

Título: UMA AVALIAÇÃO DE IMPACTO E RETORNO ECONÔMICO DO PLANO SETORIAL DE QUALIFICAÇÃO (PLANSEQ)

Área 2: Economia Social | Subárea: Economia do Trabalho

Autor: Francis Carlo Petterini

Minicurrículo: Doutor em Economia pelo CAEN/UFC. Professor dos Cursos de Ciências Econômicas e Finanças da UFC – Campus de Sobral.

E-mail: petterini@ufc.br

Endereço: Rua Anahid Andrade, 471 - Praça Senador Figueira; CEP 62.011-000; Centro - Sobral – CE.

Telefax: +55 (88) 3613 2829

Título: UMA AVALIAÇÃO DE IMPACTO E RETORNO ECONÔMICO DO PLANO SETORIAL DE QUALIFICAÇÃO (PLANSEQ)

Resumo: O PLANSEQ possui um orçamento anual de cerca de R\$ 300 milhões (recursos do FAT). Entretanto, as avaliações de impacto e retorno econômico de seus programas são raras (ou inexistentes). O artigo preenche esta lacuna ao estudar o caso do programa da CDL de Fortaleza, realizado em 2009, onde o impacto estimado sob a empregabilidade dos tratados foi de 11,1%. Como não havia dados primários para a construção de um grupo de controle propôs-se o uso da PNAD/IBGE e da PED/DIEESE-SEAD, gerando outra contribuição para futuras avaliações. Por fim, a análise aponta que o investimento de R\$ 1,1 milhão deve gerar de R\$ 5,1 milhões de retorno para a sociedade. **Palavras-chave:** avaliação de impacto e retorno econômico; qualificação de mão-de-obra.

Abstract: The PLANSEQ has an annual budget around R\$ 300 million. Nevertheless, impact evaluations and economic return analysis of its programs are rare (or nonexistent). The article fills this gap by studying the case of the CDL program in Fortaleza, conducted in, where the impact estimated on the employability was of 11.1%. As there were no primary data to construct a control group was proposed the use of PNAD/IBGE and PED/DIEESE-SEAD, generating another contribution to future evaluations. The analysis also shows that the investment of R\$ 1.1 million can to generate R\$ 5.1 million for the society. **Key-words:** impact and economic return evaluation; labor qualification.

1. Introdução

Um dos principais componentes da política brasileira de emprego, trabalho e renda é o Plano Nacional de Qualificação (PNQ)¹. O plano atua em três frentes: (i) “programas territoriais” em parceria com estados, municípios e entidades sem fins lucrativos, buscando atender demandas por qualificação identificadas com base na territorialidade; (ii) “programas especiais” em parceria com entidades do movimento social e organizações não-governamentais, buscando construir metodologias e tecnologias de qualificação; e, (iii) “programas setoriais”, ou PLANSEQs, em parceria com sindicatos, empresas, movimentos sociais ou governos municipais e estaduais que precisam atender demandas emergenciais, estruturantes ou de qualificação mais específicas.

O PLANSEQ, em seu conjunto, possui um orçamento anual de cerca de R\$ 300 milhões, oriundo do Fundo de Amparo ao Trabalhador (FAT) – ver MTE (2010). Mas apesar de muitos programas já terem sido realizados com seus recursos, quase não existem (ou inexitem) avaliações de impacto e retorno econômico documentadas na literatura². Assim, o artigo procura preencher esta lacuna a partir de um estudo de caso.

O caso em questão é o seguinte: em 2008, a Câmara de Dirigentes Lojistas de Fortaleza (CDL) celebrou um convênio com Ministério do Trabalho e Emprego (MTE) para um programa na área de comércio e serviços. O objetivo foi promover a qualificação de 1,5 mil trabalhadores, ao longo de 2009, através da realização de cursos para um público alvo formado por: (i) jovens que buscam 1º emprego; e, (ii) trabalhadores sem ocupação,

¹ No Brasil, tal política é denominada por Sistema Público de Emprego, Trabalho e Renda (SPETR), que é essencialmente um conjunto de nove instrumentos: salário mínimo; seguro desemprego; intermediação da mão-de-obra, SINE; carteira de trabalho; PNQ; abono salarial; políticas de juventude; programas de geração de emprego e renda; e, programa de microcrédito produtivo orientado. Detalhes em Brasil (2010).

² A única avaliação encontrada foi Oliveira e Rios-Neto (2007), que estudam o impacto de um programa do Plano Nacional de Qualificação Profissional (PLANFOR), outro componente do PNQ, na Região Metropolitana de Belo Horizonte, em relação a duração da taxa de desemprego dos participantes.

preferencialmente mulheres chefes de famílias com mais de 40 anos. Todos deveriam possuir, ou estar cursando, o ensino médio.

Para participar, os interessados faziam suas inscrições nas principais unidades do SINE de Fortaleza. A seleção foi realizada pelos seguintes critérios: (i) ordem de inscrição; (ii) checagem de documentos comprobatórios; e, (iii) um teste simples de matemática e redação.

Os cursos, todos de 200 horas/aula, característica dos PLANSEQs, foram realizados nas áreas de: operador de caixa; auxiliar de escritório; almoxarife; assistente administrativo; repositor de mercadorias; vendedor; e, promotor de vendas.

O programa envolveu R\$ 1,1 milhão. Dinheiro público. Então é importante que se faça uma avaliação de impacto e retorno econômico deste investimento. Ou seja, verificar se o programa obteve o impacto desejado e, em caso afirmativo, se o investimento valeu a pena para a sociedade.

Assim, a principal contribuição do artigo é a apresentar as estimativas de impacto deste programa sobre a empregabilidade na área de comércio e serviços de Fortaleza. Além disso, traduz-se o impacto na empregabilidade em renda futura dos participantes e apresenta-se uma análise de seu custo/benefício social.

Outra contribuição do artigo é a discussão de como pode ser montado um grupo de controle para os tratados através da Pesquisa Nacional por Amostra de Domicílios (PNAD), do IBGE, ou da Pesquisa de Emprego Desemprego (PED), da parceria DIEESE-SEAD. Pois, como ainda é comum no Brasil, nas diretrizes originais do programa não houve uma preocupação em mensurar os contra factuais.

Além desta introdução, o trabalho está estruturado da seguinte forma: na seção 2 se apresenta os grupos de tratamento e controles; na seção 3 a estratégia de identificação; na seção 4 se apresenta uma análise descritiva da base de dados; na seção 5 se expõe os resultados estimados e uma análise de robustez; na seção 6 se discute o retorno econômico do programa; e, por fim, a seção 7 apresenta as considerações finais.

2. Tratados e Controles

Para avaliar o impacto do PLANSEQ através do programa da CDL utilizam-se três fontes de dados, uma primária (para os tratados) e duas secundárias (para os controles).

Quanto a primária, a organização do programa guardou os registros dos 1.518 alunos. Cerca de um mês após o término da última turma, em janeiro de 2010, a organização buscou contatar os alunos por telefone, perguntando se estavam ou não empregados. Conseguiram localizar 1.284 deles, e mais de 2/5 respondeu que estavam empregados e, em geral, comentavam que deviam isso ao programa³. Assim organizou-se a base de dados primária, tendo como indicador de impacto a empregabilidade dos egressos.

Infelizmente, como ainda é comum na elaboração e monitoramento de programas no Brasil, a organização não guardou os registros das inscrições em espera e dos não selecionados. Dessa forma, perdeu-se a possibilidade de construir um grupo de controle através de dados primários, forçando intrepidamente a se usar bases de dados secundárias para tal.

Além disso, como será apresentado adiante, também se deixou de coletar um conjunto de informações potencialmente importantes para estimar o impacto do programa. Mas, como argumentado por Ravallion (2001), não será por conta destas limitações que será impossível uma avaliação. É preciso usar a criatividade.

³ Até porque muitos logistas ligados a CDL procuram diretamente os egressos para contratação.

Assim, uma estratégia foi usar a PNAD. Mas isto, é claro, tem seus custos na aferição de impacto. E o mais notório deles é que a PNAD foi aplicada em setembro de 2009, ou seja, há uma diferença de quatro meses entre as respostas dos grupos de tratamento e controle. Como no mês de janeiro o comércio está aquecido, possivelmente se superestimar o impacto na empregabilidade nessa estratégia. Mas, apesar disso, se manterá essa possibilidade até o final do trabalho, na esperança de ser uma referência útil para outras avaliações do PLANSEQ.

O filtro aplicado na PNAD, naturalmente, iniciou por selecionar todas as pessoas entrevistadas na Região Metropolitana de Fortaleza (RMF). A seguir, dado que todos os tratados estavam desempregados no fim de 2008, selecionaram-se aqueles que na semana de referência... (i) tinham emprego⁴; e, (ii) estavam ininterruptamente empregados a menos de 12 meses⁵; e, (iii) este emprego era em um setor relacionado aos cursos oferecidos ao tratamento⁶. Ou então... (i) não tinham emprego⁷; e, (ii) se mantiveram desempregados ao longo do ano⁸; e, (iii) tomaram providências para sair do desemprego ao longo do ano⁹.

O passo seguinte foi filtrar idade e escolaridade de acordo com o grupo de tratamento, considerando que o tratado mais jovem possuía 15 anos, o mais velho 56 e que todos possuíam escolaridade máxima equivalente ao nível médio (completo ou não)¹⁰. Assim, ao final obteve-se um grupo de controle com 441 observações (aproximadamente 1/4 empregados).

A segunda estratégia para montar o grupo de controle foi usar a PED/DIEESE-SEAD aplicada pelo Instituto de Desenvolvimento do Trabalho (IDT), órgão ligado a Secretaria do Trabalho e Empreendedorismo do Estado do Ceará. Ao contrário da PNAD, pela PED é possível se fazer um desenho de controle para janeiro de 2010 e específico para Fortaleza – mais do que para a RMF como um todo – então parece ser uma base de dados mais adequada para o caso.

Devido às diferenças de questionário, os filtros seguintes diferem entre PNAD e PED. Nesse sentido, na PED inicia-se selecionando a idade e a escolaridade adequada ao grupo de tratamento, o município de residência e o pertencimento a População Economicamente Ativa (PEA)¹¹. E, em seguida, se a pessoa ocupada trabalha no ramo de comércio ou serviço ininterruptamente a menos de 12 meses; e se a pessoa desocupada tomou providências para sair do desemprego ao longo do ano¹². Ao final obteve-se um grupo de controle com 551 observações (aproximadamente 1/3 empregados).

3. Estratégia de identificação

Para estimar o impacto sobre a empregabilidade dos indivíduos (que se inscreveram no programa) é preciso saber o que teria acontecido com eles caso não tivessem participado do programa. Isto leva ao conhecido problema do contrafactual não observado, que será tratado com as seguintes definições: Y_0 sendo um indicador binário da situação (pós-programa) para o indivíduo, caso ele não tivesse se inscrito no programa, onde $Y_0 = 1$ aponta “empregado” e $Y_0 = 0$ “desempregado”; Y_1 sendo um indicador binário da situação (pós-programa) caso ele

⁴ Em *syntax* do *software* SPSS... (V9001 = 1 | V9002 = 2).

⁵ (V9062 = 4 & (V9611 = 0 & V9612 < 12)).

⁶ (V9907 >= 53010 & V9907 <= 53113).

⁷ V9001 = 3.

⁸ V9067 = 3.

⁹ (V9115 = 1 | V9116 = 2 | V9117 = 1 | V9118 = 2).

¹⁰ ((V0602 = 2 & (V6003 = 2 | V6003 = 4)) | (V0602 = 4 & (V6007 = 2 | V6007 = 3 | V6007 = 5 | V6007 = 7))).

¹¹ (C020 >= 15 & C020 <= 56) & C102 = 2 & A020 = 440 & (Q170 = 1 | Q170 = 2).

¹² (Q170 = 1 & Q180 = 1 & (Q260 > 400 & Q260 < 500)) | (Q170 = 0 & (Q162 >= 12 & Q162 <= 1)).

tivesse se inscrito no programa, onde $Y_1 = 1$ aponta “empregado” e $Y_1 = 0$ “desempregado”; e, $T = 1$ apontando que o indivíduo se inscreveu no programa e $T = 0$ o contrário. É possível observar $Y_1|T = 1$ e $Y_0|T = 0$, mas nunca observar $Y_1|T = 0$ e $Y_0|T = 1$. Este é o *problema básico da inferência causal*, que essencialmente é um problema de dados faltantes.

A ideia é mensurar como o tratamento pode aumentar as chances de emprego de um trabalhador desempregado. Assim, é possível escrever essa relação como¹³:

$$[1] \quad \Delta = E(Y_1 - Y_0|T = 1) = \Pr(Y_1 = 1|T = 1) - \Pr(Y_0 = 1|T = 1)$$

Não existem dados para se estimar $\Pr(Y_0 = 1|T = 1)$, mas existem dados para se estimar $\Pr(Y_0 = 1|T = 0)$. Então, com uma simples manipulação algébrica, Δ pode ser reescrito como:

$$\Delta = \Pr(Y_1 = 1|T = 1) - \Pr(Y_0 = 1|T = 0) + \underbrace{\left(\Pr(Y_0 = 1|T = 0) - \Pr(Y_0 = 1|T = 1) \right)}_{\text{viés de seleção}}$$

Note que sempre que $\Pr(Y_0 = 1|T = 0) \neq \Pr(Y_0 = 1|T = 1)$ existirá um viés ao se utilizar $\Pr(Y_0 = 1|T = 0)$ no lugar de $\Pr(Y_0 = 1|T = 1)$. Este é o “viés de seleção”, que surge quando o grupo de controle é muito diferente do tratamento.

O controle pode ser muito diferente em termos de variáveis potencialmente observáveis (cor, local onde mora, se estuda na rede pública ou privada, renda familiar etc.), mas que não foram tabuladas para o grupo de tratamento, como é o presente caso, ou em termos de variáveis realmente não observáveis (força de vontade, motivação etc.).

Como estratégia para estudar esse potencial viés (e então mitigá-lo), se usará o protocolo exposto em Becker e Ichino (2002) e Khandker, Koolwal e Samad (2010). Para tanto, admite-se que cada indivíduo possui um conjunto de características observáveis, definido por X , e se define o “escore de propensão” (ao tratamento) por $\Pr(T = 1|X) = p(X)$. Assim, com uma iteração de expectativas observa-se que:

$$[2] \quad \begin{aligned} \Delta = E(Y_1 - Y_0|T = 1) &\Rightarrow \Delta_{p(X)} = E_{p(X)}[E(Y_1 - Y_0|T = 1, p(X))] \Rightarrow \\ \Delta_{p(X)|T=1} = E_{p(X)|T=1}[E(Y_1|T = 1, p(X)) - E(Y_0|T = 0, p(X))|T = 1] \end{aligned}$$

O ponto chave aqui é poder usar $E(Y_0|T = 0, p(X))$, lado esquerdo de [2], para se estimar o efeito do tratamento (equação [1]). Mas sob quais circunstâncias é possível usar $E(Y_0|T = 0, p(X))$ e garantir que o viés de seleção foi mitigado?

Rosenbaum e Rubin (1983) mostram que se a exposição ao tratamento é aleatória uma vez que se controla pelo conjunto das características observáveis X , a chamada *balancing hypothesis*, ela também será aleatória ao se controlar por $p(X)$. Este resultado é importante por questões de dimensionalidade, possibilitando se trabalhar com o escalar $p(X)$ ao invés de todo X .

Se *balancing hypothesis* puder ser aceita, então as observações com escore de propensão semelhante têm semelhante distribuição de características observáveis (e não observáveis), independente do fato de terem sido tratadas ou não. Como resultado, tem-se que $\Delta = \Delta_{p(X)|T=1}$ e que é possível mitigar o viés.

¹³ Note que Δ é o chamado efeito tratamento médio sobre os tratados. Além disso, dado que todos os tratados e controles estavam desempregados no período inicial, Δ também poderia ser interpretado como uma medida de diferenças em diferenças. Detalhes, por exemplo, em Angrist e Pischke (2008).

Formalmente, as seguintes hipóteses (de identificação) são necessárias para garantir este resultado: *ignorabilidade do tratamento*: $(Y_0, Y_1) \perp T|X$, isto é, os resultados (potenciais) são independentes do tratamento condicionado nas características observáveis; e, *suporte comum*: $0 < p(X) < 1$ para $\forall X \in \Gamma$, onde Γ é suporte da distribuição de X . Isto garante comparabilidade entre os dois grupos, ou seja, não há valor de X para o qual se possa dizer com certeza a qual grupo um indivíduo pertence.

Na prática, a estratégia de identificação via escore de propensão (e pareamento) segue os seguintes passos: (i) estima-se $p(X)$ por um modelo econométrico padrão (e.g. Probit); (ii) tendo a estimativa de $p(X)$ em mãos, fraciona-se o intervalo unitário em outros intervalos por algum critério; (iii) seleciona-se a área de suporte comum, isto é, define-se uma região onde $p(X) \in [\underline{p}, \bar{p}]$, de tal forma a excluir-se da amostra os indivíduos não conectadas ao suporte comum (detalhes em Heckman, Ichimura e Todd, 1997); (iv) em cada intervalo dentro do suporte comum, testa-se se o escore de propensão médio não difere entre tratados e controles. Caso escore de propensão médio seja diferente em algum intervalo, faz-se uma nova participação e repete-se os dois passos anteriores; (v) tendo encontrado K intervalos sob o suporte comum, onde as estimativas de $p(X)$ não difiram, testa-se se a média das características X dentro de cada intervalo não difere entre tratados e controles (*balancing hypothesis*); (vi) se o teste do passo anterior falhar é preciso voltar para o primeiro passo e tentar uma nova especificação para o escore de propensão, caso contrário, segue-se com o protocolo; (vii) uma inspeção visual pode ser feita através da comparação de densidade da estimativa de $p(X)$ sob o suporte comum, separando-se tratados e controles. É de se esperar alguma sobreposição dos histogramas, indicando semelhança entre os grupos de tratamento e controle; e, (viii) com as unidades conectadas ao suporte comum (e pareadas), pode-se estimar Δ por um modelo econométrico adequadamente especificado ou através de estimadores baseados no escore de propensão.

O sucesso do protocolo em mitigar o viés depende de não haver variáveis não-observadas que afetem os resultados (de participar do tratamento) de forma diferenciada nos grupos de tratamento e controle. A estratégia de identificação apontará se os grupos de tratamento e controle são suficientemente parecidos para mitigar o viés. Caso afirmativo, pode-se proceder a uma estimativa de impacto sobre a empregabilidade seguindo, por exemplo, a ideia exposta em Card e Sullivan (1988).

Nesse sentido, assume-se que a empregabilidade (Y), que assume o valor 1 quando o indivíduo está empregado e 0 no caso contrário, é uma função indicadora de uma variável latente $Y^* = \alpha T + X\beta + u$. Onde α é um parâmetro, T é uma variável binária que indica se o indivíduo passou pelo programa, X é um vetor linha de características observáveis, β é um vetor coluna de parâmetros e u é um índice de características não observáveis. Se ocorrer $Y^* > 0$ o indivíduo está empregado, caso contrário não. Assim, formalmente escreve-se: $Y = 1[\alpha T + X\beta + u > 0]$.

O objetivo é estimar Δ , então, note que $Y = Y_1 \cdot T + Y_0 \cdot (1 - T)$ e que, portanto, a questão é estimar $\Pr(Y = 1|T, X) = \Pr(u > -\alpha T - X\beta|T, X)$. Sob a hipótese de que a função de densidade de probabilidade de u é simétrica, e definindo $F(u)$ como sua função cumulativa de probabilidade, pode-se escrever $\Pr(Y = 1|T, X) = 1 - F(-\alpha T - X\beta) = F(\alpha T + X\beta)$.

Assumindo uma forma para $F(u)$ (e.g. a de uma normal padrão), é possível estimar o impacto Δ para cada indivíduo com um conjunto particular de características $X = X_c$ (e.g. homem, entre 20 e 25 anos, com ensino médio completo) tal que:

$$[3] \quad \Delta_c = F(\alpha + X_c\beta) - F(X_c\beta)$$

E então estimar o impacto do programa sobre a empregabilidade através de:

$$[4] \quad \Delta = \sum_c \Pr(X = X_c) \cdot \Delta_c$$

Em suma, a estratégia de identificação é encadeada na seguinte sequência: a organização do programa não se preocupou em elaborar um grupo de controle; recorre-se então a PNAD e a PED para isso; como existe o risco dos grupos de tratamento e controle serem muito diferentes, gerando viés na estimação de Δ , se recorrerá a técnica do pareamento no escore de propensão; dentre as observações pareadas (sob o suporte comum $p(X) \in [\underline{p}, \bar{p}]$) se estimará $\Pr(Y = 1|T, X)$; se definirá um conjunto de características dos indivíduos pertencentes a base de dados, de forma a construir um número finito de “tipos de pessoas” para computar $\Pr(X = X_c)$; se estimará o efeito do tratamento na empregabilidade para cada tipo de pessoa – equação [3]; se estimará Δ através de uma média ponderada pela frequência em que se observa estes tipos de pessoas – equação [4]; e, por fim, se analisará a robustez dos resultados estimados (discutido a seguir).

A estratégia de estimar o impacto usando a equação [4] possui uma vantagem e uma desvantagem. A vantagem é que construir um número finito de “tipos de pessoas”, e conduzir a avaliação desta forma, facilitará a análise de retorno econômico que será apresentada adiante. A desvantagem é que, ao arbitrar a construção destes “tipos de pessoas”, pode-se estar dando mais (ou menos) peso para alguns tipos, superestimando (ou subestimando) o impacto.

Então é necessário que se faça uma análise de robustez sobre o resultado encontrado pela estimação de [4]. Onde se propõem usar dois outros estimadores de efeito tratamento médio sobre os tratados, baseados no escore de propensão (descritos em Becker e Ichino, 2002): (i) *Nearest Neighbor Matching*; e, (ii) *Radius Matching*. É de se esperar estimativas semelhantes a encontrada através da equação [4].

4. Dados e análise descritiva

As inscrições para o programa ocorreram no início de 2009 nas principais unidades do SINE de Fortaleza. Os relatórios apontam que a seleção dos alunos acabou focalizando, prioritariamente, na ordem de chegada, e depois nos jovens até 24 anos, em busca do 1º emprego, que nunca tiveram a carteira de trabalho assinada, e que possuíam no mínimo ensino médio incompleto, e, por último, nos desempregados que já possuíam alguma experiência no mercado formal. Os cursos foram oferecidos ao longo de 2009 para várias turmas.

A Tabela 1, abaixo, apresenta média e desvio padrão das variáveis observadas por grupos de tratamento e controles. Infelizmente, a organização do programa deixou de coletar uma série de variáveis potencialmente importantes para a análise de impacto. Na verdade, se limitou a coletar apenas as reportadas abaixo.

Note, primeiramente, que 41,5% dos indivíduos do grupo de tratamento estavam empregados. Por outro lado, 24,0 e 32,7% dos grupos de controle da PNAD e da PED, respectivamente, estavam empregados. Apesar dessas aparentes diferenças na empregabilidade, pelos argumentos já apresentados, ainda não é possível atribuir sucesso ao programa. É preciso ter alguma segurança de que os grupos são comparáveis.

A proporção de homens é relativamente semelhante entre os três grupos, o que já não se pode afirmar para a proporção de concludentes do ensino médio. Quanto a esta variável, condição necessária para a inscrição no programa, 51,9% dos indivíduos do grupo de tratamento havia completado os estudos em questão, em contraste aos 78,0 e 73,7% dos grupos de controle da PNAD e da PED, respectivamente.

Possivelmente isso se deve ao fato que os principais meios de divulgação do programa foram a Internet (página da CDL, lista de e-mails e redes sociais) e o rádio. Pois havia um maior interesse pelo público jovem (conseqüentemente, ainda na escola).

Tabela 1 – Média e Desvio Padrão das variáveis coletadas por grupos

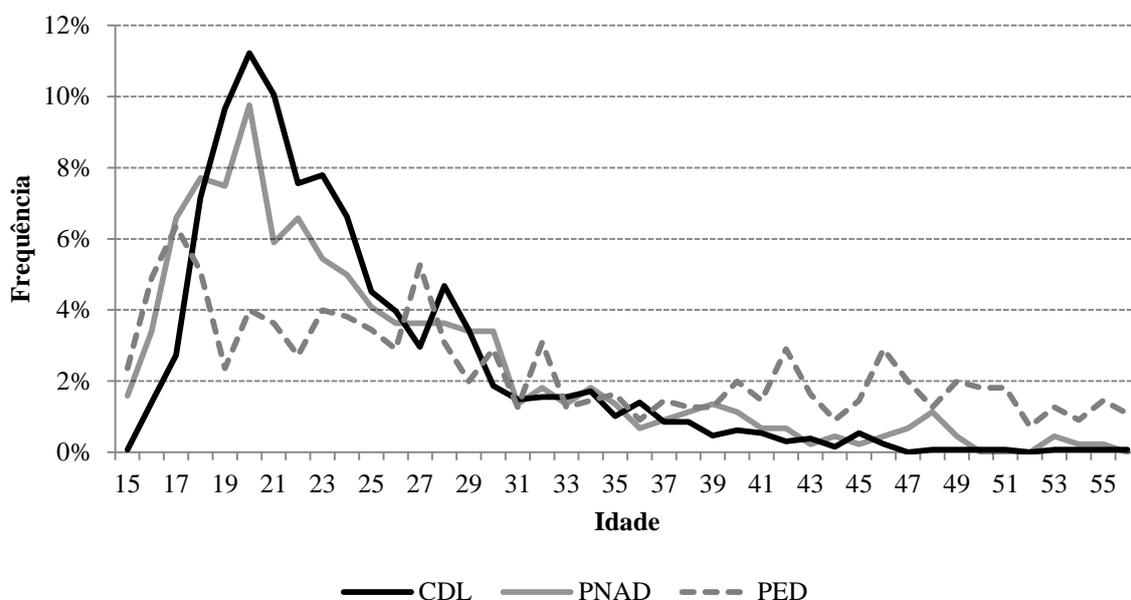
Variável	Tratamento (dados primários da CDL, 1.284 observações)		Controle 1 (PNAD, 441 observações)		Controle 2 (PED, 551 observações)	
	Média	Desvio	Média	Desvio	Média	Desvio
Empregado	41,5%	49,3%	24,0%	42,8%	32,7%	46,9%
Homem	33,3%	47,1%	39,2%	48,9%	39,6%	48,9%
Ensino Médio Completo	51,9%	50,0%	78,0%	41,5%	73,7%	44,1%
Idade (em anos)	24,30	6,31	25,05	8,00	30,65	11,87

Fonte: CDL/Fortaleza, PNAD e PED. Elaboração Própria.

A variável que deve ser observada com mais atenção aqui é a idade, particularmente por conta da construção dos “tipos de pessoas” para estimar Δ_c – equação [3]. Salienta-se que o indivíduo mais jovem da base de dados possui 15 anos, o mais velho 56. Logo, existem 42 “tipos” de idade. Os quais, associados aos 4 tipos de pessoas envolvendo as variáveis sexo e ensino médio¹⁴, geram 168 potenciais tipos. Um número pouco parcimonioso, mas que pode ser reduzido.

Considerando a densidade total das idades (sem separação entre tratamento e controles), o primeiro quartil é limitado por 19 anos (incluindo esse valor), o segundo por 22 anos (incluindo esse valor) e o terceiro por 28 anos (incluindo esse valor). Assim, usando 4 intervalos para a variável idade, definidos pelos quartis, optou-se por trabalhar com 16 “tipos de indivíduos” para estimar Δ_c . Abaixo, o Gráfico 1 ilustra a frequência das idades por grupos de tratamento e controles.

Gráfico 1 – Frequência das idades por grupos de tratamento e controles



Fonte: CDL/Fortaleza, PNAD e PED. Elaboração Própria.

¹⁴ Homem com ensino médio completo, homem com ensino médio incompleto, mulher com ensino médio completo e mulher com ensino médio incompleto.

5. Resultados estimados

A questão chave da estimativa do impacto do programa na empregabilidade é mitigar um possível viés de seleção, originado de diferenças significativas entre indivíduos dos grupos de tratamento e controles. Como mencionado anteriormente, a estratégia que será usada aqui para testar se os grupos são comparáveis é seguir o protocolo exposto em Becker e Ichino (2002) e Khandker, Koolwal e Samad (2010).

Para tanto, inicia-se estimando um modelo Probit para o escore de propensão, $p(X)$, usando como variáveis explicativas para a participação no programa, além da constante, *dummies* para: homem; ensino médio completo; e, idade estar nos intervalos (19,22], (22,28] e (28,56]. Infelizmente, estas são todas as variáveis observadas no grupo de tratamento, então, potencialmente existem variáveis importantes para a aferição do impacto que estão sendo omitidas. Mas se fará o melhor possível com os dados disponíveis. Os resultados estão expostos a seguir¹⁵, na Tabela 2, usando como controle os dados da PNAD e depois da PED.

Tabela 2 – Valores estimados para um modelo Probit para $p(X)$

Variável explicativa	Controle 1: PNAD			Controle 2: PED		
	Parâmetro	Desvio	Estatística z	Parâmetro	Desvio	Estatística z
Homem	-0,15	0,07	-2,17	0,07	-2,74	0.01
Ensino médio completo	-1,08	0,09	-11,58	-0,53	0,08	-6.42
Idade (19,22]	0,68	0,11	6,35	0,76	0,11	6.96
Idade (22,28]	0,90	0,11	7,98	0,48	0,10	4.63
Idade (28,56]	0,73	0,12	6,10	-0,17	0,11	-1.57
Constante	0,82	0,08	10,40	0,70	0,08	9.25
Pseudo R ²	0,0892			0,0956		

A discussão sobre os parâmetros estimados para o escore de propensão não é essencial, mas sim o protocolo que segue a estimação. Nesse sentido, observa-se que o suporte comum ficou definido pelos intervalos $[p, \bar{p}] = [0,3408; 0,9573]$ para o controle via PNAD e $[p, \bar{p}] = [0,4266; 0,9271]$ para o controle via PED. Como todos os indivíduos, em ambos os controles, tinham seu escore de propensão estimado conectado ao suporte comum, não se excluiu nenhuma observação da amostra nos passos seguintes.

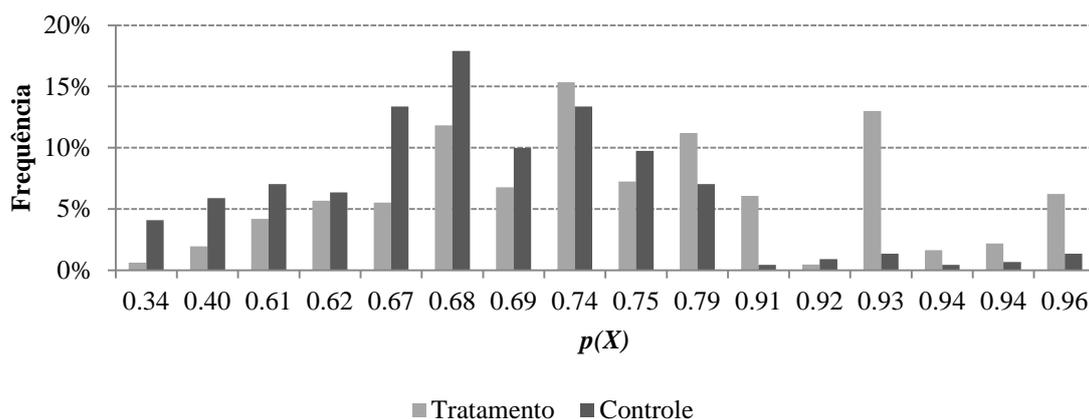
Assim, fraciona-se o intervalo unitário em K blocos de igual dimensão, 6 para o controle PNAD e 8 para o controle PED, de forma a garantir que a média do escore de propensão estimado não é diferente entre os grupos de tratamento e controle (dentro de cada bloco). Com a especificação reportada na Tabela 2 este passo é satisfeito.

Seguindo o protocolo, dentro de cada um dos blocos testou-se a diferença de média entre tratados e controles em percentual de homens, de indivíduos com ensino médio completo e de indivíduos nas faixas etárias trabalhadas. Não houve diferenças significativas, apontando que os grupos de tratamento e controles não diferem significativamente nas características observadas.

A seguir, faz-se uma inspeção visual através da comparação de densidade da estimativa de $p(X)$ sob o suporte comum, separando-se tratados e controles. O Gráfico 2 apresenta o exercício para o controle com os dados da PNAD, e o Gráfico 3 para a PED.

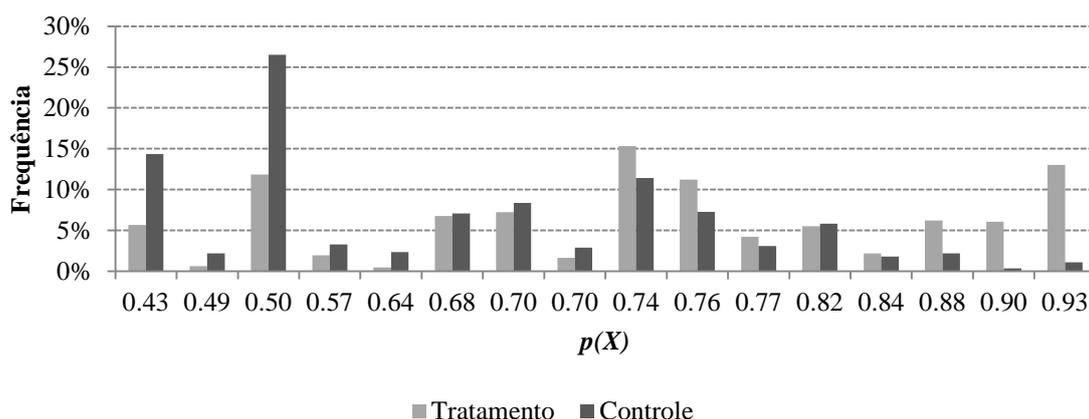
¹⁵ Usou-se a rotina *pscore* do software STATA.

Gráfico 2 – Densidade de $p(X)$ estimado para os grupos de tratamento e controle PNAD



Desconsiderando alguns casos, particularmente nas extremidades, nota-se uma boa sobreposição dos histogramas, indicando que existe uma razoável semelhança entre os grupos de tratamento e controles. O mais relevante aqui é que não há valor de X para o qual se possa dizer com certeza a qual grupo um indivíduo pertence. E que, portanto, os grupos podem ser comparados, acreditando-se que o “viés de seleção” está sendo mitigado.

Gráfico 3 – Densidade de $p(X)$ estimado para os grupos de tratamento e controle PED



Ainda nos Gráficos acima, note que os histogramas apresentam 16 picos. É claro, isso não ocorre por acaso, conforme discutido anteriormente, tratam-se dos com 16 “tipos de indivíduos” que podem ser formados com as *dummies* usadas como variáveis explicativas.

Acreditando que os grupos de tratamento e controles são suficientemente parecidos, o passo seguinte é estimar a equação [4], para $\Pr(Y = 1|T, X) = F(\alpha T + X\beta)$. Para tanto, se utilizará um modelo Probit para a empregabilidade, usando como variáveis explicativas, além da constante, *dummies* para: participação no tratamento; homem; ensino médio completo; e, idade nos intervalos (19,22], (22,28] e (28,56]. Os resultados estão expostos abaixo, na Tabela 3. Os baixos valores dos Pseudo R^2 se devem as poucas variáveis explicativas.

Primeiro observe que, para o controle via PNAD, com exceção das faixas de idade inferiores a 28 anos, os demais parâmetros estimados mostraram-se significativos ao nível de confiança padrão de 95%. Os resultados também apontam que homens e indivíduos que completaram o ensino médio também têm mais chances de estarem empregado. Além disso, o fato de um indivíduo ter passado pelo programa aumenta as suas chances de emprego, em magnitude maior do que das outras variáveis explicativas.

Para o controle via PED os resultados são semelhantes, com exceção para as *dummies* das idades. Mas certamente o que mais chama a atenção na comparação dos resultados é a

magnitude do parâmetro estimado para a *dummy* de tratamento, 0,56 com a PNAD e 0,30 com a PED. O que já era esperado por conta da diferença de data na aplicação dos questionários. E atribui-se essa diferença a questões de sazonalidade.

Tabela 3 – Valores estimados para um modelo Probit para $\Pr(Y = 1|T, X)$

Variável explicativa	Controle 1: PNAD			Controle 2: PED		
	Parâmetro	Desvio	Estatística z	Parâmetro	Desvio	Estatística z
Tratamento	0,56	0,08	7,26	0,30	0,07	4,42
Homem	0,28	0,07	4,31	0,42	0,06	6,61
Ensino médio completo	0,19	0,08	2,49	0,14	0,07	1,93
Idade (19,22]	-0,10	0,09	-1,11	0,17	0,09	1,84
Idade (22,28]	-0,14	0,10	-1,44	0,22	0,10	2,24
Idade (28,56]	-0,25	0,11	-2,21	0,19	0,10	1,83
Constante	-0,85	0,10	-8,91	-0,89	0,09	-10,20
Pseudo R ²	0,0317			0,0301		

Tendo os resultados estimados para $\Pr(Y = 1|T, X)$, a Tabela 4, abaixo, usando como controle os dados da PNAD, apresenta os resultados estimados para o impacto na empregabilidade dos participantes do programa por “tipo de indivíduo” (Δ_c) e o geral (Δ), além de apontar a quantidade e a variedade de tipos na amostra.

Tabela 4 – Resultados estimados para o impacto na empregabilidade dos participantes por “tipo de indivíduo” e geral – Δ_c e Δ – equações [3] e [4] – controle via PNAD

Sexo	Ensino Médio	Faixa Etária	Tratamento		Controle		Δ_c	$\Pr(X = X_c)$
			$\Pr(Y = 1 T, X)$	Número de observações	$\Pr(Y = 1 T, X)$	Número de observações		
Mulher	Incompleto	[15,19]	0,386	144	0,197	31	0,189	0,101
		(19,22]	0,347	167	0,169	6	0,178	0,100
		(22,28]	0,332	80	0,159	6	0,173	0,050
		(28,56]	0,296	21	0,136	2	0,160	0,013
	Completo	[15,19]	0,462	25	0,255	26	0,207	0,030
		(19,22]	0,422	71	0,223	59	0,198	0,075
		(22,28]	0,405	197	0,211	59	0,194	0,148
		(28,56]	0,366	152	0,182	79	0,184	0,134
Homem	Incompleto	[15,19]	0,497	93	0,284	43	0,213	0,079
		(19,22]	0,456	78	0,250	2	0,206	0,046
		(22,28]	0,440	28	0,237	3	0,202	0,018
		(28,56]	0,399	6	0,206	4	0,193	0,006
	Completo	[15,19]	0,574	8	0,353	18	0,221	0,015
		(19,22]	0,534	54	0,316	31	0,218	0,049
		(22,28]	0,517	87	0,301	44	0,216	0,076
		(28,56]	0,476	73	0,266	28	0,210	0,059
$\Delta = \sum \Pr(X = X_c) \cdot \Delta_c$							0,1961	

Observe que 14,8% da amostra é formada por mulheres com ensino médio completo e na faixa de 22 a 28 anos (é o tipo que mais aparece), nas quais o programa aumentou em 19,4% suas chances de estarem empregadas (caso tenham participado). E o tipo que menos aparece na amostra são homens com ensino médio incompleto e mais de 28 anos, nos quais o programa aumentou em 19,3% suas chances de estarem empregados.

Para todos os tipos, as chances de estarem empregados são maiores para aqueles que participaram do programa. E essa diferença oscila entre 19% e 20%, sem grandes variações, de forma que o impacto médio na empregabilidade daqueles que participaram do programa é estimado em 19,6%.

Já a Tabela 5, abaixo, apresenta os resultados estimados para o impacto na empregabilidade dos participantes por “tipo de indivíduo” (Δ_c) e o geral (Δ) usando como controle os dados da PED, além de apontar a quantidade e a variedade de tipos na amostra.

Tabela 5 – Resultados estimados para o impacto na empregabilidade dos participantes por “tipo de indivíduo” e geral – Δ_c e Δ – equações [3] e [4] – controle: PED

Sexo	Ensino Médio	Faixa Etária	Tratamento		Controle		Δ_c	$\Pr(X = X_c)$
			$\Pr(Y = 1 T, X)$	Número de observações	$\Pr(Y = 1 T, X)$	Número de observações		
Mulher	Incompleto	[15,19]	0,280	144	0,187	40	0,092	0,100
		(19,22]	0,340	167	0,237	6	0,103	0,094
		(22,28]	0,357	80	0,251	12	0,106	0,050
		(28,56]	0,346	21	0,241	16	0,104	0,020
	Completo	[15,19]	0,329	25	0,228	18	0,102	0,023
		(19,22]	0,394	71	0,283	32	0,111	0,056
		(22,28]	0,411	197	0,299	63	0,113	0,142
		(28,56]	0,400	152	0,288	146	0,111	0,162
Homem	Incompleto	[15,19]	0,433	93	0,318	46	0,115	0,076
		(19,22]	0,502	78	0,382	2	0,120	0,044
		(22,28]	0,520	28	0,400	10	0,120	0,021
		(28,56]	0,508	6	0,388	13	0,120	0,010
	Completo	[15,19]	0,490	8	0,371	12	0,119	0,011
		(19,22]	0,558	54	0,437	17	0,121	0,039
		(22,28]	0,576	87	0,455	39	0,121	0,069
		(28,56]	0,564	73	0,443	79	0,121	0,083
$\Delta = \sum \Pr(X = X_c) \cdot \Delta_c$							0,1111	

Assim como no caso da PNAD, usando a PED, para todos os tipos as chances de estarem empregados são maiores para aqueles que participaram do programa. E essa diferença oscila entre 9% e 12% entre os grupos, de forma que o impacto médio na empregabilidade daqueles que participaram do programa é estimado em 11,1%. Menor que no caso da PNAD, em grande parte por conta do controle da questão sazonal e do foco específico em Fortaleza.

Mas como já argumentado, ao arbitrar a construção dos “tipos de pessoas”, pode-se estar superestimando (ou subestimando) o impacto Δ – devido a ponderação em $\Pr(X = X_c)$. Então é interessante que se faça uma análise de robustez sobre o resultado apontados nas Tabela 4 e 5. Para tanto, a Tabela 6, abaixo, mostra outras duas estimativas de efeito tratamento médio sobre os tratados, baseados no escore de propensão.

Note que a técnica *Nearest Neighbor* apresenta um Δ estimado de 0,210, pouco acima do 0,196 apontado na Tabela 4 para o caso do controle via PNAD. E de 0,125 para o caso do controle via PED, próximo ao valor de 0,111 reportado anteriormente. De forma semelhante, a técnica *Radius Matching* apresenta um Δ estimado de 0,199 e 0,109 para os casos da PNAD e PED, respectivamente, também próximas dos valores reportados anteriormente.

Tabela 6 – Resultados estimados de outros estimadores de efeito tratamento médio sobre os tratados, baseados no escore de propensão

Controle e Técnica		Δ estimado	Desvio Padrão	Estatística <i>t</i>
PNAD	<i>Nearest Neighbor</i>	0,210	0,038	5,454
	<i>Radius Matching</i>	0,199	0,026	7,720
PED	<i>Nearest Neighbor</i>	0,125	0,040	3,131
	<i>Radius Matching</i>	0,109	0,026	4,213

Sendo todas estimativas da Tabela 6 estatisticamente significantes aos níveis de confiança padrão, acredita-se que os resultados apontados nas Tabelas 4 e 5 são estimativas apropriadas para o impacto do programa. E que a diferença das magnitudes das estimativas entre os grupos de controle utilizados se deve essencialmente as questões de sazonalidade.

Assim, os impactos estimados são um indicativo de sucesso do programa. Todavia, é preciso lembrar que o programa custou R\$ 1,1 milhão em recursos públicos. Então, a questão seguinte é: esses impactos seriam grandes o suficiente para justificar os custos do programa?

6. Retorno econômico

A exemplo de Vasconcellos, Biondi e Menezes-Filho (2009), e outras avaliações do gênero, este trabalho também conduzirá uma avaliação de retorno econômico para o programa com base no impacto estimado na empregabilidade dos participantes. A ideia é basicamente transformar as estimativas de Δ em um valor monetário para cada participante. E, depois, transformar esse valor numa projeção de benefícios para a sociedade que, associado ao custo de R\$ 1,1 milhão, se transformará em um fluxo de caixa. Então será possível calcular o Valor Presente Líquido (VPL) e a Taxa Interna de Retorno (TIR) do programa.

Assim, a primeira pergunta a se fazer é: quanto vale, para a sociedade, um aumento de 19,6% ou 11,1% na chance de um participante do programa encontrar emprego? Para responder esta pergunta, volte a observar as Tabelas 4 e 5, apresentadas anteriormente, e note que para cada “tipo de pessoa” da amostra se possui uma estimativa de $\Pr(Y = 1|T, X)$, para $T = 0$ e $T = 1$.

Para a coluna dos tratados, nas Tabelas 4 e 5, $\Pr(Y = 1|T, X)$ equivale a $\Pr(Y_1 = 1|T = 1, X)$ – o factual dos tratados. E para a coluna dos controles, $\Pr(Y = 1|T, X)$ equivale a $\Pr(Y_0 = 1|T = 0, X)$ – o factual dos controles. Mas como se está aceitando que $\Pr(Y_0 = 1|T = 0, X) \approx \Pr(Y_0 = 1|T = 1, X)$, porque se aceitou que os grupos de tratamento e controle são suficientemente parecidos, assim o factual dos controles pode servir de contrafactual dos tratados, para cada tipo de pessoa com as características X . E então, dada a conveniência para a abordagem algébrica, fica clara a escolha de se estimar a equação [5].

Este raciocínio permite construir uma estimativa de quantos indivíduos empregados (e desempregados) estariam no grupo de tratamento, caso não tivessem sido tratados. E, por consequência, de quantos empregos podem ser devidos exclusivamente ao tratamento (e não pelas características dos indivíduos).

Para tanto, multiplica-se a estimativa de $\Pr(Y_0 = 1|T = 1, X)$ pelo número de observações das unidades tratadas com determinado “tipo”. Este resultado é apresentado adiante, na Tabela 7.

Tabela 7 – Estimativa de empregos devidos ao tratamento e expectativa de duração do impacto por tipo de trabalhador

Sexo	Ensino Médio	Faixa Etária	Expectativa de duração do impacto (em anos)	Factual				Controle 1: PNAD				Controle 2: PED			
				$Pr(Y_i = 1 T = 1, X)$	Empregados (A)	Desempregados	Total (B)	Contrafactual			Empregos devidos ao tratamento (A-D)	Contrafactual			Empregos devidos ao tratamento (A-F)
								$Pr(Y_i = 1 T = 1, X)$ (C)	Empregados (D=B*C)	Desempregados		$Pr(Y_i = 1 T = 1, X)$ (E)	Empregados (F=B*E)	Desempregados	
Mulher	Incompleto	[15,19]	30	0,386	56	88	144	0,197	28	116	28	0,187	27	117	29
		(19,22]	30	0,347	58	109	167	0,169	28	139	30	0,237	40	127	18
		(22,28]	25	0,332	27	53	80	0,159	13	67	14	0,251	20	60	7
		(28,56]	20	0,296	6	15	21	0,136	3	18	3	0,241	5	16	1
	Completo	[15,19]	30	0,462	12	13	25	0,255	6	19	6	0,228	6	19	6
		(19,22]	30	0,422	30	41	71	0,223	16	55	14	0,283	20	51	10
		(22,28]	25	0,405	80	117	197	0,211	42	155	38	0,299	59	138	21
		(28,56]	20	0,366	56	96	152	0,182	28	124	28	0,288	44	108	12
Homem	Incompleto	[15,19]	35	0,497	46	47	93	0,284	26	67	20	0,318	30	63	16
		(19,22]	35	0,456	36	42	78	0,250	20	58	16	0,382	30	48	6
		(22,28]	30	0,440	12	16	28	0,237	7	21	5	0,400	11	17	1
		(28,56]	25	0,399	2	4	6	0,206	1	5	1	0,388	2	4	0
	Completo	[15,19]	35	0,574	5	3	8	0,353	3	5	2	0,371	3	5	2
		(19,22]	35	0,534	29	25	54	0,316	17	37	12	0,437	24	30	5
		(22,28]	30	0,517	45	42	87	0,301	26	61	19	0,455	40	47	5
		(28,56]	25	0,476	35	38	73	0,266	19	54	16	0,443	32	41	3
Estimativa do total de empregos devidos ao tratamento											252				143

Usando como controle a amostra da PNAD, na coluna apontada por (D) estima-se o número de pessoas, por tipo, que estariam empregadas se não tivesse ocorrido o programa (produto dos valores das colunas B e C). E na coluna marcada com (A-D) aponta-se a estimativa dos empregos devidos ao tratamento por tipo de indivíduo, e no total: 252. Então, argumenta-se a estimativa de que 252 das 535 pessoas tratadas, que foram entrevistadas em janeiro de 2010 e se declararam empregadas, não estariam empregadas se não fosse pelo treinamento.

De forma análoga, em outro cenário, mais realista por conta da questão sazonal e do foco em Fortaleza, usando como controle a amostra da PED, na coluna apontada com um (F) estima-se o número de pessoas, por tipo, que estariam empregadas se não tivesse ocorrido o programa (produto dos valores das colunas B e E). E na coluna mais a direita, marcada com (A-F), aponta-se a estimativa dos empregos devidos ao tratamento por tipo de indivíduo, e no total: 143.

Mas quanto vale 252 ou 143 pessoas produtivas para se possa comparar com o custo de R\$ 1,1 milhão? Para responder isso são necessárias novas hipóteses. A primeira diz respeito a quanto uma dessas pessoas recebe de salário e reverte isso na economia (em consumo, por exemplo). E a segunda diz respeito ao tempo que esse recebimento de salário irá perdurar.

Nesse sentido, a hipótese que se fará é que cada uma das 252 ou 143 pessoas em questão receberá R\$ 5.580,00 ao ano, a valores de 2009. O que equivale a 12 vencimentos de R\$ 465,00 – um salário mínimo da época.

Quanto ao tempo em que esse recebimento irá perdurar, isso depende intrinsecamente a idade da pessoa. A hipótese é de que os jovens de 1º emprego receberão seus vencimentos pelos próximos 35 anos, se homens, e 30 anos, se mulheres (período legal de contribuição previdenciária para se aposentarem). Já a hipótese para os outros perfis, definida de forma *ad-hoc*, é exposta na quarta coluna da Tabela 7, intitulada por “expectativa de duração do impacto”.

Os custos do programa são apresentados na Tabela 8, abaixo. Como era de se esperar, 45,3% das despesas (R\$ 498 milhões) se deram com pessoal e encargos.

Tabela 8 – As despesas do programa – R\$ 1.000

Item de despesa	Valor	%
Despesas e encargos com professores e coordenação	498	45,3
Vale-transporte dos alunos	211	19,2
Alimentação dos alunos	169	15,4
Divulgação	104	9,5
Material didático	79	7,2
Material de consumo	39	3,5
Total	1.100	100,0%

Fonte: CDL/Fortaleza. Elaboração Própria.

Os outros dois grandes grupos de despesa foram com os vales-transportes distribuídos aos alunos, consumindo 19,2% dos recursos, e com a alimentação deles durante o curso, consumindo 15,4% dos recursos¹⁶. As demais despesas foram com divulgação (9,5%), material didático (7,2%) e material de consumo (3,5%).

Associando os R\$ 5.580 ao ano, a valores de 2009, com as expectativas de duração do impacto por cada tipo de indivíduo e com o custo de R\$ 1,1 milhão, o VPL do programa é estimado em R\$ 9,8 milhões para o impacto medido via PNAD, e R\$ 5,1 milhões para o impacto medido via PED (a taxa de 12% a.a.). E a TIR é estimada em 127% e 72% a.a., para os controles via PNAD e PED, respectivamente. Logo, pelo impacto estimado e pelas hipóteses feitas para a análise de custo/benefício, conclui-se que o programa foi um bom investimento para a sociedade.

¹⁶ Para minimizar a evasão, durante os cursos, os alunos recebiam vale-transporte e alimentação, além do material didático.

7. Considerações finais

O PLANSEQ possui um orçamento anual de cerca de R\$ 300 milhões, oriundo do FAT. E apesar de muitos programas do PLANSEQ já terem sido realizados no país, dificilmente se encontram avaliações de impacto e retorno econômico documentadas na literatura. Este artigo buscou suprir esta lacuna estudando o programa da CDL de Fortaleza, realizado em 2009.

Um ponto crucial da análise foi a ausência de dados primários para a construção do grupo de controle. O que foi contornado através do uso dos dados da PNAD e da PED, cuja semelhança com o grupo de tratamento foi testada através de um pareamento por escore de propensão. Como ainda é comum a falta de cuidado em se produzir grupos de controle, a solução de uso de bases de dados secundárias para o caso do PLANSEQ e suas futuras avaliações também são vistas como uma contribuição deste artigo.

Observou-se que, no cenário mais realista, o programa da CDL teve um impacto de 11,1% na empregabilidade dos tratados. E a análise de custo/benefício apontou que o investimento de R\$ 1,1 milhão deve gerar R\$ 5,1 milhões de retorno para a sociedade, a valores de 2009. Ou seja, parece que o programa foi um bom negócio para a sociedade. Todavia, estes resultados positivos foram encontrados para um PLANSEQ em particular, o que não significa que os outros também estão sendo bom negócio para a sociedade.

8. Bibliografia

- Angrist, Joshua D.; Pischke, Jörn-Steffen (2008). **Mostly Harmless Econometrics: An Empiricist's Companion**. Princeton University Press.
- Becker, Sascha; Ichino, Andrea (2002). **Estimation of Average Treatment Effects Based on Propensity Scores**. *Stata Journal*, v. 2, n. 4, p. 358–77.
- Brasil (2010). **Balanco de Governo: 2003 a 2010**. Presidência da República, disponível em: <http://www.secom.gov.br/sobre-a-secom/publicacoes/balanco-de-governo-2003-2010/livro-1>.
- Card, David; Sullivan, Daniel (1988). **Measuring the Effect of Subsidized Training Programs on Movements In and Out of Employment**. *Econometrica*, Vol. 56, No 3, pp. 497-530.
- Heckman, James J.; Ichimura, Hidehiko; Todd, Petra (1997). **Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme**. *Review of Economic Studies* 64 (4): 605–54.
- Khandker, Shahidur R.; Koolwal, Gayatri B.; Samad, Hussain A. (2010). **Handbook on Impact Evaluation: Quantitative Methods and Practices**. The International Bank for Reconstruction and Development.
- MTE (2010). **Relatório de avaliação do Plano Plurianual 2008-2011**. MTE. Ver: <http://portal.mte.gov.br/data/files/8A7C812D2E7318C8012EA676FB23135E/Caderno%20Setorial%20Aval%202009%20Vers%C3%A3o%20Final.pdf>.
- Oliveira, Ana Maria H. C.; Rios-Neto, Eduardo L. G. (2007). **Uma avaliação experimental dos impactos da política de qualificação profissional no Brasil**. *Revista Brasileira de Economia*, vol. 61, no.3, Rio de Janeiro.
- Ravallion, Martin (2001). **The mystery of vanishing benefits: Ms Speedy Analyst's introduction to evaluation**. *The World Bank Economic Review*, v. 15, n. 1, p. 115-140.
- Rosenbaum, Paul R.; Rubin, Donald B. (1983). **The Central Role of the Propensity Score in Observational Studies for Causal Effects**. *Biometrika*, v. 70, n. 1, p. 41-55.
- Vasconcellos, Lígia; Biondi, Roberta L.; Menezes-Filho, Naércio A. (2009). **Avaliando o impacto do programa Escrevendo o Futuro no desempenho das escolas públicas na Prova Brasil e na renda futura dos estudantes**. *Anais do XXXVII Encontro Nacional de Economia da ANPEC*, Foz do Iguaçu, Paraná.