

## **Uma avaliação econômica do ProUni contrastando seu impacto na massa salarial dos egressos com o gasto tributário do programa**

Vinicius Augusto Lima de Almeida\*  
[vinicius.almeida@aol.com](mailto:vinicius.almeida@aol.com)

Francis Carlo Petterini\*  
[f.petterini@ufsc.br](mailto:f.petterini@ufsc.br)

\*PPGEco/UFSC

**Resumo:** O Programa Universidade para Todos (ProUni) é uma política pública educacional que concede isenção fiscal para Instituições de Ensino Superior privadas em contrapartida da oferta de bolsas de estudo para estudantes de baixa renda. Até o ano de 2017, estima-se que mais de 2 mil IES aderiram ao programa, ofertando mais de 2 milhões de bolsas desde 2005. O presente artigo utiliza pela primeira vez microdados educacionais e do mercado de trabalho para realizar uma avaliação de impacto do programa no mercado formal de trabalho e calcular o seu retorno econômico. Os resultados encontrados indicam que, em média, um diplomado pelo ProUni recebe 27,5% a mais do que o seu contrafactual. O retorno econômico agregado foi positivo e estimado em R\$ 38 bilhões desde 2005, ainda que o payback oscile entre os cursos. As discussões a partir da análise econômica do ProUni permitem avaliar melhoramentos no desenho do programa, como a melhor focalização da oferta de bolsas em cursos de graduação que gerem tanto um maior retorno no mercado de trabalho individualmente quanto, através do número de bolsas ofertadas, um maior retorno social.

**Palavras-chave:** ProUni, avaliação de impacto, retorno econômico, gasto tributário.

**Abstract:** The Programa Universidade para Todos (ProUni) is an educational public policy that grant fiscal incentives to private higher education institutions for supplying scholarships for low income students. By 2017 almost 2 thousand institutions joined in the program, supplying over than 2 millions scholarships since 2005. The paper uses for the first time individual data from education and labor market to estimate an impact evaluation of Prouni on formal labor market and calculate its economic return. The results indicates that, on average, a egressed from Prouni receives higher salaries (27,5%) than his counterfactual, which indicates that Prouni can increase the qualification of low-income students, providing better opportunities on labour market. The economic return was positive and estimated in R\$ 38 billions since 2005, varying according to the graduation course. The results suggests ways to improve the program's design, like the better focalization of scholarships towards graduations that generate higher returns on formal labor market to students, promoting a better social return.

**Keywords:** ProUni, impact evaluation, economic return, tax expenditure.

Área de submissão: Desenvolvimento econômico  
JEL: I28, J21, J24.

## 1) Introdução

Nas últimas décadas houve um expressivo incremento de matrículas no sistema de ensino superior brasileiro, embora isso tenha se configurado de forma diferente entre os tipos de Instituições de Ensino Superior (IES) (Durham 2005; Corbucci 2007, 2014; Tachibana et al. 2015). Nas IES públicas ocorreu uma expansão súbita entre o fim dos anos 2000 e meados dos anos 2010, a reboque do Programa de Apoio aos Planos de Reestruturação e Expansão das Universidades Federais (REUNI) e da remodelagem da rede de educação profissional e tecnológica (Faveri et al. 2018; Barbosa et al. 2019). Já nas IES privadas aconteceu um crescimento regular desde o fim dos anos 1990, em função de vários fatos correlatos, como, por exemplo: muitas organizações perceberam a lucratividade do setor e passaram a criar novas faculdades e cursos por todo o país; o ensino a distância (EaD) foi popularizado; o crédito estudantil foi fortalecido com o Fundo de Financiamento Estudantil (FIES); e, o número de bolsas de estudo foi aumentado com o Programa Universidade Para Todos (ProUni) (TCU, 2009; CGU, 2015; Corbucci et al. 2016).

Sucintamente, o ProUni é um sistema de concessão de isenções fiscais para IES privadas que ofertam bolsas para estudantes de baixa renda. Em números gerais, desde 2005, quando começou, estima-se que a política já custou mais de R\$ 13 bilhões, em valores atuais; e, que beneficiou mais de 500 mil alunos.

Apesar da magnitude considerável, há poucas tentativas de avaliação de impacto do ProUni. Dentre o que existe, Lépine (2018) e Becker e Mendonça (2019) se destacam ao comparar a proficiência dos beneficiários com outros alunos semelhantes ao público-alvo, usando os microdados do Exame Nacional de Desempenho dos Estudantes (Enade). As duas pesquisas indicam que o programa teria impacto positivo nas notas dos bolsistas, sugerindo que eles tendem a conseguir dedicar mais tempo à sua educação quando comparados com contrapartes de controle. Logo, isso poderia melhorar o capital humano nacional, gerando um benefício coletivo que justificaria o gasto tributário.

Todavia, essas pesquisas não chegam a contrastar o impacto com a isenção fiscal decorrente do ProUni, dando fechamento a uma análise econômica. Talvez porque seja difícil converter os resultados do Enade em unidades monetárias, a fim de compará-los com o gasto tributário em um fluxo de caixa, para então calcular, por exemplo, um Valor Presente Líquido (VPL) do investimento que a sociedade faz ao abrir mão de receita.

Além disso, em contraponto à proficiência, o salário seria uma medida de impacto mais aderente ao objetivo do programa, qual seja: incentivar o acesso ao ensino superior e a diplomação de pessoas de baixa renda, a fim de diminuir desigualdades sociais (Corbucci 2007, 2014). Porque, da perspectiva da teoria econômica e da evidência empírica, é o impacto no salário que melhor se relaciona com o incremento de produtividade, e este, por fim, é que externaria o aumento do capital humano para a coletividade, culminando na redução das desigualdades sociais (Mincer 1996; Barbosa-Filho e Pessoa 2010; Katovich e Maia 2018). Ademais, ao se avaliar o impacto através do salário, não existe necessidade de converter unidades de medida para comparar com o gasto tributário, o que favorece uma análise econômica.

Com esses pontos em mente, o objetivo geral do artigo é avançar em relação aos trabalhos de Lépine (2018) e Becker e Mendonça (2019), promovendo uma avaliação econômica do ProUni considerando seu impacto no salário dos egressos. Especificamente, a ideia é estimar quantidades de unidades monetárias que não existiriam na economia nacional sem o mecanismo de geração das bolsas, a fim de aplicar ferramentas de matemática financeira para comparar diretamente o impacto com a isenção fiscal.

Além disso, nota-se que uma das críticas do ProUni é que as bolsas são concentradas em poucos cursos – de fato, quase metade são para Administração, Direito, Pedagogia e Ciências Contábeis. Então, ao se estimar o impacto por curso, um objetivo secundário é fomentar

sugestões de melhoria no desenho da política, indicando quais formações poderiam ser priorizadas para beneficiar a sociedade.

Com esses objetivos, três bases de microdados foram cruzadas para operacionalizar a análise: a lista dos beneficiários entre 2005 e 2011, obtida no sistema do ProUni do Ministério da Educação (SISPROUNI, MEC); os registros do Censo da Educação Superior (CES), para identificar características dos estudantes, dos cursos, das faculdades etc., e montar grupos de comparação; e, a Relação Anual de Informações Sociais (RAIS, identificada), entre 2010 e 2018, para rastrear os egressos ex-bolsistas (e outros estudantes do público-alvo) formalmente empregados, e descobrir os respectivos salários. Ao cabo, tal procedimento resultou na observação de várias informações de 346.810 ex-bolsistas do ProUni.

Ao definir os diplomados e os evadidos do programa como grupos de tratamento e de controle, respectivamente, técnicas econométricas das literaturas de economia do trabalho e de avaliação de políticas públicas foram aplicadas à análise dos dados, e o resultado geral é que: um egresso representativo, nos primeiros anos depois de formado, ganha 20% a mais do que ganharia o seu contrafactual sem formação. Esse impacto é de cerca de 50% em formações como Farmácia ou Medicina; por outro lado, em muitos cursos o impacto é estatisticamente zero.

Ao converter esses impactos em efeitos na massa de salários, contrasta-se isso com os registros contábeis de renúncia fiscal, e então se estima que: a Taxa Interna de Retorno (TIR) do programa seria de 16% ao ano; o VPL social já seria maior que R\$ 38 bilhões; e, cada geração de diplomados pelo ProUni equaciona o gasto tributário com sua coorte em cerca de 10 anos. Além disso, embora formações como Administração e Ciências Contábeis tenham um VPL individual menor que Farmácia e Medicina, por exemplo, nota-se que o VPL social dos primeiros é substancialmente maior porque os beneficiários são muitos nesses cursos.

Além dessa introdução, o artigo tem a seguinte forma: nas Seções 2 e 3 o programa e o raciocínio da análise são detalhados, respectivamente; na Seção 4 se discute a econometria da aferição de impacto; na Seção 5 a construção da base de dados e as estatísticas descritivas são apresentadas; na Seção 6 têm-se os resultados econométricos; na Seção 7 faz-se a avaliação econômica do programa; e, por fim, a Seção 8 trás a conclusão.

## **2) ProUni**

O ProUni foi institucionalizado em 2005 pela Lei Nº 11.096, concedendo abatimentos no Imposto de Renda das Pessoas Jurídicas (IRPJ), na Contribuição Social sobre o Lucro Líquido (CSLL), na Contribuição Social para Financiamento da Seguridade Social (Cofins) e no Programa de Integração Social (PIS), para as IES privadas que oferecem bolsas de estudo (integrais e/ou parciais) em cursos de nível superior (na forma de graduação e/ou sequencial, sendo presenciais e/ou em EaD). Apesar das nuances legais, a regra geral é dar isenção total quando a oferta atinge a proporção de uma bolsa integral (por curso) a cada 10 alunos pagantes (em cada curso da IES).

A Tabela 1 mostra o número de instituições participantes e elegíveis ao ProUni no período 2005-2017, ilustrando que cerca de metade delas se interessam pelo programa. Provavelmente, essa fração não é maior porque a maioria das IES já obtivera isenção total ou parcial de tributos antes de 2005, através de outros instrumentos (TCU 2009; Haas e Pardo, 2017).

Os candidatos a uma bolsa devem: prestar o Exame Nacional do Ensino Médio (Enem), sem zerar a redação; e, fazer a inscrição no SISPROUNI escolhendo instituições, cursos e turnos de preferência. Na sequência, o sistema faz uma pré-seleção em função dessas informações e das bolsas disponíveis. Caso o estudante seja pré-selecionado, ele deve comprovar que: ainda não concluiu um curso superior; tem renda familiar mensal de até 1,5 salário mínimo (SM) per capita para as bolsas integrais, ou até 3 SM para as bolsas parciais; e,

que frequentou escola pública ou foi ex-bolsista de escola privada. Também há um percentual de bolsas destinadas aos estudantes com deficiência, e aos professores da rede pública de educação básica; e, uma ajuda de custo para alunos de cursos em turno integral. Uma vez no programa, para não ser desligado, o bolsista deverá cumprir frequência e aprovações mínimas nas disciplinas cursadas.

**Tabela 1: Número de IES participantes e elegíveis ao ProUni (2005-2017)**

IES	Participantes [a]	Elegíveis [b]	% a/b
2005	928	1.777	52,2
2006	1.096	2.018	54,3
2007	1.203	2.095	57,4
2008	1.243	2.201	56,5
2009	1.287	2.270	56,7
2010	1.290	2.284	56,5
2011	1.283	2.352	54,5
2012	1.309	2.383	54,9
2013	1.130	2.398	47,1
2014	1.157	2.358	49,1
2015	1.141	2.373	48,1
2016	1.106	2.494	44,3
2017	1.226	2.695	45,5

Fonte dos dados: MEC. Elaboração própria.

A Tabela 2 apresenta uma estimativa do gasto tributário com o ProUni entre os anos 2005 e 2017, com valores de 2020, usando dados da Receita Federal do Brasil (RFB). Também se mostra o número de bolsas ocupadas ao ano – regularmente, entre 65% e 75% delas são integrais. As somas dos valores das colunas [a] e [b] são da ordem de R\$ 13,5 bilhões e 2,2 milhões de bolsas.

**Tabela 2: Renúncia fiscal e bolsas ocupadas do ProUni ao ano (2005-2017)**

Ano	Renúncia (R\$ milhões*) [a]	Bolsas (milhares) [b]	1.000xa/b
2005	598,3	95,6	6.259,3
2006	604,7	109,0	5.548,3
2007	578,1	105,5	5.478,1
2008	688,1	124,6	5.524,3
2009	1.002,5	161,3	6.217,2
2010	1.195,7	152,6	7.834,8
2011	922,3	170,7	5.404,7
2012	1.244,0	176,7	7.041,8
2013	1.202,7	177,2	6.787,0
2014	909,0	223,4	4.069,1
2015	1.379,1	252,7	5.464,5
2016	1.643,0	238,9	6.875,6
2017	1.603,1	240,6	6.661,6

\*Valores de 2020 corrigidos pelo IPCA. Fonte dos dados: RFB e MEC. Elaboração própria.

Considerando que uma bolsa é ocupada por uma mesma pessoa por períodos subsequentes, e tomando por hipótese que se leva em média quatro anos para uma diplomação, divide-se 2,2 milhões por quatro para concluir que o programa teve mais de 500 mil beneficiários no espaço em tela. Sendo que, pelos números apresentados por CGU (2015), nota-se que há até 30% de casos de evasão ou de desligamento, e que isso aumenta a

rotatividade dos ocupantes de uma bolsa, então se percebe que o número de beneficiários pode ter sido maior.

Fazendo-se a razão dos valores das colunas [a] e [b], conclui-se que há um gasto tributário anual aproximado de R\$ 6 mil por bolsa, em valores de 2020. Além disso, nota-se CGU (2015) faz uma análise semelhante desse valor usando uma metodologia mais apurada, que controla uma série de inconsistências nos dados, mas ainda assim a estimativa do gasto anual por bolsa fica próxima disso.

Outro ponto para se registrar é que existe algum debate entre apoiadores e críticos do programa (Saraiva et al., 2011). Entre os primeiros, defende-se que é uma política bem focalizada, porque promove o acesso de estudantes de baixa renda de acordo com o desempenho no Enem, que é uma avaliação padronizada e de abrangência nacional. Por outro lado, os críticos argumentam que o ProUni apenas permite o acesso a uma educação superior, mas as bolsas seriam consubstanciadas em um subconjunto restrito de cursos, e às vezes em instituições mal avaliadas pelo próprio MEC.

**Tabela 3: Percentual de bolsas ofertadas por curso entre as 30 formações com mais beneficiários no período 2005-2017**

#	Curso	%	% acumulado	#	Curso	%	% acumulado
1	Administração	21,49	21,49	16	Engenharia Civil	1,50	76,37
2	Direito	12,21	33,70	17	Nutrição	1,41	77,78
3	Pedagogia	8,81	42,51	18	Eng. Produção	1,35	79,13
4	Ciências Contábeis	5,51	48,02	19	Arq. e Urbanismo	1,22	80,35
5	Enfermagem	4,23	52,25	20	Medicina	1,06	81,41
6	Educação Física	3,26	55,51	21	Eng. Elétrica	0,97	82,38
7	Psicologia	3,22	58,73	22	História	0,93	83,31
8	Sist. de Informação	2,52	61,25	23	Logística	0,87	84,18
9	Fisioterapia	2,28	63,53	24	Odontologia	0,86	85,04
10	Letras	2,18	65,71	25	Marketing	0,86	85,90
11	Serviço Social	2,07	67,78	26	Eng. Mecânica	0,85	86,75
12	Farmácia	2,04	69,82	27	Turismo	0,76	87,51
13	Com. Social	1,95	71,77	28	Economia	0,72	88,23
14	Biologia	1,60	73,37	29	Biomedicina	0,72	88,95
15	Ciências da Comp.	1,50	74,87	30	Matemática	0,66	89,61

Fonte dos dados: MEC. Elaboração própria.

De fato, como se apresenta na Tabela 3, apenas os cursos de Administração e de Direito responderam por 1/3 de todas as bolsas ofertadas entre 2005 e 2017; acrescentando-se Pedagogia e Ciências Contábeis, esses quatro englobaram praticamente metade das bolsas. Por outro lado, em face da regra de isenção fiscal, cada IES é incentivada a oferecer bolsas para todos os cursos em que realiza seleção, na proporção de suas matrículas. Portanto, tal concentração pode ser simplesmente um reflexo da oferta geral concentrada em poucos cursos – e pode não caracterizar nenhuma medida de má fé ou algo assemelhado. A listagem da tabela segue com os 30 cursos que responderam por quase 90% da oferta entre 2005 e 2017, sendo que a listagem completa conta com centenas de cursos entre graduações e níveis sequenciais.

Quanto ao argumento dos críticos em relação à qualidade, pode-se explorar o Índice Geral de Cursos (IGC) calculado pelo MEC a fim de sintetizar esse ponto – seus os valores vão de um (pior) a cinco (melhor). Nesse sentido, a média do IGC nas instituições participantes é de 3,19, e nas particulares não participantes é de 3,11; mas a diferença não é estatisticamente diferente de zero. Ou seja, por esse indicador, não há evidência de que na média uma IES participante do ProUni tenha qualidade diferente de uma não participante.

Em termos de marco lógico de política pública, no sentido de Gertler et al. (2018), salienta-se que as bolsas são o produto do ProUni. Dado esse produto, existem os seguintes resultados de curto prazo esperados sobre o público-alvo: aumento das taxas de entrada e de diplomação no ensino superior; aceleração do tempo de formação; e, dilatação do tempo dedicado aos estudos em detrimento ao trabalho, com consequente melhora na proficiência.

De fato, Lépine (2018) e Becker e Mendonça (2019) verificam uma melhora de proficiência nos beneficiários, sobretudo para quem recebe bolsa integral; e, Andriola e Barrozo-Filho (2020) notam que o tempo de formação se acelera. Ademais, embora não se tenha encontrado nada específico sobre o ProUni, Dynarski (2000, 2003), Stinebrickner e Stinebrickner (2003) e Darolia (2014) encontram evidências de que esses resultados de curto prazo ocorrem em programas semelhantes nos EUA.

Em um prazo mais dilatado, o programa visaria uma diminuição das desigualdades sociais, ao incentivar o acesso ao ensino superior e a diplomação de pessoas de baixa renda familiar (Corbucci 2007, 2014). Então, os indicadores de impacto do ProUni envolveriam o aumento da empregabilidade e dos salários dos beneficiários, e isso, por sua vez, refletiria em mais bem-estar para a coletividade.

### 3) Avaliação econômica

A avaliação econômica é dividida entre: aferição de impacto; e, contraste do impacto com os custos. Seguindo a praxe, define-se resultados potenciais  $w_{1i}$  e  $w_{0i}$  como salários factuais e contrafactuais, respectivamente, do egresso  $i$  diplomado no ProUni. Ou seja,  $w_{1i}$  é um salário observado e  $w_{0i}$  é um salário que seria observado caso  $i$  não tivesse sido diplomado; e,  $w_{1i} - w_{0i}$  é o impacto do programa na perspectiva individual. Então,  $B = \sum_i (w_{1i} - w_{0i})$  é uma massa salarial causada pelo programa, que se sucede por períodos pós-diplomação de uma coorte, e representa o impacto do programa na perspectiva da sociedade, porque se trata de dinheiro que será consumido, investido, tributado etc. Essa massa pode ser contrastada com os custos do programa, a fim de dar fechamento a uma análise econômica. Porém, é impossível computar  $B$  diretamente porque não se observa o salário contrafactual  $w_{0i}$  – trata-se do problema fundamental da inferência causal de uma política (Gertler et al., 2018).

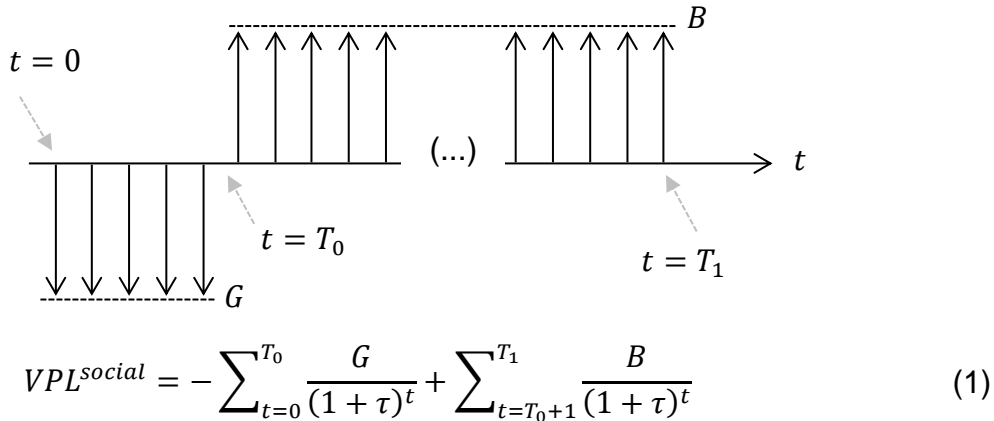
Para contornar esse problema, uma estratégia começa por aplicar a lógica de Mincer (1970), definindo:  $d = 0$  e  $d = 1$  como planos genéricos de escolarização;  $w_d$  para o salário com o respectivo plano; e,  $s_d$  e  $n_d$  como períodos de estudo e de dedicação ao trabalho a partir desses planos, tal que  $s_1 > s_0$  e  $s_d < n_d$  (o que implica que apenas o período de trabalho após a finalização do plano de escolarização é considerado na modelagem).

Assim, em tempo contínuo, o valor presente do plano  $d$  no mercado de trabalho é  $V_d = \int_{s_d}^{n_d} w_d \exp(-\gamma t) dt$ , onde  $\gamma > 0$  é uma taxa de desconto intertemporal. Como  $d = 1$  envolve mais escolarização que  $d = 0$ , a diferença  $V_1 - V_0$  é o VPL do investimento pessoal em  $d = 1$  vis-à-vis  $d = 0$ . Logo, a igualdade  $V_1 - V_0 = 0$  implica em dois resultados. Primeiro, fazendo  $\gamma = \ln(1 + r)$ , o termo  $r$  representa uma TIR do esforço pessoal em obter mais escolarização. Segundo, com alguma álgebra se verifica que  $w_1 \exp(-\gamma s_1) (1 - \exp(-\gamma n_1)) = w_0 \exp(-\gamma s_0) (1 - \exp(-\gamma n_0))$ .

Com esse último resultado, se  $n_1 = n_0$ , ou se  $n_1 \neq n_0$  mas ambos forem suficientemente grandes (i.e., as pessoas trabalham por muitos anos), ocorre  $w_1/w_0 \cong \exp(\gamma(s_1 - s_0))$ . Sem perda de generalidade, pode-se escrever  $s_1 = 1$  e  $s_0 = 0$ , significando, em particular, ter ou não ter se diplomado, respectivamente. Convenientemente, isso implica que:  $d = 1$  e  $d = 0$  passam a significar o mesmo que  $s_1 = 1$  e  $s_0 = 0$ ; e, a relação  $w_1/w_0 \cong \exp(\gamma)$  é verdadeira. Logo, definindo uma tautologia  $w \equiv w_0(w_1/w_0)^d$  para qualquer salário  $w$  observado (seja factual ou contrafactual), e associando-a ao último resultado, encontra-se  $w \cong w_0(\exp(\gamma))^d$ , que implica em  $\ln w = \ln w_0 + \gamma d + u$ , sendo  $u$  um termo de erro acrescentado para transformar o sinal de aproximação em sinal de igualdade.

Ou seja, com essa modelagem se chega a uma versão da equação Minceriana com a qual  $\gamma$  pode ser estimado, e então interpretado como a TIR do esforço de diplomação através da relação  $r = \exp(\gamma) - 1$ . Além disso, desde que  $d$  indique apropriadamente o pertencimento a grupos de tratamento e de controle, nessa estrutura  $\gamma$  também pode ser interpretado como “efeito tratamento médio sobre os tratados” (pela diplomação), ou ATT pela sigla da nomenclatura em inglês, cujos detalhes são didaticamente apresentados em Wooldridge (2010, p. 605).

**Figura 1: Diagrama do fluxo de caixa social do programa**



O problema do salário contrafactual não observado é então contornado com o resultado  $w_{0i} \cong w_{1i}/\exp(\gamma) = w_{1i}/(1 + r)$ , porque ocorre  $B \cong \sum_i (w_{1i} - w_{1i}/(1 + r)) = (r/(1 + r))\sum_i w_{1i}$ . Ou seja,  $\sum_i w_{1i}$  é a massa salarial observada dos diplomados no ProUni, cuja fração  $r/(1 + r)$  é estimável. Para dar fechamento a um fluxo de caixa social, nota-se que todo ano há um gasto tributário  $G$  com cada coorte, por um período representativo de  $T_0$  anos até a obtenção dos diplomas. Se a coorte mantém o benefício por  $T_1 - T_0$  anos, então esse fluxo seria representado por um diagrama como o da Figura 1. Assim, o VPL do programa na perspectiva social seria  $VPL^{social}$  conforme a fórmula (1), para alguma taxa de referência  $\tau$ ; com isso, se pode estimar a TIR e o *payback* para a coletividade.

#### 4) Abordagem econométrica

A estratégia de identificação do impacto do programa parte de uma equação Minceriana na forma  $\ln w_i = \ln w_{0i} + \gamma d_i + u_i$ , em que:  $i$  indexa  $N$  indivíduos;  $w_i$  representa o salário observado;  $w_{0i}$  é um termo de salário potencial;  $\gamma$  é a medida de ATT;  $d_i = 1$  e  $d_i = 0$  indicam pertencimento aos grupos de tratamento (diplomados com o ProUni) ou de controle (não diplomados semelhantes ao público-alvo), respectivamente; e,  $u_i$  é o termo de erro.

Seguindo a praxe da literatura de economia do trabalho, reescreve-se  $\ln w_{0i} = \text{cte} + \sum_{l=1}^L \beta_l x_{li}$ , em que  $\text{cte}$  é uma constante e cada  $\beta_l$  e  $x_{li}$  representam  $L$  parâmetros e covariadas do salário, respectivamente – e.g., a experiência e o seu quadrado (Taber, 2001; Heckman et al., 2006; Barbosa-Filho e Pessoa 2010). Portanto, a equação de base do exercício é:

$$\ln w_i = \text{cte} + \gamma d_i + \sum_{l=1}^L \beta_l x_{li} + u_i \quad (2)$$

Dada uma amostra  $\{w_i, d_i, x_{1i}, \dots, x_{Li}\}_{i=1}^N$ , a estimação dos parâmetros dessa equação deve considerar a endogeneidade de  $d_i$ , porque o salário é influenciado não só por características observáveis  $x_{li}$ , mas também por variáveis não observáveis incorporadas em  $u_i$ , e que podem ser correlacionadas com  $d_i$ . Genericamente, a literatura chama esse componente não observado de “habilidade”.

Wooldridge (2010, p. 61) faz uma apresentação didática dessa questão, tal que se pode rescrever o erro como  $u_i = \delta o_i + v_i$ , onde  $\delta$  é um parâmetro que associa a variável omitida  $o_i$  com  $d_i$  em (2), e  $v_i$  é o termo de erro ao não se omitir  $o_i$ . Com alguma álgebra, demonstra-se que a convergência em probabilidade do estimador de mínimos quadrados ordinários (OLS) de  $\gamma$ , denotado de  $\hat{\gamma}$ , é tal que  $\text{plim } \hat{\gamma} = \gamma + \delta \times \text{Cov}(d_i, o_i) / \text{Var}(d_i)$ , onde Cov e Var são covariância e variância, respectivamente.

Portanto, se  $\delta > 0$  (salários mais altos são associados com maiores habilidades), e se há mais pessoas habilidosas entre os diplomados ( $\text{Cov}(d_i, o_i) > 0$ ), ocorre que  $\hat{\gamma}$  superestima  $\gamma$ . Por outro lado, se  $o_i$  representa uma experiência não observada também se espera  $\delta > 0$  (porque salários mais altos devem estar associados com pessoas mais experientes), mas se há menos pessoas experientes entre os diplomados, tem-se  $\text{Cov}(d_i, o_i) < 0$  e então  $\hat{\gamma}$  subestima  $\gamma$ . Em suma, determinar o sinal e a magnitude de tal viés costuma ser complicado, e para mitigar esse problema é comum aplicar variáveis instrumentais, normalmente através de mínimos quadrados em dois estágios (2SLS).

Assim, na literatura de economia do trabalho há uma profícua discussão sobre os potenciais instrumentos associados com a escolaridade, mas que não com a habilidade. Sendo que, nos casos de análise do ensino superior, o mais comum parece ser o uso do nível de educação dos pais; mas, infelizmente, na tabulação dos dados (discutida na próxima seção) não foi possível isso.

Por outro lado, como comentado anteriormente, o MEC disponibiliza o IGC como uma medida da qualidade dos cursos, e também é possível identificar o tempo de existência das IES. Então, conjectura-se que maiores qualidade e idade da IES podem motivar o estudante a se diplomar, seja porque o curso ficaria mais interessante, seja porque a expectativa de empregabilidade seria maior em face da conclusão de cursos mais bem avaliados e em instituições mais conhecidas. Também se considera que tais variáveis nada têm a ver com a habilidade intrínseca de um estudante.

Então, sob a hipótese de que a chance de diplomação seria maior em cursos de melhor qualidade e maior idade da IES, e que tais variáveis não estariam relacionadas com a habilidade do estudante, ocorre que estas serviriam como instrumentos. Como será discutido adiante, não se rejeita a hipótese de que as chances de diplomação são maiores em instituições mais bem avaliadas e/ou antigas, o que corrobora com a perspectiva dessa instrumentalização.

Outra fonte de viés na estimação dos parâmetros de uma equação como (2) [por OLS] é que os salários só são observados entre pessoas empregadas, e estas são potencialmente mais habilidosas que as desempregadas. Portanto, ao desconsiderar observações com *missing* de salários, estaria havendo um viés semelhante a uma omissão de variável explicativa importante. Nesse sentido, a literatura costuma aplicar a chamada “correção de Heckman” para mitigar o problema, em que a equação (2) é estimada de forma condicional ao status de empregado ou desempregado – podendo ou não usar instrumentalização.

Em termos sucintos, uma forma de operacionalizar essa correção, comumente chamada de Heckit, é a seguinte: define-se uma *dummy* “estar empregado”, e se estima a probabilidade de observar alguém empregado em face de covariadas; com essas estimativas, computando-se a chamada “razão inversa de Mills”, ou *IMR* pela sigla em inglês; por fim, estima-se a regressão (2) acrescentando a *IMR* como covariada [por OLS ou 2SLS]. Tal procedimento (e suas nuances) está detalhadamente apresentado em Mullahy (1998) ou Wooldridge (2010, p. 563-564), e tem a vantagem de mitigar vieses de seleção amostral na estimação de  $\gamma$ .

Por fim, também existe o potencial viés de auto-seleção na questão da definição do grupo de controle, o que é largamente discutido na literatura de avaliação de políticas públicas (Gertler et al., 2018). De forma resumida, o problema é que uma parte considerável das pessoas mais habilidosas do público-alvo podem se auto-selecionar para participar do programa, então  $\gamma$  não refletiria o ATT do programa, mas o efeito da alta habilidade intrínseca



ao grupo de tratamento. Portanto, o grupo de controle deve ser definido com o intuito de mitigar mais essa fonte de viés.

Observa-se então que o impacto esperado do programa ocorre através da diplomação do beneficiário, que por sua vez melhoraria sua condição no mercado de trabalho. Além disso, como observado em CGU (2015), há um significativo número de bolsistas que evadem ou são desligados do programa antes da diplomação, e depois não tem matrícula encontrada em nenhuma IES por anos subsequentes. Portanto, um grupo de controle composto por essas pessoas pertenceria ao mesmo público-alvo, cuja única diferença aparente com os tratados seria a diplomação, e então isso mitigaria o viés de auto-seleção.

Ainda assim, pode-se imaginar que restaria uma variável omitida dada pela “resiliência” daqueles que continuam estudando até obter a diplomação. Nos termos discutidos acima, se  $o_i$  é essa variável, ocorre  $Cov(d_i, o_i) > 0$  (porque resiliência está associada com maiores chances de diplomação); mas se  $o_i$  também não for observada pelo mercado de trabalho, é possível imaginar que  $\delta = 0$  e, portanto, esse eventual viés seria desprezível. De toda forma, para lidar com esse ponto, a ideia é promover um pareamento por score de propensão (PSM) a fim de mitigar esse potencial viés decorrente da resiliência (Gertler et al. 2018).

Tendo em vista o discutido aqui, se repetirá os exercícios de estimação da equação (2) usando: os estimadores tipicamente usados na literatura (OLS, 2SLS e Heckit com/sem instrumentalização através do IGC e da idade da IES); os ex-bolsistas não diplomados como grupo de controle; e, amostras com e sem PSM. Com base na estimação de  $\gamma$  com esses procedimentos, o posterior exercício de matemática financeira considerará cenários em que o ATT possa estar subestimado ou superestimado.

## 5) Dados

A construção da base de dados partiu do registro dos bolsistas por coorte de ingresso entre 2005 e 2011, no SISPROUNI, contendo: o CPF e o id-MEC; o tipo da bolsa (parcial ou integral); a identificação da IES, do curso, do turno e da modalidade (presencial ou Ead); e, o gênero, a data de nascimento e o identificador se o aluno possui alguma deficiência física. Com o id-MEC, as coortes foram cruzadas no CES do respectivo ano, onde se observa: o gênero, a idade e se é bolsista do ProUni; a situação da matrícula (regular, em fase de conclusão, trancada, descontinuada, transferida e as indicações de formatura ou falecimento); e, características do curso e da IES. Então, fez-se um processo de deduplicação e eliminação de inconsistências – e.g., quando uma pessoa é dita “mulher bolsista” em uma base e “homem não bolsista” em outra. Tal procedimento culminou na observação de 346.810 beneficiários, que é coerente com a estimativa discutida anteriormente de que o ProUni teve mais de 500 mil bolsistas diferentes entre 2005-2018.

Na sequência, esses indivíduos foram rastreados ao longo das edições do CES até 2018, para classificar os diplomados e os que não estavam diplomados, nem falecidos e nem foram mais encontrado em nenhuma IES – sendo estes definidos como “evadidos”. Dentre os indivíduos observados, 84,17% e 15,83% estavam diplomados (grupo de tratamento) e evadidos (grupo de controle), respectivamente, conforme discutido na última seção.

Em seguida, o CPF foi cruzado na RAIS (identificada), onde se observa o salário no mercado de trabalho formal. Nesse ponto, considerou-se que o tempo representativo de diplomação seria de quatro anos, e haveria mais um ano extra para a pessoas se encaixar em um emprego; logo, a coorte que ingressou em 2011 potencialmente estaria no mercado de trabalho com um nível superior em 2016. Dado que a RAIS identificada mais recente que estava disponível no momento da tabulação era a de 2018, a coorte de 2011 foi rastreada em 2016, 2017 e 2018; a coorte de 2010 em 2015, 2016, 2017 e 2018;... ; e, a coorte de 2005 em 2010, 2011,...., 2017 e 2018.

---

### Tabela 4: Observações nos grupos de tratamento e de controle por coorte de ingresso

**no ProUni e presença na RAIS pós ano esperado para a diplomação**

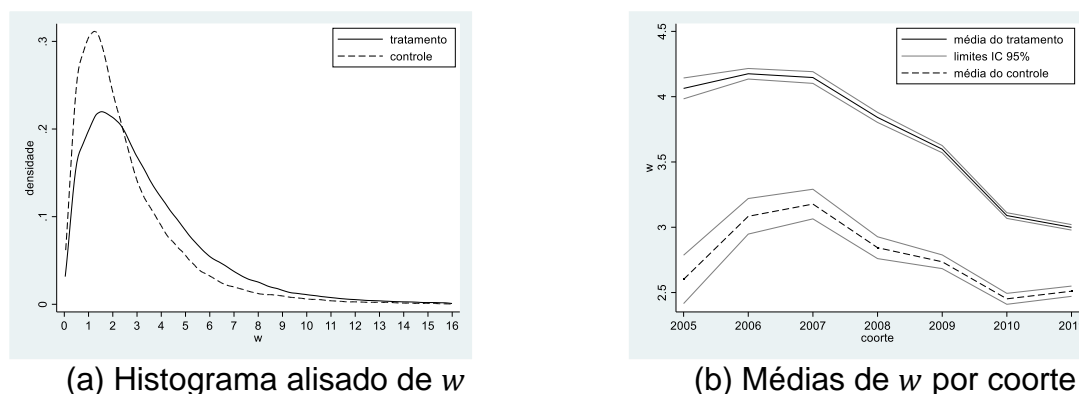
Coorte	Tratamento (Diplomados)		Controle (Evadidos)	
	Não encontrados na RAIS	Encontrados na RAIS	Não encontrados na RAIS	Encontrados na RAIS
2005	1.685 15,3%	9.340 84,7%	182 19,8%	739 80,2%
2006	3.911 12,1%	28.444 87,9%	565 21,7%	2.041 78,3%
2007	3.607 14,1%	22.063 85,9%	1.141 31,3%	2.505 68,7%
2008	4.143 14,3%	24.896 85,7%	1.665 29,8%	3.926 70,2%
2009	7.122 13,6%	45.160 86,4%	2.439 21,5%	8.918 78,5%
2010	11.476 16,7%	57.245 83,3%	3.405 23,7%	10.975 76,3%
2011	15.040 20,7%	57.785 79,3%	4.352 26,5%	12.040 73,5%
Total	46.984 16,1%	244.933 83,9%	13.749 25,0%	41.144 75,0%

Fonte dos dados: MEC e RAIS. Elaboração própria.

A Tabela 4 apresenta o número de observações nos grupos de tratamento e de controle por coorte de ingresso no ProUni e presença na RAIS, entre o ano posterior ao esperado para a diplomação e 2018. Em todas as coortes o percentual de encontrados na RAIS é maior entre os diplomados (tratamento) do que entre os evadidos (controle). No total, 83,9% dos tratados apresentou registro formal de emprego, frente a 75,0% no caso dos controles – uma diferença de 8,9 pontos percentuais (p.p.). Portanto, nota-se maior empregabilidade entre os diplomados.

A partir da RAIS, definiu-se:  $w_{ia}$  como a remuneração média mensal do indivíduo  $i$  no ano  $a$ , em salários mínimos (SM), nos termos do dicionário de microdados; e,  $w_i$  como a média de  $w_{ia}$  nos anos em que o indivíduo  $i$  é observado. Portanto,  $w_i$  se configura como um indicador do salário médio mensal do indivíduo  $i$  no mercado de trabalho formal no período após a diplomação (no caso do tratamento) ou a potencial diplomação (no caso do controle). Assim, no passo final da tabulação se fez um arranjo transversal dos dados por coorte.

**Figura 2: Histograma alisado do indicador de salário mensal em SM ( $w$ ) e intervalos de confiança (95%) das médias do salário por coorte, entre grupos de tratamento e controle**



Fonte dos dados: MEC e RAIS. Elaboração própria.

A Figura 2(a) apresenta um histograma alisado (i.e., densidade *kernel*) de  $w_i$ , entre tratados e controles, onde se observa que as maiores frequências se concentram próximas de 2 SM/mês (R\$ 2.090, considerando o SM de R\$ 1.045 em de 2020), e que a cauda da densidade é mais longa entre os tratados, ilustrando que sua média salarial é maior que a média entre os controles (de fato, a primeira é 3,5 e a segunda é 2,6). Já a Figura 2(b) apresenta esses números por coorte, ilustrando que, nos anos imediatamente seguintes à diplomação, os tratados tendem a ganhar algo entre 0,5 e 1,5 SM/mês a mais que os controles em todas as coortes – R\$ 523 e R\$ 1.568, respectivamente, em valores de 2020.

**Tabela 5: Salário médio ( $w$  em SM mensais) por cursos selecionados, entre grupos de tratamento ( $d = 1$ ) e controle ( $d = 0$ ), e estimativa prévia do valor potencialmente criado pelo programa ( $B$ )**

Curso	$w$ médio		Diferença (c) = (a)-(b)	Diplomados (d)	Prévia de $B$ (c) × (d)
	$d = 1$ (a)	$d = 0$ (b)			
Administração	3,31	2,55	0,76	60.629	46.078
Direito	3,62	3,09	0,53	37.000	19.610
Ciências Contábeis	3,58	2,59	0,99	15.919	15.760
Enfermagem	4,13	2,94	1,19	13.197	15.704
Sistemas de Informação	4,49	2,83	1,66	6.739	11.187
Farmácia	4,93	3,59	1,34	6.445	8.636
Medicina	9,76	7,84	1,92	3.564	6.843
Engenharia Civil	4,34	3,09	1,25	4.579	5.724
Ciência da Computação	4,35	3,09	1,26	4.048	5.100
Comunicação Social	3,11	2,34	0,77	5.623	4.330
Engenharia Elétrica	5,56	4,04	1,52	2.808	4.268
Engenharia de Produção	4,53	3,46	1,07	3.934	4.209
Engenharia Mecânica	5,22	3,54	1,68	2.477	4.161
Psicologia	2,75	2,36	0,39	9.899	3.861
Nutrição	2,83	2,01	0,82	4.389	3.599
Economia	4,67	2,90	1,77	1.962	3.473
Arquitetura e Urbanismo	2,78	2,05	0,73	3.756	2.742
Química	3,96	2,47	1,49	1.747	2.603
Comércio Exterior	4,17	2,83	1,34	1.243	1.666
Engenharia de Computação	4,81	3,18	1,63	980	1.597
Marketing	3,10	2,42	0,68	2.330	1.584
Fisioterapia	2,39	2,18	0,21	7.052	1.481
Biologia	2,66	2,37	0,29	4.710	1.366
Agronomia	3,88	2,68	1,20	1.127	1.352
Engenharia de Automação	4,69	3,75	0,94	1.414	1.329
Matemática	3,60	2,90	0,70	1.885	1.320
Engenharia Ambiental	3,43	2,73	0,70	1.659	1.161
Biomedicina	2,99	2,48	0,51	2.218	1.131
Análise de Sistemas	3,86	3,16	0,70	1.536	1.075
Medicina Veterinária	2,68	2,17	0,51	2.038	1.039

Fonte dos dados: MEC e RAIS. Elaboração própria.

A Tabela 5 mostra os salários médios entre cursos selecionados, separando nas colunas (a) e (b) os grupos de comparação. A coluna (c) apresenta a diferença de salários, (a)-(b), tal que os cursos em que essa diferença não se mostrou estatisticamente diferente de zero não foram listados – é o caso de Pedagogia, por exemplo. Nota-se que em Administração as

médias dos tratados e dos controles são de 3,31 e 2,55 SM/mês, respectivamente; e, que a diferença foi de 0,76 (ou R\$ 794 em valores de 2020). Já em Medicina as médias dos tratados e dos controles são de 9,76 e 7,84 SM/mês, respectivamente; e, que a diferença foi de 1,92 (ou R\$ 2.006 em valores de 2020).

Há pelo menos dois pontos para se registrar em relação aos números acima. Primeiro, que esse exercício indica que o programa tem impacto diferente entre os cursos. Segundo, que a média salarial dos evadidos em certos cursos é maior que a média salarial dos formados em outros cursos – o exemplo disso está nos casos de Medicina (7,84) vis-à-vis Administração (3,31). Imagina-se que isso decorra de variáveis não observadas (e.g., habilidade), no sentido de que na média um egresso de Medicina possa ter mais dessas características que um de Administração. Mas independente de isso ser ou não verdade, nota-se que é fundamental na estimação da equação (2) que se controle pelo curso o indivíduo fez.

A coluna (d) mostra o número de diplomados (tratados) observados por curso na base de dados. Assim, lembrando que o montante  $B = \sum_i (w_{1i} - w_{0i})$  representa um valor potencialmente criado pelo programa, em que  $w_{1i}$  e  $w_{0i}$  representam salários factuais e contrafactuais dos diplomados, a coluna (e) apresenta uma estimativa prévia de  $B$  através do produto entre as colunas (c) e (d). A premissa desse exercício é que  $(c) = (a)-(b)$  seria um indicador inicial para  $w_{1i} - w_{0i}$ , o que não é necessariamente verdade, por conta das várias fontes de viés descritas na seção anterior, mas serve para dar uma noção prévia do impacto do ProUni na massa salarial dos egressos.

Nos termos acima, o curso de Administração geraria maior benefício social, num montante perto de 46 mil SM ao mês, pelos parâmetros da ilustração – algo como R\$ 577 milhões ao ano, em valores de 2020. Complementarmente, o exercício indica que cursos como Medicina geram benefícios individuais relativamente maiores (1,92 SM/mês), mas da perspectiva coletiva geram benefícios relativamente menores (6,8 mil SM/mês), porque Medicina possui um número muito menor de bolsas que, por exemplo, Administração.

Quanto a outras covariadas  $x_{li}$  dos salários, cabe registrar que a idade média nos grupos de tratamento e controle é de 23,5 e 25,5 anos, respectivamente – essa diferença é estatisticamente diferente de zero, e mostra que os evadidos (controles) tendem a ser mais velhos. Os homens representam 58,5% das observações, e ganham na média 20,6% a mais que as mulheres, independente de pertencerem ao grupo de tratamento ou de controle. O número de bolsistas portadores de deficiência é menor que 0,2%; cerca de 3/4 das bolsas são de valor integral; perto de 2/3 dos beneficiários estudaram no turno da noite; cerca de metade dos bolsistas estudaram em universidades, mais do que em faculdades; e, 1/10 dos cursos são em EaD. Embora a base de dados disponha de algumas outras variáveis explicativas para potencialmente se aplicar na equação (2), foram essas as que recorrentemente se mostraram estatisticamente significantes nos exercícios econométricos a serem discutidos adiante.

## 6) Resultados estimados

A Tabela 6 apresenta os parâmetros estimados para a equação (2) e as regressões auxiliares. As covariadas  $x_{li}$  são: *dummies* para coortes, cursos e UF do indivíduo (propositadamente omitidas para poupar espaço); indicador de tratamento/controle  $d$  (ou  $\hat{d}$  quando  $d$  é instrumentalizado em um primeiro estágio); idade da pessoa e seu quadrado, como *proxy* da experiência não observada; *dummies* (=1) para homem, recebimento de bolsa integral, curso feito em EaD, curso presencial feito no turno da noite, curso feito em uma IES classificada como universidade (vis-à-vis uma faculdade ou instituição assemelhada) e registro de deficiência física; *dummies* para as notas do IGC (iguais a 2, 3, 4 ou 5); idade da IES; e, IMR no caso de Heckit.

**Tabela 6: Resultados dos parâmetros estimados para a equação (2) e as regressões auxiliares**

Covariada \ Variável dependente [e procedimento]	ln w	d	ln w	emp	ln w	ln w	ln w	ln w	ln w	ln w
	(i) OLS	(ii) 2SLS 1º estágio	[usando toda a base de dados] (iii) 2SLS 2º estágio	(iv) Probit p/ estimar IMR	(v) Heckit c/ d	(vi) Heckit c/ $\hat{d}$	(vii) OLS	[usando só as observações pareadas] (viii) 2SLS 2º estágio	(ix) Heckit c/ d	(x) Heckit c/ $\hat{d}$
d ou $\hat{d}$	0,279*** (0,012)		0,741*** (0,183)		0,278*** (0,011)	0,718*** (0,181)	0,294*** (0,017)	0,729*** (0,013)	0,225*** (0,013)	0,721*** (0,016)
Idade	0,067*** (0,006)	-0,034*** (0,003)	0,084*** (0,016)	0,004*** (0,001)	0,069*** (0,006)	0,085*** (0,015)	0,087*** (0,016)	0,095*** (0,019)	0,091*** (0,016)	0,098*** (0,015)
Idade <sup>2</sup>	-0,022*** (0,001)	0,005 (0,004)	-0,024*** (0,003)	-0,040*** (0,002)	-0,043*** (0,004)	-0,045*** (0,004)	-0,027*** (0,004)	-0,028*** (0,003)	-0,053*** (0,004)	-0,052*** (0,003)
Homem	0,233*** (0,007)	-0,050*** (0,002)	0,255*** (0,016)	0,005 (0,006)	0,230*** (0,007)	0,251*** (0,017)	0,253*** (0,018)	0,262*** (0,027)	0,250*** (0,019)	0,258*** (0,025)
Bolsa Integral	0,021 (0,012)	0,036*** (0,009)	0,003 (0,012)	0,002 (0,006)	0,022 (0,012)	0,005 (0,012)	-0,001 (0,012)	0,002 (0,013)	0,001 (0,014)	0,002 (0,016)
EaD	-0,076* (0,035)	-0,115** (0,041)	-0,033 (0,025)	0,049 (0,041)	-0,045 (0,038)	-0,004 (0,023)	-0,004* (0,001)	-0,004 (0,024)	-0,003 (0,023)	-0,004 (0,022)
Noturno	0,022** (0,008)	0,004 (0,004)	0,020** (0,006)	0,166*** (0,007)	0,095*** (0,017)	0,096*** (0,017)	0,026*** (0,014)	0,024*** (0,013)	0,017*** (0,014)	0,011*** (0,014)
Universidade	0,046** (0,016)	-0,001 (0,010)	0,040 (0,021)	0,016** (0,006)	0,052** (0,017)	0,046* (0,021)	0,033 (0,021)	0,024 (0,022)	0,041 (0,021)	0,029 (0,022)
Deficiência	-0,013 (0,031)	0,001 (0,015)	-0,013 (0,031)	-0,187*** (0,049)	-0,098** (0,034)	-0,103** (0,031)	-0,044 (0,041)	-0,038 (0,037)	-0,153** (0,044)	-0,143** (0,045)
IGC 2		0,051*** (0,014)								
IGC 3		0,191*** (0,013)								
IGC 4		0,207*** (0,020)								
IGC 5		0,219** (0,041)								
Idade da IES		0,010*** (0,001)								
IMR					1,151*** (0,224)	1,212*** (0,263)			1,476*** (0,287)	1,399*** (0,290)
<i>(constantes e dummies para coortes, cursos e UF foram acrescentadas, mas propositadamente omitidas aqui)</i>										
R <sup>2</sup> (*Pseudo-R <sup>2</sup> )	0,121	0,047	0,111	0,073*	0,122	0,112	0,113	0,086	0,114	0,087

Desvio padrão robusto entre parênteses; \*\*\* p < 0,01 , \*\* p < 0,05 e \* p < 0,10 .

Na coluna (i) estão os resultados de OLS usando todas as observações disponíveis, onde a estimativa do ATT do programa é  $\hat{\gamma} = 0,279$ . Uma vez que a variável dependente é o logaritmo natural do salário e que  $d$  é uma *dummy*, esse valor representa uma semi-elasticidade, então, na média (controlada pelas covariadas, *ceteris paribus*) a interpretação é que um diplomado ganha 27,9% a mais que um não diplomado.

O sinal dos parâmetros associados com a idade é coerente com o que é esperado em uma equação Minceriana: o salário aumenta com o acúmulo de experiência até certa idade, e depois se reduz. Quanto aos parâmetros das outras covariadas, nota-se que, na média (controlada pelas covariadas, *ceteris paribus*): homens ganham 23,3% a mais; quem cursou EaD ganha 7,6% a menos; e, quem frequentou o turno da noite e estudou em uma universidade ganha 2,2% e 4,6% a mais, respectivamente. Por fim, nessa regressão o  $R^2$  foi de 0,121.

Na coluna (ii) estão os resultados de primeiro estágio do 2SLS, regredindo  $d$  com as covariadas do OLS mais as *dummies* para os níveis do IGC e a idade da IES. Assim, nota-se que homens, pessoas mais velhas e quem está em EaD apresentam menores chances de diplomação; e, quem recebe bolsa integral tem maiores chances. Além disso, níveis mais altos de qualidade e mais idade da IES se mostram associados com maiores chances de diplomação, o que depõe em favor de seu uso como instrumentos – muito embora, evidentemente não se possa avaliar a ausência de correlação dessas variáveis com a habilidade não observada.

A partir desse primeiro estágio, define-se  $\hat{d}$  como a projeção linear de  $d$  nos instrumentos e demais covariadas. Portanto,  $\hat{d}$  é interpretável como “probabilidade linear de diplomação”. Nesses termos, a coluna (iii) apresenta os resultados de segundo estágio do 2SLS, em que  $d$  é substituído por  $\hat{d}$  no lado direito da igualdade da equação (2). Consequentemente,  $\hat{\gamma} = 0,741$  significa que um aumento de um p.p. em  $\hat{d}$  aumenta o salário em 0,741%. Logo, a diplomação significa o descolamento de  $\hat{d}$  em 100 p.p., e o aumento de 74,1% no salário esperado, sugerindo que o resultado de OLS para o ATT pode estar subestimado. Os demais parâmetros dessa regressão são muito parecidos com os do OLS.

Na coluna (iii) se definiu uma *dummy emp* = 1 para quem está empregado (i.e., tem  $w$  observado), sendo variável dependente em um Probit regredido nas covariadas do OLS, a fim de estimar a IMR a ser aplicada no Heckit – seguindo o procedimento detalhado em Mullahy (1998) ou Wooldridge (2010, p. 563-564). Nota-se então que a chance de estar empregado aumenta até certa idade, e depois se reduz; e, que é mais (menos) provável observar alguém empregado se esta pessoa estudou a noite (tem uma deficiência). Assim, as colunas (iv) e (v) apresentam os resultados para o Heckit usando  $d$  e  $\hat{d}$ , respectivamente, onde se observam essencialmente os mesmos resultados de OLS e 2SLS.

Por fim, as colunas (vii), (viii), (ix) e (x) apresentam os resultados estimados seguindo os mesmos procedimentos descritos acima, mas usando apenas as observações pareadas por PSM. O protocolo de pareamento seguiu os passos descritos em Gertler et al. (2018): estima-se um Probit sendo  $d$  a variável dependente, usando as covariadas do OLS; computam-se os escores de propensão; define-se um suporte comum e selecionam-se unidades de tratamento/controle semelhantes em escore de propensão, usando o critério do vizinho mais próximo; faz-se um teste de balanceamento para avaliar a qualidade do pareamento; e, então descartam-se unidades não pareadas. Ao cabo, se observam essencialmente os mesmos resultados estimados com as outras especificações.

Dada essa conclusão, repete-se a estimação do OLS usando todas as unidades de observação restritas por curso, em que o ATT aferido [entre colchetes] pode ser apresentado do maior ao menor nos seguintes termos: Farmácia [0,50]; Medicina [0,49]; Sistemas de Informação [0,44]; Enfermagem [0,43]; Engenharia Civil [0,41]; Engenharia Mecânica [0,39]; Economia [0,39]; Ciências da Computação [0,37]; Nutrição [0,36]; Engenharia de Produção [0,35]; Ciências Contábeis [0,33]; Engenharia Elétrica [0,31]; Administração [0,27];

Comunicação Social [0,26]; Arquitetura [0,25]; Marketing [0,25]; Direito [0,19]; Psicologia [0,19]; Biologia [0,19]; e, Fisioterapia [0,10]. Nos demais cursos o ATT estimado foi de menos de 0,10 ou estatisticamente zero.

## 7) Avaliação econômica

Com base no discutido até aqui, retoma-se a ideia do  $VPL^{social}$  expressa na fórmula (1), e se considera  $\gamma = 0,2$ . Esse ATT apreciado é a estimativa de OLS arredondada para baixo, a despeito dos números de 2SLS sugerirem que isso representa um valor subestimado. Portanto, parte-se de uma hipótese conservadora.

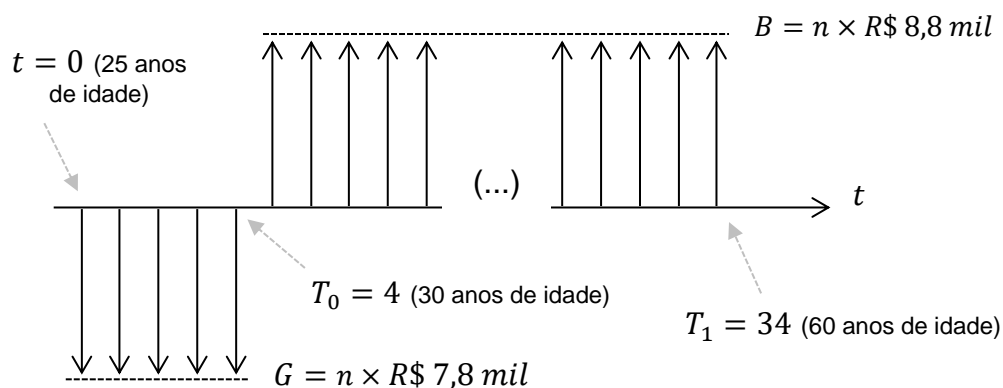
Como  $r = \exp(\gamma) - 1$ , ocorre  $B \cong (r/(1+r)) \times \Sigma_i w_{1i} = 18\% \times \Sigma_i w_{1i}$ , onde  $\Sigma_i w_{1i}$  é a massa salarial dos diplomados. Ou seja, estima-se que 18% dessa massa é causado pelo ProUni, e 82% existiria mesmo na ausência do programa. Portanto,  $B$  representa o benefício coletivo da política, na medida em que é dinheiro criado através do mecanismo das bolsas, para ser consumido, investido, tributado etc.

Como visto anteriormente, o salário médio dos diplomados nos primeiros anos pós-graduação é de 3,5 SM/mês. Então, para  $n$  formados, ocorre  $\Sigma_i w_{1i}/n = 3,5 \times R\$ 1.045 \times 13,3 = R\$ 48,6 \text{ mil/ano}$ , considerando o SM de 2020 (R\$ 1.045) e 13,3 salários anuais (i.e., doze meses mais o décimo terceiro e o adicional de férias). Consequentemente,  $B = n \times R\$ 8,8 \text{ mil/ano}$ .

Sabe-se também que gasto tributário anual para se gerar uma bolsa é de cerca de R\$ 6 mil, em valores de 2020. Mas uma bolsa é potencialmente ocupada por mais de um beneficiário, já que existe algo como 30% de casos de evasão ou de desligamento do programa (por falta de frequência e/ou excesso de reprovações). Portanto, é mais realista considerar  $G = n \times 1,3 \times R\$ 6 \text{ mil/ano} = n \times R\$ 7,8 \text{ mil/ano}$ . Pela mesma razão, também é mais razoável se considerar que uma bolsa precisa de cinco anos para gerar uma diplomação, implicando em  $T_0 = 4$ , embora a formação representativa seja de quatro anos.

Para dar fechamento ao fluxo de caixa social, como representado na Figura 3, falta definir o tempo pelo qual um egresso representativo consegue manter o benefício. Então, lembra-se que um bolsista inicia o curso em torno de 25 anos de idade. Portanto, é realista considerar que sua eventual diplomação será perto de 30 anos de idade. Tendo em conta que, atualmente, a idade mínima para a aposentadoria é de 57 anos, para mulheres, e de 62 anos, para os homens, considera-se que um diplomado trilhará de forma representativa até pelo menos os 60 anos de idade. Consequentemente, é razoável considerar que o benefício seria mantido por algo como 30 anos, o que implica em  $T_1 = 34$ . Então, tem-se que:

**Figura 3: Diagrama do fluxo de caixa social do programa usando os resultados estimados**



$$VPL^{social} = n \times \left( - \sum_{t=0}^4 \frac{7,8}{(1+\tau)^t} + \sum_{t=5}^{34} \frac{8,8}{(1+\tau)^t} \right) \quad (3)$$

A partir de (3),  $VPL^{social} = 0$  implica em uma TIR tal que  $\tau = 16\%$ . Grosso modo, isso significa que se o governo possui uma alternativa de política que gere impactos socioeconômicos de retorno maior que 16% ao ano, e não se dispõe de recursos extras, faz-se razoável considerar a possibilidade de uma diminuição do ProUni em favor dessa alternativa.

Para calcular o  $VPL^{social}$  em si, e um *payback*, é necessário estipular uma taxa de referência  $\tau$ . Nesse sentido, considera-se  $\tau = 5\%$ , sendo equivalente a taxa SELIC atual. A ideia é que, ao invés de gerar gasto tributário e criar bolsas, seria possível executar a cobrança fiscal das IES, e depositar esse valor em fundo remunerado pela SELIC. Assim, em alternativa ao mecanismo de geração de bolsas, o governo poderia simplesmente distribuir esses recursos aos indivíduos do público-alvo ao longo do tempo. Portanto, o  $VPL^{social}$  representaria o quanto se ganha com a existência do ProUni vis-à-vis essa alternativa, mantidas as outras premissas constantes.

Com essa configuração, computa-se  $VPL^{social} = n \times R\$ 76 \text{ mil}$  e um *payback* de 10 anos; e, como o programa já gerou  $n > 500 \text{ mil}$  diplomados, conclui-se que esse investimento já gerou mais de R\$ 38 bilhões desde 2005. Não obstante, é claro que os valores mudam em diferentes cenários, mas com base em tudo o que foi discutido até aqui, esses valores se mostram bem aderentes na realidade.

**Tabela 7: Resultados do contraste do gasto tributário com o impacto na massa salarial dos bolsistas egressos por curso selecionado**

Curso	$\hat{\gamma}$	$\frac{r}{1+r}$	$\frac{\sum_i W_{1i}}{n}$	$\frac{B}{n}$	TIR	$\frac{VPL^{social}}{n}$	Payback
Farmácia	0,50	0,39	4,93	25,45	39,25	204,40	6,43
Medicina	0,49	0,39	9,76	49,67	56,13	424,46	5,28
Sist. Info.	0,44	0,35	4,49	20,94	34,99	163,49	6,25
Enfermagem	0,43	0,35	4,13	19,02	32,98	145,99	6,94
Eng. Civil	0,41	0,33	4,34	19,08	33,05	146,54	6,93
Eng. Mecânica	0,39	0,32	5,22	22,35	36,37	176,22	6,64
Economia	0,39	0,32	4,67	19,78	33,79	152,93	6,86
Computação	0,37	0,31	4,35	17,56	31,38	132,74	7,89
Nutrição	0,36	0,31	2,83	11,40	23,49	76,76	7,63
Eng. Produção	0,35	0,30	4,53	17,70	31,54	134,02	7,91
C. Contábeis	0,33	0,28	3,58	13,21	26,04	93,27	7,86
Eng. Elétrica	0,31	0,27	5,56	19,72	33,73	152,41	6,87
Administração	0,27	0,24	3,31	10,27	21,78	66,53	8,78
Comunicação	0,26	0,23	3,11	9,37	20,33	58,31	9,81
Marketing	0,25	0,22	3,10	8,99	19,69	54,86	9,60
Arquitetura	0,25	0,22	2,78	8,09	18,13	46,70	9,96
Direito	0,19	0,17	3,61	8,25	18,41	48,12	9,85
Psicologia	0,19	0,17	2,75	6,13	14,31	28,91	11,69
Biologia	0,19	0,17	2,66	6,05	14,13	28,15	9,61
Fisioterapia	0,10	0,10	2,39	3,12	6,55	1,52	23,92



Também já evidenciou que o ATT pode mudar consideravelmente entre as diferentes formações. Nesse sentido, a Tabela 7 apresenta os resultados da análise econômica para os cursos em que ocorre  $\hat{\gamma} > 0,1$ , computando-se os componentes do  $VPL^{social}$  seguindo as premissas anteriores. Nas primeiras colunas recuperam-se as estimativas descritas anteriormente, e apresenta-se a fração da massa salarial que pode ser atribuída ao ProUni. Nota-se então que, por exemplo, esse número seria de 39 e de 10% em Farmácia e Fisioterapia, respectivamente.

A coluna seguinte aponta o salário médio observado entre os formados pelo respectivo curso,  $\Sigma_i w_{1i}/n$ , em SM mensais. Nota-se então que uma pessoa formada em Medicina ganha praticamente o dobro de quem é formado em Farmácia, Sistemas de Informação ou Enfermagem; e, cerca de quatro vezes mais de quem é Formado em Fisioterapia, Biologia ou Psicologia.

Conectando os valores anteriores em conformidade com o que foi descrito acima, a próxima coluna mostra a estimativa do benefício social médio do investimento no programa de bolsas,  $B/n$ , em SM mensais. Claramente um diplomado em Medicina apresenta um valor praticamente igual ao dobro (ou mais) de quem é formado em Farmácia, Sistemas de Informação ou Enfermagem; por consequência direta dos salários relativamente maiores. O que se reflete em maiores TIR e VPL médios, e menores *payback*.

Os achados da Tabela 7 permitem avaliar a política de concessão de bolsas do Prouni a partir de alguns pontos. Primeiramente, a despeito das IES contemplarem os beneficiários com uma variedade de cursos disponíveis, somente uma parcela deles apresenta verdadeiramente um grande impacto em termos de viabilidade econômica. Nesse sentido, a opção de cursos com baixo retorno social deveria ser repensada em benefício da adoção de bolsas para estudantes pobres em cursos com maior possibilidade de inclusão social. Por outro lado, o programa concentrou, como um todo, a maioria das bolsas em cursos listados na Tabela 7, indicando que em linhas gerais, o programa proveu contemplou acesso e democratização do ensino com retorno econômico.

Cabe destacar que há uma considerável heterogeneidade na empregabilidade, remuneração e permanência no mercado formal de trabalho a depender da graduação escolhida pelos estudantes. A partir das informações da Tabela 7, é possível visualizar a existência de cursos que apresentam elevado impacto com baixo retorno social (notadamente Medicina) – uma vez que contemplaram uma quantidade limitada de estudantes – ou a existência de cursos com menor impacto (Administração) que, por englobarem um grande contingente de beneficiários, gerou um impacto social agregado muito superior. Com a eventual possibilidade de se alocar as bolsas preferencialmente entre os cursos com maior retorno social, seria de grande importância acompanhar a distribuição das bolsas tendo em vista o trade-off entre a alocação que permite maiores retornos individuais e menor impacto social ou cursos que geram menor impacto individual e, devido ao grande contingente, um grande impacto social.

## 8) Conclusão

O Programa Universidade Para Todos foi instituído em 2005 e promove a concessão de bolsas em Instituições de Ensino Superior para estudantes de baixa renda com a contrapartida de isenção de impostos federais para as IES privadas. Até o ano de 2017, mais de 2 mil instituições do Ensino Superior participaram do programa e ofertaram mais de 2 milhões de bolsas de estudos, parciais e integrais, a um custo fiscal aproximado de R\$ 13 bilhões.

A despeito da sua relevância como política de acesso ao ensino superior, até o presente momento os estudos que avaliaram o impacto do Prouni no mercado formal de trabalho utilizaram questionários limitados por IES ou Unidade da Federação. O presente artigo contribui ao realizar uma avaliação de impacto em escala nacional do Prouni no mercado formal de trabalho, considerando em especial a remuneração e empregabilidade dos seus

egressos. Ademais, uma vez que o programa é financiado por gastos tributários, o artigo propõe uma avaliação de retorno econômico contrastando o benefício social gerado pelo programa – na forma do incremento de remuneração observado entre os beneficiários que concluíram os estudos com uma bolsa do programa frente aos beneficiários evadidos – contra os gastos tributários concedidos pelo Governo Federal anualmente.

Os achados do presente artigo apontam para impactos positivos do Prouni como política pública no mercado de trabalho, através de impactos duradouros e positivos na carreira profissional dos seus egressos. A análise de resultados potenciais permite afirmar que a performance dos egressos do Prouni no mercado de trabalho seria pior – no sentido de baixa remuneração e ocupação – na ausência da política pública.

Também são encontradas evidências da viabilidade econômica do Prouni, responsável por ter gerado um total de R\$ 38 bilhões desde 2005. O impacto da escolha dos estudantes entre os cursos é heterogêneo, a julgar pelo payback de cada uma das graduações. Há cursos que apresentam um elevado impacto na vida dos estudantes e que geram, em função do reduzido número de bolsas, um impacto social moderado enquanto há categorias de cursos que, a despeito de gerarem um menor impacto individual, é efetivo ao gerar consideráveis impactos sociais através de um elevado número de bolsas ofertadas. E, ao mesmo tempo, também existem opções de cursos que não são viáveis economicamente, gerando pouco ou nenhum benefício para os estudantes no mercado de trabalho dos anos seguintes. Entende-se que o Inep e a Secretaria de Ensino Superior do MEC teriam papel fundamental em desenvolver ferramentas e regramentos que permitam para os beneficiários a escolha de cursos que representem maiores possibilidades de ascensão social dos estudantes e superação da pobreza.

Há algumas limitações e desenvolvimentos futuros para a presente agenda de pesquisa. Primeiramente, cabe avaliar o matching no mercado de trabalho formal entre os egressos do Prouni, a fim de avaliar se trabalhadores estão ocupados em áreas correlatas às áreas de formação. Uma vez que consideramos nos grupos de tratamento e controle egressos e evadidos do Prouni, seria fundamental compreender em que medida a habilidade do estudante afeta o seu desempenho no mercado de trabalho. Pesquisas futuras buscarão lidar com a habilidade ao avaliar o desempenho dos estudantes no Exame Nacional do Ensino Médio e no Exame Nacional de Desempenho dos Estudantes.

## 9) Referências

- Andriola, W. B e Barrozo-Filho, J. L. “Avaliação de políticas públicas para a educação superior: o caso do Programa Universidade Para Todos (ProUni)”, *Revista da Avaliação da Educação Superior* 25 (2020): 594-621. <https://doi.org/10.1590/S1414-40772020000300005>
- Barbosa, M. P.; Petterini, F. C.; Ferreira, R. T. “Expansion of Brazilian Federal Universities: is it possible to raise economic impacts?”, *Revista de Administração Contemporânea* 24, (2019): 1-24. <https://doi.org/10.1590/1982-7849rac2020190230>
- Barbosa-Filho, F. H.; Pessoa, S. A. “Educação e crescimento: o que a evidência empírica e teórica mostra?”. *Revista Economia* 11 no.2 (2010): 265-303, (2010). [http://www.anpec.org.br/revista/vol11/vol11n2p265\\_303.pdf](http://www.anpec.org.br/revista/vol11/vol11n2p265_303.pdf)
- Becker, K. L.; Mendonça, M. J. “Avaliação de impacto do ProUni sobre a performance acadêmica dos estudantes”. *Texto para Discussão Ipea* 2512: 1-38, (2019). [https://www.ipea.gov.br/portal/images/stories/PDFs/TDs/td\\_2512.pdf](https://www.ipea.gov.br/portal/images/stories/PDFs/TDs/td_2512.pdf)
- Brasil. “Relatório de Auditoria Operacional: Programa Universidade para Todos (ProUni) e Fundo de Financiamento ao Estudante do Ensino Superior (FIES)”. *Tribunal de Contas da União*, (2009). [http://portal2.tcu.gov.br/portal/page/portal/TCU/comunidades/programas\\_governo/auditorias](http://portal2.tcu.gov.br/portal/page/portal/TCU/comunidades/programas_governo/auditorias)
- Brasil. “Relatório de Avaliação da Execução de Programas de Governo Nº 35: ProUni”. *Controladoria Geral da União* (2015). <https://eaud.cgu.gov.br/relatorios/>

- Corbucci, P. R. "Desafios da educação superior e desenvolvimento no Brasil". *Texto para Discussão Ipea* 1287: 1-35, (2007). [https://www.ipea.gov.br/portal/images/stories/PDFs/TDs/td\\_1287.pdf](https://www.ipea.gov.br/portal/images/stories/PDFs/TDs/td_1287.pdf)
- Corbucci, P. R. "Evolução do acesso de jovens à educação superior no Brasil". *Texto para Discussão Ipea* 1950: 1-40, (2014). [https://www.ipea.gov.br/portal/images/stories/PDFs/TDs/td\\_1950.pdf](https://www.ipea.gov.br/portal/images/stories/PDFs/TDs/td_1950.pdf)
- Corbucci, P. R; Kubota, L. C.; Meira, A. P. B. "Reconfiguração Estrutural da Educação Superior Privada no Brasil: nova fase da mercantilização do ensino". *Texto para Discussão Ipea* 2256: 1-42, (2016). [https://www.ipea.gov.br/portal/images/stories/PDFs/TDs/td\\_2256.pdf](https://www.ipea.gov.br/portal/images/stories/PDFs/TDs/td_2256.pdf)
- Darolia, R. "Working (and studying) day and night: heterogeneous effects of working on the academic performance of full-time and part-time students". *Economics of Education Review* 38: 38-50, (2014). <https://doi.org/10.1016/j.econedurev.2013.10.004>
- De Negri, J. A.; Castro, P. F. D.; Souza, N. R. D.; Arbache, J. S. "Mercado Formal de Trabalho: Comparação entre os Microdados da RAIS e da PNAD". *Texto para Discussão Ipea* 840: 1-29, (2001). [http://repositorio.ipea.gov.br/bitstream/11058/2155/1/TD\\_840.pdf](http://repositorio.ipea.gov.br/bitstream/11058/2155/1/TD_840.pdf)
- Durham, E. Educação superior, pública e privada (1808-2000). In: BROCK, C., SCHWARTZMAN, S. (orgs.). "Os desafios da educação no Brasil". Rio de Janeiro: Nova Fronteira 197-240 (2005). <http://www.schwartzman.org.br/simon/desafios/7superior.pdf>
- Dynarski, S. "Does aid matter? Measuring the effect of student aid on college attendance and completion". *American Economic Review* 93 no.1: 279-288, (2003). <https://www.jstor.org/stable/3132174>
- Dynarski, S. "Hope for Whom? Financial Aid for the Middle Class and Its Impact on College Attendance," *National Tax Journal* 53, no. 3: 629-662, (2000). <https://www.journals.uchicago.edu/doi/10.17310/ntj.2000.3S.02>
- Faveri, D. B.; Petterini, F. C.; Barbosa, M. P. "Uma avaliação de impacto da política de expansão dos IFs nas economias dos municípios brasileiros". *Planejamento e Políticas Públicas* 50: 125-147, (2018). <https://www.ipea.gov.br/ppp/index.php/PPP/article/view/742/464>
- Gertler, P. J.; Martinez, S.; Premand, P.; Rawlings, L. B.; Vermeersch, C. M. "Impact evaluation in practice". *The World Bank*, (2016). <https://www.worldbank.org/en/programs/sief-trust-fund/publication/impact-evaluation-in-practice>
- Haas, C. M.; Pardo, R. D. S. "Programa Universidade para Todos (ProUni): efeitos financeiros em uma instituição de educação superior privada". *Avaliação: Revista da Avaliação da Educação Superior (Campinas)* 22, no.: 718-740, (2017). <https://doi.org/10.1590/S1414-40772017000300008>
- Heckman, James J.; Lochner, Lance J.; Todd, Petra E. "Earnings functions, rates of return and treatment effects: The Mincer equation and beyond". *Handbook of the Economics of Education* 1: 307-458, (2006). [https://doi.org/10.1016/S1574-0692\(06\)01007-5](https://doi.org/10.1016/S1574-0692(06)01007-5)
- Katovich, E. S.; Maia, A. G. "A relação entre produtividade de trabalho e salário no Brasil: uma análise setorial". *Nova Economia* 28, no. 1: 7-38, (2018). <https://doi.org/10.1590/0103-6351/3943>
- Lépine, A. "Financial aid and student performance in college: evidence from Brazil." *Brazilian Review of Econometrics* 38, no. 2: 221-261(2018). <https://doi.org/10.12660/bre.v38n22018.75505>
- Mincer, Jacob. "Economic development, growth of human capital, and the dynamics of the wage structure". *Journal of Economic Growth* 1, no. 1: 29-48, (1996). <https://link.springer.com/content/pdf/10.1007/BF00163341.pdf>
- Mincer, Jacob. "The distribution of labor incomes: a survey with special reference to the human capital approach". *Journal of economic literature* 8, no. 1: 1-26, (1970). <https://www.jstor.org/stable/2720384>

- Mullahy, J. "Much ado about two: reconsidering retransformation and the two-part model in health econometrics". *Journal of health economics* 17, no. 3: 247-281, (1998). [https://doi.org/10.1016/S0167-6296\(98\)00030-7](https://doi.org/10.1016/S0167-6296(98)00030-7)
- Saraiva, Luiz Alex Silva et al. "A efetividade de programas sociais de acesso à educação superior: o caso do ProUni". *Revista de Administração Pública* 45: 941-964, (2011). <https://doi.org/10.1590/S0034-76122011000400003>
- Stinebrickner, R.; Stinebrickner, T. R. "Working during school and academic performance". *Journal of Labor Economics* 21, no. 2: 473-491, (2003). <https://doi.org/10.1086/345565>
- Taber, Christopher R. "The rising college premium in the eighties: Return to college or return to unobserved ability?" *The Review of Economic Studies* 68, no. 3: 665-691, (2001). <https://www.jstor.org/stable/2695901>
- Tachibana, T. Y.; Menezes-Filho, N. A.; Komatsu, B. K. "Ensino superior no Brasil". *Policy Paper INSPER* 4: 1-53, (2015). <https://www.insper.edu.br/wp-content/uploads/2018/09/Ensino-superior-no-Brasil.pdf>
- Wooldridge, J. M. "Econometric analysis of cross section and panel data". *MIT press* (2010). <https://mitpress.mit.edu/books/econometric-analysis-cross-section-and-panel-data-second-edition>