

ISSN 1519-4612

Universidade Federal Fluminense

TEXTOS PARA DISCUSSÃO

UFF/ECONOMIA

Universidade Federal Fluminense

Faculdade de Economia

Rua Tiradentes, 17 – Ingá – Niterói (RJ)

Tel.: (0xx21) 2629-9699 Fax: (0xx21) 2629-9700

<http://www.proac.uff.br/econ/>

Editor: Luiz Fernando Cerqueira; lcer@uol.com.br; lfcerqueira@economia.uff.br.

**O ensino médio técnico e os
rendimentos do trabalho**

Marina Aguas*

Danielle Carusi Machado**

**TD 318
Dezembro/2015**

© * IBGE, Rio de Janeiro, Brazil, e-mail: mariffaguas@gmail.com.

** Professora Adjunta da Faculdade de Economia da UFF e pesquisadora do CEDE/UFF.
E-mail: dani_carusi@hotmail.com.

RESUMO

Qualificação profissional é um tipo de investimento em capital humano cujo objetivo é fornecer ao trabalhador conhecimentos, competências e práticas em uma área específica, de forma a garantir uma melhora de seu desempenho no mercado de trabalho. Investimento em qualificação profissional pode reduzir as barreiras à entrada no primeiro emprego, assim como melhorar a qualidade do “casamento” entre trabalhador e empregador. Consequentemente, trabalhadores mais produtivos podem se beneficiar de salários mais altos, gerando impactos positivos na economia como um todo. Este artigo investiga a relação entre a educação técnica profissional e o nível de rendimentos no Brasil. Usando os dados da PNAD de 2007, a análise empírica é feita com o uso de três metodologias: MQO, efeito tratamento e propensity score matching. Os resultados mostram que ter educação técnica profissional gera um aumento da renda do trabalho de 20% a 24% e que a decisão de participar deste tipo de qualificação profissional não parece estar correlacionada com características produtivas não observadas destes trabalhadores.

PALAVRAS CHAVES: mercado de trabalho, renda do trabalho, qualificação profissional.

ABSTRACT

Vocational education is a type of human capital investment that aims to provide workers with knowledge, know-how, skills, specific competences and practices in a specific area, which is expected to improve their performance in the labor market. Investing in vocational education can also reduce the barriers to an individual's entry in their first-time job, as well as improve the quality of the job-worker match. Consequently, more skilled and productive workers can benefit from higher wages, promoting positive impacts for the entire economy. This paper investigates whether secondary vocational education is related to labor earnings in Brazil. Making use of data from the 2007 PNAD, the empirical analysis is conducted through three different methodologies: OLS, Treatment Effect and Propensity Score Matching. Evidence shows that attainment of a secondary vocational education is associated with an increase in labor earnings between 20% and 24% in all these methods. Evidence also suggests that the decision to attend such programs does not seem to be correlated with unobserved productive characteristics of workers.

KEYWORDS: labor market, income, vocational training

JEL: C21, J24, J31.

1. Introdução

Apesar dos avanços recentes na área da educação, a população brasileira ainda é caracterizada por baixos níveis de escolaridade. Segundo a Pesquisa Nacional por Amostra de Domicílios (PNAD/IBGE), em 2012, 60% das pessoas com mais de 25 anos de idade não haviam completado o ensino médio, e quase a metade sequer chegou a concluir o ensino fundamental. Entre as pessoas de 18 a 25 anos, grupo que deve estar se preparando para entrar ou que ingressou recentemente no mercado de trabalho, 45% não completaram o ensino médio. Essa pouca escolaridade da população e, em particular, dos trabalhadores, deve levar a uma baixa produtividade, que se converte em baixos rendimentos do trabalho e contribui para que grande parte dos indivíduos viva em situação de pobreza.

Neste contexto, ganham ainda mais força programas de ensino ligados à educação profissionalizante e ao treinamento profissional¹, que podem ser formas mais rápidas de elevar o capital humano individual voltado para o mercado de trabalho. Esse tipo de investimento é especialmente importante para os indivíduos que evadiram o sistema básico educacional antes do seu término, e para aqueles que o finalizaram e gostariam de continuar se qualificando, mas não tiveram vontade ou oportunidade de seguir adiante no ensino superior.

No Brasil, esse ramo educacional é rígido, atualmente, pela Lei 9.394 de 1996, denominada Lei de Diretrizes e Bases da Educação Nacional (LDB), que estabelece três segmentos de cursos: qualificação profissional, educação profissional técnica de nível médio e tecnológica de graduação e de pós-graduação. Cada uma dessas modalidades é estruturada para um público alvo determinado, possui objetivos e pré-requisitos específicos e, geralmente, são cursos com duração menor do que os cursos regulares de ensino². Essas modalidades não substituem a educação regular ou formal, no entanto, se tornam especialmente importante ao capacitar pessoas que por alguma razão abandonaram o sistema de ensino regular, ou que mesmo permanecendo nele, gostariam de ampliar sua qualificação para o trabalho. Logo, pode ser vista como uma forma mais rápida de capacitar a força de trabalho e aumentar sua produtividade individual, especialmente em países com uma grande parte da população com baixo nível de instrução.

O aumento do capital humano dos trabalhadores também pode extrapolar os benefícios individuais, na medida em que gera impactos para a economia como um todo. Do ponto de vista dos empregadores, uma força de trabalho mais qualificada e treinada aumenta a produtividade da firma, reduz os riscos e custos de contratação de um indivíduo cuja experiência profissional prévia não é de conhecimento da empresa, assim como reduz os custos de treinar um novo funcionário. A sociedade também se favorece com melhorias na capacitação profissional de sua população, visto que a maior produtividade dos participantes pode acarretar externalidades positivas, o aumento da renda pode gerar uma maior contribuição no pagamento de taxas e impostos, assim como deve possibilitar a elevação do consumo. Além disso, gastos sociais destinados ao pagamento de seguro desemprego podem ser reduzidos se a formação educacional propiciar uma redução na taxa de desocupação.

¹ De um modo geral, a educação profissional (EP) pode ser descrita como qualquer tentativa educacional organizada e sistemática que pode ser desenvolvida em escolas, instituições especializadas ou no próprio ambiente de trabalho. Tem como objetivo preparar os indivíduos para o mercado de trabalho, garantindo melhores perspectivas profissionais.

² No Brasil a educação regular é composta pelo ensino fundamental, ensino médio e ensino superior de graduação.

Com todos esses benefícios, diversos estudos procuram analisar se programas de aprendizado para o mercado de trabalho são realmente eficientes de forma a justificar os custos de realizá-los. Adicionalmente, as avaliações de impacto desses diversos cursos de educação profissional e treinamento sobre variáveis de desempenho do mercado de trabalho, em diferentes países, contribuem na elaboração e planejamento de políticas públicas, assim como permite avaliar e escolher quais os melhores formatos de cursos a serem oferecidos. Nos trabalhos realizados para os Estados Unidos, conclui-se que o impacto do treinamento profissional é, geralmente, modesto. Todavia, percebe-se uma grande heterogeneidade nos impactos dependendo das características dos participantes e do tipo de treinamento efetuado (Card et al. 2011, Heckman et al. 1999, Greenberg et al. 2003). As evidências de estudos europeus baseadas em uma grande diversidade de métodos para a avaliação dos programas de treinamento são menos claras (Kluve et al. 2007). Evidências sobre a eficácia do treinamento em países em desenvolvimento são relativamente limitadas. Betcherman et al. (2004) fazem uma resenha de 19 artigos que avaliam programas de treinamentos para jovens. Desses 19 artigos, 5 se referem a estudos para a América Latina. Enquanto para países desenvolvidos os impactos de treinamentos sobre a empregabilidade e os rendimentos se mostram neutros ou negativos, dependendo do estudo, para América Latina há indícios de efeitos positivos sobre a empregabilidade e sobre os rendimentos.

No Brasil, ainda são poucos os estudos empíricos que buscam avaliar o impacto dos programas de educação profissional. Hermeto e Rios-Neto (2007) analisaram o Plano Nacional de Qualificação Profissional (PLANFOR), a partir de um estudo experimental conduzido pelo CEDEPLAR entre 1996 e 2000. Os resultados sugeriam um impacto positivo sobre o desemprego para os indivíduos que participaram do PLANFOR, no sentido de reduzir o número e a duração dos episódios de desemprego. Severninni e Orellano (2010), usando dados da PPV/IBGE de 1996, mostram que pessoas que frequentaram cursos técnicos oferecidos em empresas tiveram sua probabilidade de desemprego reduzida e um impacto positivo sobre os rendimentos. Já Vasconcellos et al. (2010) estimam um impacto positivo de se cursar ensino médio técnico no salário recebido pelos indivíduos que realizaram essa formação. Já Barros et al. (2011), utilizando dados da PNAD de 2007, mostram que, em média, para a região Sudeste, os trabalhadores com educação técnica recebem uma remuneração 11% superior que a dos trabalhadores sem educação técnica, mas com os mesmos atributos. Reis (2014), com dados da Pesquisa Mensal de Emprego (PME/IBGE), encontra que o treinamento aumenta a empregabilidade em 1,1 pontos percentual, assim como em 6,8% o rendimento do trabalho mensal e em 8% o rendimento por hora trabalhada.

O objetivo deste artigo é investigar como é a relação entre os rendimentos do trabalho e a educação profissional técnica de nível médio para o Brasil. A escolha desse segmento, considerado intermediário, se deve ao fato de possuir uma legislação própria que cria diretrizes e regulamenta os cursos oferecidos pelas instituições de ensino. Logo, para receber o título de técnico é preciso ter o grau do ensino médio, frequentar e concluir com êxito cursos reconhecidos e estruturados pelo MEC. Dessa forma, tal modalidade se diferencia dos cursos de qualificação profissional, que ainda estão em processo de regulamentação, assim como não entra no leque de cursos profissionais do ensino superior. Objetiva-se, então, avaliar o retorno salarial de concluir um curso técnico de nível médio em relação às pessoas com o mesmo nível de instrução, mas que não realizaram essa formação.

A análise empírica tem como inspiração a equação de rendimentos especificada em Mincer (1974), que após algumas adaptações passa a incorporar a educação profissional como variável explicativa no contexto educacional. O salário por hora trabalhada se

torna, então, uma função da educação, tanto formal quanto profissional, da experiência e de características pessoais e regionais. Essa análise é realizada para as pessoas com idade entre 25 e 55 anos e que são moradoras das áreas metropolitanas ou autorepresentativas do Brasil.

A estimação da equação de rendimentos é feita através de três métodos distintos: Mínimos Quadrados Ordinários (MQO), Efeito-Tratamento (EF) e Propensity Score Matching (PSM). O MQO é o método mais usual de estimação, porém o estimador resultante só será não viesado caso a hipótese de que o valor esperado do termo não observado, condicionado nas variáveis explicativas, seja igual a zero, hipótese que normalmente é considerada forte na literatura econômica. Dessa forma, aplicam-se os dois outros métodos que se propõem a corrigir o possível viés de seleção, presente no fato de que escolhas educacionais não são aleatórias. Essa correção é feita com base em fatores observáveis, na medida em que esses métodos não conseguem captar a seletividade não observada, problema clássico da relação entre educação e salário. Logo, o artigo faz a comparação dos resultados obtidos entre os três métodos.

Os resultados mostram que a realização de curso técnico gera um impacto positivo e significativo sobre a remuneração do trabalho, independentemente da técnica utilizada. Pela estimação por MQO, esse efeito representa um incremento salarial de 22%. Já pelo método de ET, o ganho salarial esperado entre aqueles que realizam a capacitação é 21% maior do que entre os que não a realizaram. Nesse método, os resultados não sugerem a existência de viés de seleção amostral. Por fim, o método de PSM exhibe aumentos salariais que variam aproximadamente entre 22% e 24%, de acordo com o método de matching escolhido. Tais valores são similares aos encontrados no método de MQO e corroboram a ideia de ausência de seletividade nas características observáveis da amostra.

2. Dados e estatísticas descritivas

A partir dos dados da Pesquisa Nacional por Amostra de Domicílios (PNAD) de 2007, foi selecionada uma amostra de pessoas ocupadas na semana de referência e com rendimento do trabalho positivo. Entre estas, optou-se por investigar os indivíduos que moravam em regiões metropolitanas ou auto-representativas, que reportaram idade entre 25 e 55 anos e escolaridade igual ao ensino médio completo. Estas escolhas foram feitas na busca de reduzir problemas de heterogeneidade entre indivíduos e regiões, na medida em que grupos específicos devem apresentar características não observáveis muito diferentes. A restrição da análise ao grupo de pessoas que concluiu o ensino médio, e não entrou no ensino superior, tem como objetivo avaliar o retorno do nível técnico entre as pessoas que poderiam se beneficiar dele e que apresentam características mais semelhantes. Essa sub-amostra corresponde a 45% da população entre 25 e 55 anos de idade em 2007.

Com o suplemento sobre educação profissional desse ano na PNAD, foi possível identificar as pessoas que frequentaram algum curso de nível técnico, assim como aquelas que concluíram estes cursos. Também foram definidas variáveis relativas às características individuais, como sexo, idade e cor, e ao contexto regional. Adicionalmente, como parte do método de estimação de Efeito-Tratamento, foram usadas duas variáveis de identificação que se fundamentam em características observáveis dos indivíduos da amostra para tentar corrigir a seletividade dos dados. Essas variáveis tentam capturar como o local de moradia e seu entorno influenciam a decisão das pessoas em realizar cursos técnicos. A Tabela 1 apresenta a descrição dessas variáveis.

Tabela 1: Lista de variáveis

Variáveis	Descrição
Nível técnico	Igual a 1 se a pessoa realizou algum curso de nível técnico, 0 caso contrário.
Concluiu nível técnico	Igual a 1 se a pessoa realizou e concluiu algum curso de nível técnico, 0 caso contrário.
Percentual de pessoas que concluíram cursos de nível técnico no estrato de moradia	Para cada pessoa i com idade J e moradora do estrato T calcula-se o n° de pessoas, moradoras do estrato T , com idade maior ou igual a J , excluindo o próprio i , que concluíram curso técnico, dividido pelo n° de pessoas com idade maior ou igual a J , excluindo o próprio i , neste estrato T .
Percentual de pessoas que não realizaram cursos de nível técnico por falta de oferta de cursos no estrato de moradia	Para cada pessoa i com idade J e moradora do estrato T calcula-se o n° de pessoas, moradoras do estrato T , com idade maior ou igual a J , excluindo o próprio i , que não realizaram curso técnico devido a falta de oferta de curso na área de moradia, dividido pelo n° de pessoas com idade maior ou igual a J , excluindo o próprio i , neste estrato T .

Fonte: elaboração própria.

Com relação às características da amostra, as As características individuais dos grupos com e sem curso técnico, na

Tabela 3, se revelam muito próximas. O percentual de homens e mulheres é praticamente o mesmo, e o percentual de brancos é um pouco maior para o grupo que realizou curso de nível técnico, assim como sua idade média. Além disso, o nível de ocupação com rendimento do trabalho positivo entre os indivíduos que cursaram nível técnico alcança um patamar acima dos 80%, e representa uma diferença de 10 pontos percentuais para mais em relação ao grupo sem tal qualificação.

Por fim, é importante analisar o valor médio do rendimento do trabalho principal por hora trabalhada para os indivíduos na amostra, conforme a realização ou não de curso técnico de nível médio. Percebe-se que as pessoas que frequentaram curso de nível técnico recebem um maior salário por hora se comparadas àquelas que não o frequentaram. O salário/hora de quem nunca frequentou curso técnico é de aproximadamente 70% do salário/hora de quem frequentou.

Tabela 2 e Tabela 3 apresentam uma exposição do perfil dos indivíduos. Nota-se que 14% das pessoas entre 25 a 55 anos e com exatamente o ensino médio completo já frequentaram algum curso de nível técnico, e 12,7% concluíram este curso, o que equivale a um aproveitamento de 91,3%. Além disso, chama atenção o fato de que entre as pessoas que concluíram o nível técnico, 35% nunca chegaram a trabalhar na área profissional de especialização. Com relação aos indivíduos sem curso de nível técnico, é importante destacar que a maioria não o realizou por falta de interesse e uma minoria deixou de fazê-lo devido à ausência de oferta de cursos ou vagas.

As características individuais dos grupos com e sem curso técnico, na

Tabela 3, se revelam muito próximas. O percentual de homens e mulheres é praticamente o mesmo, e o percentual de brancos é um pouco maior para o grupo que realizou curso de nível técnico, assim como sua idade média. Além disso, o nível de ocupação com rendimento do trabalho positivo entre os indivíduos que cursaram nível técnico alcança um patamar acima dos 80%, e representa uma diferença de 10 pontos percentuais para mais em relação ao grupo sem tal qualificação.

Por fim, é importante analisar o valor médio do rendimento do trabalho principal por hora trabalhada para os indivíduos na amostra, conforme a realização ou não de curso técnico de nível médio. Percebe-se que as pessoas que frequentaram curso de nível técnico recebem um maior salário por hora se comparadas àquelas que não o frequentaram. O salário/hora de quem nunca frequentou curso técnico é de aproximadamente 70% do salário/hora de quem frequentou.

Tabela 2: Características pessoais segundo as Grandes Regiões e o Brasil

Variável	Brasil
Pessoas de 25 a 55 anos ocupadas e com rendimento do trabalho positivo (em 1000 pessoas, valores expandidos)	31.417
Realizou Curso de Nível Técnico (% de pessoas)	13,9
Concluiu Curso de Nível Técnico	91,4
Concluiu Curso de Nível Técnico e trabalha ou já trabalhou na área	64,0
Motivo de não realizar curso de NT: (% de pessoas)	
Cursa alguma modalidade de EP ou cursou QP	34,9
Falta de oferta de curso	2,4
Falta de interesse	36,8
Outros motivos	11,9

Fonte: elaboração própria com base nos microdados da PNAD 2007 do IBGE.

Tabela 3: Característica das pessoas ocupadas e com rendimento/hora positivo, segundo a realização de curso técnico

Variável	Não cursou nível técnico	
	Cursou nível técnico	Não cursou nível técnico
Mulher	47,3	45,8
Branco	57,3	53,5
Idade média	37,5	36,4
% de pessoas entre 25 e 55 anos, ocupadas e com rendimento do trabalho positivo, no total da população desta faixa etária	80,5	70,9
Rendimento médio por hora trabalhada (em R\$ de 2007)	8,82	6,21

Fonte: elaboração própria com base nos microdados da PNAD 2007 do IBGE.

3. Metodologia

A investigação sobre o impacto da educação de nível técnico sobre os rendimentos do trabalho tem como principal motivação a equação de rendimentos desenvolvida por Mincer (1974). Nesta, os rendimentos são uma função de fatores explicativos associados à escolaridade e à experiência. Com base nesse modelo, busca-se captar o efeito relacionado à realização de algum curso de nível técnico sobre o salário, através da inclusão de uma variável binária que indica se o indivíduo cursou ou não tal forma de educação profissional. Logo, o modelo a ser estimado é:

$$y_i = \ln R_i = f(nt_i, x_i, z_i) + u_i, i=1, 2, \dots, n (1).$$

Onde $\ln R_i$ indica o logaritmo natural dos rendimentos por hora do indivíduo i , nt_i é a variável binária que indica se a pessoa i realizou ou não curso de nível técnico, x_i denota experiência, z_i são características individuais e u_i é um erro estocástico.

Inicia-se a estimação através da aplicação do método MQO. Esses resultados, no entanto, podem ser questionados na medida em que existe uma provável fonte de viés de seleção advinda da escolha de fazer ou não um curso técnico. Nesse sentido, um indivíduo que possua características não observáveis como, por exemplo, uma maior habilidade, um forte empenho ou um retorno esperado da escolaridade mais alto, deve investir mais em seu capital humano e assim ter uma maior probabilidade de realizar um curso técnico. Provavelmente, essas características por si só já estariam relacionadas a um salário mais elevado. Entretanto, elas também podem influenciar as decisões sobre as escolhas educacionais. Logo, a estimação da equação de salário por MQO não resultaria em um valor correto do coeficiente para variável de nível técnico.

Com o propósito de buscar saídas para o problema de seleção amostral, a segunda etapa de estimação utiliza os métodos de EF e PSM que são muito comuns na literatura econômica de avaliação de impacto de programas e políticas públicas. Tais métodos buscam medir o efeito causal de um “tratamento” ou “programa” genérico sobre uma variável de resultado e podem ser aplicados neste estudo ao considerar a realização de um curso de nível técnico como um “programa” ou “tratamento” e o valor do rendimento/hora do trabalho como a variável de resultado. A forma ideal de se avaliar o impacto desse “programa” seria calcular, simultaneamente para cada indivíduo, a diferença entre o resultado salarial alcançado quando ele realiza o curso técnico e quando não o realiza. Essa situação, todavia, não é plausível de ser observada, sendo necessário utilizar algum método que estime o resultado contra factual do não tratamento. Os métodos EF e PSM propõem formas diferentes para estimar esse resultado contra factual e são apresentados a seguir.

3.1. Modelo de Efeito-Tratamento³:

Esse método estima o efeito da variável endógena binária, nt_i , sobre a variável contínua observada y_i , salário por hora trabalhada, condicionada nas variáveis independentes exógenas, x_i e w_i . A variável nt_i indica se a pessoa i recebeu ou não o tratamento, ou seja, se realizou ou não curso técnico. A decisão binária de obter o tratamento é modelada como resultado de uma variável latente não observada, nt_i^* , que supostamente

³ Para maiores informações ver Maddala (1983), Barnow et al. (1981) e Cong & Drukker (2000).

é uma função linear das variáveis explicativas exógenas, w_i , e de um componente de erro aleatório, e_i . Ou seja,

$$\begin{aligned} y_i &= x_i \beta + \delta nt_i + u_i \quad (1) \\ nt_i^* &= w_i \gamma + e_i \quad (2) \\ nt_i &= \begin{cases} 1, & \text{se } nt_i^* > 0 \\ 0, & \text{caso contrário} \end{cases} \quad (3) \end{aligned}$$

Onde, e e u possuem distribuição normal bivariada com média zero e correlação ρ .

Cabe destacar que, apesar desse método ter uma especificação paramétrica e, com isso, a identificação possa ser alcançada mesmo se $W = X$, na prática, o estimador irá apresentar um mau desempenho na ausência de uma restrição de exclusão em W . Dessa forma, X compreende as variáveis que indicam o sexo, a idade, a idade ao quadrado, a cor, a unidade da federação de moradia e a situação do domicílio de cada indivíduo, e W inclui todas as variáveis de X e as duas variáveis de identificação apresentadas na seção de dados.

Essas variáveis de identificação têm como objetivo retratar como é a percepção sobre cursar ou não o ensino de nível técnico no entorno de residência de cada indivíduo i . Acredita-se que quanto mais elevada a proporção de pessoas com idade maior ou igual à de i que concluíram um curso de nível técnico, maior será a influência sobre sua decisão de cursar tal modalidade. Da mesma forma, quanto maior o percentual de pessoas que não fizeram tais cursos por conta de falta de oferta, menor será a probabilidade de i realizá-lo, dado que provavelmente também terá se deparado com as mesmas dificuldades.

Considerando o problema de seleção entre os indivíduos que cursaram nível técnico e aqueles que não o fizeram, pode-se representar a equação de rendimentos dos indivíduos tratados por:

$$E(y_{1i} | nt_i = 1) = x_i \beta + \delta + E(u_i | nt_i = 1) = x_i \beta + \delta + \rho \sigma_u \lambda(w_i \gamma) \quad (4)$$

E a equação dos indivíduos não tratados, por:

$$E(y_{0i} | nt_i = 0) = x_i \beta + \rho \sigma_u \left[\frac{-\phi(w_i \gamma)}{1 - \Phi(w_i \gamma)} \right] \quad (5)$$

Onde λ é a razão inversa de Mills e os parâmetros são estimados em dois estágios.

Logo, a diferença nos ganhos esperados entre os participantes do curso de nível técnico e os não participantes pode ser escrita como:

$$E(y_{1i} | nt_i = 1) - E(y_{0i} | nt_i = 0) = \delta + \rho \sigma_u \left[\frac{\phi(w_i \gamma)}{\Phi(w_i \gamma) \{1 - \Phi(w_i \gamma)\}} \right] \quad (6)$$

Onde ϕ é a função densidade normal padrão e Φ a sua função distribuição acumulada. Se a correlação entre os termos de erro, ρ , é zero, o problema se reduz a uma estimação por MQO e a diferença é simplesmente δ . No entanto, acredita-se que ρ seja positivo, pois existem fatores não observados que influenciam a decisão de realizar um curso de nível técnico. Dessa forma, o método de MQO deve superestimar o efeito do tratamento, δ .

3.2. Método de Matching

O método de *matching* é uma abordagem não paramétrica que procura reproduzir *ex post* um experimento. Isso ocorre na medida em que é feita a escolha de um grupo de comparação entre os indivíduos não tratados, denominado grupo de controle, tal que o grupo selecionado seja o mais semelhante possível, em termos das suas características observáveis, ao grupo que recebeu o tratamento. A ideia é que essa comparação reduza o viés de seletividade que está relacionado a não aleatoriedade da decisão de realizar um curso de nível técnico. Dessa forma, todas as diferenças entre os indivíduos que fizeram o curso e os que não o fizeram seriam capturadas por seus atributos observáveis, remanescendo apenas a diferença entre o seu estado de tratamento. Logo, o resultado médio da variável de interesse resultante desse novo grupo constitui uma contrapartida amostral para a omissão de informações sobre os resultados que os indivíduos que realizaram o curso de nível técnico experimentariam, em média, caso não o tivessem realizado.

Criar uma correspondência ou *matching* entre os dois grupos que considere diversas características observáveis dos indivíduos é uma tarefa que se torna inviável em grandes amostras. Surge, então, um método que propõe resumir em uma única variável índice, denominada *propensity score*, as características individuais pré-tratamento e permitir assim a correspondência entre os dois grupos. O *propensity score* (PS) é definido por Rosenbaum e Rubin (1983) como a probabilidade condicional de receber o tratamento, dadas às características observadas pré-tratamento, ou seja:

$$p(X) \equiv Pr(nt = 1|X) = E(nt|X)(7)$$

Rosenbaum e Rubin (1983) mostram que, se a exposição ao tratamento é aleatória dentro dos grupos definidos por X , então este também será dentro dos grupos definidos por $p(X)$. Como resultado, se o PS de um indivíduo i , $p(X_i)$, é conhecido, então o efeito médio do tratamento sobre os tratados (ATT) pode ser estimado da seguinte forma:

$$\begin{aligned} ATT &\equiv E[Y_{1i} - Y_{0i}|nt_i = 1] = E\{E[Y_{1i} - Y_{0i}|nt_i = 1, p(X_i)]\} \\ &= E\{E[Y_{1i}|nt_i = 1, p(X_i)] - E[Y_{0i}|nt_i = 0, p(X_i)] | nt_i = 1\}(8) \end{aligned}$$

Onde Y_{1i} e Y_{0i} são os resultados potenciais nas duas situações contra factuais de tratamento e não tratamento, respectivamente.

Duas hipóteses são necessárias para derivar as equações acima:

1. Balanceamento das variáveis pré-tratamento, dado o PS. Ou seja, dado um determinado $p(X)$, realizar ou não o curso técnico é aleatório e as unidades tratadas e não tratadas precisam ser em média observacionalmente idênticas.
2. Independência condicional da variável de resultado dado o PS: $Y_1, Y_0 \perp nt | p(X)$ (9)

Assim, o primeiro passo desse método é estimar um *probit* para obter $Pr(nt_i=1|X_i)$, ou seja, a probabilidade de cada pessoa realizar um curso de nível técnico dadas suas características observadas X . Em seguida, a amostra é separada em K grupos de acordo com um intervalo igualmente espaçado do PS. Para cada grupo, testa-se se as diferenças das variáveis X , entre os indivíduos que receberam o tratamento e os que não o receberam, são estatisticamente iguais. Caso haja diferença estatisticamente significativa em alguma variável X , são incluídos em X termos de ordem mais alta ou de interatividade. Esse procedimento é realizado até que o balanceamento seja verificado.

Como na maioria das aplicações é muito difícil observar duas unidades (tratada e não-tratada) com exatamente o mesmo valor de $p(X)$ e, assim, realizar o *matching*, foram propostos alguns métodos para resolver esse problema, sendo, os mais utilizados:

- Busca pelo vizinho mais próximo: para cada unidade tratada busca-se uma unidade de controle com um valor de PS mais próximo.
- Kernel*: todas as unidades tratadas são comparáveis de forma ponderada com cada unidade de controle existente. A ponderação leva em conta pesos que são inversamente proporcionais à distância entre $p(X)$ dos tratados e dos não tratados.
- Estratificação: consiste em dividir as observações em intervalos de acordo com o valor de $p(X)$, de tal modo que para cada intervalo os grupos de tratados e de não-tratados tenham em média o mesmo PS.
- Radius matching*: cada indivíduo, i , no grupo de tratamento é comparado com outros do grupo de não tratados, cujo escore de propensão está situado dentro de determinado um raio a partir da probabilidade estimada de i .

4. Resultados

A Tabela 4 exibe os resultados para a estimação da equação (1) por MQO. Na primeira e na terceira coluna, o indicador de educação profissional é a variável frequentou curso de nível técnico. Essas duas colunas se diferenciam na medida em que são incluídas variáveis binárias para captar o efeito específico de cada estrato geográfico na segunda especificação do modelo. Nas demais colunas, a variável de educação profissional passa a ser a conclusão de curso técnico de nível médio. Os resultados mostram que realizar algum curso de nível técnico significa obter um salário/hora aproximadamente 21% maior do que aquele recebido pelas pessoas que não fizeram tal curso, mesmo tendo todos os pré-requisitos para fazê-lo. Esse efeito é significativo para o nível de 1% e varia pouco quando são incluídas as variáveis binárias para os estratos. Além disso, é possível observar que a conclusão do curso gera um efeito ligeiramente mais elevado do que a sua simples realização, 22%, lembrando que 91% das pessoas que ingressaram em tais cursos de fato os concluíram.

Tabela 4: Estimação por MQO

Variáveis	Ln (Rendimento do trabalho principal por horas habitualmente trabalhadas)			
	MQO		MQO com controle por estratos geográficos	
Frequentou curso de nível técnico	0,211*** (0,016)		0,213*** (0,017)	
Concluiu curso de nível técnico		0,221*** (0,017)		0,224*** (0,017)
Observações	22.262	22.262	22.262	22.262
R ²	0,187	0,187	0,212	0,212

Erros padrão robustos entre parênteses. *** $p < 0,01$, ** $p < 0,05$, * $p < 0,1$.

Fonte: elaboração própria com base nos micro dados da PNAD 2007 do IBGE.

A Tabela 5 apresenta os resultados da estimação das equações (3) e (4) pelo método EF em dois estágios. O 1º estágio desse método estima a probabilidade que cada indivíduo tem de realizar um curso técnico. Isso é feito através de um *probit*, incluindo como variáveis explicativas informações sobre a pessoa e sobre o seu entorno de moradia. Acredita-se que o indivíduo possa ser influenciado em sua decisão de cursar ou não o nível técnico de acordo com o número relativo de pessoas a sua volta que se beneficiaram da realização de tais cursos. Ao mesmo tempo, se o entorno sofre com a falta de oferta de tais cursos, o indivíduo também passa a considerar essa informação na hora da sua tomada de decisão.

Tabela 5: Estimação Efeito-Tratamento

Variáveis	Log Rendimento/hora	Frequentou curso de nível técnico	hazard	Log Rendimento/hora	Concluiu curso de nível técnico	hazard
Frequentou curso de NT	0,343** (0,133)					
Concluiu curso de NT				0,357** (0,145)		
% de pessoas que concluíram cursos de NT no estrato de moradia		0,038*** (0,004)			0,036*** (0,004)	
% de pessoas que não realizaram cursos de NT por falta de oferta de cursos no estrato de moradia		-0,051*** (0,012)			-0,048*** (0,012)	
Lambda			-0,072 (0,073)			-0,073 (0,078)
Constante	0,306*** (0,104)	-1,782*** (0,255)		0,386*** (0,108)	-2,888*** (0,266)	
Observações	22.262	22.262	22.262	22.262	22.262	22.262
Diferença nos ganhos esperados entre os participantes de NT e os não participantes deste curso		0,2074*** (0,010)			0,2176*** (0,010)	

Erros padrão entre parênteses. ** p<0,01, * p<0,05, * p<0,1

Fonte: elaboração própria com base nos micro dados da PNAD 2007 do IBGE.

Os resultados para o 1º estágio indicam que as variáveis de entorno exibem o comportamento esperado na medida em que a decisão de fazer curso técnico é positivamente associada ao percentual de pessoas que concluíram tais cursos no entorno de moradia do indivíduo, e negativamente associada ao percentual de pessoas que não realizaram cursos técnicos no entorno de moradia por motivo de ausência de cursos ou vagas. Ambos os efeitos são significativos para o nível de 1%. O 2º estágio estima a equação de salário levando em consideração todas as variáveis explicativas e um termo de seletividade proveniente do primeiro estágio, representado pela variável *lambda*. As covariadas que indicam a realização do curso técnico ou sua conclusão são significativas e apresentam um efeito positivo sobre o salário/hora. Já o termo de correção para a presença de seleção se mostra negativo, porém não é significativo tanto para a especificação que considera a realização do curso, quanto para a que está interessada na sua conclusão. Logo, os resultados sugerem a não existência de seleção entre as pessoas da amostra quando são utilizadas as duas variáveis de identificação definidas. Para completar a análise é preciso calcular a diferença nos ganhos esperados entre aqueles que cursaram nível técnico e os que não cursaram, conforme a equação (8). Esse resultado indica uma diferença positiva e significativa em torno de 20,7% para aqueles que realizaram curso técnico de nível médio e de 21,8% para os que o concluíram. Tais valores são ligeiramente inferiores aos resultados obtidos do método de MQO.

Da mesma forma que o método de EF, a metodologia de PSM propõe uma saída para o problema de seletividade da decisão de investir em capacitação de nível técnico. Todavia, é uma abordagem não paramétrica que seleciona através de um OS um grupo de comparação entre as pessoas que não realizaram curso técnico, tal que esse grupo represente uma situação contra factual do não tratamento. A Tabela 6 apresenta os resultados para o efeito médio do tratamento sobre os tratados (ATT), conforme visto na equação (10). A questão chave dessa metodologia é realizar a correspondência entre o grupo de tratamento e o grupo de controle da melhor forma possível. Assim, as linhas desta tabela indicam as quatro formas de *matching* que foram utilizadas.

Tabela 6: Estimação por Propensity Score Matching

Variáveis	ATT					
	Frequentou curso de nível técnico	Nº tratados	Nº controle	Concluiu curso de nível técnico	Nº tratados	Nº controle
Vizinho mais próximo	0,221 (0,016)	3083	12793	0,231 (0,017)	2793	12443
Kernel	0,232 (0,014)	3083	19151	0,235 (0,015)	2793	19442
Estratificação	0,209 (0,001)	3083	19151	0,219 (0,016)	2793	19442
Radius matching raio=0,01	0,229 (0,015)	3083	19100	0,225 (0,015)	2793	19442
Radius matching raio=0,001	0,228 (0,015)	3083	19100	0,220 (0,015)	2792	19372
Radius matching raio=0,0001	0,234 (0,015)	3058	17482	0,229 (0,016)	2775	18020

Fonte: elaboração própria com base na PNAD 2007 do IBGE.

A 1ª forma de correspondência é o método do vizinho mais próximo. Este procura, para cada indivíduo tratado, um indivíduo correspondente no grupo de controle que mais se aproxima do valor da sua probabilidade de fazer algum curso de nível técnico. O valor do ATT, que é a média das diferenças de cada grupo, indica que realizar tal curso provoca um aumento de 22% nos rendimentos por hora de trabalho, e a sua conclusão um impacto positivo de 23%. Algumas fragilidades do método são, por exemplo, o fato de não considerar as informações de todas as pessoas do grupo de controle e de poder realizar uma correspondência de baixa qualidade se o vizinho mais próximo estiver muito distante em relação ao $p(x)$,

O método de Kernel possui um ganho em relação ao primeiro método, pois considera na correspondência todas as pessoas da amostra de uma maneira ponderada. Essa ponderação atribui um maior peso para as observações do grupo de controle que tem uma probabilidade mais próxima de fazer o curso em relação às observações do grupo de tratamento. O resultado do efeito médio do tratamento sobre os tratados é positivo e próximo a 23%, tanto para a variável de realização do curso quanto para a que indica a sua conclusão.

O método de *matching* por estratificação divide grupos de tratados e de não-tratados segundo um intervalo de $p(x)$, de tal forma que a média do PS seja a mesma em cada intervalo. Em seguida, calcula-se a diferença entre as médias da variável de resultado dos tratados em relação ao grupo de controle. O resultado mostra um ATT de 20,9% quando a variável é a frequência em curso técnico e 21,9% quando esta variável passa a ser a conclusão do curso.

No último método, *radius matching*, cada unidade tratada é comparável apenas com as unidades de controle cujo PS estimado está em uma vizinhança predefinida da unidade tratada. Se o raio é configurado para ser muito pequeno, possivelmente em algumas unidades tratadas pode não haver *matching*, já que a vizinhança pode não conter unidades de controle. Por outro lado, quanto menor for o tamanho da vizinhança, melhor a qualidade dos resultados. Nesse sentido, foram utilizados três tamanhos de raio e, no entanto, os valores de ATT para esses três tamanhos não tiveram grandes variações. Para a variável de frequência em curso de nível técnico o incremento de renda variou entre 22,8%, considerando o maior raio, e 23,4%, considerando o menor raio. Já no caso da conclusão de curso esses valores variaram, respectivamente, entre 22,5% e 22,9%.

Assim, os quatro métodos utilizados para realizar a correspondência entre o grupo tratado e o grupo de controle geraram resultados similares entre si e, principalmente, valores próximos aos coeficientes obtidos pela estimação de MQO, que não considera a

presença de seleção amostral. No mesmo sentido, os resultados estimados pelo método de efeito-tratamento também condizem com os resultados de MQO.

5. Considerações finais

Com a hipótese de ausência de seletividade amostral, a estimação por MQO da equação de rendimento por hora resulta em um incremento de renda em torno de 21% para as pessoas que frequentaram cursos técnicos e de 22% para aquelas que os concluíram. Ambos os resultados se mostraram significativos ao nível de 1%. Para que a hipótese de seletividade na decisão de frequentar um curso técnico seja considerada, foram utilizados os métodos de EF e de PSM. Os resultados para o método de EF não sugerem a existência de viés de seletividade e o ganho esperado nos rendimentos do trabalho entre aqueles que cursaram nível técnico e os que não o cursaram ficam próximos aos valores exibidos na estimação por MQO. Já o método de PSM, com sua abordagem não paramétrica, conclui que essa forma de capacitação profissional tem um efeito positivo que varia entre 20,9% a 23,5% conforme o método de matching escolhido.

Logo, os diferentes métodos utilizados mostraram efeitos positivos e significativos da variável de nível técnico sobre o rendimento do trabalho por hora. Esses efeitos, no entanto, são muito próximos e não há evidências de seletividade nos dados da amostra. Vale destacar que os métodos de estimação que buscam corrigir o possível viés de seleção dos dados são baseados em hipóteses fortes e seus resultados devem ser interpretados com cautela. No caso do método de EF, chama-se atenção para a hipótese de que toda heterogeneidade entre os indivíduos com relação à decisão de fazer um curso técnico é captada pelas variáveis de identificação. Já no método de PSM, supõe-se que, condicionada na probabilidade de participação, a diferença média dos resultados entre o grupo de tratamento e o de controle é toda explicada pela realização do curso técnico.

6. Referências bibliográficas

- BARROS, R. P; FRANCO, S.; GOSNER, D.; MENDONÇA, R.; ROSALÉM, A., Educação técnica e distribuição de renda no Espírito Santo, Texto de Discussão 33, Instituto Jones dos Santos Neves, 2011.
- BARNOW, B. S., G. G. CAIN, and A. S. GOLDBERGER., Issues in the analysis of selectivity bias. *Evaluation Studies Review Annual* 5: 43–59. 1981.
- BECKER, S. O.; ICHINO, A., Estimation of Average Treatment Effects Based on Propensity Score. *The Stata Journal*, Vol.2, número 4, p.358-377, 2002.
- Betcherman, G. K. Olivas, and A. Dar., Impacts of active labour market programs: new evidence from evaluations with particular attention to developing and transition countries. *World Bank Social Protection Discussion Paper* 0402, 2004.
- CARD, D., IBARRARÁN, P., REGALIA, F., ROSAS-SHADY, D., SOARES, Y.. The Labor Market Impacts of Youth Training in the Dominican Republic. *Journal of Labor Economics* 29 (2), pp. 267-300, 2011.
- CONG, R., & D. DRUKKER., Treatment effects model. *Stata Technical Bulletin* 55: 25-33. 2000
- GREENBERG, D., MICHALOPOULOS, C., & ROBINS, P. K. A meta-analysis of government-sponsored training programs. *Industrial and Labor Relations Review*, 57, 31–53, 2003.

- HECKMAN, J. J., R. J. LALONDE, AND J. A. SMITH, “The Economics and Econometrics of Active Labor Market Programs”, In O. Ashenfelter and D. Card (Eds.), Handbook of Labor Economics, vol. III A, pp. 1865–2097. Amsterdam, 1999.
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA - IBGE. Microdados e metodologia da Pesquisa Nacional por Amostras de Domicílios, 2007. Disponível em: <http://www.ibge.gov.br/>.
- KLUVE, J. The effectiveness of European active labor market programs. Labour Economics, 17(6), 2010, 904-918.
- MADDALA, G. S., Limited-Dependent and Qualitative Variables in Econometrics. Cambridge, MA: Cambridge University Press.1983.
- MINCER, J., Schooling, experience and earnings. New York: Columbia University Press, 1974.
- OLIVEIRA, A.; RIOS NETO, E.. Uma avaliação experimental dos impactos da política de qualificação profissional no Brasil: o caso do PLANFOR na Região Metropolitana de Belo Horizonte. Revista Brasileira de Economia, 61(3), 2007.
- REIS, M., Vocational training and labor market outcomes in Brazil. The B.E. Journal of Economics Analysis and Policy. A sair. (2014)
- ROSENBAUM, P. R.; D. B. RUBIN., The central role of the propensity score in observational studies for causal effects. Biometrika 70(1): 41–55, 1983.
- SEVERNINI, E. R.; ORELLANO, V. I. F., O efeito do ensino profissionalizante sobre a probabilidade de inserção no mercado de trabalho e sobre a renda no período pré-PLANFOR. Revista ANPEC, v. 11, p. 155-174, 2010.
- VASCONCELLOS, L., LIMA, F. C., FERNANDES, J. G. E MENEZES FILHO, N. A. Avaliação Econômica do Ensino Médio Profissional. Relatório de avaliação nº 14, Programa Avaliação Econômica de Projetos Sociais, Fundação Itaú Social, 2010.