

# O IMPACTO DO ENSINO MÉDIO TÉCNICO SOBRE A RENDA: UMA COMPARAÇÃO ENTRE BRASIL E MINAS GERAIS

Alan André Borges da Costa<sup>\*</sup>  
Sibelle Cornélio Diniz<sup>1</sup>

## Resumo

O objetivo deste trabalho é analisar o retorno do ensino médio técnico sobre o salário dos indivíduos a nível Brasil e Minas Gerais. Na parte metodológica foi utilizado o *propensity score matching* e posteriormente calculou-se o efeito tratamento médio sobre os tratados (ATT). Ademais, foram estimadas regressões com o intuito de obter o efeito tratamento médio (ATE). Os resultados mostram que o impacto do ensino médio técnico é positivo, tanto para ATT quanto para ATE, e ocorre com maior intensidade em Minas Gerais.

Palavras-Chave: ensino médio técnico, *propensity score matching*, microeconometria.

Área: Economia Mineira.

---

<sup>\*</sup>U Professor na Universidade Federal de Ouro Preto (UFOP) e Fundação Itaú Social.

<sup>1</sup> Doutoranda em economia no Cedeplar/UFMG

# O IMPACTO DO ENSINO MÉDIO TÉCNICO SOBRE A RENDA: UMA COMPARAÇÃO ENTRE BRASIL E MINAS GERAIS

Alan André Borges da  
Costa\*

Sibelle Cornélio Diniz<sup>2</sup>

## 1 – Introdução

A despeito da queda recente do desemprego, da pobreza e da desigualdade no Brasil, acompanhada do crescimento da formalização do emprego, parcela da população encontra-se ainda “presa” a situações de insuficiência de renda e articulação inexistente ou precária ao mercado de trabalho (IPEA, 2011).

Estudos realizados indicam que os esforços, no nível federal, de inserção dessa população no mercado de trabalho, são, de modo geral, de baixa eficácia (IPEA, 2011 e Teixeira, 2004). Entre os diversos obstáculos, encontram-se a escassez de recursos e as dificuldades de gestão e de integração entre as políticas; por outro lado, é notável a incapacidade de boa parte das ações de atingirem o público alvo.

De fato, a situação de pobreza e exclusão social é configurada por múltiplas dimensões objetivas e subjetivas, individuais e coletivas, que impõem grandes desafios ao seu enfrentamento, mesmo em uma conjuntura de crescimento econômico. A multidimensionalidade da pobreza e da exclusão dá forma a várias armadilhas que tendem a prender os indivíduos e as famílias nessa condição, mesmo quando eles conseguem avançar em algumas dimensões da sua vida (Gordon, 2000). Dentre as dificuldades vivenciadas, encontram-se: deficiências educacionais, inadequação aos processos formais de educação e de trabalho, problemas familiares, questões nutricionais e de saúde, indisponibilidade para o trabalho por comprometimentos com a família (exemplo: filhos e/ou familiares doentes), distância entre a moradia e o local de trabalho entre outros.

Na década de 2000, sobretudo nos anos mais recentes, observou-se uma modificação desta tendência, com aceleração do ritmo de crescimento de importantes segmentos da indústria e dos serviços mais qualificados, refletindo numa nova dinâmica de formalização da economia<sup>3</sup> e redução da desigualdade de renda e da pobreza. Contudo, este quadro não implicou a eliminação da insuficiência de renda e do desemprego; parte da população se mostra deslocada do atual momento de expansão econômica, permanecendo em condição de pobreza enquanto persistem as mais diversas formas de produção não assalariadas e informais:

*(...) Não obstante, é necessário lembrar que, por melhor que seja essa conjuntura, ela não se mostrou capaz de reverter problemas historicamente acumulados por um processo de desenvolvimento tardio. O mercado de trabalho brasileiro permanece criando ocupações muito distintas, em termos de formas de contratação,*

---

\*Professor na Universidade Federal de Ouro Preto (UFOP) e Fundação Itaú Social.

<sup>2</sup> Doutoranda em Economia no Cedeplar/UFMG.

<sup>3</sup> Segundo a Pesquisa Mensal de Emprego, do IBGE, a taxa de desemprego aberto no Brasil recuou a 7% em 2010, a participação dos empregados sem carteira assinada, conta-própria e não remunerados reduziu-se a 37,1% do total de ocupados, e a remuneração média real mensal cresceu 3,3% (IPEA, 2011).

*níveis de remuneração, patamares de jornada, garantias de saúde e segurança, mecanismos de estabilidade, meios de representação coletiva, entre outros aspectos que se referem ao acesso a direitos trabalhistas e previdenciários. (IPEA, 2011, p. 201)*

O sistema público federal de emprego, tal como idealizado na Constituição, se propunha a reduzir o desemprego decorrente de falta de informações e de desequilíbrios entre qualificações requeridas pelos empregadores e habilidades dos trabalhadores: “A lógica é que os trabalhadores demitidos teriam parte de sua renda repostada pelo seguro-desemprego, obteriam auxílio para reempregar-se por meio dos serviços de orientação e intermediação de mão de obra e, se necessário, poderiam ampliar ou atualizar suas habilidades por meio de programas de treinamento e qualificação” (IPEA, 2011, p. 211). Além disso, a partir dos anos 2000, o sistema passa a estimular ou induzir a geração de postos de trabalho por meio da concessão de crédito facilitado a empresas e/ou trabalhadores auto-ocupados, associados ou cooperativados (programas de microcrédito e de apoio à economia solidária). O que se observa, no entanto, é que o sistema falha no atendimento à população mais vulnerável, tendo baixo alcance e efetividade para este público, além da pouca integração entre as ações.

No fim da década de 1990, cerca de 0,5% do PIB brasileiro era destinado a programas de capacitação profissional. No governo Fernando Henrique Cardoso, foi criado o Planfor (Plano Nacional de Formação Profissional), que seria implementado pelos Planos Estaduais de Qualificação (PEQ's). O Planfor conferia papel estratégico às Comissões Estaduais e Municipais de Emprego, bem como às Universidades e entidades sindicais, do setor privado e do Sistema S, e buscava priorizar o público mais vulnerável. Entretanto, como aponta Teixeira (2004), os cursos oferecidos demonstraram baixa capacidade de atender às necessidades do mercado, uma vez que não se respaldavam em diagnósticos das necessidades locais. Ademais, verificou-se a baixa integração com outras políticas, a inexistência de estruturas de acompanhamento e avaliação e baixa participação das comissões de emprego (IPEA, 2006; Peixoto, 2008).

Em 2003, foi proposto o Plano Nacional de Qualificação (PNQ), buscando contemplar cursos de maior duração e incorporando tópicos como empoderamento e cidadania. Previa, ainda, o fortalecimento do planejamento, do monitoramento e da avaliação da política. As ações do PNQ são desenvolvidas pelos Planos Territoriais de Qualificação (PlanTeQs), circunscritos a um território e articulados às políticas públicas locais, e pelos Planos Setoriais de Qualificação (PlanSeQs), projetos e ações de qualificação social e profissional de caráter estruturante, setorial ou emergencial. Teixeira (2004) e IPEA (2006) apontam para os ganhos de focalização do programa, que ampliou a participação de mulheres, jovens e desempregados. No entanto, este também se mostrou incapaz de incorporar mais incisivamente os indivíduos de baixa escolaridade e do setor informal. Além disso, em 2005, os programas de qualificação contaram com apenas 10% dos recursos gastos em 1998, durante o Planfor.

Souza (2009) avalia, com base na PNAD (Pesquisa Nacional por Amostra de Domicílios) e PME (Pesquisa Mensal de Emprego) 2007, que os retornos salariais dos cursos de qualificação profissional são mínimos para os trabalhadores de menor escolaridade. Indicam, ainda, que é pequena a parcela da população que realizou os cursos e se encontra ocupada. Outros estudos recentes (FGV, 2010; DIEESE, 2010) apontam grande participação do setor privado na oferta de treinamento profissional no Brasil – a maior parte dos entrevistados que realizaram cursos de capacitação o fizeram com recursos próprios. Tal configuração tende a acarretar em menor taxa de capacitação profissional entre os mais pobres, e com isso constituir-se em um mecanismo de

manutenção das desigualdades no mercado de trabalho brasileiro: “Nesse caso, o acesso mais amplo à formação dependeria na verdade da rede dos sistemas de aprendizagem e dos institutos de educação técnica e profissionalizante – que em geral não acolhem os trabalhadores menos qualificados e/ou de menor renda” (IPEA, 2011, p. 217). Ademais, FIS (2010) calculou recente o impacto de realizar o ensino médio técnico e chegou a conclusão que o aumento de renda é aproximadamente 12,5%.

Este trabalho contribui neste sentido e propõe estimar o impacto da conclusão do ensino médio para o Brasil como um todo e para um recorte que consiste no Estado de Minas Gerais. Além desta introdução a seção 2 apresenta a metodologia, a seção 3 os resultados e a 4 as considerações finais.

## 2 – Metodologia e dados

Sabe-se que a avaliação de impacto constitui essencialmente um problema de inferência causal, em geral, deseja-se isolar o efeito de alguma variável,  $T$ , sobre um indicador de interesse,  $y$  (IMBENS E WOOLDRIDGE, 2009). Aplicando ao caso do ensino técnico de ensino médio tem-se a seguinte questão: quais as condições em que se pode determinar se existe algum impacto do ensino técnico médio ( $T$ ) sobre o salário mensal ( $y$ )?

Para responder esta questão é necessário verificar como foram selecionados os indivíduos, isto é, como foi realizada a seleção dos alunos para cursar o ensino médio técnico. Esta seleção poderá ocorrer de forma aleatória ou não aleatória. No primeiro caso o impacto é calculado utilizando um simples teste de média<sup>4</sup>. Já no segundo caso a escolha da metodologia está condicionada se a auto-seleção dos indivíduos foi realizada com base em características observáveis ou não observáveis<sup>5</sup> para o pesquisador (IMBENS E WOOLDRIDGE, 2009; ANGRIST E PISCHKE, 2009). As próximas subseções abordam os modelos econométricos e a hipótese de identificação que possibilita estimar o impacto no caso de seleção em observáveis.

### 2.1 – Modelo de resultados potenciais, contrafactual e viés de seleção

Seja  $y$  o salário mensal e  $T$  uma variável *dummy* que assume o valor um se o indivíduo cursou o ensino médio técnico e zero caso contrário. Segundo Rubin (1974), para cada indivíduo  $i = 1, \dots, N$  existem dois resultados potenciais dados por  $Y_{i1}$  e  $Y_{i0}$ . O primeiro resultado,  $Y_{i1}$ , seria o salário mensal do indivíduo  $i$  caso este tivesse finalizado o ensino médio técnico. Já o segundo resultado,  $Y_{i0}$ , mostra o valor do salário mensal para o mesmo indivíduo  $i$  sem ter cursado ensino médio técnico. Repare que cada indivíduo possui dois resultados potenciais:

---

<sup>4</sup> Para mais sobre o teste de média ver Imbens e Wooldridge (2009, p.19)

<sup>5</sup> Para o caso de seleção em características não observáveis pode-se utilizar o estimador de diferenças em diferenças. Para mais ver Imbens e Wooldridge (2009, p.67) e Angrist e Pischke (2009, p.221).

$$\text{Resultado potencial do indivíduo } i = \begin{cases} Y_{i1} & \text{com ensino médio técnico} \\ Y_{i0} & \text{sem ensino médio técnico} \end{cases} \quad (1)$$

Apesar de ser do interesse do pesquisador, não é possível observar os dois resultados ao mesmo tempo, ou seja, tem-se apenas

$$Y = Y_{i1}T + Y_{i0}(1 - T) \quad (2)$$

Portanto, empiricamente estão disponíveis apenas  $Y_{i1}|T = 1$  e  $Y_{i0}|T = 0$ , mas nunca  $Y_{i1}|T = 0$  ou  $Y_{i0}|T = 1$ . Neste sentido fica claro que a avaliação de impacto torna-se um problema de dados ausentes (*missing data*), pois não é possível saber qual seria o salário do indivíduo tratado na condição de não tratado e vice-versa<sup>6</sup>.

Para resolver este problema é necessário utilizar o que a literatura denomina como contrafactual, ou seja, procura-se a resposta para a seguinte pergunta: o que teria ocorrido com os salários dos indivíduos caso não tivessem cursado o ensino médio técnico? É fácil ver que o contrafactual ideal seria observar o mesmo indivíduo  $i$  com e sem tratamento no mesmo período de tempo. No entanto, isto não é possível e a melhor solução é imitar este processo utilizando um grupo de controle que seriam os indivíduos semelhantes aos tratados, mas que não cursaram ensino médio técnico. (KHANDKER, KOOLWAL e SAMAD, 2010).

Segundo Imbens e Wooldrige (2009) e Caliendo e Hujer (2006) é de suma importância a escolha do grupo de controle a fim de evitar viés de seleção. Para ilustrar este fato considere os seguintes resultados de interesse

$$ATE = \mathbb{E}[Y_{i1} - Y_{i0}] = \mathbb{E}[Y_{i1}] - \mathbb{E}[Y_{i0}] \quad (3)$$

$$ATT = \mathbb{E}[Y_{i1} - Y_{i0}|T = 1] = \mathbb{E}[Y_{i1}|T = 1] - \mathbb{E}[Y_{i0}|T = 1] \quad (4)$$

em que ATE é o efeito tratamento médio (*Average Treatment Effects*) e ATT é o efeito tratamento médio sobre os tratados (*Average Treatment Effects on the Treated*). Repare que parte da equação (4) não pode ser calculada, pois não é possível obter o valor do salário médio do grupo de controle caso este fosse tratado ( $\mathbb{E}[Y_{i0}|T = 1]$ ). No entanto, se a seguinte igualdade  $\mathbb{E}[Y_{i0}|T = 1] = \mathbb{E}[Y_{i0}|T = 0]$  for válida, pode-se utilizar os indivíduos que não fizeram o ensino médio técnico como um grupo de controle adequado. Segundo Imbens e Wooldrige (2009) esta igualdade será válida apenas no caso de experimentos aleatórios. Logo, a diferença de médias entre tratado e controle no caso de experimentos não aleatórios não será válida devido ao viés de

---

<sup>6</sup> Na literatura de avaliação Holland (1986) denominou este problema como “problema fundamental da inferência causal”

seleção<sup>7</sup>. Por fim, pode-se dizer que “*Selection bias arises because participants and non-participants are selected groups that would have different outcomes, even in absence of the programme. It might be caused by observable factors, like age or skill differences, or unobservable factors like motivation*” (CALIENDO E HUIJER, 2006, p.4). Para resolver este problema é necessário refletir se a auto-seleção dos indivíduos ao ingressar no ensino médio técnico foi realizada com base em características observáveis (idade, região, gênero, etc.) ou não observáveis (motivação, interesse, habilidade etc.). Assim, torna-se necessária uma hipótese de identificação a fim de reduzir ou eliminar o viés de seleção.

## 2.2 – Hipóteses de identificação

A hipótese de identificação está dividida em duas subseções. A primeira parte aborda a identificação do impacto no caso da estimação por regressão bem como as variáveis utilizadas para reduzir ou eliminar o viés de seleção. Na segunda parte é apresentada a metodologia de *Propensity Score Matching* que possibilita identificar o impacto através de uma melhor seleção do grupo controle.

### 2.2.1 – Identificação no caso de regressão

O viés de seleção ocorre devido às diferenças nas características observáveis e não observáveis entre os alunos que cursaram o ensino médio técnico e aqueles que fizeram o ensino médio científico. Estas características confundem os resultados observados e, neste caso, tem-se que o tratamento,  $T$ , não é independente de  $(Y_1; Y_0)$ , ou seja, não é possível saber se a mudança no salário do indivíduo  $i$  foi devido a conclusão do ensino médio ou aos seus atributos como idade, experiência, região que reside etc. Assim, torna-se necessário utilizar alguma hipótese de identificação a fim de eliminar a influência destes fatores.

Para recuperar os parâmetros de interesse utiliza-se a hipótese de identificação denominada como ignorabilidade do tratamento (ROSENBAUM e RUBIN, 1983)

**Hipótese 1 (ignorabilidade):** condicional em  $x$ ,  $T$  e  $(Y_1; Y_0)$  são independentes, isto é,  $(Y_1; Y_0) \perp T | x$

**Hipótese 2:** diante da hipótese 1 tem-se:

$$i) E[y_{1i} | x, T] = E[y_{1i} | x]$$

$$ii) E[y_{1i0} | x, T] = E[y_{1i0} | x]$$

---

<sup>7</sup> O viés de seleção pode ser visto do seguinte modo: defina a diferença média entre tratado e controle como  $G = E[Y_{11} | T = 1] - E[Y_{10} | T = 0]$ . A diferença entre G e ATT é denominada viés de seleção sendo dada por:  $viés = E[Y_{10} | T = 1] - E[Y_{10} | T = 0]$ .

Portanto a utilização de um vetor de características observáveis,  $\mathbf{x}$ , possibilita a identificação do impacto do ensino médio técnico sobre o salário dos indivíduos. Ademais, assume-se que as características não observáveis como motivação, interesse, habilidade etc. não afetam o resultado de interesse. Assim, diante das hipóteses 1 e 2 tem-se (WOOLDRIDGE, 2002, p.608):

$$ATT(\mathbf{x}) = \mathbb{E}[y_{1t} - y_{i0} | \mathbf{x}, T = 1] = \mathbb{E}[y_{1t} - y_{i0} | \mathbf{x}] = ATE(\mathbf{x}) \quad (5)$$

Para calcular o impacto utilizando a equação (5) basta estimar a seguinte regressão (WOOLDRIDGE, 2002, P.611)

$$\mathbb{E}[y | T, \mathbf{x}] = \gamma + \beta_1 T + \mathbf{h}_0(\mathbf{x})\beta_0 \quad (6)$$

em que  $\gamma, \beta_1$  e  $\beta_0$  são parâmetros a serem estimados e  $\mathbf{h}_0(\mathbf{x})$  é uma matriz de tamanho  $n \times k$  de características individuais. Repare que a diferença condicional de médias dada por (5) é exatamente o parâmetro  $\beta_1$  da equação (6). As variáveis utilizadas em  $\mathbf{h}_0(\mathbf{x})$  serão apresentadas na subseção base de dados.

Apesar de útil esta regressão fornece apenas a diferença média, condicional em  $\mathbf{x}$ , entre os dois grupos. Não necessariamente esta é a melhor maneira de formar um grupo de controle, pois compara-se todos os indivíduos tratados com todos os indivíduos não tratados e talvez estes não sejam comparáveis. Ao invés de realizar este procedimento pode-se utilizar um pareamento (*matching*) com as características contidas em  $\mathbf{x}$  para encontrar os indivíduos com maior semelhança aos tratados. A princípio todas as variáveis contidas em  $\mathbf{h}_0(\mathbf{x})$  podem ser utilizadas, no entanto, quanto maior o número de variáveis menor a possibilidade de encontrar um “*match*”. Este problema ficou conhecido na literatura de avaliação de impacto como “problema da dimensionalidade” e para contorna-lo pode-se utilizar uma medida que resume todas as características dos indivíduos. Tal medida é dada pela probabilidade condicional do indivíduo  $i$  fazer ensino médio técnico (*Propensity Score*) e a metodologia que proporciona uma melhor seleção do grupo de controle, e consequentemente o cálculo do impacto com menor viés, é denominada *Propensity Score Matching*.

### 2.2.2 – Identificação utilizando *Propensity Score Matching*

A técnica de *Propensity Score Matching* (PSM) está relacionada a construção do grupo de controle a partir das características dos indivíduos. A ideia é resumir toda a informação contida na matriz  $\mathbf{x}$  em uma única medida dada pela probabilidade do indivíduo  $i$  concluir o ensino médio técnico. Calculada a probabilidade basta parear os dois grupos utilizando diversas medidas de distância.

Segundo Rosenbaum e Rubin (1983) o *propensity score* é definido como

$$p(\mathbf{x}) \equiv \text{pr}(T = 1|\mathbf{x}) \quad (7)$$

Ao substituir a matriz  $\mathbf{x}$  pela equação (7) pode-se utilizar um único escalar,  $p(\mathbf{x})$ , para realizar o matching, pois “*exposure to treatment is random within cells defined by  $\mathbf{x}$ , it is also random within cells defined by the values of the one-dimensional variable  $p(\mathbf{x})$* ” (BECKER E ICHINO, 2002). Diante disto surge a hipótese de ignorabilidade condicional no *propensity score*

$$\text{Hipótese 3 (ignorabilidade no propensity score): } (Y_{11}, Y_{10}) \perp T | p(\mathbf{x}) \quad (8)$$

Se a hipótese 3 é satisfeita indivíduos com o mesmo *propensity score* possuem a mesma distribuição de características observáveis independente de cursarem o ensino médio técnico. Ou seja, condicional em  $p(\mathbf{x})$  o salário mensal observado é aleatório e, portanto, o grupo de tratamento e controle são idênticos.

Para finalizar a identificação no caso do *propensity score* é necessário verificar a sobreposição da distribuição dos dois grupos, sendo assim, tem-se, em conjunto com a hipótese 3, a hipótese de ignorabilidade forte (RUBIN, 1977 e ROSENBAUM E RUBIN, 1983)

$$\text{Hipótese 4 (overlap): supondo válidas as hipóteses 1-3: } 0 < \text{pr}(T = 1|\mathbf{x}) < 1 \quad \forall \mathbf{x} \quad (9)$$

A hipótese 4 diz que não é possível prever com certeza a probabilidade do indivíduo  $i$  realizar ensino médio técnico e, portanto, haverá para cada tratado uma contraparte no grupo de controle. Neste sentido percebe-se que o impacto será calculado apenas na região onde existe sobreposição (*overlap*) das duas distribuições. Apresentadas as hipóteses de identificação resta saber estimar a função  $p(\mathbf{x})$  e estimar o impacto através do *matching*, o que será abordado nas próximas seções.

## 2.3 – Estimadores utilizando *matching*

Supondo que as hipóteses 1-4 são válidas o impacto do ensino médio técnico sobre os salários dos indivíduos,  $\tau_{ATT}^{PSM}$ , pode ser estimado como

$$\tau_{ATT}^{PSM} = \mathbb{E}_{(p(\mathbf{x})|T=1)} \{ \mathbb{E}[y_{11}|T=1, p(\mathbf{x})] - \mathbb{E}[y_{10}|T=0, p(\mathbf{x})] \} \quad (10)$$



A equação (10) mostra que o estimador de PSM é a diferença média dos salários entre tratados e controle, que estão na região de suporte comum, ponderada pelo *propensity score* estimado. Apesar de útil a equação (10) é de difícil estimação, pois é pouco provável dois indivíduos terem o mesmo valor de *propensity score*. Assim, ao invés de realizar o *matching* por indivíduo utilizam-se os métodos de estratificação e vizinho mais próximo (BECKER e ICHINO, 2002).

### 2.3.1 – Estimando o ATT por estratificação

Este método é baseado na mesma estratificação utilizada ao estimar o *propensity score* e, portanto, em cada estrato as variáveis estão balanceadas, ou seja, em média os indivíduos que concluíram o ensino médio técnico possuem o mesmo *propensity score* daqueles que não fizeram ensino médio técnico.

Seja  $T$  o conjunto de indivíduos que concluíram o ensino médio técnico (tratados) e  $C$  os indivíduos que não concluíram ensino médio (controle) nesta modalidade. Os salários para cada grupo são respectivamente representados por  $y_i^T$  e  $y_i^C$ . A diferença média entre os dois grupos, levando em consideração, os estratos é dada por

$$\tau_q^S = \frac{\sum_{i \in I(q)} y_i^T}{N_q^T} - \frac{\sum_{j \in I(q)} y_j^C}{N_q^C} \quad (11)$$

em que  $I(q)$ ,  $N_q^T$ ,  $N_q^C$  são respectivamente o conjunto de indivíduos situados no estrato  $q$  e o número de pessoas no grupo de tratamento e controle nestes mesmos blocos. A equação (10) pode ser reescrita e o estimador do ATT para o caso de estratificação será calculado como (BECKER E ICHINO, 2002, p.361-364)

$$\tau^S = \sum_{q=1}^Q \tau_q^S \frac{\sum_{i \in I(q)} D_i}{\sum_{vi} D_i} \quad (12)$$

em que  $Q$  é o número de blocos e  $\frac{\sum_{i \in I(q)} D_i}{\sum_{vi} D_i}$  é um ponderador dos estratos dado pela fração das unidades tratadas dentro de cada bloco. Assim, fica claro que o ATT é dado pela simples diferença de médias, após o *matching*, entre tratado e controle ponderado pela quantidade de indivíduos tratados em cada estrato. Apesar do *matching* por estrato ser útil pode-se melhorar o resultado ao considerar que o indivíduo com maior proximidade, ao invés de todo um estrato,

será o par ideal para realizar o emparelhamento. Neste sentido surge o *matching* por vizinho mais próximo.

### 2.3.2 – Estimando o ATT por vizinho mais próximo

A ideia geral deste método consiste em realizar o *matching* entre tratados e controles utilizando o *propensity score* mais próximo. Este procedimento pode ser realizado com ou sem reposição e neste caso uma observação do grupo de controle pode ser utilizada mais de uma vez para ser pareada com o indivíduo do grupo de tratamento. Para definir o que venha a ser “próximo” é necessário utilizar algum tipo de medida como, por exemplo, a distância euclidiana. Neste sentido torna-se imprescindível algum tipo de formalização que possibilite o cálculo do ATT.

Denote como  $C(i)$  o conjunto de indivíduos no grupo de controle que foram pareados com o grupo de tratamento utilizando o *propensity score* estimado. Os indivíduos pareados serão dados por

$$C(i) = \underset{j}{\operatorname{arg\,min}} \|p_i - p_j\| \quad (13)$$

em que  $\operatorname{min}$  é a função mínimo e  $\|\cdot\|$  é a distância euclidiana. Percebe-se que o *match* é realizado utilizando a menor distância entre o *propensity score* do indivíduo  $i$  e o indivíduo  $j$ . Denote o número de indivíduos pareados com o grupo de tratamento ( $i \in T$ ) como  $N_i^C$  e o peso da

estimação no cálculo do impacto como  $w_{ij} = \frac{1}{N_i^C}$  se  $j \in C(i)$  e  $w_{ij} = 0$  caso contrário. O estimador de ATT pode ser calculado conforme a equação (14) (BECKER E ICHINO, 2002, p.361-363)

$$\tau^{NN} = \frac{1}{N^T} \sum_{i \in T} y_i^T - \frac{1}{N^T} \sum_{j \in C} w_j y_j^C \quad (14)$$

em que  $N^T$  é a quantidade de indivíduos no grupo de tratamento e  $w_j = \sum_i w_{ij}$  é o peso. Novamente fica claro que o ATT é a diferença média entre tratado e controle utilizando como ponderador a medida  $w_j$ . A vantagem de utilizar este método ao invés da estratificação é que existem outras possibilidades para parrear o indivíduo já que é considerada uma distância mínima e o mesmo indivíduo pode ser utilizado mais de uma vez. Os desvios padrão de  $\tau^S$  e  $\tau^{NN}$  foram

obtidos através de *bootstrap* com 1000 replicações<sup>8</sup>. As medidas apresentadas nesta seção, regressão e PSM, foram calculadas a nível Brasil e para o subconjunto dado pelo Estado de Minas Gerais<sup>9</sup>.

## 2.4 – Base de dados e regressões estimadas

Os microdados utilizados são advindos da Pesquisa Nacional por Amostra de Domicílios (PNAD) realizada pelo Instituto Brasileiro de Geografia e Estatística (IBGE) para o ano de 2007. Neste ano foram coletadas as informações básicas e suplementares, sendo que esta última é um amplo questionário acerca da educação profissional e aspectos complementares da educação de jovens e adultos que possibilita obter informações sobre a conclusão (ou não) do ensino médio a nível técnico. A pesquisa abrange todo o território nacional e a data de referência é de setembro de 2007.

As variáveis utilizadas, disponíveis na pesquisa básica, estão representadas no quadro 1

**Quadro 1 – Variáveis utilizadas para estimação**

Variável	Descrição
In(Renda)	Logaritmo do rendimento mensal do trabalho principal
Urbana	Um se reside na região urbana e zero caso contrário
Metropolitana	Um se reside na região metropolitana e zero caso contrário
Agrícola	Um se o trabalho principal é agrícola e zero caso contrário
Escolaridade	Anos formais completos de estudo
Experiência	Anos de experiência
Experiência2	Quadrado da variável anos de experiência
Sindicato	Um se é associado ao sindicato e zero caso contrário
Branco	Um se o indivíduo se declara branco ou amarelo e zero caso contrário
Masculino	Um se o indivíduo é do sexo masculino e zero caso contrário
Sul	Um se o indivíduo é reside na região sul e zero caso contrário
Sudeste	Um se o indivíduo é reside na região sudeste e zero caso contrário
Centro Oeste	Um se o indivíduo é reside na região centro oeste e zero caso contrário
Norte	Um se o indivíduo é reside na região norte e zero caso contrário

Fonte: elaboração própria a partir da PNAD (2007)

Nota: a variável anos de experiência foi gerada como:  $experiencia = idade - idade\ começou\ a\ trabalhar$

<sup>8</sup> A variância de ambos os estimadores também podem ser calculadas analiticamente (BECKER E ICHINO, 2002, p.364-363):

$$var(\tau^S) = \frac{1}{N^T} \left[ var(y_i^T) + \sum_{q=1}^Q \frac{N_q^T N_q^T}{N^T N_q^C} var(y_j^C) \right]$$

p.364-363):

$$var(\tau^{MN}) = \frac{1}{N^T} var(y_i^T) + \frac{1}{(N^T)^2} \sum_{j \in C} (w_j)^2 var(y_j^C)$$

. Os resultados não variaram significativamente ao utilizar  $var(\tau^S)$  e  $var(\tau^{MN})$ .

<sup>9</sup> Além dos estimadores apresentados pode-se utilizar também o matching via Kernel e vizinho mais próximo com raio. Para mais ver Imbens e Wooldridge (2009), Becker e Ichino (2002) e Caliendo e Hujer (2006).

A pesquisa suplementar não especifica diretamente se o indivíduo concluiu ou não o ensino médio técnico. No entanto, é possível gerar a variável indicadora de tratamento a partir das informações contidas na pesquisa. Antes de explicitar o procedimento é necessário definir o que venha ser educação profissional.

Segundo a PNAD (2007) a educação profissional é um conjunto de atividades para formação ou aperfeiçoamento profissional sendo que estas estão divididas em três modalidades: qualificação profissional, técnico de nível médio e curso de graduação tecnológica.

O curso de qualificação profissional é qualquer curso de formação para exercício de uma atividade profissional. Estes cursos emitem certificado de participação e são ofertados para todos os níveis de escolaridade. O objetivo é qualificar o indivíduo e não aumentar seus anos de educação formal. Como exemplos deste tipo de curso têm-se: informática, idioma, corte e costura, culinária etc.

O curso técnico de nível médio pode ser ofertado concomitante ao ensino médio ou após a conclusão. Este tipo de curso é regido por legislação própria e diretriz curricular específica, só podendo ser ministrado por escola devidamente credenciada pelo poder público. Ademais, cursos deste nível fornecem diploma de técnico.

Já o curso de graduação tecnológica tem como pré-requisito a conclusão do ensino médio e está focado nas demandas do mercado de trabalho e tecnológica. Este tipo de curso é regido por legislação própria e diretriz curricular específica, só podendo ser ministrado por escola devidamente credenciada pelo poder público. Ao formar o aluno torna-se tecnólogo.

A variável que indica se o indivíduo cursou ensino médio técnico  $T = 1$  ou não  $T = 0$  foi gerada conforme o procedimento adotado pela FIS (2010, p.12):

- 1 – Excluem-se da amostra indivíduos sem ensino médio (variável v4745).
- 2 – Excluem-se da amostra indivíduos com ensino médio completo na modalidade educação de jovens e adultos (variável v2518).
- 3 – Para quem responde que frequenta algum tipo de educação profissional (variável v2601):
  - 3.1 – Caso frequente ensino médio profissional: como ainda não terminou o curso é excluído da amostra (variável v2605).
  - 3.2 – Caso frequente qualificação profissional não é possível saber se concluiu anteriormente ensino médio profissional, portanto, é excluído da amostra (variável v2605)
  - 3.3 – Caso frequente graduação tecnológica, é possível saber se frequentou anteriormente algum outro curso de educação profissional:
    - 3.3.1 – Caso não tenha frequentado algum outro curso de educação profissional, a variável que indica se fez ensino médio profissional assume valor igual a zero (variável v2621)
    - 3.3.2 – Caso tenha frequentado algum outro curso de formação profissional, o questionário pede apenas para o respondente indicar qual o curso de educação profissional mais importante que ele frequentou. Neste caso, como não é possível ter certeza se o indivíduo frequentou ou não

curso de ensino médio profissional, o indivíduo é excluído da amostra (variável v2621).

4 – Para quem responde que não frequenta algum tipo de educação profissional (variável v2601):

4.1 – Caso não tenha frequentado anteriormente cursos de educação profissional, a variável de interesse assume valor igual a zero (variável v2603).

4.2 – Caso tenha frequentado anteriormente cursos de educação profissional (variável v2603):

4.2.1 – Se frequentou ensino médio profissional e concluiu, a variável assume valor igual a 1 (variável v2639)

4.2.2 – Se frequentou ensino médio profissional e não concluiu a variável assume valor igual a zero (variável v2639)

4.2.3 – Se frequentou qualificação profissional ou graduação tecnológica, não é possível saber se também frequentou ensino médio profissional, e, portanto, o indivíduo é excluído da amostra. (variável v2648)

Após as exclusões, a amostra foi restringida para indivíduos maiores de 14 anos e com renda do trabalho principal maior que zero. A amostra final é constituída de 45.653 indivíduos sendo que 5.035 concluíram o ensino médio técnico e 40.618 o ensino médio comum para nível Brasil. Para Minas Gerais tem-se 4.189 indivíduos sendo que 556 concluíram o ensino médio técnico e 3.633 o ensino médio comum.

### 3 – Resultados

#### 3.1 – Estatísticas descritivas

As tabelas 1 e 2 mostram as estatísticas descritivas básicas para as variáveis utilizadas: média, desvio padrão, mínimo e máximo. A tabela 1 traz as estatísticas a nível Brasil para os grupos de tratamento e controle, já a tabela 2 exibe as mesmas estatísticas, mas para com o recorte realizado para Minas Gerais

**Tabela 1 – Estatísticas descritivas para os grupos de tratamento e controle a nível Brasil em 2007**

Variáveis	Média	Desvio padrão	Mínimo	Máximo
Controle				
Renda	1430.092	2337.281	1	120.000
Urbana	0.9486433	.2207262	0	1
Metropolitana	.4844799	.4997633	0	1
Agrícola	.0360901	.1865166	0	1
Escolaridade	12.2836	1.735739	11	15

Experiência	18.29498	12.42454	0	90
Experiência2	489.0721	589.406	0	8100
Sindicato	.1938627	.3953271	0	1
Branco	.5997944	.4899441	0	1
Masculino	.4393586	.4963133	0	1
Sul	.1422502	.3493096	0	1
Sudeste	.3584864	.47956	0	1
Centro-oeste	.1154241	.3195359	0	1
Norte	.1202892	.3253022	0	1
Tratados				
Renda	1559.871	1885.322	1	45000
Urbana	.9672159	.1780861	0	1
Metropolitana	.4883995	.4999074	0	1
Agrícola	.0184707	.1346593	0	1
Escolaridade	12.23739	1.666937	11	15
Experiência	20.58043	11.82024	0	73
Experiência2	563.2461	570.7072	0	5329
Sindicato	.3100659	.4625638	0	1
Branco	.6049092	.4889113	0	1
Masculino	.4855414	.4998329	0	1
Sul	.2027572	.402087	0	1
Sudeste	.3808003	.4856244	0	1
Centro-oeste	.0985205	.2980422	0	1
Norte	.0983524	.2978156	0	1

Fonte: elaboração própria a partir dos dados da PNAD (2007)

**Tabela 2 – Estatísticas descritivas para os grupos de tratamento e controle a nível Minas Gerais em 2007**

Variáveis	Média	Desvio padrão	Mínimo	Máximo
Controle				
Renda	1320.808	2027.199	1	30.000
Urbana	.9523996	.2129401	0	1
Metropolitana	.4062684	.4911839	0	1
Agrícola	.0451473	.2076553	0	1
Escolaridade	12.27132	1.729634	11	15
Experiência	18.30913	12.55921	0	90
Experiência2	492.9177	615.8237	0	8100
Sindicato	.1672653	.3732604	0	1
Branco	.6205681	.4852932	0	1
Masculino	.4419197	.4966639	0	1
Tratados				
Renda	1359.755	1319.576	1	10.000
Urbana	.992378	.0870368	0	1
Metropolitana	.4512195	.4979945	0	1
Agrícola	.0233813	.1512473	0	1

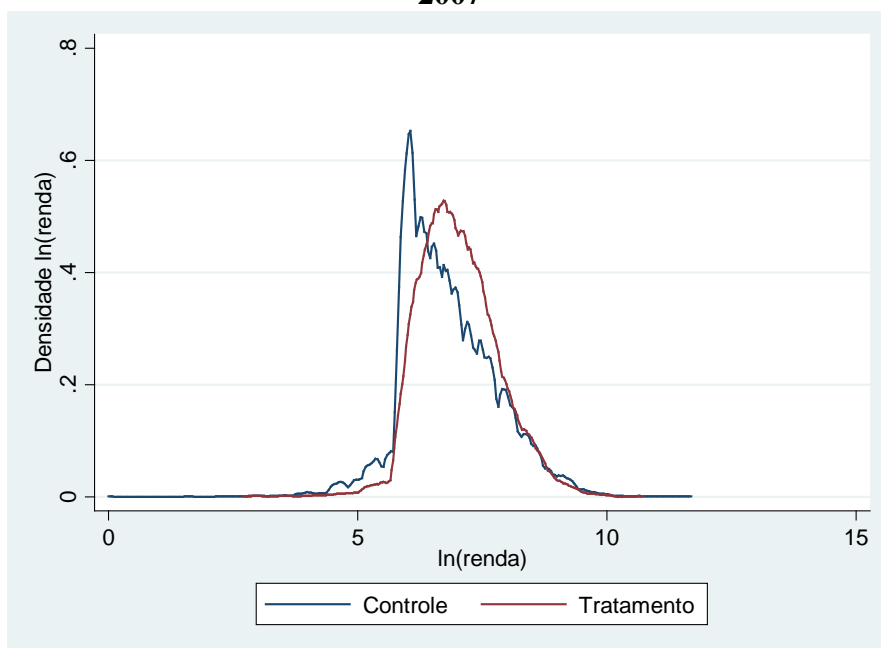
Escolaridade	12.04573	1.560303	11	15
Experiência	20.32988	12.00216	0	63
Experiência2	557.1071	570.5417	0	3969
Sindicato	.2849741	.4517924	0	1
Branco	.5838415	.4932966	0	1
Masculino	.5167683	.5001001	0	1

Fonte: elaboração própria a partir dos dados da PNAD (2007)

Analisando a tabela 1 percebe-se que a renda média do grupo de tratamento é maior que a renda média do grupo de controle revelando uma diferença de R\$ 129,78. Além da maior renda o grupo de tratamento também possui maior homogeneidade, pois seu desvio-padrão é menor quando comparado ao controle, ou seja, os indivíduos que concluíram o ensino médio estão mais próximos da média do seu grupo do que os que fizeram ensino médio comum.

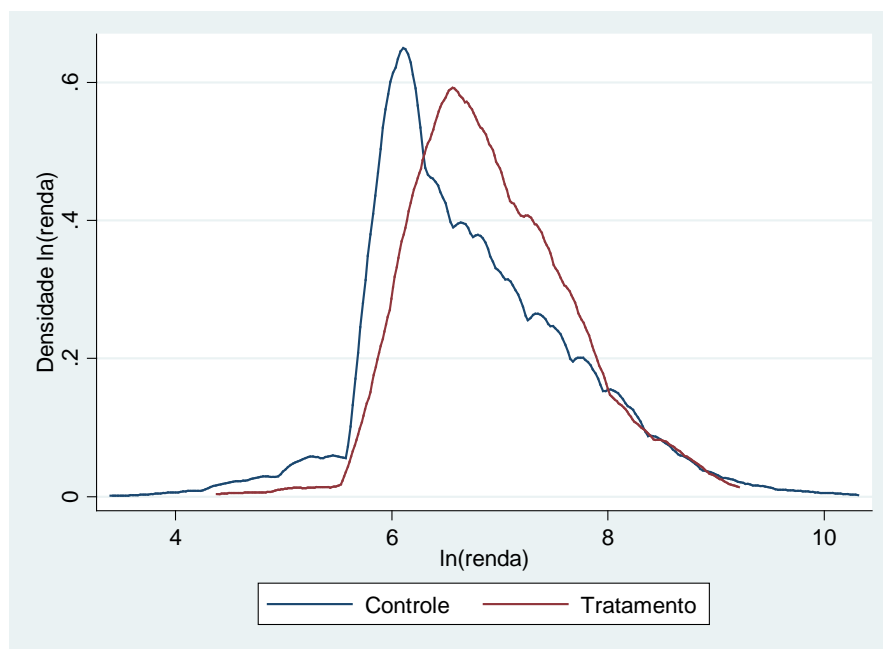
A tabela 2 mostra padrão semelhante entre os dois grupos, mas a diferença média entre tratado e controle no Estado de Minas Gerais é um pouco menor sendo apenas de R\$38,95. A variabilidade de Minas Gerais segue padrão análogo ao Brasil dado que o desvio padrão para os tratados é menor do que o grupo de controle. Estes resultados também podem ser verificados pelas Densidades de Kernel estimadas e apresentadas nos gráficos 1 e 2.

**Gráfico 1 – Densidade Kernel para os grupos de tratamento e controle a nível Brasil em 2007**



Fonte: elaboração própria a partir dos dados da PNAD (2007)

**Gráfico 2 – Densidade Kernel para os grupos de tratamento e controle a nível Minas Gerais em 2007**



Fonte: elaboração própria a partir dos dados da PNAD (2007)

Apesar da diferença positiva para o grupo de tratados, não é possível atribuir este resultado inteiramente ao ensino técnico médio, pois a entrada nesta modalidade não é realizada de forma aleatória. O que pode ter contribuído para este resultado? Percebe-se que a maioria das variáveis listadas nas tabelas 1 e 2 levam a maior renda do grupo de tratamento. Por exemplo, os indivíduos que fizeram ensino médio técnico possuem uma diferença significativa (ver teste de média nas tabelas 6-9) nas variáveis que influenciam positivamente o salário mensal: morar em região urbana, metropolitana e agrícola, maior escolaridade, maior experiência, maior proporção de brancos, indivíduos do sexo masculino etc. Logo, a diferença verificada não é completamente advinda do ensino médio técnico e outros métodos, como regressão e PSM, são necessários para levar em consideração estas diferenças.

### 3.2 – Resultados econômicos e PSM

Ao estimar a regressão obtêm-se os resultados apresentados na tabela 3. Todas as variáveis foram significativas a 1% e os sinais estão conforme o esperado na literatura.

**Tabela 3 – Regressões MQO e Probit a nível Brasil e Minas Gerais**

Variável	MQO Brasil	Probit Brasil	MQO MG	Probit MG
T	0.147*** (0.001)		0.151*** (0.001)	
Urbana	0.273*** (0.001)	0.030*** (0.000)	0.253*** (0.003)	0.112*** (0.000)
Metropolitana	0.194***	-0.004***	0.243***	0.015***



	(0.000)	(0.000)	(0.001)	(0.000)
Agrícola	-0.172***	-0.042***	-0.151***	-0.016***
	(0.001)	(0.000)	(0.004)	(0.001)
Escolaridade	0.0199***	-0.009***	0.184***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)
Experiência	0.037***	0.004***	0.038***	0.012***
	(0.000)	(0.000)	(0.000)	(0.000)
Experiência2	-0.001***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Sindicato	0.248***	0.057***	0.256***	0.081***
	(0.000)	(0.000)	(0.001)	(0.001)
Branco	0.207***	-0.012***	0.184***	-0.005***
	(0.000)	(0.000)	(0.001)	(0.001)
Masculino	0.422***	0.006***	0.516***	0.016***
	(0.000)	(0.000)	(0.001)	(0.000)
Sul	0.257***	0.031***		
	(0.000)	(0.000)		
Sudeste	0.271***	0.037***		
	(0.000)	(0.000)		
Centro Oeste	0.409***	-0.007***		
	(0.000)	(0.000)		
Norte	0.281***	0.002***		
	(0.001)	(0.000)		
Constante	2.914***		3.223***	
	(0.001)		(0.004)	
Metropolitana*Sul		0.057***		
		(0.000)		
Metropolitana*Sudeste		-0.028***		
		(0.000)		
Metropolitana*Centro		0.026***		
		(0.001)		
Metropolitana*Norte		-0.009***		
		(0.001)		
Experiência*Escolaridade		0.000***		-0.001***
		(0.000)		(0.000)
Observações	45.653	45.653	4.189	4.189

Fonte: elaboração própria a partir da PNAD (2007)

Nota: Desvio-padrão entre parênteses. Os resultados do Probit são os efeitos marginais.

\* Significativo a 10%; \*\* Significativo a 5%; \*\*\* Significativo a 1%

A diferença média entre os indivíduos que concluíram o ensino médio técnico e aqueles que fizeram o ensino médio comum é de 14,7% a nível Brasil e 15,1% a nível Minas Gerais. Assim como a diferença média é maior para o Estado de Minas outras variáveis também possuem um efeito mais intenso como, por exemplo, morar em região metropolitana e agrícola, maior escolaridade, experiência, sindicalizados e gênero masculino.

A equação de participação foi estimada através de um modelo probit e uma especificação menos parcimoniosa, incluindo termos de interação, foi utilizada de modo que a hipótese de equilíbrio (*Balancing Hypothesis*) entre tratamento e controle fosse satisfeita. Observa-se que morar em região metropolitana reduz a probabilidade de fazer em ensino médio técnico a nível Brasil, mas aumenta a chance de realizar este tipo de curso em Minas Gerais. As demais variáveis, apesar de apresentarem valores diferentes para os dois casos, possuem a mesma direção em termos de sinal.

Após estimar o *propensity score* a amostra foi dividida em blocos, por probabilidade de participação, e dentro de cada bloco foi realizado um teste de médias (Becker e Ichino, 2002, p.4). Para a amostra que compreende o território nacional foram utilizados 8 blocos e para Minas Gerais, devido a menor amostra, apenas 4. Os resultados destes testes estão apresentados nas tabelas 6-11 e fica claro que o *matching* foi realizado com sucesso.

As tabelas 4 e 5 mostram os resultados do ATT utilizando respectivamente *matching* por vizinho mais próximo e estratificação.

**Tabela 4 – ATT utilizando *matching* por vizinho mais próximo a nível Brasil e Minas Gerais**

Tratamento	Controle	ATT	Desvio-padrão	Teste t
Brasil				
5.035	20.433	0.156	0.015	10,655
Minas Gerais				
556	1397	0.157	0.044	3,607

Fonte: elaboração própria a partir dos dados da PNAD(2007)

**Tabela 5 – ATT utilizando *matching* por estratificação a nível Brasil e Minas Gerais**

Tratamento	Controle	ATT	Desvio-padrão	Teste t
Brasil				
5035	45653	0.145	0.012	12.042
Minas Gerais				
556	3.633	0,154	0,035	4,451

Fonte: elaboração própria a partir dos dados da PNAD(2007)

Os resultados mostram que a diferença média dos indivíduos que concluíram o ensino médio técnico e o ensino médio comum é de aproximadamente 15% para Brasil e Minas Gerais sendo que todos estes resultados são significativos a 1%. O padrão verificado nas regressões repete-se ao calcular o ATT, pois os resultados para Brasil e Minas Gerais são próximos, mas com um maior impacto para este último.

Este resultado está próximo aos encontrados por FIS(2010), impacto 12,5%, e Souza (2009), impacto de aproximadamente 21%. Em termos de políticas públicas percebe-se que os cursos técnicos tornam-se uma nova possibilidade de investimento, pois mostra um aumento significativo de renda para aqueles que conseguirem sua conclusão.

#### **4 – Considerações Finais**

O objetivo deste trabalho foi verificar a relação entre o ensino médio técnico e salários, mais especificamente, estimar o impacto da conclusão deste tipo de ensino sobre os salários a nível Brasil e Minas Gerais.

Esta relação já foi estimada a nível nacional e Souza(2009) e FIS (2010) encontraram um impacto positivo respectivamente de 21% e 12,5% aproximadamente. No entanto, nenhum destes autores estimou a relação para Minas Gerais.

Foi utilizada como metodologia regressões com o intuito de obter o efeito tratamento médio (ATE) e também a formação de grupos de controle utilizando *propensity score matching* a fim de calcular o efeito tratamento médio sobre os tratados (ATT).

Os resultados mostram que o impacto do ensino médio a nível Brasil e Minas Gerais são próximos a 15%, ou seja, para aqueles indivíduos que frequentaram o ensino médio técnico o salário mensal aumenta em média 15% quando comparados com os indivíduos que frequentaram apenas o ensino médio comum. Logo, políticas de incentivo a qualificação profissional podem elevar o bem estar dos indivíduos dado os ganhos privados verificados.

Este trabalho contribui para um primeiro panorama do impacto do ensino médio a nível Minas Gerais, no entanto, existem algumas oportunidades a serem desenvolvidas posteriormente: pode existir endogeneidade na variável escolaridade e os resultados podem ser viesados, logo, a metodologia de variável instrumental seria a mais correta a ser empregada; os resultados supõem que o impacto é o mesmo para todos os setores econômicos (serviços, indústria etc.) e isto não necessariamente é verdade e por fim pode-se calcular o impacto por quantil e verificar qual o efeito do ensino médio técnico em vários níveis da distribuição.

## Referências

ANGRIST, Joshua A; PISCHKE, Jorn-Steffen. **Mostly Harmless Econometrics: an empiricist's companion**. New Jersey : Princeton University Press, 2009.

BECKER, Sascha O; ICHINO, Andrea. **Estimation of Average Treatment Effects Based on Propensity Score**. The Stata Journal, Vol.2, número 4, p.358-377, 2002.

CALIENDO, Marco e HUIJER, Reinhard. **The Microeconomic Estimation of Treatment Effects: an overview**. Advances in Statistical Analysis Vol.90, número1, p199-215, 2006.

DEHEJIA, R. H., WAHBA, S. **Causal effects in nonexperimental studies: reevaluating the evaluation of training programs**. Journal of the American Statistical Association, 1999. v.94, n.448, p.1053-1062, Sept.

DEPARTAMENTO INTERSINDICAL DE ESTATÍSTICA E ESTUDOS SOCIOECONÔMICOS (DIEESE). MINISTÉRIO DO TRABALHO E EMPREGO (MTE). **Intermediação de mão-de-obra. Anuário Sistema Público de Emprego, Trabalho e Renda**. São Paulo; Brasília, 2009.

FUNDAÇÃO GETULIO VARGAS (FGV). **A educação profissional e você no mercado de trabalho**. Rio de Janeiro, 2010.

FIS, Fundação Itaú Social. **Relatório de Avaliação Econômica – Avaliação Econômica do Ensino Médio Profissional**, 2010.

GORDON, D. **Poverty and social exclusion in Britain**. York: Joseph Rowntree Foundation, 2000.

HECKMAN, J., ICHIMURA H., TODD P. **Matching as an econometric evaluation estimator: evidence from evaluating a job training program**. Review of Economic Studies, 1997. v.64(4), n.221, p.605-654, Oct.

HOLLAND, P. **Statistics and Causal Inference**. Journal of the American Statistical Association. Vol.81, p945-970, 1986.

IMBENS, Guido W; WOOLDRIDGE, Jeffrey M. **Recent Developments in the Econometrics of Program Evaluation**. Journal of Economic Literature. Vol.47 p.5-86, 2009.

IPEA. **Políticas Sociais - acompanhamento e análise** nº 13, Edição especial 2006. Cap. Trabalho e renda.

IPEA. **Políticas Sociais - acompanhamento e análise** - Cap. 6: Trabalho e renda nº 19, 2011..

KHANDKER, Shahidur R; KOOLWAL, Gayatri B; SAMAD, Hussain A. **Handbook on Impact Evaluation : quantitative methods and practices**. Whashington: The World Bank, 2010.

PEIXOTO, Patrícia E. **Do Planfor ao PNQ: uma análise comparativa sobre os Planos de Qualificação no Brasil**. Vitória, 2008. Dissertação (Mestrado em Política Social), Universidade Federal do Espírito Santo.

Pesquisa Nacional por Amostra de Domicílios (PNAD). **Notas metodológicas – Pesquisa básica e suplementar de educação profissional e aspectos complementares da educação de jovens e adultos**, 2007.

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

RUBIN, Donald B. **Estimating Causal Effects of Treatment in Randomized and Non-Randomized Studies**. Journal of Educational Psychology. Vol. 66, p.688-701, 1974.

RUBIN, Donald B. **Assignment to Treatment Group on the Basis of a Covariate**. Journal of Educational Statistics, Vol.2, número1, p.1-26, 1977.

SOUZA, Juliana de Souza. **Ensino Profissionalizante no Brasil e seu Impacto no Rendimento dos Indivíduos: uma análise de Propensity Score Matching para 2007**. Paraná: UFPR, 2009.

TEIXEIRA, Miguel Ângelo T. **Qualificação profissional e empregabilidade – uma avaliação da política pública de trabalho e renda (PPTR) frente aos desafios do mercado de trabalho**.

Dissertação (Mestrado em Ciências Sociais), Pontifícia Universidade Católica, Belo Horizonte, 2004.

WOOLDRIDGE, Jeffrey M. **Econometric Analysis of Cross Section and Panel Data**. Massachusetts: The Mit Press, 2002.

## Anexo 1

**Tabela 6 – Diferença entre tratado e controle antes do *matching* a nível Brasil**

Variável	Tratado	Controle	Valor-p
Urbana	0.967	0.949	0.000***
Metropolitana	0.488	0.485	0.564
Agrícola	0.019	0.036	0.000***
Escolaridade	12.237	12.284	0.050**
Experiência	20.580	18.295	0.000***
Experiência2	563.250	489.070	0.000***
Sindicato	0.310	0.194	0.000***
Branco	0.605	0.599	0.443
Masculino	0.486	0.439	0.000***
Sul	0.203	0.142	0.000***
Sudeste	0.381	0.359	0.001***
Centro oeste	0.099	0.116	0.000***
Norte	0.098	0.120	0.000***
Metropolitana*Sul	0.122	0.069	0.000***
Metropolitana*Sudeste	0.179	0.194	0.007**
Metropolitana*Centro	0.042	0.041	0.752
Metropolitana*Norte	0.024	0.033	0.000***
Experiência*escolaridade	254.260	229.610	0.000***

Fonte: elaboração própria a partir da PNAD (2007)

Nota: \* Significativo a 10%; \*\* Significativo a 5%; \*\*\* Significativo a 1%

**Tabela 7 – Diferença entre tratado e controle após o *matching* a nível Brasil**

Variável	Tratado	Controle	Valor-p
Urbana	0.967	0.972	0.175
Metropolitana	0.476	0.489	0.196
Agrícola	0.016	0.014	0.558
Escolaridade	12.246	12.207	0.256
Experiência	20.612	20.377	0.312
Experiência2	559.770	546.640	0.236
Sindicato	0.319	0.305	0.138
Branco	0.599	0.608	0.433
Masculino	0.517	0.514	0.731
Sul	0.208	0.212	0.568
Sudeste	0.372	0.379	0.428
Centro oeste	0.099	0.093	0.320
Norte	0.101	0.105	0.485
Metropolitana*Sul	0.124	0.129	0.412
Metropolitana*Sudeste	0.171	0.181	0.203
Metropolitana*Centro	0.039	0.038	0.754
Metropolitana*Norte	0.024	0.023	0.690
Experiência*escolaridade	254.870	251.260	0.241

Fonte: elaboração própria a partir da PNAD (2007)

Nota: \* Significativo a 10%; \*\* Significativo a 5%; \*\*\* Significativo a 1%

**Tabela 8 – Diferença entre tratado e controle antes do *matching* a nível Minas Gerais**

Variável	Tratado	Controle	Valor-p
Urbana	.96722	.94864	0.000***
Metropolitana	.4884	.48448	0.564
Agrícola	.01847	.03609	0.000***
Escolaridade	12.237	12.284	0.050**
Experiência	20.58	18.295	0.000***
Experiência2	563.25	489.07	0.000***
Sindicato	.31007	.19386	0.000***
Branco	.60491	.59979	0.443
Masculino	.48554	.43936	0.000***
Experiência*escolaridade	254.26	229.61	0.000***

Fonte: elaboração própria a partir da PNAD (2007)

**Tabela 9 – Diferença entre tratado e controle após o *matching* a nível Minas Gerais**

Variável	Tratado	Controle	Valor-p
Urbana	.99452	.99269	0.705
Metropolitana	.4479	.44059	0.808
Agrícola	.02194	.02011	0.833
Escolaridade	12.055	12.007	0.620
Experiência	20.258	20.254	0.996

Experiência <sup>2</sup>	550.01	540.23	0.766
Sindicato	.29616	.28336	0.641
Branco	.57404	.57038	0.903
Masculino	.55759	.58135	0.428
Experiência*escolaridade	244.77	244.83	0.994

Fonte: elaboração própria a partir da PNAD (2007)

**Tabela 10 – Limite inferior e quantidade de indivíduos por bloco a nível Brasil**

<b>Limite inferior pscore</b>	<b>T=0</b>	<b>T=1</b>	<b>Total</b>
0.025	1.820	65	1.885
0.050	7.808	544	8.352
0.075	10.313	1.025	11.338
0.100	9.852	1.182	11.034
0.125	4.824	803	5.627
0.150	4.520	977	5.497
0.200	1.441	419	1.860
0.300	40	20	60
<b>Total</b>	<b>40.618</b>	<b>5.035</b>	<b>45.653</b>

Fonte: elaboração própria a partir da PNAD (2007)

**Tabela 11 – Limite inferior e quantidade de indivíduos por bloco a nível Minas Gerais**

<b>Limite inferior pscore</b>	<b>T=0</b>	<b>T=1</b>	<b>Total</b>
0.009	183	3	186
0.050	1.088	101	1.189
0.100	2.081	336	2.417
0.200	281	116	397
<b>Total</b>	<b>3.633</b>	<b>556</b>	<b>4.189</b>

Fonte: elaboração própria a partir da PNAD (2007)