

REVISTA

**CADERNOS DE
FINANÇAS
PÚBLICAS**

2022
EDIÇÃO ESPECIAL

Uma avaliação econômica do ProUni contrastando a massa salarial dos egressos com o gasto tributário do programa

Vinícius Augusto Lima de Almeida

Economista da Confederação Nacional de Municípios

Francis Carlo Petterini Lourenço

Professor da UFSC

RESUMO

O ProUni é uma política pública federal de isenções fiscais para Instituições de Ensino Superior privadas que concedem de bolsas de estudo para alunos de baixa renda. O programa já custou mais de R\$ 13 bilhões, e formou cerca de 500 mil estudantes. De forma inédita, essa pesquisa rastreou milhares de egressos na RAIS (identificada), observando que um graduado pelo ProUni recebe, na média, 30% a mais do que um equivalente não graduado. Convertendo esse impacto em massa salarial, a fim de contrastá-la com o gasto tributário em um fluxo de caixa, conclui-se que, em um cenário conservador, o Valor Presente Líquido da política já seria de mais de R\$ 38 bilhões. Com base em análises desagregadas por cursos, recomenda-se a focalização das bolsas em graduações que potencializam o benefício coletivo.

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

Classificação JEL: I28, J21, J24.

SUMÁRIO

RESUMO	2
1. Introdução.....	4
2. ProUni.....	6
3. Avaliação econômica.....	9
3.1 Adaptação da equação Minceriana	9
3.2 Estratégia econométrica	10
3.3 Valor presente líquido social.....	12
4. Dados.....	13
5. Resultados econométricos.....	18
6. Matemática financeira	21
7. Conclusão	24
8. Referências.....	25

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; Corbucci 2014; 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 permitiu a graduação de cerca 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. Ao que parece, porque não precisam trabalhar para pagar a faculdade. Logo, isso poderia melhorar o capital humano nacional, gerando um benefício coletivo futuro 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 da monografia é avançar em relação aos traba-

lhos 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 inferir 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 estimado com a isenção fiscal observada.

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 da monografia é fomentar sugestões de melhoria no desenho da política, indicando se algumas 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 (SIS-PROUNI, 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: nos primeiros anos depois de formado, na média, um egresso ganha 30% a mais do que ganharia o seu contrafactual sem formação. Esse número é maior em formações como Farmácia ou Medicina; por outro lado, em muitos cursos o efeito é estatisticamente zero.

Ao converter esses impactos em massa salarial, contrasta-se isso com os registros contábeis de renúncia fiscal, e então, em um cenário conservador, 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 equacionaria 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 o número de beneficiários é muito grande nesses cursos.

Conclui-se que é possível melhorar a focalização das bolsas, a fim de potencializar o benefício coletivo do programa. Em específico, e de forma diferente do que é feito atualmente, propõe-se a criação de uma regra que condicione o montante de isenção fiscal das IES à oferta de bolsas em graduações onde há evidências de impacto na empregabilidade dos egressos. Em especial, essa regra poderia ser modificada anualmente, com base que for observado na RAIS de anos imediatamente anteriores.

Além dessa introdução, a monografia tem a seguinte forma: na Seção 2 o ProUni é detalhado; na Seção 3 se discute a estratégia de avaliação econômica; na Seção 4 a construção da base de dados e as estatísticas descritivas são apresentadas; na Seção 5 têm-se os resultados econométricos; na Seção 6 discute-se a matemática financeira do programa; e, por fim, a Seção 7 detalha a conclusão.

2. ProUni

O ProUni foi instituído em 2005 pela Lei N° 11.096 (alterando a Lei N° 10.891, de 9 de julho de 2004), 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, com ou sem fins lucrativos, que oferecem bolsas de estudo (integrais ou parciais, estas de 25% ou 50%) em cursos de nível superior (na forma de graduação ou sequencial, sendo presenciais ou em EaD). Apesar das nuances legais, a regra básica é dar isenção total quando a oferta atinge a proporção mínima de uma bolsa integral (por curso) a cada 11 alunos pagantes (em cada curso da IES).

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 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, usando 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 um e meio salário mínimo (SM) per capita para as bolsas integrais, ou até três 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, e em determinadas circunstâncias. Uma vez no programa, para não ser desligado, o bolsista deverá cumprir frequência e aprovações mínimas nas disciplinas cursadas.

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

Ano	Renúncia (R\$ milhões*) [a]	Bolsas (milhares) [b]	1.000×a/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.

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.

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 (pelo menos uma não sequencial), 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 que meio milhão. Nesse sentido, mesmo que não se tenha encontrado nenhuma estatística oficial sobre o número de formados através do ProUni, por essa lógica é razoável afirmar que cerca de 500 mil estudantes já foram graduados pelo programa até hoje.

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 primários, mas ainda assim a estimativa do gasto anual por bolsa fica próxima disso. Inclusive, sob a hipótese que se leva em média quatro anos para uma diplomação, nota-se que uma graduação média custaria ao erário algo como R\$ 24 mil, ou R\$ 500 por mês, em valores de 2020 (talvez um pouco mais devido à evasão).

Outro ponto para se registrar é que existe algum debate entre apoiadores e críticos do programa (Saraiva et al., 2011). Entre os primeiros, o principal argumento é que se trata de uma política bem focalizada em termos de público-alvo, porque promove o acesso de estudantes de baixa renda ao ensino superior, de acordo com o desempenho (meritocracia) no Enem, que é uma avaliação padronizada e de abrangência nacional.

Por outro lado, muitos críticos argumentam que o ProUni permite o acesso a uma educação superior em um subconjunto restrito de cursos. 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.

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	Arquitetura	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	Comunicação 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 Computação	1,50	74,87	30	Matemática	0,66	89,61

Fonte dos dados: MEC. Elaboração própria.

Todavia, 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; pode não caracterizar nenhuma medida de má fé ou algo assemelhado; e, talvez seja simplesmente uma falha no desenho da política. A listagem da Tabela 3 segue com os 30 cursos que responderam por quase 90% da oferta entre 2005 e 2017, sem diferenciar entre os presenciais e os em EaD, sendo que o repertório completo conta com centenas de cursos entre graduações e níveis sequenciais.

Eventualmente, alguns críticos do ProUni também declaram que o programa beneficia instituições de qualidade duvidosa. Quanto a esse argumento, 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; e, 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 de uma política pública trata da estimação do seu impacto sobre o público-alvo, e o contraste disso com os seus custos. Seguindo a praxe da literatura, nessa monografia se define resultados potenciais w_{1i} e w_{0i} como salários factuais e contrafactuais, respectivamente, do egresso i diplomado no ProUni. Posto de outra forma, w_{1i} é um salário observado e w_{0i} é um salário que seria observado caso i não tivesse concluído o ensino superior. Assim, $w_{1i} - w_{0i}$ representa um impacto do programa – i.e., um efeito do tratamento “estar mais qualificado”, na perspectiva individual do salário de um beneficiário.

Portanto, $B = \sum^i (w_{1i} - w_{0i}) = n \times (\bar{w}_{1i} - \bar{w}_{0i})$ é uma massa salarial causada pelo programa, considerando n beneficiários e as médias aritméticas de w_{1i} e w_{0i} representadas por \bar{w}_{1i} e \bar{w}_{0i} , respectivamente. Assim, $\bar{w}_{1i} - \bar{w}_{0i}$ representa um “efeito tratamento médio”.

Esse montante B 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). Então, antes de tudo é preciso uma estratégia para identificar $\bar{w}_{1i} - \bar{w}_{0i}$.

3.1. Adaptação da equação Minceriana

Para contornar o do problema fundamental da inferência causal, a estratégia adotada nessa monografia 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

$V_d = \int_{s_d}^{nd} w_d \exp(-\gamma t) dt$, 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, Mincer (1970) demonstra que a seguinte igualdade é verdadeira: $w_1 \exp(-\gamma s_1) (1 - \exp(-\gamma n_1)) = w_0 \exp(-\gamma s_0) (1 - \exp(-\gamma n_0))$.

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 $w \cong w_0 (\exp(\gamma))^d$. Isso, por sua vez, 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 adaptada para o contexto dessa monografia.

3.2. Estratégia econométrica

Partindo da equação Minceriana, reescreve-se $\ln w_0 = \text{cte} + \sum_{i=1}^L \beta_i x_i$, em que “cte” é uma constante e cada β_i e x_i 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). Assim, dada uma amostra $\{w_i, d_i, x_{1i}, \dots, x_{Li}\}_{i=1}^N$, em que i indexa indivíduos, sendo que $n < N$ deles foram diplomados pelo ProUni, a equação de base do exercício econométrico é:

$$\ln w_i = \text{cte} + \gamma d_i + \sum_{i=1}^L \beta_i x_{ii} + u_i \quad (1)$$

Dessa forma, desde que d_i indique apropriadamente o pertencimento a grupos de tratamento e de controle, γ também pode ser estimado e 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). Notadamente, nesse caso particular, γ deve ser interpretado em termos de semi-elasticidade do salário, porque a variável dependente está em logaritmo e d_i é uma dummy.

A estimação dos parâmetros da equação (1) deve considerar a endogeneidade de d_i , porque o salário é influenciado não só por características observáveis x_{ii} , 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 δ é um parâmetro que associa a variável omitida o_i com d_i em (1), 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 γ , denotado de $\hat{\gamma}$, é tal que

plim $\hat{\gamma} = \gamma + \delta \times \text{“Cov”}(d_i, o_i) / \text{“Var”}(d_i)$, onde “Cov” e “Var” indicam 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 (“Cov” $(d_i, o_i) > 0$), ocorre que $\hat{\gamma}$ superestima γ . 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 “Cov” $(d_i, o_i) < 0$ e então $\hat{\gamma}$ subestima γ .

Portanto, 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 (no caso, d_i), mas que não com a habilidade. Sendo que, nos trabalhos 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, de forma semelhante aos trabalhos de Dale e Krueger (2002) e Borgen (2014), 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, e o teste de Durbin-Wu-Hausman corrobora com a perspectiva dessa instrumentalização.

Outra fonte de viés na estimação dos parâmetros de uma equação como (1) 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, potencialmente 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 (1) é estimada de forma condicional ao status de (formalmente) empregado ou (formalmente) 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 (formalmente) empregado”, e se estima a probabilidade de observar alguém (formalmente) 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 γ .

Por fim, e não menos relevante, 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 γ 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.

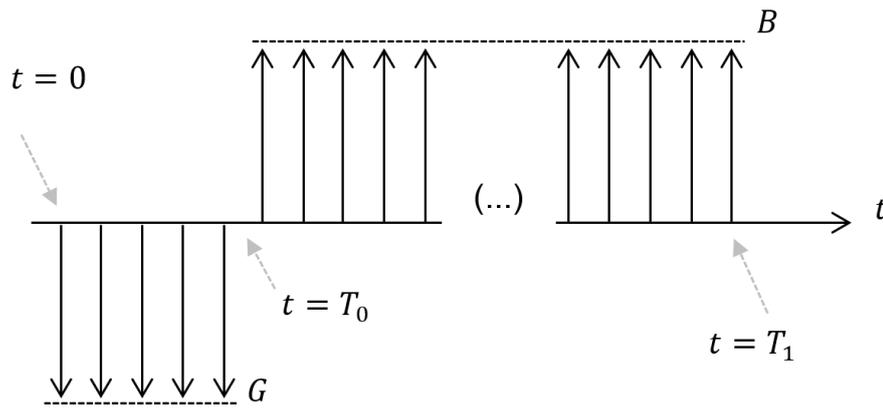
Ainda assim, pode-se imaginar que restaria uma variável omitida dada pela “resiliência” daqueles que continuam estudando até obter a diplomação. Nos termos discutidos acima, se o_i é essa variável, ocorre “Cov” (d_i, o_i) > 0 (porque resiliência está associada com maiores chances de diplomação); mas 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 escore 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 (1) usando: os estimadores tipicamente usados na literatura (OLS, 2SLS e Heckit com/sem instrumentalização através do IGC e da idade da IES); o teste Durbin-Wu-Hausman para avaliar a endogeneidade de d_i ; os ex-bolsistas não diplomados serão usados como grupo de controle; e, amostras com e sem PSM. Com base na estimação de γ com esses procedimentos, o posterior exercício de matemática financeira ainda considerará um cenário em que o ATT possa estar superestimado.

3.3. Valor presente líquido social

Tendo em mãos uma estimativa de γ , 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 computável.

Figura 1: Diagrama do fluxo de caixa social do programa



$$VPL^{social} = - \sum_{t=0}^{T_0} \frac{G}{(1 + \tau)^t} + \sum_{t=T_0+1}^{T_1} \frac{B}{(1 + \tau)^t} \quad (2)$$

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 (2), para alguma taxa de referência τ ; com isso, se pode estimar a TIR e o payback para a coletividade.

4. 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.

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.

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.

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 as pessoas se encaixarem 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. Assim, 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.

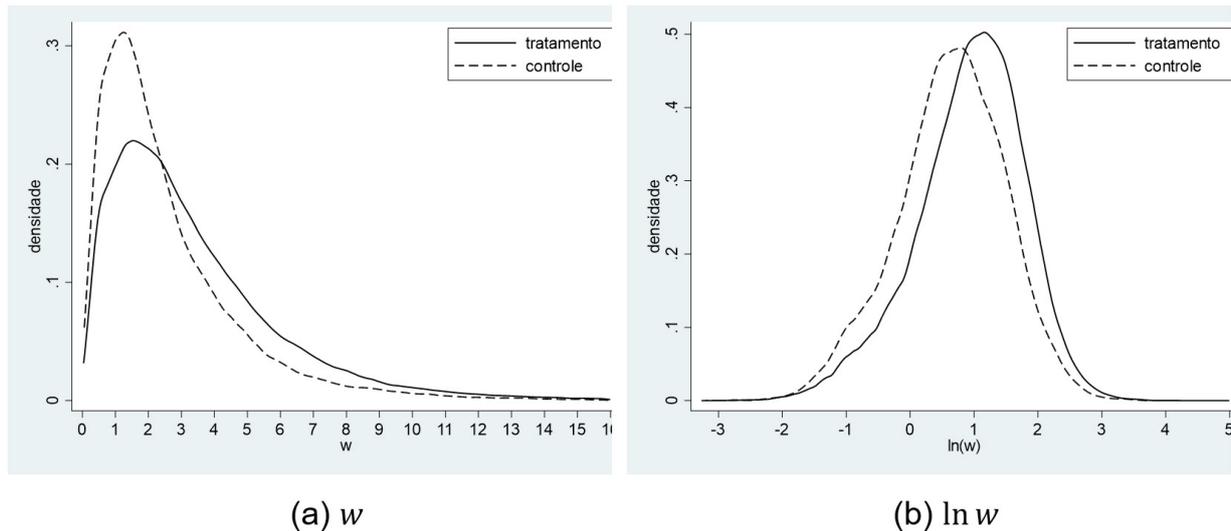
Em todas as coortes o percentual de encontrados na RAIS é maior entre os diplomados (tratamento) do que entre os evadidos (controle). No total, 83,9% dos tratados apresentou registro formal de emprego, frente a 75,0% no caso dos controles – uma diferença de 8,9 pontos percentuais (p.p.). Portanto, nota-se 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.

A Figura 2(a) apresenta um histograma alisado (i.e., densidade kernel) de w_i , entre tratados e controles. A média observada entre os tratados é de 3,50 SM, e entre os controles é de 2,65 SM; a diferença de médias, portanto, é de 0,85 SM ou R\$ 888 em 2020 [com intervalo de 95% confiança entre

0,82 e 0,88 SM]. A Figura 2(b) é análoga, considerando $\ln w_i$. Nesse sentido, a média observada entre os tratados é de 0,93, e entre os controles é de 0,63; a diferença de médias, portanto, é de 0,30 [com intervalo de 95% confiança entre 0,29 e 0,31]. Conseqüentemente, tem-se uma estimativa preliminar de γ , sem considerar qualquer covariada na equação (1), representando um ATT próximo de 30% em termos de semi-elasticidade do salário ao tratamento.

Figura 2: Histograma alisado do indicador de salário mensal em SM (w) entre grupos de tratamento e controle



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 importante na estimação da equação (1) que se controle pelo curso o indivíduo fez.

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 coluna (d) mostra o número de diplomados (tratados) observados por curso na base de dados. Assim, lembrando que o montante $B = \sum_i (w_{1i} - w_{0i}) = n \times (\bar{w}_{1i} - \bar{w}_{0i})$ 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 $\bar{w}_{1i} - \bar{w}_{0i}$, 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, em um 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.

Nessa linha de raciocínio, a Tabela 6 repete o exercício anterior substituindo cursos por um conjunto de UF – não todas, apenas onde a diferença (a)-(b) se mostrou estatisticamente diferente de zero. Dessa forma, nota-se que apesar do benefício individual da diplomação ser potencialmente maior no Distrito Federal (1,25 SM), o resultado coletivo seria substancialmente maior em São Paulo, porque há muitos beneficiários no estado. Em suma, os números das Tabelas 5 e 6 mostram que, da perspectiva coletiva, mesmo que a diplomação em determinado curso (e/ou local) tenha um impacto individual pequeno, da perspectiva coletiva esse impacto pode ser grande em face de um grande número de beneficiários.

Quanto a outras covariadas x_{ii} dos salários, cabe registrar que a média de idade 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.

Tabela 6: Salário médio (w em SM mensais) por UF selecionadas, entre grupos de tratamento ($d=1$) e controle ($d=0$), e estimativa prévia do valor potencialmente criado pelo programa (B)

UF	w médio		Diferença (c) = (a)-(b)	Diplomados (d)	Prévia de B (c) × (d)
	d=1 (a)	d=0 (b)			
São Paulo	3,87	2,79	1,08	96.823	104.307
Minas Gerais	3,39	2,56	0,83	38.781	32.377
Rio Grande do Sul	3,36	2,64	0,72	28.280	20.496
Paraná	3,43	2,63	0,80	22.780	18.137
Rio de Janeiro	3,37	2,55	0,82	16.109	13.143
Distrito Federal	4,30	3,05	1,25	6.016	7.505
Goiás	3,42	2,65	0,77	8.774	6.769
Bahia	2,99	2,59	0,41	13.904	5.632
Amazonas	3,17	2,31	0,86	5.298	4.570

Mato Grosso	3,53	2,75	0,78	5.262	4.112
Santa Catarina	3,26	2,60	0,66	6.024	4.000
Pernambuco	3,20	2,59	0,61	5.539	3.356
Espírito Santo	3,28	2,51	0,77	4.230	3.248
Mato Grosso do Sul	3,55	2,54	1,00	3.146	3.156
Ceará	3,21	2,51	0,70	4.082	2.850
Rondônia	3,29	2,36	0,93	3.038	2.834
Pará	3,09	2,53	0,56	4.316	2.409
Piauí	2,76	2,28	0,48	2.300	1.107
Rio Grande do Norte	2,14	1,92	0,22	4.785	1.040

Fonte dos dados: MEC e RAIS. Elaboração própria.

5. Resultados econométricos

A Tabela 7 apresenta os parâmetros estimados para a equação (1) e as regressões auxiliares. Além do indicador de tratamento/controle d (ou d^* quando d é instrumentalizado em um primeiro estágio), as covariadas x_{it} são: dummies para coortes, cursos e UF do indivíduo (propositadamente omitidas para poupar uma discussão tediosa, que pouco teria para contribuir com a análise); 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 noturno, curso feito em uma IES classificada como universidade (vis-à-vis uma faculdade ou instituição assemelhada) e registro de deficiência física. No caso do primeiro estágio de 2SLS, acrescenta-se dummies para as notas do IGC (iguais a 2, 3, 4 ou 5) e a idade da IES; e, no caso de Heckit, acrescenta-se a IMR.

Na coluna (i) estão os resultados de OLS usando todas as observações disponíveis, onde a estimativa do ATT é $\gamma^{OLS} = 0,279$. Lembrando que, uma vez que a o salário está em logaritmo e que d é uma dummy, esse valor representa uma semi-elasticidade. Portanto, na média (controlada pelas covariadas, ceteris paribus), a interpretação é que um diplomado ganha 27,9% a mais que um não diplomado – perto dos 30% discutidos anteriormente nas estatísticas descritivas.

O sinal dos parâmetros associados com a idade é coerente com o que é esperado em uma equação Minceriana: o salário aumenta com o acumulo 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 estudou em EaD apresentam menores chances de diplomação; e, quem recebeu 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 estimativa da “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 (1). Consequentemente, $\gamma^{2SLS} = 0,741$ representa a diferença de expectativas $E(\ln w | \hat{d}+1, \text{covariadas}) - E(\ln w | \hat{d}, \text{covariadas})$, e não é diretamente interpretável como uma semi-elasticidade.

Para comparar a magnitude das estimativas de OLS e de 2SLS, uma alternativa é computar $E(\ln w | \hat{d}^{\text{máx}}, \text{covariadas}) - E(\ln w | \hat{d}^{\text{mín}}, \text{covariadas}) = \gamma^{2SLS} \times (\hat{d}^{\text{máx}} - \hat{d}^{\text{mín}}) = 0,453$, para os valores máximo e mínimo de \hat{d} [e lembrando que no caso de uma semi-elasticidade isso seria equivalente a definir $\hat{d}^{\text{máx}} = 1$ e $\hat{d}^{\text{mín}} = 0$]. Como esse valor é maior que γ^{OLS} , nota-se que o resultado de OLS para o ATT deve estar subestimado. De fato, no teste de Durbin-Wu-Hausman, cuja estatística foi de 335,26 para 127 covariadas, rejeita-se a hipótese nula de mesmas estimativas geradas por OLS e 2SLS, indicando a endogeneidade de d e sugerindo que OLS subestima o ATT – o que é um resultado coerente, por exemplo, com a discussão apresentada por Barbosa-Filho e Pessoa (2010).

Na coluna (iii) se define uma dummy $\text{emp} = 1$ para quem está empregado (i.e., tem w observado na RAIS), 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.

Tabela 7: 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	
	[usando toda a base de dados]						[usando só as observações pareadas]				
	(i) OLS	(ii) 2SLS 1º estágio	(iii) 2SLS 2º estágio	(iv) Probit p/ estimar IMR	(v) Heckit c/ d	(vi) Heckit c/ d [^]	(vii) OLS	(viii) 2SLS 2º estágio	(ix) Heckit c/ d	(x) Heckit c/ d [^]	
d ou 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 ²	-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)	
(constantes e dummies para coortes, cursos e UF foram acrescentadas, mas propositadamente omitidas aqui)											
R ² (*Pseudo-R ²)	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 sequência, 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) (e os resultados estimados são propositadamente omitidos, como em caso anterior, para poupar uma discussão tediosa, que pouco teria para contribuir com a análise): 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/controlado 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 última 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.

6. Matemática financeira

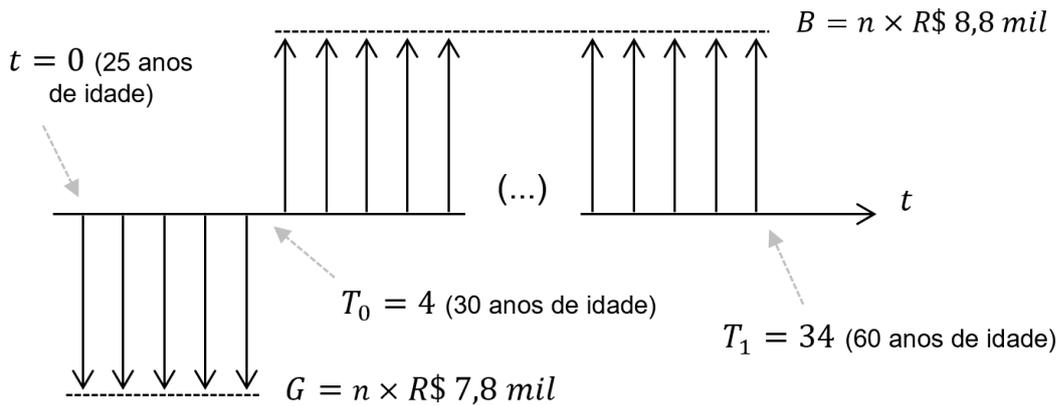
Com base no discutido até aqui, retoma-se a ideia do VPL^{social} expressa na fórmula (2), e se considera $\gamma=0,2$. Esse ATT apreciado é a estimativa de OLS arredondada para baixo, a despeito dos números de 2SLS sugerirem que isso representa um valor subestimado. Portanto, parte-se de uma hipótese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

Como $r = \exp(\gamma) - 1$, ocorre $B \cong (r/(1+r)) \times \sum_i w_{ii} = 18\% \times \sum_i w_{ii}$, onde $\sum_i w_{ii}$ é a massa salarial dos diplomados. Ou seja, estima-se que 18% dessa massa seria causado pelo ProUni, e 82% existiria mesmo na ausência do programa. Portanto, B representa o benefício coletivo da política, na medida em que seria 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-formação é de 3,50 SM/mês. Então, para n formados, ocorre $\sum_i w_{ii}/n = 3,5 \times R\$ 1.045 \times 13,3 = R\$ 48,6$ 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, $B = n \times R\$ 8,8$ 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 perto de 30% de casos de evasão ou de desligamento do programa (por falta de frequência e/ou excesso de reprovações). Portanto, é mais realista considerar algo como $G = n \times 1,3 \times R\$ 6$ mil/ano = $n \times R\$ 7,8$ mil/ano. Por conta desses atritos na alocação das bolsas, também é mais razoável considerar cinco anos de isenção fiscal para gerar uma diplomação, implicando em $T_0=4$, mesmo se a formação representativa for de quatro anos.

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



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

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 tralhará de forma representativa até pelo menos os 60 anos de idade. Conseqüentemente, é razoável considerar que o benefício seria mantido por algo como 30 anos, o que implica em $T_1 = 34$, e na fórmula do VPL^{social} dada pela equação (3).

Dessa forma, $VPL^{social} = 0$ 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 (ou extinção) do ProUni em favor dessa alternativa.

Para calcular o VPL^{social} em si, e um payback, é necessário estipular uma taxa de referência τ . 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 $VPL^{social} = n \times R\$ 76$ mil e um payback de 10 anos; e, como se estima que o programa já gerou $n > 500$ 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 razoavelmente aderentes na realidade.

Também já evidenciou que o ATT pode mudar consideravelmente entre as diferentes formações. Nesse sentido, a Tabela 8 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.

Tabela 8: Resultados do contraste do gasto tributário com o impacto na massa salarial dos bolsistas egressos por curso selecionado

Curso	$\hat{\gamma}$	$r/(1+r)$	$(\sum_i w_{ii})/n$	B/n	TIR	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

A coluna seguinte aponta o salário médio observado entre os formados pelo respectivo curso, $\sum_i w_{ii}/n$, 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. Nota-se então que 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.

As estimativas expressadas na Tabela 8 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 um impacto significativo 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 no mercado de trabalho. Por outro lado, contrastando os resultados expressos nessa seção com os números anteriormente expostos na Tabela 3, nota-se o programa concentrou a maioria das bolsas em cursos listados na Tabela 8, indicando que, em linhas gerais, o ProUni contemplou acesso e democratização do ensino com retorno econômico positivo.

Cabe destacar que há uma considerável heterogeneidade na empregabilidade, remuneração e permanência no mercado formal de trabalho, a depender da graduação escolhida pelos estudantes. Portanto, a partir das informações da Tabela 8, é possível visualizar a existência de cursos que apresentam elevado impacto individual com baixo retorno social (notadamente Medicina) – uma vez que contemplaram uma quantidade limitada de estudantes – e também a existência de cursos com menor impacto individual (como Administração) que, por englobarem um grande contingente de beneficiários, geram um impacto social agregado bastante elevado.

Com a eventual possibilidade de se redefinir as diretrizes do ProUni, e alocar as bolsas preferencialmente entre os cursos com maior retorno econômico e social, seria de grande importância acompanhar regularmente a distribuição desses benefícios, tendo em vista do dilema entre uma alocação que permite maiores retornos individuais e menores impactos sociais, vis-à-vis os cursos que geram menor impacto individual e, devido ao grande contingente, um grande impacto social.

7. Conclusão

Desde 2005, quando foi institucionalizado, o ProUni já custou mais de uma dúzia de bilhões de reais em gastos tributários, e permitiu a graduação de perto de meio milhão de alunos. Nessa monografia se promoveu uma avaliação econômica inédita do programa, considerando seu impacto no salário dos egressos, através do cruzamento de três bases de microdados (SISPROUNI, CES e RAIS identificada).

Ao definir os diplomados e os evadidos como grupos de tratamento e de controle, respectivamente, estimou-se que, nos primeiros anos depois de formado, na média, um egresso ganha 30% a mais do que ganharia o seu contrafactual sem formação. Em um exercício de matemática financeira com hipóteses conservadoras, estimou-se que: a TIR do ProUni seria de 16% ao ano; o VPL social já seria maior que R\$ 38 bilhões; e, cada geração de diplomados equacionaria o gasto tributário com sua coorte em cerca de 10 anos. Portanto, há evidência de que o programa gera um benefício coletivo que justificaria sua existência.

Por outro lado, nas análises desagregadas por cursos, notou-se que existem formações que não geram impacto na empregabilidade dos beneficiários. Além disso, existem graduações como Administração e Ciências Contábeis que têm um VPL individual menor que Farmácia e Medicina, por exemplo, mas o VPL social dos primeiros é substancialmente maior porque o número de beneficiários

é muito grande nesses cursos. Disso, percebeu-se que é possível melhorar a focalização das bolsas, a fim de potencializar a aplicação do gasto tributário.

Em específico, sob a hipótese de que os devidos dispositivos legais sejam apropriadamente analisados e apreciados, propõe-se uma futura investigação de um ajuste na redação da Lei N.º 11.096 (que regulamenta o programa). Qual seja: criar algum mecanismo que bloqueie parcialmente as isenções fiscais da IES que ofertam bolsas em cursos onde não há evidência de aumento de empregabilidade dos egressos; e, por meio cruzado desse bloqueio, em um jogo de soma zero, financiar o aumento dos benefícios das IES que ofertam bolsas em cursos onde há forte evidência de aumento de empregabilidade dos egressos – sem geração de custos extras para o erário. Nesse contexto, o modelo de pesquisa aqui apresentado poderia ser aprimorado e regularmente replicado por alguma agência de governo (e.g., INEP ou IPEA), a fim de fomentar a criação e a manutenção desse mecanismo.

Entende-se também que essa regra poderia ser modificada com alguma frequência, talvez anualmente, com base no que for mais recentemente observado na RAIS. Assim, acredita-se que isso representaria uma espécie de gestão por resultados, que potencializaria o benefício coletivo do ProUni.

8. 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

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

Borgen, N. T. “College quality and hourly wages: Evidence from the self-revelation model, sibling models and instrumental variables”. *Social Science Research* 48: 121-134, (2014). <http://dx.doi.org/10.1016/j.ssresearch.2014.05.010>

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

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

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

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

Dale, S. B.; Krueger, A. B. “Estimating the payoff to attending a more selective college: an application of selection on observables and unobservables”. *The Quarterly Journal Of Economics* 117 (4): 1491-1527 (2002) <https://www.jstor.org/stable/4132484>

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

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 INS-
PER 4: 1-53, (2015). [https://www.insper.edu.br/wp-content/uploads/2018/09/Ensino-superior-no-
-Brasil.pdf](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](https://mitpress.mit.edu/books/econometric-analysis-cross-section-and-panel-data-second-edition)