

QUALIDADE E EQUIDADE NA EDUCAÇÃO: A HIPÓTESE DA CURVA DE KUZNETS EDUCACIONAL

Leonardo Neves*
Eduardo Almeida**

Resumo: Este artigo objetivou compreender o relacionamento entre eficácia e qualidade educacionais, considerando, por base teórica a função-produção educacional. O trabalho consistiu na estimação de um painel de efeitos fixos, tendo o desvio-padrão das notas médias na Prova Brasil de 2007 e 2009 por variável dependente, controlada para um conjunto de variáveis explicativas socioeconômicas e educacionais, em médias municipais. A hipótese central testada foi a existência de uma relação em formato de *U* invertido entre a equidade e eficácia educacionais, seguindo a proposição da Curva de Kuznets, adaptada para mensurar a qualidade educacional. Os resultados indicaram que, embora não tenha se confirmado a hipótese da Curva de Kuznets, não foi verificada a existência de um *trade-off* entre desigualdade e qualidade educacionais.

Palavras-chave: Desigualdade educacional; Eficácia educacional; Curva de Kuznets Educacional; Prova Brasil; Efeitos fixos.

JEL Classification: A20; I20; I21.

1. INTRODUÇÃO

Becker (1993) e Becker e Tomes (1976) definem a formação de talentos gerados pela educação como o principal determinante da formação de capital humano, que se caracteriza enquanto um dos mais importantes fatores para o desenvolvimento socioeconômico de um país. A oferta educacional universal, assim, caracteriza um importante agente para melhorar o desenvolvimento humano de uma população, bem como reduzir a iniquidade interna, uma vez que a desigualdade de rendimento dos indivíduos é a realização última do processo de formação de habilidades concretizada pelo sistema educacional (Roemer, 1998). Tais habilidades são fundamentais, não apenas por sua importância *per se*, mas como instrumentos para aquisição de novos talentos (Soares, 2007). Neste sentido, a qualidade da oferta educacional permite um processo multiplicador de habilidades.

Ainda que a literatura nacional e internacional venha mostrando interesse pela questão da desigualdade educacional, o relacionamento desta com a eficácia do sistema educacional ainda é objeto de controvérsia. Estudos buscaram identificar este padrão de relacionamento emprestando a abordagem teórica que relaciona renda *per capita* e crescimento econômico (Kuznets, 1955), denominada *Curva de Kuznets Educacional* (Doravante, CKE) (Ram, 1990). Esta abordagem relaciona o nível de escolarização dos indivíduos, em anos de estudo, com medidas de dispersão desta escolarização, notadamente, desvio-padrão e/ou índice de Gini ou Theil. Esta abordagem, desta forma, prevê a existência de um *trade-off* entre escolarização e equidade implicando em um aumento da primeira a taxas decrescentes, acompanhado de um aumento da iniquidade educacional. Este comportamento mantém-se até que a função alcance um ponto crítico, passando a decrescer a taxas crescentes, de modo que o aumento do nível de escolarização passa a atuar no sentido de reduzir a desigualdade associada. Lin e Tang (2008) afirmam que a principal justificativa por trás da hipótese da CKE reside no tratamento da escolarização enquanto principal medida para o capital humano. Uma vez que a escolarização média está altamente correlacionada ao nível de renda e, conseqüentemente, a distribuição desta escolaridade determina a distribuição da renda, o comportamento verificado entre nível de renda e sua distribuição, proposto por Kuznets (1955), reflete um comportamento análogo entre nível de escolarização média e sua distribuição entre os indivíduos (Lin e Tang, 2008).

* Programa de Pós-Graduação em Economia Aplicada/UFJF.

** Programa de Pós-Graduação em Economia Aplicada/UFJF.

Por sua vez, Sahlberg (2011a; 2011b; 2012) sugere que a relação entre a qualidade e desigualdade educacionais assume um comportamento complementar, de modo que a equidade interna ao sistema educacional é, por si só, um agente gerador de melhoria na eficácia educacional. Nesta perspectiva, a equidade interna implica na amenização dos efeitos promotores de desigualdade gerados pelas diferenças no *background* familiar, promovendo uma apropriação mais equânime dos benefícios do sistema escolar, levando a uma melhoria da qualidade escolar (Sahlberg, 2012).

Este trabalho tem por objetivo, desta forma, identificar a presença de uma relação entre eficácia e equidade educacionais, em nível municipal, emprestando das duas abordagens supracitadas a hipótese de um relacionamento não linear entre as duas variáveis. Assumindo que a equidade age no sentido de melhorar a qualidade educacional, à medida que se reduz a iniquidade escolar, assume-se que esta relação segue uma forma quadrática. Este comportamento parabólico foi adaptado do trabalho de Kuznets (1955), que teorizou uma relação sob a forma de *U* invertido entre o crescimento econômico e a concentração de renda nos países. A hipótese da CKE implica que desigualdade educacional tende a aumentar na medida em que se verifica um aumento da eficácia educacional. Esta iniquidade aumenta a taxas decrescentes e alcança um ponto de máximo. A partir deste ponto, entretanto, o efeito da eficácia passa a reduzir a desigualdade, de modo que a função passa a decrescer a taxas crescentes.

A validade da hipótese da CKE foi testada mediante um painel de dados municipais para as informações contidas nos questionários e nas notas das escolas na Prova Brasil de 2007 e 2009, para a quinta série do ensino fundamental. Caso a hipótese nula do coeficiente de interesse não seja rejeitada, há evidência estatística de que a hipótese da CKE é válida. Caso contrário, o coeficiente será estatisticamente diferente de zero e rejeitar-se-á a hipótese.

A base de dados utilizada é proveniente dos resultados e questionários da Prova Brasil para os anos de 2007 e 2009, e de variáveis para controle municipal advindas de outras fontes. A estratégia empírica consiste em um painel de dados dos municípios brasileiros, considerando escolas municipais, regredindo o desvio-padrão das proficiências médias, por município, contra uma série de variáveis de controle e o nível médio de proficiência em Matemática, bem como seu termo quadrático.

2. REVISÃO DE LITERATURA

Muitos estudos vêm utilizando medidas de dispersão a fim de identificar o padrão de desigualdade educacional, especialmente considerando os anos de escolarização dos indivíduos e taxas de matrículas, por medidas de desenvolvimento de habilidades. Rosthal (1978) estimou uma série de indicadores de desigualdade educacional para os EUA, Maas e Criel (1982) calcularam o índice de Gini para 16 países africanos, Sheret (1988) o fizeram para Papua e Nova Guiné, todos por meio das taxas de matrículas dos alunos nas escolas. Barro e Lee (1997) e Lopez *et al* (1998) estimaram o índice de Gini para todos os estados norte-americanos. Collins (2009) estimou o coeficiente de Gini, com base em matrículas de estudantes para as 24 províncias do Camboja. Castello e Domenech (2002), a partir dos dados de Barro e Lee (1997) estimou o índice de Gini para 108 países. Thomas *et al* (2000) o fizeram para 85 países, entre 1960 e 1990, expandindo seu estudo para 140 países, de 1960 a 2000 (Thomas *et al*, 2002).

Considerando dados brasileiros, Soares (2007) utilizou resultados em testes de proficiência para calcular o índice de Gini educacional. Felício e Fernandes (2005) realizaram a decomposição estática do índice de Theil considerando a proficiências dos estudantes nos exames do SAEB. Lorel (2008), Bagolin e Júnior (2003) realizaram o cálculo do índice de Gini educacional para regiões e estados brasileiros.

Ainda que diversos estudos tenham abordado os determinantes da desigualdade educacional e sua relação com o nível de proficiência dos estudantes, o estabelecimento de uma relação entre equidade e eficácia educacionais não tem sido devidamente abordada. Poucos estudos utilizaram medidas objetivas de mensuração da desigualdade, sendo sido realizados especialmente para verificar a validade da CKE, seja por meio do desvio-padrão da quantidade

de anos de estudo médio. Lodoño (1990) e Ram (1990) sugeriram a existência de um padrão kuznetsiano para a distribuição da educação, de tal sorte que, à medida que aumenta o nível médio de escolarização, ocorre um aumento inicial da desigualdade educacional, seguida de um comportamento descendente.

Uma discreta, mas crescente literatura internacional tem verificado a hipótese da CKE. Estes trabalhos têm utilizado, por medida de desigualdade, o desvio-padrão dos anos de escolarização ou índices relativos de dispersão, como os coeficientes de Gini e Theil. Considerando a taxa média de matrículas de estudantes e o coeficiente de Gini medido por meio destas, Maas e Criel (1982) identificaram um relacionamento negativo entre as duas variáveis. Thomas *et al* (2000, 2002), Checchi (2004) e Castello e Domenech (2002), por meio do indicador de Gini para anos de escolaridade dos indivíduos e a escolarização médias verificadas em suas respectivas amostras, identificaram uma correlação negativa.

Tendo por medida de desigualdade o desvio-padrão da escolarização, Ram (1990), Lodoño (1990), Thomas *et al* (2002, 2000), Gregório e Lee (2002), Lorel (2008) e Lin e Tang (2008) confirmaram a hipótese da CKE para os dados. Por sua vez, Lin (2007), Hojo (2009) e Morrison e Murtin (2010), utilizaram o índice de Gini da escolarização dos indivíduos por medida de desigualdade, refutando a hipótese da CKE. Estes resultados corroboram a proposição de Thomas *et al* (2002), que afirma que a CKE tende a ser confirmada quando a desigualdade é medida pelo desvio-padrão, e refutada quando medida por índices relativos, tais como Gini e Theil. Outros trabalhos identificaram relações alternativas entre escolarização e sua distribuição por meio do índice de Gini (Thomas *et al*, 2000; Castelló e Domenech, 2002; Wail *et al*, 2011; Barro e Lee, 1997, 2001 e Lopez *et al*, 1998).

Thomas *et al* (2002) identificaram uma série de fatos estilizados sobre a CKE: *i*) em países pobres com níveis baixos de escolarização, mas distribuição mais equânime dos anos de escolaridade média da população, aumentos nos anos de escolarização tendem a aumentar o desvio-padrão. No entanto, quando considerado o índice de Gini o resultado diverge de tal modo que, neste cenário, tem-se uma redução da desigualdade educacional; *ii*) todavia, em países cujo nível de escolarização é alto, o aumento médio dos anos de escolarização da população tende a reduzir o desvio-padrão da escolarização; *iii*) em ambos os casos supracitados o índice de Gini se reduz.

Tendo, também, por medida de desigualdade o índice de Gini, calculado para 146 países, quinquenalmente, entre os anos de 1950 e 2010, considerando o nível de escolarização, Wail *et al* (2011) desenvolveram um painel, condicional ao sexo dos indivíduos. Os resultados indicaram que a desigualdade educacional tem diminuído consideravelmente para grande parte dos países da amostra, embora esta tendência não tenha apresentado padrão uniforme. Dentre os determinantes apontados para esta heterogeneidade, os autores apontam o desenvolvimento diferenciado entre os países e suas especificidades, de modo que as características de cada região influenciaram na tendência de dispersão apresentada. Outra importante conclusão residiu na identificação de uma relação negativa entre o índice de Gini e o tempo médio de escolarização dos indivíduos. Resultados similares foram encontrados por Thomas *et al* (2000), também utilizando informações nacionais. De acordo com os autores pode-se verificar a existência de um comportamento convergente entre nível de escolarização e redução da iniquidade escolar. A conclusão, embora aborde a quantidade educacional acumulada e não a formação de habilidades cognitivas dos indivíduos, apresenta indícios da inexistência de um *trade-off* equidade e eficácia educacional.

Lorel (2008) utilizou estratégias empíricas diversas para capturar um desenho geral da problemática. Por meio do cálculo do índice de Gini para estados e regiões brasileiras, identificou uma importante redução das desigualdades educacionais, tendo por medida o nível de escolarização dos indivíduos. O autor, ainda, lançou mão da verificação da hipótese da Curva de Kuznets Educacional tendo por medida o desvio-padrão educacional, confirmando sua existência para os dados. No que se refere à relação entre a escolaridade média e o índice de Gini, esta se mostrou negativa, indicando que o aumento da escolarização tende a reduzir a iniquidade educacional para as regiões. Ao considerar os resultados em termos de matrículas em escolas, o autor identificou uma redução forte da desigualdade.

Poucos estudos abordaram os efeitos das características regionais sobre o nível de proficiência e desigualdade no sistema educacional, conforme explicitado por Scorzafave e Ferreira (2011). Curi (2010) afirma, ainda, que embora o desempenho da escola tenha a maior parte de sua variabilidade explicada pelo desempenho total médio das escolas da amostra, uma parte desta variação advém de determinantes regionais.

Desta forma, este trabalho não apresenta precedentes na verificação de uma relação entre qualidade e desigualdade educacionais, em nível municipal, verificando a procedência de uma forma funcional em formato de “U” invertido.

3. ASPECTOS METODOLÓGICOS

3.1. DADOS EM PAINEL

A disposição de um conjunto de dados sob a forma de um painel contém observações individuais ao longo de um dado período de tempo, abarcando, assim, as dimensões de corte transversal e temporal. O método de dados em painel de dados apresenta duas grandes vantagens em relação à abordagem com dados em corte transversal. Em primeiro lugar, Hsiao (2002) afirma que um painel de dados, por permitir um número de observações maior, implica no aumento da quantidade de variáveis independentes, por fornecer um número maior de graus de liberdade. A redução da colinearidade entre as variáveis explicativas apresenta, ainda, ganhos para as estimações. Logo, uma amostra maior garante as propriedades assintóticas desejáveis à estimação, apresentando ganhos de eficiência (Wooldridge, 2002).

Em segundo lugar, o nível de proficiência e desvio-padrão dos municípios, enquanto média do desempenho dos estudantes em nível municipal, pode ser influenciada por uma gama de características não observadas nos dados, podendo ser identificados determinantes do desempenho que não são capturados pelo modelo, desde incentivos gerados dentro da escola para melhoria do rendimento, impulso familiar para bom desempenho, características de urbanização e facilidade ao acesso escolar gerado pelo município, até capacidades cognitivas inatas ao aluno, que explicam a variação das notas nos testes. A estimação de um modelo que não leve em consideração a presença destes efeitos não observados pode levar a viés nas estimativas, decorrentes da omissão de variáveis explicativas relevantes à explicação da variação da variável dependente. Este viés é verificado quando por meio de efeitos endógenos do termo de erro com qualquer variável de controle. A modelagem por meio de *pooled* de dados assume, desta forma, de acordo com Wooldridge (2002), a função primordial de solucionar os danos gerados pelo viés de variável omitida, através da identificação do padrão de correlação dos termos não observados com as variáveis independentes observadas. Assim, o *pooled* de dados com controle para características não observadas permite contornar as possíveis violações das hipóteses de identificação.

Para identificar a presença de efeitos não observados faz-se uso do teste de Breusch-Pagan, que permite a identificação da existência de variáveis não observadas, não identificáveis na abordagem de dados em corte transversal³. Caso indique a existência de características não observáveis da escola, adota-se um modelo de efeitos não observados, contornando os problemas relativos à omissão de variáveis relevantes.

De acordo com Wooldridge (2002), em um painel de efeitos não observados pode apresentar características fixas ou aleatórias, considerando a dimensão temporal. Assim, as características da escola não observadas nos dados originais podem ser fixas na defasagem temporal de análise ou variáveis, de um momento inicial até outro. O primeiro

³ O teste utiliza os procedimentos matemáticos do multiplicador de Lagrange e a abordagem da máxima verossimilhança, que testa a probabilidade de uma amostra específica ser representativa da efetiva população de interesse. O teste tem por hipótese nula, $\sigma_v^2 = 0$, donde σ_v^2 é a variância do modelo. Se a hipótese nula for rejeitada, pode-se afirmar a existência de efeitos não observados, c.c. a adoção de um modelo POLS garante as boas propriedades dos estimadores (BREUSCH e PAGAN, 1980).

caso refere-se ao modelo de efeitos fixos, enquanto o segundo é concernente à abordagem de efeitos aleatórios. Para identificar o tipo efeitos não observados das escolas, utilizou-se o teste de Hausman⁴, checando se o termo de efeitos não observados está correlacionado com os vetores de variáveis de controle. Wooldridge (2002) aponta que caso um modelo seja estimado por efeitos fixos quando a especificação correta aponta para a presença de efeitos aleatórios, os estimadores serão consistentes, porém, não serão eficientes. Se o teste de Hausman identifica a presença de variáveis não observadas invariantes no tempo e correlacionadas com o conjunto de variáveis independentes, ocorre não independência entre o termo de erro e o vetor de variáveis de controle, tal que a esperança do erro condicionada ao vetor de variáveis explicativas é diferente de zero. A hipótese de identificação da abordagem de efeitos fixos é que

$$E[U_{it}|X_i, c_i] = 0, \quad t = 1, 2, \dots, T \quad (10)$$

Em que u é o termo de erro aleatório, X é o vetor de variáveis de controle e c é termo referente ao efeito não observado. A equação 8 indica a independência entre as características observáveis e não observáveis, não correlacionadas com as variáveis de controle do modelo, sendo respeitada a hipótese da ortogonalidade⁵.

O efeito não observado, uma vez invariante no tempo, deve ser removido por meio de uma transformação aritmética de sua média, como pode ser verificado na Equação 9.

$$Y_{it} - \bar{Y}_{it} = (X_{it} - \bar{X}_{it})\beta + (c_i - \bar{c}_i) + (u_{it} - \bar{u}_{it}) \quad (11)$$

Em que $\bar{Y}_{it} = (1/T) \sum_{t=1}^T Y_{it}$; $\bar{X}_{it} = (1/T) \sum_{t=1}^T X_{it}$; $\bar{u}_{it} = (1/T) \sum_{t=1}^T u_{it}$; e $\bar{c}_i = (1/T) \sum_{t=1}^T c_i$.

A estimação de $Y_{it} - \bar{Y}_{it}$ contra $X_{it} - \bar{X}_{it}$ será consistente por OLS. Este estimador é conhecido por *Within*.

Caso os efeitos sejam aleatórios, estes devem ser retirados de uma amostra *i.i.d.*, de modo que os estimadores para ambas as abordagens sejam iguais quando garantidas as propriedades assintóticas desejáveis (Wooldridge, 2002)⁶. Assume-se que $E(c_i|X_i) = E(c_i) = 0$, ou seja, o termo não observado e o conjunto de variáveis explicativas são independentes entre si, sendo $c \sim iid(0, \sigma_c^2)$.

A estimação de uma modelo cujos efeitos não observados são tratados por aleatórios não prescinde da homocedasticidade do efeito c . Tomando

$$\varepsilon_{it} = c_i + u_{it} \quad (12)$$

e

$$E(\varepsilon_{it}^2) = E(c_i^2) + 2E(c_i u_{it}) + E(u_{it}^2) = \sigma_c^2 + \sigma_u^2 \quad (13)$$

$$\forall t \neq s, E(\varepsilon_{it}, \varepsilon_{is}) = E((c_i + u_{it})(c_i + u_{is})) = E(c_i^2) = \sigma_c^2 \quad (14)$$

Dadas as relações acima apresentadas, existe Ω , tal que Ω é a matriz de variância de ε_i , definida positiva e $E(X_i' \Omega^{-1} X_i)$ é não singular.

Se as condições acima são verificadas, $\exists \hat{\beta} | \hat{\beta} = (\sum_{t=1}^N X_i' \Omega^{-1} X_i)^{-1} (\sum_{t=1}^N X_i' \Omega^{-1} Y_i)$ e $\hat{\beta}$ é consistente. Este estimador é o método de Mínimos Quadrados Generalizados, que é o melhor estimador quando verificados efeitos aleatórios se, e somente se, a matriz de variância dos componentes não observados é conhecida. Caso contrário, deve-se estimar uma matriz Ω , produzindo, então, um estimador $\tilde{\Omega}$, consistente, possibilitando estimar $\hat{\beta}$. Este método é

⁴ O teste, desenvolvido por Hausman (1978), tem por hipótese nula $Cov(\varepsilon_i, X_{it}) = 0$. Se a hipótese nula for rejeitada, pode-se afirmar a melhor especificação do modelo por efeitos fixos, c.c. a adoção de um modelo de efeitos aleatórios garante as boas propriedades dos estimadores.

⁵ $E(u) = 0$ é pressuposto para a garantia da ortogonalidade, bem como $Cov(c, X) = 0$ (Wooldridge, 2002).

⁶ Os trabalhos comumente realizados com microdados, entretanto, tendem a apresentar um número muito maior de observações longitudinais do que temporais, de modo a se verificar diferença estatisticamente significativa entre os estimadores de efeitos fixos e aleatórios (Wooldridge, 2002).

conhecido por Mínimos Quadrados Generalizados Exequíveis.

3.2. ESPECIFICAÇÃO DO MODELO

Foi utilizada uma abordagem econométrica de painel de dados de municípios para os anos de 2007 e 2009. Com base na especificação adotada por Ram (1990) e incorporando os efeitos espaciais, o modelo foi especificado como se segue:

$$DP_{it} = \beta_0 + Y_{it}'\beta_2 + Y_{it}^2'\beta_3 + A_{it}'\beta_4 + c_i + \varepsilon_{it} \quad (15)$$

Em que DP_{it} é o desvio-padrão dos resultados dos estudantes do município i no ano t . Y_{it} é a média da proficiência do município i no ano t , em matemática; β_0 é a constante do modelo; A_{it} é o vetor de variáveis de controle referentes às características dos estudantes, escolas, professores e município, do município i no ano t ; Y_{it}^2 é o termo quadrático da proficiência média do município; c_i é o efeito específico não observado do município i ; ε_{it} é o termo de erro aleatório *iid*, tal que $\varepsilon_{it} \sim N(0, \sigma^2)$.

O termo quadrático de proficiência educacional é incluído para testar a hipótese de uma forma funcional parabólica para o relacionamento entre proficiência e desigualdade educacionais municipal, que implica que a iniquidade cresce a taxas decrescentes para valores mais baixos em termos de proficiência, atingindo um ponto de máximo, a partir do qual a melhoria do rendimento acadêmico médio municipal implica na queda da desigualdade escolar *intra*-municipal.

Neste sentido, a relação entre o nível de proficiência médio dos alunos, por município e a desigualdade educacional municipal pode apresentar comportamentos distintos, em função dos sinais dos coeficientes, como descrito abaixo.

1) $\beta_2 > 0$; $\beta_3 = 0$. Neste caso, a relação dar-se-á de forma monotonamente crescente e linear, de modo que uma maior proficiência implica em maior iniquidade;

2) $\beta_2 < 0$; $\beta_3 = 0$. Neste caso, a relação será monotonamente decrescente e linear, o que significa que um aumento na proficiência leva a redução da desigualdade;

3) $\beta_2 > 0$; $\beta_3 < 0$. Neste caso, verifica-se uma Curva de Kuznets tradicional no formato de U invertido;

4) $\beta_2 < 0$; $\beta_3 > 0$. Neste caso, “U” tradicional indicando que, à medida que aumenta proficiência, a desigualdade associada diminui, atingindo um ponto de mínimo, aumentando a taxas decrescentes a partir deste ponto.

No que se refere às hipóteses de identificação, a especificação contemporânea exige uma série de restrições, apresentadas por Todd e Wolpin (2003) como segue:

i) Os insumos são invariantes no tempo, tal que os determinantes contemporâneos do desempenho acadêmico dos estudantes representem toda a história de seu desenvolvimento, ou a proficiência depende apenas de insumos contemporâneos;

ii) As características cognitivas inatas dos indivíduos são não correlacionadas aos insumos contemporâneos.⁷

⁷ Uma hipótese de identificação adicional deve ser presumida: os efeitos não observados devem ser uniformemente distribuídos entre as observações. Considerando um choque verificado entre os anos compreendidos na amostra em qualquer variável especificada, a consistência das estimações somente será garantida se este choque afetar todas as observações, não se efetivando em apenas uma parte das observações, gerando, desta forma, vies.

3.3. Dados

A amostra abrangeu a agregação das variáveis em nível de escola, transformadas em médias municipais, bem como um conjunto de variáveis já dispostas em nível municipal, conforme apresentado na Tabela 1.A, em anexo. Foram excluídas escolas localizadas em áreas rurais, de modo que foi garantida uma maior homogeneidade da amostra. A base de dados é oriunda de fontes diversas, a saber, Prova Brasil⁸, IBGE, Finbra, Ministério da Justiça e Federação das Indústrias do Rio de Janeiro (FIRJAN). A variável dependente é o desvio-padrão e notas dos alunos, por média escolar, na Prova Brasil⁹ para 2007 e 2009. A adoção do desvio-padrão das proficiências segue a proposição de Ferreira e Gignoux (2011), que afirmam que a variância e o desvio-padrão são medidas de dispersão de resultados de proficiência preferíveis a medidas relativas, como índices de Gini ou Theil¹⁰.

As variáveis de controle para as características dos alunos, professores e escolas também foram extraídas dos dados da Prova Brasil para os respectivos anos, por meio de informações obtidas pela aplicação de questionários direcionados, e encontram-se descritas na Tabela 1.A, em anexo. As variáveis de controle das características dos municípios foram obtidas por diversas fontes: as informações sobre os gastos municipais em ensino fundamental foram extraídas da Finbra; Índice FIRJAN de Desenvolvimento Humano Municipal, *proxy* para desenvolvimento humano; os dados elaborados pelo IBGE para Produto Interno Bruto *per capita* municipal, *proxy* para renda municipal; a taxa de homicídios para cada cem mil habitantes, fornecida pelo Ministério da Justiça, *proxy* para nível de criminalidade municipal¹¹; e, também através do IBGE, as variáveis referentes às políticas municipais de educação.

As estatísticas descritivas da amostra encontram-se na Tabela 2.A, em anexo. Considerando a presença de dados perdidos na amostra, 4.992 municípios apresentaram informações para os níveis de proficiência e desvio-padrão dos resultados dos alunos nos exames de matemática da Prova Brasil de 2007. As médias foram, respectivamente, 187,25 e 24,98, cujos desvios-padrão atingiram 20,31 e 1,43. As descrições para 2009 mostraram-se similares. Embora tenha incorrido mais perdas de dados, sendo 4.962 municípios portadores de informações sobre as proficiências e desvio-padrão de seus estudantes, as médias das variáveis, respectivamente, 200,28 e 25,41, apresentaram um aumento sensível, indicando tanto uma melhoria de desempenho quanto um pequeno aumento na dispersão das notas entre os municípios. A Tabela 3.A, em anexo, apresenta as estatísticas descritivas dos dados em *pooled*.

A Figura 2 apresenta a relação entre o desvio-padrão das proficiências dos alunos e os níveis de proficiência e termo quadrático desta, respectivamente, em *Pooled* de municípios. O relacionamento entre a eficácia e a qualidade educacional, em nível municipal, sem o controle para o conjunto de variáveis explicativas, indica uma função parabólica. Porém, verifica-se um relacionamento positivo, em termos do coeficiente quadrático. Na medida em que se verifica uma proficiência superior, a dispersão das notas tende a cair. Esse comportamento mantém-se até um ponto em que há uma inflexão (em torno do ponto médio entre 200 e 250, ou seja, 225) e o aumento das notas produz um consequente incremento na dispersão das mesmas. Este comportamento sugere a presença de um *trade-off* em pontos acima e abaixo do valor do desempenho mínimo satisfatório (225 pontos) para a prova de Matemática da Prova Brasil. Neste ponto não se verifica um *trade-off* entre eficácia e equidade educacionais nos dados. A próxima seção apresentará os resultados desta relação considerando as variáveis de controle e efeitos não observados para verificar se este comportamento se mantém.

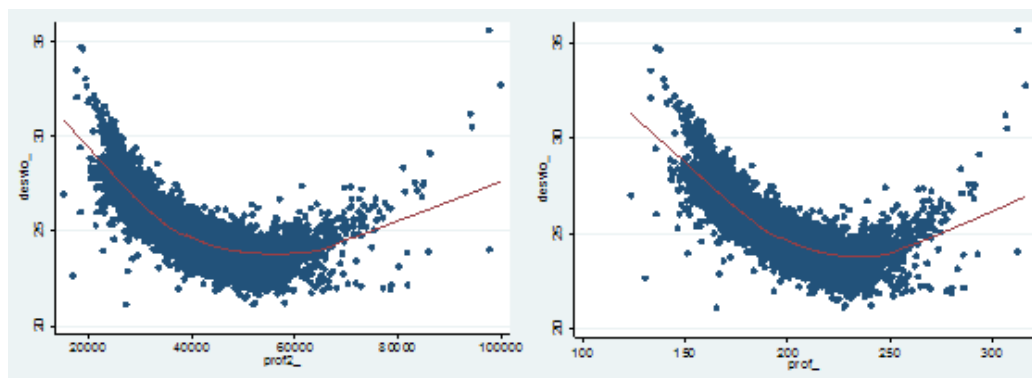
⁸ Agradecimentos ao Econs/UFJF pelo apoio no fornecimento dos microdados provenientes dos relatórios da Prova Brasil.

⁹ Disponível no sítio eletrônico do Instituto Nacional de Educação Anísio Teixeira (INEP) do Ministério da Educação (MEC).

¹⁰ Mais detalhes em Ferreira e Gignoux (2011).

¹¹ Obtido por meio do sítio eletrônico do Instituto de Pesquisa Econômica Aplicada (IPEA).

Figura 2. Relações entre desvio-padrão e proficiência média e termo quadrático da proficiência média (*pooled*)



Fonte: Elaborado pelos autores.

4. RESULTADOS E DISCUSSÃO

Cabe a esta seção a discussão dos resultados dos modelos econométricos estimados. Serão apresentados os resultados da análise de robustez dos modelos de *pooled* estimados por OLS, efeitos aleatórios (estimador de Mínimos Quadrados Generalizados) e efeitos fixos (estimador *Within*), estimados com as variáveis referentes às características dos estudantes (POLS 1, RE 1 e FE 1), com a adição das características dos professores (POLS 2, RE 2 e FE 2), considerando, posteriormente, as variáveis relativas à unidade escolar (POLS 3, RE 3 e FE 3), e por fim, com a adição das variáveis de características municipais (POLS 4, RE 4 e FE 4). Os modelos foram estimados considerando erros padrão robustos, para evitar perda de consistência das estimativas geradas pela presença de heterocedasticidade nos dados. As propriedades assintóticas desejadas são garantidas pelo alto número de observações, que dá a amostra propriedades de amostra infinita.

Primeiramente foi realizado o teste de Breusch-Pagan para verificação da presença de efeitos não observados no modelo especificado. A Tabela 1, apresenta o resultado do teste de Breusch-Pagan. Os modelos estimados sem controle para dependência espacial foram estimados, primeiramente, por OLS. A Tabela 1 apresenta os resultados das estimações para os quatro modelos expostos anteriormente. O coeficiente estimado do teste mostrou-se significativo em 1%, de tal sorte que a hipótese nula, que afirma a nulidade da variância do efeito não observado, foi rejeitada a 99% de confiança. Este resultado indica que há indícios da presença de efeitos não observados no modelo, tal que a estimação por POLS apresentam viés e inconsistência.

Procedida a verificação da presença de efeitos não observados, foi realizado o teste de Hausman para identificação do tipo de efeito não observado prevalente no modelo, a saber, aleatórios ou fixos, apresentado na Tabela 1. O coeficiente estimado apresentou rejeição da hipótese nula, ou seja, o termo de erro, não está relacionado a variável explicativa não observada, a 1% de significância. Deste modo, a melhor especificação considera os efeitos não observados como fixos, o que encontra respaldo na literatura.

Em relação aos modelos estimados por OLS, verificou-se um aumento do coeficiente de ajustamento (R^2) das estimações que consideram apenas as características dos estudantes (POLS 1) até aquela que considera todas as variáveis consideradas neste estudo (POLS 4). Dos quatro modelos estimados tem-se um aumento contínuo do R^2 ($0,764 < 0,785 < 0,79 < 0,817$), de modo que o modelo que considera as variáveis municipais tem 81,7% da variabilidade da variável dependente é explicada pela variabilidade *intra*-amostra das variáveis explicativas. Analogamente, os resultados dos testes F para cada modelo indicam que as variáveis especificadas são conjuntamente

significativas para explicar a variação da variável dependente.

Foram realizadas estimações considerando efeitos não observados, aleatórios e fixos, para os diferenciais das variáveis. Em relação às estimações considerando a presença de efeitos aleatórios, verifica-se um comportamento similar em relação às estatísticas de qualidade e ajustamento. O R^2 apresentou contínua melhoria até a estimação do último modelo ($0,2668 < 0,3383 < 0,3495 > 0,3483$), atingindo maior coeficiente para o modelo RE 3.

As estimações para efeitos fixos indicaram um aumento do R^2 de 0,319 para 0,493, do modelo FE 1 para FE 4, sendo a progressão contínua ao longo das estimações ($0,319 < 0,432 < 0,450 < 0,493$), tal que, ao considerar as variáveis municipais, 49,3% da variação da variável dependente é explicada pela variação *intra*-amostra das variáveis explicativas. Esta melhoria, ao serem adicionados novos conjuntos de variáveis, é também verificado pelo comportamento decrescente do Critério AIC ($16.460,14 > 11.468,35 > 9.989,114 > 3.891,052$), apresentando o maior *gap* quando especificadas as variáveis municipais¹².

Tabela 1.A - Resultados das estimações (OLS, Efeitos Aleatórios, Efeitos Fixos)

(Continua)												
Variáveis	POLS(1)	POLS(2)	POLS(3)	POLS(4)	RE(1)	RE(2)	RE(3)	RE(4)	FE(1)	FE(2)	FE(3)	FE(4)
Prof	-0.384*** (0.0101)	-0.396*** (0.0108)	-0.389*** (0.0115)	-0.332*** (0.0133)	-0.384*** (0.00958)	-0.396*** (0.0104)	-0.389*** (0.0111)	-0.331*** (0.0125)	-0.302*** (0.0136)	-0.338*** (0.0141)	-0.343*** (0.0143)	-0.339*** (0.0225)
Prof2	0.000844*** (2.49e-05)	0.000868*** (2.66e-05)	0.000853*** (2.85e-05)	0.000709*** (3.32e-05)	0.000844*** (2.36e-05)	0.000868*** (2.54e-05)	0.000853*** (2.75e-05)	0.000708*** (3.11e-05)	0.000696*** (3.35e-05)	0.000767*** (3.49e-05)	0.000780*** (3.56e-05)	0.000762*** (5.65e-05)
Sexo	0.986*** (0.112)	0.851*** (0.118)	0.867*** (0.121)	0.640*** (0.118)	0.986*** (0.111)	0.851*** (0.117)	0.867*** (0.120)	0.644*** (0.119)	0.885*** (0.179)	0.488** (0.200)	0.410** (0.204)	0.776*** (0.275)
Cor	-0.0304 (0.0655)	-0.000867 (0.0675)	-0.112 (0.0705)	-0.189*** (0.0728)	-0.0304 (0.0624)	-0.000867 (0.0658)	-0.112 (0.0688)	-0.188*** (0.0725)	0.405*** (0.154)	0.252 (0.162)	0.275 (0.173)	0.117 (0.216)
TV	0.448*** (0.169)	0.255 (0.175)	0.359** (0.180)	0.367* (0.193)	0.448*** (0.158)	0.255 (0.169)	0.359** (0.172)	0.359* (0.185)	1.444*** (0.417)	0.473 (0.441)	0.654 (0.453)	-0.0362 (0.620)
Carro	-0.694*** (0.0754)	-0.882*** (0.0790)	-0.978*** (0.0820)	-0.644*** (0.0905)	-0.694*** (0.0716)	-0.882*** (0.0769)	-0.978*** (0.0801)	-0.648*** (0.0900)	1.288*** (0.244)	0.992*** (0.270)	0.798*** (0.293)	0.516 (0.382)
Esc_mãe	-0.848*** (0.204)	-0.683*** (0.199)	-0.707*** (0.202)	-0.886*** (0.207)	-0.848*** (0.201)	-0.683*** (0.197)	-0.707*** (0.201)	-0.883*** (0.209)	-1.412*** (0.351)	-0.808** (0.367)	-0.738* (0.389)	-0.579 (0.522)
Esc_pai	0.0611 (0.190)	0.314 (0.192)	0.264 (0.197)	0.00847 (0.197)	0.0611 (0.189)	0.314* (0.191)	0.264 (0.196)	0.00597 (0.199)	1.838*** (0.295)	2.061*** (0.314)	1.952*** (0.330)	0.784 (0.478)
Reunião	-0.218*** (0.0815)	-0.201** (0.0841)	-0.0718 (0.0880)	-0.0866 (0.0865)	-0.218*** (0.0777)	-0.201** (0.0813)	-0.0718 (0.0855)	-0.0920 (0.0857)	-0.607*** (0.177)	-0.583*** (0.193)	-0.611*** (0.201)	-0.438 (0.272)
Reprovado	-0.265*** (0.0722)	-0.273*** (0.0745)	-0.304*** (0.0773)	-0.309*** (0.0790)	-0.265*** (0.0659)	-0.273*** (0.0702)	-0.304*** (0.0737)	-0.300*** (0.0774)	0.471*** (0.171)	0.509*** (0.186)	0.567*** (0.194)	0.521** (0.229)
Abandono	1.028*** (0.201)	1.172*** (0.214)	1.156*** (0.213)	0.761*** (0.195)	1.028*** (0.200)	1.172*** (0.214)	1.156*** (0.212)	0.762*** (0.194)	0.574** (0.229)	0.937*** (0.264)	0.846*** (0.268)	1.023*** (0.359)
Dever	-0.548*** (0.0959)	-0.247** (0.103)	-0.235** (0.107)	-0.135 (0.105)	-0.548*** (0.0944)	-0.247** (0.101)	-0.235** (0.106)	-0.132 (0.104)	-0.557*** (0.170)	-0.272 (0.193)	-0.0482 (0.198)	0.300 (0.265)
Corrige		-0.178 (0.112)	-0.266** (0.115)	-0.283** (0.114)		-0.178 (0.110)	-0.266** (0.113)	-0.284** (0.113)		-0.200 (0.198)	-0.304 (0.201)	-0.251 (0.279)
Superior		-0.149*** (0.0311)	-0.153*** (0.0313)	-0.126*** (0.0310)		-0.149*** (0.0310)	-0.153*** (0.0313)	-0.127*** (0.0311)		-0.168*** (0.0439)	-0.179*** (0.0460)	-0.148*** (0.0571)

Tabela 2.A - Resultados das estimações (OLS, Efeitos Aleatórios, Efeitos Fixos)

(Continua)												
Variáveis	POLS(1)	POLS(2)	POLS(3)	POLS(4)	RE(1)	RE(2)	RE(3)	RE(4)	FE(1)	FE(2)	FE(3)	FE(4)
Salário		0.571*** (0.0224)	0.561*** (0.0229)	0.381*** (0.0262)		0.571*** (0.0220)	0.561*** (0.0225)	0.385*** (0.0263)		0.908*** (0.0336)	0.933*** (0.0349)	0.628*** (0.0482)
Experiência		-0.150***	-0.137***	-0.0578		-0.150***	-0.137***	-0.0602		-0.289***	-0.275***	-0.229***

¹² O Critério AIC apresenta ganhos, em termos de análise de robustez em relação ao R^2 por considerar os ganhos em termos de variabilidade em dados fora da amostra. Novamente, todos os modelos apresentaram estatísticas F significativas em 1%.

	(Conclusão)									
	(0.0348)	(0.0369)	(0.0400)		(0.0346)	(0.0365)	(0.0392)	(0.0562)	(0.0633)	(0.0792)
Conteúdo	-0.0846***	-0.0789***	-0.0800**		-0.0846***	-0.0789***	-0.0804**	0.0238	0.0252	-0.0157
	(0.0272)	(0.0285)	(0.0319)		(0.0273)	(0.0286)	(0.0319)	(0.0445)	(0.0473)	(0.0640)
Biblioteca		-0.0906***	-0.0735***			-0.0906***	-0.0741***		0.00407	-0.0111
		(0.0238)	(0.0249)			(0.0229)	(0.0246)		(0.0505)	(0.0608)
Reforço		-0.0347	-0.00520			-0.0347	-0.00509		0.0822	0.0953
		(0.0268)	(0.0277)			(0.0262)	(0.0277)		(0.0539)	(0.0645)
Projeto		0.0544	0.131***			0.0544	0.131***		0.136	0.222*
		(0.0427)	(0.0412)			(0.0417)	(0.0411)		(0.0836)	(0.116)
Homicídios			0.00343***				0.00347***			0.00388***
			(0.000452)				(0.000438)			(0.00124)
PIB			-0.00310**				-0.00300**			0.0123**
			(0.00131)				(0.00145)			(0.00614)
IFDM			-0.520***				-0.532***			2.097***
			(0.149)				(0.145)			(0.467)
Gastos			0.0434***				0.0438***			0.0435***
			(0.00273)				(0.00274)			(0.00650)
Secretaria			0.172***				0.173***			0.193***
			(0.0179)				(0.0169)			(0.0391)
Sistema			-0.00104				-0.000434			0.00340
			(0.0164)				(0.0160)			(0.0417)
Plano			0.191***				0.193***			0.255***
			(0.0182)				(0.0172)			(0.0376)

Tabela 14.A - Resultados das estimações (OLS, Efeitos Aleatórios, Efeitos Fixos)

Variáveis	POLS(1)	POLS(2)	POLS(3)	POLS(4)	RE(1)	RE(2)	RE(3)	RE(4)	FE(1)	FE(2)	FE(3)	FE(4)
Conselho				0.0473*** (0.0183)				0.0479*** (0.0181)				0.158*** (0.0510)
Constante	67.45*** (1.018)	68.80*** (1.086)	67.99*** (1.152)	62.10*** (1.284)	67.45*** (0.966)	68.80*** (1.041)	67.99*** (1.110)	62.04*** (1.210)	55.34*** (1.414)	60.29*** (1.482)	60.48*** (1.484)	58.66*** (2.263)
Obs	9,812 22557.04	8,745 19183.24	8,217 17870.96	6,455 12587.13	9,812	8,745	8,217	6,455	9,812 16460.14	8,745 11468.35	8,217 9989.114	6,455 3891.052
R-quadrado												
Within	0.764	0.785	0.790	0.817	0.2668	0.3383	0.3495	0.3483	0.319	0.432	0.45	0.493
Between					0.8935	0.8559	0.8489	0.8529	0.6046	0.6065	0.588	0.5117
Overall					0.7644	0.7851	0.79	0.817	0.5283	0.5815	0.5825	0.5236
Breusch-Pagan	62.64											
Hausman	286.58											
No. de municípios	9812	8745	8264	6484	5,084	4,953	4,855	4,525	5,084	4,953	4,855	4,525

Fonte: Elaborado pelo autor a partir de dados da Prova Brasil 2007/2009, utilizando o pacote Stata 12.

Notas: * p-valor < 0.1; ** p-valor < 0.05; *** p-valor < 0.01. Erros-padrão robustos entre parênteses

Em relação às estimativas dos parâmetros das variáveis especificadas, considerando o modelo mais adequado, de acordo com a análise de robustez e os testes de Hausman e Breusch-Pagan, qual seja o modelo de efeitos fixos considerando todo o conjunto de variáveis, em níveis dos alunos, professores, escolas e municípios, os resultados das estimações podem ser verificados na Tabela 1.

As variáveis de interesse, proficiência média municipal (*Proficiência*) e seu termo quadrático (*Prof2*), apresentaram significância estatística em níveis tradicionais para suas estimativas. Ambas tiveram seus coeficientes estatisticamente diferentes de zero no nível de 1% de significância. Em relação à primeira, o coeficiente estimado foi negativo, na ordem de 0,34. O coeficiente do termo quadrático da proficiência assumiu um valor positivo, porém com

magnitude tendente a zero. Destes resultados depreendem-se dois possíveis cenários: primeiramente, verifica-se um formato quadrático na forma de “U” tradicional. Isto significa que, analogamente ao comportamento verificado na Figura 2, tem-se um *trade-off* entre eficácia e equidade escolares para os dados. À medida que se verifica níveis médios superiores de proficiência, o desvio-padrão dos resultados diminui até um ponto de mínimo (nota 222,441)¹³. A partir deste ponto tem-se um aumento da dispersão que acompanha os níveis superiores em termos das notas médias dos municípios.

O segundo cenário considera o coeficiente do termo quadrático da proficiência igual a zero, em função da magnitude ínfima da estimativa. Deste resultado conclui-se uma relação linear, decrescente monotônica, implicando na inexistência do *trade-off* entre eficácia e equidade educacionais, nos dados. Na medida em que os resultados municipais médios dos estudantes no exame de Matemática da Prova Brasil aumentam, o desvio-padrão se reduz, indicando uma convergência entre qualidade e equidade.

5. Considerações finais

Este estudo teve por objetivo testar a hipótese nele levantada de um relacionamento entre desigualdade e qualidade educacionais na forma de um “U” invertido, nos moldes da chamada Curva de Kuznets Educacional. Esta hipótese consideraria a existência de um *trade-off* entre qualidade e equidade educacionais apenas parcial, de modo que altos níveis de qualidade levariam a baixos níveis de desigualdade entre os indivíduos. A estratégia empírica adotada foi um painel de efeitos fixos, regredindo o desvio-padrão médio municipal dos resultados dos estudantes no exame de Matemática da Prova Brasil, regredindo contra um conjunto de variáveis de controle e o nível de proficiência associado, bem como seu termo quadrático.

Os resultados encontrados refutam a hipótese de um formato similar a Curva de Kuznets para o relacionamento entre qualidade e desigualdade educacionais. As estimativas do modelo permitiram duas interpretações possíveis para a função relacional que considera as duas variáveis: *i*) um formato de “U” tradicional; *ii*) uma curva monotonicamente decrescente, caso se considere nulo o coeficiente da termo quadrático da proficiência. Caso se aceite *i*, tem-se um *trade-off* parcial entre eficácia e equidade, de sorte que níveis baixos de desempenho acadêmico tendem a produzir níveis altos de desigualdade, ao passo que níveis altos de qualidade produzem níveis igualmente altos de iniquidade. Esta inclinação, porém, seria muito baixa, dada a baixa magnitude da estimativa do termo quadrático, indicando uma relativa estabilidade entre eficácia e equidade. Este cenário aponta uma variação baixa, ainda que existente, entre os fenômenos, tal que exista um *trade-off*, embora muito pouco relevante para que seja considerado em termos de atuação direta da autoridades.

Se a opção *ii* for aceita, tem-se um comportamento negativo entre a qualidade e desigualdade educacionais, de modo que alta qualidade tenda a produzir baixos níveis de desigualdade entre os indivíduos. Este resultado implica na refutação da hipótese da existência de um *trade-off* entre qualidade e desigualdade, de modo que melhorar continuamente a qualidade educacional que levaria a uma consequente redução da desigualdade, considerando apenas alunos da 5ª série do Ensino Fundamental.

ABSTRACT: This paper aims understanding the relation between educational effectiveness and inequality, with an educational production function framework. The study consisted in the estimation of a panel fixed effects, assuming the standard deviation of the average scores in the Prova Brasil 2007 and 2009 as the dependent variable, controlled for a set of socioeconomics and educational explanatory variables, aggregated in municipal averages. The central hypothesis tested was that there is an inverted U-shaped relationship between educational equity and effectiveness, as described by the Educational Kuznets Curve, adapted to measure the educational quality. The results indicated that, although not confirmed the hypothesis of Kuznets curve, it was found that there is no trade-off between educational inequality and quality.

Key words: Educational inequality; educational effectiveness; Educational Kuznets Curve; Prova Brasil; Fixed effects.

¹³ A otimização da função estimada indicou que o ponto de mínimo é 222,441.

JEL Classification: A20; I20; I21.

REFERÊNCIAS

- ALBERNAZ, Â.; FERREIRA, F.; FRANCO, C. Qualidade e Equidade na Educação Fundamental Brasileiro. **PPE**, vol. 33 N.3. 2002.
- BAGOLIN, I. P.; JÚNIOR, S. A desigualdade da distribuição da educação e crescimento no Brasil: índice de Gini e anos de escolaridade. **Estudos do CEPE**, Santa Cruz do Sul-RS, v. 18, p. 7-31, 2003.
- BARRO, Robert J.; LEE, Jong-Wha. International Measures of Schooling Years and Schooling Quality. **American economic Review**, 86, 218-23, 1997.
- BARRO, R.; LEE, J.W. International data on educational attainment: updates and implications. **Oxford Economic Papers**. n.3, 2001.
- BECKER, G. **Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education**, 3rd Ed. (U. Chicago Press); 1st Ed publ. in 1964, 1993.
- BECKER, G.; TOMES, N. **Children dowments and the quantity and quality of children**. 1976.
- BRASIL. Ministério da Educação. Inep. **O Plano de desenvolvimento da educação**. 2007.
- BREUSCH, T.; PAGAN, A. The Lagrange multiplier test and its applications to model specification in econometrics. **Review of Economic Studies**, v. 47, p. 239-253, 1980.
- CASTELLÓ, A.; DOMÉNECH, Rafael. Human capital inequality and economic growth: some new evidence. **The Economic Journal**, v. 112, n. 478, p. C187-C200, 2002.
- COLEMAN, J. S.; CAMPBELL, E. Q.; HOBSON, C. J.; MCPARTLAND, J.; MOOD, A. M.; WEINFELD, F. D. Equality of educational opportunity. Washington, D.C.: **National Center for Educational Statistics**, 1966.
- COLLINS, J. **Reconstructing Access in the combodian education system**. IN: HOLSINGER. D.; JACOB, W. (eds). **Inequality in Education: Comparative and International Perspectives**. 2009.
- CURI, E.; CAPES, Financiamento. Prova Brasil de matemática: revelações, possibilidades de avanços nos saberes de alunos de 4ª série/5º ano e indicativos para formação de professores. **Programa Observatório da Educação**. Edital, n. 038, 2010.
- GREGORIO, J.; LEE, J. Education and income distribution: New evidence from cross-country data. **Review of Income and Wealth**. 48, 395-416, 2002.
- FELICIO, F.; FERNANDES, R. Efeito da Qualidade da Escola Sobre o Desempenho Escolar: Uma Avaliação do Ensino Fundamental no Estado de São Paulo. **Encontro Nacional de Economia**. Anpec, 2005.
- FERREIRA, Francisco; GIGNOUX, Jérémie. **The measurement of educational inequality: Achievement and opportunity**. 2011.
- HAUSSMAN, J. Specification Tests in Econometrics. **Econometrica**, v. 46, n. 6, p. 1251-71, 1978.
- HOJO, M. Inequality in Japanese Education Estimation Using the Gini Education Coefficient. **The Japanese Economy**,

36(3): 3–27, 2009.

HSIAO, C. **Analysis of Panel Data**. New York, 2. ed. Cambridge University Press, 2002, 384p.

KUZNETS, Simon. Economic Growth and Income Inequality. **American Economic Review**, 45, no. 1: 1–28, 1955.

LIN, Chun-Hung A. Education Expansion, Educational Inequality, and Income Inequality: Evidence from Taiwan, 1976–2003. **Social indicators research**, v. 80, n. 3, p. 601-615, 2007.

LIN, A.; TANG, K. Human Capital inequality and the Kuznets curve. **The Developing Economies**, XLVI-1: 26–51, 2008.

LODOÑO, J. Kuznetsian Tales with attention to human capital. IN: Lundenberg, M., Squire, L. **Growth and Inequality: extracting the lessons for policy makers**. The World Bank. 1990.

LOPEZ, R.; THOMAS, V.; WANG, Y. **Addressing the education puzzle: the distribution of education and economic reform**. The World Bank, 1998.

LOREL, Benoit. Assessing Brazilian educational inequalities. **Revista Brasileira de Economia**, v. 62, n. 1, p. 31-56, 2008.

LUZ, P. Análise de políticas públicas: conceitos básicos. **PublicationsOboulo.com**, 2006.

MAAS, J.; CRIEL, C. Distribution of Primary School Enrolments in East Africa. **World Bank Staffing Paper**, v. 511, 1982.

MORRISSON, C; MURTI, F. **The Kuznets curve of education: a global perspective on education inequalities**. Centre for the Economics of Education, LSE, 2010.

ORGANIZATION FOR ECONOMIC COOPERATION AND DEVELOPMENT, **Equity in Education Thematic Review: Country Analytical Report**, Paris: 2005.

RAM, R. Educational expansion and schooling inequality: International evidence and some implications. **The Review of Economics and Statistics**, p. 266-274, 1990.

RIVKIN, S.; HANUSHEK, E.; KAIN, J. Teachers, Schools, and Academic Achievement. **Econometrica**. vol. 73, No.2, p417-58, 2005.

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

ROSTHAL, Richard A. **Measures of Disparity**. A Note. 1978.

SAHLBERG, P. Paradoxes of educational improvement: The Finnish experience. **Scottish Educational Review**, 43 (1), 3-23, 2011.

SAHLBERG, P. **Finnish Lessons: What Can the World Learn From Educational Change in Finland?**. Teachers College Press, 2011.

SAHLBERG, P. Quality and equity in Finnish schools. **School Administrator**, p. 27-30, 2012.

SCORZAFAVE, L. G.; FERREIRA, R. A. Desigualdade de Proficiência no Ensino Fundamental Público Brasileiro: Uma Análise de Decomposição. **Revista Economia**, 2011.

SHERET, M. Equality trends and comparisons for the education system of Papua New Guinea. **Studies in Educational**

Evaluation, 14(1), 1988.

SOARES, J. F. Melhoria do desempenho cognitivo dos alunos do ensino fundamental. **Cadernos de Pesquisa**, v. 37, n. 130, p. 135-160, 2007.

SOARES, J. F.; ANDRADE, RJ de. Nível socioeconômico, qualidade e equidade das escolas de Belo Horizonte. **Ensaio: avaliação e políticas públicas em educação**, 14(50), p. 107-126, 2006.

THOMAS, V.; WANG, Y.; FAN, X. A new dataset on inequality in education: Gini and Theil indices of schooling for 140 countries, 1960-2000. **World Bank. Washington, DC. Processed**, 2002.

THOMAS, V.; WANG, Y.; FAN, X. **Measuring education inequality: Gini coefficients of education**. World Bank Publications, 2000.

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

WAIL, B.; HANCHANE, S.; KAMAL, A. A New Data Set of Educational Inequality in the World, 1950-2010: Gini Index of Education by Age Group. **Available at SSRN 1895496**, 2011.

WOOLDRIDGE, J. **Econometric Analysis of cross section and panel data**. The MIT Press, 2002.

Tabela 1.A - Descrição das variáveis utilizadas

Variável	Descrição	Fonte
Sexo	Porcentagem de alunos do sexo masculino	Prova Brasil
Cor	Porcentagem dos alunos autodeclarados brancos	Prova Brasil
TV	Porcentagem de alunos cujas famílias possuem ao menos um aparelho televisor em cores	Prova Brasil
Carro	Porcentagem de alunos cujas famílias possuem ao menos um automóvel	Prova Brasil
Esc_mae	Porcentagem de alunos cujas mães possuem nível superior completo	Prova Brasil
Esc_pai	Porcentagem de alunos cujos pais possuem nível superior completo	Prova Brasil
Reprovado	Porcentagem de alunos que foram ao menos uma vez reprovados	Prova Brasil
Abandono	Porcentagem de alunos que ao menos uma vez abandonaram um ano letivo	Prova Brasil
Dever	Porcentagem de professores que sempre passam dever de casa	Prova Brasil
Corrige	Porcentagem dos professores que sempre corrigem o dever de casa em sala, quando passam	Prova Brasil
Salário	Porcentagem de professores que ganham ao menos a categoria média dos questionários	Prova Brasil
Conteúdo	Porcentagem dos professores que completaram ao menos 80% do conteúdo	Prova Brasil
Superior	Porcentagem dos professores com formação superior em matemática	Prova Brasil
Experiência	Porcentagem dos professores com ao menos cinco anos de experiência	Prova Brasil
Projeto	Escolas com projeto pedagógico	Prova Brasil
Reforço	Escolas com programa de reforço	Prova Brasil
Biblioteca	Escolas que possuem biblioteca	Prova Brasil
Homicídios	Taxa de homicídios a cada cem mil habitantes (por município)	M. Saúde
IFDM	Índice Firjan de Desenvolvimento Municipal (por município)	FIRJAM
PIB	Produto Interno Bruto municipal <i>per capita</i> (por município)	IBGE
Secretaria	Se o município possui secretaria municipal de educação	IBGE
Sistema	Se o município possui sistema de ensino próprio	IBGE
Plano	Se o município possui Plano Municipal de Educação	IBGE
Conselho	Se o município possui Conselho Municipal de Educação	IBGE
Gastos	Total de gastos municipais em educação básica <i>per capita</i>	FINBRA

Fonte: Elaborado pelos autores.

Tabela 2.A - Estatísticas descritivas

Variável	2007				2009			
	Obs	Média	Desvio-padrão	Coef. Var	Obs	Média	Desvio-padrão	Coef. Var
Desvio	4992	24.98	1.43	0.06	4962	25.41	1.68	0.07
Prof	4992	187.25	20.31	0.11	4962	200.28	25.20	0.13
Sexo	4985	0.50	0.09	0.19	4952	0.51	0.07	0.13
Cor	4987	0.36	0.17	0.48	4952	0.37	0.17	0.46
TV	4987	0.93	0.06	0.06	4952	0.94	0.05	0.05
Carro	4987	0.34	0.19	0.56	4952	0.36	0.20	0.56
Esc_mãe	4985	0.09	0.06	0.70	4961	0.09	0.05	0.56
Esc_pai	4986	0.09	0.07	0.87	4960	0.08	0.04	0.57
Reunião	4987	0.60	0.11	0.18	4961	0.60	0.11	0.19
Reprovado	4985	0.37	0.15	0.40	4961	0.36	0.14	0.40
Abandono	4881	0.09	0.08	0.97	4961	0.08	0.06	0.70
Dever	4941	0.78	0.12	0.15	4961	0.77	0.09	0.12
Corrige	4983	0.83	0.10	0.13	4961	0.83	0.09	0.11
Superior	4543	0.11	0.31	2.72	4700	0.09	0.22	2.40
Salario	4912	0.42	0.41	0.96	4410	0.81	0.34	0.42
Experiência	4914	0.83	0.24	0.29	4713	0.83	0.29	0.35
Conteúdo	4911	0.29	0.32	1.10	4707	0.33	0.38	1.13
Projeto	4970	0.91	0.24	0.26	4361	0.93	0.22	0.24
Reforço	4963	0.76	0.37	0.48	4389	0.79	0.35	0.45
Biblioteca	4982	0.68	0.39	0.58	4349	0.67	0.40	0.60
Homicídios	3328	23.21	18.71	0.81	3328	23.95	18.53	0.77
IFDM	5506	0.62	0.10	0.17	5506	0.64	0.10	0.15
PIB	5507	5.13	5.88	1.15	5507	9.36	9.84	1.05
Gastos	4563	4.30	2.36	0.55	4563	8.18	6.15	0.75
Conselho	5507	0.68	0.47	0.69	5507	0.79	0.41	0.51
Plano	5507	0.31	0.46	1.49	5507	0.57	0.50	0.88
Sistema	5507	0.43	0.49	1.16	5507	0.52	0.50	0.96
Secretaria	5507	0.26	0.44	1.67	5507	0.43	0.50	1.14

Fonte: Elaborado pelos autores.

Tabela 3.A - Estatísticas descritivas (*Pooled*)

Variável	Obs	Média	Desvio-padrão	Coef. Variação
Desvio	9954	25.19	1.57	0.06
Prof	9954	193.74	23.79	0.12
Sexo	9937	0.50	0.08	0.16
Cor	9939	0.36	0.17	0.47
TV	9939	0.94	0.06	0.06
Carro	9939	0.35	0.20	0.56
Esc_Mãe	9946	0.09	0.06	0.64
Esc_pai	9946	0.08	0.06	0.75
Reunião	9948	0.60	0.11	0.18
Reprovado	9946	0.37	0.15	0.40
Abandono	9842	0.08	0.07	0.85
Dever	9902	0.78	0.11	0.14
Corrige	9944	0.83	0.10	0.12
Superior	9243	0.10	0.26	2.62
Salário	9322	0.61	0.42	0.69
Experiência	9627	0.83	0.27	0.32
Conteúdo	9618	0.31	0.35	1.12
Biblioteca	9331	0.67	0.39	0.59
Reforço	9352	0.77	0.36	0.47
Projeto	9331	0.92	0.23	0.25
Homicídios	8835	17.25	19.35	1.12
PIB	11014	7.24	8.37	1.16
IFDM	11013	0.63	0.10	0.16
Gastos	10070	6.06	4.89	0.81
Secretaria	11014	0.35	0.48	1.37
Sistema	11014	0.47	0.50	1.05
Plano	11014	0.44	0.50	1.13
Conselho	11014	0.73	0.44	0.60

Fonte: Elaborado pelos autores.