

**UNIVERSIDADE FEDERAL DE JUIZ DE FORA
FACULDADE DE ECONOMIA
MESTRADO EM ECONOMIA APLICADA**

**DOIS ENSAIOS SOBRE OS DETERMINANTES DA DESIGUALDADE
EDUCACIONAL BRASILEIRA A PARTIR DE DADOS
LONGITUDINAIS**

Igor Vieira Procópio

Orientador: Prof. Dr. Ricardo da Silva Freguglia

Coorientadora: Prof^a. Dr^a. Flávia Chein

**JUIZ DE FORA
FEVEREIRO/2012**

IGOR VIEIRA PROCÓPIO

**DOIS ENSAIOS SOBRE OS DETERMINANTES DA DESIGUALDADE
EDUCACIONAL BRASILEIRA A PARTIR DE DADOS
LONGITUDINAIS**

Dissertação elaborada pelo discente Igor Vieira Procópio como exigência do Curso de Mestrado em Economia Aplicada da Universidade Federal de Juiz de Fora, como requisito parcial à obtenção do Título de Mestre.

**JUIZ DE FORA
FEVEREIRO/2012**

Procópio, Igor Vieira.

Dois ensaios sobre os determinantes da desigualdade educacional brasileira a partir de dados longitudinais / Igor Vieira Procópio. – 2012. 100 f.

Dissertação (Mestrado em Economia Aplicada)–Universidade Federal de Juiz de Fora, Juiz de Fora, 2012.

1.Economia. 2. Desigualdade social - Educação. I. Título.

CDU 33

TERMO DE APROVAÇÃO

IGOR VIEIRA PROCÓPIO

DOIS ENSAIOS SOBRE OS DETERMINANTES DA DESIGUALDADE EDUCACIONAL BRASILEIRA A PARTIR DE DADOS LONGITUDINAIS

Dissertação elaborada pelo discente Igor Vieira Procópio como exigência do Curso de Mestrado em Economia Aplicada da Universidade Federal de Juiz de Fora, como requisito parcial à obtenção do Título de Mestre.

Orientador: Prof. Dr. Ricardo da Silva Freguglia
Universidade Federal de Juiz de Fora

Co-orientadora: Prof^a. Dr^a. Flávia Chein
Universidade Federal de Juiz de Fora

Prof. Dr. José Simão Filho
Universidade Federal de Juiz de Fora

Prof. Dr. Naércio Aquino Menezes Filho
Insper/Universidade de São Paulo

Juiz de Fora, 15 de fevereiro de 2012.

AGRADECIMENTOS

Agradeço em primeiro lugar aos meus pais, por sempre investirem e apoiarem meus estudos e pela dedicação e amor em todos os momentos da minha vida. À minha irmã pelo apoio, incentivo e pelo exemplo de dedicação aos estudos.

Ao meu orientador, professor Ricardo Freguglia, pela excelente orientação com ideias e críticas, passadas em inúmeras conversas formais e informais, e por sempre confiar no sucesso desta empreitada.

À professora Flávia Chein, minha coorientadora, que contribuiu de forma expressiva na elaboração deste trabalho, com críticas e sugestões decisivas.

Ao professor Lourival Batista, principal incentivador para que ingressasse no mestrado.

Aos membros da banca, professores Naércio Menezes-Filho e José Simão, pelas críticas e sugestões que contribuíram para enriquecer este trabalho.

Ao CNPQ e INEP, pelo apoio financeiro através do "Observatório da Educação-2010: Projeto 34 - Análise da Evolução da Educação Básica no Brasil a partir de indicadores de fluxo e proficiência".

Ao CAED/UFJF pela disponibilização do banco de dados GERES, essencial para a elaboração deste trabalho.

À UFJF pelo apoio financeiro através do PROQUALI.

Aos meus amigos, que me apoiaram, mesmo tendo tido o convívio dificultado nesses anos.

Aos meus colegas da turma do mestrado de 2010, que convivi e aprendi muito durante esses dois anos.

E para finalizar, agradeço a todos que contribuíram e torceram pelo sucesso na realização deste trabalho.

RESUMO

O objetivo desta dissertação é investigar os principais determinantes da desigualdade educacional brasileira, tendo em vista a relação entre nível educacional e sucesso individual no mercado de trabalho e em outros aspectos socioeconômicos. Utiliza-se um banco de dados proveniente do Projeto GERES – Geração Escolar 2005– onde uma amostra de crianças foi acompanhada durante os quatro primeiros anos do ensino fundamental. Partindo da Teoria do Ciclo de Vida da Formação de Habilidades (Cunha e Heckman, 2007) analisa-se o papel da escola e da pré-escola no nível de proficiência das crianças. Paralelamente, também é investigada a relação entre origem social e nível de proficiência, sob a ótica da Teoria de Igualdade de Oportunidades (Roemer, 1998). A dissertação consiste em dois ensaios organizados sob a forma de capítulos, além de um capítulo introdutório e de um capítulo de conclusão. O primeiro ensaio investiga o papel da escola no processo de geração/redução das desigualdades de habilidades entre os indivíduos, com base na hipótese de que a desigualdade de habilidades em que as crianças iniciam o ensino fundamental limita o papel da escola na formação de habilidades. Especificamente, o primeiro estudo objetiva (i) medir a evolução da desigualdade e do efeito das escolas sobre as oportunidades na formação de habilidades e (ii) encontrar evidências empíricas para o Brasil do processo de dinâmica do aprendizado. O primeiro ponto é realizado comparando a evolução do aprendizado para grupos com características socioeconômicas distintas e através da decomposição da desigualdade de proficiência com a utilização do índice de Theil-T. Já no segundo ponto, utilizam-se técnicas de painel dinâmico para as estimações. Os resultados encontrados indicam uma relação entre nível de habilidades e origem socioeconômica. Além disso, há evidências de um efeito dinâmico da aprendizagem, dado o impacto do nível de habilidades de períodos anteriores sobre o desempenho presente. O segundo ensaio estima o efeito da pré-escola sobre o nível de proficiência de alunos do primeiro ciclo do ensino fundamental. O efeito da pré-escola foi estimado através de uma função de produção educacional considerando o processo dinâmico de aprendizado, conforme Cunha e Heckman (2007) e Todd e Wolpin (2003). Os resultados encontrados indicam o efeito cumulativo dos insumos e a presença de heterogeneidade individual. Para Matemática o efeito direto encontrado foi positivo e significativo. Já para português, os coeficientes de pré-escola apresentaram resultados não significativos, indicando que não há efeito direto sobre as habilidades em português. O efeito indireto, manifestado através da proficiência inicial, foi positivo e crescente ao longo dos anos para Matemática. Por outro lado, os estudantes com maiores habilidades em linguagem no início do ensino fundamental aumentam suas diferenças comparativamente aos menos habilidosos, mas esse efeito se inverte ao longo dos anos, indicando que a desigualdade de habilidades em Português pode ser reduzida durante o período escolar. Considerando os resultados dos dois ensaios, conclui-se que parte da desigualdade educacional brasileira é em função da diferença dos insumos escolares recebidos pelas crianças. No entanto, as habilidades com que as crianças iniciam o período escolar limitam o papel da escola em reduzir as desigualdades de habilidades, sendo, portanto, necessários investimentos na primeira infância, como, por exemplo, uma política de ampliação das matrículas na pré-escola.

Palavras-chave: 1. Desigualdade de oportunidades; 2. Formação de habilidades; 3. Primeira infância; 4. Efeito escola; 5. Brasil.

Código JEL: I21, I24, J24, J13

ABSTRACT

The aim of this thesis is to investigate the most important determinants of the Brazilian education inequality, due to the relationship of educational attainment and individual social and economic success. The data set used comes from the GERES Project – Educational Generation 2005 – in which a sample of children was tracked during the first four years of elementary school. Following the Theory of the Life Cycle of Skills Formation (Cunha and Heckman, 2007) this study analyzes the role of school and preschool in the children's proficiency achievement. This study also investigates the relationship among social origin and children's proficiency achievement, based on the Theory of Equality of Opportunity (Roemer, 1998). This thesis is composed by two independent essays which are organized as chapters, besides an introductory chapter and a chapter of conclusion. The first essay investigates the schools' role in the generation/ reduction of inequalities of students' skills. The central assumption is that the initial level of skills inequality limits the effects of schools' intervention in skills formation. Specifically, this study intends to (i) measure the evolution of inequality and the effect of schools on the opportunities in skills formation and (ii) find empirical evidence for Brazil' dynamic process of learning. The first goal is performed by comparing the evolution of learning for groups with different socioeconomic characteristics and by the decomposition of inequality of proficiency using the Theil-T index-, while the second point applies techniques of dynamic panel, based on the model of technology of skill formation (Cunha and Heckman, 2007). The results indicate a positive relationship between skill level and socioeconomic background. They also show evidence of the dynamic effects of learning, highlighting the role of previous skill level. These findings shed some light on the limited role of the regular schooling reducing inequalities of skills and the importance of investments in early childhood education. The second essay estimates the preschool effect on the child's proficiency level of students from the first four years of elementary school. Following Cunha & Heckman (2007) and Todd & Wolpin (2003), the preschool effect is estimated through an educational production function that incorporates a dynamic process of learning. The empirical analysis shows evidence of cumulative effect of inputs and the presence of individual heterogeneity. In the case of Mathematics the direct effect of preschool on child's proficiency was positive and significant, while in Portuguese, there is no direct effect. Furthermore, the indirect effect shown by the coefficient of initial proficiency is positive and increases over the years for Math. On the other hand, the students with higher skills in language at the beginning of elementary school expand their differences in relation to less skilled students, but this effect inverts over the years, indicating that the inequality of skills in Portuguese can be reduced during the scholar period. Finally, the overall results point that the educational inequality is partially explained by the difference in the school resources. However, the initial level of children's skills limits the role of schools in reducing the inequality of abilities. Therefore, investing in the young children - like an increasing in preschool attendance - can contribute to reduce the educational inequality.

Key words: 1. Inequality of opportunities; 2. Skill formation; 3. Early childhood; 4. School effects; 5. Brazil.

JEL classification: I21, I24, J24, J13

LISTA DE FIGURAS

Gráfico 1.1 – Índice de Theil-T e taxa bruta de contribuição da educação	15
Gráfico 1.2 – IOH (2001 – 2009)	18
Gráfico 1.3 – IOH - Completar a 6ª série na Idade Correta	19
Gráfico 2.1 – Acesso à Escola por Faixa Etária	28
Gráfico 2.2 – Evolução da Proficiência Média dos alunos do GERES.....	44
Gráfico 2.3 – Evolução da Proficiência Média por Raça – Matemática	44
Gráfico 2.4 – Evolução da Proficiência Média por Raça – Português	44
Gráfico 2.5 – Evolução da Proficiência Média por Nível Socioeconômico – Matemática	45
Gráfico 2.6 – Evolução da Proficiência Média por Nível Socioeconômico – Português	45
Gráfico 2.7 – Evolução da Proficiência Média por Município – Matemática	45
Gráfico 2.8 – Evolução da Proficiência Média por Município – Português	45
Gráfico 2.9 – Evolução do Índice de Theil e Taxa Bruta – Matemática	46
Gráfico 2.10 – Evolução do Índice de Theil e Taxa Bruta – Português	46
Gráfico 2.11 – Evolução da Taxa Marginal – Matemática	47
Gráfico 2.12 – Evolução da Taxa Marginal – Português	47
Gráfico 2.13 – Lowess – Proficiência Final <i>versus</i> Proficiência Inicial – Matemática	47
Gráfico 2.14 – Lowess – Proficiência Final <i>versus</i> Proficiência Inicial – Português ..	47
Gráfico 3.1 – Proficiência Média – Brasil (2000-2009)	63
Gráfico 3.2 – Matrícula na Educação Infantil – Brasil (2001-2009)	65
Gráfico 3.3 – Proficiência em Matemática (2005-2008)	79
Gráfico 3.4 – Proficiência em Português (2005-2008)	79
Gráfico 3.5 – Escolaridade da Mãe e Matrícula na Pré-Escola	79

LISTA DE QUADROS

Quadro 2.1 - Taxa ótima de investimento de acordo com os parâmetros da tecnologia de formação de habilidades	36
Quadro 2.2 - Relação dos insumos escolares avaliados na literatura nacional	39
Quadro 2.3 - Relação da Variáveis utilizadas.....	40
Quadro 3.1 - Resumo dos principais resultados da literatura nacional	69
Quadro 3.2 - Relação da Variáveis utilizadas.....	70

LISTA DE TABELAS

Tabela 2.1 -	Exclusões da Amostra Geres e Atrito.....	38
Tabela 2.2 -	Estatísticas Básicas das Variáveis Invariantes no Tempo	42
Tabela 2.3 -	Estatísticas Básicas das Variáveis Variantes no Tempo	43
Tabela 2.4 -	MQO – Especificação Contemporânea	56
Tabela 2.5 -	Estimações de Painel Dinâmico	58
Tabela 3.1 -	Médias e Desvios Padrão	77
Tabela 3.2 -	MQO – Proficiência em Matemática	80
Tabela 3.3 -	Especificação Cumulativa – MQO sem Valor Adicionado – Proficiência em Matemática	81
Tabela 3.4 -	Especificação Cumulativa – MQO com Valor Adicionado – Proficiência em Matemática	82
Tabela 3.5 -	Especificação Cumulativa – Primeira Diferença sem Valor Adicionado – Proficiência em Matemática	83
Tabela 3.6 -	Especificação Cumulativa – Primeira Diferença com Valor Adicionado – Proficiência em Matemática	83
Tabela 3.7 -	MQO – Proficiência em Português	85
Tabela 3.8 -	Especificação Cumulativa – MQO sem Valor Adicionado – Proficiência em Português	86
Tabela 3.9 -	Especificação Cumulativa – MQO com Valor Adicionado – Proficiência em Português	86
Tabela 3.10 -	Especificação Cumulativa – Primeira Diferença sem Valor Adicionado – Proficiência em Português	87
Tabela 3.11 -	Especificação Cumulativa – Primeira Diferença com Valor Adicionado – Proficiência em Português	87
Tabela AI.1 -	Decomposição do Theil-T para grupos educacionais – PNAD’s 2001 a 2009	95
Tabela AII.1 -	IOH para o Brasil de 2001 a 2009 (PNAD’s)	100

SUMÁRIO

1 – INTRODUÇÃO.....	11
1.1 – Referências.....	21
2 – DESIGUALDADE DE OPORTUNIDADES E O PAPEL DA ESCOLA NA FORMAÇÃO DE HABILIDADES.....	25
2.1 – Introdução.....	26
2.2 – Modelo de Formação de Habilidades	29
2.3 – Base de Dados.....	37
2.4 – Estatísticas Descritivas.....	40
2.5 – Estratégia Empírica.....	48
2.6 – Resultados	52
2.7 – Considerações Finais.....	57
2.8 – Referências.....	59
3 – EFEITO DA PRÉ-ESCOLA SOBRE O NÍVEL DE PROFICIÊNCIA: Uma Análise a Partir de Dados Longitudinais.....	62
3.1 – Introdução.....	63
3.2 – Revisão da Literatura.....	66
3.3 – Banco de Dados e Variáveis	69
3.4 – Estratégia Empírica	71
3.5 – Estatísticas Descritivas.....	76
3.6 – Resultados.....	79
3.7 – Considerações Finais.....	88
3.8 – Referências.....	89
4 - CONCLUSÃO	91
ANEXO I	93
ANEXO II	96

1 – INTRODUÇÃO

A elevada desigualdade de renda brasileira é um fenômeno vastamente documentado na literatura econômica. Desde os anos de 1960 quando se iniciou o levantamento sistemático de informações censitárias no país, os índices de desigualdade encontrados se situam entre os mais altos do mundo (Fishlow, 1972; Langoni, 1973; Bonelli e Sedlacek, 1989; Hoffman, 1989; Lam e Levison, 1990; Ramos, 1993; Barros e Mendonça, 1996; Deininger e Squire, 1996; Ferreira e Litchfield, 1996 e 1999; Ferreira, 2000). Apesar da queda nos índices de desigualdade apresentados na última década, o Brasil continua figurando entre os países com maiores desigualdade de renda (Barros *et al.*, 2006).

O consenso existente sobre os altos níveis de desigualdade de renda no Brasil, contudo, não se refletiu em um consenso sobre os determinantes desta desigualdade e nem na necessidade ou não de políticas públicas para a sua redução. Como observam Barros *et al.* (2008), apesar da pobreza e desigualdade de renda serem conceitos relacionados entre si, o apelo para redução de ambos é baseado em argumentos distintos. Enquanto que a redução da pobreza recebe apoio universal e incondicional, sendo inclusive a primeira meta do desenvolvimento do milênio, a redução da desigualdade nunca obteve suporte incondicional. Segundo Ferreira (2000), o crescimento econômico seria a melhor alternativa para a redução da pobreza, e as políticas redistributivas poderiam interferir negativamente sobre o crescimento.

Barros *et al.* (2008) destacam que os argumentos que dão suporte para a redução da desigualdade variam em razão do nível e da natureza da desigualdade. Quando a desigualdade é extremamente alta e fortemente relacionada à origem familiar, gênero e etnia, o consenso sobre a necessidade da redução da desigualdade é mais fácil de ser alcançado. Entretanto, quando a desigualdade é resultado de diferentes escolhas, nível de esforço ou até mesmo de talento inato entre pessoas que tiveram acesso às mesmas oportunidades e recebido o mesmo tipo de tratamento, então o suporte para a realização de políticas compensatórias é mais fraco.

Ao analisar apenas a desigualdade de resultados observada, dois aspectos importantes para o debate não se revelam de imediato e merecem maior atenção. O primeiro aspecto se refere à necessidade de se entender em que período do ciclo de vida as desigualdades são geradas. Em relação ao segundo aspecto, deve-se distinguir entre fatores que são de responsabilidade do indivíduo, como esforço e escolhas autônomas, e fatores que os indivíduos não têm controle, como origem familiar, local de nascimento, gênero.

Sobre o primeiro aspecto, o período em que as desigualdades são geradas, diversos estudos empíricos têm apontado evidências de que fatores pré-mercado de trabalho são determinantes de parcela significativa da desigualdade de rendimentos. Em geral, consideram-se dois períodos, um período pré-mercado de trabalho, onde os indivíduos se formam para atuar no mercado de trabalho, e o segundo período, que é o próprio mercado de trabalho, onde os indivíduos competem por posições. O questionamento central que surge nessa periodização do ciclo de vida é se o mercado de trabalho é gerador das desigualdades, ou apenas revela e transforma uma desigualdade que foi gerada antes da entrada dos indivíduos no mercado de trabalho.

As habilidades geradas antes da entrada do indivíduo no mercado de trabalho determinam a produtividade do indivíduo e conseqüentemente seus rendimentos. Testes de proficiência realizados em crianças e adolescentes apresentam forte ligação com o sucesso futuro no mercado de trabalho (Leibowitz, 1974; Murnane, Willette e Levy, 1995; Neal e Johnson, 1996; Keane e Wolpin, 1997; Cameron e Heckman, 1998; Hanushek e Rivkin, 2006).

No Brasil, de uma forma geral, pode-se dizer que a partir dos anos de 1960, com o início de pesquisas censitárias sistemáticas, duas correntes, aparentemente antagônicas, dominaram o debate sobre a desigualdade de renda brasileira. De um lado o trabalho de Fishlow (1972) e do outro lado Langoni (1973). Enquanto que o primeiro entendia que o principal determinante da desigualdade de renda brasileira era o funcionamento do mercado de trabalho, o último defendia que as disparidades educacionais determinavam a maior parcela da desigualdade.

Barros e Mendonça (1996) desenvolveram um arcabouço teórico com o objetivo de integrar os dois argumentos. Com o uso de uma situação metafórica, os autores comparam o processo de geração e reprodução da desigualdade com uma sequência de corridas (trajetória de vida de uma geração). Cada corrida possui duas etapas. Na primeira etapa os participantes se preparam para a competição enquanto que na segunda etapa ocorre a competição. Em função da heterogeneidade dos indivíduos (habilidade inata) e dos recursos investidos (públicos e privados) na preparação, ao final da primeira etapa os indivíduos se encontram em diferentes níveis de preparação, ou seja, a competição se iniciará em desigualdade de condições. Ao final da segunda etapa cada indivíduo recebe um prêmio e a desigualdade nestes prêmios é chamada de desigualdade de resultados.

A primeira etapa é associada ao período da infância e adolescência onde ocorre a formação do capital humano do indivíduo. O nível de capital humano será diferente entre os indivíduos de acordo com suas diferenças inatas e dos recursos investidos¹. Portanto, esta etapa está fortemente associada à origem familiar e ao sistema educacional. A segunda etapa está associada ao mercado de trabalho. Basicamente, o arcabouço teórico desenvolvido por Barros e Mendonça (1996) permite separar os fatores que geram desigualdade dos fatores que revelam e/ou transformam a desigualdade e como estes fatores atuam em cada etapa. Apesar dos autores encontrarem resultados que indicam que tanto as diferenças nas condições educacionais quanto o funcionamento do mercado de trabalho geram desigualdade, a parcela atribuída às condições educacionais é mais relevante.

Ferreira (2000), a exemplo de Barros e Mendonça (1996) analisa os determinantes da desigualdade no Brasil construindo um modelo teórico que possibilita identificar quais fatores têm maiores impactos na geração da desigualdade de renda. O autor, de certa forma, faz uma extensão do modelo de Barros e Mendonça (1996) ao considerar cinco grupos de fatores que determinam a distribuição de renda. O primeiro seriam as características inatas dos indivíduos, tais como raça, gênero e inteligência, além da riqueza inicial. O segundo se refere a diferenças em relação a características individuais adquiridas, como nível educacional e experiência profissional. O terceiro grupo engloba os mecanismos do mercado de trabalho que podem atuar gerando, revelando e/ou transformando desigualdades. O quarto grupo de fatores refere-se aos mercados de capital, enquanto que o último grupo se refere, basicamente, à composição dos domicílios. Nas conclusões de seu trabalho o autor destaca que evidências empíricas sugerem que a educação é a variável de maior poder explicativo para a desigualdade de renda brasileira. A importância da educação na distribuição da renda do trabalho já era analisada em Becker(1974) que, através da formalização da teoria do capital humano, propiciou o surgimento de estudos empíricos em diversos países apontando para a contribuição da educação nos diferenciais observados de renda.

A importância da educação na desigualdade de rendimentos pode ser medida através da decomposição da desigualdade. Na sequência utiliza-se o índice de Theil-T para ilustrar o comportamento da contribuição da educação na desigualdade no Brasil durante a década de 2000. A utilização do índice de Theil se justifica devido à sua característica de permitir uma decomposição exata da desigualdade em efeitos entre e intra grupos, de modo que seja possível medir a contribuição de determinados grupos socioeconômicos na desigualdade total.

¹ Os recursos investidos podem ser divididos basicamente em três categorias: investimento da família, investimentos públicos e escolhas do indivíduo.

O índice de Theil-T é definido pela equação (1).

$$T = \sum_{i=1}^N \left[y_i \cdot \ln \left(y_i \cdot \frac{N}{Y} \right) \right] \quad (1)$$

onde y_i é o rendimento do indivíduo i , N é a quantidade total de pessoas e Y a soma do rendimentos de todos os indivíduos da população estudada.

Se a população for dividida em grupos de acordo com características em comum, é possível decompor o índice de Theil em uma parcela que pode ser atribuída à desigualdade de rendimento entre os grupos, Theil Entre, e a outra parcela que mensura a desigualdade dentro dos grupos, Theil Intra. A parcela referente à desigualdade entre os grupos é obtida de acordo com a equação (2)².

$$T_E = \sum_{g=1}^G \frac{Y_g}{Y} \cdot \ln \left(\frac{Y_g}{Y} \cdot \frac{N}{N_g} \right) \quad (2)$$

onde Y_g representa o total de rendimentos apropriado pelo grupo g , N_g o total da população dentro do grupo g e G indica o número de grupos. A taxa bruta de contribuição indica em termos percentuais quanto da desigualdade geral de renda pode ser atribuída à desigualdade entre os diferentes grupos da população. Para calcular a taxa bruta basta dividir o valor do Theil-Entre pelo Theil total, conforme equação (3).

$$\text{Taxa bruta} = \frac{T_E}{T} \quad (3)$$

Para a análise da contribuição da desigualdade educacional na desigualdade de renda, a população foi dividida em grupos de acordo com seus anos de estudo. Portanto, com esta análise é possível medir quanto da desigualdade total de renda pode ser atribuída à diferenças de rendimento entre os grupos educacionais.

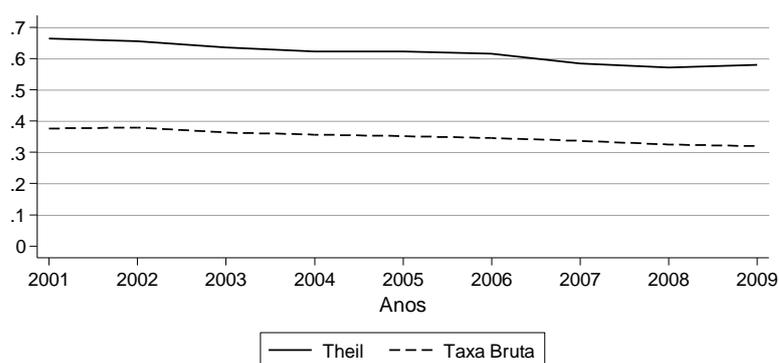
O gráfico 1.1 apresenta os valores para o índice Theil-T e para a taxa bruta de contribuição da educação para os anos de 2001 a 2009 utilizando dados da Pesquisa Nacional por Amostra de Domicílios – PNAD³.

² No corpo do texto só são apresentadas as fórmulas para a construção do Theil-Entre. No apêndice 1 se encontram todas as fórmulas para cálculo para o índice de Theil-T e sua decomposição, além da tabela com todos os valores usados para construir o gráfico.

³ Foi utilizada a variável rendimento mensal de todos os trabalhos da PNAD.

Os resultados mostram uma redução da desigualdade de rendimentos no período analisado (linha cheia), acompanhada de uma redução da contribuição das desigualdades educacionais (linha tracejada). Mas apesar dessa queda, a contribuição da educação na desigualdade de rendimentos não fica abaixo dos 30% em nenhum momento da série. Esses resultados vão ao encontro dos resultados dos trabalhos mencionados anteriormente. As desigualdades educacionais explicam grande parte da desigualdade salarial no Brasil.

Gráfico 1.1 - Índice de Theil e Taxa Bruta de Contribuição da Educação



Fonte: PNAD (2001-2009)

A segunda deficiência da análise direta da desigualdade de resultados observada é a não distinção dos determinantes da desigualdade entre fatores de responsabilidade individual e fatores em que o indivíduo não tem controle.

A Teoria de Igualdade de Oportunidades aborda explicitamente esta questão. O ideal desta teoria é que os fatores em que o indivíduo não tem controle não deveriam determinar seus resultados no mercado de trabalho. A teoria de igualdade de oportunidades tem sua origem na filosofia política, mais precisamente nas teorias sobre justiça distributiva, e se encontra no meio do caminho entre as teorias liberais, em que toda a responsabilidade sobre o sucesso no mercado de trabalho é atribuído ao indivíduo e, portanto, não deveria existir política pública de redistribuição de renda, e as teorias igualitárias, onde considera-se que o indivíduo não tem responsabilidade por seu resultado no mercado de trabalho, considerando a necessidade de políticas públicas redistributivas.

O debate sobre a responsabilidade individual nos resultados econômicos vem ganhando destaque nas discussões de políticas públicas voltadas para a redistribuição de renda. “A Theory of Justice” (1971) de John Rawls pode ser considerado o primeiro trabalho sobre justiça distributiva que separa os determinantes dos resultados em fatores de

responsabilidade do indivíduo e fatores fora do controle do indivíduo. Seguindo Rawls, destacam-se os trabalhos de Dworkin (1981a e 1981b), Arneson (1989 e 1990) e Cohen (1989). Entre os economistas, Sen (1979) em seu trabalho “Equality of What?” parte do pressuposto que toda teoria normativa sobre ordenamento social associa justiça à igualdade de algo. O que diferencia estas teorias é o que se deve igualar. Sen defende que o essencial é a formação das capacidades do indivíduo e que a atenção das políticas públicas deve ser na distribuição dos fatores determinantes destas capacidades. Roemer (1996 e 1998) apresenta uma definição formal do conceito de igualdade de oportunidades, estabelecendo de maneira abstrata a distinção entre os fatores sob responsabilidade do indivíduo e as circunstâncias fora do seu controle.

A igualdade de oportunidades pode ser analisada sob duas perspectivas (Fleurbaey e Peragine, 2009). A perspectiva *ex ante*, que busca avaliar a desigualdade de oportunidades no acesso a fatores determinantes do desenvolvimento individual (Van de Gaer, 1993; Kranich, 1996; Bourguignon *et al.*, 2003; Ferreira e Gignoux, 2008; Peragine e Serlenga, 2007; Banco Mundial, 2006; Barros *et al.*, 2008). E a perspectiva *ex post*, que analisa a desigualdade de resultados e busca separar entre os determinantes do resultado, fatores de responsabilidade do indivíduo e fatores fora do controle do indivíduo (Roemer, 1993 e 1998; Fleurbaey, 1995; Checchi e Peragine, 2009). A distinção entre as abordagens é fundamental para a formulação de políticas públicas. Seguindo a abordagem *ex ante*, as políticas públicas de igualdade de oportunidades devem garantir acesso aos fatores determinantes do desenvolvimento individual de forma equânime. Já de acordo com a perspectiva *ex post*, as políticas públicas devem analisar a desigualdade de resultados e identificar os fatores justos e os fatores injustos e redistribuir os resultados para que a desigualdade seja apenas em função de fatores justos.

Seguindo a perspectiva *ex ante*, Barros *et al.* (2008) desenvolveram um índice para medir a desigualdade de oportunidades nos serviços básicos para as crianças. De acordo com os autores, o índice é inspirado na função de bem estar social proposta por Sen (1976). Um processo de desenvolvimento no qual uma sociedade tem como objetivo suprir igualmente as oportunidades requer o acesso às oportunidades básicas pelas crianças como uma meta universal. O Índice de Oportunidades Humanas – IOH – é composto por dois elementos: (i) a taxa de cobertura de um serviço básico, que indica as oportunidades disponíveis; e (ii) pela desigualdade na distribuição destas oportunidades entre grupos, índice de dissimilaridade. Portanto, o índice é sensível tanto a um aumento na taxa de cobertura quanto a uma melhora da distribuição das oportunidades condicionada às circunstâncias.

O Índice de Oportunidades Humanas foca na taxa de cobertura e na desigualdade de acesso às oportunidades básicas para crianças em função de três fatores principais:

- i) de um ponto de vista empírico, o princípio da igualdade de oportunidades como “nivelar o campo de jogo” pode ser operacionalizado pela mensuração do acesso, para as crianças, aos bens e serviços básicos que são críticos para o pleno desenvolvimento infantil. Para crianças, o acesso define a oportunidade, pois não se pode cobrar das crianças que elas realizem esforços próprios para ter acesso a estes bens e serviços (diferentemente dos adultos);
- ii) de um ponto de vista de políticas públicas, evidências apontam que intervenções para equalizar as oportunidades no início do ciclo de vida do indivíduo são mais eficiente e menos custosas do que intervenções em um período mais avançado;
- iii) focar nas crianças contribui para colocar a discussão sobre desigualdade de oportunidades no centro do debate sobre políticas públicas. Como destacado no Relatório do Desenvolvimento Mundial 2006 (Banco Mundial 2006), no dia do nascimento as crianças não podem ser consideradas responsáveis pelas circunstâncias de sua família, apesar do fato destas circunstâncias afetarem a vida que estas crianças terão.

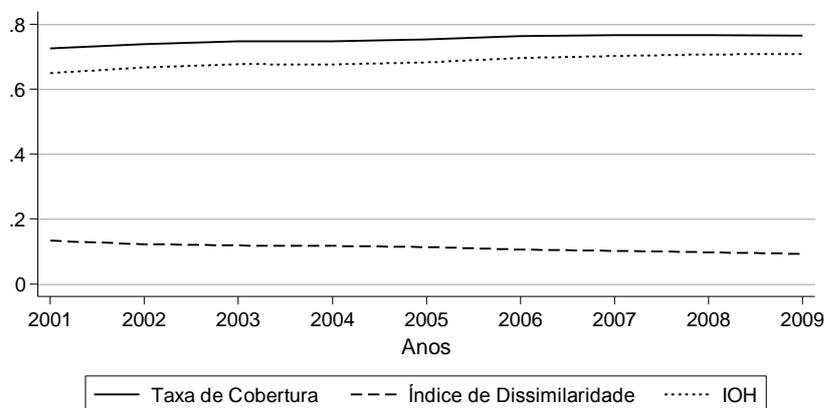
Barros *et al.* (2008) utilizam a metodologia desenvolvida para o IOH para a análise do acesso de crianças até dezesseis anos às seguintes oportunidades básicas: completar a sexta série na idade correta, atendimento escolar nas idades de 10 a 14 anos, acesso a água tratada, saneamento e eletricidade no domicílio. Para caracterizar as crianças em grupos de acordo com suas características circunstanciais, os autores utilizaram as seguintes variáveis: gênero, educação dos pais, renda per capita, número de irmãos, presença dos pais no domicílio e área de residência (rural/urbano).

O gráfico 1.2 apresenta os resultados para o IOH, composto pelas cinco oportunidades básicas definidas pelos autores, para os anos de 2001 a 2009 utilizando dados da PNAD⁴. São apresentados os valores da taxa de cobertura, do índice de dissimilaridade e do IOH. Para a taxa de cobertura e IOH quanto mais próximo de um se encontrar o valor melhor, enquanto que para o índice de dissimilaridade quanto mais próximo de zero melhor. Nota-se uma evolução favorável nos indicadores. A taxa de cobertura cresce até meados da década e depois se estabiliza, mas o IOH continua tendo uma evolução em função da melhora na distribuição

⁴ No apêndice 2 encontram-se a metodologia de construção do índice e tabelas contendo todos os resultados.

das oportunidades evidenciada pela queda no índice de dissimilaridade. O IOH começa em torno de 0,65 e cresce até 0,71.

Gráfico 1.2 - IOH (2001-2009)



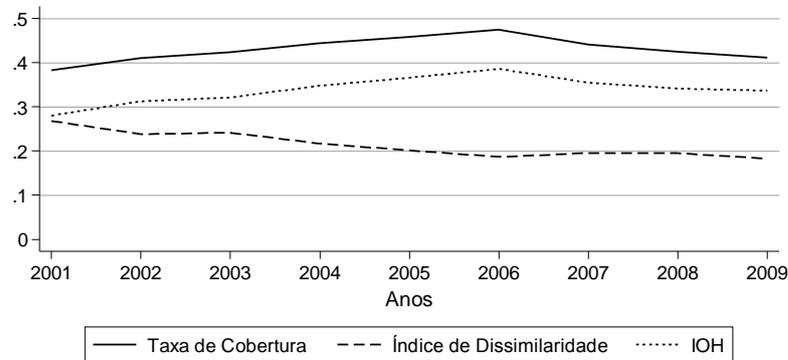
Fonte: PNAD (2001-2009)

Das cinco oportunidades básicas destacadas em Barros *et al.* (2008), duas se referem a oportunidades educacionais, sendo uma referente a acesso ao sistema educacional, atendimento escolar nas idades de 10 a 14 anos, e uma mais correlacionada com qualidade da educação, completar a sexta série na idade correta. Focando na qualidade da educação, os resultados são mais preocupantes do que os resultados apresentados no gráfico 1.2. O gráfico 1.3 apresenta os resultados apenas para os índices referentes à probabilidade de completar a sexta série na idade correta. O IOH para esta oportunidade não passa de 0,5 em nenhum momento da série. Apesar de uma elevação do índice até 2006, após esta data a trajetória se reverte e o índice tem uma queda, chegando em 2009 em valor próximo ao de 2002. O comportamento da evolução do IOH está muito colado na evolução da taxa de cobertura, que também apresenta uma queda a partir de 2006. O índice de dissimilaridade, que representa a desigualdade de oportunidades decresce quase que monotonicamente no período, indicando que apesar da redução geral das oportunidades, a desigualdade se reduziu.

Considerando os argumentos e as evidências sobre a importância da educação na desigualdade de renda, a literatura sobre Economia da Educação ganha relevância ao estudar os determinantes da formação das habilidades individuais. Entre estes determinantes, destacam-se as características familiares, a habilidade genética e o papel da escola. Em relação aos dois primeiros fatores existe um consenso sobre sua importância, no entanto, a

importância do papel da escola, e, principalmente que características da escola contribuem para um melhor desenvolvimento, ainda é alvo de muita discussão na literatura.

Gráfico 1.3 - IOH - Completar a 6ª série na idade correta (2001-2009)



Fonte: PNAD (2001-2009)

A tradição de se avaliar o papel da escola no aprendizado tem como trabalho seminal o relatório Coleman (1966). Neste trabalho foram avaliadas escolas públicas e privadas dos Estados Unidos. Os resultados apresentados geraram um pessimismo sobre o papel da escola no desenvolvimento da criança, pois não foram encontrados resultados significativos em relação às características das escolas no aprendizado. Após este trabalho, diversos outros estudos foram realizados e os resultados encontrados variam, não sendo possível concluir sobre a importância das características escolares. Segundo Todd e Wolpin (2003), a principal explicação para os diferentes resultados encontrados na literatura é a diferença nos métodos utilizados.

Os estudos sobre os determinantes da formação de habilidades amparados pela tradicional teoria do capital humano consideram apenas um período na formação de habilidades individuais. Em linhas gerais, a tradicional teoria do capital humano não se preocupa com o momento em que os insumos escolares são providos, o que importa é a quantidade de insumos, medidos através dos anos de escolaridade. Cunha *et al.* (2006) desenvolvem um modelo que considera que existe um ciclo de vida na formação de capital humano.

De acordo com este modelo, algumas habilidades tem um período determinado para serem formadas. A ausência de investimentos neste período pode prejudicar todo o desenvolvimento futuro da criança. Se o período for um período crítico, a ausência de

investimento neste período inviabiliza futuros investimentos compensatórios. Se o período for considerado um período sensível, é possível remediar este baixo investimento, no entanto, investimentos posteriores ao período sensível terão baixo retorno⁵. Portanto, de acordo com os autores, existe um *trade-off* entre eficiência e equidade em investimentos para adolescentes, mas não para crianças mais novas.

Considerando estes argumentos, políticas voltadas para a redução das desigualdades de habilidades não podem ser direcionadas apenas para a melhoria do sistema educacional regular. A desigualdade com que as crianças ingressam na escola só poderia ser compensada pelas escolas em detrimento da eficiência. Neste sentido, o objetivo desta dissertação é avaliar o papel da escola no processo de geração/redução das desigualdades e identificar o efeito de intervenções nas crianças antes da entrada no ensino fundamental, mais especificamente o efeito da pré-escola.

Para atingir o objetivo geral é utilizado um banco de dados ainda pouco explorado pela literatura. O banco faz parte do projeto Geração Escolar 2005 – Geres. O Geres é um projeto que foi desenvolvido em parceria pelas Universidade Federal de Minas Gerais (UFMG), Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio), Universidade Estadual do Mato Grosso do Sul (UEMS), Universidade Federal da Bahia (UFBa), Universidade Estadual de Campinas (UNICAMP) e Universidade Federal de Juiz de Fora (UFJF).

O Geres acompanhou alunos de 2005 a 2008 de escolas públicas e privadas de cinco cidades brasileiras, Belo Horizonte (MG), Rio de Janeiro (RJ), Campo Grande (MS), Salvador (BA) e Campinas (SP), que em 2005 estavam na primeira série do Ensino Fundamental. Foram realizados testes anuais de proficiência e aplicados questionários para avaliar as características das escolas, professores, diretores, pais e alunos. Portanto, o banco de dados utilizado na dissertação é um banco longitudinal que acompanha as crianças durante o primeiro ciclo do Ensino Fundamental (1^a a 4^a série).

Esta característica longitudinal dos dados permite avaliar a evolução do aprendizado bem como estimar o efeito da escola e pré-escola com hipóteses menos restritivas do que as aplicadas para bancos em corte transversal.

A dissertação está estruturada na forma de ensaios. No primeiro ensaio é avaliado o papel da escola. Através de técnicas de painel dinâmico avalia-se o efeito do nível de habilidade passado da criança e de determinadas características da escola no aprendizado, bem como o efeito das características individuais e familiares. Também é avaliada a

⁵ Para maiores detalhes sobre os períodos críticos e sensíveis, ver Cunha e Heckman (2007 e 2009).

desigualdade de oportunidades no aprendizado através da análise do efeito das características dos pais e do acesso às escolas. Como principais resultados destaca-se a dependência do aprendizado durante o período escolar aos níveis de habilidades com que as crianças iniciam a escola, o que limita o papel da escola da redução das desigualdades de habilidades; destaca-se também, a relação entre níveis de habilidades e origem familiar, indicando desigualdade de oportunidades na formação de habilidades.

O segundo ensaio avalia o efeito da pré-escola sobre o nível de proficiência futuro das crianças. A estimação é realizada através de uma técnica mista entre a estratégia de valor adicionado com o estimador de primeira diferença. Os resultados encontrados indicam que as crianças que fizeram pré-escola possuem um maior aprendizado, em relação às crianças que não fizeram pré-escola, para a proficiência em matemática. Já para português, os coeficientes de pré-escola apresentaram resultados não significativos, indicando que não há efeito direto sobre as habilidades em português. A dissertação é finalizada com as conclusões gerais no capítulo quatro.

1.1– Referências

ARNESON, R. Equality of Opportunity for Welfare. **Journal of Political Philosophy**, v.7, n.4, 1999.

ARNESON R. J. Liberalism, Distributive Subjectivism, and Equal Opportunity for Welfare. **Philosophy and Public Affairs**, v. 19, p.158-194, 1990.

BARROS, R. P.; MENDONÇA, R.S.P. Os Determinantes da Desigualdade no Brasil. **Economia brasileira em perspectiva**. Rio de Janeiro: IPEA, p.421-474, 1996.

BARROS, R. P. *et al.* Consequências e Causas Imediatas da Queda Recente na Desigualdade de Renda Brasileira. In: **PARCERIAS ESTRATÉGICAS**. Análise sobre a pesquisa nacional por amostra de domicílios (Pnad 2004). Brasília: Centro de Gestão e Estudos Estratégicos, n. 22, p. 89-119. Edição especial. 2006

BECKER, G. S. **Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education**. University of Chicago Press: Chicago, IL., 1974.

BONELLI, R E SEDLACEK, G. L. Distribuição de Renda: Evolução no Último Quarto de Século. In Sedlacek, G. L. e Barros, R. P. **Mercado de Trabalho e Distribuição de Renda: Uma Coletânea**, (Rio de Janeiro: IPEA; Série Monográfica #35), 1989.

BOURGUIGNON, F., FERREIRA, F. E MENENDEZ, M. Inequality of Out-Comes and Inequality of Opportunities in Brazil. Policy Research Working Paper Series 3174, The World Bank, Washington DC, 2003

CAMERON, S., E HECKMAN. J.J. Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts. **Journal of Political Economy**, v. 106, n. 2, p. 262–333, 1998.

CHECCHI, D. E PERAGINE, V. (2009). Inequality of Opportunity in Italy. **Journal of Economic Inequality**, v.8, n.4, p.429-450, 2010.

COHEN, G. A. On the Currency of Egalitarian Justice. **Ethics**, v. 99, p. 906-944, 1989.

COLEMAN, J. S.; CAMPBELL, E. Q.; HOBSON, C. J.; MCPARTLAND, J.; MOOD, A. M.; WEINFELD, F. D. Equality of educational opportunity. Washington, D.C.: National Center for Educational Statistics, 1966.

CUNHA, F., HECKMAN, J. J., LOCHNER, L. J. E MASTEROV, D. V. Interpreting the Evidence on Life Cycle Skill Formation. In **Handbook of the Economics of Education**, capítulo. 12, ed. Eric A. Hanushek e Frank Welch, p. 697–812. Amsterdam: North-Holland, 2006.

DEININGER, K. E SQUIRE, L. A New Data Set Measuring Income Inequality. **World Bank Economic Review**, v. 10, p. 565-591, 1996.

DWORKIN, R. What Is Equality? Part 1: Equality of Welfare. **Philosophy & Public Affairs**, v.10, p. 185–246, 1981a.

DWORKIN, R. What Is Equality? Part 2: Equality of Resources. **Philosophy & Public Affairs**, v.10, p. 283-345, 1981b

FERREIRA, F. H. G. Os Determinantes da Desigualdade de Renda no Brasil: Luta de Classes ou Heterogeneidade Educacional?. Texto para Discussão n. 415, Departamento de Economia PUC-Rio, 2000.

FERREIRA, F. H. G.; J. A. LITCHFIELD. Growing Apart: Inequality and Poverty Trends in Brazil in the 1980s. LSE – STICERD – DARP Discussion Paper No.23, London, 1996.

FERREIRA F.H.G.; LITCHFIELD, J.A. Educacion o Inflacion?: Explicando igualdad e n Brasil en 22 Decada de los Ochenta. Ch.4 in **Cardenas and Lustig**, Pobreza y igualdad e n America Latina, (Bogota: Tercer Mundo Editores), 1999.

FERREIRA, F. H. G. E GIGNOUX, J. The Measurement of Inequality of Opportunity: Theory and an Application to Latin America. The World Bank Policy Research WP # 4659, 2008.

FISHLOW, A. Brazilian Size Distribution of Income. **American Economic Review**, American Economic Association, v. 62, n. 2, p. 391-402, 1972.

FLEURBAEY, M. Three Solutions for the Compensation Problem. **Journal of Economic Theory**, v. 65, p. 505–521, 1995.

FLAURBAEY, M. E PERAGINE, V. Ex ante versus Ex post Equality of Opportunity. Mimeo, 2009.

HANUSHEK, E. A., AND S. G. RIVKIN. School Quality and the Black-White Achievement Gap. Working Paper no. 12651, NBER, Cambridge, MA, 2006

HOFFMAN, R. Evolução da Distribuição da Renda no Brasil, Entre Pessoas e Entre Famílias, 1979/86, in Sedlacek, G. L. and R. Paes de Barros, **Mercado de Trabalho e Distribuição de Renda: Uma Coletânea**, (Rio de Janeiro: IPEA; Série Monográfica #35), 1989

KEANE, M. P., AND K. I. WOLPIN. Career Decisions of Young Men. **Journal of Political Economy**, v. 105, p. 473–522. 1997

KRANICH, L. Equitable Opportunities: an Axiomatic Approach. **Journal of Economic Theory**, v. 71, p. 131-147, 1996.

LAM, D. E LEVISON, D. Idade, experiência, escolaridade e diferenciais de renda: Estados Unidos e Brasil. **Pesquisa e Planejamento Econômico**, v. 20, n.2, 1990.

LANGONI, C. G. **Distribuição da Renda e Desenvolvimento Econômico do Brasil**. Rio de Janeiro, Editora Expressão e Cultura, 1973

LEIBOWITZ, A. Home Investments in Children. **Journal of Political Economy**, v. 82, n. 2, 1974.

MURNANE, R. J.; J. B. WILLETT,; F. LEVY. The Growing Importance of Cognitive Skills in Wage Determination. **The Review of Economics and Statistics**, v. 77, n. 2, p.251-266, 1995

NEAL, D.; W. JOHNSON. The Role of Pre-market Factors in Black-White Wage Differences. **Journal of Political Economy**, v.104, p.869–895, 1996.

PERAGINE, V. The Distribution and Redistribution of Opportunity. **Journal of Economic Surveys**, v.13, p. 37-69, 1999.

PERAGINE, V.; SERLENGA, L. **Higher Education and Equality of Opportunity in Italy**. University of Bari, v.6, 2007

RAMOS, L. **A Distribuição de Rendimentos no Brasil: 1976/85**. IPEA: Rio de Janeiro, 1993.

RAWLS, J. **A Theory of Justice**. Cambridge, MA: Harvard University Press, 1971.

ROEMER, J. E. A Pragmatic Theory of Responsibility for the Egalitarian Planner. **Philosophy and Public Affairs**, v.22, p.146–166. 1993.

ROEMER, J. E. **Theories of Distributive Justice**. Harvard University Press, Cambridge, 1996

ROEMER, J. E. **Equality of Opportunity**. Harvard University Press, Cambridge. 1998.

SEN, A. Poverty: An Ordinal Approach to Measurement. **Econometrica**, v. 44, p. 219-232, 1976.

SEN, A. Equality of what? The Tanner Lecture on Human Values. Stanford University, 1979.

VAN DE GAER, D. Equality of Opportunity and Investment in Human Capital. Ph.D. dissertação, Catholic University of Louvain, 1993

WORLD BANK. **World Development Report 2006: Equity and Development**. The World Bank and Oxford University Press, Washington, DC, 2006

2 - DESIGUALDADE DE OPORTUNIDADES E O PAPEL DA ESCOLA NA FORMAÇÃO DE HABILIDADES

RESUMO

Este capítulo investiga o papel da escola no processo de geração/redução das desigualdades de habilidades entre os indivíduos, com base na hipótese de que a desigualdade de habilidades em que as crianças iniciam o ensino fundamental limita o papel da escola na formação de habilidades. É utilizado um banco de dados longitudinal fornecido pelo projeto Geres, onde uma amostra de crianças é acompanhada durante os quatro primeiros anos do Ensino Fundamental. Especificamente, o presente estudo objetiva (i) medir a evolução da desigualdade e do efeito das escolas sobre as oportunidades na formação de habilidades e (ii) encontrar evidência empírica para o Brasil do processo de dinâmica do aprendizado. O primeiro ponto é realizado comparando a evolução do aprendizado para grupos com características socioeconômicas distintas e através da decomposição da desigualdade de proficiência com a utilização do índice de Theil-T. Já no segundo ponto, utiliza-se o modelo da tecnologia de formação de habilidades (Cunha e Heckman, 2007) como base teórica e técnicas de painel dinâmico para as estimações. Os resultados encontrados indicam uma relação entre nível de habilidades e origem socioeconômica. Além disso, há evidências de um efeito dinâmico da aprendizagem, dado o impacto do nível de habilidades de períodos anteriores sobre o desempenho presente. Este resultado alerta para o papel limitado da escola regular na redução das desigualdades de habilidades e reforça a importância de investimentos na educação infantil.

Palavras chave: Desigualdade de Oportunidades; Formação de Habilidades; Efeito Escola.

Código JEL: I21, I24

ABSTRACT

This paper investigates the schools' role in the generation / reduction of inequalities of students' skills. The central assumption is that the initial level of skills inequality limits the effects of schools' intervention in skills formation. The data come from a longitudinal database provided by the project Geres, where a sample of children is monitored during the first 4 years of elementary school. Specifically, this study aims to (i) measure the evolution of inequality and the effect of schools on the opportunities in skills formation and (ii) find empirical evidence for Brazil' dynamic process of learning. The first point is performed by comparing the evolution of learning for groups with different socioeconomic characteristics and by the decomposition of inequality of proficiency using the Theil-T. While the second point uses the model of technology of skill formation (Cunha and Heckman, 2007) as a theoretical basis and techniques for dynamic panel. The results indicate a relationship between skill level and socioeconomic background and also show evidence of the dynamic effects of learning, highlighting the role of the skill level of prior periods. These findings shed some light on the limited role of the regular school in reducing inequalities of skills and reinforce the importance of investments in early childhood education.

Key words: Inequality of Opportunities; Skills Formation; School Effect.

2.1 – Introdução

O acesso à educação é considerado um dos aspectos fundamentais para o desenvolvimento individual e de uma nação. Indivíduos mais escolarizados têm, por exemplo, melhores oportunidades no mercado de trabalho, menor probabilidade de ficar desempregado e menor taxa de participação em crimes. Um país com uma população com alto nível de escolaridade apresenta maiores taxas de crescimento econômico, melhores taxas de inovação e facilidade de adaptação em relação a inovações introduzidas por outros países. Diversos estudos empíricos têm apontado evidências de que fatores pré-mercado de trabalho são determinantes de parcela significativa da desigualdade de rendimentos. As habilidades geradas antes da entrada do indivíduo no mercado de trabalho determinam a produtividade do indivíduo. Testes de proficiência realizados em crianças e adolescentes apresentam forte ligação com o sucesso futuro no mercado (Leibowitz, 1974; Murnane, Willett, e Levy, 1995; Neal e Johnson, 1996; Keane e Wolpin, 1997; Cameron e Heckman, 1998; Hanushek e Rivkin, 2006; Cunha e Heckman, 2007, 2009).

No Brasil destacam-se o trabalho seminal de Langoni (1973), e os trabalhos de Barros e Mendonça (1996) e Ferreira (2000), apontando a desigualdade educacional como o principal determinante da desigualdade salarial brasileira. Esses estudos trazem evidências de que a desigualdade de salários apresentada no mercado de trabalho é um reflexo da diferença de produtividade dos trabalhadores, e que esta diferença não é gerada no mercado de trabalho. Barros e Mendonça (1996) desenvolvem um modelo dividindo o ciclo de vida do trabalhador em dois períodos, um período pré-mercado de trabalho, onde as habilidades são formadas, e um segundo período correspondendo ao mercado de trabalho, onde as habilidades são transformadas em produtividade e remuneradas através dos salários. Os autores concluem que a maior parcela da desigualdade é gerada no primeiro período do ciclo de vida, ou seja, na fase de formação das habilidades. Portanto, os trabalhadores entram no mercado de trabalho em desigualdade de condições. A divisão proposta por Barros e Mendonça (1996) contribui para o entendimento da geração de desigualdades salariais, evidenciando o papel da desigualdade de habilidades.

Constata-se, pois, a relevância em se estudar políticas educacionais a partir da análise do nível de aprendizado das crianças. O objetivo deste capítulo é investigar a dependência do nível de habilidades das crianças dentro da escola em relação às habilidades passadas e aos insumos escolares recebidos. Para isso, é utilizado um banco de dados pouco explorado na literatura, fornecido pelo projeto Geração escolar 2005 – GERES. Este é um banco

longitudinal que acompanha crianças da 1ª a 4ª séries do ensino fundamental nos municípios de Belo Horizonte (MG), Rio de Janeiro (RJ), Campinas (SP), Campo Grande (MT) e Salvador (BA). Considerando a dinâmica do aprendizado e o caráter longitudinal dos dados, são utilizadas técnicas de painel dinâmico para a estimação dos parâmetros da função de produção educacional.

Os principais resultados encontrados apontam para uma desigualdade de oportunidades educacionais, evidenciando uma forte relação entre níveis de habilidades das crianças e sua origem social. Em relação à dinâmica do aprendizado, os resultados das estimações mostram a importância dos níveis de habilidade passados na formação das habilidades presentes.

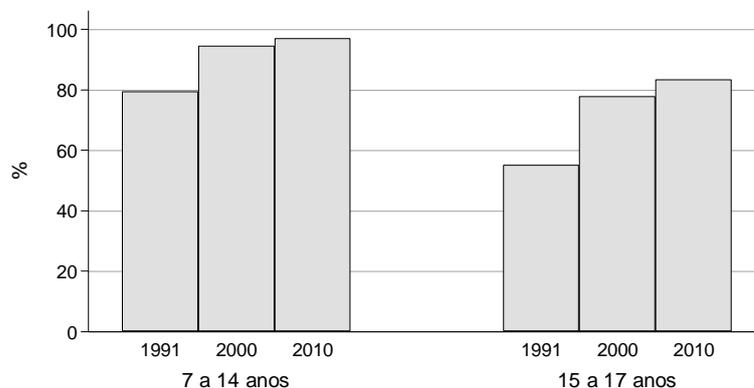
Deve-se enfatizar que o presente capítulo parte da premissa de que estudar os determinantes da formação de habilidades é essencial para a formulação de políticas públicas de redução da desigualdade de rendimentos. Os principais determinantes para a formação de habilidades são os fatores genéticos, as características familiares e os insumos escolares (Todd e Wolpin, 2003).

A eficácia dos insumos escolares é alvo de grande debate na literatura e grande divergência de opiniões e resultados encontrados nos trabalhos empíricos. O papel da escola na promoção da igualdade educacional tem como primeiro estudo o Relatório Coleman de 1966, que investigou se o desempenho dos alunos norte americanos estava relacionado com características da escola. Os resultados deste estudo apontaram para a inexistência desta relação e que o principal determinante do aprendizado das crianças estava relacionado com o *background* familiar. No entanto, após este relatório, diversos estudos foram sendo desenvolvidos e trazendo evidências contrárias, ou seja, que a escola faz sim diferença (ver Soares e Brooke, 2008).

O acesso à escola no Brasil vem crescendo em todos os níveis de ensino. Apesar de na prática o acesso ao ensino fundamental ainda não ter atingido os 100% das crianças de 7 a 14 anos, existe a meta de universalização e garantia legal de acesso a todos que buscarem a matrícula. Nos outros níveis também vem ocorrendo um aumento no número de matrículas. O gráfico 1 exemplifica esta evolução com alguns dados. São apresentados os números da evolução do percentual de crianças de 7 a 14 anos na escola e dos adolescentes de 15 a 17 anos no período de 1991 a 2010. Para as crianças de 7 a 14 anos, nota-se que em 1991 praticamente 80% frequentavam a escola e que este percentual se aproxima dos 100% em 2010. Já para os adolescentes de 15 a 17 anos, apesar do percentual de matrícula ser inferior, a

evolução no período foi mais acentuada, passando de aproximadamente 55% em 1991 para um pouco mais de 83% em 2010.

Gráfico 2.1 - Acesso à Escola por Faixa Etária



Fonte: Censo IBGE (1991, 2000 e 2010)

No entanto, apesar do inegável avanço com o aumento nas taxas de matrícula em todos os níveis de ensino no Brasil, a qualidade da educação no Brasil ainda está longe dos países desenvolvidos. Resultados do Programa Internacional de Avaliação de Alunos – Pisa – mostram que o Brasil, apesar de ser um dos países que mais avançaram entre 2000 e 2009, ainda se encontra nas últimas posições no *ranking* dos países avaliados. Em 2009, o Brasil ocupava a 53ª posição em uma lista com 65 países.

O problema da qualidade da educação deve ser analisado sob dois aspectos, os valores médios da proficiência das crianças brasileiras são baixos na comparação internacional, e existe uma desigualdade no aprendizado relacionado às características familiares e regionais. À parcela da desigualdade relacionada com o ambiente em que a criança se desenvolve, dá-se o nome de desigualdade de oportunidades na formação de habilidades, que pode ser entendida como uma desigualdade de acesso aos fatores determinantes da formação de habilidades em função de características de origem familiar e local de residência. A Teoria da Igualdade de Oportunidades (Roemer, 1998) tem como ideal que os fatores fora do controle dos indivíduos não sejam determinantes de seu sucesso socioeconômico, por consequência, a desigualdade seria resultado de diferentes escolhas e nível de esforço individual.

Para se estudar a formação de habilidades, este capítulo é baseado no modelo desenvolvido por Cunha e Heckman (2007). Neste modelo, define-se a formação de habilidades através de um processo dinâmico, em que o nível de habilidades adquiridos em uma etapa serve como insumo para a formação de habilidades do período subsequente. Cunha

e Heckman (2007 e 2009), Cunha *et al.* (2006) argumentam que desigualdades de habilidades entre as pessoas se iniciam logo no início do ciclo de vida e persistem. Crianças que recebem baixos investimentos no início da vida apresentam dificuldades de aprendizado em etapas mais avançadas do ciclo de vida.

Logo, se as habilidades desenvolvidas no período escolar forem dependentes de habilidades passadas, ou seja, do nível de habilidades em que as crianças ingressam na escola, a escola não será capaz de diminuir o *gap* entre os alunos que iniciam com níveis de habilidades mais elevados para aqueles que iniciam com níveis mais baixos. Desta forma, as escolas, se fossem todas homogêneas em suas características que determinam o aprendizado, iriam revelar e perpetuar uma desigualdade gerada em um período pré-escolar. No entanto, as escolas não são homogêneas, existindo escolas que contribuem mais para o aprendizado das crianças do que outras. Se houver uma seleção não aleatória das escolas em relação às crianças de melhores níveis de habilidades, a escola poderá contribuir para um aumento da desigualdade de habilidades nas crianças caso as crianças mais habilidosas forem selecionadas para as escolas de melhor qualidade.

Além dessa seção introdutória, o capítulo está estruturado em mais cinco seções. A segunda seção apresenta o modelo de formação de habilidades e a estratégia adotada para a estimação da função de produção educacional. A seção três traz uma descrição do banco de dados utilizado. A quarta seção apresenta as estatísticas básicas e alguns resultados preliminares, a seção cinco os resultados econométricos. A sexta seção conclui com as considerações finais.

2.2 - Modelo de Formação de Habilidades

Esta seção apresenta um modelo para a estimação da função de produção educacional com base na teoria do ciclo de vida de formação de habilidades, desenvolvida em Cunha e Heckman (2007). A principal característica deste modelo é incorporar que o aprendizado é um processo dinâmico, onde o nível de conhecimento em uma etapa serve como insumo para a formação de habilidades da etapa posterior. Os autores desenvolvem um modelo econômico para explicar as evidências encontradas na literatura de psicologia, educação e neurociência sobre o desenvolvimento de crianças.

O modelo desenvolvido permite explicar alguns fatos evidenciados na recente literatura empírica; i) as diferenças de habilidades entre indivíduos e entre grupos socioeconômicos se iniciam logo no início do ciclo de vida e persiste; ii) existe evidência de

períodos críticos e sensíveis no desenvolvimento da criança; iii) apesar dos baixos retornos de intervenções voltadas para adolescentes em situação de desvantagem, a literatura empírica mostra altos retornos para investimentos compensatórios em crianças novas em situação de desvantagem.

A tecnologia da formação de habilidades é representada por uma função de produção que determina a relação entre os insumos e o nível de habilidades. A formação de habilidades é um processo dinâmico, onde o nível de habilidades de uma etapa serve como insumo para a formação de habilidades da próxima etapa. Os autores destacam a existência de períodos sensíveis e críticos. Um período é dito sensível para determinada habilidade se neste período o investimento tem um retorno maior do que em qualquer outro período. Um período é considerado crítico quando um investimento só exibe retorno neste período.

A equação definida em (1) representa a forma geral da tecnologia de formação de habilidades. A função f_t relaciona os insumos I_t (vetor de investimentos), u_0 (habilidade inata) e θ_t (habilidades no início do período), com o produto θ_{t+1} (habilidade ao final do período t). A função f_t pode assumir diferentes formas de acordo com o período t .

$$\theta_{t+1} = f_t(u_0, \theta_t, I_t) \quad (1)$$

Assume-se que f_t seja estritamente crescente e estritamente côncava em I_t e duas vezes continuamente diferenciável em todos os seus argumentos. A tecnologia (1) pode ser reescrita em forma recursiva, ou seja, em função dos investimentos até o período t conforme a equação (2).

$$\theta_{t+1} = m_t(u_0, I_1, \dots, I_T), t=1, \dots, T \quad (2)$$

Neste modelo, destacam-se duas características da tecnologia de produção. A complementaridade dinâmica, que ocorre quando $\partial^2 f_t(h, \theta_t, I_t) / \partial \theta_t \partial I_t > 0$, ou seja, quando o estoque de habilidades adquiridos no período anterior aumenta a produtividade do investimento no período presente. E a auto-produtividade, que ocorre quando $\partial f_t(h, \theta_t, I_t) / \partial \theta_t > 0$, ou seja, o nível de habilidades de um período contribui para a formação de habilidades do período seguinte. Atuando juntas, a complementaridade dinâmica e a auto-produtividade produzem efeitos multiplicadores. Esse é o mecanismo pelo qual habilidades produzem habilidades. Estes aspectos implicam um *trade-off* entre equidade e eficiência para

investimentos tardiamente realizados em crianças, mas não para investimentos no início do ciclo de vida.

Para simplificar a exposição do modelo, consideram-se dois períodos na formação de habilidades da criança. Com dois períodos, a equação (2) passa a ser escrita conforme a equação (3).

$$\theta_3 = m_2(u_0, I_1, I_2) \quad (3)$$

Para formalizar estes conceitos, Cunha e Heckman (2007) utilizam um modelo de gerações sobrepostas, onde a família é o agente tomador de decisão. A família consiste em um pai e seu filho. Os pais investem em seus filhos por altruísmo. Este tipo de modelagem onde se considera a família como agentes tomadores de decisão foi introduzida por Becker (1974). Os modelos anteriores desenvolvidos na literatura econômica (Becker e Tomer, 1986; Aiyagari *et al*, 2002; Benabou, 2002) consideravam apenas um período no desenvolvimento da criança e implicitamente definiam os investimentos em diferentes fases da infância como sendo substitutos perfeitos, ou seja, baixos níveis de investimento nos anos iniciais da criança poderiam ser compensados por investimentos posteriores.

Cunha e Heckman (2007) mostram que os modelos que consideram apenas um período na formação de habilidades não distinguem entre investimentos iniciais e investimentos tardios. Nesses modelos, implicitamente, os investimentos em diferentes períodos são considerados substitutos perfeitos, ou seja, não importa o momento em que os investimentos são realizados, apenas a soma de investimentos em todos os períodos.

Considerando que os investimentos em diferentes períodos são substitutos perfeitos, a função geral definida para dois períodos conforme (3), assume a forma conforme (4).

$$\theta_3 = m_2(h, \theta_1, \gamma I_1 + (1-\gamma)I_2) \quad (4)$$

onde: $\gamma=1/2$.

O caso oposto ocorre quando os investimentos são considerados complementares perfeitos, como mostra a equação (5).

$$\theta_3 = m_2(h, \theta_1, \min\{I_1, I_2\}) \quad (5)$$

A tecnologia em (5), conhecida com Leontief, exige que os investimentos nos dois períodos sejam iguais para uma estratégia ótima, ou seja, um investimento no primeiro período não tem efeito no produto final se não houver investimentos no segundo período.

Neste caso extremo, crianças de famílias em situação de desvantagem que recebem baixos investimentos no primeiro período não poderão ser compensadas com altos investimentos no segundo período.

As funções apresentadas em (4) e (5) são casos específicos, apresentados para ilustrar os casos de substitutibilidade ou complementaridade perfeita. Um caso mais geral, que captura os aspectos da substitutibilidade e da complementaridade é a tecnologia CES (elasticidade de substituição constante). Esta tecnologia representa a tecnologia de formação de habilidades definidas por Cunha e Heckman (2007). A equação (6) apresenta a estrutura geral da tecnologia CES.

$$\theta_3 = h = m_2(h, \theta_1, [\gamma(I_1)^\phi + (1-\gamma)(I_2)^\phi]^{\frac{1}{\phi}}) \quad (6)$$

para $\phi \leq 1$ e $0 \leq \gamma \leq 1$. O parâmetro de participação CES γ é o multiplicador de habilidades. A elasticidade de substituição $1/(1-\phi)$ representa a medida de quão fácil é a substituição entre os investimentos nos dois períodos. Quando ϕ é pequeno, baixos níveis de investimentos no primeiro período são difíceis de serem compensados por investimentos no segundo período. No caso extremo, quando $\phi \rightarrow -\infty$ a tecnologia (6) converge para a tecnologia Leontief apresentada na equação (5). A complementaridade dinâmica, ou seja, a complementaridade entre os investimentos dos dois períodos tem uma dupla face. Se forem realizados baixos investimentos no primeiro período, quanto maior o grau de complementaridade mais difícil é compensar este baixo nível de investimentos com investimentos no segundo período. Mas um investimento alto no primeiro período perde seus efeitos se não for seguido por altos investimentos no segundo período.

No modelo desenvolvido por Cunha e Heckman (2007), o vetor de investimentos I_t pode conter qualquer investimento na formação de habilidades. Neste capítulo, este vetor de investimentos é separado em investimentos direto da família e investimentos em insumos escolares. Ainda com base na simplificação do modelo em dois períodos considera-se que o primeiro período corresponde ao período antes da criança iniciar o ensino fundamental, e o segundo período corresponde ao período escolar⁶. Portanto, no primeiro período o investimento na formação de habilidades é composto apenas por investimentos diretos da família, ou seja, $I_1 = F_1$, enquanto que no segundo período ocorrem investimentos tanto da

⁶ No Brasil a matrícula da criança na escola só é obrigatória a partir do Ensino Fundamental.

família quanto da escola, $I_2 = (S_2, F_2)$, sendo que F representa os investimentos familiares e S os insumos escolares

A equação (7) apresenta o modelo geral da relação entre os insumos familiares e escolares através de uma tecnologia CES.

$$I_2 = (\delta S^\eta + (1-\delta)F^\eta)^{\frac{1}{\eta}} \quad (7)$$

Onde δ representa o parâmetro de produtividade e η o parâmetro de substitutibilidade.

Para simplificar a apresentação do modelo, assume-se que S e F são substitutos perfeitos, ou seja, $\eta=1$, portanto, (7) passa a ser escrita de acordo com a equação (8).

$$I_2 = \delta S + (1-\delta)F \quad (8)$$

Plugando (8) em (6) o modelo completo é definido conforme (9).

$$\theta_3 = h = m_2(h, \theta_1, [\gamma(I_1)^\phi + (1-\gamma)[\delta S + (1-\delta)F]^\phi]^{\frac{1}{\phi}}) \quad (9)$$

O parâmetro γ é o multiplicador de habilidades e revela que o investimento no primeiro período não age apenas diretamente sobre θ_3 , mas também aumenta a produtividade do investimento do segundo período aumentando θ_2 através dos investimentos do primeiro período. Portanto, I_1 afeta diretamente θ_2 que por sua vez aumenta a produtividade de I_2 na produção de θ_3 . O multiplicador de habilidades γ captura o efeito líquido de I_1 sobre θ_3 através da própria produtividade e da complementaridade direta.

Considerando essa relação de investimentos entre os dois períodos e o papel da escola, crianças que receberam baixos investimentos no primeiro período terão menor capacidade de aprendizado na escola em função de entrarem com menor nível de habilidade. De acordo com este argumento, mesmo crianças que ingressam em escolas com mesma qualidade, se o nível de habilidades ao entrarem na escola for diferente, o aprendizado destas crianças será diferente. As crianças com maior nível de habilidades terão uma maior produtividade em relação aos insumos escolares e, portanto, terão um maior aprendizado em relação às crianças com menor habilidade. Assim, de acordo com esta teoria as escolas não serão capazes de reduzir a desigualdade de habilidades entre as crianças.

Considerando a tecnologia expressa na equação (9), é possível mostrar como a taxa entre investimentos no primeiro e no segundo períodos varia como uma função de ϕ e γ em consequência das escolhas dos pais em diferentes estruturas de mercado, ou seja, em mercados de crédito perfeito e em mercados com restrição de crédito.

Seja w os salários e r a taxa de juros. No início da fase adulta os pais “repassam” o nível inicial de habilidades das crianças, θ_l . Assim que se tornam adultos os pais recebem uma herança b . As variáveis de estado dos pais são sua habilidade, h , os recursos financeiros (herança), b , e o nível inicial de habilidade dos filhos, θ_l . Sejam c_1 e c_2 o consumo das famílias no primeiro e segundo períodos do ciclo de vida das crianças, respectivamente. Os pais decidem como alocar os recursos entre consumo e investimento em diferentes períodos, assim como quanto deixar de herança para o filho, b' , que pode ser positiva ou negativa. A restrição orçamentária é dada pela equação (10).

$$c_1 + I_1 + \frac{c_2 + S + F}{(1+r)} + \frac{b'}{(1+r)^2} = wh + \frac{wh}{(1+r)} + b \quad (10)$$

Seja β o fator de desconto intertemporal da utilidade e δ um parâmetro que define o altruísmo dos pais em relação a seus filhos. Seja $u(\cdot)$ a função de utilidade. A formulação recursiva do problema dos pais é dada pela equação (11)⁷.

$$V(h, b, \theta_l) = \max \{ u(c) + \beta u(c_2) + \beta^2 \delta [V(h', b', \theta_l')] \} \quad (11)$$

O problema das famílias é maximizar (11) sujeito a (10) e (9).

Quando $\phi = 1$, os investimentos no primeiro e segundo períodos são substitutos perfeitos. A estratégia ótima de investimento depende, portanto, do multiplicador CES, γ , e da taxa de juros, r . Essas duas forças atuam em direções opostas, o multiplicador de habilidades atua incentivando os pais a investirem cedo, enquanto que a taxa de juros atua no sentido oposto. É ótimo investir cedo se $\gamma > (1-\gamma)(1+r)$.

Quando $\phi \rightarrow -\infty$, a função de produção CES converge para uma Leontief e a estratégia ótima de investimento é $I_1 = I_2$. Neste caso, investimento no primeiro período é essencial. Ao mesmo tempo, investimentos no segundo período são necessários para manter o efeito do investimento no segundo período. Por questões de eficiência, desvantagens no

⁷ O sobrescrito (\cdot) indica características dos filhos, enquanto que as variáveis sem o sobrescrito são variáveis dos pais.

primeiro período irão se perpetuar e investimentos compensatórios no segundo período são economicamente ineficientes.

Para $-\infty < \phi < 1$, as condições de primeira ordem são necessárias e suficientes dada a concavidade da tecnologia em termos de I_1 e I_2 . Para uma solução interior, a taxa entre investimentos no primeiro e segundo período é dada pela equação (12).

$$\frac{I_1}{I_2} = \left[\frac{\gamma}{(1-\gamma)(1+r)} \right]^{\frac{1}{1-\phi}} \quad (12)$$

Quando $\phi \rightarrow -\infty$ a taxa entre os investimentos não é sensível a variações em γ . Quando $\phi = 0$ a função (8) se torna:

$$\theta_3 = h = m_2(h, \theta_1, (I_1)^\gamma + (I_2)^{1-\gamma}) \quad (13)$$

Em suma, a taxa ótima entre os investimentos é perto de zero para baixos valores de γ , mas explode para infinito quando γ se aproxima de 1. O quadro 2.1 apresenta uma descrição do comportamento da taxa ótima de investimento entre os períodos de acordo com os valores do parâmetro de complementaridade e de auto-produtividade. O comportamento é resumido através de quatro situações. Se os investimentos possuem alta complementaridade, mas baixa auto-produtividade a taxa ótima tende para 1; se os investimentos possuem alta complementaridade e baixa auto-produtividade a taxa ótima também tende para 1; se os investimentos possuem baixa complementaridade e alta auto-produtividade a taxa ótima tende para 0; e, se os investimentos possuem baixa complementaridade e baixa auto-produtividade a taxa ótima tende para infinito. Portanto, se os investimentos possuem alta complementaridade, essa complementaridade domina e força a igualdade entre os investimentos. No entanto, se a complementaridade for baixa e se a auto-produtividade também for baixa os investimentos tendem a se concentrar no último período, mas se a auto-produtividade for alta, os investimentos tendem a se concentrar no primeiro período.

Em relação à taxa ótima entre os investimentos no segundo período, ou seja, entre os recursos da família e os insumos escolares, como foi assumido que os dois insumos são substitutos perfeitos, a taxa ótima de investimento no segundo período vai depender do preço dos investimentos e do fator de produtividade δ .

Quadro 2.1 – Taxa ótima de investimento de acordo com os parâmetros da tecnologia de formação de habilidades.

	Baixa auto-produtividade: $\gamma < \frac{(1+r)}{(2+r)}$	Alta auto-produtividade: $\gamma > \frac{(1+r)}{(2+r)}$
Alto grau de complementaridade: $\phi < 0$	$\frac{I_1}{I_2} \rightarrow 1$ quando $\phi \rightarrow -\infty$	$\frac{I_1}{I_2} \rightarrow 1$ quando $\phi \rightarrow -\infty$
Baixo grau de complementaridade: $0 \leq \phi \leq 1$	$\frac{I_1}{I_2} \rightarrow 0$ quando $\phi \rightarrow 1$	$\frac{I_1}{I_2} \rightarrow \infty$ quando $\phi \rightarrow 1$

Fonte: Cunha e Heckman (2007).

O comportamento da taxa ótima entre os investimentos resumido do quadro 1 considera um mercado de crédito perfeito. Cunha e Heckman (2007) apresentam três fontes de restrições que podem se aplicar as famílias e crianças: (i) a impossibilidade das crianças escolherem seus pais, que é conhecida com o acidente do nascimento; (ii) a impossibilidade dos pais em financiar os investimentos dos filhos através de empréstimos em relação à renda futura dos filhos; (iii) a impossibilidade dos pais em realizar empréstimos em relação às suas próprias rendas para financiar o investimento nos filhos.

Em um modelo com mercado de crédito perfeito, o nível de investimento ótimo não é afetado pelos salários e dotações dos pais, e nem pelos parâmetros que caracterizam a função de utilidade. No entanto, mesmo considerando um mercado de crédito perfeito, os investimentos dos pais dependem de seu nível de habilidades, pois estas características afetam o retorno do investimento. Do ponto de vista da criança, este fato representa uma falha de mercado devido ao acidente do nascimento.

Considerando um mercado de créditos imperfeito, o volume de investimentos na formação de habilidades depende do nível de riqueza das famílias. Com isso, crianças de famílias com menor nível de riqueza serão prejudicadas em seu desenvolvimento se não houver políticas compensatórias. Como existe uma relação positiva entre escolaridade dos pais e nível de riqueza, em um mercado de créditos imperfeito as crianças terão uma dupla restrição de crédito.

Se a complementaridade dinâmica e a auto-produtividade forem verificadas na prática, a formação de habilidades no período escolar (segundo período) irá depender do nível de

habilidades do período anterior, ou seja, havendo uma desigualdade de habilidades entre as crianças no momento em que se inicia o período escolar, os investimentos em insumos escolares só serão capazes de reduzir a desigualdade se for abandonada a busca pela eficiência.

As restrições de crédito geram um ciclo vicioso na formação de habilidades ao longo do ciclo de vida. Crianças que tem pais com menor escolaridade iniciam o ciclo de vida com menores níveis de habilidades. A baixa escolaridade dos pais faz com que os investimentos realizados no desenvolvimento da criança sejam menos produtivos. Além dos investimentos serem menos produtivos, em função da restrição em relação ao nível de riqueza, o volume de investimentos é menos elevado. Portanto, a desigualdade de habilidades se inicia no momento do nascimento e se amplifica ao longo do ciclo de vida.

2.3 – Base de Dados

Neste capítulo utiliza-se a base de dados proveniente do Projeto Geres – Geração Escolar – 2005. Este projeto acompanha alunos de uma amostra de escolas durante os quatro primeiros anos do Ensino Fundamental. Portanto, é um estudo longitudinal onde é possível acompanhar a evolução do nível de habilidades das crianças através do nível de proficiência, ou seja, é possível acompanhar o aprendizado da criança ao longo do tempo. Esta é a primeira pesquisa, no Brasil, a acompanhar uma amostra de alunos durante a primeira fase do Ensino Fundamental, logo, a análise de seus resultados é uma contribuição para a literatura sobre avaliação educacional no Brasil.

Os objetivos do Geres são: (i) identificar as características escolares que maximizam a aprendizagem dos alunos e que minimizam o impacto da origem social sobre o aprendizado; (ii) identificar os fatores escolares que diminuem a probabilidade de repetência dos alunos; e (iii) identificar aquelas características da escola que reduzem a probabilidade do absenteísmo.

O Geres acompanhou alunos de 2005 a 2008, e foram aplicados testes anuais para estimar os níveis de proficiência em matemática e português. Vale ressaltar que em 2005 foram aplicados dois testes, um no início do ano, para ser utilizado como diagnóstico, e outro no final do ano. A aplicação dos testes, portanto, corresponde a um total de cinco ondas. Além dos testes, foram aplicados questionários para as famílias, escolas, professores e diretores. Os questionários aplicados às famílias e às próprias crianças coletaram informações sobre o nível socioeconômico e características individuais das crianças, os questionários para as escolas

buscavam obter informações sobre características das escolas (infra estrutura), dos professores, diretores e turmas.

O acompanhamento longitudinal foi realizado para os alunos que em 2005 estavam matriculados na primeira série do Ensino Fundamental. As escolas da amostra estão localizadas em cinco grandes cidades brasileiras – Belo Horizonte (MG), Rio de Janeiro (RJ), Campo Grande (MS), Salvador (BA) e Campinas (SP). O universo amostral estabelecido para o Projeto Geres consta dos alunos das turmas regulares do período diurno das escolas urbanas públicas com 20 ou mais alunos de 2ª série em 2003 e das escolas particulares com 10 ou mais alunos e com no máximo 3 turmas de 2ª série também em 2003, localizadas nos municípios de Belo Horizonte, Campinas, Campo Grande, Rio de Janeiro e Salvador.

Ao todo foram cadastrados 39.342 alunos no Projeto Geres, sendo que destes, 3.804 crianças não participaram de nenhuma etapa, ou seja, 35.538 crianças participaram de pelo menos uma etapa de avaliação. As escolas de Salvador não participaram da última onda e, portanto, não foram consideradas neste estudo. A tabela 2.1 apresenta estas informações.

Tabela 2.1 – Exclusões da Amostra Geres e Atrito

		Onda 1	Onda 2	Onda 3	Onda 4	Onda 5	Fizeram as 5 ondas
Amostra Geres	Alunos	19.024	20.295	21.373	22.704	17.715	8.920
	Escolas	303	303	298	301	233	231
Excluindo Salvador	Alunos	15.680	16.721	17.272	18.725	17.715	8.920
	Escolas	248	248	245	248	233	231

Fonte: Geres (2005-2008).

Apesar do caráter longitudinal do banco de dados obtido pelo Geres, os questionários não foram aplicados em todos os anos, com exceção do questionário de professores. Em função disso, não é possível acompanhar a evolução das características das escolas e nem do nível socioeconômico das famílias. Esta deficiência do banco de dados prejudica o acompanhamento da evolução dos investimentos, tanto em insumos escolares quanto em insumos familiares, nas crianças.

Visando minimizar esta deficiência, foram acrescentadas informações sobre as escolas no Censo Escolar, realizado anualmente pelo Instituto Nacional de Estudos e Pesquisas Anísio Teixeira – Inep. Através do identificador das escolas foi possível juntar as informações dos dois bancos de dados. Esta junção dos bancos permitiu acompanhar a evolução das características das escolas ao longo dos quatro anos.

Apesar de buscar informações das escolas nos dados do Censo Escolar, uma informação importante para a análise da evolução do aprendizado é o salário dos professores.

Através do identificador das escolas, buscou-se junto ao *site* do Inep⁸ o CNPJ das escolas. As escolas públicas não possuem CNPJ próprios, portanto, para estas foi utilizado o CNPJ dos municípios, dos estados, e da União, para as escolas municipais, estaduais e federais respectivamente. Através do CNPJ foi possível buscar a média de salários no banco de dados da RAIS-Migra⁹, tendo com isso para cada escola privada a média de salários de seus professores. Para as escolas públicas a média de salários corresponde a média por município e rede de ensino.

Para a definição dos insumos escolares que deveriam ser avaliados, foi realizada uma revisão em artigos que estimaram uma função de produção educacional com dados sobre o Brasil. O quadro 2.2 apresenta uma relação dos insumos escolares mais utilizados e os artigos em que estes insumos foram avaliados.

Com base no quadro 2.2 foram selecionadas as variáveis utilizadas na função de produção educacional estimada neste capítulo. O quadro 2.3 apresenta as variáveis utilizadas e o banco de dados de origem¹⁰.

Quadro 2.2 – Relação dos insumos escolares avaliados na literatura nacional

Insumos Escolares	Artigos
Computador para fins pedagógicos	Franco e Menezes Filho (2009); Biondi e Felício (2007)
Computador para uso dos alunos	Menezes Filho (2006); Franco e Menezes Filho (2009); Biondi e Felício (2007)
Internet na escola	Franco e Menezes Filho (2009); Biondi e Felício (2007); Curi e Menezes Filho (2009)
Laboratório de Ciências	Franco e Menezes Filho (2009); Biondi e Felício (2007); Menezes Filho e Pazzelo (2006); Curi e Menezes Filho (2009); Felício e Vasconcellos (2007)
Biblioteca	Franco e Menezes Filho (2009); Biondi e Felício (2007); Tavares et al (2009); Curi e Menezes Filho (2009); Felício e Vasconcellos (2007)
Capacitação de Professores	Barbosa Fernandes (2001); Franco e Menezes Filho (2009); Gomes Neto e Hanushek (1994); Biondi e Felício (2007); Menezes Filho (2006); Macedo (2004); Tavares et al (2007)(2009)Felício e Vasconcellos (2009)
Absenteísmo do professor	Soares (2003); Tavares et al (2009); Biondi e Felício (2007)
Experiência do Professor	Franco e Menezes Filho (2009); Gomes Neto e Hanushek (1994); Menezes Filho e Pazello (2006); Felício e Vasconcellos (2007)

(continua...)

⁸ Mais especificamente no *site* do Data Escola: <http://www.dataescolabrasil.inep.gov.br/dataEscolaBrasil>

⁹ Banco de dados administrativo do Ministério do Trabalho e Emprego com características longitudinais dos trabalhadores formalmente empregados no mercado de trabalho brasileiro.

¹⁰ As variáveis apresentadas no quadro 2.2 que não foram utilizadas neste capítulo não estavam disponíveis nos bancos de dados utilizados ou possuíam muitas observações contendo *missing*.

Quadro 2.2 – Relação dos insumos escolares avaliados na literatura nacional
(continuação)

Insumos Escolares	Artigos
Formação do Professor – Nível superior	Barbosa e Fernandes (2001); Alves (2008); Franco e Menezes Filho (2009); Biondi e Felício (2007); Macedo (2004); Tavares et al (2009); Severnini (2006); Felício e Vasconcellos(2007)
Rendimento do Professor	Menezes Filho e Pazzelo (2006); Tavares et al (2009); Curi e Menezes Filho (2009); Felício e Vasconcellos (2007)
Rotatividade do professor	Biondi e Felício (2007)
Tamanho da turma	Franco e Menezes Filho (2009); Sueli e Martins (2008); Gomes Neto e Hanushek (1994); Biondi e Felício (2007); Menezes Filho (2006)
Experiência do Diretor	Biondi e Felício (2007); Tavares et al (2009); Curi e Menezes Filho (2009)

Fonte: elaboração própria.

Quadro 2.3 – Relação das Variáveis utilizadas

Variáveis	Banco de dados de origem/Ano
Variáveis de alunos e família	
Proficiência nas disciplinas matemática e português	Geres/2005-2008
Raça e gênero	Geres/2008
Nível socioeconômico – Escolaridade da mãe	Geres/2008
Série freqüentada	Geres/2005-2008
Município de residência	Geres/2005-2008
Variáveis de Professor	
Escolaridade e experiência	Geres/2005-2008
Salário do Professor	Rais-Migra/2005-2008
Variáveis de Escola	
Existência de laboratório de informática, laboratório de ciências e biblioteca	Censo Escolar/2005-2008
Rede de ensino – Pública/Privada	Geres/2005-2008

Fonte: elaboração própria.

2.4 - Estatísticas Descritivas

Esta seção apresenta as estatísticas básicas das variáveis utilizadas na função de produção educacional e a evolução do aprendizado das crianças com base na amostra utilizada proveniente do projeto Geres. Também são apontadas evidências da relação entre esta evolução e determinados fatos estilizados da literatura.

As tabelas 2.2 e 2.3 apresentam as estatísticas descritivas básicas das variáveis definidas na seção 2.3. A tabela 2.2 apresenta os valores para as variáveis que não variam ao

longo do tempo, enquanto que a tabela 2.3 os valores para aquelas que variam¹¹. As informações de proficiência representam valores em relação à escala de proficiência do Geres¹² (faça referência a essa escala). O nível socioeconômico é uma variável formada por informações sobre a escolaridade da mãe, a ocupação do pai e presença de bens no domicílio¹³. O salário dos professores está em valores reais de 2008. As demais variáveis são *dummies* e, portanto, o valor médio representa o percentual de cada categoria. Algumas variáveis, como raça e gênero da criança, e salário do professor, não possuem informação para todos os indivíduos. Para não perder essas informações, foi criada uma variável *dummy* indicando a ausência de resposta em cada variável.

A tabela 2.2 apresenta as estatísticas das variáveis invariantes no tempo. Nesta categoria encontram-se as variáveis de características das crianças e nível socioeconômico da família e informação do local de residência. As crianças da amostra são em sua maioria não brancas, já em relação ao gênero existe um balanceamento. O valor médio da variável de nível socioeconômico não possui uma interpretação adequada, mas o desvio padrão mostra que existe uma grande variação entre as famílias. Em relação ao município de residência, 30% das crianças moravam em Campinas, 27% em Belo Horizonte, 27% no Rio de Janeiro e 15% em Campo Grande.

A tabela 2.3 apresenta as estatísticas para as variáveis não constantes ao longo do tempo. São as informações de escolas e professores, além da variável de proficiência dos alunos. Para as informações variantes no tempo, a variabilidade (desvio padrão) geral pode ser decomposta em uma parcela entre indivíduos e uma parcela intra indivíduos ao longo do tempo. A última coluna da tabela apresenta o percentual da variação geral atribuída a cada parcela. Por exemplo, (explique aqui um exemplo com números da tabela). Esta informação sobre a variação entre e principalmente da variação intra é utilizada para a estratégia de identificação dos modelos estimados na seção 2.6.

Nota-se que o valor médio e o desvio padrão da proficiência em matemática são superiores aos da proficiência em português. O salário dos professores, por sua vez, apresenta

¹¹ Os resultados são apresentados em duas tabelas, pois, algumas estatísticas apresentadas para as variáveis que variam ao longo do tempo não podem ser aplicadas nas variáveis fixas ao longo do tempo.

¹² As proficiências foram estimadas utilizando-se modelos da Teoria da Resposta ao Item – TRI. A metodologia utilizada permite que as proficiências dos alunos do GERES possam ser comparadas com as proficiências estimadas no SAEB (Soares, 2009).

¹³ As dimensões escolaridade da mãe, ocupação do pai e presença de bens de capital no domicílio foram sintetizadas em uma única dimensão, nível socioeconômico, com o uso da TRI, para maiores detalhes ver Soares e Oliveira (2005).

apenas 10% de variação ao longo dos anos, mas um alto desvio padrão entre os indivíduos, de mais de 50% do valor médio.

Tabela 2.2 – Estatísticas Básicas das Variáveis Invariantes no Tempo

Variáveis	Observações	Média	Desvio Padrão	Mínimo	Máximo
Características das crianças					
Nse (nível socioeconômico)	43025	0.11	0.62	-1.45	1.66
Branco	43080	0.34	0.47	0	1
Não branco	43080	0.66	0.47	0	1
Raça não respondida	43080	0.01	0.11	0	1
Meninos	43080	0.49	0.50	0	1
Meninas	43080	0.49	0.50	0	1
Gênero não respondido	43080	0.02	0.12	0	1
Municípios					
Belo Horizonte	43080	0.29	0.45	0	1
Campinas	43080	0.30	0.46	0	1
Campo Grande	43080	0.15	0.35	0	1
Rio de Janeiro	43080	0.27	0.45	0	1

Fonte: Geres, Censo Escolar, RAIS-Migra (2005-2008)

Os gráficos de 2.2 a 2.8 apresentam a evolução média da proficiência, em matemática e em português, dos alunos do GERES ao longo dos quatro anos do primeiro ciclo do ensino fundamental, considerando também a avaliação diagnóstica. O gráfico 2.2 apresenta o resultado para todos os alunos da amostra utilizada de forma agregada, enquanto que os demais gráficos apresentam a evolução separada para determinados grupos de características dos alunos. A divisão da evolução do nível de habilidades por grupos busca apresentar evidências de que fatores fora do controle da criança, como raça, nível socioeconômico e município de residência, influenciam no seu aprendizado e, portanto, que há uma desigualdade de oportunidades na formação de habilidades.

Analisando o gráfico 2.2 percebe-se uma diferença no comportamento das duas curvas. Enquanto o crescimento da proficiência em matemática se dá a taxas crescentes, o crescimento da proficiência em português ocorre a taxas decrescentes. As crianças entram na escola com praticamente o mesmo nível de habilidades nas duas disciplinas – aproximadamente 100 pontos na escala de proficiência do Geres. No entanto, após quatro anos, a média de proficiência em matemática passa a ser de quase 250 pontos, enquanto que a de português fica em torno de 170 pontos. Em função desta diferença entre a evolução das duas proficiências, torna-se importante a análise separada para a proficiência em matemática e para a proficiência em português.

Tabela 2.3 – Estatísticas Descritivas para as Variáveis Variantes no Tempo

Variáveis		Média	Desvio Padrão	Mínimo	Máximo	% de variação ¹
Proficiência em matemática	geral	173.02	70.91	15.31	396.55	
	entre		44.46	44.88	281.95	0.39
	intra		55.24	-0.96	342.86	0.61
Proficiência em português	geral	141.92	33.37	47.80	222.96	
	entre		22.84	79.35	196.09	0.47
	intra		24.33	39.15	213.17	0.53
Salário do professor	geral	2,157.63	1,259.40	0.00	6,731.06	
	entre		1,196.95	0.00	5,120.12	0.90
	intra		391.82	-264.71	5,645.58	0.10
<i>Dummy</i> de informação de salário	geral	0.02	0.13	0.00	1.00	
	entre		0.11	0.00	0.80	0.76
	intra		0.06	-0.78	0.82	0.24
Professor com ensino superior (<i>dummy</i>)	geral	0.70	0.46	0.00	1.00	
	entre		0.26	0.00	1.00	0.32
	intra		0.38	-0.05	1.45	0.68
Experiência do professor (<i>dummy</i> – mais de 5 anos) ²	geral	0.44	0.50	0.00	1.00	
	entre		0.30	0.00	1.00	0.38
	intra		0.39	-0.31	1.19	0.62
Escolas Especiais	geral	0.08	0.27	0.00	1.00	
	entre		0.27	0.00	1.00	1.00
	intra		0.01	-0.42	0.75	0.00
Escolas Estaduais	geral	0.24	0.43	0.00	1.00	
	entre		0.43	0.00	1.00	0.99
	intra		0.04	-0.51	0.99	0.01
Escolas Municipais	geral	0.51	0.50	0.00	1.00	
	entre		0.50	0.00	1.00	0.99
	intra		0.04	-0.24	1.26	0.01
Escolas Privadas	geral	0.17	0.37	0.00	1.00	
	entre		0.37	0.00	1.00	1.00
	intra		0.03	-0.58	0.92	0.00
Laboratório de informática (<i>dummy</i>)	geral	0.73	0.45	0.00	1.00	
	entre		0.36	0.00	1.00	0.67
	intra		0.26	-0.02	1.48	0.33
Laboratório de ciências (<i>dummy</i>)	geral	0.29	0.45	0.00	1.00	
	entre		0.43	0.00	1.00	0.92
	intra		0.13	-0.46	1.04	0.08
Biblioteca (<i>dummy</i>)	geral	0.74	0.44	0.00	1.00	
	entre		0.33	0.00	1.00	0.57
	intra		0.29	-0.01	1.49	0.44
Série do aluno	geral	2.18	1.15	1.00	5.00	
	entre		0.17	1.20	3.20	0.02
	intra		1.14	0.38	4.43	0.98

Fonte: Geres, Censo Escolar, RAIS-Migra (2005-2008)

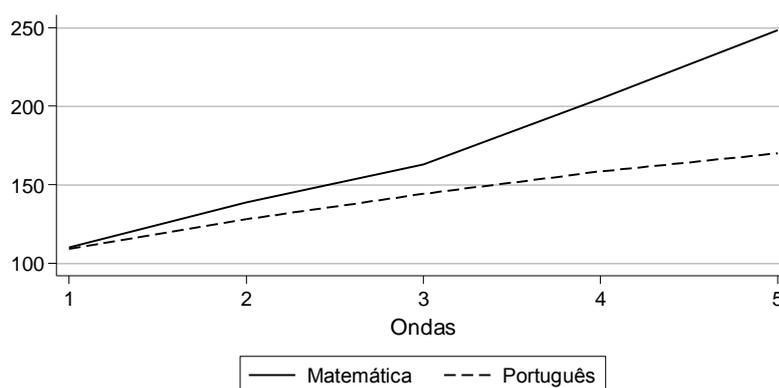
Nota 1: O percentual de variação é construído com o valor da variação intra(entre) elevado ao quadrado dividido pela variação total elevada ao quadrado.

Nota 2: A experiência do professor foi dividida em dois grupos: professores com até 5 anos de experiência na função e professores com mais de 5 anos de experiência.

Os gráficos 2.3 e 2.4 apresentam a evolução das proficiências em matemática e português, respectivamente, dividindo-se as crianças em brancas e não brancas. As curvas para cada raça apresentam um formato semelhante. Para matemática, a distância entre os grupos aumenta ao longo dos anos, ou seja, a diferença entre o nível de conhecimento dos

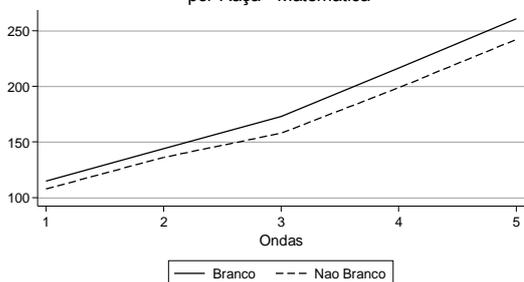
brancos vai aumentando em relação ao dos não brancos. Já para português a distância entre os dois grupos permanece praticamente a mesma ao longo dos anos, sendo que em todo o período o grupo dos brancos possui média de proficiência mais elevada. Este resultado descritivo indica que o sistema escolar, além de não estar sendo capaz de reduzir a desigualdade entre as raças no conhecimento em português, está contribuindo para um aumento destas desigualdades nas habilidades de matemática.

Gráfico 2.2 - Evolução da Proficiência Média dos Alunos do Geres



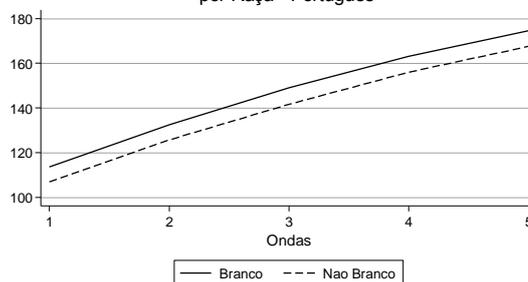
Fonte: Geres (2005-2008)

Gráfico 2.3 - Evolução da Proficiência Média por Raça - Matemática



Fonte: Geres (2005-2008)

Gráfico 2.4 - Evolução da Proficiência Média por Raça - Português

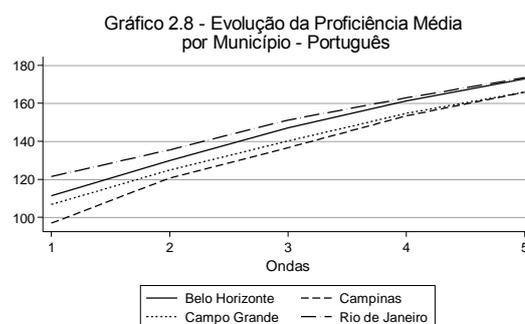
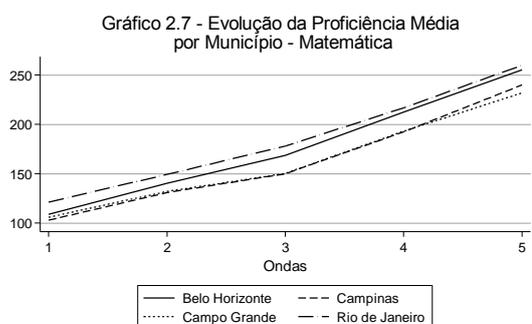
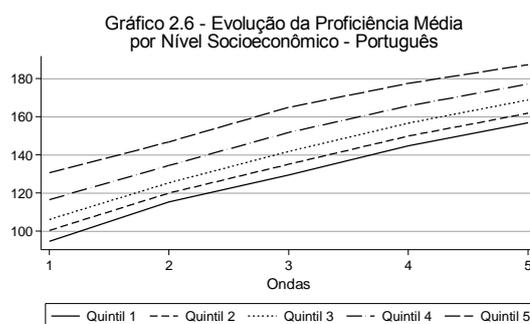
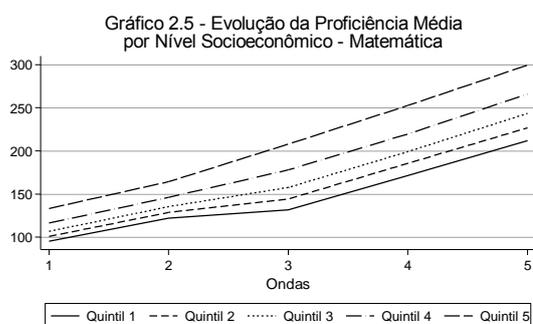


Fonte: Geres (2005-2008)

Os gráficos 2.5 e 2.6 apresentam a evolução dos níveis de proficiência para cada quintil de nível socioeconômico. Da mesma forma como acontece com a raça, o formato das curvas é semelhante para cada quintil e a distância entre as curvas na proficiência em matemática vai se ampliando, ou seja, a desigualdade entre os grupos socioeconômicos se amplia ao longo dos primeiros anos do Ensino Fundamental.

Os gráficos 2.7 e 2.8 apresentam a evolução da proficiência média por município em matemática e português, respectivamente. Destaque para os municípios de Belo Horizonte e

Campo Grande que se iniciam praticamente com o mesmo nível de proficiência. No entanto, na última onda, esta distância é visualmente perceptível, ou seja, as escolas de Belo Horizonte apresentam um maior nível de aprendizado do que as escolas do Campo Grande,. Este efeito é mais acentuado na proficiência em matemática. O município do do Rio de Janeiro inicia com maiores níveis de proficiência e se mantém assim em todas as ondas, tanto para português quanto matemática.

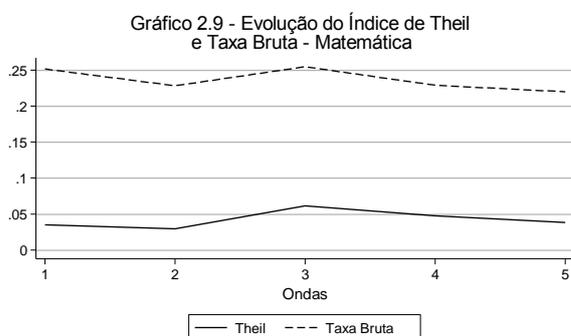


Em suma, os gráficos 2.3 a 2.8 apresentam evidências de que a raça da criança, o nível socioeconômico da família e o município de residência tem relação com o aprendizado das crianças e, portanto, são características que contribuem para a desigualdade de habilidades. Para dar robustez a esses resultados, é apresentada, na sequência, uma análise da decomposição do índice de desigualdade de Theil-T para quantificar a contribuição destas características na desigualdade de habilidades.

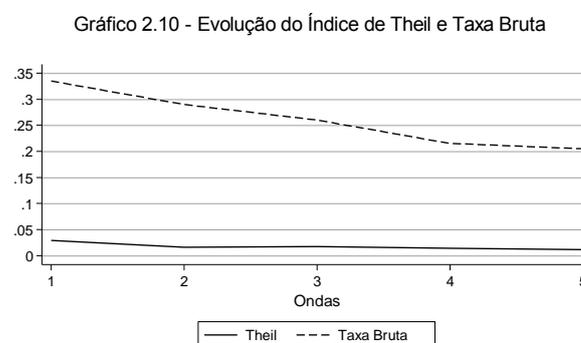
O índice de Theil permite mensurar a desigualdade e decompô-la em parcelas referentes à desigualdade entre determinados grupos e dentro destes grupos. À contribuição da desigualdade entre grupos dá-se o nome de taxa bruta, que indica em termos percentuais o quanto da desigualdade total pode ser atribuída à característica que se está controlando. Os

grupos podem ser formados também por combinação de duas ou mais características, como por exemplo, raça e nível socioeconômico. Neste caso, a taxa bruta indica a contribuição conjunta destas características na desigualdade.

Os gráficos 2.9 e 2.10 apresentam a evolução do índice de Theil e da taxa bruta conjunta para a raça da criança, o nível socioeconômico e o município de residência¹⁴. Nota-se que, na proficiência em matemática, a desigualdade aumenta ligeiramente, mas a taxa bruta diminui, ou seja, o sistema escolar está contribuindo para o aumento da desigualdade, mas reduzindo o que pode ser entendido como efeito da origem familiar ou circunstâncias (referente à contribuição conjunta da raça, nível socioeconômico e município de residência). Já para a proficiência em português, tanto o índice de Theil quanto a taxa bruta se reduzem, ou seja, a desigualdade de habilidades está se reduzindo depois do início do Ensino Fundamental e também o efeito das circunstâncias, consideradas conjuntamente.



Fonte: Geres (2005-2008)

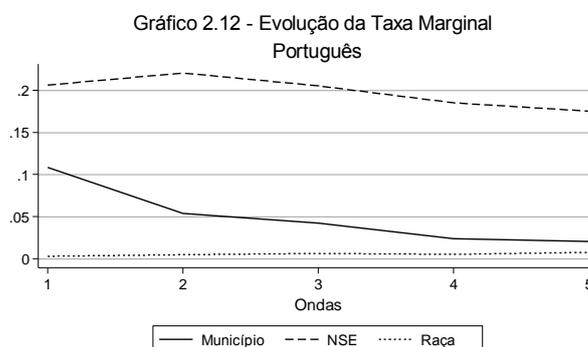
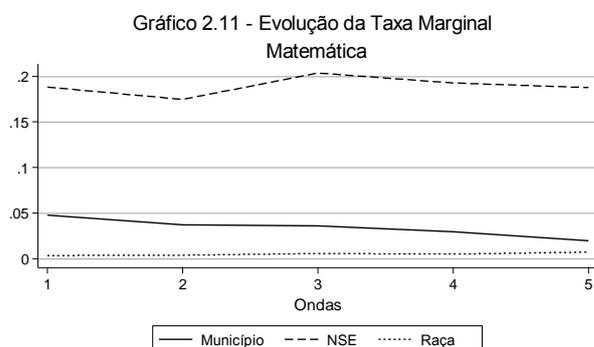


Fonte: Geres (2005-2008)

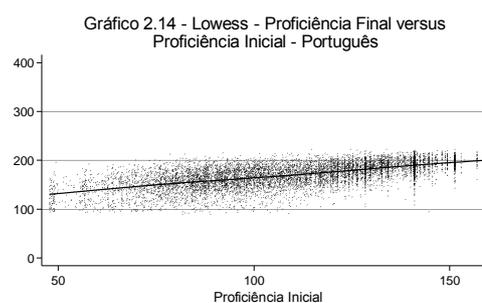
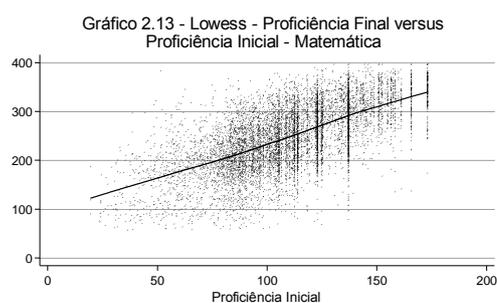
Como a taxa bruta só fornece informações sobre o efeito agregado dos três fatores que estão sendo controlados, calcula-se também a taxa marginal, que mede o efeito isolado de cada variável. Os gráficos 2.11 e 2.12 apresentam a evolução da contribuição marginal de cada variável. Percebe-se que para a proficiência em matemática apenas o efeito do município é decrescente. Ou seja, a redução na taxa bruta foi influenciada pela redução do efeito do município na desigualdade. As outras duas variáveis apresentam crescimento de sua contribuição. O efeito da desigualdade entre as raças cresce de forma monotônica, enquanto a contribuição do nível socioeconômico oscila, mas termina a série com valores próximos aos valores iniciais. Estes resultados corroboram aqueles apresentados nos gráficos de 2.3 a 2.8, onde já havia indícios de que as desigualdades entre as raças e nível socioeconômico estavam

¹⁴ Estas variáveis foram selecionadas, pois representam características que influenciam o nível de conhecimento e não estão sob controle das crianças.

aumentando enquanto se reduzia a desigualdade entre os municípios. Para a proficiência em português (Gráfico 2.12) a contribuição do nível socioeconômico também se reduz, indicando que para esta habilidade o período escolar ameniza os efeitos da origem social.



De acordo com a teoria do ciclo de vida da formação de habilidades, o estoque de habilidades ao final de um período serve com insumo da função de produção de habilidades do período seguinte. Este estoque de habilidades atua em dois sentidos, diretamente na produção de habilidades do período seguinte e indiretamente através do aumento da produtividade dos investimentos em capital humano. Os gráficos 2.13 e 2.14 apresentam o comportamento das habilidades em matemática e português, respectivamente, no último teste feito pelos alunos (quinta onda) como função do nível de habilidades no início do período escolar (avaliação diagnóstica feita na primeira onda). Nos dois casos a proficiência final é uma função crescente em relação à proficiência inicial. No entanto, a proficiência em matemática cresce a taxas maiores do que a proficiência em português, o que indica uma maior persistência do conhecimento em matemática.



Em linhas gerais, os resultados apresentados nesta seção dão uma ideia do comportamento da formação de habilidades em português e matemática. Algumas características são parecidas para as duas disciplinas. Em ambas, a desigualdade no início do período escolar obrigatório é associada às características familiares, o que indica uma desigualdade de oportunidades na formação destas habilidades. A evolução dos índices de desigualdade mostra uma redução na desigualdade de proficiência em português para as crianças durante os quatro anos analisados e um aumento da desigualdade na proficiência em matemática. Por fim, mostra-se que a proficiência em matemática apresenta uma dependência maior em relação aos níveis de proficiência passados do que a proficiência em português.

2.5 – Estratégia Empírica

Considerando o modelo de Cunha e Heckman (2007), Todd e Wolpin (2003) apresentam diversas formas de se estimar a função de produção educacional de acordo com a disponibilidade de informações no banco de dados. Os principais desafios para a estimação da função de produção educacional são: i) a ausência de bancos de dados contendo toda a história de investimentos familiares e escolares; ii) a habilidade genética específica de cada criança não é observada; iii) os resultados de testes padronizados (proficiência) são uma medida imprecisa do nível de conhecimento das crianças.

Assim como Andrabi *et al.* (2011), que partem da abordagem de Todd e Wolpin (2003) e incorporam técnicas de painel dinâmico para a análise do aprendizado de crianças no Paquistão, o presente capítulo aplica tais técnicas para a estimação dos parâmetros da função de produção educacional para os dados do projeto Geres.

A estimação do modelo seguirá uma especificação linear, conforme Todd e Wolpin (2003, 2007), e a função f_t será a mesma para todos os períodos. Deste em ponto em diante, o nível de habilidades no final do período t será representado por y^*_t , em substituição a θ_{t+1} . Esta alteração de notação facilita a exposição da estratégia quando se utiliza variáveis defasadas. Com esta notação, y^*_t representa o nível de habilidades ao final do período t , enquanto X_t representa os insumos investidos durante o período t . Para simplificar a notação os insumos escolares e familiares são incluídos em um único vetor $X_{it} = (F_{it}, S_{it})$.

A equação (14) apresenta o modelo linear na forma de erro.

$$y^*_{it} = X_{it}\beta_1 + X_{it-1}\beta_2 + \dots + X_{it}\beta_t + \sum_{t=1}^T \gamma_t u_{it0} + \sum_{t=1}^T v_{it} \quad (14)$$

onde y^*_{it} representa o verdadeiro nível de conhecimento, sem erros de medida, γ_t representa o efeito da habilidade fixa não observada u_{i0} , e v_{it} um choque de produtividade. Esta especificação implicitamente impõe uma restrição aos parâmetros, assume que um insumo aplicado em um período tem o mesmo efeito no aprendizado do que o mesmo insumo aplicado em outro período¹⁵.

Os três desafios para a estimação dos parâmetros da função de produção são abordados na sequência. O primeiro é a ausência de banco de dados com toda a história de investimentos, portanto, em um primeiro momento, assume-se que a habilidade não observada não é correlacionada com os insumos e que não há erro de medida.

A estratégia mais simples seria considerar que os insumos passados não tem efeito no nível presente de habilidades. Esta estratégia é conhecida como especificação contemporânea e a forma funcional ficaria de acordo com a equação (15), ou na forma de erro composto conforme (16).

$$y^*_i = X_i\beta + \gamma u_{i0} + v_i \quad (15)$$

$$y^*_i = X_i\beta + n_i, \quad n_i = \gamma u_{i0} + v_i \quad (16)$$

Para que as estimativas de (16) por OLS fossem consistentes, ter-se-ia que assumir que os insumos passados não determinam o aprendizado presente ou que os insumos passados não são correlacionados com os insumos contemporâneos, e que a habilidade inata não seja correlacionada com o nível de insumos. De acordo com o modelo da formação de habilidades, toda a história de investimento tem que ser considerada. E é de se esperar que a escolha dos insumos esteja correlacionada com a habilidade fixa não observada, pois em um contexto de maximização da utilidade os investimentos dependem do retorno esperado.

A estratégia mais utilizada na literatura de economia de educação para suprir a ausência da história de investimentos, é a utilização de uma defasagem da medida de proficiência. Assume-se que esta defasagem capta a história dos investimentos, ou seja, $E[y_{it} | X_{it}, X_{it-1}, \dots, X_{i1}, y_{it-1}] = E[y_{it} | X_{it}, y_{it-1}]$. Esta hipótese é conhecida como completude dinâmica. No entanto, se o efeito da aprendizagem for persistente ao longo do tempo, tal hipótese pode não se sustentar. Para testar esta hipótese a estimação deve conter mais de uma defasagem da variável dependente.

Esta especificação é conhecida como valor adicionado. A forma funcional geral é dada pela equação (17).

¹⁵ Ver Tood e Wolpin (2003) para uma forma mais flexível.

$$y^*_{it} = \alpha y^*_{it-1} + X_{it}\beta + \gamma u_{i0} + v_{it} \quad (17)$$

A estimação de (17) por MQO produzirá estimativas inconsistentes, mesmo que a habilidade não observada não seja correlacionada com os insumos, pois, necessariamente estará correlacionada com a variável defasada. Portanto, na presença de efeitos não observados a variável defasada sempre será endógena.

O segundo desafio na estimação dos parâmetros da função de produção educacional é controlar pelo efeito da habilidade não observada. Assume-se que o efeito da habilidade não observada é o mesmo em todos os períodos, ou seja, γ não depende de t . Para eliminar o efeito fixo de habilidade, realiza-se uma transformação na equação (17), tirando a primeira diferença. A equação resultante é representada por (18), sem a presença de u_{i0} .

$$y^*_{it} - y^*_{it-1} = \alpha(y^*_{it-1} - y^*_{it-2}) + (X_{it} - X_{it-1})\beta + v_{it} - v_{it-1} \quad (18)$$

A equação (18) pode ser representada através da notação de diferenças, conforme (19).

$$\Delta y^*_{it} = \alpha \Delta y^*_{it-1} + \Delta X_{it}\beta + \Delta v_{it} \quad (19)$$

Estimar (19) diretamente através de MQO produz estimativas inconsistentes, pois, o termo Δv_{it} é correlacionado com Δy^*_{it-1} . Arellano e Bond (1991) utilizam y^*_{it-2} como instrumento para Δy^*_{it-1} , pois por construção é correlacionado com Δy^*_{it-1} , mas não é correlacionado com Δv_{it} .

Uma hipótese que implicitamente estava sendo imposta nos modelos descritos até este momento, é que o vetor de insumos X_{it} é estritamente exógeno, ou seja, $E(X_{it}v_{is}) = 0$ para qualquer $s, t = 1, 2, \dots, T$. No contexto educacional, esta é uma hipótese pouco realista, pois implica dizer que nem as famílias e nem as escolas irão reagir a choques de produtividade na escolha dos insumos do próximo período.

O estimador proposto por Arellano e Bond permite relaxar esta hipótese. É exigida uma hipótese menos restritiva conhecida como exogeneidade sequencial. A exogeneidade sequencial exige que $E(X_{it}v_{is}) = 0$ para $s \geq t$, mas permite que $E(X_{it}v_{is}) \neq 0$ para $s < t$. Neste caso, o vetor $(X_{i1}, X_{i2}, \dots, X_{is-1})$ é um instrumento válido para X_{it} . Se os insumos forem endógenos, ou seja, correlacionados com o erro idiossincrático do mesmo período, a correção é realizada utilizando mais uma defasagem do vetor de insumos.

O terceiro desafio para estimar os parâmetros na função de produção educacional é considerar erros de medida na variável dependente. Os testes de proficiência apresentam apenas uma aproximação do verdadeiro nível de habilidades, portanto, a variável observada de habilidades possui erro de medida, ou seja, $y_{it}^* = y_{it} + \varepsilon_{it}$.

Portanto, (18) passa a ser escrita como:

$$(y_{it} + \varepsilon_{it}) - (y_{it-1} + \varepsilon_{it-1}) = \alpha[(y_{it-1} + \varepsilon_{it-1}) - (y_{it-2} + \varepsilon_{it-2})] + (X_{it} - X_{it-1})\beta + v_{it} - v_{it-1} \quad (20)$$

Rearranjando os termos e colocando em função das diferenças:

$$\Delta y_{it} = \alpha \Delta y_{it-1} + \Delta X_{it} \beta + (\Delta v_{it} + \Delta \varepsilon_{it} - \alpha \Delta \varepsilon_{it-1}) \quad (21)$$

A segunda defasagem da proficiência proposta como instrumento anteriormente é correlacionada por construção com $\Delta \varepsilon_{it-1}$. Andrabi *et al.* (2011) propõem 3 alternativas para resolver esta endogeneidade. A primeira e mais direta é utilizar uma terceira defasagem da variável de proficiência como instrumento, a segunda é instrumentalizar utilizando a proficiência de uma disciplina diferente, e a terceira é corrigir o erro de medida analiticamente utilizando informações sobre erro padrão da medida de proficiência fornecida pela TRI.

No banco de dados utilizado neste trabalho a informação do erro padrão não foi disponibilizada, portanto a terceira alternativa proposta por Andrabi *et al.* (2011) não será testada. Serão testadas as duas primeiras alternativas.

O uso do estimador desenvolvido por Arellano e Bond (1991) pode não ser o mais adequado quando a variável dependente apresenta uma persistência ao longo do tempo. Neste caso, é preferível utilizar o estimador desenvolvido por Blundell e Bond (1998). Este estimador utiliza instrumentos adicionais realizando a estimação um sistema, com a equação em primeira diferença e a equação original em nível. As variáveis da equação em nível são instrumentalizadas com a defasagem de suas próprias primeiras diferenças.

A correção da endogeneidade no estimador de Arellano e Bond (1991) e Blundell e Bond (1998) é realizada através do uso de instrumentos. Para esta estratégia produzir estimativas consistentes, os instrumentos utilizados devem ser exógenos. Quando o número de instrumentos utilizados é superior ao número de parâmetros a serem estimados, a validade dos instrumentos pode ser testada através do teste de Sargan, em que sob a hipótese nula os instrumentos são válidos.

Realizar a estimação do modelo dinâmico através de Mínimos Quadrados Ordinários ou Efeitos Fixos produz estimativas inconsistentes. No entanto, os valores das estimativas

servem para checar os valores encontrados com a estimação usando a metodologia de Arellano e Bond (1991) ou Blundell e Bond (1998). De acordo com Roodman (2009) e Bond (2002), as estimativas produzidas por Mínimos Quadrados Ordinários são enviesadas para cima enquanto que as estimativas obtidas pelo uso do estimador *within* envia os resultados para baixo. Portanto, bons valores para as estimativas devem estar no intervalo entre estas duas estimações.

2.6 – Resultados

Esta seção apresenta os resultados dos exercícios econométricos. A tabela 2.4 apresenta os resultados para a estimação da especificação contemporânea por Mínimos Quadrados Ordinários - MQO. A estimação foi realizada considerando as informações para a última onda. Para que as estimativas sejam consistentes, é necessário assumir que os insumos passados não são correlacionados com os insumos presentes e que a heterogeneidade individual também não é correlacionada com os insumos presentes. Esta estimação desconsidera a dinâmica do aprendizado. Apesar da teoria da formação de habilidades e de evidências empíricas não sustentarem estas hipóteses, os resultados da estimação contemporânea sem controle da heterogeneidade individual são apresentados como base de comparação.

As variáveis utilizadas são as mesmas apresentadas nas tabelas 2.3 e 2.4 da seção 2.4 (estatísticas descritivas). Para as variáveis sobre raça e gênero da criança, salário e experiência do professor, nem todos os indivíduos tinham a informação. Para não perder observações, foi acrescentada uma variável *dummy* para cada uma das variáveis informando se havia ou não resposta para o indivíduo. Foram realizadas três estimações com especificações diferentes, incluindo controles de forma gradativa para as proficiências em matemática e português. As três especificações foram controladas pela série do aluno e município de residência.

No modelo 1, foram utilizadas apenas características da criança e da família (nível socioeconômico, raça e gênero). Os resultados mostram uma forte dependência do aprendizado em relação ao nível socioeconômico e há um diferencial positivo dos brancos em relação aos não brancos, tanto para a proficiência em matemática quanto para português. Em relação ao gênero nenhuma diferença significativa foi encontrada para a proficiência em matemática e um diferencial positivo em favor das meninas na proficiência em português.

Tabela 2.4 – MQO – Especificação contemporânea

Variável dependente - proficiência	Matemática			Português		
	(1)	(2)	(3)	(1)	(2)	(3)
Nível socioeconômico	44.958*** (0.996)	25.552*** (1.243)	25.059*** (1.245)	15.780*** (0.370)	9.310*** (0.466)	9.032*** (0.466)
Branco	6.565*** (1.275)	5.124*** (1.231)	4.854*** (1.229)	2.815*** (0.475)	2.363*** (0.461)	2.219*** (0.459)
Feminino	1.548 (1.187)	1.236 (1.143)	1.015 (1.140)	5.782*** (0.442)	5.701*** (0.428)	5.611*** (0.426)
Escola Federal		54.223*** (2.974)	60.004*** (3.212)		18.625*** (1.114)	20.845*** (1.200)
Escola Estadual		9.533*** (1.515)	3.816** (1.945)		1.550*** (0.568)	-0.664 (0.727)
Escola Municipal			Omitida			
Escola Privada		43.198*** (2.201)	39.418*** (2.674)		13.528*** (0.825)	12.028*** (0.999)
Laboratório de informática		-8.123*** (1.711)	-7.831*** (1.717)		-2.115*** (0.641)	-2.056*** (0.642)
Laboratório de Ciências		4.193** (1.862)	4.231** (1.888)		1.592** (0.698)	1.609** (0.706)
Biblioteca		6.784*** (1.874)	6.272*** (1.903)		2.384*** (0.702)	2.427*** (0.711)
Salário do Professor			-0.004*** (0.001)			-0.001*** (0.000)
Salário do Professor não informado			-11.180*** (4.272)			-3.363** (1.597)
Professor com Ensino Superior			0.677 (1.731)			1.808*** (0.647)
Experiência do Professor – menos de 5 anos			-8.263*** (2.540)			-4.854*** (0.949)
Experiência do Professor – 5 a 10 anos			-4.023** (2.002)			-2.272*** (0.748)
Experiência do Professor – 11 a 15 anos			-4.856*** (1.744)			-1.890*** (0.652)
Experiência do Professor – mais de 15 anos			Omitida			
Experiência do Professor – não informada			-9.607*** (2.302)			-3.725*** (0.861)
Constante	11.988* (7.177)	2.117 (7.272)	39.063*** (9.387)	69.097*** (2.671)	65.342*** (2.724)	81.317*** (3.509)
Observações	8,605	8,605	8,605	8,605	8,605	8,605
R-quadrado	0.335	0.384	0.388	0.350	0.389	0.396

Fonte: elaboração própria.

Nota 1: *** significância a 1%, ** significância a 5%, * significância a 10%

Nota 2: desvio padrão entre parênteses

Nota 3: todas as estimações as foram controladas por efeitos fixos de município e pela série do aluno.

O modelo 2 mostra as estimativas da especificação acrescentando informações sobre a rede de ensino da escola (pública/privada) e da infraestrutura (existência de laboratório de informática, laboratório de ciências e biblioteca). A magnitude das variáveis individuais se reduz, evidenciando uma associação entre características da criança com características das escolas. A informação sobre a rede de ensino mostrou uma contribuição positiva das escolas especiais, privadas e estaduais em relação às escolas municipais para as duas disciplinas analisadas, o mesmo ocorrendo para as escolas com laboratório de ciências e biblioteca. As estimativas também mostram uma contribuição negativa da existência de laboratório de informática.

No terceiro modelo foram adicionadas informações sobre os professores (salário, experiência e escolaridade). A magnitude das variáveis individuais se reduz novamente, no entanto de forma menos acentuada, e também há uma redução dos coeficientes das variáveis de escola. O salário do professor apresenta um efeito negativo, enquanto que o aumento na escolaridade e experiência possuem um efeito positivo. Essa análise é a mesma para as duas disciplinas. De uma forma geral, a principal diferença entre as duas disciplinas é a magnitude dos parâmetros, sendo sistematicamente superior para a proficiência em matemática.

Os próximos resultados são obtidos utilizando a especificação da dinâmica do aprendizado. São apresentados resultados para a estimação por MQO, por efeitos fixos e pelo estimador de Blundell e Bond (1998) para modelos com uma defasagem da variável dependente, duas e três defasagens. A inclusão gradativa de defasagens tem o objetivo de mostrar o comportamento da dinâmica do aprendizado.

Conforme mencionado na seção 2.5 (Estratégia Empírica), em modelos dinâmicos, as estimativas por MQO e efeitos fixos são enviesadas. No entanto, servem como limites para as estimativas de Blundell e Bond, pois as estimativas por MQO são enviesadas para cima, enquanto que as estimativas por efeitos fixos são enviesadas para baixo. Portanto, os valores para os coeficientes das defasagens da variável dependente devem estar no intervalo entre as estimativas de MQO e efeitos fixos. Em todas as estimações as variáveis utilizadas são as mesmas para a estimação do modelo 3 da tabela 2.4.

O estimador de Blundell e Bond permite considerar três situações na relação entre as variáveis explicativas e o termo de erro (já expurgado dos efeitos fixos). As variáveis podem ser exógenas, pré-determinadas (correlacionadas com o erro do período anterior) e endógenas (correlacionadas com erro do mesmo período). Nas estimativas apresentadas na tabela 2.5, as variáveis de características individuais foram consideradas exógenas. O mesmo tratamento foi

dado para as *dummies* de município, e para o salário do professor, laboratório de ciências e rede de ensino da escola¹⁶. A informação de série cursada pelo aluno foi considerada como pré-determinada, pois o resíduo do período anterior certamente está relacionado com a série do período atual. E por fim, as variáveis de escolaridade e experiência do professor, e existência de laboratório de informática e biblioteca foram consideradas como endógenas. Como exemplo de uma possível fonte de endogeneidade, suponha uma política aplicada no período atual que simultaneamente altere estas características dos professores e da escola e que também altere outras características que podem influenciar no aprendizado e que não estão sendo controladas pelo modelo, como alguma nova prática didática. Em uma situação como esta, os coeficientes das variáveis observadas podem estar refletindo, na verdade, a alteração na prática didática não observada.

Feitas estas considerações, os resultados das estimações dinâmicas são apresentados na sequência. A primeira estimação contém apenas uma defasagem da proficiência. As colunas de 1 a 3 da tabela 2.5 apresentam os resultados da estimação por Mínimos Quadrados Ordinários (MQO), Efeitos Fixos (EF) e Blundell e Bond (BB), respectivamente¹⁷. No painel A, são apresentados os resultados para a proficiência em matemática. O valor do coeficiente da variável defasada na estimativa por BB é de 0,740 (significativo a 1%) e está no intervalo entre as estimativas de MQO e EF. No entanto, o teste de Sargan rejeita o modelo apresentando um valor muito alto para a estatística de teste. Este resultado indica que o uso dos instrumentos não é válido. Isto pode ter ocorrido em função do processo dinâmico da aprendizagem não ser captado com o uso de apenas uma defasagem.

O próximo passo foi considerar duas defasagens para a variável dependente. Novamente foram realizadas três estimativas de acordo com o acontecido para apenas uma defasagem. Os resultados são apresentados nas colunas de 4 a 6 da tabela 2.5. Os valores para as variáveis defasadas foram de 0,398 para a primeira defasagem e 0,426 na segunda defasagem, ambas significativas a 1% (resultados de Blundell e Bond, coluna 6). Estes resultados vão ao encontro do argumento de um processo dinâmico da aprendizagem de forma mais persistente, ou seja, a segunda defasagem da proficiência contribui para explicar a proficiência atual. Da mesma forma como ocorreu para o uso de uma única defasagem, os valores dos coeficientes da variável defasada estão no intervalo entre as estimativas de MQO

¹⁶ Estas três variáveis, embora possam existir argumentos para serem consideradas pré-determinadas ou endógenas, foram tratadas como exógenas devido à baixa variação ao longo do tempo, o que impede a identificação se usadas defasagens destas como instrumentos.

¹⁷ Na parte inferior da tabela, são apresentadas as defasagens utilizadas como instrumentos para as equações em diferença e em nível e os valores do teste de Sargan.

e EF¹⁸. No entanto, o teste de Sargan rejeita novamente o modelo. Apesar de o teste continuar rejeitando o modelo, o valor da estatística de teste teve uma forte redução, apontando indícios que o uso de mais uma defasagem contribui para aproximar o modelo do verdadeiro modelo de formação de habilidades.

Na sequência foi acrescentada uma terceira defasagem da proficiência. Os resultados encontrados para os coeficientes foram significativos indicando que a persistência do aprendizado permanece mesmo considerando três períodos de defasagem. Neste modelo o valor da estatística do teste de Sargan se reduziu novamente, embora tenha continuado a rejeitar o modelo.

A coluna 10 refaz a estimação com três defasagens da proficiência, mas considera que há uma persistência do efeito de insumos passados sobre a proficiência. Portanto, ao invés de incluir duas defasagens dos insumos endógenos, inclui apenas a terceira defasagem. Isso porque se houver uma correlação serial dos erros, pode haver uma correlação entre os insumos defasados com o termo de erro. Com esta estimação, o resultado do teste de Sargan não rejeitou o modelo conforme demonstra o valor da probabilidade do teste. Os coeficientes das defasagens da variável dependente estão entre os valores encontrados para a estimativa por MQO (coluna 7) e por EF (coluna 8)¹⁹. Portanto, considera-se que o uso de três defasagens consegue captar a dinâmica do processo de aprendizado. Apesar da segunda defasagem não apresentar coeficiente com valor significativo, a terceira defasagem volta a ser significativa.

Os resultados confirmam que o aprendizado é um processo dinâmico, ressaltando a importância que os níveis de habilidades passados têm no aprendizado presente. Este resultado indica que mesmo que as crianças entrassem todas em escolas iguais em suas características, a desigualdade de habilidades aumentaria durante os anos escolares, pois as crianças que iniciassem com níveis de habilidades mais elevados teriam um ritmo maior de aprendizado.

Quanto aos efeitos de insumos escolares e de características individuais sobre o aprendizado, embora a incorporação do componente dinâmico tenha reduzido tais efeitos, como apontado pelos testes F da tabela 2.5, os mesmos permanecem significativos a 1%.

O painel B da tabela 2.5 mostra os resultados para as estimativas de painel dinâmico para a proficiência em português. A inclusão das defasagens de forma gradativa apresenta o

¹⁸ O valor da estimativa pontual da segunda defasagem é superior ao encontrado pela estimativa de MQO, mas esta dentro do intervalo de confiança.

¹⁹ Considerando o intervalo de confiança.

mesmo resultado das estimativas de matemática. Os resultados da coluna 10 são considerados consistentes de acordo com o teste de Sargan. Os coeficientes das defasagens da variável dependente são positivos e significativos, indicando a persistência do aprendizado também em português.

2.7 - Considerações Finais

O capítulo contribui para as políticas de avaliação educacional e de formulação de políticas de redução das desigualdades utilizando uma análise da dinâmica do aprendizado e dos efeitos de investimentos realizados anteriormente à entrada da criança na escola. Para tanto, a partir do modelo teórico de Cunha e Heckman (2007) para tecnologia da formação de habilidades e com o uso de informações longitudinais do GERES, um banco de dados educacional com ainda pouco explorado na literatura nacional, são estimados parâmetros de uma função de produção educacional.

Os resultados encontrados a partir de técnicas de estimação de painel dinâmico mostram evidências empíricas da persistência do aprendizado. Este resultado indica que a escola tem um papel limitado na redução de desigualdades de habilidades, dado a dependência do aprendizado a habilidades passadas. Ou seja, grande parte do que a criança consegue aprender no sistema de ensino regular é dependente dos investimentos em capital humano que foram realizados na primeira infância, antes da sua entrada na escola.

Os resultados dos efeitos dos insumos escolares não podem ser considerados conclusivos, em virtude da não estabilidade de seus coeficientes com alterações de especificações dos modelos. No entanto, existe uma indicação de efeitos positivos de recursos escolares, como biblioteca e laboratório de ciências, além do salário dos professores.

Como contribuição para a formulação de políticas públicas destaca-se a necessidade de intervenções em crianças de famílias de baixo nível socioeconômico antes mesmo da idade escolar. O baixo nível de habilidades medido quando as crianças entram na escola persiste durante o período escolar, principalmente para matemática.

Tabela 2.5 – Estimações de Painel Dinâmico

	(1-MQO)	(2-EF)	(3-BB)	(4-MQO)	(5-EF)	(6-BB)	(7-MQO)	(8-EF)	(9-BB)	(10-BB)
PAINEL A - MATEMÁTICA										
L.profiiciência	0.840*** (0.005)	0.175*** (0.006)	0.740*** (0.011)	0.638*** (0.007)	-0.119*** (0.007)	0.398*** (0.028)	0.578*** (0.008)	-0.475*** (0.011)	0.644*** (0.086)	0.592*** (0.142)
L2.profiiciência				0.390*** (0.009)	-0.022*** (0.008)	0.426*** (0.020)	0.321*** (0.010)	-0.036*** (0.010)	0.069 (0.044)	0.084 (0.057)
L3.profiiciência							0.254*** (0.013)	0.115*** (0.015)	0.320*** (0.047)	0.333*** (0.064)
Teste F características individuais	85.40***	-	201.60***	53.74***	-	237.50***	20.44***	-	32.65***	18.05***
Teste F características escolares	83.26***	4.06***	243.79***	44.58***	2.14**	352.50***	12.08***	0.28	47.30***	28.47***
Teste F características dos professores	2.95***	4.88***	19.34***	2.79**	3.25***	46.62***	1.20	1.32	39.39***	40.67***
R-quadrado	0.792	0.767	-	0.783	0.750	-	0.779	0.674	-	-
e(X ²)	-	-	2337	-	-	471.5	-	-	53.64	17.92
e(probabilidade)	-	-	0.0000	-	.-	0.0000	-	-	0.0036	0.528
PAINEL B - PORTUGUES										
L.profiiciência	0.718*** (0.004)	0.036*** (0.006)	0.509*** (0.012)	0.608*** (0.006)	-0.166*** (0.007)	0.509*** (0.041)	0.552*** (0.007)	-0.466*** (0.009)	0.580*** (0.091)	0.520*** (0.117)
L2.profiiciência	-	-	-	0.249*** (0.006)	-0.115*** (0.006)	0.247*** (0.022)	0.219*** (0.008)	-0.196*** (0.009)	0.235*** (0.058)	0.226*** (0.067)
L3.profiiciência	-	-	-	-	-	-	0.099*** (0.007)	-0.000 (0.008)	0.124*** (0.029)	0.151*** (0.031)
Teste F características individuais	114.05***	-	413.81***	59.17***	-	167.97***	27.93***	-	23.44***	16.97***
Teste F características escolares	48.38***	4.43***	274.26***	13.34***	1.24	66.59***	4.54***	2.82**	29.81***	23.31***
Teste F características dos professores	4.37***	6.91***	24.26***	3.74***	14.64***	8.52***	1.98*	12.95***	12.46***	18.84***
R-quadrado	0.773	0.725	-	0.762	0.626	-	0.764	0.552	-	-
e(X ²)	-	-	1314	-	-	281.3	-	-	55.08	23.29
e(probabilidade)	-	-	0.0000	-	-	0.0000	-	-	0.0024	0.2250
Observações	34,040	34,040	34,040	25,460	25,460	25,460	16,956	16,956	16,956	16,956
Número de alunos	-	8,605	8,605	-	8,605	8,605	-	8,605	8,605	8,605
Instrumentos Equação em diferença	-	-	Y _{t-2} , X _{t-2}	-	-	Y _{t-3} , X _{t-2}	-	-	Y _{t-4} , X _{t-2}	Y _{t-4} , X _{t-3}
Instrumentos Equação em nível	-	-	ΔY _{t-2} , ΔX _{t-2}	-	-	ΔY _{t-3} , ΔX _{t-2}	-	-	ΔY _{t-4} , ΔX _{t-2}	ΔY _{t-4} , ΔX _{t-3}

Fonte: elaboração própria.

Nota 1: *** significância a 1%, ** significância a 5%, * significância a 10%

Nota 2: desvios padrão entre parênteses

2.8 - Referências

ALVES, F. Políticas educacionais e desempenho escolar nas capitais brasileiras. **Cadernos de Pesquisa**, v. 38, n. 134, p. 413-440, 2008.

ANDRABI, T., DAS, J., KHWAJA, A. I. E ZAJONC, T. Do Value Added Estimates Add Value? Accounting for Learning Dynamics. **American Economic Journal: Applied Economics**, v. 3, n. 3, p. 29-54, 2011

ANDRADE, D. F E MIRANDA, E. C. M. GERES: Relatório do Plano Amostral. GERES-2005. Disponível em: < <http://www.geres.ufmg.br>>. Acesso em: 22/05/2011, 2004.

ARELLANO, M. E BOND, S. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. **Review of Economic Studies**, v. 58, p. 277-297, 1991.

BARBOSA, M.; F. E FERNANDES, C. A Escola brasileira faz diferença? Uma investigação dos efeitos da escola na proficiência em matemática dos alunos da 4ª série. In: FRANCO, C. (org.) **Avaliação, ciclos e promoção na educação**. Porto Alegre: Artmed, 2001.

BARROS, R. P.; MENDONÇA, R.S.P. Os Determinantes da Desigualdade no Brasil. **Economia brasileira em perspectiva**. Rio de Janeiro: IPEA, p. 421-474., 1996.

BIONDI, R. L.; FELÍCIO, F. DE. Atributos escolares e o desempenho dos estudantes: uma análise em painel de dados do Saeb. INEP-MEC, 2007.

BLUNDELL, R.; S. BOND. Initial conditions and moment restrictions in dynamic panel data models. **Journal of Econometrics**, v. 87, p. 115-143, 1998.

BROOKE, N.; SOARES, J.F. (Orgs.). **Pesquisa em Eficácia Escolar: Origens e Trajetórias**. Belo Horizonte: Editora UFMG, 2008.

CAMERON, S. V.; HECKMAN, J.J. Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males. **Journal of Political Economy**, University of Chicago Press, v. 106, n. 2, p. 262-333, Abril, 1998.

COLEMAN, J. S.; CAMPBELL, E. Q.; HOBSON, C. J.; MCPARTLAND, J.; MOOD, A. M.; WEINFELD, F. D. Equality of educational opportunity. Washington, D.C.: National Center for Educational Statistics, 1966.

CUNHA, F.; HECKMAN, J. J.; LOCHNER, L.; MASTEROV, D. Interpreting the Evidence on Life Cycle Skill Formation. In **Handbook of the Economics of Education**, edited by E. A. Hanushek and F. Welch. Amsterdam: North-Holland, 2006

CUNHA, F.; HECKMAN, J. The Technology of Skill Formation. **American Economic Review**, v. 97, n. 2, p. 31-47, 2007.

CUNHA, F.; HECKMAN, J. The Economics and Psychology of Inequality and Human Development. NBER Working Paper n. 14695, 2009.

CURI, A. Z.; MENEZES-FILHO, N. A. A relação entre educação pré-primária, salários, escolaridade e proficiência escolar no Brasil. XXXIV Encontro Nacional de Economia – ANPEC, 2009.

FELÍCIO, F.; VASCONCELLOS, L. O Efeito da Educação Infantil sobre o Desempenho Escolar medido em Exames Padronizados. XXXV Encontro Nacional de Economia – ANPEC, 2007.

FERREIRA, F. H. G. Os Determinantes da Desigualdade de Renda no Brasil: Luta de Classes ou Heterogeneidade Educacional?. Texto para Discussão n. 415, Departamento de Economia PUC-Rio, 2000.

FRANCO, A. M. P. E MENEZES-FILHO, N. A. Os determinantes do aprendizado com dados de um painel de escolas do Saeb. XXXVII Encontro Nacional de Economia – ANPEC, 2009.

GERES. Estudo Longitudinal sobre a Qualidade e Equidade no Ensino Fundamental Brasileiro, 2005.

GOMES-NETO, J. B.; HANUSHEK, E. A. Causes and Consequences of Grade Repetition: Evidences for Brazil. **Economic Development and Cultural Change**, v. 43, n. 1, p. 117-48 1994.

HANUSHEK, E. A. E RIVKIN, S. G. School Quality and the Black-White Achievement Gap. Working Paper no. 12651, NBER, Cambridge, MA, 2006.

IBGE - Instituto Brasileiro de Geografia e Estatística. Censo demográfico, 1991, 2000 e 2010. Rio de Janeiro.

INEP - Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. Censo Escolar, 2005-2008.

KEANE, M. P. E WOLPIN, K. I. The Career Decisions of Young Men. **Journal of Political Economy**, v. 105, n. 3, p. 473-522, 1997.

LANGONI, C. G. **Distribuição da Renda e Desenvolvimento Econômico do Brasil**. Rio de Janeiro, Editora Expressão e Cultura, 1973

LEIBOWITZ, A. Home Investments in Children. **Journal of Political Economy**, v. 82, n. 2, p. 111-131, 1974.

MACEDO, G. A. Fatores Associados ao Rendimento Escolar de Alunos da 5ª série (2000) - uma abordagem do valor adicionado. CEDEPLAR – UFMG, 2004.

MACHADO, A. F.; MORO, S.; MARTINS, L.; RIOS, J. Qualidade do Ensino em Matemática: Determinantes do Desempenho de Alunos em Escolas Públicas Estaduais Mineiras. **Economia**, v. 9, n. 1, p. 23-45, 2008.

MENEZES-FILHO, N. A. Os determinantes do desempenho escolar do Brasil. São Paulo, 2006.

MENEZES-FILHO, N. A.; PAZELLO, E. Do Teachers' wages matter for proficiency? Evidence from a funding reform in Brazil. **Economics of Education Review**, v.26, p.660-672, 2007.

MINISTÉRIO DO TRABALHO E EMPREGO. Raismigra. Brasília: MTE, 2005/2008.

MURNANE, R. J.; J. B. WILLETT,; F. LEVY. The Growing Importance of Cognitive Skills in Wage Determination. **The Review of Economics and Statistics**, v.77, n.2, 1995

NEAL, D. A.; JOHNSON, W. R. The Role of Premarket Factors in Black-White Wage Differences. **Journal of Political Economy**, v.104, n.5, p.869-95, 1996.

PISA – Programme for International Student Assessment, 2009.

ROEMER, J. E. **Equality of Opportunity**. Harvard University Press, Cambridge, 1998.

ROODMAN, D. How to do xtabond2: An introduction to difference and system GMM in Stata. **The Stata Journal**, v.9, n.1, 2009.

SOARES, T. M. Influência do professor e do ambiente em sala de aula sobre a proficiência alcançada pelos alunos avaliados no Simave-2002. **Estudos em Avaliação Educacional**, n. 28, 2003.

SEVERNINI, E. R. **A relação entre violência nas escolas e proficiência dos alunos**. Dissertação (Mestrado em Economia) – Departamento de Economia, PUC-RIO, 2007.

TAVARES, P. A.; CAMELO, R. S.; KARMISKI, P. R. A falta faz falta? Um estudo sobre o absenteísmo dos professores da rede estadual paulista de ensino e seus efeitos sobre o desempenho escolar. XXXVII Encontro Nacional de Economia – ANPEC, 2009.

TODD, P.; WOLPIN, K. I. On the Specification and Estimation of the Production Function for Cognitive Achievement. **Economic Journal**, v.113, p.3–33, 2003.

TODD, P.; WOLPIN, K. The Production of Cognitive Achievement in Children: Home, School, and Racial Test Score Gaps. **Journal of Human Capital**, v.1, n.1, p.91-136, 2007.

3 - EFEITO DA PRÉ-ESCOLA SOBRE O NÍVEL DE PROFICIÊNCIA: UMA ANÁLISE A PARTIR DE DADOS LONGITUDINAIS

RESUMO

Este capítulo estima o efeito da pré-escola sobre o nível de proficiência de alunos do primeiro ciclo do ensino fundamental, utilizando um banco de dados longitudinal produzido pelo Projeto GERES. O uso do GERES permite a avaliação da evolução da aprendizagem, bem como o controle de características não observadas dos estudantes. O efeito da pré-escola foi estimado através de uma função de produção educacional considerando o processo dinâmico de aprendizado, conforme Cunha e Heckman (2007) e Todd e Wolpin (2003). A estratégia empírica utilizada foi uma adaptação da estimação por valor adicionado e especificação cumulativa. Os resultados encontrados indicam o efeito cumulativo dos insumos e a presença de heterogeneidade individual. Para Matemática o efeito direto encontrado foi positivo e significativo. Já para português, os coeficientes de pré-escola apresentaram resultados não significativos, indicando que não há efeito direto sobre as habilidades em português. O efeito indireto, manifestado através da proficiência inicial, foi positivo e crescente ao longo dos anos para Matemática. Por outro lado, os estudantes com maiores habilidades em linguagem no início do ensino fundamental aumentam suas diferenças comparativamente aos menos habilidosos, mas esse efeito se inverte ao longo dos anos, indicando que a desigualdade de habilidades em Português pode ser reduzida durante o período escolar.

Palavras Chaves: Pré-escola, GERES, Dinâmica do Aprendizado, Proficiência, Função de Produção Educacional

ABSTRACT

This chapter estimates the preschool effect on the child's proficiency level based on a database built by GERES Project, which is a longitudinal study that monitors a sample of students from 1st grade for four years. The use of GERES allows the evaluation of the evolution of learning, controlling for unobserved characteristics of the students that affect proficiency or performance. Following Cunha & Heckman (2007) and Todd & Wolpin (2003), the preschool effect is estimated through an educational production function that incorporates a dynamic process of learning. The empirical strategy is an adaptation of value added and cumulative specification estimates. Our findings show evidence of cumulative effect of inputs and the presence of individual heterogeneity. In the case of Mathematics the direct effect of preschool on child's proficiency was positive and significant, while in Portuguese, there is no direct effect. Furthermore, the indirect effect shown by the coefficient of initial proficiency is positive and increased over the years for Math. On the other hand, the students with higher skills in language at the beginning of elementary school expanded their differences in relation to less skilled students, but this effect inverts over the years, indicating that the inequality of skills in Portuguese can be reduced during the scholar period.

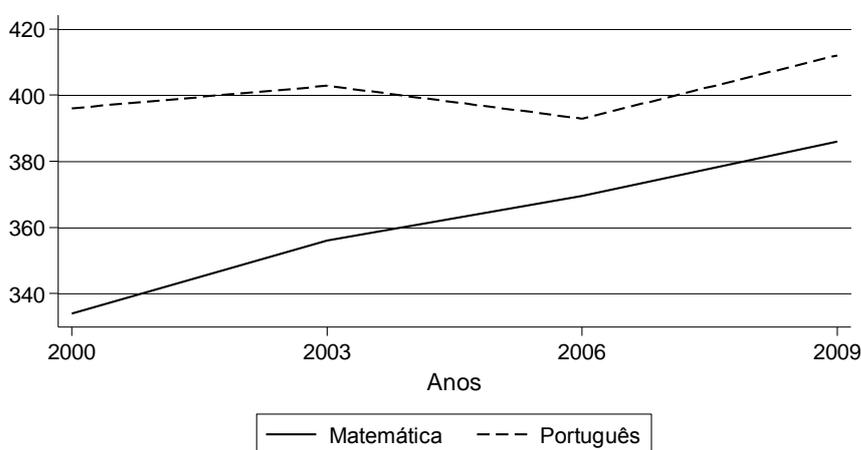
Key Words: Preschool, GERES, Dynamic Process of Learning, Proficiency, Educational Production Function

3.1 – Introdução

Resultados do Programa Internacional de Avaliação de Alunos (Pisa) de 2009 mostram que o Brasil evoluiu em relação aos resultados da edição da avaliação anterior, realizada em 2006. No entanto, essa evolução não foi suficiente para tirar o Brasil da parte inferior do ranking mundial. Em um total de sessenta e cinco países avaliados em 2009, o Brasil ocupou a 53ª posição no ranking de leitura e a 57ª posição no ranking de matemática. Estes resultados mostram a necessidade de políticas públicas voltadas para a melhoria da educação no Brasil. Dentre os diversos tipos de políticas nessa área, como valorização e capacitação de professores, melhor infraestrutura das escolas, aumento da jornada escolar, destacam-se as políticas voltadas para as crianças mais novas, abaixo dos 6 anos de idade.

O gráfico 3.1 mostra evolução do Brasil durante a década passada, nos anos onde foram aplicados a avaliação internacional. Nota-se um crescimento monotônico para a proficiência em matemática, saindo de uma média de 334 pontos no ano de 2000 para 386 em 2010. Para a proficiência em leitura, apesar do incremento entre os anos de 2000 e 2010, esse crescimento foi menos acentuado, inclusive havendo uma queda entre os anos de 2003 e 2006.

Gráfico 3.1 - Proficiência Média - Brasil (2000-2009)



Fonte: PISA (2000-2009)

Na comparação com os países da América Latina participantes da avaliação o Brasil aparece à frente de Argentina e Colômbia, mas atrás de México, Uruguai e Chile. A China, país em desenvolvimento e pertencente ao BRIC, apresentou o melhor resultado entre todos

os países participantes da edição do PISA de 2009, tanto para leitura quanto para matemática. A comparação com estes países mostra que, apesar da evolução brasileira entre 2000 e 2009 ser inferior apenas à evolução de Chile e Luxemburgo, em termos absolutos o país ainda precisa evoluir muito para se aproximar dos países com melhores resultados.

Como estratégia para melhorias no nível de aprendizado das crianças vem ganhando destaque a necessidade de investimentos na primeira infância, período que antecede a entrada da criança no ensino regular obrigatório, ou seja, abaixo dos seis anos de idade. Avanços recentes na neurociência têm demonstrado a importância de investimentos e estímulos nos primeiros anos de vida para o desenvolvimento cerebral e capacidade de aprendizado. Nesse sentido, diversos estudos avaliam o efeito de intervenções realizadas na primeira infância no sucesso do aprendizado escolar e no mercado de trabalho. Considerando estas evidências, a frequência à pré-escola é apontada como um mecanismo de estimular o desenvolvimento cognitivo e não cognitivo das crianças, principalmente para crianças em situação de vulnerabilidade.

Cunha e Heckman (2007, 2009, 2010) mostram que investimentos educacionais na primeira infância focados em crianças de famílias de baixo nível socioeconômico apresentam altas taxas de retorno em contraste com investimentos realizados em adolescentes. De acordo com os autores, existe um *trade off* entre equidade e eficiência em investimentos realizados em adolescentes, mas não para investimentos realizados na primeira infância.

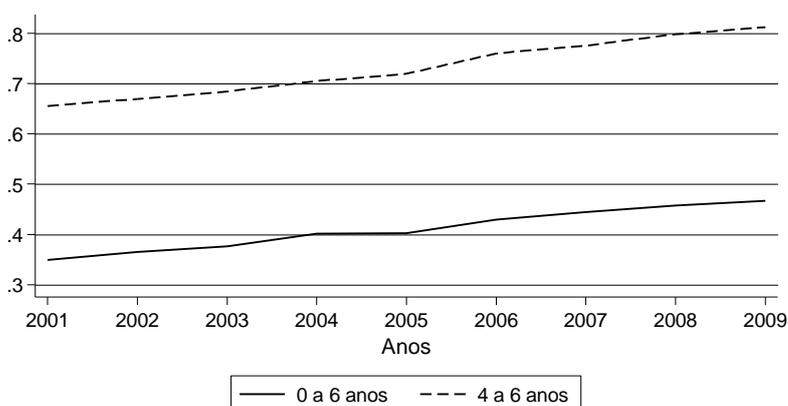
A literatura internacional mostra a existência de diversos programas voltados para a promoção do desenvolvimento infantil, principalmente para crianças de famílias desfavorecidas (CITAR). Nos Estados Unidos destacam-se o *High/Scope Perry PreSchool Project*, o *Carolina Abecedarian Project*, o *Early Training Project*, o *Head Start*. Na América Latina os programas mais conhecidos são o Oportunidades do México, o *Proyecto de Desarrollo Infantil* da Bolívia e a construção de pré-escolas na Argentina.

No Brasil, a análise do impacto de investimentos em primeira infância é voltada para o efeito da pré-escola e creches. A matrícula à pré-escola não é obrigatória e a oferta pública não é universalizada como ocorre no ensino fundamental. Dados do Banco Mundial (2006) mostram que crianças de famílias de classes mais pobres têm um menor percentual de frequência à educação infantil. Considerando as evidências dos efeitos positivos no desenvolvimento infantil em crianças que frequentam a pré-escola, este acesso desigual tem um efeito perverso na geração de desigualdades.

Algumas ações estão sendo tomadas pelo governo brasileiro. Um primeiro passo foi a redução da idade obrigatória para matrícula no ensino fundamental de 7 para 6 anos. Outra iniciativa foi a criação do Fundo de Manutenção e Desenvolvimento da Educação Básica e de Valorização dos Profissionais da Educação (FUNDEB) que atende os níveis educacionais da creche ao ensino médio.

Dados da Pesquisa Nacional por Amostra de Domicílios (PNAD) mostram um aumento no percentual de crianças de 6 anos ou menos frequentando a escola na última década. O Gráfico 3.2 mostra a evolução da frequência para a faixa etária de 0 a 6 anos, fase denominada de primeira infância, e da faixa específica da pré-escola, 4 a 6 anos. Os resultados mostram um crescimento de forma monotônica, sendo para que a faixa da pré-escola os valores partem de 65% no início da década para um pouco mais de 81% no ano de 2009. Já para a primeira infância como um todo, a frequência era em torno de 35% em 2001 e atinge 47% em 2009.

Gráfico 3.2 - Matrícula na Educação Infantil - Brasil (2001-2009)



Fonte: PNAD/IBGE (2001-2009)

Uma série de estudos recentes avalia o efeito da pré-escola nos testes de proficiência de alunos do ensino fundamental e ensino médio (Curi e Menezes Filho, 2009; Felício e Vasconcelos, 2007; Calderini e Souza, 2009; Felício *et al.*, 2009; Rodrigues *et al.*, 2010). Os resultados encontrados indicam uma relação positiva entre frequentar a pré-escola e o sucesso escolar. Este capítulo visa contribuir para a análise do impacto de investimentos na primeira infância utilizando um banco de dados ainda pouco explorado na literatura nacional, o projeto Geração Escolar 2005 – GERES. O GERES é um estudo longitudinal que acompanhou uma amostra de alunos da 1ª série do ensino fundamental durante quatro anos. A utilização de um

estudo longitudinal permite avaliar a evolução do aprendizado, controlando por características não observadas dos alunos que afetam a proficiência ou desempenho. Tal estratégia é apontada por pesquisadores de avaliação educacional como a melhor opção para avaliar os determinantes do aprendizado.

O capítulo está estruturado da seguinte maneira. Além desta introdução, a seção 3.2 traz uma breve revisão da literatura internacional e nacional sobre avaliação dos efeitos de programas de educação infantil. A seção 3.3 apresenta o banco de dados utilizado e a estratégia de identificação do efeito da pré-escola. A seção 3.4 mostra algumas estatísticas descritivas e a seção 3.5 os resultados econométricos. Por fim, a seção 3.6 traz as considerações finais.

3.2 – Revisão da Literatura

Estudos da neurociência têm demonstrado a importância dos primeiros anos de vida da criança na formação e desenvolvimento cerebral (Knudsen, 2004; Casella et al., 2011). Avanços em equipamentos de ressonância magnética permitem enxergar como o cérebro reage a estímulos, permitindo detectar que tipos de estímulos e em que fase do ciclo de vida os resultados são mais favoráveis para o desenvolvimento cerebral.

Diante destes avanços, no entendimento do funcionamento cerebral, Cunha e Heckman (2007) desenvolveram um modelo econômico para a formação de habilidades ao longo do ciclo de vida do indivíduo. Este modelo representa através de uma função de produção a dinâmica do aprendizado, enfatizando a importância dos investimentos realizados nos primeiros anos de vida para o sucesso educacional e no mercado de trabalho. O entendimento do processo de formação de capacidades individuais amplia o conhecimento sobre as origens da desigualdade entre pessoas e contribui para o desenho de políticas de redução das desigualdades.

Cunha *et al* (2007) e Cunha e Heckman (2007, 2009) mostram a importância de investimentos em crianças desfavorecidas e que o retorno destes investimentos é alto quando aplicados nos primeiros anos de vida. Cunha e Heckman (2007) apresentam, ainda, evidências e argumentos teóricos de que não existe *trade off* entre equidade e eficiência nos investimentos na primeira infância. Políticas públicas voltadas para a promoção da formação de capital humano devem ser focalizadas em crianças de famílias pobres, devido ao baixo investimento privado destas famílias na formação de habilidades de seus filhos.

A literatura internacional apresenta uma vasta análise de programas governamentais focados no desenvolvimento infantil. Barnett (1992), Currie (2001), Currie e Blau (2007), Cunha (2006) apresentam uma revisão dos principais programas e resultados encontrados. De forma geral os resultados apontam uma relação positiva entre os programas e ganhos de habilidades.

Os Estados Unidos se destacam por alguns programas implementados através de seleção aleatória de seus participantes, visando a obter avaliações mais precisas de seus resultados. O *High/Scope Perry PreSchool* foi implementado durante os anos de 1962 e 1967 em cerca de 120 crianças de baixa renda em idade entre 3 e 4 anos. As crianças foram divididas em grupos de tratamento e controle. O grupo de tratamento recebeu meio dia de pré-escola, visitas semanais durante oito meses por ano ao longo de dois anos. O programa coletou informações das crianças dos dois grupos dos 3 ao 15 anos e depois com 19, 27 e 40 anos. Os resultados da avaliação mostraram que o grupo tratado apresentou baixa taxa de repetência, alto desempenho em testes cognitivos, altas taxas de conclusão do ensino fundamental, além de menor envolvimento em crime, altos salários e baixa utilização de programas de apoio governamental na idade adulta.

O *Carolina Abecedarian Project* selecionou um grupo de 111 crianças de famílias em situação de risco, com pais de baixa escolaridade, baixos salários e alto nível de problemas de saúde. Foram selecionadas crianças nos primeiros meses de vida e separadas aleatoriamente em dois grupos. O grupo tratado teve atenção intensiva por oito horas por dia, cinco dias na semana até completarem 5 anos. Quando as crianças completaram 5 anos houve outro sorteio para a formação de outro grupo de tratamento até os 8 anos. No final do programa havia, portanto, quatro grupos. Um grupo que recebeu o tratamento nos dois períodos, um grupo que recebeu o tratamento no primeiro período mas não no segundo, outro grupo que não recebeu tratamento no primeiro período mas recebeu no segundo e um grupo que não recebeu tratamento em nenhum período. Esta divisão em quatro grupos permitiu avaliar se o tratamento recebido no primeiro período continua fazendo efeito mesmo se interrompido.

Estes dois programas tem em comum o fato de serem de baixa escala, o que não permite a generalização de seus resultados. Um programa de maior escala foi o *Head Start*, que teve início em 1965. O programa teve como alvo crianças de 3 a 5 anos, garantindo acesso a pré-escola a crianças desfavorecidas. Os resultados das avaliações mostraram efeitos positivos no ingresso à universidade e menores taxas de criminalidade.

Para a América Latina destacam-se os programas Oportunidades do México, *Proyecto Integral de Desarrollo Infantil* da Bolívia e a análise da construção de pré-escolas na Argentina. Para o programa Oportunidades, os resultados indicam uma melhora nas atividades motoras das crianças, aumento na probabilidade de matrícula na pré-escola e expectativa de anos de estudos para o grupo de tratados (Gertler e Fernald, 2004; Behrman, Parker e Todd, 2004; Schady, 2006). A ampliação de oferta de pré-escola na Argentina contribui para a melhora no desempenho em língua espanhola para as crianças beneficiadas (Berlinsk *et al.*, 2006).

Para o Brasil, os estudos que avaliam efeitos de intervenções na primeira infância focam na análise do efeito da pré-escola nos níveis de proficiência dos alunos no ensino fundamental e médio e também em resultados do mercado de trabalho.

Curi e Menezes-Filho (2009) estimam o efeito da pré-escola na conclusão dos ciclos escolares, no desempenho escolar medido por testes de proficiência e nos salários futuros. Utilizam dados da PPV e do SAEB 2003. Estes autores utilizaram os estimadores MQO e *Probit* e encontram uma relação positiva entre pré-escola e resultados futuros, embora não significativos em todas as especificações.

Felício e Vasconcellos (2007) avaliam o efeito da pré-escola sobre a proficiência em matemática dos alunos da 4ª série das escolas de rede pública usando dados do SAEB 2003 e Prova Brasil 2005. Através de uma análise de *matching* para os dados individuais do SAEB 2003 e de efeitos fixos para um painel de escolas com dados do SAEB 2003 e Prova Brasil 2005, as autoras concluem que a pré-escola tem efeitos positivos sobre a proficiência.

Calderini e Souza (2009), com base em dados da Prova Brasil 2005, utilizam uma análise por variáveis instrumentais para avaliar o efeito da pré-escola sobre a proficiência em matemática e português dos alunos da 4ª série e encontram resultados positivos nessa relação.

Felício, Menezes e Zoghbi (2009) analisam a relação entre idade de entrada na escola e medidas de alfabetização nas crianças da 2ª série do município de Sertãozinho-SP utilizando dados da Provinha Brasil. Através de uma análise por estimadores de *matching* concluíram que alunos que ingressaram com 5 anos de idade ou menos na escola obtiveram notas maiores em relação às crianças que entram na escola com 6 anos ou mais. Outro resultado que merece destaque é o fato de não terem sido encontradas diferenças significativas entre alunos que frequentaram um ano de Educação Infantil e alunos que frequentaram mais anos.

Pazello e Almeida (2010) utilizaram dados das PNAD's de 1982 a 2007 e trabalham com dados agregados de sucessivas gerações educacionais, relacionando a proporção de

indivíduos de uma determinada geração que frequentou pré-escola e o seu desempenho escolar futuro. As autoras concluem que ingressar mais cedo na escola tem efeitos positivos sobre a conclusão do ensino fundamental, mas não tem efeitos significativos sobre a conclusão dos ensinos médio e superior. O quadro mostra um resumo dos principais resultados encontrados, bem como o banco de dados utilizado e método aplicado.

Quadro 3.1 – Resumo dos principais resultados da literatura nacional

Autores	Base de Dados	Tipo de estimação	Principais resultados
Curi e Menezes-Filho (2009)	SAEB 2003	Mínimos Quadrados Ordinários	Resultados positivos mas não significativos
Felício e Vasconcellos (2007)	SAEB 2003 e Prova Brasil 2005	<i>Matching</i> e efeitos fixos para painel de escolas	Resultados positivos
Calderini e Souza (2009)	Prova Brasil 2005	Variáveis Instrumentais	Resultados positivos
Felício <i>et al</i> (2009)	Provinha Brasil	<i>matching</i>	Resultados positivos
Rodrigues <i>et al</i> (2010)	SAEB 2005	Variáveis Instrumentais	Resultados positivos

Fonte: elaboração própria.

3.3 - Banco de Dados e Variáveis

Neste capítulo, utilizam-se os dados provenientes do GERES, uma pesquisa longitudinal realizada para avaliar a qualidade e a equidade no ensino fundamental brasileiro. O projeto foi coordenado por seis instituições de ensino superior brasileiras: Universidade Federal da Bahia – UFBA, Universidade Federal de Juiz de Fora – UFJF, Pontifícia Universidade Católica do Rio de Janeiro – PUC Rio, Universidade Federal de Minas Gerais – UFMG e Universidade de Campinas – Unicamp, Universidade Estadual de Mato Grosso do Sul - UEMS e contou com financiamento da Fundação Ford e do PRONEX — Programa Núcleo de Excelência/CNPq.

O GERES avaliou trezentas e três escolas de cinco cidades brasileiras ao longo de quatro anos. As cidades avaliadas foram Salvador (BA), Belo Horizonte (MG), Rio de Janeiro (RJ), Campinas (SP) e Campo Grande (MS). Uma amostra de alunos foi acompanhada ao longo destes quatro anos e foram aplicadas cinco ondas de testes com estes alunos. Os alunos selecionados estavam cursando, no primeiro ano do projeto, o segundo ano do ensino fundamental ou a série equivalente quando o sistema de ensino ainda era com um ensino fundamental de oito anos. No início do primeiro ano foi aplicado um teste com o intuito de servir como linha de base e ajuste para os demais testes. Após esta primeira onda, os testes foram aplicados no final de cada ano. Em função de mudanças de escolas não foi possível

acompanhar todos os alunos inicialmente cadastrados. O município de Salvador só participou das quatro primeiras ondas.

Apesar do caráter longitudinal do banco de dados obtido pelo Geres, os questionários não foram aplicados em todos os anos, com exceção do questionário de professores. Em função disso, não é possível acompanhar a evolução das características das escolas e nem do nível socioeconômico das famílias. Esta deficiência do banco de dados prejudica o acompanhamento da evolução dos investimentos, tanto em insumos escolares quanto em insumos familiares, nas crianças. Visando minimizar esta deficiência, buscaram-se informações sobre as escolas no Censo Escolar, realizado anualmente pelo Instituto Nacional de Estudos e Pesquisas Anísio Teixeira – Inep. Uma variável considerada importante na análise do aprendizado é o salário do professor. Esta informação foi obtida na RAIS-Migra²⁰. Através do CNPJ das escolas foi possível buscar o salário médio dos professores de cada escola. Para as escolas públicas a média de salário corresponde à média por município e rede de ensino.

A seleção das variáveis utilizadas como insumos na função de produção educacional foi baseada em diversos estudos sobre o efeito de características dos professores e das escolas reunidos no *site* da Fundação Itaú Social, em particular no projeto Fatores Associados ao Sucesso Educacional. O quadro apresenta 2 as variáveis utilizadas e o banco de dados de origem.

Quadro 3.2 – Relação da Variáveis utilizadas

Variáveis	Banco de dados de origem/Ano
Proficiência nas disciplinas matemática e português	Geres/2005-2008
Variáveis de alunos e família	
Raça	Geres/2008
Nível socioeconômico	Geres/2008
Variáveis de Professor	
Escolaridade e experiência	Geres/2005-2008
Salário do Professor	Rais-Migra/2005-2008
Variáveis de Escola	
Existência de laboratório de informática, laboratório de ciências e biblioteca, e acesso a internet	Censo Escolar/2005-2008
Rede de ensino – Pública/Privada	Geres/2005-2008

Fonte: elaboração própria.

²⁰ Banco de dados administrativo do Ministério do Trabalho e Emprego com características longitudinais dos trabalhadores formalmente empregados no mercado de trabalho brasileiro.

3.4 - Estratégia Empírica

A estimação da função de produção educacional considerando a teoria do desenvolvimento infantil demanda informações sobre toda a história de investimentos na criança, desde os primeiros anos até o período escolar, além de uma medida da habilidade inata da criança. Em função da limitação dos bancos de dados existentes em não possuir todas estas informações, diversas estratégias vêm sendo adotadas na literatura. Todd e Wolpin (2003) fazem uma revisão dos principais métodos utilizados, bem como a necessidade de informações e as hipóteses de identificação necessárias para cada método obter estimativas consistentes. Esta seção apresenta brevemente o modelo geral utilizado em Todd e Wolpin (2003) e a estratégia empírica utilizada neste capítulo.

A equação (1) apresenta a forma geral, onde A_{ia} representa o nível de habilidades do indivíduo i na idade a , $F_i(a)$ representa a história de investimentos da família até a idade a , $S_i(a)$ os insumos escolares até a idade a , u_{i0} indica a habilidade inata, adquirida no momento do nascimento, e f_a é uma função que relaciona os insumos para a formação de habilidades. Na forma geral esta função pode variar com a idade.

$$A_{ia} = f_a \left[F_i(a), S_i(a), u_{i0} \right] \quad (1)$$

Para simplificar a exposição da estratégia, os insumos escolares e familiares serão representados por um único vetor $X = (F, S)$.

Considerando que a tecnologia de formação de habilidades é a mesma para todas as idades, $f_a = f$. E considerando que f é uma função linear e aditiva separável em seus argumentos, (1) pode ser reescrita conforme (2):

$$A_{ia} = X_{ia} \alpha_1^a + X_{i,a-1} \alpha_2^a + X_{i,a-2} \alpha_3^a + \dots + X_{i1} \alpha_a^a + \beta_a u_{i0} + \varepsilon_i(a) \quad (2)$$

onde α representa o efeito de X no nível de habilidades, o sobrescrito indica o efeito do insumo na idade a e o subscrito indica a distância entre o período onde o insumo foi aplicado e o período em que se está medindo a proficiência²¹. De acordo com Todd e Wolpin (2003) a equação (2) representa o verdadeiro modelo.

²¹ O subscrito 1 indica o efeito do insumo aplicado ao longo do ano na proficiência medida ao final do ano. O subscrito 2 indica o efeito do ano anterior e assim por seguinte.

As principais dificuldades para a estimação de (2) são: i) escassez de bancos de dados contendo informações sobre toda a história de investimentos na criança; ii) inexistência de uma medida da habilidade inata adquirida no momento do nascimento; iii) possível correlação entre a habilidade inata e o nível dos investimentos.

A correlação apontada no item (iii) se verifica se considerarmos o comportamento otimizador das famílias. Nesse cenário, a escolha do nível de investimento na formação de habilidades das crianças depende do nível de habilidades inata.

Considerando especificamente o caso da pré-escola, a matrícula na educação infantil não é obrigatória no Brasil, esta decisão é de livre escolha dos pais. Este fato pode produzir uma correlação entre as habilidades não observadas da criança e a frequência à pré-escola. Rodrigues *et al.* (2010) apontam quatro possíveis canais onde a correlação entre a decisão de matricular na pré-escola e o valor nos testes de proficiência pode ocorrer; (i) se as famílias que matriculam seus filhos na pré-escola forem as famílias que valorizam mais o investimento em capital humano, as crianças que fizerem pré-escola irão ter maiores estímulos familiares para estudar e portanto poderão obter maiores níveis de proficiência não em função da pré-escola; (ii) é possível que os pais com maiores custos de oportunidade para o tempo decidam matricular seus filhos na pré-escola, para explorar suas vantagens no mercado de trabalho, considerando que algumas destas vantagens são importantes na formação de capital humano da criança (educação dos pais); (iii) se os pais que matriculam as crianças na pré-escola são aqueles com menores habilidades para cuidar de seus filhos, é possível que parte da diferença entre os níveis de proficiência das crianças que fizeram pré-escola e das que não fizeram seja em parte pela inabilidade dos pais em promoverem a formação de capital humano dos filhos; e (iv) é possível que as crianças com maiores dificuldades de desenvolvimento é que sejam matriculadas nas pré-escolas, o que fariam com que as crianças com piores habilidades fossem matriculadas na pré-escola. Se ocorrer uma dessas hipóteses, a simples diferença entre os níveis de proficiência das crianças que fizeram pré-escola para as que não fizeram não irão refletir o efeito causal da pré-escola. Nos dois primeiros casos, esta diferença estará superestimada enquanto que nos dois últimos estará subestimada.

Para lidar com as dificuldades de se estimar (2) diretamente, algumas estratégias vêm sendo propostas na literatura (Todd e Wolpin, 2003). A primeira estratégia é ignorar a dinâmica do aprendizado e considerar que apenas os insumos contemporâneos determinam o aprendizado, esta estratégia é conhecida como especificação contemporânea. A segunda estratégia, conhecida como valor adicionado, é possível de ser aplicada quando existe pelo

menos uma medida defasada da proficiência. Esta medida defasada capta a história de investimentos.

A estratégia mais simples que se pode adotar em termos de quantidade de informações é a especificação contemporânea, no entanto, é a que exige um maior número de restrições nas hipóteses de identificação. Esta especificação considera apenas os insumos aplicados durante o ano em que a proficiência está sendo medida e adota as seguintes hipóteses: i) apenas os insumos contemporâneos importam, ou os insumos passados não estão correlacionados com os insumos presentes; ii) os insumos contemporâneos não são correlacionados com a habilidade inata não observada. A especificação contemporânea assume a forma definida na equação 3:

$$A_{ia} = X_{ia}\alpha + \varepsilon_i(a) \quad (3)$$

Nesta especificação o resíduo inclui todos os insumos passados, a habilidade inata e erros de medida. Se as hipóteses (i) e (ii) definidas no parágrafo anterior forem satisfeitas a estimação por MQO produzirá estimativas consistentes.

Em função da limitação de banco de dados contendo toda a história de investimentos e de evidências de que as hipóteses da especificação contemporânea não se sustentam, diversos estudos passaram a utilizar uma estratégia conhecida como valor adicionado. Para esta estratégia, é necessária a existência de uma medida defasada da proficiência que se assume captar toda a história passada dos investimentos. A equação (4) apresenta a forma funcional da especificação de valor adicionado. Conforme pode ser notado comparando-se as equações (3) e (4), a única diferença do valor adicionado para a especificação contemporânea é a presença da proficiência defasada.

$$A_{ia} = \gamma A_{i,a-1} + X_{ia}\alpha + \varepsilon_{ia} \quad (4)$$

Apesar desta estimação considerar que o processo de aprendizado é um fenômeno dinâmico, e, conseqüentemente se aproximar do verdadeiro modelo de formação de habilidades, a estimação deste modelo por MQO exige que não haja a presença de habilidades não observadas, mesmo que as habilidades não observadas não sejam correlacionadas com o nível de insumos do vetor X . Se houver a presença de habilidades não observadas na formação de habilidades, estas habilidades serão correlacionadas com a medida defasada de proficiência.

Apresentadas as limitações das especificações contemporânea e de valor adicionado, Todd e Wolpin (2003) passam a expor as propriedades da estimação do modelo cumulativo. Se a escolha dos insumos não for correlacionada com a habilidade inata não observada, a estimação de (2) por OLS produz estimativas consistentes. No entanto, se a escolha dos pais e da escola for correlacionada com o nível de habilidades as estimativas serão inconsistentes.

Para controlar pela correlação entre a habilidade não observada e os insumos, é necessário pelo menos duas observações em diferentes períodos de tempo. Utiliza-se uma estimação de efeitos fixos para expurgar o efeito das habilidades inatas não observadas. Todd e Wolpin (2003) detalham as hipóteses necessárias para obter estimativas consistentes utilizando a estimação do modelo cumulativo com efeitos fixos.

O banco do Geres não possui informações sobre toda a história de investimentos, possui informação sobre os quatro primeiros anos do período escolar e uma avaliação logo no início do período escolar, avaliação diagnóstica. Esta avaliação é utilizada para suprir a ausência de informações de investimentos no período anterior ao período escolar. Portanto, considerando as informações disponíveis no banco do Geres e as estratégias apresentadas anteriormente, neste capítulo é adotada uma estratégia mista, com a especificação cumulativa e do valor adicionado.

Define-se a proficiência medida na quinta onda (último ano avaliado no Geres) como A_{ia} . Observa-se, portanto o vetor $A = (A_{ia}, A_{i,a-1}, A_{i,a-2}, A_{i,a-3}, A_{i,a-4})$ e o vetor $X = (X_{ia}, X_{i,a-1}, X_{i,a-2}, X_{i,a-3})$. Além destes dois vetores de informação, também é possível saber se a criança fez pré-escola. A pré-escola é um insumo aplicado antes de $X_{i,a-3}$. Portanto, $A_{i,a-4}$ capta a informação do vetor $X' = (X_{i,a-4}, X_{i,a-5}, \dots, X_{i1})$ e se a criança fez pré-escola. Logo, assume-se que o efeito do vetor X' na proficiência, medida durante o período escolar, se manifesta através de $A_{i,a-4}$. Esta é uma adaptação da hipótese de completude dinâmica. Em termos formais,

$$E[A_{ia} | X_{ia}, X_{i,a-1}, \dots, X_{i1}, u_{i0}] = E[A_{ia} | X_{ia}, X_{i,a-1}, X_{i,a-2}, X_{i,a-3}, A_{i,a-4}, u_{i0}]$$

Considerando os quatro anos do ensino fundamental em que as crianças são acompanhadas no Geres, o modelo de formação de habilidades para cada ano pode ser visualizado através do sistema de equações definido em (5).

$$\begin{aligned} A_{ia} &= X_{ia}\alpha_1^a + X_{i,a-1}\alpha_2^a + X_{i,a-2}\alpha_3^a + X_{i,a-3}\alpha_4^a + \gamma^a A_{i,a-4} + \beta^a u_{i0} + \varepsilon_{ia} \\ A_{i,a-1} &= X_{i,a-1}\alpha_1^{a-1} + X_{i,a-2}\alpha_2^{a-1} + X_{i,a-3}\alpha_3^{a-1} + \gamma^{a-1} A_{i,a-4} + \beta^{a-1} u_{i0} + \varepsilon_{i,a-1} \\ A_{i,a-2} &= X_{i,a-2}\alpha_1^{a-2} + X_{i,a-3}\alpha_2^{a-2} + \gamma^{a-2} A_{i,a-4} + \beta^{a-2} u_{i0} + \varepsilon_{i,a-2} \\ A_{i,a-3} &= X_{i,a-3}\alpha_1^{a-3} + \gamma^{a-3} A_{i,a-4} + \beta^{a-3} u_{i0} + \varepsilon_{i,a-3} \end{aligned} \quad (5)$$

De acordo com o apresentado em (5), o insumo pré-escola teria seu efeito apenas sobre o nível de proficiência da avaliação diagnóstica e, portanto não seria possível identificá-lo com esta estratégia. No entanto, estudos demonstram que o fato da criança fazer pré-escola aumenta as capacidades cognitivas e não cognitivas e estas capacidades continuam tendo efeitos no aprendizado. Por exemplo, o aumento de habilidades não cognitivas como motivação e/ou concentração pode fazer com que o aprendizado das crianças no período escolar seja maior em relação ao aprendizado de crianças que não fizeram pré-escola. A nova hipótese de completude dinâmica pode ser escrita conforme (6).

$$E[A_{ia} | X_{ia}, X_{i,a-1}, \dots, X_{i1}, \text{Pre}_i, u_{i0}] = E[A_{ia} | X_{ia}, X_{i,a-1}, X_{i,a-2}, X_{i,a-3}, A_{i,a-4}, \text{Pre}_i, u_{i0}] \quad (6)$$

Com essa nova hipótese definida, o sistema de equações definido em (5) passa a ser escrito pelo novo sistema definido em (7).

$$\begin{aligned} A_{ia} &= X_{ia}\alpha_1^a + X_{i,a-1}\alpha_2^a + X_{i,a-2}\alpha_3^a + X_{i,a-3}\alpha_4^a + \gamma^a A_{i,a-4} + \delta^a \text{Pre}_i + \beta^a u_{i0} + \varepsilon_{ia} \\ A_{i,a-1} &= X_{i,a-1}\alpha_1^{a-1} + X_{i,a-2}\alpha_2^{a-1} + X_{i,a-3}\alpha_3^{a-1} + \gamma^{a-1} A_{i,a-4} + \delta^{a-1} \text{Pre}_i + \beta^{a-1} u_{i0} + \varepsilon_{i,a-1} \\ A_{i,a-2} &= X_{i,a-2}\alpha_1^{a-2} + X_{i,a-3}\alpha_2^{a-2} + \gamma^{a-2} A_{i,a-4} + \delta^{a-2} \text{Pre}_i + \beta^{a-2} u_{i0} + \varepsilon_{i,a-2} \\ A_{i,a-3} &= X_{i,a-3}\alpha_1^{a-3} + \gamma^{a-3} A_{i,a-4} + \delta^{a-3} \text{Pre}_i + \beta^{a-3} u_{i0} + \varepsilon_{i,a-3} \end{aligned} \quad (7)$$

Com o sistema de equações definido em (7) é possível estimar o efeito dos insumos aplicados durante o período escolar, o efeito do nível de habilidades inicial da criança e o efeito direto da pré-escola em cada ano que a criança é avaliada. No entanto, o fato da criança ter feito pré-escola pode aumentar o nível de habilidades em que a criança inicia o ensino fundamental, ou seja, parte do efeito da pré-escola é captado pela medida de proficiência na avaliação diagnóstica, efeito indireto. Portanto, o efeito obtido com esta estratégia pode estar subestimando o verdadeiro efeito.

Todd e Wolpin (2003) adotam a hipótese de que o efeito dos insumos não varia com a idade, ou seja, $\alpha_x^a = \alpha_x^{a'}$. Esta simplificação reduz o número de parâmetros. Neste capítulo esta hipótese é relaxada, permitindo que os insumos tenham efeitos diferentes em cada idade. Dessa maneira, pode-se verificar, por exemplo, se um “bom professor” tem um impacto no aprendizado de uma criança na 1ª série diferente do impacto no aprendizado na 4ª série.

Para aplicar esta estratégia é realizada uma interação entre os insumos e uma *dummy* indicativa de qual onda se está medindo a proficiência. Portanto, por exemplo, α^l mede o

efeito de um insumo aplicado na onda a sobre a proficiência na onda a , α^{a-1}_1 mede o efeito do mesmo insumo aplicado na onda $a-1$ sobre a proficiência da onda $a-1$ e α^a_2 mede o efeito de um insumo aplicado na onda $a-1$ sobre a proficiência medida na onda a . O sistema definido em (8) apresenta a modificação em (7) com a inclusão da interação com a *dummy* de onda.

A estimação do sistema apresentado em (8) por MQO produzirá estimativas inconsistentes se a habilidade inata não observada for correlacionada com os demais insumos. Para controlar por essa possível endogeneidade é utilizada a estratégia de efeitos fixos, através do uso do estimador de primeira diferença. Para expurgar o efeito das habilidades observadas do sistema definido em (8) é necessário assumir a hipótese que seus efeitos não variam com a idade, ou seja, ($\beta^a = \beta^{a-1}$), neste caso a diferenciação elimina a habilidade não observada.

$$\begin{aligned}
A_{ia} &= (X_{ia} * O_a)\alpha_1^a + (X_{i,a-1} * O_a)\alpha_2^a + (X_{i,a-2} * O_a)\alpha_3^a + (X_{i,a-3} * O_a)\alpha_4^a + \gamma^a(A_{i,a-4} * O_a) + \\
&\quad + \delta^a(\text{Pre}_i * O_a) + \beta^a u_{i0} + \varepsilon_{ia} \\
A_{i,a-1} &= (X_{i,a-1} * O_{a-1})\alpha_1^{a-1} + (X_{i,a-2} * O_{a-1})\alpha_2^{a-1} + (X_{i,a-3} * O_{a-1})\alpha_3^{a-1} + \gamma^{a-1}(A_{i,a-4} * O_{a-1}) + \\
&\quad + \delta^{a-1}(\text{Pre}_i * O_{a-1}) + \beta^{a-1} u_{i0} + \varepsilon_{i,a-1} \\
A_{i,a-2} &= (X_{i,a-2} * O_{a-2})\alpha_1^{a-2} + (X_{i,a-3} * O_{a-2})\alpha_2^{a-2} + \gamma^{a-2}(A_{i,a-4} * O_{a-2}) + \\
&\quad + \delta^{a-2}(\text{Pre}_i * O_{a-2}) + \beta^{a-2} u_{i0} + \varepsilon_{i,a-2} \\
A_{i,a-3} &= (X_{i,a-3} * O_{a-3})\alpha_1^{a-3} + \gamma^{a-3}(A_{i,a-4} * O_{a-3}) + \delta^{a-3}(\text{Pre}_i * O_{a-3}) + \beta^{a-3} u_{i0} + \varepsilon_{i,a-3}
\end{aligned} \tag{8}$$

3.5 – Estatísticas descritivas

Esta seção apresenta algumas estatísticas descritivas dos dados do Geres. A tabela 3.3 apresenta a média e o desvio padrão das variáveis utilizadas em separado tanto para os alunos que fizeram pré-escola quanto para os alunos que não fizeram.

Pode-se notar que os alunos que fizeram pré-escola têm em média um nível maior de habilidades durante os quatro primeiros anos do período escolar. Em matemática a média de proficiência é de 192,4 para os alunos que fizeram pré-escola e 166,4 para os que não fizeram, uma diferença de 26 pontos na escala de proficiência, ou 0,37 desvio padrão. Para português os valores médios são de 151,8 e 139,4, para quem fez pré-escola e para quem não fez, respectivamente, o que corresponde a uma diferença de 0,42 desvio padrão. Esta diferença não pode ser interpretada com efeito da pré-escola em função da não aleatoriedade na seleção à pré-escola.

Em relação às características dos alunos, percebe-se que aproximadamente 35% dos alunos eram brancos, mas esse percentual de brancos é menor para os que não fizeram, 28%,

ou seja, as crianças brancas tiveram mais acesso a pré-escola. Em relação ao gênero há certa homogeneidade, com uma pequena participação mais elevada das meninas. O nível socioeconômico apresenta uma desigualdade maior entre os dois grupos (frequentaram e não frequentaram pré-escola). As crianças que fizeram pré-escola têm um nível socioeconômico mais alto do que as que não fizeram, o que indica uma seleção positiva em relação à frequência à pré-escola.

As informações dos professores mostram que os professores dos alunos que fizeram pré-escola possuem em média maiores salários, mais tempo de profissão e um percentual maior de professores com ensino superior. Em relação às escolas, 82% dos alunos que fizeram pré-escola ingressaram em escolas públicas e 95% dos que não fizeram pré-escola foram para escolas públicas. Os recursos escolares, laboratório de informática e biblioteca, são bem distribuídos entre os grupos, apenas a existência de laboratório de ciências é mais significativa para o grupo de alunos que fez pré-escola, 30%, contra 17% para os que não fizeram. A última linha da tabela mostra que 87% das crianças da amostra fizeram pré-escola.

Tabela 3.1 – Médias e Desvios-Padrão

Variáveis	Pré-Escola (A)	Sem Pré-Escola (B)	Diferença (A – B)
Proficiência em Matemática	152.002 (29.551)	139.460 (28.136)	12.542***
Proficiência em Português	192.718 (69.646)	166.698 (63.965)	26.020***
Nível Socioeconômico	0.164 (0.619)	-0.216 (0.472)	0.380***
Branco	0.349 (0.477)	0.282 (0.450)	0.067***
Raça não declarada	0.006 (0.080)	0.010 (0.102)	-0.004***
Meninos	0.492 (0.500)	0.533 (0.499)	-0.042***
Meninas	0.500 (0.500)	0.455 (0.498)	0.046***
Gênero não declarado	0.008 (0.091)	0.012 (0.110)	-0.004**
Escola Pública	0.813 (0.390)	0.948 (0.221)	-0.136***
Laboratório de Informática	0.723 (0.448)	0.753 (0.431)	-0.031***
Laboratório de Ciências	0.303 (0.460)	0.169 (0.375)	0.134***

(continua...)

Tabela 3.1 – Médias e Desvios-Padrão (*continuação*)

Variáveis	Pré-Escola (A)	Sem Pré-Escola (B)	Diferença (A – B)
Biblioteca	0.739 (0.439)	0.721 (0.448)	0.017**
Experiência do professor - 0 a 5 anos	0.060 (0.237)	0.058 (0.235)	0.001
Experiência do professor - 6 a 10 anos	0.116 (0.321)	0.120 (0.325)	-0.004
Experiência do professor – 11 a 15 anos	0.159 (0.366)	0.153 (0.360)	0.006
Experiência do professor – mais de 15 anos	0.519 (0.500)	0.468 (0.499)	0.052***
Experiência do professor – não declarada	0.146 (0.353)	0.201 (0.401)	-0.056***
Professor com ensino superior	0.706 (0.456)	0.665 (0.472)	0.041***
Salário dos professores	2,160.15 (1,288.73)	2,065.95 (1,027.83)	94.197***
Salário dos professores – não declarado	0.024 (0.152)	0.012 (0.110)	0.011***
Belo Horizonte	0.283 (0.451)	0.302 (0.459)	-0.019**
Campinas	0.301 (0.459)	0.244 (0.430)	0.057***
Campo Grande	0.124 (0.330)	0.297 (0.457)	-0.173***
Rio de Janeiro	0.292 (0.454)	0.156 (0.363)	0.135***

Fonte: Geres (2005-2008)

Nota1: erros padrão em parênteses.

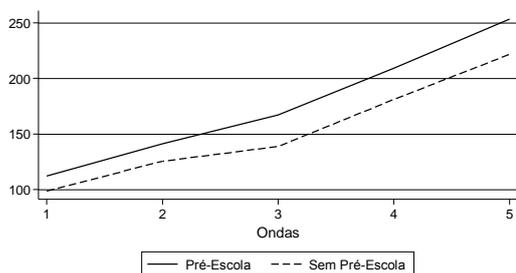
Nota 2: ***Significante a 1%. **Significante a 5%. *Significante a 10%.

De uma forma geral, os resultados apresentados na tabela 3.3 indicam que os alunos que fizeram pré-escola tiveram um maior nível de investimentos durante os primeiros quatro anos do ensino fundamental, tanto nos insumos escolares, quanto nos insumos privados, considerando o nível socioeconômico das famílias como *proxy*.

O gráfico 3.5 mostra a relação entre frequência da criança à pré-escola e escolaridade da mãe. A relação é estritamente crescente, ou seja, quanto maior a escolaridade da mãe, maior o percentual de filhos matriculados na pré-escola. Considerando as evidências de um efeito positivo da pré-escola no aprendizado, este fato implica na geração de um círculo

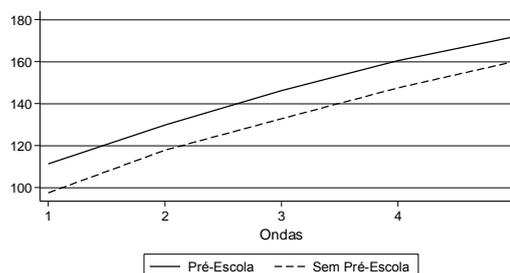
vicioso, onde filhos de pais com baixo nível de escolaridade terão uma probabilidade menor de atingir sucesso escolar.

Gráfico 3.3 - Proficiência em Matemática (2005-2008)



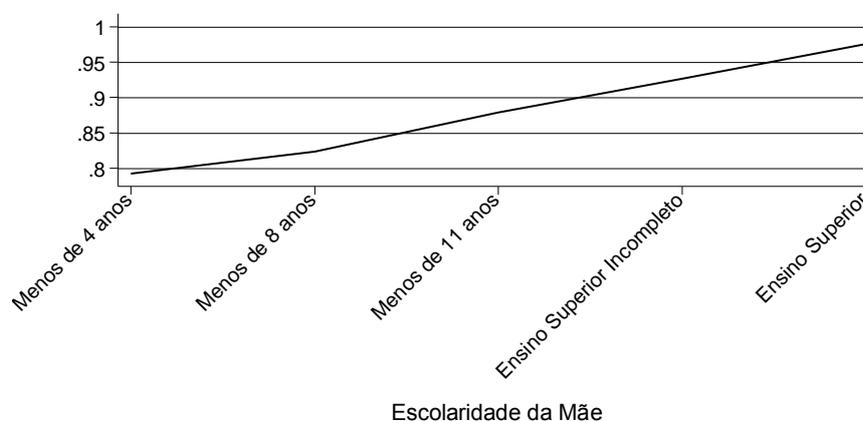
Fonte: Geres(2005-2008)

Gráfico 3.4 - Proficiência em Português (2005-2008)



Fonte: Geres(2005-2008)

Gráfico 3.5 - Escolaridade da Mãe e Matrícula na Pré-Escola



Fonte: Geres(2005-2008)

3.6 – Resultados

Esta seção apresenta os resultados econométricos do capítulo. São apresentados primeiro os resultados para as estimações do efeito da pré-escola na proficiência em matemática e em seguida para a proficiência em português. A apresentação para cada disciplina seguirá a seguinte ordem. Uma primeira estimação exploratória com a especificação contemporânea por Mínimos Quadrados Ordinários separadamente para cada onda. Depois são apresentadas as estimações cumulativas sem considerar o valor adicionado e considerando o valor adicionado. Estas especificações são estimadas pelo estimador MQO e Primeiras Diferenças.

Para cada especificação testada são utilizados quatro modelos com a inclusão gradativa de controles. O *modelo 1* inclui como controle apenas características da criança e da família, mais precisamente raça do aluno e nível socioeconômico da família. No *modelo 2* acrescenta-se a rede de ensino da escola, se é pública ou privada. No *modelo 3* são utilizadas características de infraestrutura da escola, existência de laboratório de informática, laboratório de ciências, biblioteca e acesso a internet. E por fim, no *modelo 4* acrescentam-se variáveis com características dos professores, experiência, salário e escolaridade.

A tabela 3.2 apresenta os primeiros resultados para a proficiência em matemática. São reportados apenas os coeficientes da variável de interesse, pré-escola. Esta primeira especificação mostra separadamente o efeito da pré-escola em cada onda. Nota-se que o efeito na onda 5 é maior que na onda 2, indicando que o efeito da pré-escola persiste durante os primeiros anos do ensino fundamental e que a distância na proficiência entre quem fez pré-escola e quem não fez aumenta.

Tabela 3.2 – MQO – Proficiência em Matemática

Variável dependente – Proficiência em Matemática	Onda 2	Onda 3	Onda 4	Onda 5
Modelo 1				
Pré-Escola	6.573*** (0.979)	12.055*** (1.615)	10.650*** (1.799)	13.090*** (1.956)
Observações	8,476	8,476	8,476	8,476
R-quadrado	0.219	0.248	0.229	0.228
Modelo 2				
Pré-Escola	6.472*** (0.965)	11.866*** (1.592)	10.495*** (1.783)	12.940*** (1.938)
Observações	8,476	8,476	8,472	8,476
R-quadrado	0.240	0.269	0.243	0.242
Modelo 3				
Pré-Escola	6.019*** (0.949)	10.803*** (1.555)	9.803*** (1.784)	12.229*** (1.927)
Observações	8,452	8,381	8,079	8,334
R-quadrado	0.274	0.312	0.272	0.254
Modelo 4				
Pré-Escola	5.158*** (0.941)	10.076*** (1.544)	9.088*** (1.772)	9.906*** (1.888)
Observações	8,452	8,381	8,079	8,334
R-quadrado	0.290	0.323	0.283	0.289

Nota 1: Modelo 1: raça e nível socioeconômico; Modelo 2: modelo 1 + escola pública/privada; Modelo 3: modelo 2 + laboratório de informática, laboratório de ciências, biblioteca, internet; Modelo 4: modelo 3 + salário, escolaridade e experiência dos professores.

Nota 2: ***Significante a 1%. **Significante a 5%. *Significante a 10%.

Nota 3: Erros padrão em parênteses.

Comparando-se os quatro modelos, nota-se que a magnitude do coeficiente de pré-escola vai diminuindo à medida que se inclui controles relacionados à escola e aos professores. Este fato indica que existe uma correlação positiva entre a variável de pré-escola e os insumos escolares, ou seja, as crianças que fizeram pré-escola formam um grupo que recebe melhores insumos escolares. Para as estimativas nesta especificação serem consistentes é necessário assumir que não existe correlação entre os insumos presentes e os insumos passados e nem entre os insumos presentes e a habilidade inata não observada.

A tabela 3.3 mostra os primeiros resultados para a especificação cumulativa. De uma forma geral, os coeficientes apresentam menor magnitude em relação aos coeficientes apresentados na tabela 3.2. Isto demonstra que a hipótese necessária para a estimação pela especificação contemporânea de não correlação entre insumos presentes e insumos passados não se verifica, ou seja, a formação de conhecimento é um processo dinâmico, conforme apontado pela Teoria do Ciclo de Vida da Formação de Habilidades de Cunha e Heckman (2007).

Tabela 3.3 – Especificação Cumulativa - MQO sem Valor Adicionado – Proficiência em Matemática

Variável dependente – Proficiência em Matemática	(1)	(2)	(3)	(4)
Pré-escola*onda2	6.573*** (1.630)	6.472*** (1.613)	5.992*** (1.577)	5.131*** (1.539)
Pré-escola*onda3	12.055*** (1.630)	11.881*** (1.613)	10.518*** (1.578)	9.498*** (1.542)
Pré-escola*onda4	10.650*** (1.630)	10.495*** (1.613)	9.245*** (1.580)	8.285*** (1.544)
Pré-escola*onda5	13.090*** (1.630)	12.944*** (1.613)	11.299*** (1.580)	7.811*** (1.548)
Observações	33,904	33,904	33,904	33,904
R-quadrado	0.513	0.523	0.549	0.572

Nota 1: Modelo 1: raça e nível socioeconômico; Modelo 2: modelo 1 + escola pública/privada; Modelo 3: modelo 2 + laboratório de informática, laboratório de ciências, biblioteca, internet; Modelo 4: modelo 3 + salário, escolaridade e experiência dos professores.

Nota 2: ***Significante a 1%. **Significante a 5%. *Significante a 10%.

Nota 3: Erros padrão em parênteses.

A tabela 3.4 mostra os resultados para a especificação mista, cumulativa com valor adicionado. A inclusão de proficiência da avaliação diagnóstica tem o objetivo de captar o efeito dos insumos aplicados antes do período escolar e não observados no banco de dados do GERES. Para esta especificação obter estimativas consistentes é necessário que não haja efeito de habilidades não observadas na formação de habilidades.

Tabela 3.4 – Especificação Cumulativa - MQO com Valor Adicionado – Proficiência em Matemática

Variável dependente – Proficiência em Matemática	(1)	(2)	(3)	(4)
Pré-escola*onda2	2.641** (1.342)	2.661** (1.338)	2.378* (1.324)	2.148 (1.310)
Pré-escola*onda3	6.066*** (1.342)	6.097*** (1.338)	5.258*** (1.325)	4.828*** (1.313)
Pré-escola*onda4	4.258*** (1.342)	4.276*** (1.338)	3.517*** (1.327)	3.289** (1.315)
Pré-escola*onda5	6.578*** (1.342)	6.624*** (1.338)	5.247*** (1.327)	3.413*** (1.317)
Proficiência inicial*onda2	0.753*** (0.017)	0.741*** (0.018)	0.726*** (0.018)	0.718*** (0.018)
Proficiência inicial*onda3	1.147*** (0.017)	1.124*** (0.018)	1.086*** (0.018)	1.069*** (0.018)
Proficiência inicial*onda4	1.225*** (0.017)	1.208*** (0.018)	1.166*** (0.018)	1.141*** (0.018)
Proficiência inicial*onda5	1.248*** (0.017)	1.227*** (0.018)	1.219*** (0.018)	1.140*** (0.018)
Observações	33,904	33,904	33,904	33,904
R-quadrado	0.671	0.673	0.683	0.691

Nota 1: Modelo 1: raça e nível socioeconômico; Modelo 2: modelo 1 + escola pública/privada; Modelo 3: modelo 2 + laboratório de informática, laboratório de ciências, biblioteca, internet; Modelo 4: modelo 3 + salário, escolaridade e experiência dos professores.

Nota 2: ***Significante a 1%. **Significante a 5%. *Significante a 10%.

Nota 3: Erros padrão em parênteses.

Os resultados encontrados apresentam menor magnitude em relação aos apresentados na tabela 3.3. Esta queda de magnitude indica uma relação positiva entre os insumos passados, não observados, e os insumos aplicados durante o período escolar.

Para controlar por possíveis correlações entre o nível de investimentos e a habilidade inata não observada as próximas duas especificações são aplicações baseadas em estimadores de primeira diferença. Nas estimações por Primeira Diferença é necessário omitir uma das *dummies* interativas entre pré-escola e onda e também entre proficiência inicial e onda. Foram omitidas as interações com a onda 2 nos dois casos. Portanto, os coeficientes das demais interações devem ser analisados como a diferença em relação à *dummy* omitida. Na tabela 3.5, os resultados mostram que o efeito da pré-escola sobre a proficiência em matemática é maior nas ondas três, quatro e cinco do que em relação ao efeito na onda dois.

A tabela 3.6 apresenta os resultados para a estimação de efeitos fixos com valor adicionado. Focando a análise no modelo (4), nota-se que apenas o coeficiente da onda três é significativo, ou seja, os alunos que fizeram pré-escola possuem um ganho de habilidades maior do que os alunos que não fizeram pré-escola até a onda três, depois a diferença entre os dois grupos se estabiliza.

Tabela 3.5 – Especificação Cumulativa – Primeira Diferença sem Valor Adicionado – Proficiência em Matemática

Variável dependente – Proficiência em Matemática	(1)	(2)	(3)	(4)
Pré-escola*onda2			omitida	
Pré-escola*onda3	5.482*** (1.163)	5.399*** (1.162)	4.742*** (1.161)	4.296*** (1.161)
Pré-escola*onda4	4.078** (1.645)	4.011** (1.643)	3.356** (1.642)	3.130* (1.643)
Pré-escola*onda5	6.517*** (2.015)	6.426*** (2.012)	5.351*** (2.012)	4.392** (2.014)
Observações	25,428	25,428	25,428	25,428
R-quadrado	0.553	0.554	0.560	0.563

Nota 1: Modelo 1: raça e nível socioeconômico; Modelo 2: modelo 1 + escola pública/privada; Modelo 3: modelo 2 + laboratório de informática, laboratório de ciências, biblioteca, internet; Modelo 4: modelo 3 + salário, escolaridade e experiência dos professores.

Nota 2: ***Significante a 1%. **Significante a 5%. *Significante a 10%.

Nota 3: Erros padrão em parênteses.

Tabela 3.6 – Especificação Cumulativa – Primeira Diferença com Valor Adicionado – Proficiência em Matemática

Variável dependente – Proficiência em Matemática	(1)	(2)	(3)	(4)
Pré-escola*onda2			omitida	
Pré-escola*onda3	3.425*** (1.150)	3.427*** (1.149)	2.982*** (1.150)	2.748** (1.150)
Pré-escola*onda4	1.617 (1.626)	1.606 (1.625)	1.189 (1.627)	1.171 (1.628)
Pré-escola*onda5	3.936** (1.992)	3.924** (1.991)	2.938 (1.993)	2.271 (1.996)
Proficiência inicial*onda2			omitida	
Proficiência inicial*onda3	0.394*** (0.015)	0.383*** (0.015)	0.363*** (0.015)	0.354*** (0.016)
Proficiência inicial*onda4	0.472*** (0.021)	0.467*** (0.021)	0.446*** (0.022)	0.448*** (0.022)
Proficiência inicial*onda5	0.494*** (0.026)	0.486*** (0.026)	0.496*** (0.027)	0.490*** (0.028)
Observações	25,428	25,428	25,428	25,428
R-quadrado	0.565	0.566	0.570	0.573

Nota 1: Modelo 1: raça e nível socioeconômico; Modelo 2: modelo 1 + escola pública/privada; Modelo 3: modelo 2 + laboratório de informática, laboratório de ciências, biblioteca, internet; Modelo 4: modelo 3 + salário, escolaridade e experiência dos professores.

Nota 2: ***Significante a 1%. **Significante a 5%. *Significante a 10%.

Nota 3: Erros padrão em parênteses.

Comparando os resultados da tabela 3.6, especificação cumulativa com valor adicionado e efeitos fixos, com os resultados da tabela 3.5, especificação cumulativa com efeitos fixos e sem valor adicionado, nota-se que a inclusão da proficiência inicial como variável independente fez com que os coeficientes das *dummies* interativas de pré-escola reduzissem sua magnitude na onda dois e perdesse a significância nas ondas três e quatro. Isto

indica que não dá para assumir a hipótese de não correlação entre os insumos aplicados no período escolar e os insumos investidos antes da entrada da criança na escola, e nem a hipótese de não correlação entre a decisão dos pais em matricular seus filhos na pré-escola e os outros investimentos dos pais na formação de habilidades de seus filhos antes deles entrarem na escola.

Os coeficientes das variáveis interativas entre proficiência inicial e onda de aplicação do teste, com valores positivos e significativos, mostram que o desenvolvimento no aprendizado em matemática durante o período escolar depende do nível de habilidades que a criança inicia no ensino fundamental. Os coeficientes aumentam de magnitude ao longo dos anos, o que está de acordo com o modelo desenvolvido por Cunha e Heckman (2007). Neste modelo o aprendizado em uma etapa depende do nível de aprendizado na etapa anterior. Isto implica que a desigualdade de habilidades em matemática das crianças tende a aumentar no período escolar.

As próximas tabelas apresentam o resultado para a proficiência em português. A tabela 3.7 mostra os resultados para a estimação por MQO da especificação contemporânea para os quatro anos do projeto Geres. Nesta estimação, o efeito da pré-escola é positivo e significativo para todas as ondas analisadas e para todos os modelos. O mesmo ocorreu para a proficiência em matemática, no entanto, a magnitude do efeito é menor para português e com valores próximos em todas as ondas, em matemática o efeito aumentava ao longo dos anos. Da mesma forma do que para matemática, a inclusão de controles reduz o efeito da pré-escola, o que indica a correlação entre o nível de insumos aplicados durante o período escolar e a variável de pré-escola.

A tabela 3.8 apresenta os resultados para a especificação cumulativa sem valor adicionado estimado por MQO. Os mesmos comentários realizados para os resultados da tabela 3.7 são válidos. Da mesma forma que para matemática, os resultados são positivos e significativos, mas a magnitude é inferior e mais estável ao longo dos anos.

Para que os resultados apresentados na tabela 3.8 sejam consistentes, os insumos investidos antes da entrada da criança na escolar devem ser ortogonais aos insumos investidos no período escolar. Com a estimação utilizando a estratégia do valor adicionado é possível testar esta hipótese. Se os coeficientes da proficiência inicial, que capta a informação sobre os investimentos no período anterior à escola, forem significativos, não é possível assumir a hipótese de ortogonalidade entre insumos do período escolar e os insumos passados. A tabela 3.9 mostra que os coeficientes são positivos e significativos. Portanto, parte do efeito da pré-

escola apresentado na tabela 3.8 é de fato efeito das habilidades em que a criança ingressa na escola.

Tabela 3.7 – MQO – Proficiência em Português

Variável dependente – Proficiência em Português	Onda 2	Onda 3	Onda 4	Onda 5
Modelo 1				
Pré-escola	5.201*** (0.654)	5.753*** (0.786)	5.898*** (0.781)	5.206*** (0.746)
Observações	8,476	8,476	8,476	8,476
R-quadrado	0.259	0.234	0.210	0.202
Modelo 2				
Pré-escola	5.123*** (0.642)	5.670*** (0.777)	5.831*** (0.774)	5.158*** (0.741)
Observações	8,476	8,476	8,472	8,476
R-quadrado	0.286	0.251	0.224	0.213
Modelo 3				
Pré-escola	4.853*** (0.627)	5.063*** (0.761)	5.479*** (0.781)	4.947*** (0.738)
Observações	8,452	8,381	8,079	8,334
R-quadrado	0.327	0.292	0.244	0.225
Modelo 4				
Pré-escola	4.312*** (0.623)	4.765*** (0.757)	5.179*** (0.776)	3.910*** (0.716)
Observações	8,452	8,381	8,079	8,334
R-quadrado	0.340	0.300	0.256	0.275

Nota 1: Modelo 1: raça e nível socioeconômico; Modelo 2: modelo 1 + escola pública/privada; Modelo 3: modelo 2 + laboratório de informática, laboratório de ciências, biblioteca, internet; Modelo 4: modelo 3 + salário, escolaridade e experiência dos professores.

Nota 2: ***Significante a 1%. **Significante a 5%. *Significante a 10%.

Nota 3: Erros padrão em parênteses.

A tabela 3.10 mostra os resultados para estimação controlando por efeitos fixos, sem valor adicionado. Em todos os modelos especificados e para todas as ondas os efeitos são não significativos.

A tabela 3.11 apresenta os resultados para a estimação do valor adicionado controlando por efeitos não observados. Novamente os resultados das *dummies* de pré-escola foram não significativos. Vale destacar os resultados para a proficiência inicial. De forma diferente do que aconteceu com os resultados da tabela 3.6 para a proficiência em matemática, o efeito da proficiência inicial diminui ao longo dos anos, se tornando inclusive negativo nas ondas 4 e 5²². Este resultado mostra que para português o nível de habilidades com que a

²² No modelo 4 o coeficiente da onda 4 embora negativo foi não significativo.

criança inicia a escola não determina o sucesso escolar, ou seja, mesmo havendo uma desigualdade de habilidades em português no momento em que as crianças iniciam o ensino fundamental, esta desigualdade tende a diminuir se os insumos escolares forem distribuídos de forma eficiente.

Tabela 3.8– Especificação Cumulativa – MQO sem Valor Adicionado – Proficiência em Português

Variável dependente – Proficiência em Português	(1)	(2)	(3)	(4)
Pré-escola*onda2	5.201*** (0.744)	5.123*** (0.736)	4.875*** (0.720)	4.330*** (0.705)
Pré-escola*onda3	5.753*** (0.744)	5.676*** (0.736)	5.033*** (0.720)	4.745*** (0.706)
Pré-escola*onda4	5.898*** (0.744)	5.827*** (0.736)	5.497*** (0.721)	5.048*** (0.707)
Pré-escola*onda5	5.206*** (0.744)	5.165*** (0.736)	4.697*** (0.721)	3.379*** (0.709)
Observações	33,904	33,904	33,904	33,904
R-quadrado	0.444	0.456	0.484	0.507

Nota 1: Modelo 1: raça e nível socioeconômico; Modelo 2: modelo 1 + escola pública/privada; Modelo 3: modelo 2 + laboratório de informática, laboratório de ciências, biblioteca, internet; Modelo 4: modelo 3 + salário, escolaridade e experiência dos professores.

Nota 2: ***Significante a 1%. **Significante a 5%. *Significante a 10%.

Nota 3: Erros padrão em parênteses.

Tabela 3.9 – Especificação Cumulativa – MQO com Valor Adicionado – Proficiência em Português

Variável dependente – Proficiência em Português	(1)	(2)	(3)	(4)
Pré-escola*onda2	1.501** (0.597)	1.537** (0.597)	1.573*** (0.596)	1.378** (0.590)
Pré-escola*onda3	1.802*** (0.597)	1.824*** (0.597)	1.601*** (0.596)	1.458** (0.591)
Pré-escola*onda4	2.363*** (0.597)	2.378*** (0.597)	2.384*** (0.597)	2.048*** (0.592)
Pré-escola*onda5	2.043*** (0.597)	2.062*** (0.597)	1.829*** (0.597)	1.069* (0.593)
Proficiência inicial*onda2	0.606*** (0.009)	0.596*** (0.009)	0.582*** (0.009)	0.576*** (0.009)
Proficiência inicial*onda3	0.647*** (0.009)	0.640*** (0.009)	0.624*** (0.009)	0.616*** (0.009)
Proficiência inicial*onda4	0.579*** (0.009)	0.573*** (0.009)	0.562*** (0.009)	0.556*** (0.009)
Proficiência inicial*onda5	0.518*** (0.009)	0.515*** (0.009)	0.513*** (0.009)	0.481*** (0.010)
Observações	33,904	33,904	33,904	33,904
R-quadrado	0.644	0.645	0.649	0.657

Nota 1: Modelo 1: raça e nível socioeconômico; Modelo 2: modelo 1 + escola pública/privada; Modelo 3: modelo 2 + laboratório de informática, laboratório de ciências, biblioteca, internet; Modelo 4: modelo 3 + salário, escolaridade e experiência dos professores.

Nota 2: ***Significante a 1%. **Significante a 5%. *Significante a 10%.

Nota 3: Erros padrão em parênteses.

Tabela 3.10 – Especificação Cumulativa – Primeira Diferença sem Valor Adicionado – Proficiência em Português

Variável dependente – proficiência em português	(1)	(2)	(3)	(4)
Pré-escola*onda2			omitida	
	0.552	0.542	0.272	0.294
Pré-escola*onda3	(0.500)	(0.500)	(0.502)	(0.502)
	0.697	0.692	0.684	0.622
Pré-escola*onda4	(0.708)	(0.707)	(0.710)	(0.711)
	0.004	0.015	-0.149	-0.251
Pré-escola*onda5	(0.867)	(0.866)	(0.869)	(0.872)
Observações	25,428	25,428	25,428	25,428
R-quadrado	0.476	0.477	0.480	0.482

Nota 1: Modelo 1: raça e nível socioeconômico; Modelo 2: modelo 1 + escola pública/privada; Modelo 3: modelo 2 + laboratório de informática, laboratório de ciências, biblioteca, internet; Modelo 4: modelo 3 + salário, escolaridade e experiência dos professores.

Nota 2: ***Significante a 1%. **Significante a 5%. *Significante a 10%.

Nota 3: Erros padrão em parênteses.

Tabela 3.11 – Especificação Cumulativa – Primeira Diferença com Valor Adicionado – Proficiência em Português

Variável dependente – proficiência em português	(1)	(2)	(3)	(4)
Pré-escola*onda2			omitida	
	0.301	0.276	0.055	0.083
Pré-escola*onda3	(0.500)	(0.500)	(0.502)	(0.503)
	0.862	0.830	0.809	0.698
Pré-escola*onda4	(0.708)	(0.708)	(0.710)	(0.712)
	0.542	0.503	0.270	0.093
Pré-escola*onda5	(0.867)	(0.867)	(0.870)	(0.872)
Proficiência inicial*onda2			omitida	
	0.041***	0.044***	0.039***	0.040***
Proficiência inicial*onda3	(0.007)	(0.007)	(0.008)	(0.008)
	-0.027***	-0.023**	-0.022**	-0.014
Proficiência inicial*onda4	(0.010)	(0.010)	(0.011)	(0.011)
	-0.088***	-0.081***	-0.075***	-0.069***
Proficiência inicial*onda5	(0.012)	(0.013)	(0.013)	(0.014)
Observações	25,428	25,428	25,428	25,428
R-quadrado	0.480	0.481	0.483	0.485

Nota 1: Modelo 1: raça e nível socioeconômico; Modelo 2: modelo 1 + escola pública/privada; Modelo 3: modelo 2 + laboratório de informática, laboratório de ciências, biblioteca, internet; Modelo 4: modelo 3 + salário, escolaridade e experiência dos professores.

Nota 2: ***Significante a 1%. **Significante a 5%. *Significante a 10%.

Nota 3: Erros padrão em parênteses.

3.7 – Considerações finais

O objetivo do capítulo foi estimar o efeito de frequentar pré-escola no nível de proficiência em matemática e português. Para atingir o objetivo foi utilizado um banco de dados longitudinal que acompanhou uma amostra de crianças de quatro municípios brasileiros ao longo dos quatro primeiros anos do ensino fundamental. As crianças avaliadas foram as que em 2005, primeiro ano do projeto, frequentavam na 1ª série do ensino fundamental.

O efeito da pré-escola foi estimado através de uma função de produção educacional considerando o processo dinâmico de aprendizado. A estratégia empírica utilizada foi uma adaptação da estimação por valor adicionado, para considerar o efeito cumulativo dos insumos e a presença de heterogeneidades não observadas. Para verificar como as hipóteses utilizadas nas estratégias mais restritivas, especificação contemporânea e valor adicionado, não se sustentam, foram apresentados os resultados para todas as especificações de forma gradual.

Os resultados encontrados indicam que a especificação mais adequada é a que considera o efeito cumulativo dos insumos e controla pela heterogeneidade individual não observada. O efeito da pré-escola difere entre as proficiências de matemática e português. Para matemática o efeito direto encontrado foi positivo e significativo. Em relação ao efeito indireto, manifestado através da proficiência inicial, o efeito encontrado além de positivo, é crescente ao longo dos anos.

Já para português, os coeficientes de pré-escola apresentaram resultados não significativos, indicando que não há efeito direto sobre as habilidades em português. Em relação aos efeitos indiretos, os alunos com habilidades mais elevadas no início do ensino fundamental ampliam suas diferenças em relação aos alunos menos habilidosos, mas este efeito se reverte ao longo dos anos, indicando que a desigualdade de habilidades em português pode ser reduzida durante o período escolar.

Para a formulação de políticas públicas de redução das desigualdades e aumento do nível de escolaridade da população, os resultados do capítulo corroboram as evidências da necessidade de investimentos na primeira infância, principalmente para a elevação das habilidades em matemática. Considerando que a frequência à pré-escola é maior para crianças de famílias de nível socioeconômico mais elevado, uma política de educação infantil voltada para famílias mais pobres pode contribuir para diminuir o efeito do ciclo vicioso de formação de habilidades, reduzindo o efeito da origem social sobre o aprendizado das crianças.

3.8 – Referências

ALMEIDA, R. B.; PAZELLO, E. T. O efeito da pré-escola sobre o desempenho escolar futuro dos indivíduos. XXXVIII Encontro Nacional de Economia – ANPEC, 2010.

BARNETT, W. S. Benefits of Compensatory Preschool Education. **Journal of Human Resources**, v.27, n.2, p.279-312, 1992.

BARNETT, W. S. Long-Term Effects of Early Childhood Programs on Cognitive and School Outcomes. **Future of Children**, v.5, n.3, 1995.

BEHRMAN, J.; PARKER, S. W.; TODD, P. E. Medium Term Effects of the Oportunidades Program Package, Including Nutrition, on Education of Rural Children Age 0-8 in 1997. Technical Document, n.9 on the Evaluation of Oportunidades, 2004

BERLINSKI, S.; GALIANI, S.; MANACORDA, M. Giving children a better start: Preschool attendance and school-age profiles. **Journal of Public Economics**, v.92, p.1416-1440, 2008.

BERLINSKI, S.; GALIANI, S.; GERTLER, P. The effect of pre-primary education on primary school performance. **Journal of Public Economics**, v.93, p.219-234, 2009.

BLAU, D.; CURRIE, J. Preschool, daycare, and afterschool care: Who's minding the kids? In E. Hanushek, & F. Welch (Eds.), **Handbook of the Economics of Education**, v.2, capítulo 20, Amsterdam: North-Holland, p.1163-1278, 2006

CALDERINI, S. R.; SOUZA, A. P. Pré-escola no Brasil: Seu Impacto na Qualidade da Educação Fundamental. XXXVII Encontro Nacional de Economia – ANPEC, 2009.

CASELLA, E. B.; AMARO JR, E.; COSTA, J. C. (2011). As Bases Neurológicas da Aprendizagem da Leitura. In: **Aprendizagem Infantil: Uma abordagem da neurociência, economia e psicologia cognitiva**. Coord. Araújo, A. Rio de Janeiro, 2011.

CUNHA, F.; HECKMAN, J. The Technology of Skill Formation. **American Economic Review**, v.97, n.2, p.31–47, 2007.

CUNHA, F.; HECKMAN, J. The economics and psychology of inequality and human development. NBER Working Paper n. 14695. 2009.

CUNHA, F.; HECKMAN, J. Investing in Our Young People. NBER Working Paper n.16201, 2010.

CUNHA, F.; HECKMAN, J.; LOCHNER, L. J.; MASTEROV, D. V. Interpreting the evidence on life cycle skill formation. In E. A. Hanushek, & F. Welch (Eds.), **Handbook of the Economics of Education**, capítulo 12. Amsterdam: North-Holland, p.697-812, 2006

CUNHA, F.; HECKMAN, J.; SCHENNACH, S. M. Estimating the technology of cognitive and noncognitive skill formation. **Econometrica**, v.78, n.3, p.883-931, 2010.

CURI, A.; MENEZES-FILHO, N. A. A Relação entre Educação Pré-Primária, Salários, Escolaridade e Proficiência Escolar no Brasil. **Estudos Econômicos**, v.39, n.4, p.811-850, 2009.

CURRIE, J. Early Childhood Education Programs. **Journal of Economic Perspectives**, American Economic Association, v.15, n.2, p.213-238, 2001.

FELÍCIO, F.; VASCONCELLOS, L. O Efeito Da Educação Infantil Sobre O Desempenho Escolar Medido Em Exames Padronizados. XXXV Encontro Nacional de Economia – ANPEC, 2007.

FELICIO, F.; TERRA, R.; ZOGHBI, A. C. (2012). The effects of early childhood education on literacy scores using data from a new Brazilian assessment tool. **Estudos Econômicos**, v.42, n.1, 2012.

IBGE - Instituto Brasileiro de Geografia e Estatística. Pesquisa Nacional por Amostra de Domicílios – PNAD, 2001-2009. Rio de Janeiro.

GERES. Estudo Longitudinal sobre a Qualidade e Equidade no Ensino Fundamental Brasileiro, 2005.

GERTLER, P.; FERNALD, L. The Medium Term Impact of Oportunidades on Child Development in Rural Areas. Não publicado, University of California at Berkeley, 2004.

KNUDSEN, E. I. Sensitive Periods in the Development of the Brain and Behavior. **Journal of Cognitive Neuroscience**, v.16, n.1, p.1412-1425, 2004.

PISA – Programme for International Student Assessment, 2009.

RODRIGUES, C.G; PINTO, C. X. C.; SANTOS, D. D. The impact of daycare attendance on mat test scores for a cohort of 4 graders in brazil. Working paper, 2010.

TODD. P. E.; WOLPIN, K. I. On The Specification and Estimation of The Production Function for Cognitive Achievement. **Economic Journal**, v.113, p. 3-33, 2003

SCHADY, N. Early Childhood Development in Latin America and the Caribbean. **Economía**, v.6, n.2, p.185-213, 2006

4 – CONCLUSÃO

Esta dissertação teve como objetivo investigar os determinantes da desigualdade educacional brasileira. A motivação para estudar a desigualdade educacional surge a partir dos estudos que apontam o capital humano como principal fator explicativo da desigualdade de renda e do fato do Brasil apresentar umas das maiores desigualdade de renda do mundo. Dois aspectos da desigualdade são examinados. O primeiro aspecto está relacionado ao momento do ciclo de vida em que as desigualdades são geradas. O segundo se refere à divisão dos determinantes da desigualdade em fatores sob o controle do indivíduo, considerados justos, e fatores fora do controle individual, considerados injustos.

O primeiro capítulo teve o objetivo de realizar uma revisão da literatura abordando a discussão sobre a relação entre educação e desigualdade de renda, evidenciando os principais artigos internacionais e nacionais sobre o tema. Para isso, usou-se a decomposição do índice de Theil-T e o IOH. Os resultados do Theil, com a finalidade de verificar a contribuição da educação na desigualdade de rendimentos apontaram que, apesar da contribuição da educação na desigualdade de rendimentos ter se reduzido, juntamente com a própria desigualdade de renda, durante a década de 2000, essa contribuição ficou acima dos 30% em todo o período. Os resultados do IOH, que tem a finalidade de medir as oportunidades básicas das crianças e adolescentes até 16 anos, mostraram uma melhora das oportunidades básicas durante a década passada, acompanhada por uma redução da desigualdade de acesso a estas oportunidades. O IOH é um índice composto por cinco subíndices, sendo a probabilidade de concluir a sexta série na idade correta o subíndice que melhor reflete a qualidade da educação. Em função disso, também foram apresentados os valores deste subíndice isoladamente. Os resultados indicam uma baixa oportunidade para as crianças completarem a sexta série na idade correta. O comportamento ao longo da década de 2000 mostra uma pequena melhora das oportunidades e uma diminuição da dependência entre conseguir completar a sexta série na idade correta e origem familiar.

Diante dessas evidências iniciais, de uma relação positiva entre educação e desigualdade de renda e baixas oportunidades educacionais no Brasil, os capítulos 2 e 3, ensaios independentes, analisam o papel da escola e da pré-escola, respectivamente, sobre o aprendizado dos alunos durante os quatro primeiros anos do ensino fundamental.

O segundo capítulo investigou o papel da escola na formação de habilidades e no processo de geração/redução das desigualdades. Este capítulo utilizou o modelo desenvolvido

em Cunha e Heckman (2007) que mostra que o aprendizado é um processo dinâmico, ou seja, o nível de habilidades de um período contribui para o aprendizado nos períodos subsequentes. Os autores mostram a importância de investimentos na formação de habilidades das crianças antes mesmo da entrada na escola, principalmente até os 4 anos de idade.

Os resultados do capítulo mostraram a dependência que o aprendizado durante o período escolar tem em relação ao nível de habilidades com que a criança inicia na escola, limitando o papel da escola na redução de desigualdades de habilidades. Também são apresentadas evidências que o nível de habilidades das crianças tem uma dependência em relação à origem familiar, o que indica uma desigualdade de oportunidades na formação de habilidades.

Dado que o capítulo 2 mostrou a importância das habilidades com que a criança inicia a escola, o capítulo 3 investigou a relação entre a pré-escola e o nível de proficiência dos alunos. Foi utilizada uma adaptação da técnica de valor adicionado para controlar por possíveis habilidades fixas não observadas e incluir o efeito cumulativo dos insumos. Os resultados indicaram uma relação positiva e persistente ao longo dos anos para a proficiência em matemática e uma relação positiva para português, porém não persistente ao longo dos anos.

Como contribuição para a literatura sobre o assunto e para a formulação de políticas públicas para a redução das desigualdades, a dissertação fortalece os argumentos para a melhoria do nível educacional da população brasileira. O capítulo 2 mostrou que, apesar dos insumos escolares contribuírem para o aprendizado dos alunos, e, portanto, devem ser alvo de políticas, a baixa habilidade que alguns alunos iniciam na escola limita o papel da escola na redução das desigualdades. Considerando a necessidade de se elevar o nível de habilidades que as crianças iniciam no período escolar obrigatório, o capítulo 3 mostrou que a pré-escola tem efeitos positivos, principalmente sobre as habilidades em matemática.

ANEXO I

O índice Theil-T (T) tem como característica permitir a decomposição da desigualdade em fatores entre e intra grupos. O Theil-intra (T_{Ig}) indica o quanto da desigualdade dentro de determinado grupo contribui para a desigualdade geral. Já o Theil-entre (T_{Eg}) indica a concentração de renda dentro do grupo, se o valor for positivo, significa que o grupo detém uma parcela da renda maior do que sua participação relativa na população e caso contrário se for negativo. Para cada grupo definido é calculado, portanto, um valor para o Theil-intra e para o Theil-entre. A soma do Theil-intra com o Theil-entre de todos os grupos apresenta o mesmo valor para o Theil-T da população como um todo.

A equação (1) apresenta a fórmula do índice de Theil-T para a desigualdade geral.

$$T = \sum_{i=1}^N \left[y_i \cdot \ln \left(y_i \cdot \frac{N}{Y} \right) \right] \quad (1)$$

onde y_i é a renda do indivíduo i , $Y = \sum_{i=1}^N y_i$ é a renda total da população, N é o total da população.

As equações (2) e (2a) apresentam a fórmula para o Theil-intra para cada sub-grupo definido e o para Theil-intra agregado, respectivamente. O Theil-intra agregado informa quanto da desigualdade geral é explicada pela desigualdade dentro dos grupos.

$$T_{Ig} = \frac{Y_g}{Y} \cdot \sum_{i \in g} \left[\frac{y_i}{Y_g} \cdot \ln \left(y_i \cdot \frac{N_g}{Y_g} \right) \right] \quad (2)$$

$$T_I = \sum_{g=1}^G T_{Ig} \quad (2a)$$

onde $Y_g = \sum_{i \in g} y_i$ é a renda agregada do grupo g , N_g é a população do grupo g com $g = \{1, 2, \dots, G\}$, sendo G o número de sub-grupos definido.

As equações (3) e (3a) apresentam a fórmula para o Theil-entre para cada sub-grupo definido e o para Theil-entre agregado, respectivamente. O Theil-entre agregado, de maneira semelhante ao Theil-intra agregado, informa quanto da desigualdade geral é explicada pela desigualdade entre os grupos. Através do Theil-entre agregado é possível calcular a taxa

bruta, que mostra, em termos percentuais, o quanto da desigualdade total é explicada por diferenças entre os grupos. A equação (4) define a taxa bruta.

$$T_{Eg} = \frac{Y_g}{Y} \cdot \ln \left(\frac{Y_g}{Y} \cdot \frac{N}{N_g} \right) \quad (3)$$

$$T_E = \sum_{g=1}^G T_{Eg} \quad (3a)$$

$$\text{Taxa bruta} = \frac{T_E}{T} \quad (4)$$

O índice de Theil para a desigualdade total pode ser recomposto através da soma dos elementos desagregados, conforme a equação (5):

$$T = \sum_{g=1}^G T_{Ig} + \sum_{g=1}^G T_{Eg} \quad (5)$$

Tabela AI.1 – Decomposição do Theil-T para grupos educacionais – PNAD's 2001 a 2009

	2001	2002	2003	2004	2005	2006	2007	2008	2009
Theil	0.643	0.633	0.615	0.602	0.603	0.598	0.565	0.555	0.553
Taxa Bruta	0.359	0.363	0.347	0.340	0.335	0.329	0.319	0.308	0.309
Sem instrução									
Elemento-Intra	0.370	0.375	0.363	0.349	0.357	0.328	0.366	0.365	0.360
Theil-Intra	0.014	0.013	0.012	0.012	0.012	0.009	0.011	0.011	0.010
Theil-Entre	-0.037	-0.035	-0.033	-0.032	-0.030	-0.027	-0.026	-0.026	-0.024
Primeiro ciclo do fundamental incompleto									
Elemento-Intra	0.431	0.377	0.418	0.365	0.353	0.383	0.390	0.362	0.337
Theil-Intra	0.026	0.022	0.023	0.020	0.018	0.018	0.018	0.016	0.014
Theil-Entre	-0.043	-0.041	-0.038	-0.036	-0.035	-0.033	-0.031	-0.028	-0.028
Primeiro ciclo do fundamental completo									
Elemento-Intra	0.435	0.433	0.403	0.427	0.436	0.393	0.379	0.361	0.357
Theil-Intra	0.036	0.035	0.031	0.033	0.032	0.027	0.025	0.022	0.021
Theil-Entre	-0.033	-0.033	-0.031	-0.029	-0.029	-0.029	-0.027	-0.026	-0.026
Segundo ciclo do fundamental incompleto									
Elemento-Intra	0.360	0.385	0.372	0.374	0.354	0.394	0.373	0.383	0.317
Theil-Intra	0.036	0.039	0.036	0.036	0.033	0.036	0.033	0.033	0.026
Theil-Entre	-0.047	-0.047	-0.045	-0.043	-0.043	-0.042	-0.040	-0.038	-0.039
Segundo ciclo do fundamental completo									
Elemento-Intra	0.424	0.385	0.419	0.390	0.384	0.444	0.376	0.389	0.349
Theil-Intra	0.037	0.031	0.035	0.031	0.030	0.034	0.032	0.031	0.026
Theil-Entre	-0.013	-0.017	-0.017	-0.019	-0.019	-0.018	-0.021	-0.021	-0.020
Ensino médio incompleto									
Elemento-Intra	0.390	0.496	0.395	0.388	0.377	0.376	0.342	0.335	0.320
Theil-Intra	0.019	0.023	0.018	0.018	0.017	0.017	0.015	0.016	0.015
Theil-Entre	-0.014	-0.015	-0.017	-0.018	-0.018	-0.019	-0.018	-0.020	-0.019
Ensino médio completo									
Elemento-Intra	0.427	0.423	0.427	0.398	0.397	0.398	0.355	0.366	0.349
Theil-Intra	0.098	0.099	0.104	0.100	0.102	0.104	0.093	0.098	0.095
Theil-Entre	0.038	0.030	0.024	0.020	0.012	0.007	0.000	-0.005	-0.008
Superior incompleto									
Elemento-Intra	0.439	0.419	0.389	0.385	0.432	0.403	0.364	0.388	0.356
Theil-Intra	0.033	0.033	0.030	0.030	0.036	0.034	0.030	0.033	0.030
Theil-Entre	0.040	0.040	0.035	0.035	0.038	0.034	0.026	0.027	0.023
Superior completo									
Elemento-Intra	0.409	0.382	0.391	0.417	0.426	0.411	0.433	0.416	0.467
Theil-Intra	0.112	0.107	0.110	0.116	0.120	0.120	0.127	0.123	0.144
Theil-Entre	0.344	0.350	0.338	0.329	0.328	0.328	0.319	0.308	0.313
Não especificado									
Elemento-Intra	0.291	0.292	0.330	0.265	0.253	0.315	0.269	0.242	0.390
Theil-Intra	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Theil-Entre	-0.002	-0.002	-0.002	-0.002	-0.002	-0.001	-0.001	-0.001	-0.001

Fonte: elaboração própria.

ANEXO II

Barros *et al.* (2008) desenvolveram um índice para medir a desigualdade de oportunidades nos serviços básicos para as crianças. De acordo com os autores, o índice é inspirado na função de bem estar social proposta por Sen (1976). Um processo de desenvolvimento no qual uma sociedade tem como objetivo suprir igualmente as oportunidades básicas requer o acesso às oportunidades básicas pelas crianças como uma meta universal. O Índice de Oportunidades Humanas – IOH – é composto por dois elementos: (i) a taxa de cobertura de um serviço básico, que indica as oportunidades disponíveis; e (ii) pela desigualdade na distribuição destas oportunidades entre grupos definidos com base nas suas circunstâncias exógenas. Portanto, o índice é sensível tanto a um aumento na taxa de cobertura quanto a uma melhora da distribuição das oportunidades condicionada às circunstâncias.

Neste estudo os autores definem oportunidades básicas como um subconjunto de bens e serviços para as crianças, como acesso à educação e serviços de saneamento básico, por exemplo. Os autores ressaltam que a provisão universal de oportunidades básicas para crianças encontra um suporte quase consensual entre as sociedades, em comparação com políticas redistributivas para adultos.

O Índice de Oportunidades Humanas foca na taxa de cobertura e na desigualdade de acesso às oportunidades básicas em função de três fatores principais:

- i) de um ponto de vista empírico, o princípio da igualdade de oportunidades como "nivelar o campo de jogo" pode ser operacionalizado pela mensuração do acesso, para as crianças, aos bens e serviços básicos que são críticos para o pleno desenvolvimento infantil. Para crianças, o acesso define a oportunidade, pois não se pode cobrar das crianças que elas realizem esforços próprios para ter acesso a estes bens e serviços (diferentemente dos adultos);
- ii) de um ponto de vista de políticas públicas, evidências apontam que intervenções para equalizar as oportunidades no início do ciclo de vida do indivíduo são mais eficiente e menos custosas do que intervenções em um período mais avançado;
- iii) focar nas crianças contribui para colocar a discussão sobre desigualdade de oportunidades no centro do debate sobre políticas públicas. Como destacado no Relatório do Desenvolvimento Mundial 2006 (Banco Mundial 2006), no dia do nascimento as crianças não podem ser consideradas responsáveis pelas

circunstâncias de sua família, apesar do fato destas circunstâncias afetarem a vida que estas crianças terão.

Para a construção do índice de desigualdade de oportunidades Barros *et al.* (2008a) assumem as seguintes hipóteses: (i) a relação entre o resultado, circunstâncias e esforço é considerada aditiva separável; (ii) o nível de esforço e as circunstâncias são estocasticamente independentes; (iii) a relação entre esforço e resultado é meritocrática; (iv) todas as circunstâncias que afetam o resultado e estão fora do controle dos indivíduos são observadas. Como ressaltam os próprios autores, estas hipóteses dificilmente se sustentam na realidade, mas devido à ausência de informações para ajustar o índice a estas questões, os autores trabalham com estas limitações e explicitam as consequências de assumir estas hipóteses²³.

Para a elaboração de uma medida para o índice, os autores conceituam a desigualdade de oportunidades como a desigualdade na média dos resultados entre os grupos definidos em termos de suas circunstâncias comuns. Adotando este conceito, é possível estimar a desigualdade de oportunidades por dois tipos de contrafactual. Pode-se definir a desigualdade de oportunidades como a desigualdade de resultados que permanece se toda a desigualdade dentro dos grupos com as mesmas circunstâncias for eliminada (abordagem direta); ou pode ser definida através da diferença entre a desigualdade geral de resultados e o nível que ainda permanece após eliminar a desigualdade entre os grupos (abordagem indireta). Para o caso de resultados discretos, como é o caso do acesso aos serviços básicos, a abordagem direta apresenta benefícios e é, portanto, a abordagem utilizada pelos autores.

O Índice de Oportunidades Humanas é construído em duas etapas. Primeiro é necessário construir uma medida da desigualdade na distribuição dos resultados em função das circunstâncias. Em seguida, calcula-se o IOH utilizando informações da desigualdade na distribuição e da taxa de cobertura.

Para o cálculo do índice de desigualdade de oportunidades é preciso ter informações sobre se os indivíduos têm acesso ou não aos serviços básicos e informações sobre suas circunstâncias.

A fórmula do índice de desigualdades é apresentada a seguir:

$$D = \frac{E|P(I=1|x) - P(I=1)|}{2P(I=1)} \quad (1)$$

²³ Ver Barros *et al.* (2008a) subseção 1.3.

onde $I = 1$ indica se a criança possui acesso a determinado serviço e 0 caso contrário, e x é um vetor de variáveis indicando as circunstâncias das crianças. D é conhecido com índice de dissimilaridade e mede a distancia entre duas distribuições.

Desde que $P(I=1) = E(P(I=1|x))$, pode-se reescrever D da seguinte maneira:

$$D = \frac{E|P(I=1|x) - E(P(I=1|x))|}{2E(P(I=1|x))}, \quad (2)$$

que é a expressão utilizada para estimar D . O procedimento para a estimação de D é realizado em três etapas. Na primeira etapa é estimada a probabilidade condicional de se ter acesso à determinado serviço público. Esta medida de probabilidade pode ser obtida através de uma variedade de procedimentos paramétrico, não-paramétrico ou semi-paramétricos. Os autores utilizam uma regressão logística separável, conforme apresentado a seguir:

$$Ln\left(\frac{P(I=1|x_1, \dots, x_m)}{1 - P(I=1|x_1, \dots, x_m)}\right) = \sum_{k=1}^m h_k(x_k) \quad (3)$$

que é equivalente a expressão:

$$P(I=1|x_1, \dots, x_m) = \frac{\exp\left(\sum_{k=1}^m h_k(x_k)\right)}{1 + \exp\left(\sum_{k=1}^m h_k(x_k)\right)} \quad (4)$$

onde x_k representa um vetor de variáveis com k -dimensões indicando as circunstâncias, ou seja, $x = (x_1, \dots, x_m)$. O segundo passo é, dados os coeficientes estimados na etapa anterior, estimar os valores das probabilidades de ser ter acesso aos serviços básicos condicionado à suas circunstâncias.

$$p = \frac{\exp\left(\beta_0 + \sum_{k=1}^m x_{ki} \beta_k\right)}{1 + \exp\left(\beta_0 + \sum_{k=1}^m x_{ki} \beta_k\right)} \quad (5)$$

O terceiro e último passo consiste em calcular a equação a seguir:

$$\bar{p} = \sum_1^n w_i p_i . \quad (6)$$

onde $w_i = \frac{1}{n}$ ou algum peso amostral.

Após estes três passos, utiliza-se o seguinte estimador para D:

$$D = \frac{1}{2\bar{p}} \sum_{i=1}^n w_i |p_i - \bar{p}| \quad (7)$$

Os autores demonstram que D é um estimador consistente para D.

$\bar{p} = \sum_{i=1}^n w_i p_i$ e p é o valor ajustado da regressão logística onde se calcula a probabilidade da criança ter acesso a determinada oportunidade condicionada às suas circunstâncias.

Com o índice de desigualdade calculado, estima-se a medida geral de oportunidade (r), o IOH, que é dada pela seguinte equação:

$$r = \bar{p}(1 - D) . \quad (8)$$

Barros *et al* (2008) utilizam a metodologia desenvolvida para o IOH para a análise do acesso de crianças até dezesseis anos às seguintes oportunidades básicas: completar a sexta série na idade correta, atendimento escolar nas idades de 10 a 14 anos, acesso a água tratada no domicílio, saneamento e eletricidade no domicílio. Para caracterizar os tipos dos indivíduos os autores utilizaram como controle as seguintes variáveis: gênero, educação dos pais, renda per capita, número de irmãos, presença dos pais no domicílio e área de residência (rural/urbano).

Tabela AII.1 – IOH para o Brasil de 2001 a 2009 (PNAD's)

Índices	2001	2002	2003	2004	2005	2006	2007	2008	2009
Acesso a serviços de eletricidade									
IOH	0.90	0.92	0.92	0.92	0.92	0.93	0.95	0.96	0.97
Dissimilaridade	0.04	0.04	0.03	0.04	0.03	0.03	0.02	0.02	0.01
Taxa de cobertura	0.94	0.95	0.95	0.95	0.95	0.96	0.97	0.98	0.98
Acesso a serviços de saneamento									
IOH	0.46	0.48	0.49	0.49	0.49	0.51	0.55	0.55	0.55
Dissimilaridade	0.20	0.19	0.18	0.19	0.19	0.17	0.16	0.15	0.15
Taxa de cobertura	0.58	0.59	0.60	0.60	0.61	0.62	0.66	0.65	0.65
Acesso a serviços de água tratada									
IOH	0.65	0.67	0.68	0.67	0.67	0.68	0.69	0.70	0.71
Dissimilaridade	0.14	0.13	0.12	0.14	0.14	0.13	0.13	0.12	0.11
Taxa de cobertura	0.76	0.77	0.78	0.77	0.77	0.78	0.79	0.80	0.80
Atendimento escolar para crianças de 10 a 14 anos									
IOH	0.95	0.96	0.96	0.96	0.96	0.96	0.97	0.97	0.97
Dissimilaridade	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Taxa de cobertura	0.96	0.97	0.97	0.97	0.97	0.97	0.97	0.98	0.98
Probabilidade de completar a 6ª série na idade correta									
IOH	0.28	0.31	0.32	0.35	0.37	0.39	0.35	0.35	0.34
Dissimilaridade	0.27	0.24	0.24	0.22	0.20	0.19	0.20	0.18	0.18
Taxa de cobertura	0.38	0.41	0.42	0.44	0.46	0.47	0.44	0.42	0.41

Fonte: elaboração própria.