

textos para discussão

128 | Outubro de 2018

Uma solução automatizada para avaliações quantitativas de impacto: primeiros resultados do MARVIm

Daniel Grimaldi
Arthur Pinto
Breno Albuquerque
Felipe Buchbinder
João Paulo Pereira
Leandro Ortiz
Marcus Tortorelli
Ricardo Martini

Presidente do BNDES

Dyogo Henrique de Oliveira

Diretoria de Estratégia e Transformação Digital

Ricardo Ramos

Área de Planejamento Estratégico

Mauricio dos Santos Neves

textos para discussão

128 | Outubro de 2018

**Uma solução automatizada
para avaliações quantitativas
de impacto: primeiros
resultados do MARVIm**

**Daniel Grimaldi
Arthur Pinto
Breno Albuquerque
Felipe Buchbinder
João Paulo Pereira
Leandro Ortiz
Marcus Tortorelli
Ricardo Martini**

Resumo

O MARVIm (Modelo Automatizado em R para Verificação de Impacto) foi desenvolvido como uma ferramenta de primeira resposta, ágil, a eventuais demandas de avaliações de impacto de interesse do BNDES. Como uma forma de apresentar as potencialidades do MARVIm, implementou-se um conjunto de avaliações com o objetivo de investigar uma mesma questão: como as diferentes intervenções do BNDES impactam no curto prazo a *performance* das firmas beneficiadas? Dez diferentes tipos de intervenção foram avaliados com o uso de microdados da Relação Anual de Informações Sociais (Rais) e da Centralização de Serviços Bancários S.A. (Serasa) durante os anos de 2008-2011. As variáveis de interesse analisadas foram sempre faturamento bruto, número médio de empregados, produtividade do trabalho, resultado líquido e investimento. De maneira geral, houve evidência positiva de impacto do BNDES sobre o faturamento, o emprego e o investimento das MPMEs apoiadas com linhas para aquisição de BK (fundamentalmente BNDES Finame) e Cartão BNDES. Não foram identificados impactos relevantes sobre o resultado líquido ou a produtividade do trabalho dos beneficiados. Contudo, os resultados sobre a produtividade provavelmente demoram mais do que dois anos para serem sentidos, requerendo um intervalo de análise maior do que o que foi aqui implementado.

Abstract

MARVIm (Automated Model in R for Verification of Impact) was developed as an agile first response tool to eventual demands of impact assessments of interest to BNDES. As a way to showcase the potential of MARVIm, we have implemented a series of assessments to investigate the same issue: how do the different interventions from BNDES affect the short-term performance of clients? Ten different types of intervention were assessed with the microdata from the Annual Social Information Report (Rais) and the Centralized Banking Services S.A. (Serasa) during the years 2008-2011. The variables of interest analyzed were always gross revenues, average number of employees, labor productivity, net profit and investment. In general, there was positive evidence of impact from BNDES on revenue, employment and investment of micro, small and medium-sized companies (MSME) supported with lines for BK acquisition (mainly BNDES Finame) and BNDES Card. Relevant impacts on the net income or labor productivity of clients were not identified. However, the results on the productivity will probably take more than two years to be observed, requiring an analysis interval longer than that implemented herein.

Sumário

1. Introdução	9
2. A base de dados	10
3. O Modelo Automatizado em R para Verificação de Impacto (MARVIm)	12
4. Impactos de curto prazo sobre as firmas beneficiadas	32
5. Considerações finais e os próximos avanços para o MARVIm	41
Referências	42
Apêndice 1: Detalhamento da base de dados utilizada no MARVIm	45
Apêndice 2: Identidades contábeis na base da Serasa	53

Daniel Grimaldi é mestre em Economia pela FEA-USP e economista do BNDES; Arthur Pinto é mestre em Economia pelo Cedeplar-UFMG e economista do BNDES; Breno Albuquerque é doutorando em Economia pela EPGE-FGV e economista do BNDES; Felipe Buchbinder é doutor em administração pela FGV e engenheiro do BNDES; João Paulo Pereira é graduado pela PUC-RJ e economista do BNDES; Leandro Ortiz é graduado pela Unimonte e administrador do BNDES; Marcus Tortorelli é mestre em Engenharia Elétrica pela PUC-RJ e analista de sistemas do BNDES; e Ricardo Martini é mestre em Economia pelo Cedeplar-UFMG e economista do BNDES. Os autores agradecem a Vinícius Carrasco, por ter motivado o desenvolvimento do MARVIm, e a Fabio Giambiagi, Paulo Faveret e Victor Pina pela revisão cuidadosa feita em versões anteriores do trabalho. Este trabalho é de exclusiva responsabilidade dos autores, não refletindo, necessariamente, a opinião do BNDES.

1. Introdução

Nos últimos anos, avaliações de impacto de políticas públicas têm recebido crescente atenção da sociedade – como parte de um movimento mais amplo de cobranças por maior eficiência do gasto público. Os recursos que financiam uma dada política são escassos e sempre podem financiar políticas alternativas. Como as necessidades da sociedade são, por sua vez, virtualmente ilimitadas, alocar recursos para o seu melhor uso é de suma importância.

Por essa razão, toda política deve ser avaliada, com cômputo de seus custos (inclusive os associados à melhor alternativa de aplicação – o custo de oportunidade) e benefícios. A avaliação também cumpre o fundamental papel de gerar informação para que políticas sejam aperfeiçoadas. Como importante instrumento de política pública, impõe-se ao BNDES o desafio de dar escala a suas atividades de monitoramento e avaliação, compreendendo o senso de urgência e o rigor metodológico que o tema exige.

Seguindo a definição presente em Lazzarini e outros (2015b), há a seguinte escala crescente de confiabilidade para as técnicas de estimação de impacto: (i) avaliações de nível 1 são aquelas em que é feita uma mensuração sem aleatorização, apenas comparando os tratados com dados agregados locais e/ou setoriais; (ii) avaliações de nível 2 são aquelas em que um grupo de controle envolvendo indivíduos com características semelhantes aos tratados é estabelecido por técnicas que emulam um experimento; e (iii) avaliações de nível 3 são aquelas em que ocorre definição aleatória de grupos de tratamento e controle – uma aleatorização.

Foi com base nisso que a unidade de monitoramento e avaliação do BNDES desenvolveu um modelo de avaliação baseado em *data science*, denominado Modelo Automatizado em R para Verificação de Impacto (MARVIm). Como ficará claro ao longo do texto, essa ferramenta consiste fundamentalmente em:

- um conjunto de *scripts* (programados no *software* R) que implementam automaticamente técnicas para avaliações de impacto de nível 2, baseadas em diferenças-em-diferenças com pareamento por escore de propensão (*p-score*);¹ e
- um extenso conjunto de dados que, previamente tabulado, permite rapidamente consolidar informações que atendam a uma ampla gama de possíveis avaliações de impacto.

Espera-se que o MARVIm seja uma ferramenta de primeira resposta a eventuais demandas de avaliações de interesse do BNDES, uma vez que o tempo gasto para integrar um conjunto básico de dados e técnicas será bastante reduzido, conferindo ao Banco agilidade para construir uma análise preliminar de impacto.

¹ Em breve o código do MARVIm estará disponível em plataformas adequadas de compartilhamento. Por ora, se houver interesse, entrar em contato com daniel.grimaldi@bndes.gov.br.

Essa agilidade na análise preliminar é requisito fundamental para que o BNDES atenda à demanda crescente por comprovações da efetividade de seus financiamentos, tendo em vista a multiplicidade de linhas de crédito utilizadas pela instituição. Adicionalmente, ao se estabelecer um conjunto consistente e padronizado de rotinas para a elaboração de avaliações de impacto, os resultados poderão ser comparados de maneira sistemática, permitindo uma visão panorâmica da atuação do Banco.

Como uma forma de apresentar as potencialidades do MARVIm, implementou-se um conjunto de avaliações com o objetivo de investigar uma mesma questão: como as diferentes intervenções do BNDES impactam a *performance* das firmas beneficiadas no curto prazo?

Para tanto, as empresas apoiadas pelo BNDES foram agrupadas segundo a lógica de atuação dos financiamentos. Cada grupo foi avaliado separadamente e, em todos os casos, buscou-se a identificação de impactos sobre nível de emprego, faturamento bruto, resultado líquido e produtividade do trabalho. Outras dimensões de desempenho foram também investigadas em casos específicos – exportações e investimento. Além disso, via de regra, a janela de investigação de impacto foi de até um ano após o financiamento.

Os principais resultados obtidos foram a identificação de impactos positivos e robustos sobre o faturamento das firmas apoiadas por meio do Cartão BNDES e do BNDES Finame. O emprego também foi positivamente afetado pelo Cartão BNDES, e pelo BNDES Finame, no caso das micro, pequenas e médias empresas (MPME). Por fim, resultados positivos sobre o investimento foram encontrados no BNDES Automático – Investimento.

Além de resultados, essa primeira aplicação do modelo levantou diversas questões e explicitou algumas limitações da metodologia – que precisarão ser enfrentadas futuramente. Inferir causalidade é algo bastante complexo e não pode ficar a cargo de rotinas automatizadas. Dessa forma, os resultados do MARVIm devem sempre ser interpretados em conjunto com outras evidências.

Para apresentar detalhadamente a experiência de implantação do MARVIm e os resultados da primeira rodada de utilização da ferramenta, este texto se divide em cinco seções, além desta introdução. A seção seguinte apresenta a base de dados (detalhada no Apêndice 1), enquanto a terceira seção detalha o MARVIm. A quarta seção expõe os resultados e a última seção discute possíveis avanços dessa agenda de trabalho.

2. A base de dados

Diante do desafio de entregar um extenso conjunto de avaliações de impacto, o passo inicial foi construir uma base de dados que cobrisse amplo espectro de empresas. Isso era condição necessária, ainda que não suficiente, para analisar a

efetividade de intervenções tão distintas quanto o Cartão BNDES e o BNDES Finem, por exemplo.

Trabalhos recentes de avaliação de impactos do BNDES fizeram uso de bases de dados que concentram informações sobre empresas de grande porte, frequentemente de capital aberto, como a Economática e o Valor 5000 – ver, por exemplo, Lazzarini e outros (2015b) e Bonomo, Brito e Martins (2014). A vantagem de utilizar essas bases é que as informações tendem a ser bastante confiáveis. As desvantagens, contudo, são muitas.

Em primeiro lugar, um recorte limitado a empresas de capital aberto traz um claro viés para a análise. Isso porque a avaliação se concentra sobre firmas de maior porte e que, tipicamente, sofrem menos com problemas de restrição de crédito. É razoável esperar que o efeito do financiamento do BNDES sobre tais empresas seja menor, já que estas têm, de fato, acesso a outras fontes de recursos. Assim, ao concentrar a avaliação sobre firmas de maior porte, corre-se o risco de avaliar apenas um conjunto particular que tende a subestimar os efeitos das intervenções do BNDES.

Em segundo lugar, tais bases de dados limitam também o número de intervenções do BNDES que podem ser avaliadas, pois diversos produtos têm como foco firmas de menor porte – Cartão BNDES, por exemplo. Portanto, o uso de informações restritas a empresas de capital aberto também prejudica o esforço de construir uma visão panorâmica do impacto do BNDES, na medida em que concentra a análise em firmas com características particulares e em um conjunto limitado de produtos ou programas.

Trabalhos recentes superaram essa limitação fazendo uso dos dados identificados nas pesquisas anuais do Instituto Brasileiro de Geografia e Estatística (IBGE) – ver Cavalcanti e Vaz (2017) e Machado e outros (2016), por exemplo. Esse é certamente o caminho mais adequado a ser explorado no futuro. O uso de tais informações, entretanto, exige a submissão de projetos, e a execução das estimativas precisa ser feita na Sala de Acesso Restrito (SAR) do IBGE. Em função de todos os procedimentos envolvidos no uso da SAR, o tempo entre o desenho da avaliação e a extração de resultados pode facilmente atingir a escala de meses. Não serve, portanto, ao objetivo de construir com agilidade uma primeira resposta.

O presente trabalho apoiou-se em três bases de dados: (i) a primeira, denominada Relação Anual de Informações Sociais (Rais), traz dados sobre emprego e setor de atividade das empresas brasileiras; (ii) a segunda, da Secretaria de Comércio Exterior (Secex), aporta informações a respeito de exportações das empresas brasileiras; (iii) por fim, a base da Centralização de Serviços Bancários S.A. (Serasa) consolida dados derivados de demonstrativos contábeis. A Tabela 1 resume as informações contidas em cada uma dessas fontes. Para mais informações a respeito dessas fontes primárias e de seu processo de consolidação, ver o Apêndice 1.

Tabela 1. Resumo das fontes utilizadas

Base de dados	Período de cobertura	Tipo de informação
Rais Empresa	1999-2015	Dados sobre emprego
Secex	1997-2012	Dados sobre exportação
Serasa	2005-2012	Dados contábeis

Fonte: Elaboração própria.

3. O Modelo Automatizado em R para Verificação de Impacto (MARVIm)

Ao analisar o conjunto de trabalhos produzidos recentemente na literatura, é possível observar duas categorias de análises. A primeira, que será denominada de “macroavaliações”, reúne trabalhos que adotaram uma análise mais agregada a respeito dos efeitos do Banco.² Tais iniciativas têm como mérito a vontade de construir uma visão panorâmica acerca da ação do BNDES. Contudo, essa abordagem agrega instrumentos de apoio muito distintos – a lógica de apoio de um BNDES Finem é bastante distinta da lógica de um BNDES Progeren, por exemplo. Um risco consequente disso é a perda de poder de identificação das estratégias adotadas.

Na segunda categoria, que será denominada de “microavaliações”, estão aqueles estudos que se dedicaram a mensurar o impacto de um programa ou produto específico do BNDES. Nesse tipo de trabalho, tem-se a situação oposta à do caso anterior. Por um lado, ganha-se na estratégia de identificação de impacto ao considerar as especificidades envolvidas em cada intervenção; por outro, perde-se a visão panorâmica e a capacidade de generalizar os resultados obtidos para uma análise mais ampla do Banco.³

Normalmente, as microavaliações – ou avaliações de impacto – seguem um método artesanal. Cada avaliação desejada é trabalhada de forma única, sendo a base de dados e os métodos de estimação elaborados para atender isoladamente o caso investigado. A vantagem dessa abordagem é que em todo o processo busca-se entender e tratar as especificidades de cada instrumento de apoio analisado.⁴

Contudo, existem também desvantagens, e a principal delas é o tempo de resposta. Se tudo é construído com base na lógica da intervenção que se pretende

² Exemplos desse tipo de avaliação podem ser encontrados em Lazzarini e outros (2015) e Bonomo, Brito e Martins (2014).

³ Exemplos dessas avaliações podem ser encontrados em Machado e outros (2016), Cavalcanti e Vaz (2017) e Gadenne (2017).

⁴ Alguns exemplos de avaliações de impacto já implementadas sob essa lógica no BNDES podem ser localizados em Machado, Parreiras e Peçanha (2011), Machado e Roitman (2015) e Machado e outros (2016).

analisar, o processo é muito mais demorado. Leva-se bastante tempo até que alguma resposta empírica possa ser produzida a respeito da efetividade de qualquer medida. Outra importante desvantagem é a quantidade de retrabalho. As bases são tabuladas apenas para uma avaliação, e os métodos são programados para implementação apenas em um contexto específico. Com isso, novas demandas requerem novo processo de tabulação, novos *scripts* de programação etc. A ausência de rotinas consistentes, por sua vez, também dificulta a comparação sistemática de resultados.

Foi pensando nessas desvantagens que a equipe do Departamento de Avaliação e Promoção de Efetividade (DEAPE) do BNDES iniciou, no segundo semestre de 2016, o desenvolvimento do MARVIm, que, conforme disposto na introdução, consiste fundamentalmente em:

- um conjunto de *scripts* (programados no *software* R) que implementem automaticamente técnicas para avaliações de impacto de nível 2 (aquelas que buscam a construção de contrafactuais) baseadas em diferenças-em-diferenças com escore de propensão (*p-score*); e
- um extenso conjunto de dados que, previamente tabulado, permita rapidamente consolidar informações que atendam a uma ampla gama de possíveis avaliações de impacto.

Dessa forma, será possível operar com razoável agilidade um conjunto básico de dados e técnicas para construir uma primeira resposta para questões de avaliação de impacto, o que permite a elaboração de uma visão panorâmica da atuação do BNDES. Não há pretensão, contudo, de construir a visão final sobre cada tema, dado que inferir causalidade é algo bastante complexo e precisa ser encarado como tal.⁵ Em particular, inferir causalidade com base em resultados estatísticos – dirigidos fundamentalmente por correlação – é algo que deve ser feito com muita cautela⁶ e não pode ficar simplesmente a cargo de rotinas automatizadas.

3.1 O problema fundamental da avaliação de impacto

O problema fundamental da inferência causal é apresentado por Angrist e Pischke (2009) como um problema de *missing data*. Para deixar claro esse ponto, os autores fazem uso do que ficou conhecido na literatura como modelo de resultados potenciais (ou Rubin Causal Model):

$$\text{Resultado potencial} = \begin{cases} Y_{1i} & \text{se } D_i = 1 \\ Y_{0i} & \text{se } D_i = 0, \end{cases}$$

⁵ Para uma discussão mais fundamental a respeito de modelos e inferência causal, ver Pearl (2009).

⁶ Inferências estatísticas são fartas em exemplos que ficaram conhecidos na literatura como Paradoxo de Simpson – casos em que uma aparente correlação surge para uma população, mas o inverso vale em subpopulações. Tais exemplos mostram como os resultados devem ser cuidadosamente interpretados. Para uma discussão mais ampla a esse respeito, ver, por exemplo, Pearl, Glymour e Jewell (2016).

em que D_i assume valor 1, se a unidade i foi exposta a determinado tratamento (uma linha de crédito do BNDES, por exemplo), e 0, caso contrário.⁷ Y_{1i} , por sua vez, representa o resultado que uma determinada variável da unidade i assumiria caso fosse submetida ao tratamento, e Y_{0i} constitui o resultado da mesma variável, para a mesma unidade, na ausência do tratamento. Nesse caso, o impacto do tratamento pode ser representado simplesmente por $Y_{1i} - Y_{0i}$. Contudo, o problema fundamental é que, em um dado momento do tempo, a unidade i só pode assumir *status* de tratada ou não tratada, não sendo possível, portanto, observar Y_{1i} e Y_{0i} simultaneamente.

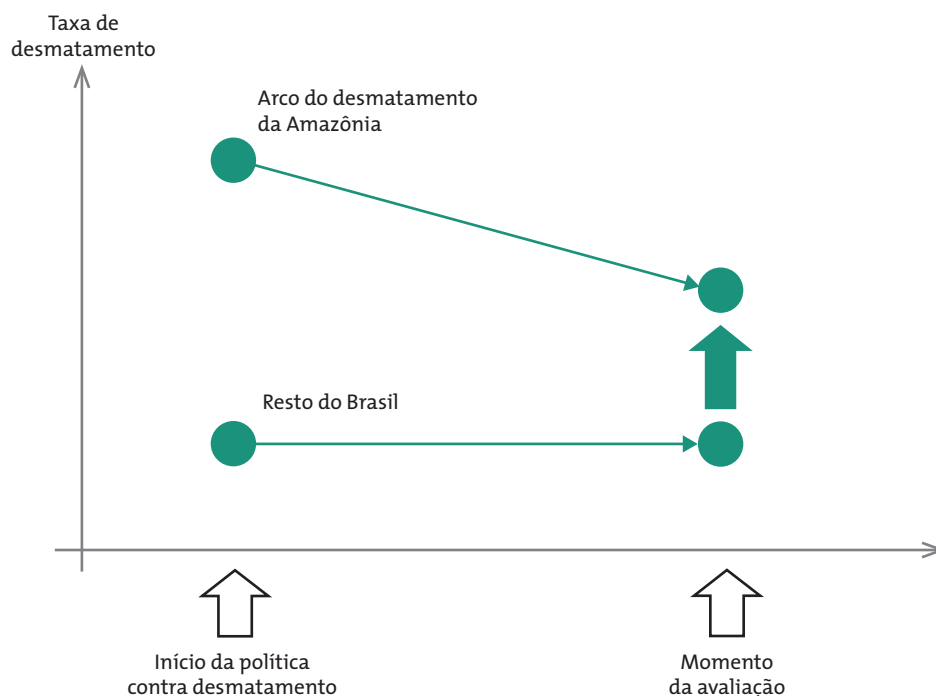
Quando o acesso ao tratamento é definido de maneira aleatória em uma população, o problema de *missing data* é em parte superado, abandonando-se o efeito do tratamento sobre determinada unidade e investigando-se o efeito médio do tratamento sobre um grupo de unidades. Nesse caso, pode-se substituir $Y_{1i} - Y_{0i}$ pela diferença de média entre unidades tratadas e não tratadas. Isso porque, nesse caso, $\sum Y_i^{D_i=1} / N^{D_i=1} - \sum Y_i^{D_i=0} / N^{D_i=0}$ é um estimador não viesado para $\Sigma[Y_i | D_i = 1] - E[Y | D_i = 1]$ – o efeito médio do tratamento sobre os tratados (*average treatment effect on the treated*).

Intuitivamente, se o tratamento é alocado de modo aleatório, os dois grupos formados são estatisticamente iguais em todas as características, exceto o *status* de tratamento. Pode-se inferir que qualquer diferença de desempenho entre os grupos é causada pelo tratamento. Contudo, comumente em avaliações de impacto de políticas públicas, o tratamento não é concedido de maneira aleatória. No caso de uma política voltada para a redução do desmatamento clandestino, por exemplo, é natural que as áreas prioritariamente atendidas tenham taxas de desflorestamento bem superiores às demais. Nesse cenário, que pode ser ilustrado pela Figura 1, uma diferença de média entre regiões atendidas pela política – arco do desmatamento da Amazônia, por exemplo – e regiões não atendidas – restante do Brasil – leva à falsa conclusão de que o tratamento tem impacto positivo sobre o desmatamento. Esse é o problema que ficou conhecido na literatura como viés de seleção da política pública.

Outra análise ingênua seria acompanhar a variação no desmatamento da região apoiada durante o período considerado. Essa abordagem não resolve o viés de seleção, conforme ilustrado na Figura 2. Nesse caso, a política é efetiva porque o desmatamento nas regiões atendidas é menor do que seria na ausência da política. O problema é que aquilo que teria ocorrido na ausência da política é algo que não pode ser observado, uma vez que não aconteceu. A mera comparação da variação no desmatamento no arco do desmatamento da Amazônia mostra um aumento do desmatamento, sugerindo o fracasso de uma política que, de acordo com a Figura 2, foi efetiva. Nessa análise, novamente, observam-se os efeitos do viés de seleção da política pública.

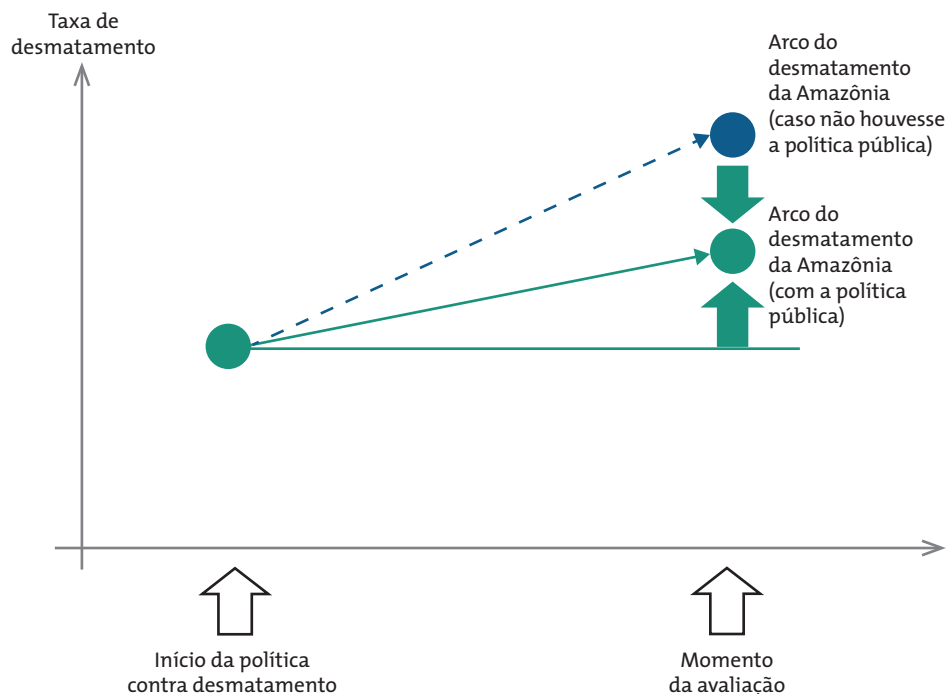
⁷ Na literatura de avaliação de impacto, define-se como unidades “tratadas” aquelas que foram beneficiárias de determinada intervenção, que pode ser uma política pública, um mecanismo de incentivo, um tratamento médico etc.

Figura 1. Comparação ingênua entre tratados e não tratados



Fonte: Elaboração própria.

Figura 2. Comparação ingênua entre antes e depois do tratamento



Fonte: Elaboração própria.

Por sua capacidade de lidar com o viés de seleção resultante de diferenças em características observáveis, as técnicas baseadas em pareamento (*matching*) tornaram-se bastante populares, sendo largamente aplicadas para avaliações de

impacto (CALIENDO; KOPEINIG, 2008). Intuitivamente, o que essas técnicas fazem é selecionar, entre as unidades não tratadas, as mais parecidas com aquelas que receberam tratamento – e, portanto, mais adequadas como unidades de controle ou contrafactual. Considerando o exemplo apresentado, seria o equivalente a identificar municípios com características muito parecidas – vizinhos dentro do arco do desmatamento, por exemplo –, em que apenas um deles foi beneficiado pela política de desmatamento.

3.2 Os estimadores do MARVIm

O problema fundamental da inferência causal ocorre quando se pretende medir o impacto de intervenções do BNDES. Em todas as operações realizadas pelo Banco, dois importantes processos de seleção acontecem: as próprias empresas, com base em suas características intrínsecas, optam por procurar ou não o apoio do Banco para implementar seus planos; e o BNDES – ou um repassador por ele autorizado, no caso das operações indiretas – decide, com base em características observáveis das firmas solicitantes, se realiza ou não a operação.⁸ Nesses processos, é razoável supor que uma parcela importante do viés de seleção pode ser controlada por variáveis observáveis.⁹ No entanto, existem também características que condicionam a predisposição das firmas em buscar melhores formas de financiamento e que não podem ser diretamente observadas – empreendedorismo do dono, por exemplo.

Para lidar com o problema de viés de seleção descrito, o MARVIm utiliza estimadores que combinam ajuste de regressão com pareamento e diferenças-em-diferenças para tentar inferir resultados robustos para os impactos das intervenções analisadas. Para compreender melhor como isso ocorre, considere-se o seguinte contexto:

- um *policymaker* deseja avaliar o impacto de determinada política pública sobre uma dimensão Y das empresas apoiadas;
- a *performance* em Y depende de um conjunto X de características observáveis das empresas (que pode ou não afetar também a probabilidade de acessar o tratamento);
- um conjunto V (que pode ter intersecção com X) afeta também a probabilidade de uma firma ser beneficiária dessa política pública, gerando, portanto, um viés de seleção;
- a variável D_i define a condição de uma firma i com respeito à política pública, conforme o modelo de resultados potenciais definido anteriormente;

⁸ Vale notar que, nesse segundo processo de seleção, a capacidade de pagamento das firmas apoiadas é um elemento crucial, haja vista o peso da análise de risco de crédito.

⁹ O faturamento bruto das empresas, por exemplo, pode ser observado na base de dados e condiciona o acesso a determinadas operações com o BNDES.

- existe um conjunto N_t de firmas beneficiárias e um conjunto N_c de firmas que não tiveram acesso à política.

Nesse cenário, a dificuldade do avaliador pode ser descrita pelo problema de estimar corretamente o parâmetro τ da equação 1.¹⁰

$$y_i = \alpha + \tau D_i + \beta X_i + \epsilon \quad \forall \text{ firma}_i \quad (1)$$

Em virtude das características (observáveis ou não) da empresa, o candidato natural a estimador de impacto – a diferença simples de desempenho entre tomadores e não tomadores de recursos – retorna o efeito verdadeiro do tratamento adicionado de um viés.¹¹ É, portanto, um estimador ingênuo para a equação 1.

Grosso modo, a estimação pode ser feita por Mínimos Quadrados Ordinários (OLS, do inglês Ordinary Least Squares). Contudo, esse método não é adequado para inferências quando as distribuições das variáveis do conjunto X são muito desbalanceadas ou assimétricas, pois os resultados tornam-se muito sensíveis à especificação da forma funcional. Adicionalmente, o efeito médio do tratamento é pouco preciso para espaços em X em que há pouca representatividade de tratados ou controles, e isso contamina a estimativa média de impacto obtida por $\hat{\tau}_{obs}$.¹²

O MARVIm contorna esse problema por meio de técnicas baseadas em escores de propensão (*p-scores*). O primeiro estágio de estimadores desse tipo equivale sempre a implementar a regressão que busca explicar a probabilidade de uma determinada firma ser apoiada pela política pública que se deseja avaliar. Isso é feito por meio de um modelo de regressão logística definido, conforme a equação 2.

$$e(v) \equiv Pr(D_i = 1 | V_i = v) = \frac{\exp(v'\gamma)}{1 + \exp(v'\gamma)} + \epsilon \quad (2)$$

Uma das formas mais tradicionais de aumentar a comparabilidade entre tratados e controles dentro da amostra é parear unidades por meio do algoritmo de

¹⁰ Os estimadores aplicados supõem, implicitamente, que, para grupos suficientemente homogêneos, a resposta ao tratamento de tratados e controles é idêntica – ou seja, supõe-se que $\beta_t = \beta_c = \beta$. Existem, contudo, formas alternativas de estimação que flexibilizam também essa hipótese – ver, por exemplo, Hirano e Imbens (2001).

¹¹ São exemplos de características observáveis: porte, participação no mercado externo e custo da mão de obra. São não observáveis todas aquelas sobre as quais o modelo não tem informação – por exemplo, motivação dos administradores.

¹² “(...) *In observational studies we often find substantial differences between covariates distribution in the two treatment arms. Such lack of covariate balance creates two problems. First, it can make subsequent inferences sensitive to ostensibly minor changes in the methods and specifications used. For example, adding an interaction or quadratic term to a linear regression specification can change the estimated average treatment effect substantially when covariate distributions are far apart. Second, lack of balance can make the inferences imprecise. For covariate values with either few treated or few controls, it may be difficult to obtain precise estimates for treatment effects, and this, in turn, may make the estimates of overall treatment effects imprecise*” (IMBENS; RUBIN, 2015, p. 337).

vizinho mais próximo. Nessa abordagem, escolhe-se para cada firma do conjunto N_t uma outra firma que está em N_c cujo p -score é semelhante. Ou seja, para cada firma apoiada será encontrada uma outra que não foi beneficiada pela política, mas que apresentou probabilidade de tratamento muito parecida com a da primeira. A equação 1 pode ser estimada por OLS dentro dessa subamostra, obtendo-se um estimador $\hat{\tau}_{psm}$.¹³

Uma vantagem dessa técnica é que ela tenderia a construir grupos de controle e tratamento bastante homogêneos com respeito a V , reduzindo o viés de seleção em observáveis. Entretanto, uma crítica comum a esse algoritmo de pareamento é que ele reduz muito a amostra, descartando uma quantidade expressiva de informação. Assim, o ganho em comparabilidade é obtido muitas vezes à custa do aumento na variância do estimador, prejudicando a inferência estatística (CALIENDO; KOPEINIG, 2008).

Uma forma de contornar essa limitação é aplicar estimadores que usam o p -score para atribuir peso às unidades analisadas – *propensity score weighting* (HIRANO; IMBENS, 2001). Nesse caso, não se restringe a análise apenas às firmas pareadas. Nesse método, a equação 1 é estimada por Mínimos Quadrados Ponderados (WLS, do inglês Weighted Least Square), e o peso de cada observação w_i é dado por uma função inversa do p -score – equação 3 –, seguindo o que foi exposto por Imbens e Rubin (2015). A intuição desse segundo estimador – que será definido por τ_{psw} – é que tendem a ter maior peso na estimação as unidades mais “comparáveis”, isto é, aquelas que estão em áreas de maior densidade na distribuição do p -score.

$$w_i = \begin{cases} \frac{\frac{N_t}{e(v_i)}}{\sum_{i \in N_t} \frac{1}{e(v_i)}} & \text{se } D_i = 1 \\ \frac{\frac{N_c}{1 - e(v_i)}}{\sum_{i \in N_c} \frac{1}{1 - e(v_i)}} & \text{se } D_i = 0 \end{cases} \quad (3)$$

Adicionalmente, seguindo Heckman, Ichimura e Todd (1998), aplica-se um método simples para expurgar o viés de seleção por variáveis não observáveis fixas no tempo: calcula-se o impacto considerando como variável de impacto $\Delta_Y \equiv Y_{it} - Y_{it-1}$. A intuição é que, caso uma empresa tenha alguma característica não observável e que condicione sua *performance* em Y_p , ela provavelmente já se fazia notar em $t-1$. Logo, ao avaliar o impacto sobre Δ_Y , exclui-se naturalmente a

¹³ Uma forma alternativa de estimar o impacto do tratamento é implementar um estimador simples de diferenças de médias para Y entre tratados e controles na amostra pareada. Imbens e Rubin (2015), por exemplo, defendem que a combinação de pareamento com o ajuste por regressão em covariadas é uma abordagem mais robusta e que permite controlar por heterogeneidade adicional, podendo existir mesmo dentro da amostra pareada.

influência de quaisquer fatores fixos no tempo que porventura não estejam contemplados em X . Por fim, tanto τ_{psm} quanto τ_{psw} são aplicados sobre uma amostra que respeita as restrições de suporte comum, seguindo Imbens e Rubin (2015). A forma como isso é feito ficará mais clara na próxima subseção.

3.3 A qualidade do balanceamento

Quando as distribuições de covariadas são muito similares entre tratados e controles, como se esperaria em um contexto de experimento aleatório, estimadores simples como $\hat{\tau}_{ols}$ podem ser considerados adequados. Contudo, em situações nas quais há grande heterogeneidade, há também um viés de seleção e, portanto, uma prescrição para o uso de estimadores mais robustos – como $\hat{\tau}_{psm}$ e $\hat{\tau}_{psw}$.

A confiabilidade das técnicas de *matching* em lidar com tal viés dependerá diretamente de sua capacidade de obter uma amostra adequadamente balanceada. É por isso que uma etapa crucial para os estimadores definidos é avaliar quão diferentes são as distribuições de covariadas (V) entre os grupos de tratados e controles, antes e depois do pareamento.¹⁴

Mesmo em situações nas quais a amostra pareada está, na média, suficientemente balanceada entre tratados e controles, é possível que existam regiões do espaço de covariadas em que sejam observadas poucas unidades tratadas. Como consequência, a inferência causal é muito menos confiável em tais regiões e deve ser evitada.¹⁵ Ou seja, o impacto da política só pode ser estimado para faixas do *p-score* em que se observa adequada densidade de empresas beneficiadas e não beneficiadas.¹⁶

É em razão disso que todo *matching* implementado pelo MARVIm é submetido a uma etapa que avalia sua qualidade. Essa etapa consiste em: (i) identificar a região de suporte comum, excluindo unidades que estejam fora dela; e (ii) calcular uma série de estatísticas que permitam avaliar o grau de balanceamento entre tratados e controles antes e depois do pareamento.

A literatura sugere diversas maneiras de avaliar a região de suporte comum. Para alguns autores, essa análise deveria ser feita prioritariamente por meio de uma inspeção visual das distribuições de *p-score* para tratados e controles (LECHNER, 2001). Por outro lado, Caliendo e Kopeinig (2008) defendem que a inspeção visual deve ser complementada por métricas formais para auxiliar o avaliador.

¹⁴ Formalmente, essa necessidade é apresentada como hipótese de “ignorabilidade forte” (*unconfoundedness*).

¹⁵ Formalmente, essa restrição está definida pela hipótese de suporte comum, crucial para a validade da estratégia de identificação (BLUNDELL; DIAS, 2002).

¹⁶ Intuitivamente, imagine-se que, para um dado conjunto de características $X_i = x$, a probabilidade de uma empresa ser beneficiada pela política é igual a 100%. Para esse tipo de empresa, portanto, não há como achar uma outra que seja “comparável” para servir de contrafactual porque, por definição, todas aquelas com características semelhantes também foram beneficiadas pela política.

A mais comum dessas métricas é denominada Comparação de Mínimo e Máximo (CALIENDO; KOPEINIG, 2008). Nessa abordagem, a região de suporte comum é limitada, abaixo, pelo menor *p-score* estimado para unidades tratadas e, acima, pelo maior *p-score* obtido para unidades de controle. Todas as observações fora desse intervalo são, portanto, desconsideradas na análise.

O problema dessa abordagem é que, se as distribuições tiverem caudas muito finas, podem ainda existir regiões dentro dos limites definidos pela comparação de máximo e mínimo com densidade de controles (ou de tratados) muito baixa – são, portanto, regiões onde não há adequada comparabilidade. Para superar esse problema, Imbens e Rubin (2015) sugerem um procedimento de *trimming* com base no *p-score*, que exclui de maneira sequencial valores extremos (muito próximos de 0 ou de 1), encurtando as caudas das distribuições nas regiões de baixa densidade.

Especialmente por se tratar de um método automatizado, o MARVIm não poderia depender exclusivamente de uma inspeção visual. Por isso, ele combina a exclusão por mínimo e máximo com *trimming* para definir a região de suporte comum.¹⁷ Vale destacar que, durante esse processo, realiza-se o descarte de observações para as quais parece não haver adequada comparabilidade. Ao se fazer isso, prejudica-se a validade externa da avaliação. No contexto deste trabalho, isso implica capacidade limitada de extrapolar para outras empresas os impactos obtidos na subamostra especificada. Entretanto, aumenta-se a confiança para inferência causal – ganha-se, portanto, validade interna. Ou seja, no *trade-off* entre capacidade de generalização dos resultados e confiabilidade das estimativas, optou-se por privilegiar o último.

Definida a região de suporte comum, é preciso avaliar também se as amostras estão balanceadas quanto a suas covariadas. Para tanto, o modelo implementa quatro medidas de sobreposição das distribuições que são sugeridas por Imbens e Rubin (2015). A primeira é denominada diferença normalizada (definida pela equação 4) e permite inferir se a diferença entre a localização (média) das distribuições de tratados e controles é grande a ponto de causar vieses nos estimadores mais simples, como $\hat{\tau}_{ols}$. Além de comparar a posição das distribuições, cabe avaliar se as dispersões são semelhantes. Isso é feito pela razão logarítmica dos desvios-padrão, definida na equação 5.¹⁸

$$\hat{\Delta}_{ct} = \frac{\bar{X}_t - \bar{X}_c}{\sqrt{(S_c^2 + S_t^2)/2}} \quad \text{em que, } S_{d=c,t}^2 \equiv \frac{1}{N_d - 1} \sum_{i:D_i=d} (X_i - \bar{X}_d)^2 \quad (4)$$

$$\hat{\Gamma}_{ct} = \ln(S_t) - \ln(S_c) \quad (5)$$

¹⁷ Depois da exclusão por comparação de mínimo e máximo, a amostra restante é submetida a um procedimento de *trimming*, seguindo Imbens e Rubin (2015).

¹⁸ Usa-se a razão logarítmica porque ela é tipicamente mais semelhante a uma distribuição normal.

Outra abordagem possível é avaliar a fração das unidades tratadas (controles) que seriam consideradas *outliers* pela distribuição da amostra de controles (tratados). Para entender a implementação dessa análise, considere-se que $\hat{F}_c(\cdot)$ e $\hat{F}_t(\cdot)$ são as funções das distribuições empíricas de X_i para cada uma das subamostras, enquanto $\hat{F}_c^{-1}(\cdot)$ e $\hat{F}_t^{-1}(\cdot)$ são suas inversas. Para qualquer nível de significância α é possível calcular $\hat{\Pi}_c^\alpha$ e $\hat{\Pi}_t^\alpha$, definidos pelas equações 6 e 7. O MARVIm implementa essas medidas considerando um α de 0.05. A avaliação de $\hat{\Pi}_c^\alpha$ e $\hat{\Pi}_t^\alpha$ permite inferir se as caudas das distribuições têm densidades semelhantes.

$$\hat{\Pi}_c^\alpha = 1 - (\hat{F}_c(\hat{F}_t^{-1}(1 - \alpha/2)) + \hat{F}_c(\hat{F}_t^{-1}(\alpha/2))) \quad (6)$$

$$\hat{\Pi}_t^\alpha = 1 - (\hat{F}_t(\hat{F}_c^{-1}(1 - \alpha/2)) + \hat{F}_t(\hat{F}_c^{-1}(\alpha/2))) \quad (7)$$

Observando a forma como a etapa de pareamento afeta essas quatro métricas, é possível avaliar a sobreposição entre as amostras de tratados e controles e, com isso, julgar a qualidade do balanceamento. No contexto de avaliações de política pública, tradicionalmente as distribuições de unidades apoiadas e não apoiadas serão bastante heterogêneas *a priori*. Espera-se, contudo, que o *matching* seja capaz de torná-las mais homogêneas. Quando isso não ocorre, tem-se um indicativo de que mesmo esses métodos não estão sendo capazes de controlar adequadamente o viés de seleção.

3.4 Superando um *trade-off*: seleção automatizada de variáveis

Um grande desafio para a aplicação dos métodos de estimação de impacto descritos neste estudo é a especificação das covariadas das equações 1 e 2, definidas, respectivamente, pelos vetores X e V . Essa etapa é crucial porque, como já explicitado, uma especificação equivocada poderia afetar decisivamente os estimadores utilizados.

É evidente que o ideal seria fazer a escolha das variáveis caso a caso, entendendo as particularidades de cada política pública a ser avaliada. Agregar políticas muito heterogêneas, como fazem os trabalhos definidos como macroavaliações, tende a enfraquecer a estratégia de identificação associada aos métodos de *matching*. Isso ocorre porque comparar firmas semelhantes em características que não explicam o acesso à política pública será, na melhor das hipóteses, inócuo para controlar o viés de seleção – e, na pior delas, pode introduzir o viés.

Os elementos que condicionam uma empresa a buscar, por exemplo, capital de giro são totalmente distintos daqueles que a levam a buscar crédito para expandir a planta. Misturar os dois casos em uma mesma equação de *p-score* pode produzir coeficientes pouco confiáveis para a equação 2 – concluindo-se, eventualmente, que empresas muito endividadas buscam crédito para expansão. Como consequência, os estimadores de impacto não serão confiáveis.

Fazer artesanalmente a especificação para cada caso, ainda que seja desejável, é inviável quando se deseja avaliar um conjunto amplo de intervenções em curto espaço de tempo – como se propõe a fazer o MARVIm. A saída encontrada para enfrentar esse *trade-off* foi utilizar técnicas de aprendizagem automatizada (*machine learning*) para que a seleção pudesse ser feita pelo próprio modelo em cada intervenção analisada. Desde que cada intervenção seja adequadamente definida como um conjunto razoavelmente uniforme de políticas de financiamento, a especificação pode ser automatizada, sem prejuízos à estratégia de identificação de impacto.

Essa escolha de variáveis foi implementada no MARVIm por meio de um algoritmo de inclusão por ajuste de regressão (*forward selection*). Para compreender como funciona esse procedimento, considere-se que V precisa ser selecionado por meio de um conjunto inicial de dimensão k , que abarca todas as variáveis disponíveis. A especificação da equação 2 começa pela estimação de k regressões, em que cada uma delas tem como explicativa um dos k possíveis termos.

Adiciona-se primeiro à especificação o termo que tiver demonstrado a maior capacidade para explicar a probabilidade de acesso à política – a métrica usada para tal corte é a razão de verossimilhança calculada em cada uma das k regressões. Esse processo é executado iterativamente, com a inclusão de covariadas a V parando apenas quando nenhuma variável disponível for capaz de melhorar o poder explicativo da estimação, ou se todo o conjunto k já tiver sido utilizado.¹⁹ Em seguida, são consideradas todas as interações possíveis entre os k^* termos selecionados. Repete-se então o procedimento anterior escolhendo aquelas variáveis que apresentarem maior poder explicativo e descartando as demais.²⁰

Procedimento semelhante é adotado para a definição de X . Mas, nesse segundo caso, a variável explicada é a dimensão de impacto (Y), e a regressão-base – aquela a partir da qual o modelo irá adicionar covariadas – já conta com uma *dummy* que define o *status* de acesso à política pública (D_i) desde o ponto de partida.

3.5 O passo a passo de uma avaliação no MARVIm

Com o conhecimento dos principais procedimentos executados pelo MARVIm, é possível detalhar o passo a passo de uma avaliação feita com base nesse modelo. Existem alguns passos que antecedem a execução automatizada. Primeiramente, é preciso determinar qual intervenção se pretende avaliar – no caso do BNDES, normalmente se considera uma linha, produto ou programa.

¹⁹ Utiliza-se um ponto de corte para o ganho esperado na razão de verossimilhança, a partir do qual a contribuição da variável adicional passa a ser considerada negligenciável. Para mais detalhes, ver Imbens e Rubin (2015).

²⁰ Serão analisadas, na prática, $k^*(k^* + 1)/2$ possíveis variáveis de segunda ordem, incluindo-se aí os termos quadráticos (interação de uma variável com ela mesma).

O passo seguinte é elaborar uma base que identifique todas as empresas que formam o conjunto de tratadas, bem como o ano e o valor envolvido em cada caso. Outro importante *input* do modelo é a base de análise, que nesse exercício consolida dados da Rais, Secex e Serasa.²¹ Existe também uma série de parâmetros que precisam ser escolhidos pelo avaliador e que são apresentados no Box 1. É importante destacar que cada um desses parâmetros permite customizar algum aspecto da avaliação, sempre com o objetivo de testar diferentes hipóteses a respeito dos impactos esperados. Analisar os resultados obtidos por diferentes configurações é etapa fundamental para a robustez das conclusões.

Definidos esses parâmetros, o MARVIm inicia a etapa automatizada por meio do cruzamento entre as duas fontes de dados fornecidas como *input*. As empresas financiadas pela intervenção e que não forem encontradas na base de análise serão automaticamente excluídas da avaliação. A título de exemplo, as tabelas 2 e 3 trazem estatísticas básicas geradas pelo modelo ao fim desse cruzamento de dados para o caso da avaliação dos instrumentos do BNDES de financiamento à aquisição de bens de capital.²²

Tabela 2. Representatividade da amostra – número de empresas

Ano	Obs. de análise	Obs. financiadas	Obs. encontradas	Sobreposição
2006	29.152	24.014	799	3.33%
2007	51.985	36.634	4.309	11.76%
2008	66.119	47.961	6.003	12.52%
2009	102.477	86.330	6.366	7.37%
2010	151.164	137.838	8.794	6.38%
2011	183.808	173.133	7.568	4.37%
2012	132.646	131.908	246	0.19%

Fonte: Elaboração própria, com base no MARVIm.

Tabela 3. Representatividade da amostra – valor do financiamento*

Ano	Valor financiado	Valor financiado encontrado	Sobreposição
2006	13.058.460.668	1.067.393.639	8.17%
2007	20.879.496.423	9.132.670.050	43.74%
2008	26.086.549.947	14.058.830.423	53.89%
2009	27.622.465.909	12.304.975.762	44.55%
2010	53.315.358.274	22.931.094.603	43.01%
2011	58.172.211.967	22.172.188.878	38.11%
2012	15.211.571.828	1.348.124.136	8.86%

Fonte: Elaboração própria, com base no MARVIm.

* Valores financiados em R\$.

²¹ Para mais detalhes, ver Apêndice 1.

²² São consideradas nesse recorte todas as empresas que receberam crédito no âmbito da linha Aquisição de BK (seja pelo BNDES Finame, seja pelo BNDES Finem) mais todas aquelas que adquiriram máquinas e equipamentos por meio do Cartão BNDES.

Box 1. Parâmetros do MARVIm

(a) Primeiro tratamento:

Define como momento do tratamento de uma firma apenas o primeiro ano em que ela tiver recebido o apoio. Permite medir o impacto apenas nos anos seguintes ao acesso inicial das firmas.

(b) Tratamento único:

Funciona como um filtro adicional que restringe as unidades tratadas apenas àquelas que tiverem sido tratadas uma única vez. Tem como objetivo separar firmas que tiverem sido tratadas diversas vezes (um *status* que se define tecnicamente como sendo de tratamento continuado) daquelas que acessaram a política apenas uma vez.

(c) Período de tratamento:

Serve para restringir a avaliação de impacto. A estimação pode ser feita para cada um dos anos disponíveis ou considerar todo o intervalo. Note que limitar a análise a anos específicos mostra como os resultados se alteram com respeito às condições particulares de cada momento do tempo – em alguns casos, existe fundamentação teórica para supor que o efeito da intervenção pode depender de questões conjunturais.^a

(d) Janela de pareamento:

Define o número de anos antes do tratamento que devem ser considerados para o pareamento das observações. Janelas de pareamento maiores exigem que as firmas pareadas tenham, para as variáveis selecionadas, médias próximas durante um intervalo de tempo maior.

(e) Janela de *performance*:

Define o número de anos depois do tratamento que devem ser considerados para a mensuração do resultado. O aumento dessa janela pode ser usado para investigar se o impacto do tratamento se dissipa ao longo do tempo.

(f) *Lag*:

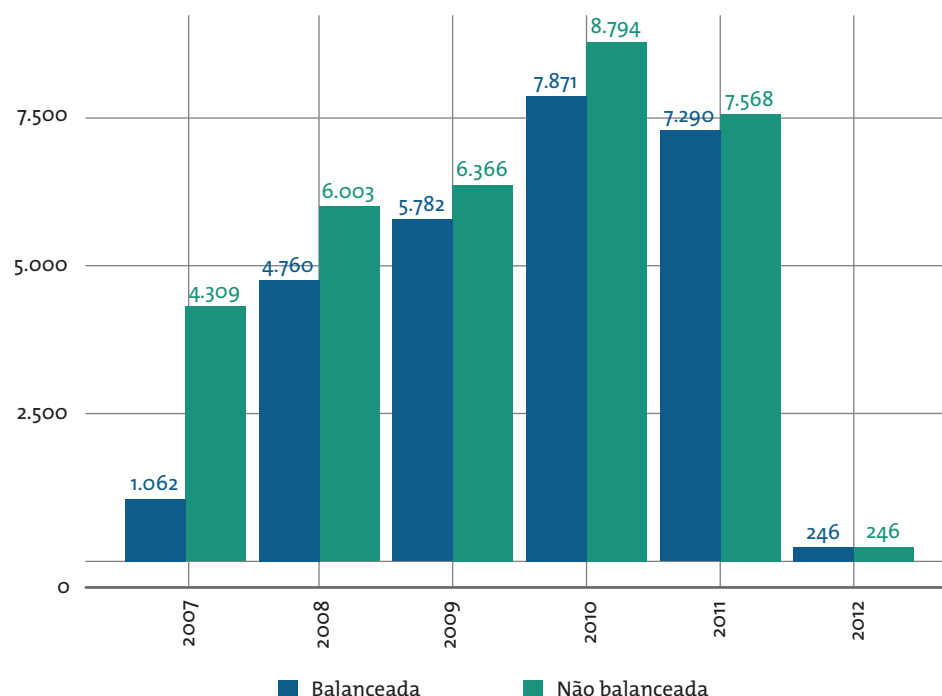
Permite defasar as variáveis X e Y (com relação à data fornecida no arquivo de unidades tratadas), deslocando toda a análise. Esse argumento não altera, contudo, o momento do tempo que define unidades tratadas.

^a Ver Machado e outros (2016), por exemplo, com o caso do BNDES PSI.

Nesse caso em particular, é possível notar que apenas uma pequena parcela das empresas apoiadas pelo BNDES foi encontrada também na base de análise – no melhor ano, a sobreposição foi de 12,52%. Essa informação é importante porque, quanto maior o percentual de firmas apoiadas encontradas, maior tende a ser a capacidade de generalização dos resultados. Adicionalmente, a Tabela 3 mostra que a sobreposição das amostras foi maior na dimensão de valor financiado. Isso indica que, entre as unidades tratadas, foi possível encontrar na base de análise aquelas que tinham proporcionalmente maior fatia dos financiamentos. Há, portanto, uma indicação de viés de amostragem, reforçando a ideia de que o resultado obtido para essas empresas pode não ser generalizável para o conjunto de firmas apoiadas pelo BNDES no âmbito da aquisição de bens de capital.

O passo seguinte do modelo consiste na aplicação de um filtro por setor de atividade, restringindo a análise apenas às divisões da Classificação Nacional de Atividades Econômicas (Cnae) nas quais existe alguma firma financiada.²³ Feito esse recorte, o modelo aponta a quantidade de empresas financiadas em cada momento do tempo, bem como sua representatividade no número total de empresas restantes na base de dados.

Gráfico 1. Quantidade de empresas apoiadas (por período)



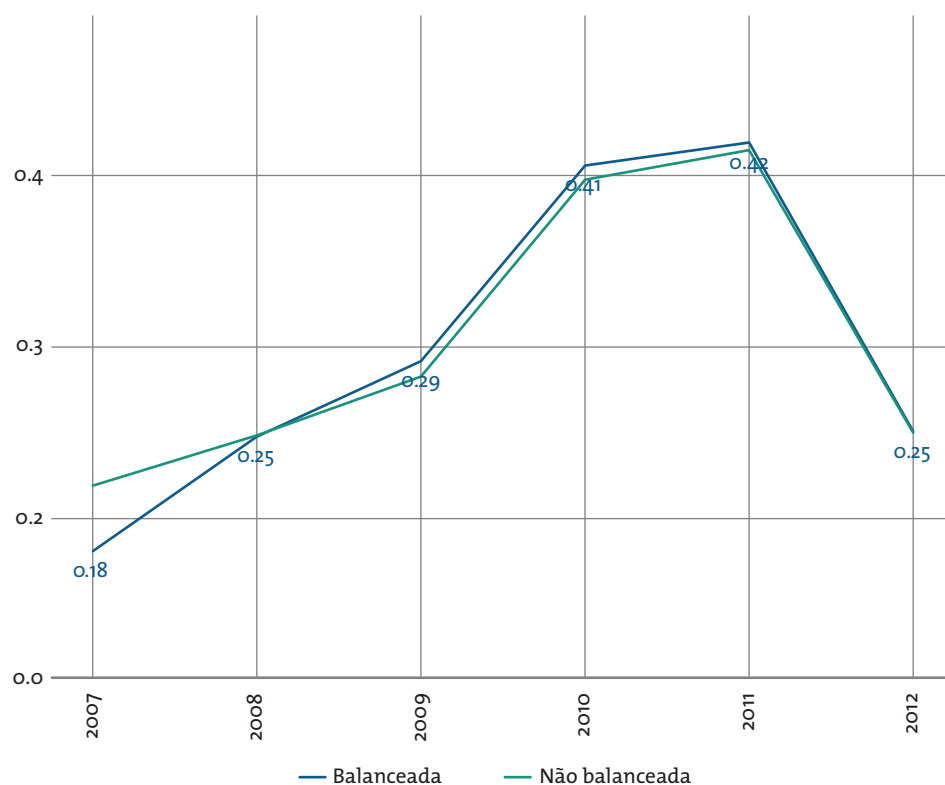
Fonte: Elaboração própria, com base no MARVIm.

²³ Esse filtro pré-pareamento serve para aumentar a comparabilidade das amostras. De certa forma, entende-se implicitamente que firmas cujos setores de atividades não contam com nenhuma empresa beneficiada não devem ser elegíveis para a intervenção que se deseja avaliar.

Novamente considerando o caso escolhido para exemplo, é possível notar no Gráfico 1 que, no ano de 2010, havia 8.794 firmas apoiadas pelo BNDES, sendo que foi possível obter informações sobre 7.871 também em 2009 – período pré-tratamento. Na prática, o impacto só pode ser estimado para essas observações, denominadas balanceadas.

De forma geral, nesse exemplo, há um considerável número de empresas apoiadas, o que afasta a possibilidade de que os estimadores sofram com inconsistências decorrentes de micronumerosidade.²⁴ Já pelo Gráfico 2, é possível notar que 2011 é o ano em que o conjunto de empresas tratadas atinge a maior participação dentro da amostra total disponível de empresas – representando cerca de 42%.

Gráfico 2. Densidade de empresas apoiadas (total de empresas após filtro por Cnae – %)



Fonte: Elaboração própria, com base no MARVIm.

O passo seguinte do processo é estimar, por meio da regressão logística, os *p-scores* para cada observação da base. O modelo implementa os passos descritos previamente para a especificação da equação 2. Nesse caso em particular e considerando 2010 o ano do tratamento,²⁵ foram escolhidas 11 ($k^* = 11$) variáveis explicativas, entre elas: anos de estudo médio da mão de obra, percentual de empregados com nível

²⁴ Exceto para os anos de 2007 e 2012, quando a amostra cai drasticamente. Esse comportamento é consequência da base da Serasa utilizada pelo MARVIm, e é por isso que, neste trabalho, as análises se concentraram sempre no período 2008-2011.

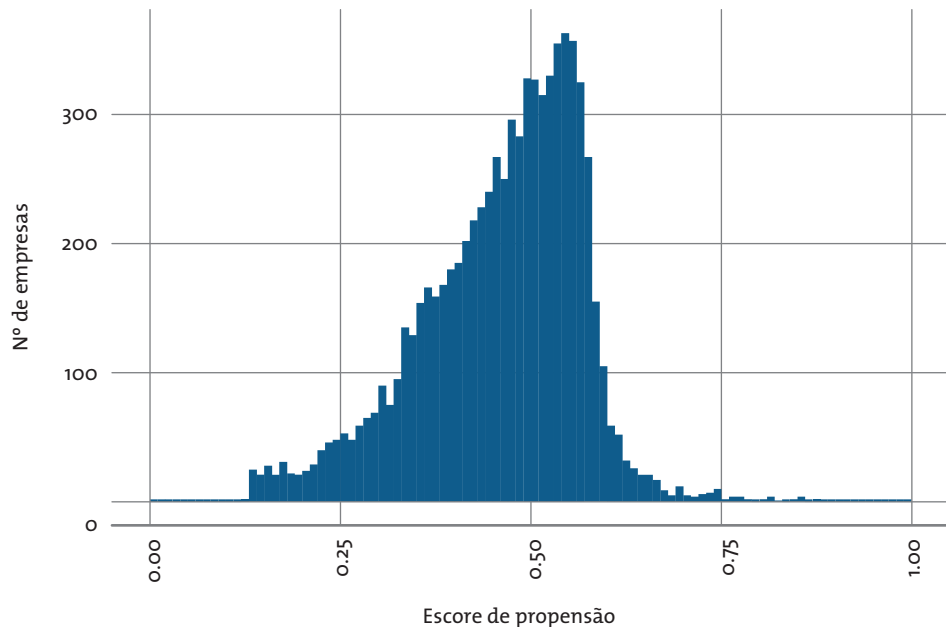
²⁵ Portanto, foram consideradas firmas tratadas todas aquelas unidades balanceadas que receberam do BNDES um financiamento para a aquisição de bens de capital em 2010.

superior, margem Ebitda (do inglês, *earnings before interest, taxes, depreciation and amortization*), ativo financeiro e ativo imobilizado. A equação do *p-score* contou também com outros 12 termos iterados, totalizando 23 covariadas para *V*.

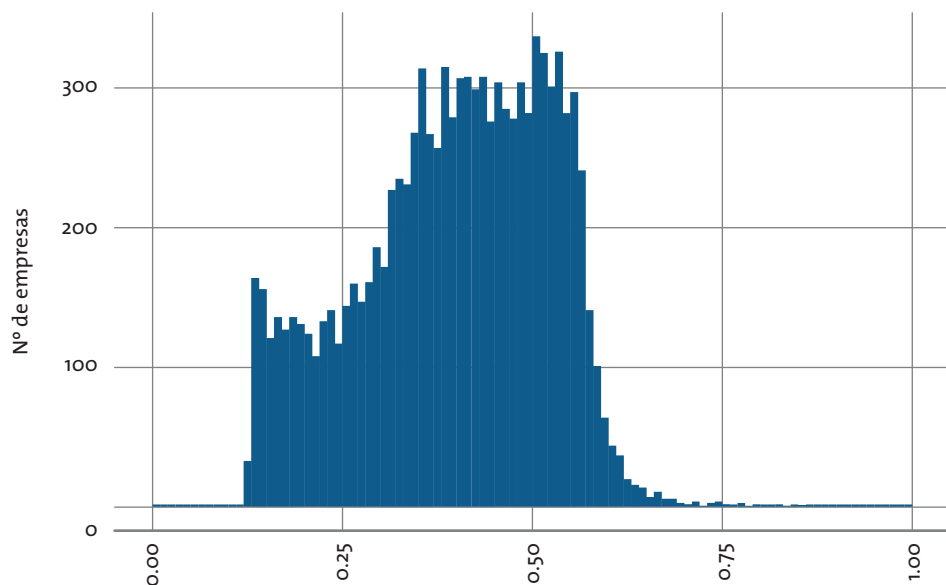
Ao utilizar os resultados estimados nesse primeiro estágio, o modelo restringe novamente a base de dados, excluindo, dessa vez, as empresas que não fazem parte da região de suporte comum – definida conforme os procedimentos descritos na subseção 3.3. A inspeção visual do Gráfico 3 é uma das etapas importantes para

Gráfico 3. Distribuição dos *p-scores* para tratados e controles (aquisição de bens de capital)

3a. Tratados



3b. Não tratados

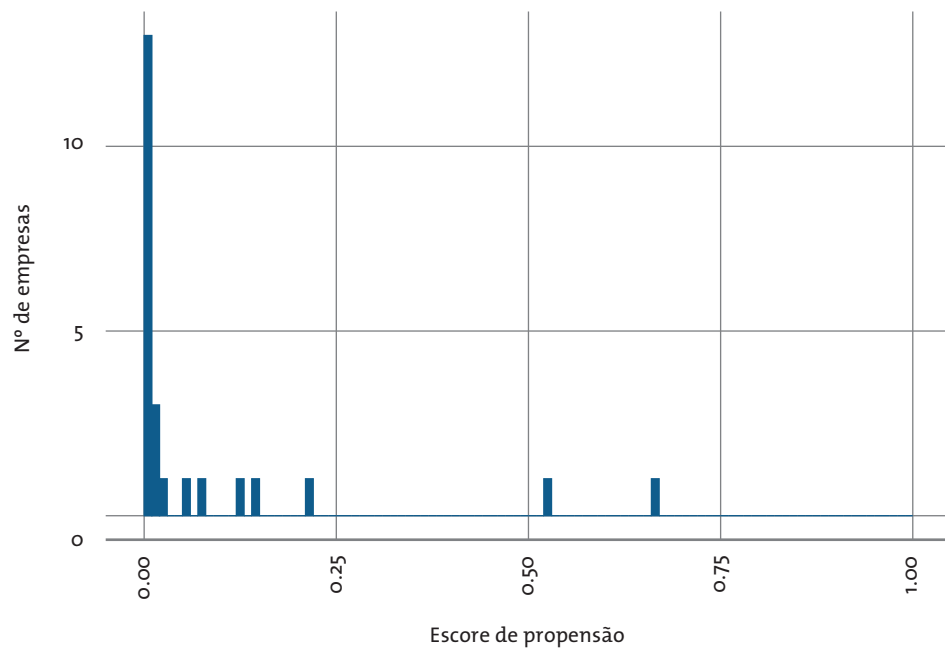


Fonte: Elaboração própria, com base no MARVIm.

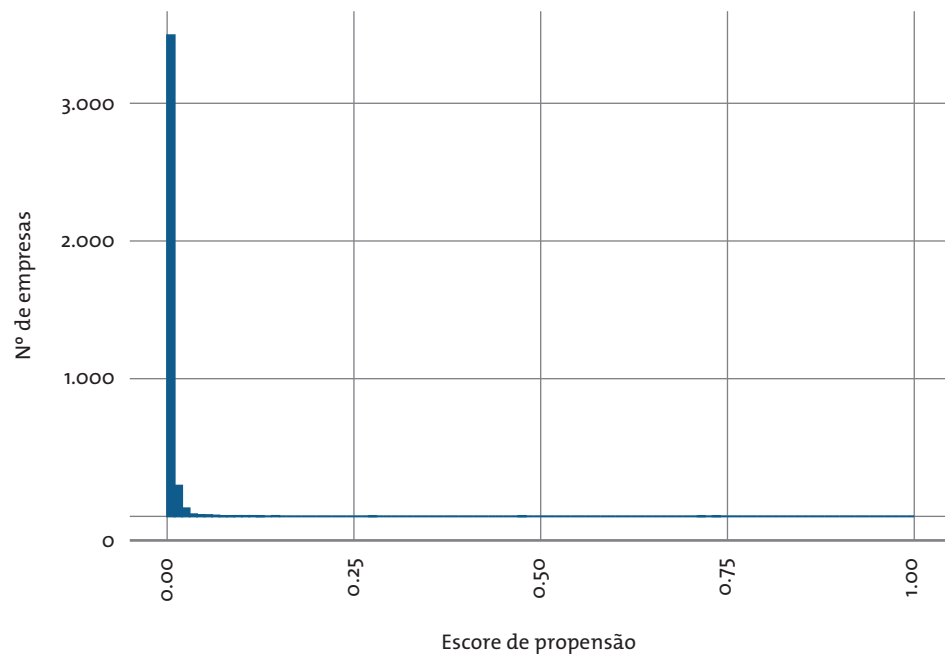
julgar a qualidade do balanceamento. Nesse caso, é possível notar que há razoável sobreposição entre as distribuições de p -score para tratados e não tratados. O alcance das distribuições também é bom. Elas têm densidades elevadas para trechos que vão desde aproximadamente 0,15 até cerca de 0,6 – o alcance total chega a quase 0,9. Pela inspeção visual, há bons indicativos de que a estimação do primeiro estágio não sofre de problemas de micronumerosidade e que deve permitir um pareamento bem balanceado.

Gráfico 4. Distribuição dos p -scores para tratados e controles (Profarma)

4a. Tratados



4b. Não tratados



Fonte: Elaboração própria, com base no MARVIm.

Para que fique claro esse argumento, o Gráfico 4 traça as mesmas curvas em uma avaliação feita para empresas apoiadas pelo BNDES Profarma. A observação visual, nesse caso, deixa evidente o problema de micronumerosidade, com a distribuição dos tratados sendo formada fundamentalmente por “saltos” que a fazem parecer uma função discreta. Além disso, a curva dos não tratados é excessivamente concentrada em *p-scores* muito baixos, tendo baixa representatividade para faixas superiores a 0,1. No exemplo do Profarma, portanto, a mera inspeção visual já é capaz de sinalizar a baixa confiabilidade das estimativas obtidas pelos métodos implementados no MARVIm.

O passo seguinte do modelo é o pareamento, implementado pelo algoritmo de vizinho mais próximo, conforme descrito na subseção 3.2. Retornando ao exemplo dos financiamentos para aquisição de bens de capital e restringindo a análise ao ano de 2010, a observação das tabelas 4 e 5 permite avaliar como o pareamento afetou as médias, os desvios-padrão e as métricas formais de balanceamento discutidas anteriormente.²⁶ É possível notar que, após o pareamento, a amostra de controle é, na média, muito mais parecida com a amostra de empresas apoiadas.

Considere-se, a título de ilustração, o percentual de funcionários graduados. Originalmente, as empresas financiadas pelo BNDES apresentavam um valor médio de 0,035, contra 0,077 no grupo de controle. Depois do pareamento, os valores passaram, respectivamente, para 0,030 e 0,031. Também as métricas formais apresentaram valores bem-comportados depois do pareamento. Exceto pelo ativo imobilizado e pela margem Ebitda, cujas distribuições dos controles continuam apresentando caudas mais pesadas que o esperado, as amostras pareceram bem balanceadas depois do pareamento. De forma geral, existem evidências favoráveis à comparabilidade entre os grupos depois do pareamento – cenário bastante diferente daquele observado para a amostra inicial.

Tabela 4. Balanceamento da amostra: antes do pareamento

Variável	\bar{X}_c	\bar{X}_t	S_c	S_t	$\hat{\Lambda}_{ct}$	$\hat{\Gamma}_{ct}$	$\hat{\Pi}_c^{0.05}$	$\hat{\Pi}_t^{0.05}$
Anos de estudo (média)	11.393	10.489	1.92	1.662	(0.5028)	(0.147)	0.143	0.0340
Ativo financeiro	17,713.135	19,619.951	210,869.98	291,493.202	0.0075	0.324	0.054	0.0459
Capital de giro	10,705.961	10,378.757	168,999.98	309,305.417	(0.0013)	0.604	0.069	0.0368
Exportações	4,055.320	7,511.185	68,385.25	193,146.139	0.0239	1.038	0.021	0.0281
Faturamento líquido	101,294.101	136,671.606	489,924.89	1,756,666.559	0.0274	1.277	0.053	0.0466
Funcionários graduados (%)	0.077	0.035	0.14	0.076	(0.3804)	(0.589)	0.097	0.0066
Idade da firma	20.938	21.762	13.38	12.971	0.0626	(0.031)	0.046	0.0182
Ativo imobilizado	36,957.113	67,099.533	336,076.33	1,783,988.545	0.0235	1.669	0.109	0.0316
Margem Ebitda	0.061	0.239	3.43	4.023	0.0477	0.159	0.101	0.0231

(Continua)

²⁶ O MARVIm permite definir um ano em particular para a análise ou definir uma estimação em que as diferentes observações, em diferentes anos, são alinhadas e empilhadas como em um *event-study*. Nesse segundo caso, as estimações contam com *dummies* para controlar efeitos específicos de cada ano.

(Continuação)

Variável	\bar{X}_c	\bar{X}_t	S_c	S_t	$\hat{\Delta}_{ct}$	$\hat{\Gamma}_{ct}$	$\hat{\Pi}_c^{0.05}$	$\hat{\Pi}_t^{0.05}$
Massa salarial (dez.)	739.373	806.761	3,681.08	8,726.994	0.0101	0.863	0.067	0.0387
Passivo financeiro	18,069.140	23,599.779	151,643.56	397,877.348	0.0184	0.965	0.068	0.0419
Produtividade do trabalho	249.073	46.857	12,142.47	754.531	(0.0235)	(2.778)	0.127	0.0131
<i>p-score</i>	0.370	0.458	0.15	0.120	0.6462	(0.211)	0.134	0.0677
<i>p-score-lin</i>	(0.680)	(0.142)	1.43	1.356	0.3861	(0.052)	0.134	0.0677

Fonte: Elaboração própria, com base no MARVIm.

Para o exemplo apresentado, a análise de qualidade do balanceamento se comportou conforme o esperado e, nesse sentido, em concordância com as hipóteses necessárias para a confiabilidade dos métodos aplicados pelo MARVIm. O passo seguinte do modelo é a especificação da equação 1, que é novamente implementado automaticamente por *forward selection*, para cada medida de resultado que se deseja analisar. Voltando ao exemplo retratado, e considerando o impacto sobre o emprego médio, o modelo seleciona 17 covariadas para compor o termo X da equação 1. Entre elas, constam: idade da firma, ativo total, anos médios de estudo da mão de obra, resultado financeiro, margem Ebitda etc. Diversas dessas variáveis já estavam contempladas no primeiro estágio. Contudo, conforme apontam Hirano e Imbens (2001), a combinação de métodos – utilizando o *p-score* com o ajuste por regressão – é benéfica porque ajuda a controlar eventual heterogeneidade que ainda persista na amostra, mesmo depois do pareamento.

Tabela 5. Balanceamento da amostra: depois do pareamento

Variável	\bar{X}_c	\bar{X}_t	S_c	S_t	$\hat{\Delta}_{ct}$	$\hat{\Gamma}_{ct}$	$\hat{\Pi}_c^{0.05}$	$\hat{\Pi}_t^{0.05}$
Anos de estudo (média)	10.570	10.43	1.562	1.594	(0.0902)	0.021	0.051	0.049
Ativo financeiro	7,831.482	9,340.13	37,838.253	62,303.355	0.0293	0.499	0.050	0.050
Capital de giro	8,459.948	9,542.78	51,505.816	55,709.031	0.0202	0.078	0.065	0.039
Exportações	1,919.933	2,524.68	23,016.223	35,800.955	0.0201	0.442	0.020	0.030
Faturamento líquido	72,347.447	85,231.70	230,542.584	345,613.487	0.0439	0.405	0.047	0.055
Funcionários graduados (%)	0.031	0.03	0.064	0.062	(0.0153)	(0.036)	0.027	0.023
Idade da firma	21.645	21.78	13.330	12.946	0.0102	(0.029)	0.042	0.018
Ativo imobilizado	18,915.839	25,690.08	94,519.554	178,101.454	0.0475	0.634	0.103	0.036
Margem Ebitda	0.171	0.20	0.350	0.314	0.0784	(0.108)	0.077	0.032
Massa salarial (dez.)	518.508	578.54	1,388.072	1,719.307	0.0384	0.214	0.058	0.046
Passivo financeiro	10,092.434	12,341.07	48,589.652	72,949.119	0.0363	0.406	0.068	0.044
Produtividade do trabalho	48.414	44.79	618.300	743.648	(0.0053)	0.185	0.081	0.032
<i>p-score</i>	0.447	0.46	0.099	0.107	0.1332	0.072	0.039	0.066
<i>p-score-lin</i>	(0.228)	(0.17)	0.435	0.467	0.1245	0.070	0.039	0.066

Fonte: Elaboração própria, com base no MARVIm.

Com a especificação da equação 1 definida, o MARVIm é capaz de estimar o impacto da intervenção utilizando as técnicas definidas na subseção 3.2. A Tabela 6 traz, para um conjunto selecionado de variáveis, os resultados obtidos com cada estimador para o caso do financiamento à aquisição de bens capital em 2010, considerando o impacto nesse mesmo ano.

Para que se possa julgar a magnitude dos efeitos, vale notar que a amostra inicial de empresas com informações encontradas na base de análise do MARVIm totalizou, nesse exercício, 8.794 Cadastros Nacionais da Pessoa Jurídica (CNPJ) e representava R\$ 22,9 bilhões em financiamentos – ver tabelas 2 e 3. O desembolso médio foi de R\$ 2,6 milhões. O impacto estimado sobre o nível de emprego foi positivo e estatisticamente significativo em todos os casos. O mesmo ocorreu com o faturamento bruto. Ou seja, as técnicas aplicadas apontaram impacto positivo do financiamento do BNDES à aquisição de máquinas e equipamentos sobre o faturamento bruto e o nível de emprego das firmas beneficiárias. Considerando as estimativas mais conservadoras, essa intervenção teria gerado, na média, 12 empregos adicionais por empresa, bem como uma expansão do faturamento bruto da ordem de R\$ 1,5 milhão – em ambos os casos, considerando o período de até um ano após a contratação da operação no BNDES.

Tabela 6. Impacto médio estimado (variáveis selecionadas)

Variável de impacto	$\hat{\tau}_{ols}$	$\hat{\tau}_{psm-mm}$	$\hat{\tau}_{psw}$
Emprego	12,65***	13,30***	13,14***
Faturamento bruto	1.493***	2.366***	2.317***
Resultado líquido	980,55	306,25	558,68*
Produtividade do trabalho	(80,37)	(16,26)	(14,58)

Fonte: Elaboração própria, com base no MARVIm.

* Significante a 10%. ** Significante a 5%. *** Significante a 1%.

Nos casos do resultado líquido (lucro líquido) e da produtividade do trabalho, não foi possível obter impactos estatisticamente significantes. A interpretação, nesses casos, é de que não há diferença de *performance* das firmas apoiadas quando comparadas com seus contrafactuais. Ou seja, o financiamento para aquisição de bens de capital não impactou lucro nem produtividade do trabalho, ao menos no curto prazo que define esse exercício.

É evidente que esses são os resultados obtidos para uma especificação particular, que considerou apenas empresas financiadas em 2010. Contudo, a vantagem da automação é justamente a capacidade de refazer com grande agilidade o mesmo exercício com diversas especificações distintas, o que permite considerar um conjunto mais amplo de evidências para se concluir a respeito da efetividade das intervenções. Aproveitando-se da flexibilidade do método desenvolvido

neste estudo, a seção seguinte analisará de forma mais completa uma série de intervenções, buscando conhecer o impacto de curto prazo sobre as empresas apoiadas pelo BNDES.

4. Impactos de curto prazo sobre as firmas beneficiadas

Nesta seção, o MARVIm é utilizado para realizar avaliações sobre o impacto de uma série de instrumentos de apoio do BNDES em uma janela de tempo ainda próxima do tratamento (um ano, no máximo). Os resultados serão denominados, então, de impactos de curto prazo. O objetivo é mostrar como essa ferramenta pode ser empregada para extrair lições sobre a atuação do Banco, ao analisar as especificidades de cada intervenção. O Quadro 1 exhibe os instrumentos de apoio que foram analisados. Como o modelo escolhe as variáveis explicativas autonomamente em cada caso e as intervenções foram definidas de forma a agregar tipos de financiamento razoavelmente homogêneos, garante-se maior consistência para a estratégia de identificação.

Por padrão, as variáveis de interesse analisadas foram: o faturamento bruto, o número médio de empregados, a produtividade do trabalho e o resultado líquido. O fato de as mesmas variáveis de resultado terem sido escolhidas na avaliação de diferentes tipos de intervenções objetiva apenas ilustrar o potencial do MARVIm. O uso da ferramenta está restrito às variáveis utilizadas e, com uma base de dados mais ampla, seria possível avaliar outras dimensões de desempenho.

É importante notar que, em alguns casos, os normativos que regem os instrumentos avaliados sequer fazem menção à expectativa de impacto sobre as variáveis elencadas. Ou seja, o efeito avaliado não foi explicitado pelo definidor do instrumento, ainda que possa parecer desejável do ponto de vista econômico. Mesmo assim, o conjunto de avaliações realizadas permite ampliar a compreensão de aspectos ligados ao:

- desempenho das firmas – possivelmente aspectos motivadores da busca de apoio no BNDES;
- comportamento do emprego – considerando que, historicamente, o *funding* do BNDES contou com grande participação de recursos oriundos do Fundo de Amparo ao Trabalhador (FAT);
- comportamento da produtividade – dimensão comumente apontada como determinante do crescimento de longo prazo da economia;
- canal de transmissão do apoio até o objetivo.

Em algumas das análises, houve o interesse particular de investigar o impacto sobre outras dimensões, como, por exemplo, exportações (presente no objetivo), no caso das linhas do Exim.

Explorando a flexibilidade da ferramenta, foram feitos diferentes recortes para testar minimamente a robustez dos resultados. Como via de regra, as análises de impacto foram feitas considerando-se apenas anos específicos (por exemplo, 2009) e em uma estimativa agregada do período todo (2007 a 2011). Exploraram-se também possibilidades de haver impacto mensurável em até um ou dois anos depois do tratamento – a análise se concentra, portanto, nos resultados de curto prazo do apoio. Ademais, realizaram-se tais análises prioritariamente para duas bases de dados, uma antes e outra depois da exclusão de firmas consideradas *outliers*.

Quadro 1. Definições das intervenções avaliadas

Intervenção	Alvo	Exceto	Tratamento
Operações diretas	Empresas que contrataram na modalidade direta com o BNDES	Fundos de investimento nos quais o Banco possui cotas e operações não reembolsáveis	Contratação da operação
BNDES Profarma	Empresas que contrataram operações no BNDES Profarma	–	Contratação da operação
BNDES Exim Pré-embarque	Empresas que contrataram operações no BNDES Exim Pré-embarque	–	Contratação da operação
BNDES Exim Pós-embarque	Subfornecedores contratados por empresas apoiadas pelo produto para exportar serviços de engenharia	Operações de financiamento a aeronaves e beneficiários diretos dos serviços de engenharia	Liberação dos recursos
Cartão BNDES – fornecedores de software	Empresas que forneceram <i>software</i> via Cartão BNDES	–	Captura da transação
Cartão BNDES – clientes	Empresas com Cartão BNDES que realizaram transações no produto	–	Captura da transação
Fabricantes de bens de capital	Empresas que forneceram bens de capital por meio do BNDES (Cartão BNDES+ BNDES Finame)	–	Captura da transação (Cartão BNDES) + contratação da operação (BNDES Finame)
Aquisição de bens de capital	Empresas que adquiriram bens de capital financiados pelo BNDES (Cartão BNDES + Finame + Finem)	Finame Fabricantes Produção e capital de giro	Captura da transação (Cartão BNDES) + liberação da operação (BNDES Finame + BNDES Finem)
BNDES Automático – Investimento	Empresas que contrataram operações no BNDES Automático para investimentos	Crédito agrícola, giro, refinanciamentos e reestruturação financeira	Contratação da operação
BNDES Progeren	Empresas beneficiárias do BNDES Progeren	–	Liberação dos recursos

Fonte: Elaboração própria.

As tabelas 7 a 10 resumem os resultados obtidos, reportando apenas os casos em que a aplicação das técnicas de pareamento permitiu um balanceamento satisfatório da amostra – os demais casos são comentados nas subseções subsequentes. Como era de se esperar, o MARVIm se mostrou muito mais adequado para lidar com intervenções com um número elevado de beneficiados – caso da aquisição de bens de capital, por exemplo.

De forma geral, foram considerados robustos os casos nos quais houve certa estabilidade no sinal e na significância estatística dos coeficientes nas diferentes especificações. Analogamente, foram reportados impactos positivos ou negativos nos casos em que o sinal do coeficiente estimado se manteve estável nas diferentes especificações.

Tabela 7. Aquisição de bens de capital – MPME

Estimação	Ano	Média para tratados	Coefficientes	Significância	Impacto estimado (%)
Faturamento (R\$ MM)	2009	30,94	[1,43 ; 1,19]	***	[4,62 ; 3,83]
	2010	40,05	[2,37 ; 1,49]	***	[5,91 ; 3,73]
	Empilhado	39,30	[1,69 ; 0,85]	***	[4,31 ; 2,15]
Emprego (unid.)	2009	198,37	[9,42 ; 9,15]	***	[4,75 ; 4,61]
	2010	231,79	[13,30 ; 12,65]	***	[5,74 ; 5,46]
	Empilhado	228,21	[10,04 ; 9,23]	***	[4,40 ; 4,04]
Investimento (taxa)	2009	0,040	[0,02 ; 0,00]	***	[50 ; 1]
	2010	0,068	[0,07 ; 0,03]	**	[103 ; 44]
	Empilhado	0,051	[-6,59 ; -92,35]	-	-

Fonte: Elaboração própria, com base no MARVIm.

Notas: Para os coeficientes, são reportados o maior e o menor coeficiente estimado.

* indica significância a 10%, ** indica significância a 5%, *** indica significância a 1%.

Tabela 8. Aquisição de bens de capital – grande

Estimação	Ano	Média para tratados	Coefficientes	Significância	Impacto estimado (%)
Faturamento (R\$ MM)	2009	1.121,19	[78,81 ; 59,95]	**	[7,03 ; 5,35]
	2010	1.372,39	[91,98 ; 3,12]	-	-
	Empilhado	1.677,48	[88,89 ; 54,01]	***	[5,30 ; 3,22]
Emprego (unid.)	2009	2.919,04	[44,68 ; -45,43]	-	-
	2010	3.187,01	[69,98 ; 8,10]	-	-
	Empilhado	3.352,92	[77,88 ; 45,52]	-	-
Investimento (taxa)	2009	0,052	[0,02 ; -0,00]	-	-
	2010	0,039	[0,03 ; 0,01]	*	[76,37 ; 25,46]
	Empilhado	0,043	[0,00 ; -0,23]	-	-

Fonte: Elaboração própria, com base no MARVIm.

Notas: Para os coeficientes, são reportados o maior e o menor coeficiente estimado.

* indica significância a 10%, ** indica significância a 5%, *** indica significância a 1%.

Tabela 9. BNDES Automático – Investimento

Estimação	Ano	Média para tratados	Coefficientes	Significância	Impacto estimado (%)
Faturamento (RS MM)	2009	176,18	[21,45 ; -1,74]	-	-
	2010	283,08	[22,23 ; -1,67]	-	-
	Empilhado	232,24	[12,72 ; 8,27]	-	-
Emprego (unid.)	2009	604,32	[9,71 ; -0,40]	-	-
	2010	609,63	[28,71 ; -0,48]	-	-
	Empilhado	743,16	[26,84 ; -2,29]	-	-
Investimento (taxa)	2009	0,013	[0,01 ; 0,00]	***	[75,90 ; 30,36]
	2010	0,012	[0,01 ; 0,01]	***	[81,25 ; 81,25]
	Empilhado	0,013	[0,01 ; 0,00]	***	[79,06 ; 39,53]

Fonte: Elaboração própria, com base no MARVIm.

Notas: Para os coeficientes, são reportados o maior e o menor coeficiente estimado.

* indica significância a 10%, ** indica significância a 5%, *** indica significância a 1%.

Tabela 10. Cartão BNDES

Estimação	Ano	Média para tratados	Coefficientes	Significância	Impacto estimado (%)
Faturamento (RS MM)	2009	30,53	[1,23 ; 0,86]	***	[4,03 ; 2,82]
	2010	37,84	[1,54 ; 0,59]	***	[4,06 ; 1,55]
	Empilhado	39,74	[1,26 ; 0,26]	***	[0,66 ; 0,03]
Emprego (unid.)	2009	186,08	[13,26 ; 11,80]	***	[7,13 ; 6,34]
	2010	210,86	[13,29 ; 11,88]	***	[6,30 ; 5,63]
	Empilhado	215,41	[10,55 ; 9,28]	***	[4,90 ; 4,31]
Investimento (taxa)	2009	0,029	[0,02 ; 0,01]	**	[69,30 ; 34,65]
	2010	0,053	[0,02 ; 0,01]	***	[37,98 ; 18,99]
	Empilhado	0,038	[-1,90 ; 0,1]	-	-

Fonte: Elaboração própria, com base no MARVIm.

Notas: Para os coeficientes, são reportados o maior e o menor coeficiente estimado.

* indica significância a 10%, ** indica significância a 5%, *** indica significância a 1%.

Em cada tabela, a partir da quarta coluna e até a sexta, têm-se os coeficientes estimados – maior e menor entre os estimadores, suas significâncias e o *range* de impacto estimado em percentual. O processo de análise para cada intervenção será detalhado no restante da seção de resultados. Por ora, a observação das referidas tabelas faz com que alguns resultados em particular se destaquem nessa análise panorâmica implementada sobre a efetividade do BNDES.

Impactos robustos e positivos sobre as firmas apoiadas foram observados em alguns contextos. O crédito para aquisição de bens de capital afetou positivamente o faturamento bruto, o nível de emprego e o investimento das MPMEs.²⁷ O Cartão BNDES, outra intervenção concentrada em MPMEs, também parece afetar o faturamento, o emprego e a taxa de investimento.²⁸ Por fim, no BNDES

²⁷ Outros estudos já haviam identificado impacto positivo do BNDES Finame sobre o investimento das firmas. Ver, por exemplo, Machado e outros (2016) e Cavalcanti e Vaz (2017).

²⁸ Outros estudos também já haviam identificado impacto positivo do Cartão BNDES sobre o emprego dos beneficiados. Ver, por exemplo, Machado, Parreiras e Peçanha (2011).

Automático – Investimento, o próprio investimento é positivamente afetado. De maneira geral, parece que a adicionalidade do Banco sobre MPMEs se mostrou maior do que sobre as grandes empresas – o que corrobora a ideia de que o primeiro grupo sofreria mais com restrição de crédito.

Chama a atenção o fato de que nenhuma das intervenções pareceu impactar a produtividade do trabalho dos beneficiados. A trajetória recente da produtividade na economia brasileira tem sido uma preocupação frequente. É natural que o BNDES seja convocado, nos próximos anos, a desenvolver políticas que promovam ganhos nessa dimensão. As intervenções analisadas neste estudo não parecem ser efetivas para esse fim. Uma parte da explicação pode estar na janela de impacto utilizada. Os resultados sobre a produtividade provavelmente demoram mais do que dois anos para se fazer sentir, requerendo um intervalo de análise maior do que o implementado. Pode não haver impacto no curto prazo, mas sim no longo prazo.²⁹

As subseções seguintes discutem brevemente a qualidade do balanceamento e o comportamento dos resultados estimados em cada uma das avaliações realizadas.

4.1 Operações diretas

Foram analisadas setecentas observações, relativas a 371 empresas tratadas, correspondentes a 53% do total desembolsado nos produtos.³⁰ Mesmo depois do balanceamento da amostra, notam-se diferenças significativas entre as médias de empresas tratadas e não tratadas de variáveis que guardam certa correlação com o porte das empresas (notadamente ativo financeiro, imobilizado e patrimônio líquido). Com isso, a região de suporte comum ficou mais estreita que o desejável – com baixa densidade para *p-scores* acima de 0,01. Em alguma medida, isso ocorreu devido à baixa densidade de tratados com relação à base de análise total.

As diferentes análises realizadas sugerem um efeito positivo, embora não robusto (não mantém significância estatística em todas as especificações, em especial quando controlado pelo porte da empresa), da contratação do crédito sobre faturamento bruto, quantidade média de empregos e investimento. Há a necessidade de realização de análises mais aprofundadas sobre o tema (preferencialmente com bases de maior abrangência temporal, buscando aumentar a região de suporte comum e possibilitar a avaliação de efeitos de mais longo prazo).

²⁹ Trabalho recente dos autores Cavalcanti e Vaz (2017) identifica impacto positivo do BNDES Finance sobre a produtividade das firmas apoiadas. Esse resultado, contudo, vale apenas nos casos em que o apoio do BNDES é permanente.

³⁰ Alguns recortes testados, contudo, apresentaram amostras menores – como no caso da análise das firmas tratadas no ano de 2009, no qual a amostra se reduz a 182 empresas.

4.2 BNDES Profarma

No período observado, o BNDES Profarma beneficiou um total de 81 firmas (em média, 12 por ano). Apenas 42 empresas (51,8% do total de beneficiárias) foram de fato encontradas nas bases de dados compiladas para a avaliação. Para o período em análise, o valor total financiado pelo BNDES para esse programa foi de R\$ 2,9 bilhões, dos quais R\$ 2,3 bilhões, ou 79%, foram encontrados nas bases de dados. Nessas condições, o limitado número de unidades tratadas encontradas impediu que a análise pudesse ser realizada para variados recortes.

O balanceamento mostrou-se eficaz em aproximar as médias de empresas tratadas e não tratadas, e a região de suporte comum entre os dois grupos de empresas chegou a quase 0,75. Todavia, a distribuição em ambos os casos é bastante assimétrica à esquerda, com a maior parte das observações centradas em zero. A média de anos de estudo dos trabalhadores vinculados à empresa parece ser a variável observada que mais contribui para o valor do *p-score*, ou seja, para o *status* de tratamento da firma.

As diferentes análises realizadas não apontam um efeito significativo da contratação de crédito sobre qualquer das variáveis analisadas, seja em nível, seja em diferenças. O que fica claro é que, em razão do reduzido número de tratados na amostra, a avaliação sofre com o problema de micronumerosidade.

4.3 BNDES Exim Pré-embarque

Considerando a totalidade do período entre 2006 e 2012, foram encontradas na base de análise 875 empresas, de um total de 1.483, que tiveram operações de crédito contratadas, responsáveis por 47% do total desembolsado no produto. O balanceamento se mostrou eficaz em aproximar as médias de empresas tratadas e não tratadas. Mesmo assim, variáveis típicas de porte, como o ativo total, permaneceram com grande desbalanceamento entre os grupos de tratados e controles – sendo os grupos de controle compostos por empresas, na média, menores. A região de suporte comum entre os dois grupos de empresas ficou acima de 0,9. Todavia, a distribuição é bastante assimétrica à esquerda, com a maior parte das observações próxima a zero.

Foram realizadas diferentes análises, incluindo a variável exportação. Na estimação agregada de todo o período, os resultados apontaram para um impacto positivo no valor exportado e no faturamento bruto. Esse resultado não foi robusto, contudo, na análise de anos específicos. Dessa forma, não é possível concluir a respeito da efetividade do apoio sobre as variáveis analisadas.

4.4 BNDES Exim Pós-embarque

Assim como no caso do Pré-embarque, dada a natureza do produto, foi analisado também o valor exportado. Utilizaram-se duas abordagens, uma procurando efeito

a partir do ano da liberação e outra considerando que o efeito poderia anteceder o ano de liberação em um ou dois anos, uma vez que a empresa pode ser fornecedora para uma segunda empresa que efetivamente realizou a obra no exterior e ter sido impactada antes de a segunda receber a liberação.

A região de suporte comum entre os grupos de tratados e controles ficou limitada ao máximo de 0,25. Para a primeira abordagem, foram analisados os anos de 2008 e 2009, pela maior concentração de empresas financiadas. Nesses dois anos, não foi encontrada significância para a variável exportação. Já o faturamento bruto apresentou significância e consistência de sinal em 2008. Porém, para 2009, o resultado encontrado não teve significância estatística.

Na segunda abordagem, analisou-se apenas o ano de 2009. Considerando o possível efeito com um ano de antecedência, o valor exportado continuou com estimadores de baixa significância. O mesmo comportamento foi observado para o faturamento bruto, assim como quando se considerou o efeito com dois anos de antecedência.

4.5 Cartão BNDES – clientes

As análises foram prioritariamente realizadas utilizando a base de dados limitada a firmas com faturamento até R\$ 300 milhões (MPMEs) – cujas características são mais compatíveis com as empresas tratadas. De todas as firmas apoiadas pelo produto, foram encontradas 5.789 na base de análise. Essas empresas realizaram 12.501 transações (7% do total financiado pelo produto no período), de um total de 107.288 observações disponíveis. Esse corte permitiu melhorar o *matching* entre as empresas tratadas e não tratadas. Embora com maior densidade em *p-scores* até 0,5, a região de suporte comum se estendeu até *p-scores* de 0,85, evidências de um bom pareamento.

As diferentes análises realizadas apontam para um efeito positivo, significativo e robusto (em todas as especificações) do apoio via Cartão BNDES sobre a quantidade média de empregos e sobre o investimento da firma. Além disso, quando controlado pelo porte da empresa, há efeito robusto e estatisticamente significativo sobre o faturamento bruto. Há, ainda, indicação de efeito negativo, porém não robusto (com variações de sinal), sobre a produtividade do trabalho.

4.6 BNDES – fornecedores de *softwares*

Durante o período observado, o Cartão BNDES realizou 3.284 transações envolvendo fornecedores de *software*, que totalizaram aproximadamente R\$ 662,0 milhões. Dessas transações, apenas 117 (ou 3,6%) foram encontradas na base de análise do MARVIm. Essa amostra foi responsável por R\$ 128,9 milhões (ou 19,5%) do valor total das transações no período.

O balanceamento mostrou-se eficaz em aproximar as médias de empresas tratadas e não tratadas. A região de suporte comum entre os dois grupos de empresas superou 0,50. A distribuição, em ambos os casos, é bastante assimétrica à esquerda, com a maior parte das observações centrada no intervalo entre 0 e 0,10. Entre as variáveis observadas para a construção do *p-score* a receber o tratamento, verificou-se que a mais importante é a média de anos de estudo dos trabalhadores vinculados à empresa em cada ano de referência.

As diferentes análises realizadas não apontam um efeito significativo da contratação de crédito sobre a maioria das variáveis de interesse (faturamento bruto, número médio de empregados e resultado líquido), seja em nível, seja em diferenças. Em relação à produtividade do trabalho, a análise revelou efeito negativo em nível, ainda que estatisticamente não significativo em diferenças.

4.7 Fabricantes de bens de capital

Em todos os exercícios realizados, não foi possível encontrar um padrão claro entre a intervenção e algumas variáveis que, supostamente, revelariam impactos (faturamento, emprego, lucro e produtividade). Por exemplo, ao analisar o impacto para anos específicos, obtiveram-se resultados antagônicos nas variáveis de faturamento (impacto positivo em 2007 e negativo em 2009) e de número médio de empregados (positivo em 2008 e negativo em 2009). Para demais variáveis e períodos, os resultados são ora não conclusivos, ora estatisticamente não significantes.

O pequeno número de firmas tratadas encontradas na base de análise (aproximadamente 10%), o desbalanceamento entre o número de observações nos grupos de tratamento e de controle e a concentração de *p-score* no primeiro quartil do domínio dessa variável podem contribuir para a ausência de identificação de impacto nos resultados do presente exercício.

4.8 Aquisição de bens de capital

Conforme já antecipado, essa análise contou com um elevado número de empresas – fundamentalmente em função dos produtos BNDES Finame e Cartão BNDES. Isso permitiu a implantação de um detalhamento adicional: a avaliação de impacto dividiu a amostra por faixas de porte, separando MPMEs de grandes empresas.³¹ No primeiro grupo, foi possível trabalhar com 25.735 observações tratadas, que representavam cerca de 25% do total do financiamento atribuído à intervenção. No segundo grupo, a amostra contou com 2.210 observações, que abarcavam 35% do total financiado. Em ambos os casos, cerca de 50% das observações estavam concentradas nos anos de 2010 e 2011.

³¹ A separação seguiu a regra de porte do BNDES, que toma como grandes as empresas com faturamento bruto superior a R\$ 300 milhões.

Os histogramas de *p-score* apresentaram sistematicamente grande sobreposição entre tratados e não tratados, bem como uma densidade expressiva em intervalos adequados. As métricas formais também apontaram para amostras balanceadas depois do pareamento. Ou seja, as análises implementadas para essa intervenção passaram no crivo de qualidade do balanceamento.

Para as MPMEs, foi possível identificar impacto positivo, estatisticamente significativo e robusto sobre o faturamento bruto, número médio de empregados e investimento. O ano de 2010 apresentou também impacto positivo sobre o resultado líquido, mas esse resultado deixou de ser estatisticamente significativo em 2011. O impacto sobre a produtividade foi sistematicamente negativo, mas em nenhum dos casos se mostrou estatisticamente significativo.

No caso da amostra com empresas grandes, foi possível obter impacto robusto apenas sobre faturamento.

4.9 BNDES Automático – Investimento

A base de tratamento nessa análise inclui informações sobre 5.603 intervenções realizadas entre 2006 e 2012. Destas, apenas 485 foram encontradas na base de dados de referência, totalizando 8,6% das intervenções realizadas (ou 11% do valor contratado) no período. As amostras finais contêm entre 14 e 89 observações por ano, um número relativamente baixo de graus de liberdade para análises.

Para avaliar o BNDES Automático – Investimento, em função da natureza do instrumento, entre as várias especificações, optou-se também por utilizar uma base de dados restrita às empresas que reportaram informações sobre o imobilizado no ano de tratamento e no anterior. O pareamento não obteve bons resultados, principalmente porque houve forte concentração de *p-scores* no primeiro decil da distribuição. Ainda assim, observou-se impacto positivo no investimento das firmas apoiadas, não sendo possível verificar impacto sobre as demais variáveis.

4.10 BNDES Progeren

No caso da análise realizada na base completa e considerando a totalidade do período entre 2007 e 2011, foram analisadas 1.715 empresas que receberam o BNDES Progeren, correspondentes a 60% do total desembolsado nas linhas desse produto. O balanceamento mostrou-se eficaz em igualar as médias de empresas tratadas e não tratadas. Apesar disso, a região de suporte comum ficou mais estreita que o desejável – com baixa densidade para *p-scores* acima de 0,25.

As diferentes análises realizadas convergem no sentido de não apontarem um efeito significativo da liberação de crédito sobre as variáveis analisadas. Não obstante, tais resultados não permitem concluir pela ausência de impacto do programa estudado, uma vez que a região de suporte se mostrou demasiado estreita para que tais resultados sejam considerados robustos.

5. Considerações finais e os próximos avanços para o MARVIm

A lógica de eficiência na política pública tem merecido crescente atenção da sociedade. Por um lado, esse movimento pode ser entendido como sinal de amadurecimento econômico nacional, na medida em que as ações do Estado brasileiro passam a reconhecer mais claramente a lógica de que os recursos são escassos e devem, portanto, ser usados de forma a trazer o maior benefício possível. Por outro, ela impõe a todo agente público a missão de avaliar constantemente para prestar contas a respeito de sua atuação.

Esse movimento gera para instituições públicas, como o BNDES, uma necessidade: aumentar a escala de suas atividades de monitoramento e avaliação. Existem diversas iniciativas gestadas no DEAPE para responder a isso. O MARVIm é uma delas e possibilita a construção de informações a respeito da efetividade do BNDES por meio de avaliações de impacto quantitativas. É evidente que os resultados da ferramenta não pretendem substituir o esforço de elaborar avaliações de impacto artesanais, nem de buscar metodologias ainda mais robustas para a atribuição de causalidade – o ideal a ser perseguido está determinado por experiências de aleatorização.

O MARVIm demonstrou potencial para aplicar de forma ágil um conjunto de métodos de prateleira bastante utilizados em avaliação de impacto. É indiscutível que os resultados aqui obtidos merecem maior investigação, mas eles já oferecem um interessante conjunto de evidências capazes de traçar uma visão panorâmica, ainda que incompleta, a respeito da atuação do BNDES.

A interpretação mais geral dos resultados aponta para maior adicionalidade do apoio nos casos em que há maior concentração em beneficiários de menor porte, como é o caso da aquisição de bens de capital para MPMEs, do Cartão BNDES e, de certa forma, do BNDES Automático – Investimento. Houve, contudo, diversos casos em que não se obteve impacto ou em que não se observou robustez nos resultados. Em todos os casos, cabe a ressalva de que este estudo não se propõe a ser uma posição definitiva sobre a efetividade dos instrumentos avaliados, mas apenas uma primeira impressão, uma vez que o MARVIm apresenta diversas limitações.

Em particular, ficou clara a dificuldade dos métodos implementados em lidar com formas de atuação do BNDES que atingem um conjunto pequeno de empresas. Isso se explica porque os resultados não são conclusivos para boa parte da atuação direta do Banco. Uma maneira de contornar essa limitação seria por meio de bases de dados mais completas. Toda a análise apresentada neste estudo depende diretamente dos dados da Serasa – aos quais o BNDES não tem mais acesso desde 2012. Métodos quantitativos são intensivos em dados. É por isso que, para dar escala e confiabilidade a avaliações quantitativas desenvolvidas internamente, a

instituição precisa considerar: investimentos adicionais para aquisição de bases mais recentes e completas, atualizando a base da Serasa, por exemplo; e a construção de acordos que viabilizem o acesso ao amplo conjunto de dados produzidos por outras instituições do Estado brasileiro, como o IBGE e a Secex.

Mesmo com as melhores bases disponíveis, algumas políticas do BNDES têm escala muito pequena, do ponto de vista de número de beneficiários, para serem avaliadas por meio dos métodos apresentados no presente trabalho. É por isso que existe também a necessidade de incorporar ao MARVIm técnicas alternativas, que consigam lidar com casos de micronumerosidade.

Para que a aplicação de métodos quantitativos de avaliação possa abarcar de forma mais completa a atuação do Banco, é fundamental que se desenvolva a capacidade de captar impactos também sobre a localidade. Afinal, boa parte do financiamento da instituição tem como foco impactos que não podem ser captados sobre as empresas beneficiadas – isso vale, por exemplo, para toda a carteira de infraestrutura.

É por isso que os passos seguintes na agenda de desenvolvimento do MARVIm incluem duas importantes medidas. Em primeiro lugar, o DEAPE está desenvolvendo também a automatização de técnicas de controle sintético, mais adequadas à busca de causalidade em situações em que o número de beneficiários é baixo. Em segundo lugar, está sendo elaborada uma segunda base de dados contendo informações a respeito dos municípios brasileiros. Isso permitirá a avaliação de medidas cujo impacto se dá sobre a localidade. Concluídas essas medidas, o MARVIm será capaz de executar métodos quantitativos de avaliação de impacto que alcancem, com agilidade e rigor metodológico, a maior parte da atuação do BNDES.

Referências

AMBROZIO, A. M. H. P. *et al.* Credit scarcity in developing countries: an empirical investigation using Brazilian firm-level data. *Economia*, [S.l.], v. 18, n. 1, p. 73-87, Jan.-Apr. 2017. ISSN 1517-7580. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1517758016301412>>. Acesso em: 1º ago. 2018.

ANGRIST, J. D.; PISCHKE, J.-S. *Mostly harmless econometrics: an empiricist's companion*. Princeton: Princeton University Press, 2009.

BLUNDELL, R.; DIAS, M. C. Alternative approaches to evaluation in empirical microeconomics. *Portuguese Economic Journal*, [S.l.], v. 1, n. 2, p. 91-115, 2002.

- BONOMO, M.; BRITO, R.; MARTINS, B. *Macroeconomic and financial consequences of the after crisis government-driven credit expansion in Brazil*. Brasília: Banco Central do Brasil, 2014. (Working paper series, 378).
- CALIENDO, M.; KOPEINIG, S. Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys*, [S.l.], v. 22, n. 1, p. 31-72, 2008.
- CAVALCANTI, T.; VAZ, P. H. Access to long-term credit and productivity of small and medium firms: a causal evidence. *Economics Letters*, [S.l.], v. 150, p. 21-25, Jan. 2017.
- FILZMOSER, P.; GARRETT, R. G.; REIMANN, C. Multivariate outlier detection in exploration geochemistry. *Computers & Geosciences*, [S.l.], v. 31, p. 579-587, 2005.
- GADENNE, L. Tax me, but spend wisely: Sources of public finance and government accountability. *American Economic Journal: Applied Economics*, [S.l.], v. 9, n. 1, p. 274-314, Jan. 2017. Disponível em: <<http://www.aeaweb.org/articles?id=10.1257/app.20150509>>. Acesso em: 1º ago. 2018.
- HECKMAN, J. J.; ICHIMURA, H.; TODD, P. Matching as an econometric evaluation estimator. *The Review of Economic Studies*, [S.l.], v. 65, n. 2, p. 261-294, Apr. 1998.
- HIRANO, K.; IMBENS, G. W. Estimation of causal effects using propensity score weighting: an application to data on right heart catheterization. *Health Services and Outcomes Research Methodology*, [S.l.], v. 2, n. 3-4, p. 259-278, Dec. 2001.
- IMBENS, G. W.; RUBIN, D. B. *Causal inference in statistics, social, and biomedical sciences*. New York: Cambridge University Press, 2015.
- LAZZARINI, S. G. *et al.* What do state-owned development banks do? Evidence from BNDES, 2002–2009. *World Development*, v. 66, p. 237-253, 2015.
- _____. *Guia de avaliação de impacto socioambiental para utilização em negócios e investimentos de impacto: guia geral com foco em verificação de adicionalidade*. [S.l.]: Insper Metricis, 2018.
- LECHNER, M. A note on the common support problem in applied evaluation studies. *Univ. of St. Gallen Economics*, 2001 (Discussion paper).
- MACHADO, L. *et al.* Effects of BNDES PSI on investment of Brazilian industrial companies: impact estimates based on a matching approach. In: XLII ENCONTRO NACIONAL DE ECONOMIA, 2014, Natal. *Anais...*, [S.l.], Anpec, 2016.

_____.; PARREIRAS, M. A.; PEÇANHA, V. R. Avaliação de impacto do uso do Cartão BNDES sobre o emprego nas empresas de menor porte. *Revista do BNDES*, Rio de Janeiro, n. 36, p. 5-42, dez. 2011.

_____.; ROITMAN, F. B. Os efeitos do BNDES PSI sobre o investimento corrente e futuro das firmas industriais. *Revista do BNDES*, Rio de Janeiro, n. 44, p. [89]-122, dez. 2015.

PEARL, J. *Causality: models, reasoning, and inference* [S.l.]: Cambridge University Press, 2009.

_____.; GLYMOUR, M.; JEWELL, N. P. *Causal inference in statistics: a primer*. [S.l.]: John Wiley & Sons, 2016.

Apêndice 1: Detalhamento da base de dados utilizada no MARVIm

As fontes originais

Os dados da Rais são oriundos das informações coletadas pelo Ministério do Trabalho (MTE). Estão entre as principais variáveis presentes neste estudo: setor de atividade da firma, ano de sua abertura, número de trabalhadores, qualificação (grau de instrução) da mão de obra, gastos com salários e rotatividade dessa mão de obra.³²

A base da Secex foi cedida pelo Ministério da Indústria, Comércio Exterior e Serviços (Mdic) diretamente ao BNDES. Ela contém informações sobre exportação de todas as firmas brasileiras entre 1997 e 2012. A partir de 2012, contudo, o Mdic passou a fornecer ao BNDES apenas as informações das empresas efetivamente financiadas pelas linhas de financiamento à exportação – inviabilizando o uso dessa informação para fins de avaliação de impacto.

A base da Serasa abrange dados contábeis levantados privadamente pela Serasa Experian e adquiridos pelo BNDES.³³ Essa mesma base de dados já foi utilizada recentemente para investigar se firmas brasileiras sofrem com problemas de restrição de crédito – ver Ambrozio *et al.* (2016). Contudo, seu uso traz alguns problemas: ela não tem um desenho amostral bem definido, e os dados não decorrem de balanços necessariamente auditados. Logo, ela está mais sujeita a erros de informação do que as bases que lidam apenas com empresas de capital aberto. Além disso, informações contábeis mais detalhadas, obrigatórias para empresas que auditam seus resultados, podem não ser encontradas para todas as empresas da base da Serasa – é preciso, portanto, lidar com *missing data*.

Apesar disso, ela traz um conjunto amplo de empresas e, entre as bases disponíveis, é a única alternativa à sala de sigilo do IBGE com representatividade para empresas de menor porte. Considerando o objetivo de traçar um panorama a respeito do impacto do BNDES sobre o conjunto de empresas brasileiras, as vantagens de utilizar a Serasa são mais relevantes do que os problemas a ela relacionados. Além disso, os problemas informacionais podem ser mitigados por meio de um tratamento adequado dos dados, conforme ficará claro adiante.

³² Os registros identificados da Rais são dados sigilosos e seu acesso foi viabilizado por meio de um acordo de cooperação com o MTE. Essa base é entregue de forma consolidada uma vez por ano ao BNDES. Inicialmente, são duas bases distintas: Rais Estabelecimento – com dados básicos sobre os estabelecimentos empregadores de mão de obra – e Rais Trabalhadores – com dados sobre os trabalhadores vinculados a esse conjunto de estabelecimentos. Esses dois arquivos são combinados em uma única base, denominada Rais Empresa, que consolida informações sobre os estabelecimentos (Radical CNPJ) e sobre os empregados a ele vinculados.

³³ A Área de Crédito do BNDES comprou os registros referentes ao período de 2009 a 2013 para fins de análise de risco de crédito. Durante a aquisição, os registros eram disponibilizados para uma janela de quatro anos. A base de dados conta, portanto, com observações para o período de 2005 até 2012. Em razão de o número de observações antes de 2007 e depois de 2011 cair substancialmente, este trabalho utilizou apenas informações para o período entre 2007 e 2011.

A Tabela A.1 expõe o número de observações disponíveis em cada base de dados para o período de 2005 até 2012 – anos para os quais a Serasa apresenta alguma informação. Vale destacar que a base da Secex é, por definição, censitária – as empresas que não aparecem como exportadoras tiveram montante exportado igual a zero no período disponível. A Rais é de preenchimento obrigatório; por isso, ainda que possa não ser informada por algumas empresas, ela tem uma cobertura bastante ampla. Fica claro que a Serasa é a principal restrição ativa, tanto em relação ao período quanto do ponto de vista de número de empresas. Apesar disso, quando comparada com Economática ou com Valor 5000, ela ainda tem uma cobertura bem superior.³⁴

Tabela A.1. Número de observações em cada fonte de dados (período selecionado)

Ano	Rais Empresa	Secex	Serasa
2005		21.252	178
2006	5.801.402	20.591	6.990
2007	5.954.049	20.888	22.447
2008	6.159.380	20.407	27.305
2009	6.403.610	19.822	25.478
2010	6.554.448	20.038	24.817
2011	6.777.123	20.113	20.168
2012	6.791.307	18.508	1.060

Fontes: MTE, Mdic e Serasa Experian.

A consolidação da base final

Definidas as fontes, o passo seguinte foi consolidar os diferentes dados em uma única base, capaz de suportar um amplo conjunto de avaliações de impacto. Nesse sentido, uma série de decisões precisou ser tomada. Em primeiro lugar, diante da importância dos dados contábeis para a mensuração adequada da *performance* econômica das empresas, optou-se por restringir a base final a empresas que estivessem presentes na Serasa. Na prática, essa decisão limitou a análise a um teto de aproximadamente vinte mil a 25 mil empresas por ano e ao período compreendido entre 2007 e 2011 – ver Tabela A.1.

A base da Serasa contém um extenso conjunto de contas e indicadores de balanços. Optou-se por fazer uma seleção prévia das informações que seriam utilizadas na base de dados final. Dois procedimentos foram adotados: o primeiro consistiu em uma seleção *ad hoc* de variáveis de interesse, e o segundo, em uma análise da taxa de preenchimento dessas variáveis, na qual foram descartados os campos não informados, na média, por mais de 10% da amostra. O Gráfico A.1 mostra como esse segundo critério impactou a construção da base final.

³⁴ A Economática engloba informações para cerca de setecentas empresas por ano, enquanto o Valor 5000 traz dados das cinco mil maiores empresas brasileiras.

Essa segunda etapa teve dois objetivos. Do ponto de vista técnico, buscava-se lidar indiretamente com problemas informacionais. As contas mais comuns do balanço (com maior taxa de preenchimento) têm maior probabilidade de serem preenchidas de forma correta por todas as empresas, enquanto informações mais complexas tendem a ser preenchidas corretamente apenas por empresas mais estruturadas – aquelas que fazem auditoria do balanço, por exemplo. Do ponto de vista prático, pretendia-se minimizar o número de *missings* na base de dados.

Uma base com informações faltantes prejudica essa comparação, porque modelos com diferentes conjuntos de variáveis seriam também modelos com número de observações distintos. Como o MARVIm desconsidera observações com informação incompleta para fins de seleção de modelo, incluir variáveis com baixa taxa de preenchimento poderia resultar em estimações com poucas observações. Logo, esse recorte era importante também para manter o maior número possível de empresas na base de dados de análise.

Depois de restringir a base de dados às variáveis que tinham alta taxa de preenchimento, também se excluíram variáveis que fossem formadas por outras identidades contábeis já selecionadas. Por exemplo, ativo total é formado pela soma do ativo circulante com o ativo não circulante. Manter as três variáveis na base de dados implicaria o fato de que pelo menos uma delas seria redundante e, na melhor das hipóteses, aumentaria o custo computacional para seleção de variáveis explicativas. O Apêndice 2 traz as variáveis que foram descartadas por esse motivo.

Além dos filtros aplicados sobre a base da Serasa, a base da Rais Empresa limitou-se àquelas observações que não haviam declarado Rais negativa – o que significa declarar que a empresa não contou com nenhum empregado vinculado durante todo o ano de referência.³⁵ Implementados esses filtros, as três bases de dados foram unidas para dar origem à base utilizada pelo MARVIm.

O conjunto final de dados foi consolidado com 31 variáveis e 113.714 observações. Os gráficos A.2 e A.3 mostram a distribuição dessas empresas segundo a divisão Cnae e o ano de observação, respectivamente. O Gráfico A.2 mostra que os dados envolvem grande diversidade de divisões Cnae, mas apenas 28 divisões têm mais de mil observações – e essas mesmas divisões respondem por cerca de 85% do total de observações. Já o Gráfico A.3 reforça a ideia de que a base de dados consolidada deve ser trabalhada prioritariamente no período compreendido entre 2007 e 2011, intervalo no qual existem aproximadamente vinte mil empresas por ano, mesmo depois da aplicação dos filtros.

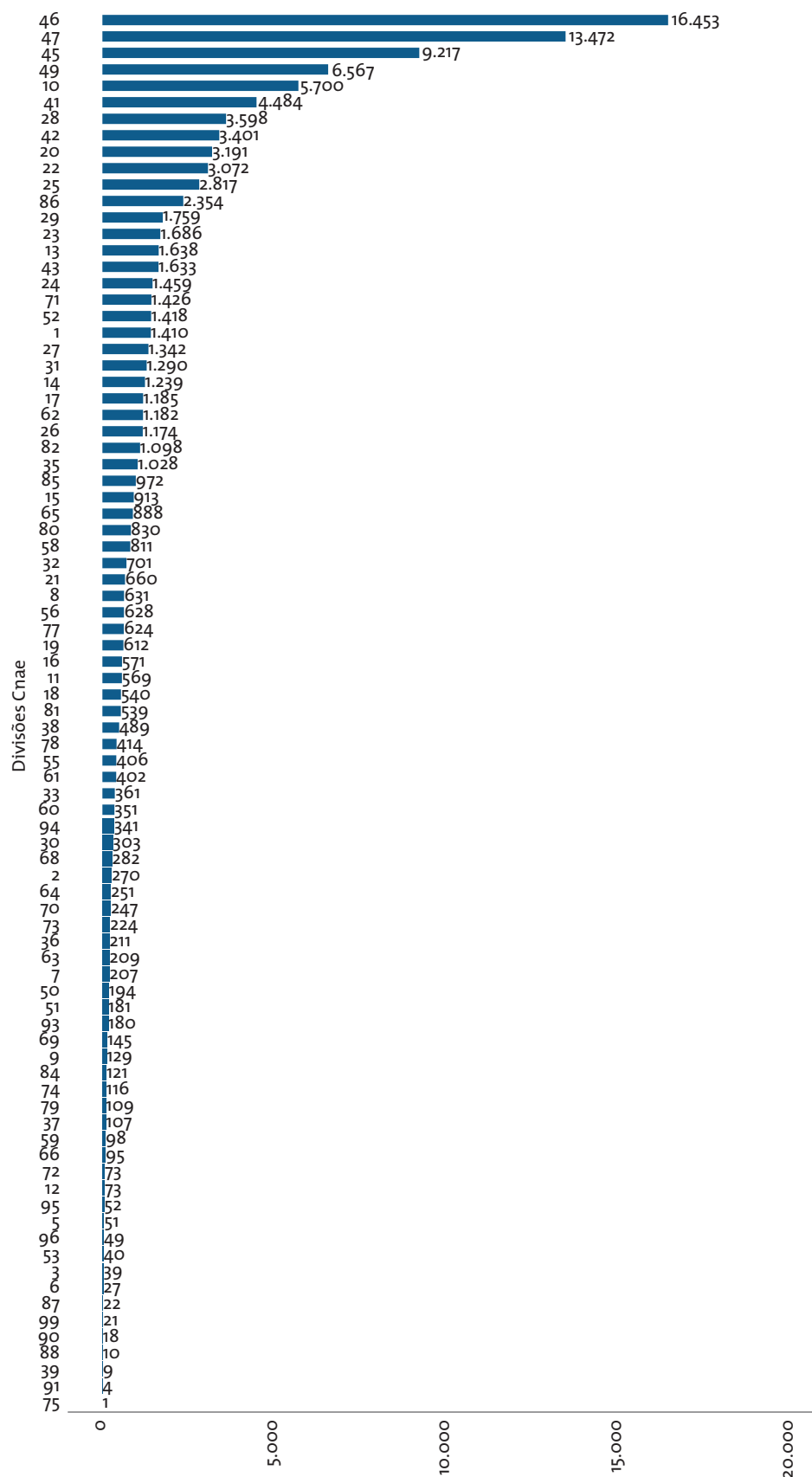
³⁵ Boa parte delas diz respeito a microempreendedores individuais. O fato é que tais firmas têm um incentivo ainda menor em apresentar dados formais de balanço, e a chance de tais dados serem capturados pela Serasa era muito baixa.

Gráfico A.1. Taxa de preenchimento
por variáveis da Serasa (total de informantes no ano – %)



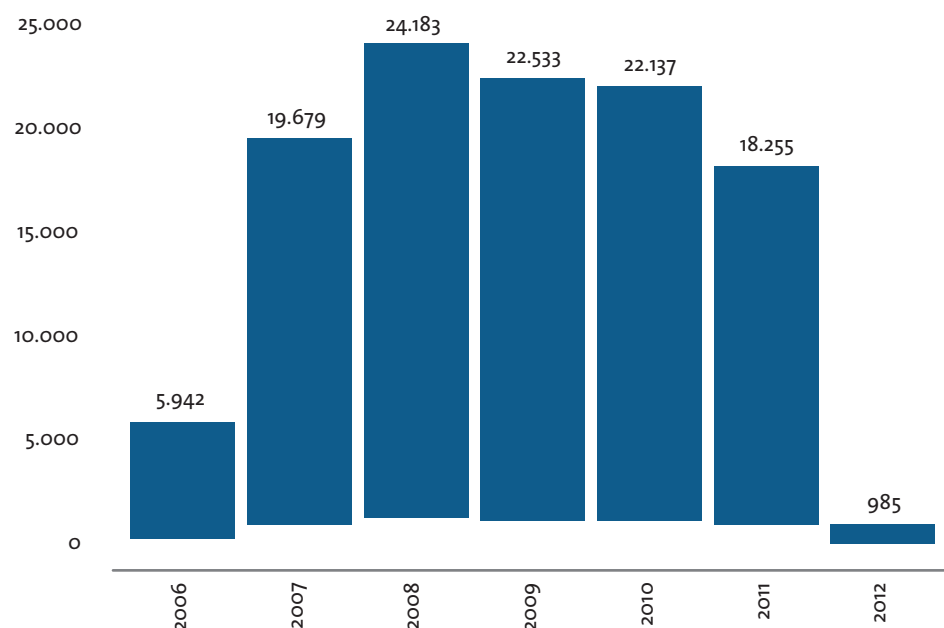
Fonte: Serasa Experian.

Gráfico A.2. Número de observações (por código de atividade econômica)



Fonte: Elaboração própria, com base no MARVIm.

Gráfico A.3. Número de empresas (por ano)



Fonte: Elaboração própria, com base no MARVIm.

A Tabela A.2 contém estatísticas descritivas básicas para a base consolidada. É possível notar que as variáveis de balanço (com origem na Serasa) têm, normalmente, desvio-padrão bastante elevado. Alguns valores máximos e mínimos atraem a atenção. Se alguns casos podem ser interpretados apenas como um sinal da diversidade das empresas – fato que motivou o uso da base da Serasa –, outros levantam dúvidas acerca da qualidade da informação – por exemplo, margem Ebitda. Ou seja, novamente a preocupação a respeito de problemas informacionais aparece.

Tabela A.2. Estatísticas descritivas para a base consolidada

Variável	Média	Mediana	Desvio-padrão	Máximo	Mínimo
Ativo financeiro	17.343	1.511	326.864	59.222.000	(234)
Ativo operacional	33.781	7.512	358.981	49.296.000	0
Ativo total	156.715	17.821	3.418.570	564.951.000	0
Capital de giro	11.973	2.490	300.999	41.106.000	(60.441.000)
Custos	86.417	20.547	962.465	167.882.000	0
Faturamento líquido	116.243	28.677	1.369.786	217.346.000	1
Faturamento bruto	136.875	32.823	1.721.009	276.935.000	1
Ebitda	39.885	3.347	1.077.028	152.988.000	(8.316.000)
Ativo imobilizado	48.012	3.151	1.415.637	279.822.000	(71.888)
Passivo financeiro	18.868	2.030	220.585	32.211.000	0
Passivo operacional	20.283	3.482	403.401	79.487.000	0
Passivo total	156.715	17.821	3.418.570	564.951.000	0
Patrimônio líquido	73.823	5.984	2.085.513	343.439.000	(1.856.008)
Resultado de atividade	10.271	1.472	288.469	43.044.000	(3.441.000)
Resultado financeiro	(2.272)	(293)	79.995	10.033.000	(10.673.000)

(Continua)

(Continuação)

Variável ¹	Média	Mediana	Desvio-padrão	Máximo	Mínimo
Resultado bruto	29.827	6.880	468.546	64.367.000	(1.233.328)
Resultado líquido	4.042	556	195.651	28.751.000	(10.352.000)
Exportações	5.060	0	109.935	19.299.192	0
Número de empregados (dez.)	410	126	1.737	117.959	1
Número de empregados (média)	399	122	1.678	114.029	0
Massa salarial (dez.)	749	181	6.116	903.376	0
Anos de estudo (média)	10,98	11,11	1,90	16,00	1,30
Funcionários graduados (%)²	0,06	0,02	0,12	1,00	0,00
Idade da firma	21	18	13	120	0
Margem Ebitda	0,08	0,12	11,65	356,07	(3.408,00)
Resultado não operacional³	3.235	1	102.990	16.574.000	(7.917.000)
Produtividade do trabalho⁴	231	60	5.426	1.413.660	(43.866)

Fonte: Elaboração própria, com base no MARVIm.

¹ Todas as variáveis monetárias são apresentadas em R\$ mil, exceto o valor exportado, que está em US\$ mil.

² Representa o percentual de mão de obra na firma que faz parte do grupo ocupacional 2 – profissionais das ciências e das artes.

³ Construída como a diferença entre o resultado operacional da firma e a soma do resultado da atividade com o resultado financeiro.

⁴ Construída como a razão entre o resultado bruto – ou seja, o faturamento líquido menos o custo de mercadorias vendidas ou serviços prestados – e o número médio de empregados no ano.

**Tabela A.3. Estatísticas descritivas
para a base consolidada (após detecção de outliers)**

Variável ¹	Média	Mediana	Desvio-padrão	Máximo	Mínimo
Ativo financeiro	4.132,54	1.041	15.033,01	660.358	(234)
Ativo operacional	13.254,84	5.721	29.812,45	1.053.156	0
Ativo total	33.829,09	13.026	89.346,69	2.700.855	19
Capital de giro	4.963,15	1.978	25.097,78	780.376	(1.474.765)
Custos	33.392,04	15.874,50	71.608,52	2.123.353	0
Faturamento líquido	44.394,75	22.685,50	87.242,48	2.783.368	2
Faturamento bruto	51.074,72	25.984	100.045,10	3.209.614	2
Ebitda	9.078,73	3.113,50	37.369,54	1.480.128	(1.178.784)
Ativo imobilizado	10.168,62	2.151	39.744,97	2.136.797	(19.298)
Passivo financeiro	5.561,79	1.391,50	21.541,17	1.309.431	0
Passivo operacional	6.862,40	2.553	20.956,86	1.817.245	0
Passivo total	33.829,09	13.026	89.346,69	2.700.855	19
Patrimônio líquido	12.716,16	4.277,50	44.088,84	1.830.141	(1.400.732)
Resultado da atividade	3.463,22	1.506	12.289,38	420.126	(528.224)
Resultado financeiro	(839,44)	(228)	5.876,04	241.554	(370.983)
Resultado bruto	11.002,72	5.738	22.950,50	873.496	(273.775)
Resultado líquido	1.063,02	479	9.826,11	262.617	(658.414)
Exportações	1.395,44	0	30.526,95	5.225.717,00	0
Número de empregados (dez.)	248,83	97	822,7	73.247	1
Número de empregados (média)	239,12	93,75	769,82	72.012,33	0,08
Massa salarial (dez.)	340,17	131,58	929,33	53.775,89	0

(Continua)

(Continuação)

Variável ¹	Média	Mediana	Desvio-padrão	Máximo	Mínimo
Anos de estudo (média)	10,89	11	1,91	16	1,3
Funcionários graduados (%)	0,05	0,02	0,12	1	0
Idade da firma	18,33	15	12,19	79	0
Margem Ebitda	0,15	0,13	4,89	225	(787,41)
Resultado não operacional³	0,27	0	1,07	8	(4)
Produtividade do trabalho⁴	225,66	62,29	1.844,26	177.204,00	(13.565,71)

Fonte: Elaboração própria, com base no MARVIm.

¹ Todas as variáveis monetárias são apresentadas em R\$ mil, exceto o valor exportado, que está em US\$ mil.

² Representa o percentual de mão de obra na firma que faz parte do grupo ocupacional 2 – profissionais das ciências e das artes.

³ Construída como a diferença entre o resultado operacional da firma e a soma do resultado da atividade com o resultado financeiro.

⁴ Construída como a razão entre o resultado bruto – ou seja, o faturamento líquido menos o custo de mercadorias vendidas ou serviços prestados – e o número médio de empregados no ano.

Como fica claro na terceira seção, o próprio método de seleção de unidades contrafactuais tenderia a descartar observações muito extremas – ou a reduzir bastante o peso delas. Ou seja, a própria técnica de avaliação de impacto proposta mitiga boa parte desse problema informacional. A despeito disso, e para não depositar confiança excessiva apenas no MARVIm, foi implementada uma análise para detecção multidimensional de *outliers*.³⁶ Esse processo categorizou como *outliers* mais de 54 mil observações – ou aproximadamente 48% das 113.714 originais. Foi, então, gerada uma nova base de dados com as 58.872 observações restantes.

Essa base adicional foi utilizada para testar a robustez de todos os resultados reportados na quarta seção, aumentando a confiança de que as conclusões expostas não decorrem de problemas informacionais da base da Serasa. A Tabela A.3 exibe as estatísticas descritivas dessa base depois da detecção de *outliers* e, como esperado, o filtro reduziu bastante o desvio-padrão das variáveis, de tal forma que os máximos e mínimos passaram a apresentar valores menos extremos.

³⁶ Para mais detalhes a respeito do método implementado, ver Filzmoser, Garrett e Reimann (2005).

Apêndice 2: Identidades contábeis na base da Serasa

Existem seis diferentes “detalhamentos” de ATIVO ou PASSIVO na base da Serasa: TOTAL, CIRCULANTE, NÃO CIRCULANTE, FINANCEIRO, OPERACIONAL e PERMANENTE. Contudo, não havia grau de liberdade suficiente para utilizar todos eles, pois, na base da Serasa, valem as seguintes igualdades:

1. $TOTAL = CIRCULANTE + NÃO\ CIRCULANTE$
2. $NÃO\ CIRCULANTE = PERMANENTE$
3. $CIRCULANTE = FINANCEIRO + OPERACIONAL$

Diante disso, optou-se por preservar os seguintes detalhamentos: TOTAL, FINANCEIRO E OPERACIONAL. Por meio desses três, é possível encontrar todos os demais pelas equações 1, 2, e 3. Além disso, a variável SALDO DE TESOURARIA equivale à diferença entre ATIVO FINANCEIRO e PASSIVO FINANCEIRO. Em razão disso, ela também foi excluída da base final de análise.

Além disso, existiam duas variáveis diferentes que pareciam fazer menção à mesma conta: FATURAMENTO LÍQUIDO e FATURAMENTO BRUTO. Apesar da alta correlação entre elas (aproximadamente 0,99), notou-se que o FATURAMENTO LÍQUIDO era superior ao FATURAMENTO BRUTO com excessiva frequência. Já o FATURAMENTO LÍQUIDO era sempre inferior ao FATURAMENTO BRUTO. Com isso, julgou-se que a variável correta era FATURAMENTO LÍQUIDO.

Coordenação Editorial

Gerência de Editoração e Memória
do BNDES

Projeto Gráfico

Fernanda Costa e Silva

Produção Editorial

Expressão Editorial

Editoração Eletrônica

Expressão Editorial

Editado pelo
Departamento de Comunicação
Outubro de 2018



www.bndes.gov.br