais; a segunda mede como a educação será avaliada no mercado de trabalho.

O restante desta revisão será organizado da seguinte forma: primeiro apresentarei resultados de pesquisas que estavam interessadas em analisar os determinantes dos resultados educacionais de forma geral; depois me concentrarei em trabalhos interessados especificamente no efeito da renda familiar, dado que este é também o ponto principal de interesse desta pesquisa.

2.2.1. Os determinantes dos resultados educacionais

Primeiro é importante notar que os resultados educacionais estudados variam muito: incluem desde a frequência a escola, evasão e repetência, até testes de conhecimento e decisões de estudo após a escola obrigatória. Os fatores que influenciam os resultados educacionais podem ser divididos em dois grandes blocos: políticas públicas determinadas pelo governo e decisões e características da família e da criança ou jovem.

Existe um corpo de trabalhos interessado especificamente no papel da escola (ou insumos escolares) sobre os resultados escolares do aluno. Estes trabalhos restringem sua análise às

¹ Assim nomeada a partir da análise do autor em MINCER (1974).

crianças que freqüentam a escola e tentam determinar se fatores como a qualificação dos docentes ou a razão aluno-professor, que são em geral determinados pelo grau de investimento público, influenciam o desempenho do aluno.

Os primeiros trabalhos mostravam não haver uma forte influência da escola sobre os resultados de seus alunos, portanto concluem que estes resultados dependem prioritariamente das características demográficas, familiares e individuais. HANUSHEK (1986) resume vários estudos a respeito da eficiência das escolas públicas sobre os resultados educacionais nos EUA e conclui que parece não haver relação sistemática entre a melhora dos indicadores relacionados com a qualidade da escola, entre eles, aumento do gasto público e diminuição da razão aluno-professor, e os resultados educacionais (testes, freqüência escolar, continuar após idade obrigatória, evasão).

CARD e KRUEGER (1996) fazem uma compilação de trabalhos mais recentes, e mostram que a maioria destes trabalhos encontra uma relação positiva e significativa entre a qualidade da escola e a freqüência escolar e a renda posterior. Em seu estudo, comparando os Estados norte-americanos da Carolina do Norte e do Sul, CARD e KRUEGER também encontram que o aumento relativo de recursos escolares e diminuição da razão aluno-professor para negros no Sul aumentaram a freqüência escolar e renda posterior deste grupo.

Outro corpo de trabalhos concentra-se em determinar quais as características familiares e pessoais que influenciam os resultados escolares. Estão mais interessados em saber o efeito conjunto destas variáveis, e as possíveis inter-relações entre elas, do que em isolar o efeito de uma variável específica.

HAVEMAN e WOLFE (1995) comentam trabalhos que analisam diversos indicadores educacionais para os EUA. Entre as características dos pais, destaca-se seu capital humano, em geral medido pelo número de anos completos de educação: está incluído

em todos os estudos, e é sempre significativo e com efeito não desprezível; além disso, em geral, o capital humano da mãe é o mais relevante.

O nível de renda da família também é em geral considerado. A renda é na quase totalidade dos estudos positivamente relacionada com o resultado educacional, e é estatisticamente significativa na maioria dos estudos. O efeito da renda, porém, é pequeno: ele é maior se a medida de renda refere-se a um período mais longo de tempo ou se poucas variáveis explicativas são incluídas, o que sugere estar captando efeito de outras variáveis quando há poucos controles. As elasticidades estimadas variam muito, porém a definição de renda também varia. Cabe observar que a educação dos pais é muitas vezes interpretada como uma medida de renda permanente, sua inclusão portanto capta parte do efeito da renda. Existe alguma evidência de que importa a fonte de renda: a renda do salário tem efeito positivo, porém a renda advinda de programas de bem-estar tende a ter menor impacto sobre a educação, em alguns casos inclusive negativo.

Outras variáveis em geral estudadas são: viver com apenas um dos pais, em geral com efeito negativo e significativo; o efeito da mãe trabalhar varia muito entre estudos; número de irmãos e presença de livros em casa têm efeito significativo e quantitativamente importante. A maioria dos estudos não encontra efeito significativo para raça, depois de se controlar renda e *background* familiar (em geral medido por educação ou ocupação dos pais). O efeito da vizinhança é positivamente correlacionado com educação, mas é pouco significativo.

A análise destas pesquisas deve levar em conta que elas desconsideram a endogeneidade das variáveis explicativas, o que afeta a possibilidade de se poder afirmar que as correlações com o resultado educacional obtidas são relações causais. Por exemplo, a oferta de trabalho da mãe pode não ser exógena; outras variáveis, como renda familiar e educação dos pais, podem estar relacionadas a habilidades não-observadas.

Existem vários estudos para o caso brasileiro. A maioria destes estudos parece estar mais interessada em determinar quais os fatores que influenciam o resultado educacional, e não no efeito específico de cada um dos fatores. Estes trabalhos enfatizam as diferenças entre subgrupos, por exemplo, meninos e meninas, diferentes regiões, e diferentes grupos etários.

Antes de apresentar alguns destes estudos, comentarei alguns trabalhos que analisam a educação na América Latina.

MENEZES *et alli* (2000) analisam as decisões de estudo e trabalho dos jovens na América Latina e Caribe na década de 90. Com microdados de 18 países, utilizam variáveis micro- e macroeconômicas para explicar estas decisões, e mostram que as variáveis microeconômicas são suficientes para explicar grande parte das decisões. Entre elas, a educação dos pais e o número de filhos são as mais importantes, a renda familiar têm efeito significativo, mas de menor magnitude.

Em relação ao atraso escolar, SCHIEFELBEIN e WOLFF (1993) analisam a evolução da repetência no ensino primário na América Latina. Ela é alta, principalmente na 1ª série do 1º grau, mas apresenta queda de 1980 a 1989. Segundo os autores, os grupos que mais repetem são: alunos da 1ª série do primário; de famílias mais pobres; vivendo na zona rural; crianças com uma segunda língua (em geral, língua indígena materna); as meninas repetem menos, mas a diferença vem diminuindo.

BARROS e LAM (1996) comparam os resultados educacionais do Estado de São Paulo e da região Nordeste de jovens de 14 anos em 1982 a partir dos dados da PNAD. Mostram que apesar da frequência escolar ser similar nas duas regiões, a média de anos completos de educação em São Paulo é bem maior. Os autores analisam a importância da educação e da renda dos pais para explicar tal diferença, e encontram, que apesar de terem efeito significativo, estes fatores familiares explicam menos de 20% da diferença de educação entre regiões. Eles concluem que é necessário uma

melhora na oferta de ensino para diminuir a desigualdade entre regiões, pois as características familiares explicam muito pouco desta desigualdade.

BARROS *et alii* (2001) analisam para o Brasil os determinantes do desempenho educacional, medido como o número de anos completos de estudo, com base em 2 bancos de dados, PNAD e PPV - Pesquisa sobre Padrões de Vida. Os autores restringem sua análise à parte urbana das regiões SE e NE, no ano de 1996, para jovens de 11 a 25 anos. Eles dividem os determinantes da educação em: disponibilidade e qualidade de serviços educacionais; retorno no mercado de trabalho local; recursos familiares; recursos da comunidade; características pessoais; e região de residência.

A qualidade da educação foi medida pela média de escolaridade dos professores de cada comunidade e por indicadores de infra-estrutura e qualidade da escola. A oferta de escola, medida pela razão entre o número de escolas e a população em idade escolar, tem efeito positivo e significativo, mas a qualificação do professor aparece com efeito não conclusivo: em relação a professores de 1º grau o efeito é positivo, mas para o 2º grau o efeito é negativo. O custo de oportunidade foi estimado a partir do salário médio de indivíduos de mesmas características e da probabilidade de emprego; seu efeito é sempre negativo sobre a escolaridade. Características pessoais e da comunidade aparecem com os sinais esperados. A educação dos pais é o mais importante determinante da escolaridade. O impacto da renda domiciliar *per capita* é positivo e significativo.

Os autores analisam a possibilidade dos efeitos serem distintos para diferentes grupos. Eles refazem as estimativas em separado para algumas subamostras, de acordo com idade, sexo, e região. O efeito da renda é semelhante para ambos os sexos e em ambas as regiões, porém cresce com a idade do jovem. Este resultado pode estar relacionado a dois efeitos distintos: o efeito coorte, tal que gerações mais velhas dependiam mais da renda familiar para estudar; e em segundo lugar, o efeito do

custo de oportunidade de se trabalhar: para crianças de 11 anos, o custo deve ser baixo se comparado a jovens de 25 anos. O efeito da educação dos pais é distinto para todas as subamostras: maior no Nordeste; maior para homens; e maior para os mais velhos. Além disso, a educação da mãe é mais importante para a escolaridade da filha, e a do pai, para a do filho.

KASSOUF (2000) utiliza dados da PNAD de 1995 para analisar as decisões de trabalho e educação no Brasil para crianças de 7 a 14 anos morando com os dois pais. Vou concentrar meus comentários nas estimativas referentes à frequência escolar. Suas estimativas são feitas em separado para as subamostras homem/ mulher e rural/ urbano; seus resultados também mostram diferentes dinâmicas de decisão para meninos e meninas.

Dada a colinearidade entre renda familiar e educação dos pais, KASSOUF prefere analisar estimativas usando ou uma ou outra variável explicativa, pois suas estimativas com os dois conjuntos de variáveis tornam-se imprecisas. O salário do pai assim como a educação da mãe são sempre significativos, independentemente de especificação ou subamostra. O salário da mãe torna-se significativo se são excluídas as variáveis educação dos pais e idade; a educação do pai torna-se significativa se excluída a renda familiar. A renda não-salarial nunca é significativa.

Em relação à diferenciação entre meninos e meninas, o efeito da renda de ambos os pais é maior para meninos. A educação da mãe tem maior efeito que a do pai, e é maior para meninos. O efeito da educação dos pais é maior para a zona rural.

GOMES-NETO e HANUSHEK (1994) analisam a repetência baseados nos dados do projeto EDURURAL do Banco Mundial para o Nordeste Rural. Eles analisam os fatores que explicam a repetência no NE rural utilizando um painel com 3 anos de informação, 1981, 1983 e 1985. A probabilidade de repetência é estimada incluindo como variável explicativa os resultados de testes. As notas dos alunos importam para explicar a repetência, mas as características

familiares (sexo, idade e educação dos pais) não são significativas, um resultado surpreendente se comparado a outros estudos. Este resultado, porém, pode ser função de uma alta colinearidade entre as notas obtidas e estas características.

Os autores também concluem que a não existência de oferta de séries superiores à 2ª série do ensino fundamental na escola aumenta a probabilidade de repetência. Em geral, considera-se que a oferta de ensino fundamental é suficiente para a demanda no Brasil, mas pode ainda ser um problema em núcleos regionais mais pobres. Como visto acima, BARROS *et alli* (2000) acharam efeito significativo da quantidade de escolas existentes sobre a educação dos jovens.

O capítulo 3 apresentará estimativas das determinantes da educação no Brasil, especificamente, estimará o efeito das características familiares e individuais sobre frequência escolar e estar na série ideal.

2.2.2. Renda familiar e resultados educacionais

Os estudos acima comentados estavam interessados em determinar os diversos fatores relacionados com os resultados educacionais. Em sua maioria consideram a renda familiar entre as variáveis explicativas, mas não estão interessados especificamente em seu efeito. Em geral encontrou-se que a renda tem efeito positivo e em vários casos significativo sobre os resultados escolares estudados.

O capítulo 5 irá analisar de forma específica o efeito da renda familiar sobre os dois resultados escolares: frequência e estar na série ideal.

Para analisar o papel da renda sobre resultados escolares, pode-se medir o efeito 'bruto' da renda. Isto é, medir o efeito da renda em si somado ao efeito dos fatores que a renda pode estar representando, isto é, fatores correlacionados com a renda.

Muitos trabalhos estimam apenas o efeito 'bruto' da renda. É uma análise válida, ela comprova ou não que famílias com maior renda possuem filhos com melhor resultado educacional. Para a definição de políticas públicas, porém, seria importante definir por que canal a renda importa. Se a renda importa porque pais mais ricos têm melhor condição de ajudar seus filhos na escola, uma política que crie aulas de reforço para alunos com dificuldades pode ser interessante, por outro lado, dar dinheiro para estas famílias não vai melhorar sua condição de ajudar os filhos com lição de casa. Por outro lado, se a questão principal captada pela renda refere-se ao custo direto da escola, subsídios para famílias mais pobres podem ser uma alternativa. No caso do Brasil, onde a maioria das crianças está em escola pública, o custo de oportunidade pode ser o custo mais relevante. Neste caso, dar dinheiro para as famílias pode ter efeito direto sobre suas decisões em relação a educação e trabalho dos filhos.

Trabalhos que estimam o efeito 'bruto' da renda acham em sua maioria correlações positivas e significativas. BEHRMAN e KNOWLES (1999) compilam estudos para diversos países e a maioria apresenta associação significativa entre renda e escola. As estimativas mais altas referem-se a regiões de baixa renda. A maioria dos estudos tem como variável dependente o número de anos completos de escola.

Os autores, então, apresentam uma análise para o Vietnã, enfocando vários indicadores de educação: idade de início da escola, séries completas por ano de escola (repetência), última série completada e teste de aptidão. Usam dados de 1996 para crianças de 6 a 17 anos. As médias dos indicadores são apresentadas por quintil de renda e são significativamente diferentes entre si; estimativas das elasticidades da renda confirmam sua importância para explicar os indicadores educacionais. Crianças mais pobres entram na escola mais velhas, repetem mais, completam menos anos de escola e têm pior resultado em testes.

Como as estimativas do efeito da renda não incluem controles, elas são afetadas pelo efeito de variáveis relacionadas tanto com o indicador escolar como também com a renda, como por exemplo, educação dos pais e preferências.

Outros estudos, mais próximos ao apresentado no capítulo 5, procuram estimar o efeito 'líquido' da renda sobre a educação, tentando determinar se existe uma relação causal entre renda e educação. Como já dito, este resultado é importante para políticas públicas, pois se o efeito captado por renda for na verdade apenas relacionado com habilidades familiares ou conexões sociais, uma política de dar renda para famílias mais pobres estaria equivocada em seu objetivo de melhorar a educação.

Os trabalhos apresentados para a América Latina e Brasil, apesar de não enfocarem a renda, já estimam um efeito da renda que é parcialmente líquido, pois incluem diversas variáveis de controle. Porém não verificam se a correlação estimada da renda é função só dela ou também de variáveis não-controladas.

A vasta literatura que busca determinar o retorno da educação no mercado de trabalho procura estimar o efeito causal da educação sobre a renda futura. No caso dos trabalhos aqui discutidos, buscar-se-ia determinar o efeito da renda dos pais sobre a educação dos filhos, um passo anterior, pois será esta educação que por sua vez influenciará a renda futura da criança. No caso da estimativa do retorno da educação, em geral considera-se que a habilidade inata do indivíduo é o fator não-observado mais importante. A habilidade afeta tanto a decisão de anos de estudo como a renda futura no mercado de trabalho. Portanto as estimativas do retorno da educação incluem o efeito da habilidade. No caso da relação entre renda familiar e educação dos filhos, a habilidade da família e suas preferências em relação à educação parecem ser os fatores não-observados mais relevantes: afetam a renda dos pais e também os indicadores educacionais dos filhos, na medida em que a decisão de frequência escolar e de uso do tempo das crianças é basicamente dos pais. A habilidade da

criança, no entanto, também pode ser fonte de problema nas estimativas, pois também pode estar relacionada à renda dos pais como sugere FEINSTEIN (1998).

FEINSTEIN, usando dados para a Grã-Bretanha, mostra que existe relação entre renda dos pais e habilidade da criança. Mais especificamente, o autor mostra que a renda dos pais explica testes de aptidão para crianças aos 10 anos de idade, condicionando em testes de aptidão realizados aos 5 anos de idade. Isto é, supondo que as crianças têm mesma habilidade antes de começar a escola e têm mesmas características sociais, a renda dos pais tem correlação com a habilidade adquirida após a entrada na escola. Os efeitos de variáveis relacionadas à classe social são maiores em magnitude que o efeito da renda sobre os testes de aptidão, mas é interessante notar que mesmo controlando por estas variáveis (educação e idade dos pais, número de irmãos, etc) o efeito-renda ainda permanece significativo. Se os testes medem de alguma forma a habilidade da criança, mostra-se que a habilidade da criança está relacionada com educação e renda dos pais, reforçando, no caso da estimativa da renda familiar em relação à educação, a necessidade de se isolar o efeito das habilidades não-observadas do efeito da renda.

Utilizando a mesma estratégia empírica seguida por este trabalho, porém analisando a matrícula no ensino superior nos EUA, ACEMOGLU e PISCHKE (2001) exploram a variação diferenciada ocorrida na distribuição de renda ao longo do tempo e entre regiões para identificar o efeito 'líquido' da renda sobre a matrícula. Os autores estimam efeito significativo para a renda.

Como comentado na apresentação dos trabalhos teóricos, a preocupação inicial destes autores era distinguir entre famílias com restrição e sem restrição de renda para investir em educação. Esperar-se-ia que o efeito da renda seria significativo apenas para o primeiro grupo; o efeito-renda para o grupo como um todo seria uma média dos efeitos por grupo de renda. ACEMOGLU e PISCHKE estimam uma regressão permitindo que o efeito-renda varie por

quartil de renda, mas não encontram nenhum resultado conclusivo: estimativas imprecisas e com resultado diferente do esperado (mais ricos com maior efeito marginal).

Para o Brasil, CARVALHO (2001) procura identificar o efeito causal da renda familiar sobre educação e trabalho infantis. Comentarei suas estimativas para a frequência escolar. Sua estratégia de identificação consiste em usar uma mudança na regra da previdência rural em 1992 que gerou um expressivo aumento de renda para a população rural com direito a aposentadoria. O autor utiliza dados da PNAD de antes da reforma (1989 e 1990), e depois dela (1992 e 1993); e uma faixa etária de 10 a 14 anos. É importante ressaltar que o autor analisa apenas o efeito da renda advinda do benefício da aposentadoria, não considerando a renda do trabalho em sua análise.

CARVALHO (2001) apresenta seus resultados em separado para meninos e meninas. Apenas a estimativa do efeito marginal da renda para meninas é significativa, e seu valor é 0.045. Para meninos, nenhuma estimativa é significativa. As mesmas estimativas são feitas separando-se a renda nos benefícios recebidos por mulher e por homem. Neste caso, apenas se mantém significativa a estimativa para meninas relativa a benefícios advindos de uma mulher. O autor faz o mesmo exercício para vários subgrupos para testar a robustez de seu resultado. Separando a amostra de meninas em 2 grupos etários, o efeito do benefício só é significativo para o grupo mais velho, de 12 a 14 anos; e aparece o mesmo padrão anterior, apenas o benefício advindo de uma mulher é significativo.

É difícil comparar os resultados obtidos por KASSOUF (2000) e CARVALHO (2001), apesar de ambos estimarem o efeito da renda sobre a frequência escolar, e usarem amostras semelhantes. CARVALHO, como comentado, utiliza apenas a renda obtida com o benefício da previdência, enquanto KASSOUF usa a renda salarial, e separa a renda que não advém do salário em outra variável (nunca significativa em suas estimativas). Em relação ao receptor da

renda, no caso de KASSOUF refere-se aos pais da criança, mas em CARVALHO é mais provável que os receptores do benefício sejam os avós. Talvez a única comparação refira-se à renda de outras fontes de KASSOUF com o benefício utilizado por CARVALHO: e ambos tendem a não ser significativos. Como já comentado, HAVEMAN e WOLFE (1995) em sua compilação de artigos mostram que a fonte de renda é importante para determinar sua significância, e que a renda do trabalho é a mais importante.

O quinto capítulo vai mostrar estimativas para o efeito da renda familiar advinda de todas as fontes e incluindo todos os membros da família (com exceção da própria criança), sendo que a fonte mais importante é a renda do trabalho, e os receptores de renda mais relevantes são os pais. A estratégia utilizará a variação na distribuição de renda no tempo e entre regiões, como em ACEMOGLU e PISCHKE (2001), para isolar o efeito da renda de variáveis não-observadas.

capítulo 3: FATORES DETERMINANTES DA EDUCAÇÃO NO BRASIL

Como exposto na revisão da literatura, pode-se dividir os fatores que influenciam os resultados educacionais em dois grandes blocos, as políticas públicas e as características da família e da criança. As decisões das famílias são condicionais às possibilidades e restrições que as políticas públicas criam, e dependem também das próprias características das famílias, aí também consideradas as características das crianças.

Primeiro serão apresentadas as principais políticas públicas voltadas para a educação no período analisado, décadas de 80 e 90. Em seguida serão apresentadas algumas das características da família e da criança que a literatura revisada sugere como importantes para explicar os resultados educacionais e que serão utilizadas para as estimativas que se seguem.

Nas estimativas, as políticas públicas estarão apenas representadas por *dummies* de tempo, e em certa medida por *dummies* regionais, primeiro pelo maior interesse deste trabalho em avaliar as características familiares, e segundo, por não haver informação das políticas nos dados da PNAD.

A apresentação das características familiares será feita por ano, e não por coorte, como foram antes apresentados os indicadores educacionais. Esta decisão leva em conta a impossibilidade de se identificar concomitantemente os efeitos das variáveis coorte, ano e idade; desta forma, preferiu-se utilizar como determinantes para as estimativas as variáveis ano e idade. O ano está intimamente relacionado com as políticas públicas, que como veremos, mudam bastante no decorrer do tempo, afetando várias gerações a cada ano. A idade é variável relevante

na decisão de estudo e trabalho, portanto também importante de se considerar explicitamente.

3.1. Políticas públicas

O desempenho educacional do Brasil sempre foi muito criticado, pois mesmo o país apresentando gastos com educação comparáveis a países desenvolvidos, se relativos ao PIB, os resultados educacionais têm estado aquém do que se poderia considerar razoável, mesmo se comparado a seus pares na América Latina, como vimos no primeiro capítulo.

As críticas até meados da década de 90 referem-se constantemente ao mau uso dos recursos destinados à educação, e a uma alocação de recursos que favoreceu o ensino superior em detrimento do ensino básico. É claro o círculo vicioso causado pela combinação escola pública de baixa qualidade e ensino superior público e gratuito de boa qualidade. Só os melhores alunos conseguem uma vaga no ensino público superior, e os bons alunos em sua maioria vêm de escolas privadas, e portanto pertencem a camadas mais ricas da população.

PLANK *et alli* (1996) comentam as questões político-institucionais que consideram mais relevantes para explicar o mau desempenho educacional do país. Eles afirmam que não foi a falta de recursos ou incompetência que geraram este resultado, mas o fato de objetivos privados terem prevalecido sobre o interesse público. Consideram como fatores principais o clientelismo, que aloca recursos em função de objetivos políticos e eleitoreiros; a força política das escolas privadas em receber recursos públicos; e a alocação de grande parte dos recursos para as universidades federais.

VASCONCELLOS (1998) também analisa o círculo vicioso causado pelo favorecimento das universidades na alocação de recursos. Os recursos educacionais, advindos de impostos pagos por todos,

geram ensino de má qualidade para a maioria e de boa qualidade para os poucos, em geral mais ricos, que conseguem chegar às universidades públicas. Desta forma fica perpetuada a distribuição de renda: famílias de baixa renda e baixa educação só conseguem para seus filhos educação de má qualidade, dificultando a aquisição de educação superior. A pouca e má qualidade da educação gera baixos salários e a perpetuação das desigualdades sociais. GARCIA *et alli* (2001) estimam a relação entre desigualdade de renda e alocação de gastos educacionais utilizando uma amostra de 84 países e dados de meados da década de 80. Mostram que quanto maior a desigualdade de renda do país, maior o gasto com ensino superior relativamente ao gasto com ensino fundamental, resultado que contribuiria para perpetuar a desigualdade de renda entre gerações.

Os trabalhos que procuram explicar a renda do trabalho no Brasil encontram na educação, medida em número de anos de estudo completos, seu principal determinante. Procurando explicar a desigualdade de renda, BARROS (1997) compila estudos mostrando que se não houvesse desigualdade de educação, a desigualdade de renda cairia entre 35% e 50%. Portanto, garantir o acesso à educação e incentivar a permanência na escola podem realmente ter conseqüências muito benéficas a longo prazo em relação à renda e sua distribuição.

Várias mudanças institucionais e de objetivos políticos em relação à educação têm contribuído para alterar este quadro de desigualdade na oferta de ensino público. As políticas públicas passaram a descentralizar decisões e dar maior atenção ao ensino básico. Apesar do objetivo de universalização do ensino fundamental já constar como prioridade desde os anos 70, é sua retomada como meta a ser atingida em 10 anos na Constituição de 1988 que dá novo impulso a sua concretização. Além disso, a parcela de impostos que a União deve aplicar em educação aumenta de 13% para 18% da arrecadação em função da Constituição.

A Lei de Diretrizes e Bases da Educação Nacional - LDB¹, sancionada em 1996, também trata dos recursos financeiros públicos destinados à educação, definindo o mínimo que cada esfera do governo deve aplicar em educação. A União deve aplicar pelo menos 18%, como já determinado pela Constituição, e Estados e Municípios, 25% da receita de seus impostos na manutenção e desenvolvimento do ensino público.

O Fundo de manutenção e desenvolvimento do ensino fundamental e de valorização do magistério - Fundef² reitera a vinculação de recursos das três esferas do governo e determina que, a partir de 1998, 60% dos recursos devem ir para o ensino fundamental. A União define um custo mínimo por aluno que assegure a qualidade do ensino e, no caso de Estados e municípios não possuírem recursos suficientes, a União fica responsável por complementar estes recursos, com transferência direta para as escolas. Os municípios passam a ser os maiores interessados em ter o maior número de alunos, pois os recursos são medidos de acordo com o número de alunos efetivamente matriculados. Em 1998 todos os Estados adotaram o valor mínimo por aluno de R\$315; 8 Estados precisaram de complementação da União para atingir o mínimo por aluno, 7 deles da região Nordeste, e 1 da região Norte.

A LDB também introduz inovações em relação aos níveis de ensino e determina maior flexibilidade e fortalecimento da autonomia de Estados e municípios. A obrigatoriedade de matrícula é mantida dos 7 aos 14 anos de idade, mas já consta na lei a intenção de estender a obrigatoriedade à faixa etária correspondente ao ensino médio. A educação infantil, oferecida em creches (crianças de até 3 anos) ou pré-escolas (de 4 a 6 anos) é enfatizada e incluída como parte integrante da educação básica (que inclui também o ensino fundamental e o ensino médio). Cabe aos municípios a oferta de educação infantil e ensino

¹ Lei n. 9394 de 20/12/1996.

² Emenda constitucional 14/96 de 1996.

fundamental, e aos Estados a oferta do ensino médio. Torna-se flexível a organização dos ensinos fundamental e médio, podendo ser em séries ou ciclos, mas mantém-se a duração mínima de 8 e 3 anos, respectivamente; foi aumentada a carga horária mínima anual para 800 horas, num mínimo de 200 dias letivos. Além disso, a LDB trata da melhora da qualificação dos professores, determinando um mínimo de qualificação a ser atingido a médio prazo para cada nível de ensino.

CASTRO *et alli* (2000) comentam a política de descentralização da execução das políticas educacionais, que se deu principalmente pela transferência de recursos financeiros, mas também incluiu a descentralização de programas do Ministério da Educação e do Desporto - MEC, como os programas de merenda escolar, de oferta gratuita de livros didáticos, e de recursos para material e capacitação profissional.

Questiona-se muito nos últimos anos a qualidade das escolas, e os baixos resultados encontrados nas avaliações do Exame nacional do ensino médio - Enem e do Sistema de avaliação do ensino básico - Saeb³ confirmam a pertinência desta preocupação. Este aspecto da política educacional, apesar de naturalmente importante, não será aqui analisado, pois o trabalho não pretende analisar a dinâmica interna da escola, mas apenas os indicadores quantitativos gerados. Faz-se a hipótese simplista de que qualquer escola é melhor que nenhuma, e que sempre há algum ganho na passagem de série, portanto estar freqüentando e estar na série ideal são fatos positivos.

A política educacional na década de 90, portanto, passou a dar mais importância para o ensino básico, o que fica patente não só pelas definições de funcionamento do ensino, mas também pela melhor e maior alocação de recursos financeiros. Isto não

³ Exames de avaliação, o Enem é realizado no final do ensino médio podendo ser feito por todos os concluintes; o Saeb é realizado em uma amostra de escolas pelas 4^a, 8^a séries do ensino fundamental e pelas 3^a séries do ensino médio.

significa, porém, que o país já alcançou todos os seus objetivos em relação à educação. Ainda é grande o atraso escolar dos alunos e é baixa a escolaridade dos jovens, como vimos no primeiro capítulo; além disso, mesmo que em menor número, ainda existem crianças fora da escola. A análise da situação educacional contida no Plano Nacional de Educação⁴ atesta que as crianças fora da escola não são consequência de falta de vagas, mas da precariedade do ensino ofertado e de condições de exclusão e marginalidade social em que vivem alguns segmentos da população.

Dada a relevância da educação para a geração de renda, são justificadas políticas públicas que criem incentivos para promovê-la. Várias medidas de assistencialismo ocorrem vinculadas à escola, como a merenda escolar, a distribuição de uniformes e material, e o transporte escolar. Além disso, vários programas de bolsa-escola e programas visando eliminar o trabalho infantil têm sido realizados ao longo dos últimos anos.

Como o objetivo último deste trabalho é analisar o efeito da renda sobre a educação, os programas de bolsa-escola são de natural interesse, pois justamente usam a renda como incentivo à educação.

A bolsa-escola é considerada no contexto de programas de renda-mínima. O projeto original de renda mínima, apresentado como projeto de lei em 1991⁵ pretendia que a renda mínima atingisse a todos os cidadãos com mais de 25 anos e renda inferior a um determinado patamar, e sugeria que os idosos fossem os primeiros beneficiados. O debate em relação a este projeto e as primeiras experiências de bolsa-escola, basicamente de Brasília e Campinas iniciadas em 1995, mudaram o foco de atenção do projeto de renda mínima para famílias com crianças.

AMARAL *et alli* (1998) resumem várias experiências municipais de programas de renda mínima e bolsa-escola. Todos os projetos

⁴ Lei n. 10172 de 09/01/2001.

apresentam um critério de seleção baseado na renda familiar, em geral *per capita* e exigem a matrícula e frequência escolar de dependentes até 14 anos; além disso exigem um tempo mínimo de residência no município para evitar migração em função do projeto. Alguns projetos incluem também crianças de 0 a 6 anos como beneficiárias. O valor pago de renda mínima está sempre relacionado à linha de pobreza, podendo ser valor fixo para qualquer família, ou variar de acordo com a renda familiar.

Em 1997 foi sancionada uma lei que autoriza a União a co-financiar, em até 50%, programas municipais de renda mínima vinculados à educação; e em 2001 novo impulso foi dado aos programas de bolsa-escola com a criação do Programa Nacional de Renda mínima vinculada à educação - 'bolsa-escola'⁶. Os requisitos para que os programas de renda mínima municipais sejam apoiados pela União são que a renda familiar *per capita* seja inferior a valor fixado a nível nacional, e que haja crianças na família entre 6 e 15 anos, matriculadas em estabelecimento de ensino fundamental regular, com frequência escolar igual ou superior a 85%. O pagamento é feito preferencialmente à mãe, e a participação da União limita-se a R\$15 por criança até o limite de 3 crianças por família. Os municípios com prioridade são os que já estavam assistidos pela lei de 1997, seguidos dos pertencentes aos 14 Estados de menor índice de desenvolvimento humano - IDH, e pertencentes a micro-regiões com IDH igual ou inferior a 0,5⁷.

O programa bolsa-escola, portanto, atenderia diversos objetivos, dar renda mínima a famílias necessitadas, garantir a permanência das crianças na escola, e diminuir o trabalho infantil. No médio e longo prazos, a maior escolaridade alcançada por estas crianças geraria melhores oportunidades de trabalho e de renda, tirando-as do ciclo de pobreza que suas famílias se

⁵ Programa de garantia de renda mínima, apresentado pelo senador Eduardo Suplicy.

⁶ Lei n. 9533 de 10/12/1997; e lei n. 10219 de 11/04/2001.

encontram. Melhores oportunidades para os mais pobres num país que se encontra entre os mais desiguais do mundo é um objetivo importante a ser alcançado.

O programa de renda mínima, vinculando o recebimento da renda à frequência escolar, gera uma correlação forçosamente positiva entre renda e educação. É interessante, contudo, saber se o aumento de renda familiar traria o aumento de escolaridade naturalmente, isto é, se as preferências das famílias já não determinariam por si só a frequência e permanência na escola, dado o aumento de renda.

3.2. Características familiares e individuais

As variáveis relativas à família e à criança que serão investigadas são: renda familiar; anos completos de educação do pai e da mãe; número de filhos morando com a família; famílias com um só pai morando no domicílio; idade, sexo e raça da criança⁸. Outras características relacionadas à família são: região de residência e se é área rural ou urbana.

A dinâmica de decisão em relação à educação é muito diferente entre crianças e jovens, como atestaram os estudos para o Brasil que analisaram as decisões por faixa etária. Para os mais velhos, o custo de oportunidade de não trabalhar torna-se mais importante para as decisões. Não analisarei as decisões da família em relação ao trabalho dos filhos, mas dada a diferença de processos de decisão, não seria correto trabalhar com a faixa etária de 7 a 17 anos supondo o mesmo modelo de decisão para todos.

O restante do trabalho, portanto, restringirá a análise às crianças de 7 a 14 anos, faixa etária para a qual a escola é obrigatória.

⁷ O IDH varia de 0 a 1, sendo 1 o resultado ideal.

⁸ Maiores detalhes das variáveis no anexo A2.

A unidade pesquisada na PNAD é o domicílio, mas preferi utilizar a unidade familiar para esta pesquisa, pois a família está mais diretamente relacionada com as decisões de educação dos filhos. De qualquer forma, a maioria dos domicílios pesquisados possui apenas uma família.

A variável de maior interesse para este trabalho é a renda, pois sua relação com os indicadores educacionais será examinada em maior detalhe no quinto capítulo. A renda familiar considera a renda de todas as fontes dos pais, filhos, outros parentes e agregados morando com a família. Não está incluída a renda da própria criança. Trata-se da renda mensal da semana de referência da entrevista. A renda inclui, portanto, choques transitórios, como o desemprego. Por outro lado, o fato da renda incluir não só a renda do trabalho e também incluir a renda de várias pessoas pode minimizar o efeito dos choques transitórios, pois estes não devem afetar todas as pessoas da família em um mesmo momento, e as rendas de outras fontes dão um caráter menos transitório para a medida de renda.

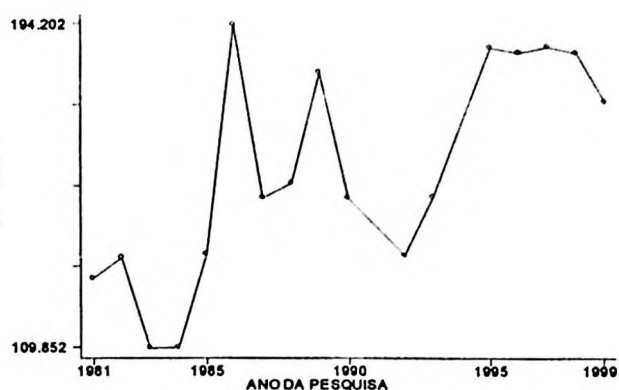
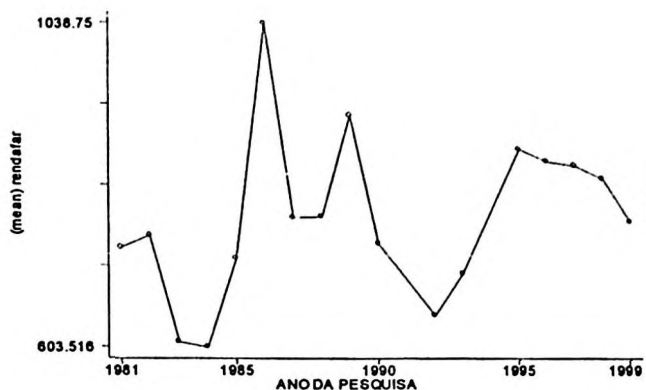
A renda nominal foi corrigida para valores de setembro de 1999 com base no índice nacional de preços ao consumidor - INPC do IBGE. Como o período de análise inclui um período de praticamente hiperinflação e 5 mudanças de moeda, mesmo a correção mais acurada dos valores nominais deve ser vista com cautela. Mesmo dentro de um mesmo ano, sendo os valores informados referentes a semana de referência, pode haver ruído na comparação de rendas. Por exemplo, as negociações salariais promovidas pelos sindicatos ocorrem uma vez ao ano, mas em diferentes meses dependendo do setor ou localização.

Os gráficos 1a e 1b trazem a média da renda familiar total e a média da renda *per capita* (exclusive renda da própria criança), considerando-se famílias com renda positiva. Analisando-se os gráficos, pode-se reconhecer em seus ciclos os períodos de recessão e *boom* econômicos.

Gráficos 1. Renda familiar total e *per capita* (R\$)

(a) renda total

(b) renda *per capita*



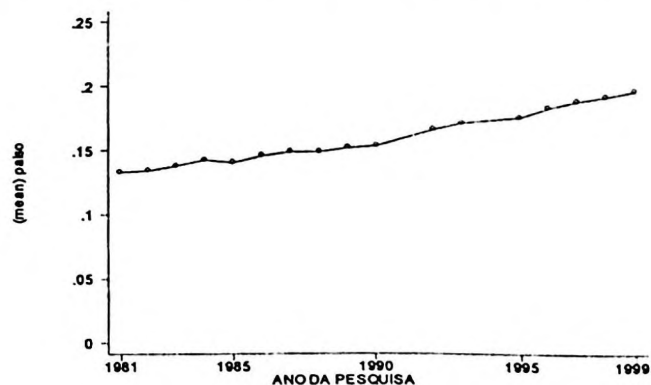
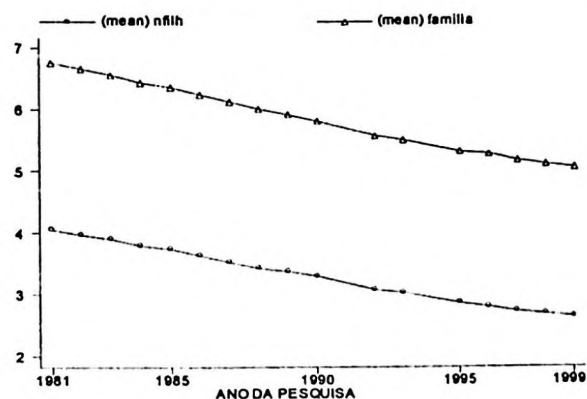
Nota. renda em reais (R\$) de 1999.

Pode-se notar que ambas as tendências, da renda total e da renda *per capita*, apresentam as mesmas oscilações, porém a renda *per capita* apresenta menores quedas e melhores recuperações, por exemplo quando recupera-se em 1995, recupera-se a um nível mais alto que o pico de 1989 e mantém-se constante no tempo, a exceção do último ano. Comparando-se, portanto, os anos extremos, não há praticamente melhora de renda total, mas há melhora da renda *per capita*. Tal fato está relacionado com a queda do número de pessoas na família, como mostram os gráficos 2.

Gráficos 2. Tamanho da famílias

(a) Número de pessoas e de filhos

(b) Famílias com só 1 dos pais

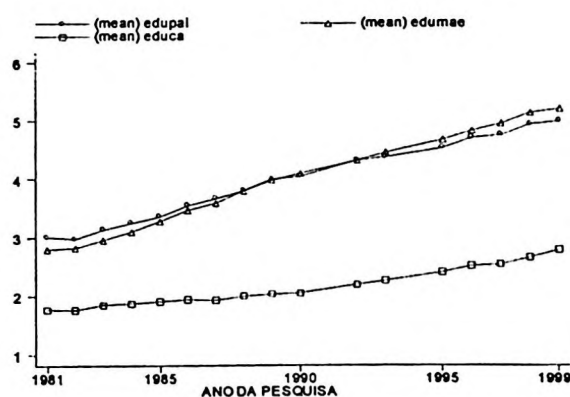


A queda no número médio de pessoas na família e no número de filhos reflete a queda na taxa de natalidade dos últimos anos. Também aumentaram o número de famílias com um só pai presente, na maioria dos casos (87,6%) é o pai (e não a mãe) que não mora no domicílio. Isto comprova o papel importante da mãe como núcleo familiar, justificando que o programa de bolsa-escola considere preferencialmente a mãe como responsável.

Menor número de filhos pode ser vantajoso para a educação, pois permite maior disponibilidade de tempo e recursos dos pais para com a criança. Por outro lado, a presença de apenas um pai no domicílio pode dificultar o sucesso escolar, pois é mais provável que a criança precise ajudar mais nos afazeres domésticos e mesmo com trabalho, e provavelmente menor será a disponibilidade de tempo do pai presente.

O aumento do nível educacional dos pais, medido pelo número de anos de estudo completo⁹ é mostrado no gráfico 3. Cresce bastante a média educacional dos pais, e a média das mães ultrapassa a média dos pais na década de 90.

Gráfico 3. Anos de estudo - pais e filhos



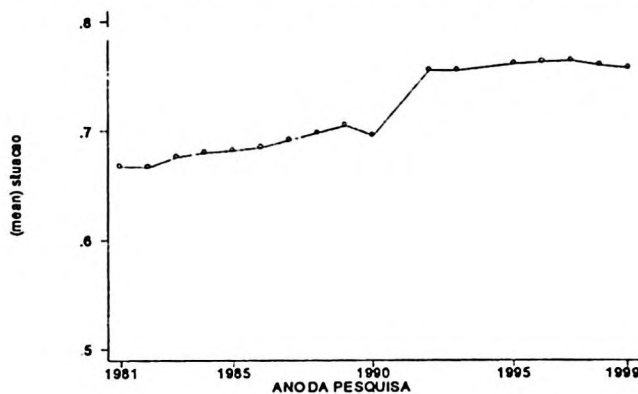
Nota. edupai/edumae - educação dos pais; educa - educação do filho.

⁹ Varia de 0 a 16 anos, explicação mais detalhada em anexo A2.

Para comparação, o gráfico 3 inclui também a média educacional atingida pelos filhos até o ano da entrevista, a média em si não é tão interessante, pois as crianças em sua maioria continuarão estudando, mas a tendência de aumento de educação no tempo mostra a melhora também da escolaridade dos filhos. Esta melhora não é primordialmente função de uma maior idade média, pois o aumento da idade média dos filhos é muito pequeno, de 10,48 para 10,60 no período de 1981 a 1999.

As regiões mais populosas são o Sudeste e Nordeste, as regiões mais ricas e mais pobres, respectivamente. A tabela 1 mostra a distribuição da população entre regiões. Aumenta a proporção de famílias vivendo na zona urbana, como mostra o gráfico 4¹⁰.

Gráfico 4. População urbana (proporção)



As variáveis individuais utilizadas são raça, idade e sexo; suas médias e proporções para o período são apresentadas na tabela 1. A idade média no período é 10,47 anos¹¹ e a população dividi-se basicamente entre brancos e pardos, com 5,3% de negros.

¹⁰ A quebra de tendência no gráfico está relacionada à mudança de uso de Censo demográfico, que é a base de informação para definir zona urbana e rural, do Censo de 1980 para o de 1991 a partir do ano de 1992, incorporando apenas neste ano as mudanças de definição de áreas rurais para urbanas que ocorreram ao longo da década de 80.

¹¹ Idade no mês de junho de cada ano.

Tabela 1. Características individuais - 1981 a 1999

Região	%	SE 40,29	NE 33,42	CO 7,51	N 4,11	S 14,67
Raça	%	branca 49,93	negra 5,3	parda 44,32	amarela 0,46	
Sexo	%	homem 50,43	mulher 49,57			
Idade média		10,47	d-padrão	2,29		

A seguir, na tabela 2, são apresentadas as correlações entre os indicadores 'frequência escolar' e 'estar na série ideal' e cada um de seus possíveis determinantes. Os gráficos dos indicadores e de seus determinantes no tempo já deram uma idéia do sinal destas correlações.

Tabela 2. Correlações entre os indicadores educacionais e seus determinantes

	<i>Frequência</i> ¹	<i>Ideal</i>
Renda <i>per capita</i> ²	0.131	0.259
Educação do pai	0.237	0.411
Educação da mãe	0.245	0.415
Número de filhos	-0.148	-0.230
Urbano	0.221	0.224
Branco	0.127	0.252
1 só pai	-0.025	-0.050
idade	-0.011	-0.332
homem	-0.023	-0.062

Nota. 1: inclui crianças de 7 anos ou mais em pré-escola no ano de 1981 e de 1983 a 1986; 2: apenas renda positiva.

Renda, educação dos pais, e ser branco são positivamente correlacionados com os indicadores educacionais. A correlação destas variáveis com estar na série ideal, entretanto, é praticamente o dobro de com estar frequentando. O mesmo vale para número de filhos, viver com um só pai e ser homem, com correlação negativa com os indicadores educacionais, mas sendo a correlação muito mais forte para ideal. A diferente magnitude dos coeficientes está refletindo o fato de que a grande maioria das crianças está na escola, de modo que as características familiares já não influem tanto nesta decisão. Estar na série

ideal, contudo, ainda continua sendo muito influenciado por estas características, pois inclui, além do fato de se estar freqüentando, o empenho e capacidade necessários para passar de uma série escolar para a outra.

Apenas a correlação com morar em área urbana tem o mesmo sinal (positivo) e a mesma magnitude para ambos os indicadores. A correlação com idade é negativa e alta com ideal, mas é praticamente inexistente com freqüência; este fato reflete a relação não linear entre freqüência e idade, com poucas crianças freqüentado aos 7 e aos 14 anos, e também o fato de 'ideal' refletir o desempenho acumulado ao longo dos anos.

Uma correlação importante entre as variáveis observadas, que já foi comentada na revisão da literatura é a correlação entre renda e educação dos pais. De fato, os seus valores para a amostra utilizada são 0,508 para educação do pai e 0,484 para educação da mãe, em relação à renda real *per capita* (em reais).

A seguir são apresentadas as estimativas do efeito destes determinantes sobre os indicadores educacionais. Veremos, então, o efeito de cada um deles sobre a educação, em especial, o efeito da renda familiar *per capita*.

3.3. Estimativas

Para as estimativas que se seguem a amostra será redefinida, em função de o objetivo último deste trabalho ser determinar o efeito da renda familiar sobre o resultado educacional do filho. Desta forma serão tiradas as observações que não sejam consideradas filhos na família e observações que não tenham renda familiar positiva.

A partir desta nova amostra, serão retiradas 5% das observações de cada ponta da distribuição de renda, para eliminar

outliers¹². A tabela 3 mostra a composição das observações antes da exclusão.

Tabela 3. Característica das observações

<i>Condição na família (%) :</i>		<i>Informação de renda (%) :</i>	
Filho	93,45	Renda>0	96,83
Outro parente	5,68	= 0 / sem inform.	3,17
Chefe/cônjuge	0,19		
outros	0,68		

O fato de se cortar jovens da amostra que não são filhos pode causar um viés de seleção, eles são em maior proporção que a amostra total mulheres não-brancas e apresentam menores médias para os indicadores educacionais, porém o objetivo do trabalho é analisar a decisão de educação na relação entre pais e filhos, portanto a análise só faz sentido para a amostra de filhos.

A seleção que ocorre em função da renda poderia também afetar as estimativas: os sem informação podem pertencer em maior quantidade ao extremo mais pobre, e portanto para o qual a renda seria uma restrição importante, mas são menos de 4% da amostra. Os outros 10% cortados, 5% de cada ponta da distribuição, podem ter dinâmicas de decisão em relação à educação muito diversas, porém preferiu-se evitar maior erro de medida, que poderia ser causado por sua inclusão.

A renda média cai de R\$156,99 para R\$118,79 em função dos cortes, basicamente em função dos cortes dos extremos da renda (pois o corte relativo a condição de família e informação de renda levam a um pequeno aumento da renda média da amostra), portanto famílias muito ricas estão sendo cortadas, como pode-se avaliar pelos valores da renda para alguns percentis, apresentados na tabela 4 (a partir da amostra só de filhos e com renda positiva). As rendas mais altas da amostra, acima de R\$1000,00 para o 99º percentil, passam para a faixa entre 400 e

¹² Explicação de escolha deste corte de amostra no anexo A2.

700 reais, enquanto as rendas mais baixas dobram de valor em função do corte dos extremos de renda.

Tabela 4. Renda *per capita* (R\$) por percentis

ano	p1	p5	mediana	p95	p99
1981	8.456	16.193	72.411	451.844	1101.956
1985	7.384	15.513	69.499	490.838	1194.519
1990	5.838	13.344	71.489	556.032	1334.477
1995	8.596	18.911	96.706	706.341	1611.769
1999	8.333	18.571	91.666	658.333	1500.000

A amostra original com crianças de 7 a 14 anos continha 1.175.331 observações, das quais 9,43% foram excluídas em função da condição na família e da falta de informação de renda. Destas, 10% da amostra foi excluída, 5% de cada extremo da distribuição de renda, resultando numa amostra final para as estimativas de 959.264 observações¹³.

As variáveis dependentes são 'frequência escolar'¹⁴ e 'estar na série ideal', as duas variáveis são binárias. Suas médias não se alteram de forma significativa em função dos cortes de amostra: a média de frequência passa de 83,50% para 84,45%; e a probabilidade de se estar na série ideal, de 38,06% para 38,24%.

As estimativas abaixo, por mínimos quadrados ordinários - MQO¹⁵ utilizam os seguintes regressores: renda familiar, educação do pai, educação da mãe, sexo, morar em área urbana, idade, número de filhos, presença de um só dos pais e raça; além de ano, região de residência e quartil de renda. Foram criadas categorias *missing* para as observações sem informação de raça e sem informação da educação dos pais (na maioria dos casos em função da ausência de um dos pais no domicílio), desta forma não se

¹³ A abertura do número de observações entre anos e regiões está no anexo A2, tabela A2.

¹⁴ Por definição anterior, a frequência incluiria os que já concluíram o ensino de 2º grau, porém não há nenhum caso até os 14 anos de idade.

¹⁵ Estimativas PROBIT serão apresentadas no capítulo 5 para comparação.

perdem as observações nem a possibilidade de usar estas variáveis como controles.

Todas as variáveis, com exceção de renda, são conjuntos de *dummies*; desta forma possibilita-se uma relação funcional mais geral com o indicador educacional. O grande número de observações permite o uso de *dummies* para as variáveis de controle, pois a diminuição de graus de liberdade não compromete a precisão dos resultados estimados. Os erros-padrão são corrigidos usando-se o estimador de variância considerando as células região-ano-quartil de renda como *clusters*, portanto admitindo que as observações dentro de cada célula podem não ser independentes.

Como foram utilizadas *dummies*, para cada nível das variáveis categóricas foi estimado um coeficiente. Desta forma, o coeficiente não representa o efeito marginal a cada nível, mas o efeito total de cada nível em relação ao nível cuja *dummy* foi omitida da regressão.

Não foi controlado o fato de haver irmãos entre as observações; eles são tratados como variáveis independentes. A decisão de oferta de trabalho da criança também não é levada em consideração. Um modelo mais completo deveria incluir a dinâmica relacionada às decisões de oferta de trabalho da família e sua relação com decisões de escola e renda familiar.

Primeiro é apresentada a estimativa para a variável dependente frequência escolar. O R^2 da regressão é 0.149, portanto é baixo o poder explicativo das características familiares e individuais em relação à frequência escolar. Na tabela 5 estão apresentados alguns coeficientes estimados, em seguida, para facilitar a visualização, os outros regressores são apresentados em forma de gráfico (gráficos 5), com o valor dos coeficientes estimados no eixo da ordenada. Os gráficos também mostram o intervalo de confiança de 95% para cada coeficiente estimado.

O efeito marginal da renda é positivo e significativo, assim como o efeito da educação dos pais. Ser negro ou pardo (em relação a branco), homem, viver no meio rural, nas regiões NE e S (em relação à região SE) têm efeito negativo sobre a frequência escolar. Pertencer ao 3º ou 4º quartil de renda tem efeito positivo em relação a pertencer ao 1º quartil (mais pobre).

Tabela 5. Regressão para 'frequência escolar'

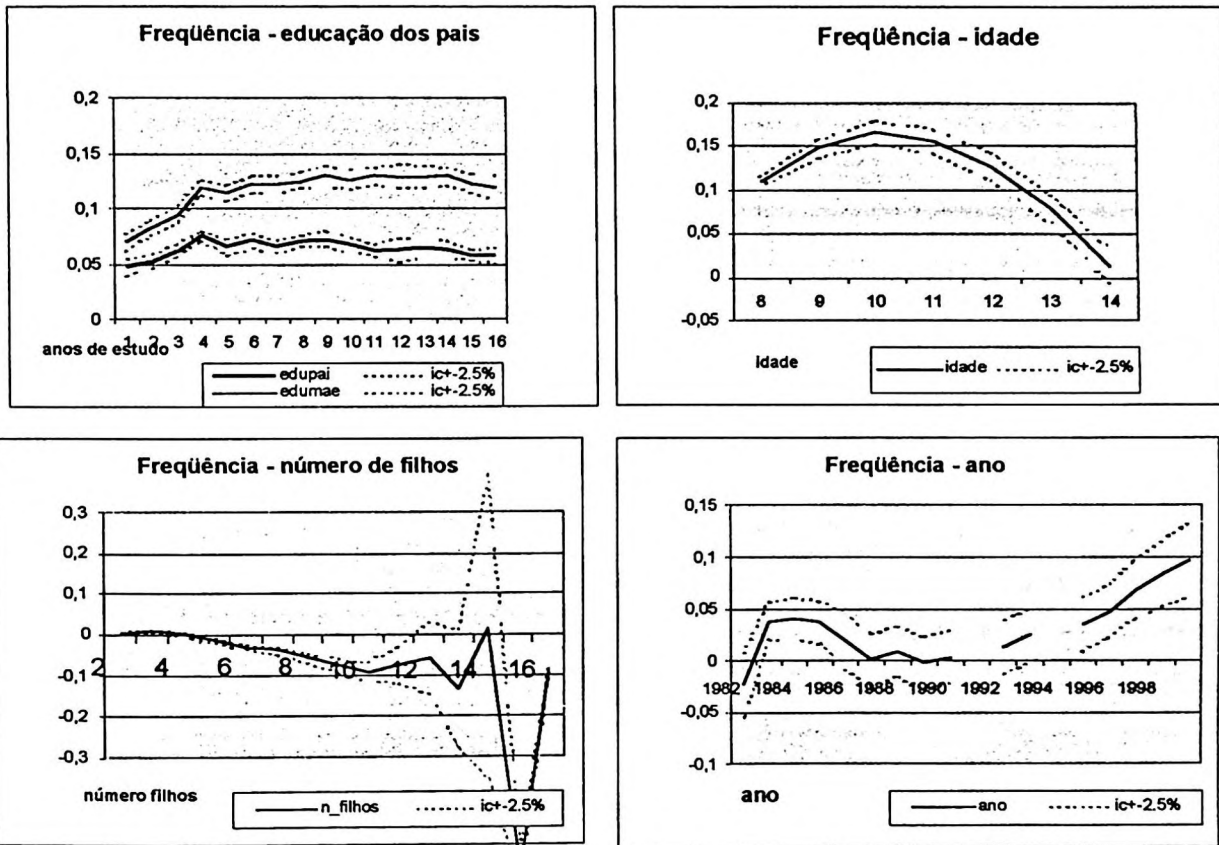
Frequência	Coef.	erro.p.	t	P> t	[95% Int. Conf.]
ln(renda)	0,019	0,003	6,850	0,000	0,013 0,024
homem	-0,018	0,002	-9,210	0,000	-0,022 -0,014
só 1 pai	-0,058	0,016	-3,570	0,000	-0,090 -0,026
área urbana	0,085	0,003	24,720	0,000	0,078 0,091
Raça:					
negra	-0,026	0,003	-8,610	0,000	-0,032 -0,020
amarela	0,006	0,007	0,890	0,375	-0,008 0,021
parda	-0,008	0,002	-4,870	0,000	-0,011 -0,005
Região:					
N	-0,004	0,004	-0,890	0,375	-0,012 0,004
NE	-0,025	0,005	-4,800	0,000	-0,035 -0,015
S	-0,010	0,004	-2,690	0,007	-0,017 -0,003
CO	-0,003	0,004	-0,800	0,423	-0,010 0,004
Quartil de renda:					
quartil 2	0,008	0,006	1,440	0,150	-0,003 0,020
quartil 3	0,025	0,006	4,180	0,000	0,013 0,036
quartil 4	0,035	0,008	4,250	0,000	0,019 0,051
constante	0,459	0,018	25,600	0,000	0,424 0,495
R ²	0,149				
n. observações	959.264				

Nota: outros coeficientes nos gráficos 5; omitidos coeficientes de *dummies* de *missing* e n.filhos=22; as regressões utilizam a peso amostral informado pela PNAD.

O efeito da educação dos pais é positivo e significativo em relação a nenhum ano de estudo. Há, no entanto, uma inflexão nas curvas dos coeficientes apresentadas nos gráficos 5, indicando

que a partir do antigo primário completo não há efeito adicional de mais um ano de estudo sobre a frequência escolar, isto é, o efeito de mais anos de estudo é praticamente o mesmo de ter o primário completo em relação a nenhum ano de estudo. O efeito da educação da mãe é maior que o efeito da educação do pai, como outras pesquisas já haviam sugerido.

Gráficos 5. Coeficientes estimados para frequência escolar



Nota. outros coeficientes na tabela 5.

Número de filhos só passa a ser estatisticamente significativo e negativo a partir de 5 filhos. Para 2 e 3 filhos, o efeito é positivo e significativo, o que pode sugerir um ganho de escala em se mandar mais filhos para a escola, diminuindo custos por filho, como o custo de transporte. Para mais de 10

filhos, o erro-padrão torna-se muito grande, provavelmente em função das poucas observações com número tão alto de filhos.

Os coeficientes relacionados à idade apresentam uma forma côncava, como esperado, pois em relação às crianças de 7 anos de idade, o efeito é maior para as idades centrais, dos 9 aos 12 anos, quando a grande maioria está na escola, e cai para os mais velhos, entretanto estes ainda têm maior probabilidade de estar na escola do que as crianças de 7 anos.

Os anos, que captam o efeito das políticas públicas, também apresentam efeitos diferenciados: em relação ao ano base, 1981, o início da década de 80 contribui mais para ter a criança na escola; no final da década a frequência equivale a de 1981; e a década de 90 apresenta um efeito positivo e significativo sobre a frequência escolar.

Abaixo a estimativa para a variável dependente 'ideal', apresentada na tabela 6 e gráficos 6. O R^2 é 0,360, bem mais alto que o relativo à frequência. Como as correlações já mostravam, o efeito das características familiares sobre estar na série ideal é bem maior que sobre estar frequentando a escola. O resultado da estimativa para ideal, condicional em se estar frequentando (isto é, para uma subamostra apenas com crianças na escola) é bem similar ao não condicional, gerando apenas um R^2 um pouco mais alto, 0,392.

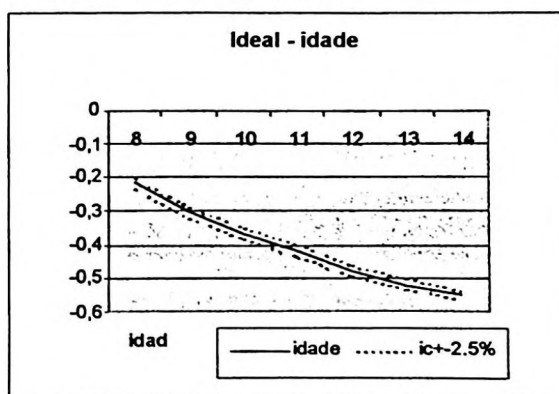
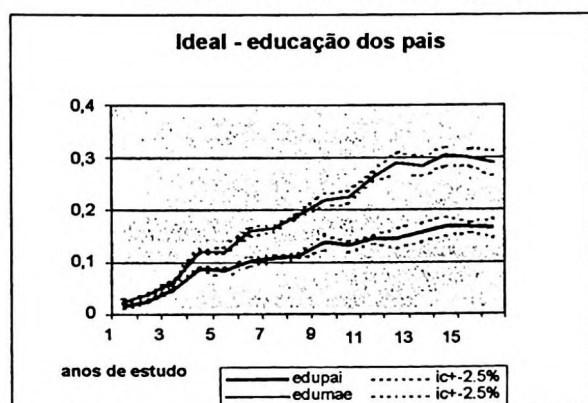
Renda, ser mulher, morar na região SE ou S, morar com os dois pais e na área urbana têm efeito positivo sobre estar na série ideal. Pertencer à raça branca, e em especial à raça amarela, também gera maior probabilidade de se estar na série ideal. Diferentemente das correlações calculadas, morar com os dois pais apresenta um coeficiente de menor magnitude para estar na série ideal em relação ao coeficiente para frequência escolar. Além disso, o coeficiente estimado para área urbana é maior para frequência, e a correlação era praticamente a mesma em relação à frequência e à série ideal.

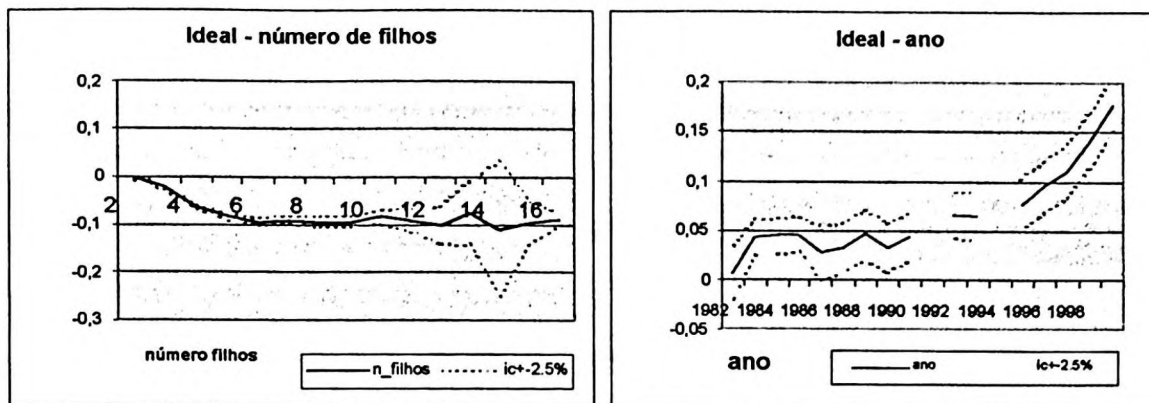
Tabela 6. Regressão para 'série ideal'

Ideal	Coef.	erro.p.	t	P> t	[95% Int. Confiança]	
ln (renda)	0,070	0,003	22,950	0,000	0,064	0,076
homem	-0,065	0,002	-34,150	0,000	-0,069	-0,062
só 1 pai	-0,024	0,009	-2,600	0,010	-0,042	-0,006
área urbana	0,042	0,002	22,250	0,000	0,038	0,046
Raça:						
negra	-0,090	0,004	-21,630	0,000	-0,099	-0,082
amarela	0,075	0,014	5,230	0,000	0,047	0,103
parda	-0,055	0,002	-25,370	0,000	-0,059	-0,050
Região:						
N	-0,116	0,003	-35,890	0,000	-0,122	-0,109
NE	-0,087	0,004	-20,820	0,000	-0,096	-0,079
S	0,044	0,004	11,350	0,000	0,037	0,052
CO	-0,049	0,003	-17,340	0,000	-0,054	-0,043
Quartil de renda:						
quartil 2	-0,015	0,005	-2,720	0,007	-0,026	-0,004
quartil 3	-0,009	0,006	-1,330	0,183	-0,021	0,004
quartil 4	0,004	0,008	0,460	0,649	-0,012	0,020
constante	0,323	0,016	19,610	0,000	0,291	0,356
R ²	0,360					
n. observações	959.264					

Nota. omitidos coeficientes de *dummies* de *missing* e *n.filhos=2*; outros coeficientes apresentados nos gráficos 6.

Gráficos 6. Série ideal - coeficientes estimados





Nota. outros coeficientes na tabela 6.

O efeito dos pais possuírem ensino fundamental ou mais é bem maior sobre estar na série ideal do que sobre estar freqüentando a escola. Além disso, existe efeito marginal em relação a cada ano adicional de estudo dos pais. O efeito da educação da mãe continua sendo maior.

Em relação ao número de filhos, a partir de 3 filhos, o efeito é negativo e significativo. O efeito da idade também é negativo, e aumenta com o tempo, o que é razoável, pois quem deixa de estar na série ideal não tem como voltar a ela.

A partir de 1983, o efeito captado pelas *dummies* de ano são sempre positivos em relação à base 1981, mas só crescem ano a ano no final da década de 90.

Da mesma forma que as variáveis de controle foram utilizadas como conjuntos de *dummies* para captar formas funcionais mais gerais entre estas variáveis e os indicadores educacionais, pode-se também esperar que o efeito da renda sobre os indicadores educacionais seja diferente de acordo com o nível de renda. Como já discutido na revisão da literatura, seria razoável esperar que a renda fosse mais importante para famílias mais pobres, não sendo uma restrição para os mais ricos. O capítulo 5 mostrará

estimativas que confirmam diferentes efeitos da renda familiar sobre os indicadores educacionais de acordo com o quartil de renda ao qual a família pertence.

O próximo capítulo analisará as mudanças no tempo na relação entre as características familiares e demográficas estudadas e os indicadores educacionais.

capítulo 4: ANÁLISE DE DECOMPOSIÇÃO NO TEMPO

A capítulo anterior apresentou as estimativas referentes ao efeito de características familiares e demográficas sobre dois indicadores educacionais: a frequência escolar e estar na série ideal. Estas estimativas envolveram praticamente um período de 20 anos, o que traz uma maior riqueza de informações e permite estimativas mais precisas e de modelos com formas funcionais mais gerais. Por outro lado, dado o longo período de informações, é possível que a composição da amostra em relação a estas características e o efeito destas tenham mudado no decorrer do tempo. Desta forma, as estimativas não deixam de ser válidas, mas representam antes um resultado médio para o período.

Este capítulo pretende verificar se e como as características analisadas modificaram sua importância para explicar os resultados educacionais. Já vimos que os indicadores educacionais melhoraram muito durante os anos de 80 e 90, analisarei agora quanto deste avanço é fruto de uma variação das características observadas e/ou de uma mudança do impacto destas características sobre os indicadores educacionais. Usarei a decomposição de Oaxaca para analisar o peso dos efeitos da composição da amostra e de seus coeficientes estimados sobre as mudanças dos indicadores educacionais no tempo.

Para fazer esta análise serão utilizados dois períodos, os mais distantes entre si no tempo que os dados permitem: início da década de 80 (de 1981 a 1985) e fim da década de 90 (de 1995 a 1999). O uso de períodos maiores de um ano têm a vantagem de minimizar o efeito de choques de curto prazo. Para cada período

serão estimadas regressões iguais de MQO¹. Os coeficientes estimados representam o impacto que a variável tem sobre o resultado educacional. Aplicarei os coeficientes de 80 na população de 90, e os coeficientes de 90 na população de 80. Desta forma, teremos a informação de qual seria o resultado educacional se o impacto das características não mudasse no tempo, mas a população mudasse (efeito-composição); e no segundo caso, a informação do efeito de uma mudança na magnitude do impacto destas características para a mesma população (efeito-coeficiente).

As variáveis explicativas são as mesmas utilizadas no capítulo anterior com exceção de raça. Não utilizarei 'raça' pois os primeiros anos de 80 não possuem esta informação para a maioria da população. A não inclusão desta variável causa uma pequena alteração nos coeficientes das outras variáveis, porém na maioria dos casos os coeficientes não saem do intervalo de confiança de 95% (exceção para as *dummies* das regiões N e S quando 'série ideal' é a variável dependente).

Como se viu no capítulo anterior, a renda familiar média *per capita* aumentou do início da década de 80 para o final da década de 90. A educação dos pais aumentou, principalmente a da mãe, que inclusive passou a ser maior que a do pai. Mais crianças moram na zona urbana. O número de filhos por família diminuiu e a idade média aumentou um pouco.

Estas trajetórias indicam que o efeito-composição sozinho pode levar a um aumento das probabilidades médias dos indicadores educacionais, frequência escolar e estar na série ideal.

Abaixo são apresentadas as distribuições dos principais determinantes da educação para os dois períodos considerados, desta forma temos uma informação mais detalhada da evolução destes determinantes, se compara à informação apenas da média no

¹ o uso de MQO não restringe a variável dependente ao intervalo [0,1], mas para a análise em questão basta saber a direção dos efeitos relativos à composição e aos coeficientes, e esta estará corretamente estimada.

tempo. Além disso, como as regressões utilizam *dummies* para cada nível das variáveis, não são só as médias que são consideradas na regressão. Estas distribuições foram calculadas a partir da amostra para as estimativas, portanto apenas com filhos em famílias com renda positiva e sem 5% de cada ponta da distribuição de renda.

As distribuições são mostradas abertas para os quartis de renda, pois em seguida serão também apresentados resultados por quartil de renda. Como os quartis de renda foram criados dentro de cada região e ano, somar as populações por quartis das diferentes regiões pode gerar resultados de difícil interpretação, pois as características de um quartil em uma região podem equivaler a outro quartil em outra região. Por exemplo, os indicadores de educação e a renda do NE no quartil mais rico (quartil 4) são semelhantes na média ao quartil anterior (quartil 3) da região SE. Em função disto, apresentarei apenas as distribuições para os dois quartis mais distantes, mais pobre (1) e mais rico (4). Cabe lembrar que as estimativas que utilizam quartis não sofrem deste problema, pois incluem sempre controles de região e de ano.

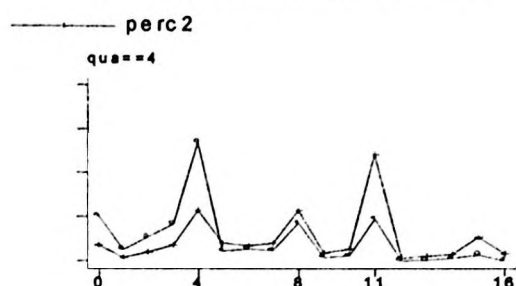
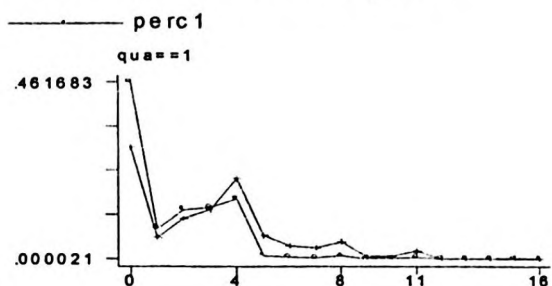
Apresento apenas a distribuição da escolaridade da mãe, gráfico 1, pois a escolaridade do pai mostra distribuição e evolução no tempo similares. Há melhora para ambos os quartis de renda comparando-se os dois períodos. Diminui bastante a proporção de mães com zero anos de estudo; para o quartil mais pobre, o nível de escolaridade concentra-se no primário completo; já para o mais rico, aumentou bastante a proporção de mães com ensino médio completo.

O número de filhos, apresentado no gráfico 2, diminui para todos os quartis, sendo a moda de 3 filhos para os quartis mais pobres, e de 2 filhos para os mais ricos.

Gráficos 1. Distribuição dos anos de escolaridade da mãe

a - Quartil mais pobre

b - Quartil mais rico

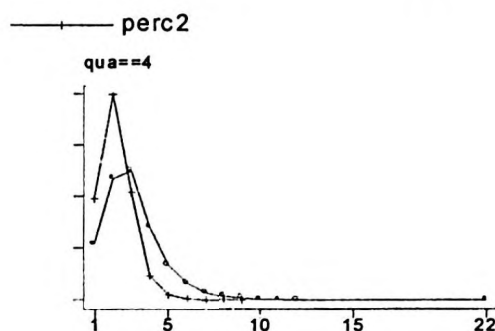
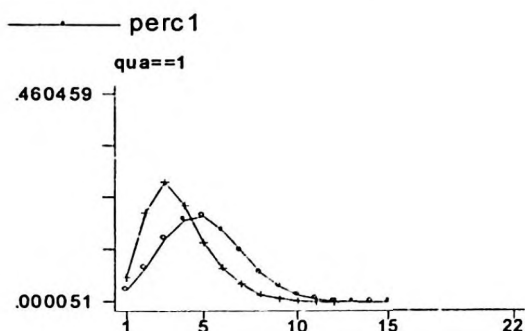


Nota: perc1 - 1981 a 1985; perc2 - 1995 a 1999.

Gráficos 2. Distribuição do número de filhos

a - Quartil mais pobre

b - Quartil mais rico



Nota: perc1 - 1981 a 1985; perc2 - 1995 a 1999.

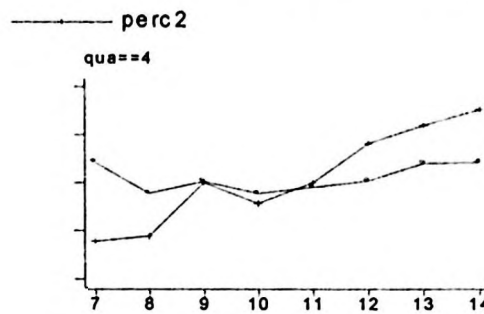
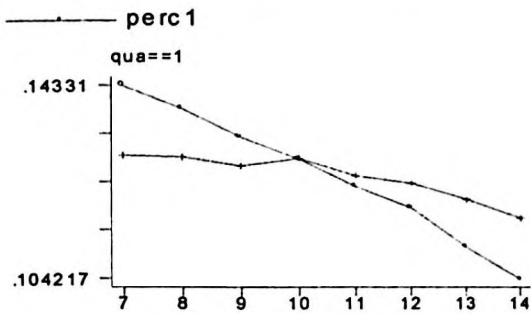
A distribuição da idade dos filhos, apresentada no gráfico 3, também se altera no tempo, seguindo a diminuição da taxa de natalidade: aumenta relativamente o número de crianças mais velhas. Mantém-se, porém, a diferença significativa entre quartis: mais crianças mais jovens nos quartis mais pobres.

A distribuição da renda familiar *per capita* desloca-se para valores mais altos de renda, melhorando um pouco a distribuição (menos assimétrica), como mostrado no gráfico 4. O valor médio total em reais aumenta para R\$ 145,44 no final de 90, o aumento foi maior para o quartil mais rico, de 59,4%, e menor para o quartil mais pobre, 35,1%.

Gráficos 3. Distribuição de idades

a - Quartil mais pobre

b - Quartil mais rico

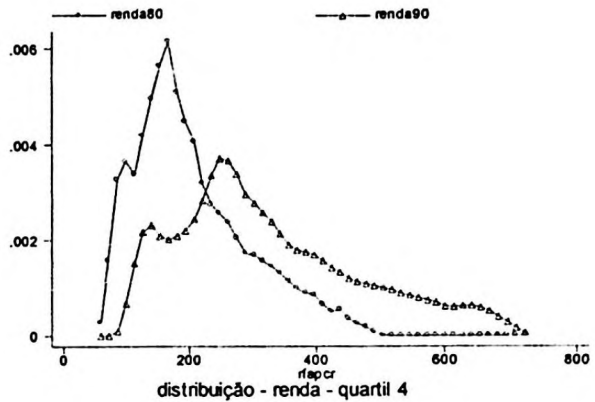
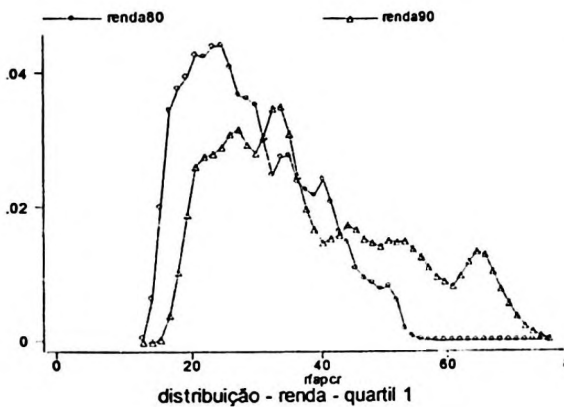


Nota: perc1 - 1981 a 1985; perc2 - 1995 a 1999.

Gráficos 4. Densidade da renda familiar *per capita*

a - Quartil mais pobre

b - Quartil mais rico



Nota: estimativa não-paramétrica, utilizando densidade de Epanechnikov, número de pontos = 50, janela calculada automaticamente pelo programa Stata².

Aumentou a população urbana, como vimos, e este aumento se deu principalmente para os quartis mais pobres. Nos últimos anos 59,6% das pessoas do quartil mais pobre e 92,5% das pessoas do mais rico moram em área urbana, como mostra a tabela 1. Quanto à região de residência, não muda a proporção da população entre as 5 regiões.

² Sejam $m = \min(\sqrt{(\text{variância})}, (\text{faixa entre quartis}/1,1349))$ e $n = \text{número de observações}$. A janela h é calculada como $h = (0,9m)/n^{1/5}$.

Tabela 1. População urbana (%)

<i>período\quartil</i>	<i>1</i>	<i>4</i>	<i>total</i>
1981 a 85	48,1	86,9	68,3
1995 a 99	59,6	92,5	77,7

Aumentou o número de famílias com a presença de um só pai, principalmente para famílias pertencentes ao quartil mais pobre, 18,6% destas famílias vivem com um só pai, comparado a 12,5% das famílias do quartil mais rico. A proporção de brancos não se alterou no tempo, proporção crescente a cada quartil de maior renda.

As estimativas continuam utilizando conjuntos de *dummies* como regressores, portanto a constante estimada representa a probabilidade de frequência, ou de se estar na série ideal, do indivíduo escolhido como base: com renda zero, menina, de 7 anos de idade, vivendo com os dois pais, sem irmãos, na área rural da região Sudeste, com pais sem escolaridade e pertencente ao quartil mais pobre de renda.

Os coeficientes estimados dos regressores mostram quanto as características em questão são importantes para diferenciar a probabilidade de estar na escola (ou de atraso) de um indivíduo com o indivíduo-base. Por exemplo, se a mãe tivesse 4 anos de escolaridade, a probabilidade de estar na escola subiria de 37,5% (valor da constante) para 51,1% (constante somada ao coeficiente da *dummy* de escolaridade da mãe igual a 4 anos), como mostra a tabela 2 a seguir. A magnitude dos coeficientes está relacionada à maior ou menor diferenciação de resultado entre o indivíduo escolhido como base e outro; se cai o coeficiente, significa que diminui esta diferença.

As probabilidades de frequência escolar e de se estar na série ideal estimadas representam o resultado médio da população, isto é, representam a probabilidade de um indivíduo que tivesse os valores médios das variáveis explicativas. Como as variáveis explicativas são *dummies*, à exceção de renda, suas médias

equivalem à percentagem de observações que apresenta determinado nível da variável categórica que gerou as *dummies*. Desta forma, toda a distribuição das características está sendo contemplada no cálculo da probabilidade média da população, e não só suas médias.

Se a composição destas características muda, muda a probabilidade média dos indicadores educacionais; como vimos acima, as características mudam de um período para outro na direção de aumentar a probabilidade estimada. Mas não é só a composição da população que pode mudar no tempo, a magnitude do impacto de suas características sobre os indicadores educacionais também se altera; aqui, o impacto estimado (ou seja, o coeficiente estimado) refere-se à diferença de probabilidade em relação ao indivíduo-base.

A primeira decomposição irá calcular a probabilidade de se estar na escola (ou na série ideal) supondo que a população de 80, com suas características familiares e individuais, convive com o impacto das características válido em 90, isto é, será calculada a seguinte probabilidade média, relativa às características j :

$$\text{prc8p9} = \sum b_j . X_j ,$$

onde:

prc8p9 : composição 80, coeficientes 90;

b_j : coeficiente (inclusive constante) estimado para variável j para 90;

X_j : valor médio da variável explicativa j em 80.

O outro contrafactual refere-se a manter os coeficientes de 80 e considerar as características da população de 90, portanto supor que o impacto das características é o mesmo que no início de 80. Para tanto será calculada a probabilidade:

$$\text{prc9p8} = \sum b_j . X_j ,$$

onde:

prc9p8: composição 90, coeficientes 80;

b_j : coeficiente (inclusive constante) estimado para variável j para 80;

X_j : valor médio da variável explicativa j em 90.

4.1. Decomposição para a amostra total

O exercício de decomposição é apresentado primeiramente para a frequência escolar. Não serão apresentados os coeficientes de todos os regressores, porém estão comentadas as principais variações de um período para o outro.

A tabela 2 apresenta alguns dos coeficientes estimados para a regressão de frequência escolar para os dois períodos. A probabilidade de frequência aumenta de 79,6% para 91,1%, mas o poder explicativo do modelo diminui de um período a outro, passando de 15,9% a 10,0% (R^2). Isto mostra que as características familiares e individuais utilizadas perdem força para explicar a frequência escolar, o que já era de se esperar, pois a grande maioria das crianças está na escola no último período.

A maioria das variáveis diminuiu seu efeito, medido pelo coeficiente estimado, de um período para outro. Destaca-se a diminuição dos efeitos da escolaridade dos pais, da renda e de forma mais acentuada do efeito de morar em área urbana, todos efeitos positivos sobre a probabilidade de frequência. Ser homem e viver com um só dos pais aumentam sua importância, portanto também contribuem para diminuir a probabilidade média de frequentar escola. A variação no tempo do efeito da idade difere entre idades: até 11 anos diminui o efeito em relação aos 7 anos, e a partir daí aumenta. As *dummies* para os dois quartis mais ricos diminuem seu impacto em relação ao quartil mais pobre.

A constante aumentou bastante, em 22,5 pontos percentuais. Isto significa que o indivíduo com as características-base aumentou a probabilidade de estar na escola. Está-se comparando o indivíduo de 7 anos de idade de 80 com o de 90, este aumento está sinalizando a melhora de freqüência escolar e de entrada na idade correta.

Tabela 2. Regressões para freqüência escolar - alguns coeficientes

variável dependente: freqüência escolar				
	81 a 85	e.padrão	95 a 99	e.padrão
ln(renda)	.023	.005	.014	.003
homem	-.010	.004	-.019	.003
urbano	.125	.005	.030	.004
pai só	-.062	.021	-.070	.015
edu=4 mãe	.136	.007	.084	.007
edu=11 mãe	.158	.009	.100	.007
idade=10	.166	.010	.138	.011
nº filho=3	.005	.003	.011	.002
constante	.375	.022	.600	.022
R2	.159		.100	
freq. média	.796		.911	

Nota. mesma especificação do modelo do capítulo 3, tabela 5 e gráfico 5, com exceção de raça.

Nem todas as mudanças de coeficiente acima citadas são significativas, considerando-se um intervalo de confiança de 95%: apenas as variáveis de escolaridade dos pais, área de residência urbana e a constante apresentam variação significativa entre períodos. O cálculo de decomposição, entretanto, considera todos os coeficientes estimados, não levando em conta o grau de precisão das estimativas.

A tabela 3 traz as probabilidades de freqüência escolar, as estimadas para cada período, e as calculadas pela decomposição de Oaxaca. A probabilidade média de freqüência escolar no início de 80 era de 79,6%, e os efeitos composição e coeficiente geram isoladamente aumento similar da freqüência: a probabilidade média passaria para 85,6% e 86,9%, respectivamente. Os 2 efeitos juntos

geraram uma probabilidade de freqüência de 91,1% no final da década de 90.

Grande parte do efeito gerado pela variação dos coeficientes é fruto da variação na constante estimada, que mede antes o nível de probabilidade do indivíduo-base escolhido, e não da mudança da importância relativa das características familiares observadas para explicar diferenças em relação a este indivíduo-base. Se este indivíduo-base não tivesse mudado sua probabilidade, a mudança de impacto das características levaria antes a uma diminuição da probabilidade média de se freqüentar escola, e não a um aumento.

Tabela 3. Decomposição de Oaxaca - freqüência escolar

	coeficiente 81 a 85	coeficiente 95 a 99
composição 81 a 85	.796	.869
composição 95 a 99	.856	.911

A tabela 4 apresenta o resultado das regressões para os 2 períodos relativo à variável dependente 'série ideal'. O poder explicativo do modelo estimado mantém-se praticamente constante, em cerca de 34% (R2). A probabilidade média de se estar na série ideal passa de 29,3% no primeiro período para 49,1% no segundo.

A maioria dos coeficientes estimados também variaram de um período a outro no sentido de diminuir a probabilidade média de se estar na série ideal, com exceção da constante e dos níveis mais baixos de educação dos pais. Diminuiu o impacto positivo da renda e de viver em zona urbana, e aumentou o impacto negativo de sexo, número de filhos e viver com um só dos pais. O efeito das *dummies* de quartil diminui bastante de magnitude. O impacto da educação dos pais variou de um período a outro de forma diferenciada de acordo com o nível educacional: aumentou o efeito da educação do pai até 4 anos de escolaridade e da educação da mãe de 1 ano de escolaridade em relação a nenhuma escolaridade, e diminuiu o efeito da educação em níveis superiores a estes.

As variações significativas, considerando-se o intervalo de confiança de 95%, referem-se a sexo, área urbana e idade, além de alguns quartis de renda e regiões. As outras variações de coeficientes não saíram do intervalo de confiança de 95%, entre elas, a renda e escolaridade dos pais.

Tabela 4. Regressões para série ideal - alguns coeficientes

variável dependente: série ideal				
	81 a 85	e.padrão	95 a 99	e.padrão
ln(renda)	.078	.007	.071	.004
homem	-.047	.003	-.082	.003
urbano	.054	.003	.022	.005
pai só	-.027	.007	-.063	.033
edu=4 mãe	.131	.004	.128	.004
edu=11 mãe	.284	.005	.271	.009
idade=10	-.443	.012	-.302	.015
nº filho=3	-.017	.004	-.023	.004
constante	.291	.028	.349	.025
R2	.346		.342	
ideal média	.293		.491	

Nota. mesma especificação do modelo do capítulo 3, tabela 6 e gráfico 6, com exceção de raça.

A constante aumentou em 5 pontos percentuais a probabilidade de se estar na série ideal. Isto significa que o indivíduo com as características-base aumentou a probabilidade de estar na série ideal em 5%. Este aumento capta, portanto, o aumento de frequência escolar aos 7 anos de idade, o que mostra o aumento de crianças entrando na idade certa.

Da mesma forma que para 'frequência escolar', foram calculados dois contrafactuais para a probabilidade de se estar na série ideal: mantendo-se os coeficientes estimados de 80 para a população de 90; e supondo a mesma composição de população de 80 no final de 90. Foram também estimadas as regressões e calculados os contrafactuais para uma amostra condicional em se estar frequentando escola. Trata-se de uma tentativa de determinar se a dinâmica relacionada a não se frequentar escola, e portanto não poder estar na série ideal, é ou não diferente da

dinâmica relacionada com o atraso escolar de quem frequenta escola.

Como as características familiares consideradas são as mesmas para frequência e estar na série ideal, o efeito da variação da composição das características também é de contribuir para aumentar a probabilidade de se estar na série ideal.

Na tabela 5 estão todas as probabilidades de se estar na série ideal, as estimadas para cada período, e as calculadas como acima descrito. Pode-se ver que o aumento da probabilidade se deveu tanto ao efeito-composição quanto ao efeito-coeficiente, sendo o primeiro efeito um pouco maior, levando à probabilidade de 41,2%, contra a probabilidade de 37,8% gerada pelo efeito-coeficiente. Os dois efeitos levaram a probabilidade média de se estar na série ideal para 49,1% no final de 90. Para a amostra condicional, os dois efeitos são ainda mais parecidos.

Desconsiderando a variação da constante, ocorreria um efeito dos coeficientes menor, mas ainda positivo, sobre a probabilidade de se estar na série ideal de 80 para 90.

Todas as probabilidades são mais altas para a amostra condicional, o que já era esperado, dado que todas as observações desconsideradas não estavam na série ideal. O aumento de probabilidade de um período a outro é maior para a amostra não-condicional, confirmando a importância do aumento da frequência escolar para a melhora do indicador de série ideal.

Tabela 5. Decomposição de Oaxaca - estar na série ideal

	Amostra total		Amostra condicional	
	coeficientes		coeficientes	
	81 a 85	95 a 99	81 a 85	95 a 99
composição 81 a 85	.293	.378	.368	.451
composição 95 a 99	.412	.491	.458	.539

Nota. amostra condicional: subamostra apenas com crianças frequentando escola.

4.2. Decomposição por quartil de renda

É possível que os indicadores educacionais evoluam de forma diferenciada entre os quartis de renda. Será apresentada agora a mesma decomposição de Oaxaca para os quartis de renda mais rico e mais pobre. Primeiro será considerado que as variáveis explicativas têm o mesmo efeito sobre a educação independentemente do nível de renda, e olhar-se-á por quartil apenas o efeito da variação de composição de suas características. Ou seja, com os coeficientes estimados para a amostra total, serão calculadas as probabilidades médias de cada quartil de renda, para os períodos de 80 e 90.

Em seguida será permitido que as características afetem a educação de forma diferenciada por quartil (serão estimados os coeficientes para cada subamostra composta por um quartil de renda), e verificar-se-á se o impacto das características e sua variação no tempo são diferentes entre quartis.

Antes de apresentar o resultado das estimativas, é interessante olhar a evolução dos indicadores educacionais por quartil de renda no tempo.

Como mostra a tabela 6, o aumento de frequência escolar evoluiu de forma diferenciada entre regiões e quartis de renda: a frequência dos quartis mais pobres cresce mais rápido que dos quartis mais ricos, mesmo porque os quartis mais ricos já estavam bem próximos da universalização do ensino.

Há grande diferença de média de frequência se comparadas as médias por região e quartil. O quartil mais pobre da região NE tinha apenas 62,1% de crianças na escola no período de 1981 a 1985, comparado a 77,7% entre 1995 e 1999. No mesmo período, o quartil mais rico do SE aumentou sua frequência de 94,7% para 97,8%.

Tabela 6. Probabilidade média por região e quartil

Frequência escolar					
1981 a 1985					
região/quartil	1	2	3	4	média
SE	0,743	0,805	0,872	0,947	0,846
N	0,812	0,855	0,881	0,925	0,869
NE	0,621	0,673	0,753	0,881	0,720
S	0,704	0,773	0,844	0,916	0,800
CO	0,682	0,763	0,841	0,930	0,799
média	0,691	0,755	0,830	0,923	0,796

1995 a 1999					
região/quartil	1	2	3	4	média
SE	0,893	0,928	0,956	0,978	0,941
N	0,848	0,893	0,936	0,963	0,908
NE	0,777	0,834	0,885	0,948	0,854
S	0,892	0,935	0,965	0,985	0,942
CO	0,867	0,909	0,945	0,973	0,923
média	0,849	0,896	0,934	0,970	0,911

Na tabela 7 a mesma abertura é apresentada para as probabilidades de se estar na série ideal por quartil de renda. Também nota-se a grande defasagem entre a região NE e os melhores resultados das regiões SE e S.

Assim como as médias dos indicadores escolares são diferentes entre os níveis de renda, e evoluíram de forma diferenciada no tempo, é possível que as características observadas tenham efeito diferenciado sobre os indicadores de cada quartil.

Como vimos, as características de cada quartil evoluem de forma a melhorar os indicadores educacionais, por exemplo melhora para todos os níveis de renda a escolaridade dos pais e a renda familiar *per capita*, e diminui o número de filhos. Os níveis destas variáveis continuam, contudo, apresentando diferenças entre quartis.

Tabela 7. Probabilidade média por região e quartil

Série ideal					
1981 a 1985					
região/quartil	1	2	3	4	média
SE	0,203	0,271	0,366	0,564	0,357
N	0,162	0,195	0,249	0,386	0,248
NE	0,105	0,127	0,181	0,349	0,179
S	0,234	0,303	0,421	0,601	0,373
CO	0,144	0,196	0,278	0,476	0,267
média	0,169	0,221	0,308	0,497	0,293
1995 a 1999					
região/quartil	1	2	3	4	média
SE	0,413	0,527	0,642	0,758	0,594
N	0,240	0,290	0,412	0,596	0,380
NE	0,163	0,218	0,313	0,549	0,296
S	0,470	0,609	0,692	0,818	0,639
CO	0,327	0,412	0,541	0,712	0,495
média	0,321	0,419	0,533	0,699	0,491

A tabela 8 mostra que o efeito-composição explica mais o aumento da probabilidade de se estar na escola para os quartis mais ricos. Para diferentes estimativas por quartil, o efeito-composição leva a uma probabilidade média da frequência de 96,6% enquanto o efeito dos coeficientes leva a 94,4%. De forma contrária, o efeito-coeficiente gera maior variação do que o efeito-composição para os quartis mais pobres, gerando uma probabilidade de 80,3% contra 75,6% do efeito-composição.

Tabela 8. Decomposição de Oaxaca - frequência escolar por quartil

	Mesmas estimativas		Estimativas por quartil	
	coeficientes		coeficientes	
	81 a 85	95 a 99	81 a 85	95 a 99
Quartil 1 (25% mais pobres)				
composição 81 a 85	.706	.807	.691	.803
composição 95 a 99	.761	.851	.756	.849
Quartil 4 (25% mais ricos)				
composição 81 a 85	.925	.938	.923	.944
composição 95 a 99	.963	.967	.966	.970

De novo, a variação da constante tem papel fundamental, principalmente para os quartis mais ricos. Portanto, o que se vê é uma diminuição da importância das características observadas (em relação ao indivíduo-base). O poder de explicação dos modelos, em concordância com a diminuição de importância das variáveis estudadas, diminui para todos os quartis de renda, sendo esta diminuição maior quanto mais rico o quartil de renda. Não há muita diferença entre os resultados usando coeficientes estimados para a amostra total ou para cada quartil.

Tabela 9. Decomposição - série ideal por quartil

	Mesmas estimativas		Estimativas por quartil	
	coeficientes		coeficientes	
	81 a 85	95 a 99	81 a 85	95 a 99
Quartil 1 (25% mais pobres)				
composição 81 a 85	.170	.238	.169	.251
composição 95 a 99	.255	.322	.227	.321
Quartil 4 (25% mais ricos)				
composição 81 a 85	.491	.559	.497	.563
composição 95 a 99	.628	.686	.642	.699

A tabela 9 mostra os resultados para 'série ideal'. O efeito-composição é maior para o quartil mais rico independentemente dos coeficientes serem estimados com a amostra total ou por quartil. Para o quartil mais pobre, muda a importância relativa entre efeitos composição e coeficiente quando as estimativas são feitas para cada quartil em separado. No caso das estimativas em separado, o efeito-coeficiente torna-se mais importante, como era para frequência escolar, gerando probabilidade de 25,1% contra 22,7% do efeito-composição. O fator mais importante a contribuir para esta diferença de resultados para o quartil mais pobre é a renda. O efeito da renda para a amostra total diminui no tempo, mas quando estimado em separado, aumenta no tempo para os mais pobres, e diminui para mais ricos.

Vimos, portanto, que ambos, efeito-composição e efeito-coeficiente na medida em que inclui a variação da constante, contribuíram para a melhora dos indicadores educacionais. As políticas públicas ocorridas entre um período e outro são captadas por esta variação da constante, que no caso da frequência escolar é responsável pelo fato do efeito total da variação de coeficientes ser positivo.

É importante notar que a diminuição dos coeficientes da maioria das variáveis explicativas, que a princípio poderia sugerir um incentivo à piora dos indicadores educacionais, antes indica a menor importância destas características, fato relacionado a alta proporção de crianças na escola no último período.

É importante também notar que a evolução dos indicadores educacionais para o quartil mais pobre foi bem maior, e mesmo observando que o efeito-composição em relação ao efeito-coeficiente foi menos importante para este quartil, ele causa um aumento relativo nos indicadores educacionais maior que o efeito-composição causa sobre o quartil mais rico: no caso da frequência escolar, por exemplo, o aumento gerado pelo efeito-composição foi de 7,8% para o quartil mais pobre e de 4,1% para o quartil mais rico.

O próximo capítulo concentrará sua análise no papel da renda familiar sobre os indicadores educacionais, procurando determinar se o efeito positivo e significativo até agora estimado é causal ou apenas reflete o efeito de variáveis não-observadas.

capítulo 5: FREQUÊNCIA E ATRASO ESCOLAR E SUA RELAÇÃO COM A RENDA FAMILIAR

Pretende-se determinar neste capítulo se existe uma relação causal entre renda familiar e educação. A causalidade entre duas variáveis, diferentemente da correlação, significa que uma ação em uma delas causa um efeito particular em outra, e que se esta ação não tivesse ocorrido, o efeito também não o teria. A observação de correlação entre os indicadores educacionais e a renda familiar, vista nos capítulos anteriores, não é garantia de causalidade, pois outros fatores podem ter afetado as duas variáveis simultaneamente.

A comprovação de existência de causalidade entre renda e educação mostraria que o uso de políticas públicas que afetam a renda, com o objetivo de melhorar os indicadores educacionais, tem boa probabilidade de sucesso. Caso contrário, programas como o bolsa-escola estariam apenas cumprindo um papel de redistribuição de renda, não fosse a obrigatoriedade das famílias de manter os filhos na escola para receber o benefício.

Como já comentado, a renda pode estar captando o efeito sobre os indicadores educacionais de variáveis relacionadas a habilidades e preferências da família. Como a decisão a respeito da educação das crianças é basicamente dos pais, suas habilidades afetam tanto sua renda, como também os indicadores educacionais das crianças. Além disso, pode haver também relação entre a renda familiar e a habilidade da própria criança, como atesta o estudo de FEINSTEIN (1998) para a Grã-Bretanha. A direção do viés causado pelo problema de variável omitida depende de como esta variável omitida afeta renda e educação; no caso de habilidades, elas

afetam tanto a renda como a educação positivamente, e portanto o viés será no sentido de aumentar o efeito estimado da renda.

A inclusão de outras variáveis observadas como controles na regressão é necessária para atenuar o viés de variável omitida. Porém, o problema permanece pois fatores como habilidade e preferências por educação continuam no termo de erro.

Além disso, pode haver erros de medida na renda, resultantes do período de alta inflação e das mudanças de moeda. As estimativas utilizando MQO apresentariam um viés de atenuação em função da existência de erro de medida na renda.

O método a ser utilizado neste trabalho para examinar a existência de causalidade procura encontrar grupos similares que tenham diferentes níveis de renda por razões que não afetam diretamente a educação dos filhos. Desta forma é possível comparar os indicadores educacionais destes grupos e inferir a relação de causalidade entre renda e educação.

É interessante ter em mente o efeito de qual renda está sendo estimado. A renda informada na PNAD, como visto, inclui tanto salários como renda de outras fontes, ambos referentes a um determinado mês do ano, e portanto inclui choques, transitórios ou não. A inclusão de controles correlacionados à renda permanente, como educação dos pais, e do controle da posição na distribuição de renda, com quartil de renda, leva a que o coeficiente estimado da renda exacerbe o efeito destes choques. Não será possível distinguir entre variações transitórias ou permanentes, mas de qualquer forma os indicadores educacionais podem ser afetados tanto por choques transitórios como permanentes na renda, mesmo que a literatura enfatize a restrição de crédito relacionada à renda permanente como origem da relação entre renda e educação.

5.1. Estratégia empírica

5.1.1. Apresentação do instrumento

O método de variáveis instrumentais será usado para isolar o efeito direto da renda e determinar portanto se existe um efeito causal entre renda e educação. Um dos maiores problemas com o uso deste método é encontrar um 'bom' instrumento. O instrumento, para gerar uma estimativa consistente, deve ser correlacionado com a variável endógena, no caso, renda familiar, e não ser correlacionado com o termo de erro. Além disso, deve ter variação suficiente para identificar o efeito da renda sobre a educação.

A estratégia a ser adotada será explorar as variações na distribuição de renda ao longo dos anos. Como veremos, os períodos de alta inflação e os diversos planos de estabilização geraram variações na renda que dificilmente estão relacionadas com as características familiares não-observadas, e portanto podem ser consideradas variações exógenas. A distribuição da renda será medida pela diferença de renda média entre quartis de renda, criados por região de residência e ano.

Vários autores utilizam estratégia semelhante, recorrendo a algum evento exógeno, que permita definir grupos que tenham a variável endógena afetada de forma diferente por este evento. Estas diferentes variações na variável endógena são então usadas para explicar a variável de interesse. A princípio, estes eventos poderiam gerar apenas 4 grupos para uma análise de diferenças em diferenças: os tratados e não-tratados antes e depois do evento, porém é sempre interessante aumentar o número de grupos de forma a gerar maior variação no instrumento. O instrumento, neste caso, é formado pelo conjunto de todas as interações entre as *dummies* das variáveis que definem os grupos.

CARVALHO (2001), como apresentado no capítulo 2, utiliza a reforma da previdência rural no Brasil para estimar a relação da renda familiar com o trabalho infantil e com a frequência escolar. O instrumento é construído como o conjunto das

interações das características que incluem a família no grupo afetado pela reforma com uma *dummy* de tempo (antes e depois da reforma).

DUFLO (2001), para determinar a relação entre anos de educação e renda do trabalho na Indonésia, usa como instrumento para anos de estudo um programa de construção de escolas na década de 70 que privilegiou regiões do país onde a matrícula inicial na escola era baixa. ANGRIST e KRUEGER (1991), também para medir a relação entre anos de educação e renda, exploram a compulsoriedade da escola até os 16 anos nos EUA, que obriga jovens com mês de aniversário depois do início do ano letivo a ficar mais um ano na escola. Os autores representam a compulsoriedade da escola com *dummies* de quadrimestre de nascimento interagidas com *dummies* de ano de nascimento. A princípio, a variação exógena da educação causada pela compulsoriedade da escola já é dada pelo mês de nascimento. Porém, dado que o efeito da lei de compulsoriedade de escola diminui no tempo, é possível usar a interação com coorte para aumentar a variação do instrumento.

ACEMOGLU e PISCHKE (2001) utilizam a mesma estratégia empírica deste trabalho para explicar o efeito da renda dos pais sobre a matrícula em ensino superior nos EUA. Eles exploram o fato de que a renda variou de forma diferenciada entre quartis de renda e entre regiões no tempo, favorecendo grupos mais ricos. Os autores acham um forte efeito-renda, contrariando estimativas de CARNEIRO e HECKMAN (2002), que alegam que a restrição de crédito só existe para uma pequena parcela pobre da população norte-americana. ACEMOGLU e PISCHKE não incluem as características familiares como controles em suas estimativas. Eles supõem que o instrumento utilizado é suficiente para isolar o efeito direto da renda sobre a matrícula. Considero necessário para o caso brasileiro, no entanto, o controle das características familiares observadas para garantir que o instrumento não seja correlacionado com o termo de erro.

Aqui também será utilizada a variação diferenciada da renda entre quartis e regiões no tempo. São criadas 340 células = 17 anos x 5 regiões x 4 quartis de renda. Como mostrado no anexo A2, tabela A2, não há problema com tamanho de amostra, pois o número de observações por célula é grande: média de 2821 observações por célula, sendo que a menor célula possui 394 observações. Quanto mais células criadas, maior a variação do instrumento, mas diminui o número de observação por células. Além disso, só aumenta a variação do instrumento se houver variação diferenciada entre os novos subgrupos criados. Este é o caso das regiões de residência, e talvez fosse verdade em relação aos Estados de residência, pois algumas regiões são muito desiguais internamente. Porém, com a inclusão dos Estados, o número de células iria para mais de 1000, dificultando a implementação computacional deste instrumento.

O uso do conjunto das interações entre as *dummies* de quartil, de região e de ano como instrumentos permite colocar, se for necessário, as 3 *dummies* (região, quartil e ano) como controles na equação principal. No caso aqui estudado, as diferenças de renda entre regiões, entre quartis e entre anos são provavelmente relacionadas a variáveis omitidas que também afetam a educação. As *dummies* de ano controlam não só efeitos macroeconômicos, mas também representam as políticas educacionais, que como vimos mudaram muito ao longo do período analisado. As regiões apresentam diferentes níveis de desenvolvimento, que também afetam a renda e o resultado educacional. Desta forma, a diferença entre regiões e entre anos, que são representadas genericamente por *dummies*, estão relacionadas tanto à renda como à educação, e portanto precisam ser incluídas como variáveis de controle na regressão principal. A posição na distribuição de renda, representada pelo quartil de renda a que a família pertence, também pode estar relacionada a características não-observadas, como habilidades. Esta posição

portanto deve também ser utilizada como controle na regressão principal, por meio das *dummies* de quartil.

Será utilizado o método de mínimos quadrados em 2 estágios (2SLS¹), que equivale ao método de variáveis instrumentais. O primeiro estágio gera a estimativa da renda prevista como função do instrumento e dos controles exógenos. O segundo estágio, a regressão principal, estima a educação em função da estimativa de renda gerada no primeiro estágio e os mesmos controles.

Como já explicado, o instrumento é constituído pelo conjunto de interações entre as *dummies* de tempo, de região e de quartil de renda, mas a parte excluída da regressão principal refere-se apenas à variação diferenciada no tempo entre quartis e regiões, pois as *dummies* de nível - região, ano e quartil - estarão na regressão principal como controles.

BLUNDELL *et alii* (1998), interessados em determinar o efeito do salário sobre a oferta de trabalho, utilizam método semelhante, e formalizam as condições para a implementação deste instrumento. Considere o modelo (a ser estimado):

$$\text{edu}_{it} = a + b \cdot \ln w_{it} + \sum c_j \cdot F_j + u_{it}$$

onde edu: educação (frequência escolar ou série ideal)
lnw: ln(renda)
F_j: característica familiar j
u: termo de erro
a, b, c_j: coeficientes
i - indivíduo, t - ano.

1. restrições de exclusão:

$$E(u_{it} \mid \text{quartil, região, tempo}) = D_q + D_r + D_t$$

onde D: *dummy*
t- ano, r - região, q - quartil.

As diferenças não-explicadas na média do indicador educacional entre grupos (após controle das variáveis familiares

¹ *Two stage least squares.*

observadas) são explicadas apenas por um efeito grupo permanente (região e quartil) e por um efeito tempo aditivo. Esta condição garante a não-correlação entre o resíduo e o instrumento (interações das *dummies* de quartil, região e ano).

2. condição de posto:

$$[E(\ln w_{it}|q,r,t) - E(\ln w_{it}|q) - E(\ln w_{it}|r) - E(\ln w_{it}|t)]^2 \neq 0$$

Para que esta condição seja satisfeita, a renda familiar tem que variar de forma diferenciada entre quartis e regiões ao longo do tempo, ou seja, após controlar os efeitos de quartil, região e tempo ainda existe variância na renda.

A hipótese de identificação implícita nestas condições é de que as variáveis não-observadas (e portanto não controladas) podem variar entre quartis e entre regiões e também ao longo do tempo, mas não podem variar entre quartis de forma diferenciada entre regiões ao longo do tempo.

A estimação, portanto, compara grupos que, afora os níveis das variáveis ano, quartil e região, e controles familiares adicionais, foram afetados em sua renda de forma diferenciada no tempo, e esta variação diferenciada é advinda de uma fonte exógena à educação.

O estimador gerado equivale a um estimador de dados agrupados como demonstra ANGRIST (1991), isto é, a um estimador de mínimos quadrados ponderados gerado a partir de uma amostra formada com as médias das variáveis dos grupos. A ponderação refere-se ao número de observações por grupo. Se houvesse apenas dois períodos e dois grupos, o estimador seria a estimador de diferenças em diferenças. Como são mais grupos e períodos, ANGRIST (1991) refere-se a este estimador de 2SLS como um estimador de Wald eficiente.

Se não houvesse variáveis de controle além das *dummies* que definem os grupos, o resultado da estimativa com instrumento (IV) seria idêntico ao da estimativa sobre as médias das variáveis. No caso de haver controles adicionais, estes também são reduzidos às suas médias quando os dados são agrupados, e portanto sua variação dentro das células também é ignorada na estimação por mínimos quadrados ponderados. Se a variação dos controles dentro das células for importante, então os resultados pelo método em 2 estágios com os dados individuais e por mínimos quadrados ponderados com os dados agrupados serão diferentes. Neste caso é preferível o resultado com os dados individuais, que controla corretamente pelas características familiares e demográficas, pois leva também em conta a variação destas características dentro dos grupos.

Para verificar a validade do instrumento serão realizados testes como sugerido em BOUND *et alii* (1995). Será feito um teste F de significância dos instrumentos na regressão do 1º estágio, ou seja, regressão de renda familiar em função dos instrumentos e das variáveis de controle exógenas. Também será apresentado o resultado do R^2 parcial desta regressão. Para testar a existência de correlação entre o termo de erro e o instrumento, será feito um teste de superidentificação: será usada a estatística de teste J de Hansen, que é robusta para heterocedasticidade.

5.1.2. Especificações e questões econométricas

Serão apresentados diferentes especificações de forma a melhor analisar a endogeneidade da renda. Serão utilizadas as mesmas variáveis de controle dos capítulos anteriores para controlar basicamente por efeitos macroeconômicos e características familiares e individuais.

As *dummies* de região, ano e quartil procuram representar o efeito de variáveis agregadas, como as políticas públicas, sobre unidades microeconômicas. Segue-se então sugestão de MOULTON

(1990) de corrigir a matriz de variância-covariância em função da possível existência de correlação entre os erros dentro destes grupos (*clusters*) formados por região, ano e quartil de renda.

A primeira regressão não inclui controle algum e equivale a estimar o efeito 'bruto' da renda sobre os resultados educacionais, isto é, mede-se o efeito da própria renda e de todas as variáveis a ela relacionadas. A segunda regressão inclui *dummies* de ano, de região e de quartil de renda.

O segundo bloco de regressões inclui as características familiares e individuais como controles. Devido à importância já verificada da educação dos pais, primeiro são apresentadas as regressões sem controlar pela educação dos pais, para se medir quanto importante é sua inclusão para corrigir o efeito da renda. A colinearidade entre educação e renda pode aumentar o erro-padrão das estimativas destas variáveis, mas a amostra é grande suficiente para que a precisão das estimativas não seja afetada. A inclusão das características familiares, em especial educação dos pais, pode diminuir a importância das *dummies* de quartil.

Como as variáveis dependentes são variáveis binárias, o uso de um modelo de probabilidade linear (como o método de MQO) não causa viés, entretanto, ele afeta o erro-padrão e o resultado estimado não fica restrito ao intervalo de probabilidade, 0 a 1. Portanto, serão estimados para comparação modelos PROBIT, que usam a função densidade acumulada da normal para transformar o resultado estimado em valores no intervalo 0,1. É bom lembrar que o método PROBIT não resolve os problemas de erro de medida e de viés de variável omitida.

O método de variáveis instrumentais se aplica também a modelos não-lineares, como o PROBIT. Porém não é possível usar o método em 2 estágios, pois as estimativas só seriam consistentes se a função fosse linear na variável endógena². Na falta de programa estatístico pronto para o cálculo pelo método de

² DAVIDSON e MACKINNON (1993), capítulo 7.

variável instrumental de um modelo não linear, só será utilizado o método de variáveis instrumentais para o modelo linear.

As estimativas serão apresentadas para o período de 1981 a 1999 e para três subperíodos: 1981 a 1986, 1987 a 1992 e 1993 a 1999. As estimativas apresentadas no capítulo 4 já mostravam a variação da magnitude do efeito-renda no tempo. Cada período foi definido de forma a conter número parecido de anos, e conter, cada um, uma variação importante da distribuição de renda, de forma a garantir as propriedades do instrumento dentro de cada período³. O primeiro período inclui o Plano Cruzado, que leva a uma estabilização temporária da moeda, acompanhada de uma grande recuperação econômica. O segundo período equivale ao período de mais altas taxas de inflação. O terceiro inclui o Plano Real, que trouxe a efetiva estabilização da moeda e com ela uma melhora real da renda dos mais pobres.

Considerando que o viés de variável omitida seja a questão mais importante relacionada à estimativa por MQO, é possível prever a direção do viés gerado. Se a variável omitida tiver relação positiva com a renda familiar e com o resultado educacional como se supõe, o viés de MQO é para cima, isto é, o efeito marginal estimado por MQO é maior do que o efeito verdadeiro. Por outro lado, a presença de erro de medida na variável sempre causa um viés de atenuação no coeficiente estimado. Portanto, pode-se ter que as duas correções em direções contrárias se compensem, não sendo possível saber de antemão para que lado irá a estimativa IV.

Por fim, é bom ter em mente qual o problema gerado se as hipóteses para o uso do instrumento não forem válidas. BOUND et *alli* (1995) comentam os problemas de inconsistência e de viés de amostra finita. Uma correlação entre o instrumento e o termo de erro pode levar a uma estimativa IV mais inconsistente que a de MQO. Mesmo que não haja correlação entre o instrumento e o termo

³ Por isso foi feita uma pequena alteração dos períodos em relação aos períodos utilizados no capítulo 4.

de erro, com amostra finita (em que haveria alguma correlação do instrumento com o erro, mesmo que assintoticamente ela fosse zero), a estimativa IV apresentará viés na direção da esperança da estimativa de MQO, e o viés torna-se pior se a correlação do instrumento com a renda for fraca e se a amostra for pequena. O uso de maior número de instrumentos pode tornar o estimador de IV mais eficiente (mais correlacionado com a renda), porém também tende a piorar o viés de amostra finita.

5.1.3. Efeitos de grupo e efeitos heterogêneos

A interpretação das estimativas IV deve ser feita com cuidado na presença de efeito da renda média dos grupos e de efeitos marginais heterogêneos da renda entre indivíduos.

Na presença de efeito da renda média dos grupos em adição ao efeito da renda individual, a estimativa com instrumento será maior que a estimativa de MQO. Como o instrumento basicamente reduz a informação da renda às médias por grupos formados por região-ano-quartil, se a renda média de cada grupo tiver um efeito próprio sobre a educação, esta variável será confundida com a renda individual, e seus efeitos serão somados na estimativa com instrumento. BARROS *et alli* (2001) estimam um efeito positivo, mas não-significativo, da renda *per capita* municipal sobre o número de anos completos de estudo; o efeito é de cerca de 1/3 do efeito da renda domiciliar.

Na presença de efeitos marginais heterogêneos, a estimativa de IV mede na verdade uma média ponderada dos efeitos individuais, conhecida como LATE (*local average treatment effect*⁴). O peso da ponderação refere-se ao incremento relativo na renda induzido pelo instrumento, isto é, indivíduos com suas rendas mais afetadas pela variação na distribuição de renda

⁴ Efeito tratamento médio local. O termo LATE é uma alusão ao termo ATE, efeito tratamento médio, que seria o resultado de MQO no caso de efeitos heterogêneos, o acréscimo 'local' enfatiza que a média estimada por IV será relativa ao grupo mais afetado pelo tratamento.

receberão maior peso para o cálculo da estimativa IV. Uma hipótese necessária para esta interpretação é a monotonicidade, como mostram ANGRIST e IMBENS (1995), isto é, a hipótese de que a renda de todas as observações seja afetada na mesma direção pelo instrumento. No caso de CARVALHO (2001), por exemplo, é claro quem é o grupo mais afetado pelo instrumento (reforma previdenciária): famílias no meio rural com a presença de idosos e crianças têm um aumento de renda, portanto, na presença de efeitos heterogêneos, a estimativa estará ponderando mais este grupo. A variação da distribuição de renda utilizada neste trabalho afeta a renda de diferentes grupos de formas diferentes no tempo, o método IV, portanto, mesmo na presença de efeitos heterogêneos, deve gerar uma estimativa média que valha para uma população mais abrangente que a estimativa de CARVALHO (2001).

5.2. Evolução da renda e da educação

Como vimos anteriormente⁵, o aumento de frequência escolar e diminuição de atraso escolar evoluíram de formas diferentes entre regiões e quartis de renda. A região NE praticamente alcança as outras regiões ao longo do tempo em relação à frequência. A frequência escolar dos quartis mais pobres cresce mais rápido que a dos quartis mais ricos, mesmo porque os quartis mais ricos já estavam bem próximos da universalização do ensino. Todas as regiões aumentam também a probabilidade de estar na série ideal, porém os níveis se mantêm bem diferentes entre regiões durante todo o período. Finalmente, a probabilidade de estar na série ideal cresce mais para os quartis mais pobres, porém este efeito é resultado apenas das regiões SE, S e CO, que mais que dobraram seus indicadores. Resta saber se esta variação diferenciada dos indicadores educacionais no tempo foi função de uma variação também diferenciada da renda familiar.

⁵ Tabelas 7 e 8 do capítulo 1 (indicadores educacionais por região e ano) e tabelas 6 e 7 do capítulo 4 (indicadores por quartil e região).

A tabela 1 resume a evolução da renda⁶ aberta por quartil e região para os três períodos que serão usados nas estimativas, 1981 a 1986, 1987 a 1992, e 1993 a 1999.

Tabela 1. Renda familiar per capita média, R\$ de 1999

1981-1986					
região/qua	1	2	3	4	Total
SE	34.92	69.11	119.21	257.97	124.42
N	36.65	67.73	111.70	240.49	114.37
NE	23.75	39.66	64.23	160.56	66.10
S	36.46	71.10	118.80	252.29	110.32
CO	32.77	60.67	101.59	239.57	105.18
Total	31.25	59.02	100.77	228.12	101.96
1987-1992					
região/qua	1	2	3	4	Total
SE	33.83	71.68	128.38	293.83	138.54
N	32.18	64.13	112.59	267.51	116.94
NE	20.69	36.56	63.10	180.76	69.65
S	35.94	76.18	131.64	298.24	125.94
CO	30.81	61.34	108.31	273.08	117.18
Total	29.44	59.85	107.03	260.63	112.71
1993-1999					
região/qua	1	2	3	4	Total
SE	43.93	91.38	159.53	355.57	170.26
N	34.15	65.54	115.31	297.08	124.91
NE	26.41	46.14	78.82	228.30	87.69
S	46.09	96.64	167.86	368.43	163.62
CO	37.95	75.04	131.26	326.18	140.17
Total	37.35	75.16	132.75	318.06	139.41

Nota. Famílias com renda positiva, entre os 5º e 95º percentis da distribuição de renda.

Do primeiro para o segundo período caiu a renda média dos quartis mais pobres para todas as regiões; o segundo quartil mais pobre manteve sua média de renda; o maior aumento de renda neste período foi do quartil mais rico. Todos os quartis em média recuperaram-se na década de 90. Comparando-se, porém, os períodos extremos, o quartil mais pobre apresenta um aumento de renda de apenas 20%, enquanto o quartil mais rico aumentou 39% o valor de sua renda. Há diferenças também dentro dos quartis: os quartis mais pobres da região N não recuperaram a renda média do 1º

⁶ O anexo A4 traz a evolução da renda ano a ano por quartil (tabela A1) e por região (tabela A2).

período e o quartil mais pobre da região NE aumentou a renda em apenas 10%. Estas diferenças podem ser em grande parte explicadas pela incapacidade dos mais pobres de se defender da alta inflação durante a década de 80 e início de 90. Além disso, os vários planos de estabilização tiveram diferentes efeitos distributivos, seja em função de suas regras de correção inflacionária, seja em função de terem causado recessão ou recuperação econômica. O diferente desenvolvimento econômico das regiões também ajuda a explicar tais diferenças.

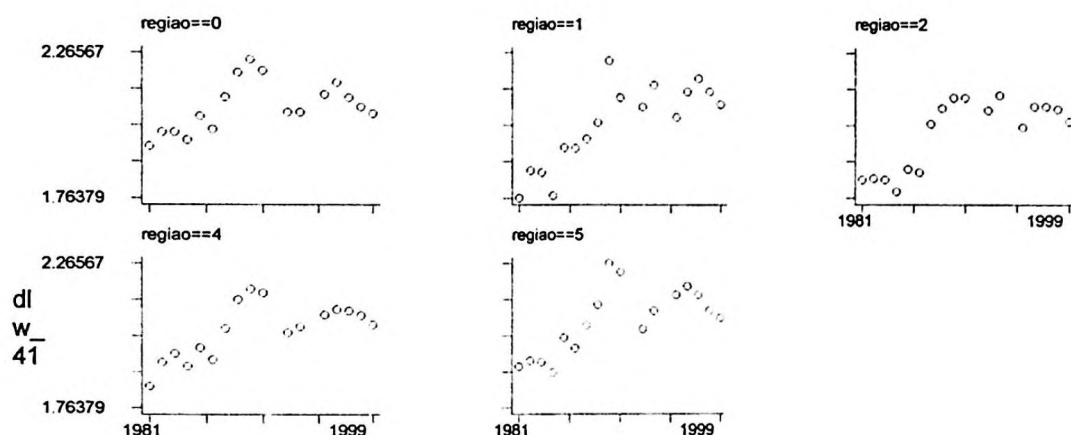
À primeira vista, as trajetórias de frequência e atraso escolar comparadas à trajetória da renda familiar não apresentam um padrão similar. Os resultados educacionais crescem de forma constante ao longo do tempo, enquanto a renda oscila bastante, apesar também de ter crescido se comparados os anos extremos. Estas trajetórias, contudo, incluem não só o efeito advindo da relação destas variáveis, mas podem estar relacionadas a outros fatores. As políticas educacionais da década de 90, como vimos, com certeza influenciaram a evolução dos resultados educacionais no período, independentemente do efeito da renda familiar.

A seguir pretende-se mostrar de forma mais clara a relação entre as variações na frequência escolar (e em estar na série ideal) e na desigualdade de renda. Pode-se falar em desigualdade de renda, dado que são as variações entre quartis que estão sendo consideradas.

Será utilizado o procedimento de diferenças em diferenças, que compara a variação dos grupos no tempo para ver se esta ocorreu de forma diferenciada entre grupos. Serão construídas medidas de desigualdade baseadas nos quartis, e a evolução no tempo desta desigualdade será apresentada por região. A desigualdade da educação será calculada como a diferença da proporção de crianças frequentando a escola (ou na série ideal) entre os quartis das pontas da distribuição. A desigualdade de renda será medida pelo logaritmo da razão entre as rendas médias dos quarto e primeiro quartis para cada região e ano.

A evolução da desigualdade de renda no tempo por região é mostrada no gráfico 1. O SE é a região que apresenta menor crescimento da desigualdade no tempo; N e NE, apesar de começarem com os níveis de desigualdade mais baixos, aumentam sua desigualdade, especialmente nos anos 80. No final da década de 80, as mais altas taxas de inflação registradas no período podem explicar a piora geral nas distribuições de renda. Durante a década de 90, a desigualdade mantém-se relativamente constante, com alguma melhora nos últimos anos.

Gráfico 1. Desigualdade da renda (ln) entre quartis 4 e 1 no tempo por regiões



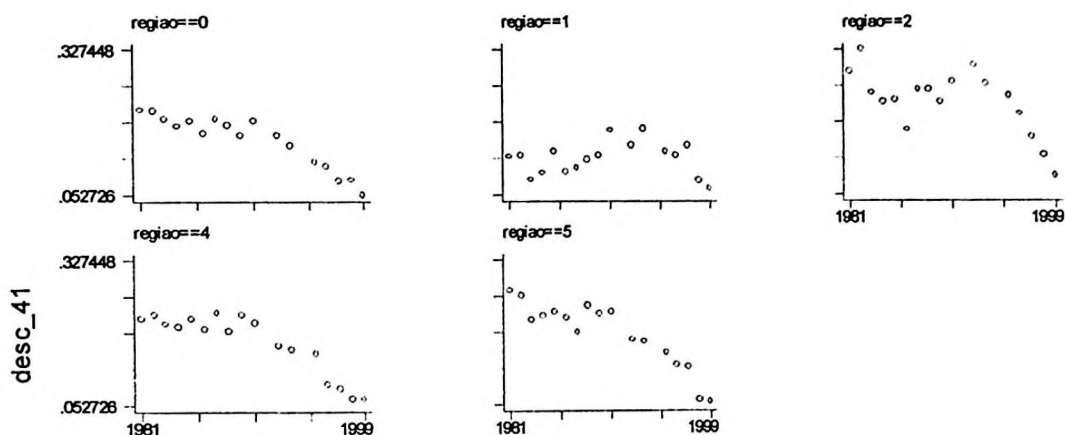
Nota. 0 - SE, 1 - N, 2 - NE, 4 - S, 5 - CO.

A comparação das trajetórias da frequência escolar e da probabilidade de se estar na série ideal por regiões é mostrada nos gráficos 2 e 3.

As regiões SE e CO apresentam uma queda constante na diferença de frequência escolar entre quartis, enquanto para as outras regiões a diminuição da diferença entre quartis começa apenas na década de 90. O mais alto nível de diferença observada refere-se aos primeiros anos de 80 no NE e CO. Em relação a estar na série ideal, as regiões SE e S apresentam uma diminuição na

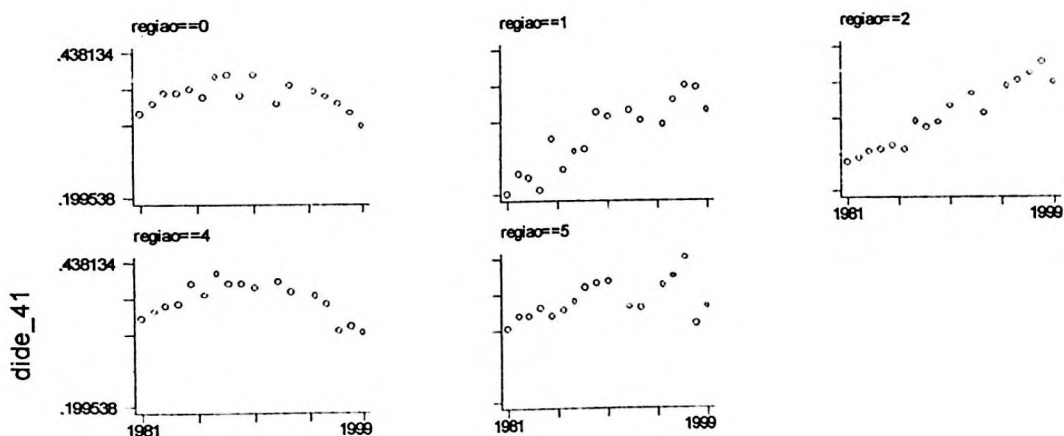
distância entre quartis na década de 90, basicamente voltando ao nível do início do período. Esta distância aumenta nas regiões N e NE durante todo o período. É interessante lembrar que a probabilidade média de estar na série ideal melhora continuamente no tempo, portanto a piora da desigualdade entre quartis indica que a melhora ocorreu predominantemente para o quartil mais rico.

Gráfico 2. Diferença da probabilidade de frequência escolar entre quartis 4 e 1



Nota. 0 - SE, 1 - N, 2 - NE, 4 - S, 5 - CO.

Gráfico 3. Diferença de probabilidade de estar na série ideal entre quartis 4 e 1



Nota. 0 - SE, 1 - N, 2 - NE, 4 - S, 5 - CO.

A piora da distribuição de renda nos anos 80 coincide com a constância na diferença da frequência escolar, enquanto a manutenção e queda da desigualdade de renda nos anos 90 está relacionada à diminuição na desigualdade da frequência escolar entre quartis; tais variações geram uma relação positiva entre renda e escola. Não é tão clara esta relação entre renda e a probabilidade de estar na série ideal pois, apesar da desigualdade de renda se manter relativamente constante ou cair na década de 90 para todas as regiões, a diferença de atraso escolar continua a crescer para as regiões N e NE.

Olhar outras medidas de distribuição pode ser interessante, de forma a saber de que parte da distribuição vem a variação da desigualdade. No anexo A4 são apresentados os mesmos gráficos comparando-se as médias dos indicadores dos quartis 4 e 1 (por região e ano) com a mediana da renda e com a média dos indicadores educacionais. Desta forma tem-se a desigualdade de renda dos mais ricos e dos mais pobres, ambos em relação ao indivíduo mediano (ou médio no caso da educação). Por exemplo, o aumento da desigualdade de renda nas regiões N e NE ocorre basicamente entre os extratos mais ricos da população (comparação entre média do quartil mais rico e mediana da população). A queda constante da desigualdade de frequência escolar só ocorre entre os extratos mais ricos das regiões SE e CO, e a queda na diferença de atraso só ocorre entre os extratos mais ricos das regiões SE e S.

Para confirmar a existência de variação diferenciada de renda estima-se uma regressão da desigualdade da renda em função de uma tendência temporal criada por região, controlando-se por *dummies* de ano e de região, e observando se as estimativas destas tendências diferem entre regiões.

A tabela 2 mostra os coeficientes das tendências por região. A variável dependente é a desigualdade de renda entre os quarto e primeiro quartis, e as variáveis de tendência foram construídas

pela interação das *dummies* de região e uma variável categórica, representando os anos, que varia de 1 a 17.

Comparadas à região SE, todas as regiões apresentam estimativas diferentes, sendo os coeficientes do NE e N, e do S e CO semelhantes entre si. O teste de Wald apresentado na tabela rejeita a hipótese de que todos os coeficientes são iguais.

Tabela 2. Tendência regional da desigualdade de renda

	Coef.	E.padrão	teste
SE	.005	.002	F(4, 60) = 12,84 p-value = 0,00
N	.018	.002	
NE	.015	.002	
S	.009	.002	
CO	.011	.002	

dummies ano e região incluídas

Nota. desigualdade = log da razão das renda dos quartis 4 e 1; tendência = variável categórica por região: de 1 a 17 para os anos de 1981 a 1999, e 0 se outra região; teste F: se coeficientes das tendências iguais entre si.

5.3. Estimativas

Foram utilizadas diferentes especificações para determinar o efeito marginal da renda familiar sobre a educação. As especificações usam os seguintes controles, todos em forma de *dummies*, além de renda familiar (seu logaritmo natural):

1: nenhum controle

2: ano, região, quartil de renda

3: sexo, área rural/urbana, idade, número de filhos, raça, viver com um só pai, ano, região, quartil

4: educação do pai, educação da mãe, sexo, área rural/urbana, idade, número de filhos, raça, viver com um só pai, ano, região, quartil

As mesmas especificações são estimadas⁷ por MQO e PROBIT para o período todo, e também são apresentadas as estimativas de

⁷ Os pesos da PNAD não foram utilizados nas estimativas deste capítulo, não há, porém, mudança significativa nos coeficientes estimados em relação aos resultados apresentados no capítulo 3.

MQO para cada um dos três períodos escolhidos, 1981 a 1986, 1987 a 1992, e 1993 a 1999. O procedimento de variáveis instrumentais e os testes são apresentados para a especificação 4.

Tabela 3. Efeito marginal da renda familiar sobre frequência escolar

especificação	MQO				PROBIT ^a
	total	81-86	87-92	93-99	total
1	.100 (.004)	.111 (.006)	.103 (.005)	.068 (.006)	.104 (.004)
2	.067 (.003)	.089 (.005)	.070 (.004)	.043 (.003)	.084 (.003)
3	.043 (.003)	.058 (.005)	.041 (.004)	.030 (.003)	.059 (.003)
4	.020 (.003)	.024 (.004)	.016 (.004)	.014 (.003)	.027 (.002)
sem cluster	(.001)	(.002)	(.003)	(.002)	-
IV - 4 ^b	-.027 (.029)	-.055 (.056)	-.003 (.031)	.003 (.062)	-
teste F ^c	173.89 [.0]	141.80 [.0]	124.18 [.0]	97.75 [.0]	
R ² parcial ^d	.070	.033	.044	.036	
Hansen ^e	219.835 [.768]	36.570 [.708]	24.780 [.531]	43.106 [.635]	
N. obs. ^f	959264	441380	236235	281649	959250

Nota:

erro-padrão entre parênteses; utilizado estimador de variância robusto para cluster, sendo o cluster definido por ano-região-quartil; p-value entre colchetes; não foram utilizados os pesos amostrais da PNAD.

a - os efeitos marginais do PROBIT são calculados para os valores médios das variáveis explicativas; b - o coeficiente IV estimado para a amostra total não inclui o ano de 1981 (por problemas computacionais em função do tamanho da amostra); c - significância conjunta dos instrumentos excluídos na regressão de 1º estágio; d - R² parcial relativo ao 1º estágio; e - teste de superidentificação; f - PROBIT não utiliza observações que prevêm perfeitamente o resultado, daí diferença de número de observações entre MQO e PROBIT.

A tabela 3 resume os resultados para frequência escolar. Entre especificações, MQO e PROBIT seguem o mesmo padrão: o efeito marginal da renda diminui a cada inclusão de controles adicionais: o efeito bruto da renda é de 0.10, e cai para 0.02

com todos os controles. Apenas a inclusão da educação dos pais faz cair à metade o coeficiente da renda: continua significativo, mas com maior erro-padrão.

A título de exemplo, são apresentados os erros-padrão da especificação 4 sem a correção por *cluster*. Como esperado, aumenta o erro-padrão quando é feita a correção por *cluster*, as estimativas, porém, continuam significativas.

Entre especificações diminui o efeito das *dummies* de quartil de renda em função da introdução de características familiares, mas as *dummies* dos dois quartis mais ricos continuam significativas (em relação ao quartil mais pobre), indicando que elas ainda estão servindo de *proxy* para variáveis não-observadas.

Na abertura por períodos aparece o mesmo padrão de queda do efeito da renda entre especificações, além disso, o efeito da renda cai no tempo em todas as especificações. Vimos no capítulo 4 que não só a renda, mas a maioria das características familiares perdeu poder explicativo no tempo, fato relacionado a alta proporção de crianças na escola na final da década de 90.

O teste F da significância do instrumento (excluído) é sempre significativo, e o R^2 parcial, apesar de relativamente baixo, confirma a existência de correlação entre o instrumento e a renda. O teste de superidentificação indica não haver correlação entre o instrumento e o termo de erro.

O resultado com o método de variáveis instrumentais mostra que a renda familiar não tem nenhum efeito causal sobre a frequência escolar, indicando que variáveis não-observadas, como habilidades e preferências, foram responsáveis pela estimativa significativa de MQO. O uso do instrumento, como já discutido, também ajuda a resolver o viés de atenuação relacionado ao erro de medida da renda. Há razões para se acreditar que o erro de medida é alto, e que este tipo de instrumento (que praticamente reduz a informação da renda a médias de grandes grupos) ajuda a minorar o problema de erro de medida. Este ajuste levaria a um

aumento da estimativa, o que deixa mais robusto o resultado não significativo estimado.

Este resultado refere-se a crianças de 7 a 14 anos, portanto em idade escolar, quando o custo de oportunidade de estar na escola é baixo e a maioria das escolas é pública. É possível, entretanto, que a renda familiar tenha influência sobre a frequência de jovens em idade de estarem freqüentando o ensino médio ou superior, pois aumenta o custo de oportunidade de não se trabalhar e as faculdades são em grande parte pagas.

O fato de não se ter encontrado um efeito-renda significativo para a amostra total não impede que a renda seja fator restritivo para famílias mais carentes. A renda pode afetar de forma diferenciada famílias com maior ou menor condição de financiar a educação dos filhos. No caso da existência de efeitos heterogêneos, o efeito estimado equivale ao efeito médio da população, é possível, portanto, que exista algum subgrupo para o qual o efeito-renda ainda é significativo.

Já vimos que o efeito marginal varia entre os períodos analisados, é possível que este efeito também varie por nível de renda dentro de cada período. Para verificar a possibilidade de diferentes efeitos marginais por grupos de renda, estimei uma regressão de MQO com 4 variáveis renda, criadas pela interação da renda com as *dummies* de quartil. Desta forma, supõe-se que o efeito das variáveis de controle é igual para todos os grupos de renda e que só o efeito-renda varia por grupo de renda. A tabela 4 apresenta os resultados estimados por MQO e no canto da tabela são apresentados os efeitos marginais estimados por PROBIT. No caso do PROBIT foi usada a mesma especificação da tabela 3, e os efeitos foram calculados para as médias das variáveis explicativas de cada quartil de renda.

Os coeficientes estimados da renda diminuem quanto mais rico o quartil de renda, e o efeito diminui no tempo para os quartis mais ricos. Ele nunca é significativo para o quartil mais rico, e

deixa de ser significativo no último período também para os quartis 2 e 3. O quartil mais pobre não só sempre apresenta efeito significativo, como seu valor cresce no último período. As estimativas PROBIT confirmam a maior importância da renda para os mais pobres.

Tabela 4. Regressões considerando efeito-renda diferenciado por quartil - frequência escolar

especificação	MQO				PROBIT
	total	81-86	87-92	93-99	total
qua 1	.049 (.006)	.037 (.008)	.036 (.009)	.047 (.012)	.042
qua 2	.043 (.008)	.045 (.013)	.035 (.010)	.025 (.015)	.034
qua 3	.022 (.007)	.032 (.011)	.022 (.008)	.006 (.011)	.024
qua 4	-.002 (.004)	.010 (.006)	-.004 (.004)	-.006 (.005)	.011
N. obs.	959264	441380	236235	281649	959250

Nota. estimativas referentes à especificação 4; erro-padrão entre parênteses; peso amostral da PNAD não utilizado; MQO - consideradas 4 variáveis renda, uma para cada quartil, com valor 0 caso contrário; PROBIT - 1 variável renda, efeitos marginais calculados nas médias de cada quartil; diferença de número de observações entre MQO e PROBIT em função de PROBIT não usar observações que prevêm o resultado perfeitamente.

Como as estimativas não foram calculadas com variável instrumental, seria possível que o efeito diferenciado encontrado fosse em função de variáveis não-observadas e não da renda. Infelizmente a estratégia utilizada aqui para isolar o efeito da renda necessita da variação entre quartis, que, por construção, não existe nas variáveis renda criadas por quartil. De qualquer forma a estimativa para o quartil mais pobre no último período é de 0,047, bem maior que a estimativa média por MQO, de 0,014.

Agora voltemos nossa atenção para o atraso escolar. A magnitude dos efeitos marginais da renda referentes a estar na série ideal é bem mais alta que dos efeitos referentes à

frequência escolar, Como já constatado no capítulo 3. As estimativas caem com a inclusão de mais controles, de 0.195 sem nenhum controle para 0.070, mas não é clara a queda do efeito-renda no tempo, como mostra a tabela 5.

Tabela 5. Efeito marginal da renda familiar sobre estar na série ideal

especificação	MQO				PROBIT ^a
	total	81-86	87-92	93-99	total
1	.195 (.005)	.176 (.009)	.188 (.008)	.202 (.008)	.205 (.006)
2	.188 (.006)	.199 (.011)	.176 (.010)	.168 (.008)	.195 (.005)
3	.151 (.005)	.164 (.010)	.134 (.009)	.137 (.008)	.179 (.006)
4	.070 (.003)	.078 (.005)	.058 (.004)	.062 (.004)	.085 (.003)
IV - 4 ^b	.109 (.021)	.144 (.029)	.096 (.031)	.214 (.025)	-
teste F ^c	173.89 [.0]	141.80 [.0]	124.18 [.0]	97.75 [.0]	
R ² parcial ^d	.070	.033	.044	.036	
Hansen ^e	218.134 [.792]	31.337 [.886]	21.595 [.711]	47.797 [.440]	
N. obs. ^f	959264	441380	236235	281649	959219

Nota. mesmas observações da tabela 3.

Em relação às variáveis não-observadas, estar na série ideal envolve mais fatores que em relação à simples frequência. Se a habilidade dos pais é importante para ajudar a criança nas suas tarefas escolares, ela está relacionada com o desempenho escolar, que afetará a possibilidade de repetência e, só mais tarde, a possibilidade de evasão. A política educacional (inclusive política de aprovação) e a qualidade da escola (e dos professores) importam também de forma mais direta para o atraso do que para a frequência escolar. O atraso escolar inclui não só alunos repetentes e evadidos, mas também alunos que entraram

atrasados na escola. Portanto, a relação da renda com estar na série ideal pode ocorrer em função da possibilidade da família garantir uma vaga escolar para seu filho na idade correta, fato que pode estar relacionado diretamente com a renda da família como também com a oferta de vagas da comunidade.

O resultado das estimativas IV para atraso é bem distinto dos estimados para frequência: aumenta a magnitude do efeito-renda e apesar do aumento do erro-padrão, as estimativas são significativas. O aumento das estimativas poderia estar relacionado à correção de erro de medida, porém as estimativas para frequência mostraram que esta correção não é tão importante.

Além disso, pode haver variáveis não-observadas que afetam estar na série ideal e renda familiar em sentido contrário, e portanto poderiam causar um viés para baixo nas estimativas com MQO. É de se esperar, porém, que habilidades familiares e qualidade da escola afetem educação e renda no mesmo sentido, isto é, tenham relação positiva com renda familiar e educação. A oferta de trabalho da criança pode gerar efeitos distintos sobre renda familiar e educação: em períodos de 'boom' econômico, o aumento de oferta de trabalho, e também de renda, deve elevar a renda familiar e ao mesmo tempo aumentar a probabilidade da criança trabalhar, diminuindo sua dedicação à escola, o que poderia levar à repetência. Este argumento, porém, parece mais importante para crianças mais velhas.

Por fim, não podemos deixar de considerar a hipótese de que o instrumento não seja bom para tratar renda no caso da variável dependente ser atraso escolar. Como mais fatores não-observados são importantes, algum deles pode estar gerando correlação do termo de erro com o instrumento, apesar do teste de superidentificação não indicar tal correlação.

O efeito da renda dos grupos pode ser importante para explicar o resultado com instrumento. Como já explicado, este método de IV soma ao efeito da renda familiar o efeito da renda

média do grupo. Se a renda média do grupo estiver relacionada com a qualidade da escola, é possível que seu efeito seja importante para explicar o atraso escolar.

A qualidade da escola e a política de aprovação podem ter variado de forma diferenciada entre região e quartil de renda no tempo. A política de progressão continuada (não-repetência) nas escolas, por exemplo, não foi projeto nacional como foi a ênfase na universalização do ensino na década de 90, o que pode contribuir para estes resultados. Neste caso, o problema estaria mais relacionado ao último período, de 1993 a 1999, que é justamente o período que apresenta estimativa IV mais diferenciada.

Estes resultados indicam a importância de se incluir na análise do atraso escolar variáveis relacionadas à escola, ou pelo menos à comunidade em que estas se encontram. Como visto, BARROS *et alli* (2001) encontram efeito positivo e significativo para a qualidade da educação de 1º grau, medida pela oferta de escola e qualificação dos professores, e encontram efeito positivo, apesar de não-significativo, para a renda média das comunidades.

A tabela 6 mostra os efeitos marginais das rendas médias por quartil. Em relação a estar na série ideal, o efeito estimado é bem maior para os quartis mais ricos, relação que se mantém no tempo. A distinção ocorre entre o quartil mais pobre e os outros, o que pode sugerir que para os mais pobres importa estar na escola, e que a preocupação com desempenho começaria apenas a partir de determinado nível de renda. Como já explicado, não foi possível usar o método de variáveis instrumentais com a abertura da renda, e portanto não se separou o efeito das rendas de variáveis não-observadas.

Tabela 6. Regressões considerando efeito-renda diferenciado por quartil - série ideal

especificação	MQO				PROBIT
	total	81-86	87-92	93-99	total
qua 1	.032 (.004)	.026 (.006)	.015 (.006)	.036 (.005)	.056
qua 2	.069 (.008)	.050 (.006)	.049 (.008)	.076 (.008)	.073
qua 3	.094 (.006)	.083 (.006)	.085 (.008)	.096 (.007)	.089
qua 4	.084 (.003)	.110 (.005)	.078 (.005)	.062 (.005)	.094
N. obs.	959264	441380	236235	281649	959219

Nota. mesmas observações da tabela 4.

Vimos, portanto, que a renda familiar não tem influência direta sobre a frequência escolar de crianças de 7 a 14 anos. Não foi, entretanto, descartada a possibilidade de que a renda importe para parcelas mais pobres da população.

Ao contrário da relação com frequência, estimou-se um efeito marginal da renda sobre estar na série ideal significativo. Mesmo que as estimativas IV estejam comprometidas por não garantirem a independência entre o instrumento e o termo de erro, as estimativas por MQO indicam um efeito marginal da renda maior sobre atraso que sobre frequência escolar, o que sugere não só a maior influência de habilidades e preferências familiares, mas também a importância para a análise de outras variáveis não-observadas, como políticas de aprovação e qualidade da escola.

COMENTÁRIOS FINAIS

Apresentou-se a evolução da educação no Brasil no período de 1981 a 1999, principalmente em relação à frequência e atraso escolar. Foram analisados os principais determinantes desta evolução, entre eles, as políticas públicas relacionadas à educação e características familiares.

Mostrou-se que a evolução dos indicadores educacionais no tempo é tanto função das políticas e características familiares e demográficas, como também da mudança do impacto destes fatores sobre a educação. A composição das características familiares varia no sentido de melhorar os indicadores educacionais. Por outro lado, a maioria destes características, entre elas a renda familiar, diminui seu efeito sobre os indicadores educacionais, o que indica a diminuição de sua importância, principalmente para explicar a frequência escolar.

Por fim, analisou-se especificamente o efeito da renda familiar sobre os indicadores educacionais. Para tanto, utilizou-se o método de variáveis instrumentais, baseando-se na variação da renda no tempo entre quartis de renda e regiões. Mostrou-se que a renda não tem efeito direto sobre a frequência escolar, mas as estimativas indicam que ela tem influência sobre o atraso escolar, fato que merece análise mais pormenorizada no futuro.

A não-existência de efeito-renda sobre a frequência escolar não tira o mérito de programas de renda mínima como o bolsa-escola. Os dois objetivos destes programas são importantes de serem atingidos, a redistribuição de renda e a permanência das crianças na escola. Se não há relação causal, é importante manter a compulsoriedade da frequência escolar, e sua fiscalização, como condição para recebimento da ajuda financeira. Por outro lado, os

resultados encontrados indicam que é possível que a renda importe para a decisão de frequência escolar para famílias mais pobres.

A influência da renda familiar sobre o atraso escolar encontrada sugere que outras variáveis importam para determinar o atraso escolar, além das características familiares. Foi sugerido que a qualidade da escola, relacionada com a renda familiar e com a renda da região de residência, pode ser importante para explicar os resultados encontrados.

anexo A1: DESCRIÇÃO DOS DADOS DA PNAD

O banco de dados utilizado é a Pesquisa Nacional por Amostra de Domicílios - PNAD, realizada anualmente pelo Instituto Brasileiro de Geografia e Estatística - IBGE. Os dados são coletados todo mês de setembro. Não há pesquisa em anos de realização de censos demográficos, 1991 no período de interesse, e não houve pesquisa também em 1994. A abrangência geográfica atual da pesquisa é usada desde 1981, por isso escolheu-se trabalhar com os dados a partir deste ano, e até 1999. A amostra da PNAD inclui todas as regiões do Brasil, inclusive as áreas rurais. A exceção é a parte rural da região Norte, mas ela representa apenas 2,3% da população do país¹, portanto sua omissão não deve afetar a análise.

As grandes vantagens em se utilizar a PNAD para analisar a educação são sua abrangência temporal e geográfica. Outras pesquisas, como a Pesquisa Mensal de Emprego - PME ou a Pesquisa sobre Padrões de Vida - PPV², trariam outras vantagens para a pesquisa, a primeira por ter um painel, isto é, ter informação de 2 anos por domicílio, a segunda por trazer mais informações relativas a educação. A PME, porém, restringe-se a 6 regiões metropolitanas, que dificilmente podem representar o resto do país no que tange à trajetória da educação. Já a PPV foi realizada apenas no ano de 1996 e restringe sua pesquisa às regiões Nordeste e Sudeste.

A PNAD é uma pesquisa domiciliar, e como tal restringe suas perguntas aos moradores do domicílio. Portanto se algum parente não morar no domicílio, não é possível obter informação sobre

¹ Dado do Censo demográfico de 2000 realizado pelo IBGE.

² Ambas pesquisas também realizadas pelo IBGE.

ele. De especial relevância para este trabalho, não há informação sobre os pais que não moram no domicílio.

Abaixo a apresentação de algumas variáveis utilizadas.

A educação refere-se ao último nível educacional cursado ou em andamento. As perguntas a respeito de educação na PNAD são feitas em duas etapas. Primeiro questiona-se se o indivíduo frequenta algum tipo de estabelecimento de ensino, se frequentar, responderá perguntas a este respeito. Se não frequentar, será perguntado se já frequentou algum estabelecimento de ensino, em caso afirmativo, será perguntado sobre o último frequentado. Mais especificamente, o indivíduo que frequenta estabelecimento de ensino responde qual tipo de ensino é: escola de 1º ou 2º grau (regular ou supletivo), ensino superior ou pós-graduação, e outros cursos: pré-escola ou creche, alfabetização de adultos, curso pré-vestibular. No caso em que se aplica, é perguntada a série do curso que está sendo frequentada. Para o indivíduo que não frequenta, mas já frequentou, é perguntado qual o último curso e série frequentado e completado.

De 1981 a 1990 não é possível identificar pessoas que evadiram sem completar nenhum ano de estudo, eles estão alocados junto com os que nunca frequentaram. Já a partir de 1992 é possível identificar estes casos. Este fato, porém, não afeta os indicadores de educação aqui estudados, frequência e estar na série ideal.

Nos anos analisados foram usados diferentes critérios de classificação para crianças em idade escolar, 7 anos ou mais, que responderam estar frequentando pré-escola ou creche. Em 1982 elas foram alocadas na categoria alfabetização de adultos; neste caso, não houve problema em identificá-las. Em 1981 e de 1983 a 1986, porém, elas são alocadas na 1ª série do 1º grau; isto causa um aumento incorreto na proporção de crianças na 1ª série, e não há informação disponível para distinguir estas crianças das que realmente responderam estar na 1ª série do 1º grau. A partir de

1987 o critério do IBGE permitiu crianças até 9 anos de idade na classificação pré-escola, portanto de 1987 a 1990, crianças a partir de 10 anos, que responderam estar na pré-escola, estão alocadas na 1ª série. E a partir de 1992, novo critério passou a realocar crianças para a 1ª série apenas a partir dos 12 anos de idade.

Houve também uma reclassificação das crianças de 5 anos de idade, idade a partir da qual a PNAD questiona sobre educação. Nos anos de 1981 a 1986 as crianças nesta idade que já estavam cursando o ensino fundamental foram reclassificadas para pré-escola. Como esta idade não é importante para a análise deste trabalho, simplesmente definiu-se a amostra a partir de 6 anos de idade.

A PNAD pesquisa todos os tipos de renda de todos os membros da família. A renda refere-se à semana de referência da entrevista, em setembro: renda mensal de todos os trabalhos (se mais de um), pensões, ganhos de juros, entre outros. Para o período todo utilizado neste trabalho, é pesquisada a renda de pessoas de 10 anos ou mais de idade.

Não há informação de raça nas PNADs de 1981 e 1983, e apenas para parte da amostra em 1984 e 1985. Em 1984, a informação existe apenas para mulheres de 15 a 54 anos; em 1985, para parte das crianças de 0 a 17 anos de 10 Estados.

anexo A2: DEFINIÇÃO DE VARIÁVEIS E DA AMOSTRA

O interesse deste trabalho restringe-se a crianças e jovens em idade escolar, isto é, que estão em idade de cursar escola de 1º e de 2º grau. Considerando a idade de início de escola obrigatória, 7 anos, e o tempo mínimo de escola sendo 11 anos, o universo de análise refere-se basicamente a crianças e jovens de 7 a 17 anos. Algumas tabelas, porém, incluem crianças de 6 anos, pois já existe proporção significativa de crianças nesta idade na escola de 1º grau. As regressões restringem-se à faixa etária de 7 a 14 anos.

Abaixo as explicações para as variáveis construídas a partir dos dados da PNAD.

- coorte/ idade

Para a construção das coortes, utilizou-se um ano de julho a junho. A idade utilizada durante o trabalho refere-se ao mês de junho de cada ano, em função da suposição de que crianças que completam 7 anos no 1º semestre deveriam entrar na escola neste ano. A informação de educação da PNAD refere-se a setembro; considerou-se que a informação dada em setembro não seria diferente da informação de junho, por exemplo, considerou-se que quem informou não frequentar mais escola, já não frequentava 3 meses antes.

A princípio o uso da idade de junho não traz maiores consequências, porém, como alguns critérios de classificação da PNAD são definidos de acordo com a idade (em setembro), algumas informações ficam afetadas. Por exemplo, crianças de 7 anos em pré-escola realocadas para a 1ª série: a idade utilizada será a de junho, portanto 6 anos para os nascidos de julho a setembro, e

portanto as proporções de crianças de 6 anos em pré-escola e 1ª série também ficam comprometidas pela reclassificação da PNAD.

Abaixo, na tabela A1, estão o número de observações da amostra por coorte e idade num total de 1.747.445 observações.

Tabela A1. Número de observações - PNAD

coorte	idade											
	6	7	8	9	10	11	12	13	14	15	16	17
1964												11287
1965											11440	11744
1966										11558	11565	11345
1967									11297	11509	11515	11166
1968								11494	11593	11393	11171	11091
1969							11561	11690	11689	11458	11169	6255
1970						11640	11890	11894	11510	11705	6229	6444
1971					11373	11876	11727	11658	11611	6332	6297	6275
1972				11737	11982	11841	11541	11529	6454	6420	6265	6306
1973			11728	12012	12059	12057	11796	6597	6611	6600	6397	6422
1974		11649	11856	11987	11832	11706	6512	6535	6518	6404	6328	
1975	12164	12477	12099	12297	12141	6778	6837	6758	6684	6867		6584
1976	12717	12639	12277	12400	6919	6884	6856	6840	6867		6529	6504
1977	12684	12579	12350	6984	7003	6974	6989	6741		6820	6828	
1978	13108	13231	7318	7294	7413	7232	7192		7051	7218		7129
1979	13367	7295	7422	7184	7149	7414		7074	7418		7158	7122
1980	7616	7972	7572	7439	7548		7534	7678		7756	7687	7664
1981	7862	7450	7610	7442		7426	7655		7646	7427	7423	7484
1982	7931	7875	7925		7800	7867		8080	8056	8093	7785	7753
1983	7755	7687		7725	7757		7971	7730	7964	7849	7771	
1984	7116		7280	7360		7427	7366	7569	7567	7528		
1985		7135	7023		7309	7192	7381	7374	7436			
1986	7349	7452		7579	7250	7605	7462	7606				
1987	7343		7388	7250	7376	7256	7590					
1988		7293	7039	7531	7346	7481						
1989	7129	7014	7303	7236	7236							
1990	6583	7096	6773	6906								
1991	6849	6787	6693									
1992	6747	6853										
1993	6725											

Nota: calculado usando peso amostral da PNAD.

- educação

Dado o interesse nos ensinos fundamental e médio, não se considera a trajetória educacional do jovem após o 2º grau; aos que já completaram o 2º grau são atribuídos 11 anos completos de educação, independentemente de terem continuado estudando ou não. Apesar de importante e de apresentar aumento expressivo de frequência no período estudado, crianças na pré-escola são consideradas como não frequentando escola.

O cálculo de número de anos completos de estudo foi feito a partir das informações da PNAD. Varia de 0 a 16 anos. 16 anos de estudo foram atribuídos para pessoas que freqüentavam ou completaram pós-graduação e para pessoas que tinham superior com mais de 4 anos. Esta definição segue critério da PPV. Se o curso não era seriado, o critério foi conservador, atribuindo apenas o nível de educação adquirido com certeza: no caso de freqüência de supletivo, se era 1º grau, foram atribuídos 0 anos de estudo; se 2º grau, 8 anos de estudo; classes de alfabetização ficaram com 0 anos de estudo.

- renda familiar

A renda familiar considerada inclui as rendas do chefe, cônjuge, outro parente, agregado e filhos (à exceção da renda do filho em questão). Deixa-se de fora a renda de pensionistas e empregados domésticos e seus parentes, pois o vínculo com a família é apenas comercial e de trabalho.

A renda nominal informada a cada ano foi corrigida para valores de setembro de 1999, usando o Índice nacional de preços ao consumidor - INPC do IBGE.

Não se inclui a renda da criança pois isto traria um viés na estimativa do efeito da renda sobre as variáveis educacionais. O aumento de renda do filho pode estar correlacionado com uma diminuição de freqüência e de se estar na serie ideal, pois está relacionado com a criança passar a trabalhar (caso contrário a renda é zero). Já a renda da família em geral tem correlação positiva com a educação. Somar as 2 rendas, portanto, pode levar a estimativas menores que o efeito da renda da família. Tais estimativas foram testadas e confirmam esta hipótese.

A renda utilizada para as regressões é a renda familiar *per capita*. Como nas estimativas feitas está incluso o número de crianças como variável de controle, na prática, usar renda *per capita* ou não gera pouca diferença.

- raça

As categorias que serão usadas são: branco, pardo, negro, e amarela (aí incluídos indígenas, quando há informação a partir de 1992, seguindo critério da PPV).

- número de filhos

Soma as pessoas cuja condição na família é 'filho' e que possuem 17 anos de idade ou menos, inclusive a criança observada. Varia de 0 a 22 filhos.

- região

Para ter a mesma classificação todos os anos, o estado do Tocantins foi mantido como pertencente ao estado de Goiás, e portanto, pertencente a região Centro-Oeste (e não Norte).

- amostra para estimativas:

Como os estimativas principais do trabalho estão interessadas na relação entre renda familiar e resultado educacional dos filhos, observações sem informação de renda familiar (ou com informação de renda familiar igual a zero) e observações de jovens de 7 a 14 anos que não eram filhos foram desconsideradas para o cálculo das regressões.

Supõe-se que as informações de valores extremos nas caudas da distribuição de renda têm maior erro de medida. Foram testados 3 cortes de outliers: 1% de cada ponta da distribuição, 5% e 10%. As estimativas do efeito da renda aumentam quanto maior o corte. Este resultado parece confirmar o problema de erro de medida, pois este causaria um viés de atenuação nas estimativas. Decidi utilizar um corte de 5% em cada ponta da distribuição.

Abaixo, na tabela A2, é apresentado o número de observações por região e ano da subamostra utilizada para as estimativas.

Tabela A2. Número de observações usadas nas estimativas

ano	região					Total
	SE	N	NE	S	CO	
1981	21138	1576	16173	8758	3705	51349
1982	21154	1627	16181	8563	3869	51394
1983	21653	1706	16387	8416	3970	52133
1984	21766	1800	16939	8463	3931	52900
1985	22922	2031	17098	8562	4163	54777
1986	23513	2072	17830	8634	4332	56380
1987	23938	2204	18100	8654	4226	57122
1988	24721	2210	18379	8924	4388	58622
1989	25571	2331	18911	9328	4528	60671
1990	24511	2569	18945	8994	4511	59529
1992	24951	2702	17023	8703	4556	57935
1993	25123	2866	17529	9000	4672	59190
1995	24610	2858	18475	8809	4686	59437
1996	24082	2867	17775	8739	4611	58075
1997	23461	2793	17593	8768	4729	57344
1998	23212	2817	17119	8539	4566	56254
1999	23039	2773	17327	8510	4505	56154
Total	399365	39802	297785	148364	73948	959264

Nota: calculado usando peso amostral da PNAD.

anexo A3: OUTROS INDICADORES EDUCACIONAIS

Abaixo são apresentadas as proporções de crianças, em cada categoria relacionada à escola, por idade e por coorte (ano de nascimento, de julho a junho): estar freqüentando (já apresentada no capítulo 1, tabela 5), nunca ter freqüentado escola e ter evadido. Nunca ter freqüentado escola está aberto entre os que realmente nunca freqüentaram estabelecimento de ensino, os que freqüentam pré-escola, e os que freqüentam classes de alfabetização de adultos.

As proporções estão em formato de percentagem (%).

Tabela A1. Nunca freqüentou escola

coorte	idade												
	6	7	8	9	10	11	12	13	14	15	16	17	
1964													10.620
1965												10.397	10.426
1966										9.938	10.358	9.178	
1967									9.907	9.408	9.396	9.787	
1968								10.119	9.743	8.652	9.116	9.309	
1969							10.286	9.352	8.479	8.531	9.004	8.890	
1970						11.859	10.103	7.553	7.734	8.459	8.743	8.204	
1971					12.562	9.645	7.675	6.887	7.429	7.531	7.563	7.603	
1972				16.746	10.924	8.252	7.306	6.843	6.593	6.891	7.384	6.916	
1973			21.809	13.650	8.983	7.960	7.279	7.354	6.783	7.196	7.319	7.985	
1974		36.015	18.253	11.454	9.396	8.026	6.507	6.623	6.306	6.627	7.352		
1975	62.840	29.725	15.575	10.223	8.524	7.512	6.303	5.937	6.218	6.318			5.003
1976	53.570	25.442	14.438	10.594	7.920	7.098	6.667	5.758	6.424			4.486	3.964
1977	48.430	22.644	13.804	9.365	7.854	6.774	6.352	6.080		3.921	4.360		
1978	44.596	20.754	11.217	8.978	6.456	6.340	5.988		3.661	3.856			3.889
1979	41.900	18.739	10.821	8.073	6.674	6.306		3.951	3.415			3.888	3.614
1980	38.184	17.278	10.459	8.363	7.062		4.365	3.425		3.034	3.520	3.082	
1981	35.876	15.586	11.265	8.630		4.337	3.001		2.938	3.115	2.809	2.366	
1982	34.919	17.031	11.186		4.831	3.574		2.527	2.422	2.284	2.031	1.931	
1983	34.348	17.223		6.170	3.662		2.838	2.395	1.765	1.812	1.640		
1984	34.959		8.563	4.840		3.312	2.573	2.049	1.641	1.453			
1985		12.544	7.527		3.438	2.645	1.595	1.354	1.373				
1986	29.671	11.448		4.406	2.980	2.228	1.525	1.335					
1987	26.032		6.021	4.060	2.537	1.526	1.179						
1988		9.431	5.887	2.720	1.501	1.015							
1989	21.131	8.341	4.426	1.799	1.350								
1990	20.389	7.461	2.965	1.494									
1991	19.403	5.481	2.131										
1992	16.108	4.107											
1993	13.890												

Pode-se ver uma grande diminuição de crianças que nunca estiveram na escola (ou na pré-escola) na tabela A1, inclusive um grande aumento de crianças de 6 anos já entrando no sistema escolar. Já a partir da coorte de 1982 tem-se que menos de 5% das crianças de 10 anos ou mais nunca freqüentou a escola. Vale notar que esta tabela indica a entrada tardia no sistema escolar, pois mesmo para as gerações mais jovens há uma queda progressiva, dos 6 aos 10 anos, na proporção dos que nunca freqüentaram.

Aumentos na mesma coorte da proporção que nunca freqüentou de uma idade para outra não podem ocorrer; os aumentos calculados para a década de 80 devem ser em função de estarem aí incluídos os evadidos com zero anos de estudo; em 90, é erro de medida. A sensível melhora do ano de 1990 para 1992 (diagonais começando na coorte 1984 e 1986) é em parte função da reclassificação dos evadidos.

Tabela A2. Frequentar pré-escola ou creche

coorte	idade						
	6	7	8	9	10	11	12
1971					0.000	0.916	0.000
1972				0.000	1.746	0.000	0.000
1973			0.000	2.636	0.000	0.000	0.000
1974		0.000	4.144	0.000	0.000	0.000	0.000
1975	14.504	8.914	0.000	0.000	0.000	0.000	0.000
1976	28.375	0.000	0.000	0.000	0.000	0.000	0.000
1977	21.753	0.000	0.000	0.000	0.000	0.000	0.000
1978	25.645	0.000	0.000	2.545	0.000	0.000	0.000
1979	28.961	0.000	5.329	2.645	0.000	0.000	
1980	30.702	12.217	5.151	2.746	0.000		0.000
1981	42.701	12.353	6.383	2.288		0.951	0.000
1982	43.826	13.491	6.074		2.177	1.236	
1983	43.585	13.156		4.165	2.464		0.000
1984	44.781		7.078	4.525		0.604	0.000
1985		15.590	8.657		2.026	0.917	0.000
1986	46.974	16.695		3.529	1.641	0.652	0.000
1987	50.196		7.021	3.399	1.454	0.465	0.000
1988		14.138	6.767	2.983	1.046	0.290	
1989	52.591	13.697	5.241	1.639	0.569		
1990	52.316	12.918	4.324	1.669			
1991	54.232	11.098	3.153				
1992	53.602	10.331					
1993	50.758						

Nota: os zeros são função do critério de reclassificação da PNAD.

Tabela A3. Freqüentar alfabetização de adultos

coorte	idade					
	12	13	14	15	16	17
1964						0.481
1965					0.389	0.412
1966				0.529	0.552	0.452
1967			0.201	0.528	0.449	0.290
1968		0.000	0.550	0.228	0.384	0.437
1969	0.000	0.662	0.321	0.381	0.310	0.183
1970	0.857	0.350	0.275	0.351	0.185	0.223
1971	0.237	0.237	0.341	0.235	0.127	0.149
1972	0.052	0.132	0.078	0.248	0.197	0.072
1973	0.079	0.076	0.110	0.050	0.078	0.087
1974	0.051	0.089	0.173	0.082	0.072	
1975	0.029	0.028	0.109	0.055		0.148
1976	0.000	0.056	0.140		0.110	0.193
1977	0.000	0.100		0.280	0.195	
1978	0.000		0.215	0.267		0.091
1979		0.184	0.164		0.155	0.226
1980	0.054	0.197		0.184	0.169	0.264
1981	0.032		0.231	0.128	0.171	0.190
1982		0.134	0.220	0.302	0.203	0.241
1983	0.034	0.118	0.289	0.173	0.210	
1984	0.056	0.160	0.119	0.138		
1985	0.000	0.122	0.242			
1986	0.017	0.123				
1987	0.086					

Tabela A4. Evadidos

coorte	idade												
	6	7	8	9	10	11	12	13	14	15	16	17	
1964												45.359	
1965												38.705	47.356
1966											32.010	41.212	47.269
1967									24.041	33.895	41.567	47.831	
1968								15.972	24.912	33.207	40.777	47.553	
1969							10.596	17.767	25.892	34.662	41.523	48.856	
1970						5.556	10.206	17.370	25.332	33.392	41.388	49.460	
1971					3.024	5.920	10.478	17.182	24.762	34.603	41.975	47.227	
1972				1.935	2.973	5.951	10.833	17.252	26.083	35.509	40.557	47.567	
1973			1.130	2.080	3.190	6.311	10.638	17.882	26.836	32.396	39.405	45.375	
1974		0.585	1.166	1.728	3.148	5.849	11.290	17.368	24.796	32.946	39.575		
1975	0.106	0.388	1.072	2.232	3.267	5.959	10.415	16.290	23.452	30.929		44.615	
1976	0.024	0.568	1.371	1.859	3.613	5.632	10.239	16.909	22.188		38.125	43.789	
1977	0.060	0.582	1.402	2.350	2.919	5.596	10.868	15.677		30.421	36.944		
1978	0.040	0.568	1.077	2.030	3.622	5.796	9.583		21.623	27.750		39.742	
1979	0.089	0.810	1.502	1.902	3.452	5.688		14.964	19.816		32.232	36.903	
1980	0.124	0.811	1.313	1.941	3.201		9.217	12.385		23.885	28.806	33.729	
1981	0.115	0.558	1.180	1.910		5.363	6.456		16.493	20.897	25.506	30.861	
1982	0.182	0.661	1.133		3.640	5.291		10.693	14.618	18.819	21.913	27.480	
1983	0.134	0.561		2.337	3.324		6.390	8.759	12.178	16.547	20.867		
1984	0.068		2.110	2.063		3.984	5.308	6.770	9.956	14.177			
1985		1.360	1.876		2.930	3.586	4.027	5.535	8.290				
1986	0.470	1.167		2.092	3.036	2.730	3.690	5.175					
1987	0.395		1.660	1.747	1.947	2.166	2.738						
1988		0.989	1.579	1.667	1.478	1.383							
1989	0.376	1.026	0.926	1.165	1.052								
1990	0.361	0.860	1.017	1.015									
1991	0.350	0.647	0.682										
1992	0.272	0.397											
1993	0.318												

A tabela A4 inclui os jovens que afirmaram não estar freqüentando estabelecimento de ensino, porém já freqüentaram. Note que as porcentagens se acumulam de uma idade para a outra (a cada idade somam-se os evadidos no último ano com os que tinham evadido anteriormente). Na década de 80 a evasão era praticamente o dobro do que o apresentado nos últimos anos: nos últimos 2 anos de pesquisa, menos de 10% dos jovens até 14 anos evadiram; proporção que dobra aos 16 anos de idade. Deve-se lembrar que a evasão na década de 80 está subestimada, pois não considera evadidos com zero anos de estudo.

Abaixo a proporção que já concluiu o ensino médio, estes estão somados às medidas de freqüência e série ideal usadas no decorrer do trabalho.

Tabela A5. Já concluíram ensino médio

coorte	idade		
	15	16	17
1964			2.358
1965		0.168	2.530
1966	0.000	0.101	2.875
1967	0.000	0.191	2.896
1968	0.000	0.173	2.671
1969	0.000	0.148	2.732
1970	0.000	0.244	2.897
1971	0.000	0.289	2.831
1972	0.049	0.369	3.224
1973	0.008	0.455	2.740
1974	0.041	0.190	
1975	0.007		3.260
1976		0.354	3.163
1977	0.040	0.478	
1978	0.084		3.726
1979		0.563	4.493
1980	0.107	0.676	3.885
1981	0.058	0.413	5.105
1982	0.060	0.599	5.345
1983	0.041	0.489	
1984	0.065		

Nota. incluídos nas tabelas de freqüência escolar e série ideal, capítulo 1.

É interessante olhar a magnitude do atraso escolar. Condicional em se estar freqüentando, mostrarei a proporção de crianças que estão mais de 1 ano atrasadas. Separar os que estão apenas 1 ano atrasados dos que estão mais atrasados pode dar uma

idéia da gravidade do atraso. Se a maioria estivesse apenas um ano atrasada, poder-se-ia considerar o atraso uma questão menos importante, talvez apenas de adequação do desenvolvimento da criança com o conteúdo escolar, porém, a maior parte dos repetentes ainda na escola está mais de um ano atrasada. A tabela A6 mostra as proporções com os mais de 1 ano atrasado, condicional em se estar freqüentando: mais de 50% dos mais velhos (canto superior direito) estão nesta situação.

A tabela A7 mostra a média educacional atingida por cada coorte e idade, independentemente de estarem cursando ou não.

Tabela A6. Mais de uma série atrasada, condicional em estar freqüentando

coorte	idade									
	9	10	11	12	13	14	15	16	17	
1964										66.452
1965								68.214	65.293	
1966							66.735	68.578	66.121	
1967						66.824	67.316	67.320	65.568	
1968					66.210	66.991	66.359	66.767	66.240	
1969				60.623	65.571	65.513	67.524	67.842	64.453	
1970			56.029	60.263	64.969	67.026	65.314	65.787	64.700	
1971		45.188	54.649	58.504	63.574	64.057	62.484	63.494	64.940	
1972	32.683	45.501	52.643	58.931	62.605	64.175	62.306	65.361	61.538	
1973	32.360	44.378	52.661	57.519	62.114	62.785	62.823	62.717	64.764	
1974	32.891	44.639	52.000	55.413	62.652	61.423	64.275	65.750		
1975	32.079	43.084	49.494	55.205	59.962	59.136	63.254		62.409	
1976	28.713	41.558	48.247	54.252	57.815	60.254		62.939	63.182	
1977	27.293	40.743	47.407	51.773	58.292		60.924	61.584		
1978	27.004	38.236	46.714	52.502		58.935	60.083		59.148	
1979	25.392	36.863	45.321		55.665	58.189		58.940	58.391	
1980	24.354	37.006		48.150	54.125		56.939	57.470	58.389	
1981	23.444		40.648	47.212		52.935	54.073	55.844	52.515	
1982		31.904	40.583		49.641	51.688	53.237	52.825	51.053	
1983	19.934	31.199		42.484	48.223	50.379	50.382	50.463		
1984	19.653		39.217	41.570	47.570	47.542	47.246			
1985		30.006	35.832	38.894	44.206	45.214				
1986	17.648	27.852	35.142	38.719	41.299					
1987	16.880	26.535	30.583	33.428						
1988	16.973	23.259	27.284							
1989	13.625	19.810								
1990	11.941									

Tabela A7. Média de anos completos de educação

coorte	idade							
	10	11	12	13	14	15	16	17
1964								5.007
1965						4.622	5.014	
1966					4.192	4.542	5.109	
1967					3.643	4.151	4.650	5.068
1968				3.114	3.658	4.291	4.680	5.098
1969			2.576	3.129	3.763	4.270	4.747	5.210
1970		1.965	2.562	3.217	3.804	4.354	4.809	5.242
1971	1.473	2.015	2.706	3.321	3.891	4.476	4.918	5.315
1972	1.467	2.150	2.738	3.373	3.943	4.510	4.920	5.457
1973	1.575	2.164	2.784	3.373	3.987	4.496	5.076	5.378
1974	1.568	2.171	2.878	3.410	4.069	4.557	4.977	
1975	1.633	2.242	2.872	3.522	4.114	4.587		5.535
1976	1.677	2.298	2.892	3.586	4.154		5.112	5.616
1977	1.684	2.335	2.969	3.564		4.769	5.216	
1978	1.772	2.373	2.998		4.262	4.807		5.889
1979	1.806	2.409		3.697	4.306		5.475	6.087
1980	1.798		3.102	3.790		5.055	5.671	6.238
1981		2.548	3.211		4.592	5.253	5.813	6.523
1982	1.910	2.563		3.947	4.683	5.355	6.079	6.808
1983	1.940		3.368	4.115	4.817	5.511	6.322	
1984		2.667	3.450	4.136	4.958	5.762		
1985	2.014	2.784	3.565	4.321	5.136			
1986	2.103	2.825	3.609	4.462				
1987	2.146	2.973	3.817					
1988	2.305	3.127						
1989	2.412							

anexo A4: OUTRAS ESTATÍSTICAS DA EVOLUÇÃO DA RENDA E DA EDUCAÇÃO

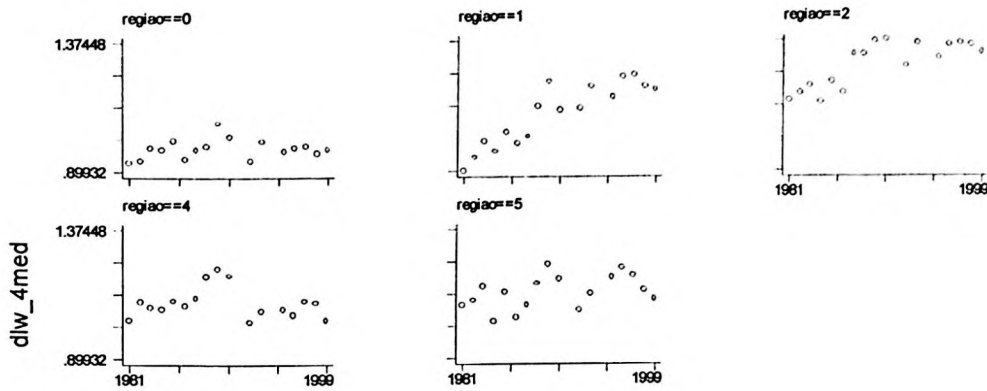
Tabela A1. Renda familiar *per capita* média, R\$ de 1999

ano/qua	1	2	3	4	Total
1981	31.26	58.94	98.95	215.31	97.63
1982	31.26	60.09	101.96	226.80	101.24
1983	25.64	47.86	80.96	187.87	82.87
1984	26.26	48.18	81.65	184.53	83.40
1985	29.44	55.92	96.89	226.53	99.48
1986	43.14	81.94	140.77	318.06	144.06
1987	31.07	62.36	109.36	258.60	113.61
1988	27.80	57.13	102.59	251.62	108.67
1989	31.62	63.89	116.70	302.00	127.00
1990	27.96	57.27	103.45	256.86	109.08
1992	28.74	58.57	102.76	232.42	104.68
1993	29.86	58.79	102.82	252.17	110.17
1995	39.31	79.48	140.59	335.22	147.10
1996	38.32	78.86	141.37	339.59	148.37
1997	39.19	79.47	140.94	339.72	147.84
1998	39.43	79.15	139.62	332.40	145.88
1999	38.16	75.50	131.92	311.16	137.72
Total	32.88	65.08	114.24	271.14	118.78

Tabela A2. Renda familiar *per capita* média, R\$ de 1999

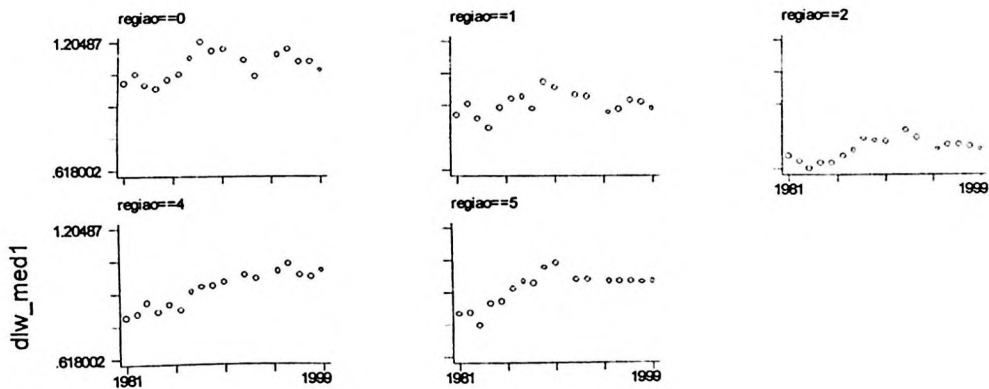
ano/região	SE	N	NE	S	CO	Total
1981	119.66	102.94	64.52	106.86	92.33	97.63
1982	124.79	112.01	65.85	108.15	100.66	101.24
1983	101.74	92.04	54.61	88.08	81.72	82.87
1984	100.42	92.46	54.95	93.35	86.11	83.40
1985	120.15	114.49	63.12	111.22	103.54	99.48
1986	175.65	162.26	91.76	153.40	160.56	144.06
1987	139.60	123.51	71.22	125.01	119.46	113.61
1988	137.29	108.25	66.69	115.81	108.97	108.67
1989	156.04	145.72	76.31	142.20	133.84	127.00
1990	133.70	117.32	67.11	122.02	121.03	109.08
1992	125.56	93.50	66.58	123.87	102.61	104.68
1993	128.75	103.11	70.67	136.15	112.80	110.17
1995	182.44	133.94	91.43	170.72	144.57	147.10
1996	184.57	133.51	90.90	170.28	148.49	148.37
1997	182.88	129.70	90.91	173.67	148.60	147.84
1998	177.45	127.95	93.00	170.53	148.64	145.88
1999	167.49	121.36	89.13	161.20	138.07	137.72
Total	145.27	119.64	74.86	133.83	121.92	118.78

Gráfico A1. Diferença de renda (ln) entre quartil 4 (mais rico) e mediana



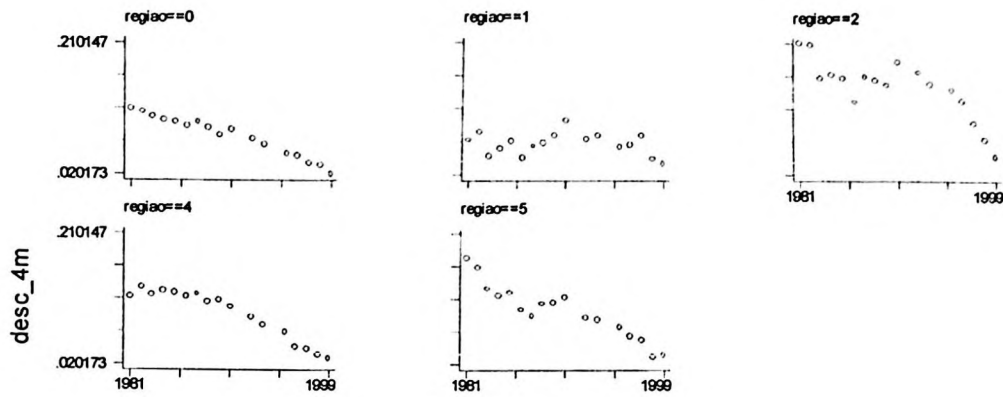
Nota: 0 - SE, 1 - N, 2 - NE, 4 - S, 5 - CO.

Gráfico A2. Diferença de renda (ln) entre mediana e quartil 1 (mais pobre)



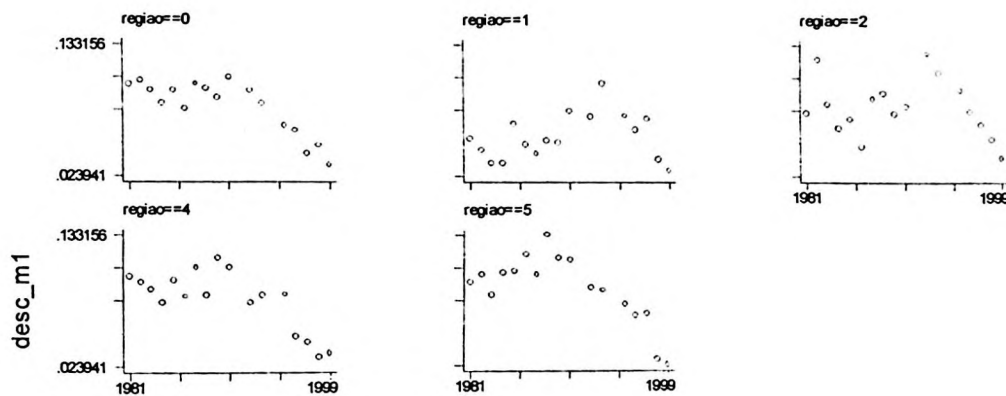
Nota: 0 - SE, 1 - N, 2 - NE, 4 - S, 5 - CO.

Gráfico A3. Diferença entre as probabilidades de frequência escolar do quartil 4 e da amostra total



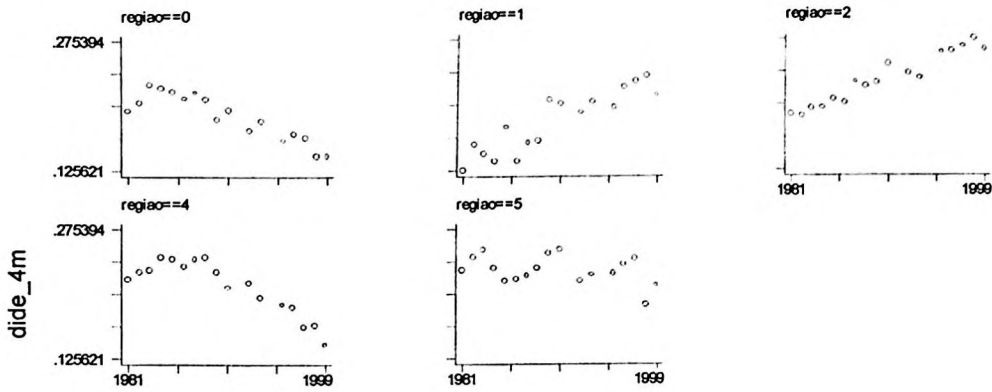
Nota: 0 - SE, 1 - N, 2 - NE, 4 - S, 5 - CO.

Gráfico A4. Diferença entre as probabilidades de frequência escolar da amostra total e do quartil 1



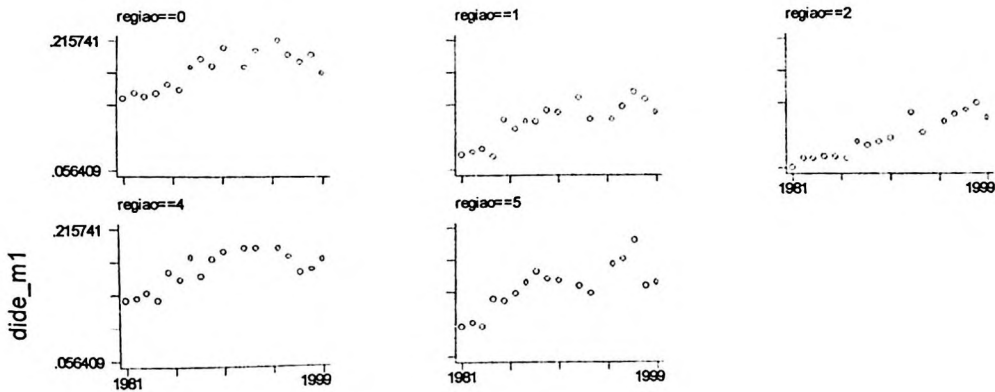
Nota: 0 - SE, 1 - N, 2 - NE, 4 - S, 5 - CO.

Gráfico A5. Diferença entre as probabilidades de estar na série ideal do quartil 4 e da amostra total



Nota: 0 - SE, 1 - N, 2 - NE, 4 - S, 5 - CO.

Gráfico A6. Diferença entre as probabilidades de estar na série ideal da amostra total e do quartil 1



Nota: 0 - SE, 1 - N, 2 - NE, 4 - S, 5 - CO.

Bibliografia

- ACEMOGLU, D., PISCHKE, J. Changes in the wage structure, family income, and children's education. *European Economic Review*, v. 45, 2001.
- AMARAL, C., SANTOS, E., BLANCO, F., RAMOS, S. Programas de renda mínima e bolsa-escola - concepção, gestão e financiamento. *Planejamento e Políticas Públicas*, n. 17, IPEA, 1998.
- ANGRIST, J. Grouped-data estimation and testing in simple labor-supply models. *Journal of Econometrics*, v. 47, 1991.
- 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, Jun. 1995.
- ANGRIST, J., KRUEGER, A. Does compulsory school attendance affect schooling and earnings?. *The Quarterly Journal of Economics*, v. CVI n. 4, 1991.
- ANGRIST, J., KRUEGER, A. Empirical strategies in labor economics. In: ASHENFELTER, O., CARD, D. (ed.), *Handbook of Labor Economics*, v. 3, Elsevier, Amsterdam: North Holland, 1999. cap. 23.
- BANCO MUNDIAL - Latin American and Caribbean Studies, *Closing the gap in education and technology*, 2002. Endereço eletrônico: <http://www.worldbank.org>. cap. 5.
- BARROS, R., LAM, D. Income and educational inequality and children's schooling attainment. In: BIRDSALL, N., SABOT, R. (ed.), *Opportunity foregone: education in Brazil*. Inter-american development bank. Washington D.C., 1996. cap. 12.
- BARROS, R., MENDONÇA, R., SANTOS, D., QUINTAES, G. Determinantes do desempenho educacional no BRASIL. Texto para discussão n. 834, IPEA, Out. 2001.
- BARROS, R., Os determinantes da desigualdade no BRASIL, *Programa de seminários acadêmicos*, Seminário n. 22/97 - 2/10/1997, IPE-USP, 1997.
- BECKER, G. *Human capital and the personal distribution of income: an analytical approach*, Woytinsky Lecture, University of Michigan Press, Ann Arbor, 1967. Republicado em BECKER, G. *Human Capital*, 2ª ed., Nova York: NBER, 1975.

- BECKER, G., TOMES, N. Human capital and the rise and fall of families. *Journal of labor economics*, v. 4 n. 3 parte 2, 1986.
- BEHRMAN, J., KNOWLES, J. Household income and child schooling in Vietnam, *The World Bank Economic Review*, v. 13 n. 2, Mai. 1999.
- BIRDSALL, N., SABOT, R. (ed.) *Opportunity foregone: education in BRASIL*. Inter-american bank of development, 1996.
- BLANDEN, J., GOODMAN, A., GREGG, P., MACHIN, S. Changes in intergenerational mobility in Britain. *Discussion Paper*, n. 517, Centre for Economic Performance-LSE, Jan. 2002.
- BLUNDELL, R., DUNCAN, A., MEGHIR, C. Estimating labor supply responses using tax reforms. *Econometrica*, v. 66 n. 4, Jul. 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, Jun. 1995.
- BRASIL. Constituição (1988). Emenda Constitucional n. 14, de 1996. Fundo de manutenção e desenvolvimento do ensino fundamental e de valorização do magistério - Fundef. Disponível em <http://prolei.cibec.inep.gov.br>.
- BRASIL. Lei n. 10172, de 9 de janeiro de 2001. Plano Nacional de Educação. Disponível em <http://prolei.cibec.inep.gov.br>.
- BRASIL. Lei n. 10219, de 11 de abril de 2001. Programa Nacional de Renda Mínima vinculada à educação - "Bolsa escola". Disponível em <http://prolei.cibec.inep.gov.br>.
- BRASIL. Lei n. 9394, de 20 de dezembro de 1996. Lei de Diretrizes e Bases. Disponível em <http://prolei.cibec.inep.gov.br>.
- CARD, D. Estimating the return to schooling: progress on some persistent econometric problems. *Econometrica*, v. 69 n. 5, Set. 2001.
- CARD, D. The causal effect of education on earnings. In: ASHENFELTER, O., CARD, D. (ed.) *Handbook of Labor Economics* v. 3, Elsevier, Amsterdam: North Holland, 1999. cap. 30.
- CARD, D., KRUEGER, A. School resources and student outcomes: an overview of the literature and new evidence from North and South Carolina. *Journal of Economic Perspectives*, v. 10 n. 4, 1996.
- CARNEIRO, P., HECKMAN, J. The evidence on credit constraints in post-secondary schooling. *The Economic Journal*, v. 112, Out. 2002.
- CARVALHO, I. Household income as a determinant of child labor and school enrolment in Brazil: evidence from a social security

- JOHNSTON, J., DINARDO, J. *Econometric Methods*. 4. ed. McGraw-Hill, 1997.
- KASSOUF, A. Trabalho infantil: escolaridade x emprego. In: XXVIII Encontro Nacional de Economia (ANPEC), 2000. Campinas: UNICAMP, 2000. Disponível em <http://www.anpec.org.br>.
- MENEZES, N., FERNANDES, R., PICCHETTI, P., NARITA, R. The choice between school and work in Latin America. In: XXVIII Encontro Nacional de Economia (ANPEC), 2000. Campinas: UNICAMP, 2000. Disponível em <http://www.anpec.org.br>.
- MINCER, J. *Schooling, experience and earnings*. Nova York: Columbia University Press, 1974.
- 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 issue 2, 1990.
- PLANK, D., AMARAL SOBRINHO, J., XAVIER, A. Why Brazil lags behind in Educational development. In: BIRDSALL, N, SABOT, R. (ed.) *Opportunity foregone: education in Brazil*, Inter-american development bank. Washington D.C., 1996. cap. 4.
- PNUD, *Informe sobre Desarrollo Humano*, 2000. Endereço eletrônico: <http://www.undp.org>.
- SCHIEFELBEIN, E., WOLFF, L. Repetition and inadequate achievement in Latin America's primary schools: a review of magnitudes, causes, relationships and strategies. *Estudos em avaliação educacional*, n. 7, 1993.
- SHEA, J. Does parents' money matter?. *Journal of Public Economics*, v. 77, 2000.
- VASCONCELLOS, L. *Uma contribuição ao estudo da exclusão na educação como causa da concentração de renda no BRASIL*. 1998. Dissertação (Mestrado em Economia) - EAESP-FGV, São Paulo.