

IGUALDADE DE OPORTUNIDADES: ANALISANDO O PAPEL DAS CIRCUNSTÂNCIAS NO DESEMPENHO DO ENEM

Fernanda Leite Santana(PPGE/ UFPB)
Lauro César Nogueira (PPGE/ UFPB)
Erik Alencar de Figueiredo(PPGE/ UFPB)

RESUMO:

Este trabalho visa avaliar como as circunstâncias sociais influenciam o resultado econômico dos alunos que prestam o Exame Nacional do Ensino Médio (ENEM). A ideia é saber se a origem familiar interfere no seu desempenho assumindo que os mesmos são responsáveis por uma parcela da diferença de seus resultados. Para tanto, emprega-se duas estratégias empíricas a partir dos microdados do ENEM e do SAEB. Isto é, fazem-se estimações não paramétricas, em especial, e em seguida estima-se parametricamente uma Função de Produção Educacional (FPE) baseado em Brunello e Checchi (2006). Os resultados apontam forte diferencial no esforço empregado entre os tipos para se obter um bom desempenho no ENEM, tanto no Brasil como na Região Nordeste. Em síntese, a origem importa no desempenho dos indivíduos.

Palavras-chaves: Desigualdade de Oportunidades; Densidade de Kernel; ENEM

ABSTRACT:

This paper aims to assess how social circumstances influence the economic outcome of the students providing the National Secondary Education Examination (ENEM). The idea is whether family background affects the performance assuming that they are responsible for a portion of the difference in their results. To do so, it employs are two empirical strategies from microdata ENEM and Saeb. That is, they are non-parametric estimation, in particular, then it is estimated parametrically a Feature Production Education (FPE) based on Brunello and Checchi (2006). The results show strong differential effort between employee types to obtain a good performance in ENEM, both in Brazil and in the Northeast. In summary, the origin of matter in the performance of individuals.

Keywords: Inequality of Opportunity; Kernel Density; ENEM

1 INTRODUÇÃO

Nas últimas décadas, correntes filosóficas e econômicas tem buscado compreender a dinâmica da desigualdade social, no entanto, conforme destacados pelos denominados igualitários de oportunidades, exceto que se esteja disposto a aceitar grandes perdas de eficiência, a igualdade não pode ser alcançada porque os agentes são de certo modo responsáveis pelos seus próprios resultados.

O final do século XX foi marcado pelo desenvolvimento de teorias igualitárias, tanto na filosofia, como nas ciências sociais. No entanto, podemos destacar algumas das principais contribuições na literatura, em especial, Harsanyi (1953), na formulação do véu da ignorância e do utilitarismo; Rawls (1971), com a inovadora abordagem da teoria de justiça distributiva; Sen (1980), acrescentando as noções de funcionamento e capacidades; e Ronald Dworkin (1981), tornando-se referência para as abordagens igualitárias. Recentemente, destacam-se alguns trabalhos, em especial, Roemer (1998) e Fleurbaey (1995), apontando para a existência de fatores caracterizados pela responsabilidade dos indivíduos na formulação das normas de justiça social e, por conseguinte, se opondo as teorias de desigualdade de resultados.

Esta última abordagem conhecida como a desigualdade de oportunidades¹, retrata o resultado econômico como fruto de dois fatores: *i*) fatores de responsabilidade (relacionadas ao esforço dos indivíduos); e *ii*) fatores de não responsabilidade (*background* familiar – raça, sexo, renda, educação dos pais etc). As primeiras são denominadas de características de responsabilidade, de forma que, toda desigualdade social oriunda destas é considerada justa. As segundas, por sua vez, são conhecidas por características de não responsabilidade e apontadas como origens de desigualdade social injusta, e, portanto, para esta corrente teórica devem ser compensadas.

Nesse rol de discussões, há diversos estudos que apontam para a influência de variáveis que fogem do controle individual nos resultados econômicos individuais². Os resultados encontrados na maioria dos trabalhos da literatura sugerem que a influência dos fatores de não responsabilidade responde por pouco mais de 30% no resultado final do indivíduo. Esta influência pode se tornar mais decisiva em cenários com baixos níveis de renda e educação. Neste contexto, a presente pesquisa busca avaliar como as variáveis de não responsabilidade determinam o desempenho educacional dos estudantes brasileiros, ou seja, como a desigualdade de oportunidade afeta o desempenho dos alunos que prestaram o Exame Nacional do Ensino Médio (ENEM). Sabe-se, também, que o ENEM é o passaporte da maioria das instituições de ensino superior no Brasil, em outras palavras, o acesso às oportunidades. Logo, investigar tais determinantes é de fundamental importância para compreender como a desigualdade de oportunidades dificulta a mobilidade intergeracional dos estudantes brasileiros que concluem o ensino médio.

A ideia é saber se a origem familiar interfere nos resultados individuais dos alunos e para tanto, assumimos que os indivíduos são responsáveis por uma parcela da diferença nos resultados, como aquelas originadas da pura escolha, embora, não por outras, como as originadas pelo *background* social, discriminação sexual, racial, entre outras. Para isso, distingue-se os indivíduos por tipos – características sociais – que prestaram o exame, em seguida, emprega-se duas abordagens de estimações, em especial, métodos não paramétricos para apurar prováveis diferenças.

Para tanto, o trabalho está organizado em cinco seções além desta introdução. Na seção 2, descreve-se a questão conceitual da teoria da igualdade de oportunidades. Na seção 3,

¹ Ver Roemer (1998).

² Ver Torbjorn *et. al.* (2010), Gamboa (2011), Figueiredo & Ziegelmann (2013).

apresenta-se a metodologia e a descrição dos dados. A seção 4 avalia-se os principais resultados e por fim as considerações finais.

2 IGUALDADE DE OPORTUNIDADES

A igualdade de oportunidades constitui um princípio básico para atenuar desigualdades entre indivíduos. De acordo com Lefranc, Pistolesi e Trannoy (2009), os resultados dependem de um conjunto de fatores determinísticos e aleatórios relevantes para o sucesso ou fracasso do agente econômico. Para analisar a igualdade de oportunidades necessita-se identificar as variáveis ou características aos quais os indivíduos são designados responsáveis por elas, e também as variáveis que não o são, objetivando assim, calcular o conjunto de oportunidades dos indivíduos.

Para tornar o conceito mais claro, considerem uma pequena ilustração: Antônio (A) e Bruno (B) possuem um conjunto de oportunidades idêntico, ou seja, pertencem a famílias de mesma classe social, frequentaram as mesmas escolas, são da mesma raça, sexo, etc. Após a conclusão do ensino médio, ambos decidem trabalhar na mesma profissão. Todavia, (A) trabalha duro e destina poucas horas ao lazer, enquanto que (B) destina maior parte do tempo ao lazer. Tais escolhas imerge uma desigualdade de renda, visto que: (A) obtém um maior nível de renda proveniente do trabalho do que (B). Nesta situação genuína de pura escolha dos agentes, a desigualdade existente é um problema social? Para os igualitários de resultado, a resposta é sim, pois qualquer desigualdade social é indesejável. Por outro lado, os igualitários de oportunidades, consideram a resposta não, dado que, a diferença nos resultados é nitidamente devido a uma escolha ótima de preferência individual.

Agora, considerem o caso de duas jovens: Carol (C), pertencente a uma família rica e tradicional; e Daniela (D), uma jovem com origem de família bastante humilde. Ambas desejam tornarem-se cirurgiãs plásticas, entretanto, sabe-se que para conseguir ingressar em tal profissão os gastos são substanciais, de modo que, a família de (D) não possuem recursos suficientes, o que a faz decidir tornar-se odontóloga. Além disso, sabe-se que a renda de uma cirurgiã plástica é bem superior a de uma odontóloga. Nesse contexto, dado que as duas jovens ao trabalhar exercem um nível de esforço semelhante, (C) torna-se rica rapidamente, enquanto que (D) apresenta um padrão de vida apenas considerável. Neste caso, assim como os igualitários de resultados, os igualitários de oportunidades consideram a desigualdade observada injusta ou indesejável.

Os casos referidos acima nos ilustram duas importantes questões. A primeira diz respeito às diferenças do conjunto de oportunidades enfrentadas pelos indivíduos, enquanto que a segunda corrobora sobre a influência das variáveis no resultado individual. As variáveis relacionadas as responsabilidades como, por exemplo, a quantidade de horas trabalhadas são denominadas de esforço e são estas determinantes para o diferencial nos resultados econômicos individuais. Contudo, existe uma gama de variáveis que fogem ao escopo da responsabilidade do indivíduo, como por exemplo, a origem, raça, sexo, *background* familiar, etc, variáveis denominadas de não responsabilidade ou circunstâncias.

No caso de (A) e (B), temos circunstâncias semelhantes, com esforço diferenciado. Por sua vez, o caso de (C) e (D), embora (D) se esforce da mesma forma que (C), as circunstâncias a impede de obter melhores resultados. A literatura que trata igualdade de oportunidades segue dois princípios: *i*) Princípio de Compensação; e *ii*) Princípio de Recompensa. Em termos práticos, temos que (A) dever ser recompensado pelo esforço e (D) deve ser compensada devido à circunstância.

De acordo com O'Neill *et al* (2001), e assumindo que não há políticas compensatórias, admite-se que o conjunto de oportunidades de um indivíduo é representado por S_x , sendo o mesmo determinado por um vetor de características de não responsabilidade, x . Tal fato faz

com que os resultados do indivíduo dependa do nível de esforço empregado, ou seja, das características de responsabilidade individual condicionado ao conjunto de oportunidade que o indivíduo possui.

Assim, os indivíduos podem optar por diferentes escolhas e diferentes resultados, estes podem ser resumidos por $z = y[e, x]$, em que z representa a utilidade ou renda (no caso da pesquisa, o desempenho no ENEM) ao longo do tempo; e e representa o esforço individual empregado. Assume-se, também, que a função de distribuição de e é contínua e postula-se duas suposições:

- **SINC** (*Strictly Increasing*): $z = y[e, x]$ é estritamente crescente em e . Tal pressuposto é bastante plausível, dado que, quanto maior o nível de esforço maior a utilidade resultante. Definindo $F_z^*(z|x)$ e $F_e^*(e|x)$ a função de distribuição acumulada (fda) de z e e , ambas condicionadas a x , respectivamente. Logo, podemos expressar SINC por:

$$F_z^*(y[e, x]|x) = F_e^*(e|x) \quad (1)$$

A equação 1 nos indica que o nível de esforço de um indivíduo (e) condicionado ao seu tipo - conjunto de oportunidades - será menor que $\alpha - th$ percentil da distribuição do seu esforço se, e somente se, o resultado for inferior a $\alpha - th$ percentil.

- **IND** (independente): $F_e^*(e|x)$ é independente de x . Esta suposição significa que não são assumidas diferenças na fda do esforço entre tipos diferentes, ou seja, entre indivíduos com diferentes características de não responsabilidade pertencentes ao mesmo percentil. Tal pressuposto é igualmente plausível, visto que os indivíduos não poderiam ser responsáveis pelo nível de esforço empregado, caso estes dependessem das circunstâncias aos quais estivessem inseridos. A equação 1 adicionada a suposição IND implica no Axioma de Identificação de Roemer (RIA).
- **RIA** (*Roemer's Identification Axiom*): $F_z^*(y[e, x]|x) = F_e^*(e|x) \Rightarrow e' = e''$. A RIA implica que dois indivíduos com diferentes conjuntos de oportunidades, mas com o mesmo percentil da distribuição dentro do seu tipo, exerce o mesmo nível de esforço.

Com base nestas suposições, define-se $\pi = F_z^*(z|x)$ como a fda do resultado de z condicionado as características de não responsabilidade, x . Analogamente, assume-se que esta função é estritamente crescente em z . $F_z^{-1}(\pi|x)$ expressa o resultado obtido - desempenho dos alunos no ENEM - pelo indivíduo do tipo x e que estava no $100 * \pi - th$ percentil da fda do resultado dentro do seu tipo. De acordo com a RIA, $F_z^{-1}(\pi|x)$ equivale a observar $y[p, x]$. Entretanto, $F_z^{-1}(\pi|x)$ fornecerá informações das características de responsabilidade e não responsabilidade dos indivíduos. De acordo com O'Neill *et al* (2001), isto possibilita desenhar o resultado - desempenho no ENEM com uma função de $\pi \in [0,1]$ para diferentes valores de x . Sendo o conjunto de oportunidades para um particular tipo de x determinado por alguns resultados do tipo x , e podendo ser obtido ao variar suas características de responsabilidade e ou π . Deste modo, o conjunto de oportunidade do indivíduo do tipo x será:

$$S_x = (z, \pi) \in (\mathbb{R}^+ \times [0,1])z = F_z^{*-1}(\pi|x) \quad (2)$$

em que \mathbb{R}^+ representa o conjunto dos números reais não negativos. Se F_z^{*-1} estiver disponível, pode-se descrever o conjunto de oportunidades para diferentes tipos de indivíduos e também qual a extensão das opções diferentes ou níveis de esforço que produzem diferentes resultados.

Usualmente, O'Neill *et al* (2001) assume que a sociedade determina de algum modo os elementos de x . Considera-se, no entanto, que x é uma variável multidimensional composta por elementos como raça, sexo, *background* familiar e habilidade inata. Os dados, assim exigidos, são complexos e difíceis de serem delimitados, em que pese a estrutura de mobilidade intergeracional que leva em conta o desempenho dos filhos (z) em função das características do pais (x).

Há diversos métodos para obter o conjunto de oportunidades, por exemplo, os empregados por Roemer (1998); Betts & Roemer (1998); Van de Gaer *et al* (1998). Contudo, para evitar problemas com viés de especificação³, optou-se em utilizar as curvas de densidade acumulada do resultado de z condicionada as características de não responsabilidade de x , o que permitirá analisar as diferenças de oportunidades enfrentadas entre os vários tipos de conjuntos. Dessa forma, a estimação dar-se-á pelo método não paramétrico kernel bivariado, o qual será descrito na próxima seção.

3 METODOLOGIA

A seguir apresenta-se resumidamente as ferramentas utilizadas para atingir os objetivos propostos.

3.1 Estimador de Kernel Bivariado de Densidade Acumulada

Conforme Cameron & Trivedi (2005), o estimador de densidade de Kernel, introduzido por Rosenblatt (1956) é uma generalização do histograma e que faz uso de uma função de ponderação alternativa expressa por:

$$\hat{f}_{(x_0)} = \frac{1}{Nh} \sum_{i=1}^n k\left(\frac{x_i - x_0}{h}\right) \quad (3)$$

A ponderação da função $K(\cdot)$ é denominada função de kernel. O parâmetro h é um parâmetro de alisamento chamado de largura da banda e a densidade é estimada através da avaliação de $\hat{f}_{(x_0)}$ sob uma vasta gama de valores de x_0 utilizado na formação de um histograma. Usualmente a avaliação se dá nos valores amostrais de x (x_1, \dots, x_N), ou seja, a estimativa de densidade de Kernel torna-se mais suave do que a de um histograma.

A função Kernel $K(\cdot)$ é uma função contínua, simétrica em torno de zero, e satisfaz as condições de contorno limitado adicionais. Com base em Lee (1996), $K(\cdot)$ satisfaz:

1. $K(z)$ é contínua e simétrica em torno de 0;
2. $\int K(z)dz = 1$;
3. Ou (a) $k(z) = 0$ se $z \geq z_0$ para algum z_0 ; ou (b) $|z|K(z) \rightarrow 0$ com $|z| \rightarrow 0$;
4. $\int K(z)dz = k$, onde k é uma constante.

De acordo com Cameron & Trivedi (2005), na prática, as funções de Kernel funcionam melhor se atender a condições (3a), em vez de apenas satisfazer a condição mais fraca (3b). Além disso, restringir para o intervalo $[-1,1]$ em vez de $[-z_0, z_0]$, simplesmente para uma normalização para que $K(z)$ seja restrita a $z \in [-1,1]$.

³ Variáveis relacionadas umas com as outras pode causar viés de especificação nas estimações. Ver detalhes em Cameron & Trivedi, 2005.

Atendidas as condições acima, denotando o vetor de características de não responsabilidade x – variáveis de circunstância – e o resultado obtido pelo indivíduo (z), a distribuição de z pode ser expressa como:

$$f[z|x] = \frac{f[z|x]}{f_x[x]} \quad (4)$$

em que $f_x[x]$ representa a distribuição marginal das circunstâncias e $f[z|x]$ a distribuição conjunta de z e x . Para estimar (4), substitui-se o numerador e o denominador por estimativas. A distribuição marginal das circunstâncias é estimada utilizando a técnica adaptada de densidade de kernel para distribuições univariadas:

$$\hat{f}_{xA}[x] = \frac{1}{nh_x} \sum_{i=1}^n \frac{1}{w_i} k \left[\frac{x - x_i}{w_i h_x} \right] \quad (5)$$

A distribuição conjunta das circunstâncias e resultados obtidos (numerado da equação (4)) é obtida por:

$$\hat{f}_A[z|x] = \frac{1}{nh_z h_x} \sum_{i=1}^n \frac{1}{w_i^2} k \left[\frac{z - z_i}{w_i h_z} \right] \left[\frac{x - x_i}{w_i h_x} \right] \quad (6)$$

O estimador de kernel adaptativo ajusta a largura da janela tornando-a mais estreita quando a densidade é maior, e amplia quando a densidade é menor, preservando os detalhes em que os dados são abundantes e reduzindo o ruído quando os dados são esparsos. A janela local dos fatores usados são determinados por:

$$w_i = \left[\frac{\hat{f}_g}{\hat{f}_k[z_i, x_i]} \right]^{1/2} \quad (7)$$

em que $\hat{f}_k[z_i, x_i]$ é o estimador kernel de janela fixa de $f[z, x]$; e \hat{f}_g é a média geométrica de $\hat{f}_k[z, x]$. Conforme destaca O'Neill et al (2001), a operacionalização deste procedimento envolve uma estratégia de estimação em dois passos. Primeiro, $\hat{f}_k[z, x]$ é estimado utilizando uma janela de largura fixa, obtida com uma largura inicial através de *Scott's optimal bandwidth*⁴.

O segundo passo, destina-se a obter a densidade utilizando pesos envolvidos na construção da densidade final das equações (5) e (6). Assume-se que a kernel é multiplicativa, conforme Tiede (1998), tornando-a em uma simples expressão de fda. Dessa forma, substitui os termos da equação (4) pelas estimativas das equações (5) e (7), obtendo-se a distribuição condicional:

$$\hat{F}_{zA}^*[z|x] = \frac{\sum_{i=1}^n \frac{1}{w_i} K \left[\frac{x - x_i}{w_i h_x} \right] G \left[\frac{z - z_i}{w_i h_z} \right]}{\sum_{i=1}^n \frac{1}{w_i} K \left[\frac{x - x_i}{w_i h_x} \right]} \quad (8)$$

⁴ Ver Scott (1991).

em que $G(z) = \int_{-\infty}^z K(t)dt$ representa a fda da função de kernel. O conjunto de oportunidades para o indivíduo dado o seu tipo x_i , pode assim ser estimado por:

$$\hat{S}_{x_i} = \{(z, \pi) \in (\mathbb{R}^+ \times [0,1]) | z = \hat{F}_{zA}^{*-1}[\pi|x]\} \quad (9)$$

3.2 Função de Produção Educacional

A Função de Produção de Educação, denominada de FPE, analisa a forma como os diversos insumos do processo educacional podem afetar os resultados educacionais dos indivíduos, servindo de respaldo para a escolha dos principais determinantes educacionais, e com isso proporcionando um guia para a aplicação de políticas públicas alocativas. De acordo com Hanuschek (1970, 1979 e 2007), a FPE é dada por:

$$A_{it} = g(F_i^t, P_i^t, I_i S_i^t) \quad (10)$$

em que A_{it} é o vetor educacional realizado pelo estudante i -th no tempo t ; F_i^t é o vetor de características individuais e familiares do i -th estudante acumulado no período t ; P_i^t é o vetor do corpo discente (influências de pares), ou seja, variáveis socioeconômicas e *background* familiar de outros estudantes na escola acumulados no período t ; I_i é o vetor de dotações iniciais do i -ésimo indivíduo; e S_i^t é o vetor de insumos escolares relevantes para o i -th estudante acumulados no período t .

A partir desse modelo, Albernaz, Ferreira e Franco (2002) estimaram a função de produção educacional para o ensino fundamental através de um modelo linear hierárquico, utilizando como base de dados o SAEB para as 8ª Série de 1999. A conclusão dos autores é de que cerca de 80% da variância do desempenho médio entre as escolas, deve-se a diferenças na composição sócio-econômica de seus alunos.

Cabe aqui ressaltar um aspecto relevante para a estimação de (10). Refere-se omissão de variáveis como, por exemplo, a habilidade do indivíduo. Infelizmente esta é uma variável de difícil mensuração e a sua negligência na FPE torna os regressores do Ordinary least square (OLS) endógenos. Além disso, temos problemas de simultaneidade entre renda e educação. Contudo, apenas para fins de dectar relações, seguimos a proposta de Brunello e Checchi (2006), estimando um modelo log-linear da seguinte expressão:

$$y_i = \alpha_i + \beta_1 X_i + \beta_2 Q_i + \varepsilon_i \quad (11)$$

em que y_i é o log dos resultados individuais do ENEM; α_i são as dummies de sexo, localização, dependência administrativa em que o indivíduo estudou; X é um vetor com as variáveis de não responsabilidade, incluindo o *background* familiar; Q é a variável que mede a qualidade da educação representada pelo desempenho do SAEB por escola; ε_i é o vetor de erros. As informações da qualidade da escola é justificada pelo fato de normalmente associar-se que indivíduos com baixo *background* familiar podem estudar em escolas com baixa qualidade, sendo assim, sobestimar os outros parâmetros da regressão.

3.3 Descrição dos Dados

As densidades de Kernel foram obtidas tendo como base os microdados do (ENEM) captados pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira

(INEP), no ano de 2010, composto por 2.937.253 estudantes de todas as regiões do Brasil. O Enem foi criado em 1998 com a finalidade de avaliar o desempenho dos estudantes brasileiros ao final do ensino médio. A partir de 2009, foram implementadas algumas mudanças para contribuir com a democratização das oportunidades de acesso, visto que a partir desse ano, o exame passou a ser utilizado como mecanismo de seleção para o ingresso de nível superior.

Além disso, os microdados são divididos em duas bases de dados: *i*) variáveis relacionadas aos inscritos no exame; *ii*) variáveis do questionário socioeconômico⁵ (características dos indivíduos e de sua família). A junção destas bases de dados possibilita compilar informações a respeito a variáveis de não responsabilidade e de resultados individuais, necessários ao atendimento do objetivo da pesquisa. Desta maneira, as variáveis selecionadas são sintetizadas na Tabela 1:

Tabela 1 – Descrição das Variáveis

Variável	Descrição
Resultados individuais	Somatório dos pontos obtidos nas provas que contemplam quatro áreas de conhecimento: Linguagens, códigos e suas tecnologias; Matemática e suas tecnologias; Ciências da Natureza e suas tecnologias; e Ciências Humanas e suas tecnologias.
Renda familiar	Somatório da renda familiar que são distribuídas em cinco classes de renda, que vai desde as famílias que possuem renda de até um salário mínimo até as famílias com renda acima de nove salários mínimos, isto é: i) Renda familiar mínima: Até um salário mínimo; ii) Renda familiar baixa: 1 a 3 salários mínimos; iii) Renda familiar média: 3 a 6 salários mínimos; iv) Renda familiar Alta: 6 a 9 salários mínimos; v) Renda familiar Máxima: acima de 9 salários mínimos.
Escolaridades dos pais	Agregadas em quatro estratos: i) Pais com ensino fundamental incompleto e analfabetos: menos de 4 anos de estudos; ii) Pais com ensino fundamental e/ou médio: de cinco a onze anos de estudos; iii) Pais com ensino superior ou pós-graduação: mais de onze anos de estudos.
Dependência escolar	Se o aluno cursou predominantemente em escola pública ou privada. Cabe ressaltar aqui, que optou-se por excluir as escolas indígenas e escolas quilombolas pela falta de dados;
Sexo –	Feminino ou masculino;
Raça	Divididos em brancos e não brancos (aqui abrangendo pretos, pardos, amarelos e indígenas);
Localização da moradia	Zona rural ou zona urbana

Fonte: Elaboração dos autores, a partir de dados do ENEM.

⁵ Consideraram-se apenas os indivíduos que responderam ao questionário socioeconômico.

Por sua vez, a estimação da FPE, através de (10), foram obtidas com as variáveis descritas na Tabela 1 adicionada de uma variável que denota qualidade educacional. Para obter esta variável utilizou-se os microdados do Sistema Nacional de Avaliação da Educação Básica (SAEB). A escolha pelo SAEB se deve ao objetivo da base que é de obter informações sobre os níveis de aprendizagem demonstrados pelos alunos agregados por unidade escolar e nas respectivas redes de ensino e não pela avaliação de cada aluno individualmente.

4 RESULTADOS EMPÍRICOS

Nesta seção, serão apresentados os resultados da presente pesquisa. Primeiramente, realiza-se uma breve análise das estatísticas descritivas dos dados, pois os mesmos nos fornece importantes indicações que caracterizam o perfil do aluno que prestaram o exame. Posteriormente, controem-se as densidades de Kernel sob perspectiva nacional e regional e interações de circunstâncias. Por fim, infere-se sobre as relações das desigualdades de oportunidades no Brasil. A Tabela 2 sumariza as principais características dos estudantes, como também estatísticas descritivas dos dados.

Tabela 2 – Análise Descritiva

Estudantes por Variáveis de Não Responsabilidade				
Sexo		Raça		
Masculino	Feminino	Branco	Ñ Branco	
1.187.420	1.749.833	1.326.571	1.610.682	
Alunos Conforme Renda Familiar				
Mínima	Baixa	Média	Alta	Máxima
709.298	1.439.917	492.999	129.143	165.896
Tipo de Escola		Escolaridade da Mãe		
Pública	Privada	Analfabeta	Nível Médio	Nível Superior
2.374.253	563.000	911.794	1.538.368	487.091
Regiões dos Alunos e Nota Média				
Centro Oeste	Nordeste	Norte	Sul	Sudeste
247.592	895.005	260.453	402.020	1.132.183
2591	2561	2533	2679	2721

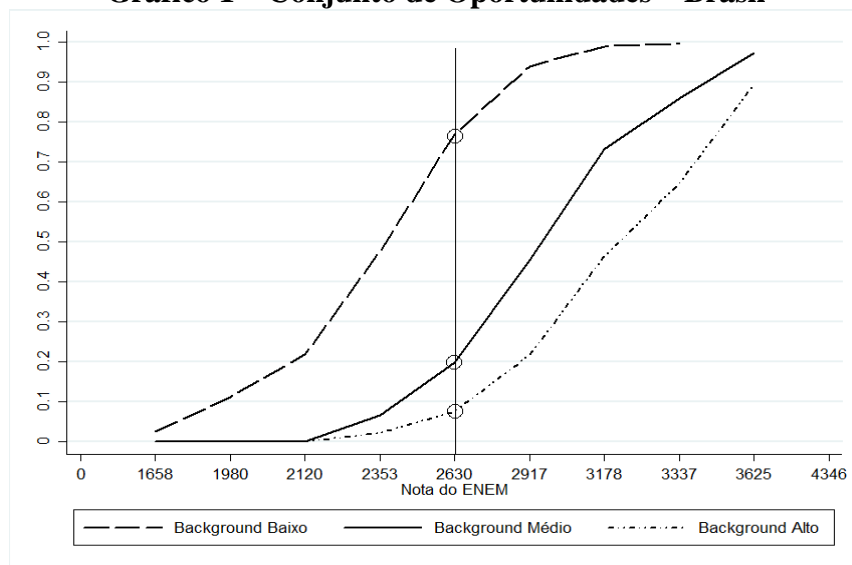
Fonte: Elaboração dos autores, a partir de dados do ENEM.

Verifica-se, também, que aproximadamente 60% dos que prestaram o exame são mulheres; mais de 80% vem de escola pública; aproximadamente 71% dos candidatos possuem renda familiar de no máximo três salários mínimos, onde 33% destes tem renda familiar de até 1 salário mínimo. Além do mais, constata-se que 31% dos candidatos são filhos de mães que estudaram no máximo quatro anos de estudos, ou seja, possuem no máximo o ensino fundamental.

Em termos regionais, constata-se que a nota média da Região Nordeste está aquém das Regiões Sul, Sudeste e Centro-Oeste e um pouco acima da Região Norte. Além disso, averigua-se que 38% dos candidatos moram na região sudeste e outros 30% moram no nordeste do país, seguidos de 13%, 8,83% e 8,43% das Região Sul, Norte e Centro-Oeste, respectivamente.

No que tange o conjunto de oportunidades nacional, verifica-se um elevado índice de desigualdade de oportunidades, expressos no esforço – em termos de probabilidade – necessários para que o indivíduo com baixo *background* social atinja a média nacional. Denota-se baixo *background* social, aqueles indivíduos, filhos de mães analfabetas, com renda familiar menor do que 1 (um) salário mínimo e que estudou predominantemente em instituição pública. O inverso se aplica aos indivíduos com alto *background* social, ou seja, estudantes filhos de mães com nível superior, renda familiar acima de 9 salários mínimos e que estudou predominantemente em escola particular. O Gráfico 1 apresenta tais resultados para o Brasil.

Gráfico 1 – Conjunto de Oportunidades – Brasil



Fonte: Elaboração dos autores, a partir de dados do ENEM.

A curva a direita representa os alunos que estão associadas às circunstâncias favoráveis e a curva a esquerda, por sua vez, os alunos associados as situações desfavoráveis. De acordo com os resultados, a probabilidade de um indivíduo com baixo *background* social atingir a nota média do ENEM (média de 2.630), é quase oito vezes maior do que um indivíduo com alto *background* social. Além do mais, a probabilidade de indivíduos com baixo *background* estar entre os 5% que obtém as melhores notas é de 0,26%, enquanto que os indivíduos com alto *background* é de 35%.

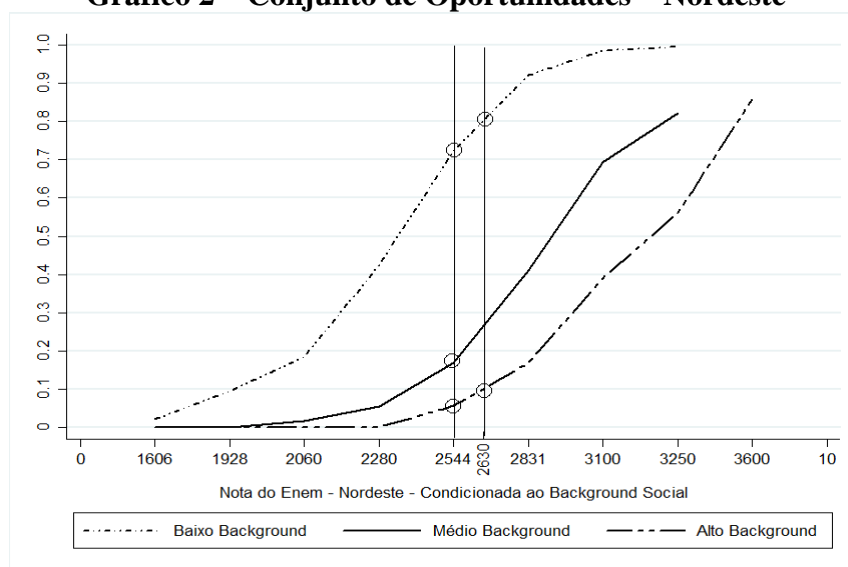
Em síntese, é praticamente inviável para aqueles penalizados pelas circunstâncias obter os melhores resultados, isto é, nos parece que alguns alunos são sumariamente excluídos do estrato social superior, enquanto que outros parece não precisar de esforçar para obter melhores resultados. Em outras palavras, o conjunto de oportunidades do primeiro grupo (baixo) é cerca de 135 vezes menor do que o segundo grupo (alto). A priori, parece existir um imenso abismo social expresso na desigualdade educacional do país.

Aplicando o mesmo instrumental para a Região Nordeste, os resultados apresentam diferenças substanciais. Considerando o mesmo cenário nacional (nota igual a média), o esforço deve ser ainda mais superior, cerca de 8 vezes maior. Agora, se considerarmos as probabilidades de obterem notas acima da média nacional, verifica-se que os estudantes nordestinos, filhos de mãe analfabeta, oriundos de escolas públicas e com renda familiar menor do que 1 (um) salário mínimo é de apenas 27%. Enquanto que um estudante com situação social oposta é de aproximadamente 95%. Em outras palavras, um estudante com baixo *background* social deve ser esforçar cerca de 3,5 vezes mais do que um estudante

nordestino com alto *background* para chegar ao resultado acima da média nacional. O Gráfico 2 mostra o conjunto de oportunidade para a Região Nordeste.

Para ter-se uma ideia do nível da desigualdade de oportunidades da Região Nordeste, quando analisamos as probabilidades de estar entre os 5% que obtiveram os melhores desempenhos no ENEM, verifica-se resultados semelhantes ao cenário nacional. A chance de um indivíduo pertencente ao percentil social inferior ser inserido, é cerca de 0,4%. Por outro lado, se considerarmos os estudantes nordestinos do percentil social superior, verifica-se que a possibilidade do mesmo de estar entre os 5% é de aproximadamente 44%. Em síntese, o conjunto de oportunidades de um nordestino com baixo *background* social é 114 vezes menor do que indivíduos com alto *background*.

Gráfico 2 – Conjunto de Oportunidades – Nordeste



Fonte: Elaboração dos autores, a partir de dados do ENEM.

Destaca-se, ainda, que apesar dos resultados da Região Nordeste apontar para um menor nível de diferença de oportunidades em detrimento ao cenário nacional, não necessariamente significa menor desigualdade educacional. Notem que os resultados referem-se ao conjunto de oportunidade intragrupo – neste caso, analisamos apenas Região Nordeste – e não entre grupos (considerando as demais Regiões).

Tabela 3 – Probabilidade Condicional Acumulada – Nordeste

Percentil	Background Baixo	Background Médio	Background Alto
1%	0.019908	.	.
5%	0.093828	.	.
10%	0.184095	0.017223	.
25%	0.422787	0.054784	.
50%	0.723526	0.168852	0.055756
75%	0.921523	0.410143	0.167656
90%	0.986239	0.693236	0.390577
95%	0.996139	0.823402	0.560271
99%	.	.	0.858971

Fonte: Elaboração dos Autores, a partir de dados do ENEM.

E se compararmos os resultados da Região Nordeste com os da Região Sudeste? Bem, quando se analisa os estudantes sudestinos, constata-se que a possibilidade do estudante de baixo *background* social estar entre os 10% que obtém melhores resultados é de aproximadamente 3% contra 1,4%, se o mesmo for aluno nordestino. Em termos de conjunto de oportunidades entre grupos, averigua-se que o estudante sudestino pertencente ao estrato social inferior tem aproximadamente 21,5 menor possibilidade de estar entre os 10% melhores, contra 45,5 menor, caso o estudante seja nordestino. O fato do aluno morar no nordeste reduz em quase 100% as chances do mesmo atingir os 10% melhores desempenhos nacionais, comparados aos alunos da Região Sudeste. A Tabela 3 apresenta tais probabilidades.

Outra questão importante a ser destacada, diz respeito aos fatores – variáveis utilizadas nesta pesquisa. A Tabela 4 traz os principais determinantes da Desigualdade de oportunidades. Conforme exposto, a renda familiar apresentou-se como o fator mais preponderante na desigualdade de oportunidade educacional, seguindo da escolaridade da mãe, escolaridade do pai, tipo de escola frequentada pelo aluno, raça, localidade onde reside e sexo.

Tabela 4 – Probabilidade Condicional Acumulada – Nordeste

Variáveis Circunstanciais	Tipo			% Probabilidades		
	Analfabeto	Médio	Superior	Analfabeto	Médio	Superior
Pai	0,63	0,47	0,22	0,37	0,53	0,78
Mãe	0,64	0,49	0,26	0,36	0,51	0,74
Renda	Mínima	Média	Máxima	Mínima	Médio	Máxima
	0,71	0,33	0,12	0,29	0,67	0,88
Escola	Pública		Privada	Pública		Privada
	0,57		0,20	0,43		1,00
Raça	N Brancos		Brancos	N Brancos		Brancos
	0,57		0,42	0,43		1,00
Sexo	Feminino		Masculino	Feminino		Masculino
	0,54		0,45	0,46		1,00
Localização onde Mora	Rural		Urbana	Rural		Urbana
	0,64		-	0,36		0,52

Fonte: Elaboração dos Autores, a partir de dados do ENEM.

Observa-se que a influência da mãe é ligeiramente superior a do pai como determinante. No entanto, a possibilidade de um estudante filho de mãe com nível superior ou pós-graduação – mais de onze anos de estudo – obter um resultado acima da média é mais do que o dobro comparado com o estudante filho de mãe analfabeta. Dentre as variáveis, a renda é sem dúvidas o fator mais preponderante na determinação do resultado, já que alunos que pertencem a faixa de renda mínima tem que se esforçar praticamente três vezes mais do que os indivíduos pertencentes a faixa máxima de renda para obterem resultados acima da média.

Em linhas gerais, os resultados mostram-se ainda mais expressivos quando avaliamos percentis superiores. Por exemplo, a possibilidade de um aluno pertencer ao nível mínimo de renda estar entre os 5% melhores no exame é de 45 vezes menos que os alunos que fazem parte da faixa superior da renda. No entanto, deve-se ressaltar que a renda em geral carrega em si outras características.

Contudo, com a finalidade de iniciar a busca por respostas para questões como: quais fatores estão associados a esta desigualdade de oportunidades? E o que poderia ser feito para atenuar tais desigualdades? Nesse preceito, estimou-se, a princípio por OLS⁶, a função de produção da educação expressa pela Equação (10). Todavia, destaca-se que a fragilidade da estimação dar-se-á, principalmente, pela falta de dados que contemplem toda a expressão. Entretanto, atualmente buscam-se mecanismos que possibilitem corrigir o viés. Isto sem dúvidas é a grande tarefa para futuros trabalhos que contemplem o tema em questão.

Embora, assumam-se aqui que a regressão por OLS não acure precisamente os parâmetros, eles confirmam a direção dos efeitos enfatizados nos resultados estimados não parametricamente. Designadamente, os resultados contidos na Tabela 5 apontam respectivamente o tipo de escola, a renda familiar e educação dos pais como principais fatores na determinação do resultado no exame, mas também contemplam raça, sexo e localização dos indivíduos.

Tabela 5 – Estimação por OLS

Variáveis	OLS	Robusta
Lognota	Coeficientes	Coeficientes
Anos de Estudos do Pai	0.01236***	0.01236***
Anos de Estudos da Mãe	0.01254***	0.01254***
Renda Familiar	0.02268***	0.02268***
Localização	0.001825**	0.001825**
Tipo de Escola	0.03000***	0.03000***
Qualidade da Escola	0.00137***	0.00137***
Cor	0.02159***	0.02159***
Sexo	0.00479***	0.00479***
Constante	7.37245***	7.37245***

Fonte: Elaboração própria a partir de dados do ENEM/SAEB.

***1%; **5% significância

De acordo com a Tabela 5, observa-se o impacto positivo, de se obter melhores resultados no ENEM, para aqueles que moram na área urbana, são do sexo masculino e de cor branca. Além do mais, quanto maior a escolaridade e renda dos pais melhores os resultados individuais. Verifica-se que os resultados coincidem com a estimação não paramétrica, bem como o esperado na literatura, de que quanto maior o conjunto de oportunidades, melhores os resultados econômicos.

Além disso, os resultados permitem desconfiar principalmente da qualidade de nossas escolas como um dos principais fatores para desigualdade apontada. Não obstante, quando se olha para a variável “tipo de escola” verifica-se que esta tem influência significativa sobre o desempenho no exame. Em resumo, estudar em escola pública é uma grande desvantagem.

⁶ Ordinary Least Squares – Maiores detalhes ver Cameron & Trivedi (2005)

Ao assumir este fato com verdadeiro, diante dos resultados empíricos apresentados destaca-se “o porquê disto”. Assim, parece plausível aceitar que a qualidade de nossas escolas pode ser a grande vilã da desigualdade de oportunidade educacional do Brasil.

Por outro lado, este resultado acarreta importantes implicações políticas educacionais, pois abona o fato de que, de certo modo, a desigualdade de oportunidade resultante da origem familiar pode ser suavizada com políticas que procurem minimizar a papel da família sobre o resultado educacional do indivíduo, melhorando a qualidade das escolas.

Em linhas gerais, os resultados (paramétricos e não paramétricos) apontam forte influência do *background* familiar no desempenho dos alunos do Enem, conforme os diversos resultados apresentados. Ou seja, verifica-se que tanto no Brasil como na Região Nordeste a origem importa. Em outras palavras, há uma grande desigualdade de oportunidades no Brasil, visto que, o desempenho escolar é afetado pelas circunstâncias que o aluno enfrenta. E estas, podem ser minimizadas com políticas voltadas e melhorar a qualidade de ensino.

5 CONSIDERAÇÕES FINAIS

Neste trabalho analisou-se o grau de igualdade de oportunidades no desempenho educacionais advindos dos fatores sociais dos estudantes brasileiros que prestaram o exame do ENEM. Procurou-se investigar se a origem familiar tem papel relevante no desempenho dos alunos que concluíram o ensino médio no Brasil e no Nordeste. A partir dos microdados do ENEM e do SAEB no ano de 2010, lançou-se mão de duas estratégias empíricas a fim de obter indicativos do papel das circunstâncias sociais no resultado econômico.

A amostra foi composta por 2.937.253 observações, dos quais 60% são mulheres (40% são do sexo masculino); 80% vêm de escola pública e aproximadamente 71% dos candidatos possuem renda familiar de no máximo três salários mínimos. Para a primeira estratégia, fez-se o recorte dos indivíduos em baixo *background* social, sendo estes filhos de mães analfabetas, com renda familiar menor do que 1 (um) salário mínimo e que estudou predominantemente em instituição pública; e alto *background* social, ou seja, estudantes filhos de mães com nível superior, renda familiar acima de 9 salários mínimos e que estudou predominantemente em escola particular.

Através do estimador de Kernel Bivariado de Densidade Acumulada, verificou-se que variáveis como: renda familiar, escolaridade dos pais, tipo de escola, etc., são fatores essenciais na determinação da desigualdade de oportunidades. Pelos resultados (em termos de probabilidade) parecem inviável aqueles penalizados pelas circunstâncias obter melhores resultados. Por exemplo, um indivíduo com baixo *background* precisa esforçar cerca de 99,38% a mais do que um indivíduo com alto *background* para estar entre os 5% com melhores notas. Destaca-se, também, que tais resultados tornam-se mais expressivos quando analisados para a Região Nordeste.

Sabendo que a maior parte dos estudantes possuem características que os classificam em baixo *background* social, conclui-se, *a priori*, que praticamente inexistente mobilidade intergeracional educacional. Os resultados nos levam a crer que alguns estudantes brasileiros são sumariamente excluídos do estrato social superior. Neste contexto e na procura de investigar a direção dos efeitos dessas variáveis, aplicou-se uma Função de Produção educacional baseada em Brunello e Checchi (2006), em que se utilizou o resultado do exame como uma *proxy* do resultado econômico individual. Além disso, inseriu-se uma *proxy* para a qualidade educacional.

Analogamente aos resultados não paramétricos, verificou-se que quanto melhores o conjunto de oportunidade (expressos no *background* familiar – escolaridade e renda dos pais), maiores são os desempenhos no ENEM. Além disso, observou-se um impacto positivo para

aqueles indivíduos que moram em área urbana, são do sexo masculino, são considerados brancos e que estudaram em escolas com melhores resultados no SAEB.

Todavia, os resultados encontrados abrem caminho para outras questões relacionadas à igualdade de oportunidades como o papel do sistema de cotas segundo os princípios de recompensa e compensação. Adicionalmente, abre espaço para discussões a respeito das políticas voltadas a melhoria da qualidade das escolas como alternativa para suavizar o papel da família sobre resultado educacional, visto que o Brasil ainda apresenta cenário com baixo nível de renda e educação.

Referencias

Albanaz, Â., Ferreira, F. H., & Franco, C. **Qualidade e Equidade na Educação Fundamental**. Texto para Discussão n° 455, Rio de Janeiro, 2002.

Betts, J.R. and J. Roemer. **Equalizing Opportunity Through Educational Finance Reform**. Mimeo, 1998.

Bourguignon, F., Ferreira, F. H., & Menéndez, M. **Inequality of Opportunity in Brazil**. Review of Income and Wealth, pp. 585-618, 2007.

Brunello, G., & Checchi, D. **Does School Tracking Affect Equality of Opportunity? New International Evidence**. IZA DP No. 2348, 2006.

CAMERON, A. C., TRIVEDI, P. K.; **Microeconometrics Methods And Applications**. Published In The United States Of America By Cambridge University Press, New York, 2005.

Dworkin, R. **What is Equality? Part 2: Equality of Resources**. Philosophy and Public Affairs, Vol. 10, No. 4, pp. 283-345, 1981.

ENEM. (s.d.). Microdados. Acesso em 2012, disponível em INEP: <http://portal.inep.gov.br/>

Figueiredo, E. A., Ziegelmann, F. A. **Estimation of Opportunity Inequality in Brazil using Nonparametric Local Logistic Regression**, 2013.

Fleurbaey, M. **Equal opportunity or equal social outcome**. Economics and Philosophy, pp. 25-55, 1995.

Haegeland, T., Kirkeboen L. J., Raaum, O., Salvanes, K. G. **Why children of college graduates outperform their schoolmates. A study of cousins and adoptees**. Discussion papers n° 628, Statistics Norway, Research Department. 2010.

Hanushek, E. A. **The Production of Education, Teacher Quality, and Efficiency**. In: **Do Teachers Make a Difference?** (pp. 79-99). U.S. Office of Education. 1970.

Hanushek, E. A. **Schools and the Equal Opportunity Problem**. In: L. Woessmann, & P. E. Peterson (Eds.). Cambridge: MIT Press, 2007.

Hanushek, E. A. **Conceptual and Empirical Issues in the Estimation of Educational Production Functions**. The Journal of Human Resources, pp. 351-388, 1979.

Harsanyi, J. C. **Cardinal Utility in Welfare Economics and in the Theory of Risk-taking**. Journal of Political Economy, University of Chicago Press, vol. 61, pp. 434, 1953.

Lee, M. J. **Methods of Moments and Semiparametric Econometrics for Limited Dependent Variable Models.** Berlin/New York, Springer-Verlag, 1996.

Lefranc, A. & Pistoiesi, N. & Trannoy, A. **Equality of opportunity and luck: Definitions and testable conditions, with an application to income in France.** Journal of Public Economics, Elsevier, vol. 93(11-12), pp. 1189-1207, 2009.

O'Neill, D., Sweetman, O., & Van DeGaer, D. **Equality of opportunity and kernel density estimation: an application to intergenerational mobility.** A research Annual, pp. 259–274, 2001.

Rawls, J. **A Theory of Justice.** Cambridge: Review. 1971.

Roemer, J. **Equality of Opportunity.** Harvard University Press. 1998.

SAEB. (s.d.). Microdados. Acesso em 2012, disponível em INEP: <http://portal.inep.gov.br/>

Sen, A. **The Welfare Basis of Real Income Comparisons: A Reply.** Journal of Economic Literature, American Economic Association, vol. 18(4), pp. 1547-52, 1980.

Trede, M. **Making mobility visible: a Graphical Device.** Economic Letters, Vol. 59, pp. 77-82, 1998.

Van de gaer, D, E.Schokkaert, and M. Martinez. **Measuring Intergenerational Mobility and Equality of Opportunity.** Economics Department Working, 1998.

Waltenberg, F. D. **Inequality of Opportunity in Educational Achievement in Latin America: Evidence from PIZA 2006-2009.** Discussion paper n^o 49. Center for Studies on Inequality and Development, 2011.