

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA  
ESCOLA NACIONAL DE CIÊNCIAS ESTATÍSTICAS

Programa de Mestrado em Estudos Populacionais e Pesquisas Sociais

**INVESTIMENTO PRODUTIVO EM EDUCAÇÃO NO BRASIL:  
educação infantil para crianças pobres**

**Iara Azevedo Vitelli Viana**

**Rio de Janeiro**

**Abril 2012**

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA  
ESCOLA NACIONAL DE CIÊNCIAS ESTATÍSTICAS

**Iara Azevedo Vitelli Viana**

**INVESTIMENTO PRODUTIVO EM EDUCAÇÃO NO BRASIL:  
educação infantil para crianças pobres**

Dissertação apresentada ao Programa de Pós-Graduação em Estudos Populacionais e Pesquisas Sociais da Escola Nacional de Ciências Estatísticas do Instituto Brasileiro de Geografia e Estatística como requerimento parcial para a obtenção do Grau de Mestre em Estudos Populacionais e Pesquisas Sociais

**Orientador: Prof. Dr. José Eustáquio Diniz Alves**

**Co-orientadora: Profa. Dra. Denise Britz do Nascimento Silva**

**Rio de Janeiro**

**Abril 2012**

VIANA, Iara Azevedo Vitelli

Investimento produtivo em educação no Brasil: educação infantil para crianças pobres / Iara Azevedo Vitelli Viana. - - 2012.

122 f. : il.

Orientador: José Eustáquio Diniz Alves.

Coorientadora: Denise Britz do Nascimento Silva.

Dissertação (Mestrado) - - Escola Nacional de Ciências Estatísticas, Programa de Pós-graduação em Estudos Populacionais e Pesquisas Sociais, Rio de Janeiro, BR-RJ, 2012.

1. Educação infantil. 2. Desempenho escolar. 3. Modelo Hierárquico. 4. STATA. I. Alves, José Eustáquio Diniz, orient. II. Silva, Denise Britz do Nascimento, coorient. III. Escola Nacional de Ciências Estatísticas (Brasil). IV. IBGE. V. Título.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA  
ESCOLA NACIONAL DE CIÊNCIAS ESTATÍSTICAS

**Iara Azevedo Vitelli Viana**

**INVESTIMENTO PRODUTIVO EM EDUCAÇÃO NO BRASIL:  
educação infantil para crianças pobres**

Dissertação apresentada ao Programa de Pós-Graduação em Estudos Populacionais e Pesquisas Sociais da Escola Nacional de Ciências Estatísticas do Instituto Brasileiro de Geografia e Estatística como requerimento parcial para a obtenção do Grau de Mestre.

**Banca examinadora:**

---

**Prof. Dr. José Eustáquio Diniz Alves**

Orientador – ENCE/IBGE

---

**Prof<sup>a</sup>. Dr<sup>a</sup>. Denise Britz do Nascimento Silva**

Co-orientadora – ENCE/IBGE

---

**Prof<sup>a</sup>. Dr<sup>a</sup>. Maysa Sacramento de Magalhães**

ENCE/IBGE

---

**Prof. Dr. Simon Schwartzman**

IETS

*“Smart investments start by addressing a major  
root cause of inequality - families and  
disadvantaged early childhoods.”  
(James Heckman)*

## AGRADECIMENTOS

É certo que não se chega a lugar nenhum sozinho, e muitos foram os que fizeram parte da minha trajetória até aqui. Primeiro agradeço a minha mãe, que sempre acredita e incentiva os meus sonhos, sejam lá quais forem eles. Uma mulher admirável, forte e alegre, que me criou para confiar na educação. Agradeço ao meu pai, pelo amor, pelo exemplo do trabalho e pelas grandes oportunidades que tive na vida. Aprendi, com ele, que nunca devemos nos esquecer das questões microeconômicas nas discussões sobre o Bem-estar Social.

Ao Renan, pela sempre agradável companhia nos bons e maus momentos, pelas discussões construtivas, pelas horas intermináveis de conversas sobre trabalho, renda, educação, direitos sociais e tudo aquilo que se passa pela cabeça de uma economista estudando populações. Mas, sobretudo, pelo amor e carinho.

À toda a minha família, que sempre torce por mim: às minhas avós, tias e primos. Aos amigos (e agregados) da Economia, pelo incentivo, admiração e pelos momentos de diversão. Aos colegas do mestrado, pelo caminho compartilhado, e aos professores da ENCE, pelos ensinamentos.

Agradeço imensamente aos professores José Eustáquio Diniz Alves e Denise Britz do Nascimento Silva, pela orientação, apoio, críticas e incentivos. Este trabalho também é de vocês.

Por fim, não há como esquecer a Diretora e Professoras da escolinha infantil que frequentei quando criança: “E enquanto canta meu pequeno Rouxinol, inspira-me”.

## RESUMO

Este trabalho objetiva discutir a importância do investimento em um sistema de educação infantil abrangente e de qualidade. A evidência sugere que a educação iniciada ainda nos primeiros anos de vida permite enfrentar diversos problemas sociais decorrentes dos baixos níveis de educação e habilidades, como desemprego, baixa renda, violência e falta de planejamento familiar, impactando no desenvolvimento socioeconômico e na mobilidade social intergeracional ascendente.

Esse processo é discutido sob a ótica da teoria do ciclo dinâmico do desenvolvimento das habilidades e como ela se relaciona com o surgimento das desigualdades socioeconômicas entre os indivíduos. O ciclo de formação das habilidades diz que muitas das habilidades essenciais para o sucesso socioeconômico futuro dos indivíduos são desenvolvidas apenas na infância, e são necessárias para o desenvolvimento de outras habilidades ao longo da vida.

O trabalho se desenvolve em duas etapas. Primeiro é feita uma análise do efeito da frequência à educação infantil sobre resultados escolares na 4ª e 8ª série (ciclo de 8 anos) utilizando a base de dados da prova SAEB 2005. A análise foi feita utilizando modelagem hierárquica de dois níveis. Os resultados mostram que o efeito da educação infantil sobre a nota do teste é positiva, todavia os alunos da 8ª série possuem um efeito reduzido, possivelmente devido à perda das vantagens decorrentes da educação infantil ao longo da vida acadêmica.

Em seguida, utilizando dados da PNAD 2009, é feita uma análise de quantificação de crianças em situação de risco do desenvolvimento de suas habilidades, considerando fatores descritos pela literatura como adversos a esse desenvolvimento. Alguns desses fatores são: o grande número de filhos na família; baixa renda familiar; baixos níveis de educação e das habilidades dos pais; e arranjos familiares não tradicionais, representado majoritariamente por mães solteiras e ausência do pai. A análise mostrou que o número de crianças entre 0 e 6 anos em situação de risco ao seu desenvolvimento é grande. Por exemplo, existem cerca de 2.2 milhões de crianças, 11,54%, que moram apenas com a mãe, que possui baixa escolaridade. Esse quadro é pior para a população pobre, 17% das crianças com renda domiciliar per capita inferior a  $\frac{1}{4}$  de salário mínimo moram apenas com a mãe, que possui baixa escolaridade.

Na análise dos dados do SAEB e dos dados da PNAD foram utilizadas metodologias de estimação que consideram o desenho amostral complexo dessas pesquisas. Utilizou-se o software STATA para os procedimentos de estimações.

**Palavras-chave:** Educação Infantil; Desempenho Escolar, Eficácia Escolar; Modelo Hierárquico; Amostragem complexa.

## ABSTRACT

This study focus on the importance of investing in an embracing and high quality system of early childhood education. The evidence suggests that early education is effective in reducing social problems due to low levels of abilities and skills, such as unemployment, low income, violence and lack of family planning, promoting socioeconomic development and intergenerational upward social mobility.

This process is discussed based on the theory of dynamic cycle skills formation and relation with socioeconomic inequalities. The cycle theory indicates that many of the required essential skills to a successful socioeconomic adult life are developed just in childhood, and these ones beget the development of other skills throughout life.

The study presents two complementary analytical approaches based on Brazilian survey data. An analysis of preschool attendance effectiveness on school results of the 4th and 8th grades based on Brazilian educational data (SAEB 2005). This statistical analysis was carried out using a two-level hierarchical model. The results show the positive effect of early childhood education on the test score. However, for students in 8th grade, preschool attendance has a smaller effect, possibly due to the attenuation of early childhood education benefits throughout e according to academic life.

In addition, using data from a National household survey (PNAD 2009), the number of children at skills development risk, considering adverse environmental factors described in the literature, is estimated. Some of these adversities are: large number of children in the family, low income, parenting low education and skills levels, and non-traditional family arrangements represented mostly by single mothers and no father. The evidence shows the large number of children with 0 to 6 years old at risk. There are, for example, roughly 2.2 million (11.54%) children living only with their mother, who in turn has little education. The picture is even worse within the poor, 17% of the children from households with per capita income less than  $\frac{1}{4}$  of the minimum wage live only with the mother.

The analysis and estimation procedures took into account the complex survey data structure and were implemented using STATA.

**Keywords:** Early childhood education; School achievement, school effectiveness; Hierarchical model; Complex survey data.

## SUMÁRIO

LISTA DE TABELAS .....	x
LISTA DE QUADROS .....	xii
LISTA DE GRÁFICOS E FIGURAS .....	xiii
INTRODUÇÃO.....	15
1 A JUSTIFICATIVA DA EDUCAÇÃO INFANTIL.....	24
1.1 Ciclo de formação das habilidades: surgimento e determinantes das desigualdades.....	25
1.2 Eficiência econômica da educação infantil .....	32
1.3 Modelo para formação das habilidades .....	34
1.4 Evidências da produtividade da educação infantil .....	37
1.4.1 Disponibilidade de bases de dados.....	38
1.4.2 Evidências internacionais.....	40
1.4.3 Evidências brasileiras .....	42
2 FONTES DE DADOS E METODOLOGIA .....	45
2.1 Fontes de dados .....	45
2.1.1 SAEB: frequência à educação infantil e proficiência escolar .....	45
2.1.2 PNAD: características familiares e domiciliares e frequência à educação infantil	47
2.2 Análise de dados provenientes de amostras complexas .....	49
2.2.1 Amostragem complexa no SAEB .....	53
2.2.2 Amostragem complexa na PNAD.....	55
2.3 Modelagem estatística .....	56
2.3.1 Modelo hierárquico de dois níveis .....	56
2.3.2 Aplicação de modelos hierárquicos aos dados do SAEB.....	62
2.3.3 Incorporação do plano amostral do SAEB na estimação do modelo hierárquico de dois níveis .....	65
3 EDUCAÇÃO INFANTIL E PROFICIÊNCIA ESCOLAR: SAEB 2005 .....	69
3.1 Características dos alunos de 4ª e 8ª séries e frequência à educação infantil nos dados SAEB 2005.....	69
3.2 Estimação dos modelos hierárquicos.....	76

4 CARACTERÍSTICAS SOCIOECONÔMICAS FAMILIAR E DOMICILIAR DAS CRIANÇAS DE 0 A 6 ANOS E FREQUENCIA À EDUCAÇÃO INFANTIL E ENSINO FUNDAMENTAL.....	87
4.1 Acesso à educação infantil e ensino fundamental e diferenciais socioeconômicos e regionais .....	87
4.1.1 Acesso à educação para crianças de 0 a 6 anos pela população pobre.....	93
4.2 Características familiares adversas.....	94
COMENTÁRIOS FINAIS .....	109
REFERÊNCIAS BIBLIOGRÁFICAS .....	113
APÊNDICE A – INCLUSÃO NO STATA DA AMOSTRAGEM COMPLEXA NAS ESTIMAÇÕES UTILIZANDO A PNAD 2009 .....	117
APÊNDICE B: ESPECIFICAÇÃO DOS MODELOS HIERÁRQUICOS ESTIMADOS....	118
APÊNDICE C – ROTINAS DE ESTIMAÇÃO DOS MODELOS HIERÁRQUICOS NO STATA: REESCALONAMENTO DE PESOS E INCLUSÃO DA AMOSTRAGEM COMPLEXA .....	120
C.1 Reescalonamento dos pesos:.....	120
C.2 Estimações no <i>GLLAMM</i> : inclusão da amostragem complexa .....	121

## LISTA DE TABELAS

TABELA 2.1: Variáveis sobre o início da educação escolar nos questionários dos respondentes de 4ª e 8ª séries do SAEB 2005 .....	46
TABELA 2.2: Conceituação dos intervalos do coeficientes de variação (CV) das estimativas .....	51
TABELA 2.3: Participação no levantamento SAEB 2005.....	54
TABELA 2.4: Variáveis do SAEB 2005 utilizadas na estimação dos modelos hierárquicos ..	63
TABELA 2.5: Efeito das variáveis utilizadas nos modelos hierárquicos estimados.....	64
TABELA 2.6: Componentes necessários para o reescalonamento de pesos amostrais na análise multinível segundo a metodologia de Pfeffermann (1998) .....	66
TABELA 2.7: Pesos amostrais disponíveis na base de dados do SAE 2005 .....	66
TABELA 3.1: Porcentagem de estudantes de 4ª e 8ª séries por características selecionadas e respectivos coeficiente de variação (CV) e efeito do plano amostral ampliado (EPA <sub>A</sub> ), Brasil 2005 .....	72
TABELA 3.2: Porcentagem de estudantes de 4ª e 8ª séries que iniciaram a vida escolar no maternal/pré-escola por características selecionadas e respectivos coeficiente de variação (CV) e efeito o plano amostral ampliado (EPA <sub>A</sub> ), Brasil 2005 .....	74
TABELA 3.3: Resultados estimados dos modelos hierárquicos 1 a 3 para os estudantes de 4ª série, Brasil 2005 .....	77
Tabela 3.4: Resultados estimados dos modelos hierárquicos 1 a 3 para os estudantes de 8ª série, Brasil 2005 .....	79
TABELA 3.5: Resultados estimados os modelos hierárquicos 4 e 5 para os estudantes de 4ª e 8ª séries, Brasil 2005 .....	83
TABELA 4.1: Número de crianças frequentando os níveis relativos à educação infantil <sup>1</sup> por idade, Brasil 2009 .....	88

TABELA 4.2: Taxa de frequência de crianças de 0 a 6 anos à educação infantil e ensino fundamental por região, quintil de renda, cor e situação do domicílio considerando os grupos de idade, Brasil 2009 .....	91
TABELA 4.3: Taxa de frequência de crianças de 0 a 6 anos à educação infantil e ensino fundamental por população não pobre e pobre, regiões do Brasil, 2009.....	94
TABELA 4.4: Variáveis adversas no desenvolvimento das habilidades das crianças de 0 a 6 anos descritas na literatura e variáveis correspondentes utilizadas da PNAD 2009 .....	96
TABELA 4.5: Porcentagem e total de crianças de 0 a 6 anos que apresentam características familiares ou domiciliares adversas ao desenvolvimento das habilidades e respectiva taxa e total de frequência à educação infantil e ensino fundamental, população total, pobre e não pobre, Brasil 2009 .....	98
TABELA 4.6: Porcentagem e total de crianças de 0 a 6 anos que apresentam duas características familiares ou domiciliares adversas ao desenvolvimento das habilidades e respectiva taxa de frequência e total de frequência à educação infantil e ensino fundamental, população total, Brasil 2009.....	105
TABELA 4.7: Porcentagem e total de crianças de 0 a 6 anos que apresentam duas características familiares ou domiciliares adversas ao desenvolvimento das habilidades e respectiva taxa de frequência e total de frequência à educação infantil e ensino fundamental, população pobre, Brasil 2009.....	106
TABELA 4.8: Porcentagem de crianças de 0 a 6 anos por quantidade de características adversas ao desenvolvimento das habilidades, população total, pobre e não pobre, Brasil 2009 .....	108

## LISTA DE QUADROS

QUADRO 2.1: Variáveis sobre educação infantil na PNAD 1995 a 2009 .....	48
QUADRO 2.2: Critérios para definição dos estratos de interesse no SAEB 2005 .....	54
QUADRO 2.3: Plano amostral da pesquisa PNAD.....	56

## LISTA DE GRÁFICOS E FIGURAS

FIGURA 1.1: Taxa de retorno do investimento em capital humano.....	32
FIGURA 2.1: Exemplo de dados com estrutura hierárquica de 3 níveis .....	57
GRÁFICO 3.1: Porcentagem de estudantes da 4ª por idade, Brasil 2005.....	70
GRÁFICO 3.2: Porcentagem de estudantes da 8ª por idade, Brasil 2005.....	71
GRÁFICO 3.3: Percentil correspondente à nota média dos estudantes de 4ª e 8ª séries por quintil de renda, Brasil 2005 .....	75
GRÁFICO 3.4: Percentil correspondente à nota média dos estudantes de 4ª e 8ª séries das escolas públicas, por início da educação escolar, Brasil 2005 .....	76
GRÁFICO 3.5: Estimativa do efeito da escola sobre o efeito da frequência à educação infantil no desempenho escolar para os alunos de 4ª série, com seu respectivo intervalo de confiança (95%), estimado a partir dos resíduos de nível 2, $u_{1j}$ .....	84
GRÁFICO 3.6: Estimativa do efeito da escola sobre o efeito da frequência à educação infantil no desempenho escolar para os alunos de 8ª série, com seu respectivo intervalo de confiança (95%), estimado a partir dos resíduos de nível 2, $u_{1j}$ .....	85
GRÁFICO 4.1: Taxa de frequência e frequência absoluta de crianças de 0 a 6 anos à educação infantil e ensino fundamental, Brasil 1995 a 2009 .....	89
GRÁFICO 4.2: Taxa de frequência de crianças de 0 a 3 anos e de 4 a 6 anos à educação infantil e ensino fundamental, Brasil 1995 a 2009 .....	90
GRÁFICO 4.3: Taxa de frequência de crianças de 0 a 6 anos à educação infantil e ensino fundamental nos estados brasileiros, 2009 .....	92
GRÁFICO 4.4: Taxa de frequência de crianças de 0 a 6 anos à educação infantil e ensino fundamental por idade e grupos de idade, população pobre e população não pobre, Brasil 2009 .....	93

GRÁFICO 4.5: Taxa de frequência de crianças de 0 a 6 anos à educação infantil e ensino fundamental por número de filhos na família, população total, pobre e não pobre, Brasil 2009 .....	99
GRÁFICO 4.6: Taxa de frequência de crianças de 0 a 6 anos à educação infantil e ensino fundamental por número de componentes domiciliares por dormitório, população total, pobre e não pobre, Brasil 2009 .....	99
GRÁFICO 4.7: Taxa de frequência de crianças de 0 a 6 anos à educação infantil e ensino fundamental, por tipo de família, população total, pobre e não pobre, Brasil 2009.....	100
GRÁFICO 4.8: Taxa de frequência de crianças de 0 a 6 anos à educação infantil e ensino fundamental por escolaridade da mãe, população total, pobre e não pobre, Brasil 2009.....	101
GRÁFICO 4.9: Taxa de frequência de crianças de 0 a 6 anos à educação infantil e ensino fundamental por condição de emprego e desemprego da mãe, população total, pobre e não pobre, Brasil 2009.....	102
GRÁFICO 4.10: Taxa de frequência de crianças de 0 a 6 anos à educação infantil e ensino fundamental por existência de rede de esgoto <sup>1</sup> , população total, pobre e não pobre, Brasil 2009.....	103
GRÁFICO 4.11: Taxa de frequência de crianças de 0 a 6 anos à educação infantil e ensino fundamental por existência de microcomputador com internet no domicílio, população total, pobre e não pobre, Brasil 2009.....	103
GRÁFICO 4.12: Taxa de frequência de crianças de 0 a 6 anos à educação infantil e ensino fundamental por existência de máquina de lavar no domicílio, população total, pobre e não pobre, Brasil 2009.....	104

## INTRODUÇÃO

A educação é a base para a cidadania e o progresso de qualquer nação, o futuro de um país depende da boa formação de suas crianças e jovens, futuros trabalhadores e cidadãos. O esforço educacional dos pais, da família, da sociedade e do país é fundamental para garantir qualidade no ensino desde as tenras idades. É por meio da educação das jovens gerações que um país aumenta a produtividade da sua força de trabalho, garante melhor qualidade de vida para todos os habitantes e possibilita a mobilidade social ascendente para as camadas da população excluídas do processo de desenvolvimento humano. A educação infantil tem sua importância reconhecida nesse processo e tem tomado diversos debates nacionais e internacionais, que observam ser esse um dos mecanismos iniciais no processo de desenvolvimento de uma sociedade.

Na base do reconhecimento da educação infantil não está somente a importância que essa etapa possui no processo educacional, está também o direito que a criança possui ao cuidado e ao estímulo ao seu desenvolvimento. O reconhecimento do “direito fundamental à infância”, segundo o qual as crianças não devem trabalhar, e sim aprender e brincar, alçou esses indivíduos a um patamar de destaque na sociedade contemporânea, conferindo-lhes o status de sujeitos de direitos, que, por conseguinte, passaram a ser incluídos pelo poder público em políticas que visam o bem-estar da população em geral. Nesse contexto, o trabalho ora exposto tem por finalidade discutir o processo econômico e social que a educação infantil de qualidade pode proporcionar, inclusive no que toca aos efeitos da redução de desigualdades e, principalmente, na interrupção do ciclo da transferência intergeracional de pobreza.

A educação infantil consiste na educação que ocorre antes do ensino primário, normalmente voltada para o desenvolvimento de crianças de 0 a 6 anos. Em muitos países crianças de 6 anos já fazem parte do nível primário de ensino, como é o caso do Brasil desde a lei 11.274 de 2006. Segundo a LDB, Lei de Diretrizes e Bases da Educação Nacional (BRASIL, 1996), “a educação infantil, primeira etapa da educação básica tem como finalidade o desenvolvimento integral da criança até seis anos<sup>1</sup> de idade, em seus aspectos físico, psicológico, intelectual e social, complementando a ação da família e da comunidade”.

Bujes (2003) descreve a evolução das concepções sobre a educação infantil da seguinte maneira, primeiro acreditou-se que um dos seus principais objetivos era externalizar

---

<sup>1</sup> A LDB é anterior à lei 11.274/2006, e não foi revista.

as potencialidades que o educando possuía. Alguém que não tinha sucesso educacional e profissional se justificava por não ter potencial para tal. Um pouco mais tarde percebeu-se que não se tratava de um déficit de potencialidades, mas sim de um déficit de experiências, os problemas estavam nas carências de ordem econômica, social e cultural. Posteriormente concluiu-se, no entanto, que o problema poderia ser outro, “um destino social que, na base, condenava as crianças dos meios pobres ao fracasso, a uma vida dura, a um futuro já delineado no ponto de partida” (BUJES, 2003).

As teorias correntes sobre a educação infantil e o desenvolvimento das habilidades corroboram as teorias econômicas sobre Capital Humano. Ambas são uma alternativa para as teorias que diziam que os indivíduos possuem uma capacidade geneticamente determinada que explicaria desigualdades socioeconômicas. Atualmente, já é de conhecimento que o Capital Humano, e as habilidades (que são expressões do Capital Humano), são resultados de processos de aquisição de competências, influenciados apenas parcialmente pela genética.

Habilidades são aqui entendidas como conhecimentos, capacidades, aptidões, etc.. O desenvolvimento das habilidades de um indivíduo ocorre durante toda a vida, iniciando-se no nascimento (ou ainda antes disso), em um processo de ciclo dinâmico, ou seja, habilidades geram habilidades. É ressaltado que a infância é crucial no desenvolvimento desse processo, sendo responsável por habilidades que não podem ser geradas em outros momentos da vida. Esse é o embasamento teórico que sugere que a educação infantil tem papel fundamental no futuro de um indivíduo, pois, quando não ocorre apropriadamente, determinadas habilidades não são geradas, comprometendo o sucesso socioeconômico do indivíduo no futuro.

Internacionalmente, pesquisadores como James Heckman<sup>2</sup> (CUNHA e HECKMAN, 2010; CUNHA e HECKMAN, 2011; CUNHA *et al.*, 2005; HECKMAN, 2006; HECKMAN, 2008; HECKMAN e MASTEROV, 2004; HECKMAN e RAUT, 2007) advogam a importância de programas de educação infantil como solução dos diversos problemas que afetam os Estados Unidos, e que podem facilmente ser estendidos a outros países. Por exemplo, no Brasil dados da PNAD para o ano de 2009 (IBGE, 2010a) mostram que a evasão escolar entre o total de adolescentes de 15 a 17 anos é de 15%. Entre os adolescentes dessa faixa etária que frequentavam a escola, apenas metade se encontrava no nível adequado à idade, o ensino médio. Como consequência da má escolarização, a população de 25 anos ou

---

<sup>2</sup>James J. Heckman é professor de Economia da Universidade de Chicago e professor de Ciências e Sociedade da Universidade de Dublin. Foi ganhador do Prêmio Nobel de Economia em 2000 pelo trabalho na área de econometria e microeconomia. É pesquisador da economia da educação infantil e possui a fundação *The Heckman Equation*, que busca disseminar os valores do investimento em educação infantil.

mais de idade possui em média 7 anos de estudos, número que não compreende nem a conclusão do ensino fundamental. Ainda, dados da PeNSE<sup>3</sup> 2009 (IBGE, 2009) mostram que cerca de 25% das adolescentes no 9º ano do ensino fundamental que já tiveram relação sexual disseram não ter usado preservativo na última relação sexual. Dados da PNAD 2009 revelam que 6% das adolescentes de 15 a 17 anos tiveram filhos nesse ano (cerca de 283.000 mulheres).

Esses são alguns problemas sociais que pesquisadores que tratam do tema da eficiência econômica da educação infantil afirmam que podem ser superados com investimentos em indivíduos o mais cedo possível. Uma educação na primeira infância de qualidade, levando à preparação de pessoas mais eficientes economicamente, tem efeitos em fatores socioeconômicos como emprego, renda, escolaridade, gravidez indesejada, atitudes violentas e etc., resultando em aumento da renda nacional, redução dos níveis de pobreza e menor necessidade de programas sociais compensatórios para jovens e adultos.

A atenção à educação infantil vem crescendo aceleradamente no mundo inteiro, seja em decorrência da necessidade de alguém que cuide das crianças possibilitando a ambos os pais trabalharem fora (importante para o aumento da renda familiar), seja pela invocação da importância que a educação infantil tem no desenvolvimento de habilidades. Dados da UNESCO referentes aos anos de 2010 e 2011, UNESCO (2012), mostram que grande parte dos países do Leste Europeu e Leste Asiático possuem uma taxa de matrícula na pré-escola<sup>4</sup> maior que 80%. Países como Cuba e Porto Rico possuem taxas superiores a 90% e países do Oeste Europeu, como Bélgica, Alemanha, França e Espanha, possuem a universalização do ensino pré-primário.

No Brasil, dados do Censo 2010 (IBGE 2012a) mostram que 85% das crianças de 4 a 6 anos se encontram matriculadas em escolas ou na educação infantil, porcentagem bastante relevante, que mostra o comprometimento do país com a educação das crianças. Todavia, é importante ressaltar que a educação infantil tem seus efeitos positivos realizados quando acontece com a qualidade necessária para o correto desenvolvimento das habilidades das crianças, e bem se sabe que muitas das instituições de ensino infantil do país ainda possuem o papel anacrônico do assistencialismo (alguém que cuide das crianças enquanto os pais trabalham), ou lhes faltam o apoio institucional, pedagógico e financeiro para desenvolver uma proposta de ensino eficiente.

---

<sup>3</sup> Pesquisa Nacional de Saúde do Escolar

<sup>4</sup> A faixa etária da pré-escola varia em cada país, podendo ser de 3 a 5, 3 a 6 ou 4 a 6.

Nesse contexto se insere a discussão sobre quem deve ser o principal responsável pela educação de crianças entre 0 e 6 anos, se a família ou o Estado. Sendo essa a fase mais importante do desenvolvimento individual, exige bastante daquele que se responsabiliza pela educação das crianças. É claro que muitas famílias necessitam da participação do Estado no oferecimento de creches ou pré-escolas públicas para que se torne possível aos seus componentes a realização de trabalho remunerado, essencial para a reprodução doméstica, mas não se pode deixar de preponderar à justificativa educacional na atenção a crianças de 0 a 6 anos.

Ademais, sendo essa uma política que tem entre os objetivos a focalização na população em desvantagem social, cabe ressaltar a situação limitadora que muitas dessas famílias apresentam no que concerne às condições de prover os meios adequados para o cuidado e a educação das crianças, além dos baixos níveis de conhecimentos que esses pais possuem sobre o processo educacional infantil (e a impossibilidade de adquirir esses conhecimentos). O Brasil, como um país de desigualdades sociais e ainda de muitos pobres, necessita de políticas de qualidade voltadas para o bem-estar de crianças em situação de risco social, visando um melhor desempenho econômico no futuro, e assim, a quebra da reprodução intergeracional da pobreza. Isso determina, em parte, o foco governamental que deve existir na expansão e qualificação da educação infantil para crianças de famílias de baixa renda.

O Brasil demorou a atingir níveis altos de escolarização na pré-escola, considerando que a própria educação infantil tenha sido historicamente relegada pelo poder público. Essa etapa da escolarização teve sua importância reconhecida apenas a partir de reformas legais e institucionais que ocorreram com a promulgação da Constituição Brasileira de 1988. Todavia, como cita Campos *et al.* (2006), “velhas concepções, preconceitos provenientes de uma história de colonização e escravidão que ainda marca o presente, rotinas e práticas herdadas de tradições assistencialistas, convivem e resistem às propostas mais generosas que presidiram as novas diretrizes legais, baseadas em uma visão da criança como sujeito de direitos, orientada pelos conhecimentos da Psicologia do Desenvolvimento e informada pelas noções de respeito à diversidade”.

O Art. 208, I e IV, da Constituição fala do dever do Estado com a educação infantil:

Art. 208. O dever do Estado com a educação será efetivado mediante a garantia de:

~~I – ensino fundamental, obrigatório e gratuito, inclusive para os que a ele não tiveram acesso na idade própria;~~

~~I — ensino fundamental, obrigatório e gratuito, assegurada, inclusive, sua oferta gratuita para todos os que a ele não tiveram acesso na idade própria; (Redação dada pela Emenda Constitucional nº 14, de 1996)~~

I - educação básica obrigatória e gratuita dos 4 (quatro) aos 17 (dezessete) anos de idade, assegurada inclusive sua oferta gratuita para todos os que a ela não tiveram acesso na idade própria; (Redação dada pela Emenda Constitucional nº 59, de 2009)

(...)

~~IV — atendimento em creche e pré-escola às crianças de zero a seis anos de idade;~~

IV - educação infantil, em creche e pré-escola, às crianças até 5 (cinco) anos de idade; (Redação dada pela Emenda Constitucional nº 53, de 2006) (Constituição Federal de 1988)

Até 2006 a redação desse artigo previa o atendimento em creche e pré-escola às crianças de zero a seis anos de idade. Com a mudança da duração do ensino fundamental de 8 anos para 9 anos, e a incorporação das crianças de 6 anos nesse nível (prevista pela lei 11.274/2006), a Constituição passa a delimitar a idade para atendimento em creche e pré-escola de 0 a 5 anos. Em 2009 a Emenda Constitucional nº 59 altera o inciso I do Art. 208 e dispõe que a obrigatoriedade da educação entre 4 e 17 anos deverá ser implementada progressivamente até o ano de 2016, nos termos do Plano Nacional de Educação. De sorte que, em 2006 as crianças de 6 anos passaram a fazer parte do grupo de atendimento escolar obrigatório, e até o ano de 2016 serão incluídas as crianças de 4 e 5 anos. Ainda, tanto o Art. 30, XI, como o Art. 211, § 2º da Constituição confere aos municípios a responsabilidade pela oferta da educação infantil e diz da competência da União de legislar sobre as diretrizes e bases da educação nacional (Art. 22, XXIV)

A formalização da educação infantil é reiterada pela LDB (BRASIL, 1996), que regulariza o sistema educacional, com base na Constituição. A LDB diz sobre o direito à educação e dever de educar, no Art. 4º “O dever do Estado com a educação escolar pública será efetivado mediante a garantia de: (...) IV - atendimento gratuito em creches e pré-escolas às crianças de zero a seis anos de idade”. Sobre a educação infantil incube a União, em colaboração com os Estados, Distrito Federal e Municípios, das suas diretrizes (Art. 9, IV) e reitera a obrigação municipal pela oferta desse nível educacional, sendo a prioridade o ensino fundamental (Art. 11, V). Deve-se destacar o Art. 5º, §2º “em todas as esferas administrativas, o Poder Público assegurará em primeiro lugar o acesso ao ensino obrigatório [ensino fundamental], nos termos deste artigo, contemplando em seguida os demais níveis e modalidades de ensino, conforme as prioridades constitucionais e legais”. O Art. 21, I, inclui

a educação infantil como parte do sistema básico de educação, juntamente com o ensino fundamental e médio. Ao colocar a educação infantil como parte da educação básica, a LDB reorienta a responsabilidade sobre a oferta de creches e pré-escolas para o âmbito dos órgãos de educação do governo. A Lei ainda versa sobre como deve ser oferecida, “I - creches, ou entidades equivalentes, para crianças de até três anos de idade; II - pré-escolas, para as crianças de quatro a seis anos de idade” (Art. 30).

Seguindo uma tendência de formalização cada vez maior das políticas para educação, em 2001 é aprovado o primeiro PNE, Plano Nacional de Educação, para o período de 2001 a 2010, previsto na Constituição Federal (BRASIL, 2001). O primeiro PNE previa a ampliação do atendimento na educação infantil e a extensão da escolaridade obrigatória para crianças de seis anos de idade (que efetivamente ocorreu em 2006). Na proposta do PNE 2011-2020<sup>5</sup> (BRASIL, 2010) existem apenas 20 metas fixadas, número bastante reduzido se comparado com o PNE anterior. A Meta 1 se refere à universalização do ensino escolar da população de 4 e 5 anos até 2016, e ampliação do atendimento das crianças até 3 anos, de forma a atender 50% dessa população.

O reconhecimento da importância educacional e superação da justificativa assistencialista nas políticas e instituições voltadas para as crianças de 0 a 6 anos ocorre concomitantemente com os estudos que comprovam a eficiência econômica da educação infantil. A influência positiva nos resultados econômicos e sociais futuros em indivíduos que participaram de políticas focadas em crianças antes que atinjam a idade ideal para entrarem no ciclo educacional básico foi evidenciada em vários estudos. Pesquisas internacionais de longo prazo sobre programas de educação para crianças em desvantagens sociais mostraram que intervenções ainda cedo são eficientes, pois aumentam a nota em testes e a probabilidade de completar o ciclo educacional básico, reduzem retenção, tempo em classes especiais de educação, a criminalidade e a delinquência entre os jovens (CUNHA *et al.*, 2005). No Brasil as análises não são de longo prazo, por falta de dados. Elas se baseiam em dados proveniente de testes de avaliação da educação básica brasileira, como a Prova Brasil e SAEB, ambos do INEP. Klein (2007), Curi e Menezes (2009) e Nogueira *et al.* (2009) encontraram resultados positivos para a frequência à pré-escola, comparado com a não frequência.

---

<sup>5</sup> Até o final de 2011 o PNE 2011-2020 ainda não havia sido aprovado (Projeto de Lei 8035/2010) pela inconstitucionalidade de algumas emendas e a incompatibilidade com as normas orçamentárias e financeiras.

Com as evidências da eficiência da educação infantil, fica clara a necessidade de expansão do atendimento e qualidade desse nível de ensino, aliado a políticas que envolvam a família, fonte primária das desigualdades. De acordo com Ciasca e Mendes (2009), a educação infantil de qualidade deve envolver habilidades de ordem física, afetiva social, cognitiva, ética, estética e de relação intra e interpessoal. Segundo as palavras de Ciasca e Mendes (2009), a escola que recebe a criança

“deve oportunizar um ambiente físico e social onde a criança perceba-se acolhida e segura para enfrentar desafios; à medida que tais desafios se ampliam, possibilitam-lhe aumentar o conhecimento de si mesma, dos outros e do meio em que vive, ao mesmo tempo em que contribuem para o desenvolvimento de habilidades essenciais, como: autonomia, criatividade, expressividade e solidariedade”.

Araújo (2011) recomenda que a educação de qualidade deve tomar como referência os estudos sobre a neurobiologia para se repensar a política educacional, e estabelece inúmeras recomendações de ordem pedagógica e de acompanhamento da educação das crianças.

Cria-se, então, uma nova demanda: a avaliação do sistema de ensino infantil, prática discutida entre pedagogos e educadores pelo caráter especial que é a medição de avanços e desenvolvimentos em crianças de 0 a 6 anos. Didonet (2006 *apud* CIASCA e MENDES, 2009) afirma “que o modelo de avaliação escolhido deve estar estreitamente articulado com os objetivos que se quer alcançar, ou seja, a coerência entre avaliação e finalidades da educação infantil é imprescindível (...)”. De sorte que se faz extremamente necessária a contínua discussão das melhores práticas de avaliação, dado que a eficiência da educação infantil depende de sua qualidade.

Entre modelos nacionais de avaliação da educação infantil existe a Prova ABC, Avaliação Brasileira do Final do Ciclo da Alfabetização, idealizada pelo movimento independente Todos Pela Educação com o Instituto Paulo Montenegro/Ibope, a Fundação Cesgranrio e o Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP). A primeira edição da prova foi aplicada em 2011 a 6000 alunos do 4º anos do ensino fundamental (ou 3ª série no ciclo de 8 anos) de escolas municipais, estaduais e particulares de todas as capitais do País. Os resultados, divulgados em agosto do mesmo ano, mostraram o mau desempenho deste nível educacional, em grande medida, determinado pela baixa qualidade ou inexistência da educação infantil. Os responsáveis pela pesquisa apontam, ainda, que “o direito à Educação é o mesmo para todas as crianças e as diferenças de desempenho encontradas, mesmo considerando o nível socioeconômico, ou melhor, sobretudo considerando o socioeconômico, deveriam ser muito menores” (TODOS PELA EDUCAÇÃO *et al.*, 2011).

A implantação do ensino infantil de qualidade (aliada à ainda inacabada expansão do atendimento) tem sido foco de diversas políticas públicas no Brasil. Aqui cabe ressaltar que o processo de transição demográfica possibilita maiores oportunidades nessa área. O Censo de 2010 (IBGE, 2012a) mostrou que a taxa de fecundidade brasileira já se encontra abaixo de 2,1 filhos por mulher, patamar considerado abaixo do nível de reprodução. A média nacional é de 1,86 filhos por mulher, sendo no Norte igual a 2,42, única região brasileira a apresentar taxa acima do valor de reposição. A acentuada queda na fecundidade ocorrida nas últimas décadas resultou em mudanças na estrutura etária populacional que acompanha a transição demográfica. Em determinado ponto desse processo o número de crianças que nascem é cada vez menor, podendo resultar em decréscimo populacional. No Brasil ainda não há diminuição populacional, mas já pode ser percebido um número cada vez menor de nascimentos. O Censo de 2000 (IBGE, 2012b) acusou a existência de 16.375.728 crianças de 0 a 4 anos e o Censo de 2010 (IBGE, 2012a) acusou a existência de 13.796.159 crianças nessa mesma faixa etária, uma diminuição de cerca de 16% em 10 anos. Soares (2008) mostra que o Brasil conta com um bônus demográfico educacional absoluto na medida em que a população em idade escolar está diminuindo de tamanho em termos quantitativos, significando que com a mesma quantidade de recursos se pode investir em qualidade, já que o número de crianças nos primeiros grupos etários da pirâmide populacional está diminuindo.

Assim, mostra-se a importância da análise dessa nova questão socioeconômica emergente no Brasil, o investimento em educação infantil. A existência de um sistema de educação infantil de qualidade permite enfrentar diversos problemas sociais decorrente dos baixos níveis de educação e habilidades, como o desemprego, a baixa renda, violência e falta de conhecimento para colocar em prática a autodeterminação reprodutiva. É, portanto, de extrema importância que o assunto seja inserido nas pautas das discussões de políticas sociais que buscam o bem-estar da população.

Esse trabalho se desenvolve em três etapas. Na primeira, capítulo 1 são discutidas as teorias que sugerem a importância do investimento em educação infantil em um contexto de maior produtividade econômica e melhores resultados socioeconômicos. É discutida a teoria do ciclo dinâmico do desenvolvimento das habilidades e como ela se relaciona com o surgimento das desigualdades socioeconômicas entre os indivíduos.

Na segunda etapa, capítulo 3, os dados SAEB 2005<sup>6</sup> são utilizados para investigar a evidência da influência positiva da educação infantil sobre os resultados socioeconômicos futuros de um indivíduo, aqui representado pela proficiência dos alunos de 4ª e 8ª série na prova de matemática de 2005. O impacto positivo sobre a educação é verificado com base na Teoria de Produção Educacional (FPE), que relaciona uma série de insumos no processo educacional com o seu resultado. São inúmeros os trabalhos que utilizam como resultado da educação o desempenho dos alunos em testes aplicados, como é feito aqui, e, como insumos desse resultado fatores como características individuais e escolares. Aqui, o principal insumo avaliado é quando ocorreu o início da escolarização. Essa análise é feita utilizando uma modelagem hierárquica de dois níveis, considerando o aluno e suas características no primeiro nível e a escola no segundo nível.

Não são analisados neste trabalho fatores relativos à qualidade da educação infantil. Isso ocorre devido à inexistência de dados que possibilitem a relação desse contexto no âmbito da proposta empenhada. O que se busca é a relação da frequência à educação infantil e os resultados educacionais posteriores, podendo haver alguma discussão a respeito da qualidade, porém, sem seja esse o objetivo.

Por fim, na última parte da discussão, capítulo 4, objetiva-se analisar informações sobre o acesso à educação de crianças de 0 a 6 anos e as características socioeconômicas familiares dessas crianças. É mostrado qual o público do investimento educacional em questão e qual a população que apresenta maior risco no desenvolvimento das habilidades considerando fatores adversos apontados pela literatura. Nessa análise são utilizados dados da PNAD 2009, buscando o retrato mais atual possível da frequência à educação infantil.

---

<sup>6</sup> Escolheu-se o ano de 2005 por ser o ano mais recente disponibilizado em microdados.

## 1 A JUSTIFICATIVA DA EDUCAÇÃO INFANTIL

Alguns dos principais problemas sociais, como gravidez não desejada na adolescência, evasão escolar, criminalidade e desemprego, podem ser relacionados com determinadas habilidades individuais, sejam inteligência básica, habilidades sociais como saber se relacionar com outras pessoas, habilidades adquiridas como conhecimentos, autocontrole e persistência. Existem diversas políticas públicas que são vertidas para o desenvolvimento educacional dos indivíduos, e possuem entre seus objetivos, ainda que não explícito, impactar sobre a resolução dos problemas sociais mencionados anteriormente, entre outros. Isso porque já é comprovado que o desenvolvimento do conhecimento afeta positivamente nos resultados socioeconômicos futuros, por isso o interesse na área. Podemos citar como exemplos dessas políticas os programas voltados para redução da evasão escolar ou programas de inclusão profissional através de cursos técnicos.

Essas políticas educacionais existentes comumente se focam no desenvolvimento das habilidades cognitivas, que podem ser compreendidas como aquelas relacionadas à aquisição de conhecimentos em geral ou o chamado “ensino formal”, como raciocínio lógico, memória, capacidade de resolução de problemas, conceituações etc., e ignoram as habilidades não-cognitivas, que são relacionadas à formação social e emocional, como saúde física e mental, atenção, motivação, perseverança, autoconfiança, aversão ao risco, autocontrole, preferências de lazer e etc., também importantes no processo educacional.

É de amplo conhecimento que na infância acontece parte do processo de desenvolvimento dessas habilidades, e alguns pesquisadores discutem que as habilidades não-cognitivas, desenvolvidas ainda nos primeiros anos de vida, é preponderante para a aquisição das demais habilidades no futuro. Assim, para o desenvolvimento eficiente das habilidades individuais é necessário que a educação comece o mais cedo possível, considerando que investimentos na políticas educacionais que se baseiam na interferência em pessoas adolescentes ou adultas teriam eficácia reduzida quando comparadas com as interferências na educação infantil no que tange a amenização de problemas sociais.

A importância do desenvolvimento das habilidades ainda na infância e a relação econômica e social envolvida é exatamente o objeto dos estudos sobre educação infantil pelos professores James Heckman, Gosta Esping-Andersen<sup>7</sup> e outros pesquisadores. Esses estudos

---

<sup>7</sup> Gosta Esping-Andersen é professor de sociologia na Universidade de Pompeu Fabra, Espanha. Seus trabalhos abrangem dinâmicas do curso de vida, estratificação social e políticas sociais, entre elas, educação infantil.

dispõem sobre a importância de se investir em crianças em desvantagens sociais o mais precocemente possível como mecanismo de aumentar o desempenho econômico e social delas no futuro, o que implica na redução na incidência dos diversos problemas sociais, como os anteriormente citados. O presente capítulo discorre sobre os argumentos utilizados pelos estudiosos da área para embasar a teoria do desenvolvimento infantil.

### **1.1 Ciclo de formação das habilidades: surgimento e determinantes das desigualdades.**

Estudos da neurobiologia comprovaram que a estrutura cerebral apresenta seu maior desenvolvimento nos primeiros anos de vida do ser humano. O embasamento teórico das afirmações que sugerem a importância do investimento em educação infantil parte do pressuposto de que as habilidades não-cognitivas são desenvolvidas ainda nesse período, e são primordiais para o desenvolvimento futuro de outras habilidades. Ou seja, o ciclo de formação das habilidades individuais necessárias para a acumulação de capital humano é considerado dinâmico, o que pode ser resumido por meio da seguinte estrutura de Heckman e Masterov (2004): habilidades geram habilidades, motivação gera motivação.

Assim, quando um indivíduo não tem seu desenvolvimento educacional suficientemente trabalhado na infância, maiores são as chances de que ele viva em situação de desvantagem socioeconômica. Dado que as habilidades individuais são determinadas desde cedo, Heckman (2008) enfatiza: *“If a child is not motivated to learn and engage early on in life, the more likely it is that when the child becomes an adult, he or she will fail in social and economic life”*. Dessa forma, o investimento levado a efeito na educação infantil conduz a um incentivo individual por maior aprendizado, além de facilitar a formação em idades mais avançadas.

A discussão de quando se deve começar a investir em educação infantil é extensa e controversa, não sendo amplamente retratada nesse trabalho. Concerne dizer que a maior parte dos debates sobre o assunto consideram que os investimentos feitos antes dos 3 anos são mais eficientes no desenvolvimento futuro das habilidades individuais. Outros, mais minuciosos, consideram que o período de maior sensibilidade ocorre entre o sexto mês de vida fetal até o primeiro ano de vida, apesar do pequeno número de pesquisas sobre a eficiência das intervenções no período pré-natal. Para maiores discussões sobre o assunto, veja Doyle *et al.*(2007).

Na discussão da formação das habilidades deve-se considerar que a distinção entre o que é natural e o que é criado pelo meio é obsoleta. Se habilidade é unicamente determinada por fatores genéticos, intervenções públicas em educação que buscam melhorar o desempenho socioeconômico dos indivíduos seriam limitadas. Cunha e Heckman (2010) consideram que as habilidades são produzidas e a expressão genética é altamente influenciada pelas condições do ambiente. Ou seja, as habilidades são suscetíveis às influências do ambiente e também possuem um fator genético, não sendo possível controlar cada um desses dois fatores, por exemplo, em modelos lineares usualmente utilizados.

Ambas as habilidades mencionadas aqui, cognitivas e não-cognitivas, são importantes no processo do desenvolvimento intelectual humano. Vários trabalhos e pesquisas já comprovaram que as habilidades cognitivas são importantes determinantes do sucesso socioeconômico dos indivíduos no futuro, assim como as habilidades não-cognitivas o são. Perseverança, auto-confiança, motivação e auto-controle são exemplos de habilidades não-cognitivas importantes nos resultados socioeconômicos dos indivíduos. Esping-Andersen (2007a), por exemplo, salienta a importância das habilidades não-cognitivas para o desempenho do indivíduo, argumentando que as habilidades cognitivas são importantes no desempenho escolar, mas apenas escondem outra dimensão do capital humano diferente da educação formal.

A análise das habilidades não-cognitivas foi sempre limitada pela dificuldade de mensuração e pela quase inexistência de pesquisas que procurassem captá-la. Atualmente algumas pesquisas buscam mensurar essa habilidade através de variáveis de autopercepção sobre auto-estima, motivação, persistência, otimismo e etc.. Esses estudos apontam novos caminhos para a causalidade da habilidade não-cognitiva.

As habilidades aqui discutidas são determinadas desde cedo, decorrentes da educação que o indivíduo recebe em sua infância. Se essa educação não foi suficientemente adequada para a aquisição de novas habilidades e conhecimentos, no futuro tem-se um indivíduo potencialmente em desvantagem social. Isso decorre do fato de que essas habilidades influenciarão os resultados socioeconômicos futuros dos indivíduos, podendo ter relação direta com os resultados negativos como desemprego, pobreza e delinquência. O mecanismo para que esses indivíduos tenham a chance de serem bem-sucedidos seria prover a eles os meios para que possam desenvolver suas habilidades desde cedo. Esping-Andersen (2007b) diz: *“Since cognitive (and non-cognitive) abilities influence school success and, subsequently, adult’s life chances, the policy challenge is to ensure a strong start for all children”*. Heckman e Raut (2007) corroboram dizendo: *“The most effective intervention for the children*

*of poor SES should be directed at preschool stage, so that these children are prepared for school and college*”, e Heckman (2008) conclui, *“Early interventions promotes schooling, reduces crime, foster workforce productivity, and reduces teenage pregnancy”*.

Considerando o ciclo de formação das habilidades individuais, os diferenciais, tanto cognitivos quanto não cognitivos, entre indivíduos e entre grupos socioeconômicos surgem nas idades iniciais do desenvolvimento infantil. Crianças em ambientes socialmente favoráveis recebem investimento para a formação das habilidades desde novos, enquanto que crianças pobres não recebem, daí o motivo da diferença. As famílias pobres têm menor probabilidade de possuírem os recursos sociais e econômicos necessários para proverem às crianças o desenvolvimento inicial que as possibilite ter oportunidades de sucesso em seu futuro.

Essas famílias sem recursos dependem unicamente da oferta pública de serviços como alternativa para educação dos filhos, todavia, essa nem sempre é acessível ou de boa qualidade, e, somando a falta de formação educacional dos pais, essas crianças têm o desenvolvimento das suas habilidades ameaçado logo cedo. A defasagem criada, por sua vez, afetará todo o desenvolvimento futuro do indivíduo, dado o ciclo de formação dinâmico das habilidades. Portanto, as crianças em desvantagens sociais apresentarão diferenças negativas que surgem cedo e que tendem a persistir por toda a vida.

Diversas vezes a escola e sua estrutura foram indicadas como determinantes dessas diferenças. Todavia, pesquisas mostram que a escola tem um pequeno papel na determinação dos diferenciais de habilidades, assim como em aumentá-los ou diminuí-los. Em 1966, Coleman *et al.* (*apud* KAHLENBERG, 2001) conduziram um estudo para avaliar a disponibilidade de oportunidades iguais de educação para diferentes raças, cores, religiões e nacionalidades. Através da função de produção educacional objetivaram concluir que o desempenho escolar desigual dos negros era decorrente de uma insuficiência de recursos em suas escolas. O estudo acabou mostrando que são as famílias, e não as escolas, a maior fonte de desigualdades no desempenho dos estudantes. Isso porque são elas as responsáveis pela educação inicial das crianças, fase em que as habilidades cognitivas e não cognitivas são determinadas.

Hanushek e Luque (2002) estudando escolas de diversos países corroboram os achados de Coleman *et al.* (1966, *apud* KAHLENBERG, 2001) sobre a ineficiência no investimento em escolas para intervenção tardia na diminuição das desigualdades. *“(...) policies such as improving teacher education or reducing class size (...) have proven ineffective in the United States. (...) and limited elsewhere.”* (HANUSHEK e LUQUE, 2002).

Heckman e Masterov (2004) argumentam que as escolas trabalham o que lhes é oferecido pelos pais, ou seja, as escolas lidam com as habilidades dos alunos que foram desenvolvidas no convívio familiar, ou *background* familiar, tendo a própria escola menor influência sobre o desenvolvimento das habilidades das crianças. Colocando de outra forma pelos autores: “*Successful schools are those that teach children from functioning families*” (HECKMAN e MASTEROV, 2004).

Cunha *et al.* (2005) mostram diferenciais nas habilidades cognitivas e não cognitivas que surgem desde cedo, sendo essas diferenças altamente correlacionadas com as características familiares, como educação dos pais. Assim, quando são controladas essas características, as diferenças são quase eliminadas. Observa-se que a existência de diferença de habilidades na idade inicial da educação (6 anos) permanece ao longo do desenvolvimento da criança, mostrando que as diferenças que surgem cedo tendem a permanecer. Quando há o controle pela educação da mãe, a proficiência da mãe no teste AFQT<sup>8</sup> e estrutura familiar as diferenças quase desaparecem, comprovando a correlação que há entre essas variáveis e o desenvolvimento das habilidades individuais.

Crescer em uma família que ofereça ambiente propício ao desenvolvimento torna as crianças mais aptas à vida socioeconômica futura. Por outro lado, ambientes familiares adversos, que apresentam fatores de risco social, são desfavoráveis. Um ambiente familiar adverso durante os primeiros anos de vida refletirá mais tarde no baixo desenvolvimento das habilidades não-cognitivas como afeto, expressão das emoções e a maneira de se relacionar com outros indivíduos. Heckman e Masterov (2004) consideram que é a família o maior responsável pelo desenvolvimento das habilidades e motivações necessárias para a construção de estudantes e trabalhadores de sucesso. Os autores colocam: “*While the debate is not settled as to which features of adverse family environments are most harmful to the success of children, there is uniform agreement that poor environments adversely affect child outcomes*” (HECKMAN E MASTEROV, 2004).

Rutter (1981 *apud* SILVA *et al.*, 2008) contribui com a discussão ao circunscrever os cinco fatores familiares considerados adversos ao desenvolvimento infantil: i) número excessivo de filhos na família; ii) baixa renda; iii) discórdia conjugal; iv) número excessivo de pessoas residentes na mesma casa; e v) psicopatologia familiar. Heckman e Masterov (2004)

---

<sup>8</sup> AFQT: *Armed Forces Qualification Tests* mede o conhecimento de nível educacional médio (ou *highschool*) em uma série de disciplinas acadêmicas. É o teste utilizado pelas forças armadas americanas no recrutamento de pessoas.

citam ainda, baixos níveis de educação e das habilidades dos pais, falta de estímulo cognitivo e emocional às crianças, falta de conhecimento pelos pais de como educar uma criança e arranjos familiares não tradicionais, representado majoritariamente por família monoparental feminina. Fica claro que países como o Brasil, que possuem incidência dos fatores familiares considerados adversos em grande número, estão em risco maior ainda concernente ao desenvolvimento infantil.

Cunha *et al.* (2005) apresentam uma discussão sobre a eficiência dos esforços remediativos que buscam suplantar a adversidade causada pelo ambiente familiar não favorável. É possível compensar parcialmente a exposição ao ambiente não favorável se intervenções de qualidade forem feitas cedo o suficiente na vida da criança. Os esforços que se mostram mais eficazes, citam os autores, são os que suplementam os recursos escassos das famílias dessas crianças, já que a família é uma das principais fontes de desigualdade socioeconômicas (trata-se da reprodução intergeracional da pobreza, crianças filhas de pais pobres tem maiores chances de continuarem pobres).

Experimentos que enriquecem o ambiente de crianças em desvantagem social demonstram a causalidade existente entre o ambiente no qual a criança é exposta até os cinco anos de vida e os resultados socioeconômicos futuros<sup>9</sup>. Esses experimentos concluem que intervenções que compensam parcialmente os efeitos adversos que a família propicia à criança podem reverter alguns dos danos causados pela desvantagem social em que ela se encontra. Alguns desses experimentos serão citados no capítulo seguinte.

Esping-Andersen (2007a) argumenta que as habilidades são positivamente correlacionadas com os incentivos dados às crianças nos seus primeiros anos de vida pelos pais. Portanto, as políticas públicas deveriam considerar o ambiente familiar como centro da intervenção. O autor cita quatro mecanismos determinantes para uma mudança nas oportunidades das famílias: i) capital cultural familiar; ii) renda familiar; iii) dedicação dos pais à educação dos filhos; iv) estrutura familiar.

O primeiro mecanismo exposto é o do capital cultural, ou o conhecimento que há no meio familiar. Esse conhecimento não é apenas uma combinação de renda e escolaridade, mas diz respeito a toda uma bagagem cultural que a família possui. Famílias mais pobres costumam não possuir esses conhecimentos necessários para ajudarem seus filhos no desenvolvimento das habilidades, tendo a escola que se responsabilizar por esse incentivo.

---

<sup>9</sup> Essa causalidade pode, ainda, ser usada para contradizer o argumento da determinação genética das habilidades.

O segundo mecanismo citado se contrapõe ao pressuposto clássico de que a oferta universal da educação em todos os níveis poderia cancelar os efeitos negativos da falta de recursos da família para aquisição de capital humano pelos filhos, sem necessidade de intervir na renda familiar. O autor considera um pressuposto otimista, pois vários estudos comprovam a grande elasticidade da renda familiar nos resultados socioeconômicos futuros dos filhos em países como França, Estados Unidos e Reino Unido. Todavia, coloca o autor, se a renda fosse mesmo tão determinante, esperaria-se um maior efeito das políticas de bem estar social que se baseiam na redistribuição de renda (o autor considera o valor da transferência em relação ao seu real impacto sobre a educação, contrapondo à possível conclusão de que as transferências existentes não são em valor suficiente para gerar o resultado educacional esperado). Por outro lado, não há dúvidas de que o efeito de políticas contra a pobreza sobre os resultados educacionais são reais.

Em estudo sobre a população pobre da cidade do Recife, cadastrada no CadÚnico do Programa Bolsa Família, Alves e Cavenaghi (2009) mostram que o fator que mais diferencia as famílias em situação de pobreza é a presença de crianças de 0-6 anos matriculadas na pré-escola ou creche. Ou seja, a educação infantil em creches e pré-escolas, além de representar um maior capital humano para as crianças, também contribui para uma maior inserção dos pais no mercado de trabalho, contribuindo, conseqüentemente, para o aumento da renda familiar.

Ter uma maior renda, sem dúvida, tem impacto sobre a educação direcionada aos filhos. Uma maior renda pode ser devido ao fato de ambos os pais trabalharem, nesse caso sua vantagem seria sobreposta pelo pouco tempo dedicado aos filhos. Esping-Andersen (2007a) ressalta nessa discussão que o impacto da renda da mãe sobre os resultados futuros da criança são controversos. Há evidências consideráveis de que educação externa no primeiro ano de vida da criança pode ter efeitos prejudiciais. Todavia, crianças de famílias em situação de risco social teriam vantagens, dado que a oferta de trabalho entre mulheres desse grupo social é menor, e, portanto, há um maior número de mulheres que ficam com seus filhos durante o primeiro ano de vida deles. Isso não é realizado porque, considera o autor, a qualidade da educação ofertada pelos pais depende do seu nível educacional, e mães com baixo nível econômico possuem alta probabilidade de possuírem também baixos níveis educacionais. Assim sendo, é positiva a presença constante da mãe no primeiro ano de vida se ela possuir educação em nível suficiente para sobrepor uma educação externa. Ainda, ter mãe empregada a partir do primeiro ano de vida não tem efeitos prejudiciais, porém, somente se a educação externa dada à criança for de qualidade e as condições de emprego da mãe favoráveis.

Concernente a esse assunto, o autor discute a dificuldade imposta pelos pais na interferência pública na educação dos filhos. As famílias são tratadas como um santuário pela civilização ocidental, e, portanto, seria um desafio aos formuladores de políticas públicas permitirem que as famílias alcancem seus objetivos privados, ao mesmo tempo em que produzam bens públicos e externalidades sociais, decorrentes de intervenções de políticas sociais;

O último mecanismo apresentado se relaciona à estrutura familiar. Esping-Andersen (2007c) diz que as famílias estão muito mais instáveis, e cresce o número de crianças em lares de mãe solteira/separada ou divorciada. O autor cita duas razões pelas quais crescer em um lar de família monoparental feminina apresenta adversidade, o primeiro se relaciona ao fato de que há maior probabilidade de uma mãe solteira ser pobre, e, portanto, da criança viver em um ambiente com recursos escassos. O segundo motivo se relaciona ao menor tempo que os pais se dedicam à criança, dado que há a presença apenas da mãe em casa, e, sendo apenas a mãe a responsável por sustentar o lar, ela passa mais tempo dedicada ao trabalho, e menos aos filhos.

O ciclo de formação dinâmico das habilidades cognitivas e não-cognitivas mostra que investimentos na formação dos indivíduos devem ocorrer o mais cedo possível, levando à adolescentes, jovens e adultos mais capazes social e profissionalmente. “*Our analysis suggests that social policy should be more active in attempting to alter noncognitive traits, including values, especially for children from disadvantaged environments who receive poor discipline and little encouragement at home*” (CUNHA *et al.*, 2005). Quando aplicados corretamente, esses investimentos podem levar à diminuição dos problemas sociais relacionados à falta de habilidades, como já mencionados anteriormente.

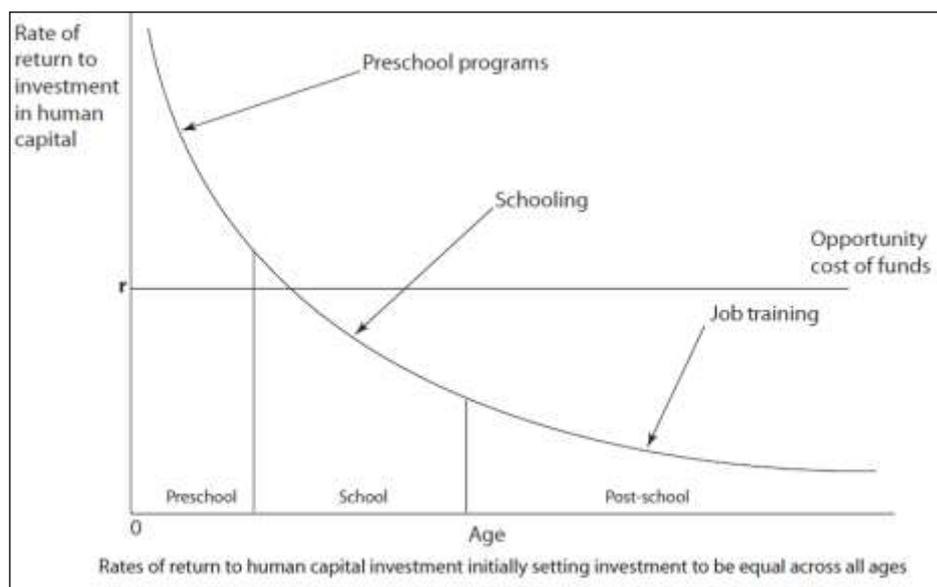
Temos, portanto, que as habilidades dos indivíduos são as grandes responsáveis pelas diferenças nos seus resultados socioeconômicos. A família tem papel crucial na formação dessas habilidades, contribuindo com a genética, mas principalmente com a oferta do ambiente no qual o indivíduo irá crescer, que se relaciona diretamente com o desenvolvimento individual. Assim, um foco no replanejamento de políticas voltadas para crianças e suas famílias tem grande importância, pois, quanto mais equitativa for a educação inicial, menores serão as desigualdades no futuro, uma vez que ela desenvolve habilidades essenciais para que as crianças se tornem adultos economicamente eficientes.

## 1.2 Eficiência econômica da educação infantil

Como citado anteriormente, a educação infantil tem grande importância no desenvolvimento inicial das habilidades individuais, que por sua vez é preponderante para o desenvolvimento das habilidades no futuro. No estudo da importância da educação infantil, o argumento maior dado por Heckman (2006) para se investir em indivíduos nos primeiros anos da sua infância, considerando o ciclo de formação dinâmico das habilidades, é o da eficiência econômica. Intervenções ainda cedo na vida de crianças se justificam pela estimativa de ter altos benefícios em relação ao custo e altas taxas de retorno quando comparadas a intervenções em idades mais avançadas. Quanto mais tarde for a intervenção na vida de pessoas em desvantagem social, mais caro será remediar a situação.

A FIGURA 1.1 demonstra a alta taxa de retorno para as idades menores a um investimento constante. *Ceteris paribus*, a taxa de retorno de um investimento na educação infantil é maior do que a taxa de retorno desse mesmo investimento se feito em idades mais avançadas. O investimento ideal seria aquele que igualasse a taxa marginal de retorno do investimento com o custo de oportunidade (representado por  $r$  na figura) em todos os períodos e para todas as pessoas, assumindo que esses investimentos sejam factíveis. Ou seja, a estratégia ótima de investimento é empregar relativamente maiores recursos na infância, e menores nas idades adultas, dado os retornos desiguais.

**FIGURA 1.1:** Taxa de retorno do investimento em capital humano  
(investimento igual em todas as idades)



Fonte: CUNHA et al (2005)

A justificativa para o maior retorno do investimento nas idades iniciais relaciona-se à natureza autoprodutiva e à dinâmica complementar do desenvolvimento das habilidades. A *autoprodutividade* pressupõe que as habilidades adquiridas em um estágio da vida persistem nos estágios futuros, assim, as habilidades produzidas em certo estágio ampliam as habilidades adquiridas em estágios seguintes. A *complementaridade dinâmica*, por sua vez, pressupõe que as habilidades geradas em um estágio resultam em um aumento da produtividade dos investimentos feitos em estágios futuros. Ou seja, os níveis de habilidades de cada estágio repercutem no desenvolvimento das habilidades futuras, sendo, portanto, sinérgicas. Juntos, *autoprodutividade* e *complementaridade* produzem um *efeito multiplicador das habilidades*, que explica porque habilidades geram habilidades.

Constata-se que a formação dinâmica das habilidades tem implicação direta em como os investimentos feitos nos indivíduos ao longo de sua vida devem ser alocados visando o desenvolvimento de suas habilidades da maneira mais eficiente, como retratado na FIGURA 1.1. Cunha e Heckman (2010) reconhecem, todavia, que, dado esse caráter autoprodutivo e complementar dinâmico da educação, os ganhos provenientes do investimento nos primeiros anos de vida podem ser atenuados com o passar do tempo caso não haja investimentos contínuos na educação desses jovens, podendo resultar na não realização dos impactos positivos futuros da educação infantil.

Como já mencionado, o processo de formação das habilidades ocorre através de multiestágios. Investimentos em cada estágio produzem habilidades naquele estágio, assim, investimentos qualitativamente diferentes podem ser utilizados para desenvolvimento das diversas habilidades. Alguns estágios podem ser mais produtivos na produção de determinada habilidade e alguns investimentos podem ser mais produtivos em determinados estágios. Derivam-se dois conceitos, o de *período sensível* e *período crítico*. O *período sensível* se caracteriza por ser mais produtivo na produção de determinada habilidade, e *período crítico* se caracteriza por ser ele sozinho efetivo na produção da habilidade. Intervenções como mostradas por Cunha *et al.* (2005) comprovam que o QI, a mensuração de algumas habilidades cognitivas, está sujeito a acumulação em um período crítico. As pesquisas mostraram que o QI não muda a partir do 10 anos, ou seja, as intervenções para aumento do QI são efetivas, porém por um período limitado.

Argumenta-se ainda que quando se investe em adolescentes ou jovens cujas habilidades não foram corretamente desenvolvidas na primeira infância, há um *trade-off* entre eficiência e equidade. É caro intervir em pessoas nessas condições, tendo, as políticas públicas, que optar por atender todos os jovens e não ser eficiente, ou ser eficiente e não

atender a todos. Em qualquer situação as taxas de retorno são baixas comparadas com o investimento feito (o que não significa que não há retornos), tornando difícil ser justificado economicamente. Cunha *et al.* (2005) analisam diversos programas americanos voltados para adolescentes em situação de risco como o *Big Brothers/Big Sisters*, *Philadelphia Future Sponsor-A-Scholar*, e *Quantum Opportunity Program*. Todos apresentaram efeitos positivos após o final do programa, seja melhores notas na escola, maiores probabilidades de conclusão do ensino básico ou menores taxas de desemprego. Todavia deve-se ressaltar que esses tipos de programas possuem impactos limitados. O autor salienta que esses programas apenas aliviaram, e não reverteram, as adversidades impostas pelo ambiente familiar de baixa qualidade. Os investimentos feitos na primeira infância conseguem algo raro nas políticas públicas, atender aos requisitos de eficiência e equidade ao mesmo tempo, ou seja, promover a justiça social e ser economicamente produtivos.

Heckman e Masterov (2004) salientam que as habilidades individuais não são todas produzidas na infância, e que outras instituições não possuem impacto neutro nos indivíduos. Escolas e empresas, por exemplo, afetam também a eficiência individual. Ainda, não é alegado que a educação infantil é determinante única do desempenho socioeconômico futuro, e que, portanto, não se deve escassear os recursos para educação tardia investindo apenas em educação infantil. O que se quer dizer é que se deve priorizar a educação das crianças, direcionando a elas grande parte dos investimentos. A melhor maneira de incrementar os resultados socioeconômicos futuros de uma pessoa é investir no ambiente no qual a criança é exposta logo cedo.

### **1.3 Modelo para formação das habilidades**

Cunha *et al.* (2005) usam modelos econômicos para representar as evidências mostradas na Figura 1.1 e discutidas nas seções anteriores. Nessa seção serão apresentados os modelos e definidos os conceitos de *complementaridade*, *autoprodutividade* e *efeito multiplicador de habilidades*, e como essas noções ajudam a demonstrar as evidências empíricas da importância da educação infantil. Os modelos apresentados não serão aplicados à dados, são apenas utilizados para embasar a discussão teórica.

*Complementaridade*, *autoprodutividade* e *efeito multiplicador* são conceitos importantes para a tecnologia de formação das habilidades, apresentada em Cunha *et al.* (2005). Essa tecnologia demonstra porque são baixos os retornos dos investimentos realizados tardiamente na infância ou na jovem adolescência de um indivíduo, enquanto são altos os

retornos dos investimentos feitos no início da infância. A tecnologia se aplica considerando indivíduos provenientes de famílias em condições adversas e situação de risco social, que não conseguem proporcionar ambientes adequados para o crescimento e educação dos filhos no início da infância. Sumarizando, a tecnologia mostra porque intervenções são mais eficientes se feitas no início da infância comparado com intervenções tardias, e porque não há *trade-off* entre equidade e eficiência nos investimentos precoces, sendo esse *trade-off* presente nos investimentos tardios.

O modelo apresentado pressupõe dois períodos na vida da criança,  $t=1$  e  $t=2$ . Chamaremos  $t=1$  de primeira infância e  $t=2$  de período posterior, sendo o desenvolvimento das habilidades em  $t=1$  considerado a educação infantil. A vida adulta é um terceiro período. Assume-se a existência de dois tipos de habilidades,  $S^C$  e  $S^N$ . Por exemplo,  $S^C$  pode ser habilidade cognitiva e  $S^N$  habilidade não-cognitiva. O investimento dos pais nas habilidades  $k$  ( $k=S, C$ ) dos filhos no período  $t$  ( $t=1, 2$ ) é  $I_t^k$ . No modelo apresentado os pais fazem as decisões a respeito dos filhos e tem controle sobre os investimentos efetuados. Assume-se que cada indivíduo nasce com a condição inicial  $S_0 = (S_0^C, S_0^N)$ . Em cada estágio,  $S_t = (S_t^C, S_t^N)$  denota o vetor de habilidades em estoque.

A *complementaridade direta* da habilidade  $l$  adquirida em  $t=1$  sob o investimento  $I$  para produção da habilidade  $k$  em  $t=2$  é:

$$\frac{\partial^2 S_2^k}{\partial I_2^k \partial S_1^l} = \frac{\partial f_2^k(S_1, I_2^k)}{\partial I_2^k \partial S_1^l} > 0$$

Os estoques de habilidades em  $t=1$  ajudam a promover a aquisição de habilidades em  $t=2$  através do aumento da produtividade dessa aquisição. É justamente essa complementaridade que explica porque os retornos educacionais são maiores nos estágios futuros do ciclo de vida para crianças mais habilidosas (com maiores níveis de  $S_1$ ). Crianças com melhores habilidades cognitivas e não cognitivas são mais eficientes no aprendizado futuro de habilidades cognitivas e não cognitivas.

A *autoprodutividade* de  $S_1^C$  ocorre quando:

$$\frac{\partial S_2^C}{\partial S_1} = \frac{\partial f_2^C(S_1, I_2^C)}{\partial S_1} > 0$$

Altos níveis de estoque de habilidade em um período proporcionam altos níveis de estoque de habilidade no período seguinte. O efeito conjunto da *complementaridade* e *autoprodutividade* explicam a maior produtividade do investimento em crianças em desvantagem social. Jovens adolescentes em desvantagem possuem baixos níveis de habilidades e, portanto, o efeito da *complementaridade* e *autoprodutividade* são baixos.

Considerando  $h$  o nível de capital acumulado quando a criança adentra a fase adulta, e dependente dos estoques de  $S_t$ ,  $t=1,2$  e que as habilidades criadas são devidas aos investimentos realizados, a tecnologia das habilidades na fase adulta,  $h$ , como função dos investimentos  $I_t$ ,  $t=1,2$  é:

$$h = \{\gamma I_1^\emptyset + (1 - \gamma) I_2^\emptyset\}^{\frac{1}{\emptyset}} \quad (1.1)$$

Aqui é apresentado o conceito de *efeito multiplicador de habilidades*. O parâmetro  $\gamma$  é um multiplicador que capta o efeito de  $I_1$  em  $h$  através da *autoprodutividade* (habilidade adquirida em um período persiste em períodos futuros) e *complementaridade* (habilidade produzida em um estágio aumenta a produtividade dos investimentos em períodos subsequentes).  $I_1$  afeta a acumulação de  $S_1$ , esse estoque, por sua vez, afeta a produtividade de  $I_2$  na formação de  $S_2$ . Quando  $\gamma = 1$ , então todo o capital humano é produto do investimento durante a primeira infância. Quando  $\gamma = 0$ , então nada se produz na primeira infância.

O parâmetro  $\emptyset$  é a elasticidade de substituição entre  $I_1$  e  $I_2$  na produção de  $h$ . Ele informa qual a facilidade de substituição entre as habilidades para criar  $h$ , e representa o grau de complementaridade (substituibilidade) entre os investimentos primários e tardios na produção de habilidades. Assim,  $\emptyset$  informa qual a facilidade de compensação em  $t_2$  pelos baixos níveis de investimento em  $t_1$ .

→ Quando  $\emptyset$  é pequeno, investimentos baixos em  $t_1$  não são facilmente remediados através de investimentos em  $t_2$ . No caso extremo de  $\emptyset \rightarrow -\infty$ , temos que os *inputs* são complementares perfeitos, e absolutamente não é possível a remediação.

→ Quando  $\emptyset = 1$ , temos o caso de substitutos perfeitos, e é possível compensar por déficits de investimentos em educação infantil através do investimento futuro, apesar de poder não ser economicamente eficiente.

Considerando os indivíduos pobres, os investimentos realizados em educação infantil são presumidamente baixos,  $I_1$  é abaixo da média. Os programas que visam remediar essa adversidade investem nesses jovens no período posterior a  $t_1$ , fazendo  $I_2$  relativamente alto. Dependendo do valor do parâmetro  $\emptyset$ , é possível, através da tecnologia de produção, remediar uma infância com baixos investimentos, todavia pode não ser economicamente eficiente.

A equação (1.1) nos mostra que as habilidades do indivíduo quando adulto são decorrentes da combinação de habilidades cognitivas e não-cognitivas adquiridas anteriormente. Cunha *et al.* (2005) pontuam que as diferentes atividades desenvolvidas no mercado de trabalho requerem uma combinação em diferentes proporções dessas habilidades. Quando há déficits em determinadas habilidades existem dois mecanismo remediativos, um seria o investimento compensatório na habilidade deficitária, o outro mecanismo seria incentivar esses indivíduos a buscar atividades em que não são requeridas as habilidades em falta. A lógica seria que pessoas motivadas, mas não proficientes, são tão eficientes quanto pessoas proficientes, mas desmotivadas.

O que essa análise mostra é a não existência do *trade-off* entre equidade e eficiência no investimento em educação infantil. No entanto, a intervenção em um período posterior em crianças que não tiveram uma educação infantil adequada depende da função (1.1), a tecnologia de produção das habilidades. Se  $I_1$  e  $I_2$  forem complementares perfeitos, por exemplo, a compensação em  $I_2$  pelos baixos níveis de  $I_1$  não é possível. Se forem substitutos perfeitos, e a intervenção em  $t=2$  pode, tecnologicamente, compensar os déficits em habilidades. Todavia deve-se considerar o alto custo que pode ter essa compensação, levando a uma ineficiência econômica. Cunha e Heckman (2010) colocam: “*dynamic complementarity, self-productivity of human capital, and multiplier effects imply an equity-efficiency trade-off for late child investments but not for early investments*”.

O modelo descrito foi simplificado de Cunha *et al.* (2005), e é o mais simples na análise da eficiência da educação infantil. Um modelo mais rico pode ser encontrados em Cunha e Heckman (2010), que incorporam as habilidades dos pais na determinação do estoque de habilidades dos filhos e efeito de políticas governamentais nas escolhas dos pais.

#### **1.4 Evidências da produtividade da educação infantil**

Cunha *et al.* (2005) resumizam como os principais achados na literatura sobre educação infantil os oito pontos a seguir: i) habilidades têm importância na aquisição de um nível de educação formal maior e nos retornos socioeconômicos futuros; ii) habilidade é

multidimensional, tanto habilidade cognitiva quanto habilidade não cognitiva têm impacto no sucesso socioeconômico do indivíduo. Ambas as habilidades são influenciadas por genes e pelo ambiente; iii) diferenças no nível de habilidade desenvolvida entre indivíduos e entre grupos sociais aparecem ainda cedo, na infância; iv) é possível compensar parcialmente pelo ambiente familiar adverso ao qual o indivíduo foi exposto em sua infância; v) diferentes tipos de habilidades parecem ser manipulados em diferentes idades; vi) quanto mais tarde for feito o investimento com objetivo remediativo, menos eficiente ele vai ser; vii) os retornos econômicos para investimentos em educação infantil são altos; e viii) a complementaridade entre investimentos em educação infantil e investimentos mais tarde pressupõe a continuidade dos investimentos para que o gasto em educação infantil seja efetivado. Vários são os estudos que comprovam estes achados. A seguir são apresentados alguns, com uma ressalva inicial sobre a disponibilidade de bases de dados que permitam o estudo adequado do tema no âmbito brasileiro.

#### **1.4.1 Disponibilidade de bases de dados**

A influência positiva nos resultados econômicos e sociais de políticas focadas em crianças de 0 a 6 anos foi evidenciada em vários estudos. A maioria deles se baseia no desempenho de crianças ainda na escola, não fazendo o acompanhamento até a juventude ou vida adulta. Essas evidências, que podem ser consideradas incompletas, decorrem da escassez de dados adequados para a análise dos impactos da educação infantil. Os problemas com os dados para a análise aqui discutida decorrem de dois fatos. O primeiro se relaciona com as habilidades não cognitivas, variável importante na análise do desenvolvimento das habilidades individuais. O segundo problema decorre da extensão longitudinal dos estudos sobre o assunto.

Apesar de alguns estudos existentes, as habilidades não cognitivas são habitualmente ignoradas quando se trata sobre o assunto, especialmente devido à falta de informações confiáveis sobre como mensurá-la. Sobre o assunto, Cunha *et al.* (2005) citam que nenhum fator individual emergiu como dominante na literatura sobre habilidades não cognitivas, e que é improvável que algum vá aparecer devido ao caráter múltiplo das habilidades não-cognitivas.

Por outro lado, as habilidades cognitivas são frequentemente medidas em testes avaliativos (por exemplo, nos modelos do SAEB) e são alvos de políticas públicas que buscam melhorar os indicadores do aprendizado resultante dessas pesquisas. Saúde física e

mental, atenção, motivação, perseverança, auto-confiança e outras são algumas das habilidades fundamentais na performance socioeconômica dos indivíduos que não são mensurados nesses testes mas que possuem grande influencia nos seus resultados. Temos, portanto, que a análise dos determinantes da educação, assim como dos gargalos a serem enfrentados pelas políticas públicas acabam sendo prejudicados.

O outro fato que gera problemas nas análises sobre educação infantil decorre da extensão longitudinal das pesquisas existentes. Para melhor análise dos efeitos da educação infantil seria adequada a utilização de pesquisas do tipo painel que acompanhassem o indivíduo desde a infância até o momento em que fosse possível ter uma medida do seu resultado socioeconômico, que seria, no mínimo, após ele entrar no mercado de trabalho.

No cenário internacional se encontram pesquisas específicas, ou que possibilitam este tipo de análise, com maior facilidade. É o caso dos programas *Chicago Child-Parent Centers (CPC)*, *Abecedarian program* e *High/Scope Perry Preschool*, que serão analisadas à frente. Além dessas pesquisas, existem outras desenvolvidas pelo governo americano, como *National Longitudinal Survey of the Youth (NLSY)*, um conjunto de pesquisas realizadas pelo *Bureau of Labor Statistics*, do governo americano que objetivam reunir informações em diversos pontos no tempo sobre atividades no mercado de trabalho e outros eventos significativos, como os educacionais. A NLSY79, por exemplo, reúne informações sobre 12.686 homens e mulheres nascidos entre 1957 e 1964, que foram entrevistados anualmente até o ano de 1994, e depois bianualmente. A pesquisa abrange ainda os filhos tidos pelas mulheres participantes da pesquisa. No Reino Unido, a maior pesquisa sobre impactos da educação infantil é o *Effective Provision of Pre-School Education (EPPE) Project*, iniciado em 1997, que acompanha o desenvolvimento de 3000 crianças desde os 3 anos e está previsto para continuar as pesquisas até 2013.

No Brasil não existem investigações desenvolvidas para esse tipo de análise, todavia algumas pesquisas de alcance nacional possibilitam alguma análise de longo prazo, como seria o caso da Pesquisa de Padrão de Vida<sup>10</sup> (IBGE), analisada por Curi e Menezes-Filho (2009). Cunha e Heckman (2011) criticam o fato de que no Brasil existem diversos programas em larga escala sem qualquer contrapartida de avaliação, devendo ser outro o modelo a ser seguido, o de coleta de dados que permita acompanhar o desenvolvimento infantil do período

---

<sup>10</sup> A Pesquisa de Padrão de Vida (PPV) foi uma pesquisa domiciliar realizada pelo IBGE entre março de 1996 e março de 1997, nas regiões do Nordeste e Sudeste do Brasil. Objetivava obter informações sobre os determinantes do bem estar social de diferentes grupos sociais.

pré-natal até a fase adulta. Completam, “Estes dados devem ter informações detalhadas sobre o ambiente familiar, o *status* intelectual e emocional dos pais, os recursos que as famílias dispõem para investir na educação dos filhos e as características das escolas e dos professores com quem as crianças passam boa parte do tempo” (CUNHA e HECKMAN, 2011). Araújo (2011) recomenda que o estímulo a “desenvolvimento de pesquisas através da formulação de experimentos e levantamento sistemático de dados que explorem o relacionamento da dimensão neurobiológica e psicológica com a educação” e “Determinar os custos e os benefícios de diferentes intervenções voltadas para o desenvolvimento infantil”.

#### **1.4.2 Evidências internacionais**

Três programas americanos se destacam nas análises internacionais por terem acompanhado indivíduos em suas pesquisas durante longos períodos, o *Chicago Child-Parent Centers (CPC)*, o *Abecedarian program* e *High/Scope Perry Preschool* (CUNHA e HECKMAN, 2011). Todos os três programas focalizaram crianças em desvantagens socioeconômicas, por exemplo, pais com baixos níveis de escolaridade, renda, capacidade cognitiva, níveis elevados de comportamento patológicos entre outras características adversas. Os programas também envolveram os pais de alguma maneira, objetivando envolvê-los no processo educacional de seus filhos e para que esses também implementassem o currículo do programa em casa.

O primeiro desses programas, *Chicago Child-Parent Centers (CPC)*, existe desde 1967, mas os dados para a pesquisa foram coletados apenas a partir de 1980. É um programa de meio período, 3 horas diárias durante o ano escolar, que ocorre em escolas públicas de Chicago. Os participantes entram com 3-4 anos e ficam no programa por um máximo de 6 anos. A análise ocorreu por um método quase experimental (*matching*) em uma amostra de 989 crianças, que foram acompanhadas até os 21 anos de idade.

O segundo programa considerado, *Abecedarian program*, foi um programa que oferecia cuidados às crianças em horário integral, desde os primeiros meses de vida. As 111 crianças que entraram no programa entre 1972-77 foram acompanhadas até 21 anos de idade. Esse era um programa experimental, portanto foram formados quatro grupos para avaliação, crianças que i) não sofreram nenhum tipo de intervenção pelo programa; ii) sofreram intervenção até os 5 anos de idade; iii) sofreram intervenção dos 5 aos 8 anos; e iv) sofreram intervenção durante toda a infância até os 8 anos. O tratamento constou de aulas diárias, em

tempo integral, durante os 12 meses do ano, suplementado por visitas regulares à casa das crianças pelos professores durante os três primeiros anos do ensino primário.

Finalmente, o *High/Scope Perry Preschool* foi um programa também experimental de meio período em escolas públicas em Ypsitanti, Michigan. Atendeu 58 crianças, e acompanhou outras 65 como comparação. As crianças entraram no programa entre 1961 e 1965 com 3-4 anos de idade e permaneceram no programa por dois anos. O tratamento constou de aulas diárias com 2,5 horas de duração durante o ano escolar, suplementado com uma visita semanal à casa do aluno pelo professor. O acompanhamento ocorreu anualmente até os 15 anos de idade, e depois aos 19, 27 e 40 anos de idade.

Os resultados das análises dos três programas mostraram que intervenções ainda cedo em crianças em desvantagem socioeconômica são eficientes. Os participantes apresentaram aumento da nota em testes cognitivos e aumento da probabilidade de completar o ciclo educacional básico, redução da retenção e tempo em classes especiais de educação, e diminuição da criminalidade e delinquência.

O programa *Perry Preschool* encontrou efeitos substanciais na educação até os 27 anos de idade, além de maior probabilidade entre os que participaram do programa de estarem empregados, de ganhar mais e serem menos dependentes de programas de bem-estar social. No entanto, o ganho nos testes de QI desapareceu ainda na 2ª série. O programa *Abecedarian* encontrou um importante resultado entre os testes de QI. Aqueles que participaram do programa desde os primeiros meses de vida apresentaram resultados maiores de QI que perduraram até os 21 anos de vida (última idade analisada). Esses resultados se concentraram entre mulheres, mas reforçaram a idéia de que quanto mais cedo a intervenção acontece, melhores são os resultados futuros.

Outros resultados mostram que, no programa *Perry Preschool*, enquanto 45% dos participantes no grupo de controle concluíram o ensino médio, o mesmo percentual para o grupo de tratamento foi de 66%. No Projeto *Abecedarian*, 51% do grupo de controle concluiu o ensino médio, enquanto que 67% do grupo de tratamento chegou ao fim do ciclo básico de educação. Por fim, ainda no Projeto *Abecedarian*, 13% das crianças do grupo de controle concluíram o ensino superior, sendo que no grupo de tratamento o percentual foi 3 vezes maior.

Hidalgo e Urzúa (2010) fizeram uma análise da efetividade dos investimentos em educação infantil no “mundo em desenvolvimento” utilizando dados de uma pesquisa longitudinal iniciada no Chile em 2007 pela *Junta Nacional de Jardines Infantiles (JUNJI)*. A pesquisa objetivava avaliar o efeito da participação nos centros de educação infantil sobre o

resultado do desenvolvimento das crianças. Foram analisadas 482 crianças que tinham idades de 2-14 meses no ano de 2007, separadas em grupo tratamento (331 crianças) e grupo controle (151 crianças). Os dados foram coletados em abril de 2007 e em outubro de 2007, 7 meses depois. Os resultados mostram um efeito inicial da educação infantil negativo (comparado com as crianças que não frequentavam o centro infantil). Os autores justificam através da psicologia: a criança mostrará um extremo estresse por ter deixado de conviver com mãe, e, a princípio, negará qualquer outro cuidador, tendo efeito sobre o nível de habilidades. Todavia, ao longo do tempo, o efeito médio da frequência ao centro infantil é positiva em todas as áreas do desenvolvimento infantil.

Hidalgo e Urzúa (2010) citam quatro outros programas recentes de avaliação de políticas públicas voltadas para a educação infantil na América Latina. *Proyecto Integral de Desarrollo Infantil (PIDI)* na Bolívia, a expansão do sistema de educação infantil na Argentina e Uruguai, e *Hogares Comunitarios de Bienestar Familiar (HCBF)* na Colômbia. As avaliações, realizadas por outros autores, tiveram resultados positivos sobre a aquisição de habilidades, tais como maior aquisição de habilidades em linguagem, melhores notas em testes de espanhol e matemática, melhores notas na educação primária, menor taxa de retenção, menor taxa de evasão na adolescência, e outros.

### **1.4.3 Evidências brasileiras**

No Brasil, a existência de dados para análises como essas feitas para os programas americanos é limitada. Avaliações de intervenções no início da infância se baseiam em dados de frequência à educação infantil proveniente de testes de avaliação da educação básica brasileira, como a Prova Brasil e SAEB<sup>11</sup>, ambos do INEP, ou de outras pesquisas que porventura captam essa frequência, como a PPV. Há falta de pesquisas que acompanhem o indivíduo por um período de tempo maior, ideal para avaliação do impacto duradouro dessas intervenções.

Entre os estudos disponíveis, Klein (2007) faz uma avaliação do impacto da frequência à educação infantil no Brasil utilizando dados do SAEB 2003. Para os alunos na 4ª série foi possível saber se a entrada na escola ocorreu no maternal ou pré-escola (diferentes

---

<sup>11</sup> SAEB é uma prova aplicada pelo INEP que avalia conhecimentos em língua portuguesa e matemática dos alunos na 4ª e 8ª séries do ensino fundamental (ciclo de 8 anos) e 3º ano do ensino médio nas escolas públicas e privadas, urbanas e rurais.

níveis da educação infantil). Os resultados encontrados para Brasil mostram que apesar de parecer que quanto mais cedo se entra na escola melhor o desempenho no teste, os resultados para a 4ª série não mostram diferenças significativas entre quem entrou no maternal e quem entrou na pré-escola, mostrando possível desqualificação da educação maternal. No entanto, para todos os níveis avaliados, 4ª série, 8ª série e 3º ano, a média em matemática de quem entrou no sistema de ensino somente na 1ª série ou depois é significativamente menor do que a média de quem frequentou a educação infantil. A análise pelo índice socioeconômico (ISE) mostrou que há diferenças entre escolas públicas e privadas. Os resultados médios para as escolas públicas, que apresentam ISE mais baixo, são menores do que os resultados para as escolas particulares. A melhor média para as escolas públicas, correspondente ao quintil superior de quem teve ingresso na pré-escola, se iguala ao resultado do 5º quintil de quem ingressou na 1ª série ou ao 3º quintil de quem ingressou na pré-escola no ensino privado.

O autor utiliza ainda uma regressão multinível para a análise das notas médias em matemática no teste SAEB considerando o início da vida escolar do indivíduo: se na educação infantil, 1ª, 2ª ou 3ª série. A regressão foi controlada ainda por cor e pelo índice socioeconômico (ISE). O segundo nível da regressão hierárquica inclui duas variáveis amplamente utilizadas na literatura: média do ISE da turma e média do atraso escolar dos alunos da turma. O controle por cor mostrou maior média entre os que se autodeclararam brancos, em comparação aos que se autodeclararam pretos. Considerando a regressão multinível, atraso escolar na série e ISE baixo são fatores negativos na média do teste, sendo que há uma porcentagem maior de atrasados entre aqueles que iniciaram a educação na 1ª série. A análise concluiu que “a pré-escola faz diferença e que pode atenuar as desigualdades socioeconômicas. Mas a pré-escola atual não é suficiente para igualar o desempenho dos alunos de diferentes níveis socioeconômicos”.

Curi e Menezes-Filho (2009) fazem um estudo sobre os impactos da educação infantil nos resultados socioeconômicos futuros baseados nas pesquisas SAEB 2003 e PPV. Os autores avaliam o desempenho escolar dos alunos medidos pelo testes de proficiência SAEB, condicionado à frequência à educação infantil para o ano de 2003. Eles utilizaram o método estatístico de Mínimos Quadrados Ordinários (MQO) e encontraram o mesmo efeito positivo da pré-escola para os alunos de 4ª e 8ª séries do ensino fundamental e 3º ano do ensino médio, sendo esse efeito maior para os alunos do ensino fundamental. Considerando o teste de matemática, frequentar a educação infantil aumenta em 7,5% o resultado do teste para os alunos da 4ª série, para a 8ª série a melhora é de 3% e para o 3º ano é de 1%. A análise econométrica para os resultados na prova indicou, ainda, características familiares e do aluno

como maiores determinantes do desempenho escolar, resultado de acordo com os argumentos internacionais sobre a importância da família na educação infantil, como cita o autor.

Curi e Menezes-Filho (2009) utilizam a PPV para avaliar a relação entre educação pré-primária, escolaridade e salário. A análise da conclusão dos ciclos escolares (fundamental I, fundamental II, ensino médio e ensino superior) condicionadas à frequência à educação infantil, mostrou, como esperado, efeito positivo da frequência à escola de crianças entre 0 e 6 anos. O efeito marginal da pré-escola sobre a conclusão do ensino fundamental I é de 4%, o efeito da creche não é significativo. O efeito da pré-escola sobre a conclusão do fundamental II é em torno de 18%, sendo o efeito da creche também não significativo. A conclusão do 3º ano sofre efeito da frequência à creche e pré-escola, respectivamente 13% e 24%. E, por fim, também creche e pré-escola influenciam a conclusão do ensino superior, efeito de 19% e 5,5%. Já a análise dos salários mostrou que o efeito da creche não é estatisticamente significativo sobre os rendimentos, por outro lado o efeito da pré-escola é significativo, ou seja, a pré-escola tem efeitos positivos no salário.

Outro estudo sobre o assunto foi realizado por Nogueira *et al.* (2009) com 299 famílias residentes em Belo Horizonte, pertencentes aos estratos médio e baixo da população e que escolarizavam seus filhos nas redes de ensino federal, estadual, municipal e privada. Utilizou-se no estudo a proficiência em testes de Língua Portuguesa. Uma simples análise estatística fornece o resultado positivo da educação infantil sobre os estudantes analisados. O aluno que frequentou a pré-escola teve um resultado na proficiência em média 11,04 pontos maior. O grupo das crianças que iniciaram essa escolarização antes dos 2 anos apresenta uma diferença na proficiência média de 21,36 pontos em relação aos que a iniciaram com 6 anos. Quando comparamos quem ingressou antes dos 2 anos e quem ingressou entre 3 e 4 anos aparece ainda uma relação relevante com a proficiência, embora menor: uma diferença nos pontos médios de 13,86, mostrando que a idade com que se inicia a escolarização tem impacto na proficiência.

As análises citadas mostraram que a frequência à pré-escola faz diferença nos resultados de testes cognitivos, como discutido inicialmente. Essa área de estudo no Brasil é ainda recente, os trabalhos disponíveis que analisam o desempenho da pré-escola brasileira estão sendo desenvolvidos há relativo pouco tempo, havendo ainda grandes possibilidades de estudo nessa área.

## **2 FONTES DE DADOS E METODOLOGIA**

Na seção 2.1 deste capítulo descrevem-se as bases de dados utilizadas para as análises propostas, SAEB e PNAD. Na seção 2.2 é considerada a amostragem complexa na utilização desses microdados. E, por fim, na seção 2.3 é apresentada a modelagem estatística utilizada para a análise da relação entre frequência à educação infantil e resultado educacional.

### **2.1 Fontes de dados**

#### **2.1.1 SAEB: frequência à educação infantil e proficiência escolar**

O Teste SAEB faz parte do Sistema Brasileiro de Avaliação da Educação Básica (SAEB). De acordo com o INEP (2008),

o SAEB foi fundamentado na idéia de sistema, com a perspectiva de desencadear um processo de avaliação, por meio de levantamento periódico de dados, para captar a evolução do quadro educacional brasileiro, a partir de dois pressupostos básicos: o rendimento dos alunos reflete a qualidade do ensino ministrado; e nenhum fator determina, isoladamente, a qualidade do ensino.

Além da mensuração da proficiência em testes, são aplicados aos alunos questionários sobre fatores contextuais, permitindo aos pesquisadores da área relacionar o rendimento escolar com possíveis determinantes. O SAEB, portanto, monitora o desempenho dos sistemas de ensino da educação básica, oferecendo informações que permitem a identificação de determinantes e deficiências para se atingir o sucesso educacional brasileiro.

O teste SAEB ocorre por amostragem desde 1990 e ocorre a cada dois anos avaliando conhecimentos em língua portuguesa e matemática de estudantes de 4<sup>a</sup> e 8<sup>a</sup> séries do ensino fundamental (ciclo de 8 anos) ou 5<sup>a</sup> e 9<sup>a</sup> ano (ciclo de 9 anos) e de alunos do 3<sup>o</sup> ano do ensino médio. São avaliados estudantes matriculados em escolas públicas (federais, estaduais e municipais) e privadas da área urbana. O teste engloba ainda os alunos da área rural da região Nordeste, de Minas Gerais e de Mato Grosso do Sul. Em 2005 foram incluídos na população da 4<sup>a</sup> série os alunos das escolas rurais não federais. Pela grande abrangência territorial, é feito por amostragem, oferecendo resultados apenas para Brasil, Grandes Regiões e Unidades da Federação.

Desde 2005 ocorre também a Prova Brasil, que assim como o teste SAEB, acontece a cada dois anos. A Prova engloba escolas da rede pública nas áreas urbanas, sendo quase universal, havendo, portanto, uma expansão do alcance dos resultados da avaliação. Por

utilizar a mesma metodologia de avaliação do teste SAEB, em 2007 as duas avaliações foram realizadas em conjunto. Na análise aqui proposta optou-se por utilizar o teste SAEB pelo fato desse apresentar resultados mais abrangentes considerando o território nacional (áreas rurais) e por abarcar escolas do ensino privado. Escolheu-se o ano de 2005 por ser o ano mais recente disponibilizado em microdados.

A nota da proficiência obtida pelos estudantes no teste SAEB são apresentadas em uma escala que varia de zero a 500. Cada uma das disciplinas tem uma interpretação específica da escala, que é única para as três séries avaliadas. As notas apontam os distintos graus de desenvolvimento de habilidades, competências e aquisição de conhecimentos pelos estudantes ao longo dos anos de estudo. Como a escala é cumulativa, as notas de estudantes de séries distintas não podem ser comparadas diretamente. Essa comparação só é possível considerando a mesma série, e a mesma disciplina, pois se entende que alunos em uma mesma série devem ter desenvolvimento de habilidades, competências e aquisição de conhecimentos semelhantes.

A pergunta em evidência na análise aqui proposta diz respeito ao início da frequência escolar. No questionário 2005, as respostas possíveis a essa pergunta para os alunos da 4ª série e 8ª série são diferentes (TABELA 2.1). Para análise foi considerado apenas, para ambas as séries, se o início das atividades escolares ocorreu antes ou depois do ensino fundamental.

**TABELA 2.1:** Variáveis sobre o início da educação escolar nos questionários dos respondentes de 4ª e 8ª séries do SAEB 2005

Questionário da 4ª série	
A042M028	Quando você começou a estudar?
	A No maternal
	B Na pre-escola
	C Na 1ª série
	D Na 2ª série
Questionário da 8ª série	
A082M032	Quando você começou a estudar?
	A Na pre-escola
	B Na 1ª série
	C Na 2ª série
	D Na 3ª série

Fonte: INEP (2008)

Os dados SAEB foram utilizados no capítulo 3 para investigar a evidência da influência positiva da educação infantil sobre os possíveis resultados socioeconômicos futuros de um indivíduo, aqui representado pela proficiência dos alunos de 4ª e 8ª série na prova de matemática de 2005. É importante considerar que testes de proficiência são comumente confundidos com testes de habilidades, todavia não são medidas puras senão um conjunto da inteligência cognitiva, não cognitiva e insumos do meio ambiente. Não é possível controlar totalmente as diferenças que existem nas habilidades não cognitivas entre grupos socioeconômicos e raciais utilizando um ajuste econométrico, segundo Cunha e Heckman (2010). Todavia essa diferença pode ser amplamente reduzida utilizando variáveis que controlem habilidade e educação da mãe, renda e estrutura familiar, pois o nível de habilidades das crianças é altamente correlacionado com os fatores do *background* familiar.

Outro ponto importante a se considerar, como o próprio INEP (2008) aconselha, é a cautela na identificação dos fatores associados ao desempenho dos alunos. “Não é estabelecida uma relação direta e, de certo modo, mecânica, entre diferentes variáveis para determinação desses fatores” (INEP, 2008). O instituto recomenda, ainda, a utilização das variáveis em análises que considerem a abordagem estatística em modelos multiníveis<sup>12</sup>, que exercem o melhor controle desses diferentes fatores.

Por fim, não foram incluídos na análise os alunos do 3º ano do ensino médio considerando um possível confundimento no “efeito escola” proveniente de mudanças de escolas entre o ensino fundamental e médio, que podem afetar o resultado do teste SAEB, porém não podem ser controlados nesta pesquisa.

### **2.1.2 PNAD: características familiares e domiciliares e frequência à educação infantil**

A Pesquisa Nacional por Amostra de Domicílios, PNAD, realizada anualmente pelo IBGE, investiga informações em domicílios particulares e coletivos, sobre assuntos socioeconômicos e demográficos relativos aos moradores e à habitação. Entre os assuntos recorrentes e mais relevantes estão características gerais como sexo, idade, cor/raça e educação, além de informações sobre trabalho e rendimento. Também são pesquisados temas com periodicidade variável, de acordo com demandas, como migração, fecundidade, nupcialidade e outros, que são apresentados na forma de suplementos. As informações levantadas pela PNAD constituem um importante instrumento para formulação, validação e

---

<sup>12</sup> A seção 2.3 explica o conceito de modelos multiníveis.

avaliação de políticas públicas voltadas para o desenvolvimento socioeconômico e a melhoria das condições de vida dos brasileiros.

A PNAD foi implantada progressivamente a partir de 1967, sendo aplicada trimestralmente até o primeiro trimestre de 1970, e, a partir de então, passou a ser aplicada anualmente (com exceção dos anos censitários 1970, 1980, 1991, 2000). Desde 1973 é aplicada em todo o Brasil, com exceção das áreas rurais da região Centro-Oeste, incluídas em 1982, e Norte, incluídas em 2004. Por razões especiais não foi realizada em 1974, 1975 e 1994. A amostra da pesquisa permite fornecer estimativas com precisão controlada apenas para estados e regiões metropolitanas (Belém, Fortaleza, Recife, Salvador, Belo Horizonte, Rio de Janeiro, São Paulo, Curitiba e Porto Alegre)<sup>13</sup>.

Considerando a questão de interesse, educação infantil, a PNAD possui desde 1992 questão sobre frequência à *escola* ou à *creche*. Até 1995, as questões sobre educação eram aplicadas apenas para as pessoas de 5 anos ou mais de idade, a partir de então as questões passaram a ser investigadas para todos os moradores do domicílio (QUADRO 2.1).

**QUADRO 2.1:** Variáveis sobre frequência atual à educação infantil na PNAD 1995 a 2009

	1995-1999	2000 <sup>1</sup>	2001-2006	2007-2009
Creche/pré-escolar	■			
Creche		■	■	■
Pré-escolar ou classe de alfabetização		■		
Pré-escolar			■	
Classe de alfabetização				■
Maternal, jardim de infância, etc				■

<sup>1</sup> Censo 2000

Fonte: IBGE (2011)

Entre 1995 e 1999 a frequência à *creche* era pesquisada juntamente com a *pré-escola*. No questionário da amostra do censo de 2000 foi captada a frequência à *creche* ou *pré-escolar/classe de alfabetização* separadamente. A PNAD de 2001 (em relação à PNAD de 1999) ampliou a pesquisa investigando separadamente *creche* e *pré-escolar* e inquirindo a respeito da rede de ensino (*pública* ou *particular*). E, por fim, em 2007, passou-se a pesquisar a frequência à *creche*, *classe de alfabetização* ou *maternal/jardim de infância/etc*. A PNAD de 2009 inova pesquisando a respeito da dependência administrativa (federal, estadual ou

<sup>13</sup> Para detalhes sobre o assunto veja IBGE (2010b)

municipal). Desde 2007 a educação infantil é classificada pelo IBGE entre *creche* e *pré-escolar*, sendo que *pré-escolar* compreende *maternal/jardim de infância/etc.* e *classe de alfabetização*. Considerou-se como *creche* o estabelecimento, juridicamente regulamentado ou não, destinado a dar assistência a crianças nas primeiras idades.

Foram utilizados os dados da PNAD de 1995 a 2009 para captar a frequência à educação infantil por crianças de 0 a 6 anos, e os dados da PNAD 2009 foram utilizados para descrever o contexto socioeconômico desses indivíduos, considerando características que a literatura considera adversas ao desenvolvimento das habilidades das crianças de 0 a 6 anos, apresentadas anteriormente no capítulo 1. As variáveis selecionadas para essa análise são melhor descritas no capítulo 4.

## **2.2 Análise de dados provenientes de amostras complexas**

A análise de uma pesquisa amostral é fundamentada na inferência sobre a estimação pontual do parâmetro de interesse, na precisão da estimativa pontual e no intervalo de confiança dessa estimativa. As pesquisas sociais normalmente utilizadas por pesquisadores da área social, como as que são produzidas pelo IBGE e pelo INEP, possuem plano amostral complexo, isto é, apresentam: i) probabilidades distintas de seleção das unidades; ii) conglomeração das unidades; e iii) estratificação. Essas três características da amostragem complexa influenciam a inferência estatística, levando à necessidade da inclusão desses fatores na análise dos dados.

Usualmente, pesquisadores utilizam somente os pesos amostrais individuais nas análises dos dados provenientes de pesquisas com amostragem complexa, o que corrige as estimativas pontuais de acordo com as probabilidades distintas de seleção. Porém, a consideração da conglomeração e estratificação é evitada, seja pelo desconhecimento, ou pela dificuldade do manejo do plano amostral, o que leva a subestimação da variância das estimativas dos parâmetros. É facilmente possível levar em conta o plano amostral da pesquisa na estimação dos parâmetros de modelos estatísticos através de alguns softwares estatísticos. O STATA, por exemplo, possui o conjunto do comando *svy*, criado especialmente para o manuseio de dados amostrais complexos.

Na análise de dados amostrais complexos, se não considerarmos o desenho do plano amostral trataremos os dados como provenientes de pesquisas censitárias ou amostras aleatórias simples. Se utilizarmos os pacotes usuais para análise amostral nos softwares estatísticos, trataremos os dados como independente e identicamente distribuídos (IID), ou

seja, provenientes de pesquisas com amostragem aleatória simples com reposição (SILVA e PESSOA, 1998). Quando incluído o peso amostral individual na análise temos como resultado estimativas que são corretas pontualmente, pois os estimadores dos parâmetros são influenciados pelos pesos distintos das observações. No entanto, como citado anteriormente, somente a utilização dos pesos geralmente não é suficiente para evitar a subestimação (ou superestimação) da variância das estimativas, resultado da influência da estratificação e conglomeração sobre as respectivas estimativas de variância.

Neste trabalho, todas as estimativas apresentadas são obtidas com métodos adequados para a análise de dados provenientes de uma pesquisa amostral complexa. São considerados os pesos, a estratificação e a conglomeração da amostra. Assim, as estimativas dos parâmetros de interesse possuem valores pontuais e correspondentes desvios padrão corretos. Abaixo são especificadas algumas medidas que serão apresentadas na avaliação das estimativas obtidas.

### **Coefficiente de variação**

A partir das estimativas de variância corretas é possível obter a precisão das estimativas a com base no coeficiente de variação (CV). O coeficiente de variação é uma avaliação do erro relativo nas estimativas que são provenientes de pesquisa amostral, e é dado por:

$$CV(\hat{\beta}) = \frac{\sqrt{\hat{V}(\hat{\beta})}}{\hat{\beta}}$$

onde  $\hat{\beta}$  é a estimativa de interesse, e  $\hat{V}(\hat{\beta})$  a variância estimada da estimativa  $\hat{\beta}$  ( $\sqrt{\hat{V}(\hat{\beta})}$  seria o erro padrão estimado). Assim, quanto menor a variância, menor será o coeficiente de variação, e mais precisa será a estimativa (TABELA 2.2). Os resultados deste trabalho serão apresentados juntamente com seus respectivos coeficientes de variação, como medida de precisão dos valores. Caso o erro padrão represente até 15% do parâmetro estimado, esse será considerado com boa precisão e adequado às análises objetivadas.

**TABELA 2.2:** Conceituação dos intervalos do coeficientes de variação (CV) das estimativas

Intervalos de valores do CV	Indicador	Conceito
0%	z	Exata
0 a 5%	a	Ótima
5 a 15%	b	Boa
15 a 30%	c	Razoável
30 a 50%	d	Pouco precisa
Mais de 50%	e	Imprecisa

Fonte: CARVALHO, 2006

### Efeito do plano amostral

Outra medida pertinente a planejamento amostral que será também utilizada é o Efeito do Plano Amostral Ampliado ( $EPA_A$ )<sup>14</sup>. O  $EPA_A$  mede o impacto da especificação incorreta do plano amostral sobre a estimativa de variância de  $\hat{\beta}$  em um modelo estatístico. É dado por, sob a hipótese de observações IID:

$$EPA_A(\hat{\beta}, v_0) = \frac{\hat{V}_{VERDADEIRO}(\hat{\beta})}{\hat{E}_{VERDADEIRO}(v_0)}$$

onde  $\hat{V}_{VERDADEIRO}(\hat{\beta})$  é a variância de  $\hat{\beta}$  sob o plano amostral verdadeiro, e  $\hat{E}_{VERDADEIRO}(v_0)$  é a variância esperada de  $\hat{\beta}$  sob a hipótese de Amostra Aleatória Simples com reposição, AAS ( $v_0 = \hat{V}_{AAS}(\hat{\beta})$ ).

Quanto mais afastado de 1 for o valor do  $EPA_A$ , maior será o impacto na estimação de  $\hat{\beta}$  por não se utilizar o plano amostral correto. Assim, conclui-se que quanto maior o  $EPA_A$ , a precisão da estimativa de  $\hat{\beta}$  sob hipótese IID é subestimada e, portanto, menor a probabilidade real de cobertura do intervalo de confiança caso o plano amostral não seja considerado na estimação. Caso contrário, quanto menor o  $EPA_A$ , a precisão da estimativa de  $\hat{\beta}$  é superestimada e maior a probabilidade real de cobertura do intervalo de confiança. Resumindo:

<sup>14</sup> O efeito do plano amostral considerado é o ampliado, ou *efeito da especificação incorreta do plano amostral (misspecification effect, meff)*. Para mais informações, veja (SILVA e PESSOA, 1998)

$EPA_A = 1 \rightarrow$  significância nominal igual a real

$EPA_A < 1 \rightarrow$  significância nominal menor que a real

$EPA_A > 1 \rightarrow$  significância nominal maior que a real

### Intervalo de confiança

A partir de estimativa pontual  $\hat{\beta}$  de um parâmetro  $\beta$ , pode-se construir um intervalo de confiança aproximado (ao nível de significância  $1-\alpha$ ) para o parâmetro com base na distribuição assintótica de:

$$t_0 = \frac{\hat{\beta} - \beta}{\sqrt{\sigma_0}}$$

onde  $\sigma_0$  é a variância do estimador  $\hat{\beta}$  considerando as observações IID, ou seja, as observações têm distribuição Normal padrão,  $t_0 \sim N(0;1)$ . Assim, o intervalo de confiança com significância aproximada de  $(1-\alpha)$  é dado por:

$$IC_{(1-\alpha)} = [\hat{\beta} - z_{\frac{\alpha}{2}} * \sqrt{\sigma_0}, \hat{\beta} + z_{\frac{\alpha}{2}} * \sqrt{\sigma_0}],$$

onde  $P(Z > z_\alpha) = \int_{z_\alpha}^{+\infty} \varphi(t) dt = \alpha$ , onde  $\varphi$  é a função de densidade da distribuição normal padrão.

No entanto, considerando a amostragem complexa (que não possui dados IID), a estatística de teste passa a ser:

$$t^*_0 = \frac{\hat{\beta} - \beta}{\sqrt{\hat{V}(\hat{\beta})}}$$

que possui a distribuição assintótica  $t^*_0 \sim N[0; EPA(\hat{\beta}; \sigma_0)]$ . Dependendo do valor do  $EPA(\hat{\beta}; \sigma_0)$ , o intervalo de confiança baseado na distribuição assintótica de  $t^*_0$  pode ser bem distinto daquele baseado na distribuição assintótica obtida sob a hipótese de observações IID. Assim, tem-se que o intervalo de confiança sob amostragem complexa é igual a:

$$IC_{(1-\alpha)} = [\hat{\beta} - z_{\frac{\alpha}{2}} * \sqrt{\sigma_0 * EPA}, \hat{\beta} + z_{\frac{\alpha}{2}} * \sqrt{\sigma_0 * EPA} ].$$

### 2.2.1 Amostragem complexa no SAEB

A metodologia de amostragem do SAEB é definida de maneira estratificada em três níveis, (estrato, escola e turmas), com probabilidades desiguais de seleção, baseada no Censo Escolar de 2004<sup>15</sup>.

O planejamento da amostra se inicia com a definição dos estratos, grupos de unidades amostrais mutuamente exclusivos. O primeiro critério para a estratificação refere-se à série em que o aluno está matriculado, 4<sup>a</sup> e 8<sup>a</sup> séries do ensino fundamental (ciclo de 8 anos) e 3<sup>o</sup> ano do ensino médio. O segundo critério importante é o de Unidade da Federação (26 Estados mais o Distrito Federal). A partir de então, em todas as combinações dos dois critérios anteriores é aplicado o terceiro critério, dependência administrativa (federal, estadual, municipal e particular<sup>16</sup>), o quarto critério, de localização dentro do estado (capital ou interior) e o último critério, referente à alocação segundo o tamanho da escola (uma ou duas turmas, três ou mais turmas). O QUADRO 2.2 descreve os estratos definidos. No SAEB 2005, ao combinar todos os critérios de estratificação, haveria 688 possíveis combinações. Todavia, devido ao número baixo ou nulo de unidades (escolas e/ou alunos) em determinadas combinações, foram decididas algumas agregações de estratos. Por fim, foram utilizados 438 estratos no desenho amostral da pesquisa.

O tamanho total da amostra de escolas e turmas para cada estrato é calculado e alocado proporcionalmente ao tamanho do número de turmas em cada escola<sup>17</sup>. Antes de efetuar a seleção das unidades a serem avaliadas, as escolas são ordenadas de acordo com os códigos de CEP correspondentes. Por fim, as escolas são sorteadas pelo método denominado Amostragem Sequencial de Poisson (ASP), e as turmas por Amostragem Aleatória Simples (AAS). A TABELA 2.3 apresenta a participação no levantamento.

<sup>15</sup> A coleta do SAEB é feita tempo depois da coleta do Censo Escolar, o que pode acarretar variações nos cadastros. É selecionada uma amostra reserva para o caso de escolas que deixaram de existir ou apresentaram turmas discordantes da pesquisa.

<sup>16</sup> Quando houve estratos em tamanho reduzido, reuniu-se estadual e municipal em uma única categoria intitulada “pública”.

<sup>17</sup> Para maiores detalhes, sugere-se INEP (2008).

**QUADRO 2.2:** Critérios para definição dos estratos de interesse no SAEB 2005

Critério	Número	Descrição
<b>Urbano</b>		
Série	3	4 <sup>as</sup> , 8 <sup>as</sup> , 3 <sup>ª</sup>
Estado	27	26 estados + DF
Rede	4	Federal, Estadual, Municipal e Particular
Localização	2	Capital ou Interior
<b>Rural</b>		
Série	1	4 <sup>as</sup>
Região	5	Regiões geográficas
Rede	4	Federal, Estadual, Municipal e Particular
Localização	2	Capital ou Interior

Fonte: INEP (2008)

**TABELA 2.3:** Participação no levantamento SAEB 2005

Estrato	Escolas participantes	Turmas participantes	Alunos com aplicação do teste	
			Matemática	Português
4 <sup>a</sup> série	2.837	3.607	41.783	39.594
8 <sup>a</sup> série	1.955	2.706	33.189	33.164

Fonte: INEP (2008)

Verifica-se, com base na descrição do plano amostral do SAEB, que o uso de estratificação e de probabilidades desiguais de seleção fazem com que seja necessário levar em conta o desenho amostral da pesquisa nos procedimentos de análise. Riether e Rauter (2000) mostram que quando os pesos e o desenho amostral são desprezados em uma análise do SAEB, obtêm-se estimativas incorretas. Quando considerados apenas os pesos, as estimativas pontuais se tornam corretas, mas os respectivos erros-padrão são subestimados. Conseqüentemente, os dados provenientes de amostras complexas como a do SAEB devem ser ponderados utilizando-se os pesos amostrais e procedimentos de expansão específicos da teoria da amostragem. Assim, o número total e as quantidades de interesse estimados para escolas, turmas e alunos correspondem aos obtidos no Censo Escolar 2004 (após aplicação das regras que definiram o universo de seleção do SAEB 2005).

### 2.2.2 Amostragem complexa na PNAD

A PNAD é uma pesquisa amostral com probabilidades desiguais de seleção e que apresenta um, dois ou três estágios de conglomeração (municípios, setores censitários e domicílios), dependendo do estrato. A definição e seleção da amostra são realizadas utilizando a malha setorial do censo mais recente (a PNAD da década de 90 utiliza a malha do censo de 91 e a PNAD da década de 2000 utiliza a malha do censo de 2000).

Segundo Silva, Pessoa e Lila (2002), o processo de planejamento da amostra da PNAD é realizado em duas etapas. Na primeira, definem-se 36 estratos naturais que são divididos em dois grupos. No primeiro grupo se encontram 9 regiões metropolitanas, a saber: Porto Alegre, Curitiba, São Paulo, Rio de Janeiro, Belo Horizonte, Salvador, Recife, Fortaleza e Belém. No outro grupo se encontram os demais municípios dos estados referentes às regiões metropolitanas citadas anteriormente, mais as outras 18 UF's (SC, ES, GO, MS, MT, TO, AM, RO, RR, AC, AP, SE, AL, PB, RN, PI, MA e DF).

No segundo momento, ao primeiro grupo se aplica uma amostragem estratificada por município e conglomerada em dois estágios, sendo a unidade primária de seleção (UPA) o setor censitário e a unidade secundária de seleção (USA) o domicílio. Em cada UPA aplica-se amostragem sistemática com probabilidades proporcionais ao tamanho do setor. Os setores são ordenados segundo situação (urbano ou rural) e por um código geográfico, “conferindo um efeito de estratificação implícita por situação devido ao uso do sorteio sistemático” (Silva, Pessoa e Lila, 2002).

Ao segundo grupo é aplicada uma amostragem conglomerada em três estágios, sendo a unidade primária de seleção (UPA) o município, a unidade secundária de seleção (USA) o setor censitário e a unidade terciária de seleção (UTS) o domicílio. Em cada município, é aplicada a amostragem sistemática com probabilidade proporcional ao tamanho populacional. Alguns municípios possuem probabilidade de seleção igual a 1 e são chamados *autorrepresentativos*, os demais são denominados *não autorrepresentativos*. No segundo estágio, setores censitários são selecionados com amostragem sistemática e probabilidades proporcional ao número de domicílios.

A seleção dos municípios e setores censitários em ambos os grupos é realizada apenas uma vez na década, sendo apenas os domicílios selecionados anualmente. Para isso, antes da fase de seleção dos mesmos é feita uma listagem dos domicílios existentes dentro do setor sorteado, para fins de atualização do cadastro. Esse cadastro é complementado pelas unidades domiciliares existentes em conjuntos residenciais, edifícios e favelas com 30 ou mais

unidades residenciais, que tenham surgido, nestes municípios, após a realização do Censo Demográfico de 2000 (nomeado de Cadastro de Projetos de Novas Construções). Os domicílios são, então, selecionados por meio de amostragem sistemática simples, com equiprobabilidade.

As probabilidades de seleção dos municípios dentro de cada um dos 36 estratos são variáveis e, portanto, a ponderação da amostra é variável. Assim, a expansão da amostra para toda a população é feita pela utilização de pesos amostrais e estimadores de razão.

O QUADRO 2.3 apresenta informações retiradas de Silva, Pessoa e Lila (2002), sendo uma boa síntese do plano amostral da PNAD e quais seriam as variáveis de ponderação da amostragem.

**QUADRO 2.3:** Plano amostral da pesquisa PNAD

	Área 1: Região metropolitana	Área 2: Municípios autorrepresentativos	Área 3: Municípios não autorrepresentativos
Estrato	Município	Município	Grupo de municípios
UPA	Setor	Setor	Município
USA	Domicílio	Domicílio	Setor

Fonte: Silva, Pessoa e Lila (2001)

Na análise dos microdados desta dissertação, a amostragem complexa da PNAD 2009 foi considerada a partir das variáveis de estrato (V4617 - STRAT) e de unidade primária de amostragem, UPA, (V4618 - PSU). Essas variáveis são encontradas no arquivo de microdados dos domicílios, sendo necessário fazer a correspondência nos microdados referentes a pessoas. Através do software STATA, o plano amostral foi incluído no cálculo das estimativas de interesse. O APÊNDICE A apresenta as rotinas utilizadas para a inclusão da amostragem complexa nos procedimentos de estimação realizados no software STATA.

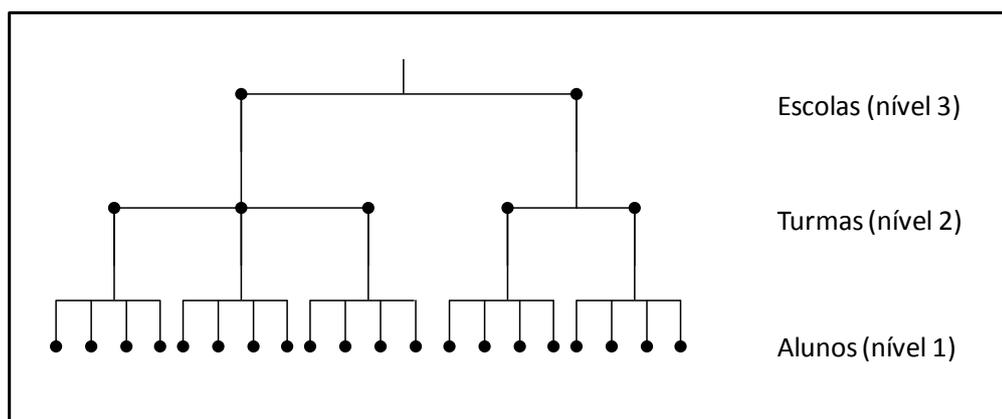
## 2.3 Modelagem estatística

### 2.3.1 Modelo hierárquico de dois níveis

Em diversas pesquisas sociais e econômicas é comum que as informações coletadas assumam uma estrutura de agrupamento, ou hierarquia. Usualmente, pesquisas educacionais que objetivam explicar desempenho, repetência ou algum outro resultado educacional

possuem tal estrutura na qual alunos podem ser agrupados em turmas, que são agrupadas em escolas, por exemplo. Cada degrau da hierarquia é denominada nível (FIGURA 2.1), podendo haver dados agrupados em dois, três ou mais níveis. Os modelos hierárquicos, ou multiníveis, são modelos estatísticos aplicados a dados que apresentam essa estrutura hierárquica, permitindo que sejam modelados nos diferentes níveis de observação.

**FIGURA 2.1:** Exemplo de dados com estrutura hierárquica de 3 níveis



As variáveis incluídas em um modelo multinível podem ser medidas em seu próprio nível ou podem ser agregadas (ou desagregadas) para serem medidas em outro nível. Por exemplo, as notas de alunos em um determinado teste podem ser agregadas para representar a média dos alunos de uma determinada turma ou escola nesse teste. O mesmo pode ser feito com o nível socioeconômico dos alunos, que quando agrupado gera o nível socioeconômico da turma ou escola. Assim, as características do nível agregado podem ser atribuídas aos indivíduos pertencentes a um dado grupo. Esse modelo leva em consideração que indivíduos são influenciados por diferentes contextos em que estão inseridos, e que esses contextos são caracterizados de forma hierárquica, com diferentes níveis.

Os modelos multiníveis devem ser usados em dados com estrutura hierárquica, primeiramente, para captar a heterogeneidade latente de cada subgrupo. No caso dos dados escolares, por exemplo, cada escola influencia seus alunos de maneira distinta, resultando em diferentes produtos educacionais. Albernaz *et al.* (2002) apontam que “ (...) é demasiado restritivo impor que os coeficientes relacionando as variáveis independentes e a variável dependente sejam constantes entre os subgrupos (no caso educacional, entre escolas)”. Segundo, pela ausência de independência das observações devido ao efeito de grupo, as observações consideradas nos vários níveis não são independentes. Esta (Essa??) é justamente a relação que os modelos multiníveis levam em consideração pois objetivam o tipo de análise

que determina o efeito direto de variáveis independentes controladas por variáveis agrupadas. Ou seja, permitem investigar se os níveis superiores servem como moderadores para as relações nos níveis inferiores. E por fim, pela estrutura de agregação dos dados em subgrupos, que faz com que a correlação média entre os indivíduos inseridos em um mesmo grupo seja maior do que a correlação média entre as variáveis de indivíduos de grupos distintos, pois indivíduos agrupados estão sujeitos à influência dos mesmos fatores.

Análises educacionais por modelos tradicionais de regressão linear de 1 nível apontam forte relação entre as características dos alunos e suas famílias e o desempenho escolar. Esses resultados obtidos revelam a pouca importância das variáveis educacionais, *vis-à-vis* as características da família. De fato, tais resultados se devem à metodologia linear utilizada, que não capta a estrutura hierárquica dos dados nem a flexibilidade da estimação de equações separadas para cada nível. Assim, o modelo hierárquico é justificado por não subestimar os efeitos escolares no desempenho dos alunos, dado que a modelagem hierárquica resulta em erros padrões adequados à estrutura dos dados e, conseqüentemente, intervalos de confiança e testes de significância mais precisos.

Albernaz *et al.* (2002), utilizando a metodologia hierárquica, apontam que 80% da variância do desempenho entre as escolas brasileiras resulta da diferença no nível socioeconômico médio dos alunos da escola. Depois de controlado o efeito socioeconômico, diferenças no desempenho escolar são devidas a quantidade e qualidade dos insumos escolares. Machado *at al.* (2006) encontram evidências contrárias na avaliação da qualidade do ensino de matemática nas escolas públicas estaduais em Minas Gerais, corroborando o fato de que o efeito da escola e do município sobre o desempenho do aluno é reduzido frente às características do aluno e do *background* familiar, mesmo utilizando o modelo hierárquico.

**O modelo hierárquico de dois níveis será utilizado nesta dissertação para avaliação do impacto da frequência à educação infantil nos resultados escolares futuros utilizando os dados do SAEB 2005.** Albernaz *et al.* (2002) pontuam a recomendação da aplicação deste tipo de análise quando considerada a função de produção educacional por ser um modelo adequado à estrutura dos dados referentes à avaliação educacional. Machado *et al.* (2006) citam a “importância em se considerar fatores medidos em diferentes níveis para se explicar o desempenho ou algum outro produto educacional, como frequência ou progressão.” O próprio INEP (2008) aconselha cautela na identificação dos fatores associados ao desempenho dos alunos, pois não é estabelecida uma relação direta e, de certo modo, mecânica, entre diferentes variáveis para determinação desses fatores. O instituto recomenda a

manipulação das variáveis em modelos multiníveis, que exercem o melhor controle desses diferentes fatores.

A modelagem hierárquica pode ser representada como uma extensão da modelagem linear de 1 nível na qual, por exemplo, interceptos e inclinações das retas de regressões estimadas podem variar por grupo. Considere um modelo de regressão linear com apenas uma variável explicativa.

Seja  $Y_i$  a variável de interesse para o  $i$ -ésimo aluno tal que,

$$Y_i = \beta_0 + \beta_1 X_i + e_i$$

Considere agora um modelo com intercepto variável entre grupos (escolas) no qual  $Y_{ij}$  é o valor da variável de interesse para o  $i$ -ésimo aluno da  $j$ -ésima escola:

$$Y_{ij} = \beta_{0j} + \beta_1 X_{ij} + e_{ij}$$

Adicionalmente, pode-se também definir um modelo com intercepto e inclinação aleatórios (que variam entre os grupos):

$$Y_{ij} = \beta_{0j} + \beta_{1j} X_{ij} + e_{ij}$$

Conceitualmente o modelo é visto como um sistema hierárquico de equações no qual especificam-se equações para cada um dos níveis em estudo, por exemplo, alunos (nível 1) e escolas (nível 2). O modelo hierárquico de dois níveis mais simples é o modelo nulo, ou **modelo de análise de variância com intercepto aleatório**, que apresenta o efeito aleatório no intercepto e nenhuma variável explicativa, de modo a estimar quanto da variação observada na variável de interesse decorre do primeiro nível e quanto decorre do segundo nível, permitindo a decomposição da variância entre os diferentes níveis do modelo.

Considerando o exemplo das notas em um teste de proficiência de alunos agrupados em escolas, seja  $Y_{ij}$  a nota do  $i$ -ésimo na  $j$ -ésima escola. O modelo de intercepto aleatório é dado por:

$$Y_{ij} = \beta_{0j} + e_{ij} \quad (2.1)$$

$$\beta_{0j} = \gamma_{00} + u_{0j} \quad (2.2)$$

tal que, substituindo 2.2 em 2.1:

$$Y_{ij} = \gamma_{00} + u_{0j} + e_{ij}$$

onde:

$\gamma_{00}$  = nota média dos alunos entre todos os grupos

$u_{0j}$  = efeito aleatório da escola j na nota média

$e_{ij}$  = erro aleatório associado ao aluno i na escola j na nota média

Usualmente se assume que os resíduos  $e_{ij}$  e  $u_{0j}$  têm distribuição normal e variância  $\sigma_e^2$  e  $\sigma_u^2$  tal que  $e_{ij} \sim N(0, \sigma_e^2)$  e  $u_{0j} \sim N(0, \sigma_u^2)$ . A medida da variância atribuída à correlação entre os alunos da mesma escola, ou seja, explicada pelas características escolares (nível 2) é chamada de coeficiente de correlação intra-classe,  $\rho$ :

$$\text{Var}(Y_{ij}) = \text{Var}(e_{ij} + u_{0j}) = \sigma_e^2 + \sigma_u^2$$

$$\rho = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2}$$

O valor do coeficiente de correlação intra-classe pode variar de 0 (não há diferenças entre os grupos) até 1 (não há diferença dentro dos grupos).

No processo de elaboração de um modelo multinível, variáveis explicativas são adicionadas ao modelo em ambos os níveis e as variáveis no primeiro nível podem ter efeitos aleatórios. Quando isso ocorre são gerados **modelos com intercepto e coeficientes (inclinação) aleatórios**.

Uma simplificação do caso geral de modelos hierárquicos de dois níveis pode ser definido considerando-se uma variável explicativa  $X$  no nível 1 e uma variável explicativa  $Z$  no nível 2, e efeitos aleatórios no intercepto e na variável explicativa de nível 1, dado pelas seguintes equações:

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + e_{ij} \quad (2.3)$$

$$\beta_{0j} = \gamma_{00} + \gamma_{01}Z_j + u_{0j} \quad (2.4)$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}Z_j + u_{1j} \quad (2.5)$$

A equação (2.3) refere-se ao nível 1, alunos. O modelo atribui a cada subgrupo um intercepto e um efeito da variável  $X$  diferentes, condicionados às características  $Z$  da escola. Consolidando as equações (2.3), (2.4) e (2.5), obtemos:

$$Y_{ij} = (\gamma_{00} + \gamma_{01}Z_j) + (\gamma_{10}X_{ij} + \gamma_{11}Z_jX_{ij}) + (u_{1j}X_{ij} + u_{0j} + e_{ij})$$

onde:

$y_{ij}$ : resultado do teste SAEB para o aluno  $i$  da escola  $j$

$X_{ij}$ : variável de início da educação escolar (1 = educação infantil e 0 = ensino médio)

$Z_j$ : variável indicadora da forma de gestão da escola (1 = pública e 0 = privada)

$\gamma_{00}$ : nota média dos alunos em escola privada

$\gamma_{01}$ : diferença média na nota dos estudantes de escola pública em relação à privada

$\gamma_{10}$ : efeito na nota média dos estudantes de escola privada que frequentaram educação infantil

$\gamma_{11}$ : diferença média na nota dos estudantes em escolas públicas que frequentaram educação infantil em relação aos estudantes em escolas privadas

$u_{0j}$ : efeito aleatório da escola  $j$  na nota média

$u_{1j}$ : efeito aleatório da escola  $j$  na nota média de quem frequentou educação infantil

$e_{ij}$ : erro aleatório associado ao aluno  $i$  da escola  $j$

Esse modelo também é comumente denominado modelo de efeitos mistos por ser composto de efeitos fixos ( $\gamma_{00}, \gamma_{01}, \gamma_{10}, \gamma_{11}$ ) e efeitos aleatórios ( $u_{0j}, u_{1j}, e_{ij}$ ). Os termos  $u_{0j}$  e  $u_{1j}$  são os erros correspondentes ao segundo nível e comum aos indivíduos do mesmo grupo. A média de cada termo de erro é igual a zero, assim como a do termo composto ( $u_{0j} + u_{1j} + e_{ij}$ ), cuja variância incorpora a variabilidade remanescente não explicada pelos efeitos fixos. Assume-se que esses termos de erros são independentes do termo de erro do primeiro nível,  $e_{ij}$ , mas que a covariância entre  $u_{0j}$  e  $u_{1j}$  é diferente de zero.

Os parâmetros  $\beta_{0j}$  e  $\beta_{1j}$  permitem que os interceptos e as inclinações das retas de regressão de cada escola variem aleatoriamente. Assim, a nota de cada estudante pode não depender apenas de suas características individuais, mas também das características das

escolas (por exemplo, se é pública ou privada). A variabilidade do coeficiente  $\beta_{ij}$ , ou o intervalo de confiança, pode ser calculada a partir da sua variância. Dado o modelo de intercepto aleatórios, representado pelas equações 2.1 e 2.2, a variância do coeficiente  $\beta_{0j}$  é:

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

O intervalo de confiança do coeficiente  $\beta_{0j}$  entre os subgrupos (por exemplo as notas médias de cada escola) é dado por:

$$\beta_{0j} \pm 1,96 * \sqrt{var(\beta_{0j}) * EPA(\beta_{0j})}$$

Por fim, o método de estimação usualmente utilizado para estimar modelos hierárquicos é o método de Máxima Verossimilhança (ML), que é geralmente robusto e produz estimativas assintoticamente eficientes e consistentes.

### 2.3.2 Aplicação de modelos hierárquicos aos dados do SAEB

Na análise multinível da proficiência no teste SAEB 2005 condicionada à frequência à educação infantil serão analisados modelos de efeitos fixos e aleatórios, considerando variáveis explicativas de primeiro e segundo nível. As variáveis escolhidas para o modelo (TABELA 2.4) tentam refletir o melhor e mais objetivamente possível as características de *background* familiar e do aluno, observando as limitações impostas pela base de dados.

Os modelos escolhidos para análise consideraram poucas características das escolas na determinação do desempenho escolar, buscando apenas incluir variáveis de controle. Aqui não se deseja determinar os fatores e características da escola associados ao bom desempenho do aluno, mas sim identificar quais seriam as características dos alunos que determinam seu desempenho (mais especificamente, qual o efeito da educação infantil) controlando-se o efeito de escola.

A variável *renda* foi construída utilizando um sistema de categorização semelhante ao Critério Brasil (ABEP, 2011), excluindo-se a utilização da variável de escolaridade da mãe devido ao grande número de valores faltantes na base SAEB 2005. Os respondentes obtiveram um escore a partir da posse de televisão, rádio, videocassete, carro, máquina de lavar roupa, geladeira, freezer, banheiros e empregadas (variáveis disponíveis no SAEB) e

foram classificados por quintis de renda, e não em classes sociais tal como propõe o Critério Brasil.

**TABELA 2.4:** Variáveis do SAEB 2005 utilizadas na estimação dos modelos hierárquicos

Variável	Descrição	Possíveis repostas
<b>Nível 1</b>		
<i>inicio</i>	Quando o respondente começou a estudar	1 - maternal/pré-escola 0 - 1ª/2ª/3ª série
<i>sexo</i>	Sexo do respondente	1 - feminino 0 - masculino
<i>cor_raça</i>	Cor/raça declarada pelo respondente	1 - não branco 0 - branco
<i>repetência</i>	Ocorrência de repetência 1 ou mais vezes	1 - sim 0 - não
<i>exp_particular</i>	Estudou alguma vez em escola particular	1 - sim 0 - não
<i>dever</i>	Faz dever de matemática	1 - não 0 - sim
<i>trabalha</i>	Trabalha	1 - sim 0 - não
<i>renda</i>	Quintil de renda	1 - 1º, 2º e 3º quintil 0 - 4º e 5º quintil
<i>esc_mãe</i>	Escolaridade da mãe	1 - inferior ao Ensino Médio 0 - ao menos Ensino Médio
<i>esc_pai</i>	Escolaridade do pai	1 - inferior ao Ensino Médio 0 - ao menos Ensino Médio
<i>dicionário</i>	Possui dicionário em casa	1 - não 0 - sim
<i>densidade_domiciliar</i>	Número de pessoas residindo no domicílio	1 - Até 5 pessoas 0 - Mais que 5 pessoas
<b>Nível 2</b>		
<i>rede</i>	Rede de ensino da escola atual	1 - pública 0 - privada
<i>local<sup>1</sup></i>	Localização da escola	1 - rural 0 - urbano
<i>regiao</i>	Região	1 - NO/NE/CO 0 - SU/SE + DF

<sup>1</sup> somente para 4ª série

Os modelos estimados são especificados no APÊNDICE B. A análise se inicia com o modelo de análise de variância com efeitos aleatórios de modo a obter a informação acerca de quanto da variação observada no desempenho escolar decorre da variação entre as escolas e quanto está associada às características do aluno. O modelo 2 inclui a variável de **início da educação escolar, *inicio***, e o modelo 3 inclui outras características *X* do aluno e as

covariáveis *Z* no nível da escola, objetivando o controle da variabilidade entre subgrupos. Todos esses modelos possuem efeito aleatório apenas no intercepto, com a finalidade de permitir a decomposição das variâncias de nível 1 e de nível 2 condicionadas a características do aluno e das escolas. A TABELA 2.5 resume os modelos estimados.

**TABELA 2.5:** Efeito das variáveis utilizadas nos modelos hierárquicos estimados

	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
<b>Nível 1</b>					
intercepto	aleatório	aleatório	aleatório	aleatório	aleatório
<i>inicio</i>	-	fixo	fixo	aleatório	aleatório
<i>sexo</i>	-	-	fixo	-	fixo
<i>cor_raça</i>	-	-	fixo	-	fixo
<i>repetência</i>	-	-	fixo	-	fixo
<i>exp_particular</i>	-	-	fixo	-	fixo
<i>dever</i>	-	-	fixo	-	fixo
<i>trabalha</i>	-	-	fixo	-	fixo
<i>renda</i>	-	-	fixo	-	fixo
<i>esc_mãe</i>	-	-	fixo	-	fixo
<i>esc_pai</i>	-	-	fixo	-	fixo
<i>dicionário</i>	-	-	fixo	-	fixo
<i>densidade_domiciliar</i>	-	-	fixo	-	fixo
<b>Nível 2</b>					
<i>rede</i>	-	-	fixo	-	fixo
<i>local</i> <sup>1</sup>	-	-	fixo	-	fixo
<i>regiao</i>	-	-	fixo	-	fixo

<sup>1</sup> somente para 4ª série

Em seguida, desenvolve-se a análise dos modelos com intercepto e coeficiente aleatório. É considerado o efeito aleatório na variável de início da educação escolar, *início* (além do intercepto). Espera-se, nesta situação, identificar aspectos relativos às diferenças no desempenho da frequência à educação infantil no teste SAEB 2005 entre as escolas. O modelo 4 inclui apenas a variável de nível 1 *início*, e o modelo 5 inclui outras características *X* do aluno e as covariáveis *Z* no nível da escola.

### 2.3.3 Incorporação do plano amostral do SAEB na estimação do modelo hierárquico de dois níveis

A estimação dos modelos hierárquicos deve considerar o desenho complexo da amostragem do SAEB 2005, como já discutido, para que não haja subestimação nas estatísticas de variância dos parâmetros que é o foco desta análise. Para tal, foi utilizado o pacote *GLLAMM*<sup>18</sup> do *software STATA*, que permite a inclusão da amostragem complexa no modelo hierárquico de dois níveis.

Estudos incluindo amostragem complexa em modelos lineares hierárquicos são ainda recentes, havendo pouca documentação sobre quais metodologias e procedimentos são adequados para essa estimação. Chantala *et al.* (2006) indicam a necessidade de que, na estimação em modelos hierárquicos, os pesos amostrais derivados da amostragem complexa sejam construídos de maneira diferente em comparação com a construção dos pesos utilizados em modelos lineares com apenas um nível. Não há, no entanto, consenso sobre qual o melhor método para a computação desses pesos.

Os pacotes estatísticos que operam análises multiníveis recomendam a utilização de reescalonamento baseado em diferentes metodologias, e cabe aos usuários se certificarem de que os pesos sejam reescalados como recomendado. A metodologia referenciada para utilização do *GLLAMM* é a descrita por Pfeiffermann (1998) segundo Rabe-Hasketh e Skrondal (2006).

A correção dos pesos utilizada nesta dissertação foi realizada através do pacote *PWIGLS* do *STATA*, que opera o reescalonamento de pesos e a estimação do modelo utilizando a metodologia adequada. Os componentes necessários para o reescalonamento são os identificadores do desenho amostral (variáveis associadas aos estágios de seleção de uma amostra conglomerada) e os pesos amostrais das Unidades Primárias de Amostragem e Unidades Finais de Amostragem (TABELA 2.6).

---

<sup>18</sup> Generalized Linear Latent and Mixed Models. [www.gllamm.org](http://www.gllamm.org)

**TABELA 2.6:** Componentes necessários para o reescalonamento de pesos amostrais na análise multinível segundo a metodologia de Pfeffermann (1998)

Nível	Variável	Descrição
1	UPA_id	Identificador da Unidade Primária de Amostragem
	UPA_peso	Peso amostral da Unidade Primária de Amostragem, no caso são as escolas. Indica quantos clusters (escolas) na população estão sendo representadas pela observação.
2	UFA_id	Identificador da Unidade Final de Amostragem
	UFA_peso	Peso amostral da Unidade Final de Amostragem dentro da UPA, no caso, peso do aluno dentro da escola. Indica quantos membros da população da UPA são sendo representados pela observação

Fonte: Chantala *et al* (2006)

**TABELA 2.7:** Pesos amostrais disponíveis na base de dados do SAEB 2005

Variável	Descrição
peso_at	Peso do aluno na turma
peso_ac	Peso do aluno por disciplina
peso_ec	Peso da turma
peso_tc	Peso da escola

Fonte: INEP, 2008

A base de dados do SAEB 2005 fornece quatro tipos distintos de pesos amostrais, peso do aluno na turma, peso do aluno por disciplina, peso da turma e peso da escola (TABELA 2.7). A variável *peso\_tc* (peso da escola) corresponde ao peso amostral da Unidade Primária de Amostragem, *UPA\_peso*, componente necessário no reescalonamento segundo Pfeffermann (1998).

Adicionalmente, o outro peso amostral necessário para o reescalonamento (segundo Pfeffermann, 1998) é aquele associado às Unidades Finais de Amostragem, *UFA\_peso*, que seria correspondente ao peso do aluno na escola. Esse peso não está disponível na base de dados do SAEB 2005, mas pode ser obtido a partir da fórmula de cálculo dos pesos usualmente utilizada para obtenção de estimativas em pesquisas amostrais conglomeradas em 2 estágios.

Seja  $W_{i,j}$  o peso final associado ao aluno que é calculado para cada disciplina (*peso\_ac*, peso do aluno por disciplina na base de dados do SAEB 2005), onde  $W_{i,j}$  pode ser obtido combinando-se o peso da escola  $W_j$  (associado à probabilidade de uma escola ser

selecionada) e o peso do aluno na escola  $W_{ij}$  (associado à probabilidade de um aluno ser selecionado dado que sua escola está na amostra do SAEB). Consequentemente:

$$W_{i,j} = W_j * W_{i|j}$$

Utilizando-se os descritores da Tabela 2.6, tem-se:

$$W_{i,j} = UPA\_peso_j * UFA\_peso_{i|j}$$

A partir da equação acima pode-se calcular o peso do aluno  $i$  na escola  $j$  :

$$UFA\_peso_{i|j} = \frac{W_{i,j}}{UPA\_peso_j}$$

O peso associado ao aluno em cada escola da amostra é essencial para permitir que o desenho amostral do SAEB seja devidamente incorporado na estimação dos modelos multiníveis utilizados nesta dissertação. Entretanto, tal informação não está diretamente disponível na base dados do SAEB mas pode ser obtida dado que *peso\_ac*, peso do aluno por disciplina na base SAEB2005 corresponde a  $W_{i,j}$  e a variável *peso\_tc* (peso da escola) corresponde ao peso amostral da Unidade Primária de Amostragem, *UPA\_peso*. Sendo assim calcula-se, a partir das informações da base de dados do SAEB 2005:

$$UFA\_peso_{i|j} = \frac{peso\_ac}{peso\_tc}$$

De posse dos componentes adequados ao reescalonamento, esse é empreendido pelo pacote *PWIGLS*, que fornece os pesos corrigidos que são empregados para a estimação dos multiníveis apresentados nesta dissertação. O passo seguinte trata da estimação dos modelos hierárquicos. Para tal utilizou-se o pacote *GLLAMM* do *software STATA*, que considera o caráter hierárquico dos dados, bem como incorpora as especificidades da amostragem complexa no procedimento de estimação. O *GLLAMM* não é, em geral, o pacote mais recomendado quando a variável resposta tem supostamente uma distribuição normal (como no caso da variável de interesse nesta dissertação). Isso decorre do fato de que o *GLLAMM* faz a estimação baseado em aproximações por quadratura, método que faz com que a estimação no

*STATA* seja demorada (dependendo do modelo estimado, pode-se levar horas ou até dias). Todavia, ressalta-se que, apesar da variável resposta dos modelos utilizados nesta dissertação ter uma distribuição normal, optou-se ainda assim pelo *GLLAMM* pela possibilidade que esse pacote possui de incluir a amostragem complexa na estimação dos modelos. No APÊNDICE C estão descritas as rotinas utilizadas para o reescalonamento dos pesos e para estimação dos modelos descrito na seção 2.3.2 incorporando o desenho amostral.

### **3 EDUCAÇÃO INFANTIL E PROFICIÊNCIA ESCOLAR: SAEB 2005**

Foi mencionado no capítulo inicial que a frequência à educação infantil tem importância significativa nos resultados socioeconômicos futuros por ser a educação referente, justamente, à fase em que as habilidades individuais começam a ser desenvolvidas. Assim, considerando ciclo dinâmico do desenvolvimento das habilidades, quando essa educação não ocorre (ou não ocorre apropriadamente), o indivíduo em questão tem comprometido todo o desenvolvimento necessário para um bom desempenho socioeconômico na fase adulta.

Este capítulo mostra a análise da relação entre a frequência à educação infantil e os resultados escolares ao fim do 1º ciclo do ensino fundamental e ao fim do ensino fundamental (4ª e 8ª séries do ciclo de 8 anos), observando a teoria do ciclo de desenvolvimento das habilidades. Como discutido no capítulo 1, a observância das limitações dos dados disponíveis para este tipo de análise é de fundamental importância na relação pretendida. Aqui utiliza-se os dados do SAEB 2005.

São utilizadas duas bases para estimação dos modelos hierárquicos nessa análise, uma referente aos respondentes da prova de matemática na 4ª série, com 37.187 observações (89% da base total), dados que representam uma população de 3.318.754 estudantes desse nível, e outra referente aos respondentes da prova de matemática na 8ª série, com 32.151 observações (96,93% da base total), correspondendo a uma população de 2.804.805 estudantes. A utilização de parte da base total se justifica pela inexistência de informações completas para cada indivíduo, ou seja, a existência de valores faltantes (*missings*). As observações selecionadas para as bases possuem informações completas para todas as variáveis utilizadas nas análises.

#### **3.1 Características dos alunos de 4ª e 8ª séries e frequência à educação infantil nos dados SAEB 2005**

O sistema de ensino brasileiro pressupõe idades ideais para a frequência de cada nível de ensino, por exemplo, as idades consideradas ideais para a frequência a 4ª e 8ª série são 10 e 14 anos. Todavia, é comum haver estudantes que possuem idades superiores à adequada para série que frequentam. Considera-se que está no nível errado aquele estudante que possui  $X+2$  anos de idade, sendo  $X$  a idade ideal para o nível de ensino. A partir da idade dos estudantes e do nível de ensino que frequentam é possível derivar o indicador de distorção idade/série. De

acordo com Rigotti e Cerqueira (2004), essa taxa mostra a inadequação da idade do aluno para a série que ele frequenta. Ela corresponde ao total de matrículas de pessoas que estão cursando determinada série em idade superior à considerada ideal sobre o total de matrículas na série em questão:

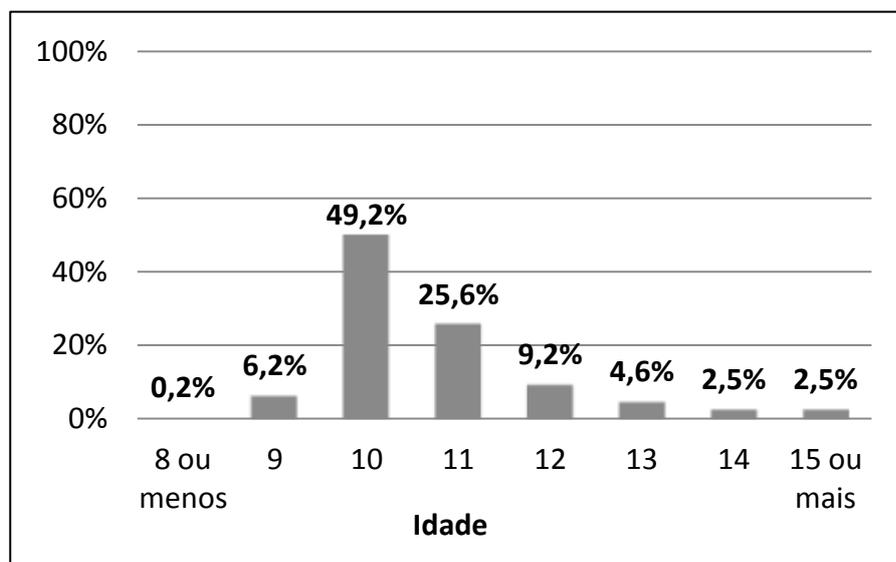
$$\text{Distorção idade/série} = \frac{M_{jsi\_sup}}{M_{js}}$$

$M_{jsi\_sup}$  = número de matrículas de pessoas com idade superior à idade adequada de estar cursando uma determinada série  $s$  do nível de ensino  $j$ ;

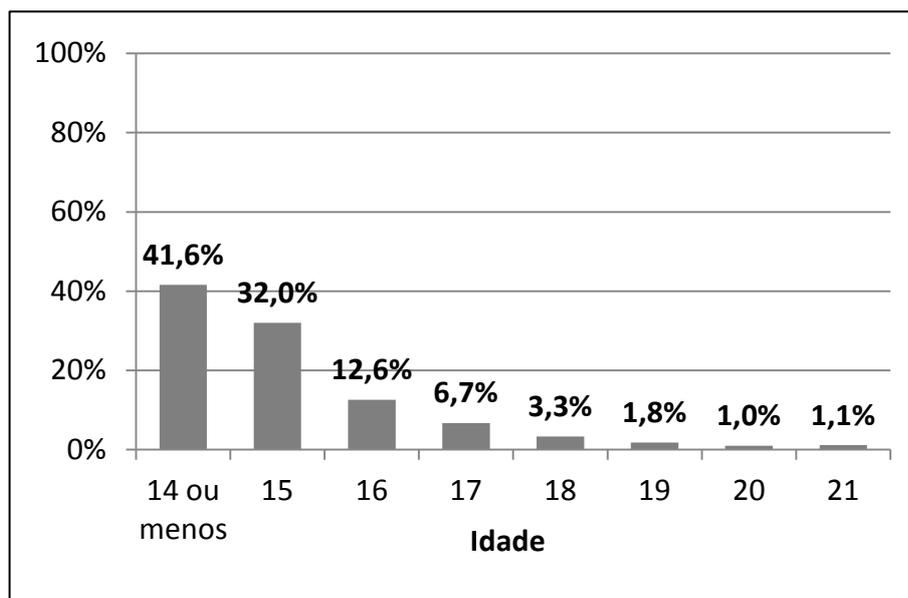
$M_{js}$  = número total de matrículas na série  $s$  do nível de ensino  $j$ .

Para calcular o referido indicador com os dados do SAEB 2005, utiliza-se o número de estudantes numa dada série e não o de matrículas na escola. Neste caso, estima-se que 18,91% dos alunos da 4ª série possuem 12 anos ou mais, sendo qualificados como atrasados para esse nível, e na 8ª série, 26,39% dos estudantes possuem 16 anos ou mais e também são considerados atrasados (GRÁFICOS 3.1 e 3.2). Esse indicador é importante por determinar problemas relacionados, principalmente, com a alta repetência em determinada série, que é um dos principais entraves no sistema de ensino brasileiro, com graves consequências para os níveis de escolaridade da população e está associado com índices de aprendizado e proficiência considerados ruins.

**GRÁFICO 3.1:** Porcentagem de estudantes da 4ª por idade, Brasil 2005



Fonte: SAEB 2005

**GRÁFICO 3.2:** Porcentagem de estudantes da 8ª por idade, Brasil 2005

Fonte: SAEB 2005

Os dados do SAEB 2005 mostram que 78,72% dos alunos da 4ª série iniciaram os estudos no maternal ou pré-escola (TABELA 3.1). Já na coorte da 8ª série, apenas 31,02% dos estudantes informam ter começado os estudos na pré-escola. Com relação às demais variáveis, para a 4ª série, verifica-se que, entre os estudantes, a maioria se considera não branco, 59,17%; aqueles que dizem ter repetido alguma série escolar representam 31,08% dos alunos; 66,16% apenas estudaram em escolas públicas ao longo da vida escolar e 73,59% reportam fazer dever de matemática sempre ou quase sempre. As variáveis de *background* familiar indicam que 83,72% das crianças não trabalham; 55,50% se encontram nos quintis inferiores de renda; 66,53% e 65,06% possuem a mãe e o pai com escolaridade inferior ao ensino médio completo, respectivamente; 91,56% dos estudantes possuem dicionário nos domicílios de moradia e 90,72% dizem conviver com até 4 outras pessoas no domicílio.

Para aqueles que cursavam a 8ª série em 2005, 54,77% se consideram não brancos; 33,03% já repetiram o ano escolar; 70,43% apenas estudaram em escolas públicas e 51,34% fazem o dever de matemática sempre ou quase sempre. As características de *background* familiar mostram que 26,64% dos estudantes não trabalham; 56,79% pertencem aos quintis inferiores de renda; 69,67% e 69,61% possuem mãe e pai com escolaridade inferior ao ensino médio, respectivamente; 94,34% dos alunos têm dicionário em casa e 81,68% convivem com até 4 outras pessoas no domicílio.

**TABELA 3.1:** Porcentagem de estudantes de 4ª e 8ª séries por características selecionadas e respectivos coeficiente de variação (CV) e efeito do plano amostral ampliado (EPA<sub>A</sub>), Brasil 2005

	4ª série			8ª série		
	%	CV	EPA <sub>A</sub>	%	CV	EPA <sub>A</sub>
<b>Início da Educação Infantil</b>						
Maternal/pré-escola	78,72	a <sup>1</sup>	10,8358	31,02	a	14,1093
1ª/2ª/3ª série	21,28	a	10,8358	68,98	a	14,1093
<b>Sexo</b>						
Masculino	51,12	a	4,2952	47,62	a	7,9641
Feminino	48,88	a	4,2952	52,38	a	7,9641
<b>Cor/raça<sup>2</sup></b>						
Branco	40,83	a	10,1324	45,23	a	22,0388
Não Branco	59,17	a	10,1324	54,77	a	22,0388
<b>Repetência</b>						
Não	68,92	a	8,9752	66,97	a	16,5934
Sim	31,08	a	8,9752	33,03	a	16,5934
<b>Experiência escola particular</b>						
Não	66,16	a	9,4169	70,43	a	9,00749
Sim	33,84	a	9,4169	29,57	a	9,00749
<b>Dever de matemática</b>						
Sim	73,59	a	9,8874	51,34	a	9,0074
Não	26,41	a	9,8874	48,66	a	9,0074
<b>Trabalha</b>						
Não	83,72	a	16,4626	26,64	a	22,1939
Sim	16,28	a	16,4626	73,36	a	22,1939
<b>Renda</b>						
Quintis superiores (4º e 5º)	44,50	a	11,4304	43,21	a	21,1362
Quintis inferiores (1º, 2º e 3º)	55,50	a	11,4304	56,79	a	21,1362
<b>Escolaridade da mãe</b>						
Ao menos o Ensino Médio completo	33,47	a	14,5557	30,33	a	10,8474
Menos que o Ensino Médio completo	66,53	a	14,5557	69,67	a	10,8474
<b>Escolaridade do pai</b>						
Ao menos o Ensino Médio completo	34,94	a	9,6218	30,39	a	11,9909
Menos que o Ensino Médio completo	65,06	a	9,6218	69,61	a	11,9909
<b>Dicionário</b>						
Sim	91,56	a	12,7911	94,34	a	14,7608
Não	8,44	a	12,7911	5,66	a	14,7608
<b>Densidade domiciliar</b>						
Até 5 pessoa	90,72	a	7,7634	81,68	a	5,9437
Mais que 5 pessoas	9,27	a	7,7634	18,32	a	5,9437
<b>Rede de ensino</b>						
Privada	11,14	a	36,9161	12,88	a	3,0595
Pública	88,86	a	36,9161	87,12	a	3,0595
<b>Localização</b>						
Urbano	88,15	a	13,8064	-	-	-
Rural	11,85	a	13,8064	-	-	-
<b>Região</b>						
SU/SE + DF	55,55	a	1,7294	60,51	a	12,7023
NO/NE/CO	44,45	a	1,7294	39,49	a	12,7023

<sup>1</sup> Intervalo do coeficiente de variação, ver seção 2.2

<sup>2</sup> Branco = Branco + Amarelos; Não branco = Negro + Pardo/mulato

Fonte: SAEB 2005

Por fim, para os alunos da 4ª série há uma dependência de 88,86% do ensino público; 88,15% dos alunos estudam em escolas que se localizam na zona urbana e 55,55% dos estudantes estão nas regiões SU/SE+DF. Na 8ª série a dependência do ensino público é de 87,12% e 60,51% dos estudantes se encontram nas regiões SU/SE+DF. Para a 8ª série não está disponível a variável de localização urbana ou rural da escola.

As variáveis escolhidas para análise, sob a consideração da frequência anterior à educação infantil, mostram que as situações de adversidade sociais são, na maioria das vezes, relacionadas com menor frequência à educação infantil (TABELA 3.2). Para a 4ª série, na rede privada de ensino, 95,68% dos alunos iniciaram os estudos no maternal ou pré-escola, enquanto que essa porcentagem na rede pública é de 76,60%, ou seja, cerca de 20 pontos percentuais a menos. Entre os alunos da 8ª série, 65,87% dos que estudam na rede privada iniciaram os estudos na pré-escola, sendo que apenas 25,86% dos estudantes da rede pública iniciaram na mesma etapa.

Aqueles que se encontram nos quintis inferiores de renda possuem taxa de frequência à educação infantil de 72,05%, enquanto que essa taxa nos quintis superiores é de 87,06%. Considerando os estudantes da 8ª série, 24,14% dos que se encontram nos quintis inferiores frequentaram educação infantil, e 40,06% dos que se encontram nos quintis superiores frequentaram esse mesmo nível.

Ainda, para ambas as séries, há menor porcentagem de frequência à educação infantil entre os alunos que se autodefinem não branco (negro/pardo ou mulato); já repetiram de série; apenas estudaram em escola pública; não fazem dever de matemática; a mãe e o pai possuem escolaridade inferior ao ensino médio; não possuem dicionário em casa; moram com mais de 4 pessoas; estudam na rede pública de ensino e a escola está localizada na zona rural. Para a 4ª série, ainda há menor frequência à educação infantil entre os estudantes que trabalham e cujas escolas se localizam nas regiões Norte, Nordeste ou Centro-Oeste. Já para a 8ª série, a menor frequência ocorre entre os estudantes que dizem não trabalhar e cujas escolas se localizam nas regiões SU/SE+DF.

Todas as estimativas foram calculadas com boa precisão, considerando o coeficiente de variação como medida de avaliação da qualidade da estimativa. As estimativas possuem coeficientes de variação de até 5%, sendo consideradas adequadas ao objetivo desta análise.

**TABELA 3.2:** Porcentagem de estudantes de 4ª e 8ª séries que iniciaram a vida escolar no maternal/pré-escola por características selecionadas e respectivos coeficiente de variação (CV) e efeito o plano amostral ampliado (EPA<sub>A</sub>), Brasil 2005

	4ª série			8ª série		
	%	CV	EPA <sub>A</sub>	%	CV	EPA <sub>A</sub>
<b>Sexo</b>						
Masculino	76,70	a	8,1472	30,92	a	11,0375
Feminino	80,85	a	8,4179	31,1	a	7,6559
<b>Cor/raça<sup>2</sup></b>						
Branco	82,33	a	7,2808	33,28	a	14,2636
Não Branco	76,24	a	9,0229	29,15	a	7,9887
<b>Repetência</b>						
Não	83,84	a	10,5204	33,26	a	3,9798
Sim	67,40	a	5,2386	26,47	a	4,9488
<b>Experiência escola particular</b>						
Não	76,11	a	6,666	19,79	a	10,3501
Sim	83,84	a	11,5815	57,76	a	8,2885
<b>Dever de matemática</b>						
Sim	81,56	a	9,2793	31,77	a	9,7175
Não	70,84	a	6,1977	30,23	a	9,8748
<b>Trabalha</b>						
Não	80,82	a	10,1311	22,19	a	6,5599
Sim	67,99	a	6,1728	34,22	a	10,6345
<b>Renda</b>						
Quintis superiores (4º e 5º)	87,06	a	11,8176	40,06	a	11,6153
Quintis inferiores (1º, 2º e 3º)	72,05	a	8,4201	24,14	a	7,925
<b>Escolaridade da mãe</b>						
Ao menos o Ensino Médio completo	89,99	a	8,8845	49,47	a	9,4459
Menos que o Ensino Médio completo	73,05	a	7,1956	22,98	a	9,5317
<b>Escolaridade do pai</b>						
Ao menos o Ensino Médio completo	89,51	a	7,4867	49,51	a	10,0326
Menos que o Ensino Médio completo	72,94	a	8,4895	22,94	a	8,7865
<b>Dicionário</b>						
Sim	80,23	a	9,6041	31,64	a	14,3499
Não	62,43	a	7,8603	20,66	a	3,1837
<b>Densidade domiciliar</b>						
Até 5 pessoa	80,35	a	9,8057	32,9	a	13,3261
Mais que 5 pessoas	62,88	a	6,5118	22,62	a	5,24
<b>Rede de ensino</b>						
Privada	95,68	a	3,08029	65,87	a	10,9815
Pública	76,60	a	7,6268	25,86	a	13,0914
<b>Localização</b>						
Urbano	80,69	a	9,6157	-	-	-
Rural	64,14	a	10,0717	-	-	-
<b>Região</b>						
SU/SE + DF	83,76	a	6,7362	29,56	a	10,7013
NO/NE/CO	72,43	a	16,8835	33,25	a	7,3161

<sup>1</sup> Intervalo do coeficiente de variação, ver seção 2.2

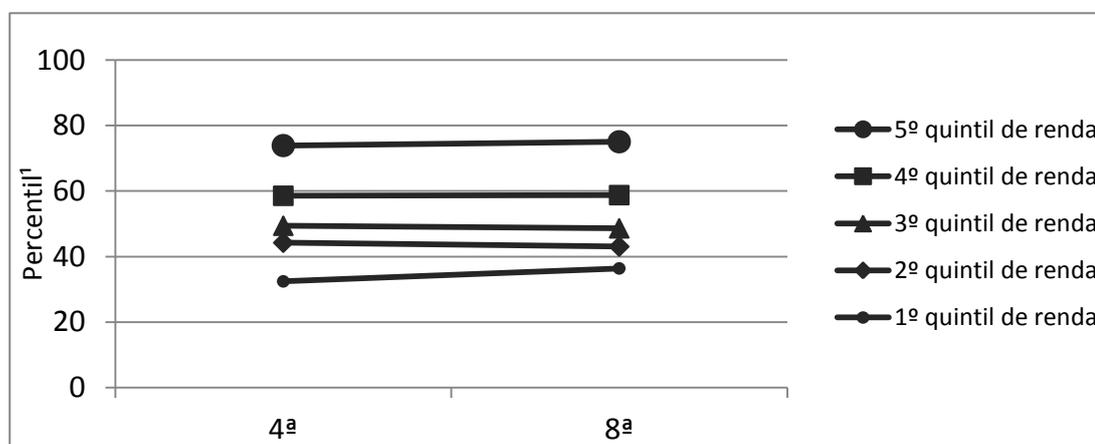
<sup>2</sup> Branco = Branco + Amarelos; Não branco = Negro + Pardo/mulato

Fonte: SAEB 2005

O  $EPA_A$ , medida do impacto da especificação incorreta do plano amostral na estimação da precisão das estimativas, apresenta valores maiores que 1 para todas as estimativas, indicando a importância da consideração do plano amostral na estimação das variâncias e respectivos coeficientes de variação. Por exemplo, a estimação da porcentagem de estudantes da 4ª série que começaram a estudar no maternal (TABELA 3.1) possui um  $EPA_A$  igual a 9,8360. Isso significa que a porcentagem estimada teria variância estimada 10,83 vezes menor caso não fosse considerado o plano amostral complexo (isto é, caso as estimativas fossem calculadas sob a hipótese de amostragem aleatória simples com reposição, AAS). Para a variável associada à rede de ensino privada, a estimativa de variância considerando o plano amostral complexo (e correto) é 35,92 vezes maior que sob a hipótese de AAS.

Como já discutido anteriormente, as diferenças de habilidades entre indivíduos surgem cedo, ainda na infância, decorrente de insumos educacionais desiguais nessa fase. Características de *background* familiar são altamente correlacionadas com o nível de habilidades das crianças. Ao analisar o percentil médio de desempenho no teste SAEB de acordo com a renda (GRÁFICO 3.3), encontra-se uma variação de 41,5 pontos entre os indivíduos no quintil superior e inferior de renda para a 4ª série e em 38,7 pontos para a 8ª série, mostrando que existem desempenhos educacionais desiguais entre rendas desiguais. Se analisada a desigualdade entre indivíduos que frequentaram e que não frequentaram a educação infantil (GRÁFICO 3.4), a diferença entre os percentis médios é de 25,5 pontos na 4ª série e de 17,7 na 8ª série.

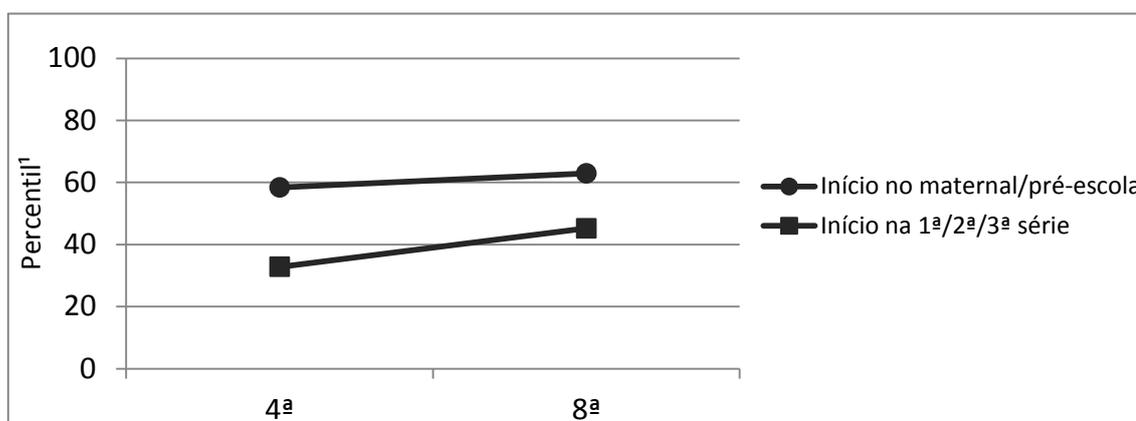
**GRÁFICO 3.3:** Percentil correspondente à nota média dos estudantes de 4ª e 8ª séries por quintil de renda, Brasil 2005



<sup>1</sup> Percentil da distribuição das notas de todos os participantes do teste SAEB

Fonte: SAEB 2005

**GRÁFICO 3.4:** Percentil correspondente à nota média dos estudantes de 4<sup>a</sup> e 8<sup>a</sup> séries das escolas públicas, por início da educação escolar, Brasil 2005



<sup>1</sup> Percentil da distribuição das notas de todos os participantes do teste SAEB

Fonte: SAEB 2005

Em ambos os gráficos anteriores há evidências da diminuição das defasagens da 4<sup>a</sup> para 8<sup>a</sup> série (senão, há um aumento de no máximo 1 ponto percentil). Positivamente, pode ser inferido que as desigualdades não estão aumentando entre uma série e outra porque está havendo uma educação com qualidade suficiente para corrigir certas desvantagens. Por outro lado, pode-se também conjecturar que as vantagens que certos indivíduos teriam (por estar no quintil superior de renda ou ter frequentado a educação infantil) foram sendo perdidas ao longo do tempo, e esses indivíduos estariam em um caminho de equalização com aqueles em desvantagens.

### 3.2 Estimação dos modelos hierárquicos

Seguindo para a análise dos modelos hierárquicos, o procedimento de modelagem estatística objetivou identificar se a frequência à educação infantil tem efeito significativo e preciso sobre o resultado educacional medido pela proficiência no SAEB 2005. As TABELAS 3.3 e 3.4 apresentam os resultados dos modelos de análise de variância com intercepto aleatório para as 4<sup>a</sup> e 8<sup>a</sup> séries.

No modelo 1, modelo nulo, o resultado do coeficiente  $\gamma_{00}$  representa o valor médio das notas no teste SAEB 2005. O valor estimado para o intercepto  $\gamma_{00}$  para a 4<sup>a</sup> série (TABELA 3.3) indica que os estudantes têm uma média estimada de 187,14 e variância estimada correspondente ao componente aleatório (segundo nível) de 754,7994. O coeficiente de correlação intraclasse ( $\rho$ ) é igual a 0,3503, o que significa que 35,03% da variância entre as

notas dos alunos decorre das diferenças no nível 2, ou seja, da heterogeneidade entre as escolas. Obtem-se, portanto, que 64,97% da variação das notas decorre das características do nível 1, aluno.

**TABELA 3.3:** Resultados estimados dos modelos hierárquicos 1 a 3 para os estudantes de 4ª série, Brasil 2005

	4ª série		
	MODELO 1	MODELO 2	MODELO 3
<b>EFEITOS FIXOS DE NÍVEL 1</b>			
Intercepto	187,1430 a <sup>1</sup> 1,9307	175,1419 a 1,6418	266,1527 a 3,1668
<b>Início escolar: educação infantil</b>		<b>15,1421 a</b> 0,7026	<b>2,9436 b</b> 0,3316
Sexo: feminino			-2,8601 b 0,3091
Cor/raça: não branco <sup>2</sup>			-1,6799 c 0,2972
Repetência: sim			-6,8126 a 0,3395
Experiência em escola particular: sim			1,1849 c 0,3432
Faz dever de matemática: não			-4,1310 b 0,3331
Trabalha: sim			-3,0475 b 0,3979
Renda: 1º, 2º e 3º quintil			-4,6975 b 0,3360
esc. da mãe: Inferior ao Ensino Médio			-1,1693 c 0,3034
esc. do pai: inferior ao Ensino Médio			-12,5116 a 0,3625
Dicionário em casa: não			-56,0137 a 0,4236
Densidade domiciliar: mais que 5 pessoas			-2,2583 c 0,4308
<b>EFEITOS FIXOS NÍVEL 2</b>			
Rede: pública			-14,2817 b 1,1836
Região: NO/NE/CO			-9,2674 b -0,9935
Local: rural			-0,9324 e -1,0209

**TABELA 3.3 (cont.):** Resultados estimados dos modelos hierárquicos 1 a 3 para os estudantes de 4ª série, Brasil 2005

	4ª série		
	MODELO 1	MODELO 2	MODELO 3
<b>EFEITOS ALEATÓRIOS</b>			
Variância do intercepto	754,7994 b 50,3989	690,2072 b 44,9830	81,8771 b 8,9665
Variância no nível 1	1400,0357 a 20,1820	1377,4564 a 19,4488	644,4012 a 17,5211
<i>(ρ)</i>	0,3503	0,3338	0,1127
Observações no nível 1	37187	37187	37187
Observações no nível 2	2811	2811	2811
População estimada	3.318.754	3.318.754	3.318.754

<sup>1</sup> Intervalo do valor do coeficiente de variação, ver seção 2.2

<sup>2</sup> Branco = Branco + Amarelos; Não branco = Negro + Pardo/mulato

Fonte: SAEB 2005

Para a 8ª série (TABELA 3.4), a nota média estimada é 244,77, sendo a variância estimada do componente aleatório igual a 897,3302. Portanto, 33,26% da variância decorre do nível 2 (entre as escolas) e 66,74% decorre das diferenças no nível 1 (características do aluno).<sup>19</sup>

O modelo de intercepto aleatório permite que cada subgrupo possua um intercepto diferente (uma nota média para cada escola), portanto, a média apresentada anteriormente é, na verdade, a média desses coeficientes (ponderada pelo plano amostral). O cálculo da variação dessas médias é feito a partir do intervalo de confiança, como apresentado na seção 2.3.1:

$$\beta_{0j} \pm 1,96 * \sqrt{\text{var}(\beta_{0j}) * \text{EPA}(\beta_{0j})}$$

onde  $\text{var}(\beta_{0j}) = \text{var}(\gamma_{00}) + \text{var}(u_{0j})$ . Como mencionado no capítulo 2, o reescalonamento dos pesos amostrais e a inclusão da amostragem complexa na estimação corrigem a variância das estimativas, que antes seriam subestimadas. Portanto, no cálculo do intervalo de confiança, como mencionado na seção 2.3.1, não é necessário a inclusão do EPA, dado que as variâncias já estão corrigidas.

<sup>19</sup> As notas médias das 4ª e 8ª séries não podem ser comparadas diretamente, pois sumarizam conhecimentos cumulativos nas séries. Ver seção 2.1.1.

De acordo com o modelo 1 estimado para a 4ª série,  $var(\gamma_{00}) = (1,9307)^2$  e  $var(u_{0j}) = 754,7994$ . Portanto,  $var(\beta_{0j}) = 758,5270$ . O intervalo de confiança com significância de 95% é igual a  $187,1430 \pm 1,96 * \sqrt{758,5270 * EPA(\beta_{0j})}$ , ou seja, as escolas possuem médias no teste SAEB 2005 para os alunos da 4ª série que variam de 133,16 a 241,12. Fazendo o mesmo cálculo para os alunos da 8ª série, tem-se uma variação das médias das escolas de 185,86 a 303,70.

**Tabela 3.4:** Resultados estimados dos modelos hierárquicos 1 a 3 para os estudantes de 8ª série, Brasil 2005

	8ª série		
	MODELO 1	MODELO 2	MODELO 3
<b>EFEITOS FIXOS DE NÍVEL 1</b>			
Intercepto	244,7771 a <sup>1</sup> 2,5288	242,5526 a 2,1759	356,0316 a 3,8441
<b>Início escolar: educação infantil</b>		<b>6,2045</b> b 0,6131	<b>0,7751</b> b 0,5534
Sexo: feminino			-10,7957 a 0,5378
Cor/raça: não branco <sup>2</sup>			-1,7133 c 0,5010
Repetência: sim			-17,7119 a 0,6300
Experiência em escola particular: sim			2,2214 c 0,6603
Faz dever de matemática: não			-8,7709 a 0,4130
Trabalha: sim			3,0189 c 0,6036
Renda: 1º, 2º e 3º quintil			-12,7688 a 0,6207
esc. da mãe: Inferior ao Ensino Médio			-7,7382 b 0,6628
esc. do pai: inferior ao Ensino Médio			-16,4293 a 0,6735
Dicionário em casa: não			-12,8748 b 1,3147
Densidade domiciliar: mais que 5 pessoas			-4,4049 b 0,5706

**TABELA 3.4 (cont.):** Resultados estimados dos modelos hierárquicos 1 a 3 para os estudantes de 8ª série, Brasil 2005

	8ª série		
	MODELO 1	MODELO 2	MODELO 3
<b>EFEITOS FIXOS NÍVEL 2</b>			
Rede: pública			-28,9945 a 1,4318
Região: NO/NE/CO			-14,1159 b 1,5128
<b>EFEITOS ALEATÓRIOS</b>			
Variância do intercepto	897,3302 c 138,5242	880,3780 c 136,3699	210,5117 b 16,6139
Variância no nível 1	1800,7737 a 30,0569	1799,1784 a 28,8413	1535,1117 a 22,0355
( <i>p</i> )	0,3326	0,3286	0,1206
Observações no nível 1	32151	32151	32151
Observações no nível 2	1952	1952	1952
População estimada	2.804.805	2.804.805	2.804.805

<sup>1</sup> Intervalo do valor do coeficiente de variação, ver seção 2.2

<sup>2</sup> Branco = Branco + Amarelos; Não branco = Negro + Pardo/mulato

Fonte: SAEB 2005

A estimação dos modelos seguintes incluem variáveis que buscam controlar essas diferenças nos resultados dos alunos incluindo no modelo variáveis adicionais de aluno (nível 1) e escola (nível 2). O modelo 2 inclui apenas a variável de início da educação escolar, foco central da análise. O resultado do modelo estimado indica que a variável indicadora de início da educação escolar está associada com uma maior nota no teste SAEB, o que é possível atestar através dos coeficientes positivos estimados para tal variável com respectivos erros padrão e coeficientes de variação. Os alunos da 4ª série que frequentaram maternal ou pré-escola possuem uma nota, em média, 15,14 pontos maior, e os alunos da 8ª série que frequentaram pré-escola possuem uma nota 6,20 pontos maior.

No caso do modelo 2, o intercepto representa a nota média no teste de matemática para os indivíduos que não frequentaram a educação infantil, e a soma da estimativa do intercepto e do coeficiente da variável *início* resulta na nota média daqueles que frequentaram a educação infantil. Obtêm-se, portanto, que para a 4ª série essas notas são iguais a 175,15 para quem não frequentou educação infantil e 190,28 para quem frequentou educação infantil. Para a 8ª série as notas são 242,55 e 248,76, respectivamente.

Como esperado, a inclusão da variável de início da educação escolar resulta em uma redução da variância de nível 1, de 1400,0357 no modelo 1 para 1377,4564 no modelo 2, e da variância de nível de 2, de 6754,7994 para 690,2072. O resíduo total muda de 2154,8351 para 2067,6636. Portanto, a porcentagem da variância original explicada pela variável *início* é igual a 4,05%, sendo o coeficiente de correlação intra-classe igual a 0,3338.

O modelo 3 inclui efeitos fixos relativos as variáveis explicativas de nível 1 (o efeito aleatório entre as escolas continua somente no intercepto) e de nível 2, objetivando a análise de variância condicionada à essas características do aluno e da escola. Observa-se que o aumento na nota daqueles que iniciaram os estudos precocemente continua sendo estatisticamente significativo e preciso, mas declina para 2,94 pontos na 4ª série e 0,77 na 8ª série.

Excetuando-se o efeito da variável *trabalha* para os alunos da 8ª série, as demais variáveis no modelo de efeitos fixos têm efeitos negativos quando consideradas as situações de risco para o desenvolvimento das habilidades. Para a 4ª série, os efeitos negativos são mais expressivos para os alunos que declararam que a escolaridade do pai é inferior ao ensino médio; que não possuem dicionário em casa; estudam na rede pública de ensino e cuja escola se localiza nas regiões Norte, Nordeste ou Centro-Oeste. Para os alunos da 8ª série, os efeitos negativos mais expressivos encontram-se nas mesmas situações dos alunos da 4ª série, além das situações de repetência, pertencer aos quintis inferiores de renda e ser do sexo feminino.

As variâncias de nível 1 e 2 diminuem expressivamente após a inclusão das variáveis de controle. No nível 1 passa a ser igual a 644,4012 e no nível 2 igual a 81,8771. Portanto, o modelo agora explica 64,87% da variância original entre as notas dos alunos.

Inferiu-se aqui que as diferenças nas performances dos estudantes podem ser explicadas tanto pelos fatores referentes às escolas quanto às características individuais. Já foi mencionado que os fatores escolares não são de interesse, sendo feito apenas o controle desse diferencial por escolas através da estrutura hierárquica e inclusão de poucas variáveis. Com relação às características individuais, estudantes de diferentes classes sociais, por exemplo, possuem diferentes retornos dessa característica de acordo com a escola em que estão inseridos. Assim, uma análise que considere que o efeito da educação infantil pode variar de acordo com a escola em que o aluno estuda capta os diferentes efeitos da variável de interesse (efeito aleatório).

O modelo 4 (TABELA 3.5) considera apenas a variável de início da educação escolar, sendo que essa passa ser considerada como um efeito aleatório (que varia entre escolas). Para a 4ª série, os estudantes que frequentaram educação infantil têm, em media, uma nota 19,29

pontos maior. Esse coeficiente pode variar dependendo da escola em que o aluno está inserido. A aleatoriedade desse coeficiente mostra que a educação infantil pode aumentar a nota no teste de matemática de -6,42 a 45,00 pontos, dependendo da escola em que o aluno estuda (intervalo de confiança à 95%). De acordo com a modelagem estatística, esse coeficiente é significativo e preciso.

Entre os alunos da 8ª série, o efeito da educação infantil é de 5,55 pontos positivos, sendo também significativo e preciso. Esse efeito pode variar de -9,54 a 20,65 entre as escolas (intervalo de confiança à 95%). Observa-se pelo possível efeito negativo da educação infantil tanto na 4ª quanto na 8ª série que, provavelmente, a vantagem da frequência à educação infantil pode ser dissipada ao longo da vida escolar do estudante (efeito baixo ou negativo), seja pela baixa qualidade com que ela ocorreu ou pela baixa qualidade da educação posterior.

O modelo 4 é conceitualmente descrito como:

$$Y_{ij} = \beta_{0j} + \beta_{1j}início_{ij} + e_{ij}$$

$$\beta_{0j} = \gamma_{00} + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + u_{1j}$$

$$Y_{ij} = \gamma_{00} + \gamma_{10}início_{ij} + u_{1j}início_{ij} + u_{0j} + e_{ij}$$

Cabe observar a distribuição dos resíduos do segundo nível correspondente ao coeficiente *início*, ( $u_{1j}$ ). A variação deste resíduo representa o efeito da escola sobre o efeito médio da frequência à educação infantil. Como já mencionado, a estimativa para o efeito médio geral da educação infantil ( $\beta_{1j}$ ) é de 19,29 pontos a mais. A partir de então, o efeito médio em uma escola  $j$  qualquer é igual ao efeito médio geral mais o efeito aleatório estimado para o nível 2, ( $u_{1j}$ ). Os GRÁFICOS 3.5 e 3.6 mostram que existem estimativas de resíduos para o segundo nível que são diferentes de zero, indicando que as escolas têm influências diferenciadas sobre o efeito da educação infantil.

**TABELA 3.5:** Resultados estimados os modelos hierárquicos 4 e 5 para os estudantes de 4ª e 8ª séries, Brasil 2005

	4ª série		8ª série	
	MODELO 4	MODELO 5	MODELO 4	MODELO 5
<b>EFEITOS FIXOS DE NÍVEL 1</b>				
Intercepto	172,2261 a <sup>1</sup> 1,6658	233,7702 a 3,2248	242,8063 a 1,7590	351,8841 a 4,0121
<b>Início escolar: educação infantil</b>	<b>19,2881</b> a 0,8601	<b>36,9125</b> b 3,2357	<b>5,5559</b> b 0,6522	<b>4,8515</b> d 2,3135
Sexo: feminino		-2,9030 b 0,3100		-10,8061 a 0,5363
Cor/raça: não branco <sup>2</sup>		-1,6500 c 0,2976		-1,7152 c 0,5241
Repetência: sim		-6,6704 a 0,3428		-17,6737 a 0,6245
Experiência em escola particular: sim		1,0817 c 0,3442		2,4361 c 0,7228
Faz dever de matemática: não		-4,1016 b 0,3331		-8,8006 a 0,4804
Trabalha: sim		-3,0518 b 0,4017		3,0523 c 0,6004
Renda: 1º, 2º e 3º quintil		-4,7082 b 0,3309		-12,7836 a 0,6251
Esc. da mãe: Inferior ao Ensino Médio		-1,1977 c 0,3014		-7,7049 b 0,6612
Esc. do pai: inferior ao Ensino Médio		-12,4092 a 0,3657		-16,4387 a 0,6723
Dicionário em casa: não		-56,1405 a 0,4174		-12,9612 b 1,3116
Densidade domiciliar: mais que 5 pessoas		-2,3151 c 0,4273		-4,4840 b 0,5699
<b>EFEITOS FIXOS NÍVEL 2 (intercepto)</b>				
Rede: pública		-6,9509 c 2,1439		-24,5575 b 2,3859
Região: NO/NE/CO		-3,9337 c 1,1521		-12,9002 b 1,8821
Local: rural		-3,9069 c 1,0158		- -
<b>EFEITOS FIXOS NÍVEL 2 (início)</b>				
Rede: pública		-11,2582 b 1,2808		-2,658 d 1,0726
Região: NO/NE/CO		-7,0102 b 0,7772		-0,7486 e 1,0347
Local: rural		-2,7885 e 1,0821		- -

**TABELA 3.5 (cont.):** Resultados estimados os modelos hierárquicos 4 e 5 para os estudantes de 4ª e 8ª séries, Brasil 2005

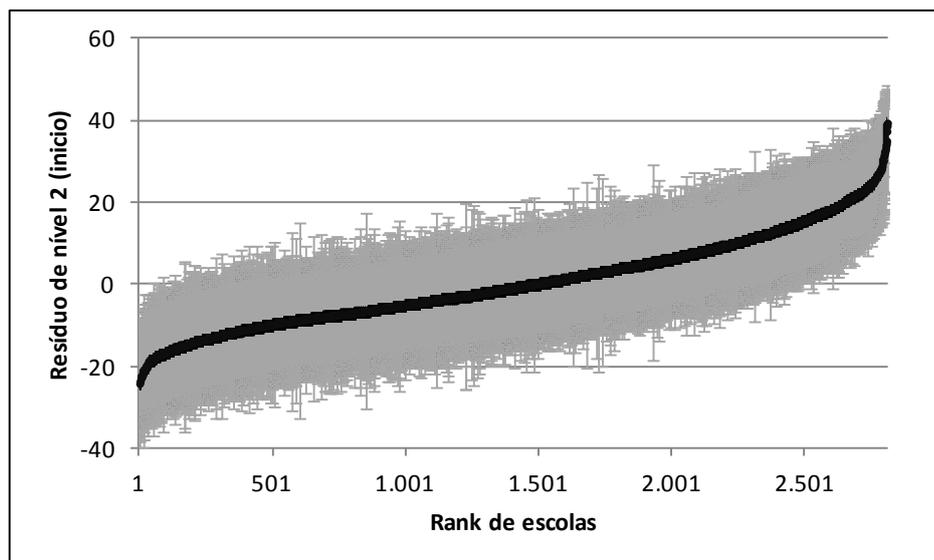
	4ª série		8ª série	
	MODELO 4	MODELO 5	MODELO 4	MODELO 5
<b>EFEITOS ALEATÓRIOS</b>				
Variância do intercepto	152,4098 c 33,8737	1,6804 e 2,2760	1029,5732 b 115,0139	144,7740 b 23,0167
Variância de <i>início</i>	171,4062 b 13,6931	28,1881 c 5,7741	58,9035 b 10,1968	11,7384 c 4,1969
Variância no nível 1	1350,0604 a 18,4400 <i>0,1014</i>	642,1521 a 17,4903 <i>0,0026</i>	1761,5891 a 25,3284 <i>0,3689</i>	1534,593 a 22,1245 <i>0,0862</i>
<i>(P)</i>				
Observações no nível 1	37187	37187	32151	32151
Observações no nível 2	2811	2811	1952	1952
População estimada	3.318.754	3.318.754	2.804.805	2.804.805

<sup>1</sup> Intervalo do valor do coeficiente de variação, ver seção 2.2

<sup>2</sup> Branco = Branco + Amarelos; Não branco = Negro + Pardo/mulato

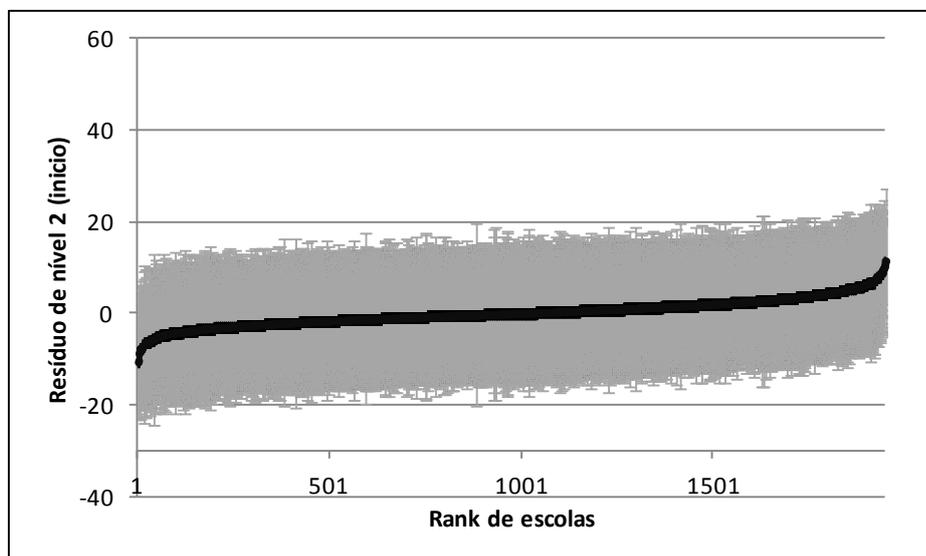
Fonte: SAEB 2005

**GRÁFICO 3.5:** Estimativa do efeito da escola sobre o efeito da frequência à educação infantil no desempenho escolar para os alunos de 4ª série, com seu respectivo intervalo de confiança (95%), estimado a partir dos resíduos de nível 2,  $u_{1j}$ .



Fonte: SAEB 2005

**GRÁFICO 3.6:** Estimativa do efeito da escola sobre o efeito da frequência à educação infantil no desempenho escolar para os alunos de 8ª série, com seu respectivo intervalo de confiança (95%), estimado a partir dos resíduos de nível 2,  $u_{1j}$ .



Fonte: SAEB 2005

Por fim, o modelo 5 considera todas as covariáveis de primeiro e segundo nível, com efeitos aleatórios tanto no intercepto quanto no coeficiente da variável *início*. Assim, esse modelo permite que o efeito médio da educação infantil na nota média do estudante varie de acordo com a escola e de acordo com as características da escola.

O efeito da frequência à educação infantil nos alunos da 4ª série é positivo, preciso e estatisticamente significativo, em média igual a uma nota 36,91 pontos maior para aqueles que, além de terem frequentado a educação infantil, estudam em escola privada, localizada na área urbana e nas regiões do Sul, Sudeste e Distrito Federal. Esse efeito pode variar de 24,73 a 49,10 com 95% de confiança, dependendo da escola em que o aluno estuda. Os alunos que estudam em escola pública têm esse efeito da educação infantil 11,26 pontos menor, em média. Se a escola está localizada na área rural esse efeito era menor em 2,78 pontos, em média. E, caso a escola esteja nas regiões Norte, Nordeste ou Centro-Oeste (Exceto o DF), o efeito da educação infantil cai 7,01 pontos.

Para a 8ª série o efeito da educação infantil não é preciso (vide coeficiente de variação entre 30% e 50% indicado pela letra *d* ao lado do coeficiente estimado), todavia é estatisticamente significativo. O efeito ainda é positivo: ter começado a estudar na educação infantil significa alcançar, em média, uma nota no teste SAEB 2005 4,85 pontos maior. Esse efeito pode variar de -3,25 a 12,95 (intervalo de confiança à 95%) dependendo da escola em

que o aluno estuda, sendo, em média, 2,65 pontos menor se o aluno estuda em escola pública, e 0,75 pontos menor se a escola se localiza nas regiões Norte, Nordeste ou Centro-Oeste. Os efeitos das covariáveis de segundo nível foram também pouco precisos, todavia estatisticamente significativos.

Os modelos mostraram que a educação infantil pode ter efeito positivo sobre o resultado escolar dos alunos de 4ª e 8ª série medido pelo teste SAEB 2005. A escola em que o aluno estuda exerce influência sobre esse efeito, podendo aumentá-lo ou diminuí-lo. Assim, fatores relativos à escola são preponderantes na realização da vantagem da frequência à educação infantil, sendo necessária, posteriormente, a identificação desses fatores.

## **4 CARACTERÍSTICAS SOCIOECONÔMICAS FAMILIAR E DOMICILIAR DAS CRIANÇAS DE 0 A 6 ANOS E FREQUENCIA À EDUCAÇÃO INFANTIL E ENSINO FUNDAMENTAL**

A educação infantil brasileira abrange crianças de 0 a 5 anos de idade desde 2006, com implementação obrigatória no ensino público e privado desde 2010. Todavia, aqui se optou por trabalhar com crianças entre 0 e 6 anos de idade por ser esse o período normalmente designado para educação infantil. Ainda que crianças brasileiras de 6 anos possam estar frequentando o ensino fundamental e não mais o ensino infantil, o desenvolvimento obtido nessa fase compreende ainda o desenvolvimento infantil. Portanto, as análises feitas adiante consideram a frequência de crianças de 0 a 6 anos à educação, em qualquer nível de ensino, seja educação infantil ou ensino fundamental, objetivando caracterizar o contexto socioeconômico familiar das crianças em idade de desenvolvimento inicial de suas habilidades.

Este estudo aborda a frequência à creche, maternal/jardim de infância/etc, classe de alfabetização ou ensino fundamental, resumida em frequência à educação infantil ou ensino fundamental, de crianças de 0 a 6 anos e o seu contexto, considerando o ano de 2009. Além da população total e alguns recortes geográficos, utilizou-se um recorte socioeconômico de população pobre, aqui definida como aquela com renda domiciliar per capita igual ou inferior à  $\frac{1}{2}$  salário mínimo.

### **4.1 Acesso à educação infantil e ensino fundamental e diferenciais socioeconômicos e regionais**

Uma análise primordial sobre a frequência aos níveis relativos à educação infantil brasileira na PNAD de 2009 mostra cerca de 7,8 milhões de crianças frequentando creche, maternal/jardim de infância/etc e classe de alfabetização (TABELA 4.1). Normalmente considera-se a educação infantil aquela voltada para crianças de 0 a 6 anos de idade, sendo que no Brasil esse nível compreende crianças de 0 a 5 anos desde 2006. Contudo, observa-se que há frequência de crianças com idades superiores a essas nesse nível. Essa peculiaridade da educação infantil deve-se ao atraso escolar que algumas crianças possuem, o que dá origem ao indicador distorção idade/série, que apresenta a adequação que há entre a idade e a série que o aluno frequenta. Adaptando a relação proposta por Rigotti e Cerqueira (2004) ao caso da educação infantil, temos:

$$\text{Distorção idade/série} = \frac{M_{ed\ inf\_sup6}}{M_{infantil}}$$

onde:

$M_{ed\ inf\_sup6}$  = crianças na educação infantil com idade superior a 6 anos.

$M_{infantil}$  = total de crianças frequentando a educação infantil;

Calculando para o ano de 2009, encontra-se uma distorção igual a 5,08% considerando a educação infantil de 0 a 6 anos. Ou seja, cerca de 5 em cada 100 alunos matriculados na educação infantil tem mais de 6 anos. Se considerada a educação infantil de 0 a 5 anos, a distorção chega a 20,81%, todavia há de se considerar a adequação incompleta ao ensino fundamental de 9 anos.

Subsistem, no entanto, quase 400.000 crianças com idade superior a 7 anos no ensino infantil, uma distorção séria que impacta toda a escolarização do indivíduo, que já entrando no ensino fundamental com certo atraso, tenderá a manter esse atraso por toda a vida. Essa distorção foi reduzida ao longo do tempo, mas não se pode negar que o avanço é insuficiente. Em 1999 o número de crianças com idade superior a 7 anos frequentando o ensino infantil era cerca de 730.000, havendo, portanto, uma redução nessa distorção de cerca de 45% em 10 anos.

Uma situação inversa é a absorção de crianças de 6 anos no ensino fundamental. Em 2009 existiam 1.567.591 crianças de 6 anos frequentam o ensino primário, o que representa 52,49% dessa faixa etária.

**TABELA 4.1:** Número de crianças frequentando os níveis relativos à educação infantil<sup>1</sup> por idade, Brasil 2009

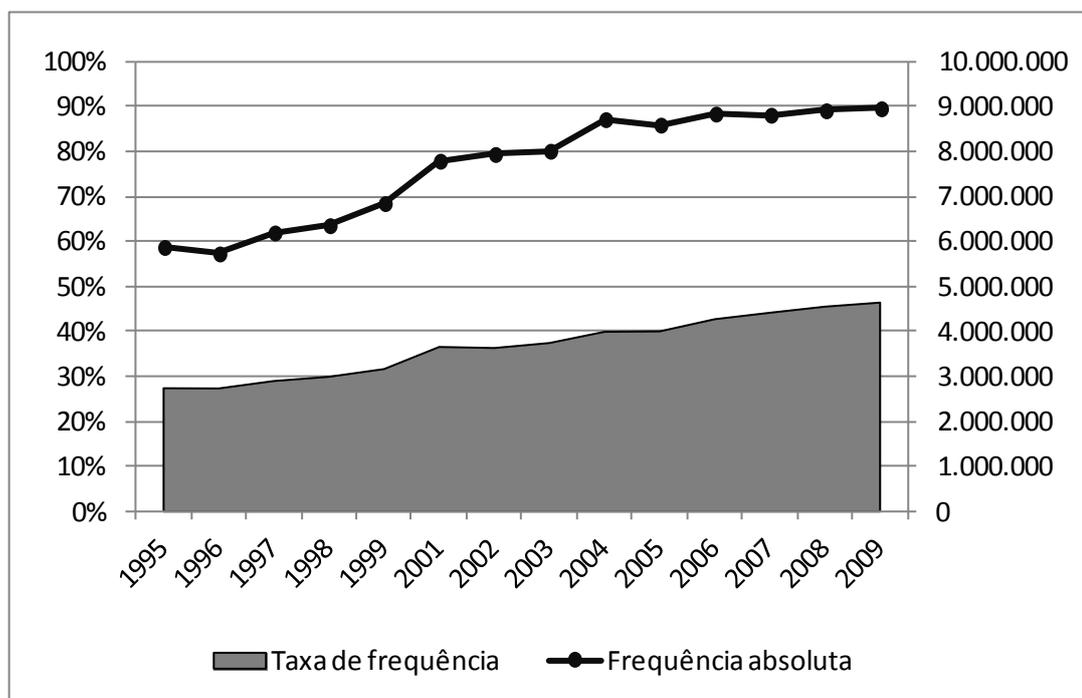
	Creche	Maternal, jardim de infância, etc	Classe de alfabetização	Total
0 a 3 anos	1.937.367	0	0	1.937.367
4 e 5 anos	556.784	2.853.502	811.531	4.221.817
6 anos	0	554.710	668.293	1.223.003
7 a 11 anos	0	99.796	295.731	395.527
<b>Total</b>	<b>2.494.151</b>	<b>3.508.008</b>	<b>1.775.555</b>	<b>7.777.714</b>

<sup>1</sup>Níveis de ensino precedentes ao ciclo do Ensino Fundamental

Fonte: PNAD 2009

A evolução do acesso à educação de crianças de 0 a 6 anos é mostrada nos GRÁFICOS 4.1 e 4.2. Entre 1995 e 2009 a frequência das crianças de 0 a 6 anos passou de 27,56% para 46,71%, representando um aumento de 69,48%. O maior aumento foi entre a faixa etária de 0 a 3 anos, a taxa de frequência passou de 7,58% em 1995 para 18,39% em 2009, um aumento de 142,61%. Entre as crianças de 4 a 6 anos a taxa passou de 53,48% para 81,28%, aumento de 51,98%. Ao longo do período analisado houve um acréscimo de 3.084.854 crianças de 0 a 6 anos em creches, pré-escolas e escolas, totalizando 8.952.545 crianças em 2009. Todavia, ainda havia 10.215.343 crianças nessa faixa sem acesso à educação.

**GRÁFICO 4.1:** Taxa de frequência e frequência absoluta de crianças de 0 a 6 anos à educação infantil e ensino fundamental, Brasil 1995 a 2009

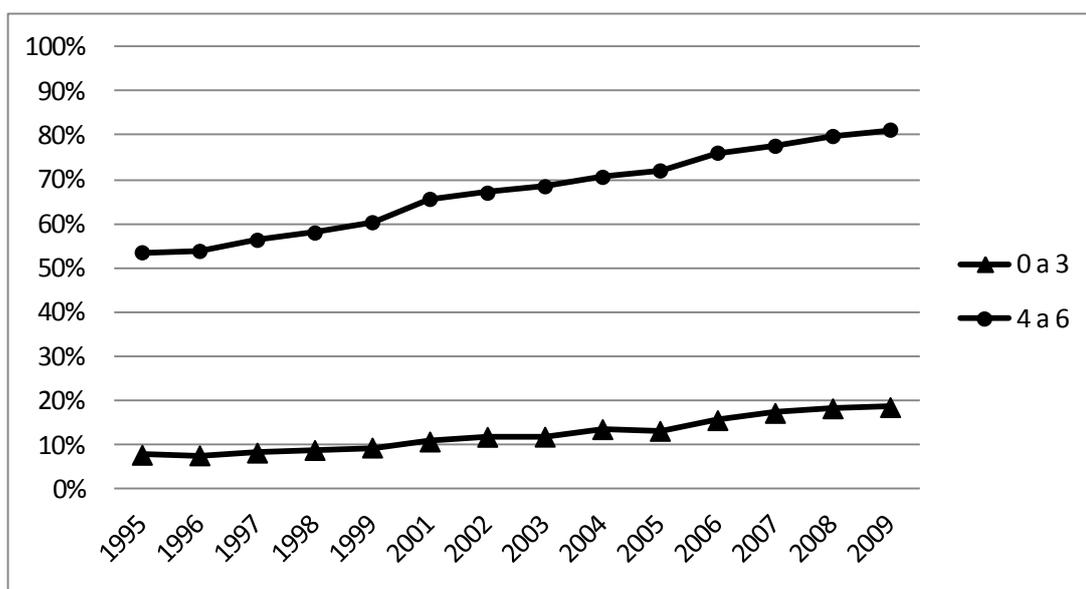


Fonte: PNAD 1995 a 2009

Em um país tão populoso, extenso e diverso quanto o Brasil é necessário verificar quem tem acesso e quais as principais diferenças nesse acesso. A primeira comparação a ser feita consiste na consideração dos disparates regionais existentes no Brasil (TABELA 4.2). A região com melhor indicador é a Sudeste, com 50,83% das crianças de 0 a 6 anos na creche, pré-escola ou escola. Divergindo dos resultados comuns na comparação de indicadores socioeconômicos regionais, o Nordeste possui a segunda maior frequência, com 47%, seguido do Sul, Centro-Oeste e Norte.

Ainda na comparação regional, fez-se necessário avaliar separadamente a frequência entre crianças de 0 a 3 e 4 a 6 anos para melhor compreender as diferenças. A frequência de 0 a 3 é maior para a região Sul (24,24%) seguida do Sudeste (22,80%), o Nordeste ocupa a terceira colocação (15,51%). No intervalo de 4 a 6 anos as maiores frequências são das regiões Nordeste (85,90%) e Sudeste (84,91%) e menor para a região Sul (70,98%).

**GRÁFICO 4.2:** Taxa de frequência de crianças de 0 a 3 anos e de 4 a 6 anos à educação infantil e ensino fundamental, Brasil 1995 a 2009



Fonte: PNAD 1995 a 2009

Outros recortes que fornecem indicadores da existência de diferenças no acesso à creche, pré-escola e escola se relacionam à renda, cor/raça e situação do domicílio. A frequência pelo quintil de renda é previsível, quanto maior a renda, maior a taxa. Na população de 0 a 6 anos a frequência é 40,03% maior no quintil mais alto em relação ao quintil mais baixo. Essa diferença é de 148,52% na população de 0 a 3 anos e de 18,94% na população de 4 a 6 anos, ou seja, a maior diferença de acesso entre os quintis de renda se encontra na população de 0 a 3 anos.

Na comparação por cor ou raça, as taxas para pardos, pretos e brancos é bastante similar, sendo um pouco maior entre os pretos. Na população de 0 a 6 anos as taxas são, respectivamente, 46,17%, 49,88% e 47,01%. Na população de 0 a 3 a frequência é maior também entre os pretos 22,07%, seguidos pelos brancos, 20,16% e pelos pardos, 16,19%. Já na população de 4 a 6 anos, a maior frequência é entre os brancos, 82,65%, seguido pelos pardos, 80,26%, e depois pelos pretos, 79,58%. A população amarela possui a maior taxa

entre a população de 0 a 6 anos, e quase totalidade das crianças de 4 a 6 frequentam creche, pré-escola ou escola. Já a população indígena possui as menores taxas, resguardadas peculiaridades culturais das comunidades indígenas. Por fim, entre a população urbana há maior taxa de frequência à creche, pré-escola e escola, na comparação com a população rural.

**TABELA 4.2:** Taxa de frequência de crianças de 0 a 6 anos à educação infantil e ensino fundamental por região, quintil de renda, cor e situação do domicílio considerando os grupos de idade, Brasil 2009

	0 a 6 anos	0 a 3 anos	4 a 6 anos
<b>Brasil</b>	<b>46,71</b> a <sup>1</sup>	<b>18,39</b> a	<b>81,28</b> a
Norte	37,22 a	8,27 b	73,50 a
Nordeste	47,00 a	15,51 a	85,90 a
Centro-oeste	41,18 a	14,72 a	73,75 a
Sul	45,83 a	24,24 a	70,98 a
Sudeste	50,83 a	22,80 a	84,91 a
<b>Renda</b>			
1º quintil	40,77 a	12,31 a	76,35 a
2º quintil	43,95 a	14,61 a	79,01 a
3º quintil	48,01 a	19,82 a	81,45 a
4º quintil	51,63 a	23,23 a	85,75 a
5º quintil	57,08 a	30,59 a	90,81 a
<b>Cor/raça</b>			
Amarelo	53,85 b	16,38 d	99,13 a
Indígena	24,39 d	2,51 e	46,79 c
Parda	46,17 a	16,19 a	80,26 a
Preta	49,88 a	22,07 b	79,58 a
Branca	47,01 a	20,16 a	82,65 a
<b>Situação domicílio</b>			
Rural	38,69 a	8,86 b	73,25 a
Urbano	48,51 a	20,47 a	83,15 a

<sup>1</sup> Intervalos do coeficiente de variação, ver seção 2.2

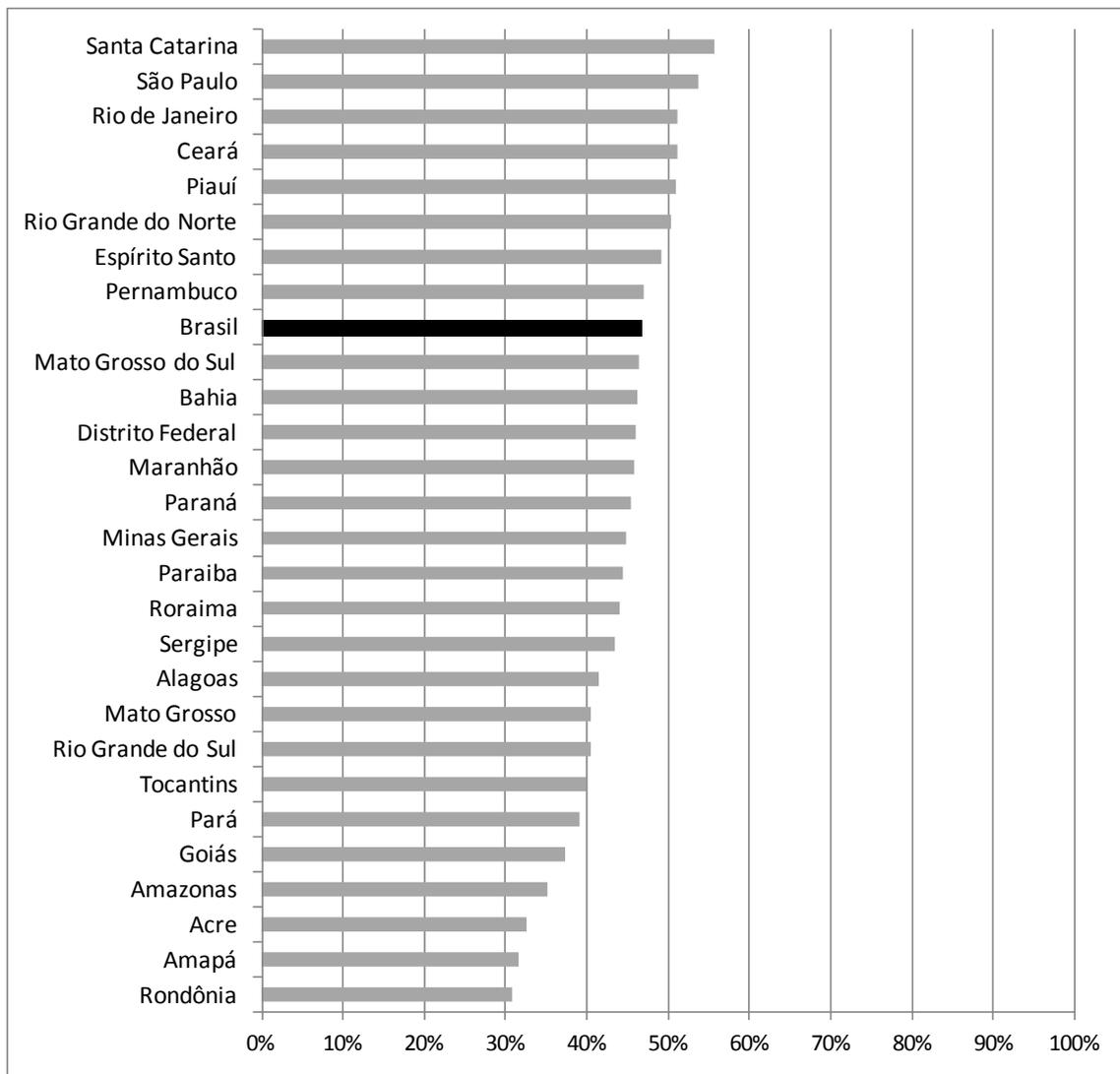
Fonte: PNAD 2009

É interessante ressaltar que a taxa de frequência entre a população preta é maior que a taxa da população branca entre a população de 0 a 6 anos, que ocorre devido à maior frequência entre a população de 0 a 3 anos. Considerando a população preta mais pobre que a população branca, pode-se inferir que a escolha pela frequência das crianças mais novas a creches ocorre pela necessidade do cuidado da criança enquanto os pais permanecem fora de casa, e impossibilidade do pagamento de alguém que realize este cuidado dentro da própria

residência. Já a população branca possui maiores recursos para manter a crianças dentro do domicílio.

O GRÁFICO 4.3 apresenta as porcentagens da frequência para os estados brasileiros. Como indicado anteriormente, os estados nordestinos tem alguns resultados expressivos na comparação nacional. Destacam-se os indicadores dos estados nordestinos Ceará, Piauí, Rio Grande do Norte e Pernambuco, que são maiores que a média nacional e maiores do que os valores para os estados do sudeste e sul, que historicamente possuem melhores indicadores desse gênero.

**GRÁFICO 4.3:** Taxa de frequência de crianças de 0 a 6 anos à educação infantil e ensino fundamental nos estados brasileiros, 2009

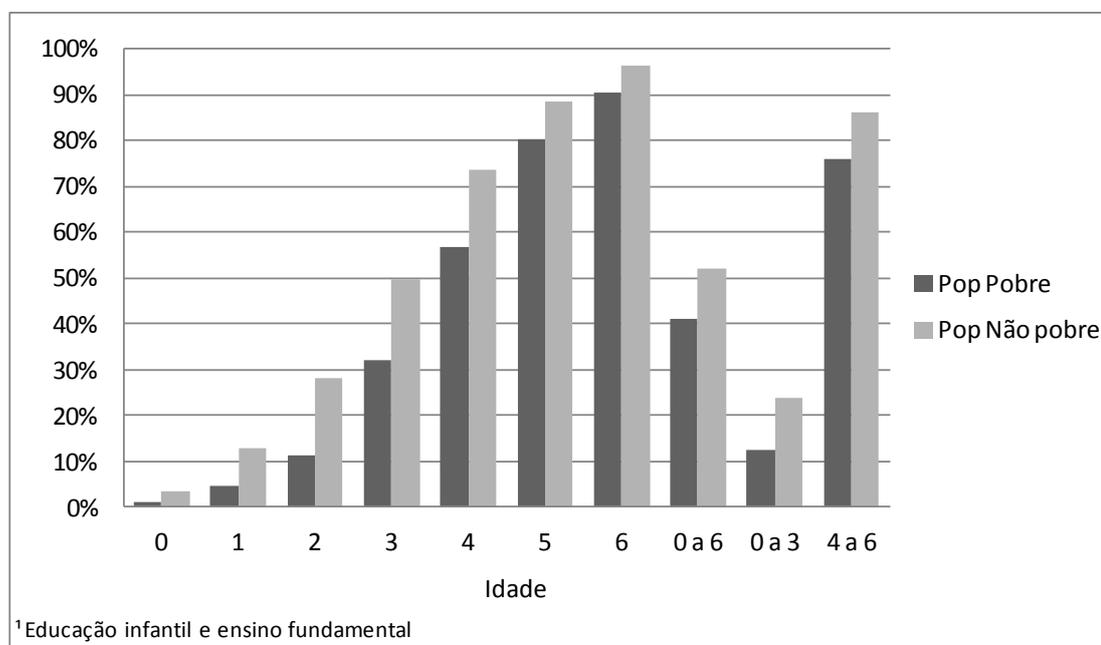


Fonte: PNAD 2009

#### 4.1.1 Acesso à educação para crianças de 0 a 6 anos pela população pobre<sup>20</sup>

Existem 9.403.853 crianças de 0 a 6 em situação de pobreza, de acordo com os critérios aqui estabelecidos para tal, o que representa 49,06% das crianças nessa faixa etária. Das crianças pobres, 41,22%, ou 3.876.387, frequentam algum nível educacional (na população não pobre, 51,99% das crianças de 0 a 6 anos estão frequentando creche, pré-escola ou escola). O GRÁFICO 4.4 apresenta a frequência relativa à creche, pré-escola e escola por idade e grupos de idade das crianças de 0 a 6 anos para a população considerada pobre e população não pobre no ano de 2009. Das crianças de 6 anos, 90,38% na população pobre e 96,48% na população não pobre frequentavam algum nível educacional em 2009, atendimento que deve-se considerar como parte do ensino fundamental. Como já mencionado, 52,49% das crianças de 6 anos frequentam o ensino primário. A frequência nas idades de 4 e 5 são também expressivas, caminhando para um atendimento universal até 2016, como previsto pelo PNE 2011-2020.

**GRÁFICO 4.4:** Taxa de frequência de crianças de 0 a 6 anos à educação infantil e ensino fundamental por idade e grupos de idade, população pobre e população não pobre, Brasil 2009



Fonte: PNAD 2009

<sup>20</sup> População pobre consiste nas pessoas com renda domiciliar per capita igual ou inferior à ½ salário mínimo.

A frequência entre a população pobre é inferior à frequência da população não pobre em todas as idades, o que mostra uma desigualdade importante no que concerne a formulação de políticas públicas, pensando na educação infantil para crianças pobres. A dependência do ensino público é de 58,36% na população não pobre, sendo 89,16% na população pobre.

Interessante desmembrar a frequência entre população pobre e não pobre para as regiões do país. 22,41% dos pobres se encontram na região Nordeste, e por isso considerada a mais pobre do país. Todavia é a segunda região com maior frequência de crianças de 0 a 6 anos em algum nível educacional (TABELA 4.2). Os resultados (TABELA 4.3) mostram que a maior frequência está sempre entre a população não pobre. A diferença na porcentagem da frequência entre a população não pobre e população pobre no Nordeste é de 10,54 pontos percentuais, menor do que a diferença na região Sul, de 16,70 pontos percentuais. Portanto, considerando a região Nordeste, havendo uma maior homogeneidade na frequência entre os dois grupos socioeconômicos e maior número absoluto de pobres, é melhor compreensível a alta taxa de frequência dessa região.

**TABELA 4.3:** Taxa de frequência de crianças de 0 a 6 anos à educação infantil e ensino fundamental por população não pobre e pobre, regiões do Brasil, 2009

	Não pobre	Pobre
Norte	41,19 a <sup>1</sup>	34,52 a
Nordeste	54,48 a	43,94 a
Sudeste	54,68 a	43,56 a
Sul	51,24 a	34,54 a
Centro-oeste	45,19 a	34,96 a

<sup>1</sup> Intervalos do coeficiente de variação, vide capítulo 2

Fonte: PNAD 2009

## 4.2 Características familiares adversas

Foi discutido no capítulo 1 que o ambiente que as famílias propiciam às crianças é grande determinante do desenvolvimento das habilidades dos indivíduos em sua infância. Os fatores adversos familiares descritos pela literatura (RUTTER (1981) *apud* SILVA *et al.* (2008), HECKMAN E MASTEROV (2004) e ESPING-ANDERSEN (2007a)) são i) grande número de filhos na família; ii) baixa renda familiar; iii) discórdia conjugal; iv) número excessivo de pessoas residentes na mesma casa; v) psicopatologia familiar; vi) baixos níveis

de educação e das habilidades dos pais; vii) baixo nível de capital cultura familiar; viii) pouca dedicação dos pais aos filhos; ix) falta de estímulo cognitivo e emocional às crianças; x) falta de conhecimento pelos pais de como educar uma criança; e xi) arranjos familiares não tradicionais, representado majoritariamente por mães solteiras e ausência do pai.

A PNAD possibilita a análise de uma grande parte desses fatores (TABELA 4.4). A variável de número de filhos na família que moravam no domicílio é utilizada na análise do grande número de filhos, e a variável de número de componentes do domicílio por dormitório permite a análise do número de pessoas residentes na mesma casa. A variável de tipologia familiar será utilizada para arranjos familiares, separados entre casal com filhos e mãe com filhos. A baixa renda foi analisada anteriormente através dos quintis e renda familiar per capita, assim como em toda a análise considerando a população pobre. As variáveis de nível de instrução e desemprego da mãe serão utilizadas como uma *proxy* para os fatores que concernem educação e habilidade dos pais. Pode-se considerar que essa também represente a falta de estímulo cognitivo e emocional às crianças, a falta de conhecimento pelos pais de como educar uma criança e baixo nível de capital cultural familiar, todavia esses são fatores muito mais complexos de serem medidos e que não podem ser diretamente relacionados ao nível de instrução da mãe. Os fatores de discórdia conjugal, psicopatologia familiar e pouca dedicação dos pais à educação dos filhos não podem ser analisados através da PNAD.

Além das análises dos fatores descritos pela literatura, analisou-se também fatores relacionados às estruturas domiciliares, considerando o contexto brasileiro da presença de precariedade nas moradias, que se relaciona amplamente com as condições ambientais propiciadas pelas famílias às crianças. Analisou-se a existência de coleta de esgoto adequada no domicílio, pensando nessa como uma variável resumo das piores condições domiciliares que uma criança pode estar exposta, que pode gerar doenças que continuamente atrapalham o desenvolvimento cognitivo e não cognitivo das crianças. Outra variável analisada foi a existência de microcomputador com acesso à internet, que possibilita o acesso à informação. E por último foi analisada a presença de máquina de lavar, que pode ser compreendida como um aparelho que permite às mães, principalmente, desprender menos tempo nas atividades de lavar roupas, podendo passar mais tempo com os filhos.

**TABELA 4.4:** Variáveis adversas no desenvolvimento das habilidades das crianças de 0 a 6 anos descritas na literatura e variáveis correspondentes utilizadas da PNAD 2009

LITERATURA	PNAD
Grande número de filhos na família	Número de filhos na família
Número excessivo de pessoas na mesma casa	Número de pessoas/dormitório
Arranjos familiares não tradicionais	Tipologia familiar
Baixa renda	População pobre
Baixos níveis de educação e habilidades dos pais	Nível de instrução da mãe
	Ocupação da mãe
	Máquina de lavar
Estrutura domiciliar	Internet
	Esgoto
Discórdia conjugal	-
Psicopatologia familiar	-
Pouca dedicação dos pais aos filhos	-
Falta de estímulo cognitivo/emocional às crianças	-
Falta de conhecimentos de como educar uma criança	-
Baixo nível de capital cultura familiar	-

A TABELA 4.5 apresenta os resultados. O excessivo número de filhos na família é adverso ao desenvolvimento ao se tornar um empecilho na alocação de recursos financeiros e no tempo despendido pelos pais na criação de cada criança. No Brasil 11,81% das crianças de 0 a 6 anos fazem parte de famílias que possuem número grande de filhos, o que significa cerca de 2.260.000 crianças. Na população não pobre apenas 3,28% das crianças estão sob essa situação de risco social, enquanto que na população pobre a situação de risco ocorre com 20,04% das crianças de 0 a 6 anos. As famílias numerosas são, na maioria, pertencentes à população pobre, 85,06% das crianças de 0 a 6 anos em famílias com 3 filhos ou mais possuem renda familiar igual ou inferior à meio salário mínimo, o que representa cerca de 1.920.000 crianças.

A frequência à creche, pré-escola e escola representa um mecanismo de compensação desse e de outros, riscos. 45,90% das crianças em famílias numerosas frequentam creche, escola ou pré-escola. A frequência na população não pobre é quase semelhante para crianças em famílias com até 3 filhos e em famílias com mais de 3 filhos, 51,71% e 52,37% (GRÁFICO 4.5). Já na população pobre há uma maior taxa de frequência entre as crianças de famílias mais numerosas, 44,77%, comparado a famílias menos numerosas, 40,15%, mas

ainda assim existem pouco mais de 1.060.000 crianças de famílias pobres e numerosas que não frequentam a educação infantil ou ensino fundamental.

Assim como muitos filhos em uma família é prejudicial ao desenvolvimento de habilidades das crianças, a existência de muitas pessoas em uma mesma casa também é. A variável utilizada nessa análise é um índice de pessoas residentes no domicílio por dormitório, sendo considerado inadequado a existência de 3 ou mais pessoas por cômodo. Existem 9.389.601 crianças nessa situação (48,99% das crianças de 0 a 6 anos), sendo que apenas 41,04% frequentam instituições de cuidado e ensino. Os pobres representam 63,91% dessa população, e a taxa de frequência entre eles é de 38,84%, o que significa que existem cerca de 3.670.000 crianças pobres morando em casas com mais pessoas do que o indicado fora da creche, pré-escola ou escola. A frequência à creche, pré-escola e escola é maior para as crianças que moram em domicílios onde há menos pessoas por dormitório (GRÁFICO 4.6), o que nos leva a pensar que haveria, portanto, menos pessoas disponíveis para o cuidado com as crianças, criando a necessidade de um atendimento externo, o contrário do que poderia acontecer em um domicílio onde moram mais pessoas.

Pressupõe-se que mulheres que moram com os filhos sozinhas, sem a presença de um companheiro, têm maior necessidade do compartilhamento das atividades, seja a creche, pré-escola e escola. Os dados da PNAD 2009 mostram que 19,79% das crianças de 0 a 6 anos vivem em famílias uniparentais femininas, na população pobre seriam 23,15% e na população não pobre seriam 16,56%. As taxas de frequência à creche, pré-escola e escola mostram que são exatamente as mães solteiras que mais utilizam esse serviço, na população pobre, por exemplo, enquanto a taxa de frequência é de 43,90% nas famílias uniparentais femininas, ela é de 40,41% nas famílias biparentais (GRÁFICO 4.7). Tem-se, portanto, 2.176.548 crianças de famílias pobres vivendo apenas com a mãe, sendo que 1.221.043 não frequentam creches, pré-escolas ou escolas.

A tipologia familiar e o número de filhos foram as únicas características analisadas em que as crianças em situação adversa possuem maior taxa de frequência às instituições de cuidado e ensino. Assim como analisado no número de filhos por família e no número de pessoas por dormitório, também na análise da tipologia familiar a taxa de frequência às instituições de cuidado e ensino é menor na população pobre, que seria a mais carente desse serviço, pelas justificativas assistenciais e educacionais das crianças.

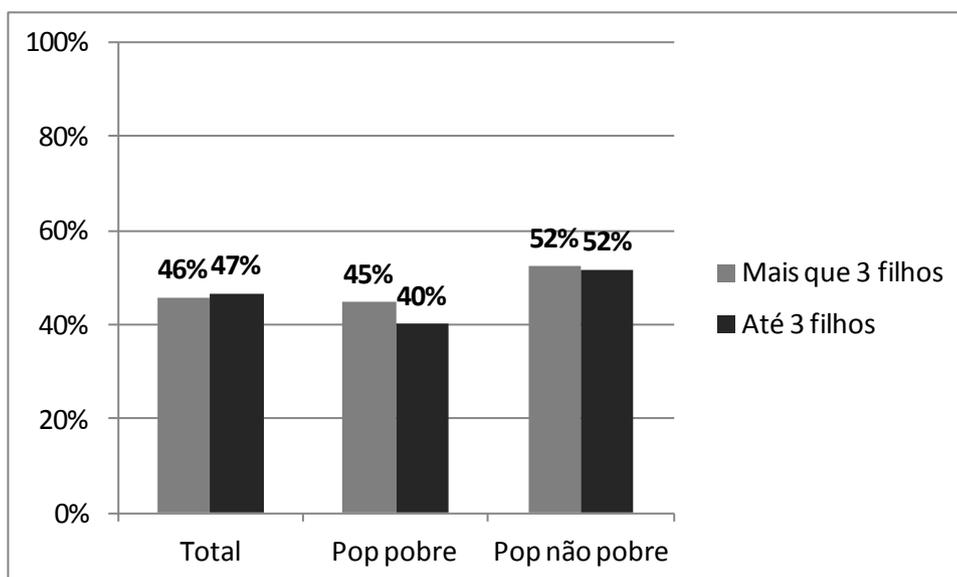
**TABELA 4.5:** Porcentagem e total de crianças de 0 a 6 anos que apresentam características familiares ou domiciliares adversas ao desenvolvimento das habilidades e respectiva taxa e total de frequência à educação infantil e ensino fundamental, população total, pobre e não pobre, Brasil 2009

	População total		População pobre		População não-pobre	
	%	Total	%	Total	%	Total
Número de filhos na família: maior que 3	11,81% a <sup>1</sup>	2.263.306	20,47% a	1.925.213	3,46% b	338.089
Frequência à creche, pré-escola e escola	45,90% a	1.038.857	44,77% a	861.918	52,37% a	177.057
Pessoas por dormitório: igual ou maior que 3	48,99% a	9.389.601	63,81% a	6.000.436	34,71% a	3.389.177
Frequência à creche, pré-escola e escola	41,04% a	3.853.488	38,84% a	2.330.660	44,93% a	1.522.833
Tipologia familiar: monoparental feminino	19,79% a	3.793.394	23,15% a	2.176.548	16,56% a	1.616.846
Frequência à creche, pré-escola e escola	48,10% a	1.824.459	43,90% a	955.529	53,74% a	868.930
Nível de instrução da mãe: inferior ao EM	55,85% a	10.706.166	75,33% a	7.083.452	37,10% a	3.622.711
Frequência à creche, pré-escola e escola	42,45% a	4.544.768	40,13% a	2.842.589	46,99% a	1.702.312
Mãe desempregada	8,80% a	1.686.397	11,46% a	1.077.916	6,23% a	608.480
Frequência à creche, pré-escola e escola	44,51% a	750.697	42,28% a	455.709	48,48% a	294.988
Não existência de esgoto	29,15% a	5.587.439	38,16% a	3.588.510	20,47% a	1.998.698
Frequência à creche, pré-escola e escola	41,76% a	2.333.315	39,45% a	1.415.667	45,91% a	917.602
Não posse de micro com internet	80,97% a	15.520.156	96,45% a	9.069.581	66,06% a	6.450.578
Frequência à creche, pré-escola e escola	52,04% a	8.077.239	42,86% a	3.887.331	54,94% a	3.544.258
Não posse de máquina de lavar	63,51% a	12.173.526	83,97% a	7.896.415	43,80% a	4.276.647
Frequência à creche, pré-escola e escola	43,34% a	5.276.006	40,97% a	3.235.161	47,74% a	2.041.671

<sup>1</sup> Intervalo do coeficiente de variação, vide capítulo 2

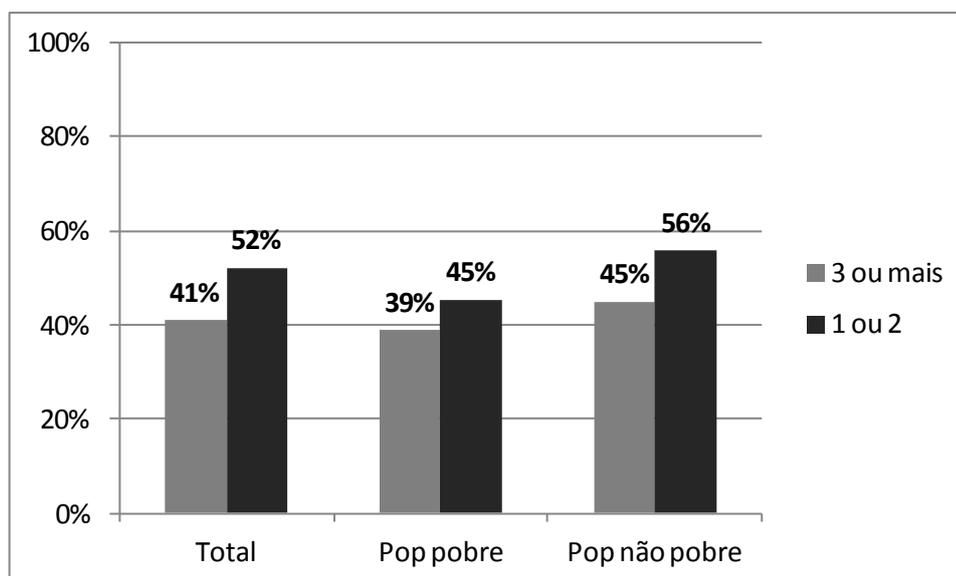
Fonte: PNAD 2009

**GRÁFICO 4.5:** Taxa de frequência de crianças de 0 a 6 anos à educação infantil e ensino fundamental por número de filhos na famílias, população total, pobre e não pobre, Brasil 2009



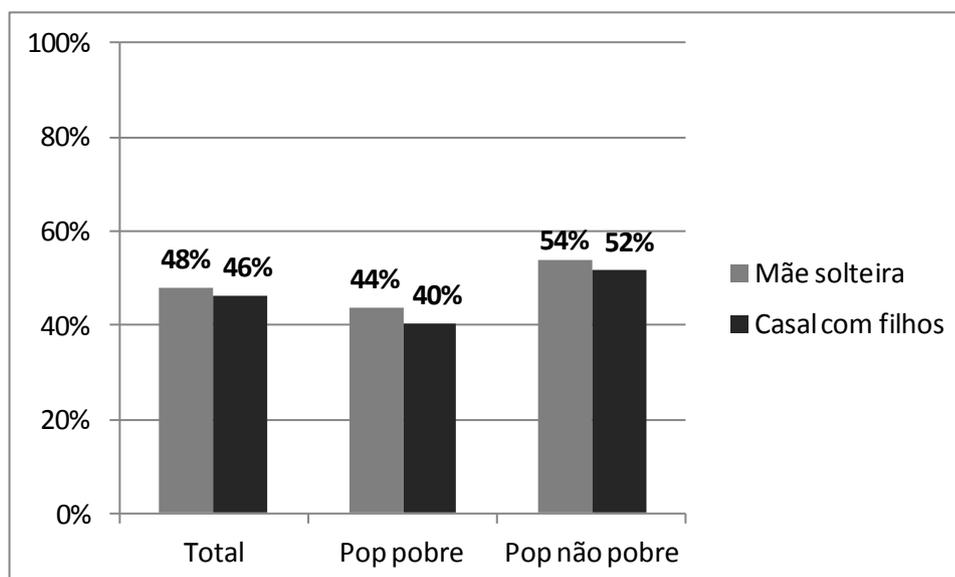
Fonte: PNAD 2009

**GRÁFICO 4.6:** Taxa de frequência de crianças de 0 a 6 anos à educação infantil e ensino fundamental por número de componentes domiciliares por dormitório, população total, pobre e não pobre, Brasil 2009



Fonte: PNAD 2009

**GRÁFICO 4.7:** Taxa de frequência de crianças de 0 a 6 anos à educação infantil e ensino fundamental, por tipo de família, população total, pobre e não pobre, Brasil 2009



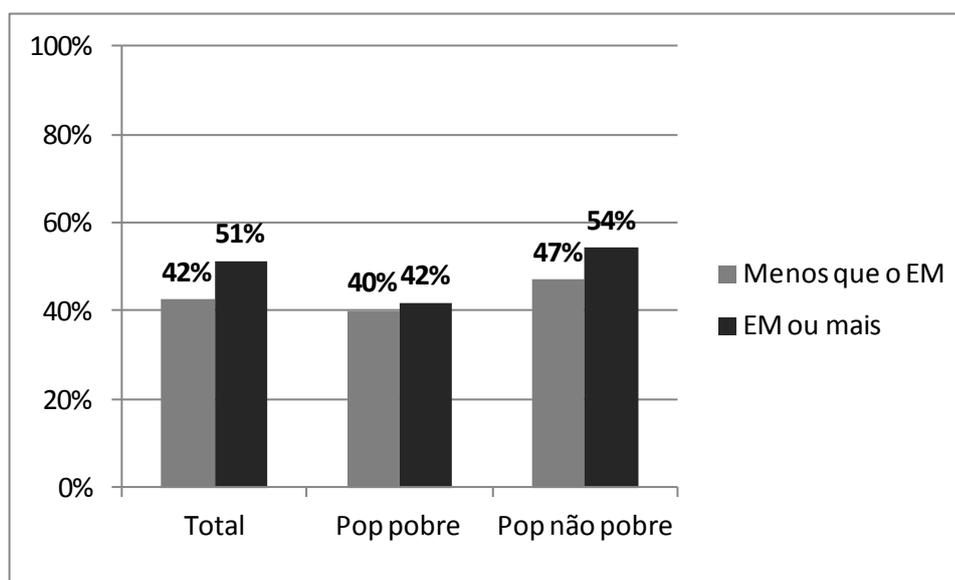
Fonte: PNAD 2009

A não frequência à creche, pré-escola e escolas tem impacto no desenvolvimento das habilidades, pois se pressupõe que aquele que se responsabilizará pelo cuidado da criança não tenha os conhecimentos/habilidades necessários para educar uma criança. O baixo nível educacional da mãe é uma característica que prediz o nível de conhecimentos em como se deve criar uma criança, influenciando no seu desenvolvimento cognitivo. Mais da metade das crianças em idade de desenvolvimento infantil possuem mãe com escolaridade inferior ao ensino médio, seriam 10.706.166 crianças. Na população pobre 75,33% estão nessa situação, sendo que 40,13% desses têm acesso à educação, e na população não pobre são 37,10% em risco, havendo o acesso de 46,99% na população não pobre. Portanto, existem 4.240.863 crianças pobres filhas de mães com baixa escolaridade que não possuem acesso a creches pré-escolas e escolas.

Outra adversidade relacionada à mãe é a sua situação de desemprego, que coloca as crianças em situação de risco social ainda maior, devido à redução na renda familiar. Nessa situação se encontram 8,8% das crianças, sendo maior a prevalência na população pobre, 11,46% do que na população não pobre, 6,23%. No caso, seriam 1.686.397 crianças sob esse risco, das quais 63,91% estão na população pobre. Das crianças de 0 a 6 anos com mães desempregadas, 42,45% frequentam instituições de cuidado e ensino, por outro lado, existem quase 1 milhão dessas crianças que são pobres e não frequentam creches ou escolas.

Para ambas as características maternas analisadas, as crianças em situação de risco social possuem menor taxa de frequência às instituições de cuidado e ensino, ainda agravado pela menor frequência na população pobre em relação à população não pobre (GRÁFICOS 4.8 E 4.9). Enquanto 42,45% das crianças filhas de mães com escolaridade inferior ao ensino médio frequentam escolas ou creches, 51,31% das crianças filhas de mãe mais escolarizadas o fazem. Essa diferença diminui na população pobre, 40,13% e 41,70% respectivamente, mas persiste na população não pobre, 46,99% e 54,17% respectivamente. Na análise da situação de desemprego da mãe, 44,51% dos filhos de mães desempregadas e 53,83% dos filhos das mães empregadas frequentam escolas e creches. A diferença é maior na população não pobre, comparada com a população pobre. A taxa de frequência é de 42,28% das crianças com mãe desempregada e 47,53% com mãe empregada em famílias pobres, e na população não pobre, 48,48% das crianças filhas de mães desempregada e 57,28% de mães empregadas.

**GRÁFICO 4.8:** Taxa de frequência de crianças de 0 a 6 anos à educação infantil e ensino fundamental por escolaridade da mãe, população total, pobre e não pobre, Brasil 2009

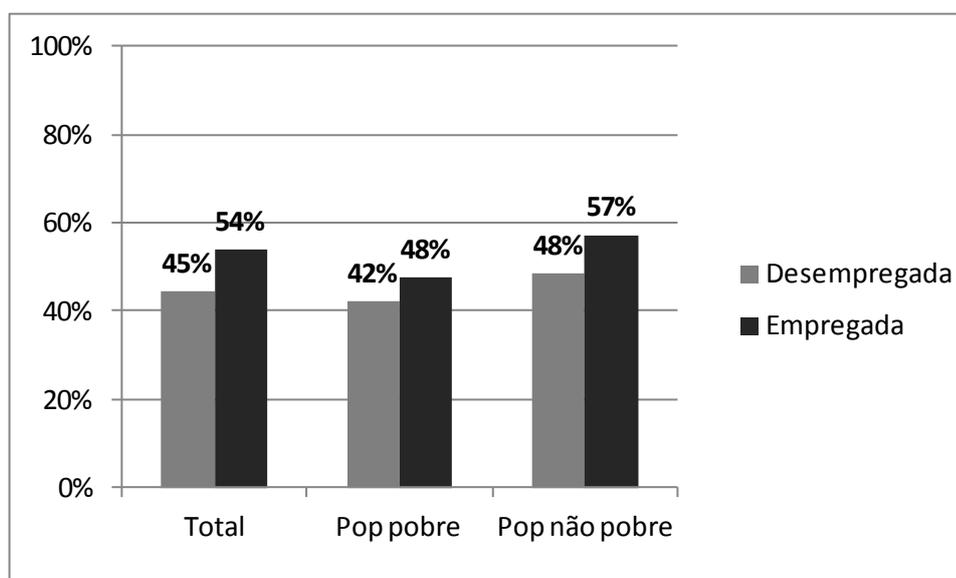


Fonte: PNAD 2009

Considera-se, também, que as características do domicílio também possuem influência sobre o desenvolvimento das habilidades das crianças. 29,15% das crianças de 0 a 6 anos moram em domicílios com escoamento precário de esgoto, o que representa 5.587.439 crianças. Entre os pobres a porcentagem é de 38,16%, 3.588.510 crianças e entre os não pobres é de 20,47%, 1.998.698 crianças. A frequência total à creches, pré-escolas e escolas é de 41,76% das crianças moradoras de domicílios com esgoto precário, contra 49,74% das

crianças que moram em domicílios com rede de esgoto adequada (GRÁFICO 4.10). Na população pobre as frequências são, respectivamente, 39,45% e 43,19%, e na população não pobre é de 45,91% e 53,75%. Assim, subsistem 2.172.843 crianças pobres em domicílios com esgotamento precário que não frequentam creches e escolas.

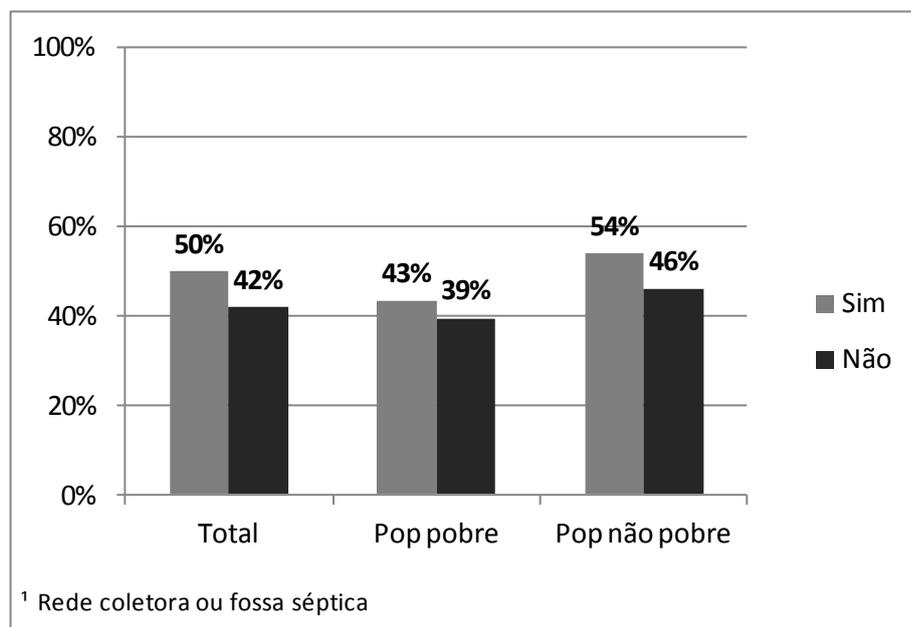
**GRÁFICO 4.9:** Taxa de frequência de crianças de 0 a 6 anos à educação infantil e ensino fundamental por condição de emprego e desemprego da mãe, população total, pobre e não pobre, Brasil 2009



Fonte: PNAD 2009

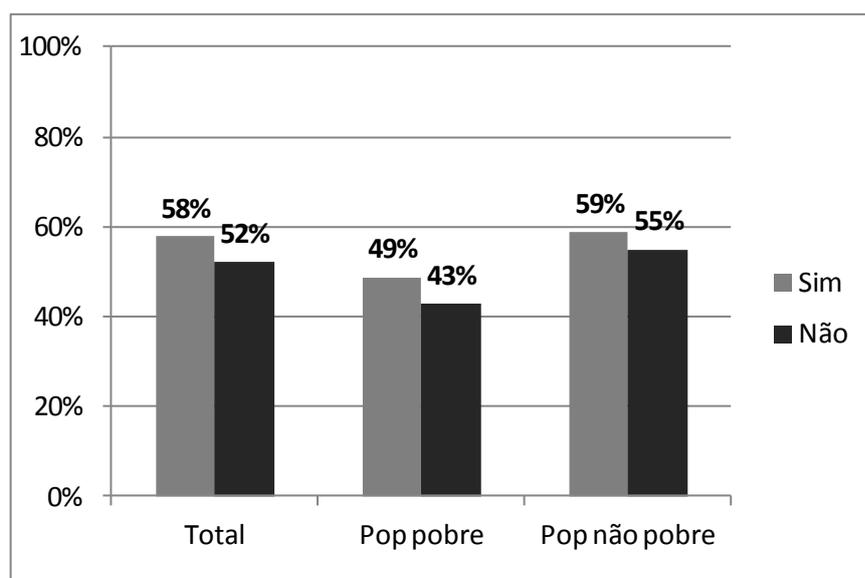
A posse de microcomputador com internet é a variável adversa analisada que possui o pior desempenho. 80,97% das crianças não possuem essa forma de acesso à informação em seu domicílio, ou seja, mais de 15 milhões de crianças de 0 a 6 anos. Entre os pobres essa porcentagem chega a 96,45% das crianças, entre a população não pobre a porcentagem é de 66,06%. A frequência à educação infantil e ensino fundamental é maior entre aqueles que possuem internet no domicílio, e também maior para a população não pobre (GRÁFICO 4.11)

**GRÁFICO 4.10:** Taxa de frequência de crianças de 0 a 6 anos à educação infantil e ensino fundamental por existência de rede de esgoto<sup>1</sup>, população total, pobre e não pobre, Brasil 2009



Fonte: PNAD 2009

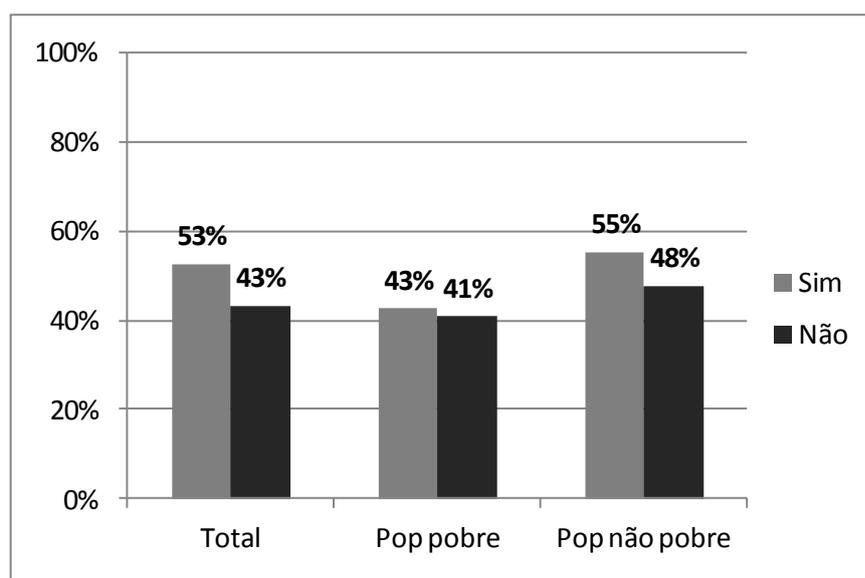
**GRÁFICO 4.11:** Taxa de frequência de crianças de 0 a 6 anos à educação infantil e ensino fundamental por existência de microcomputador com internet no domicílio, população total, pobre e não pobre, Brasil 2009



Fonte: PNAD 2009

Por fim, a inexistência de máquina de lavar roupas acontece nos domicílios de 83,97% das crianças pobres e nos domicílios de 43,80% das crianças não pobres, configurando 63,51% das crianças totais. A taxa de frequência à creche, pré-escola e escolas das crianças de 0 a 6 anos é maior nas casas que possuem esse utensílio, 52,59%, comparado com a taxa de frequência de 43,34% entre as crianças que moram em domicílios que não possuem máquina de lavar (GRÁFICO 4.12). Entre a população não pobre a taxa também é maior, 53,30% das crianças em domicílios que possuem o utensílio frequentam instituições de cuidado e ensino, contra 42,59% na população pobre.

**GRÁFICO 4.12:** Taxa de frequência de crianças de 0 a 6 anos à educação infantil e ensino fundamental por existência de máquina de lavar no domicílio, população total, pobre e não pobre, Brasil 2009



Fonte: PNAD 2009

Muitas crianças possuem a combinação de duas ou mais características adversas ao desenvolvimento das habilidades na infância, o que torna ainda mais relevante o cenário apresentado. As TABELAS 4.6 e 4.7 apresentam a porcentagem e total de crianças de 0 a 6 anos que apresentam duas características familiares ou domiciliares adversas ao desenvolvimento das habilidades, e respectiva taxa e total de frequência à educação infantil e ensino fundamental, para população total e população pobre.

**TABELA 4.6:** Porcentagem e total de crianças de 0 a 6 anos que apresentam duas características familiares ou domiciliares adversas ao desenvolvimento das habilidades e respectiva taxa de frequência e total de frequência à educação infantil e ensino fundamental, população total, Brasil 2009

	Número de filhos na família maior que 3		Pessoas por dormitório maior que 2		Tipologia familiar: mãe e filhos		Nível de instrução da mãe inferior ao EM		Desocupação da mãe		Não existência de esgoto		Não posse de micro com internet	
	%	Total	%	Total	%	Total	%	Total	%	Total	%	Total	%	Total
Número de filhos na família maior que 3 Frequência à creche, pré-escola e escola														
Pessoas por dormitório maior que 2 Frequência à creche, pré-escola e escola	9,95%	1.907.205												
Tipologia familiar: mãe e filhos Frequência à creche, pré-escola e escola	2,09%	400.609	9,42%	1.805.615										
Nível de instrução da mãe inferior ao EM Frequência à creche, pré-escola e escola	10,08%	1.932.123	33,46%	6.413.575	11,54%	2.211.974								
Desocupação da mãe Frequência à creche, pré-escola e escola	0,82%	157.177	5,05%	967.978	2,67%	511.783	5,18%	992.897						
Não existência de esgoto Frequência à creche, pré-escola e escola	4,41%	845.304	15,40%	2.951.855	5,79%	1.109.821	19,64%	3.764.573	2,38%	456.196				
Não posse de micro com internet Frequência à creche, pré-escola e escola	11,18%	2.142.970	43,36%	8.311.196	16,96%	3.250.874	52,08%	9.982.636	7,43%	1.424.174	27,09%	5.192.581		
Não posse de máquina de lavar Frequência à creche, pré-escola e escola	9,30%	1.782.614	36,15%	6.929.192	13,80%	2.645.169	42,66%	8.177.021	5,94%	1.138.573	23,89%	4.579.208	21,19%	4.061.675
	45,40%	809.307	39,70%	2.750.889	45,74%	1.209.900	41,33%	3.379.563	43,29%	492.888	40,75%	1.866.027	49,29%	2.002.000

Fonte: PNAD 2009

**TABELA 4.7:** Porcentagem e total de crianças de 0 a 6 anos que apresentam duas características familiares ou domiciliares adversas ao desenvolvimento das habilidades e respectiva taxa de frequência e total de frequência à educação infantil e ensino fundamental, população pobre, Brasil 2009

	Número de filhos na família maior que 3		Pessoas por dormitório maior que 2		Tipologia familiar: mãe e filhos		Nível de instrução da mãe inferior ao EM		Desocupação da mãe		Não existência de esgoto		Não posse de micro com internet	
	%	Total	%	Total	%	Total	%	Total	%	Total	%	Total	%	Total
Número de filhos na família maior que 3 Frequência à creche, pré-escola e escola														
Pessoas por dormitório maior que 2 Frequência à creche, pré-escola e escola	17,76%	1.670.124												
Tipologia familiar: mãe e filhos Frequência à creche, pré-escola e escola	3,75%	352.644	13,52%	1.271.401										
Nível de instrução da mãe inferior ao EM Frequência à creche, pré-escola e escola	18,35%	1.725.607	50,66%	4.763.992	17,00%	1.598.655								
Desocupação da mãe Frequência à creche, pré-escola e escola	1,55%	145.760	7,47%	702.468	3,96%	372.393	8,00%	752.308						
Não existência de esgoto Frequência à creche, pré-escola e escola	8,16%	767.354	23,41%	2.201.442	8,16%	767.354	29,73%	2.795.765	3,70%	347.943				
Não posse de micro com internet Frequência à creche, pré-escola e escola	20,14%	1.893.936	62,11%	5.840.733	22,33%	2.099.880	72,70%	6.836.601	10,67%	1.003.391	37,56%	3.532.087		
Não posse de máquina de lavar Frequência à creche, pré-escola e escola	17,46%	1.641.913	54,48%	5.123.219	19,43%	1.827.169	64,85%	6.098.399	9,06%	851.989	34,60%	3.253.733	82,56%	7.763.821
	45,02%	739.189	38,84%	1.989.858	43,78%	799.934	40,15%	2.448.507	41,47%	353.320	39,37%	1.280.995	44,28%	3.437.820

Fonte: PNAD 2009

Considerando as crianças de famílias com 3 ou mais filhos, observa-se que 84,25% moram em domicílios em que o número de pessoas por dormitório é igual ou superior a 3. Dessas, 44,88% frequentam creches, pré-escolas e escolas. Na população pobre a porcentagem de crianças com essas duas características é de 86,76%, sendo a taxa de frequência igual a 44,02%, existindo, portanto, 934.936 crianças pobres, em famílias e domicílios numerosos que não frequentam nenhum nível educacional.

Com relação ao nível de instrução da mãe, 85,35% das crianças de 0 a 6 anos em famílias com 3 ou mais filhos possuem mãe com nível inferior ao ensino médio, e na população pobre essa porcentagem é de 89,64%. As taxas de frequência são, respectivamente, 45,24% e 44,72%, havendo, portanto, 953.916 crianças consideradas pobres, em famílias numerosas e com nível de instrução da mãe baixo que não frequentam creches, pré-escolas ou escolas.

Entre as crianças em domicílios cujo número de pessoas por dormitório é inadequado (igual ou superior a 3), 68,29% possuem mãe com baixo nível de instrução, e considerando a população pobre, essa porcentagem é de 79,39%. As taxas de frequência às instituições de cuidado e ensino são bastante semelhantes, 38,97 e 38,10%, respectivamente. Há, portanto, 2.948.911 crianças pobres cuja mãe possui baixa instrução e que moram em domicílios numerosos que não frequentam creches, pré-escolas e escolas.

Ainda, as mães solteiras possuem, em 58,31% dos casos, baixa instrução. Na população pobre essa porcentagem sobe para 73,43%. Apesar dessa diferença, as taxas de frequência das crianças às creches, pré-escolas e escolas são semelhantes, 43,80% e 42,16%, respectivamente. Existem, por fim, 924.622 crianças nessas situações de adversidades que não frequentam algum nível educacional.

Por fim, a TABELA 4.8 mostra que enquanto na população pobre apenas 1,07% das crianças de 0 a 6 anos possuem nenhuma das características de adversidades analisadas aqui, na população não pobre essa porcentagem é de 19,85%. Na outra ponta da análise, enquanto 10,99% das crianças pobres possuem 5 dessas características adversas, apenas 1,61% das crianças não pobres apresentam esse mesmo número de características.

Os dados da PNAD 2009 mostraram quantas são as crianças de 0 a 6 anos que não frequentam creches, pré-escolas e escolas, considerando características consideradas adversas ao desenvolvimento das habilidades. Esse desenvolvimento é essencial para que a criança consiga continuar adquirindo outras habilidades, que no futuro serão determinantes para o seu sucesso socioeconômico. Portanto, fica claro que uma maior atenção à essa população é

necessária, através de programas que incluam a educação de crianças e envolvam a sua família.

**TABELA 4.8:** Porcentagem de crianças de 0 a 6 anos por quantidade de características adversas ao desenvolvimento das habilidades, população total, pobre e não pobre, Brasil 2009

	Pobre	Não pobre	Total
0	1,07%	19,85%	10,64%
1	6,72%	26,12%	16,60%
2	19,39%	26,20%	22,86%
3	32,43%	18,46%	25,31%
4	27,16%	7,56%	17,18%
5	10,99%	1,61%	6,21%
6	2,07%	0,19%	1,11%
7	0,17%	0,01%	0,09%
8	0,00%	0,00%	0,00%

Fonte: PNAD 2009

## COMENTÁRIOS FINAIS

Este trabalho busca discutir que a existência de um sistema de educação infantil de qualidade permite enfrentar diversos problemas sociais decorrentes dos baixos níveis de educação e habilidades, como desemprego, baixa renda, violência e falta de planejamento familiar a partir da teoria do ciclo de formação das habilidades, discutida no capítulo 1.

O ciclo de formação das habilidades justifica o investimento que deve ser feito em crianças logo após o nascimento. As habilidades são diversas, e muitas delas são desenvolvidas apenas na infância, como é o caso das habilidades não-cognitivas, aquelas relacionadas à formação social e emocional, saúde física e mental, atenção, motivação, perseverança, autoconfiança, aversão ao risco, autocontrole, etc. Essas habilidades, em conjunto com todas as outras, são determinantes do sucesso socioeconômico de indivíduos no futuro. Assim, para o desenvolvimento eficiente das habilidades individuais é necessário que a educação comece o mais cedo possível, considerando que investimentos nas políticas educacionais que se baseiam na interferência em pessoas adolescentes ou adultas teriam eficácia reduzida quando comparadas com as interferências na educação infantil no que tange a amenização de problemas sociais.

Através da análise de dados do SAEB 2005, mostra-se que a educação infantil tem efeito positivo sobre os resultados socioeconômicos futuros do indivíduo, representados aqui pelo resultado na proficiência no teste de matemática. Cabe ressaltar que os dados do SAEB não permitem a realização de análises sobre a qualidade da educação infantil, pode-se entretanto supor que a qualidade causa algum impacto no resultado do teste. Ainda, a análise adequada do efeito da educação infantil exigiria a separação entre aqueles que começaram a estudar entre a idade de 0 e 3 anos (creche), e aqueles que começaram a estudar entre as idades de 4 a 6 anos (pré-escola), captando o efeito do início da escolarização em cada fase da infância, o que não é possível realizar com os dados disponíveis para esta dissertação.

A empiria deste trabalho mostra que a educação infantil possui efeito positivo sobre o resultado educacional, medido pelo teste SAEB. Considerando um modelo hierárquico com efeito aleatório somente no intercepto, o efeito da educação infantil é de 2,94 pontos a mais na 4ª série e 0,77 pontos a mais na 8ª série. Já no modelo que permite que esse efeito varie pelas características da escola, viu-se que os alunos da 4ª série que estudam em escolas privadas, na zona urbana, nas regiões Sul, Sudeste e Distrito Federal, tem efeito da educação infantil igual a 36,91 pontos positivos. Os alunos que estudam em escola pública tem esse efeito 11,26 pontos menor, em média. Se a escola está localizada na área rural esse efeito é menor em 2,78

pontos. E, caso a escola esteja nas regiões Norte, Nordeste ou Centro-Oeste (Exceto o DF), o efeito da educação infantil cai 7,01 pontos.

Para a 8ª série o efeito de ter começado a estudar na educação infantil para os alunos de escola privada, nas regiões Sul, Sudeste e Distrito Federal é igual a uma nota em média 4,85 pontos maior. Esse efeito pode variar dependendo da escola em que o aluno estuda, sendo, em media, 2,65 pontos menor se o aluno estuda em escola pública, e 0,75 pontos menor se a escola se encontra nas regiões Norte, Nordeste ou Centro-Oeste.

O resultado sugere que o efeito da educação infantil na 4ª série não é tão expressivo quanto esperado qual se considera esse efeito como fixo. A utilização de modelos multiníveis, que permitem considerar tal efeito como aleatório levando em consideração a estrutura hierárquica dos dados, ajudam a evidenciar o efeito em questão.. Já na 8ª série, em ambos os casos, o efeito é reduzido, devendo-se considerar um possível efeito da qualidade da educação no ensino fundamental, sugerindo melhorias nesse nível. Cabe mais uma vez ressaltar as limitações dos dados utilizados neste estudo, como discutido na seção 1.4.1.

Os dados analisados mostram que a geração que entrou há pouco na população em idade ativa ainda não experimentou uma maior taxa de frequência à educação infantil. Entre os alunos da 8ª série da prova de matemática do SAEB 2005 (que possuem agora 20 anos ou pouco mais), apenas 31,02% começaram a estudar na pré-escola. A PNAD mostra que em 1995 53,48% das crianças de 4 a 6 anos frequentavam creche, pré-escola ou escola (população agora com 20 a 22 anos). Todavia houve melhoras ao longo dos anos, entre os respondentes da 4ª série do teste de matemática do SAEB houve 78,72% de frequência a maternal ou pré-escola, e a PNAD 2009 mostra frequência de 81,28% da população de 4 a 6 anos. Assim, tem-se que os recentes entrantes no mercado de trabalho possuem, em certa medida, potencial para uma despreparação profissional, mas o aumento da taxa de frequência à educação infantil aponta para uma melhora nesse quadro (se contarmos com a qualidade da educação).

As famílias são as principais fontes de debilidades sociais no desenvolvimento das habilidades. Crescer em uma família que ofereça ambiente propício ao desenvolvimento dessas habilidades torna as crianças mais aptas à vida socioeconômica futura. A literatura cita diversos fatores associados ao risco no desenvolvimento das habilidades. A quantificação das crianças em situações de adversidade social feitas no capítulo 4 mostrou que na população total 11,81%, ou 2.263.306 crianças de 0 a 6 anos, estão em risco por serem de famílias com mais que 3 filhos.

Um número grande de filhos significa aumento da razão de dependência demográfica e pode ser prejudicial ao desenvolvimento humano da família na medida em que exista uma concorrência entre as crianças pelos escassos recursos e na medida em que se reduz a disponibilidade de tempo que os pais poderiam se dedicar às suas próprias carreiras profissionais. Famílias com baixo nível de educação e renda e com muitos filhos têm maior probabilidade de serem pobres, pois podem ser vítimas da “armadilha da pobreza” (AZARIADIS, 2004). Isto é, o círculo vicioso da pobreza entre gerações tende a se manter quando a educação dos filhos não supera significativamente a educação dos pais. Como dito, 85,06% das crianças de 0 a 6 anos provenientes de famílias numerosas pertencem à população pobre, sendo que 55,23% não frequentam creches, pré-escolas ou escolas, totalizando 1.063.395 crianças.

Existem 10.706.166 crianças, ou 55,85%, que possuem mãe com escolaridade baixa: menos que o ensino médio. Ainda, 3.793.394 crianças, 19,79%, que moram apenas com a mãe. Se cruzarmos as duas variáveis, chegamos a um número de 2.211.974 crianças, 11,54%, que moram apenas com a mãe, que possui baixa escolaridade. Esse quadro é pior para a população pobre, aquela que mais precisa do suporte governamental na educação das crianças. Considerando a população pobre, essa porcentagem seria de 17%.

Heckman (2008) diz que a medida de adversidade adequada não seria exatamente a pobreza, ou falta de recursos materiais, mas sim a falta de habilidades na educação dos filhos. Não é possível aumentar generalizadamente a educação dos pais (ou seria altamente custoso), nem se pode desenvolver suas habilidades a ponto de poderem criar seus filhos da maneira mais eficiente possível. A educação infantil de qualidade garante a todas as crianças, independente de características socioeconômicas, uma formação educacional inicial de qualidade. Assim, crianças em desvantagens sociais poderiam concorrer com as outras crianças em situação de igualdade desde o início. Essa é considerada a melhor política direcionada a crianças pobres, em busca da superação da situação de risco social.

A educação infantil continua cumprindo, por vezes, a função de cuidados com as crianças enquanto os seus responsáveis estão fora exercendo outras atividades. Todavia, o caráter educacional das creches e pré-escolas já está institucionalizado e em expansão no território nacional. A educação infantil sai, portanto, do âmbito da assistência social e passa a ser responsabilidade da área educacional, levando a discussões pedagógicas para essas crianças, e não apenas a discussões relativas aos cuidados.

Os marcos legais já foram postos, a educação para crianças de 0 a 6 anos é amplamente embasada pela legislação, e o atendimento está avançando. Promover uma

educação infantil de qualidade é uma condição básica para o exercício da cidadania e para o desenvolvimento humano e sustentável da população brasileira, e para a mobilidade social intergeracional ascendente.

Finalmente, cabe ressaltar que para um completo desenvolvimento econômico e social do Brasil é necessário que a população brasileira se qualifique e, assim, aumente a sua escolaridade. Nesse processo devem ser consideradas as diversas dificuldades que são enfrentadas, como as de cunho cognitivo e não cognitivo, decorrente da baixa educação inicial que impede o indivíduo de prosseguir com a qualificação (processo discutido neste trabalho), os gargalos institucionais, como a baixa qualidade do ensino público no nível básico e o número ainda insuficiente de escolas de ensino técnico, e dificuldades de acesso devido a contextos socioeconômicos.

## REFERÊNCIAS BIBLIOGRÁFICAS

ALBERNAZ, A.; FERREIRA, F. H. G.; FRANCO, C. Qualidade e equidade no ensino fundamental brasileiro. **Pesquisa e Planejamento Econômico**, Rio de Janeiro, v. 32, nº 3, p. 453-476, 2002. Disponível em: <http://ppe.ipea.gov.br/index.php/ppe/article/viewFile/139/74>. Acesso em 12/01/2012.

ALVES, J. E. D. CAVENAGHI, S. Família, trabalho e pobreza nas políticas de transferência de renda: O caso do Programa Bolsa Família no Recife. VII ENABER, USP, São Paulo, 2009. Disponível em: <http://aplicativos.fipe.org.br/enaber/pdf/117.pdf>. Acesso em 03/02/2012.

ARAÚJO, A. (Coord.). **Aprendizagem infantil**: Uma abordagem da neurociência, economia e psicologia cognitiva. Rio de Janeiro, Academia Brasileira de Ciências, 2011.

ASSOCIAÇÃO BRASILEIRA DE EMPRESAS DE PESQUISA. **Critério de classificação econômica Brasil**. [S.l.], 2011. Disponível em: <http://www.abep.org/novo/Utils/FileGenerate.ashx?id=197>. Acesso em 13/12/2011.

AZARIADIS, Costa. The theory of poverty traps: What have we learned? In: BOWLES, Samuel. *Poverty Traps*. University of Wisconsin-Madison, Santa Fe Institute, 2004.

BARROS, R.; CARVALHO, M.; FRANCO, S.; MENDONÇA, R.; PÀDUA, E.; ROSALÉM, A.; e TSUKADA, R.. A importância da qualidade da creche para a eficácia na promoção do desenvolvimento infantil. In: **Aprendizagem infantil**: Uma abordagem da neurociência, economia e psicologia cognitiva. ARAÚJO, A. (Coord.). Rio de Janeiro, Academia Brasileira de Ciências, 2011.

BRASIL. Constituição (1988). **Constituição da República Federativa do Brasil**. Brasília, DF: Senado, 1988.

BRASIL. Lei nº 9.394, de 20 de dezembro de 1996: Lei de Diretrizes e Bases da Educação Nacional. **Diário Oficial da União**, Poder Executivo, Brasília, DF, 23 de dezembro de 1996.

BRASIL. Lei nº 10.172, de 9 de janeiro de 2001: Plano Nacional de Educação 2001-2010. **Diário Oficial da União**, Poder Executivo, Brasília, DF, 10 de janeiro de 2001.

BRASIL. Projeto de Lei nº 8035 de 20 de dezembro de 2010: Plano Nacional de Educação 2011-2020. **Congresso Nacional**, Brasília, DF, 20 de dezembro de 2010.

BUJES, M. I. E. **Infância e maquinarias**. Rio de Janeiro: Editora DP&A. 2003.

CAMPOS, M. M.; FULLGRAF, J. e WIGGERS, V.. A Qualidade da educação infantil brasileira: alguns resultados de pesquisa. **Cadernos de Pesquisa**, v. 36, n. 127, p. 87-128, jan./abr. 2006. Disponível em: <http://www.scielo.br/pdf/cp/v36n127/a0536127.pdf>. Acesso em 14/02/2012.

CARVALHO, V. S. F.. **Os Jovens e o Mercado de Trabalho**: Mudanças na Decisão de Ingresso entre 1992 e 2004. Rio de Janeiro: ENCE, 2006. 148 p. Disponível em: [http://www.lep.ibge.gov.br/ence/pos\\_graduacao/mestrado/dissertacoes/pdf/2006/valeria\\_carvalho\\_TC.pdf](http://www.lep.ibge.gov.br/ence/pos_graduacao/mestrado/dissertacoes/pdf/2006/valeria_carvalho_TC.pdf). Acesso em 13/12/11. Acesso em 23/11/2011.

CERQUEIRA, C. A. e RIGOTTI, J. I. R.. As bases de dados do INEP e os indicadores educacionais: conceitos e aplicações. In: **Introdução à demografia da educação**. Orgs: Eduardo Luiz G. Rios-Neto e Juliana de Lucena Ruas Riani (Org.). Campinas: Associação Brasileira de Estudos Populacionais - ABEP, 2004.

- CHANTALA, K.; BLANCHETTE, D.; SUCHINDRAN, C. M.. Software to Compute Sampling Weights for Multilevel Analysis. Carolina do Norte (EUA): Carolina Population Center at UNC, 2006. Disponível em: [http://www.cpc.unc.edu/research/tools/data\\_analysis/ml\\_sampling\\_weights/Compute%20Weights%20for%20Multilevel%20Analysis.pdf](http://www.cpc.unc.edu/research/tools/data_analysis/ml_sampling_weights/Compute%20Weights%20for%20Multilevel%20Analysis.pdf). Acesso em 13/12/11.
- CIASCA, M. I. F. L.; MENDES, D. L. L. L.. Estudos de avaliação da Educação Infantil. Est. Aval. Educ. São Paulo, v. 20, n. 43, p. 293-304, maio/ago 2009. Disponível em: <http://www.fcc.org.br/pesquisa/publicacoes/eae/arquivos/1494/1494.pdf>. Acesso em 14/02/2012
- CUNHA, F. e HECKMAN, J.. *Investing in our young people*. Cambridge, MA: 2010. NBER Working paper 16201. Disponível em: [http://jenni.uchicago.edu/human-inequality/papers/inv-young-rep\\_all\\_2007-01-31b\\_mms.pdf](http://jenni.uchicago.edu/human-inequality/papers/inv-young-rep_all_2007-01-31b_mms.pdf). Acesso em 25 de janeiro de 2011. 33 p.
- CUNHA, F. e HECKMAN, J.. Capital Humano. In: ARAÚJO, A. (Coord.). **Aprendizagem infantil: Uma abordagem da neurociência, economia e psicologia cognitiva**. Rio de Janeiro, Academia Brasileira de Ciências, 2011.
- CUNHA, F.; HECKMAN, J. J.; LOCHNER, L.; MASTEROV, D. V. **Interpreting the evidence on life cycle skill formation**. Cambridge, MA, NBER, 2005. Working paper 11331. Disponível em: <http://www.nber.org/papers/w11331.pdf>. Acesso em 25 de janeiro de 2011. 109 p.
- CURI, A. Z. e MENEZES-FILHO, N. A.. **A relação entre educação pré-primária, salário, escolaridade e proficiência escolar no Brasil**. Estudos Econômicos. Vol 39, nº 4. Out-Dez 2009: páginas 811-850. . Acesso em 25 de janeiro de 2011. 33 p.
- DOYLE, O.; HARMON, C.; HECKMAN, J. e TREMBLAY, R. *Early childhood intervention: rationale, timing and efficacy*. UCD Geary Institute. Dublin: 2007. Discussion Paper Series. Disponível em: [http://www.ucd.ie/geary/static/publications/workingpapers/GearyWp\\_200705.pdf](http://www.ucd.ie/geary/static/publications/workingpapers/GearyWp_200705.pdf). Acesso em 25 de janeiro de 2011. 13 p.
- ESPING-ANDERSEN, G. *Investing in children and their life chance*. Barcelona, Universitat Pompeu Fabra, 2007a. Disponível em: [http://dcpis.upf.edu/~gosta-esping-andersen/materials/investing\\_children.pdf](http://dcpis.upf.edu/~gosta-esping-andersen/materials/investing_children.pdf). Acesso em 27 de janeiro de 2011.
- ESPING-ANDERSEN, G. *Childhood investments and skill formation*. Barcelona, Springer Science, 2007b. Disponível em: [http://iyyunim.edu.haifa.ac.il/user\\_files/file/lead/files/ma\\_articles/Esping-Andersen%20\(2008\)%20childhood%20investment%20and%20skill%20%20%20%20%20%20%20%20%20formation.pdf](http://iyyunim.edu.haifa.ac.il/user_files/file/lead/files/ma_articles/Esping-Andersen%20(2008)%20childhood%20investment%20and%20skill%20%20%20%20%20%20%20%20%20formation.pdf). Acesso em 27 de janeiro de 2011.
- ESPING-ANDERSEN, G. **Equal opportunities in an increasing hostile world**. Barcelona, Universitat Pompeu Fabra, 2007c. Disponível em: [http://dcpis.upf.edu/~gosta-esping-andersen/materials/equal\\_opportunities.pdf](http://dcpis.upf.edu/~gosta-esping-andersen/materials/equal_opportunities.pdf). Acesso em 27 de janeiro de 2011.
- HANUSHEK, E. A. e LUQUE, J. A. *Efficiency and equity in schools around the world*. National bureau of economic research. Cambridge: 2002. Working paper 8949.
- HECKMAN, J. J. *Investing in disadvantaged young children is an economically efficient policy*. Forum on "Building the economic case for investing in preschool". Nova York: 2006. Disponível em: <http://www.inpathways.net/IPCNLibrary/ViewBiblio.aspx?aid=9006>
- HECKMAN, J. J. *The case for investing in disadvantaged Young children*. **CESifo DICE Report**, n. 6, v. 2. 2008. Pág. 3-8. Disponível em: [http://econpapers.repec.org/article/cesifodic/v\\_3a6\\_3ay\\_3a2008\\_3ai\\_3a2\\_3ap\\_3a3-8.htm](http://econpapers.repec.org/article/cesifodic/v_3a6_3ay_3a2008_3ai_3a2_3ap_3a3-8.htm). Acesso em 11/10/2011.

HECKMAN, J. J. e MASTEROV, D. V. *The productivity argument for investing in young children*. University of Chicago: 2004. Disponível em: <http://jenni.uchicago.edu/Invest/>. Acesso em 15/09/2009.

HECKMAN, J. J. e RAUT, L. K. *Intergenerational long term effect of preschool: estimates from a structural dynamic programming model*. Western Economic Association Meeting. 2007. Disponível em: <http://www.e.u-tokyo.ac.jp/cirje/research/workshops/micro/micropaper08/micro1014.pdf>. Acesso em 15/09/2009.

HIDALGO, G. N. e URZÚA, S. **The effects of participation in Public Child Care Centers: evidence from Chile**. In: LACEA-LAMES, 2011, Santiago, Chile. Disponível em: <http://www.webmeets.com/lacea-lames/2011/prog/viewpaper.asp?pid=441>. Acesso em 14/02/2012.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICAS. Pesquisa Nacional por Amostra de Domicílios 1995 a 2009. Disponível em: [www.bme.ibge.gov.br](http://www.bme.ibge.gov.br). Acesso em: 13/07/2011.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICAS. **Microdados Pesquisa Nacional por Amostra de Domicílios 2009**. Disponível em <http://www.ibge.gov.br/home/estatistica/populacao/trabalhoerendimento/pnad2009/microdados.shtm>. Acesso em 20/01/2011

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICAS. Pesquisa Nacional de Saúde do Escolar 2009. Rio de Janeiro, 2009. Disponível em: <http://www.ibge.gov.br/home/estatistica/populacao/pense/pense.pdf>. Acesso em 12/01/2012.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICAS. **Pesquisa Nacional por Amostra de Domicílios: Síntese de Indicadores 2009**. Rio de Janeiro, 2010b. Disponível em: [http://www.ibge.gov.br/home/estatistica/populacao/trabalhoerendimento/pnad2009/pnad\\_sintese\\_2009.pdf](http://www.ibge.gov.br/home/estatistica/populacao/trabalhoerendimento/pnad2009/pnad_sintese_2009.pdf). Acesso em 13/06/2011.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICAS. Síntese de Indicadores Sociais: uma análise das condições de vida da população brasileira. Rio de Janeiro, 2010a. Disponível em: [http://www.ibge.gov.br/home/estatistica/populacao/condicaodevida/indicadoresminimos/sinteseindicsociais2010/SIS\\_2010.pdf](http://www.ibge.gov.br/home/estatistica/populacao/condicaodevida/indicadoresminimos/sinteseindicsociais2010/SIS_2010.pdf). Acesso em 12/01/2012.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICAS. **Banco Multidimensional de Estatísticas (BME)**. 2011. Disponível em: [www.bme.ibge.gov.br](http://www.bme.ibge.gov.br). Acesso em: 13/07/2011.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICAS. **Censo Demográfico 2010: Resultados preliminares da amostra**. In: Sistema IBGE de Recuperação automática, SIDRA, 2012a. Disponível em: <http://www.sidra.ibge.gov.br>. Acesso em 11/01/2012.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICAS. **Censo Demográfico 2000: Amostra- características gerais da população**. In: Sistema IBGE de Recuperação automática, SIDRA, 2012b. Disponível em: <http://www.sidra.ibge.gov.br>. Acesso em 11/01/2012.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA (Ministério da Educação). **Microdados do SAEB 2005**. Disponível em: <http://portal.inep.gov.br/web/guest/basica-levantamentos-acessar>. Acesso em 20/01/2011.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA (Ministério da Educação). **Microdados do SAEB 2005: Manual do usuário**. Brasília, abril de 2008. 217 p. Disponível em: <http://portal.inep.gov.br/web/guest/basica-levantamentos-acessar>. Acesso em 12/01/2010.

KAHLENBERG, R. D.. Learning from James Coleman. **The public interest**. 2001. Disponível em: [http://www.nationalaffairs.com/public\\_interest/detail/learning-from-james-coleman](http://www.nationalaffairs.com/public_interest/detail/learning-from-james-coleman). Acesso em 11/01/2012

KLEIN, R.. (2007). **A pré-escola no Brasil**. Revista eletrônica ibero-americana sobre calidad, eficacia y cambio em educación. Vol 5, n° 2e. 2007: página 273 – 288.

MACHADO, A. F.; MORO, S.; MARTINS, L.; RIOS, J. Qualidade do ensino em matemática: determinantes do desempenho de alunos em escolas estaduais mineiras. In: Encontro Nacional De Economia, 34ª edição, 2006, Salvador. Disponível em: <http://www.anpec.org.br/encontro2006/artigos/A06A091.pdf>. Acesso em 25/03/2011.

NOGUEIRA, C. M. M. et al. A influência da família no desempenho escolar: estudo de dados da geração escolar 2005. In: **XIV Congresso Brasileiro de Sociologia**, 2009, Rio de Janeiro. Disponível em: [http://www.educacao.ufrj.br/artigos/n8/numero8-09\\_a\\_influencia\\_da\\_familia\\_no\\_desempenho\\_escolar\\_estudo\\_de\\_dados\\_da\\_geracao\\_escolar\\_2005.pdf](http://www.educacao.ufrj.br/artigos/n8/numero8-09_a_influencia_da_familia_no_desempenho_escolar_estudo_de_dados_da_geracao_escolar_2005.pdf). Acesso em 15/10/2010.

PESSOA, D. G. C.; SILVA, P. L. N.. Análise de dados amostrais complexos. Associação Brasileira de Estatística, São Paulo, 1998. Disponível em: <http://www.ie.ufrj.br/download/livro.pdf>. Acesso em 12/02/2012.

PFERFFERMANN, D.; SKINNER, C. J.; HOLMES, D. J.; GOLDSTEIN, H. et al.. Weighting for unequal selection probabilities in multilevel models. *Journal of the Royal Statistical Society*, v. 60(1), p.23-40, 1998.

RABE-HASKETH, S. e SKRONDAL, A.. Multilevel modeling of complex survey data. *Journal of the Royal Statistical Society*, v. 169(4), p. 805-827, 2006.

RIETHER, M. M.; Rauter, R.. A metodologia de amostragem do SAEB. *R. Bras. Est. Pedag.*, Brasília, v. 81, n. 197, p. 143-153, jan./abr. 2000.

SILVA, P. L. N.; PESSOA, D. G. C.; LILA, M. F. Análise estatística de dados da PNAD: incorporando a estrutura do plano amostral. **Ciência e Saúde Coletiva**, v. 7, n° 4, 2002. p 659-670. Disponível em: <http://www.scielo.org/pdf/csc/v7n4/14597.pdf>. Acesso em 15/06/2011.

SILVA, M. D. F.; FARIAS, M. A.; SILVARES, E. F. M e ARANTES, M. C.. Adversidade familiar e problemas comportamentais entre adolescentes infratores e não-infratores. **Psicologia em Estudo**, Maringá, v. 13, n. 4, p. 791-798, out/dez 2008. Disponível em: <http://www.scielo.br/pdf/pe/v13n4/v13n4a17.pdf>. Acesso em 14/02/2012.

SOARES, Sergei S. D. **O bônus demográfico relativo e absoluto no acesso à escola**. IPEA, Textos para Discussão 1340, Brasília, 2008. Disponível em: [http://www.ipea.gov.br/sites/000/2/publicacoes/tds/td\\_1340.pdf](http://www.ipea.gov.br/sites/000/2/publicacoes/tds/td_1340.pdf). Acesso em 14/02/2012.

TODOS PELA EDUCAÇÃO; INSTITUTO PAULO MONTENEGRO/IBOPE; FUNDAÇÃO CESGRANRIO; INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA (INEP). **Prova ABC – Resultados da Avaliação de Aprendizagem de Leitura e Matemática**. 2011. Disponível em [http://www.todospelaeducacao.org.br/arquivos/biblioteca/resultados\\_prova\\_abc\\_coletiva\\_vfinal.ppt](http://www.todospelaeducacao.org.br/arquivos/biblioteca/resultados_prova_abc_coletiva_vfinal.ppt). Acesso em 18/11/2011.

UNESCO. **Unesco Institute for Statistics**. Disponível em: <http://stats.uis.unesco.org/unesco/ReportFolders/ReportFolders.aspx>. Acesso em 12/01/2012.

## APÊNDICE A – INCLUSÃO NO STATA DA AMOSTRAGEM COMPLEXA NAS ESTIMAÇÕES UTILIZANDO A PNAD 2009

```
svyset PSU [pweight=peso], singleunit(centered) strata(strata)
```

O software escolhido para análise (*STATA*) não calcula as estimativas das variâncias se ocorre de algum estrato possuir apenas uma UPA. Deve-se indicar na definição do plano amostral qual o procedimento a ser adotado, no caso adotou-se o procedimento de igualar a média do estrato singular à média geral (`singleunit(centered)`).

As estimações seguintes devem ser feitas utilizando o prefixo `svy`. Por exemplo:

```
** Frequência à creche, pré-escola e escola por região
svy linearized: prop freq_escola_0a6, over (região)
svy linearized: prop freq_escola_0a6, over (região grupidade)
svy linearized: prop freq_escola_0a6, over (região pobre)
svy linearized: prop freq_escola_0a6, over (uf_n)

** Frequência à creche, pré-escola e escola por renda
xtile quintile=renda_familiar[aw=peso], n(5)
svy linearized: prop freq_escola_0a6, over (quartil)
svy linearized: prop freq_escola_0a6, over (quartil grupidade)

** Frequência à creche, pré-escola e escola por cor
svy linearized: prop freq_escola_0a6, over (cor_n)
svy linearized: prop freq_escola_0a6, over (cor_n grupidade)

** Frequência à creche, pré-escola e escola por situação do domicílio
svy linearized: prop freq_escola_0a6, over (situação_n)
svy linearized: prop freq_escola_0a6, over (situação_n grupidade)
```

## APÊNDICE B: ESPECIFICAÇÃO DOS MODELOS HIERÁRQUICOS ESTIMADOS

### Modelo 1

$$y_{ij} = \beta_{0j} + e_{ij}$$

$$\beta_{0j} = \gamma_{00} + u_{0j}$$

$$y_{ij} = \gamma_{00} + u_{0j} + e_{ij}$$

### Modelo 2

$$y_{ij} = \beta_{0j} + \beta_{1j}início_{1j} + e_{ij}$$

$$\beta_{0j} = \gamma_{00} + u_{0j}$$

$$y_{ij} = \gamma_{00} + \beta_{1j}início_{1j} + u_{0j} + e_{ij}$$

### Modelo 3

$$y_{ij} = \beta_{0j} + \beta_{1j}início_{1j} + \beta_{2j}sexo + \beta_{3j}cor\_raça + \beta_{4j}repetência + \beta_{5j}dever + \beta_{6j}exp\_particular + \beta_{7j}trabalha + \beta_{8j}renda + \beta_{9j}esc\_mãe + \beta_{10j}esc\_pai + \beta_{11j}dicionário + \beta_{12j}densidade\_domiciliar + e_{ij}$$

$$\beta_{0j} = \gamma_{00} + \gamma_{01}rede + \gamma_{02}região + \gamma_{03}local + u_{0j}$$

$$y_{ij} = \gamma_{00} + \gamma_{01}rede + \gamma_{02}região + \gamma_{03}local + \beta_{1j}início_{1j} + \beta_{2j}sexo + \beta_{3j}cor\_raça + \beta_{4j}repetência + \beta_{5j}dever + \beta_{6j}exp\_particular + \beta_{7j}trabalha + \beta_{8j}renda + \beta_{9j}esc\_mãe + \beta_{10j}esc\_pai + \beta_{11j}dicionário + \beta_{12j}densidade\_domiciliar + u_{0j} + e_{ij}$$

### Modelo 4

$$Y_{ij} = \beta_{0j} + \beta_{1j}início_{ij} + e_{ij}$$

$$\beta_{0j} = \gamma_{00} + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + u_{1j}$$

$$Y_{ij} = \gamma_{00} + \gamma_{10}início_{ij} + u_{1j}início_{ij} + u_{0j} + e_{ij}$$

### Modelo 5

$$Y_{ij} = \beta_{0j} + \beta_{1j}início_{ij} + \beta_{2j}sexo + \beta_{3j}cor\_raça + \beta_{4j}repetência + \beta_{5j}dever + \beta_{6j}exp\_particular + \beta_{7j}trabalha + \beta_{8j}renda + \beta_{9j}esc\_mãe + \beta_{10j}esc\_pai + \beta_{11j}dicionário + \beta_{12j}densidade\_domiciliar + e_{ij}$$

$$\beta_{0j} = \gamma_{00} + \gamma_{01}rede + \gamma_{02}região + \gamma_{03}local + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}rede + \gamma_{12}região + \gamma_{13}local + u_{1j}$$

$$Y_{ij} = \gamma_{00} + \gamma_{01}rede + \gamma_{02}região + \gamma_{03}local + \gamma_{10}início + \gamma_{11}rede * início + \gamma_{12}região * início + \gamma_{13}local * início + \beta_{2j}sexo + \beta_{3j}cor\_raça + \beta_{4j}repetência + \beta_{5j}dever + \beta_{6j}exp\_particular + \beta_{7j}trabalha + \beta_{8j}renda + \beta_{9j}esc\_mãe + \beta_{10j}esc\_pai + \beta_{11j}dicionário + \beta_{12j}densidade\_domiciliar + u_{1j}início_{ij} + u_{0j} + e_{ij}$$

## APÊNDICE C – ROTINAS DE ESTIMAÇÃO DOS MODELOS HIERÁRQUICOS NO *STATA*: REESCALONAMENTO DE PESOS E INCLUSÃO DA AMOSTRAGEM COMPLEXA

### C.1 Reescalonamento dos pesos:

No *software* deverá haver as seguintes correspondências:

- UPA\_id = psu\_id
- UPA\_peso = psu\_wt
- UFA\_id = fsu\_id
- UFA\_peso = fsu\_wt

```
net install pwigls, from (http://www.cpc.unc.edu/restools/data\_analysis/ml\_sampling\_weights/) replace

gen psu_id=mascara
gen psu_wt=peso_tc

gen fsu_id=id_aluno_n
gen fsu_wt = (peso_ac/peso_tc)

pwigls, psu_id(psu_id) fsu_id(fsu_id)psu_wt(psu_wt) fsu_wt(fsu_wt)
*psu_m1wt(method_1_level_2_adjust) fsu_m1wt(method_1_level_1_weight)
*psu_m2wt(method_2_level_2_adjust) fsu_m2wt(method_2_level_1_weight)

rename psu_m1wt pw2
rename fsu_m1wt pw1
```

O programa *PWIGLS* computa dois pesos distintos, baseados em duas metodologias diferentes. A Metodologia 1 é adequada para planos amostrais informativos no nível 1 e não informativo no nível 2. Já a Metodologia 2 é adequada para plano amostrais informativos em ambos os níveis. Para a análise foi escolhido o Método 1. Para maiores informações sobre qual método é mais indicado recomenda-se Pfeffermann (2008).

## C.2 Estimações no *GLLAMM*: inclusão da amostragem complexa

A inclusão da amostragem complexa é feita através das seguintes opções de estimação do *GLLAMM*:

- *i(mascara)*: informa a variável que define a conglomeração (cluster)
- *pweight(pw)*: informa que os dados estão dispostos em níveis, sendo que a estimação procurará por *pw1* como a variável que determina o peso no nível 1 e *pw2* como a variável que determina o peso no nível 2.
- *robust cluster(estrato)*: Após a estimação do modelo a opção retorna os erros-padrão robustos considerando a estratificação

```

*****
**MODELO 1
*****

gllamm profic_2, i(mascara) link(id) family(gau) pweight(pw)

gllamm, robust cluster(estrato)

*****
**MODELO 2
*****

xi: gllamm profic_2 i.inicio_EI_D, i(mascara) ink(id) family(gau)
pweight(pw)

gllamm, robust cluster(estrato)

*****
**Modelo 3
*****

eq f1: rede_D região_D local_D

xi: gllamm profic_2 i.inicio_EI_D i.sexo_D i.cor_D i.repetencia_D
i.exp_part_D i.dever_D i.trabalha_D i.renda_D i.escolaridade_mãe_D
i.esc_pai_D i.dicionário_D i.dens_dom_D, i(mascara)link(id) family(gau)
pweight(pw) geqs(f1)

gllamm, robust cluster(estrato)

*****
** Modelo 4
*****

gen one=1

eq one: one

```

```
eq inicio: inicio_EI_D

xi: gllamm profic_2 i.inicio_EI_D, i(mascara) link(id) family(gau)
pweight(pw) nrf(2) eqs(one inicio)

gllamm, robust cluster(estrato)

*****
**Modelo 5
*****

gen one=1

eq one: one

eq inicio: inicio_EI_D

eq f1: rede_D região_D local_D

eq f2: rede_D região_D local_D

xi: gllamm profic_2 I.inicio_EI_D i.sexo_D i.cor_D i.repetencia_D
i.exp_part_D i.dever_D i.trabalha_D i.renda_D i.escolaridade_mãe_D
i.esc_pai_D i.dicionário_D i.dens_dom_D, i(mascara) link(id) family(gau)
pweight(pw) nrf(2) geqs(f1 f2) eqs(one inicio)

gllamm, robust cluster(estrato)
```