

# Capital Humano, Desigualdade e Ganhos de Proficiência: uma análise do Projovem

Flávia Chein

Cedeplar - Universidade Federal de Minas Gerais e Departamento de Economia- Universidade Federal de Juiz de Fora

Ricardo Freguglia

Departamento de Economia- Universidade Federal de Juiz de Fora

Fernando Tavares Júnior

Departamento de Ciências Sociais/Pesquisador CAED - Universidade Federal de Juiz de Fora

July 20, 2011

## Abstract

Este artigo avalia os determinantes do desempenho escolar dos alunos do Projovem – Programa Nacional de Inclusão de Jovens. Usando dados do Sistema de Monitoramento e Avaliação do Projovem (2007), analisamos os principais fatores capazes de determinar o processo de aprendizagem no ensino suplementar vis-à-vis ao ensino regular. Nossa análise é dividida em duas partes. Primeiro, estimamos uma função de produção educacional na qual o desempenho escolar do aluno (proficiência) é explicado por variáveis de características individuais, do contexto socioeconômico, de características familiares, da região onde o aluno frequentou o programa, além da inclusão dos efeitos específicos de núcleo – local de realização do Projovem. Em seguida, adotamos o procedimento do Heckman em dois estágios para tratar do problema de viés de seleção decorrente da evasão de alunos. Os principais resultados encontrados indicam que os efeitos fixos dos núcleos do programa explicam a maior parte da agregação de proficiência no ensino do Projovem.

*This paper evaluates the determinants of students school performance in Projovem-National Youth Inclusion Program. Based on data from the System of Monitoring and Evaluation of Projovem, we analyze the main factors that determine the process of learning in Projovem vis a vis in the regular school. Our analysis is carried out in two steps. First, we estimate an education production function in which the school performance is explained by individual characteristics, social and economics context, family characteristics (mainly parents' education) and region, besides local program characteristics (captured by fixed effects). In a second exercise, we adopt the Heckman two step approach in order to take into account selection bias due to data attrition. We find that students proficiency gains in the Projovem are mainly explained by core fixed effects.*

**Keywords:** Human Capital, Proficiency, education production function, education projects, Projovem

**Palavras-Chave:** Capital Humano, Proficiência; Função de produção educacional; Projetos de educação; Projovem

**JEL Classification:** J10; J13; I2; I21

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

# 1 Introdução

Nas últimas décadas consolidaram-se as análises de fatores sociais que influenciam a produção da proficiência. Dois modelos são mais freqüentes. O primeiro observa o resultado geral dos Sistemas de Ensino, avaliando os estudantes ao final dos ciclos escolares. No Brasil, tais etapas são o 5º e 9º anos do Ensino Fundamental e 3º ano do Ensino Médio. Na década atual disseminaram-se os sistemas estaduais de avaliação da educação e consolidaram-se os sistemas federais, em especial o Sistema de Avaliação da Educação Básica, o SAEB, iniciado nos anos 90 e que passou a disponibilizar uma escala comparável de resultados que é adotada também pelos sistemas estaduais<sup>1</sup>. Tal modelo afere se os egressos do Sistema de Ensino dispõem das habilidades previstas para cada etapa da escolarização e permite identificar deficiências, experiências de sucesso e instruir novas políticas públicas.

No entanto, como nesse modelo não se conhece a proficiência inicial, não se sabe quanto cada aluno “aprendeu”, ou desenvolveu sua proficiência, no período em que freqüentou determinada escola. Portanto, mostra-se um bom mecanismo para avaliar o conjunto do sistema, mas não necessariamente é eficiente para avaliar o efeito derivado de cada unidade do sistema em relação ao desenvolvimento da proficiência, tanto em termos de qualidade quanto de equidade. Certamente, todas as unidades têm responsabilidade com os resultados finais, mas aquelas que recebem estudantes com proficiência inicial mais baixa obviamente terão que vencer um obstáculo maior.

Um segundo modelo de análise do desenvolvimento de proficiência tem seu foco na proficiência agregada, ou o quanto cada aluno aprendeu em cada etapa da escolarização, ou que freqüentou seguidamente a mesma unidade de ensino (escola). No Brasil, o projeto “GERES – Estudo Longitudinal da Geração Escolar 2005” é um exemplo deste modelo. A partir de uma amostra de escolas de diferentes redes em cinco capitais estaduais, são monitorados e avaliados os processos de ensino e o desenvolvimento da aprendizagem. Outro exemplo é o Projeto “Entre Jovens”, que avalia um modelo de intervenção educacional em diferentes escolas a partir de uma amostra intencional, equivalente e gradativa.

Além destes projetos, outras políticas públicas passaram a monitorar a proficiência inicial e final de estudantes para acompanhar a evolução da aprendizagem, direcionar intervenções, avaliar projetos pedagógicos, analisar os processos de ensino e instruir novas políticas públicas. Um exemplo é a avaliação dos anos iniciais em Minas Gerais, que gerou um conjunto específico de crianças para atendimento especial e acompanhamento ano a ano. Outro exemplo é o Programa Nacional de Inclusão de Jovens – Projovem, que prevê aplicação de Avaliação Diagnóstica e Exame Final Nacional Externo, ambos comparáveis e equalizados na Escala SAEB.

O Projovem tem como objetivo central possibilitar a retomada dos estudos de jovens entre 18 e 24 anos que completaram a quarta série, mas não a oitava série do ensino fundamental. Mesmo com o aumento da freqüência escolar na última década, uma quantidade grande de jovens ainda não concluíram, sequer, o ensino fundamental. Estes jovens, além de precisarem voltar a estudar, precisam ter boa qualidade no ensino, para poderem ter maiores perspectivas futuras em relação ao mercado de trabalho. Por isso a importância de programas como o Projovem, que possibilita a retomada dos estudos dos jovens.

No Projovem, também há avaliações padronizadas pela TRI (Teoria de Resposta ao Item) e, diferentemente do SAEB, existem dados de proficiências do mesmo aluno em dois momentos do tempo, ou seja, na fase diagnóstica e na fase final. Desta forma, é possível estudar os determinantes do desempenho educacional e mostrar o impacto do programa na proficiência dos jovens. Pode-se também comparar as notas dos alunos deste programa com as notas dos alunos do ensino regular.

---

<sup>1</sup>Ver Andrade et al. (2000), ALBERNAZ et al. (2002), Pereira (2006).

Cabe salientar que o Projovem não objetiva, unicamente, a agregação de proficiência pelos jovens, sendo também importante a sua inclusão social, conscientização política, inserção no mercado de trabalho e a possibilidade de avanço nas etapas escolares. Entretanto, a agregação de proficiência é de extrema relevância para o avanço escolar do jovem. Isto porque existem evidências de que os estudantes que são mais bem avaliados nos testes de proficiência tendem a atingir níveis mais altos de estudos. Segundo Bishop (apud CURI, 2006) as notas obtidas pelos indivíduos são altamente correlacionadas com a frequência escolar futura e a conclusão de determinada etapa do processo de ensino é significativamente relacionado a melhores desempenhos escolares. Desta forma justifica-se a investigação deste trabalho, ressaltando-se que não há na literatura nacional nenhum estudo acadêmico relacionado ao Projovem.

Este trabalho tem o objetivo de avaliar os determinantes do desempenho escolar dos alunos do Projovem. Vale ressaltar que se adota como sinônimo de desempenho a proficiência, ou seja, a nota obtida pelo aluno em testes padronizados. Os principais resultados dos determinantes da avaliação diagnóstica confirmam evidências da literatura sobre educação regular: tem maior proficiência inicial quem é mais rico (ou menos pobre), branco, sofreu menos reprovações, tem pais alfabetizados e é mulher (leitura) ou homem (cálculo). Já o resultados de agregação de proficiência do Projovem apontam que a maior parte dos ganhos de proficiência são explicados pelo efeitos de regionais e/ou de núcleo. Há também evidências de que as mulheres agregam, no Projovem, em média, 4 pontos a mais de proficiência em relação aos homens.

Além dessa seção introdutória, este artigo encontra-se organizado em cinco seções. A segunda seção aborda uma descrição do Projovem, uma breve revisão da literatura sobre os fatores associados ao desempenho escolar, e sobre os fatos estilizados no Brasil. A terceira seção contém a estratégia empírica adotada, incluindo a descrição da base de dados e o modelo empírico. Na quarta seção analisam-se os principais resultados obtidos e, por fim, na quinta seção destacam-se as conclusões do artigo.

## 2 Referencial Teórico

### 2.1 Fatores Associados ao Desempenho Escolar

Os fatores que determinam o desempenho escolar pertencem a três grandes categorias: os associados à estrutura escolar, os associados à família e aqueles relacionados ao próprio aluno. Desta forma os trabalhos empíricos procuram estimar uma função de produção educacional, onde a proficiência pode ser explicada como função destes três grupos de fatores (Hanushek, 2003).

A partir do estudo empírico seminal de Coleman et al. (1966), que analisou a qualidade das escolas públicas e privadas nos Estados Unidos, vários estudos similares foram realizados na Inglaterra e na França, com resultados evidenciando o maior poder de explicação dos fatores extra-escolares para as desigualdades educacionais observadas. Segundo Soares (2004), estes resultados acabaram gerando certo pessimismo pedagógico, pois inicialmente levava à conclusão de que a escola teria pouco impacto no desempenho dos alunos e que as diferentes maneiras de se organizar a escola seriam igualmente eficientes.

Com o desenvolvimento de outros métodos de análise, principalmente na década de 1980, muitos trabalhos passaram a estimar os efeitos dos insumos escolares por outros modelos – diferentemente dos modelos de mínimos quadrados ordinários até então utilizados –, tais como o hierárquico. Em geral, estes trabalhos controlam os efeitos do “background” familiar e das características do próprio aluno, com o objetivo de saber os efeitos das variáveis escolares sobre a proficiência dos discentes. De forma geral, a literatura internacional encontra influência significativa das variáveis escolares no desempenho dos alunos.

Na literatura internacional, a análise dos determinantes do desempenho escolar tem destaque nos estudos de Lee (2001) e Hanushek (2003). Glewwe e Kremer (2005) fazem um levantamento dos diferentes tipos de análises empíricas realizadas nos países em desenvolvimento sobre o efeito de características de professores e escolas no desempenho dos alunos. Dentre eles destaca-se o trabalho de Hanushek (1995)<sup>2</sup>, cuja conclusão é a de que os recursos mensurados – com exceção das “facilidades de infra-estrutura” – não estão sistematicamente relacionados com o desempenho dos alunos nos países em desenvolvimento.

Em linhas gerais, como mostram os resultados dos diversos trabalhos, a avaliação de programas educacionais é extremamente importante do ponto de vista do gestor de políticas públicas. Isto porque, através da investigação dos determinantes do desempenho educacional, podem-se esclarecer quais recursos são importantes para a eficácia da escola e de programas educacionais (Hanushek, 2002).

## 2.2 Fatos Estilizados no Brasil

No Brasil, a avaliação educacional e a análise do efeito escola tiveram início na década de 1990, com a disponibilização dos dados do Sistema Nacional de Avaliação da Educação Básica (SAEB) pelo Instituto Nacional de Estudos e Pesquisas Educacionais (INEP). Os estudos educacionais de maior relevância no Brasil, portanto, foram realizados em um período que já havia passado o chamado pessimismo pedagógico das décadas de 60 e 70 (Barbosa et al., 2001a). Em geral, os estudos brasileiros se concentram na identificação das variáveis que influenciam o desempenho escolar dos alunos e a equidade dentro do sistema educacional<sup>3</sup>.

Como ilustrado na literatura, a precariedade da educação brasileira é constatada tanto na qualidade quanto na equidade. Além de ocupar sistematicamente o último lugar entre todos os países avaliados pelo PISA (Programme for International Student Assessment), internamente os indicadores brasileiros de qualidade da educação estão muito abaixo do previsto, com uma tênue tendência de melhora e com desigualdade perene. Os estudantes agregam pouca proficiência ano a ano e a cada ano aumentam as desigualdades. Os dados mais atuais da série histórica do SAEB (1995-2005) mostram que a proficiência média em Língua Portuguesa no final do Ensino Fundamental caiu 24 pontos em 10 anos. Além disso, entre a quarta e a oitava séries fundamentais, a agregação média é de apenas 15 pontos ao ano, menos de um nível na escala de proficiência, que abrange 25 pontos. Em Matemática, a agregação média é ainda menor, 14 pontos por ano.

O Relatório Nacional do PISA, que permite análises contextuais mais amplas, evidencia que o baixo desempenho é agravado por sua associação com diferenças regionais e sócio-econômicas. Tais resultados reforçam as evidências sobre a baixa qualidade dos sistemas de ensino no Brasil.

Várias políticas públicas tentam reverter esse quadro. Um exemplo de relativo sucesso é a melhoria do fluxo verificada entre os anos 80 e 2000. No entanto, a priorização do fluxo comprometeu a qualidade da educação. A diminuição da desigualdade em algumas redes de ensino, com estratégias de recuperação, turno integral e apoio a estudantes em atraso, muitas vezes, também revela graves limites em relação ao alcance de metas de aprendizagem: mais alunos chegam ao final dos ciclos, mas o rendimento médio tende a ser mais baixo.

Projetos de aceleração de aprendizagem, com turnos alternativos, em geral noturnos, e organização curricular diversa, também experimentam tais desafios: conseguem diminuir a defasagem idade-série e recuperar estudantes que haviam abandonado a escola, mas seus resultados são em média 15 a 20 pontos abaixo dos alunos regulares, como observado em experiências

---

<sup>2</sup>Hanushek (1995) faz uma revisão de 96 estudos sobre os determinantes do desempenho dos alunos.

<sup>3</sup>Franco et al. (2007), Soares e Andrade (2006), Albernaz et al. (2002), Rios-Neto et al. (2002), Menezes-Filho (2007), Alves (2008), Machado et al. (2008), Barbosa et al. (2001b), Dwyer et al. (2007), Alves e Soares (2007).

em Minas Gerais e no Rio de Janeiro. Em geral, observa-se grande dificuldade de as políticas públicas conciliarem inclusão, qualidade e equidade.

Este cenário complexo motivou a investigação aprofundada dos dados do Projovem por se tratar de mais uma política que objetiva superar o quadro negativo descrito acima. Seu core está ligado ao que chamam de Projeto Pedagógico Integrado (PPI). Sem retirar os elementos característicos da escola regular (professor, aluno, sala de aulas, turno médio de 4 horas presenciais por dia, material didático, provas, etc.), o PPI propõe sua reorganização sob “novos princípios”.

## 3 Estratégia Empírica

### 3.1 O PROJOVEM

Com o objetivo de atender uma parcela grande de jovens entre 18 e 24 anos que completaram a quarta série, mas não a oitava série do ensino fundamental, foi criado em 2005, pela coordenação da Secretaria Geral da Presidência da República em parceria com o Ministério da Educação, o Ministério do Trabalho e Emprego e o Ministério do Desenvolvimento Social e Combate à Fome, o Projovem (Programa Nacional de Inclusão de Jovens: Educação, Qualificação e Ação Comunitária). A motivação do surgimento deste programa foi a constatação da existência de exclusão para uma parcela significativa de jovens entre 18 e 24 anos que não completaram o ensino fundamental. Segundo os dados da Pesquisa Nacional por Amostra de Domicílios (PNAD) de 2003 – época em que se investigava a possibilidade de criação do programa – cerca de quatro milhões de jovens com idades entre 18 e 24 anos tinham apenas entre quatro e sete anos de escolaridade. O programa busca, portanto, oferecer ao jovem a oportunidade de recuperar e continuar seus estudos, além de possibilitar novos horizontes para sua inclusão no mercado de trabalho (Relatório Parcial de Avaliação do Projovem, 2008).

Atuando em todas as 27 capitais brasileiras, o Projovem atendeu a 191.500 jovens no período de 2005 a 2007. Além do condicionante de idade para a realização do programa, em que o estudante deve ter entre 18 e 24 anos, o aluno deve ter ao menos quatro anos de estudo no ensino regular<sup>4</sup>.

### 3.2 Base de Dados

Os dados são originais do Sistema de Monitoramento e Avaliação (SMA) do Projovem. As principais bases utilizadas foram de matrícula e rendimento acadêmico (avaliações). Os principais dados de matrícula são demográficos: sexo, cor, ano de nascimento, núcleo/UF, data de início das aulas, estado civil e fecundidade. Os dados relativos ao desempenho nos testes e avaliações são expressos de duas formas: (i) Escores equalizados na escala SAEB de acordo com a Teoria da Resposta ao Item; (ii) Número total de acertos em cada teste de acordo com a Teoria Clássica dos Testes.

Para consolidação da base de dados para análise, tomou-se como referência a base total de matrículas, considerando os jovens que chegaram ao final do curso. Como o objeto de análise é a proficiência agregada como proxy de uma função de produção da educação, só é possível calcular este dado para os núcleos que tiveram aplicação da Avaliação Diagnóstica em larga

---

<sup>4</sup>É importante ressaltar que, apesar de o Projovem ter semelhanças com o programa de Educação de Jovens e Adultos (EJA), pelo fato de ser um programa voltado para a recuperação escolar de Jovens, tem objetivos diferentes deste. A principal diferença consiste no fato de o Projovem, além de ser voltado para recuperação escolar, também se preocupa com a profissionalização e a inclusão social do jovem. Ao entrar no programa, o jovem recebe uma bolsa de R\$ 100,00, além de concluir o ensino fundamental e receber treinamento profissional. Outro benefício é a oportunidade de fazer aulas de informática e de cidadania.

escala, o que ocorreu até fins de 2006. Assim, são consideradas as cinco primeiras entradas (61.402 alunos), entre julho de 2005 e dezembro de 2006, que representam a primeira fase do Projovem (59% dos concluintes) e para as quais há dados mais completos. Desses 61.402 alunos, foram eliminados aqueles alunos que não efetuaram a avaliação diagnóstica, para os quais, portanto, não é possível observar a proficiência, ou seja, nossa amostra é composta pelos alunos de núcleos com entradas até dezembro de 2006, e que realizaram a avaliação diagnóstica de matemática e português.

Após os filtros iniciais, foram feitos sucessivos testes e procedimentos de crítica de dados relativos aos resultados gerais de proficiência. Como é relatado na literatura, os processos de estimação de proficiência merecem atenção especial, com correção de lançamentos indevidos, duplicados, inconsistentes e limpeza da base, sem que isso, no entanto, afetasse suas características principais (medidas de tendência central, dispersão, normalidade, variância, etc). Entre a população de referência e a amostra com dados completos de proficiência, foram feitos ainda testes de viés.

Os fatores testados foram: ano de nascimento, gênero, cor, estado civil, última série cursada, pater-maternidade, unidade da federação (e região), além da proficiência agregada, inicial (avaliação diagnóstica), intermediária e final (exame final nacional externo), tanto em português quanto em matemática. A partir dos testes efetuados, observou-se que não há viés significativo em nenhuma das características analisadas, exceto uma leve oscilação na distribuição geográfica. Tal desequilíbrio não afetou nenhuma das demais características analisadas, o que sinalizou para uma distribuição aleatória dos demais elementos. Corrigidos pelo peso, o percentual de cada UF tornou-se equivalente aos observados na distribuição da população e, como já destacado anteriormente, as demais características permaneceram estáveis. A amostra final considerada nesse artigo totalizou 37.378 alunos que realizaram a avaliação diagnóstica. Para a agregação de proficiência, foram considerados alunos que, além das avaliações diagnósticas em Português e Matemática, também fizeram as avaliações finais. Nesse caso, a amostra para avaliação da agregação de proficiência em Matemática foi de 20.775 e de 20.776, em Português.

### 3.3 Análise Descritiva

Como foi dito anteriormente, o Projovem foi formulado como uma política pública capaz de monitorar a proficiência inicial e final de estudantes para acompanhar a evolução da aprendizagem, direcionar intervenções, avaliar projetos pedagógicos, analisar os processos de ensino e instruir novas políticas públicas. Nesse sentido, os dados do Sistema de Monitoramento e Avaliação do Projovem possibilitam analisar não apenas os determinantes da proficiência, mas também os fatores que afetam a agregação de proficiência.

Logo, com base nas informações desse Sistema de Monitoramento e Avaliação, vamos analisar as seguintes hipóteses:

1. A população atendida pelo Projovem apresenta características resultantes da baixa qualidade do sistema educacional.
2. Os jovens atendidos pelo Programa apresentam agregação de proficiência superior à média das escolas regulares
3. Os fatores sócio-econômicos diagnosticados como indutores de desigualdade educacional apresentam coeficientes negativos ou não significativos em relação à proficiência agregada.
4. Tais jovens chegam ao final do curso com proficiência equivalente à obtida pela média dos estudantes das redes municipais das respectivas capitais.

A tabela 1 descreve o desempenho médio geral dos alunos do Projovem e as demais características analisadas neste trabalho. Quanto às características individuais, verifica-se a focalização do Projovem: alunos com idade média superior a 20 anos, sendo mais de 15% casados e cerca de 46% já possuem filhos.

[Inserir tabela 1]

A proficiência dos alunos, ao ingressarem no Projovem, é, em média, de 192 pontos em Matemática e 195 em Português. De acordo com dados do SAEB, em 2005, a proficiência média em Português era de 172,3 pontos e 231,9, respectivamente na 4<sup>a</sup>. e 8<sup>a</sup>. séries do ensino fundamental. Quanto à proficiência em Matemática, dados mostram que, na 4<sup>a</sup>. série do ensino fundamental a mesma fica em torno de 182 pontos, não chegando a 250 pontos ao final da 8<sup>a</sup>. Série (239,5). Ao concluírem os estudos no Projovem, os alunos agregam, em média, cerca de 23 pontos em Matemática e 33 pontos em Português, tomando-se como referências as chamadas avaliações diagnósticas e os exames finais. Estudos para o Brasil mostram que, entre a quarta e a oitava séries fundamentais, a agregação média é de apenas 15 pontos ao ano, menos de um nível na escala de proficiência, que abrange 25 pontos. Em Matemática, a agregação média é ainda menor, 14 pontos por ano.

As figuras 1 e 2 trazem o desempenho na proficiência inicial (diagnóstica) e final dos alunos do Projovem por cor e sexo. O desempenho na avaliação diagnóstica em Português e Matemática parece ser um pouco melhor para os homens e brancos, comparativamente a mulheres e não brancos. Contudo, tais diferenciais são bastante reduzidos quando se analisa a agregação de proficiência ao final do programa. Contudo, a agregação deveria ser maior naqueles grupos com pior desempenho inicial para que houvesse uma redução da desigualdade de proficiência.

[Inserir figuras 1 e 2]

De outro lado, a análise regional do desempenho dos alunos do Projovem é apresentada na figura 3.

[Inserir figura 3]

É possível perceber, pela análise da figura 3, que as curvas de distribuição da proficiência na avaliação diagnóstica de Português e Matemática nas Regiões Sul, Sudeste e Centro-Oeste encontram-se mais à direita comparativamente às Regiões Norte e Centro-Oeste, o que indica um desempenho melhor dos alunos nas três primeiras regiões.

Já na agregação de proficiência em Matemática, o Sudeste e Centro-Oeste parecem se sair um pouco pior, com curvas de distribuição uma pouco mais à esquerda comparativamente às demais regiões. Contudo, deve-se ressaltar que, no caso de Português, a agregação parece ser maior na Região Nordeste, o que indicaria uma diminuição da desigualdade regional em termos de proficiência a partir do Projovem.

Complementarmente, a tabela 2 aponta para um desempenho em Matemática e em Português abaixo da média nacional, nos Estados do Maranhão, Ceará, Piauí, Rio Grande do Norte, Bahia, Paraíba, Pernambuco e Alagoas, na Região Nordeste, como também dos Estado do Tocantins, Roraima, Rondônia e Amapá, na Região Norte.

[Inserir tabela 2]

Ainda no que se refere aos resultados da tabela 2, cabe lembrar que, no Projovem, as Unidades de Federação referem-se, basicamente, as capitais ou regiões metropolitanas.

No que tange às características sócio-econômicas dos alunos, a tabela 1 evidencia um baixo nível sócio-econômico dos alunos, que, em sua grande maioria, apresentam acesso apenas ao nível básico de infra-estrutura domiciliar. A análise de tais características torna-se mais clara por meio de uma análise de componentes principais das diversas variáveis de nível sócio-econômico, com o intuito de reduzir o problema a apenas uma dimensão ou componente. Os resultados dessa análise são apresentados nas tabelas 3. É possível explicar mais de 20% da variância do conjunto de variáveis sócio-econômicas em um componente.

[Inserir tabela 3]

A tabela 4 apresenta a composição do primeiro vetor resultante da análise de componentes principais. Verifica-se que as variáveis que menos contribuem para o componente são a existência de calçamento na rua do domicílio, bem como a posse de mais de um automóvel. Ressalte-se que, em média, 66% dos domicílios estão localizados em ruas com calçamentos, e, apenas 4% possuem mais de um automóvel no domicílio.

[Inserir tabela 4]

A figura 4 traz uma análise de desempenho por nível de riqueza. Em relação à proficiência prévia, é possível verificar uma relação crescente da proficiência média com a riqueza para os indivíduos com mais pobres (riqueza menor que zero), com reversão do sinal para níveis positivos de riqueza. No caso da agregação de proficiência, a relação se inverte, sendo decrescente para níveis de riqueza negativa e crescente para níveis de riqueza positiva.

[Inserir figura 4]

### 3.4 Modelo Empírico

A análise da produção e do desenvolvimento da proficiência em determinada área do conhecimento retrata a aquisição gradativa de competências e habilidades fundamentais para o domínio cognitivo necessário à superação dos desafios cotidianos que mobilizam tais saberes. Para estimar a proficiência, cada estudante é testado através de itens preparados especificamente para este tipo de análise, de acordo com os princípios que instruem a Teoria da Resposta ao Item. Tais itens expressam situações problema unidimensionais que avaliam a habilidade e/ou a competência de o estudante resolver tal desafio ao mobilizar conhecimentos específicos. São controladas as probabilidades de acerto ao acaso (“chute”), inibidos os empecilhos de interpretação do problema exposto, anuladas as “armadilhas ou pegadinhas”, entre outros fatores que podem relacionar o acerto ou o erro ao item a outra competência que não a avaliada. Tal percurso teórico e tecnológico levou a proficiência a ser interpretada como o fenômeno observável mais próximo da atividade fim da educação, a aprendizagem, superando as análises tradicionais que se dedicam aos resultados gerais dos Sistemas de Ensino: a escolarização. Por isso, a proficiência é mais relevante para a análise dos efeitos da educação na economia e no comportamento social do que os anos de estudo concluídos.

Neste sentido, a forma geral dada à função de produção educacional que será estimada neste trabalho segue o padrão da literatura (Hanushek, 2003), como pode ser observada na equação (1).

$$y = F(c, m, s, r, e) \tag{1}$$

onde:

$c$  = características individuais;

$m$  = características familiares;



$s$  = socioeconômicas;  
 $r$  = regionais;  
 $e$  = núcleo escolar.

O modelo proposto é descrito pela equação (2):

$$Profic_j = c_j\beta_1 + m_j\beta_2 + s_j\beta_3 + r_j\beta_4 + e_j\beta_5 + \varepsilon_j \quad (2)$$

onde:

$Profic_j$  = proficiência do jovem nas avaliações padronizadas;

$c$  = vetor de características individuais do jovem;

$m$  = vetor de características familiares do jovem;

$s$  = vetor de características socioeconômicas do jovem;

$r$  = vetor de características regionais do núcleo frequentado pelo jovem;

$e$  = efeito específico não-observado dos núcleos escolares;

$\varepsilon$  = termo de erro;

$\beta$  = vetor de parâmetros de interesse, tal que  $\beta' = (\beta'_1, \beta'_2, \beta'_3, \beta'_4, \beta'_5)$ .

A estimação da equação (2), tendo por variável dependente o resultado do aluno na avaliação diagnóstica, por Mínimos Quadrados Ordinários, sob hipóteses usuais, gera resultados consistentes e não viesados.

Entretanto, se quisermos mensurar a agregação de proficiência, é possível que a estimação por MQO da equação (2) seja inviabilizada pela existência de viés de seleção, isso porque só é possível observar a agregação de proficiência de indivíduos que realizam a avaliação diagnóstica e a avaliação final, mas parte dos alunos evadem durante o programa, não realizando a avaliação final. Com o objetivo de tratar desse problema, propõe-se estimar a (2) pelo procedimento de dois estágios do Heckman para correção do viés de seleção amostral (Heckman, 1979).

## 4 Resultados

As tabelas 5 e 6 apresentam as estimações por mínimos quadrados ordinários da equação (2) para as proficiências em Português e Matemática, respectivamente. Duas diferentes medidas de proficiência são consideradas: a proficiência na avaliação diagnóstica (colunas 1, 2 e 3) e a proficiência agregada pelos alunos do Projovem (colunas 4, 5 e 6). Como os resultados da avaliação diagnóstica são resultantes de um ensino anterior ao Projovem – em geral provenientes do sistema de ensino regular brasileiro –, é possível entender os resultados das colunas (1), (2) e (3) como sendo resultados da escola regular em contraposição aos resultados das colunas (4), (5) e (6) como resultados da agregação de proficiência decorrente de um programa especial de educação – o Projovem.

[Inserir tabelas 5 e 6]

De uma forma geral, os resultados das colunas (1), (2) e (3) das tabelas 5 e 6 reproduzem as principais evidências da literatura empírica: melhor desempenho dos brancos em Português e Matemática; mulheres têm um melhor desempenho em Português e um pior desempenho em Matemática; a reprovação afeta negativamente o desempenho, assim como o fato dos pais não serem alfabetizados.

Constata-se, ainda, um efeito significativo do componente de riqueza ou nível sócio-econômico seja na proficiência em Matemática, seja na proficiência em Português. Em ambos os casos, há uma reversão de sinal do coeficiente quando se comparam os resultados da coluna (1) com aqueles das colunas (2) e (3). Essa reversão está relacionada com a tentativa de se identificar os chamados efeitos regionais (dummies de estados brasileiros) e os efeitos de núcleo (dummies

de locais de funcionamento das turmas). Enquanto o modelo básico – expresso na coluna (1) – reporta um coeficiente positivo para o componente de riqueza, os modelos das colunas (2) e (3) – que consideram os efeitos regionais e os efeitos de núcleo, respectivamente – reportam um coeficiente de riqueza negativo. Tal fato pode estar relacionado à desigualdade sócio-econômica dos estados brasileiros pelos quais os núcleos estão distribuídos, isto é. o componente de riqueza pode estar captando o fato de alunos com menores (maiores) componentes de riqueza residirem em Estados mais (menos) pobres. Quando esta desigualdade regional é captada via inclusão das dummies de estado ou de núcleo, há indícios de um efeito negativo do componente de riqueza sobre a proficiência do aluno na avaliação diagnóstica. Essa mudança de sinal do componente de riqueza pode estar refletindo o fato da desigualdade regional ser mais forte do que desigualdade interpessoal de riqueza, ou melhor, entre indivíduos dentro de um mesmo Estado ou núcleo. Vale ressaltar a elevada significância estatística para ambos os efeitos fixos, dos estados e dos núcleos (colunas (2) e (3)). Ademais, a inclusão destes efeitos incrementa o poder de explicação do modelo, cujo R-quadrado passa para 13%, tal resultado evidencia que grande parte da proficiência pode ser explicada por efeitos regionais ou de núcleo.

Já os modelos para identificação dos determinantes da agregação de proficiência no Projovem são menos claros, como retrata a análise das colunas (4), (5) e (6). Para os resultados de Português (tabela 5), observa-se um efeito negativo da não reprovação ou de uma só reprovação comparativamente aos que reprovaram muitas vezes. Do ponto de vista das variáveis individuais, saliente-se um efeito significativo da alfabetização dos pais. São observados aumentos na agregação de proficiência para os alunos cujos pai e mãe não sabem ler e escrever e para aqueles cujo apenas o pai ou mãe sabem ler e escrever, comparativamente aos alunos que têm pai e mãe que sabem ler e escrever. Esse resultado é coerente com aquele encontrado na avaliação diagnóstica: filhos de pais analfabetos iniciariam com menor nível de proficiência e, portanto, agregariam mais. De outro lado, há um efeito positivo e significativo do componente de riqueza sobre a agregação de proficiência. Tal resultado é, de certa forma, contra-intuitivo, tendo em vista que se espera uma maior agregação de proficiência para os mais pobres, e não para os mais ricos (ou menos pobres), o que levaria a uma redução das desigualdades pela participação no Projovem.

Por outro lado, assim como encontrado nos resultados da avaliação diagnóstica, os efeitos fixos regionais e dos núcleos (colunas 5 e 6) parecem explicar uma parcela significativa da agregação da proficiência em relação ao modelo básico (coluna 4). A inclusão destes efeitos incrementa o poder de explicação do modelo, cujo R-quadrado passa para 10% na agregação de proficiência em matemática e para 16% em português.

Os resultados para agregação de proficiência em Matemática seguem um padrão similar ao de Português, com alguma especificidade. No caso da Matemática, há um efeito também positivo, mas não significativo do componente de riqueza sobre a agregação de proficiência. Assim, o nível sócio-econômico parece não estar relacionado com o rendimento do aluno.

Conforme descrito na seção 3.2 referente à base de dados do Projovem, a amostra final considerada nesse artigo totalizou 37.378 alunos que realizaram a avaliação diagnóstica. Porém, a amostra para avaliação da agregação de proficiência em Matemática foi de 20.775 e de 20.776, em Português. Esta redução do número de observações é decorrente do processo de evasão de alunos, o qual é de fato bastante significativo no ensino suplementar em geral. Portanto, desconsiderando os jovens que não chegaram ao final do curso, isto é, evadiram, pode-se deparar com um processo de seleção amostral. Como o objeto de análise é a proficiência agregada como *proxy* de uma função de produção educacional, a equação de proficiência deveria levar em conta este processo seletivo.

A tabela 7 apresenta a estimação dos determinantes da proficiência agregada (equação 2) segundo o procedimento de dois estágios de Heckman para a proficiência agregada em Português e Matemática. A hipótese de identificação utilizada é que o processo de agregação de

proficiência é não-linear, ou seja, estima-se uma equação de seleção cujas variáveis explicativas são exatamente as mesmas da equação principal, isto é, não há restrições de exclusão na equação estrutural. A coluna (1) mostra os parâmetros estimados para a proficiência agregada em Português, mantidas as mesmas variáveis explicativas do modelo de agregação de proficiência das tabelas 5 e 6, mas incluindo os efeitos fixos de núcleo. Na coluna ao lado, estão os resultados da equação de seleção, que envolve o segundo passo do procedimento de Heckman, cujo coeficiente lambda (razão inversa de Mills) é estimado ao final da tabela. A mesma estrutura é mantida para a proficiência agregada em Matemática, na coluna (2). Nas colunas (3) e (4), testa-se uma nova forma funcional da equação de proficiência, bem como da equação de seleção, as quais passam a incluir variáveis explanatórias adicionais: número de filhos e estado civil. A hipótese de identificação continua sendo baseada na não-linearidade do processo de agregação de proficiência.

[Inserir tabela 7]

Pela coluna (1), pode-se observar que a significância e o sinal dos coeficientes estimados para a agregação de proficiência de Português pouco variam em relação aos coeficientes de MQO (coluna (6) da tabela 5). O coeficiente lambda estimado não é estatisticamente significativo a níveis convencionais de significância. Fato semelhante verifica-se para a agregação de proficiência em Matemática (coluna 2), bem como para as especificações das colunas (3) e (4), que incorporam as variáveis de número de filhos e estado civil para as proficiências de Português e Matemática, respectivamente.

Dadas as limitações da especificação do modelo baseada na hipótese simples de não-linearidade, testou-se uma especificação adicional da metodologia de Heckman que inclui a variável trabalho (indicando se o aluno trabalha ou não) na equação de seleção. A hipótese de identificação, portanto, envolve a exclusão de uma variável – trabalho – em relação à equação estrutural, ou, a adoção da variável trabalho como um instrumento na estimação da probabilidade de continuar no Projovem, isto é, de realizar a prova final, dado que fez a avaliação diagnóstica. A identificação do viés de seleção baseia-se na hipótese de que a variável trabalho afeta a agregação de proficiência apenas por meio da decisão de continuar no Projovem, ou realizar a prova final.

Os principais resultados obtidos podem ser observados na tabela 8. Na coluna (1), encontram-se os resultados para a proficiência agregada em Português e na coluna (2) estão os resultados para Matemática.

[Inserir tabela 8]

Os resultados mais importantes são a significância do coeficiente lambda (razão inversa de Mills) para a equação de proficiência agregada de Português. Isto evidencia a relação do trabalho com a evasão escolar dos alunos do Projovem, captando um processo seletivo no qual a desconsideração dos alunos evadidos gera um viés nas estimativas dos determinantes da proficiência agregada em Português. O mais notório resultado está relacionado ao coeficiente significativo de 4,09 para a dummy de sexo feminino, indicando as mulheres agregam mais de 4 pontos na proficiência em Português que os homens. Comparativamente à coluna (3) da tabela 7, este mesmo coeficiente era não significativo aos níveis convencionais.

Para a equação de proficiência agregada em Matemática (coluna 2), o coeficiente lambda estimado não foi significativo, indicando que o processo seletivo não se verifica para os estudantes de matemática no Projovem.

Por fim, os resultados aqui apresentados são apenas exploratórios e merecem uma maior investigação futura.

## 5 Conclusão

Este artigo é um esforço inicial para compreensão dos determinantes da proficiência pregressa e da agregação de proficiência em projetos especiais de educação que, sob diferentes formatos, vêm, há muito, sendo implantados em países em desenvolvimento com o intuito de minimizar as históricas desigualdades de acesso ao sistema regular de ensino. Para tanto, faz uso das informações provenientes do Sistema de Monitoramento e Avaliação do Projovem.

Como salientado ao longo desse trabalho, são várias as evidências sobre a baixa qualidade dos sistemas de ensino no Brasil, bem como sobre os desafios em relação à elevada e progressiva desigualdade, ao fluxo escolar ineficiente e excludente, entre outros. Como resultado, tem-se uma baixa agregação de proficiência, uma elevada defasagem idade-série e baixa inclusão após a idade escolar.

Os resultados aqui encontrados, além de reproduzirem evidências da literatura empírica em termos de determinantes da proficiência, apontam para um importante papel das variáveis de núcleo/turma como explicação para o desempenho escolar do aluno. É importante ressaltar que, como se tem informações sobre a proficiência no momento de ingresso e no momento do exame final, pode-se obter a real informação da agregação de proficiência dos participantes do programa, ou seja, quanto cada aluno “aprendeu”, ou desenvolveu sua proficiência, no período em que frequentou o Projovem.

Na análise do processo de seletividade decorrente das evasões ocorridas durante o programa, via procedimento de dois estágios de Heckman, não há evidências de seleção amostral quando não se utilizam restrições de exclusão na equação estrutural. Como forma de avançar na estratégia de identificação de um possível viés de seleção dos alunos que realizam a avaliação final dado que realizaram a avaliação diagnóstica, resultados, testou-se uma especificação adicional da metodologia de Heckman que inclui a variável trabalho (indicando se o aluno trabalha ou não) na equação de seleção. O principal resultado evidencia a relação do trabalho com a evasão escolar dos alunos do Projovem, captando um processo seletivo no qual a desconsideração dos alunos evadidos gera viés nas estimativas dos determinantes da proficiência agregada em Português (o mesmo não se verifica para Matemática). Ressalte-se, após a correção de Heckman, o coeficiente significativo de 4,09 para a dummy de sexo feminino, indicando que as mulheres agregam mais de 4 pontos na proficiência em Português em relação aos homens.

## 6 Referências Bibliográficas

### References

- [1] ALBERNAZ, Ângela; FERREIRA, Francisco H. G; CRESO, Franco. Qualidade e equidade no ensino fundamental brasileiro. *Pesquisa e Planejamento Econômico - IPEA*, v. 32, n. 3, p. 453-476, 2002.
- [2] ALVES, Fátima. Políticas educacionais e desempenho escolar nas capitais brasileiras. *Cadernos de Pesquisa- PUC-RIO*, v. 38, n. 134, p. 413-440, maio/ago. 2008
- [3] ALVES, Maria Teresa Gonzaga; SOARES, José Francisco. Efeito-escola e estratificação escolar: o impacto da composição de turmas por nível de habilidade dos alunos. *Educação em Revista*, v. 45, p. 25-58. Jun. 2007.
- [4] ANDRADE, Dalton Francisco de; TAVARES, Heliton Ribeiro; VALLE, Raquel Cunha. Teoria da resposta ao item: conceitos e aplicações. In: *14º SINAPE, ABE - Associação Brasileira de Estatística*, São Paulo, 2000.

- [5] ANDRADE, Josemberg e LAROS, Jacob. Fatores associados ao desempenho escolar: Estudo multinível com os dados do SAEB/2001. *Psicologia: Teoria e Pesquisa*. V.23, n. 1, PP. 033042. Jan/mar. 2007.
- [6] BARBOSA, Maria Eugenia; BELTRÃO, Kaizô Iwakami; FERNANDES, Cristiano; SANTOS, Denis; SUÁREZ, Mayte; ANDRADE, Adler do Couto. O SAEB - Sistema Nacional de Avaliação da Educação Básica: objetivos, características e contribuições na investigação da escola eficaz. *Revista Brasileira de Estudos de População*. V. 18, n. 1\2, jan.\dez. 2001a.
- [7] BARBOSA, Maria Eugenia Ferrão; FERNANDES, Cristiano. A escola brasileira faz diferença? Uma investigação dos efeitos da escola na proficiência em Matemática dos alunos da 4a série. In: C. Franco (org), *Promoção, ciclos e avaliação educacional*. ArtMed, Curitiba (Brasil). 2001b.
- [8] BRASIL. Programa Nacional de Inclusão de Jovens - PROJOVEM. Relatório Parcial de Avaliação do Projoovem. Juiz de Fora MG, FADEPE: 2008
- [9] BUCHMANN, Claudia; HANNUM, Emily. Education and stratification in developing countries: a review of theories and research. *Annual Reviews Sociology*, v. 27, p. 77-102, 2001.
- [10] COLEMAN, J. S. et al.. Equality of Educational Opportunity. Washington DC: US Government Printing Office, 1966.
- [11] CURI, Andréa Zaitune; MENEZES-FILHO, Naércio Aquino. A relação entre desempenho escolar e os salários no Brasil. IBMEC, WPE-03, 2006.
- [12] DWYER, Tom; WAINER, Jacques; DUTRA, Rodrigo Silveira; COVIC, André; MAGALHÃES, Valdo B; FERREIRA, Luiz Renato Ribeiro; PIMENTA, Valdiney Alves; CLAUDIO, Kleucio. Desvendando Mitos: Os computadores e o desempenho no sistema escolar. *Educ. Soc.*, Campinas, vol. 28, n. 101, p. 1303-1328, set./dez. 2007.
- [13] FIGLIO, David N. E; KENNY, Lawrence. "Individual Teacher Incentives And Student Performance," *NBER Working Papers* 12627, National Bureau of Economic Research.
- [14] FRANCO, Creso; MANDARINO, Mônica; ORTIGÃO, Maria Isabel. O projeto pedagógico e os resultados escolares. *Pesquisa e Planejamento Econômico - IPEA*, v. 32, n. 3, p. 477
- [15] FRANCO, Creso; ORTIGÃO, Isabel; ALBERNAZ, Ângelo; BONAMINO, Alicia; AGUIAR, Glauco; ALVEZ, Fátima; SÁTYRO, Natália. Qualidade e equidade em educação: Reconsiderando o significado de "fatores intra-escolares". *Avaliação de Políticas Públicas Educacionais*, v.15, n. 55, p. 277-298, abr/jun. 2007.
- [16] HAMBLETON, R.K. (Ed.). *Applications of Item Response Theory*. Vancouver (BC): Educational Research Institute of British Columbia. 1983.
- [17] HAMBLETON, R.K., SWAMINATHAN. H., ROGERS. H.J. *Fundamentals of Item Response Theory*. London: Sage, 1991.
- [18] HAMBLETON, R.K. Principles and selected applications of item response theory. In: LINN, R. L. *Educational Measurement*. 3ed. Phoenix: American Council on Education, Oryx Press, p.147-200, 1993.
- [19] HANUSHEK, E. A. The failure of input-based schooling policies. *The Economic Journal*, v. 113, p.64-98, Feb.2003.

- [20] HANUSHEK, E. Publicly provided education. In: Alan J. Auerbach and Martin Feldstein (ed.), *Handbook of Public Economics*, Amsterdam: North-Holland, 2002, p.2047-2143.
- [21] HECKMAN, J. (1979) Sample selection bias as a specification error. *Econometrica*, 47, 153-61
- [22] KLEIN, Ruben. Como está a educação no Brasil? Ensaio: avaliação de políticas públicas da educação. Rio de Janeiro, v. 14, n.51, p. 139-172, abr./jun. 2006.
- [23] MACHADO, Ana. Flavia; MORO, Sueli; MARTINS, Ludiemy; RIOS, Juan. Qualidade do Ensino em Matemática: Determinantes do Desempenho de Alunos em Escolas Públicas Estaduais Mineiras. *Associação Nacional de Pós-graduação em Economia- ANPEC*, v.9, n.1, p.23-45, jan/abr 2008
- [24] MACHADO, Danielle Carusi; GONZAGA, Gustavo. O impacto dos fatores familiares sobre a defasagem idade-série de crianças no Brasil. *Revista Brasileira de Economia - Fundação Getulio Vargas*, v. 6, n. 4, p. 449-476, 2007.
- [25] MENEZES-FILHO, Naércio Aquino: “Os Determinantes do Desempenho Escolar do Brasil,” Mimeo, Instituto Futuro Brasil, 2007.
- [26] PASQUALI, L. Teoria da Resposta ao Item - IRT: uma introdução. In: PASQUALI, LUIZ (Org.). *Teoria e Métodos de Medida em Ciências do Comportamento*. Brasília: Laboratório de Pesquisa em Avaliação e Medida / Instituto de Psicologia / UnB: INEP, 1996. 432p.
- [27] PEREIRA, Danielle Ramos de Miranda. Fatores associados ao desempenho escolar nas disciplinas de matemática e português no ensino fundamental: uma perspectiva longitudinal. 2006. P.118. Tese (Doutorado em Demografia) - Centro de Desenvolvimento e Planejamento Regional Universidade Federal de Minas Gerais, Belo Horizonte, 2006
- [28] RIOS-NETO, Eduardo. L. G.; CÉSAR, Cibele Comini; RIANI, Juliana de Lucena Ruas. Estratificação educacional e progressão educacional por série no Brasil. *Pesquisa e Planejamento Econômico - IPEA*, v. 32, n. 3, p. 395-416, 2002.
- [29] SANTIBAÑEZ, LUCRECIA. Why we should care if teachers get A's: Teacher test scores and student achievement in México. *Economics of Education Review*. P.510-520. 2006.
- [30] SOARES, Jose Francisco. O efeito da escola no desempenho cognitivo de seus alunos. *Revista Electrónica Iberoamericana sobre Calidad. Eficacia y Cambio em Educación - REICE*, v. 2, n. 2, p. 83-104, 2004.
- [31] SOARES, Jose Francisco; ANDRADE, Renato Júdice de. Nível socioeconômico, qualidade e equidade das escolas de Belo Horizonte. *Avaliação de políticas públicas educacionais*, Rio de Janeiro, v.14, n.50, p.107-p.126, jan.\mar. 2006.
- [32] SOARES, Jose Francisco e COLLARES, Ana Cristina Murta. Recursos Familiares e o Desempenho Cognitivo dos Alunos do Ensino Básico Brasileiro. *Revista de Ciências Sociais*. Rio de Janeiro, vol. 49, n°3, PP. 615 a 418, 2006.

Tabela 1

Estatística Descritiva			
	Média	Desvio Padrão	Descrição
Sexo Feminino	0.57	0.50	Dummy indicativa do gênero feminino a partir de informações cadastrais do aluno
Proficiência AD Matemática	192.50	42.08	Proficiência na Avaliação Diagnóstica de Matemática
Proficiência AD Português	195.49	41.20	Proficiência na Avaliação Diagnóstica de Português
Proficiência Agregada Matemática	23.13	37.23	Diferença entre notas no ENFE e na AD Matemática
Proficiência Agregada Português	33.39	37.56	Diferença entre notas no ENFE e na AD Português
<b>Nível Sócio-Econômico</b>	0.00	1.51	Componente de Riqueza (comp1)
Som	0.68	0.47	Variável Dummy indicativa da existência de som no Domicílio
Freezer	0.15	0.36	Variável Dummy indicativa da existência de freezer no Domicílio
Telefone	0.45	0.50	Variável Dummy indicativa da existência de telefone no Domicílio
Calçamento	0.66	0.47	Variável Dummy indicativa da existência de calçamento na rua do Domicílio
Celular	0.66	0.48	Variável Dummy indicativa da existência da posse de celular
Geladeira	0.87	0.34	Variável Dummy indicativa da existência de geladeira no Domicílio
Computador	0.08	0.27	Variável Dummy indicativa da existência de computador no Domicílio
Automóvel	0.12	0.32	Variável Dummy indicativa da existência de automóvel no Domicílio
Dois Automóveis	0.04	0.19	Variável Dummy indicativa da existência de mais de um automóvel no Domicílio
Banheiro	0.96	0.19	Variável Dummy indicativa da existência de banheiro no Domicílio
TV	0.93	0.26	Variável Dummy indicativa da existência de TV no Domicílio
<b>Características Individuais</b>			
Idade	20.81	2.09	Idade do aluno no início do programa obtida a partir de informações cadastrais
Casado	0.16	0.37	Dummy indicativa se o aluno é casado, construída a partir do questionário do aluno
Outros	0.01	0.10	Dummy indicativa de aluno não casado, construída a partir do questionário do aluno
Filhos	0.46	0.50	Dummy indicativa se o aluno possui filhos ou não, obtida do questionário do aluno
Não-Branco	0.76	0.43	Dummy indicativa se o aluno é branco ou não, obtida do questionário do aluno
Nenhuma Reprovação	0.18	0.39	Dummy obtida do questionário do aluno sobre não ter sofrido reprovação anterior.
Uma reprovação	0.21	0.41	Dummy obtida do questionário do aluno sobre ter sofrido uma reprovação anterior.
Trabalha	0.32	0.47	Dummy obtida do questionário do aluno sobre trabalhar ou não.
Pai e Mãe não sabem ler e escrever	0.12	0.32	Dummy indicativa se pai e mãe não sabem ler e escrever
Pai ou Mãe sabe ler e escrever	0.31	0.46	Dummy indicativa se pai ou mãe sabe ler e escrever
Pai e Mãe sabem ler e escrever	0.58	0.49	Dummy indicativa se pai e mãe sabem ler e escrever

Fonte: Sistema de Monitoramento e Avaliação do Projovem

Tabela 2  
Análise Regional da Proficiência

UF	Diagnóstica		Agregada	
	Matemática	Português	Matemática	Português
AC	194.50	200.43	-5.75	8.83
AL	190.37	191.44	15.32	31.72
AM	191.55	195.77	29.01	27.15
AP	189.36	196.37	28.12	23.14
BA	185.82	186.56	20.00	58.04
CE	180.75	186.07	29.87	33.17
DF	195.96	194.91	33.98	35.58
ES	192.36	198.28	20.93	30.75
GO	205.71	211.57	21.61	54.22
MA	174.61	177.38	27.78	36.89
MG	202.97	206.59	14.29	25.19
MS	210.78	216.89	11.34	19.92
MT	196.14	199.15	11.60	21.72
PA	193.23	200.20	22.36	28.25
PB	186.86	185.27	27.00	28.04
PE	189.96	193.07	23.11	31.32
PI	182.56	190.03	20.68	33.00
PR	217.68	220.04	41.10	54.97
RJ	195.81	197.30	25.00	36.23
RN	180.93	184.91	30.59	38.64
RO	189.31	192.95	23.99	31.34
RR	181.58	188.90	21.74	30.52
RS	205.41	209.44	23.57	27.18
SC	220.08	212.00	32.04	22.12
SE	195.18	195.40	16.41	26.15
SP	201.01	202.76	23.17	26.24
TO	176.27	180.78	27.61	28.05
<b>Brasil</b>	<b>190.97</b>	<b>193.88</b>	<b>23.89</b>	<b>33.71</b>

Fonte: Sistema de Monitoramento e Avaliação do Projovem

**Tabela 3****Análise de Componentes Principais Nível Sócio-Econômico do Aluno**

Componente	Auto-Valor	Diferença	Proporção	Acumulado
Comp1	2.2747	1.0199	0.2068	0.2068
Comp2	1.2548	0.1927	0.1141	0.3209
Comp3	1.0621	0.0807	0.0966	0.4174
Comp4	0.9813	0.1063	0.0892	0.5066
Comp5	0.8750	0.0201	0.0795	0.5862
Comp6	0.8549	0.0329	0.0777	0.6639
Comp7	0.8220	0.0459	0.0747	0.7386
Comp8	0.7761	0.0555	0.0706	0.8092
Comp9	0.7206	0.0116	0.0655	0.8747
Comp10	0.7090	0.0396	0.0645	0.9391
Comp11	0.6694		0.0609	1.0000

Número de observações 37378

Fonte: Elaboração própria

**Tabela 4****Análise de Componentes Principais – Auto Vetores**

Variável	Comp1
Som	<b>0.3693</b>
Freezer	<b>0.3090</b>
Telefone	<b>0.3931</b>
Calçamento	<b>0.1997</b>
Celular	<b>0.3236</b>
Geladeira	<b>0.3692</b>
Computador	<b>0.2990</b>
Automóvel	<b>0.2952</b>
Dois Automóveis	<b>0.0355</b>
Banheiro	<b>0.2346</b>
TV	<b>0.3160</b>

Fonte: Elaboração própria

**Tabela 5: Determinantes da Proficiência em Português**

	Avaliação Diagnóstica			Agregada		
	(1)	(2)	(3)	(4)	(5)	(6)
Componente de Riqueza	0.387*** (0.146)	-0.564*** (0.149)	-0.698*** (0.152)	0.567*** (0.183)	0.693*** (0.184)	0.710*** (0.188)
Coorte de Nascimento	-1.083*** (0.103)	-1.033*** (0.101)	-0.781*** (0.103)	0.173 (0.127)	0.109 (0.124)	0.051 (0.126)
Sexo Feminino	4.035*** (0.441)	3.056*** (0.436)	3.742*** (0.441)	-1.232** (0.553)	-0.4 (0.539)	-0.655 (0.542)
Não-brancos	-5.449*** (0.496)	-2.993*** (0.504)	-2.824*** (0.507)	2.632*** (0.613)	1.119* (0.618)	1.342** (0.621)
Nenhuma reprovação	11.269*** (0.564)	11.920*** (0.562)	12.212*** (0.567)	-3.295*** (0.689)	-2.492*** (0.679)	-2.439*** (0.679)
Uma reprovação	7.930*** (0.532)	8.420*** (0.525)	8.750*** (0.525)	-2.753*** (0.657)	-2.277*** (0.642)	-1.815*** (0.641)
Pai e mãe não sabem ler e escrever	-14.662*** (0.688)	-13.770*** (0.682)	-13.611*** (0.686)	2.571*** (0.828)	3.881*** (0.821)	3.929*** (0.823)
Pai ou mãe sabe ler e escrever	-7.307*** (0.472)	-6.669*** (0.466)	-6.480*** (0.469)	1.324** (0.588)	1.925*** (0.572)	1.908*** (0.573)
Dummies de UF	não	sim	sim	não	sim	sim
Dummies de Núcleo	não	não	sim	não	não	sim
Constante	203.552*** (0.826)	201.149*** (1.005)	153.086*** (3.153)	31.661*** (1.020)	28.810*** (1.224)	66.389*** (18.448)
Observações	37378	37378	37378	20776	20776	20776
R2	0.040	0.08	0.13	0.004	0.07	0.16

Erro padrão robusto entre parênteses.

\* significante a 10%; \*\* significante a 5%; \*\*\* significante a 1%.



**Tabela 6: Determinantes da Proficiência em Matemática**

	Avaliação Diagnóstica			Agregada		
	(1)	(2)	(3)	(4)	(5)	(6)
Componente de Riqueza	0.698*** (0.147)	-0.438*** (0.149)	-0.610*** (0.153)	0.173 (0.184)	0.182 (0.188)	0.22 (0.193)
Coorte de Nascimento	-1.276*** (0.105)	-1.231*** (0.103)	-1.102*** (0.105)	-0.049 (0.127)	-0.097 (0.126)	-0.270** (0.129)
Sexo Feminino	-11.081*** (0.445)	-12.149*** (0.440)	-11.852*** (0.447)	1.014* (0.545)	0.827 (0.544)	0.524 (0.554)
Não-brancos	-6.179*** (0.504)	-3.680*** (0.513)	-3.486*** (0.521)	-1.789*** (0.608)	-1.034* (0.622)	-0.794 (0.634)
Nenhuma reprovação	14.713*** (0.578)	15.293*** (0.574)	15.551*** (0.584)	-0.072 (0.695)	-0.899 (0.694)	-1.414** (0.706)
Uma reprovação	9.477*** (0.544)	9.948*** (0.535)	10.247*** (0.539)	-1.505** (0.650)	-1.922*** (0.644)	-2.136*** (0.652)
Pai e mãe não sabem ler e escrever	-13.293*** (0.673)	-12.400*** (0.667)	-12.318*** (0.674)	2.376*** (0.843)	2.554*** (0.837)	2.545*** (0.848)
Pai ou mãe sabe ler e escrever	-6.746*** (0.478)	-6.102*** (0.471)	-5.938*** (0.477)	1.639*** (0.583)	1.524*** (0.578)	1.400** (0.586)
Dummies de UF	não	sim	sim	não	sim	sim
Dummies de Núcleo	não	não	sim	não	não	sim
Constante	203.552*** (0.826)	206.854*** (1.026)	185.793*** (0.963)	23.688*** (1.019)	23.675*** (1.269)	25.863*** (4.748)
Observações	37378	37378	37378	20775	20775	20775
R2	0.05	0.09	0.13	0.002	0.03	0.10

Erro padrão robusto entre parênteses.

\* significante a 10%; \*\* significante a 5%; \*\*\* significante a 1%.

**Tabela 7**

	Agregação de Proficiência - Heckman							
	Var. Dep.=Proficiência Agregada							
	(1) PORT	Select PORT	(2) MAT	Select MAT	(3) PORT	Select PORT	(4) MAT	Select MAT
Componente de Riqueza	0.996** (0.398)	0.0307*** (0.005)	0.451 (0.404)	0.0307*** (0.005)	0.889** (0.350)	0.0265*** (0.005)	0.469 (0.356)	0.0265*** (0.005)
Coorte de Nascimento	-0.274 (0.415)	-0.0347*** (0.004)	-0.498 (0.423)	-0.0348*** (0.004)	-0.556 (0.458)	-0.0402*** (0.004)	-0.841* (0.469)	-0.0404*** (0.004)
Sexo Feminino	1.259 (2.254)	0.190*** (0.015)	1.887 (2.293)	0.189*** (0.015)	3.392 (2.651)	0.235*** (0.017)	4.555* (2.709)	0.235*** (0.017)
Não-brancos	1.111 (0.692)	-0.0258 (0.017)	-1.034 (0.696)	-0.0257 (0.017)	0.692 (0.713)	-0.0241 (0.018)	-1.097 (0.722)	-0.0241 (0.018)
Nenhuma reprovação	-2.187*** (0.769)	0.0287 (0.019)	-1.183 (0.772)	0.0282 (0.019)	-2.069** (0.809)	0.0309 (0.020)	-1.027 (0.819)	0.0305 (0.020)
Uma reprovação	-1.834*** (0.654)	0.000175 (0.018)	-2.054*** (0.656)	0.000193 (0.018)	-1.835*** (0.692)	0.00464 (0.019)	-2.104*** (0.701)	0.00469 (0.019)
Pai e mãe não sabem ler e escrever	4.640*** (1.139)	0.0695*** (0.023)	3.037*** (1.149)	0.0693*** (0.023)	5.143*** (1.157)	0.0711*** (0.024)	3.097*** (1.173)	0.0709*** (0.024)
Pai ou mãe sabe ler e escrever	2.437*** (0.801)	0.0485*** (0.016)	1.807** (0.807)	0.0480*** (0.016)	2.492*** (0.745)	0.0388** (0.017)	1.891** (0.753)	0.0384** (0.017)
Filhos					-3.127** (1.441)	-0.119*** (0.018)	-3.853*** (1.477)	-0.119*** (0.018)
Casado					-0.498 (0.894)	0.0402* (0.022)	0.326 (0.906)	0.0399* (0.022)
Divorciado/viúvo/outros					-0.134 (3.339)	-0.154** (0.075)	-2.394 (3.388)	-0.154** (0.075)
Constante	16.03 (14.870)	0.165*** (0.042)	13.85 (15.140)	0.165*** (0.042)	13.51 (13.660)	0.217*** (0.045)	9.21 (13.950)	0.218*** (0.045)
Lambda		18.527 (21.006)	13.451 (21.409)		24.979 (20.053)		22.885 (20.495)	
Observations	37.378			35.674				

Erro padrão robusto entre parênteses.

\* significante a 10%; \*\* significante a 5%; \*\*\* significante a 1%.

**Tabela 8**  
**Agregação de Proficiência - Heckman com Identificação na Variável Trabalho**

	Var. Dep.=Proficiência Agregada			
	(1) PORT	Select PORT	(2) MAT	Select MAT
Componente de Riqueza	0.968*** (0.274)	0.0264*** (0.005)	0.358 (0.263)	0.0264*** (0.005)
Coorte de Nascimento	-0.673** (0.312)	-0.0413*** (0.004)	-0.670** (0.301)	-0.0414*** (0.004)
Sexo Feminino	4.088** (1.764)	0.223*** (0.017)	3.555** (1.704)	0.223*** (0.017)
Não-brancos	0.615 (0.705)	-0.022 (0.018)	-1.001 (0.671)	-0.022 (0.018)
Nenhuma reprovação	-1.978** (0.790)	0.0329* (0.020)	-1.155 (0.751)	0.0325 (0.020)
Uma reprovação	-1.811** (0.713)	0.00629 (0.019)	-2.123*** (0.678)	0.00635 (0.019)
Pai e mãe não sabem ler e escrever	5.360*** (1.027)	0.0733*** (0.024)	2.812*** (0.977)	0.0730*** (0.024)
Pai ou mãe sabe ler e escrever	2.608*** (0.687)	0.0396** (0.017)	1.729*** (0.654)	0.0391** (0.017)
Filhos	-3.471*** (1.064)	-0.119*** (0.018)	-3.353*** (1.026)	-0.120*** (0.018)
Casado	-0.374 (0.857)	0.0397* (0.022)	0.167 (0.815)	0.0394* (0.022)
Divorciado/viúvo/outros	-0.62 (3.161)	-0.154** (0.075)	(1.760) (3.026)	-0.154** (0.075)
Trabalho	-	-0.0673*** (0.016)	-	-0.0676*** (0.016)
Constante	9.834 (8.789)	0.249*** (0.045)	14.49* (8.519)	0.250*** (0.045)
Lambda		30.426*** (12.746)	15.077 (12.370)	
Observações			35.674	

Erro padrão robusto entre parênteses.  
\* significante a 10%; \*\* significante a 5%; \*\*\* significante a 1%.

Figura 1 Distribuição de Proficiência na AD por Cor e Sexo

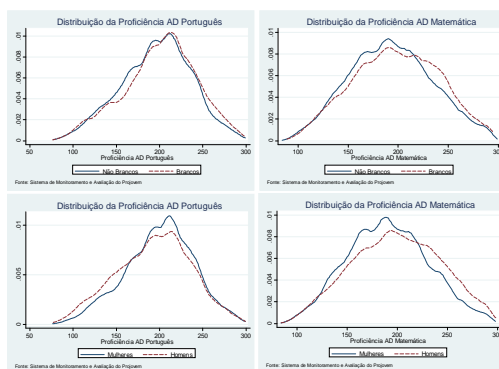


Figura 2 Distribuição de Proficiência Agregada por Cor e Sexo

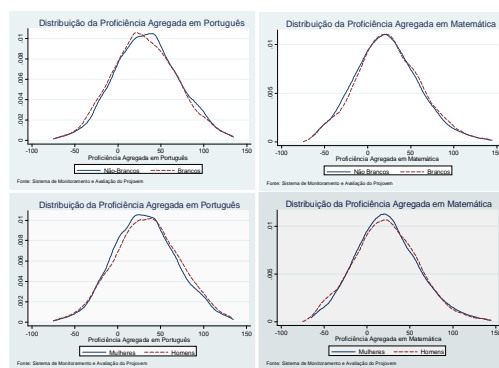


Figura 3 Distribuição Regional da Proficiência

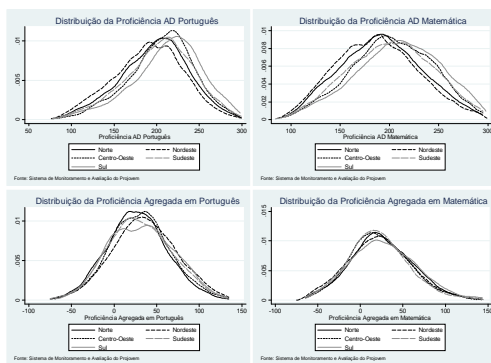


Figura 4 Distribuição de Proficiência por Nível de Riqueza

