**Pré-escola no Brasil:**

***Seu Impacto na Qualidade da Educação Fundamental*.**

**Sérgio Ricardo Calderini[[1]](#footnote-1)**

**André Portela Souza [[2]](#footnote-2)**

**Abstract**

*Aiming to obtain more consistent estimators for the preschool effect on proficiency exams, we instrumentalized the variables “attended preschool” and “schooling delay” by the “number of preschool in the municipality”, in the year that the pupil attended the preschool, and the “semester that the child was born”. We found support that the preschool average gains for the Math exam ranges from 17% to 21% (77% - 96% standard deviation gains) and for the Portuguese exam it ranges from 12% to 15% (52% to 61% standard deviation gains). This work also reveals the following score rank: firstly, the best grades were obtained by students that were never retained and entered school at the correct age; secondly, the children that suffered grade retention; and finally, with very similar results, the pupils that were delayed in school progression by events different from grade repetition, and students that were aged appropriate for their grade but entered school early.*

Key-words: Preschool, education, performance, schooling delay.

JEL Classification: I21, I28

**Resumo**

*Com o propósito de obtermos estimadores mais consistentes do efeito da pré-escola nos exames de proficiência da 4ª série, instrumentalizamos as variáveis “ter freqüentado a pré-escola” e “atraso escolar” por duas variáveis instrumentais, o “número de pré-escolas por município” na época da educação infantil do aluno e o fato da “criança ter nascido no segundo semestre do ano”. Encontramos evidências de que o efeito da pré-escola foi equivalente a ganhos em relação a média no exame de matemática entre 17% e 21% (entre 77% e 96% de ganhos em desvios padrões) e ganhos em relação a média no exame de português entre 12% e 15% ( entre 52% e 61% de ganhos em desvios padrões). Nossas análises também revelam que obtiveram melhores resultados nos exames: primeiramente os alunos que estavam na idade escolar correta e que nunca repetiram; seguidos pelos alunos que estavam com idade escolar acima da ideal por repetição; e depois com resultados bem semelhantes os alunos com idade escolar acima da ideal, mas por fatores diferentes da repetição, e os alunos que estavam com a idade escolar correta, mas que entraram antecipadamente.*

Palavras-Chaves: Pré-escola, educação, performance, atraso escolar.

Área ANPEC: **11** - Economia Social e Demografia

# 1. Introdução

Apurando-se os dados sobre os gastos públicos em educação (Tabela 1) observa-se que historicamente a fatia destinada à educação infantil no Brasil é consistentemente a menor parte do orçamento para educação. Essa constatação é agravada ao levarmos em conta os dados da Pesquisa Nacional por Amostra de Domicílios (PNAD) de 2004, que mostra que apenas 40,4% das crianças entre 0 e 6 anos de idade estavam matriculadas em alguma instituição educacional[[3]](#footnote-3).

 **Tabela 1** – Investimento Público em Educação no Brasil em porcentagem do PIB

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ***Ano*** | ***Educação Infantil*** | ***Ensino Fundamental*** | ***Ensino Médio*** | ***Ensino Superior*** |
| ***1ª a 4ª série*** | ***5ª a 8ª série*** |
| *2000* | 0,3 | 1,2 | 1,0 | 0,5 | 0,5 |
| *2001* | 0,3 | 1,2 | 1,1 | 0,6 | 0,6 |
| *2002* | 0,3 | 1,1 | 1,2 | 0,7 | 0,7 |
| *2003* | 0,3 | 1,1 | 1,2 | 0,6 | 0,6 |
| *2004* | 0,4 | 1,3 | 1,1 | 0,5 | 0,5 |

*Fonte: Levantamento feito pelo INEP baseado em dados do INEP/MEC, IPEA/DISOC e IBGE*[[4]](#footnote-4)

Essa defasagem no número de matriculados na educação infantil e a atual alocação de recursos nas diversas fases do sistema de ensino brasileiro nos motivam a verificar se uma nova alocação de recursos, com incremento nos investimentos na pré-escola, não seria justificada ao trazer uma produtividade maior ao sistema de ensino do país como um todo.

Ao avaliamos os efeitos da pré-escola nos exames de proficiência da 4ª série de 2005, através de duas variáveis instrumentais, o número de pré-escolas por município na época da educação infantil do aluno e o fato criança ter nascido no segundo semestre do ano, encontramos evidências de que o efeito da pré-escola foi equivalente a ganhos em relação a média no exame de matemática entre 17% e 21% (entre 77% e 96% de ganhos em desvios padrões) e ganhos em relação a média no exame de português entre 12% e 15% ( entre 52% e 61% de ganhos em desvios padrões). Nossas analises também revelam que obtiveram melhores resultados nos exames os alunos que estavam na idade escolar correta e nunca repetiram, seguidos pelos alunos com idade escolar acima da ideal por repetição e depois com resultados bem semelhantes os alunos com idade escolar acima da ideal, causada por motivo diverso do que a repetição (como, entrada tardia ou abandono de algum período escolar) e os alunos que estavam com a idade escolar correta, mas que entraram antecipadamente e repetiram.

# 2. O aprendizado na primeira infância[[5]](#footnote-5)

Em um trabalho sobre transferência de riqueza de pais para filhos, Becker e Tomes (1986) elaboram um modelo de investimento para o desenvolvimento das habilidades de uma criança em que os investimentos feitos em diferentes estágios do desenvolvimento infantil teriam a mesma produtividade. Essa visão, apesar de ser amplamente utilizada na literatura, implicaria em dizer que se investirmos a mesma quantia na alfabetização de uma criança de 7 anos ou de um adolescente de 16 anos teríamos o mesmo resultado: dois indivíduos igualmente capazes de ler e escrever. Entretanto, sabemos que simplesmente pelo fato dos alunos estarem em estágios diferentes de desenvolvimento os resultados esperados são distintos, pois eles possuem diferentes níveis de dificuldade para o aprendizado.

Portanto, usando uma abordagem diferente, Cunha e Heckman (2007) implementam um modelo em que investimentos em etapas diferentes do ciclo de vida de uma criança possam produzir retornos diferentes. Essa abordagem permite interpretar que há períodos ótimos para que certas habilidades de uma criança sejam desenvolvidas. A partir dessa premissa os autores definem dois períodos fundamentais para o desenvolvimento de uma criança. O primeiro é o “período sensível”, em que investimentos para a produção de certas habilidades são mais produtivos. E, o segundo é o “período crítico”, época única da vida de uma criança em que certas habilidades podem ser desenvolvidas. A interpretação do desenvolvimento de habilidades segundo uma tecnologia de múltiplos estágios faz imaginarmos que os investimentos efetuados em um estágio permitem uma “auto-produtividade”, uma vez que as habilidades adquiridas amplificam os efeitos de outras habilidades. Ademais, essa tecnologia é complementar, pois habilidades pré-adquiridas facilitam a assimilação de outras habilidades em etapas futuras, elevando a produtividade dos investimentos nos estágio subseqüentes. Como os autores colocam, essa abordagem é fundamental para a definição de políticas públicas, já que os investimentos para o desenvolvimento de habilidades em adolescentes carentes tenderiam a ser muito menos produtivos do que os investimentos na educação de crianças pobres em época de pré-escola.

A abordagem de Cunha e Heckman ganha mais destaque pois também encontra fundamentos na neurociência. Knudsen et al. (2006) fazem uma revisão da literatura sobre o assunto e concluem que as habilidades humanas são adquiridas através de uma seqüência de “períodos sensíveis”, os quais seriam épocas em que o desenvolvimento de certos circuitos neurais estariam mais plásticos, ou seja, os seres humanos possuiriam uma seqüência hierárquica de momentos ótimos em suas vidas para receber influências do ambiente. No caso dos primeiros anos de vida, os autores completam que as competências cognitivas, lingüísticas, sociais e emocionais são fortemente formadas pelas primeiras experiências de vida da criança em desenvolvimento, e todas contribuem significativamente para o sucesso profissional desse futuro adulto.

Na literatura americana, os efeitos de alguns programas[[6]](#footnote-6) de intervenção na primeira infância são comprovados por vários trabalhos e diversas análises de custo-benefício[[7]](#footnote-7). A conclusão foi de que os participantes dos programas tiveram notas maiores em exames de proficiência, concluíram o ensino médio em maior proporção, e obtiveram menores índices de atraso escolar e de envolvimento com crime e delinqüência. Para alguns dos programas Knudsen et al. (2006) e Barnett (2008) relatam taxas de retorno na ordem de 17% e elevação do QI dos indivíduos participantes com efeitos persistentes ao longo-prazo.

A literatura internacional sobre educação na primeira infância inclui também interessantes trabalhos na América Latina. Berlinski et al. (2006) utilizam a expansão da rede de pré-escolas na Argentina como tratamento para a elevação de matrículas na educação infantil, e concluem que um ano de pré-escola elevou a nota em exames de proficiência da 3ª série em 8% em relação a média ( equivalentes a 23% do desvio padrão), além de afetar positivamente a atenção e participação em sala de aula, e o esforço e a disciplina escolar. Já no Uruguai, Berlinski et al. (2008) avaliam o impacto da educação infantil entre irmãos da mesma família que freqüentaram a pré-escola com aqueles que não freqüentaram, controlando portanto para efeitos específicos e ocultos do *background* familiar. Seus resultados indicam que aos 15 anos os alunos que freqüentaram a pré-escola possuem 27 pontos percentuais a mais de chance de estarem na escola, e estimulam os autores a defenderem a pré-escola como uma política adequada para evitar a repetição escolar nos primeiros anos.

Similarmente aos trabalhos internacionais, o Banco Mundial (2001) relata no Brasil casos de programas informais de desenvolvimento da primeira infância, como o PROAPE (Programa de Alimentação de Pré-Escolar), com intervenções de saúde, alimentação e educação, que resultaram segundo o Banco Mundial em elevação no aproveitamento escolar.

Entretanto, no mesmo trabalho, o Banco Mundial (2001) cita uma pesquisa do IPEA[[8]](#footnote-8), baseada em dados de 1997, em que conclui-se que o investimento público para o desenvolvimento da primeira infância é desigual, sendo destinado em sua grande parte para famílias mais ricas.

Corroborando com o trabalho do Banco Mundial, o levantamento do Grupo de Trabalho instituído pela Portaria Interministerial nº3.219 (2006), com base nos microdados da PNAD 2004 (replicado na Tabela 2), mostra como as classes mais pobres da população são as que menos se beneficiam do sistema de ensino infantil.

Tendo em vista os resultados acima expostos, podemos ter no país um modelo de pré-escola perverso em termos de diminuição da desigualdade social, uma vez que ao privarmos as crianças mais pobres o acesso à creches e pré-escola estamos lhe retirando a possibilidade de desenvolverem habilidades na idade correta, o que lhes impediria de prosseguirem nos próximos ciclos escolares em igualdade de condições.

Com isso, valeria o questionamento empírico do real impacto da pré-escola no futuro escolar dos alunos brasileiros para suportar uma recomendação de política pública de expansão dos investimentos para o desenvolvimento da primeira infância.

**Tabela 2** - Taxa de freqüência escolar por faixas de renda domiciliar per capita - 2004

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | ***Até ½ sm\**** | ***De ½ a 2 sm*** | ***De 2 a 3 sm*** | ***A partir de 3 sm*** | ***Total*** |
| *população de 0 a 3 anos* | 8,5 | 15,5 | 25,7 | 35,3 | 13,4 |
| *população de 4 a 6 anos* | 63,1 | 74,6 | 88,5 | 92,2 | 70,5 |

*\* Salário mínimo*

*Fonte: Relatório do Grupo de Trabalho instituído pela Portaria Interministerial nº3.219 (2006), com base nos dados da PNAD 2004*

Dentro desse princípio, Curi e Menezes-Filho (2006) utilizaram a Pesquisa de Padrão de Vida (PPV) realizada entre 1996 e 1997, no nordeste e sudeste brasileiro além dos resultados do Sistema Nacional de Avaliação da Educação Básica (SAEB) de 2003, para pesquisar o efeito da educação pré-escolar sobre a escolaridade, os salários e a proficiência escolar. Através de um modelo Logit de escolha discreta os autores constataram que o fato de um indivíduo ter feito a pré-escola tem um impacto positivo e significante na conclusão dos quatro ciclos escolares (Primário, Ginásio, Colegial e Universidade), enquanto que a creche tem um efeito positivo e significante na conclusão do colegial e do ensino universitário. Adicionalmente, por regressões através de Ordinary Least Squares (OLS), verificou-se que a pré-escola tem efeito positivo e significante na renda desse indivíduo quando adulto, e no resultado dos exames de proficiência da 4ª, 8ª, e 11ª série. Apesar da pergunta ser extremamente importante, o tratamento da questão por OLS contém sérias questões sobre endogeneidade o que poderia invalidar a consistência dos estimadores obtidos. Em estudos semelhantes para a medição do retorno da educação no salário com amostras americanas, Griliches (1977) e Hausman e Taylor (1981) já apontaram que o efeito da quantidade dos anos de estudo é subestimado em regressões OLS pela ausência no modelo das variáveis não observáveis. Apesar da questão desses autores serem diferentes do tema desse trabalho, vale ressaltar que a fonte de endogeneidade é muito semelhante para as duas questões: a ausência de variáveis que meçam as habilidades dos indivíduos.

Uma vez que a inclusão dessas variáveis no modelo é dificultada pelo simples fato delas serem ocultas, uma solução possível é o uso de variável instrumental para tratarmos o problema de correlação entre o resíduo e a variável explicativa endógena, *pré-escola.* Dessa forma, nosso trabalho amplia a literatura internacional e se aprofunda nos estudos dos efeitos da educação infantil no Brasil, uma vez que melhoramos a consistência do estimador da *pré-escola* ao instrumentalizá-lo. Como proposta de variável instrumental sugerimos o uso do percentual de pré-escolas em relação ao número de crianças de 4 a 6 anos de idade de cada município, obtidos através do censo educacional de 2000. A idéia desse indicador como instrumento é que a oferta de pré-escolas pela prefeitura esteja relacionada com o fato dos alunos cursarem ou não a pré-escola.

Na seção 3 do nosso trabalho discutimos algumas estatísticas dos alunos que realizaram o exame da Prova Brasil 2005, e algumas diferenças entre os alunos que fizeram a pré-escola e aqueles que não freqüentaram o ensino infantil. Na seção 4 descrevemos o modelo econométrico adotado no trabalho, discutimos a construção das variáveis e argumentamos sobre a necessidade de instrumentalizar algumas variáveis assumidas como endógenas ao modelo. A seção 5 relata nossos resultados e apresenta nossas interpretações sobre eles. Finalmente, concluímos na seção 6, onde resumimos os principais pontos de nossa pesquisa.

# 3. Painel da educação infantil brasileira e da Prova Brasil 2005

A Prova Brasil 2005 foi aplicada durante o mês de novembro, e abrangeu toda a rede pública urbana da 4ª e 8ª séries do ensino fundamental. As provas foram aplicadas em cerca de 160 mil turmas de 41 mil escolas com mais de 30 alunos matriculados na série avaliada, em 5.398 municípios[[9]](#footnote-9). Todos alunos examinados obtiveram uma nota para a prova de matemática e outra nota para a prova de português.

Pela Tabela 3 podemos observar que na faixa etária principal de nosso estudo (crianças com cerca de 10 anos de idade), quase que 98,5% das crianças do Brasil freqüentam a escola. Dentre os que freqüentam a escola, por volta de 80% são residentes da zona urbana (Tabela 4) e dentro das áreas urbanas 85,1% dos alunos são atendidos pela rede pública (Tabela 5). Com isso observamos que a PB2005 abrangeu cerca de 67% das crianças brasileiras que idealmente estariam na 4ª série do ensino fundamental.

**Tabela 3** - Freqüência escolar brasileira por idade

|  |  |  |
| --- | --- | --- |
|  | ***Idade*** |   |
| ***Frequenta Escola*** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** | **Total** |
| **Não** | 1,95% | 1,85% | 1,53% | 1,64% | 2,06% | 3,38% | 6,05% | 10,42% | 3,60% |
| **Sim** | 98,05% | 98,15% | 98,47% | 98,36% | 97,94% | 96,62% | 93,95% | 89,58% | 96,40% |
| **Total** | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100,00% |

*Fonte: PNAD 2005*

**Tabela 4** - Distribuição por idade e por região dos alunos brasileiros

|  |  |  |
| --- | --- | --- |
|  | ***Idade*** |   |
| ***Região*** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** | **Total** |
| **Rural** | 18,54% | 19,51% | 20,11% | 19,95% | 18,94% | 18,75% | 18,51% | 18,30% | 19,09% |
| **Urbana** | 81,46% | 80,49% | 79,89% | 80,05% | 81,06% | 81,25% | 81,49% | 81,70% | 80,91% |
| **Total** | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100,00% |

*Fonte: PNAD 2005*

**Tabela 5** - Distribuição por tipo de escola e por idade nas regiões urbanas brasileiras

|  |  |  |
| --- | --- | --- |
|  | ***Idade*** |   |
| ***Tipo de Escola*** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** | **Total** |
| **Privada** | 15,39% | 15,65% | 14,89% | 13,76% | 13,63% | 12,68% | 12,74% | 14,48% | 14,17% |
| **Pública** | 84,61% | 84,35% | 85,11% | 86,24% | 86,37% | 87,32% | 87,26% | 85,52% | 85,83% |
| **Total** | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 100,00% |

*Fonte: PNAD 2005*

A Tabela 6 e a Figura 1, construídas com as informações dadas pelos alunos que realizaram a P B2005, nos fornece uma análise inicial entre os alunos que freqüentaram e os que não freqüentaram a pré-escola. A princípio, observamos uma indicação de que os alunos que não freqüentaram a pré-escola se atrasam mais na progressão escolar do que aqueles que freqüentaram a pré-escola. A Tabela 6 também compara a distribuição de idades dos alunos da 4ª série que fizeram a PB2005 com a distribuição do total de alunos brasileiros da mesma série estimada pela PNAD 2005.

**Tabela 6** - Comparação entre alunos da 4a série (2005) que fizeram e os que não fizeram a pré-escola por faixa etária

|  |  |  |
| --- | --- | --- |
|  | ***Idade*** |   |
| ***Frequentaram Pré-escola*** | **<= 8** | **9** | **10** | **11** | **12** | **13** | **14** | **>= 15** | **Total** |
| **Não** | 1,89% | 4,71% | 33,07% | 26,69% | 15,41% | 8,53% | 4,81% | 4,89% | 100,00% |
| **Sim** | 1,28% | 4,76% | 49,74% | 26,86% | 9,49% | 4,26% | 2,06% | 1,54% | 100,00% |
| **Total** | 1,43% | 4,75% | 45,54% | 26,82% | 10,98% | 5,33% | 2,75% | 2,38% | 100,00% |
| **Comparação com a População Total da 4ª série pela PNAD 2005** |
| **4 série** | 1,14% | 10,86% | 44,13% | 20,27% | 10,41% | 6,54% | 4,24% | 2,41% | 100,00% |

*Fonte: Prova Brasil 2005*

**Figura 1** – Faixa etária por grupo de alunos que freqüentaram ou não a pré-escola

|  |  |
| --- | --- |
|  |  |

Já a Tabela 7 nos fornece maiores informações sobre a relação formação dos pais e o fato do aluno ter freqüentado ou não a pré-escola, enquanto a Tabela 8 aborda a relação formação dos pais e a nota do aluno no exame de matemática da PB2005. Nos dois casos observamos – além da evidente relação da educação paterna na educação dos filhos – o fato dos maiores saltos positivos se darem quando um dos pais freqüentou a 4ª série, e posteriormente quando um dos pais freqüentou o ensino médio.

**Tabela 7** – Freqüência da pré-escola em função da formação dos pais

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | Educação Paterna |  |
|  |  | Sem educação Formal | 4ª Série | 8ª Série | E. Médio | E. Superior | Total |
| Educação Materna | Sem educação Formal | **63,99%** | **67,67%** | **70,14%** | **72,48%** | **72,17%** | **66,68%** |
| 4ª Série | **67,76%** | **72,75%** | **74,91%** | **78,91%** | **78,36%** | **73,32%** |
| 8ª Série | **70,47%** | **76,84%** | **76,88%** | **79,01%** | **79,70%** | **76,93%** |
| E. Médio | **73,94%** | **82,04%** | **81,98%** | **84,62%** | **85,25%** | **83,04%** |
| E. Superior | **70,87%** | **82,15%** | **81,74%** | **85,13%** | **83,01%** | **82,49%** |
| Total | **66,91%** | **74,11%** | **76,74%** | **81,77%** | **81,84%** | **75,88%** |

*Fonte: Prova Brasil 2005*

**Tabela 8** – Nota de matemática na PB2005 (4ª série) em função da formação dos pais

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | **Educação Paterna** |  |
|  |  | Sem educação Formal | 4ª Série | 8ª Série | E. Médio | E. Superior | Total |
| **Educação Materna** | Sem educação Formal | **168,30**  | **173,34**  | **171,41**  | **173,77**  | **173,54**  | **170,67**  |
| 4ª Série | **174,26**  | **180,31**  | **180,95**  | **188,40**  | **183,06**  | **180,46**  |
| 8ª Série | **172,28** | **182,11**  | **180,96**  | **186,64**  |  **185,17**  | **181,89**  |
| E. Médio | **176,58**  | **190,53**  | **190,34**  | **195,46**  | **197,10**  | **192,57**  |
| E. Superior | **172,66**  | **187,62**  | **187,40**  | **196,80**  |  **189,05**  | **189,08**  |
| Total | **171,42**  | **181,07**  | **182,02**  | **191,22**  | **188,46**  | **182,47**  |

*Fonte: Prova Brasil 2005*

Quanto às características de sexo e raça[[10]](#footnote-10) a Tabela 9 nos reporta que nas provas de matemática se destacam as crianças brancas, independentemente do sexo, enquanto que nas provas de português as melhores notas foram obtidas pelas meninas brancas.

**Tabela 9** – Notas na PB2005 (4ª série) em função do sexo e raça

|  |
| --- |
| Notas de matemática |
|  | Branco | Não Branco | Total |
| Meninas |  186.71  |  178.82  |  181.84  |
| Meninos |  186.70  |  179.17  |  182.01  |
| Total |  186.71  |  179.00  |  181.92  |
| Notas de português |
|  | Branco | Não Branco | Total |
| Meninas |  184.32  |  176.48  |  179.48  |
| Meninos |  174.72  |  167.61  |  170.29  |
| Total |  179.44  |  171.92  |  174.77  |

*Fonte: Prova Brasil 2005*

Abordando mais diretamente a relação ter feito, ou não, a pré-escola e a proficiência escolar na 4ª série, a Tabela 10 e a Figura 2 mostram que há uma diferença estatisticamente significante no rendimento escolar entre esses dois grupos de estudo (os que fizeram e aqueles que não fizeram a pré-escola). Para toda a mostra de alunos da 4ª série que fizeram a PB2005 observamos que em média os alunos que freqüentaram a pré-escola obtiveram 14,6 pontos a mais no exame de matemática, e 15,6 pontos a mais no exame de português.

**Tabela 10** - Comparação de notas entre alunos da 4a série (2005) que fizeram e os que não fizeram a pré-escola

|  |
| --- |
| **Notas de matemática** |
|  | **Idade** |  |
|  | **<= 8** | **9** | **10** | **11** | **12** | **13** | **14** | **>= 15** | **Total** |
| **Sem pré-escola** | 157,32 | 173,87 | 177,22 | 170,78 | 166,65 | 165,68 | 163,87 | 167,09 | 171,22 |
| **Com pré-escola** | 163,34 | 186,19 | 193,94 | 183,53 | 169,93 | 166,40 | 164,41 | 163,42 | 185,85 |
| **Total** | 161,34 | 183,11 | 190,89 | 180,34 | 168,77 | 166,11 | 164,17 | 165,31 | 182,17 |
| **Teste t : Dif = 0 (P-value)** | 0,000 | 0,000 | 0,000 | 0,000 | 0,000 | 0,001 | 0,083 | 0,000 |  |
| **Notas de português** |
|  | **Idade** |  |
|  | **<= 8** | **9** | **10** | **11** | **12** | **13** | **14** | **>= 15** | **Total** |
| **Sem pré-escola** | 148,84 | 166,98 | 170,59 | 163,00 | 157,77 | 156,19 | 154,42 | 155,76 | 163,27 |
| **Com pré-escola** | 155,94 | 180,72 | 187,97 | 176,52 | 160,84 | 156,20 | 153,76 | 152,06 | 178,95 |
| **Total** | 153,58 | 177,29 | 184,79 | 173,13 | 159,75 | 156,20 | 154,05 | 153,97 | 175,00 |
| **Teste t : Dif = 0 (P-value)** | 0,000 | 0,000 | 0,000 | 0,000 | 0,000 | 0,950 | 0,037 | 0,000 |  |

*Fonte: Prova Brasil 2005*

**Figura 2** – Nota dos alunos da 4ª série na Prova Brasil 2005, para os que freqüentaram e os que não freqüentaram a pré-escola.

|  |  |
| --- | --- |
|  |  |

*Fonte: Prova Brasil 2005*

Vale a pena notar que no período de estudo abrangido por nosso trabalho a entrada obrigatória das crianças na escola, regida pela Lei 9.394 de 1.996, determinava que as crianças eram obrigadas a iniciar o ensino fundamental (1ª série) aos 7 anos de idade. As mudanças posteriores, pela Lei 11.114 de 2005, que alteram a idade inicial para 6 anos, e pela Lei 11.274 de 2006, que estabelecem o ciclo escolar do ensino fundamental de 9 anos, não afetaram, portanto, nossa pesquisa.

# 4. O Modelo e os Dados

A aplicação da técnica de OLS para a estimação do efeito da educação pré-escolar na proficiência sofre do problema tradicional de endogeneidade, pois provavelmente há correlação entre a variável *dummy*, ter feito ou não pré-escola e o termo de erro do modelo, resultando em estimações possivelmente inconsistentes.[[11]](#footnote-11)

*Nota i = δ0 + δ1 (pré-escolai )+ δ2 (Controlesi ) + εi* [1]

A estimação da equação [1] acarreta na omissão de algumas variáveis não observáveis, que são determinantes no efeito da variável *Notai*. Conseqüentemente, os efeitos dessas variáveis ocultas são capturados pelo termo de erro *εi*. Dessa forma, temos variáveis ocultas, que na maioria das vezes estão correlacionadas com o fato do indivíduo ter feito ou não pré-escola, impactando o resíduo da equação, e resultando em uma fonte de endogeneidade.

Dessa forma, adotamos como metodologia o uso da técnica econométrica Two-Stage Least Square (2SLS), considerando para a construção das variáveis os exames de proficiência da Prova Brasil 2005 para as notas dos alunos da 4ª série, e o Censo Educacional de 2000 para a construção da variável instrumental.

 Pela Figura 3, pode-se perceber que a idéia do instrumento é voltar ao ano em que os alunos da 4ª série de 2005 estavam na pré-escola e aferir o número de pré-escola em relação ao número de crianças com idade entre 4 e 6 anos para o município desse aluno. Tal instrumento seria válido uma vez que a variação da oferta de pré-escolas entre municípios está correlacionada com o fato dos alunos do município cursarem a pré-escola e não parecer estar relacionada ao resíduo da equação de nosso modelo[[12]](#footnote-12). Como robustez do modelo testamos duas especificações diferentes: número total de pré-escolas para cada 10 crianças com idade entre 4 e 6 anos por município; e o número de pré-escolas públicas para cada 10 crianças com idade entre 4 e 6 anos por município[[13]](#footnote-13).

  **Figura 3** – Seqüência de Estudo dos Alunos Avaliados

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Ano** | **2000** | **2001** | **2002** | **2003** | **2004** | **2005** |
| Ano de Prova Brasil | - | - | - | - | - | Sim |
| Censo Educacional | Sim | Sim | Sim | Sim | Sim | Sim |
| Censo Populacional | Sim | - | - | - | **-** | - |
| *Série:* | **Pré-Escola** | **Pré-Escola** | 1ª | 2ª | 3ª | **4ª** |

Pela equação [2] teríamos o segundo estágio de nossa regressão, já corrigido para a provável endogeneidade. E a equação [3], por sua vez, seria a equação reduzida do primeiro estágio.

*Nota 2005im = β0 + β1 (pré-escola i)+ β2 (caract. criança i) + β3 (caract. familia i) +*

*β 4 (caract. município m)+ e i* [2]

*Pré-escola i = θ0 + θ1 (Instrumento m)+ θ2 (caract. criança i)+ θ3 (caract. familia i)+*

 *θ4 (caract. município m)+ ri* [3]

Pelas equações [2] e [3] podemos perceber que apesar de nosso modelo estar avaliando a nota de cada aluno (*i*), localizado no município (*m*), nossa variável instrumental será gerada apenas para o município (*m*).

Como variáveis de controle das características das crianças incluímos as seguintes informações a partir do questionário respondido pela criança antes do exame da PB2005: raça, sexo, idade em meses, e idade em meses ao quadrado. Sobre as características da família incluímos *dummies* com informações da PB sobre a educação dos pais[[14]](#footnote-14), e controles para estrutura familiar: se a criança mora só como pai, ou se a criança mora só com a mãe, ou se a criança não mora com nenhum dos pais. Para os municípios incluímos também como controle a renda per capita do município e quantidade média de anos de estudo para a população acima de 25 anos de idade, ambas as informações obtidas pelo Ipeadata.[[15]](#footnote-15)

Às regressões adicionamos, também, controles para os instrumentos. Quando o instrumento era o número total de pré-escolas para cada 10 crianças com idade entre 4 e 6 anos por município em 2000, os controles foram o número total de pré-escolas para cada 10 crianças com idade entre 4 e 6 anos por município em 2005 e o número total de escolas fundamentais para cada 10 crianças com idade entre 7 e 18 anos por município em 2005. E, quando o instrumento era o número de pré-escolas públicas para cada 10 crianças com idade entre 4 e 6 anos por município em 2000, os controles foram o número pré-escolas públicas para cada 10 crianças com idade entre 4 e 6 anos por município em 2005 e o número de escolas fundamentais públicas para cada 10 crianças com idade entre 7 e 18 anos por município em 2005. Uma vez que não temos a contagem da população por idade por município anualmente, estimamos as populações de 4 a 6 anos e 7 a 18 anos, através da população por idade por estado da PNAD 2005 e da estimação da população total dos municípios pelo IBGE.[[16]](#footnote-16)

Em uma segunda parte do trabalho avaliamos também os efeitos do atraso escolar nas notas de proficiência. Sendo a variável “*Criança com idade escolar acima da ideal”* também endógena ao modelo, incluímos um segundo instrumento: o fato da criança ter nascido no segundo semestre do ano. Ao observarmos nossa amostra de alunos verificamos que as crianças que nasceram no segundo semestre do ano fizeram menos a pré-escola do que os alunos que nasceram na primeira metade do ano. Nossa interpretação para o fato é que alguns pais de crianças com aniversário no final do ano ou os próprios educadores, ao chegar na época de matricular os filhos na pré-escola acreditam que essas crianças sejam muito novas para entrar na escola. Com isso, seguram as crianças fora da escola por um ano, até que elas tenham a idade para entrar direto no ensino fundamental. Avaliando os dados, observamos que essa diferença estatisticamente significante aparece em quase todas as unidades da federação brasileira (Anexo A1). Como freqüentar a pré-escola nos parece melhorar o rendimento escolar, acabamos encontrando que fazer o aniversário no 2º semestre acaba sendo um fato exógeno para indicar maior repetência.

Para melhor compreendermos os efeitos da pré-escola e do atraso escolar dividimos nossa população em quatro grupos (Figura 3): alunos com a idade escolar correta que nunca repetiram (Grupo A); alunos com idade escolar correta que repetiram (Grupo B) – ou seja, grupo composto por crianças que entraram antecipadamente na escola e acabaram repetindo; alunos com idade escolar acima da ideal por fatores diversos do que repetência (Grupo C) – ou seja, crianças que entraram tardiamente ou abandonaram algum ano escolar; e alunos com idade escolar acima da ideal que repetiram (Grupo D).

Nossas variáveis foram montadas com base nas informações de idade e repetência, fornecidas pelo questionário da Prova Brasil 2005. Como critério de idade escolar ideal para a 4ª série adotou-se o conceito de que a crianças deveriam ter completado 10 anos durante o segundo semestre do ano anterior (2004) ou tinham completado 10 anos no 1º semestre do ano de 2005.[[17]](#footnote-17)

**Figura 3** – Divisão dos Grupos de Alunos Avaliados

|  |  |  |
| --- | --- | --- |
|  | **Alunos com idade escolar correta** | **Alunos com idade escolar atrasada** |
| **Alunos que nunca repetiram** | Grupo**A** | Grupo**C** |
| **Alunos que repetiram** | Grupo**B** | Grupo**D** |

Analisando os resultados da proficiência escolar para grupos diferentes de alunos conseguimos observar os efeitos da pré-escola e do momento da entrada na escola como determinantes do rendimento dos alunos na 4ª série do ensino fundamental.

Abaixo (Tabela 11) incluímos as estatísticas descritivas das variáveis analisadas e o tamanho das amostras estudadas.

**Tabela 11** – Estatísticas Descritivas

|  |
| --- |
| **Estatística Descritiva dos Alunos** |
| Variável | Observações | Média | Des.Padrão | Mínima | Máxima |
| **Proficiência em Matemática** | 1,975,557 | 180.6245 | 39.90735 | 75.34725 | 330.6538 |
| **Proficiência em Português** | 1,975,557 | 173.3942 | 41.73293 | 64.52156 | 324.6158 |
| **Pré-escola** | 1,884,780 | 0.746221 | 0.435173 | 0 | 1 |
| **Aniversário no 2o Semestre** | 1976,122 | 0.523277 | 0.499458 | 0 | 1 |
| **Criança com idade escolar Atrasada** | 1,843,539 | 0.3194 | 0.466244 | 0 | 1 |
| **Atrasado não por Repetição** | 1,751,642 | 0.194346 | 0.395696 | 0 | 1 |
| **Repetiu pelo menos 1 vez** | 1,877,500 | 0.313786 | 0.464031 | 0 | 1 |
| **Idade em meses** | 1,843,539 | 134.4388 | 15.54412 | 96 | 191 |
| **Masculino** | 1,908,104 | 0.511174 | 0.499875 | 0 | 1 |
| **Não - Branco** | 1,878,693 | 0.620706 | 0.485212 | 0 | 1 |
| **Reside só com o pai** | 1,975,557 | 0.036809 | 0.188292 | 0 | 1 |
| **Reside só com a mãe** | 1,975,557 | 0.193866 | 0.395325 | 0 | 1 |
| **Não reside com o pai ou a mãe** | 1,975,557 | 0.104411 | 0.305793 | 0 | 1 |
| **Reside - sem info.** | 1,975,557 | 0.050655 | 0.219292 | 0 | 1 |
| **Educ. Mãe - Completou a 4ªSérie** | 1,976,122 | 0.249859 | 0.432931 | 0 | 1 |
| **Educ. Mãe - Completou a 8ªSérie** | 1,976,122 | 0.123806 | 0.32936 | 0 | 1 |
| **Educ. Mãe - Completou o Ens. Médio** | 1,976,122 | 0.125344 | 0.331109 | 0 | 1 |
| **Educ. Mãe - Completou o Ens. Superior** | 1,976,122 | 0.081537 | 0.273657 | 0 | 1 |
| **Educ. Mãe - Sem info.** | 1,975,557 | 0.313124 | 0.463765 | 0 | 1 |
| **Educ. Pai - Completou 4ªSérie** | 1,976,122 | 0.192026 | 0.393894 | 0 | 1 |
| **Educ. Pai - Completou a 8ªSérie** | 1,976,122 | 0.113003 | 0.316597 | 0 | 1 |
| **Educ. Pai - Completou o Ens. Médio** | 1,976,122 | 0.103332 | 0.304392 | 0 | 1 |
| **Educ. Pai - Completou o Ens. Superior** | 1,976,122 | 0.084793 | 0.278574 | 0 | 1 |
| **Educ. Pai - Sem info.** | 1,975,557 | 0.406615 | 0.491202 | 0 | 1 |
| **Estatística Descritiva dos Municípios** |
| Variável | Observações | Média | Des.Padrão | Mínima | Máxima |
| **Renda per Cap. do Munic. (R$ de 2000)** | 5,003 | 170.7567 | 96.9741 | 28.38 | 954.649 |
| **Anos de Estudo - Media Munic. (pop>25anos)** | 5,027 | 4.032266 | 1.327046 | 0 | 9.653 |
| **Pré-esc. em 2000 por 10 crianças de 4 a 6 anos** | 5,410 | 0.126571 | 0.133587 | 0.002572 | 2.440292 |
| **Pré-esc. Pública em 2000 por 10 crianças de 4 a 6 anos** | 5,410 | 0.116365 | 0.134057 | 0 | 2.440292 |
| **Pré-esc. em 2005 por 10 crianças de 4 a 6 anos** | 5,506 | 0.173364 | 0.150135 | 0 | 1.759669 |
| **Pré-esc. Pública em 2005 por 10 crianças de 4 a 6 anos** | 5,506 | 0.160631 | 0.148711 | 0 | 1.759669 |
| **Esc. Fund. em 2005 por 10 crianças de 7 a 18 anos** | 5,507 | 0.055644 | 0.051793 | 0 | 0.412715 |
| **Esc. Fund. Pública em 2005 por 10 crianças de 7 a 18 anos** | 5,507 | 0.053494 | 0.052237 | 0 | 0.412715 |
| **Numero de Observações por Grupo de Crianças e porcentagem de freqüência da pré-escola** |
| **Grupo A: Idade correta & não repetiram** | 1,060,091 | 81,9% |  |  |  |
| **Grupo B: Idade correta & repetiram** | 173,314 | 68,8% |  |  |  |
| **Grupo C: Idade atrasada & não repetiram** | 186,353 | 64,0% |  |  |  |
| **Grupo D: Idade atrasada & repetiram** | 387,224 | 64,1% |  |  |  |

# 5. Resultados Econométricos

Em nossa primeira análise observamos os efeitos da pré-escola apenas para o Grupo A de nossa amostra. A Tabela 12 mostra a robustez do efeito significante e positivo da pré-escola às varias técnicas econométricas.[[18]](#footnote-18) Na tabela com os primeiros estágios, o Modelo (1) se refere às regressões com o instrumento *Pré-escolas em 2000 por 10 crianças de 4 a 6 anos* sem controle para os números de escolas em 2005 (primeira coluna para o conjunto de regressões com IV), o Modelo (2) é relativo às regressões com o instrumento *Pré-escolas em 2000 por 10 crianças de 4 a 6 anos* , mas com os controles para os números de escolas em 2005 (segunda coluna para o conjunto de regressões com IV); o Modelo (3) é o primeiro estágio para as regressões com o instrumento *Pré-escolas Públicas em 2000 por 10 crianças de 4 a 6 anos* sem controles para o número de escolas em 2005 (terceira coluna para o conjunto de regressões com o IV) ; e finalmente o Modelo (4) inclui o primeiro estágio das regressões com o instrumento *Pré-escolas Públicas em 2000 por 10 crianças de 4 a 6 anos* , mas com os controles para o número de escolas em 2005 (quarta coluna para o conjunto de regressões com IV). Ao longo do trabalho seguiremos sempre essa lógica para a apresentação dos resultados.

Pelas análises do primeiro estágio das regressões com controle para o número de escolas em 2005 observamos que para as crianças na idade correta e que nunca repetiram o incremento de uma pré-escola no município eleva em 0,28 a chance de uma criança fazer a pré-escola enquanto que a construção de uma pré-escola pública aumenta em 0,27 a chance de uma criança freqüentar a pré-escola. Nos segundos estágios encontramos resultados fortes para o impacto da pré-escola. No exame de matemática observamos ganhos ao redor dos 30% em relação a media (equivalentes a ganhos ao redor dos 120% em desvios padrões). No exame de português os ganhos ficam acima dos 20% em relação a média (perto dos 90% de ganho em desvios padrões).

As regressões com instrumento mostram claramente estimadores muito mais elevados dos que as regressões por OLS, ou por painéis com Efeito Fixo que considerem em um dos eixos, ao invés do tempo, a escola ou a turma dos alunos. Uma possível explicação para o fato das idiossincrasias dos alunos não capturadas pelas primeiras técnicas estar enviesando os estimadores para baixo é a não inclusão de alguns fatores positivamente correlacionados com a pré-escola, mas negativamente correlacionados como o rendimento escolar, por exemplo, pais que trabalham e não possuem interesse na vida escolar dos filhos, mas os matriculam na pré-escola apenas para que tenham uma ocupação durante o dia.

**Tabela 12** – Regressões para crianças na idade escolar correta e que nunca repetiram (Grupo A)

#

# Primeiros Estágios das Regressões com IV

#

Para continuar nossa análise e verificar os efeitos da pré-escola nos outros grupos necessitamos controlar o modelo pela variável endógena “*Criança com idade escolar acima da ideal*”, isso nos exigiu a inclusão de nosso segundo instrumento “*Aniversário no 2o Semestre*”. A Tabela 13 contém os resultados do modelo que inclui os 4 grupos estudados (Grupos A, B, C e D). Ao observarmos os primeiros estágios notamos que para o grupo total de alunos a oferta de pré-escola eleva a chance das crianças freqüentarem o pré e possui relação negativa com o fato da criança estar com idade escolar acima da ideal, como já esperado. Por outro lado, o fato da criança fazer aniversário no segundo semestre mostra uma maior chance dela não fazer a pré-escola -- também como esperado -- e uma menor chance dela estar com a idade escolar acima da ideal. A interpretação do coeficiente da variável “*Aniversário no 2º semestre*” merece uma atenção especial, pois os modelos são controlados para a idade em meses das crianças, logo não é o fator idade que esta variável está captando, mas sim o fato da criança ter entrado mais tarde na escola porque os pais ou educadores acreditarem que a criança era muito nova para a vida escolar. Como nesse conjunto de regressões as crianças se atrasaram por motivos diversos (repetência, entrada tardia, ou abandono) não tínhamos um sinal esperado para a variável *Aniversário no 2º semestre*, nas regressões para *Idade Escolar acima do Ideal*.

Por sua vez, no segundo estágio continuamos observando o forte impacto da pré-escola nas notas na 4ª série, com ganhos em relação a média no exame de matemática por volta dos 18% (superiores a 80% de ganhos em desvios padrões) e ganhos em relação a média no exame de português por volta dos 14% (próximos aos 60% de ganhos em desvios padrões). Porém, a variável “*idade escolar acima da ideal*” apresenta uma interessante mudança para as regressões instrumentalizadas em relação às regressões sem instrumentos: a inversão do sinal nas duas provas e a perda de significância estatística no exame de matemática.

A fim de melhor compreender esse resultado, que poderia indicar que crianças as crianças com idade escolar acima da ideal possuem um rendimento melhor do que as crianças com idade escolar adequada, prosseguimos com outras regressões para compararmos os resultados de cada um dos 4 grupos.

**Tabela 13** – Regressões para todas as crianças (Grupos A, B, C e D)



# Primeiros Estágios das Regressões com IV



Iniciamos as comparações analisando os alunos que estão na idade escolar correta e que nunca repetiram (Grupo A) contra os alunos que estão com a idade escolar acima da ideal por repetições (Grupo D). A Tabela 14 nos mostra mais uma vez a oferta de pré-escola com correlação positiva com a pré-escola. Entretanto apenas para esses grupos o aniversário no 2º semestre se mostrou positivamente relacionado com a pré-escola. No segundo estágio, conseguimos observar que o grupo dos alunos que estão na idade correta e que nunca repetiram (Grupo A) possui um rendimento melhor do que os alunos que estão com idade escolar acima da ideal por repetição (Grupo D).

**Tabela 14** – Regressões para as crianças na idade escolar correta que nunca repetiram e crianças com idade escolar acima da ideal por repetição (Grupos A vs. D)



# Primeiros Estágios das Regressões com IV



Nossa próxima comparação se concentrou nos dois grupos de alunos com idade escolar acima da ideal, Grupos C e D. O primeiro estágio dessas regressões nos mostra novamente que a maior oferta de pré-escolas pelo município eleva a chance das crianças freqüentarem a pré-escola, que o aniversário no segundo semestre é negativamente correlacionado com a freqüência na pré-escola, e que o aniversário na segunda metade do ano diminui a chance da criança entrar atrasada ou abandonar a escola por algum período, em comparação com as crianças que entraram na data correta, mas se atrasaram por repetições. No segundo estágio observamos que a pré-escola perdeu significância para essa amostra, possivelmente pela interferência do Grupo C (crianças com idade escolar acima da ideal por razões diferentes de repetição), em que entrar atrasado na escola ou ficar um período longe da sala de aula afastam os benefícios da educação infantil. Por fim, percebemos que o grupo de alunos do Grupo D ( alunos com idade escolar acima da ideal por repetição) obteve uma performance um pouco melhor do que os alunos do Grupo C.

**Tabela 15** – Regressões para as crianças com idade escolar acima da ideal (Grupos C vs. D)



# Primeiros Estágios das Regressões com IV



No último conjunto de comparações (Grupos B e C) as regressões de primeiro estágio reportam resultados coerentes com as regressões anteriores. No segundo estágio observa-se que a *pré-ecola* não obteve significância estatística, mas novamente tivemos a interferência das peculiaridades das crianças do Grupo C. Além disso, notamos que o rendimento das crianças que entram cedo e repetem (Grupo B) é muito parecido com o rendimento das crianças que se atrasam por abandono ou porque entrarem tardiamente na escola (Grupo C), sendo até um pouco pior do que esse segundo grupo nos exames de português.

Com essas três últimas tabelas podemos concluir uma ordem no desempenho desses grupos estudados nos exames da Prova Brasil 2005: Foram melhores os alunos do Grupo A, posteriormente os alunos do Grupo D, depois os alunos do Grupo C, e por último os alunos do Grupo B.

 **Tabela 16** – Regressões para as crianças na idade escolar correta que repetiram e crianças que se atrasaram não por repetição (Grupos B vs.C)



# Primeiros Estágios das Regressões com IV



# 5. Conclusão

Com o presente trabalho observamos que ao instrumentalizarmos a variável *pré-escola* em regressões para estimarmos a proficiência escolar na 4ª série obtem-se estimadores muito mais altos do que nas regressões por OLS, indicando um possível viés de omissão para baixo nos estudos anteriores para o Brasil. Nossos resultados indicam que para toda a amostra de alunos brasileiros da 4ª série que realizaram a Prova Brasil em 2005, o efeito da pré-escola foi equivalente a ganhos em relação a média no exame de matemática entre 17% e 21% (entre 77% e 96% de ganhos em desvios padrões) e ganhos em relação a média no exame de português entre 12% e 15% ( entre 52% e 61% de ganhos em desvios padrões). O significante impacto da pré-escola foi persistente a várias especificações e as mudanças dos grupos de alunos, e vão ao encontro dos resultados obtidos em estudos internacionais (por exemplo, Knudsen et al., 2006; Berlinski et al., 2006; e Berlinski et al., 2008).

Nossas analises também revelam a importância do momento da entrada da criança na escola. Entre os nossos grupos de estudo da 4ª série, apresentaram as piores performances no exame da Prova Brasil 2005 os grupos de alunos que entraram antecipadamente e acabaram repetindo (Grupo B), e o conjunto de alunos que entraram atrasados ou que abandonaram a escola por algum período (Grupo C). O baixo resultado nos exames dos alunos que entraram antecipadamente é preocupante, pois mostra que os alunos mesmo refazendo alguma série não conseguem se recuperar. Uma vez que a primeira infância é em termos psicológicos um momento extremamente crítico (Knudsen et al., 2006), podemos especular que uma criança jovem demais que entre no ensino fundamental pode sofrer traumas sérios pela frustração de não conseguir acompanhar o andamento das aulas, e esses traumas continuarão atrapalhando o desempenho da criança ao longo da sua vida escolar.

Dada a importância do tema e os fortes resultados apresentados, acreditamos que políticas públicas de expansão e melhoria da pré-escola seriam fundamentais para elevar o rendimento e diminuir a desigualdade do sistema educacional brasileiro. Esperamos que novos trabalhos deêm continuidade a discussão, incluindo, por exemplo, pesquisas que consigam aferir o impacto da elevação da qualidade da pré-escola na qualiadade dos alunos ao longo da vida escolar.

# Bibliografia

Banco Mundial - Departamento de Desenvolvimento Humano. (2001). *Desenvolvimento da Primeira Infância: Foco sobre o Impacto das Pré-Escolas.* Brasil: Banco Mundial.

Barnett, W.S. (2008). Preschool education and its lasting effects: Research and policy implications. *Education and the Public Interest Center & Education Policy Research Unit*: Arizona State University and University of Colorado at Boulder.

Becker, G. S., & Tomes, N. (1986). Human Capital and the Rise and Fall of Families. *Journal of Labor Economics* , Vol. 4 (3), 1-39.

Berlinski, S., Galiani, S., & Gertler, P.J. (2006). The Effect of Preprimary Education on Primary School Performance. *William Davidson Institute*. Working Paper No. 838.

Berlinski, S., Galiani, S., & Manacorda, M. (2008). Giving children a better start: Preschool attendance and school-age profiles. *Journal of Public Economics*, 92, 1416-1440.

Cunha, F., & Heckman, J. (2007). The Technology of Skill Formation. *American Economic Review*, Vol. 97 (2), 31-47.

Cunha, F., Heckman, J. J., Lochner, L., & Masterov, D. V. (2005). Interpreting the Evidence on Life Cycle Skill Formation. *National Bureau Of Economic Research* , Working Paper 11331.

Curi, A. Z., & Menezes-Filho, N. A. (2006). Os Efeitos da Pré-escola sobre os Salários, a Escolaridade e a Proficiência Escolar. *XXXIV Encontro Nacional de Economia.* Salvador, BA: ANPEC.

Currie, J. (2001). Early Childhood Education Programs. *Journal of Economic Perspectives*, 15 (2), 213-238.

Feinstein, L. (2003). Inequality in the Early Cognitive Development of British Children in the 1970 Cohort. *Economica,* 70 (277), 73-97.

Griliches, Z. (1977). Estimating The Returns To Schooling: Some Econometric Problems. *Econometrica* , Vol. 45 (1), 1-22.

Griliches, Z., Hall, B., & Hausman, J. (1978). Missing data and self-selection in large panels. *Annales de I'INSEE* , 137-176.

Grupo de Trabalho instituído pela Portaria Interministerial nº3.219. (2006). *Relatório do Grupo de Trabalho Interministerial.* Brasília: Ministério da Educação - Secretaria de Educação Básica.

Hausman, J. A., & Taylor, W. E. (1981). Panel Data and Unobservable Individual Effects. *Econometrica* , Vol 49 (6), 1377-1398.

Knudsen, E., Heckman, J.J., Cameron, J.L., & Shonkoff., J.P. (2006). Economic, Neurobiological, and Behavioral Perspectives on Building America's Future Workforce. *Proceedings of the National Academy of Sciences*, 103, no. 27: 10155-10162.

Krueger, A.B. (2002). Inequality, Too Much of a Good Thing. *Paper to the Princeton University Industrial Relations Section*. Princeton, NJ: Princeton University Press, 12-23.

Schady, N. (2006). Early Childhood Development in Latin America and the Caribbean. *Journal of the Latin American and Caribbean Ecoonomic Association*, Vol. 6 (2), 185-213.

# Anexos

**A1** – Relação semestre de nascimento e freqüência da pré-escola para cada estado e o DF

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **UF** | **Aniv. 1o Sem.** | **Aniv. 2o Sem.** | **Diferença** | **Teste t : dif = 0 (P-value)** |
| **AC** | 65.80 | 66.32 | -0.52 | 0.30 |
| **AL** | 64.55 | 63.75 | 0.80 | 0.06 |
| **AM** | 63.10 | 64.35 | -1.25 | 0.00 |
| **AP** | 69.72 | 67.21 | 2.51 | 0.00 |
| **BA** | 67.08 | 66.39 | 0.69 | 0.00 |
| **CE** | 72.14 | 72.11 | 0.03 | 0.45 |
| **DF** | 81.25 | 81.53 | -0.28 | 0.27 |
| **ES** | 73.04 | 71.40 | 1.64 | 0.00 |
| **GO** | 76.09 | 74.50 | 1.59 | 0.00 |
| **MA** | 70.43 | 69.71 | 0.72 | 0.02 |
| **MG** | 82.04 | 79.81 | 2.23 | 0.00 |
| **MS** | 67.24 | 62.22 | 5.02 | 0.00 |
| **MT** | 69.92 | 68.11 | 1.81 | 0.00 |
| **PA** | 66.70 | 66.54 | 0.16 | 0.31 |
| **PB** | 72.89 | 71.52 | 1.37 | 0.00 |
| **PE** | 71.17 | 69.81 | 1.36 | 0.00 |
| **PI** | 72.53 | 73.72 | -1.19 | 0.02 |
| **PR** | 75.63 | 71.04 | 4.59 | 0.00 |
| **RJ** | 77.47 | 79.05 | -1.58 | 0.00 |
| **RN** | 73.22 | 45.13 | 28.09 | 0.00 |
| **RO** | 66.63 | 65.33 | 1.30 | 0.03 |
| **RR** | 73.12 | 72.80 | 0.32 | 0.39 |
| **RS** | 72.63 | 71.65 | 0.98 | 0.00 |
| **SC** | 80.40 | 77.22 | 3.18 | 0.00 |
| **SE** | 76.81 | 76.24 | 0.57 | 0.19 |
| **SP** | 81.38 | 80.55 | 0.83 | 0.00 |
| **TO** | 72.16 | 71.59 | 0.57 | 0.20 |
| **Total** | **75.22** | **74.05** | **1.17** | **0.00** |

**A2** – Crianças na Idade Escolar Correta e que nunca repetiram



1. EAESP - Fundação Getúlio Vargas, SP / CEPESP. [↑](#footnote-ref-1)
2. EESP - Fundação Getúlio Vargas, SP. [↑](#footnote-ref-2)
3. Esse número pode ser dividido em 13,4% para as crianças de 0 a 3 anos, e 70,5% para as crianças de 4 a 6 anos (Cálculos elaborados pelo Grupo de Trabalho instituído pela Portaria Interministerial nº3.219, 2006). [↑](#footnote-ref-3)
4. <http://www.inep.gov.br/estatisticas/gastoseducacao/PIB_nivel_ensino.htm> (Nov/2007) [↑](#footnote-ref-4)
5. Período de 0 aos 6 anos de idade. [↑](#footnote-ref-5)
6. Os programas listados foram: o Perry Preschool Experiment; o Abecedarian Project; e o Chicago Child-Parent Center Program. [↑](#footnote-ref-6)
7. Como alguns exemplos podemos citar: Cunha, Heckman, et al. (2005); Krueger (2002); Barnett (2008); Knudsen et al. (2006); e Curie (2001). [↑](#footnote-ref-7)
8. IPEA, 1999. Uma Avaliação dos Custos e dos Benefícios da Educação Pré-escolar. Rio de Janeiro. [↑](#footnote-ref-8)
9. Fonte: Portal do Inep – Prova Brasil, site: http://provabrasil.inep.gov.br. [↑](#footnote-ref-9)
10. Foram consideradas como brancas as crianças que se identificaram como brancas ou amarelas. [↑](#footnote-ref-10)
11. Veja, por exemplo, Griliches (1977); e Hausman e Taylor (1981). [↑](#footnote-ref-11)
12. A oferta de pré-escolas poderia indicar que o governo local fosse mais focado em educação, o que poderia refutar a idéia de que nosso instrumento fosse completamente exógeno ao modelo. Para contornar essa questão, incluímos como controle no modelo o número de pré-escolas em relação às crianças de 4 a 6 anos e o número de escolas fundamentais em relação às crianças de 7 a 18 anos de idade no ano da realização da PB2005. [↑](#footnote-ref-12)
13. O número de crianças com idade entre 4 e 6 anos por município em 2000 foi obtido pelo Censo Demográfico de 2000 do IBGE. [↑](#footnote-ref-13)
14. As dez *dummies* criadas para o pai e para a mãe foram: Se pai ou mãe completou pelo menos a 4ª série, mas não completou a 8ª; Se pai ou mãe completou pelo menos a 8ª série, mas não completou o Ensino Médio; Se pai ou mãe completou pelo menos o Ensino Médio, mas não completou o Ensino Superior; e se a criança não conhece a formação do pai ou da mãe. [↑](#footnote-ref-14)
15. Website: [http://www.ipeadata.gov.br](http://www.ipeadata.gov.br/). [↑](#footnote-ref-15)
16. As estimações foram feitas partindo da população com a faixa etária desejada pelo Censo Demográfico 2000, e ponderadas pela variação total da população do município pelo IBGE e ponderadas novamente pela variação da faixa etária no estado pela PNAD 2005. Abaixo demonstramos a estimação da faixa etária 4-6 anos, como exemplo:

$$Crianças Município 2005\_{4-6 anos} = Crianças Município 2000\_{4-6 anos} × \left(\frac{População Município 2005\_{Total}}{População Município 2000\_{Total}}\right) × \left(\frac{\frac{Crianças Estado 2005\_{4-6 anos}}{População Estado 2005\_{Total}}}{\frac{Crianças Estado 2000\_{4-6 anos}}{População Estado 2000\_{Total}}}\right)$$

 [↑](#footnote-ref-16)
17. Efetuar a 4a série com 10 anos de idade seria uma seqüência normal para alunos sem repetência que houvessem entrado na escola aos 7 anos de idade conforme exigia a legislação à época. [↑](#footnote-ref-17)
18. Os resultados completos, com todos os estimadores de nossos modelos podem ser obtidos junto aos autores. Como modelo das regressões, encontra-seno Anexo A2 as regressões para as crianças na Idade escolar correta e que nunca repetiram. Os primeiros estágios dos modelos estão reportados em apenas uma tabela, pois as amostras e as especificações para as regressões com as provas de matemática e com as notas de português são exatamente as mesmas. [↑](#footnote-ref-18)