



UNIVERSIDADE FEDERAL DA PARAÍBA
CENTRO DE CIÊNCIAS SOCIAIS APLICADAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA
CURSO DE MESTRADO EM ECONOMIA

WALLACE PATRICK SANTOS DE FARIAS SOUZA

**O IMPACTO DO BACKGROUND FAMILIAR SOBRE O ESFORÇO E SUA
INFLUÊNCIA NO RESULTADO ECONÔMICO DOS INDIVÍDUOS: UMA
APLICAÇÃO DO MODELO DE ROY PARA O BRASIL**

JOÃO PESSOA - PB

2014

WALLACE PATRICK SANTOS DE FARIAS SOUZA

**O IMPACTO DO BACKGROUND FAMILIAR SOBRE O ESFORÇO E SUA
INFLUÊNCIA NO RESULTADO ECONÔMICO DOS INDIVÍDUOS: UMA
APLICAÇÃO DO MODELO DE ROY PARA O BRASIL**

Dissertação apresentada ao Programa de Pós-Graduação em Economia da Universidade Federal da Paraíba – UFPB, em cumprimento às exigências de conclusão do Curso de Mestrado em Economia.

Orientador: Prof. Dr. Erik Figueiredo

JOÃO PESSOA - PB

2014

S729i Souza, Wallace Patrick Santos de Farias.

O impacto do background familiar sobre o esforço e sua influência no resultado econômico dos indivíduos: uma aplicação do Modelo de Roy para o Brasil / Wallace Patrick Santos de Farias Souza.- João Pessoa, 2014.

36f.

Orientador: Erik Figueiredo

Dissertação (Mestrado) - UFPB/CCSA

**O IMPACTO DO BACKGROUND FAMILIAR SOBRE O ESFORÇO E SUA
INFLUÊNCIA NO RESULTADO ECONÔMICO DOS INDIVÍDUOS: UMA
APLICAÇÃO DO MODELO DE ROY PARA O BRASIL**

Dissertação apresentada ao Programa de Pós-Graduação em Economia da Universidade Federal da Paraíba – UFPB, como requisito para conclusão do curso de Mestrado em Economia.

Submetida à apreciação da banca examinadora, sendo aprovada em: ____/____/____.

BANCA EXAMINADORA



Prof. Dr. Erik Alencar de Figueiredo

Universidade Federal da Paraíba (UFPB)

Orientador



Prof. Dr. José Luis da Silva Netto Júnior

Universidade Federal da Paraíba (UFPB)

Examinador Interno



Prof. Dr. Francisco Soares de Lima

Universidade do Estado do Rio Grande do Norte (UERN)

Examinador Externo



Prof. Dr. Italo Fittipaldi

Universidade Federal da Paraíba (UFPB)

Examinador Interno

*Aos meus familiares, amigos e a todos
pelo amor, por me ajudar e me
acompanhar em todas as etapas da
minha vida. Muito Obrigado.*

AGRADECIMENTOS

Primeiramente agradeço a Deus por se fazer presente em minha vida e me conceder saúde, força e perseverança para alcançar meus objetivos.

Aos meus pais, Alcivan e Omanita, a minha irmã Withaynara, minha vó Raimunda Ramos (Tanina), e meu Tio/Primo Zenildo (Dedé) por sempre me apoiar e acreditar no meu potencial.

À minha vó Antônia Anita, a todos os meus tios e tias, ao meu avô Omar, aos meus primos e primas, por todos os momentos que torceram pela conclusão dessa etapa em minha vida.

Ao meu avô Chico Leonel (In memoriam) por sempre acreditar e sempre falar que eu conseguiria concluir este curso. Sem dúvidas esta etapa é dedicada a ele.

Ao meu orientador prof. Dr. Erik Alencar de Figueiredo, por todas as oportunidades concedidas durante o curso e pelos conhecimentos repassados.

À todos os amigos de curso pelas horas de estudo e momentos compartilhados.

À minha família do coração de João Pessoa, Etevaldo, Herbert, Vinícius, Ana, Fernanda, Patrícia, Paola, pelo convívio diário e por me ajudarem a ser uma pessoa melhor.

Ao meu grande amigo e ex professor Etevaldo, por todo apoio e ensinamentos durante esse tempo.

À minha namorada Ana Cláudia Annegues, pelo convívio, os estudos e todos os ótimos momentos juntos.

Aos meus professores que desde a graduação me proporcionaram conhecimentos para a minha formação acadêmica e pessoal.

Aos membros da banca Prof. Dr. Francisco Soares de Lima e André, pela participação e contribuições a este trabalho.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior- CAPES pela oportunidade de dedicar-me exclusivamente a vida acadêmica.

Ao pessoal da secretaria do mestrado, Risomar, Carol e Ricardo, pela presteza e apoio no decorrer do curso.

A todos que fazem parte do Núcleo de Estudos em Economia Social (NEES), pelo conhecimento que compartilhamos e construção de um grupo comprometido com a pesquisa e produção, sempre com muito respeito e seriedade.

Enfim, a todos que de uma forma direta ou indireta contribuíram para a conclusão do meu curso de mestrado e realização deste trabalho. Sabendo que todo agradecimento é um reconhecimento de que, por mais que um trabalho acadêmico possa parecer individual, ele é, na verdade, fruto de muitos outros esforços individuais e coletivos.

RESUMO

Este estudo mensura a influencia que a educação dos pais exerce sobre o esforço e sobre o nível de rendimento auferido pelos indivíduos. O esforço será aproximado através do nível de escolaridade, dividindo os indivíduos em dois grupos, i.e, sem nível superior (“Baixo”) representando indivíduos que desempenham um menor esforço e com nível superior (“Alto”) indicando um esforço mais elevado. A estratégia empírica de estimação foi baseada em três passos: i) a estimação do efeito de tratamento médio convencional por diferentes métodos, seguindo Heckman e Honoré (1990) e Heckman e Vytlačil (1999); ii) os *bounds* de variação de Shaikh e Vytlačil (2004) para o impacto do tratamento e iii) um teste de robustez para as estimativas que inclui a variável instrumental taxa de desemprego como forma de endogeneizar os anos de estudo dos indivíduos. Os resultados, tendo como base os dados da Pesquisa Nacional por Amostra de Domicílios (PNAD) do ano de 1996, indicam um efeito significativo da educação do pai e da mãe, sobretudo nos indivíduos menos escolarizados. Isso pode ser explicado, entre outros fatores, pelo menor nível de renda e consequentemente um maior grau de dependência desses indivíduos do seu *background* familiar, seguindo os pressupostos da desigualdade de oportunidades. Por fim, a não utilização da variável instrumental tende a superestimar os coeficientes do efeito de tratamento, sobretudo a educação do pai em detrimento da educação da mãe.

Palavras-Chave: Desigualdade de Oportunidades, Modelo de Roy, Efeito de Tratamento.

Classificação-JEL: C14, C21, D63

ABSTRACT

This study measures the influence that parents' education has on the effort and the level of income earned by individuals. The effort will be approximated by education levels, dividing the individuals into two groups, ie, no upper level ("Low"), representing individuals who play a lesser effort, and higher level ("High") indicating a higher effort. The empirical estimation strategy was based on three steps: i) estimating the average effect of conventional treatment by different methods, following Heckman and Honore (1990) and Heckman and Vytlačil (1999), ii) the bounds of variation of Shaikh and Vytlačil (2004) for the impact of treatment and iii) a test of robustness for estimating instrumental variable that includes the unemployment rate as a way to endogenize the years of schooling of individuals. The results, based on data from the National Sample Survey (PNAD) of 1996, indicate a significant effect of education of father and mother, especially in less educated individuals. This can be explained, among other factors, by the lower level of income and consequently a higher degree of dependence on these individuals of their family background, following the assumptions of unequal opportunities. Finally, the non-use of instrumental variable coefficients tends to overestimate the effect of treatment, especially education of the father rather than the mother's education.

Key-Words: Inequality of Opportunity, Roy Model, Treatment Effect

JEL-Classification: C14, C21, D63

LISTA DE TABELAS

Tabela 1 – Estatísticas Descritivas: PNAD de 1996	28
Tabela 2 – Equação de Renda dos Indivíduos (Log do Salário por Hora)	29
Tabela 3 – Efeito de Tratamento Médio – Educação do Pai (Tratamento)	30
Tabela 4 – Efeito de Tratamento Médio – Educação da Mãe (Tratamento)	30
Tabela 5: Estimação de Bounds – Educação do Pai (Tratamento)	31
Tabela 6: Estimação de Bounds – Educação da Mãe (Tratamento)	32
Tabela 7 – Efeito de Tratamento Médio – Variável Instrumental (Taxa de Desemprego)	32

LISTA DE FIGURAS

Figura 1 – Conjunto de Oportunidades (indivíduos com nível superior)	15
Figura 2 – Conjunto de Oportunidades (indivíduos sem nível superior)	16
Figura 3: Densidades Kernel – Educação do Pai (Tratamento).....	38
Figura 4: Densidades Kernel – Educação da Mãe (Tratamento).....	38

SUMÁRIO

1. Introdução	13
2. Procedimentos Teóricos e Metodológicos	17
2.1 Desigualdade de Oportunidades: Aspectos Teóricos.....	17
2.2 O Problema de Identificação no Modelo de Roy.....	18
2.3 O Modelo de Roy Generalizado Não Paramétrico.....	19
2.4 Inferência.....	21
2.4.1 Estimador de Variáveis Instrumentais Baseados na Variância.....	21
2.4.2 Métodos <i>Pairwise-Matching</i>	22
3. Bounds sobre o Efeito de Tratamento Médio	24
3.1 Bounds de Shaikh e Vytlacil.....	24
3.2 Variável Instrumental.....	25
3.3 Inferência.....	25
3.3.1 Bounds de Shaikh e Vytlacil.....	26
4. Dados	27
5. Resultados	29
5.1 Efeito de Tratamento Médio.....	29
5.2 Estimação de Manski <i>Bounds</i>	31
5.3 Estimação com Variável Instrumental.....	32
6. Considerações Finais	34
7. Referências	35
Apêndice A	38
Densidades Kernel para o Efeito de Tratamento Médio.....	38

1. Introdução

O conceito de desigualdade de oportunidades, formulado por John Roemer (1993), surgiu como resultado das discussões filosóficas sobre teorias de justiça iniciadas em Rawls (1971). Influenciado pela ideia de compensação social estabelecida em Dworkin (1981a, 1981b),¹ Roemer (1993) postula que os resultados econômicos dos indivíduos dependem de fatores de circunstâncias (variáveis sobre as quais o indivíduo não possui controle) e do seu nível de esforço.

Seguindo esse conceito, diferente da abordagem igualitária moderna que considera a perfeita igualdade como o ideal de justiça social, tem-se que os resultados econômicos são consequência de dois grupos de variáveis:² i) as de não responsabilidade (circunstâncias), i.e. aquelas as quais estão fora do controle dos agentes econômicos, tais como, região de nascimento, raça e *background* familiar (escolaridade e ocupação dos pais) e; ii) as de responsabilidade são os fatores que os indivíduos possuem controle, isto é, anos de estudo, horas trabalhadas, decisão de migrar, entre outras. (Roemer, 1998).

Considerar-se-á a definição de esforço adotada em Borguignon et al. (2007), isto é, o esforço como uma variável observável definida a partir de um modelo paramétrico, no qual, endogenamente, depende das circunstâncias. Assim, o esforço aqui é aproximado a partir dos anos de estudo dos agentes, enquanto que dentro do conjunto de circunstâncias a ênfase é dada para a escolaridade dos pais, controlada para diversos outros fatores.

No entanto, dificuldades quanto à identificação de um modelo a ser adotado e a má definição relacionada ao conjunto de variáveis de esforço podem não traduzir a real mensuração dos resultados, quando se analisa diferencial de oportunidades entre indivíduos. Assim, a presente proposta vai além das abordagens anteriores e pretende investigar a influência do *background* familiar sobre o nível de esforço individual, dado que dois indivíduos (ou grupos) que exercem um nível de esforço aproximado podem auferir diferentes níveis de rendimentos devido às características que estão fora do seu controle.

Para tanto, o nível educacional individual, controlado por outras características, será usado como *proxy* de esforço e a educação dos pais (*background* familiar) como principal fator de circunstância dos indivíduos³. Em um primeiro momento, serão verificadas quais as circunstâncias que impactam o fato de o indivíduo ter ou não determinado nível de esforço, representado pelo nível educacional superior. Essa característica é usada para definir um nível de esforço maior (indivíduos com nível superior) ou menor (indivíduos sem nível superior).

Isso resulta em dois subgrupos a partir da amostra inicial, um com todos os indivíduos que tem nível superior e outro com indivíduos sem nível superior. Dentro de cada subgrupo serão controladas outras características de responsabilidade para que se

¹ Os estudos de Ronald Dworkin propõem um mecanismo de compensação social para aqueles com uma má dotação de talentos naturais ou má sorte na loteria do nascimento.

² Para discussões sobre esse conceito ver também: Dworkin (1981), Arneson (1989).

³ Ferreira e Veloso destacam a alta persistência entre as gerações, onde o *background* familiar corresponde a cerca de 80% da influência na renda e na educação do indivíduo.

tenha um nível de esforço aproximado. O grupo composto por agentes com mais anos de estudo representa um nível de esforço maior.

Controlando para um subconjunto de covariáveis, o qual inclui raça, local de nascimento, sexo, gênero, entre outras, será possível identificar o impacto de dois diferentes níveis de educação dos pais em indivíduos com um mesmo nível de esforço. Para isso, será empregado o Modelo de Roy,⁴ difundido em trabalhos como Heckman e Honoré (1990) e Heckman e Vytlacil (1999).⁵

O Modelo de Roy permite distinguir o efeito de uma variável (tratamento) em dois grupos, logo também é possível comparar o impacto da variável de circunstância no resultado econômico dos indivíduos dentro de conjuntos específicos de variáveis de esforço. Em suma, será construída uma amostra para verificar o impacto das circunstâncias (*background* familiar) no resultado econômico dos indivíduos, controlado para dois níveis de esforço distintos, sendo este último representado pelo nível educacional. Dessa forma é possível identificar as características predominantes em dois conjuntos de indivíduos, quais sejam, com e sem nível de educação superior.

Será usada a educação do pai e da mãe (separadamente) como tratamento, dividindo assim, cada subgrupo em duas partes: uma com os indivíduos que tem pais com nível superior (grupo de tratamento) e a outra com indivíduos que tem pais que não tem nível superior (grupo de controle). Assim, a estimação da renda (salário) para os grupos contém informações dos diferenciais de oportunidades e da interação intergeracional das circunstâncias com o esforço. Também é possível estimar a distribuição dos resultados potenciais a partir de um único corte transversal de dados sobre salários e covariáveis, sabendo apenas a qual grupo as pessoas pertencem.

A identificação do modelo segue uma série de pressupostos⁶ e sua inferência é possível a partir de uma abordagem semi-paramétrica com base nos Efeitos de Tratamento Médios Locais (LATE),⁷ no qual tem sido desenvolvidos métodos para estimar efeitos distributivos de uma mudança em um instrumento sobre o tratamento (Chen e Khan, 2007). Dessa forma, será possível gerar contrafactuais com base nos dados para características específicas, a fim de mensurar o real impacto na desigualdade de oportunidades dos indivíduos. Após a identificação, a estimação será feita por métodos com variáveis instrumentais e métodos com *Pairwise-Matching*, derivado a partir de funções *Kernel* para pares simétricos com estados de tratamento opostos.

Será também utilizada uma abordagem para identificação dos efeitos do tratamento com base em análises de sensibilidade que limitam a apresentar faixas de estimativas de parâmetros de tratamento estimado. Para tanto, será aplicado os *bounds* de Shaikh e Vytlacil (2004), que exploram hipóteses estruturais não paramétricas suaves, através da utilização de variáveis instrumentais. Exige-se que a variável instrumental afete o tratamento, mas não afete diretamente o resultado. Portanto serão

⁴ Desenvolvido originalmente por Roy (1951) para verificar distribuição de rendimentos.

⁵ Ver também Heckman e Vytlacil (2005), Carneiro et al (2011), D'Haultfoeuille e Maurel (2009) entre outros.

⁶ Ver seção 2.

⁷ Do inglês *local average treatment effects*.

utilizadas as taxas de desemprego entre os indivíduos jovens no início dos anos 1980 (entre 15 e 20 anos) como instrumento para tratar a endogeneidade dos anos de estudo.⁸

Como uma motivação inicial, a Figura 1 mostra as funções de densidades acumuladas do logaritmo dos salários dos indivíduos com nível superior, divididos entre filhos de pai com nível superior e filhos de pai sem nível superior. Ainda não é analisada nenhuma outra característica que afeta o resultado. Portanto, o intuito é apenas verificar o diferencial nas distribuições tendo um esforço aproximado dentro de cada grupo de agentes. É importante salientar também que não foi realizado nenhum teste de dominância estocástica ou significância estatística que comprove o diferencial entre as curvas.

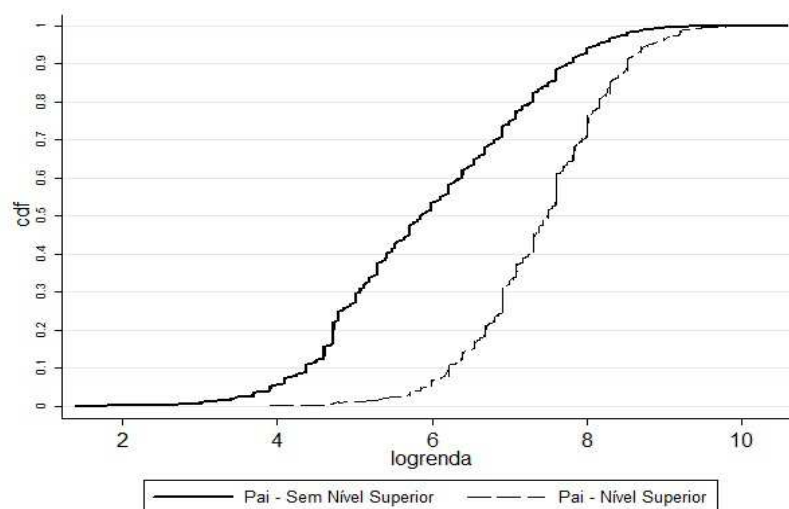


Figura 1: Conjunto de Oportunidades (indivíduos com nível superior)

Fonte: Elaboração própria a partir dos dados da PNAD

No que tange o conjunto de oportunidades, percebe-se uma acentuada disparidade entre as curvas, com a distribuição do log dos salários dos indivíduos com pais sem nível superior totalmente à esquerda. Isso indica que os indivíduos auferem rendimentos menores devido ao baixo *background* familiar e assim tem-se um resultado insatisfatório do ponto de vista social, dado que em consonância com a desigualdade de oportunidades, apenas a desigualdade devido a variáveis de não responsabilidade (também denominada de desigualdade injusta) é socialmente indesejável.

A Figura 2 abaixo apresenta as mesmas funções de densidade acumuladas para os indivíduos sem nível superior, utilizando o mesmo procedimento descrito para a Figura 1. Percebe-se também nesse caso um diferencial, porém menor entre as curvas e, portanto, as mesmas observações podem ser feitas sobre a influência da escolaridade do pai nesses indivíduos.

⁸ Na seção 3 será justificada a utilização desse instrumento.

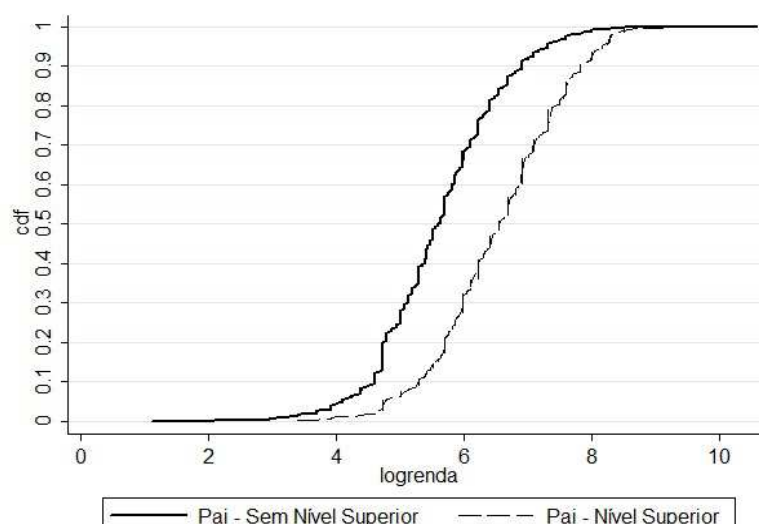


Figura 2: Conjunto de Oportunidades (indivíduos sem nível superior)

Fonte: Elaboração própria a partir dos dados da PNAD

Cabe ressaltar que outras variáveis influenciam o diferencial entre os grupos, logo alguns questionamentos são imprescindíveis para qualquer conclusão robusta. O diferencial entre as distribuições é estatisticamente significativo? Controlando para diversas características, qual o real impacto da educação dos pais (*background* familiar) no resultado econômico dos indivíduos? Como reduzir as desigualdades de oportunidades? Assim, o objetivo desse estudo tenta direcionar para respostas aceitáveis a essas perguntas.

Diante disso, o trabalho propõe uma avaliação do impacto do *background* familiar no resultado econômico sobre indivíduos que exercem um nível aproximado de esforço, incorporando as ideias de justiça social discutidas.

Para tanto, será aplicado o modelo de Roy, que divide uma amostra específica em um grupo de tratamento, que são dotados de determinada característica, e um grupo de controle, sem essa característica. Através de estimativas semi paramétricas e construção de *bounds* para o efeito de tratamento, o diferencial entre os grupos é mensurado, seguindo critérios da desigualdade de oportunidades.

2. Procedimentos Teóricos e Metodológicos

Esta seção é destinada a apresentar os procedimentos da estratégia empírica adotada no estudo. Em primeiro lugar, serão destacados os conceitos teóricos acerca da literatura de desigualdade de oportunidades ou desigualdade injusta, assim como o problema de identificação das abordagens tradicionais. Em seguida, nas subseções 2.2 e 2.3 são apresentadas, respectivamente, a forma de identificação e o Modelo de Roy Generalizado, visando tornar claros os procedimentos de análise de estudo. Os métodos de estimação referentes ao modelo adotado serão apresentados na subseção 2.4.

2.1 Desigualdade de Oportunidades: Aspectos Teóricos

A subseção analisa brevemente a desigualdade de oportunidades à luz dos desenvolvimentos recentes da literatura, assim como as dificuldades de mensuração devido a problemas de identificação dos modelos. Assim, abre-se espaço para a identificação e o emprego de uma nova abordagem do Modelo de Roy, baseado em uma variável de tratamento específica.

Para analisar a desigualdade de oportunidades é necessário identificar as variáveis pelas quais os indivíduos são responsáveis e aquelas variáveis que estão fora do seu controle, sendo possível mensurar seu conjunto de oportunidades. Em outras palavras, como já destacado anteriormente, a desigualdade de oportunidades considera o resultado do indivíduo como fruto de fatores de circunstância e esforço (Roemer, 1998).⁹

Assim, dado que os indivíduos não são responsáveis pelas suas circunstâncias, apenas a desigualdade devido a estas é considerada socialmente indesejável. Portanto, a igualdade de oportunidades segue o princípio da responsabilidade, onde os agentes devem arcar com as consequências de suas escolhas, e o princípio da compensação, onde o efeito de diferenças nas circunstâncias deve ser eliminado.¹⁰ Em suma, a igualdade de oportunidades significa corrigir as circunstâncias desiguais e manter inalteradas as diferenças de esforço.

Em consonância com a desigualdade de oportunidades, estudos têm sido desenvolvidos na tentativa de esboçar políticas ótimas de compensação (Roemer, 2003), teorias de justiça e bem-estar social (Fleurbaey e Maniquet, 2011) e métodos de mensuração da desigualdade, entre outros. Para este último, nem sempre é simples a operacionalização da desigualdade de oportunidades, dada a dificuldade verificada na definição do conceito de esforço. Borghingnon et al (2007) considera o esforço como observável, usando um modelo paramétrico no qual este depende endogenamente das circunstâncias. Fleurbaey (1998), por outro lado, considera o esforço como não observado, baseado na *Roemer's Identification Assumption* (RIA).¹¹ Em resumo, o

⁹ Outras abordagens da desigualdade de oportunidades podem ser encontradas em Ooghe et al. (2007)

¹⁰ Para uma síntese da discussão sobre normas de justiça social ver Fleurbaey e Maniquet (2011).

¹¹ A *Roemer's Identification Assumption* implica que dois indivíduos com diferentes conjuntos de oportunidades, mas com o mesmo percentil da distribuição dentro do seu tipo, exercem o mesmo nível de esforço. Para uma suposição de fator não observável usando RIA, ver Checchi e Peragine (2009).

esforço pode ser ou não considerado uma variável observável, dependendo da abordagem utilizada.

Para estimação da função de rendimentos, nem sempre é possível ter um conjunto de variáveis de circunstância e esforço bem definido. Com isso, os métodos de estimação usados podem conter um viés causado pela má identificação do modelo, assim como pela indefinição quanto ao conjunto de variáveis que correspondem ao grupo de responsabilidade e de não responsabilidade.

Na tentativa de contornar essas dificuldades, será usado o Modelo de Roy, apresentado na próxima subseção através da separação da população em um grupo de tratamento e um grupo de controle, sendo a educação do pai e da mãe, separadamente, o tratamento aplicado. É importante salientar que toda a estratégia empírica será empregada para dois subgrupos, ou seja, para o grupo dos indivíduos com e sem nível superior.

2.2 O Problema de Identificação no Modelo de Roy

Para a identificação do modelo, a presente subseção expõe os pressupostos acerca do tratamento utilizado e das condições necessárias para a utilização dos efeitos de tratamento médios locais (LATE). O tratamento é considerado como uma variável binária, e são descritas medidas de comparação do efeito do tratamento sobre a desigualdade entre o grupo tratado e não tratado.

Assume-se que o efeito do tratamento (nesse caso pais com nível superior), é representado por uma variável binária D_i , sobre o diferencial de um resultado Y_i (tal como o log dos salários) continuamente distribuído. A relação causal entre o tratamento e o resultado é caracterizada por resultados potenciais Y_{Di} que o indivíduo obteria dado o status do tratamento $D_i = D$ (Rubin, 1977). Assim, Y_{1i} é o salário potencial do agente que tem pais com nível superior, enquanto Y_{0i} é o salário potencial do agente com pais sem formação superior em cada grupo.

Os resultados potenciais podem ser determinados por um modelo de mudança endógena:

$$Y_{1i} = \mu_1(x_i) + \sigma_1(x_i)\varepsilon_{1i},$$

$$Y_{0i} = \mu_0(x_i) + \sigma_0(x_i)\varepsilon_{0i}.$$

onde x_i é um conjunto de covariáveis exógenas, $\mu_1(x_i)$ e $\mu_0(x_i)$ são as localizações condicionais, e $\sigma_1(x_i)$ e $\sigma_0(x_i)$ são as escalas condicionais. Os termos de erro normalizados ε_{1i} e ε_{0i} , têm as propriedades $E[\varepsilon_{Di} | x_i] = 0$ e $E[\varepsilon_{Di}^2 | x_i] = 1$, para $D = 0, 1$.

Dito isso, o parâmetro de interesse é o efeito do tratamento sobre a desigualdade nos grupos, mensurado pela razão de escalas condicionais, definida como

$r(x) = \sigma_1(x) / \sigma_0(x)$, dadas pelas variáveis x ou pela razão de escala média, sendo $r = E_x[\sigma_1(x) / \sigma_0(x)]$. Podemos também analisar as medidas de desigualdades baseadas na variância, identificando a razão de variância condicional como $\sigma_1^2(x) / \sigma_0^2(x)$ dadas as variáveis x e a razão da variância média $E_x[\sigma_1^2(x) / \sigma_0^2(x)]$.

Contudo, a identificação das razões de escala ou de variância não é simples, pois exige o conhecimento do resultado potencial para todos os status de tratamento, sendo que ambos os status não podem ser observados para um mesmo indivíduo. Os métodos convencionais também podem superestimar o efeito causal do tratamento sobre a desigualdade entre os grupos. Diante disso, será apresentado o estimador não paramétrico na subseção de inferência, assim como a construção de contrafactuais para solucionar o problema de identificação.

No entanto, as condições LATE podem ser usadas para identificar toda a distribuição marginal ou potencial para os grupos e para avaliar o impacto tanto da média quanto da dispersão nos resultados potenciais (Imbens e Rubin, 1997). Para identificação, os métodos LATE usam uma variável instrumental exógena Z_i condicionada às covariáveis, para induzir variação no status de tratamento. A dependência entre o instrumento e o status de tratamento é reconhecida pelo indicador de tratamento potencial D_{zi} . O instrumento pode ser uma variável binária ou uma variável continuamente distribuída.

Indivíduos com $D_{0i} = 0$ e $D_{1i} = 1$, são chamados *compliers*.¹² Chen (2008) usa como instrumento o custo da matrícula no município de residência dos alunos, sendo continuamente distribuído e exógeno a X_i . Na presente discussão dentro do âmbito LATE, será usado instrumento binário, sendo apresentado na seção 3.

Segundo os pressupostos LATE, o efeito do tratamento para *compliers* pode ser identificado apenas pela parte da variabilidade do tratamento que não é correlacionada com variáveis omitidas. Assim, o diferencial de resultados potenciais, dado a variável de tratamento binária, pode ser escrito como:

$$E[Y_{1i} - Y_{0i} | D_{1i} > D_{0i}, x_i] = \frac{E[Y_i | z_i = 1, x_i] - E[Y_i | z_i = 0, x_i]}{E[D_i | z_i = 1, x_i] - E[D_i | z_i = 0, x_i]}, \quad (1)$$

onde Y_i denota o resultado observado, e deixando Y_{1i} e Y_{0i} denotar os resultados potenciais.¹³

2.3 O Modelo de Roy Generalizado Não Paramétrico

Dado a identificação, será apresentado nesta subseção o modelo generalizado a ser testado empiricamente, alicerçado em pressupostos sobre as variáveis explicativas e sobre seus resultados. O objetivo é identificar e posteriormente mensurar o benefício

¹² São os agrupamentos de acordo com o status de tratamento

¹³ Para outras regras de identificação ver Imbens and Rubin (1997) e Abadie (2002, 2003).

marginal do tratamento, dado um conjunto de variáveis explicativas observadas e não observadas.

Sabendo que há dois resultados potenciais (Y_0, Y_1) , e uma variável binária de escolha D , onde $D=1$ se o indivíduo está no grupo de tratamento obtendo o resultado Y_1 e $D=0$ se o indivíduo não está no grupo de tratamento obtendo Y_0 , de modo que os resultados são observados. Assim, o resultado observado Y pode ser escrito pela regressão abaixo:

$$Y = DY_1 + (1-D)Y_0. \quad (2)$$

Para tanto, a estrutura do resultado é representada por

$$Y_i = \mu_i(X) + U_i, \quad (3)$$

onde X representa um vetor de variáveis observadas enquanto que U_i são fatores não observados, para $i = 0, 1$. Substituindo a equação (3) na equação (2), tem-se,

$$Y = \mu_0(X) + \{[\mu_1(X) - \mu_0(X)] + U_1 - U_0\}D + U_0. \quad (4)$$

O benefício marginal do tratamento é o efeito causal associado à mudança de uma pessoa idêntica, *ceteris paribus*, do estado “0” para o estado “1”. Portanto o benefício marginal é caracterizado pela diferença entre os resultados observados nos dois estados.

$$\beta = Y_1 - Y_0$$

$$\beta = [\mu_1(X) - \mu_0(X)] + (U_1 - U_0). \quad (5)$$

Para tanto, não é assumida nenhuma forma para as funções μ_1 e μ_0 , e a distribuição de U_0 e U_1 é desconhecida. No modelo de Roy original (1951) o custo do tratamento é zero e o vetor de variáveis não observadas segue uma normal com média zero e variância constante, $(U_0, U_1) \sim N(0, \sigma)$.

Assume-se então, que não há custos de tratamento e aplica-se o modelo desenvolvido por Heckman e Honoré (1990), com X regressores observados e hipóteses não paramétricas sobre a distribuição dos regressores não observados (U_0, U_1) , ao contrário da hipótese de normalidade do modelo original para U_i . A suposição de custo de tratamento zero é plausível, dado que o tratamento está fora do controle dos indivíduos. Em outras palavras, o tratamento depende do mercado de trabalho, estando fora da responsabilidade dos agentes. Outra justificativa é que qualquer custo já está implicitamente incorporado no benefício subjetivo do tratamento.

Além destas, são utilizadas as seguintes hipóteses no modelo:

- 1) (U_0, U_1) são independentes de X . Essa suposição assume que D é endógeno, mas os outros regressores da equação de tratamento e da equação de resultados são exógenos. Essa condição deve ser interpretada como a hipótese de independência dos termos de erro condicionados a X .
- 2) Os valores de $E|Y_1|$, $E|Y_0|$ são finitos. A hipótese satisfaz as condições de integração padrão e garante que o benefício médio seja bem definido.
- 3) $0 < \Pr(D=1|X) < 1$. É o pressuposto de um grupo de tratamento e um grupo de controle na população para X .

Dadas essas suposições, o modelo de Roy utilizado no presente trabalho é um caso particular do modelo de Heckman e Vytlacil (2005), que junto com os resultados de Vytlacil (2002) implica no Efeito de Tratamento Médio Local. No entanto, é imposta uma estrutura de separabilidade aditiva na equação de resultados (equação 3), onde X é independente do vetor de erros.¹⁴

2.4 Inferência

O processo de estimação do tratamento referente à identificação detalhada na seção anterior, tendo como parâmetros de interesse a razão de escalas condicionais e média descritas. Pode também ser usada a razão das variâncias condicionais e a razão da variância média para mensurar o efeito do tratamento baseado na variância dos grupos ou ainda a diferença entre os resultados potenciais. É proposto dois estimadores a seguir: o estimador de variáveis instrumentais¹⁵ baseados na variância e o estimador de correspondência baseado na variância, podendo ser estimados por vários métodos.

2.4.1 Estimador de Variáveis Instrumentais Baseados na Variância

Primeiramente é feita uma estimativa da variância nos resultados potenciais generalizados da equação (1) para os grupos. Abadie (2002) generaliza o diferencial de resultados da equação (1) para identificar a média de funções arbitrárias de resultados potenciais $h(Y_{1i})$ e $h(Y_{0i})$, explorando o fato de que o efeito de tratamento médio local em $h(Y_{Di})D_i$ é igual à média de $h(Y_{1i})$ para *compliers*.

Para tanto, a estimativa pode ser feita por métodos de *Kernel* para cada valor da covariável x . Assim, é construída uma função *Kernel* $\phi(.)$ para mensurar a distância entre dois pontos (i, j)

$$\phi_{ij} \equiv \phi[(x_i - x_j) / h_n] / h_n, \quad (6)$$

¹⁴ Ver outra análise a esse respeito em Vytlacil e Yildiz (2007).

¹⁵ estimador IV (Do inglês *instrumental variable*).

em que os dois pontos tem pesos maiores quanto mais próximos estiverem.

Em seguida são utilizados os pesos de *Kernel* para calcular o primeiro e o segundo momento condicional em cada valor de x , utilizando a generalização da equação (1) para o grupo de tratamento e o grupo de controle.

$$\begin{aligned}\hat{E}[h(Y_{1i}) | D_{1i} > D_{0i}, x_i] &= \frac{\hat{E}[h(Y_i)D_i | Z_i = 1, x_i] - \hat{E}[h(Y_i)D_i | Z_i = 0, x_i]}{\hat{E}[D_i | Z_i = 1, x_i] - \hat{E}[D_i | Z_i = 0, x_i]} \\ &= \frac{\sum_j \phi_{ij} h(Y_j) D_i [Z_j - \bar{Z}_{ij}]}{\sum_j \phi_{ij} D_i [Z_j - \bar{Z}_{ij}]}. \quad (7)\end{aligned}$$

$$\hat{E}[h(Y_{0i}) | D_{1i} > D_{0i}, x_i] = \frac{\sum_j \phi_{ij} h(Y_j) (1 - D_i) [Z_j - \bar{Z}_{ij}]}{\sum_j \phi_{ij} (1 - D_i) [Z_j - \bar{Z}_{ij}]}.$$

onde $\bar{Z}_j \equiv \sum_i \phi_{ij} Z_i / \sum_i \phi_{ij}$ para cada j e $h(Y) = Y$ ou Y^2 é o primeiro ou o segundo momento. Assim, $\hat{V}[Y_{Di} | D_{1i} > D_{0i}, x_i]$ é derivado.

Portanto, o estimador IV para a razão da variância condicional dado um determinado x , pode ser escrito:

$$\hat{r}_{IV}(x) = \frac{\hat{V}[Y_{1i} | D_{1i} > D_{0i}, x_i = x]}{\hat{V}[Y_{0i} | D_{1i} > D_{0i}, x_i = x]}, \quad (8)$$

onde é calculada para uma covariável específica x . Assim a equação (9) mostra a razão média de variação obtida a partir da média com todas as covariáveis.

$$\hat{r}_{IV} = \frac{1}{n} \sum_i \hat{r}_{IV}(X_i). \quad (9)$$

2.4.2 Métodos *Pairwise-Matching*

O estimador de correspondência baseado na variância usa *scores* de propensão estimados $\hat{p}(x, z)$ para procurar e atribuir pesos aos pares simétricos (i, j) com status de tratamento opostos, dados os valores de x e z .

O ideal é encontrar pares para os quais a soma das probabilidades dentro de cada par é igual a um. No entanto, dado que essas estimativas são difíceis de serem encontradas, é construída uma função *Kernel* que atribui um maior peso aos pares em que a soma das probabilidades estão mais próximas de um.

Para tanto, o peso de *Kernel* para cada par (i, j) com $(D_i, D_j) = (1, 0)$ é dado pela equação (10), onde K_{1n} é uma função *Kernel*.

$$\hat{k}_{ij} = K_{1n}(1 - \hat{p}(x_i, z_i) - \hat{p}(x_j, z_j)). \quad (10)$$

Os *scores* de propensão para um indivíduo que recebe tratamento $\hat{p}(x_i, z_i)$, e para um que não recebe tratamento $\hat{p}(x_j, z_j)$, são obtidos de forma não paramétrica, sendo ϕ_{ij} definido anteriormente na equação (6).

$$\hat{p}(x_i, z_i) = \sum_j \phi_{ij} D_j / \sum_j \phi_{ij},$$

$$\hat{p}(x_j, z_j) = \sum_i \phi_{ij} (1 - D_i) / \sum_i \phi_{ij}.$$

Em seguida, calcula-se a razão de escala para cada par, dadas os valores das covariáveis $x_i = x_j = x$. Assim, para o estimador baseado na variância é calculado como $\hat{V}(Y_i | D_i = 1, x_i = x) / \hat{V}(Y_j | D_j = 0, x_j = x)$.

Por fim, o estimador da razão de escala é a média ponderada das razões das variâncias condicionais ao longo de pares simétricos ponderados pelas estimativas de $\hat{k}_{ij}$ da equação (10), sabendo os valores de x . O estimador da razão baseado na variância é dado por:

$$\hat{r}_v(x) = \frac{\sum_{\{(i,j): x_i = x_j = x\}} D_i (1 - D_j) \hat{k}_{ij} [\hat{V}(Y_i | D_i = 1, x_i) / \hat{V}(Y_j | D_j = 0, x_j)]}{\sum_{\{(i,j): x_i = x_j = x\}} D_i (1 - D_j) \hat{k}_{ij}}. \quad (11)$$

Estes procedimentos com várias etapas semi paramétricas tem a desvantagem de necessitar da seleção de parâmetros de alisamento, o que dificulta a implementação em amostras finitas. Isso é uma consequência da generalização do modelo proposto, não assumindo funções de localização específicas. A implementação de algumas condições paramétricas, pode permitir a implementação de forma mais simples em amostras finitas.¹⁶

No entanto, se o parâmetro a ser estimado depender apenas do tratamento, e não de todas as covariáveis x , as funções de variância condicional dependerão apenas dos *scores* de propensão.

É importante salientar, que os estimadores propostos envolvem a seleção de *bandwidths*. Será usada a regra de Silverman¹⁷, que proporciona uma fórmula substituindo os parâmetros desconhecidos por *bandwidths* ótimos, com sua amostra analógica.

¹⁶ Para exemplos de imposição de formas específicas para as funções, ver Powell, Stock e Stoker (1989).

¹⁷ Ver Silverman (1986) e Jones, Marron e Sheather (1996).

3. Bounds sobre o Efeito de Tratamento Médio

A estimação por pares de correspondência pode não apresentar resultado confiável se a amostra for muito heterogênea e com isso o real impacto do tratamento pode não ser verificado. Nesta seção serão apresentados os *bounds* de Shaikh e Vytlacil (2004) com a utilização de variáveis instrumentais, assim como a justificativa para a utilização de tal variável.

3.1 Bounds de Shaikh e Vytlacil

Os pressupostos de Shaikh e Vytlacil (2004), mantendo-se a natureza não paramétrica, são mais fortes do que as impostas por Manski (1990)¹⁸.

Lembre-se que Z é solicitado para que $\Pr\{D=1|Z=1\} > \Pr\{D=1|Z=0\}$, e dadas às suposições do Modelo de Roy, implica em um resultado potencial maior para o grupo tratado. Assim, se $r(1) > r(0)$, sendo r o parâmetro que distingue o status de tratamento D .

$$\begin{aligned}\Pr\{Y=1|Z=1\} &> \Pr\{Y=1|Z=0\} \Leftrightarrow r(1) > r(0) \\ \Pr\{Y=1|Z=1\} &< \Pr\{Y=1|Z=0\} \Leftrightarrow r(1) < r(0).\end{aligned}$$

Portanto, os *bounds* resultantes sobre o efeito do tratamento médio são dados pela equação (12):

$$B_{SV}^L \leq E[Y_1 - Y_0] \leq B_{SV}^U, \quad (12)$$

onde:

$$B_{SV}^L = \Pr\{Y=1|Z=1\} - \Pr\{Y=1|Z=0\}$$

$$B_{SV}^U = \Pr\{D=1, Y=1|Z=1\} + \Pr\{D=0|Z=1\} - \Pr\{D=0, Y=1|Z=0\}$$

quando $\Pr\{Y=1|Z=1\} > \Pr\{Y=1|Z=0\}$

$$B_{SV}^L = \Pr\{D=1, Y=1|Z=1\} - \Pr\{D=0, Y=1|Z=0\} - \Pr\{D=1|Z=0\}$$

$$B_{SV}^U = \Pr\{Y=1|Z=1\} - \Pr\{Y=1|Z=0\}$$

quando $\Pr\{Y=1|Z=1\} < \Pr\{Y=1|Z=0\}$

$$\text{e } B_{SV}^L = B_{SV}^U = 0$$

quando $\Pr\{Y=1|Z=1\} = \Pr\{Y=1|Z=0\}$

Os *bounds* de Shaikh e Vytlacil sempre identificam o sinal do efeito médio do tratamento. Esses *bounds* são menores do que aqueles de Manski (1990). O trade-off entre as análises de Shaikh e Vytlacil (2004) e Manski e Pepper (2000) é que este

¹⁸ Ver Manski (1990)

último requer que se saiba a priori se $Y1 \geq Y0$ ou $Y1 \leq Y0$, enquanto o primeiro exige impor uma estrutura para ser capaz de determinar o sinal do efeito de tratamento médio da distribuição dos dados observados.

3.2 Variável Instrumental

Uma variável deve atender a dois requisitos essenciais para validade de um instrumento. Primeiro deve esta fortemente correlacionada com a aplicação do tratamento e em segundo lugar, dentro de classes de risco observadas, é não correlacionado com os resultados.

Portanto serão utilizadas as taxas de desemprego entre os indivíduos jovens (entre 15 e 20 anos) como instrumento para endogeneizar os anos de estudo, tal como Arkes (2010). A explicação para a utilização dessa variável como instrumento é o fato de que a taxa de desemprego afeta a acumulação de capital humano.

A relação da taxa de desemprego com a acumulação de capital humano se da através da interação dos efeitos renda e substituição: i) o efeito renda: a elevada taxa de desemprego cria à necessidade de complementar a renda que diminui e com isso as crianças são mais propensas a entrarem no mercado de trabalho; ii) efeito substituição: há incentivo para qualificação, dado que o custo de oportunidade de investir em capital humano é reduzido.

Dessa forma, uma elevada taxa de desemprego resulta em uma maior demanda por investimentos em qualificação para os mais jovens, pois, segundo Arkes (2010) o efeito substituição anula o efeito renda.

3.3 Inferência

Serão construídos alguns conjuntos aleatórios C_n , tal que, para cada parâmetro θ entre os *bounds* superior e inferior, temos o a região de confiança para parâmetros identificáveis.

$$\liminf \Pr\{\theta \in C_n\} \geq 1 - \alpha.$$

Assim, para cada $-1 \leq \theta \leq 1$, será construído um teste com a hipótese nula de que θ esta entre os *bounds* inferior e superior. Portanto, a região C_n é definida para o conjunto de valores de θ para os quais não consegue-se rejeitar a hipótese nula. Concretamente, define

$$C_n = \{-1 \leq \theta \leq 1 : T_n(\theta) \leq \hat{c}_n(\theta, 1 - \alpha)\},$$

onde $T_n(\theta)$ é a estatística de teste que fornece evidencias contra a hipótese nula e $\hat{c}_n(\theta, 1 - \alpha)$ é o valor crítico apropriado, que será construído usando subamostragem.

Denotando por $T_{n,b,i}(\theta)$ o teste estatístico $T_n(\theta)$, calculado usando apenas o i -ésimo conjunto de dados de tamanho b . Deixe $\hat{c}_n(\theta, 1-\alpha)$ denotar o menor $1-\alpha$ quantil da distribuição.

3.3.1 Bounds de Shaikh e Vytlačil

Considere Δ_n de acordo com a definição abaixo:

$$\Delta_n = \left\{ \frac{1}{n_1} \sum_{1 \leq i \leq n: Z_i=1} Y_i - \frac{1}{n_0} \sum_{1 \leq i \leq n: Z_i=0} Y_i \right\}.$$

Dado o desvio padrão das variáveis por s_n , a estatística de teste é apresentada nas expressões abaixo de acordo com o valor de θ . Assim, para $0 < \theta \leq 1$, $-1 \leq \theta < 0$ e $\theta = 0$, tem-se respectivamente as equações (13), (14) e (15):

$$T_n(\theta) = \left(\frac{-\Delta_n}{s_n} \right)_+^2 + \left(\frac{\Delta_n - \theta}{s_n} \right)_+^2 + \left(\frac{\theta - \delta_{2n}(1, 0)}{s_{2n}(1, 0)} \right)_+^2; \quad (13)$$

$$T_n(\theta) = \left(\frac{\Delta_n}{s_n} \right)_+^2 + \left(\frac{\theta - \Delta_n}{s_n} \right)_+^2 + \left(\frac{\delta_{1n}(1, 0) - \theta}{s_{1n}(1, 0)} \right)_+^2; \quad (14)$$

$$T_n(\theta) = \left(\frac{\Delta_n}{s_n} \right)_+^2. \quad (15)$$

4. Dados

Os dados utilizados no estudo serão fornecidos pela Pesquisa Nacional por Amostra de Domicílios (PNAD) para o ano de 1996. Será utilizado esse ano por conta do suplemento de mobilidade social, onde são sumarizadas as informações relacionadas ao *background* familiar, i.e., educação, escolaridade e ocupação dos pais dos indivíduos selecionados. Os resultados encontrados com base nesse ano não serão ultrapassados, dadas as hipóteses assumidas no modelo e também ao fato que há um repasse intergeracional de circunstâncias, tornando-as imutáveis por um período considerável às conclusões obtidas.

O vetor de variáveis de circunstâncias inclui a posição ocupacional do pai proposta por Pastore e Silva (1999);¹⁹ uma *dummy* para as regiões, que assume valor 1 se o indivíduo reside nas Regiões mais dinâmicas do país: Sul, Sudeste e Centro-Oeste; e uma *dummy* para raça, que assume valor 1 para os não-brancos (i.e. negros, pardos e indígenas). Além destas, a educação do pai (da mãe) é o fator de não responsabilidade usado como tratamento, assumindo valor 1 se o pai (mãe) tem nível superior e 0 caso contrário.

Como *proxy* de esforço individual será utilizado indivíduos com nível superior e indivíduos sem nível superior separados em dois subgrupos, na tentativa de homogeneizar amostras de indivíduos com um nível aproximado de esforço, para um determinado grupo de circunstâncias. Por fim, os rendimentos serão representados pelo salário por hora²⁰ (em logaritmo), expressos em reais daquele ano.

Por último, a taxa de desemprego será coletada para o período de junho de 1980 a dezembro de 1984, quando os indivíduos eram jovens, e assim seus pais decidiam sobre a sua qualificação.

Serão utilizadas duas amostras a partir dos dados disponíveis. A Amostra I reúne indivíduos entre 15 e 60 anos e será usada para estimar o efeito de tratamento médio local e os *bounds* sobre o efeito de tratamento médio, enquanto a Amostra II, com menos observações, será usada para a estimação com a variável instrumental desemprego, dado que a Pesquisa Mensal de Emprego (PME) do Instituto Brasileiro de Geografia e Estatística (IBGE) refere-se às taxas mensais de desemprego aberto dessazonalizadas apenas para as seis principais regiões metropolitanas brasileiras – Belo Horizonte, Porto Alegre, Recife, Rio de Janeiro, Salvador e São Paulo.

A Amostra II ainda considera apenas indivíduos com idade entre 24 e 35 anos residentes nas áreas metropolitanas onde o desemprego é calculado. Isto porque tendo a PNAD de 1996 como base das estimações, estes indivíduos tinham entre 15 e 20 anos no início dos anos 1980, onde decidiram (influenciados pelos pais) suas escolhas a respeito de qualificação e mercado de trabalho. Adicionalmente, são considerados apenas indivíduos não migrantes, supondo que estes viveram na mesma área metropolitana em sua juventude.

¹⁹ Categorias: 1) Baixo Inferior: Lavrador, pescador, lenhador, etc; 2) Baixo Superior: braçal, servente, vigias, etc.; 3) Médio Inferior: pedreiro, eletricitista, marceneiro, etc.; 4) Médio Médio: praticistas, viajantes, chefes de nível baixo, etc.; 5) Médio Superior: assessores, gerentes de alto nível, diretores, etc. e; 6) Alto: grandes proprietários, magistrados, nível superior, etc.

²⁰ Renda de todos os trabalhos dividida pelo número de horas trabalhadas.

As estatísticas descritivas estão sumarizadas na Tabela 1. Pode-se observar uma redução da renda média e na média de idade na Amostra II em relação à Amostra I, seguidos por um aumento nos anos de escolaridade médio, salientando as restrições feitas em tal amostra. As informações apresentadas são restritas apenas as principais variáveis utilizadas nas estimações.

Tabela 1 - Estatísticas Descritivas: PNAD de 1996

	Amostra I		Amostra II	
	Média	Desvio Padrão	Média	Desvio Padrão
Renda	645,09	1048,89	534,49	734,15
Idade	38,26	9,78	30,08	3,72
Anos de estudo	6,70	4,65	7,87	4,06
Educação*	14%	-	14%	-
Escolaridade do Pai*	3%	-	4%	-
Escolaridade da Mãe*	2%	-	2%	-
Raça**	42%	-	40%	-
Gênero***	62%	-	60%	-
Migração	60%	-	47%	-
Nº de Observações	52.731		20.333	

* indivíduos com nível superior

**indivíduos não brancos

***indivíduos do sexo masculino

Fonte: Elaboração própria a partir dos dados

De um modo geral, além da escolaridade dos pais, raça e gênero apresentados na Tabela, a ocupação do pai e o local de nascimento completam o conjunto de variáveis de não responsabilidade. Portanto, esse é o vetor de variáveis de circunstâncias que impactam no resultado econômico e educacional dos indivíduos usado nas estimações.

5. Resultados

Esta seção apresenta os resultados e discussões deste estudo, dada à estratégia empírica adotada. Primeiro serão mostrados, apenas como uma abordagem inicial, os resultados para o efeito de tratamento médio convencional, calculado por diferentes métodos de estimação. Na seção 5.2 serão apresentados os resultados estimados através da utilização dos *bounds* sobre o efeito de tratamento médio utilizando a Amostra I, enquanto que a seção 5.3 faz uma estimativa com a Amostra II restrita, incluindo a variável instrumental desemprego, disponível apenas para as seis principais regiões metropolitanas brasileiras como descritas na subseção de dados.

5.1 Efeito de Tratamento Médio

Esta subseção é conduzida pela estimação do efeito de tratamento médio através da utilização de diferentes métodos, apresentados como uma abordagem inicial que antecede aos principais resultados do presente estudo. A Tabela 2 mostra os resultados da equação de renda dos indivíduos usando como variável dependente o salário por hora individual, em logaritmo. A variável explicativa é uma *dummy* para a educação do pai e da mãe, respectivamente, dividida entre os que possuem e os que não possuem nível superior.

Tabela 2: Equação de Renda dos Indivíduos (Log do salário por hora)

Estimação por MQO – Diferença de Médias Simples				
	Educação do Pai		Educação da Mãe	
	Coefficiente	Constante	Coefficiente	Constante
Educação dos pais	1,4396*	2,2391*	1,2789*	2,2639*
	(0,0226)	(0,0041)	(0,0325)	(0,0044)
R ²	0,05		0,02	
Nº de observações	59.805		59.805	

Nota: *p-valor < 0,01.

Desvio padrão entre parênteses

De acordo com as estimativas por Mínimos Quadrados Ordinários, a diferença de médias entre o grupo de pai com nível superior e a média dos indivíduos com pai menos instruídos é de 1,43. Isso significa que ter pai mais educado, *ceteris paribus*, impacta quase 1,5 vezes a mais no resultado econômico do indivíduo quando comparado ao filho de pai sem nível superior. Usando como tratamento a escolaridade da mãe, o resultado é menor, porém a diferença entre os grupos ainda é bastante elevada. Os resultados foram obtidos usando o *bootstrap* com 1.000 replicações.

Entretanto, dado que não foram usadas outras variáveis de controle nesta primeira abordagem, os resultados não podem ser utilizados como os verdadeiros efeitos médios da escolaridade dos pais, servindo apenas como apresentação do tratamento a ser utilizado.

Dito isso, a Tabela 3 apresenta o Efeito de Tratamento Médio a partir de três diferentes métodos de estimação, sendo a escolaridade do pai o tratamento utilizado.

Para cada modelo foi estimado o efeito do tratamento para o grupo “Baixo” que inclui somente os indivíduos sem o nível de instrução superior e “Alto” apenas com indivíduos que possuem ao menos nível superior, controlado para o conjunto de covariáveis descritas na seção de dados. A Tabela 4 apresenta as mesmas informações, porém o Efeito de Tratamento é a escolaridade da mãe.

Tabela 3: Efeito de Tratamento Médio – Educação do Pai (Tratamento)

	Rendimento Individual (Log do Salário por Hora)					
	MQO-cf		Probit		Logit	
	Baixo	Alto	Baixo	Alto	Baixo	Alto
ETM	0,6613*	0,3285*	0,4966*	0,3490*	0,4892*	0,3385*
	(0,0913)	(0,3896)	(0,0829)	(0,0431)	(0,0892)	(0,0457)
Nº de observações	24.943	3.787	28.187	4.172	28.187	4.172

Nota: *p-valor < 0,01.

Desvio padrão entre parênteses

Tabela 4: Efeito de Tratamento Médio – Educação da Mãe (Tratamento)

	Rendimento Individual (Log do Salário por Hora)					
	MQO-cf		Probit		Logit	
	Baixo	Alto	Baixo	Alto	Baixo	Alto
ETM	0,6122*	0,2015**	0,4723*	0,2240**	0,4484*	0,1988**
	(0,1267)	(0,0795)	(0,1615)	(0,9801)	(0,1591)	(0,0887)
Nº de observações	24.822	3.777	29.278	4.205	29.278	4.205

Nota: *p-valor < 0,01. **p-valor < 0,05.

Desvio padrão entre parênteses

O primeiro modelo é o controle da função estimado por Mínimos Quadrados Ordinários (MQO-cf), introduzido por Rosenbaum e Rubin (1983) onde é adotada a suposição que os resultados potenciais condicionados as covariáveis são independentes do tratamento. Os modelos Probit e Logit foram estimados usando a reponderação do *propensity score*, usado quando o tratamento não é atribuído aleatoriamente.

Em todos os casos o efeito de tratamento apresentou sinal positivo, mas a magnitude reduziu consideravelmente em comparação com os resultados da diferença de médias simples reportados na Tabela 2. Isso indica que o controle do tratamento por um conjunto de covariáveis torna a amostra mais homogênea e pode eliminar superestimações de tais efeitos. Observa-se também que o efeito do tratamento é maior nos indivíduos com menor grau de instrução (grupo “Baixo”), dado que estes são geralmente grupos com menor nível de renda e, portanto mais suscetíveis ao *background* familiar dos pais.

Em suma, os resultados sugerem que pais com nível superior tem um impacto maior no rendimento dos filhos, sobretudo naqueles indivíduos sem nível superior, como mostra os resultados mais elevados para o grupo “Baixo”, chegando a 0,66 quando a educação do pai é usada como tratamento e 0,61 para a educação da mãe.

As próximas subseções mostram os resultados mais específicos do estudo, tendo assim um maior grau de robustez nas estimativas. No entanto, já é possível observar que o *background* familiar e as circunstâncias de um modo geral exercem um impacto considerável no esforço e, por conseguinte no resultado econômico dos indivíduos.

5.2 Estimação de Manski *Bounds*

Dada à dificuldade de se obter uma estimativa pontual precisa acerca do efeito de tratamento médio, optou-se também pela estimação incluindo *bounds* de variação, estabelecendo limites inferior e superior onde possivelmente o efeito está contido. Usando como tratamento o nível educacional dos pais, a Tabela 5 mostra o efeito de tratamento médio estimado acompanhado dos limites de variação de tal efeito, com 1000 replicações via *bootstrap*.

Tabela 5: Estimação de Bounds - Educação do Pai (Tratamento)

	Efeito de Tratamento Médio			
	Baixo		Alto	
	LBound	UBound	LBound	UBound
ETM	0,3340	0,6490	0,2541	0,4236
R ²	0,18		0,15	
Nº de observações	46.627		7.252	

Nota: Para cada grupo é apresentado as estimativas do valor mínimo (LBound) e máximo (UBound) com um intervalo de confiança de 95%.

Os resultados mostram uma menor variação entre os limites para o grupo de indivíduos com nível superior (“Alto”), assim como valores menores para tais limites em comparação ao grupo (“Baixo”), associado a um menor R². Isso corrobora os resultados da seção anterior ao mostrar que o impacto do tratamento é maior no grupo de indivíduos sem nível superior, tendo assim um maior repasse intergeracional devido à educação dos pais.

Em suma, o impacto da educação do pai é maior nos indivíduos que não possuem nível superior, porém esse impacto é associado a um maior limite de variação, o que faz desse grupo, possivelmente, mais dependente do seu *background* familiar. A Tabela 6 apresenta o mesmo raciocínio da Tabela 5, tendo como tratamento a educação da mãe.

Tabela 6: Estimação de Bounds – Educação da Mãe (Tratamento)

	Efeito de Tratamento Médio			
	Baixo		Alto	
	LBound	UBound	LBound	UBound
ETM	0,1466	0,7703	0,0348	0,3861
R ²	0,18		0,12	
Nº de observações	48.429		7.302	

Nota: Para cada grupo é apresentado as estimativas do valor mínimo (LBound) e máximo (UBound) com um intervalo de confiança de 95%.

Percebe-se na Tabela 6 a mesma situação apresentada no caso anterior, com o grupo menos escolarizado tendo limites maiores em termos de valores mínimo e máximo assim como maior variação entre estes. Porém quando se comparam os dois tratamentos utilizados (Tabela 5 com a Tabela 6), as estimativas dos *bounds* tendo a Educação da Mãe como tratamento apresenta uma variação superior à educação do pai. Isso denota que os indivíduos, em média, podem ser mais suscetíveis a serem impactados pelo nível de escolaridade da mãe.

5.3 Estimação com Variável Instrumental

A aplicação das estimações nesta subseção é possível através da utilização da Amostra II, dadas às restrições feitas aos dados. Tais restrições são devido à utilização da taxa de desemprego como variável instrumental, disponível a partir da década de 1980 apenas para as seis principais áreas metropolitanas brasileiras. A taxa de desemprego média para cada área é considerada de junho de 1980 a dezembro de 1984.

Em suma, a amostra utilizada nesta estimativa consiste em indivíduos de 24 a 35 anos de idade, não migrantes e residentes nas áreas metropolitanas de Belo Horizonte, Porto Alegre, Recife, Rio de Janeiro, Salvador e São Paulo.

Tabela 7: Efeito de Tratamento Médio – Variável Instrumental (Taxa de Desemprego)

	Tratamento			
	Educação do Pai		Educação da Mãe	
	Baixo	Alto	Baixo	Alto
ETM	0,4484* (0,0799)	0,8297 (0,7601)	0,4585*** (0,2587)	1,9957 (35,8654)
R ²	0,18	0,08	0,17	0,02
F (p-valor)	215 (0,00)	13 (0,00)	90 (0,00)	5 (0,00)
Nº de observações	10.398	1.736	10.834	1.743

Nota: *p -valor < 0,01. **p-valor < 0,05. ***p-valor < 0,10

Desvio padrão entre parênteses

As estimativas da Tabela 7 para o efeito de tratamento médio com variável instrumental seguem Manski e Pepper (2000) e Cerulli (2012). Não é possível fazer comparações diretas com as duas subseções anteriores devido à diferenciação no banco de dados. No entanto, diferentemente dos casos anteriores, verifica-se um efeito maior quando o tratamento usado é a escolaridade da mãe nos dois grupos de indivíduos, sendo os coeficientes do grupo com alto grau de instrução, não significativos.

Observando apenas os indivíduos sem nível superior (grupo “Baixo”), cujos coeficientes foram significativos, percebe-se que a endogeneidade tende a superestimar o efeito de tratamento médio. Em outras palavras, o retorno da educação dos pais no resultado econômico dos filhos tende a ser superestimado quando a variável instrumental não foi considerada, sobretudo a educação do pai em detrimento da

educação da mãe. Por fim, o teste de Wald foi realizado e indica a taxa de desemprego como instrumento válido para o retorno da educação dos pais.

6. Considerações Finais

Este trabalho analisou o impacto das circunstâncias no resultado econômico dos indivíduos para níveis específicos de esforço, utilizando dados da PNAD de 1996 devido às informações sobre a educação e ocupação dos pais (*background* familiar) dos mesmos. Os resultados foram baseados na estimação dos Efeitos de Tratamento Médio, utilizando como variável de tratamento a educação do Pai e da Mãe, condicionada a um vetor de covariáveis de circunstâncias, e como variável de resultado o salário por hora dos indivíduos (em logaritmo). Os indivíduos foram divididos entre os que têm e os que não têm nível superior.

A mensuração dos resultados seguiu três abordagens de estimação: o efeito de tratamento convencional através de diferentes métodos de estimação como uma apreciação inicial; a estimação dos *bounds* de variação do efeito de tratamento dado que as estimações pontuais muitas vezes não são confiáveis; e a estimação através da utilização da taxa de desemprego como variável instrumental, como forma de endogeneizar os anos de estudo dos indivíduos.

No geral, as estimativas mostram que o impacto da educação dos pais é maior no grupo de indivíduos sem nível superior (“Baixo”), corroborando a ideia de que por serem menos escolarizados, provavelmente tem um nível de rendimento médio menor e assim são mais suscetíveis a serem impactados pelo seu *background* familiar. Em outras palavras, estes indivíduos são mais influenciados e de certa forma, mais dependentes do nível educacional dos pais, sobretudo no momento em que eram jovens quando decidiram sobre educação ou mercado de trabalho.

Com relação ao tratamento utilizado, a educação do pai apresentou um impacto maior no rendimento dos indivíduos para a maioria dos modelos estimados. No entanto, quando é utilizada a taxa de desemprego como instrumento, a escolaridade da mãe apresenta um coeficiente maior, enfatizando a influencia cada vez maior sobre os indivíduos. São observados ainda coeficientes menores para o efeito de tratamento, constatando a superestimação destes quando não se considera a variável instrumental e dando maior robustez aos resultados.

Em suma, as estimativas evidenciam a importância da educação dos pais na escolha educacional e no rendimento dos indivíduos, evidenciando assim o repasse intergeracional de educação como um fator relevante na explicação das diferenças de renda e nível de esforço empreendido.

7. Referências

- ABADIE, A. Bootstrap tests for distributional treatment effects in instrumental variable models. *Journal of the American Statistical Association*, 97(457): 284-292, 2002.
- ABADIE, A. Semiparametric instrumental variable estimation of treatment response models. *Journal of Econometrics*, 113: 231-263, 2003.
- ARKES, J. Using unemployment rates as instruments to estimate returns to schooling. *Southern Economic Journal* 76(3), 711–722, 2010.
- ARNESON, R. Equality and Equal Opportunity for Welfare. *Philosophical Studies*, 56, 77-93, 1989.
- BOURGUIGNON, F.; FERREIRA, F. e MENÉNDEZ, M. Inequality of opportunity in Brazil. *Review of Income and Wealth*, 53, 585-618, 2007.
- CARNEIRO, P.; HECKMAN, J. J.; VYTLACIL, E. J. Estimating Marginal Returns to Education. *American Economic Review, American Economic Association*, vol. 101(6), pages 2754-81, 2011.
- CERULLI, G. Ivtreatreg: a new Stata routine for estimating binary treatment models with heterogeneous response to treatment under observable and unobservable selection, *Working Paper CNR-Ceris*, 2012.
- CHECCHI, D. & PERAGINE, V. Regional disparities and inequality of opportunity: the case of Italy. *Journal of Economic Inequality*, forthcoming, 2009.
- CHEN, S. H. e KHAN, S. Estimating the Causal Effects of Education on Wage Inequality Using IV Methods and Sample Selection Models, working paper, SUNY at Albany, 2007.
- CHEN, S. H. Estimating the variance in wages in the presence of selection and unobservable heterogeneity. *Review of Economics and Statistics*, 90(2), 2008.
- D'HAULTFOEUILLE, X. e MAUREL, A. Inference on a Generalized Roy Model, with an Application to Schooling Decisions in France. IZA Discussion Papers 4606, Institute for the Study of Labor (IZA), 2009.
- DWORKIN, R. What is equality? Part 1: Equality of welfare. *Philosophy & Public Affairs*, 10, 185-246, 1981a.
- DWORKIN, R. What is equality? Part 2: Equality of resources. *Philosophy & Public Affairs*, 10, 283-345, 1981b.
- FERREIRA, S. & VELOSO, F. Intergenerational mobility of wages in Brazil. *Brazilian Review of Econometrics*, 26, 181-211, 2006.

FLEURBAEY, M. e MANIQUET, F. *Compensation and responsibility*. In: Arrow, K. Sen, A. e Suzumura, K. (eds). *Handbooks in Economics: Social choice and welfare*, v.II. Amsterdam: Elsevier, 2011.

FLEURBAEY, M. *Equality among responsible individuals*, in J. F. Laslier, M. Fleurbaey, N. Gravel, A. Trannoy (eds.). *Freedom in economics. New perspectives in normative analysis*, London: Routledge, 1998.

HECKMAN, J. J. e E. J. VYTLACIL, E. J. Local instrumental variables and latent variable models for identifying and bounding treatment effects. *Proceedings of the National Academy of Sciences*, 96, 4730-4734, 1999.

HECKMAN, J. J. e HONERE, B. E. The empirical content of the Roy model. *Econometrica*, 58(5): 1121-1149, 1990.

HECKMAN, J. J. e VYTLACIL, E. J. Structural equations, treatment effects and econometric policy evaluation. *Econometrica*, 73 (3), 669-738, 2005.

IMBENS, G. W; e RUBIN, D. Estimating outcome distributions for compliers in instrumental variable models. *Review of Economic Studies*, 64(4): 555-574, 1997.

JONES, M. C.; MARRON, J. S.; SHEATHER, S. J. A brief survey of bandwidth selection for density estimation. *Journal of the American Statistical Association*, 91(433): 401-407, 1996.

MANSKI, C. Nonparametric bounds on treatment effects. *American Economic Review, Papers and Proceedings*, 80:319–323, 1990.

MANSKI, C. e PEPPER, J. (2000). Monotone instrumental variables: With na application to the returns to schooling. *Econometrica*, 68:997–1010, 2000.

OOGHE, E.; SCHOKKAERT, E. e VAN DE GAER, D. Equality of opportunity versus equality of opportunity sets. *Social Choice and Welfare*, 28, 209-230, 2007.

PASTORE, J. & SILVA, N. *Mobilidade social no Brasil*. São Paulo: Makron Books, 1999.

POWELL, J. L.; STOCK, J. H.; STOKER, T. M. 1989. Semiparametric estimation of index coefficients. *Econometrica*, 57: 1403-1430, 1989.

QUANDT, R. E. A New Approach to Estimating Switching Regressions. *Journal of the American Statistical Association*, 67 (338), 1973.

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

ROEMER, J. A pragmatic theory of responsibility for the egalitarian planner. *Philosophy & Public Affairs*, 10, 146-166, 1993.

ROEMER, J. *Equality of Opportunity*. New York: Harvard University Press, 1998.

ROSENBAUM, P. e RUBIN, D.B. The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*, 70, 41-55, 1983.

ROY, A. D. Some thoughts on the distribution of earnings. *Oxford Economic Papers*, 3(2): 135-146, 1951.

RUBIN, D. B. Assignment to treatment group on the basis of a covariate. *Journal of Educational Statistics*, 2: 1-26, 1977.

SHAIKH, A. e VYTLACIL, E. Limited dependent variable models and bounds on treatment effects: *A nonparametric analysis*. University of Chicago e Columbia University, 2004.

SILVERMAN, B. W. *Density Estimation*. Chapman and Hall: London, 1986.

VYTLACIL, E. J. e YILDIZ, N. Dummy endogenous variables in weakly separable models. *Econometrica*. Forthcoming, 2007.

VYTLACIL, E. J. Independence, monotonicity, and latent index models: An equivalence result. *Econometrica*, 70 (1), 331-341, 2002.

APÊNDICE A

Densidades Kernel para o Efeito de Tratamento Médio

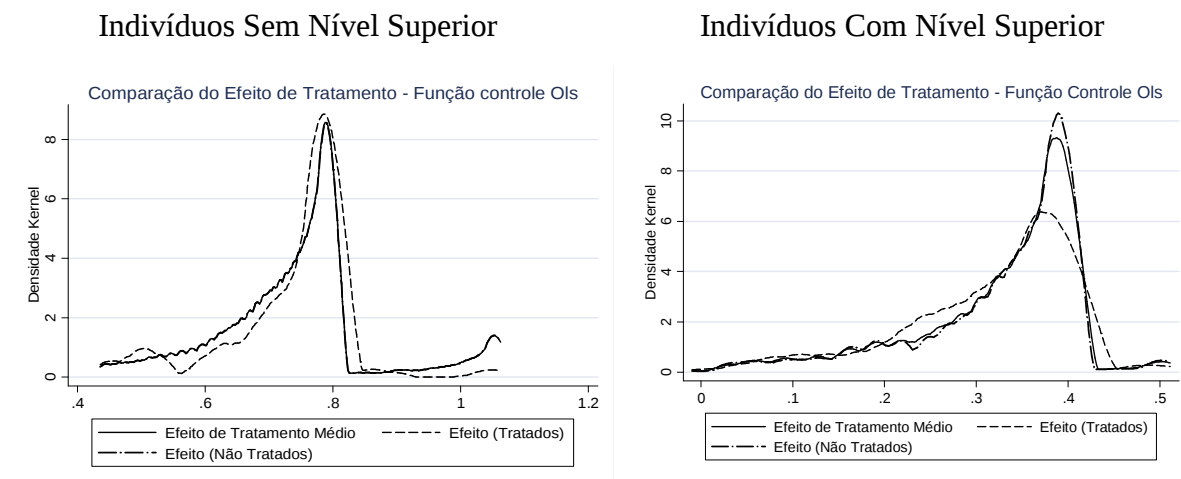


Figura 3: Densidades Kernel – Educação do Pai (Tratamento)

Fonte: Elaboração própria com base nos dados

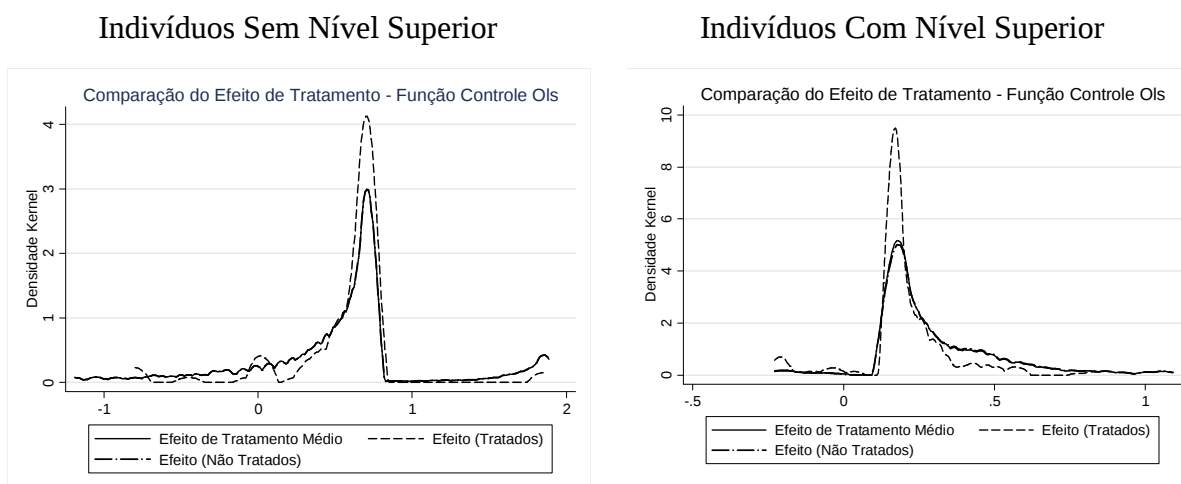


Figura 4: Densidades Kernel – Educação da Mãe (Tratamento)

Fonte: Elaboração própria com base nos dados