





ESTADO DO RIO GRANDE DO SUL

Secretaria de Planejamento, Governança e Gestão  
Departamento de Planejamento Governamental

# ESTUDOS DEPLAN

*Edição*

*Nº 10/2018*

Avaliação de Impacto de Políticas Públicas  
Conceitos, Metodologias e Experiências

*Junho/2018*



## **GOVERNO DO ESTADO DO RIO GRANDE DO SUL**

**Governador:** José Ivo Sartori

**Vice-Governador:** José Paulo Dornelles Cairoli

## **SECRETARIA DE PLANEJAMENTO, GOVERNANÇA E GESTÃO**

**Secretário:** Josué de Souza Barbosa

**Secretário Adjunto:** Melissa Guagnini Hoffmann Custódio

## **DEPARTAMENTO DE PLANEJAMENTO GOVERNAMENTAL (DEPLAN)**

**Diretor:** Antonio Paulo Cargnin

**Diretora Adjunta:** Carla Giane Soares da Cunha

## **EQUIPE EDITORIAL**

Antonio Paulo Cargnin

Juliana Feliciati Hoffmann

## **FICHA TÉCNICA:**

Juliana Feliciati Hoffmann e Ana Júlia Possamai (Coordenação e elaboração)  
César Stallbaum Conceição, Fernanda Rodrigues Vargas, Gisele da Silva  
Ferreira, Juliana Feliciati Hoffmann, Sílvia Letícia Lorenzetti, Rayssa Miczewski  
de Araújo, Vinícius Dias Fantinel (elaboração)

**REVISÃO E TRADUÇÃO:** Marlise Margô Henrich

**CAPA:** Laurie Fofonka Cunha

Estudos DEPLAN / Departamento de Planejamento Governamental - RS.  
Porto Alegre : Secretaria de Planejamento, Governança e Gestão, 2015-

v. : il.

Semestral.

Título especial a cada edição, de acordo com o assunto predominante.  
Publicado pela Secretaria de Planejamento, Planejamento, Governança  
e Gestão, 2015-

ISSN 2447-4576

1. Desenvolvimento regional – Periódico – Rio Grande do Sul. I. Rio  
Grande do Sul. Secretaria de Planejamento e Gestão. Departamento de  
Planejamento Governamental.

CDU 332.1(816.5)(05)

Bibliotecário responsável: João Vítor Ditter Wallauer – CRB 10/2016



As opiniões nesta publicação são de exclusiva responsabilidade dos autores, não exprimindo, necessariamente, o posicionamento da Secretaria de Planejamento, Governação e Gestão

## **Apresentação**

É com satisfação que a Secretaria de Planejamento, Governança e Gestão (SPGG) lança a edição de número dez da Revista Estudos Deplan. A publicação reúne um conjunto de textos elaborados com base no conteúdo apresentado ao longo do Workshop de Avaliação de Impacto de Políticas Públicas, realizado pela SPGG em parceria com o Banco Mundial no final do ano de 2017.

A Revista Estudos Deplan tem como foco a divulgação de metodologias e análises de caráter técnico-científico da área de planejamento e avaliação de políticas públicas. Com isso, temos como objetivo final contribuir para o aprimoramento das políticas públicas oferecidas pelo Estado do Rio Grande do Sul.

Iniciativas dessa ordem são fundamentais para a organização, a sistematização e a disseminação da produção técnica e para a evolução das metodologias no setor público. Permitem que se registrem o avanço e as dificuldades inerentes aos processos e se evolua do ponto de vista das metodologias, sempre com ênfase na prestação de melhores serviços à sociedade.

Josué de Souza Barbosa

Secretário de Planejamento, Governança e Gestão



## **Editorial**

Em um cenário de crescente complexidade socioeconômica e ambiental, a adoção de políticas públicas eficazes e eficientes impõe-se como agenda primordial. Nesse sentido, a avaliação de impacto opera como importante instrumento. Para além de identificar e mensurar os efeitos exclusivos de uma intervenção sobre o resultado que dela se espera, a realização da avaliação de impacto permite aperfeiçoar o desenho da intervenção e, assim, qualificar seu planejamento e sua implementação.

Tendo isso em vista, e alinhado às estratégias de modernização da gestão pública, o Departamento de Planejamento Governamental da Secretaria de Planejamento, Governança e Gestão organizou o *Workshop Avaliação de Impacto de Políticas Públicas*, com o apoio do Banco Mundial. Realizado nos dias 29 e 30 de novembro de 2017, no Auditório da SPGG, o evento teve como objetivo apresentar e difundir a cultura de avaliação entre os servidores e gestores gaúchos, com a finalidade última de aperfeiçoar as políticas públicas estaduais, por meio da adoção de uma gestão orientada por resultados.

A programação do *Workshop* contou com a presença de palestrantes especialistas na área de avaliação de impacto do Banco Mundial, de órgãos governamentais, bem como da academia e do terceiro setor. Além de buscar familiarizar os participantes em relação a conceitos básicos e metodologias de avaliação de impacto, a programação buscou contemplar o intercâmbio de diferentes experiências de avaliação em curso ou finalizadas, a fim de estimular o aprendizado coletivo.

E é justamente para contribuir nesses sentidos que apresentamos a nova edição da Revista Estudos Deplan, que reúne textos elaborados a partir da transcrição e das apresentações das palestras proferidas no *Workshop Avaliação de Impacto de Políticas Públicas*, validados pelos autores. A edição conta com quatro seções, sendo a primeira uma sensibilização sobre a importância da agenda de avaliação. Em seguida, são abordados os principais métodos experimentais e não experimentais de avaliação de impacto. Na terceira seção, por sua vez, são apresentadas diferentes avaliações de programas nas áreas de desenvolvimento e apoio ao setor privado, educação e governança. Por fim, experiências de avaliação na área pública e no terceiro setor são discutidas desde o ponto de vista das organizações.

Desejamos uma ótima leitura.

A Comissão Editorial.

## Sumário

---

|  |     |
|--|-----|
| SENSIBILIZAÇÃO E CAPACITAÇÃO SOBRE AVALIAÇÃO DE IMPACTO .....  | 8   |
| INTEGRANDO AVALIAÇÃO DE PROCESSOS E AVALIAÇÃO DE IMPACTO: UMA ABORDAGEM.....   | 13  |
| AVALIAÇÃO DE IMPACTO DE POLÍTICAS PÚBLICAS: MÉTODOS EXPERIMENTAIS.....   | 28  |
| AVALIAÇÃO DE IMPACTO DE POLÍTICAS PÚBLICAS: MÉTODOS NÃO-EXPERIMENTAIS .....  | 37  |
| AVALIAÇÃO DE IMPACTO DE PROGRAMAS DE DESENVOLVIMENTO PARA O SETOR PRIVADO: A AGENDA DO BANCO MUNDIAL .....             | 50  |
| AVALIAÇÃO DE IMPACTO DO PROJETO EXTENSÃO PRODUTIVA E INOVAÇÃO – PEPI.....  | 63  |
| O IMPERATIVO DA AVALIAÇÃO DAS POLÍTICAS DE APOIO AO SETOR PRIVADO NO BRASIL: MUITO GASTO, POUCO IMPACTO.....           | 79  |
| AVALIAÇÃO DE IMPACTO COMO INSTRUMENTO PARA O APRIMORAMENTO DAS POLÍTICAS EDUCACIONAIS .....                            | 92  |
| COMUNICAÇÃO PARA O ENGAJAMENTO FAMILIAR NA VIDA ESCOLAR DOS FILHOS: UMA AVALIAÇÃO DE IMPACTO DO PROJETO EDUQMAIS ..... | 107 |
| AVALIAÇÃO DE IMPACTO DE POLÍTICAS PÚBLICAS NA ÁREA DE GOVERNANÇA.....  | 120 |
| POLÍTICAS DE COMPRAS GOVERNAMENTAIS: UMA PROPOSTA DE AVALIAÇÃO .....   | 130 |
| EXPERIÊNCIA DO CENTRO DE LIDERANÇA PÚBLICA: UMA AVALIAÇÃO.....   | 139 |
| INICIATIVAS DE AVALIAÇÃO DE IMPACTO REALIZADAS POR ÓRGÃOS PÚBLICOS ESTADUAIS: RS, RN e CE .....                        | 146 |
| EXPERIÊNCIAS DE AVALIAÇÃO DE IMPACTO NO TERCEIRO SETOR.....  | 157 |

## SENSIBILIZAÇÃO E CAPACITAÇÃO SOBRE AVALIAÇÃO DE IMPACTO\*

Caio Piza<sup>1</sup>

**Resumo:** O texto destaca a importância da avaliação de impacto para aperfeiçoar o modo como são feitas as políticas públicas. Argumenta-se que a avaliação não consiste em um instrumento de auditoria da ação ou de agentes públicos. O objetivo de uma avaliação de impacto é identificar o efeito causal de um programa, isolando a influência de outros fatores que também podem afetar as variáveis de resultado. Dessa forma, o formulador de políticas pode utilizar os métodos de avaliação de impacto em seu favor, mediante o uso de instrumentos que podem contribuir para o aperfeiçoamento de uma política nova ou já implementada.

**Palavras-chave:** Políticas públicas. Avaliação de impacto. Efeito causal. Gestão por resultados.

**Abstract:** The text highlights the importance of impact evaluation to improve the way public policies are made. It is argued that impact evaluation does not consist of an audit instrument of the action or of public agents. Impact evaluation aims at identifying the causal effect of a program, isolating the influence of other factors that may also affect the outcome variables. In this way, the policy maker may use the impact assessment methods in his favor by using instruments that may contribute to the improvement of a new or already implemented public policy.

**Keywords:** Public policies. Impact evaluation. Causal effect. Managing for results.

---

\* Texto elaborado a partir da transcrição da fala e da apresentação em slides de palestras proferida no Workshop Avaliação de Impacto de Políticas Públicas, realizado nos dias 29 e 30 de novembro de 2017, no Auditório da Secretaria de Planejamento, Governança e Gestão (SPGG-RS), com o apoio do Banco Mundial.

<sup>1</sup> Economista do Banco Internacional para Reconstrução e Desenvolvimento (BIRD).

## INTRODUÇÃO

O objetivo deste texto é discutir o tema de avaliação de impacto, utilizando ideias simples e intuitivas que podem mudar a forma como fazemos políticas públicas. A apresentação será mais motivacional. Abordarei como usarmos avaliação de impacto na unidade onde atuo no Banco Mundial com o intuito de fazer com que todos tenham o mesmo entendimento básico do tema. Será apontado durante a apresentação como o formulador de políticas pode utilizar os métodos de avaliação de impacto a seu favor, ou seja, não como um instrumento de ‘auditoria’, mas como um instrumento que pode auxiliar no aperfeiçoamento (ou redesenho) da própria política implementada.

## ANTES DE COMEÇAR, UMA REFLEXÃO NECESSÁRIA

Para compreender o que se entende por avaliação de impacto, propõe-se inicialmente um exercício de reflexão. Digamos que, na qualidade de secretário ou subsecretário, estejamos próximos de iniciar um programa para aumentar a produtividade do setor privado por meio de apoio a pequenos negócios. Para tanto, é necessário, primeiro, desenhar o programa.

Gostaríamos, evidentemente, que o programa tivesse efeito positivo sobre as empresas apoiadas. Digamos que um dos seus objetivos seja aumentar a taxa de sobrevivência dos pequenos negócios, visto que vários pequenos negócios morrem cedo. O programa visa, então, garantir que os pequenos empreendimentos permaneçam mais tempo no mercado e, com isso, tenham mais tempo para colher os frutos da intervenção em termos de aumento de produtividade e crescimento. O SEBRAE, por exemplo, quando concebe um programa de apoio a pequenos negócios, está pensando basicamente nesses pontos.

Existe uma questão de economia política extremamente importante, pois uma vez que o programa é iniciado, torna-se muito difícil descontinuá-lo, mesmo quando se prova que ele não funciona. Portanto, iniciar um programa novo envolve uma decisão estratégica. Não se sabe se esse programa, de antemão, funcionará ou não, mas se sabe que, caso seja implementado, não será muito fácil descontinuá-lo ali adiante.

Diante dessa situação, têm-se quatro opções. A primeira opção é lançar o programa de imediato em escala ampla e torcer para que dê certo. A segunda opção é lançar o programa e checar de tempo em tempo se os indicadores de resultados estão caminhando na direção correta. Em outras palavras, realizar análise de monitoramento. Uma terceira opção seria lançar o programa em escala ampla e, depois de dois anos, fazer uma avaliação rigorosa de impacto, para saber se funcionou ou não. Por fim, uma quarta opção é lançar o programa num formato piloto. Testar esse piloto de forma rigorosa para entender como o programa funciona na prática. E, se funciona, de que forma funciona, quem se beneficia mais, se os que mais se beneficiam são de fato a população-alvo. Se for o caso, verificar se é necessário redesenhar a intervenção e realocar recursos para maximizar seu potencial de

gerar o resultado esperado. Ou seja, testar o programa, antes de lançá-lo em escala ampliada, para descobrir quais são os desafios de implementação que se encontrará na prática. Caso se desenhe um programa, se teste e descubra que não produziu o efeito que se gostaria, pelo menos a experiência serviu para aprendizagem. Pode-se tentar redesenhar a intervenção para o tipo de objetivo que se quer alcançar.

A opção quatro sintetiza os objetivos da avaliação de impacto: antes de lançar uma intervenção, testar um piloto e aprender o máximo possível com isso.

## O SIGNIFICADO DA AVALIAÇÃO DE IMPACTO

Avaliação de impacto não é análise de monitoramento, não é avaliação de resultado ou auditoria. A avaliação de impacto não se refere à avaliação de agentes públicos. A avaliação de impacto avalia um programa. Mesmo que tenha sido pensado por um formulador de políticas, ele não tem a obrigação de ter certeza de que o programa funcionará na prática.

É difícil saber qual programa de apoio ao pequeno negócio aumentará a taxa de sobrevivência das empresas e a produtividade. Há inúmeras intervenções possíveis, e muitas já sendo implementadas com esse propósito. Não se deve se sentir incompetente caso a avaliação aponte que o programa que você está liderando ou que ajudou a conceber não tem efeito. Na verdade, é natural que não tenha – surpreendente seria, de saída, o contrário. Testam-se muitas alternativas antes de se saber ao certo o que funciona. É assim na ciência, é assim na política pública.

O objetivo de uma avaliação de impacto é identificar o efeito exclusivo de uma intervenção sobre as variáveis de resultado. No exemplo acima, as variáveis de resultado são a taxa de sobrevivência das empresas e a produtividade. São duas variáveis de resultado. Por meio da avaliação de impacto, busca-se identificar o efeito causal da intervenção/programa, isolando a influência de outros fatores que também podem afetar as variáveis de resultado – como, por exemplo, o aquecimento da economia.

Suponha, a título de ilustração, que a economia brasileira esteja crescendo e que as empresas apresentem uma menor taxa de mortalidade. Uma análise de monitoramento mostraria que a política está dando certo. Mas a questão relevante é se a evolução do indicador se deve exclusivamente à implementação do programa ou não. Se a economia mudar de trajetória, o efeito permanecerá? Descobrir isso é objetivo de uma avaliação de impacto.

Portanto, análise de monitoramento e avaliação de impacto são instrumentos diferentes. Deve-se evitar que uma avaliação ocorra tarde e, idealmente, garantir que a avaliação seja realizada por agentes que não possuam nenhum compromisso ou comprometimento com o programa. O ideal seria que os governos contassem com departamentos de avaliação tal qual o que se constituiu na Secretaria de Planejamento do Governo do Estado do Rio Grande do Sul, que, a partir da concepção de um grupo de

pesquisa, tem buscado utilizar dados disponíveis para gerar esse conhecimento internamente, algumas vezes com o apoio do Banco Mundial na validação dos planos de avaliação. Parcerias equivalentes têm sido constituídas com o Serviço Brasileiro de Apoio às Micro e Pequenas Empresas (SEBRAE) do Paraná e do Rio de Janeiro.

Dessa forma, o trabalho de avaliação fica institucionalizado, o que facilita a internalização do conhecimento gerado. O formulador de política ou quem está envolvido no projeto deve se sentir representado na avaliação de impacto. As equipes de pesquisa e do governo unidos na identificação dos desafios, das etapas da intervenção que necessitam de revisão e/ou reparos, das fases passíveis de serem avaliadas; isso ajuda muitíssimo na institucionalização da cultura de avaliação.

Na prática, portanto, a avaliação de impacto é um método de testar ideias ou hipóteses. Para isso, existem métodos experimentais e *quase-experimentais* (ou não-experimentais). O método *quase-experimental* tenta se aproximar ao máximo do melhor método de avaliação que é o experimental.

## PORFÓLIO DO BANCO MUNDIAL NO BRASIL

A equipe do *Development Impact Evaluation* (DIME) trabalha com avaliação de impacto de programas em diferentes setores. No setor educacional, existem muitos estudos produzidos pelo escritório do Banco Mundial no Brasil e, alguns casos, com o apoio do DIME. No setor de agricultura, o Banco Mundial, com o apoio do DIME, está avaliando o impacto de programas nos estados no Ceará, São Paulo, Santa Catarina e Rio de Janeiro. Estamos dialogando também uma possível colaboração com o Ministério do Meio Ambiente. A agenda do setor privado é a que mais tem crescido, por meio de parcerias com SEBRAE, Banco do Nordeste, Banco Central, Ministério de Desenvolvimento da Indústria e Comércio, Associação de Educação Financeira e SERASA. Também tem crescido a agenda na área de governança, de transportes, de violência contra a mulher e do uso de entretenimento para mudança de comportamento.

## CONSIDERAÇÕES FINAIS

Neste texto, buscou-se apresentar a essência da ideia de avaliação de impacto. Avaliação de impacto não é monitoramento. A avaliação de impacto quer identificar efeito causal exclusivo de uma intervenção.

## REFERÊNCIAS

WORLD BANK. *Science for impact: better evidence for better decisions.* Washington, D.C., 2018: World Bank Group. Disponível em: <<http://documents.worldbank.org/curated/en/857631525294854313/Science-for-impact-better-evidence-for-better-decisions>. Acesso em 15 jun. 2018 >. Acesso em: 07 de maio de 2018.

## INTEGRANDO AVALIAÇÃO DE PROCESSOS E AVALIAÇÃO DE IMPACTO: UMA ABORDAGEM\*

Xavier Cirera<sup>1</sup>

**Resumo:** O texto apresenta uma discussão sobre a avaliação qualitativa e integrada de políticas públicas, como forma de avaliar possibilidades de melhoria dessas políticas. Para tanto, são discutidas a dificuldade e a complexidade da formulação e da implementação de políticas públicas, a partir de exemplos na área da inovação. Diante disso, busca-se ressaltar a necessidade de uma teoria robusta para a formulação de políticas de inovação, bem como apresentar uma proposta de abordagem de análise qualitativa de programas de Ciência, Tecnologia e Inovação (CT&I), baseada em estágios (*capabilities escalator*).

**Palavras-chave:** Políticas públicas. Política de inovação. Avaliação de políticas públicas. CT&I.

**Abstract:** The text presents a discussion about a qualitative and integrated method of policy evaluation, as a way of assessing possibilities of improvement of such policies. Difficulty and complexity of the formulation and implementation of public policies are discussed, taking as examples innovation policies. It aims at emphasizing the need for a robust theory for the formulation of innovation policies, as well as at presenting a proposal for a qualitative analysis method of Science, Technology and Innovation (ST&I) policies, based on a capabilities escalator approach.

**Keywords:** Public policies. Innovation policy. Public policy evaluation. ST&I.

---

\* Texto elaborado a partir da transcrição da fala e da apresentação em slides de palestras proferida no Workshop Avaliação de Impacto de Políticas Públicas, realizado nos dias 29 e 30 de novembro de 2017, no Auditório da Secretaria de Planejamento, Governança e Gestão (SPGG-RS), com o apoio do Banco Mundial.

<sup>1</sup> Economista do Banco Internacional para Reconstrução e Desenvolvimento (BIRD).

## INTRODUÇÃO

Fazer política pública é muito difícil e complexo. No contexto das políticas de inovação, em especial, essa complexidade envolve a tentativa de mudar os incentivos dos agentes econômicos de maneiras diversas, com muitos fatores influenciando ao mesmo tempo. Um ponto de partida é ter um bom embasamento teórico ou um marco, que oriente quais políticas são importantes, quais são os problemas e por quê.

Em face disso, o texto trata da avaliação qualitativa e integrada de políticas públicas, buscando avaliar possibilidades de melhorias. Para tanto, o Banco Mundial desenvolveu uma metodologia de avaliação conjunta e complementar à avaliação de impacto de programas específicos, a qual busca integrar avaliação de impacto e avaliação de processos.

Na primeira parte do texto, procura-se discutir a dificuldade e complexidade de formulação e implementação da política pública, a partir do exemplo das políticas de inovação. Em seguida, ressalta-se a necessidade de uma teoria para a formulação de uma política de inovação. Nesse sentido, discutem-se as chamadas escalas de capacidade nas empresas. Já na terceira parte, serão apresentadas metodologias de análise qualitativas dos programas de Ciência, Tecnologia e Inovação (CT&I). Por fim, algumas considerações finais.

## POLÍTICAS PÚBLICAS E COMPLEXIDADE: O DESAFIO DA POLÍTICA DE INOVAÇÃO E O RISCO DE FALHAS

O livro *The Innovation Paradox* (Cirera; Maloney, 2017), publicado recentemente pelo Banco Mundial, tenta descrever, a partir da teoria da inovação e da teoria neoschumpeteriana, que é alto o retorno dos investimentos em inovação nos países distantes da fronteira tecnológica, como, por exemplo, no Brasil. Diante disso, pergunta-se: se os retornos da inovação são altos, por que esses países não inovam? Por que não se adaptam às novas tecnologias? Trata-se de um paradoxo: os retornos são altos, mas empresas não fazem investimento em inovação. As explicações discutidas no livro são variadas. Destaca-se a ausência, nesses países, de elementos complementares e necessários para inovar, tais como tecnologias, facilidade de registro de empresas, monitoramento de trabalhadores, entre outros.

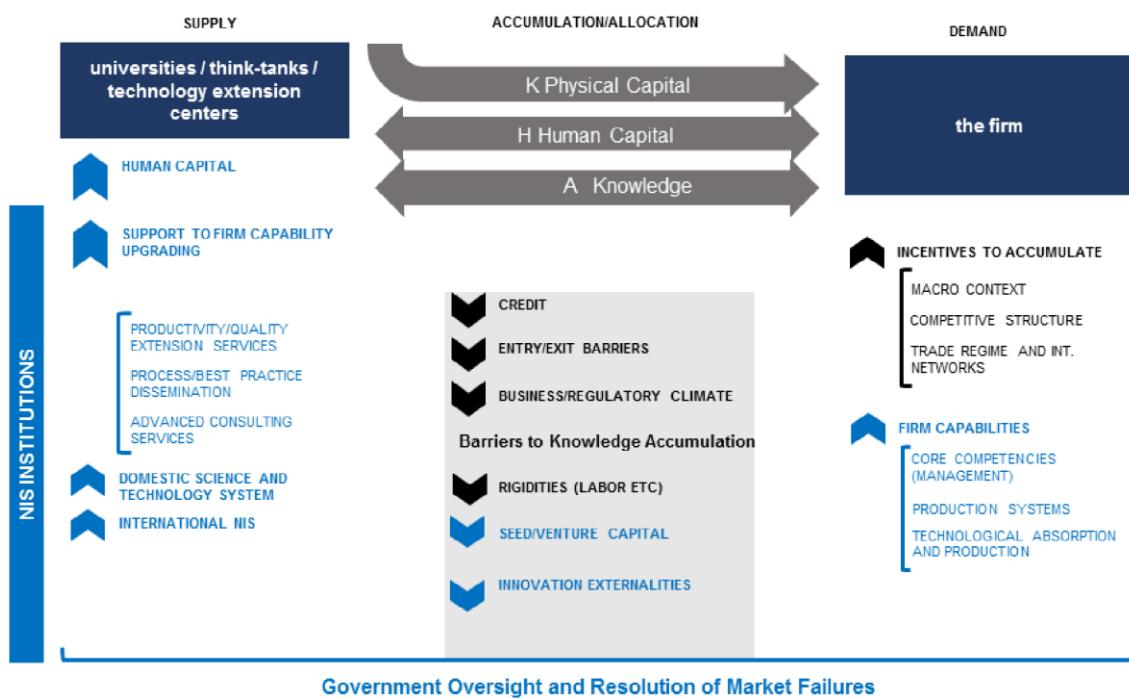
Um elemento importante é o chamado “dilema da política de inovação”, que expressa o fato de que, nos países que apresentam maiores problemas associados à falta de complementariedades e de fatores alternativos, também se tem muito pouca capacidade no governo para implementar políticas públicas. Nesse cenário de menos recursos, implementar políticas de inovação se torna algo ainda mais complexo.

A Figura 1 representa o conceito de Sistema Nacional de Inovação. Nela, é possível verificar a relação existente entre a demanda das empresas que adotam tecnologias, inovam ou introduzem mudanças tecnológicas no mercado, e a oferta de conhecimento disponível das universidades, consultorias, centros de extensão tecnológica, laboratórios, etc. Dessa forma,

percebe-se a complexidade associada à inovação, em que o desempenho de inovação das empresas depende dos demais fatores, como a acumulação e alocação de capital físico e humano e do fluxo de conhecimentos.

Questões associadas ao ambiente de negócios, ou restrições ao financiamento, por exemplo, afetam a decisão de investimento das empresas em inovação, bem como os incentivos das universidades em fazer com que o conhecimento seja aplicável à indústria. A inexistência de universidades que tenham relação com o setor privado ou as restrições do sistema financeiro em dar crédito às pequenas e médias empresas, por exemplo, são também questões que afetam a decisão em inovação – e demonstram a complexidade da formulação e implementação de uma política dessa natureza.

Figura 1 – Estrutura do Sistema Nacional de Inovação



Fonte: Maloney (2017).

## DA TEORIA À PRÁTICA EM INOVAÇÃO: ESCALA DE CAPACIDADES

Diante desse cenário de complexidade, sugere-se construir as políticas de inovação de maneira progressiva e gradual, alinhadas com as capacidades das empresas. Trata-se do que chamamos de “escala de capacidades das empresas” (*capabilities escalator*). Em países como o Brasil, não se encontram muitas “Microsoft” ou “Apple”, mas pequenas e médias empresas com pouca capacidade ou nenhuma capacidade de fazer inovação muito radical. Logo, as políticas devem ser ajustadas para complementar essas capacidades das empresas.

O ponto principal é a necessidade de avaliar políticas adequadas ao tecido produtivo existente. E, então, adotar políticas de inovação a partir de um avanço sequencial e gradual. Em vez de fomentar só a pesquisa e o desenvolvimento (P&D) avançados, por exemplo, deve-se estimular inicialmente a adoção de novas tecnologias. Ou iniciar com uma política de apoio às consultorias em questões de gestão e, na medida em que as empresas avancem, conceber outras políticas mais sofisticadas e mais maduras (Figura 2). Desse modo, não apenas as empresas aprendem, como também o governo e os formuladores de política pública.

Figura 2 – Escala de capacidades: necessidades das políticas públicas de inovação



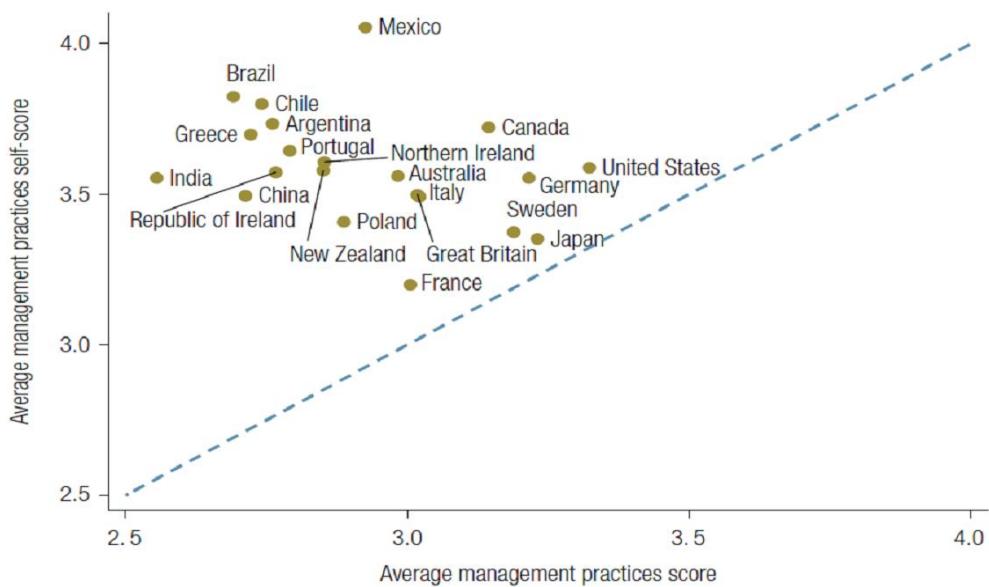
*Note:* NIS = National Innovation System; NQI = national quality infrastructure; R&D = research and development; STEM = science, technology, engineering, and mathematics.

Fonte: Cirera; Maloney(2017).

Isso mostra a importância de um marco teórico para auxiliar a formulação de políticas mais apropriadas. Muitas vezes, o que se tem é uma combinação de políticas que tentam fazer coisas de mais e não estão alinhadas às capacidades de fato existentes nas empresas. Como resultado, tem-se muito pouco impacto.

Outro elemento importante a ser enfatizado, e que é relevante para o apoio ao setor privado, é o fato de que as empresas “acham que sabem”. A *World Management Survey* (WMS), aplicada em diferentes empresas, buscou verificar a correlação entre o índice objetivo da qualidade da gestão, mensurado a partir da existência de práticas objetivas, e a percepção do empresário ou empreendedor quanto à qualidade de sua gestão (WORLD BANK, 2018). A Figura 3 apresenta os resultados dessa pesquisa. A linha pontilhada mostra os pontos em que o índice factual seria igual à percepção. Em todos os países, porém, os empresários pensam que são melhores em gestão do que realmente são – com destaque para a posição do Brasil, bem acima do nível.

Figura 3 – Gestão de fato e percepção de gestão: as empresas acham que sabem

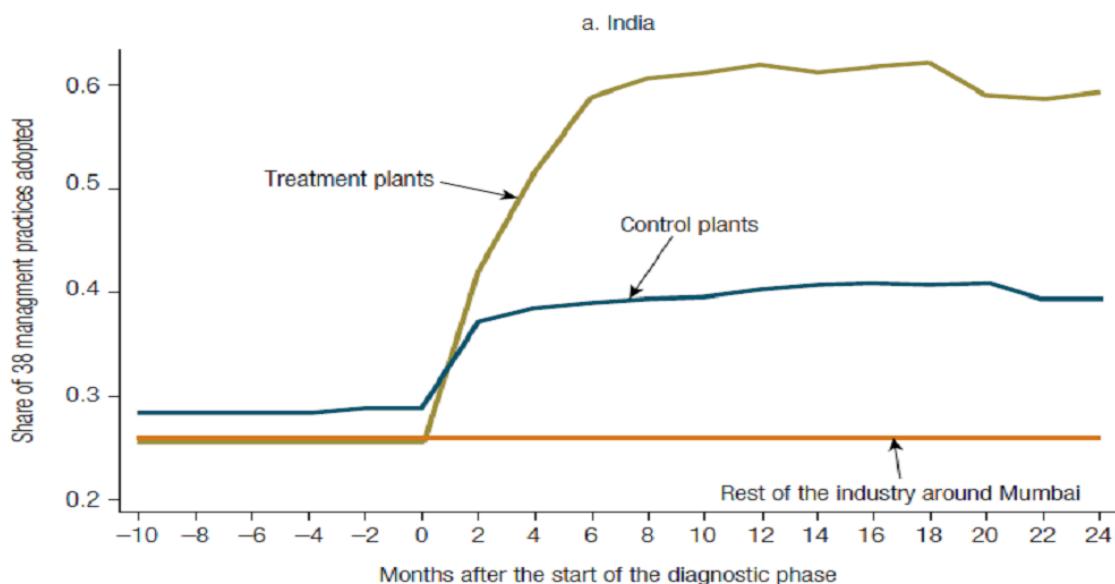


Fonte: World Bank (2018); Bloom and Van Reenen (2007).

Demonstrar isso é necessário para que as empresas tomem consciência da necessidade de adotar boas práticas internacionais de gestão e, assim, tenham condições de concorrer no mercado mundial. Além disso, há implicações em questões práticas, requerendo atenção quando da concepção, do desenho e da implementação de um novo programa. Quando um programa de assistência técnica pretende ajudar a pequena e a média empresa, as empresas acreditam não precisarem dessa ajuda. É difícil convencer as empresas e os empresários a participarem, dada a percepção de que isso é desnecessário e/ou custoso. No Quênia, por exemplo, um programa oferecia todo o processo de auditoria ambiental para as empresas, desde o diagnóstico dos problemas, até a entrega de um resumo específico sobre as soluções e/ou reformas necessárias para que a produção economize recursos. Nenhuma empresa se interessou. Esse resultado evidencia a existência de assimetria de informações, bem como a complexidade das políticas de inovação – o que torna ainda mais importantes os processos de avaliação e aprendizagem.

Cirera e Maloney (2017) destacam a necessidade de iniciar as políticas de inovação por melhorias na gestão das empresas e, então, seguir para a melhoria das tecnologias e a adoção de projetos mais complexos de P&D. Evidências coletadas na Índia e na Colômbia mostram que programas de extensão para apoio à gestão de empresas têm um impacto positivo na produtividade, desde que bem desenhados.

Figura 4 – Melhoria de práticas de gestão por meio de programas de extensão (Índia e Colômbia)



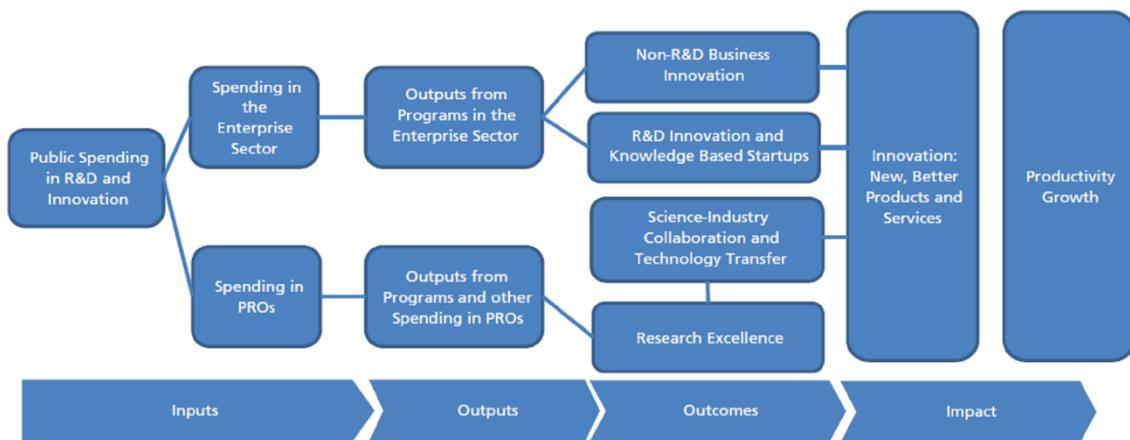
Fonte: Cirera; Maloney (2017)

## AVALIAÇÃO DO DESENHO E IMPACTO DA POLÍTICA PÚBLICA: ANÁLISE EM QUATRO ESTÁGIOS

Para dar conta das especificidades do campo da política pública de inovação, elaborou-se a metodologia de análise de gasto público da inovação, que busca avaliar a política como um todo. A ideia é qualificar a alocação dos recursos. Em outras palavras, verificar onde se poderia reduzir a intervenção e onde haveria espaço para novas políticas, pois ainda não há nada em operação.

Para tanto, construiu-se basicamente o marco lógico da intervenção. A partir das políticas do Estado, são avaliados os recursos gastos com uma política específica, das quais se tem produtos (*outputs*), e espera-se que esses produtos tenham algum impacto (Figura 5). Como exemplo, as políticas de apoio à consultoria de gestão, que têm como produtos consultorias que visam introduzir mudanças nas empresas para, enfim, obter ganhos de produtividade.

Figura 5 – Estrutura lógica da intervenção em Ciência, Tecnologia e Inovação



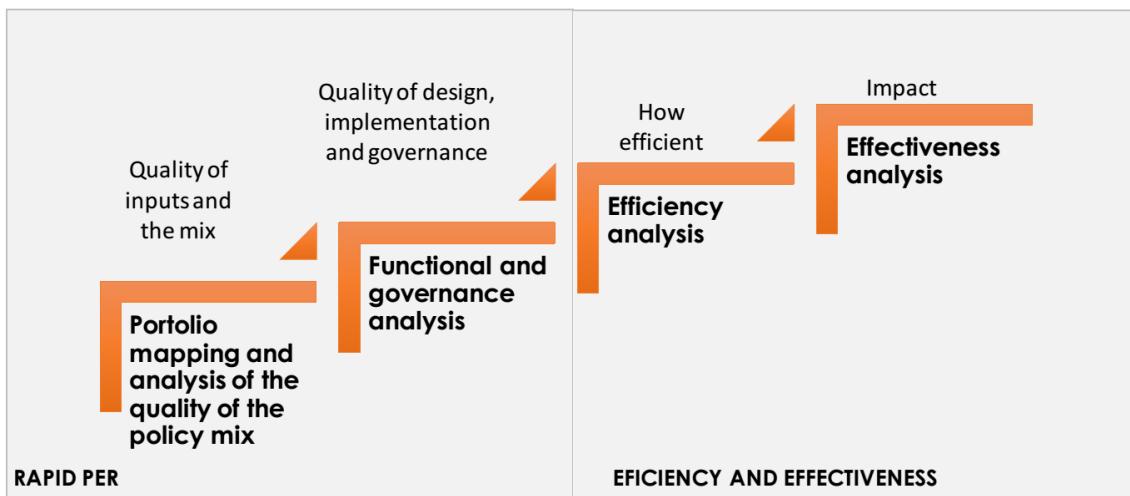
Fonte: Correa (2014).

Antes de aplicar o marco lógico, primeiramente, o Banco Mundial faz uma avaliação dos diferentes elementos da cadeia lógica, avaliando a qualidade do conjunto de política (*policymix*) e dos instrumentos existentes. Esse trabalho é feito através da observação da necessidade em termos de inovação (do território, do país ou do estado) e de como os instrumentos que estão disponíveis buscam solucionar os problemas identificados. Ocorrem muitos problemas quando não se olha de antemão a combinação desses instrumentos.

O segundo ponto compõe a chamada análise funcional de governança, em que se analisam as opções de boas práticas já existentes de desenho e de implementação das políticas. Isso facilita a análise quando, ao finalizar uma avaliação de impacto com resultado não positivo, pergunta-se se o problema foi o desenho ou a implementação do programa. Em seguida, cabe analisar a eficiência das alternativas, que está ligada tanto à qualidade do serviço a ser oferecido, quanto ao seu custo para o beneficiário – tanto em termos de tempo, quanto de esforço. Por vezes, o processo para a aplicação do programa é custoso, e as questões de eficiência devem ser consideradas. Por fim, realiza-se a avaliação de impacto propriamente dita, em que se busca analisar a efetividade do programa em produzir os resultados esperados.

A Figura 6 sintetiza essas etapas. Em síntese, essa metodologia visa retornar uns passos antes da avaliação de impacto e verificar possibilidades de melhorar a eficiência e a qualidade do desenho e da implementação dos programas – além de melhorar o conjunto de programas que estão sendo oferecidos no país.

Figura 6 – Estágios de análise da eficiência e efetividade de um programa em C&T



Fonte: Elaborado pelo autor.

Para tanto, o primeiro passo é identificar esse “portfólio de medidas” que já são oferecidas, com quais objetivos e a que custo. É necessária a construção de uma base de dados com todos os programas (para inovação, feiras de exportação, etc.), os respectivos objetivos, tipos de beneficiários e, o que é mais importante, o orçamento de cada um. No Brasil, esse exercício é bastante difícil, porque há muitas instituições públicas, semipúblicas, universidades e outras organizações que ofertam programas para a pequena e média empresa.

No entanto, isso permite informar com evidência quanto se gasta com pequenas e médias empresas, em que região ou setor, por exemplo. Permite verificar o quanto é destinado a empresas com potencial de inovação, a empresas inovadoras, a empresas que não são inovadoras, etc. Permite, também, analisar o gasto destinado ao objetivo geral de toda política de apoio ao setor privado do país. A exemplo de outros países, em que o objetivo geral era a diversificação econômica, quando analisados todos os programas existentes, observou-se que apenas 20% de todos os recursos estavam destinados à realização desse objetivo.

No Brasil, poucos programas têm como objetivo específico melhorar a produtividade, e o valor investido é baixo. Essa é uma questão que pode ser analisada. Outra questão é a existência de sobreposição de programas com o mesmo objetivo, tais como programas para incubadora ou *startups*. Do ponto de vista da política pública, é preciso fazer algum tipo de racionalização. Nesse caso, talvez fazer um programa com maior escala. O que não pode é haver diversos programas concorrendo pelos mesmos beneficiados. Esse tipo de sobreposição pode ser analisado a partir da abordagem de mapeamento de programas sintetizada no Quadro 1.

Quadro 1 – Mapeamento de programas de políticas para pequena e média empresa

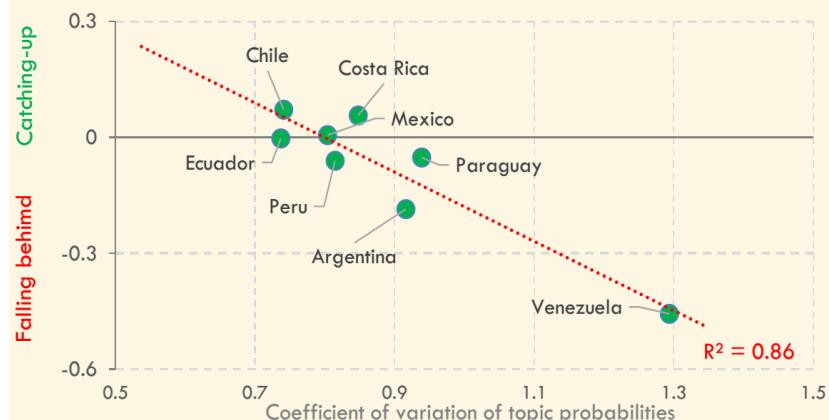
| Step  | Illustrative guiding questions   |
|---|--|
| <b>Mapping of policy instruments</b>                                      | <ul style="list-style-type: none"> <li>▪ What is the composition of the mix of policy instruments?</li> <li>▪ How have resources been allocated across the set of instruments?</li> <li>▪ How balanced are its goals?</li> <li>▪ What is the distribution of instruments across agencies and departments?</li> <li>▪ What is the profile of target beneficiaries aimed at by the mix?</li> <li>▪ To what degree has each instrument penetrated segments of micro, small and medium firms?</li> <li>▪ What is the geographical outreach of the mix of instruments?</li> <li>▪ What is the pattern of disbursement for each of the instruments over time?</li> </ul> |
| <b>Internal coherence in the composition of the policy instrument mix</b> | <ul style="list-style-type: none"> <li>▪ What is the composition of the policy mix relative to the space of all possible instruments aiming at the same objectives (i.e. international comparison)?</li> <li>▪ Is there a balance of goals and intended beneficiaries?</li> <li>▪ Is the choice of instrument consistent with its intended goals?</li> <li>▪ Is the scale of instruments and programs appropriate?</li> </ul>  |
| <b>Consistency across the composition of the policy instrument mix</b>    | <ul style="list-style-type: none"> <li>▪ Does the mix of instruments presents conflict, neutrality or complementarities?</li> <li>▪ What has been the intensity of co-deployment of instruments by target firms, regions?</li> <li>▪ Is the sequencing across instruments appropriate?</li> </ul>  |

Fonte: Elaborado pelo autor.

Por exemplo, em trabalho que buscou analisar o compromisso com a temática da inovação presente nos planos de governo ao longo de um período de 30 a 40 anos, nos países da América Latina, buscou-se verificar se haveria correlação entre a menção da inovação como prioridade e o grau de inovação apresentado no país. As Figuras 7 e 8 demonstram a existência de uma correlação positiva e consistente. Verifica-se, por exemplo, a enorme diferença entre Argentina e Chile.

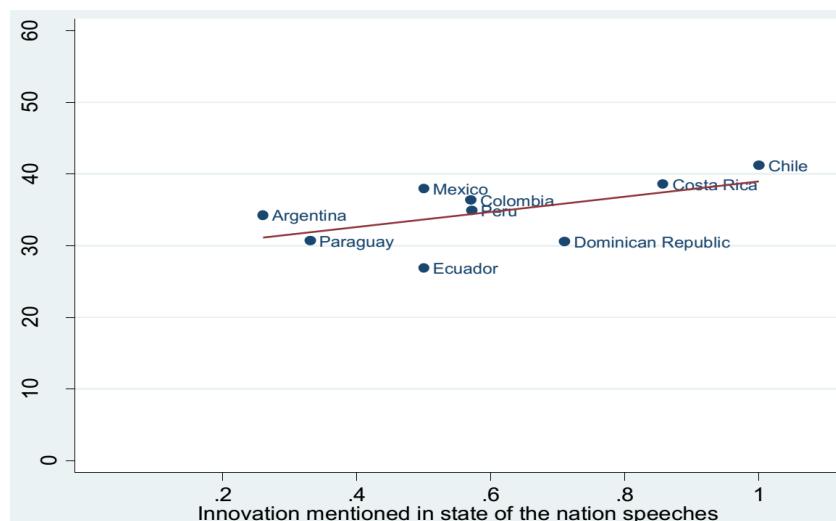
Figura 7 – Vencedores nunca desistem e desistentes nunca crescem?

Policy discontinuity and  $\Delta$  in % US GDP per capita, 1950-2010



Fonte: Calvo, Eizmendi and Reyes (2017).

Figura 8 – Comprometimento político com inovação e grau de inovação

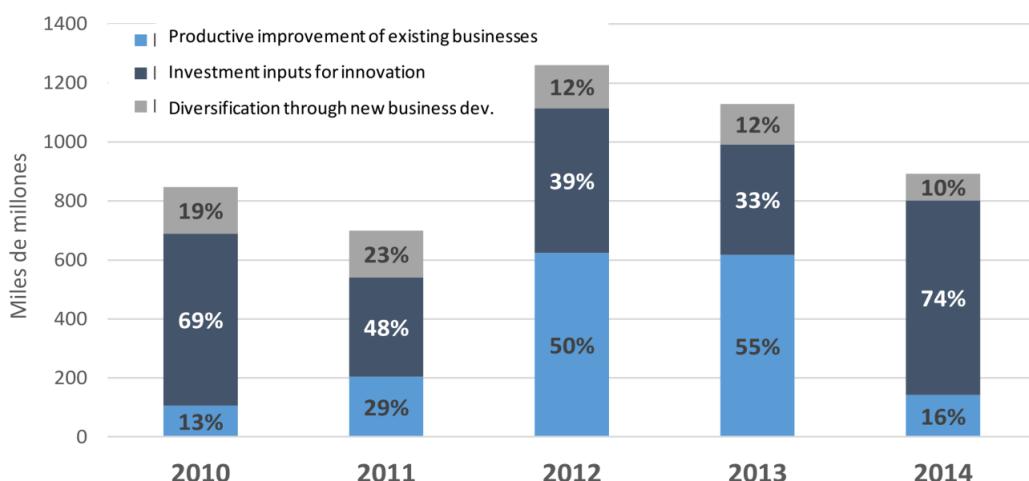


Fonte: Cirera; Maloney (2017)

Esses resultados evidenciam a importância de ser consistente nas políticas – e essa é uma discussão que está se tendo no âmbito do Programa Inovar-Auto, do Governo Federal do Brasil. Independentemente do desenho do programa, investimentos no setor de automóveis são de 5 a 10 anos. Nesse setor, portanto, as políticas não podem mudar a cada 3 ou 4 anos, porque isso causa incerteza e inibe o desenvolvimento da indústria.

A Figura 9, por exemplo, evidencia a volatilidade do gasto de um país da América Latina com política de diversificação no período de 2010 a 2014, tendo representado de 19 a 10% da despesa total com políticas para o setor privado. Contudo, a estratégia geral adotada no país afirma ter a diversificação como objetivo primordial. Então tem que ser consistente: se essa é a prioridade, é para onde a maioria dos recursos deve ser alocada.

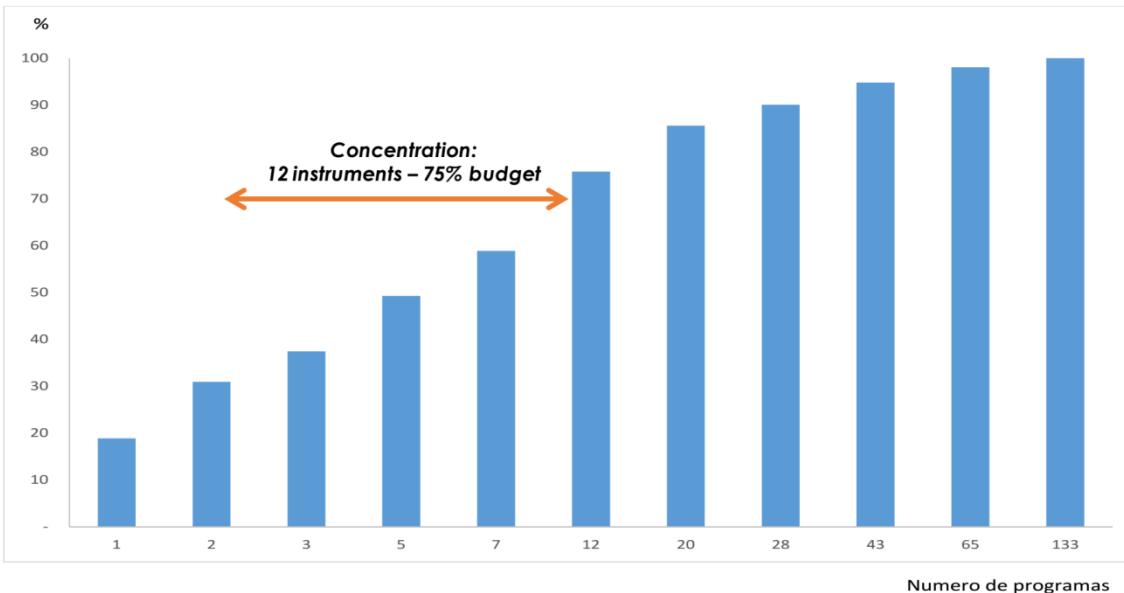
Figura 9 – Despesas públicas por objetivo de apoio ao setor privado (2010-2014)



Fonte: Elaborado pelo autor.

Quando se verificam os dados, ainda, percebe-se a concentração de recursos em alguns poucos programas, como pode ser observado na Figura 10.

Figura 10 – Muito orçamento para poucos programas

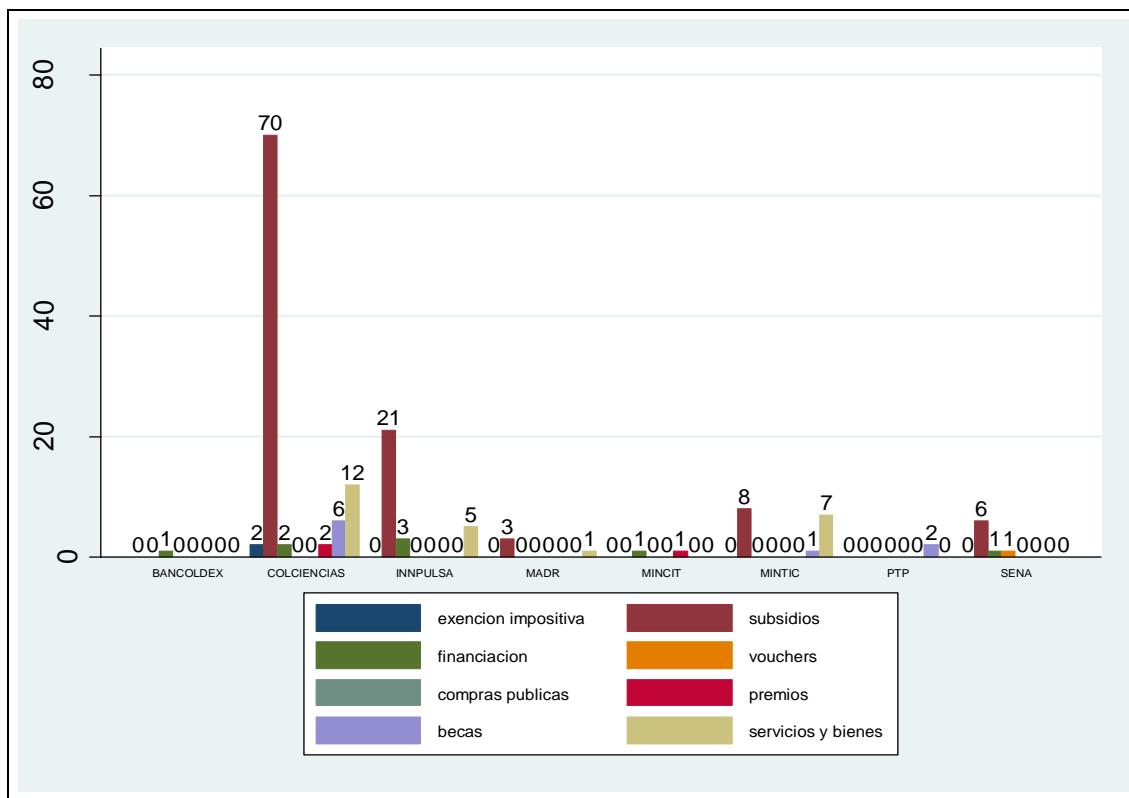


Fonte: Elaborado pelo autor.

## A IMPORTÂNCIA DO DESENHO DA INTERVENÇÃO: FORMULAÇÃO E IMPLEMENTAÇÃO

Além da identificação do rol de medidas existentes e dos recursos por elas consumidos, outro elemento importante a ser considerado é o mecanismo da intervenção em si. A Figura 11 apresenta a alta concentração de recursos em subsídios equivalentes (*matching grants*), além de outros instrumentos com menor participação. No entanto, antes de definir o instrumento, é preciso avaliar qual o problema que enfrentam as empresas, por exemplo, melhora da produtividade, informação, crédito do mercado financeiro, assistência técnica para atualizar métodos de produção, certificação, etc.? O BNDES basicamente opera por meio de crédito subsidiado. Se não foi esse o problema das empresas, porém, o crédito subsidiado não necessariamente apresentará impacto.

Figura 11 – Alta concentração de apoio direto por meio de subsídios equivalentes (*matching grants*)

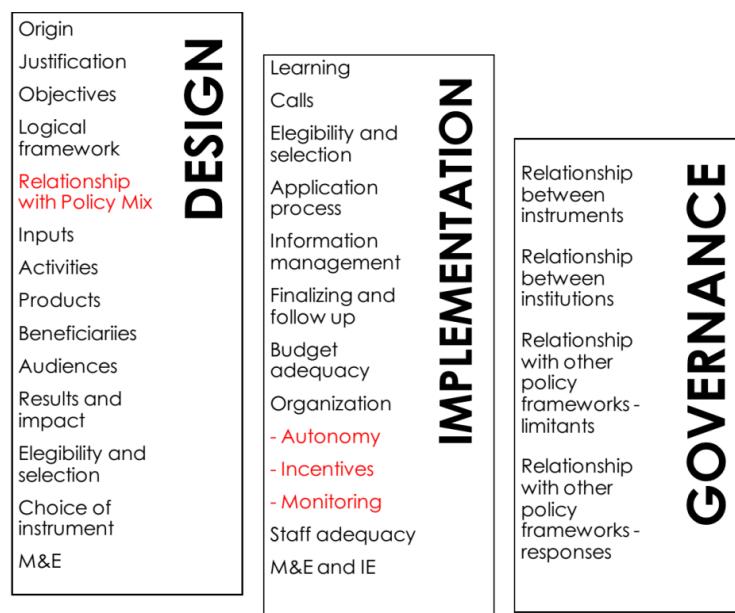


Fonte: Elaborado pelo autor.

Analisar o mecanismo de intervenção requer chamar atenção também para a importância da qualidade do desenho e da implementação dessa intervenção. Para tanto, o Banco Mundial elaborou um modelo de boas práticas para ajudar a melhorar o desenho e a implementação dos programas de apoio ao setor privado. O modelo contém uma série de elementos para apoiar o desenho e identificar a origem do programa, a necessidade que pretende atender, seu fundamento (estudos, pesquisas, etc.), sua justificativa, seus objetivos e, especialmente, seu marco lógico. Um dos problemas observados em alguns países quando começaram a fazer a avaliação de impacto era a inexistência de um marco lógico. Nesse caso, não se pode fazer um monitoramento e avaliação de impacto sem um marco lógico.

Quanto à implementação, o modelo busca facilitar a identificação do mecanismo da aprendizagem, bem como questões da prática de gestão, da coleta e monitoramento de informação, etc. Por fim, há ainda questões de governança e de coordenação a serem consideradas. A Figura 12 resume essas três dimensões do modelo e suas respectivas categorias de análise.

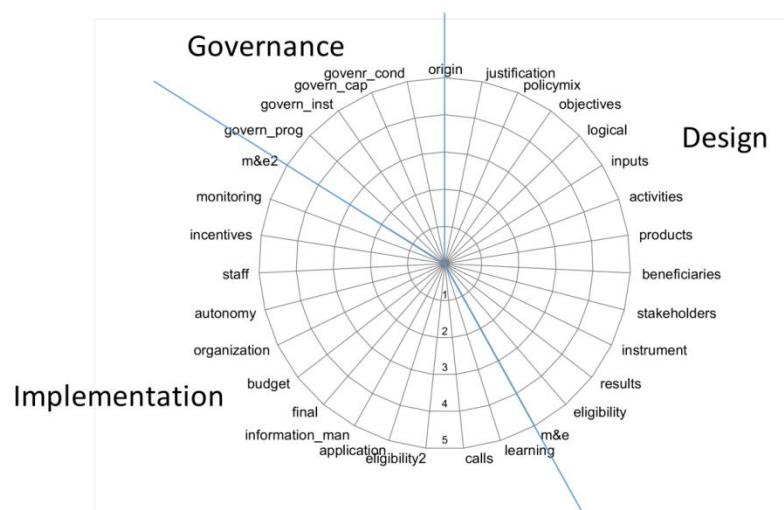
Figura 12 – Desenho, implementação e governança: dimensões e categorias de análise



Fonte: Elaborado pelo autor.

A Figura 13 apresenta um resumo de todas as dimensões, em que 1 ponto é considerado muito ruim, enquanto 5 pontos representam a melhor prática internacional. Por exemplo, um caso onde a gestão da informação é *standard*, e pode-se procurar e encontrar qualquer informação, recebe 5 pontos. Por sua vez, se não há nenhum sistema de informação, 1 ponto. Um segundo exemplo é de um programa de não tem nenhuma justificativa descrita e preservada na memória institucional. Os funcionários mudam de departamento constantemente e operam o programa sem nem saber por que o fazem. Nesse caso, 1 ponto.

Figura 13 – Metodologia de análise funcional



Fonte: Elaborado pelo autor.

O marco lógico também é um caso específico, em que toda a planificação de instrumento é ligada ao orçamento para o ano. Nesse sentido, todo o objetivo do instrumento pode ser perdido se não houver processos e sistemas dentro do governo, que se mantenham ao longo do programa. Esse é um ponto muito importante, porque isso permite que sejam feitas intervenções mais sistêmicas, onde se podem corrigir os sistemas de informação, as necessidades de treinamento, etc. para todos.

Vale dizer que a maioria das instituições tem ao menos um programa considerado boa prática de desenho, implementação ou governança. Há muita coisa que é aprendizagem dentro da instituição e que pode ser arranjada a muito baixo custo. Identificar essas questões é um primeiro – e fundamental – passo.

## **CONSIDERAÇÕES FINAIS**

O texto apresentou estratégias para uma avaliação qualitativa de políticas e programas públicos, com vistas a fazer uma avaliação integrada de toda a intervenção. Para tanto, o Banco Mundial criou alguns modelos para apoiar não só a realização das avaliações de impacto de um programa específico, mas também para ajudar a melhorar o desenho e a implementação desses programas.

No contexto das políticas de inovação, em específico, reconhece-se a dificuldade e a complexidade que envolve a formulação e aplicação de uma política pública direcionada a mudar os incentivos dos agentes econômicos, de maneira sistêmica. Sendo assim, buscou-se ressaltar ao longo deste texto a necessidade de uma teoria que fundamente de maneira consistente a formulação de uma política de inovação para o desenvolvimento, adequada ao tecido econômico existente.

Nesse sentido, ressalta-se a importância do marco lógico do programa, que é basicamente o fundamento da intervenção. Sem o marco lógico não há monitoramento e avaliação de impacto. Tampouco há clareza do que pode e como pode funcionar. Portanto, o mecanismo de aprendizagem não é ter uma avaliação de impacto em si, mas ter um processo que ajude a integrar os resultados esperados dentro do programa. Devemos investir em avaliação de impacto, mas também na adoção de boas práticas de desenho e implementação. Temos que ter a mínima clareza quanto às regras que definem o que será avaliado, quem fará o que e quem/o que entregará quais resultados. Isso serve tanto para melhorar um programa, quanto para decretar seu sim, diante da evidência de que, mesmo sendo adequadamente formulado e implementado, não produz os resultados esperados.

## **REFERÊNCIAS**

BLOOM, Nicholas; VAN REENEN, John. Measuring and Explaining Management Practices across Firms and Countries. *Quarterly Journal of Economics*, 122 (4), 2007, pp. 1351–1408.

CIRERA, Xavier; MALONEY, William F. *The innovation paradox: developing-country capabilities and the unrealized promise of technological catch-up*. World Bank Group, 2017.

CORREA, Paulo. *Public Expenditure Reviews in Science, Technology, and Innovation: A Guidance Note*. Washington, DC: World Bank, 2014

MALONEY, William. *Revisiting the National Innovation System in Developing Countries*. Washington, DC: World Bank, 2017

WORLD BANK. *World Management Survey (WMS)*. Washington, DC: World Bank, 2018. Disponível em: <http://worldmanagementsurvey.org/>. Acesso em: 18 jun. 2018.

## AVALIAÇÃO DE IMPACTO DE POLÍTICAS PÚBLICAS: MÉTODOS EXPERIMENTAIS\*

Rafael Dantas<sup>1</sup>

**Resumo:** O texto apresenta o método experimental para avaliação do impacto nas políticas públicas, o qual objetiva identificar a existência (ou não) de uma relação de causa e efeito e, assim, produzir evidência rigorosa sobre a eficácia de um programa ou uma política pública. Será dada ênfase para algumas definições acerca do público-alvo e do grupo controle, sendo este obtido por meio de diferentes técnicas de aleatorização, a fim de assegurar a validade externa e interna do resultado.

**Palavras-chave:** Avaliação do impacto. Políticas públicas. Método experimental.

**Abstract:** This text presents the experimental method to impact evaluation on public policy. Experimental methods aim at identifying the existence (or not) of a cause and effect relationships and then produce criterious evidence about the effectiveness of a program or a public policy. It is presented some definitions about target group e control group. Different randomization techniques to define control group are presented, in order to assure external and internal validation of the evaluation research.

**Keywords:** Impact evaluation. Public policy. Experimental method.

---

\* Texto elaborado a partir da transcrição da fala e da apresentação em slides de palestras proferida no Workshop Avaliação de Impacto de Políticas Públicas, realizado nos dias 29 e 30 de novembro de 2017, no Auditório da Secretaria de Planejamento, Governança e Gestão (SPGG-RS), com o apoio do Banco Mundial.

<sup>1</sup> Coordenador de Campo em Avaliação de Impacto, Banco Mundial.

## INTRODUÇÃO

O artigo trata do método experimental para avaliação do impacto das políticas públicas. O método experimental visa identificar a existência (ou não) de uma relação de causa e efeito, produzindo evidência rigorosa sobre a eficácia de um programa. Nas seções seguintes, abordaremos esse tema, destacando algumas definições sobre público-alvo e aleatorização, necessária para a obtenção de uma amostra que produza estimativas que sejam representativas.

## OBJETIVOS DA POLÍTICA PÚBLICA: COMO MEDIR?

Os objetivos que uma política pública pretende alcançar são amplos. Por exemplo, pode ser que o objetivo seja melhorar a aprendizagem dos estudantes, aumentar a produtividade das pequenas empresas, interromper a transmissão intergeracional da pobreza, ou engajar o cidadão na política. Mas será que a política pública empregada para alcançar esse objetivo específico funcionou? Quais argumentos podem ser usados para comprovar, ou não, a afirmação de que determinada política foi eficaz?

É importante contestar as afirmações de que a política funciona ou não funciona. Este questionamento trata do viés de seleção que pode haver nos resultados obtidos, uma vez que não necessariamente o programa pode ter incentivado ou estimulado as ações que levaram a tais resultados. Por isso, o objetivo de avaliação de impacto é identificar uma relação de causa e efeito. Ou seja, atribuir exclusivamente ao programa o efeito que se está observando.

## AVALIAÇÃO DE IMPACTO E O GRUPO CONTRAFACTUAL

Mas como medir o impacto do programa?

Duas abordagens comumente empregadas são avaliar o que ocorreu antes e depois da intervenção e avaliar o que ocorreu de diferente entre os participantes e os não-participantes. Veremos que essas são duas abordagens ingênuas e que não medem o impacto de uma intervenção.

Na avaliação antes-e-depois, mede-se o resultado da intervenção comparando-o com as métricas obtidas antes da implementação do programa. O problema dessa abordagem é que ela não leva em consideração o efeito do tempo nas métricas avaliadas. Considere uma política de treinamento da mão de obra com o objetivo de aumentar a absorção pelo mercado de trabalho. Algum tempo depois, observamos que os trabalhadores que participaram desse programa estão empregados. Mas poderíamos afirmar que isso se deu por conta do programa que eles participaram? Falhamos em observar que, com o tempo, outras variáveis também

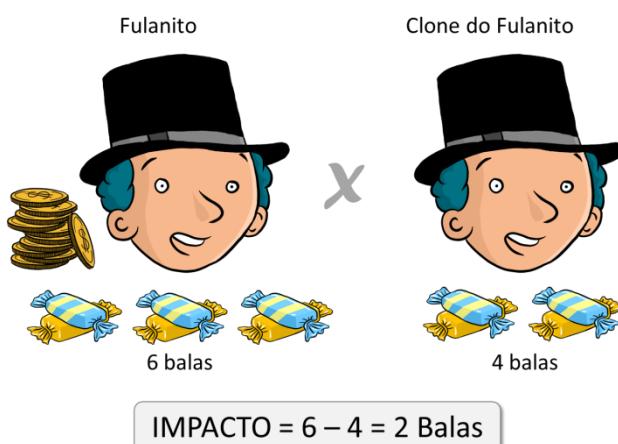
mudam. Por exemplo, pode haver uma melhoria generalizada na economia, aumentando a demanda por trabalho e sendo esse um fator para os resultados que observamos. Não conseguimos assim isolar o efeito da política citada.

Comparar participantes com não-participantes também não mensura o impacto de um programa. Considere a mesma política de treinamento da mão-de-obra e que os candidatos devem se inscrever para participar do programa. Ao final, observamos que a empregabilidade é maior dentre aqueles que se inscreveram e participaram do programa do que aqueles que não se inscreveram e não participaram. No entanto, o ato de se inscrever pode nos dizer muito mais sobre as características dos candidatos que acabaram participando do programa. Estes podem ser os candidatos mais motivados, mais esforçados ou que possuem mais informação sobre os programas disponíveis. E, por conta dessas características, eles teriam, de qualquer jeito, maiores chances de emprego do que aqueles que não se inscreveram. Ou seja, também não conseguimos isolar o impacto da política do impacto das características dos participantes e, portanto, temos uma medida viesada.

A avaliação de impacto, por outro lado, busca medir o resultado que teria ocorrido com o participante caso ele não tivesse participado do programa. Trata-se da chamada análise do contrafactual, por meio da qual se busca comparar o participante nesses dois estados, um em que ele recebe a intervenção e outro em que não recebe, ao mesmo tempo, no mesmo lugar e sob mesmas circunstâncias. Ou seja, é a situação ou evento que não aconteceu, mas poderia ter acontecido.

Por exemplo, suponha que em um experimento uma pessoa receba seis balas para consumo, e que exista, sob as mesmas condições, um clone perfeito dessa pessoa, mas que recebe quatro balas para consumo. A diferença de duas balas é identificada como impacto. A Figura 1 ilustra isso.

Figura 1 – Mundo ideal da avaliação de impacto: comparação entre clones perfeitos

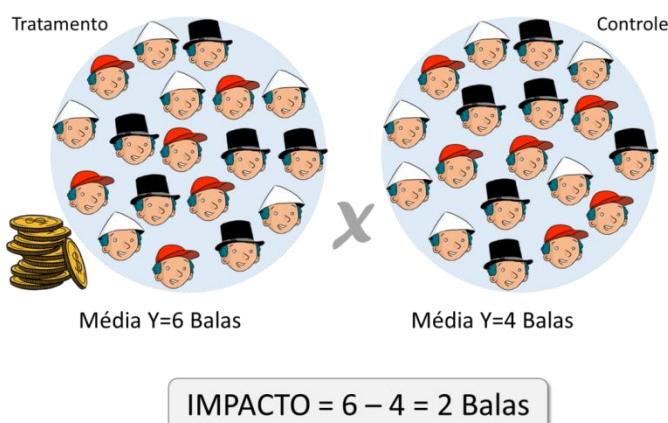


Fonte: Elaboração do autor.

Contudo, medir dois estados para um *mesmo* participante é algo impossível. Para resolver esse problema, podem-se comparar, então, grupos estatisticamente semelhantes, com e sem a intervenção.

Por meio de técnicas estatísticas, especificam-se dois grupos semelhantes entre si: o grupo tratamento que recebe a intervenção (tratamento) e um grupo que não recebe a intervenção (controle). Os dois grupos, tratamento e controle, possuem características iniciais (observáveis e não-observáveis) semelhantes. Entretanto, apenas o grupo tratamento recebe a intervenção, sendo essa a única característica que diferencia um grupo do outro. Assim, a intervenção é a única razão para diferenças observadas após a intervenção. Logo, qualquer diferença observada entre esses dois grupos pode-se atribuir à intervenção (Figura 2).

Figura 2 – Avaliação de impacto na prática: grupos estatisticamente semelhantes



Fonte: Elaboração do autor.

## MÉTODO EXPERIMENTAL: GRUPOS COMPARÁVEIS PARA ASSEGURAR VALIDADE EXTERNA E INTERNA DO RESULTADO

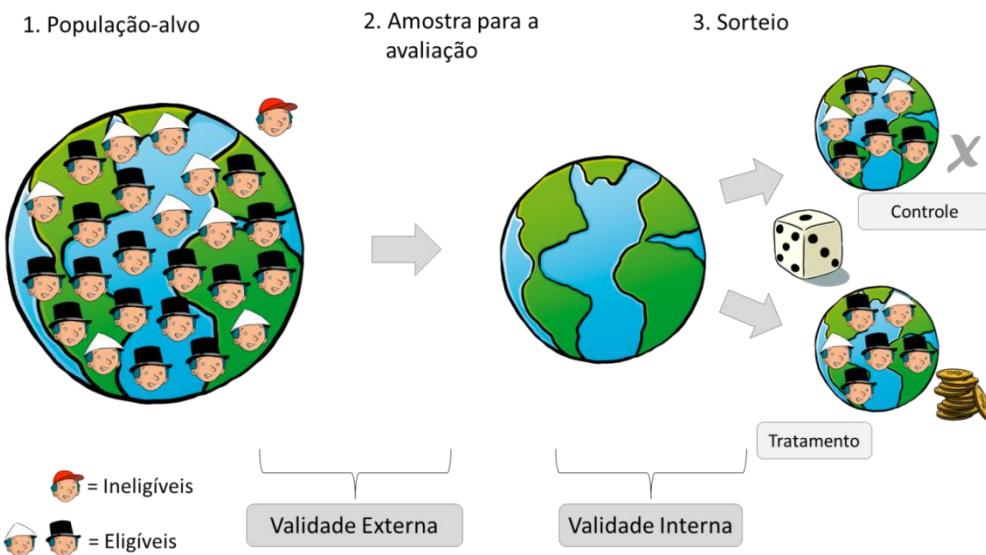
O método experimental é um dos métodos usados para construir os grupos tratamento e controle em uma avaliação de impacto. Nesse método, a definição dos participantes de cada grupo é feita aleatoriamente a partir de uma amostra da população-alvo do programa, composta pelos participantes elegíveis ao programa. A probabilidade de pertencer ao grupo tratamento ou ao grupo controle é igual para todos.

Sendo a amostra aleatória de tamanho suficientemente grande, e o grupo tratamento e controle definidos por sorteio, pode-se concluir que os grupos são comparáveis entre si, sendo a única diferença o tipo de intervenção que um deles receberá. Essa diferença indicará o impacto da política.

Para assegurar que as características, tanto observáveis quanto não-observáveis, estão igualmente distribuídas entre os dois grupos, uma amostra grande deve ser selecionada. Segundo a “lei dos grandes números”, em estatística, uma amostra grande possibilita obter estimativas não viesadas. Ou seja, as estimativas de impacto obtidas são devidas somente à diferença na intervenção.

A partir das etapas de seleção dos participantes em cada grupo, tem-se duas formas de analisar a evidência produzida, seja por validade externa, seja por validade interna. A validade externa é a capacidade de generalização dos resultados para a população-alvo. Ou seja, o impacto observado pode ser estendido à população como um todo. Por sua vez, quanto à validação interna, entende-se que os resultados obtidos evidenciam uma causalidade, pois se estão usando grupos comparáveis. A comparação entre os grupos, então, reflete o impacto naquela população (Figura 3).

Figura 3 – Seleção aleatória dos grupos de tratamento e controle



Fonte: Elaboração do autor.

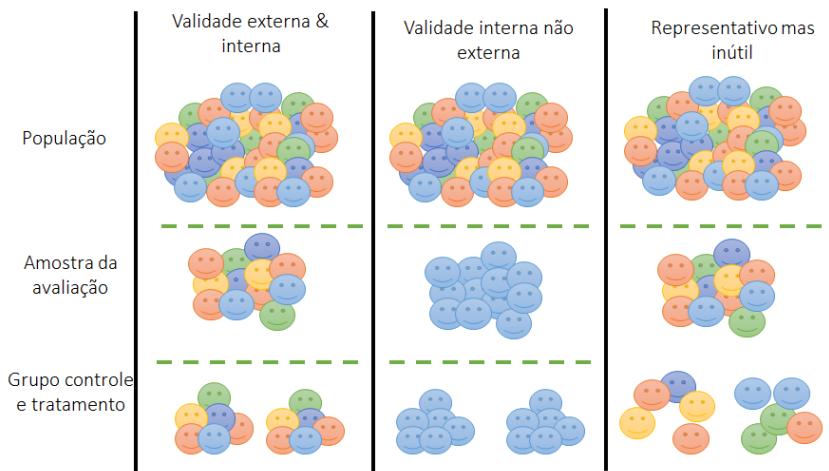
Cuidados com a validade interna e externa são necessários para analisarem-se adequadamente os resultados. A Figura 4 ilustra algumas situações. Na primeira coluna, observa-se uma amostra que é representativa da população-alvo, cujos grupos controle e tratamento foram aleatorizados, sendo representativos da população. Assim sendo, os resultados da análise detêm validade externa e interna.

Na segunda coluna, porém, é apresentado um exemplo em que se tem validade interna, mas não validade externa. A amostra selecionada não é representativa da população, porém produz resultados causais por meio da seleção aleatória dos integrantes que compõem ambos os grupos. Logo, os resultados obtidos valem apenas para essa população em azul.

Na última coluna, por fim, não há validade interna. A comparação está ocorrendo entre pessoas diferentes nos grupos, embora a amostra selecionada seja representativa da

população. Entretanto, o resultado é inútil porque se deseja generalizar um resultado que não é válido para a população como um todo.

**Figura 4 – Exemplos de validação**



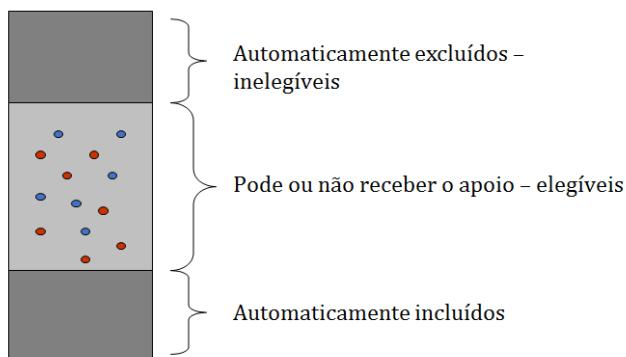
Fonte: Elaboração do autor.

## PÚBLICO-ALVO DA POLÍTICA E PÚBLICO-ALVO DA AVALIAÇÃO DE IMPACTO

Acima se falou de público-alvo e população. Mas qual é o público-alvo do programa? Considere a Figura 5 abaixo. Suponha que exista um programa cujo foco sejam as pequenas e médias empresas. Consequentemente, grandes empresas serão excluídas automaticamente, pois não são elegíveis. Contudo, pode haver um subgrupo que precise de mais assistência, e que, portanto, seja automaticamente incluído no programa. Esse grupo, porém, também será tratado como inelegível para fins de avaliação de impacto.

Por sua vez, um segundo grupo (grupo do meio da Figura 5) pode ou não receber o programa. Esse é um grupo relevante para a avaliação de impacto, sendo considerado um grupo elegível. Dentro desse grupo é que ocorrerá a seleção aleatória do grupo que receberá ou não a intervenção (tratamento e controle). Dessa forma, a avaliação de impacto irá abranger somente esse grupo, e a generalização das conclusões obtidas ocorrerá somente para ele.

Figura 5 – Definição do público-alvo



Fonte: Elaboração do autor.

## OPORTUNIDADES DE ALEATORIZAÇÃO

Compreendido como se compõe o público-alvo de uma avaliação de impacto, cabe discutir como aleatorizá-lo para compor os grupos tratamento e controle. Quando se formulam políticas públicas, há pelo menos quatro fatores que se controla e que são passíveis à aleatorização, que são: o acesso, o *timing*, o encorajamento e a intensidade.

- **Acesso (aleatorização pura):**

Podemos escolher a quem será oferecido o acesso ao programa. As possíveis unidades de aleatorização podem ser: indivíduo ou domicílio, estudantes ou escolas, bairro ou região. Para isso, é preciso ter uma amostra suficientemente grande para conseguir detectar os efeitos. Por exemplo, ao escolher estudantes em vez de escolas, obtém-se uma amostra maior. Isso decorre do fato de que, em uma, existe similaridade entre os estudantes, sendo necessária uma amostra maior de escolas para se detectar o mesmo efeito do que uma amostra de estudantes, independentemente da escola. O agrupamento reduz a efetividade do tamanho da amostra, ainda que reduza os custos de operacionalização e coleta dos dados.

Não obstante, deve ser considerado o risco de contaminação da amostra. Ao se definir os estudantes como nível de aleatorização, por exemplo, por estarem muito próximos um do outro, existe a possibilidade de alguém que esteja recebendo o tratamento influenciar o outro que não está recebendo, ou vice-versa. Nesses casos, pode-se pensar em fazer por escolas, que estão mais separadasumas das outras. Mas a regra geral é tentar aleatorizar na menor unidade possível.

- **Entrada gradual (phase-in design):**

Contexto em que todos devem receber o programa, mas por restrição orçamentária e/ou de logística, não é possível que todos recebam o programa ao mesmo tempo. Portanto, não há como ter um controle puro. Diante disso, pode-se aleatorizar quem receberá primeiro e por último o programa. Nesse cenário, quem recebe primeiro atuará como grupo

tratamento, e quem recebe depois atuará como grupo controle – mesmo que eventualmente venha a receber o programa.

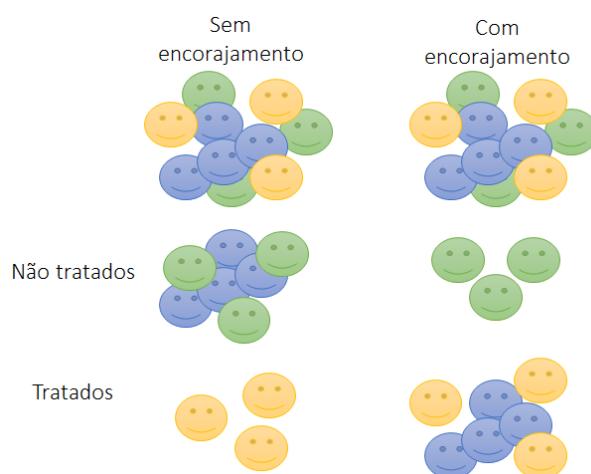
- **Encorajamento:**

Suponha que se lance um programa que ofereça instrumentos financeiros para população carente. O programa tem como objetivo a inclusão financeira e disponibiliza microcrédito, microseguros, internet *banking*, educação financeira, etc. Pode-se pensar que muitos indivíduos terão interesse no programa. No entanto, pode acontecer o inverso: baixa adesão (*take-up*) – que, na verdade, é um problema bastante frequente nas políticas públicas.

Diante desse cenário, pode-se sugerir um experimento com encorajamento. Dentre os elegíveis, aleatorizar quem receberá um incentivo para ser encorajado a participar (grupo de tratamento) e quem não receberá incentivo (grupo de controle). Esse incentivo deverá afetar apenas a probabilidade da pessoa aceitar ou não o programa, com o objetivo de aumentar adesão ao programa. O grupo relevante para a avaliação de impacto será aquele que responder ao encorajamento e atenderá ao programa.

Portanto, estão sendo considerados três grupos. Têm-se aqueles que decidirão aderir ao programa, os chamados de tratados, e os que não irão aderir, os não tratados. Há ainda aqueles que, independentemente de encorajamento ou não, nunca vão querer ser tratados. A Figura 6 ilustra esse experimento em três tipos de pessoas: os azuis, os amarelos e os verdes. Os verdes são os que não querem ser tratados, independentemente do incentivo. Os amarelos querem sempre ser tratados e sempre irão aderir ao programa, com ou sem incentivo. Já o grupo azul é composto por aqueles que, se não for dado encorajamento, não serão tratados, mas, se for dado, serão tratados. Logo, o grupo relevante para a avaliação de impacto será o grupo azul. Dessa forma, comparamos os integrantes do grupo azul que não receberam o encorajamento, com os que receberam.

Figura 6 – Esboço de encorajamento



Fonte: Elaboração do autor.

- **Variações de tratamento ou de intensidade de tratamento**

Contexto onde todos os elegíveis devem receber o programa, mas, devido à já possuir uma adesão alta, não é oportuno lançar mão do encorajamento. Diante disso, a estratégia é oferecer múltiplos tratamentos.

Por exemplo, um programa de incentivo a pequenas e médias empresas. Pode-se oferecer para um grupo um programa onde será fornecida uma informação sobre quais são as melhores práticas de negócios. Para um segundo grupo, será oferecida uma informação a mais, considerada uma variação nesse tratamento. Dessa forma, serão considerados dois tratamentos e dois grupos, se medirá o impacto da informação mais detalhada.

Outra forma de variação no tratamento é oferecer o mesmo tratamento, mas com intensidades diferentes. Por exemplo, considere um programa de transferência condicional de renda, em que para um determinado grupo seja oferecido R\$ 40,00, enquanto que para outro grupo é oferecido R\$ 60,00, e a outro grupo é ofertada uma dosagem a mais, R\$ 80,00. Com isso, procede-se a uma avaliação de impacto dessas diferentes doses do tratamento.

## CONSIDERAÇÕES FINAIS

A contestação acerca das afirmações de que a política funciona ou não funciona leva à avaliação do viés de seleção, que pode ocorrer nos resultados obtidos. Um programa pode ou não ter incentivado, ou estimulado ações que levaram a determinados resultados. A avaliação de impacto auxilia na identificação desses resultados, buscando identificar uma relação de causa e efeito atribuída exclusivamente à intervenção.

As técnicas estatísticas são meios de analisar o impacto da política, a partir da especificação de dois grupos semelhantes entre si. Sendo que o grupo tratamento o que recebe a intervenção, e o grupo controle o que não recebe. Logo, a única característica que diferencia um grupo do outro é a aplicação da intervenção.

O método experimental possibilita definir os grupos, tratamento e controle, aleatoriamente. Para isso, é importante especificar a população-alvo, que é composta pelos participantes elegíveis ao programa, com a finalidade de selecionar uma amostra aleatória de tamanho suficientemente grande. O delineamento a respeito da seleção e validação dos participantes do programa deve ser bem definido para que a construção dos grupos seja representativa e conduza a estimativas não viesadas. Este texto buscou orientar sinteticamente as formas disponíveis para compor esses grupos.

## REFERÊNCIAS

GERTLER, Paul J. et al. *Impact Evaluation in Practice*. Washington, DC: Inter-American Development Bank and World Bank, 2016.

## AVALIAÇÃO DE IMPACTO DE POLÍTICAS PÚBLICAS: MÉTODOS NÃO-EXPERIMENTAIS\*

Beatrice Zimmermann<sup>1</sup>

**Resumo:** O texto apresenta uma introdução aos métodos não-experimentais de avaliação de políticas públicas. Para tanto, ressalta-se a importância do desenho de um bom contrafactual, a fim de produzir evidências robustas a partir da avaliação. Qualquer grupo de comparação não é, necessariamente, um bom grupo de comparação, sendo necessário tomar cuidados para construí-lo. Nesse sentido, dois métodos não-experimentais serão apresentados: Diferenças-em-Diferenças (DD) e Regressão Descontínua (RDD), sendo discutidas suas características e vantagens em relação aos métodos chamados “ingênuos” de definição do contrafactual.

**Palavras-chave:** Políticas públicas. Avaliação de impacto. Métodos não-experimentais. Contrafactual.

**Abstract:** The text presents an introduction to the non-experimental methods of public policy evaluation. The importance of designing a good counterfactual is emphasized, in order to produce robust evidences. It is argued that any comparison group is not necessarily a good comparison group, so it is required attention being in order to build it. In this sense, two non-experimental methods are presented: Differences-in-Differences (DD) and Regression Discontinuity (RD). It is discussed their characteristics and advantages in contrast to the so-called *naive* methods of defining the counterfactual.

**Keywords:** Public policies. Impact evaluation. Non-experimental methods. Counterfactual.

---

\* Texto elaborado a partir da transcrição da fala e da apresentação em slides de palestras proferida no Workshop Avaliação de Impacto de Políticas Públicas, realizado nos dias 29 e 30 de novembro de 2017, no Auditório da Secretaria de Planejamento, Governança e Gestão (SPGG-RS), com o apoio do Banco Mundial.

<sup>1</sup> Coordenadora de Campo do *Development Impact Evaluation* (DIME), Banco Mundial.

## INTRODUÇÃO

O presente artigo tem como objetivo fazer uma introdução aos métodos não-experimentais de avaliação de políticas públicas. Os métodos não-experimentais, apesar de não constituírem a primeira e mais adequada das opções em termos de avaliação de impacto, precisam ser conhecidos e dominados por avaliadores de políticas públicas. Considerando que nem sempre é possível planejar um programa antes de ele ser executado, é altamente provável que o analista de política, ao longo de sua carreira, precisará trabalhar com dados gerados *a priori*, quando o projeto objeto da avaliação já ocorreu. Além disso, algumas vezes há a necessidade de se utilizar métodos não-experimentais mesmo quando um experimento foi construído, pois alguns fatores ao longo do processo (implementação, desenho, etc.) podem fazer com que a comparação simples entre os grupos de tratamento e controle não gere resultados precisos. Nesses casos, métodos estatísticos não-experimentais são essenciais para contornar essas questões.

O texto a seguir se organiza da seguinte maneira: a primeira parte expõe a lição que fundamenta a necessidade de métodos robustos de avaliação de impacto; em seguida, discute-se a importância da situação contrafactual para avaliações de impacto e se examinam estratégias (boas e ruins) para construir grupos de comparação. Por fim, detalha-se dois métodos não-experimentais – Diferenças-em-Diferenças (DD) e Regressão Descontinuada (RDD), discutindo suas características e vantagens em relação ao que se chama de métodos “ingênuos” de construção do grupo de comparação (situação contrafactual).

## AVALIAÇÃO DE IMPACTO: FATO CONCRETO – CORRELAÇÃO NÃO É CAUSALIDADE

A primeira lição fundamental aprendida por um analista de políticas públicas é que correlação não é causalidade. Como analista é preciso ter claro que aconselhar a formulação de políticas observando apenas relações não bem identificadas entre variáveis, considerando, por exemplo, correlações pouco estudadas entre elas, não é a melhor estratégia.

Para exemplificar essa questão, a Figura 1 expõe uma situação didática sobre a temática. Na figura é apresentada a evolução de duas variáveis, consumo de sorvetes e índice de ataques de tubarão, ao longo dos meses do ano. Vê-se, claramente, que essas duas variáveis possuem trajetórias bastante similares, o que nos indica que são duas variáveis correlacionadas. Entretanto, ao contrário do apresentado na mensagem de *warning*, não é possível afirmar que reduzir a quantidade de sorvetes vendida reduz o índice de ataques de tubarões. Tal análise não aponta nenhum sentido causal entre as variáveis.

Figura 1 – Correlação entre ataques de tubarão e vendas de sorvete



Fonte: Pinterest - <https://tinyurl.com/ycf3oofm>

É nesse sentido que técnicas de avaliação de políticas públicas são necessárias. O objetivo de uma avaliação de impacto é medir o efeito causal de intervenções sobre as variáveis de interesse. Para isso, o grande desafio é construir um excelente contrafactual, seja ele por método experimental ou por método não-experimental. O grupo de controle precisa ser bem construído para que mimitize o que teria acontecido com o grupo de tratados caso eles não tivessem recebido o tratamento (situação contrafactual). Precisamos comparar os resultados de grupos que possam efetivamente ser comparados (idênticos), de maneira que possamos argumentar que a única diferença entre eles é o tratamento.

Para se construir um bom contrafactual existem dois métodos: o experimental e o não-experimental. O método experimental constitui a melhor estratégia de avaliação, garantindo maior confiabilidade aos resultados ao controlar<sup>2</sup> como é definido o contrafactual. Mais especificamente, nele a definição dos grupos de tratamento e controle acontece através da utilização de algum método de aleatorização. É justamente essa capacidade que se tem de aleatorizar de alguma maneira (sorteio, ordem de chegada, etc.) que garante a comparabilidade entre os grupos e a identificação do efeito causal no método experimental.

Já no caso dos métodos não-experimentais, é necessário utilizar modelos estatísticos (sustentados em hipóteses) para poder construir um grupo de comparação que possa ser sustentado como uma boa aproximação da situação contrafactual. Dois desses métodos serão discutidos em mais detalhes na próxima seção.

<sup>2</sup> Essa capacidade de controlar como a situação contrafactual é definida faz com esse método seja conhecido pela alcunha de experimento controlado.

## CONSTRUINDO UM BOM CONTRAFACTUAL

Como exposto na seção anterior, o grande desafio das análises de avaliação de impacto é a definição do contrafactual. Em termos conceituais, uma situação contrafactual representa aquilo que teria acontecido com a unidade de avaliação (seja ela pessoa, escola, município, etc.) na situação de ela não ter sido afetada pelo tratamento (projeto ou ação que está sendo avaliado). Entretanto, na prática, esse cenário comparativo perfeito é impossível de acontecer, considerando que não temos controle sobre o tempo, isto é, a habilidade de ir e voltar temporalmente para analisar diferentes possibilidades para a mesma unidade de avaliação. Diante dessa impossibilidade, o que analistas de política fazem para identificar ou construir um bom grupo de comparação é utilizar teoremas e modelos estatísticos.

No caso experimental, a aleatorização define os grupos de tratamento e controle (situação contrafactual) de forma que, na média, as características observáveis e não-observáveis dos grupos sejam idênticas. Com isso, ao se comparar as variáveis de resultado de ambos os grupos, estará se inferindo unicamente o efeito do programa ao qual o grupo de tratamento foi exposto, dado que a única diferença entre eles nesse cenário é o tratamento. Em outras palavras, pode-se afirmar que, no método experimental, a diferença entre as variáveis de resultado dos grupos tratado e controle é exclusivamente devida à intervenção (efeito causal).

No caso dos métodos não-experimentais, por sua vez, mesmo ao se utilizar de estratégias rigorosas para construir o contrafactual, é muito difícil sustentar o argumento de que se obteve grupos idênticos para a comparação. Nesses casos, possuem-se argumentos para defender que, na média, os grupos do estudo são similares entre si. Em comparação ao método experimental, a defesa da estratégia de identificação do efeito é mais difícil aqui, considerando que se demanda muito mais hipóteses para se estimar um efeito por um método não-experimental.

Para expor com clareza algumas das estratégias mais rigorosas não-experimentais, é importante discutir dois métodos de definição da situação contrafactual considerados ingênuos<sup>3</sup>. Tais métodos propõem as seguintes duas abordagens para a problemática:(i) comparar a situação do grupo dos tratados antes e depois da intervenção; ou então (ii) comparar a situação dos tratados com a de algum grupo não-tratado. Esses dois métodos, apesar de parecerem interessantes e adequados à questão, não são eficientes. Tais estratégias, como será discutido mais a frente, carregam vieses que impedem que se possa atribuir o efeito estimado exclusivamente ao programa avaliado.

A Figura 2 apresenta um resumo das estratégias abordadas aqui. Os quatro primeiros métodos serão discutidos em mais detalhes através de um estudo de caso na próxima seção. Os métodos ingênuos serão também abordados para se identificar onde recaem suas principais falhas.

---

<sup>3</sup>O termo ingênuo é utilizado no sentido de tais estratégias não fornecerem um bom contrafactual, mesmo que a princípio pareçam constituir boas abordagens. Tais procedimentos não são sequer considerados métodos de avaliação de impacto.

Figura 2 – Ferramentas para obter o contrafactual



Fonte: Elaboração própria.

## MÉTODOS INGÊNUOS: QUALQUER GRUPO DE COMPARAÇÃO NÃO É, NECESSARIAMENTE, UM BOM GRUPO DE COMPARAÇÃO

### *Estudo de Caso*

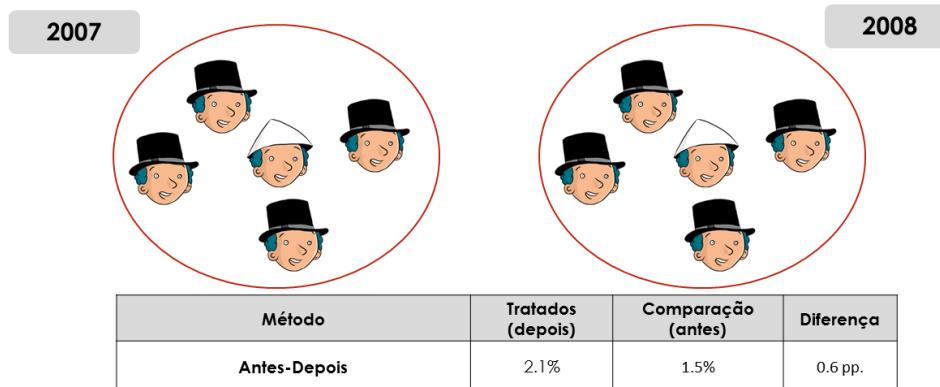
Em determinada cidade se observa que microempresas enfrentam restrições para acessar crédito. Diante disso, o gestor de política para o setor privado cria uma solução através de um programa de acesso a crédito. Por meio desse programa, o estado irá criar mecanismos para injetar capital nas microempresas que necessitem de crédito. O primeiro mecanismo criado estipula que todas as microempresas que se dirigirem à secretaria do estado e se cadastrarem no programa (preencham um formulário) terão à sua disposição R\$20.000,00 para investimento. Tal intervenção ocorreu no início de 2007. O objetivo final (variável final de interesse) desse gestor é aumentar a lucratividade dessas empresas. Após um ano da implementação da intervenção, o gestor deseja que o avaliador verifique se o programa alcançou o objetivo e estime qual foi a magnitude do impacto da intervenção.

A primeira ideia que ocorre ao analista é comparar a taxa de lucro das empresas participantes antes e depois de elas terem entrado no programa (Figura 3). Essa estratégia, à primeira vista, parece adequada para responder à pergunta do gestor, afinal as microempresas não tinham o tratamento antes e agora foram tratadas, logo comparar os dois tempos indicará o efeito do programa. Será?

Conforme mostrado na Figura 3, os indivíduos (microempresas) são os mesmos, observados no ano de 2008 e no ano de 2007 (exatamente antes da intervenção). Nessa análise, ao comparar os resultados de taxa de lucro antes e depois do programa, chega-se a um impacto de 0,6 p.p. sobre essa variável. Porém, será esse efeito resultado exclusivo da participação no programa? A resposta é não, esse não é um efeito factível e bem estimado. Temos, nesse caso, um problema conhecido como efeito tempo, isto é, inúmeras outras coisas (que não dizem respeito ao programa) podem ter afetado a variável de interesse, e o analista não as está levando em consideração. A economia, por exemplo, pode ter crescido, fazendo com que as empresas em geral tenham feito mais negócios e aumentado suas taxas de lucro.

Isso significa que o número estimado pode estar contaminado com outras coisas e não representa um bom indicativo do efeito do programa.

Figura 3 – Plano 1: comparar a taxa de lucro das empresas tratadas antes e depois do programa

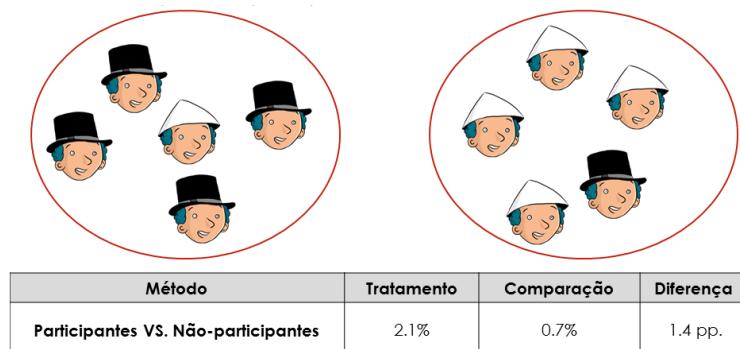


Fonte: Elaboração própria.

Sendo alertado para essa questão, o avaliador pensa, então, em outra maneira de estimar o impacto do programa. Agora ele imagina que comparar os participantes com os não-participantes seja uma forma eficiente de responder à pergunta. Ele tem à sua disposição uma base de dados que contém, além de todos os CNPJs das empresas tratadas, mais uma seleção de CNPJs de empresas que não foram tratadas. Quando aplicado esse método, comparando as médias de taxa de lucro das empresas que foram tratadas e das não que não foram, observa-se o resultado de 1,4 p.p. de diferença entre esses grupos.

Contudo, um observador cuidadoso percebe que os dois grupos comparados são bastante diferentes entre si (Figura 4). Considerando, por exemplo, que todas as características dos empresários estão resumidas nos chapéus que estão usando, na média, o grupo de tratamento teria um chapéu quase preto e o grupo que não recebeu o tratamento teria um chapéu quase branco. Isso significa que o grupo de controle não está fornecendo um bom contrafactual para o grupo de tratamento, e essa medida de impacto não é uma boa estimativa do efeito do programa. Essa situação é conhecida como problema de viés de seleção. Nesse caso, as empresas que se autosselecionaram para ser do grupo de tratamento (as características delas que as fizeram escolher ir até a secretaria se inscrever) e as que se autosselecionaram para ser do grupo controle (pois não foram até a secretaria) são muito diferentes entre si. Como por esse método essas diferenças não estão sendo levadas em consideração, a medida de efeito estimada não é acurada e não representa o impacto do programa.

Figura 4 – Plano 2: comparar a taxa de lucro das empresas que receberam crédito com a das que não receberam



Fonte: Elaboração própria.

É importante destacar que, além de características observáveis, podem existir ainda características não-observáveis que estejam afetando a variável de interesse e a decisão de ser tratado ou não. Nesse caso, por exemplo, uma característica não-observável que pode influenciar essas questões é a motivação dos empresários, indicador extremamente difícil de se medir. Nesse sentido, inclusive medidas não-observáveis podem estar impactando a estimativa de 1,4 p.p. sobre a taxa de lucro, tornando-a um resultado pouco confiável do impacto do programa.

Em resumo, os dois métodos acima tratados não conseguem estabelecer boas situações contrafactualis, de forma que não são válidos para análises de avaliação de impacto. A estratégia do antes e depois incorre no problema de outras questões poderem ter acontecido ao longo do tempo, enquanto que o método de comparar participantes e não-participantes carrega o problema de viés de seleção. Ambos os métodos conduzem a estimativas viesadas do efeito do programa.

Por fim, a análise dos métodos ingênuos demonstra uma lição importante para avaliadores de política: qualquer grupo de comparação não constitui necessariamente um bom grupo de comparação. É necessário utilizar métodos mais consistentes e rigorosos para se garantir avaliações de impacto efetivas. A seguir, dois desses métodos serão discutidos.

## **MÉTODOS NÃO-EXPERIMENTAIS (I): DIFERENÇAS-EM-DIFERENÇAS (DD)**

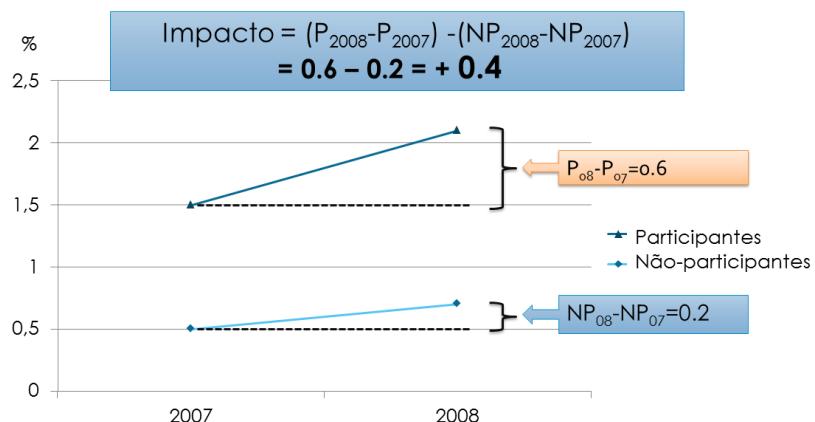
O método de Diferenças-em-Diferenças combina a dimensão temporal do antes-depois com a decisão de participação (participantes vs. não-participantes). Sobre algumas hipóteses extras, esse método consegue resolver tanto o problema do efeito tempo quanto o de viés de seleção.

Ele resolve o efeito tempo ao se assumir que o que afetou a variável de interesse ao longo do período afetou de forma “macro” ambos os grupos (tratado e não-tratado), de forma que, ao se tomar a diferença entre os grupos ( $T - C$ ) esse efeito some. O problema de viés de

seleção é contornado por esse método ao se assumir que as variáveis (observáveis ou não) que afetam a definição dos grupos de tratamento e controle (e a variável de interesse) são constantes ao longo do tempo. Partindo dessa hipótese, quando é calculada a diferença da variável de interesse entre o tempo um ( $t_1$ ) e o tempo zero ( $t_0$ ) a parcela advinda dessas variáveis some, tendo em vista que elas afetam de forma idêntica a variável de interesse ao longo do tempo.

Utilizando o estudo de caso da seção anterior e empregando o método de Diferenças-em-Diferenças, o que se calcula, em resumo, são as seguintes diferenças (Figura 5): (i) a diferença da taxa de lucro dos tratados antes e depois da data do tratamento (0,6 p.p.); (ii) a diferença da taxa de lucro do grupo de controle antes e depois da data do tratamento (0,2 p.p.); (iii) a diferença dos resultados obtidos nas diferenças calculadas em (i) e (ii) (0,4 p.p.). É desse processo de cálculo que advém o nome do método.

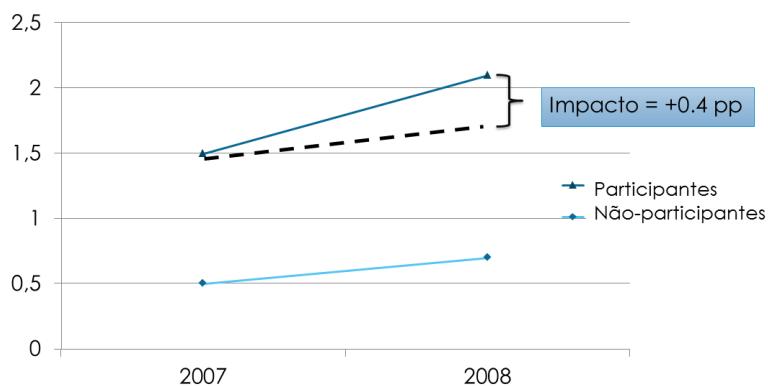
Figura 5 – Antes-depois + participantes X não-participantes: diferenças-em-diferenças



Fonte: Elaboração própria.

No entanto, além das hipóteses citadas para contornar os problemas de efeito tempo e de viés de seleção, existem ainda outras que precisam ser levadas em consideração para se usar esse método. Uma das suposições fundamentais se relaciona com o fato de, ao se aplicar o DD, estar se assumindo que os dois grupos observados são similares. Implicitamente, quando se utiliza o DD, se assume que a trajetória da variável de interesse observada para o grupo de controle seria a que se observaria para o grupo de tratamento na ausência do programa. Ou seja, assume-se que o grupo de controle representa um bom contrafactual para o grupo de tratamento. Nesse sentido, uma maneira de se entender o cálculo do efeito por esse método é imaginando que a trajetória dos não-tratados é deslocada até o nível da curva dos tratados e, então, a diferença entre essas duas trajetórias é realizada (Figura 6).

Figura 6 – Diferenças-em-Diferenças: deslocamento de trajetória

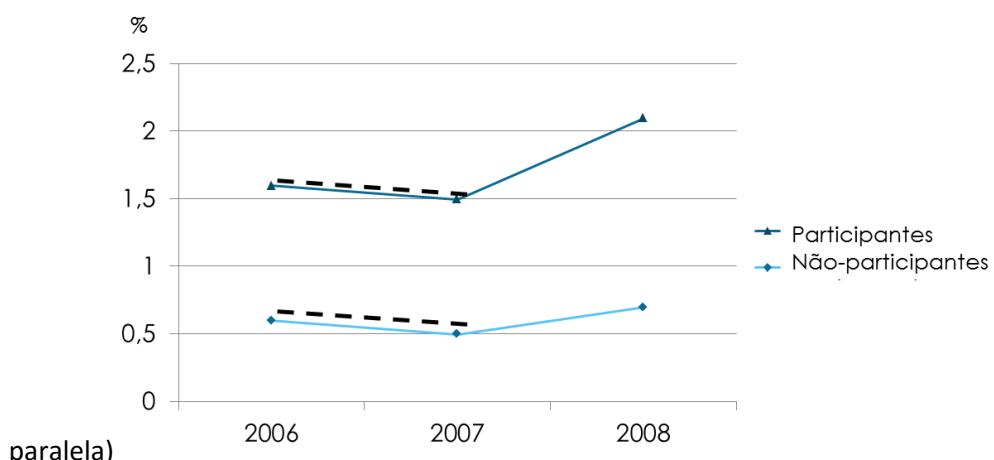


Fonte: elaboração própria.

Entretanto, para “deslocar” a trajetória do grupo de controle é necessário verificar a aderência dos dados ao que se chama de hipótese de tendência comum. Tal hipótese estabelece que o método de DD fornece uma estimativa confiável do impacto do programa se as trajetórias da variável de interesse para os grupos de tratamento e controle forem similares entre si antes da data do tratamento. Ou, de outra forma, imagina-se que, sendo os grupos bastante parecidos, a trajetória da variável de interesse fosse ser a mesma para os dois grupos caso o programa não existisse.

Para verificar a aderência das características do caso concreto a essa hipótese é fundamental ter acesso a dados históricos<sup>4</sup>. No exemplo (Figura 7), observa-se como eram as trajetórias de taxa de lucro tanto para participantes quanto para não-participantes antes de o programa acontecer (início de 2007). Nesse caso, as trajetórias são bem parecidas e, portanto, pode-se aplicar o método e assumir (defender) que o resultado expressará o efeito causal do programa.

Figura 7– Hipótese de tendência comum (ou paralela)



Fonte: elaboração própria.

<sup>4</sup>Ter acesso a dados de pelo menos um ponto anterior à intervenção é fundamental para a checagem da aderência da hipótese. Tal requerimento muitas vezes torna a aplicação do DD inviável (pouco defensável).

Por fim, cabe colocar que o método de DD pode ser ainda mais refinado se combinado com o método de *matching* ou pareamento. Esse método se baseia em um escore construído através de características observáveis para parear as observações de tratados e não-tratados, de forma a criar pares de análise similares e, portanto, comparáveis entre si. Dessa forma, a combinação do pareamento com o DD cria um grupo de controle eficiente (contrafactual defensável<sup>5</sup>) mesmo quando as trajetórias originais da variável de interesse não são completamente similares (hipótese de tendência comum pouco aderente).

## MÉTODOS NÃO-EXPERIMENTAIS (II): REGRESSÃO DESCONTÍNUA (RDD)

O método de regressão descontínua é aplicado às situações onde a regra definidora do grupo de tratamento (e, complementarmente, do de controle) é bem estabelecida e clara. Tal regra deve respeitar três características: (i) ser baseada em uma variável contínua; (ii) sua definição ser exógena (definida sem influência lógica de fatores relacionados à problemática); e deve ser exclusiva e não-manipulável (as unidades às quais se destina o programa não podem ser capazes de manipular suas informações para se adequarem à regra de elegibilidade).

Através de tal regra de elegibilidade, o agente formulador define que as observações de um dos lados desse limiar (acima ou abaixo do valor estabelecido) serão elegíveis ao tratamento, e as do lado oposto não serão. Resumidamente, a intuição do RDD é que as observações que estão bem próximas desse limiar (acima ou abaixo dele) são bastante parecidas entre si. Sendo assim, é possível se argumentar que as observações não-tratadas à margem dessa fronteira constituem um bom contrafactual para o grupo de tratamento. Importante destacar que, no RDD, toda a análise é realizada utilizando somente as observações ao redor desse threshold.

Para exemplificar, suponha que um gestor público deseje realizar uma intervenção sobre os municípios carentes em termos de educação. Para isso, o gestor define que a regra de elegibilidade será derivada da variável (contínua) Índice de Desenvolvimento Humano (IDH). Ele estabelece, então, que a intervenção será aplicada para os municípios que possuírem IDH inferior a 0,7 (*threshold*). Em outras palavras, aqui a variável definidora do tratamento é o IDH e o limiar dessa intervenção é 0,7, de forma que todas as observações que estiverem abaixo de 0,7 receberão a intervenção, e todas as que estiveram acima não receberão. A intuição da regressão descontínua é que os municípios que têm 0,699 (ou um pouco menos) de IDH, e os que têm 0,701 (ou um pouco mais) de IDH são extremamente parecidos entre si, podendo o grupo não-tratado ser considerado um bom contrafactual para o grupo de tratamento.

Para aplicar esse método ao estudo de caso anteriormente proposto, o formato em que o tratamento é oferecido precisa ser modificado. Nesse caso, o gestor decide alterar o

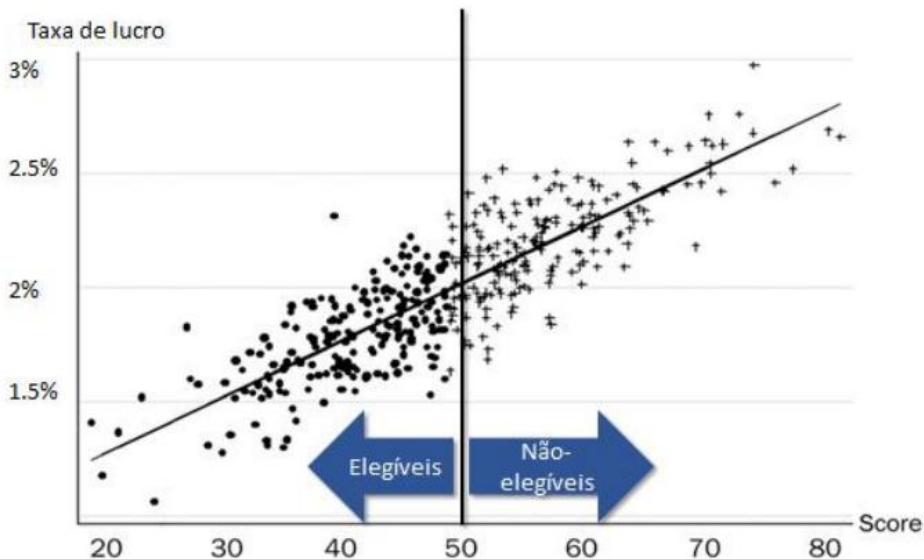
---

<sup>5</sup>Como em qualquer método, a defesa da aplicação do pareamento depende completamente da aderência de suas hipóteses ao caso concreto.

mecanismo para injetar capital nas microempresas. Agora ele decide ofertar o tratamento para as empresas conforme o escore de crédito delas. Tal escore foi calculado com base em características passadas da empresa, como idade, lucratividade, faturamento, número de empregados, etc. O escore varia de 0 a 100, onde 0 significa muita restrição a crédito, e 100 significa total acesso a crédito. O formulador da política determina discricionariamente que somente oferecerá o tratamento para as empresas que tiverem escore inferior a 50. Como o limiar escolhido é exógeno e exclusivo (somente esse programa usa), baseado numa variável contínua, e não pode ser manipulado (as empresas não conseguem trapacear dado que o escore foi calculado com base em dados passados), pode-se aplicar o RDD. A ideia, nesse caso, é comparar a taxa de lucro de empresas com escore um pouco abaixo de 50, elegíveis ao programa de crédito subsidiado, com empresas com escore um pouco acima de 50, não-elegíveis ao programa (Figura 7).

Como o nome do método alude, a estratégia de identificação no RDD se sustenta na existência de uma descontinuidade na variável de interesse após o tratamento no valor do limiar estabelecido. Entretanto, para se defender que essa descontinuidade (caso realmente verificada) decorre exclusivamente da intervenção, é importante verificar se ela não existia antes do tratamento. Se for observada alguma descontinuidade antes da intervenção, outros processos entre a variável definidora de tratamento (escore) e a variável de interesse (taxa de lucro) estão presentes e podem inviabilizar a defesa do método. No exemplo (Figura 8), percebe-se que a reta é bem suave e contínua em todo o espaço, inclusive no *threshold* 50.

Figura 8 – Regressão descontinuada: antes da intervenção

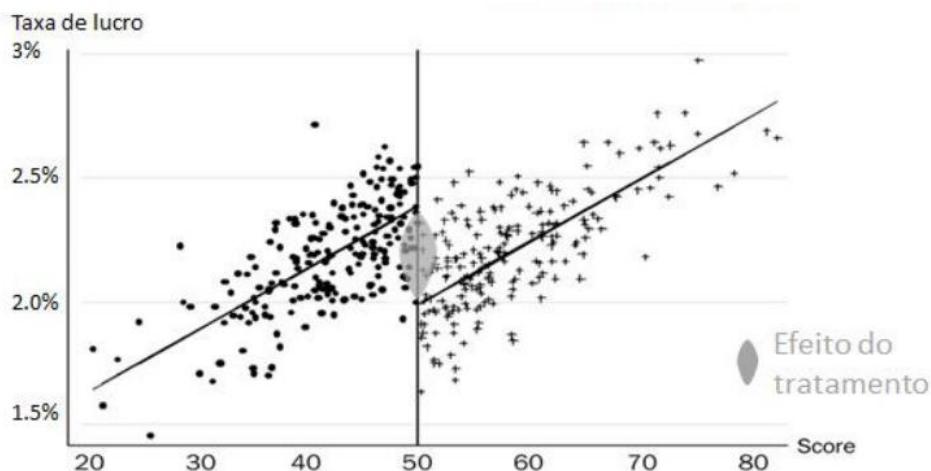


Fonte: Gertler et al. (2010).

Conforme apresentado na Figura 9, depois da intervenção se espera observar uma descontinuidade no valor de limiar, fator que permitirá avaliar o efeito do programa. É importante frisar que, no RDD, se estima um efeito local, pois somente as observações

localizadas próximas ao limiar são utilizadas na estimação. Tal definição de quão próximas do limiar as empresas devem estar para serem consideradas na estimação precisa ser avaliada com cautela. É importante garantir que as observações sejam parecidas entre si; entretanto, com o cuidado de não limitar a análise ao extremo, mantendo poucas observações e prejudicando o poder estatístico da estimativa.

Figura 9 – Regressão descontinuada: depois da intervenção



Fonte: Gertler et al. (2010).

Em resumo, o método de Regressão Descontinua é muito poderoso e é o que mais se aproxima de um método experimental em termos de poder de inferência. Porém, a elegibilidade da intervenção precisa respeitar algumas características. Precisa-se de uma variável contínua que sustente a elegibilidade para o programa. A linha de corte deve ser claramente definida, exclusiva e não-manipulável. Por fim, reitera-se que o impacto será válido apenas para aqueles indivíduos que estão próximos à linha de corte. Logo, deve se ter claro qual pergunta está se buscando responder e se a análise com as empresas próximas ao limiar propicia a resposta mais acurada.

## CONSIDERAÇÕES FINAIS

Como visto, métodos fracos geram resultados equivocados quando tratamos de Avaliação de Impacto. Não ter informação pode ser ruim; contudo, ter informação incorreta pode influenciar em sentido adverso a formulação de política pública. Resultados confiáveis só podem ser garantidos com a aplicação de métodos rigorosos, que possam ser defendidos à frente de qualquer audiência.

Em termos práticos, para decidir qual método é mais adequado ao caso concreto, é necessário analisar situação a situação. Nesse sentido, é fundamental conhecer em

profundidade o programa/intervenção que se pretende avaliar. São as regras do programa que definem em última instância qual o método mais interessante a ser utilizado. São elas que permitem argumentar a aderência do programa avaliado às hipóteses de cada método de avaliação. Muitas vezes, a complexidade dos programas faz com que seja necessário ser criativo para criar metodologias, requerendo que métodos sejam combinados a fim de encontrar resultados confiáveis, baseados em um contrafactual robusto. Cabe ao analista a responsabilidade de garantir que a identificação do efeito seja garantida através da utilização de métodos rigorosos.

## REFERÊNCIAS

GERTLER, Paul J. et al. *Impact Evaluation in Practice*. Washington, DC: Inter-American Development Bank and World Bank, 2016. Disponível em: < <https://openknowledge.worldbank.org/handle/10986/25030> >. Acesso em: 07 de maio de 2018.

## AVALIAÇÃO DE IMPACTO DE PROGRAMAS DE DESENVOLVIMENTO PARA O SETOR PRIVADO: A AGENDA DO BANCO MUNDIAL\*

Caio Piza<sup>1</sup>

**Resumo:** O texto aborda a agenda de avaliações de impacto conduzida pelo Banco Mundial na área de desenvolvimento do setor privado que, a despeito do montante de recursos investidos em programas voltados às empresas de pequeno e médio porte, é ainda limitada, especialmente nos países de renda média ou baixa. Sendo assim, além de destacar alguns trabalhos em andamento no Brasil, o texto busca discutir questões trabalhadas pelo Banco no passado e como elas têm sido tratadas na atualidade, a partir de uma agenda de pesquisa interativa, segundo a qual a avaliação deve ser útil para a operação dos programas e vice-versa.

**Palavras-chave:** Política pública. Avaliação de impacto. Programas de desenvolvimento do setor privado.

**Abstract:** The text addresses the impact evaluation agenda carried out by the World Bank (WB) on private sector development policies area. Despite the amount of resources invested in programs targeting small and medium-sized enterprises, evaluations of these investments are limited yet, especially in middle or low-income countries. Some researches in progress in Brazil are presented in this area. In addition, it is discussed how the World Bank dealt with the impact evaluation agenda in the past and how WB has been dealing with it at present. It is presented an interactive research agenda, according to which evaluation should be useful not only for assessing policy results, but also for planning its implementation.

**Keywords:** Public policy. Impact evaluation. Private Sector Development Program.

---

\* Texto elaborado a partir da transcrição da fala e da apresentação em slides de palestras proferida no Workshop Avaliação de Impacto de Políticas Públicas, realizado nos dias 29 e 30 de novembro de 2017, no Auditório da Secretaria de Planejamento, Governança e Gestão (SPGG-RS), com o apoio do Banco Mundial.

<sup>1</sup> Economista do Banco Internacional para Reconstrução e Desenvolvimento (BIRD).

## INTRODUÇÃO

O texto trata da agenda de avaliações de impacto realizadas pelo Banco Mundial na área de desenvolvimento do setor privado. Serão destacados os trabalhos em andamento no Brasil, bem como as questões que foram trabalhadas pelo Banco no passado e como elas são tratadas atualmente. Por fim, pontuam-se as questões mais importantes para essa agenda hoje.

### A AGENDA DE DESENVOLVIMENTO DO SETOR PRIVADO NO BANCO MUNDIAL

Um dos objetivos comuns da agenda de desenvolvimento do setor privado no Banco Mundial, incluindo o de inclusão financeira, é melhorar o funcionamento dos mercados, por exemplo, por meio da reforma regulatória; integrar ou facilitar o acesso a mercados e as cadeias de valor (global), sobretudo dos pequenos negócios a empresas de porte maior, na qualidade de fornecedores; apoiar o crescimento de microempresas de forma a aumentar a produtividade e gerar empregos.

A população-alvo das intervenções é, portanto, microempresas e empresas de pequeno e médio porte (MSMEs). A maior parte das empresas em países em desenvolvimento são pequenas e médias. Essas empresas são responsáveis por grande parte do emprego gerado na economia. Assim, se as MSMEs geram muitos empregos, qual a melhor forma de apoiá-las para que permaneçam ativas?

O Banco Mundial dedica muitos recursos para cobrir essa agenda, por meio de diversos projetos. Só o BIRD, entre 2006 e 2012, alocou praticamente US\$10 bi nesses projetos. A Corporação Financeira Internacional (IFC), que é o braço do setor privado do Banco, alocou mais US\$25 bi nesse período. Somam, portanto, quase US\$35 bi de dólares em programas de apoio a pequenos negócios nesse período. Na Europa, o *European Structural and Investments Funds* destinou mais de \$450 bilhões de euros para essa mesma agenda no período de 2014 a 2020 (EUROPEAN COMMISSION, 2017). No Brasil, por sua vez, 4,5% do PIB é alocado para programas de apoio ao setor privado (BANCO MUNDIAL, 2017).

Contudo, há pouca evidência quanto aos impactos ou resultados desses programas, se funcionam ou não e/ou se geram os resultados que os formuladores das políticas esperam (Cravo and Piza, 2016). Cabe, portanto, avaliar.

### PROGRAMAS DE APOIO AO SETOR PRIVADO: MUITOS RECURSOS, LIMITADAS EVIDÊNCIAS

Aqui não serão abordadas intervenções voltadas às ‘microempresas’ (conta-própria), para as quais já há bastante evidência produzida, inclusive pelo Banco Mundial. A evidência

sobre a eficácia de programas voltados às empresas de pequeno e médio porte (5, 10, 20 ou mais empregados) é ainda limitada, especialmente nos países de renda média ou baixa.

Um cuidado que se deve tomar diz respeito ao problema de validade interna da avaliação de impacto, ou seja, o efeito causal (exclusivo) de uma determinada intervenção. Mas também se busca, em alguns casos, saber quão generalizáveis são os resultados obtidos em uma avaliação. Ou seja, como formulador de política pretende-se entender se o que serviu para um grupo de empresas avaliado servirá para um grupo maior, caso se decida expandir o programa. Porém, isso nem sempre pode ser feito. A evidência que existe para países desenvolvidos não pode, simplesmente, ser transportada para países em desenvolvimento. O contexto é diferente, a qualidade de implementação dos programas é diferente, e essas questões fazem toda a diferença. Não se pode considerar simplesmente que a mesma intervenção funcionará em contextos distintos.

Integrei um grupo de trabalho que produziu uma revisão sistemática da literatura publicada entre os anos de 2000 e 2014, focando em países de renda baixa e média. A revisão levantou cerca de 40 estudos que utilizaram ou o método experimental, ou quase-experimental para avaliar o impacto de um programa específico de apoio a SMEs. Depois de três anos de análise desses estudos, conclui-se que existem pouquíssimas evidências produzidas! Ainda se sabe muito pouco sobre o que de fato funciona – o que deixa desapontados formuladores de políticas ou gestores à frente dos projetos. O que se sabe, sim, é que precisamos de mais avaliações para descobrir o que funciona e por quê.

Além dessa conclusão, o estudo verificou que, na média, programas de apoio para pequenos negócios funcionam. Mas analisados separadamente não se pode dizer muita coisa, já que há poucos estudos avaliando um mesmo tipo de intervenção. Por exemplo, não há um número elevado de estudos que estimaram o impacto de um programa baseado em crédito subsidiado, treinamento para trabalhador ou subsídio para a empresa por meio de salário subsidiado (programa em que o Estado arca com parte do salário de alguns empregados por um determinado período de tempo) que permita uma meta-análise de cada tipo de programa separadamente. É preciso mais estudos para examinar o impacto de uma determinada intervenção para concluir, no final, se a intervenção funciona ou não.

Além disso, as evidências existentes não são de boa qualidade. Há muitos estudos que utilizam ferramental de avaliação de impacto, mas não da melhor maneira possível. Quando se faz uma análise da qualidade ou da confiança do resultado que está sendo reportado, descobre-se que poucos estudos conseguiram controlar satisfatoriamente devido a possíveis fontes de viés.

Logo, é preciso fazer mais e fazer melhor: utilizar melhores dados, pensar melhor no problema e, eventualmente, fazer coleta de dados para responder perguntas que os dados administrativos não permitem responder. Somente assim poderão ser identificados os mecanismos através dos quais determinada intervenção afeta os resultados de interesse. Uma vez aprendido *como* o programa funciona para determinado grupo, pode-se eventualmente replicar essa experiência em outros lugares semelhantes.

## PROGRAMAS DE APOIO AO SETOR PRIVADO: O PROBLEMA

Empresas, independentemente do tamanho, são relutantes em adotar tecnologia que tem impacto privado, ou seja, que teria impacto positivo sobre o lucro (MCKENZIE; WOODRUFF 2014; CAMPOS *et al.* 2013). Gestão é um caso clássico. Estudos experimentais mostram que empresas que adotam melhores práticas de gestão têm um crescimento da lucratividade em 15 a 20% (BLOOM *et al.*, 2013; BRUHN *et al.* 2018). Porém, ao fazer um diagnóstico com pequenos negócios, médios negócios, grandes negócios, descobre-se que, entre as empresas de diferentes portes, há um grande número que não adota boas práticas de negócios (BLOOM *et al.*, 2012).

Diferentes são os gargalos que podem explicar esse comportamento. Pode ser falta de informação, de competitividade, de crédito, entre outros motivos. Identificar as razões, porém, não é simples.

Na prática, a estratégia padrão dos programas de apoio ao setor privado consiste em apenas focar no que é chamado de “lado da oferta”, ou seja, ações voltadas para os fatores de produção das empresas. Por exemplo, se a empresa não consegue crescer, então deve ser porque não tem capital ou acesso à mão de obra qualificada. Embora essa seja a prática mais comum, não se trata necessariamente da melhor. A avaliação do Cartão BNDES mostra isso (MACHADO *et al.*, 2011). Embora haja algum indício de que o Cartão BNDES teve efeito positivo sobre o emprego formal, a magnitude do efeito é muito baixa (menos de um trabalhador a mais, na média). Outra política comum é a de treinamento de funcionários. Quando a empresa treina um funcionário, ela não consegue, necessariamente, garantir que os benefícios daquele treinamento serão totalmente internalizados ou apropriados pela empresa, pois o funcionário pode mudar de empresa. A empresa acaba arcando com os custos, mas não, necessariamente, recebendo os benefícios. Isso faz com que as empresas acabem subinvestindo em funcionários. É um caso clássico de externalidade positiva, e a intuição é simples: se a empresa não puder extrair todo o benefício de um investimento, ela não vai investir a quantidade ótima para a sociedade. Programas focados no lado da oferta não parecem ser os mais eficazes em termos de geração de resultados. As avaliações de impacto que examinaram intervenções do ‘lado da oferta’ não encontraram grandes resultados. O que se pode fazer para melhorar esse cenário?

## ABORDAGEM INTERATIVA DE AVALIAÇÃO DE IMPACTO

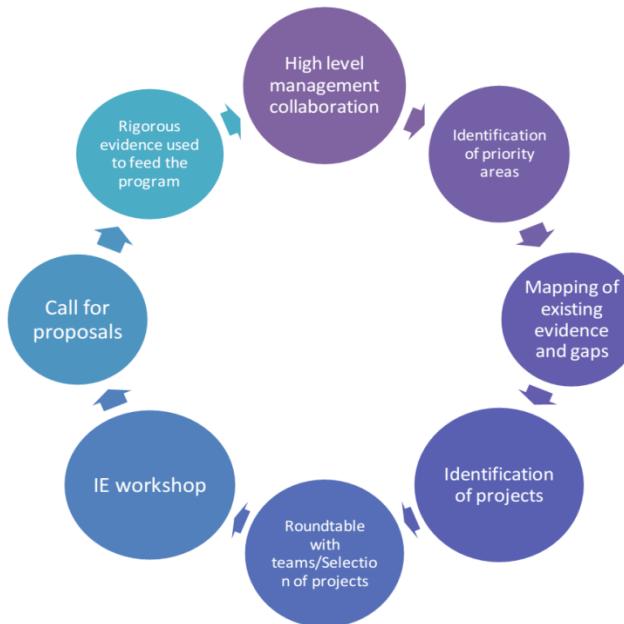
Para saber como desenhar operações ou programas baseados em evidências de alta qualidade deve-se, antes de mais nada, evitar o que já se sabe que *não funciona*. É preciso levantar informações sobre as evidências existentes sobre a política ou algo parecido com a política a ser implementada, para minimizar a chance de se incorrer no mesmo erro.

Nessa linha, a Unidade de Avaliação de Impacto do Departamento de Pesquisa do Banco Mundial (*Development Impact Evaluation – DIME*), propõe a abordagem de avaliação

seguinte. Realizamos um trabalho bem próximo e bastante colaborativo com o lado operacional do Banco e com os formuladores de política. Há intenso diálogo antes de se partir para a ação, a fim de compreender a intervenção, os passos que serão seguidos, o cronograma das atividades, as restrições e o orçamento. Uma vez que esses elementos são compreendidos, a equipe de avaliação, muitas vezes composta por membros do DIME e da instituição responsável pela intervenção, traça o plano de ação da avaliação de impacto: desenho da avaliação, como as intervenções serão implementadas, quais os protocolos que serão utilizados para verificação da qualidade da implementação e dos dados coletados, entre outros.

O modelo é baseado na abordagem interativa: a pesquisa deve ser útil para a operação e vice-versa. Isto é, as evidências existentes e geradas são utilizadas para informar a operação de maneira rigorosa, e o lado operacional, por sua vez, determina as áreas para as quais deseja obter evidências por meio de avaliações rigorosas e se mobiliza para incorporar avaliação (ões) de impacto em novas operações. Esse modelo é possível dada a proximidade e colaboração entre as equipes de pesquisa e operação. Não funciona exatamente assim em todos os setores do Banco, mas funciona muito bem, e já há dez anos, com o setor responsável pelas operações voltadas para o desenvolvimento do setor privado. A Figura 1 sintetiza essa abordagem.

Figura 1 – Abordagem interativa



Fonte: Elaboração própria, baseada no DIME Annual Report (WORLD BANK, 2018).

Essa abordagem interativa foi construída ao longo dos anos e utilizada com outros setores do Banco também, como agricultura, transportes e governança. A seguir, relata-se a construção dessa abordagem especificamente na área de apoio ao setor privado.

## APRENDIZADOS SOBRE AVALIAÇÃO DE IMPACTO DE PROGRAMAS DE APOIO AO SETOR PRIVADO

Em 2010, em Dakar, o DIME realizou um *workshop* internacional, reunindo equipes de vários países e diferentes projetos apoiados pelo Banco Mundial focados em políticas de apoio a pequenos negócios. Os projetos consistiam basicamente em apoio por meio de crédito subsidiado, que é uma prática comum, inclusive por parte do Banco até hoje. Esse *workshop* deu origem a várias avaliações de impacto desse tipo de programa, mas, curiosamente, quase todas as tentativas de avaliação fracassaram. Verificou-se, por exemplo, que muitas firmas não demandam crédito subsidiado (Campos *et al.* 2013). O crédito é oferecido, mas as empresas não elaboram a proposta requerida para a concessão do mesmo. Algumas vezes, nem mesmo manifestam interesse para obter aquele crédito subsidiado. Mesmo sendo quase uma doação (sem muita contrapartida), as empresas não se empenham para obter esse recurso. Para nós, essa baixa demanda dos pequenos negócios por crédito barato foi uma grande surpresa. De fato, após sete tentativas malsucedidas de avaliação de impacto experimental de programas de crédito subsidiado, descobrimos que o que o Banco Mundial considerava dinheiro gratuito para as empresas, na verdade, não era gratuito. Isso porque a necessidade de preencher diversos formulários para aplicar para o recurso era suficientemente custosa para inibir várias empresas de fazê-lo.

Muitas empresas não sabem elaborar uma proposta de negócio, não sabem de antemão, por exemplo, o que devem colocar numa proposta. As empresas não têm objetivo claro de o que desejam fazer, quais são suas metas e como pretendem atingi-las. Isso não está escrito em lugar nenhum. As firmas também não necessariamente sabem que não sabem. Há casos em que a empresa precisa contratar um consultor para ajudá-la a elaborar uma proposta, mas, às vezes, esse consultor não está disponível no mercado ou, se está, não necessariamente fará um bom trabalho.

Descobriu-se, portanto, que elaborar bem uma proposta é um grande impedimento e dificulta muito uma adesão alta a esse tipo de intervenção. Não se pensa nessa dificuldade quando se formulam as intervenções. Esse *workshop* de avaliação produziu uma série de avaliações que não saíram do papel, mas ajudou muito a entendermos alguns dos gargalos que impediam uma alta adesão a esse tipo de operação que o Banco Mundial vem realizando há cerca de 40 anos. A partir de então, o Banco tem desenhado projetos de forma diferente. Não todos, é verdade, mas houve uma internalização desse conhecimento.

Evento semelhante foi realizado em 2011, no Rio de Janeiro, dessa vez com foco em projetos de educação financeira. Trata-se novamente de projetos do lado da oferta, pois o pressuposto era que seria possível alterar o comportamento das pessoas com relação às decisões de consumo e poupança a partir de treinamento em educação financeira. Treinamento para funcionários é extremamente difícil de ser feito, pois há baixa procura e alta taxa de desistência. Quando a equipe de trabalho encontra efeitos positivos, no geral, são de curto prazo. O funcionário recebe treinamento, sua produtividade no trabalho aumenta um pouco, mas depois volta para o nível anterior. Isso pode ser por uma questão organizacional

ou por não ter incentivo para continuar produtivo, porque não há nenhum reconhecimento por aquilo. Nesse sentido, a resposta do funcionário pode ser racional.

Contudo, dentre os aprendizados desse evento, conseguiu-se demonstrar que, se forem bem desenhados, bem intensivos e bem implementados, programas de educação financeira podem ser muito efetivos, apesar de a literatura até aquele momento apontar o contrário. Não se pode concluir a partir dos resultados de uma avaliação que qualquer programa (de educação financeira) funciona ou não funciona. Alguns programas testados podem não ter apresentado impacto, mas talvez se desenhados um pouco diferentemente, e se mais bem implementados, poderiam ter sido eficazes.

Em um terceiro evento, realizado em 2012 em Paris, o Banco Mundial voltou os olhares para o ambiente de negócio. Diz-se que um dos obstáculos para o crescimento das empresas é um ambiente desfavorável de negócios onde operam. Diante disso, várias reformas voltadas para aumentar a geração de empresas formais ou a formalização de negócios informais foram apoiadas pelo Banco. Por exemplo, em Minas Gerais, chegou-se a testar o efeito de enviar consultores para ajudar empresas com o processo de formalização, redução no custo de formalização ou simplificação de impostos. Apesar dessas medidas, o impacto na formalização foi zero, com exceção de pouquíssimos casos (Andrade et al. 2016).

Esses três eventos revelaram que as evidências existentes eram desencorajadoras. Diante disso, o Banco Mundial iniciou uma agenda diferente, dessa vez relacionada à focalização das intervenções. Ou seja, uma discussão sobre o público-alvo dos programas.

## PROGRAMAS DE APOIO AO SETOR PRIVADO: O PROBLEMA DO PÚBLICO-ALVO

A Figura 2 mostra as trajetórias de crescimento do faturamento de dois tipos hipotéticos de empresas. Observa-se que a firma A tem o faturamento crescente no tempo, mas ela já tem um nível maior do que a empresa B no ponto de partida. Talvez essa não seja a empresa que se deva apoiar, porque já vem crescendo e já é mais próspera de saída. Porém, tendemos a olhar para ela e supor que essa é um “campeão nacional”, ou seja, que a chance dessa empresa se beneficiar da intervenção é alta porque tem acesso, por exemplo, a melhores insumos ou possui uma melhor gestão.

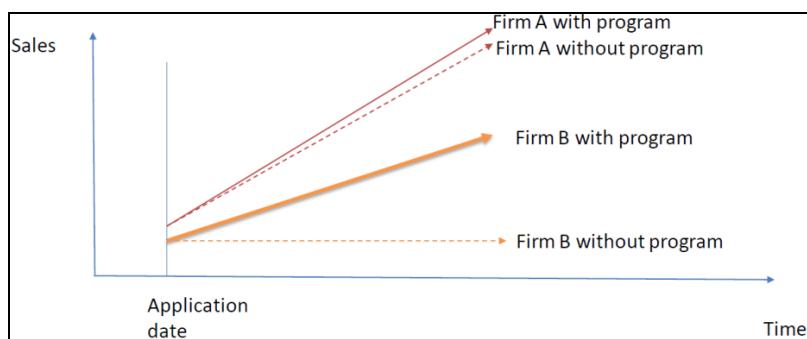
Suponha que se faça a intervenção nos dois grupos de empresas A e B a fim de identificar qual responderá melhor à intervenção. Ao comparar a firma A com o programa (linha sólida na Figura 2) com o seu cenário contrafactual sem o programa (linha pontilhada), observa-se que o programa teve efeito, porém pequeno. O programa contribuiu para deslocar pouco a trajetória.

Outra abordagem comum é a do apoio às empresas que mais precisam. Em vez de ajudar a empresa que tem o nível maior de faturamento, o apoio é para as empresas quem têm o menor nível, independentemente da trajetória. Nesse caso, não se observa a trajetória, apenas o nível inicial ('mais próspera' vs. 'menos próspera'). Nesse caso, se o programa tivesse sido direcionado para a empresa com nível menor de faturamento, o crescimento desta teria

sido enorme. Sem o programa, a empresa teria ficado estacionada. Ou seja, no caso hipotético, o crescimento só ocorreu por causa do programa.

Infelizmente é difícil saber de antemão qual grupo se beneficiará mais de uma determinada intervenção. O ideal, nesse caso, seria estratificar a amostra por tipos de empresa A e B e testar a intervenção nos diferentes subgrupos. Essa estratégia permitiria identificar para qual grupo a intervenção funcionou melhor. Com isso, antes de expandir o programa, o formulador de política identificaria qual grupo responde mais ao programa e expandiria o programa apenas para esse subgrupo. Embora essa estratégia seja totalmente viável na prática, requer uma amostra muito maior e, portanto, um orçamento maior.

**Figura 2 – Ilustração da trajetória de crescimento da firma no tempo: nível de crescimento e impacto da política**



Fonte: Mckenzie (2017b, p. 22).

Idealmente, deve-se fazer um piloto e testar para diferentes subgrupos antes de decidir como escalá-lo. Isso nem sempre é possível, mas com dados administrativos pode-se fazer alguns testes. Por exemplo, com o acesso à base de dados da Relação Anual de Informações Sociais (RAIS), pode-se subdividir a amostra em vários grupos e testar o impacto nos subgrupos.

O problema das abordagens tradicionais é que prestam pouca atenção à fase do diagnóstico. É preciso identificar quais são as restrições que precisam ser atacadas com as intervenções, de modo a aumentar a adesão ao programa e maximizar o seu impacto. Essas restrições não são conhecidas *a priori* e, por isso, é preciso testar várias hipóteses para mapear bem o problema que se quer atacar com uma determinada intervenção. Deve-se aprender quais são os mecanismos através dos quais a intervenção afeta os indicadores finais de resultados.

Por exemplo, tem-se o objetivo de melhorar a qualidade de educação. A intervenção visa melhorar a gestão da escola com um programa de treinamento voltado para o diretor. Com isso, busca-se avaliar se o programa afetará o aprendizado dos alunos, e o aprendizado será medido por meio de exames padronizados de matemática e português. Nesse caso, a primeira pergunta que deve ser feita é: por que um treinamento para diretor afetaria o aprendizado do aluno em matemática e português? Porque, se o diretor for melhor gestor,

terá maior interlocução com os professores, fazendo com que se engajem mais e, assim, faltam menos. Faltando menos, os alunos ficarão mais motivados e estudarão mais, apresentando melhores resultados. Vejam a complexidade da cadeia de resultados entre a intervenção e o aprendizado dos alunos. Isso ilustra bem quantas coisas devem acontecer para se conseguir afetar a variável de interesse, que é melhorar a nota dos alunos em matemática e português.

Há um caminho grande a ser percorrido para se chegar ao objetivo da intervenção, e esse caminho tem que ser mapeado. Deve-se identificar quais são essas etapas (elos causais) e, idealmente, medi-las. Exemplo: o diretor foi treinado? Foi. Teve mais engajamento com o professor? Teve, o que pode ser constatado pelos números de reuniões no mês, etc. Os professores estão preparando aulas de forma diferente? Sim/Não. Houve conversa com eles? Sim/Não. Nesse ponto, portanto, uma avaliação qualitativa pode ajudar. Deve-se conversar com os professores para saber o que está acontecendo, se estão faltando menos e por quê. Para além do diálogo, há dados administrativos que podem mostrar a falta dos professores, e o que mais está acontecendo na escola. Essa estratégia de mapear alguns indicadores intermediários é muito importante para aprendermos se um determinado programa funciona ou não e como.

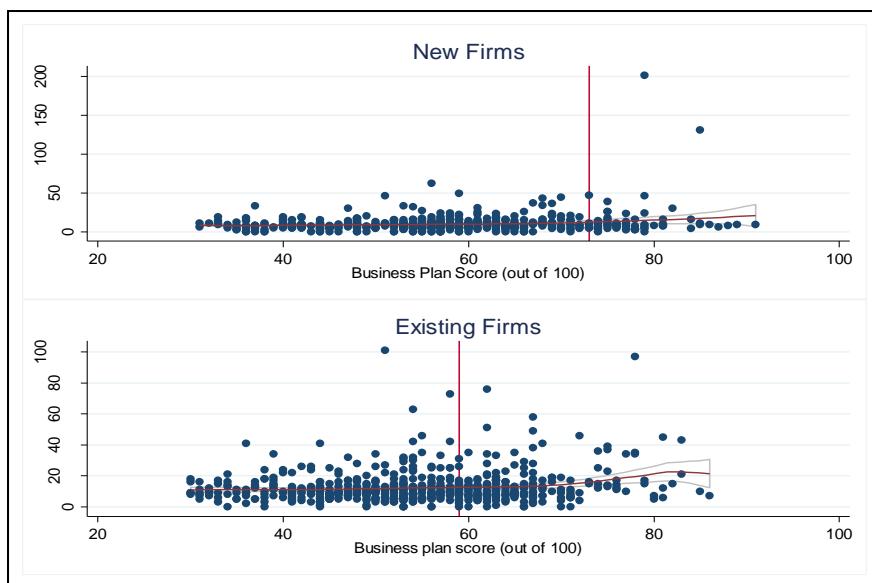
## AGENDA ATUAL: ABORDAGEM HOLÍSTICA

Atualmente, o DIME está focando a agenda de avaliação de impacto de intervenções voltadas para o desenvolvimento do setor privado em basicamente três áreas. Em primeiro lugar, estamos testando diferentes intervenções para tentar aumentar o acesso de pequenos negócios ao mercado. Quer-se aumentar o acesso a mercado e também promover a difusão de tecnologias e de conhecimento, produzindo externalidades de rede, por exemplo. Não é uma tarefa fácil, mas alguns experimentos estão sendo testados e, em breve, poderemos disseminar os resultados. Em segundo lugar, estamos preocupados com a focalização dos programas. Será que conseguimos identificar *a priori* as empresas que são mais prováveis de se beneficiarem de uma determinada intervenção? Para tanto, em vez de experimentos, tem-se usado técnicas de *machine learning* e de *big data* para analisar grandes bases de dados administrativos a fim de se realizar previsões acerca de quais subgrupos são mais prováveis de se beneficiarem de uma intervenção.

Vou utilizar uma avaliação de impacto experimental realizada pelo economista do Banco Mundial David McKenzie. A avaliação foi realizada na Nigéria e foi estruturada em torno de um programa apoiado pelo Banco Mundial que ofereceu US\$50 mil para as 500 empresas que apresentassem as propostas de negócio mais promissoras. Para isso, as empresas interessadas no prêmio tiveram de submeter uma proposta de negócio, detalhando a importância do recurso para o desenvolvimento do projeto. Alguns especialistas deram notas para cada uma das propostas submetidas. Foram identificadas aproximadamente 750 propostas promissoras, de 2.500 submetidas. Dentre essas 750, foram selecionadas 500 via sorteio para receberem o recurso. Com base nos dados, conseguiu-se prever, por exemplo, a

probabilidade de o projeto sair do papel e quem tinha maior chance de permanecer no mercado (a taxa de sobrevivência). Contudo, não havia como prever o impacto que o negócio financiado teria sobre o emprego. A empresa poderia sobreviver, mas não necessariamente crescer. Isso demonstra como é importante experimentar (McKenzie, 2017a). As Figuras 3 e 4 ilustram alguns dos resultados desse trabalho.

Figura 3 – *Nigerian business plan intervention*: impacto positivo na criação e *performance* da empresa, mas limitado impacto no emprego

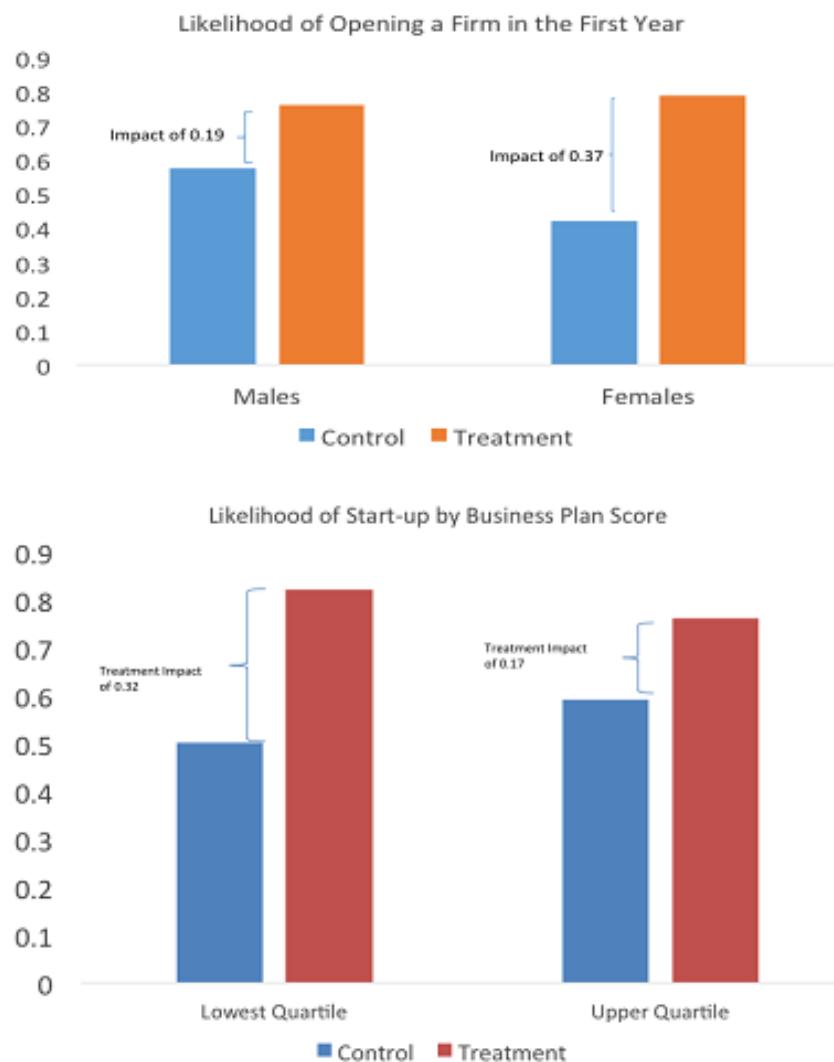


Fonte: McKenzie, 2017a.

A Figura 4, a seguir, traz ao menos dois resultados com implicação direta para a focalização da intervenção. O primeiro diz respeito ao efeito diferenciado para homens e mulheres. O primeiro gráfico mostra a proporção de empresas lideradas por homens e mulheres na ausência da intervenção (grupo de controle). Alguém poderia concluir precipitadamente que a intervenção deveria focar nas empresas encabeçadas por homens, pois tendem a abrir mais empresas naturalmente. Entretanto, o efeito do programa mostrou que as mulheres não abrem tantas empresas como os homens justamente porque não possuem capital. Com acesso ao capital, a probabilidade de abrir uma empresa é idêntica para homens e mulheres. Nesse caso, as mulheres se beneficiaram mais da intervenção.

O segundo ponto refere-se ao efeito do programa considerando a qualidade da proposta submetida. Os empresários com as piores propostas dentre as aprovadas foram os que experimentaram maior efeito do programa. As empresas que submeteram boas propostas talvez não precisassem tanto do capital quanto as piores.

Figura 4 – *Nigerian business plan intervention*: probabilidade de abertura de empresas



Fonte: McKenzie, 2017a.

Por fim, a terceira agenda visa melhorar o ambiente de negócio ou o ambiente regulatório das empresas. Nessa área, fazer experimentos pode ser mais desafiador, pois, muitas vezes, a intervenção envolve alguma reforma que afeta todos ao mesmo tempo. Há formas não-experimentais de mensuração do impacto desse tipo de intervenção e, se a adesão às novas regras decorrentes da reforma for baixa, pode-se utilizar um desenho de encorajamento. Ao invés de sortear os grupos que receberão ou não uma intervenção, aqui se sorteia um grupo para receber um incentivo para aderir às novas regras. Para um exemplo como esse, consultem o experimento conduzido no estado de Minas Gerais, que tentou aumentar a taxa de formalização dos pequenos negócios informais (ANDRADE *et al.*, 2016).

## CONSIDERAÇÕES FINAIS

O texto procurou apresentar a atual agenda de trabalho do Banco Mundial na área de desenvolvimento do setor privado, bem como a atuação do DIME na realização de avaliações de impacto dessas medidas.

Buscou-se destacar que já se sabe bastante o que não funciona em termos de intervenção nessa área, mas ainda se está longe para apontar o que fazer e como fazer para se ter resultado. Há ainda um espaço grande para experimentação, que deve ser abraçada como agenda de pesquisa. Dessa forma, busca-se gerar conhecimento robusto, que pode afetar positivamente a tomada de decisão de parceiros. O conhecimento alimenta um grupo grande de pessoas, e há muitas pessoas que demandam esse conhecimento.

## REFERÊNCIAS

- ANDRADE, G. H.; BRUHN, M.; MCKENZIE, D. A Helping Hand or the Long Arm of the Law? Experimental Evidence on What Governments Can Do to Formalize Firms. *World Bank Economic Review*, 30(1), 2016, pp. 24-54.
- BANCO MUNDIAL *Um ajuste justo*: Análise da eficiência e equidade do gasto público no Brasil. Volume 1: Síntese. Brasília: Banco Mundial, 2017. Disponível em: <<http://www.worldbank.org/pt/country/brazil/publication/brazil-expenditure-review-report>>. Acesso: 15 jun. 2018.
- WORLD BANK. *Science for impact*: better evidence for better decisions. Washington, D.C.: World Bank, 2018. Disponível em: <<http://documents.worldbank.org/curated/en/857631525294854313/Science-for-impact-better-evidence-for-better-decisions>>. Acesso: 15 jun. 2017.
- BLOOM, N.; EIFERT, B.; MAHAJAN, A.; MCKENZIE, D.; ROBERTS, J. Does Management Matter? Evidence from India. *The Quarterly Journal of Economics*, 128(1), 2013, pp. 1-51.
- BLOOM, N., SADUN, R.; VAN REENEN, J. *The Organization of Firms Across Countries*. The Quarterly Journal of Economics, 127(4), 2012, pp. 1663-1705.
- BRUHN, M.; KARLAN, D.; SCHWARZ, A. *The Impact of Consulting Services on Small and Medium Enterprises*: Evidence from a Randomized Trial in Mexico. *Journal of Political Economy*, 126(2), 2019, pp. 635-687.
- CAMPOS, F., COVILLE, A., FERNANDES, A., GOLDSTEIN, M.; MCKENZIE, D. Learning from Experiments That Never Happened: Lessons from Trying to Conduct Randomized Evaluations of Matching Grants Programs in Africa. *World Bank Policy Research Paper*, No. 6296, The World Bank, 2013.

CRAVO, T. A.; PIZA, C. *The impact of business support services for small and medium enterprises on firm performance in low- and middle-income countries: a meta-analysis.* Policy Research Working Paper Series 7664, The World Bank, 2016.

EUROPEAN COMMISSION. *SME Competitiveness.* 2017. [on line]. Disponível em <[http://ec.europa.eu/regional\\_policy/en/policy/themes/sme-competitiveness](http://ec.europa.eu/regional_policy/en/policy/themes/sme-competitiveness)>. Acesso, 15 jun. 2017.

MACHADO, L.; 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* 36, Dezembro, 2011.

McKENZIE, D.; WOODRUFF, C. *Business Practices in Small Firms in Developing Countries.* Management Science, 63(9), 2016, pp. 2967-2981.

McKENZIE, D.; WOODRUFF, C. What are we learning from business training evaluations around the developing world?. *World Bank Research Observer*, 29(1), 2014, pp. 48-82.

McKENZIE, D. Identifying and Spurring High-Growth Entrepreneurship: Experimental Evidence from a Business Plan Competition. *American Economic Review*, vol. 107(8), 2017a, pp. 2278-2307.

McKENZIE, D. *Targeting and Promoting Businesses with High-Growth Potential: Research Trends and IE opportunities.* Washington, D.C.: World Bank, 2017b. Disponível em: <<http://pubdocs.worldbank.org/en/138981499806675628/Compel-Mexico-City-Workshop-Day-3-2-David-McKenzie-Targeting-and-promoting-businesses-with-high-growth-potential.pdf>>. Acesso: 15 jun. 2018.

## AVALIAÇÃO DE IMPACTO DO PROJETO EXTENSÃO PRODUTIVA E INOVAÇÃO – PEPI\*

Juliana Feliciati Hoffmann<sup>1</sup>

Lauren Lewis Xerxenevsky

**Resumo:** O texto aborda a avaliação de impacto do Projeto Extensão Produtiva e Inovação (PEPI), considerando o período de 2011 a 2014. Foram estimados escores de propensão e o método Diferenças-em-Diferenças foi utilizado para comparar as mudanças nos indicadores de resultado ao longo do tempo entre os grupos, além da análise de sobrevivência para comparar o tempo até o fechamento da empresa. Pode-se afirmar que o PEPI impactou positivamente o número de vínculos ativos das empresas, aumentando-os em 3,9% nas empresas participantes do Projeto. Considerando o faturamento, o projeto também apresentou um efeito positivo, com um incremento de 7,95% identificado somente nas empresas de menor faturamento. Para massa salarial e produtividade das empresas, entretanto, não houve impacto significativo do Projeto. Em relação à análise de sobrevivência das empresas, identificou-se que a probabilidade de fechamento das empresas foi reduzida em 69,2% para aquelas participantes do PEPI. Além disso, conforme aumenta o nível de intensidade tecnológica, maior é a redução da probabilidade de fechamento das empresas. Os resultados estão em concordância com a literatura sobre o tema e demonstram a importância do Projeto para as pequenas e médias empresas do Estado.

**Palavras-chave:** Avaliação de Impacto. Diferenças em Diferenças. Propensity Score Matching. Apoio a Pequenas e Médias Empresas. Rio Grande do Sul

**Abstract:** The subject of this article is the impact evaluation of the Productive Extension and Innovation Project (PEIP), during the years of 2011 to 2014. It has been esteemed propensity scores and the Difference-in-Difference method was employed to compare changes in result indicators through time and in between groups. Besides that, it was done a survival analysis to compare time until the shutdown of the company. PEIP has shown a positive impact in the number of employees in companies, rising in 3,9% in companies that participated in the project. Considering the turnover, the project also presented a positive effect, raising 7,95% in companies with small turnover. Nevertheless, participating in the project showed no significant impact in salary mass and companies productivity. Regarding the survival analysis of companies, it has been spotted that the probability of shutdown was reduced in 69,2% for those who joined the PEIP. Besides that, the more it raises the technological investment, the lower probability of shutdown the company. The results are accordingly with theory about this subject and show the importance of the project to small and medium companies of the State.

**Keywords:** Impact Evaluation. Differences-in-differences. Propensity Score Matching. Support to Small and Medium Companies. Rio Grande do Sul State.

---

\* Texto elaborado a partir da transcrição da fala e da apresentação em slides de palestra proferida por Juliana Feliciati Hoffmann no Workshop Avaliação de Impacto de Políticas Públicas, realizado nos dias 29 e 30 de novembro de 2017, no Auditório da Secretaria de Planejamento, Governança e Gestão (SPGG-RS), com o apoio do Banco Mundial.

<sup>1</sup> Coordenadora do Núcleo de Monitoramento e Avaliação de Políticas Públicas (NUMA) do Departamento de Planejamento Governamental da Secretaria de Planejamento, Governança e Gestão. Estatística, doutoranda em Epidemiologia. Analista de Planejamento, Orçamento e Gestão.

## INTRODUÇÃO

Esta publicação trata de uma avaliação desenvolvida pelo Núcleo de Monitoramento e Avaliação de Políticas Públicas, criado em 2016, no Departamento de Planejamento Governamental da Secretaria de Planejamento, Governança e Gestão.

O Projeto Extensão Produtiva e Inovação (PEPI), desenvolvido pela Secretaria de Desenvolvimento, Ciência e Tecnologia, foi estruturado em 2011 e começou sua atuação em 2012. O Projeto conta com financiamento do Banco Mundial, e a avaliação de impacto era um dos indicadores de avaliação do financiamento do Banco Mundial (PROREDES), sendo uma das exigências previstas no contrato de financiamento com o Estado.

O Projeto é executado por meio de parcerias firmadas com instituições universitárias ou tecnológicas que estruturam Núcleos de Extensão Produtiva e Inovação (NEPIs) para apoio às empresas, através de assistência técnica e de treinamento, com financiamento do Banco Mundial. Um dos seus objetivos é aumentar a eficiência e a competitividade das empresas com incremento da produção, do emprego e da renda.

Até 2014, período final dessa avaliação, já tinham sido realizados em torno de 4.000 atendimentos a empresas do Estado, através de 20 NEPIs, sendo que o Projeto estava presente em 331 dos 497 municípios do RS. Portanto, a grande maioria dos municípios estava recebendo esse tipo de incentivo.

## METODOLOGIA

O primeiro passo da avaliação consistiu na identificação de indicadores de resultados, definidos a partir do objetivo do programa. Sendo assim, a avaliação buscou verificar o impacto do Projeto Extensão Produtiva e Inovação sobre os indicadores relacionados ao objetivo: “aumentar a eficiência e a competitividade das empresas com incremento da produção, do emprego e da renda”. Nesse sentido, foram definidos e validados com a equipe técnica do PEPI cinco indicadores de resultado relacionados ao objetivo, considerando a produção, o emprego, a renda, a eficiência e a competitividade das empresas, conforme detalhado no quadro a seguir. Para cada uma dessas variáveis de resultado foi feita uma análise, utilizando os indicadores descritos abaixo.

Quadro 1 – Elementos, descrição, indicadores de resultado e fonte de dados

| Elementos              | Descrição   | Indicador   | Fonte              |
|------------------------|---|---|--------------------|
| <b>Emprego</b>         | Empregos  | Quantidade de Vínculos Ativos (In)                        | RAIS/MTPS          |
| <b>Renda</b>           | Remuneração dos fatores de produção (trabalho, capital <sup>1</sup> e terra)                              | Massa salarial (em R\$ - rem. Trabalho - In)              | RAIS/MTPS<br>SEFAZ |
| <b>Produção</b>        | Produtos (bens e serviços gerados) – volumes ou valores   | Faturamento (Receita Bruta de vendas (R\$) - In)          | SEFAZ              |
| <b>Eficiência</b>      | Relação entre os resultados obtidos e os recursos empregados. Produtividade (produção/insumos empregados) | Produtividade da Mão de Obra ((Receita/nº vínculos) – In) | RAIS/MTPS<br>SEFAZ |
| <b>Competitividade</b> | Capacidade da empresa de cumprir sua missão com mais êxito do que as outras competidoras                  | Sobrevivência das firmas                                  | RAIS/MTPS<br>SEFAZ |

Fonte: elaboração própria.

No início do Projeto não houve o planejamento de uma avaliação e, portanto, nenhum método experimental foi aplicado na etapa de seleção das empresas. Foram incluídas no Projeto todas as empresas que tinham interesse em participar do programa, sem uma separação de empresas do grupo tratamento e do grupo controle.

Assim, a metodologia utilizada é um método não experimental, uma análise que permite a criação de um grupo controle, identificando as empresas que são mais parecidas com aquelas que receberam o programa.

Todas as análises foram feitas utilizando-se dados secundários. Para emprego, a fonte de dados foi a Relação Anual de Informações Sociais (RAIS), do Ministério do Trabalho. Para renda medida pela massa salarial, foram utilizados dados da RAIS e também da Secretaria da Fazenda do Estado. Para produção utilizou-se o faturamento das empresas, também com dados da Secretaria da Fazenda do Estado. Para eficiência, foi utilizada a produtividade da mão de obra, também com dados da RAIS e da Secretaria da Fazenda. Para competitividade, foi feita uma análise de sobrevivência das firmas, que será detalhada mais adiante.

Ressalta-se a questão de sigilo em relação aos dados de faturamento das empresas, o que é sempre delicado e torna esse um dado de difícil acesso. Para essa avaliação, foi feito um acordo com a Secretaria da Fazenda do Estado, e a equipe da avaliação trabalhou *in loco*, levando o material para a SEFAZ, onde foram feitas todas as análises, saindo apenas com os resultados finais. Dessa forma, foi possível ter acesso aos dados individuais das empresas.

Após a definição dos indicadores de resultado, foi iniciada a estruturação da base de dados. Inicialmente, foi utilizada a base de dados do próprio projeto, repassada pela área

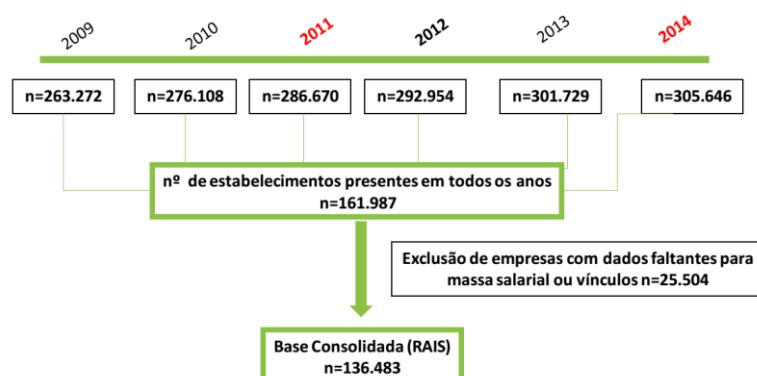
técnica responsável, contendo a listagem das empresas que estavam participando do programa. A partir, disso foi feita uma *linkagem* com a base de dados da RAIS, tanto de estabelecimentos, contendo o CNPJ das empresas e características de cada uma delas, como a área em que atua, o setor e a intensidade tecnológica, por exemplo. Foi utilizada também a base de dados de vínculos da RAIS, que contém a identificação de cada um dos empregados das empresas. Outra fonte de dados utilizada foi a base de dados da Secretaria da Fazenda do Estado. Todas essas informações foram reunidas, chegando em uma base de dados consolidada com um total de 3.302 empresas, utilizando seu CNPJ para identificar todas elas e poder fazer todos esses vínculos.

A Figura 1 apresenta a evolução do número de empresas ao longo do tempo, de 2009 a 2014. Observa-se que, em 2009, havia em torno de 260 mil e, em 2014, esse número evoluiu para 300 mil empresas na base de dados da RAIS. A partir de todas essas bases foram mantidas apenas as que permaneciam presentes em todo esse período da avaliação, resultando então em um tamanho de amostra de aproximadamente 161 mil empresas. Foram excluídas aquelas que não possuíam informação sobre massa salarial ou sobre vínculos. Assim, a base de dados das empresas do Estado, utilizada na avaliação, ficou com 136.483 empresas.

Na etapa seguinte, foram reunidas as duas bases de dados, do Projeto e da RAIS, consolidadas nos passos anteriores, conforme mostra a Figura 2. Das 3.300 empresas participantes do programa, apenas 1.705 foram identificadas na base de dados da RAIS, sendo que as demais não foram localizadas pelo CNPJ. Isso pode ter ocorrido em função de erro de digitação, por exemplo, ou qualquer outro problema de registro da informação do Projeto. Essa foi uma das dificuldades encontradas em função da avaliação ser não-experimental. Caso houvesse um planejamento no início da implementação do PEPI, certamente seria possível incluir todas as empresas que participaram, sem essa perda de aproximadamente 1.600.

Por fim, foi aplicado um filtro, mantendo somente as empresas que eram do setor industrial, por serem as mais similares àquelas que fizeram parte do programa, e também foi feita uma restrição à área de suporte comum, como será detalhado mais adiante, no método de escore de propensão.

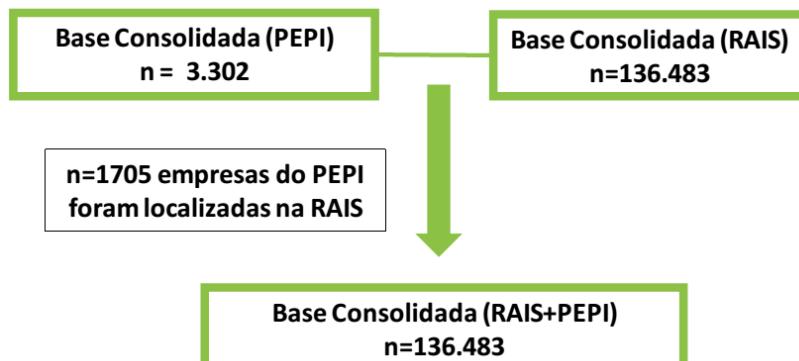
Figura 1 – Estruturação da base de dados para análise – RAIS. Estabelecimentos com informações de vínculos



Fonte: elaboração própria.

Utilizando a base de dados estruturada foi iniciada a estimativa dos escores de propensão. Essa metodologia tem como objetivo controlar o confundimento e o viés de seleção em estudos não aleatórios, buscando identificar um grupo que não participou do programa, que seja o mais parecido possível com aquele que participou. Através desse método, obtém-se a probabilidade de a empresa ser ou não alocada ao grupo tratamento, o que é chamado de propensão. Nessa análise, são consideradas várias características das empresas, de antes do início do programa em 2011, a fim de termos uma estimativa mais próxima possível dessa propensão. A partir dos escores, é realizado o pareamento ou a ponderação das empresas com base nesses valores, sendo que nessa avaliação utilizou-se a última.

Figura 2 – Estruturação da base de dados para análise – RAIS+PEPI



Fonte: elaboração própria.

O passo seguinte consistiu na estimativa do impacto do Projeto para cada um dos indicadores de resultados, através do método de diferenças-em-diferenças. Inicialmente, foi analisado o pressuposto das tendências paralelas, ou seja: espera-se que, antes do programa, o comportamento do grupo que recebeu e do grupo que não recebeu aquele programa seja muito parecido. Assim, busca-se garantir que, caso não existisse o programa, ambos os grupos teriam o mesmo comportamento.

Considerando os valores dos escores de propensão, receberam maior peso na análise as unidades do grupo controle com maior escore de propensão, ou seja, aquelas empresas que tinham uma probabilidade maior de participar do grupo tratamento. Além da estimativa do impacto através do método de diferenças-em-diferenças, também foi utilizada a análise de sobrevivência das empresas. Esse outro tipo de análise busca verificar o efeito do projeto sobre a probabilidade de sobrevivência das empresas, se havia ou não efeito do projeto no tempo em que aquela empresa permanecia aberta.

## RESULTADOS

### Empregos – Indicador: Número de vínculos ativos

A Tabela 1 resume o resultado encontrado para o impacto sobre o número de vínculos. Foi utilizada transformação logarítmica a fim de permitir a interpretação dos valores em percentual. Foi identificado um impacto de 3,9% do programa sobre o número de vínculos ( $P=0,085$ ). Assim, poderíamos dizer que o programa provoca um incremento de 3,9% no número de vínculos das empresas, sendo essa diferença estatisticamente significativa considerando 10% de significância.

Tabela 1 – Impacto do PEPI sobre o número de vínculos (ln) (n=17.786)<sup>1</sup>. Diferenças-em-diferenças ponderado pelos EP

| Variável                                       | Coeficiente | Valor-P |
|--|-------------|---------|
| programa_ano                                   | 0,039       | 0,085   |
| programa                                       | 0,136       | 0       |
| ano  | -0,045      | 0       |
| Massa salarial                                 | 0           | 0       |
| Idade (média)                                  | -0,01       | 0       |
| % sexo masculino                               | -0,081      | 0       |
| Quantidade de horas por trabalhador (média)    | 0,007       | 0,015   |
| Anos de Estudo (média)                         | -0,027      | 0       |
| Sociedade Empresária Limitada - Não            | 0           |         |
| Sociedade Empresária Limitada                  | 0,196       | 0       |
| Empresa Micro                                  | 0           |         |
| Empresa Pequena                                | 1,502       | 0       |
| Empresa Média                                  | 2,455       | 0       |
| Empresa Grande                                 | 1,831       | 0       |
| Participação no Simples - Não                  | 0           |         |
| Participação no Simples                        | -0,169      | 0       |
| Setores Calçados e Couro                       | 0,151       | 0,002   |
| Setor Esquadrias de Madeira                    | -0,112      | 0,007   |
| Setor Esquadrias de Metal                      | -0,214      | 0       |
| Setor de Vestuário                             | -0,078      | 0,04    |
| Setor Fabricação de Móveis                     | -0,031      | 0,312   |
| Baixa Intensidade Tecnológica industrial       | 0           |         |
| Média-Baixa Intensidade Tecnológica industrial | 0,02        | 0,353   |
| Média-Alta Intensidade Tecnológica industrial  | 0,042       | 0,022   |
| Alta Intensidade Tecnológica industrial        | 0,09        | 0,011   |
| Constante                                      | 2,048       | 0       |
| Valor-P (ajuste do modelo)                     | 0,0001      |         |
| R <sup>2</sup> ajustado                        | 0,727       |         |

<sup>1</sup> n= 17.786 empresas, consideradas nos anos de 2011 (antes) e 2014 (depois). Modelo ponderado pelos escores de propensão ajustado para as variáveis apresentadas na tabela e também para COREDES.

Para cada um dos indicadores de resultado foi realizada, adicionalmente, uma análise de robustez, a fim de verificar se utilizando outras metodologias as estimativas de impacto eram similares às que foram obtidas com a ponderação. Na Tabela 2, observa-se que houve convergência nos valores.

Tabela 2 – Análise de Robustez – Impacto do PEPI sobre o número de vínculos (ln) (n=17.786)<sup>1</sup>

|  | <b>Estimativa de impacto<sup>2</sup>(%)</b> | <b>Valor-P</b> | <b>R<sup>2</sup></b> |
|--|---|----------------|----------------------|
| <b>Modelo diff-diff ponderado (EP)</b>     | <b>3,9*</b>                                 | <b>0,085</b>   | <b>0,727</b>         |
| Modelo diff-diff com pareamento por Kernel | 4,0*  | 0,097          | ND                   |
| Modelo de regressão quantílica - P25       | 4,2   | 0,120          | 0,410                |
| Modelo de regressão quantílica - P50       | 4,0   | 0,351          | 0,470                |
| Modelo de regressão quantílica - P75       | 4,0   | 0,119          | 0,560                |

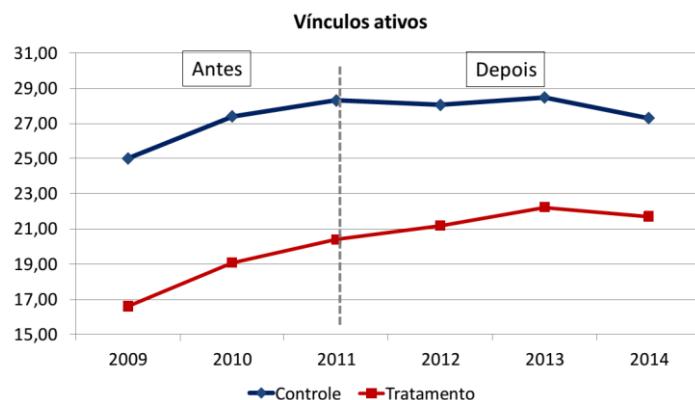
<sup>1</sup>n= 17.786 empresas, consideradas nos anos de 2011 (antes) e 2014 (depois).

<sup>2</sup> Modelos ajustados para as seguintes covariáveis: programa, ano, massa salarial, idade (média), % sexo masculino, quantidade de horas por trabalhador (média), anos de Estudo (média), Sociedade Empresária Limitada, tamanho de empresa, participação no Simples, COREDE, setores Calçados e Couro, setor esquadrias de madeira, setor esquadrias de metal, setor de vestuário, setor fabricação de móveis, intensidade tecnológica.

\* \*\*\*1% de significância, \*\* 5% de significância e \* 10% de significância

A Figura 3 resume a média anual do número de vínculos ativos (número de funcionários das empresas), para os grupos tratamento e controle, considerando-se o marco de início do programa em 2011. Observa-se que, antes de 2011, o grupo controle já apresentava uma média um pouco maior do que a do grupo tratamento. Entretanto, o objetivo desse gráfico é visualizar a tendência das retas, que se pode concluir serem paralelas e, portanto, a análise de diferenças-em-diferenças pode ser aplicada, pois esse pressuposto está satisfeito. Considerando o período após 2011, identifica-se que o grupo tratamento teve uma tendência crescente, enquanto o controle teve um número de vínculos médio mais constante ao longo dos anos.

Figura 3 – Tendências vínculos ativos (média) – RS, 2009 a 2014



Fonte dos dados: RAIS

### Renda – Indicador: Massa Salarial

Para o indicador massa salarial, também em logaritmo, observou-se um incremento de 4,7%, porém não estatisticamente significativo ( $P=0,111$ ). Assim, não é possível afirmar que esse resultado seria encontrado na população, e, portanto, não se pode afirmar que existe um impacto do PEPI sobre a massa salarial. A Tabela 3 apresenta os resultados da análise de robustez, e a Figura 4, a tendência da massa salarial média, de 2009 a 2014, bem como foi mostrado nas análises dos indicadores anteriores.

Tabela 3 - Impacto do PEPI sobre a massa salarial ( $\ln$ ) das empresas ( $n=17.786$ )<sup>1</sup>

Diferenças-em-diferenças ponderado pelos EP

| Variável                                       | Coeficiente | Valor-P |
|--|-------------|---------|
| programa_ano                                   | 0,047       | 0,111   |
| programa                                       | 0,181       | 0,000   |
| ano  | 0,285       | 0,000   |
| Vínculos ativos                                | 0,012       | 0,000   |
| % < 25 anos                                    | 0,118       | 0,146   |
| % 26 a 35 anos                                 | 0,273       | 0,001   |
| % 36 a 45 anos                                 | 0,19        | 0,032   |
| % 46 a 55 anos                                 | 0,095       | 0,310   |
| % sexo masculino                               | 0,304       | 0,000   |
| Quantidade de horas por trabalhador (média)    | 0,033       | 0,000   |
| % ensino superior                              | 0,952       | 0,000   |
| Sociedade Empresária Limitada - Não            | 0           |         |
| Sociedade Empresária Limitada                  | 0,27        | 0,000   |
| Empresa Micro                                  | 0           |         |
| Empresa Pequena                                | 1,506       | 0,000   |
| Empresa Média                                  | 1,397       | 0,000   |
| Empresa Grande                                 | -3,666      | 0,000   |
| Participação no Simples                        | -0,238      | 0,000   |
| Setor outros metais                            | 0,129       | 0,124   |
| Setor Esquadrias de Madeira                    | -0,181      | 0,001   |
| Setor Esquadrias de Metal                      | -0,219      | 0,000   |
| Setor de Vestuário                             | -0,089      | 0,051   |
| Setor Fabricação de Móveis                     | -0,059      | 0,106   |
| Baixa Intensidade Tecnológica industrial       | 0           |         |
| Alta Intensidade Tecnológica industrial        | 0,255       | 0,000   |
| Média-Alta Intensidade Tecnológica industrial  | 0,223       | 0,000   |
| Média-Baixa Intensidade Tecnológica industrial | 0,103       | 0,000   |
| Constante                                      | 6,023       | 0,000   |
| Valor-P (ajuste do modelo)                     | 0,0001      |         |
| R <sup>2</sup> ajustado                        | 0,7125      |         |

<sup>1</sup> n = 17.786 empresas, consideradas nos anos de 2011 (antes) e 2014 (depois). Modelo ponderado pelos escores de propensão ajustado para demais variáveis apresentadas, bem como COREDEs e setores da economia.

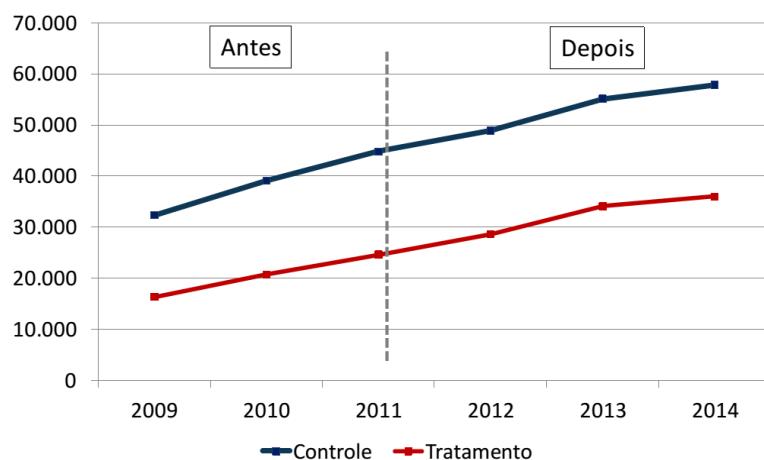
Tabela 4 – Análise de Robustez – Impacto do PEPI sobre a massa salarial (ln) das empresas  
(n=17.786)<sup>1</sup>

|  | <b>Estimativa de impacto<sup>2</sup> (%)</b> | <b>Valor-P</b> | <b>R<sup>2</sup>-ajustado</b> |
|--|--|----------------|-------------------------------|
| <b>Modelo diff-diff ponderado (EP)</b>     | <b>4,7</b>                                   | <b>0,111</b>   | <b>0,713</b>                  |
| Modelo diff-diff com pareamento por Kernel | 4,7*   | 0,089          | ND                            |
| Modelo de regressão quantílica - P25       | 5,7  | 0,194          | 0,400                         |
| Modelo de regressão quantílica - P50       | 4,7  | 0,184          | 0,470                         |
| Modelo de regressão quantílica - P75       | 2,1  | 0,510          | 0,550                         |

<sup>1</sup> n= 17.786 empresas, consideradas nos anos de 2011 (antes) e 2014 (depois).

<sup>2</sup> Modelos ajustados para as seguintes covariáveis: programa, ano, vínculos ativos, faixas etárias (%), % sexo masculino, quantidade de horas por trabalhador (média), % ensino superior, Sociedade Empresária Limitada, tamanho de empresa, COREDE, setor outros metais, setor esquadrias de madeira, setor esquadrias de metal, setor de vestuário, setor fabricação de móveis, intensidade tecnológica, setores de atividade econômica. \*\*\* 1% de significância, \*\* 5% de significância e \* 10% de significância

Figura 4 – Tendências massa salarial média – RS, 2009 a 2014



Fonte dos dados: RAIS

### Produção – Indicador: Faturamento

Para o faturamento das empresas, foi identificado um incremento de 5%, porém não significativo ( $P=0,214$ ). Assim, não se pode considerar que houve impacto no faturamento das empresas que participaram do PEPI.

Tabela 5 – Impacto do PEPI sobre o faturamento ( $\ln$ ) das empresas (n=15.191) - Diferenças-em-diferenças ponderado pelos EP

| Variável                                       | Coeficiente <sup>2</sup> | Valor-P | IC 95%        |
|--|--------------------------|---------|---------------|
| programa_ano                                   | 0,050                    | 0,214   | (-0,03;0,13)  |
| programa                                       | 0,235                    | 0,000   | (0,18;0,29)   |
| ano  | 0,180                    | 0,000   | (0,14;0,22)   |
| Vínculos ativos                                | 0,011                    | 0,000   | (0,01;0,01)   |
| % < 25 anos                                    | 0,569                    | 0,000   | (0,34;0,8)    |
| % 26 a 35 anos                                 | 0,678                    | 0,000   | (0,45;0,91)   |
| % 36 anos a 45 anos                            | 0,451                    | 0,000   | (0,21;0,7)    |
| % 46 anos a 55 anos                            | 0,277                    | 0,041   | (0,01;0,54)   |
| % sexo masculino                               | 0,385                    | 0,000   | (0,31;0,45)   |
| Quantidade de horas por trabalhador (média)    | 0,012                    | 0,017   | (0;0,02)      |
| % Ciclo 2º EF                                  | -0,082                   | 0,071   | (-0,17;0,01)  |
| % Ensino Superior                              | 0,597                    | 0,001   | (0,24;0,95)   |
| Sociedade Empresária Limitada – Não            | 0                        |         |               |
| Sociedade Empresária Limitada                  | 0,431                    | 0,000   | (0,38;0,48)   |
| Participação no Simples – Não                  | 0                        |         |               |
| Participação no Simples                        | -1,171                   | 0,000   | (-1,23;-1,11) |
| Setores Outros Metais                          | 0,237                    | 0,005   | (0,07;0,4)    |
| Setor Esquadrias de Madeira                    | -0,374                   | 0,000   | (-0,49;-0,26) |
| Setor Esquadrias de Metal                      | -0,269                   | 0,000   | (-0,39;-0,14) |
| Setor de Vestuário                             | -0,158                   | 0,003   | (-0,26;-0,05) |
| Setor Fabricação de Móveis                     | -0,241                   | 0,000   | (-0,33;-0,15) |
| Baixa Intensidade Tecnológica industrial       | 0                        |         |               |
| Média-Baixa Intensidade Tecnológica industrial | 0,391                    | 0,000   | (0,31;0,47)   |
| Média-Alta Intensidade Tecnológica industrial  | 0,238                    | 0,000   | (0,17;0,3)    |
| Alta Intensidade Tecnológica industrial        | 0,382                    | 0,000   | (0,26;0,5)    |
| Constante                                      | 11,766                   | 0,000   | (11,25;12,28) |
| Valor-P (ajuste do modelo)                     | 0,0001                   |         |               |
| R <sup>2</sup> ajustado                        | 0,640                    |         |               |

Modelo ponderado pelos escores de propensão ajustado para demais variáveis apresentadas, bem como COREDES e tamanho da empresa.

A Tabela 6 apresenta a análise de robustez da variável faturamento. Observa-se que, utilizando o método de regressão quantílica, houve uma significância estatística e que a

estimativa de impacto foi grande, de 7,95%. Assim, é possível concluir que houve um efeito positivo do projeto, embora ele só ocorra no percentil 25, ou seja, entre as empresas com menor faturamento. Então, nessas empresas há um impacto, um incremento, de 7,95% no faturamento médio das empresas, enquanto nas outras, se olharmos para todas em conjunto, não se observa nenhum resultado significativo. Com isso, podemos dizer que temos um indicativo de qual o público-alvo que está sendo mais beneficiado pelo Programa.

Tabela 6 – Análise de Robustez – Impacto do PEPI sobre o faturamento (ln) das empresas<sup>1</sup>

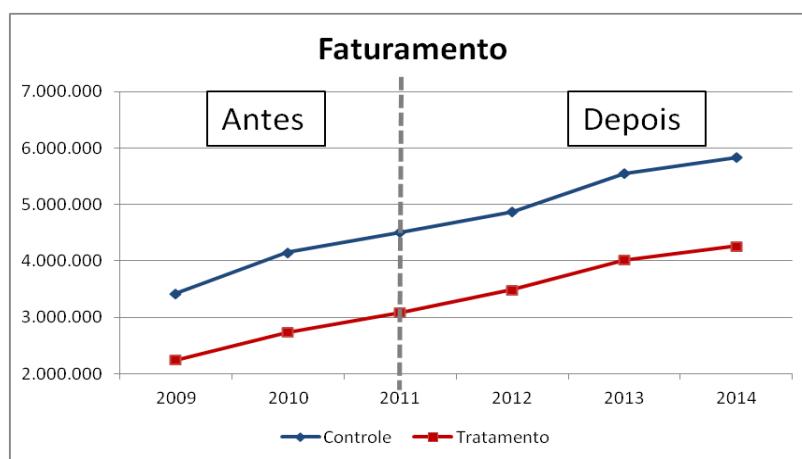
|   | Estimativa de impacto (%) <sup>2</sup> | Valor-P | R <sup>2</sup> |
|---|--|---------|----------------|
| Modelo DD ponderado (EP)                        | 5,0                                    | 0,214   | 0,640          |
| Modelo DD com pareamento por Kernel             | 6,1*                                   | 0,077   | ND             |
| Modelo de regressão quantílica por Kernel - P25 | 7,95***                                | 0,009   | 0,370          |
| Modelo de regressão quantílica por Kernel - P50 | 0,4                                    | 0,880   | 0,440          |
| Modelo de regressão quantílica por Kernel- P75  | -0,3                                   | 0,906   | 0,500          |

<sup>1</sup> n= 15.191 empresas, consideradas nos anos de 2011 (antes) e 2014 (depois).

<sup>2</sup> Modelos ajustados para as seguintes covariáveis: programa, ano, vínculos ativos, %<25 anos, %26 a 35 anos, %36 a 45 anos, %46 a 55 anos, % sexo masculino, quantidade de horas por trabalhador (média), %trab\_esc\_superior, %trab\_2 ciclo, Sociedade Empresária Limitada, tamanho de empresa, participação no Simples, COREDE, setores outros metais, setor esquadrias de madeira, setor esquadrias de metal, setor de vestuário, setor fabricação de móveis, intensidade tecnológica. EP: Escores de Propensão. DD: Diferenças-em-Diferenças. ND: Não disponível. \*\*\* 1% de significância, \*\* 5% de significância e \* 10% de significância

A Figura 5 mostra a tendência do faturamento de 2009 a 2014. Observa-se que existe tendência paralela dos dois grupos antes do início do programa.

Figura 5 – Tendências faturamento (médio) – RS, 2009 a 2014



Fonte dos dados: RAIS.

**Eficiência – Indicador: Produtividade do trabalho**

Em relação à produtividade do trabalho, conforme mostrado na Tabela 7, não é possível afirmar que existe qualquer tipo de impacto do PEPI (valor-P=0,900).

**Tabela 7 – Impacto do PEPI sobre a produtividade do trabalho (ln) das empresas (n=15.191) - Diferenças-em-diferenças ponderado pelos EP**

| Variável                                       | Coeficiente <sup>2</sup> | Valor-P | IC 95%        |
|--|--------------------------|---------|---------------|
| programa_ano                                   | 0,004                    | 0,900   | (-0,06;0,07)  |
| programa                                       | 0,091                    | 0,000   | (0,04;0,14)   |
| ano  | 0,210                    | 0,000   | (0,17;0,25)   |
| Vínculos ativos                                | -0,002                   | 0,000   | (0;0)         |
| % 26 a 35 anos                                 | 0,165                    | 0,000   | (0,08;0,25)   |
| % 56 anos a 65 anos                            | -0,249                   | 0,006   | (-0,43;-0,07) |
| % sexo masculino                               | 0,427                    | 0,000   | (0,37;0,49)   |
| % deficiência                                  | 0,722                    | 0,000   | (0,38;1,07)   |
| % Ensino Médio                                 | 0,549                    | 0,000   | (0,33;0,77)   |
| % Ciclo 1º EF                                  | 0,394                    | 0,001   | (0,16;0,63)   |
| % Ciclo 2º EF                                  | 0,330                    | 0,004   | (0,11;0,55)   |
| % Ensino Superior                              | 0,900                    | 0,000   | (0,54;1,26)   |
| Sociedade Empresária Limitada – Não            | 0                        |         |               |
| Sociedade Empresária Limitada                  | 0,194                    | 0,000   | (0,15;0,24)   |
| Participação no Simples – Não                  | 0                        |         |               |
| Participação no Simples                        | -1,012                   | 0,000   | (-1,07;-0,96) |
| Setores Outros Metais                          | 0,305                    | 0,004   | (0,1;0,51)    |
| Setores Calçados e Couro                       | -0,121                   | 0,163   | (-0,29;0,05)  |
| Setor Esquadrias de Madeira                    | -0,246                   | 0,000   | (-0,34;-0,15) |
| Setor de Vestuário                             | -0,069                   | 0,152   | (-0,16;0,03)  |
| Setor Fabricação de Móveis                     | -0,199                   | 0,000   | (-0,26;-0,13) |
| Baixa Intensidade Tecnológica industrial       | 0                        |         | (0;0)         |
| Média-Baixa Intensidade Tecnológica industrial | 0,270                    | 0,000   | (0,16;0,38)   |
| Média-Alta Intensidade Tecnológica industrial  | 0,179                    | 0,000   | (0,12;0,23)   |
| Alta Intensidade Tecnológica industrial        | 0,270                    | 0,000   | (0,16;0,38)   |
| Constante                                      | 10,891                   | 0,000   | (10,57;11,21) |
| Valor-P (ajuste do modelo)                     | 0,0001                   |         |               |
| R <sup>2</sup> ajustado                        | 0,329                    |         |               |

Modelo ponderado pelos escores de propensão ajustado para as variáveis apresentadas na Tabela 5 e também para COREDEs e tamanho da empresa.

Tabela 8 – Análise de Robustez – Impacto do PEPI sobre a produtividade do trabalho (ln) das empresas<sup>1</sup>

|   | Estimativa de impacto | Valor-P      | R <sup>2</sup> |
|---|-----------------------|--------------|----------------|
| <b>Modelo de regressão linear ponderado</b> | <b>0,004</b>          | <b>0,900</b> | <b>0,329</b>   |
| Modelo DD com pareamento por Kernel         | 0,015                 | 0,649        | ND             |
| Modelo de regressão quantílica - P25        | 0,031                 | 0,131        | 0,190          |
| Modelo de regressão quantílica - P50        | 0,007                 | 0,734        | 0,230          |
| Modelo de regressão quantílica - P75        | -0,043                | 0,108        | 0,260          |

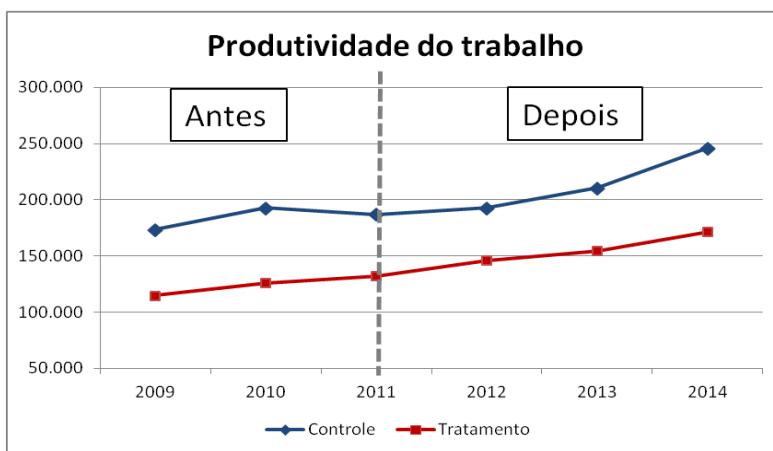
<sup>1</sup> n= 15.191 empresas, consideradas nos anos de 2011 (antes) e 2014 (depois).

<sup>2</sup> Modelos ajustados para as seguintes covariáveis: programa, ano, vínculos ativos, %26 a 35 anos, %56 a 65 anos, %deficientes, % sexo masculino, % de trabalhadores com ensino superior, % de trabalhadores com o 1º ciclo do ensino fundamental, % de trabalhadores com o 2º ciclo do ensino fundamental, Sociedade Empresária Limitada, tamanho de empresa, participação no SIMPLES, COREDE, setores outros metais, setores calçados e couro, setor esquadrias de madeira, setor de vestuário, setor fabricação de móveis e intensidade tecnológica.

EP: Escores de Propensão. DD: Diferenças-em-Diferenças. ND: Não disponível.

\*\*\* 1% de significância, \*\* 5% de significância e \* 10% de significância

Figura 6 – Tendências produtividade do trabalho (média) – RS, 2009 a 2014



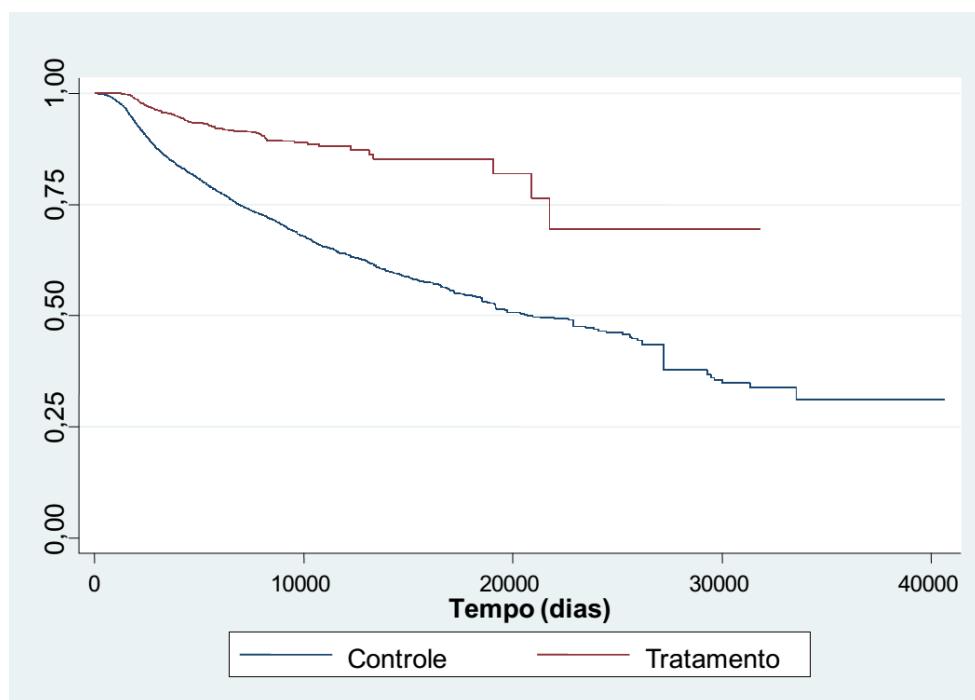
Fonte dos dados: RAIS

### Efetividade – Análise de Sobrevida das empresas

A análise realizada para avaliar efetividade difere das anteriormente apresentadas. Foi utilizada a análise de sobrevida das empresas, que considera o tempo (em dias) decorrido entre o cadastro delas na Secretaria da Fazenda e a data de fechamento das empresas, quando for o caso.

As curvas de Kaplan-Meier comparam o tempo até o fechamento das empresas entre os dois grupos, sendo que o eixo vertical mostra a probabilidade de sobrevida das empresas. Observa-se que os valores de probabilidade de sobrevida são sempre maiores naquelas empresas que participaram do programa. Com isso, temos um indicativo de que o programa pode estar aumentando a sobrevida dessas firmas.

Figura 7 – Curvas de Kaplan-Meier – comparação do tempo até o fechamento das empresas entre os dois grupos (probabilidade de sobrevida)



Fonte dos dados: Secretaria da Fazenda do Rio Grande do Sul (Sefaz-RS).

Tabela 9 – Análise de sobrevida – modelo de riscos proporcionais de Cox

| Variável  | HR    | Valor-P | IC 95%      |
|---|-------|---------|-------------|
| <b>Programa</b>                                 | 0,308 | 0,000   | (0,26;0,37) |
| <b>Baixa Intensidade Tecnológica industrial</b> |       |         |             |
| Média-Baixa Intensidade Tecnológica industrial  | 0,743 | 0,002   | (0,61;0,90) |
| Média-Alta Intensidade Tecnológica industrial   | 0,727 | 0,001   | (0,61;0,87) |
| Alta Intensidade Tecnológica industrial         | 0,499 | 0,002   | (0,32;0,78) |

A segunda etapa da análise de sobrevivência consistiu na regressão de Cox para verificar se realmente houve algum impacto na sobrevivência das empresas, considerando outras variáveis conjuntamente. Foi encontrado um resultado bem significativo ( $P<0,001$ ) para o *Hazard Ratio*, que é a medida que nos mostra qual é a redução da probabilidade de fechamento das empresas. Para obtenção da estimativa de redução da probabilidade de fechamento das empresas, usamos o conceito de probabilidade complementar e, assim, chegamos a uma redução de 69,2% (100-30,8).

Além disso, observou-se no modelo de regressão que, conforme vai aumentando a intensidade tecnológica, de média-baixa para média-alta, para alta intensidade, também vão diminuindo os valores do coeficiente, indicando que, conforme aumenta o nível de intensidade tecnológica, maior é a redução da probabilidade de fechamento das empresas. Ou seja, quanto maior a intensidade tecnológica, maior é o impacto do programa em relação ao tempo de sobrevivência das empresas (Tabela 9).

## CONSIDERAÇÕES FINAIS

Essa avaliação foi desenvolvida buscando contornar o problema do viés de autosseleção das empresas, uma vez que não foi feita uma seleção aleatória de quais empresas participaram ou não do programa. Além disso, por falta de dados primários, que consistiam apenas na base de dados dos registros administrativos do Programa, houve a necessidade de trabalhar com dados secundários que, em geral, não são de fácil obtenção pela necessidade de identificação das empresas por CNPJ.

Outra limitação decorrente da falta de planejamento prévio da avaliação foi a não localização de algumas empresas nas bases de dados secundários, em especial na base de dados da RAIS do Ministério do Trabalho, as quais acabaram não sendo incluídas nas análises.

Em relação à literatura existente sobre o tema, os resultados encontrados estão de acordo com o esperado. Alguns trabalhos relatam, entretanto, que seria necessário um período maior de tempo desde a intervenção para que seja observado algum impacto nesses indicadores. Essa poderia ser uma das justificativas para os resultados observados, que não foram significativos em alguns dos indicadores. Em função disso, a avaliação está sendo replicada utilizando dados da RAIS até 2016 e deverá ser concluída até o final de 2018.

Este texto resume parte de um relatório contendo todos os detalhes da revisão da literatura, da metodologia, dos resultados e de toda a discussão realizada com base nisso (SPGG, 2016). Outros trabalhos de avaliação estão em andamento no Núcleo de Monitoramento e Avaliação de Políticas Públicas da Secretaria de Planejamento, Governança e Gestão e devem ser concluídos ainda em 2018, entre eles: Programa de Oportunidades e Direitos Socioeducativo, da Secretaria de Desenvolvimento Social, Trabalho, Justiça e Direitos Humanos; Telemonitoramento de Veículos, da Secretaria da Modernização Administrativa e dos Recursos Humanos; Arranjos Produtivos Locais, da Secretaria de Desenvolvimento Econômico, Ciência e Tecnologia; e Comissão Interna de Prevenção a Acidentes e Violência Escolar (CIPAVE), da Secretaria da Educação.

## REFERÊNCIAS

SPGG. Secretaria de Planejamento, Governança e Gestão do Rio Grande do Sul. Avaliação de Impacto do Projeto Extensão Produtiva e Inovação. *Estudos DEPLAN*, edição nº 6, 2016. Disponível em: < <http://planejamento.rs.gov.br/upload/arquivos/201701/13103029-estudos-deplan-n6-final.pdf> >. Acesso em: 26 jun. 2018.

## O IMPERATIVO DA AVALIAÇÃO DAS POLÍTICAS DE APOIO AO SETOR PRIVADO NO BRASIL: MUITO GASTO, POUCO IMPACTO\*

Xavier Cirera<sup>1</sup>

**Resumo:** O texto tem como objetivo enfatizar a importância da avaliação de políticas públicas de apoio ao setor privado no Brasil. Muitos recursos são aplicados em programas no setor privado no país, tornando urgente a composição de uma agenda de avaliação para que se possa dialogar quanto à racionalização dessas políticas com base em evidências. Há poucas evidências quanto aos efeitos das políticas de apoio às empresas sobre a produtividade, os empregos e os preços ao consumidor, por exemplo. O efeito direto provável é, pelo contrário, regressivo na maioria dos programas, por favorecer grandes empresas já estabelecidas ou empregos formais já existentes. Nesse sentido, ressalta-se a necessidade de uma análise aprofundada desses programas, bem como de uma institucionalidade que apóie e conte com processos para difundir os resultados dessas avaliações.

**Palavras-chave:** Políticas Públicas. Avaliação de impacto. Apoio ao setor privado.

**Abstract:** The text aims at emphasizing the importance of public policy evaluation in order to support the private sector in Brazil. Many resources are invested in programs for private sector in the country. The implementation of an evaluation agenda in this area is urgent, in order to support a dialogue on the rationality of these policies, with evidence base. There is little evidence regarding the effects of business support policies on the productivity, employment and consumer prices, for example. On the contrary, the likely direct effect is regressive in most programs, once it favors large established firms and/or existing formal jobs. In this sense, the need for an in-depth analysis of these programs is highlighted, as well as the implementation of an institutionality that supports and counts on processes to disseminate the results of these evaluations.

**Keywords:** Public Policies. Impact evaluation. Support to the private sector.

---

\* Texto elaborado a partir da transcrição da fala e da apresentação em slides de palestras proferida no Workshop Avaliação de Impacto de Políticas Públicas, realizado nos dias 29 e 30 de novembro de 2017, no Auditório da Secretaria de Planejamento, Governança e Gestão (SPGG-RS), com o apoio do Banco Mundial.

<sup>1</sup> Economista Sênior do Grupo Banco Mundial.

## INTRODUÇÃO

Políticas e programas de apoio ao setor privado justificam-se com base no atendimento do interesse público. Visam tratar problemas de externalidades, de informação assimétrica e outras falhas de mercado, com vistas a promover o crescimento do emprego e aumento da produtividade. Em geral, essas intervenções apoiam certas empresas e atividades econômicas em detrimento de outras. Ao proporem uma espécie de reserva de mercado para alguns setores, concedendo-lhes uma proteção especial, impõem, porém, um constrangimento à concorrência. Sabemos que a concorrência tem um impacto positivo nas práticas de gestão, e restringir a concorrência pode ter impacto sobre a produtividade.

Portanto, ainda que possamos identificar as falhas de mercado, elas não são suficientes para justificar uma política. Muitas vezes, do ponto de vista da política pública, estamos resolvendo falhas de mercado, mas com muita pouca consideração sobre falhas de governo, ou seja, falhas na implementação. Essas falhas podem ocorrer mesmo quando se podem identificar as falhas de mercado de maneira clara, mas em muitos casos não estão identificadas no desenho da política.

Em virtude dessas questões, devemos ser muito conscientes quanto aos impactos esperados quando se desenha um instrumento. Devemos desenhar o Marco Lógico da intervenção, de modo que possamos pensar, ao menos conceitualmente, quais são os problemas que vamos atender. Mais que isso, devemos monitorar e avaliar a eficiência e a eficácia do gasto nessas intervenções, especialmente no contexto atual de crise fiscal e de problemas de financiamento. Portanto, duas coisas deve se ter em mente como muito importantes: a identificação da falha de mercado e o potencial para a falha do governo na implementação.

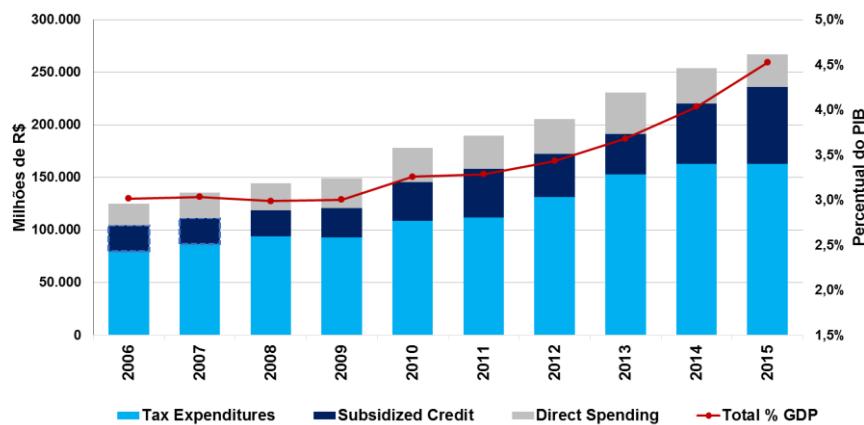
Este texto busca justamente enfatizar a importância da avaliação de políticas públicas de apoio ao setor privado no Brasil. Demonstra-se que são investidos muitos recursos nesses programas, sendo urgente compor uma agenda de avaliação e a respectiva institucionalidade, para que se possa dialogar quanto à racionalização dessas políticas, com base em evidências.

## O IMPERATIVO DA AVALIAÇÃO DAS POLÍTICAS DE APOIO AO SETOR PRIVADO NO BRASIL

Nesta seção, apresentaremos o resultado de análises da eficiência do gasto público em matéria de apoio ao setor privado, realizadas pelo escritório de Brasília do Banco Mundial. Foram analisados apenas os gastos com impactos fiscais diretos, classificados em três tipos: (1) gastos tributários e isenções fiscais; (2) crédito subsidiado; e (3) gastos gerais (apoio direto às empresas). O estudo abarcou apenas o nível federal e não incluiu: empresas estatais, políticas que afetam todas as empresas uniformemente (p. ex.: regulações, investimentos públicos em infraestrutura), tampouco políticas estaduais e municipais. Intuímos, porém, que os resultados obtidos em nível federal também se verificam em nível estadual.

A primeira conclusão a que chegamos é que o gasto com políticas de apoio ao setor privado no Brasil é muito elevado. Aumentou de 3% do PIB a 4,5% do PIB, entre 2006 e 2015 (Figura 1). Em comparação internacional, há muitos recursos que são gastos no apoio ao setor privado no Brasil, principalmente em relação aos gastos tributários. Portanto, é imperativo mostrar que os programas são efetivos no aumento da produtividade e do emprego e no crescimento das empresas.

Figura 1 – Gastos públicos relacionados às políticas federais de apoio às empresas

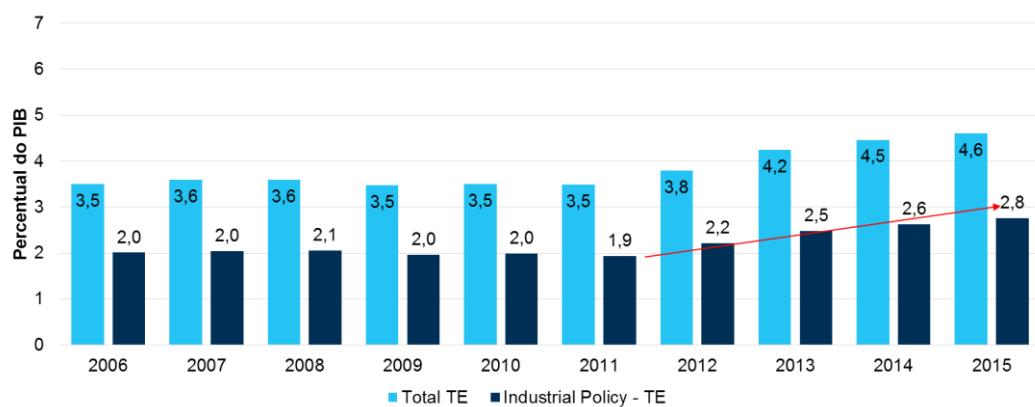


Fonte: Receita Federal do Brasil/Tesouro Nacional/SPE

Nota: Valores em milhões de reais, ano base 2015. Crédito estimado antes de 2008.

A maior parte do gasto com apoio ao setor privado é atribuída a **isenções fiscais**. Gastos tributários chegaram a 4,6% do PIB em 2015, incluindo proteção social e outras políticas, sendo 2,8% do PIB apenas os gastos tributários ligados ao setor privado (Figura 2). Esse montante deve-se majoritariamente ao regime especial do Simples Nacional, responsável por gastos de até 1,2% do PIB (Figura 3).

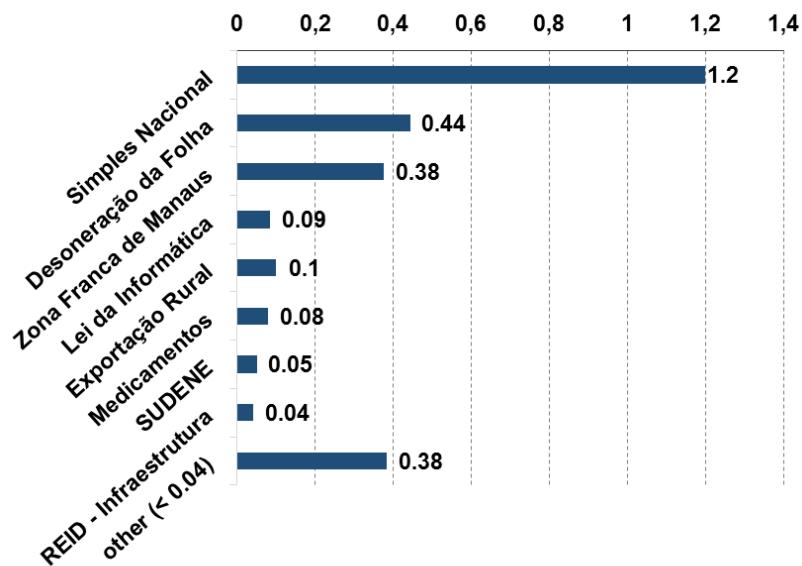
Figura 2 – Gastos tributários, em porcentagem do PIB (2006-2015)



Nota: Não incluídos os gastos tributários com ICMS em nível estadual (adicional de +1% do PIB).

Fonte: Receita Federal do Brasil

Figura 3 – Gastos tributários, em porcentagem do PIB (2015)



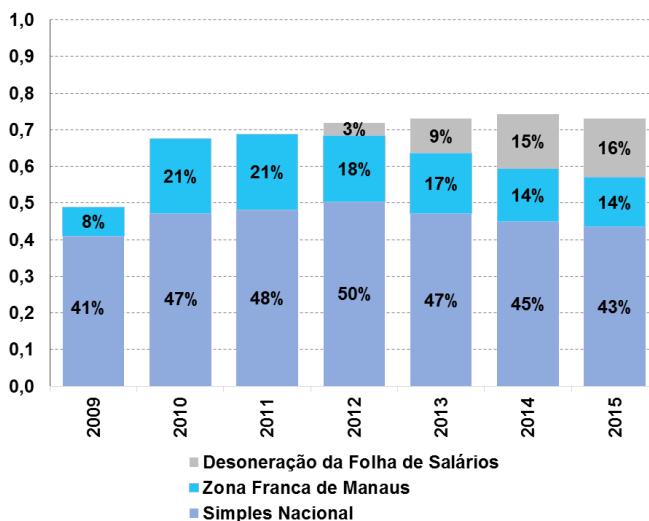
Fonte: Receita Federal do Brasil

Nota 1: "Outros" inclui a SUDAM, Inovação Tecnológica, Automotivo, Petroquímico, Etanol, Pesquisas em C&T.

Nota 2: O corte transversal inclui DTs relativas a bens de capital; a aceleração do período de depreciação dos equipamentos de capital, de 10 para 5 anos, a partir de agosto de 2012; o programa REPORTO, de abril de 2012, que isenta os impostos sobre importação e os impostos indiretos sobre investimentos em infraestrutura; o Plano Nacional de Banda Larga; a isenção de impostos indiretos para equipamentos domésticos e construções civis relacionadas a redes de telecomunicações.

A desoneração da folha também foi expandida de maneira muito ampla e para vários setores no período considerado. Mas a experiência internacional não é muito clara quanto a seu impacto. Finalmente, há a Zona Franca de Manaus (ZFM), que se trata mais de uma questão de transferência estadual do que de desenvolvimento do setor privado. Atualmente, o Simples Nacional representa quase a metade dos gastos tributários, seguido da desoneração da folha de salários e da ZFM (Figura 4).

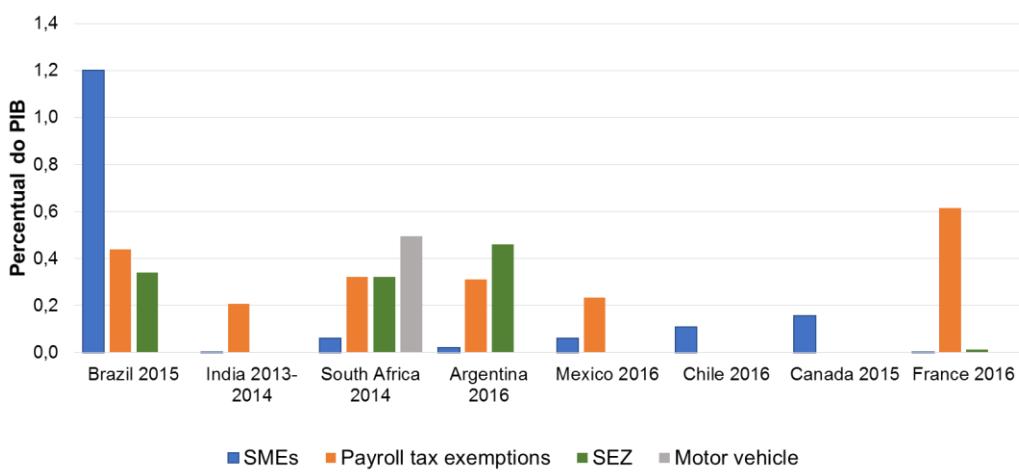
Figura 4 – Simples, Desoneração e Zona Franca de Manaus, em porcentagem do total de gastos diretos com políticas de apoio às empresas



Fonte: Receita Federal do Brasil.

Em uma comparação internacional, pode-se verificar que são muito grandes os gastos tributários no Brasil com o apoio às empresas (Figura 5). Os dados novamente excluem as esferas subnacionais e referem-se a fontes diferentes, com alguns problemas de compatibilidade na definição, mas úteis à ilustração. Verifica-se, por exemplo, que os gastos com programas voltados às pequenas e médias empresas (SMEs) e os gastos tributários com isenções na folha de pagamentos são pontos fora da curva ou *outliers*, quando comparados a outros países. No caso brasileiro, a coluna azul refere-se praticamente ao Simples Nacional e ao apoio às pequenas e médias empresas.

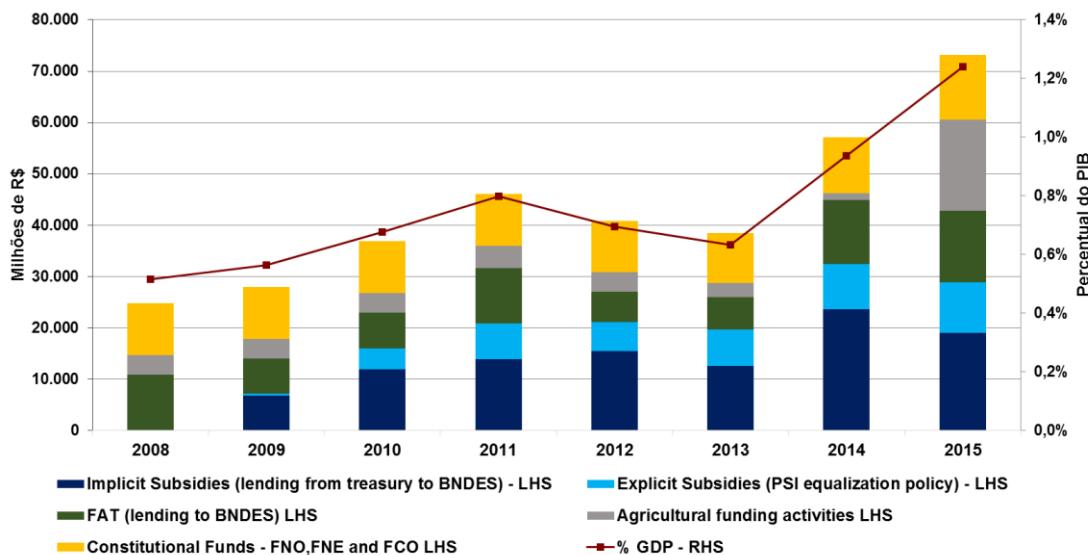
Figura 5 – Programas de gastos tributários selecionados, em porcentagem do



Fonte: Dutz *et al.* (2017).

Além das isenções fiscais, há o **crédito subsidiado**. O crédito subsidiado é muito grande no Brasil, tendo aumentado para 1,2% do PIB em 2015, majoritariamente ligado às operações do BNDES (Figura 6).

Figura 6 – Custo do Crédito Subsidiado, em porcentagem do PIB (2006-2015)



Fonte: Equipe Finance&Markets, Banco Mundial e Receita Federal do Brasil.

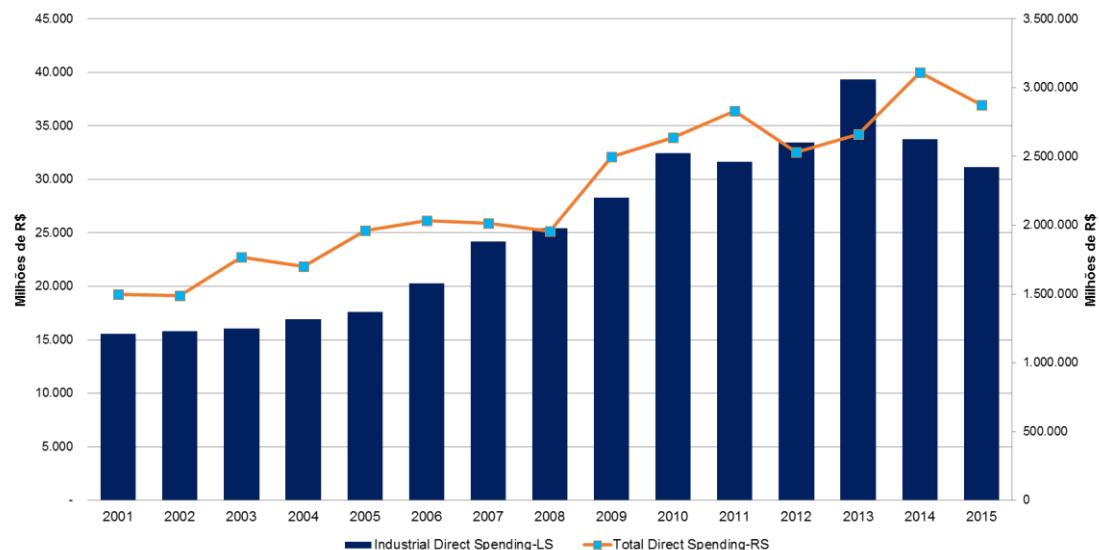
Nota 1: Valores em milhares de reais, ano base 2015.

Nota 2: Os fundos constitucionais são alocados para o Centro-Oeste (FCO), Nordeste (FNE) e Norte (FNO), para financiar empresas de todos os setores produtivos.

Nota 3: Fundos constitucionais e fundos de agricultura estimados antes de 2011.

Por sua vez, **gastos gerais**, que compreendem o apoio direto às empresas, também aumentaram (Figura 7), e já representam 0,5% PIB – a única porcentagem menor do que a de outros países.

Figura 7 – Gastos gerais (gastos diretos), em milhões de reais (2006-2015)

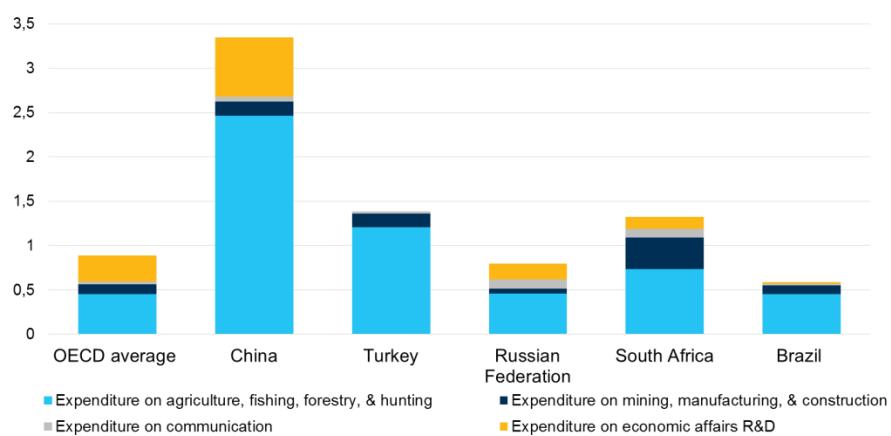


Fonte: Receita Federal do Brasil.

Nota 1: Valores em milhões de reais, ano base 2015.

Não é possível comparar com a China, onde os gastos diretos são muito significativos, mas sim com outros países em que os gastos diretos são menores (Figura 8). Entendemos que esses sejam os mecanismos de intervenção que tendem a expandir no futuro. Isto é, mais gasto direto e menos isenções tributárias e crédito subsidiado. Portanto, é nesses programas que os desenhos de políticas e as avaliações de impacto serão muito mais importantes.

Figura 8 – Gastos com assuntos econômicos, em porcentagem do PIB (2014)



Fonte: Dutz *et al.* (2017).

Não obstante o gasto significativo com políticas de apoio às empresas no Brasil, a eficácia e a eficiência dessas políticas não foram suficientemente documentadas. Faltam

avaliações sistemáticas e rigorosas quanto ao impacto desses gastos. A maioria das avaliações identificadas são quase-experimentais. É muito impactante a falta de experimentos. A criação da evidência é um elemento muito importante para facilitar as decisões de política pública, que têm que acontecer. Na seção seguinte, são apresentadas algumas evidências identificadas em estudos do Banco Mundial sobre alguns dos programas acima mencionados.

## POLÍTICAS DE APOIO ÀS EMPRESAS NO BRASIL: ALGUMAS EVIDÊNCIAS

### O Simples Nacional

Como visto, o Simples Nacional corresponde a um gasto de 1,2% do PIB brasileiro, mas há pouca evidência quanto a seu impacto. O objetivo do Simples Nacional era incentivar a formalização das pequenas e médias empresas, ao reduzir os custos operacionais por meio de simplificação de encargos fiscais e administrativos.

Porém, as evidências sobre a efetividade do Programa na formalização de empresas são inconclusivas. As evidências internacionais apontam que a maioria dos programas de formalização tem impacto limitado. Estudos realizados na Colômbia demonstraram que algum tipo de política semelhante tem um impacto positivo na formalização, mas, no geral, não muito. Os esforços para aprimorar o cumprimento da lei apresentam melhores resultados.

Além disso, percebem-se potenciais distorções na escolha de insumos intermediários, bem como consequências não-intencionais sobre a produtividade. Sabe-se, a partir de experiências de outros países, que isenções fiscais, como o Simples Nacional, podem inibir microempresários de crescerem, para evitar que passem para um regime diferente de tributação. Também há a probabilidade da carga tributária seguir um padrão de U invertido, deixando as empresas de médio porte em situação de desvantagem competitiva.

Nesse sentido, cabe ponderar a possibilidade de poupar 1,2% do PIB, retirando alguns dos incentivos do Simples, sem necessariamente elevar a carga tributária. Pelo contrário, a carga tributária deveria descer, por exemplo, por meio da integração doméstica no mercado, com alíquotas semelhantes.

### Desoneração da folha de pagamentos

Desoneração da folha de pagamentos também é um programa com custo fiscal alto, de 0,4% do PIB, que tampouco apresenta uma evidência de que está funcionando. O objetivo da intervenção é gerar empregos ou proteger os já criados. Contudo, não está claro se é efetiva. Se for, o efeito é limitado e há dúvidas se novos empregos estão, de fato, sendo criados ou se há um deslocamento do setor informal para o formal. O metaestudo constatou um impacto limitado na taxa de emprego. (DUTZ *et al.*, 2017)

Cabe observar que a intervenção foi inspirada em programas europeus, adotados em circunstâncias completamente diferentes. Há casos em que a desoneração aumentou os salários, mas a criação de novos empregos foi limitada ou inexistente (p. ex., Chile, Argentina,

Suécia e Finlândia). Na Colômbia, por sua vez, foi identificada uma redução do trabalho informal.

A eficiência do gasto, porém, é questionável, por representar um alto custo fiscal (0,4% do PIB) e um alto custo por emprego criado, equivalente ao triplo do salário do trabalhador.

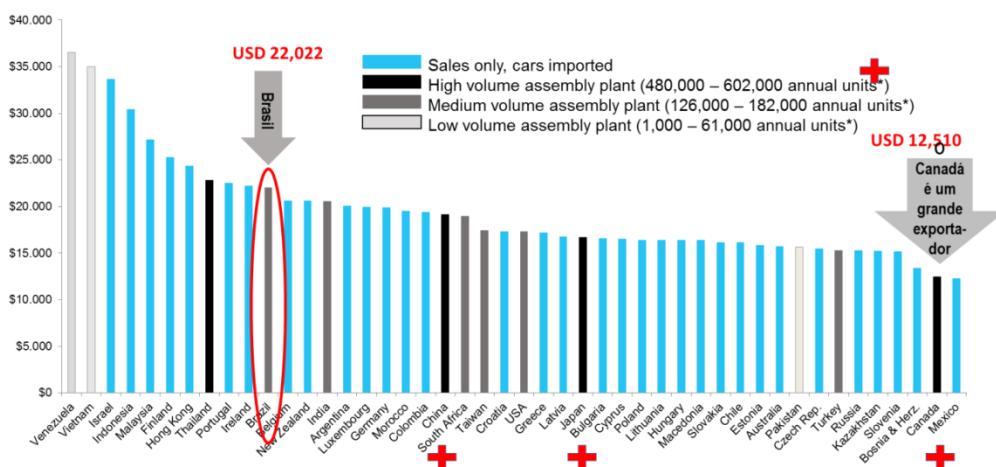
Nesse ponto se pode entender a relação com a falha do mercado. Usar isenções fiscais para fazer política para o setor privado guarda implícito no desenho da política pública uma ideia de falha de mercado que muitas vezes não é a falha de mercado principal. Por exemplo, a isenção fiscal não ajuda pequenas e médias empresas quando a falha é informação assimétrica. Uma complicação quanto à tomada de decisão sobre algumas isenções tributárias se deve a como são desenhadas. As isenções podem não ter impacto, mas são muito mais fáceis de implementar.

### Inovar-Auto

O programa Inovar-Auto, já estava a ponto de fechar, mas há uma discussão sobre o próximo programa. O programa compreende uma aglomeração de incentivos, que o torna muito complexo, produzindo distorções entre setores e entre tipos de empresas, com impacto negativo sobre a produtividade.

No caso desse programa, um elemento importante na questão da política pública não é seu custo fiscal, que não é muito elevado, mas seu custo para o consumidor. O custo de um automóvel no país é muito alto. A Figura 7 apresenta uma comparação do preço de venda anunciado para o modelo básico do Toyota Corolla 2017, que é produzido no Brasil – o que tornaria o preço mais baixo. Quando comparado ao preço de outros países, o preço pago aqui é muito alto.

Figura 7 – Preço de venda anunciado para o modelo básico do Toyota Corolla 2017



Fonte: Dutz *et al.* (2017).

Esse exemplo evidencia que as políticas de apoio ao setor privado não são neutras, mas têm um impacto muito grande nas pessoas. Quando se pretende criar emprego no setor, que é um objetivo muito importante, tem que se pensar também qual será o custo desse emprego

criado. Nesse caso, o custo mostra-se muito elevado para o consumidor. Caberia agora, portanto, pensar como apoiar o setor automotivo usando instrumentos que tenham um custo menor à população.

### **Programa de Sustentação do Investimento (PSI)**

Programas de crédito subsidiado assumem a existência de uma falha no mercado muito determinada, ligada ao mercado financeiro, ao financiamento da inovação e de outras coisas. Algumas avaliações de impacto sobre o cartão do BNDES demonstraram resultados positivos, mas, no geral, há poucos estudos.

Além da eficiência do programa, outro elemento importante que deve ser ponderado (e que nem sempre comprehende as avaliações de impacto no setor privado) é a questão da assistência fiscal. Muitas vezes focamos apenas em qual é o impacto do programa, se aumentou a produtividade, o emprego, etc. A metodologia é complexa, e os dados disponíveis não dão para mais. Mas há também o elemento da incidência, que é muito importante e que deve ter uma agenda específica no futuro. Ou seja, quem está sendo beneficiado pelo programa?

Essa questão envolve identificar quem é o beneficiário do programa, qual a diferença relativa entre empresas e entre trabalhadores. Além disso, que tipo de empresas ou que tipo de trabalhador é beneficiado e qual o impacto na desigualdade da renda. Isso porque, muito desses programas são para grandes empresas, como, por exemplo, os de crédito subsidiado. O crédito é captado por empresas médias e grandes apenas, em virtude da forma como o BNDES o concede. Esse é um ponto muito importante: como o desenho e a implementação do programa pode trocar completamente quem são os beneficiários do programa e os seus resultados.

## **PROGRAMAS DE APOIO AO SETOR PRIVADO: MUITO GASTO, POUCO IMPACTO**

Como visto, há muito pouca evidência sobre a eficácia e a eficiência dos programas de apoio ao setor privado no aumento da produtividade e na criação sustentável de empregos – e a maior parte da evidência que existe é de que não há muito impacto. Os programas maiores provavelmente não são custo-eficientes e há vários com um efeito de renda regressivo.

Além disso, há que se considerar que o desenho de muitos desses programas (como de isenções, conteúdo local, etc.) introduz muitas distorções na economia, que são contrárias à melhoria da concorrência nos mercados de produção e de insumos, da inovação, da realocação dos recursos, bem como da produtividade dentro das empresas.

As análises realizadas pelo Banco Mundial, acima sintetizadas, evidenciam que a maioria das políticas de suporte comercial beneficia as grandes empresas. No que compete aos gastos tributários, o Simples Nacional beneficia pequenas empresas formais, mas outros programas beneficiam, em grande parte, empresas de grande porte já estabelecidas (p. ex.:

ZFM e Inovar-Auto). Estudos apontam, ainda, que a desoneração da folha intensificou o padrão regressivo do sistema tributário. (DUTZ *et al.*, 2017)

Por sua vez, o crédito subsidiado beneficia, principalmente, empresas grandes e dominantes, ao passo que os gastos gerais (diretos) apresentam pouco acesso para empresas pequenas ou novas. Ademais, identificaram-se impactos negativos nos preços ao consumidor, no caso da ZFM e do programa Inovar-Auto.

Em suma, a análise aponta **muito gasto e pouco impacto**. Diante disso, propõe-se a seguinte provocação. Se os três grandes programas (Simples, desoneração da folha e ZFM) não produzem o impacto esperado, seria possível economizar ou realocar até 2% do PIB com a eliminação de programas ineficazes (Quadro 1).

Quadro 1 – Resumo das possíveis economias fiscais com os três grandes programas de apoio ao setor privado

| Ações de Política  | Eficiência | Patrimônio líquido | Economia Fiscal (% do PIB) |
|--|------------|--------------------|----------------------------|
| Reformar o Simples<br>(idealmente, a reforma faria parte de um processo mais amplo de simplificação fiscal)      | ++         | ?                  | 1,2                        |
| Revogar as desonerações da folha de pagamentos<br>(desoneração da folha)   | +          | ?                  | 0,4                        |
| Reformar a Zona Franca de Manaus<br>(idealmente, como parte de uma alternativa menos distorcida e menos onerosa) | ++         | +                  | Até 0,4                    |
| <b>Total</b>   |            |                    | <b>Até 2,0</b>             |

Fonte: Dutz *et al.* (2017).

Nesse cenário, portanto, haveria muito espaço para poupança fiscal. Não se trata de dizer que os programas devam ser parados, mas de dizer que se tem que fazer alguma coisa. Porém é preciso ter boas avaliações das políticas para tomar decisões. Afinal, 2% do PIB é muito recurso, que poderia ser gasto para aumentar o Bolsa Família ou melhorar a educação, o PRONATEC, etc. Portanto, é imperativo mais avaliações para informar o debate público, bem como para reformular as políticas, eliminá-las gradualmente ou substituí-las, enquanto se tratam os impactos na economia política.

## O DESAFIO DA INSTITUCIONALIDADE DA AVALIAÇÃO

Além das avaliações de impacto em si, há outras áreas que precisam de mais trabalho. Já falamos da necessidade de avaliação da qualidade do desenho e da implementação das políticas de apoio ao setor privado. Mas há também a questão da institucionalidade. Quais programas serão avaliados? Quem avaliará os programas? E o que vamos fazer com o resultado da avaliação de impacto (as evidências)? Diferentes países já começam a fazer mais avaliação de impacto, mas, uma vez que acaba a avaliação de impacto, não há uma institucionalidade para usar os resultados. Não há processo através do qual vamos mudar o programa ou vamos tomar algum tipo de decisão. Logo, temos que trabalhar não só por mais avaliações de impacto, mas também por uma institucionalidade clara, em que haja aprendizagem dos resultados e o uso desses resultados.

No curto prazo, a resposta é: regras e processos. Ou seja, desenhar regras claras do que se tem que avaliar ou quais são os critérios através dos quais um programa tem que ser avaliado. Por exemplo, um critério baseado no tamanho do programa: se você está gastando R\$ 30 milhões em apoio às pequenas e médias empresas, parece justo exigir que se tenha uma avaliação de impacto.

Processos também no sentido da aprendizagem. Um elemento muito importante é avaliação a médio prazo. Precisamos de avaliações no médio prazo para tomar decisões sobre continuar e/ou modificar o programa. Muitas vezes as instituições não têm esses processos, e é muito difícil convencer a pessoa política. Em virtude disso, precisamos de um processo claro que diga, por exemplo, que serão feitas avaliações do programa a cada três anos – avaliações talvez qualitativas, falando com os beneficiários, para fazer um processo de modificação.

Também cabe discutirmos o tipo de pessoal que tem que facilitar esse processo e se tem que criar uma unidade especializada ou não. Não se trata apenas de colher informação interna, mas também as melhores práticas de fora. Ocorrem muitos *workshops*, mas há uma loteria para decidir quem vai participar e depois não há nenhum tipo de disseminação interna dentro da instituição. Mais uma vez, precisamos discutir o processo, a criação ou não de uma unidade que tenha seus termos de referência, seus processos e que tenha o que fazer.

Cabe lembrar que o processo político que pode mudar mais, mas em questões do setor privado e de inovação, cujos impactos são de médio e curto prazo, precisamos de estabilidade. Um novo governo pode mudar as diretrizes, mas com regras e processos claros, haverá uma unidade que possa apoiar e com quem possa contar.

## CONSIDERAÇÕES FINAIS

Há ainda pouca evidência quanto aos efeitos das políticas de apoio às empresas em termos de produtividade, de mais e melhores empregos e de preços ao consumidor. A incidência dessas políticas também ainda não é bem compreendida. Não há evidências

rigorosas sobre a incidência dos programas (dados empresariais e familiares). O efeito direto provavelmente é regressivo na maioria dos programas. Ou seja, favorece grandes empresas já estabelecidas ou empregos formais já existentes. Sendo assim, carecemos de uma análise mais aprofundada desses programas, bem como de uma institucionalidade que apoie e conte com processos para difundir os resultados dessas avaliações.

## REFERÊNCIAS

DUTZ, M. et al. Business Support Policies in Brazil: Large Spending, Little Impact. *Background Paper*. Washington, DC: The World Bank, 2017. Disponível em: <<http://documents.worldbank.org/curated/en/458011511799140856/Business-support-policies-large-spending-little-impact>>. Acesso em: 13 mai. 2018.

## AVALIAÇÃO DE IMPACTO COMO INSTRUMENTO PARA O APRIMORAMENTO DAS POLÍTICAS EDUCACIONAIS\*

André Loureiro<sup>1</sup>

**Resumo:** A avaliação de impacto é um elemento-chave da gestão por resultados, tendência mundial vinculada à crescente demanda da sociedade por resultados concretos. Nesse sentido, este texto discute a importância da avaliação de impacto no setor educacional, a partir da experiência do Grupo Banco Mundial no Brasil. Além da análise de alguns dados sobre o setor no País, são apresentados exemplos de aplicação da Teoria da Mudança e do Modelo Lógico de Programa, ferramentas que auxiliam no esclarecimento e estabelecimento da cadeia causal de funcionamento da intervenção, até seus resultados esperados. Dessa forma, busca-se facilitar não só o planejamento, o monitoramento e a avaliação de programas educacionais, mas também o diálogo entre governo e gestores.

**Palavras-chave:** Avaliação de impacto. Gestão por resultados. Setor educacional. Teoria da Mudança. Modelo Lógico.

**Abstract:** Impact evaluation is a key element of managing for results. Managing for results is a worldwide trend that responds to society's growing demand for concrete policy outcomes. In this sense, this text discusses the importance of impact evaluation in the education sector, based on the experience of the World Bank Group in Brazil. Besides the analysis of some education data about the country, examples of the application of the Theory of Change and the Logic Model are presented. Theory of Change and the Logic Model are tools that help clarifying and establishing the causal chain of an intervention, from its beginning until its expected results. They seek to facilitate not only planning, monitoring and evaluation of education programs, but also the dialogue between government and managers.

**Keywords:** Impact assessment. Managing for results. Education Sector. Theory of Change. Logic Model.

---

\* Texto elaborado a partir da transcrição da fala e da apresentação em slides de palestras proferida no Workshop Avaliação de Impacto de Políticas Públicas, realizado nos dias 29 e 30 de novembro de 2017, no Auditório da Secretaria de Planejamento, Governança e Gestão (SPGG-RS), com o apoio do Banco Mundial.

<sup>1</sup> Economista Sênior do Banco Mundial.

## INTRODUÇÃO

O Banco Mundial apoia a realização de avaliações de impacto no setor de educação no mundo todo. A avaliação de impacto é um elemento-chave da gestão por resultados, e na educação isso não é diferente. Gestão por resultados é uma tendência mundial e está vinculada à crescente demanda da sociedade por resultados concretos. Gestores de programas educacionais são cada vez mais cobrados a prestar contas de resultados. Foi-se o tempo em que os cidadãos ficavam satisfeitos com a mera inauguração de obras, seja uma escola, seja um hospital. Atualmente, tem-se uma expectativa e uma cobrança maiores em relação ao real funcionamento dos programas ou projetos públicos e os benefícios que revertem para a sociedade. Essa preocupação de entregar resultados deve estar presente desde o momento da concepção do programa ou projeto.

Demonstrar a eficiência e a efetividade do programa é do interesse de gestor. Por exemplo, evidências apontam que um programa de escolas em tempo integral gera impactos na aprendizagem dos alunos e, no longo prazo, uma melhor inserção dos graduandos no mercado de trabalho. Contudo, trata-se de um programa relativamente caro, especialmente em um cenário de recursos escassos. Cabe, portanto, ser avaliado.

Teoricamente, qualquer programa com informações disponíveis pode ser avaliado. Porém, sugere-se focar em programas que são mais inovadores, ou seja, programas que estão inovando na forma de qualquer política ou em programas que já foram testados, avaliados e comprovados em outros lugares, mas cuja mudança de ambiente pode afetar o impacto produzido. Nesse sentido, o presente artigo aborda a importância da avaliação de impacto de maneira geral e, em específico, no setor educacional, a partir da experiência no Grupo Banco Mundial no Brasil.

Inicialmente, mostraremos alguns dados sobre a situação da educação no país e que evidenciam a necessidade de intervenção. Em seguida, apresentamos a Teoria da Mudança e o Modelo Lógico de programa, instrumentos que facilitam a construção da cadeia de resultados esperados de um programa e, assim, apoiam tanto sua implementação e seu monitoramento, quanto sua avaliação. Nesse sentido, pondera-se também quanto ao emprego de avaliações de impacto prospectivas para um maior custo-efetividade dos programas a serem implementados. Por fim, apresentamos uma experiência de avaliação na área da educação, seguida de recomendações do Banco Mundial para aumentar a qualidade do setor.

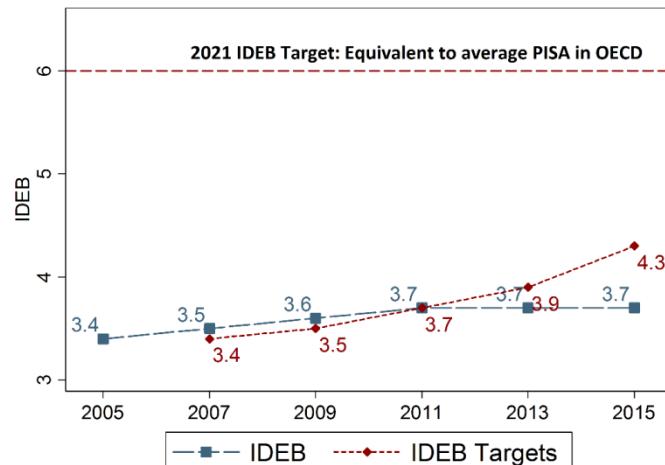
## UM RETRATO DA EDUCAÇÃO NO BRASIL

Os programas ou projetos públicos devem estar fundamentados em evidências. Logo, antes de definir e mensurar o resultado a que se deseja chegar, é necessário diagnosticar a situação como ela se apresenta na realidade.

Em se tratando da educação no Brasil, o Índice de Desenvolvimento da Educação Básica (IDEB) encontra-se estagnado já há algum tempo, tendo fugido bastante da meta que

foi estabelecida pelo Ministério da Educação (Figura 1). Comparando com a média dos países desenvolvidos no Programa Internacional de Avaliação de Estudantes (PISA), calculado pela Organização para a Cooperação e Desenvolvimento Econômico (OCDE), a meta do IDEB seria de 6 pontos em 2021. Contudo, o Brasil está muito longe desse resultado.

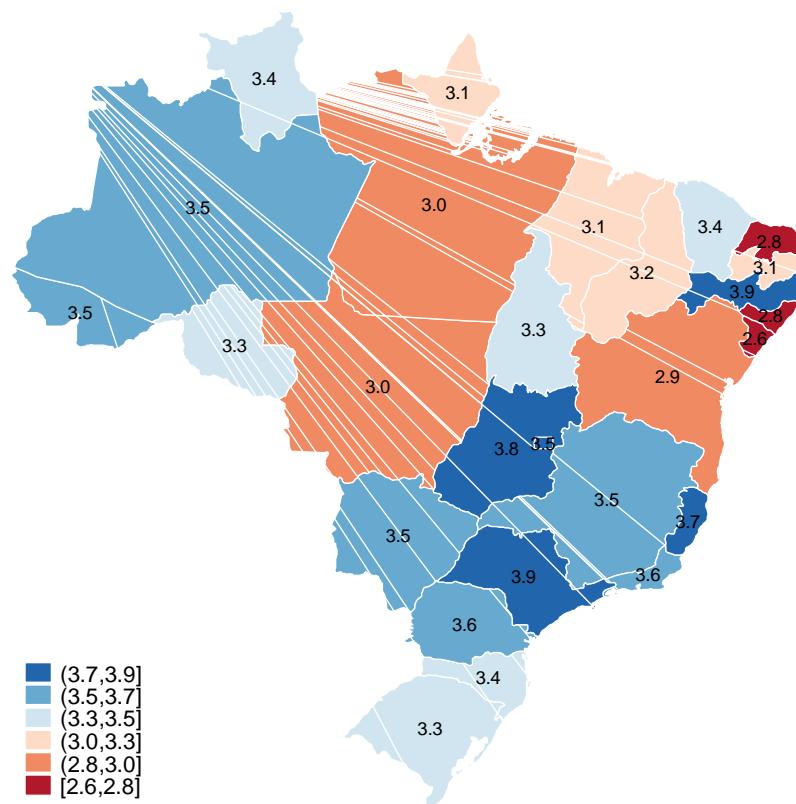
Figura 1 – IDEB do Ensino Médio: meta e média brasileiras (2005-2015)



Fonte: INEP (2015).

Não só a média brasileira no IDEB é baixa, como há uma heterogeneidade relativamente grande entre os estados. E mesmo a maior média de rede estadual, que é a de São Paulo e Pernambuco, ainda é preocupante, tendo sido de apenas 3,9 pontos em 2015 (Figura 2).

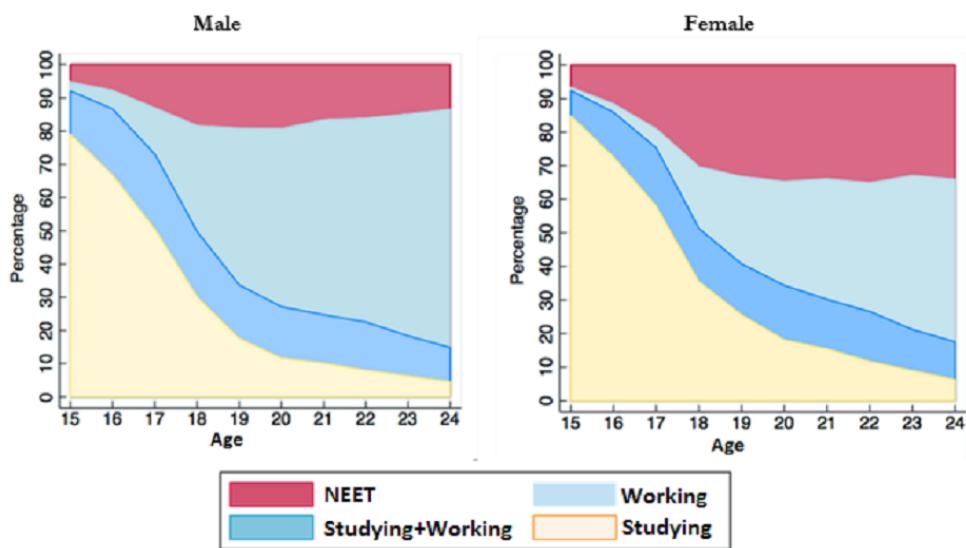
Figura 2 – IDEB: redes estaduais (2015)



Fonte: INEP (2015).

Além disso, mais de um milhão de jovens de 17 anos estão fora da escola, e boa parte destes tampouco está trabalhando (IBGE, 2015). A Figura 3 apresenta a distribuição dos jovens por idade, sexo e situação: se estudando (amarelo), trabalhando (azul escuro), estudando e trabalhando (azul claro) ou nem estudando, nem trabalhando (vermelho). Percebe-se que há uma proporção grande de jovens maiores de 16 anos que não estudam, nem trabalham – o que reflete, obviamente, em uma menor taxa de conclusão do Ensino Médio. Apenas 58% dos jovens brasileiros com 19 anos haviam concluído o Ensino Médio em 2015 (IBGE, 2015). É uma situação bastante preocupante, tendo em vista que se trata de elemento básico para a cidadania.

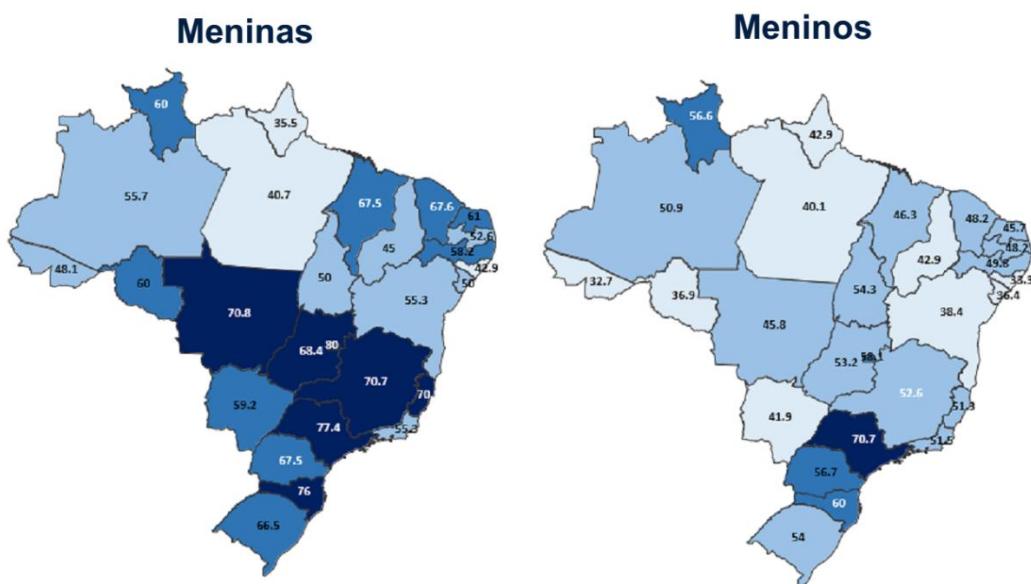
Figura 3 – Situação escola-trabalho dos jovens, por idade e sexo (2013)



Fonte: Banco Mundial (2016). Fonte dos dados brutos: IBGE (2013).

A Figura 4 evidencia que a heterogeneidade não é apenas regional, mas também de gênero. A taxa de conclusão do Ensino Médio no Estado de São Paulo, por exemplo, chega a 77,4% para meninas, sendo um pouco menor para meninos, 70,7%. No Amapá, por sua vez, pouco mais de um terço das meninas concluíram o Ensino Médio. Trata-se do único estado em que a taxa de conclusão é maior para meninos do que para meninas.

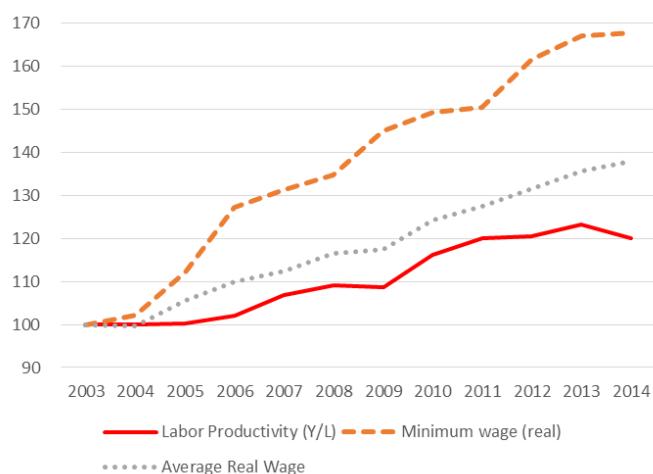
Figura 4 – Taxa de conclusão do Ensino Médio dos jovens de 19 anos, por estado e por sexo (2015)



Fonte: PNAD/IBGE, 2015.

Esse diagnóstico do Brasil não é tão diferente do de outros países em desenvolvimento, principalmente na América Latina. O país, contudo, segue especialmente estagnado nos últimos anos. E, claro, a baixa qualidade da educação brasileira tem impacto imediato na baixa produtividade do trabalho, dentre outros aspectos. No Brasil, produtividade do trabalho tem evoluído de uma forma muito lenta (Figura 5).

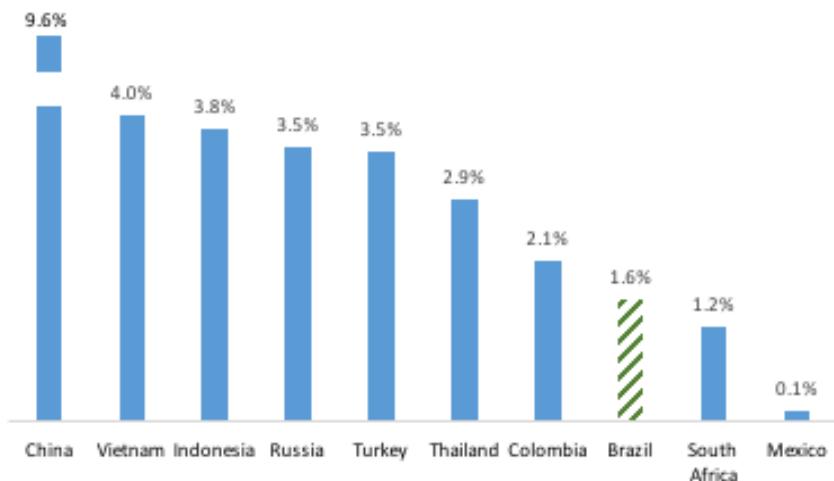
Figura 5 – Tendência da produtividade do trabalho no Brasil (2003-2014)



Fonte: Banco Mundial (2016).

Quando comparada a outros países em desenvolvimento, tais como China, Vietnã, Indonésia, Rússia, Turquia, Tailândia e Colômbia, a taxa média de crescimento da produtividade do trabalho no Brasil, entre 2001 e 2013, é bastante baixa, ficando à frente apenas da África do Sul e do México (Figura 6).

**Figura 6 – Crescimento médio anual da produtividade do trabalho nos países em desenvolvimento e no Brasil (2001-2013)**



Fonte: Banco Mundial (2016).

A partir dessas evidências é que se tem condição de pensar e planejar os programas e projetos públicos, bem como de avaliar seu impacto. Para tanto, além de dados, é fundamental compreender os mecanismos causais que estão por trás das intervenções que planejamos. Nesse sentido, na seção seguinte, apresenta-se a Teoria da Mudança, que busca apoiar essa reflexão.

## TEORIA DA MUDANÇA E MODELO LÓGICO DE PROGRAMA

Quando se planeja um programa, primeiramente, é necessário questionar o que se pretende alcançar com tal esforço. Um programa pode ter vários objetivos. Mas, para traçá-los, é necessário olhar para as evidências existentes, que apontam quais os resultados que o programa teria condições de gerar.

Para gerar esses resultados esperados, por sua vez, cabe questionar quais os resultados intermediários que devem ser gerados. Além disso, quais os produtos que precisariam ser entregues para entregar esses resultados intermediários, quais as atividades que precisam ser desenvolvidas para entregar esses produtos e, por fim, quais os insumos necessários para a realização dessas atividades. Ao responder a esses questionamentos, traçam-se, de maneira lógica, sintética e fundamentada em evidências, as etapas que é necessário percorrer para alcançar o resultado final (o objetivo) esperado por meio do

programa. Isto é, a cadeia de resultados esperada ou as hipóteses que fundamentam a Teoria da Mudança do programa.

É possível pensar a Teoria da Mudança numa estrutura de Marco Lógico, expondo a cadeia de funcionamento da intervenção que está sendo proposta, de uma forma bastante concreta, sintética e simplificada. Por meio do Marco Lógico, identificam-se as etapas que precisam ser desenvolvidas para alcançar o resultado desejado. Dessa forma, facilita-se não só o planejamento, o monitoramento e a avaliação do programa ou projeto, mas também o diálogo entre governo e gestores.

A Figura 7 apresenta uma cadeia de resultados típica. A cadeia pode ser construída de distintas formas. Porém, um dos formatos mais tradicionais é construído de trás para frente. Ou seja, primeiramente foca-se nos resultados finais que a política pretende alcançar. Tratam-se dos objetivos do programa, inclusive os objetivos de longo prazo. Em seguida, dá-se um passo para trás, e busca-se identificar quais são os resultados intermediários que precisam ser alcançados para chegar nesse resultado final. Em outras palavras, o que precisa acontecer antes para que se chegue àqueles resultados finais. E assim sucessivamente. Quais são os produtos que precisam ser entregues para alcançar esses resultados intermediários. Quais são as atividades que precisam ser desenvolvidas para entregar esses produtos. Quais são os insumos necessários para que as atividades sejam realizadas, e os produtos possam ser entregues.

Figura 7 – Teoria da Mudança e Marco Lógico: uma cadeia de resultados típica



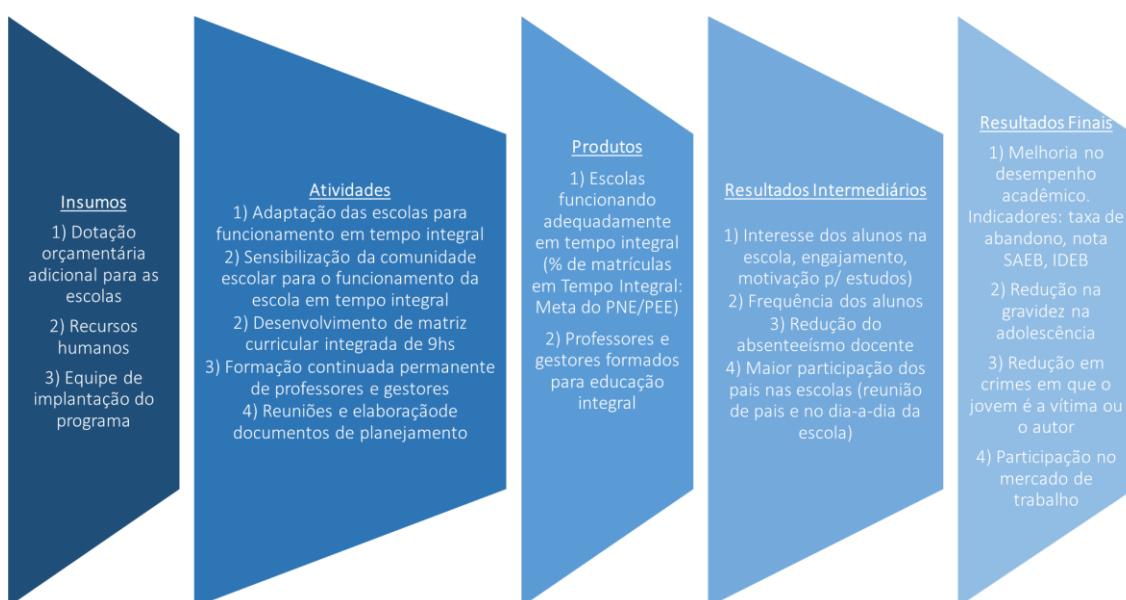
Fonte: Banco Mundial, 2016.

No caso de um programa de escolas em tempo integral, por exemplo, algumas questões merecem ser respondidas antes de efetivar sua implementação, tais como:

- ✓ Qual é o impacto de se ampliar o dia escolar no resultado acadêmico do aluno (aprendizagem /eficiência) e em outros resultados, tais como, resultado social, e resultado no trabalho?
- ✓ Que resultados intermediários seriam necessários para alcançar tais resultados?
- ✓ Que produtos precisariam ser entregues para alcançar tais resultados intermediários?
- ✓ Que atividades precisariam ser desenvolvidas para entregar tais produtos?
- ✓ Quais seriam os insumos necessários?

Com base nas respostas a essas questões, a Figura 8 sintetiza a cadeia de resultados elaborada para o Programa Federal de Fomento às Escolas de Tempo Integral de Ensino Médio.

**Figura 8 – Cadeia de Resultados: Programa Federal de Fomento às Escolas de Tempo Integral de Ensino Médio**



Fonte: Banco Mundial (2016)

O Marco Lógico do programa serve também à realização da avaliação de impacto, pois facilita a identificação das variáveis do programa e, assim, a disponibilização de instrumentos de coleta de dados sobre as etapas do programa, a custos praticamente zero. No Brasil, a educação beneficia-se de sistemas de informações bastante robustos. Se planejadas na concepção do programa, é possível fazer avaliações de impacto, tanto experimentais como quase experimentais, com certa facilidade e a custos bastante reduzidos. Ainda que na experimental possa ser necessário realizar algum tipo de coleta de dados primários, é possível

reuni-los a dados já existentes, para obter uma análise mais aprofundada de alguns aspectos do programa ou projeto de interesse.

Evidentemente, é possível coletar informações no nível de resultados finais e da intervenção e, teoricamente, fazer a avaliação. Porém, por vezes, não se consegue perceber nenhum efeito sobre o resultado final, por mais que o ambiente seja propício para que o resultado seja alcançado e que a avaliação tenha sido desenhada de maneira correta (experimental ou quase experimental, conforme a realidade do programa). Diante de um impacto zero, estaticamente não significativo, têm-se três possibilidades.

Em primeiro lugar, obviamente, pode ser que a intervenção não tenha surtido efeito. Ou seja, não houve resultado. Em segundo lugar, pode ter havido heterogeneidade de impactos, quando o impacto médio é nulo, mas se podem observar efeitos para certos subgrupos da população beneficiada. Nesse caso, em específico, pode-se optar por focalizar o programa ou, se as condições dos beneficiados efetivamente forem replicáveis ou mutáveis, gerar essas mesmas condições para os demais beneficiários. Uma terceira razão clássica para a ausência de resultados, ainda, é a falha na implementação. Um programa pode conceitualmente funcionar muito bem, mas se for mal implementado, não resultará em nenhum impacto.

Uma falha de implementação pode acontecer em um único momento da cadeia de resultados, comprometendo o restante. O programa pode ter funcionado quando da organização das atividades e da entrega dos produtos, por exemplo. Porém, a forma como os produtos foram entregues comprometeu os resultados intermediários, dissipando o possível efeito do programa. Isso não significa que a intervenção não seja potencialmente benéfica.

Nesse cenário, o Marco Lógico ajuda a identificar o que de fato não aconteceu, por conta da dificuldade de implementação, podendo ser corrigido. Portanto, traçada a cadeia de resultados, consegue-se testar, por meio da avaliação, se o programa não produziu os resultados esperados como um todo ou se foi a ausência de um resultado intermediário que inibiu o efeito esperado no resultado final. Essa verificação pode ser feita após a implementação do programa ou anteriormente, a partir da avaliação de impacto prospectiva de um piloto.

## AVALIAÇÃO DE IMPACTO PROSPECTIVA PARA UM MAIOR CUSTO-EFETIVIDADE

Testar inovações na concepção e implementação dos programas é muito importante, especialmente na educação. Um programa pode ser feito de várias formas. A partir da realização de avaliações de impacto de programas educacionais ou da avaliação de diferentes formas (modalidades, intensidades, frequências, etc.) de um mesmo programa, por exemplo, é possível fazer algumas comparações, sobretudo comparações de custo-efetividade. Isso porque se deve levar em conta que um programa pode ser até efetivo, mas talvez se consiga alcançar o mesmo resultado com um programa mais barato.

Hoje em dia, principalmente no Brasil, com a situação de restrição fiscal, essa questão se torna ainda mais premente. É necessário saber se um programa realmente produz o impacto almejado e, mesmo se produz, saber se a alternativa avaliada é mais custo efetiva. Para tanto, são realizadas avaliações de impacto prospectivas.

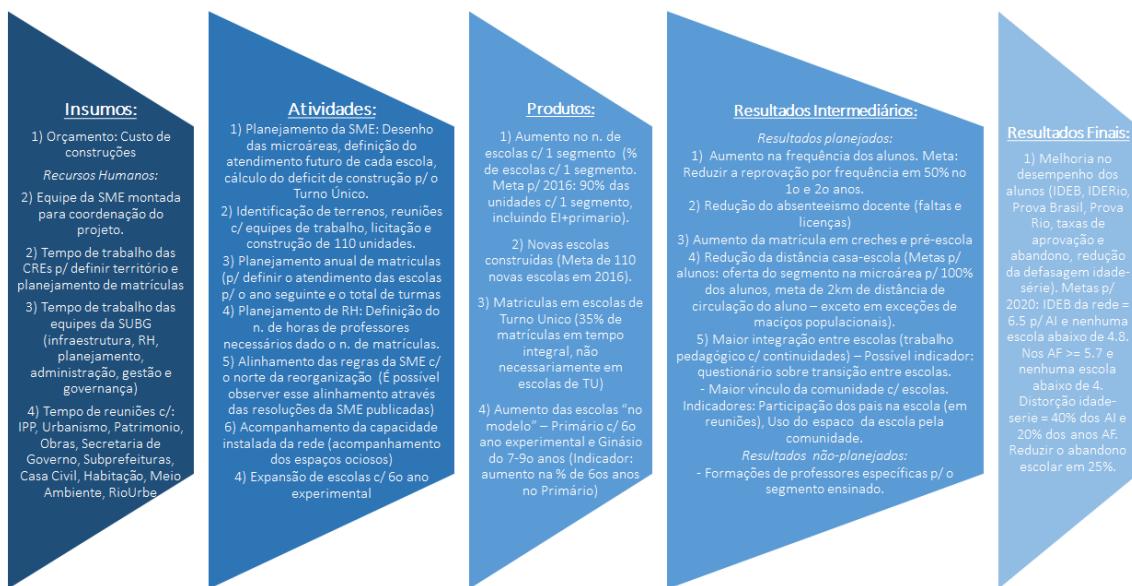
Por meio da avaliação de impacto prospectiva, faz-se o exercício de mensurar o benefício de uma alternativa ainda quando da concepção do programa. Além disso, permite fazer uma análise de incidência, por meio da qual se busca identificar quais são os grupos, os indivíduos, as firmas, por exemplo, que mais se beneficiam de uma intervenção. Com essa análise de heterogeneidade de impactos, é possível verificar em quais circunstâncias o programa funciona melhor e, por conseguinte, tentar replicar essas circunstâncias, se elas forem replicáveis, para os demais grupos, indivíduos ou firmas. Há ainda uma segunda opção, quando as circunstâncias de um programa não são transferíveis, que consiste em pensar em um programa focalizado. Esses diferentes impactos que podem ser captados por meio da avaliação de impacto prospectiva sublinham a importância dos programas-piloto.

Como visto, a compreensão dos mecanismos causais esperados é essencial para desenhar e planejar as etapas necessárias para implementar um programa ou projeto, construindo uma visão do todo – e é ainda mais útil quando se realiza uma avaliação prospectiva com base em um projeto-piloto. Por meio da avaliação do piloto, o gestor consegue tomar decisões acerca do desenho do programa, se universal ou focalizado, por exemplo, com base em evidências concretas. Além disso, consegue antever dificuldades possíveis e dar tratamento a elas de maneira antecipada, elevando a qualidade da implementação e a efetividade do programa.

## **EXPERIÊNCIA DE AVALIAÇÃO DE IMPACTO NO SETOR EDUCACIONAL**

A seguir, apresenta-se uma experiência de avaliação de impacto apoiada pelo Banco Mundial, em parceria com a Prefeitura do Rio de Janeiro (Banco Mundial, 2017). A avaliação buscou mensurar os resultados do programa de reorganização da rede de ensino público do município. Dentre a série de razões que justificavam o programa, destaca-se a significativa mudança demográfica, demonstrada pela redução na taxa de natalidade, com efeitos evidentes sobre o número de vagas ofertadas no sistema de ensino. A Figura 9 apresenta a cadeia causal esperada do programa.

**Figura 9 – Cadeia de Resultados: Reorganização da Rede de Ensino do Município do Rio de Janeiro**

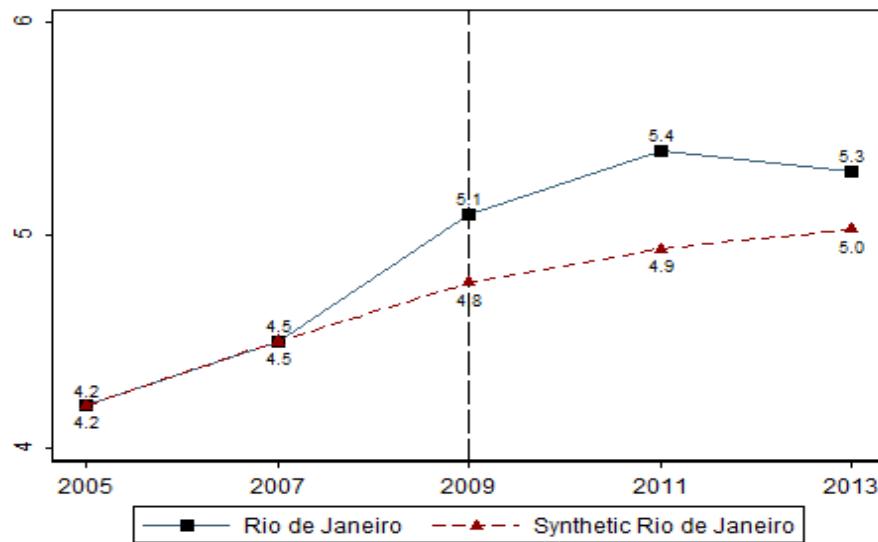


Fonte: Banco Mundial, 2017.

No desenho da avaliação, empregou-se o controle sintético, um método quase-experimental de avaliação de impacto. Por meio desse método, é possível comparar uma única unidade tratada (no caso, o município do Rio de Janeiro) com outros municípios (no caso do Brasil, 5.775 municípios). No caso, comparou-se a evolução dos resultados do IDEB do Rio de Janeiro entre 2005 e 2013 com o chamado “Rio de Janeiro Sintético”, construído a partir do que teria acontecido com o município caso não tivesse sido implementada uma série de políticas públicas lá implementadas, inclusive a de reorganização da rede e o programa de tempo integral. Por meio dessa comparação, foi possível perceber uma diferença significativa.

A Figura 10 apresenta os resultados do IDEB para os anos iniciais do Ensino Fundamental da rede municipal do município do Rio de Janeiro e o Rio de Janeiro Sintético (contrafactual) no período de 2005 a 2013. Os resultados do Rio de Janeiro foram mais altos do que aqueles observados no Rio de Janeiro Sintético. Isto é, os resultados são superiores à situação contrafactual em que o Rio de Janeiro tivesse implementado políticas similares a outros municípios. Verifica-se, em 2009, um salto nessa diferença.

Figura 10 – IDEB do Rio de Janeiro vs. Rio de Janeiro Sintético: Ensino Fundamental – Anos Iniciais (2005-2013)

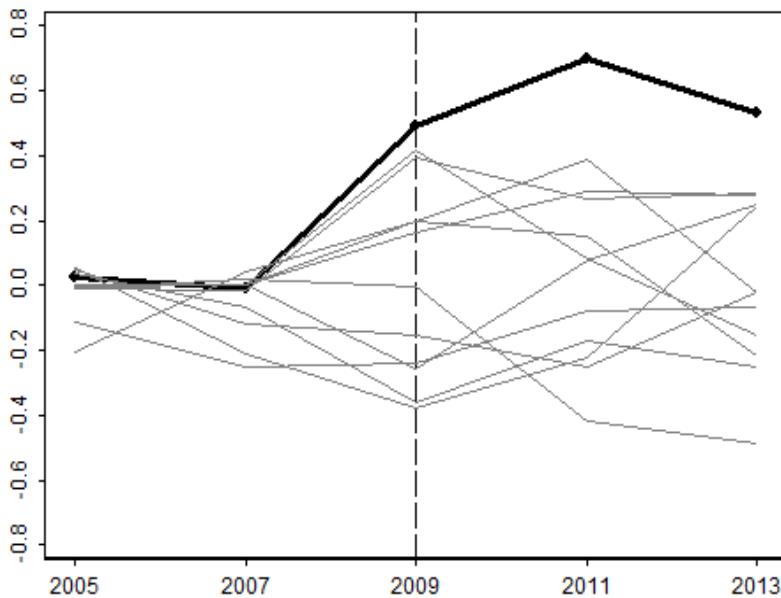


Fonte: Banco Mundial, 2017

Contudo, esse método não permite fazer nenhum tipo de inferência utilizando os métodos estatísticos tradicionais<sup>2</sup>, como o teste-t. O que é possível afirmar é se houve diferença significativa com relação a outros municípios que tiveram características iniciais e de trajetória que explicariam o IDEB semelhante ao do Rio de Janeiro. Para tanto, trata-se esse grupo colocando-o como se fosse um município que tivesse recebido a intervenção especial (placebo). Ao fazer isso, percebe-se que a diferença observada se reduz, mas ainda se mantém – uma indicação de que de fato houve algo diferente no Rio de Janeiro. A Figura 11 apresenta a diferença do IDEB para o município do Rio de Janeiro e o Rio de Janeiro Sintético (em preto) e os resultados dos testes para os 38 municípios placebo (cinza). Percebe-se que há indicação forte de que houve algo diferente.

<sup>2</sup> Não se trata de uma inferência estatística tradicional. Quando utilizado o controle sintético, a comparação, ainda que mecânica, permite afirmar que os outros municípios que teriam recebido a política (placebo) não tiveram efeito tão grande. E essa diferença é um indicativo muito forte de que a intervenção realmente tem impacto. Contudo, trata-se de um indicativo (muito forte), assim como uma inferência estatística tradicional, que também se trata de um indicativo.

Figura 11 – Teste de Placebo do IDEB: Ensino Fundamental – Anos Iniciais para o município do Rio de Janeiro, Rio de Janeiro contrafactual e municípios placebos (2005-2013)



Fonte: Banco Mundial, 2017

No entanto, ainda que se consiga perceber que tiveram alguns impactos em algumas políticas e em outras não, dado o grande conjunto de políticas (como no caso do Rio de Janeiro), é realmente difícil responder de uma forma definitiva qual, dentre todas as políticas adotadas pelo município, realmente fez a diferença. Trata-se de uma pergunta complexa.

## **QUALIDADE DA EDUCAÇÃO: RECOMENDAÇÕES DO BANCO MUNDIAL A PARTIR DE EXPERIÊNCIAS DE AVALIAÇÃO**

O que a evidência nos conta para melhorar a qualidade da educação no Brasil? O tema do *World Development Report 2018* (WDR 2018) é educação (WORLD BANK, 2017). O documento apresenta um apanhado de toda a evidência acerca de programas educacionais que funcionam e não funcionam, principalmente a partir de resultados de avaliações de impacto. Com base nessa coletânea, chegou-se a algumas conclusões, conforme seguem.

Em primeiro lugar, identificou-se que escolas excelentes são aquelas que constroem relacionamentos de ensino-aprendizagem forte em sala de aula. Em segundo lugar, é necessário ter educandos preparados para promover o desenvolvimento cerebral, por meio da nutrição e de estímulos desde a primeira infância. Em terceiro lugar, a educação requer professores capacitados e motivados e, para tanto, são necessários incentivos para atrair para o magistério pessoas talentosas e que tenham estímulos à especialização pedagógica e

temática e à formação continuada. Por fim, encontram-se mais resultados quando os recursos e a gestão estão focados no ensino-aprendizagem.

Diferentes avaliações de impacto demonstram que utilizar tecnologia na educação não necessariamente traz resultados, se não estiver realmente associada às práticas pedagógicas. No entanto, evidências apontam o resultado positivo do emprego de programas de computador que conseguem identificar claramente quais são as deficiências de cada um dos alunos e, a partir disso, focar naquelas deficiências. Como resultado, tem-se um ganho significativo geral da turma.

## CONSIDERAÇÕES FINAIS

O desenho da avaliação, bem como do sistema de monitoramento, são partes essenciais da formulação de qualquer política pública baseada em evidências. A avaliação de impacto, em específico, permite demonstrar se os programas estão atingindo os resultados pretendidos. Isto é, se um programa funciona ou não, devendo, nesse caso, pensar-se em alternativas. Caso funcione, a avaliação pode servir para que um gestor justifique a manutenção ou a adaptação do programa, pois existe evidência clara e convincente. Demonstrar a eficiência e a efetividade do programa é do interesse do gestor, especialmente em um cenário de recursos escassos. Avaliações de impacto e avaliações prospectivas podem apoiar essas decisões.

## REFERÊNCIAS

- INEP. *Índice de Desenvolvimento da Educação Básica (IDEB)*. Brasília: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep), 2015. Disponível em: <<http://portal.inep.gov.br/web/guest/IDEB>>. Acesso em: 18 abr. 2018.
- IBGE. *Pesquisa nacional por amostra de domicílios (PNAD)*. Rio de Janeiro: Instituto Brasileiro de Geografia e Estatística (IBGE), 2015. Disponível em: <[https://www.ibge.gov.br/home/estatistica/pesquisas/pesquisa\\_resultados.php?id\\_pesquisa=40](https://www.ibge.gov.br/home/estatistica/pesquisas/pesquisa_resultados.php?id_pesquisa=40)>. Acesso em 18 abr. 2018.
- WORLD BANK. *World Development Report 2018 (WDR 2018) - Learning to Realize Education's Promise*. Washington, DC: World Bank, 2017. Disponível em: <<http://www.worldbank.org/en/publication/wdr2018>>. Acesso em: 18 abr. 2018.
- GERTLER, Paul J. et al. *Impact Evaluation in Practice*. Washington, DC: Inter-American Development Bank and World Bank, 2016. Disponível em: <<https://openknowledge.worldbank.org/handle/10986/25030>>. Acesso em: 07 de maio de 2018.

## **COMUNICAÇÃO PARA O ENGAJAMENTO FAMILIAR NA VIDA ESCOLAR DOS FILHOS: UMA AVALIAÇÃO DE IMPACTO DO PROJETO EDUQMAIS<sup>\*1</sup>**

Mariana Suplicy<sup>2</sup>

**Resumo:** O artigo apresenta a avaliação de impacto do Projeto EduqMais, cujo objetivo é estimular o engajamento dos pais ou responsáveis na vida escolar dos filhos. Para tanto, buscou-se avaliar qual forma de comunicação seria a mais eficiente em produzir impacto sobre o desempenho escolar. A avaliação atestou o resultado do Projeto sobre o desempenho e a frequência escolar dos alunos, identificando o efeito nos dois eixos do EduqMais (comunicação e engajamento), sendo o segundo, mais custo-efetivo. A avaliação permitiu adequar o Projeto EduqMais, com vistas a produzir o melhor resultado.

**Palavras-chave:** Avaliação de impacto. Educação. Engajamento dos pais. Desempenho escolar.

**Abstract:** The article presents the impact evaluation of EduqMais, a project whose objective is to stimulate the engagement of parents or guardians in their children's school life. The research has sought to evaluate which form of communication would be the most efficient in producing impact on children school performance. The evaluation has attested EduqMais results on students' performance and attendance. It has been identified effect on the two axes of EduqMais (communication and engagement) –this second being more cost-effective. The impact evaluation has made possible to adapt the EduqMais, in order to produce the best result.

**Keywords:** Impact evaluaton. Education. Engagement of parentes. School performance.

---

<sup>\*</sup>Texto elaborado a partir da transcrição da fala e da apresentação em slides de palestras proferida no Workshop Avaliação de Impacto de Políticas Públicas, realizado nos dias 29 e 30 de novembro de 2017, no Auditório da Secretaria de Planejamento, Governança e Gestão (SPGG-RS), com o apoio do Banco Mundial. Versão *não* validada pela palestrante.

<sup>1</sup> A palestra baseou-se em pesquisa originalmente publicada por CUNHA, N.; LICHAND, G.; MADEIRA, R.; BETTINGER, E. What is It About Communicating With Parents?. Working Paper, Stanford University, Stanford-CA, Out. 2017. Disponível em: <[https://cepa.stanford.edu/sites/default/files/cunha\\_cover\\_paper1.pdf](https://cepa.stanford.edu/sites/default/files/cunha_cover_paper1.pdf)>. Acesso em: 28 mai. 2018.

<sup>2</sup> Gestora de operações do MGov Brasil, startup de comunicação para desenvolvimento em impacto social.

## INTRODUÇÃO

O baixo nível educacional dos alunos brasileiros é preocupante. Na avaliação do PISA (Programa Internacional de Avaliação de Alunos)<sup>3</sup> referente ao ano de 2015, dentre os 70 países analisados, o Brasil figurou em 63<sup>a</sup> posição em ciências, 59<sup>a</sup> em leitura e 66<sup>a</sup> em matemática (OCDE, 2016). Como resultado, 90% dos adolescentes brasileiros chegam ao fim do Ensino Médio sem as habilidades mínimas esperadas para essa etapa.

Em meio a essa realidade, identifica-se um desafio específico da educação, que é a falta de engajamento dos pais ou responsáveis na vida educacional dos filhos. Conforme dados de 2015 da Pesquisa Nacional de Saúde do Escolar (PeNSE), 25% dos pais/responsáveis não sabem se os filhos faltaram às aulas, 33% não perguntam para os filhos sobre os problemas deles, e 50% não têm o hábito de verificar se seu filho fez o dever de casa (BRASIL, 2016). Evidências apontam que informar e engajar impacta a aprendizagem das crianças (*i.e.* KRAFT; DOUGHERTY, 2013; BERLINSKI, S. *et al*, 2016).

O presente artigo apresenta a avaliação de impacto do Projeto EduqMais da Mgov Brasil, que visa justamente estimular o engajamento dos pais ou responsáveis na vida escolar dos filhos por meio do envio de mensagens SMS. Com tanto, objetiva-se melhorar o desempenho dos alunos, de modo a alcançarem maiores níveis de escolaridade, empregabilidade e qualidade de vida no futuro.

O experimento foi realizado com o apoio da Universidade de Stanford e da Universidade de Zurique, abrangendo uma amostra de escolas da rede estadual do Estado de São Paulo. Nas seções seguintes, apresentaremos o tema da comunicação e do engajamento escolar na vida escolar dos filhos, os mecanismos causais esperados que fundamentaram o Projeto EduqMais, o desenho da avaliação de impacto, as formas de coleta de dados e, por fim, os resultados alcançados.

## COMUNICAÇÃO E TECNOLOGIA PARA O ENGAJAMENTO FAMILIAR NA VIDA ESCOLAR DOS FILHOS

Há diferentes formas de promover e facilitar o engajamento e a participação dos pais ou responsáveis na vida educacional dos filhos. A informação é uma delas. Experiências internacionais evidenciam que a comunicação e a informação influenciam o comportamento dos pais em relação à vida escolar dos filhos e, por extensão, os filhos também mudam seu comportamento em relação à escola. Algumas experiências apostam no envio de mensagens de texto para os pais, com conteúdos diversos, tais como “Sr. Pai/Mãe, a Fulana faltou à aula de matemática ontem”.

---

<sup>3</sup> Avaliação internacional que mede o nível educacional de jovens de 15 anos por meio de provas de Leitura, Matemática e Ciências.

O impacto dessa informação sobre o comportamento dos pais pode ser dividido em duas vertentes. Uma primeira vertente diz respeito à redução do custo de monitoramento da vida escolar dos filhos. Isso porque, o envio de mensagens reduz a necessidade de os pais irem até à instituição de ensino ou questionarem seus filhos sobre sua frequência ou desempenho escolar. Uma segunda vertente, por sua vez, refere-se à percepção da relevância desse tipo de informação por parte dos pais e, consequentemente, da importância de sua atenção em relação à vida escolar do filho.

Diante das alternativas, cabe questionar qual a melhor forma para disponibilizar essas informações, considerando os custos e os benefícios quanto ao monitoramento da vida escolar dos alunos. Assim como há várias formas de promover e facilitar o engajamento e a participação dos pais, há também várias formas de comunicar-se com eles. Por um lado, oferecer uma informação correta e oportuna sobre cada aluno pode facilitar em muito o monitoramento por parte dos pais, mas exige aportes significativos de recursos. Confeccionar uniformes escolares munidos de *chip* para rastreio do aluno a fim de, por exemplo, informar os pais, via SMS, quando o aluno entrou na escola, pode ser bastante impactante, mas essa operação não é custo-efetiva.

Outra opção é fornecer o que se chama de *nudge*<sup>4</sup>. *Nudges* podem resultar em efeitos maiores no sentido esperado, uma vez que servem de estímulo à participação e ao engajamento dos pais. Além disso, permitem uma maior interatividade, com maior frequência. Isso porque não dependem da materialidade de um fato especificamente sobre uma criança ou um adolescente para, então, entrar-se em contato com seus pais. Por exemplo, se em determinada semana o aluno não tem prova, não haveria o porquê enviar uma mensagem específica sobre a prova para os pais. Porém, podem-se enviar mensagens diversas, a fim de lembrar os pais de perguntarem como o filho está indo na escola, quais conteúdos tem aprendido, de quais tem mais gostado, se tem frequentado as aulas da disciplina ‘x’, entre outras possibilidades de lembretes.

Embora menos custosos, os “empurroezinhos” podem não ser tão efetivos em termos de engajamento dos pais, se comparados à informação precisa. Cabe, pois, avaliar as diferentes abordagens.

## TEORIA DA MUDANÇA: RESULTADOS ESPERADOS DO EDUQMÁIS

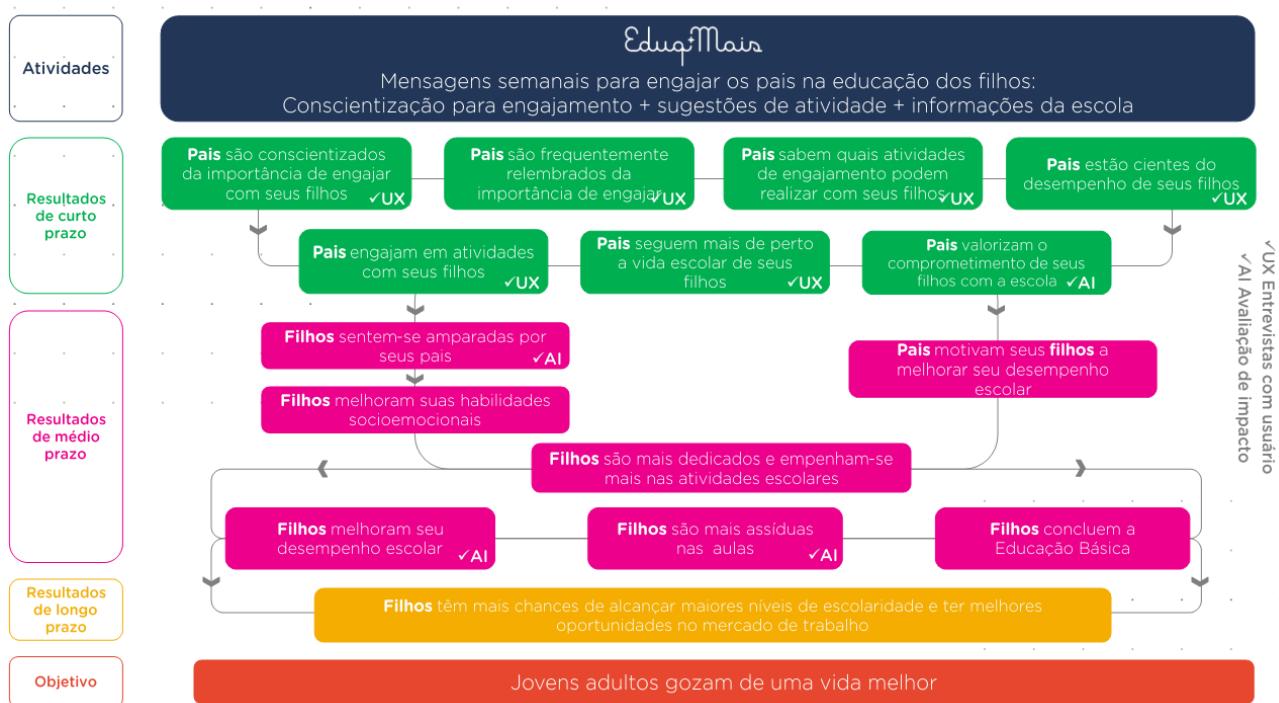
O Projeto EduqMais consiste na implantação de uma plataforma para facilitar a comunicação com os pais e incentivar hábitos de engajamento na vida escolar dos filhos. Sua formulação fundamentou-se na seguinte cadeia de resultado esperada (Figura 1), elaborada a partir da teoria da mudança<sup>5</sup>.

<sup>4</sup>Espécie de “empurrão” ou sugestões indiretas como forma de estimular reforços positivos e, assim, influenciar o comportamento e a tomada de decisão de grupos e/ou indivíduos.

<sup>5</sup> O MGov Brasil disponibiliza gratuitamente um sistema *online*, acessível por meio do endereço eletrônico: <http://mgovbrasil.com.br/teoriademudanca>. Acesso em: 28 mai. 2018.

No curto prazo, espera-se que o envio de mensagens para pais (ou responsáveis) os conscientize quanto à importância de receber informações sobre a vida escolar dos filhos e de se engajar na mesma, alterando seu comportamento. Com tanto, esses pais passariam a interessar-se mais pela vida escolar dos filhos, incentivando-os a não faltarem ou não se atrasarem para as aulas, bem como os apoioando, por exemplo, na realização das tarefas de casa. No médio prazo, espera-se que os pais motivem mais seus filhos para que melhorem seu desempenho, ao passo que os filhos, amparados pelos pais, melhorem suas habilidades socioemocionais, o que resultaria numa maior dedicação e empenho nas atividades escolares. Assim, os filhos melhorariam a assiduidade e a pontualidade nas aulas, bem como seu desempenho escolar, concluindo com êxito o ciclo de Educação Básica. Por fim, no longo prazo, o Projeto EduqMais espera que os alunos/filhos alcancem maiores níveis de escolaridade e, consequentemente, obtenham melhores oportunidades no mercado de trabalho, gozando de uma vida com melhor qualidade quando adultos.

Figura 1 – Teoria da Mudança do Projeto EduqMais



Fonte: Mgov Brasil, 2017.

## AVALIAÇÃO DE IMPACTO DO PROJETO EDUQMAIS

Em virtude das diferentes possibilidades de comunicação para promover a participação e o engajamento dos pais ou responsáveis na vida escolar dos filhos, coube avaliar qual forma de comunicação seria a mais eficiente em produzir impacto sobre o desempenho escolar dos

alunos. Para tanto, o MGov realizou um experimento, com o apoio da Universidade de Stanford e da Universidade de Zurique.

Com a finalidade de evitar vieses de autosseleção dos pais/responsáveis participantes, procedeu-se a um **experimento aleatório**. Dessa forma, buscou-se evitar que somente pais dispostos a participar do projeto fossem avaliados, dado que a simples disposição de participar do mesmo já evidenciaria uma maior preocupação com a educação dos filhos.

A avaliação ocorreu em etapas. Primeiramente, procedeu-se a um processo de adesão ao projeto na rede de ensino estadual de São Paulo<sup>6</sup>, a partir da qual foi extraída uma amostra de 287 escolas e 19.300 famílias com filhos cursando os anos finais do Ensino Fundamental (9º ano). A amostra foi composta em grande maioria por famílias cujos responsáveis eram mulheres (82%), frequentaram pelo menos o Ensino Fundamental (98%), tinham entre 25 e 44 anos (72%) e detinham renda mensal média de 1 a 3 salários mínimos (45%). A partir dessa amostra, realizou-se a aleatorização para definir os grupos de tratamento e controle, sendo cada qual subdividido em diferentes variações, a fim de verificar qual seria mais eficiente.

O Projeto EduqMais consistia na implantação de um sistema específico para facilitar a comunicação com os pais e incentivar hábitos de engajamento. Além do envio de informações sobre os alunos e de sugestões de atividades, a plataforma também detinha uma ferramenta de eventos, acessível a todas as escolas (tratamento e controle), por meio da qual os pais eram avisados sobre reuniões de pais, por exemplo. Essa possibilidade de envio de comunicações sobre eventos gerais facilitou o convencimento dos pais para que se cadastrasse no sistema – sem saber, contudo, se fariam ou não parte do tratamento. As comunicações tinham frequência semanal e foram realizadas ao longo de 23 semanas contínuas, de junho a dezembro de 2016.

Em relação ao tratamento, em específico, foram estabelecidos dois eixos do EduqMais: comunicação e engajamento<sup>7</sup>.

O **Eixo Comunicação** (Figura 2) consistia em mensagens semanais sobre frequência e desempenho dos alunos, inseridas pelas escolas na plataforma e enviadas automaticamente via SMS para os pais ou responsáveis uma vez por semana. Esse eixo subdividiu-se em diferentes intensidades de tratamento. Um grupo recebeu apenas mensagens de **conscientização**, como, por exemplo: “Faltar reduz a aprendizagem do seu filho” ou “Fazer a tarefa é importante”. Outro grupo recebeu **informações individualizadas** por aluno/filho, com informações sobre faltas, atrasos, tarefas, entre outras. Essas informações eram inseridas no sistema pelo próprio professor, gerando mensagens disparadas automaticamente, tais como: “A Júlia faltou na aula de matemática ontem”. Um terceiro grupo, ainda, recebeu **informações**

<sup>6</sup> Para a realização do experimento, a MGov pactuou com a Secretaria de Educação do Estado de São Paulo. Por meio da Secretaria, buscou-se pactuar com as diretorias regionais e, por sua vez, com as escolas. Durante o teste piloto, o EduqMais constituiu-se num projeto da própria Secretaria, chamado SMS Escola.

<sup>7</sup> Cabe aqui um esclarecimento. Por meio desse desenho, o experimento consegue apenas medir o **efeito da intenção de tratar** e não diretamente o efeito do tratamento em si. Isso porque, consegue-se monitorar o envio de mensagens, mas não se consegue monitorar se, de fato, o pai ou o responsável leu a mensagem. Por meio das chamadas automáticas de voz, procurou-se verificar isso perguntando se o pai/responsável recebeu ou não a mensagem, se leu ou não. Contudo, a resposta é autorreportada, não resolvendo por completo o problema do efeito da intenção de tratar vs. efeito do tratamento.

relativas do aluno em relação aos colegas. Por exemplo: “Seu filho faltou 3 vezes na última semana, e isso foi muito mais do que o resto da sala. O resto da sala faltou 1 vez”. Essa última modalidade buscou facilitar a percepção da relevância da informação por parte dos pais. Isso porque, saber apenas que o filho faltou na escola, por exemplo, pode não deixar claro se faltar uma vez é algo com que se deva preocupar ou não. Ou seja, não é claro se essa informação é importante ou não em um contexto maior. Outros alunos podem ter faltado no mesmo dia, por exemplo.

Figura 2 – Projeto EduqMais: Eixo Comunicação



Fonte: Mgov Brasil, 2017.

Por sua vez, o **Eixo Engajamento** (Figura 3) consistia no envio de *nudges* com mensagens genéricas, seguidas de sugestões de atividades simples que os pais poderiam realizar para apoiar o desenvolvimento socioemocional dos filhos. As atividades não exigiam conhecimento do conteúdo curricular e variavam conforme a idade dos filhos. Para a redação do conteúdo, contou-se com a curadoria de uma instituição especializada, o Instituto Ayrton Senna. As mensagens eram enviadas duas vezes por semana para cada participante.

Figura 3 – Projeto EduqMais: Eixo Engajamento



Fonte: Mgov Brasil, 2017.

Além dos diferentes tipos de mensagem, também foram testados diferentes **arranjos de envio de mensagens**, a fim de verificar quais produziriam maiores impactos. Assim, as mensagens variaram quanto:

- (i) Frequência: 0, 1, 2 ou 3 vezes na semana;
- (ii) Hora do dia: à tarde ou à noite;
- (iii) Alternâncias: horário constante ou horários variados;
- (iv) Interatividade: fluxo de mensagens que pergunta se os pais cumpriram a atividade sugerida.

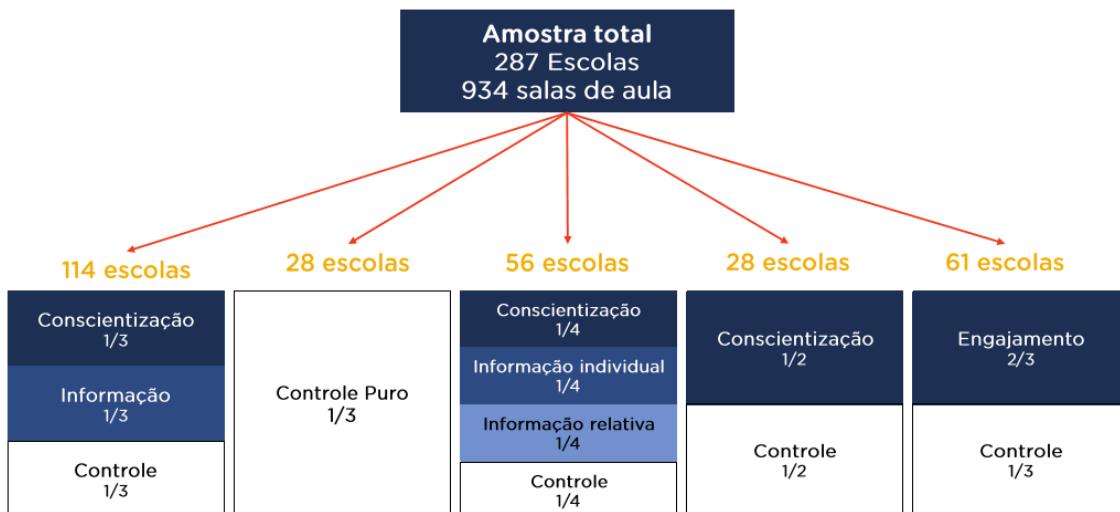
Além dos grupos de tratamento, por fim, foi composto o **grupo controle**, formado por famílias que não recebiam nenhum tipo de informação. O grupo controle também foi subdividido: por um lado, pais de alunos cujas **escolas não participavam do projeto** (controle puro) e, por outro lado, pais de alunos cujas **escolas participavam do projeto, mas não recebiam informações**. Essa subdivisão buscava tornar evidente se a mudança de comportamento de um aluno cujos pais recebiam mensagens afetaria o comportamento de alunos da mesma escola, mas cujos pais não recebiam mensagens. Isto é, buscou captar as externalidades: o efeito de quem está sendo tratado em quem não é tratado.

Para limpar os dados do controle contaminado, procedeu-se a uma técnica quase-experimental chamada diferenças-em-diferenças, que buscou captar o efeito do Eixo Engajamento a partir das diferentes tendências apresentadas antes e depois do grupo controle puro e do grupo controle contaminado. O Eixo Engajamento foi escolhido para evitar que o efeito fosse produzido não pela mudança de comportamento dos pais em relação à

participação na vida escolar dos filhos, mas dos professores – que no Eixo Comunicação inseriam informações de cada aluno na plataforma<sup>8</sup>.

A Figura 4 ilustra a divisão da amostra entre os diferentes tipos e intensidades de tratamento aplicados aos grupos.

Figura 4 – Grupos de tratamento e grupos controle



Fonte: Cunha, Lichand, Madeira e Bettinger (2017).

## O PROCESSO DE COLETA DE DADOS PARA A AVALIAÇÃO

A fim de coletar os dados primários para elaborar os indicadores e, assim, buscar validar os mecanismos causais esperados, foram utilizadas as seguintes fontes.

Primeiramente, no que se refere ao comportamento dos pais, foram realizadas chamadas automáticas de voz, no início e no final do projeto. Nessas chamadas, questionava-se, primeiro lugar, se os pais tinham conhecimento sobre o desempenho dos filhos na escola. Por exemplo, perguntava-se: “Quanto você acha que seu filho faltou no último semestre?” ou “Qual você acha que é a nota do seu filho?”.

Além das chamadas automáticas de voz, buscou-se captar também a experiência do usuário (*user experience, UX*) por meio de chamadas telefônicas diretas. Isso porque, a chamada automática de voz serve à coleta de muitos dados, por meio da digitação de respostas no teclado numérico do aparelho, mas não consegue ser específica quanto à experiência qualitativa dos pais. Já nas chamadas diretas se questionavam algumas perguntas

<sup>8</sup> Como visto, no Eixo Engajamento, as mensagens são genéricas e enviadas diretamente aos pais ou responsáveis, sem intermediação da escola.

no intuito de captar o valor percebido pelos pais em saber informações sobre a vida educacional dos filhos.

Além dessas chamadas, foi realizada uma pesquisa presencial com os alunos no final do projeto.

Em relação aos dados sobre a aprendizagem dos alunos, foram utilizados registros administrativos da Secretaria da Educação do Estado de São Paulo e os resultados do Sistema de Avaliação de Rendimento Escolar do Estado de São Paulo (SARESP), uma avaliação padronizada de larga escala realizada no estado. Essas fontes foram empregadas em virtude de se tratar de ano em que não houve Prova Brasil (MEC). O SARESP, em específico, serviu para captar efeitos de médio prazo, tendo em vista que a avaliação ocorreu num período de 6 meses e que, portanto, não haveria viabilidade de verificar resultados de longo prazo (escolaridade, empregabilidade, qualidade de vida).

Por fim, dados sobre assiduidade e aprovação dos alunos foram coletados em bancos de dados administrativos da Secretaria, bem como por meio do sistema implantado em cada escola participante do projeto.

## RESULTADOS DA AVALIAÇÃO DE IMPACTO DO PROJETO EDUQMAIS

Por meio das chamadas diretas (UX), foi possível captar comentários e percepções que serviriam para validar os mecanismos causais esperados quanto à mudança de comportamento dos pais. Por exemplo, quanto à conscientização dos pais sobre a importância de engajar-se, um pai comentou:

“Era difícil e com a história dessas mensagens, das perguntas a gente vai se interessando mais, que é um compromisso que a gente tem como os nossos filhos a tentar ajudá-los, é bom.”

Sobre serem lembrados com frequência da importância do engajamento, um dos pais comentou:

“Me mandando a mensagem, vocês estão me estimulando a fazer isso. Que às vezes, no dia a dia, a gente acaba esquecendo de fazer esse tipo de pergunta. Facilita.”

Quanto ao apoio na realização de atividades:

“Às vezes a gente fica até sem ideias, né, do que falar, e aí as mensagens chegam e ajudam”.

Sobre estar mais perto da vida escolar dos filhos:

“Eu acho bom, porque daí eu fico mais atualizada, porque eu trabalho. Quando falou que ela não tava(sic) fazendo todas as tarefas da escola, eu fui falar com ela. Aí fomos ver e não tinha terminado só um trabalho”.

Bem como em relação à conscientização quanto à importância de conhecer o desempenho escolar de seus filhos:

“Eu não posso ir na (*sic*) reunião, porque eu trabalho, sou diarista. Pra mim fica mais fácil. Eu vejo as notas dele, eu vejo o que a professora escreve dele. Então para mim está sendo bom esse projeto”.

*Feedbacks* sobre o engajamento nas atividades dos filhos também foram registrados:

“Às vezes, na correria do dia a dia, acaba se distanciando um... e com as ficas do EduqMais, eu me aproximei e percebi o quanto é importante, porque ela realmente... a gente vê a felicidade no olhinho da criança”.

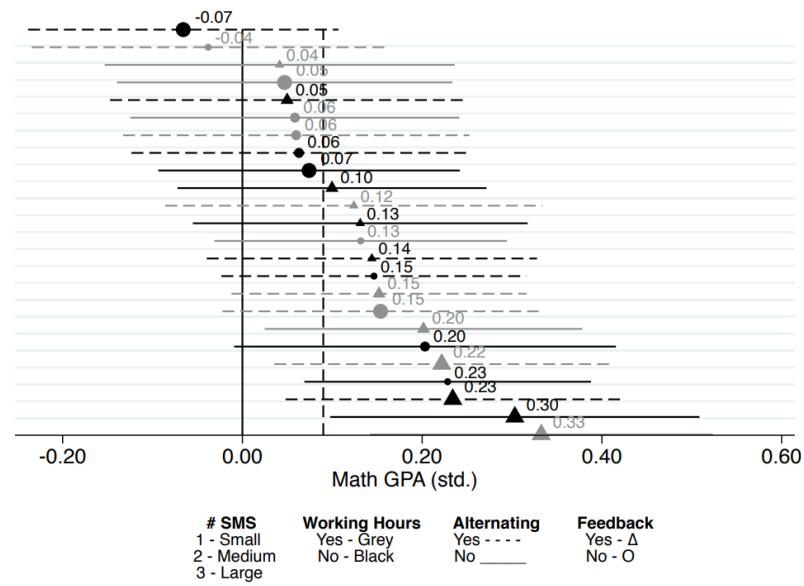
A pesquisa presencial com os alunos também identificou que 94% dos alunos com EduqMais afirmam que seus pais os incentivavam a estudar, perguntando como estava sendo a aula, acompanhando a lição ou a tarefa de casa. Trata-se de porcentagem muito maior do que a registrada no grupo que não recebeu o tratamento. Quanto à presença dos pais nas reuniões escolares, 70% dos alunos participantes do EduqMais disseram que seus pais as frequentam sempre ou às vezes, ao passo que apenas 61% dos alunos sem EduqMais afirmaram o mesmo.

Quanto ao comprometimento dos alunos, 80% dos que receberam o EduqMais consideraram importante não se atrasar para as aulas, porcentagem que cai a 77% dentre os alunos sem EduqMais.

Quanto aos resultados do desempenho escolar dos alunos, fez-se a validação do tratamento a partir dos dados do SARESP. Tanto o Eixo Comunicação, que correspondia à inserção de informações pela escola no sistema e o envio de mensagens os pais, quanto o Eixo Engajamento, que era simplesmente os empurrõezinhos e as atividades sugeridas para os pais, produziram resultado sobre a aprovação dos alunos no SARESP em Matemática e em Língua Portuguesa. Em específico, verificou-se que alunos com EduqMais tiveram desempenho 1,5% maior em avaliações padronizadas.

A Figura 5 apresenta os resultados a partir do desvio-padrão encontrado na nota de matemática alcançada pelos alunos, conforme os diferentes arranjos de envio de mensagens (frequência, hora do dia/de trabalho, hora alternada, interações). Já o Quadro 1 resume os resultados para ambas as notas em ambas as disciplinas.

Figura 5 – Avaliação Padronizada de Matemática



Fonte: Cunha, Lichand, Madeira e Bettinger (2017).

Quadro 1 – Efeito estimado em aprovação e nota

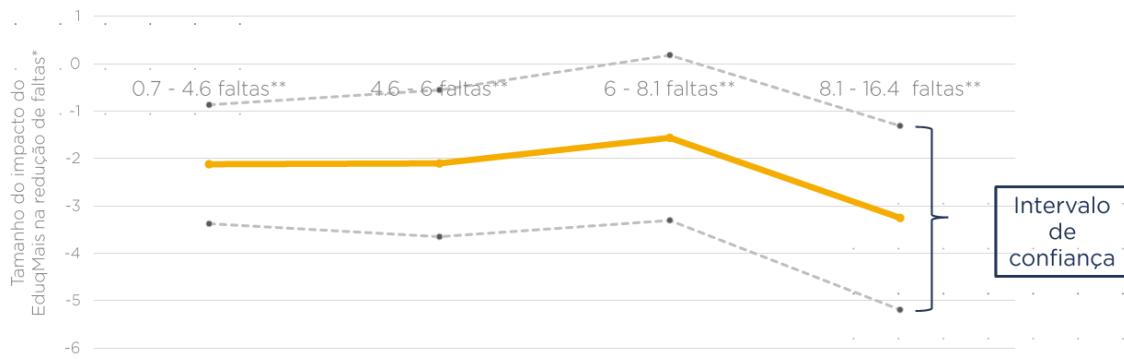
|                         | Aprovação | SARESP (em desvio padrão) |                   |
|-------------------------|-----------|---------------------------|-------------------|
|                         |           | Matemática                | Língua Portuguesa |
|                         |           |                           |                   |
| <b>Eixo Comunicação</b> | 0,03**    | 0,08*                     | 0,04              |
| <b>Eixo Engajamento</b> | 0,04**    | 0,08                      | 0,03              |

Nota:  $p$ -valor: \*\*\* <0,01 ; \*\* <0,05; \* <0,10.

Fonte:Cunha, Lichand, Madeira e Bettinger (2017).

Por sua vez, a partir dos registros administrativos da Secretaria da Educação, foi possível verificar que alunos cujos pais receberam o tratamento faltaram 10% menos às aulas. A Figura 6 apresenta os resultados sobre as faltas às aulas de Matemática no segundo semestre do ano de interesse. O efeito foi significativo, mesmo em escolas que tinham altos índices de assiduidade no início do ano.

**Figura 6 – Efeito do EduqMais sobre faltas**



Fonte: Cunha, Lichand, Madeira e Bettinger (2017).

Por fim, estimou-se o impacto do tratamento sobre o Índice de Desenvolvimento da Educação Básica (IDEB). Considerando o aumento na proficiência dos alunos, verificado por meio dos dados do SARESP e da taxa de aprovação, estimou-se que o EduqMais tem potencial de aumentar o Índice em 0,4 ponto<sup>9</sup>. Esse efeito teria sido suficiente para o Estado de São Paulo bater a meta do IDEB para os anos finais do Ensino Fundamental em 2015, que era de 5,0 pontos, tendo o Estado atingido 4,7.

## **CONCLUSÕES DA AVALIAÇÃO E A REFORMULAÇÃO DO EDUQMAIS**

Em virtude das diferentes possibilidades de informação para promover a participação e o engajamento dos pais ou responsáveis na vida escolar dos filhos, buscou-se avaliar qual forma de comunicação seria a mais eficiente em produzir impacto sobre o desempenho escolar, medido por meio de indicadores educacionais tais como aprendizado, notas, taxa de aprovação, taxa de repetência, entre outros.

Neste trabalho, apresentou-se a avaliação de impacto do Projeto EduqMais, a partir de um experimento aleatório realizado em uma amostra das escolas da rede pública estadual do Estado de São Paulo. A avaliação atestou o resultado do Projeto sobre o desempenho e o atendimento escolar dos alunos. O efeito foi identificado nos dois eixos do EduqMais (comunicação e engajamento), sendo o segundo, contudo, mais custo efetivo. Isso porque, o eixo engajamento não depende da inserção de informações em cada escola, apenas do envio de “empurroezinhos” (*nudges*) genéricos somado a sugestões de atividades.

Os resultados obtidos na avaliação conduziram à reformulação do EduqMais. Visando ao melhor custo-efetividade, passou-se a mandar dois SMS por semana, com interação. Segue,

<sup>9</sup> Para este cálculo, assume-se a aditividade dos efeitos nos eixos engajamento e comunicação. Em caso de violação dessa hipótese, considera-se que o efeito esteja entre 0,2 e 0,3 ponto no IDEB.

contudo, a possibilidade de escolas manterem o sistema, a fim de enviar informações detalhadas, se assim o desejarem.

## REFERÊNCIAS

- BERLINSKI, S. et al. *Reducing parent school information gaps and improving education outcomes: Evidence from high frequency text messaging in Chile*. 2016. Mimeo.
- BRASIL. Instituto Brasileiro de Geografia e Estatística. *Pesquisa nacional de saúde do escolar - 2015*. Rio de Janeiro: 2016. Disponível em: <<https://biblioteca.ibge.gov.br/visualizacao/livros/liv97870.pdf>>. Acesso em: 28 mai. 2018.
- CUNHA, N.; LICHAND, G.; MADEIRA, R.; BETTINGER, E. What is It About Communicating With Parents?. *Working Paper*, Stanford University, Stanford-CA, October, 2017. Disponível em: <[https://cepa.stanford.edu/sites/default/files/cunha\\_cover\\_paper1.pdf](https://cepa.stanford.edu/sites/default/files/cunha_cover_paper1.pdf)>. Acesso em: 28 mai. 2018.
- KRAFT, M. A.; DOUGHERTY, S. M. The effect of teacher–family communication on student engagement: Evidence from a randomized field experiment. *Journal of Research on Educational Effectiveness*, 6 (3), 2013, p. 199–222.
- OECD. *PISA Results 2015 - Programme for International Student Assessment*. Paris: OECD Publishing, 2016. Disponível em: <<http://www.oecd.org/pisa/>>. Acesso em: 28 mai. 2018.
- MGOV BRASIL. *Eduq+*. Comunicação para o Engajamento de Famílias. Avaliação de Impacto – 2017. Apresentação. Workshop de Avaliação de Impacto de Políticas Públicas. Governo do Estado do Rio Grande do Sul, Porto Alegre, Dezembro, 2017.

## AVALIAÇÃO DE IMPACTO DE POLÍTICAS PÚBLICAS NA ÁREA DE GOVERNANÇA\*

Vincenzo Di Maro<sup>1</sup>

**Resumo:** O texto discute o tema da avaliação de impacto na área de governança, agenda levada ao cabo pelo Programa de Avaliação de Impacto na Governança e Justiça do Setor Público do Banco Mundial. A avaliação de impacto em governança não só é factível, mas também muito promitente. Por essa razão, estudos sobre o tema vêm sendo elaborados. Exemplos nesse sentido são apresentados, demonstrando a importância de focar em pequenas mudanças nos programas, não necessariamente em grandes perguntas. Pequenas mudanças podem ter um impacto muito grande no âmbito da governança.

**Palavras-chave:** Avaliação de impacto. Governança. Gestão por resultados.

**Abstract:** The text discusses the impact evaluation in the area of governance, an agenda carried out by the World Bank's Impact Evaluation of Public Sector Governance and Justice Program. Impact evaluations are not only feasible, but also very promising. For this reason, researches on governance subjects have been elaborated. Experiences examples in this sense are presented, which demonstrate the importance of focusing on small changes in programs, and not necessarily on big questions. Small changes may have a very large impact on governance.

**Keywords:** Impact evaluation. Governance. Managing for results.

---

\* Texto elaborado a partir da transcrição da fala e da apresentação em slides de palestras proferida no Workshop Avaliação de Impacto de Políticas Públicas, realizado nos dias 29 e 30 de novembro de 2017, no Auditório da Secretaria de Planejamento, Governança e Gestão (SPGG-RS), com o apoio do Banco Mundial.

<sup>1</sup> Economista Sênior do Grupo de Pesquisa para o Desenvolvimento do Banco Internacional para Reconstrução e Desenvolvimento (BIRD). Coordena o Programa de Avaliação de Impacto na Governança e Justiça (*ieGovern*).

## INTRODUÇÃO

O presente artigo discute avanços na avaliação de impacto na área de governança. Primeiramente fornece uma introdução aos conceitos, evidencia a escassez de estudos nessa área e apresenta as atividades do Programa do Banco Mundial de Avaliação de Impacto na Governança e Justiça. São apresentadas algumas experiências de avaliação de impacto no setor. Por fim, enfatiza-se o apoio do Banco Mundial para fortalecer a avaliação de políticas públicas em governança no Brasil.

## PROGRAMA DE AVALIAÇÃO DE IMPACTO NA GOVERNANÇA E JUSTIÇA DO SETOR PÚBLICO

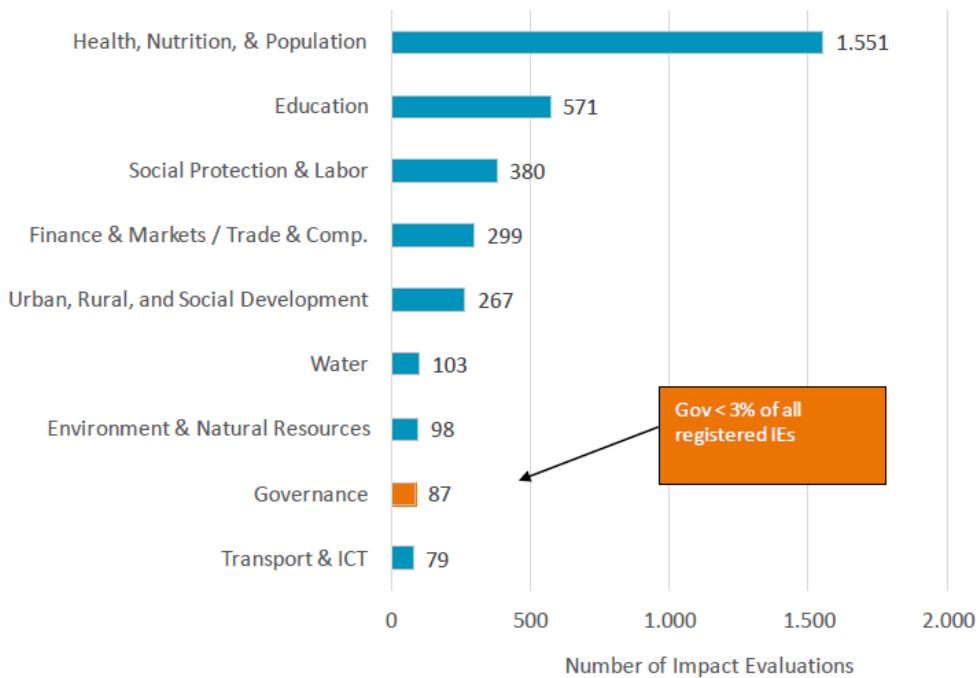
A área da governança não costuma ter muitas avaliações de impacto. Representa menos de 3% de todas as avaliações que estão registradas no 3ie<sup>2</sup>, com dados de 2015 e 2016, enquanto avaliações nas áreas de saúde, educação, emprego, mercado do trabalho, etc. são abundantes (Figura 1). Diante desse cenário, o Banco Internacional para Reconstrução e Desenvolvimento (BIRD) preocupou-se em estender o número de avaliações de impacto e estender também o campo de aplicação.

O Programa de Avaliação de Impacto na Governança e Justiça (*ieGovern*) enfoca basicamente cinco campos: servidor público, aquisições públicas, mobilização de recursos (impostos, tributos, recursos de governos locais, etc.), justiça e governos locais (orçamento participativo, etc.). A avaliação de impacto na área da justiça é igualmente escassa, assim como a pesquisa em geral nesse âmbito. O enfoque nesses cinco campos visa à qualidade da prestação de serviços, à prestação de contas, à justiça eficiente e inclusiva, à transparência, à melhoria do ambiente de negócios e dos gastos públicos.

---

<sup>2</sup>A Iniciativa Internacional para Avaliação de Impacto (3ie) administra o Registro de Avaliações de Impacto sobre Desenvolvimento Econômico Internacionais (*Registry for International Development Impact Evaluations – RIDIE*), que se concentra em avaliações de impacto relacionadas ao desenvolvimento em países de baixa e média renda (GERTLER *et al*, 2018).

Figura 1 – Registro de Avaliações de Impacto sobre Desenvolvimento Econômico Internacionais  
– RIDIE – 2015-2016



Fonte: Iniciativa Internacional para Avaliação de Impacto (3ie).

Políticas de reformas do setor público podem poupar grandes valores em termos de gasto público, podendo a economia gerada posteriormente ser alocada em outros programas ou políticas. Logo, a avaliação de impacto não necessariamente funciona apenas para programas pequenos. Em alguns casos os programas podem ser pequenos e apresentar o potencial de gerar economias bastante grandes em termos de gasto público. Essa é uma visão otimista. Existem pessoas que têm uma visão mais pessimista, que entendem ou não ser possível reformar aspectos da governança, ou que levaria tempo demasiado.

A ideia do Programa de Avaliação de Impacto é concordar com essas premissas, mas não concordar com a conclusão de que não é factível reformar. É claro que, pensando em reformar a governança de um país, a única maneira é mudar as instituições estruturais do país, e isso leva muito tempo. E talvez até requeira, como a história nos ensina, eventos históricos, revolução ou escândalos, que conseguem alterar a trajetória de algumas políticas. Nesses eventos, os escândalos são os mais frequentes, e as revoluções são muito menos frequentes: a cada 200 ou 300 anos. Portanto, o melhor caminho para reformar o que parece não reformável é trabalhar em uma série de pequenas mudanças baseadas em evidências – mudanças incrementais e graduais, mas que, juntas, resultem em uma série de políticas coordenadas, que podem gerar uma mudança geral muito grande.

Nesse sentido, a abordagem do Programa aposta muito mais sobre uma série de pequenas mudanças, em vez de falar sobre reformar instituições em uma perspectiva mais macro. O Quadro 1 traz exemplos resumidos de pequenas ações nos seguintes campos: serviço civil (serviço público), impostos, justiça e governos locais. No primeiro caso, trata-se de ações

de reformas na contratação, motivação e treinamentos de servidores públicos. São dimensões clássicas, nas quais mecanismos inovadores estão sendo testados. No que diz respeito à aquisição, a inovação se encontra especialmente nas compras públicas eletrônicas.

Por sua vez, na área de recursos públicos, existe um esforço para mudar a disposição que as pessoas têm para pagar impostos e tributos. Já no setor judicial, há um empenho em criar incentivos para servidores públicos que trabalham na justiça, incluindo juízes. Esse é um trabalho muito interessante porque não é fácil trabalhar com juízes e convencê-los a usarem dados. E é isso que se tenta fazer, ao mostrar, por exemplo, dados sobre o atraso de cada juiz por tipo de caso julgado. Um juiz poderia ser bem mais eficiente, mas simplesmente não possuir a informação quanto a seus atrasos por tipo de caso. Essa informação pode mudar a sua percepção e melhorar seu rendimento. Também há avaliações em relação à assistência judicial, que em muitos países é destinada para as pessoas mais vulneráveis. Hoje, há um alto investimento na transformação da disposição dos dados das cortes e do jurídico para dados eletrônicos. Isso gera mais eficiência e também muita informação, que pode vir a ser usada para mostrar aos próprios juízes como e para quê utilizá-las.

Por fim, registram-se avaliações relativas às áreas de governos locais, especialmente na questão de responsabilidade social, onde o cidadão está envolvido no processo de prestação de contas. Nesse sentido, um exemplo muito importante que nasceu na cidade de Porto Alegre é o Orçamento Participativo. Também há a tentativa de aprimorar o processo de descentralização, bem como, de maneira mais geral, de criar um ambiente melhor para o engajamento do cidadão nas políticas públicas.

O Quadro 1 apresenta um resumo dos tipos de intervenções sendo estudadas no Programa de Avaliação de impacto na governança.

**Quadro 1 – Tipos de intervenções sendo estudadas no Programa de Avaliação de impacto na governança**

| Civil Service | Procurement/Tax                      | Justice (De Jure)               | Subnational government |
|---------------|--------------------------------------|---------------------------------|------------------------|
| Hiring        | Electronic procurement               | Nudges for court staff (judges) | Social accountability  |
| Motivation    | Framework agreements (acordo-quadro) | Legal aid                       | Decentralization       |
| Training      | Tax morale/Compliance                | Electronic case management      | Citizen Engagement     |

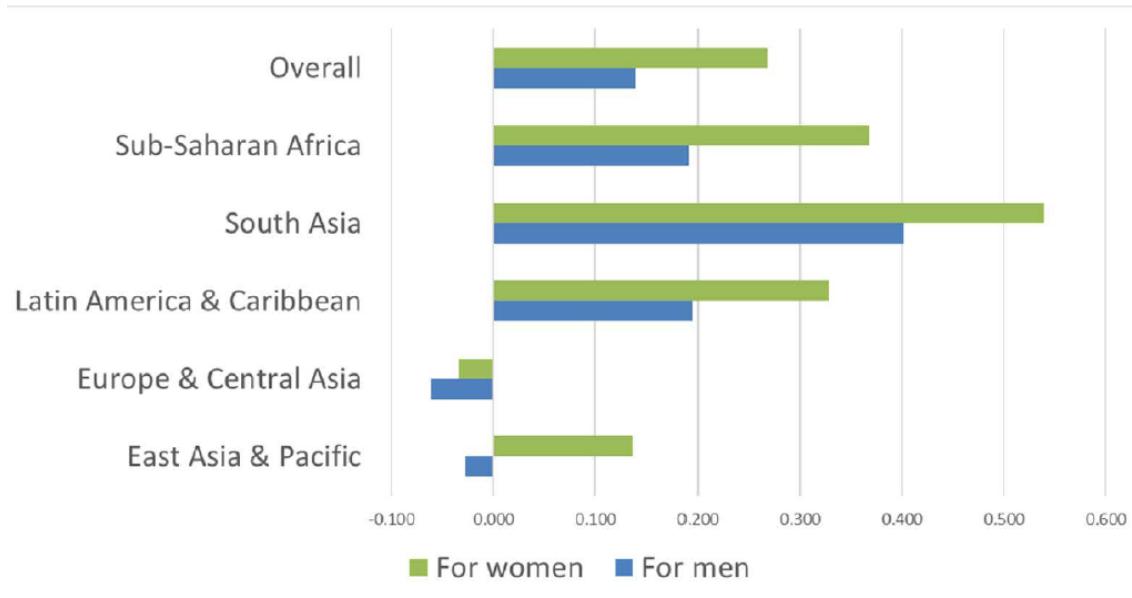
Fonte: Elaborado pelo autor.

Contudo, um grande problema na análise da governança, quando comparada aos demais setores, é a falta de dados ricos e de alta qualidade. Com o intuito de dirimir esse

obstáculo, o Programa *Development Impact Evaluation*<sup>3</sup> (DIME), juntamente com o *Governance Global Practice*, trata com empenho a questão da medição. A iniciativa *Bureaucracy Lab*<sup>4</sup>, liderada por Daniel Rogger e Zahid Hasnain, trabalha na construção de bases de dados coordenadas, em termos de indicadores de burocracia.

A Figura 2 traz um exemplo de emprego desses dados: o prêmio salarial do setor público por gênero e por região do mundo, utilizando informações agrupadas de 79 países. Evidencia-se que o prêmio salarial é maior para mulheres em todas as regiões do mundo, com exceção da Europa e da Ásia Central. A informação também está disponível por país, de forma mais detalhada.

Figura 2 – O mercado de trabalho do setor público: Prêmios salariais por sexo e região



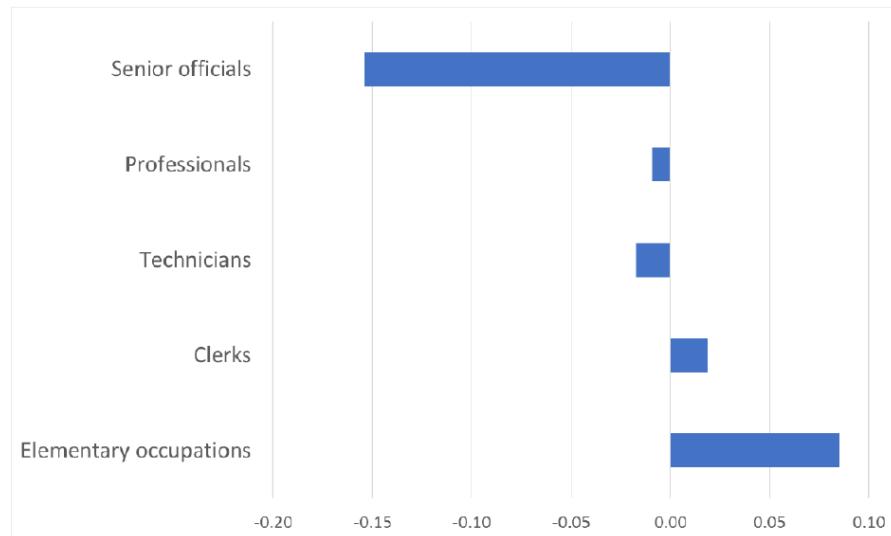
Fonte: Indicadores mundiais de burocracia, vários anos, dados de 79 países.

Já a Figura 3 apresenta o prêmio salarial do setor público desagregado por tipo de emprego, utilizando informações agrupadas dos mesmos 79 países. Verifica-se que, para executivos, profissionais e técnicos do setor público não existe prêmio salarial em relação ao setor privado (Figura 3). Os únicos grupos de ocupação que apresentam prêmio salarial em relação ao setor privado são os de funcionários públicos com empregos mais básicos ou elementares.

Figura 3 – O mercado de trabalho do setor público: Prêmios salariais por ocupação

<sup>3</sup> Criado em 2005 e relançado em 2009 como um esforço descentralizado para integrar a utilização da avaliação de impacto no Banco Mundial. (LEGOVINI, 2010)

<sup>4</sup> É uma iniciativa conjunta do *Governance Global Practice* e *Development Economics Research Group*, que visa transformar o conhecimento global sobre a administração do setor público.



Fonte: Indicadores mundiais de burocracia, vários anos, dados de 79 países.

Por fim, a Figura 4 mostra que existe variabilidade em termos de práticas de gestão e capacidade de gestão nas organizações do setor público, bem como entre os profissionais que trabalham nessas organizações. Parte dessa variabilidade é normal por conta das diferentes atribuições dos profissionais. Entretanto, é possível, por meio desse tipo de análise, criar, por exemplo, treinamentos específicos para cada especificidade de trabalho dos servidores públicos.

Figura 4 – Exemplo de diversidade nas práticas de gestão do setor público



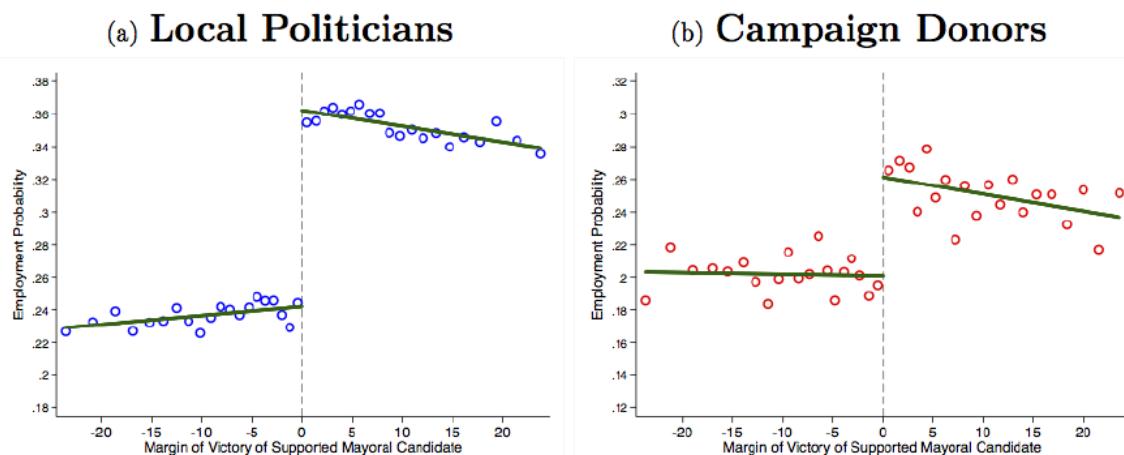
Fonte: Rasul *et al*, 2017, dados de Ghana (2015).

É muito importante que análises desagregadas desses dados sejam feitas no Brasil e nos demais países, pois isso viabiliza uma pesquisa detalhada sobre os servidores públicos. Por exemplo, seria possível verificar com consistência onde se localizam e quais as práticas de gestão do setor público, assim como verificar a variabilidade em termos de práticas de gestão de acordo com a organização e/ou secretaria.

## EXPERIÊNCIAS DE AVALIAÇÃO DE IMPACTO NA ÁREA DA GOVERNANÇA

A avaliação de impacto pode melhorar o nosso entendimento sobre diversos campos, até mesmo a respeito de campos que tradicionalmente são muito difíceis de estudar, por exemplo, a política. O recente estudo de Colonnelli *et al* (2017) mostra como o patrocínio político afeta as políticas. A partir da construção e análise de uma base de dados sobre profissionais que eram ou políticos locais, ou doadores das campanhas de políticos locais, os autores relacionaram cada doador de campanha à sua filiação partidária, demonstrando o que acontece quando o político apoiado ganha a eleição e quando perde a eleição. As análises revelaram que, se os políticos apoiados perdem a eleição, seu apoiador tem uma probabilidade de aproximadamente 20% de ser empregado no setor público local (Figura 5). A situação muda se o político apoiado ganha a eleição: essa probabilidade “salta” para 36% no caso dos apoiadores afiliados locais e para quase 30% no caso dos doadores de campanha (Figura 5).

Figura 5 – Probabilidade de emprego no setor público de políticos locais e doadores de campanhas



Fonte: Colonnelli *et al*, 2017.

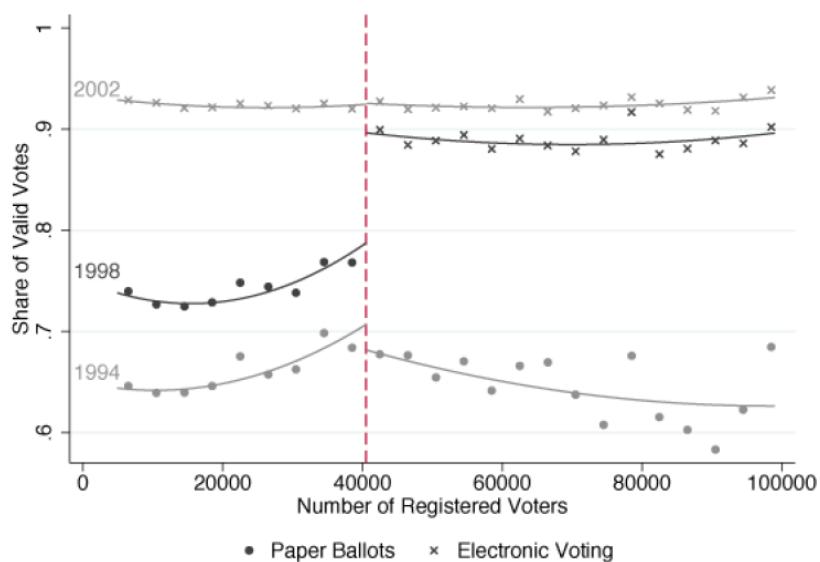
Nota: Nos 4 anos após a eleição de governos locais brasileiros, por posições de margem de vitória do candidato a prefeito apoiado sobre o adversário – eleições de 2000, 2004, 2008 e 2012 para os políticos locais e de 2004, 2008 e 2012 para os doadores de campanha.

O estudo de Colonnelli *et al* (2017) conclui que existe uma relação muito grande entre a probabilidade de conseguir ou não um emprego no setor público e ser ou não apoiador de

campanhas eleitorais. Esse resultado é preocupante, porque o estudo mostra que as pessoas que foram empregadas por causa de seu apoio ou patrocínio político não são tão qualificadas quanto as outras que não foram empregadas porque não estavam afiliadas ao partido que ganhou as eleições. O estudo indica que a contratação, nos contextos onde o patrocínio político é bastante alto, é feita baseando-se mais na afiliação partidária do que na qualidade e meritocracia do potencial servidor público. Preocupa, pois, que os empregos no setor público sejam usados como recompensa para apoiadores políticos e que onde o patrocínio político é mais alto, a educação é menor. Ou seja, nesse cenário, o montante de apoio fornecido ao partido no poder substitui as qualificações como determinantes das decisões de contratação. Especificamente, os autores verificaram que apoiadores do partido no poder são, em média, 17% mais propensos à desqualificação em termos de educação. Como consequência, há uma pior prestação de serviços públicos. Verificou-se que um aumento no patrocínio em um município está associado a pontuações mais baixas nos testes dos alunos nas escolas públicas locais. (COLONNELLI *et al*, 2017)

Outro estudo destacado é de Fujiwara (2015), que analisa as consequências da implantação da tecnologia do voto eletrônico no Brasil sobre a responsividade política e sobre a saúde infantil. Em 1998, os municípios com mais de 40 mil habitantes substituíram o voto em papel pelo voto eletrônico. O autor avalia que, com a chegada do voto eletrônico, a porcentagem de votos válidos subiu dramaticamente, de menos de 80% para quase 90% (Figura 6). Em 2002, quando as urnas eletrônicas já estavam presentes em todo o País, a percentagem de votos válidos ultrapassou 90% (Figura 6). O autor conclui que a votação eletrônica melhorou a participação política e promoveu uma emancipação de cidadãos, principalmente entre os menos instruídos.

Figura 6 – Relação entre votação em papel, votação eletrônica, percentagem de votos válidos e número de eleitores registrados nas eleições de 1994, 1998 e 2002 no Brasil



Fonte: Fujiwara, 2015.

Além disso, o estudo aponta que o incremento na participação política também promove resultados na área da saúde. Na medida em que a tecnologia dos estados se integra ao longo do tempo, os gastos do governo se deslocam para serviços de saúde, fazendo com que os gastos na área da saúde registrem um aumento de 19%. Também se verificou um aumento de 19% no número de mães com baixa escolaridade que fazem mais de sete exames pré-natais e uma redução de 6,8% no número de recém-nascidos, de mães sem o ensino primário, que nascem abaixo do peso. O estudo evidencia que os resultados positivos vão para além de um primeiro estágio, de incremento da participação política, e avançam para um segundo estágio de impactos positivos na área da saúde.

Por fim, vale mencionar o estudo não experimental de Gonçalves (2014), que avalia se o orçamento participativo nos municípios brasileiros, realizado durante o período 1990-2004, afetou o padrão dos gastos municipais das condições de vida da população. O autor conclui que os municípios que utilizam o orçamento participativo favoreceram uma alocação de gastos públicos correspondente às preferências populares, canalizando um percentual maior dos recursos para investimentos em serviços de saneamento e saúde. Isso, por fim, proporciona uma redução nas taxas de mortalidade infantil. A conclusão do estudo indica que promover uma interação mais direta entre a população e as decisões na política orçamentária pode afetar tanto a forma como os recursos locais são gastos quanto o resultado do padrão de vida da população. Não obstante, o estudo não demonstra quais são os mecanismos que explicam esse resultado positivo, em termos de gasto público.

## **CONSIDERAÇÕES FINAIS**

Uma avaliação de impacto experimental na área da governança poderia responder perguntas tais como sobre a melhoria, por exemplo, da confiança no governo, e investigar, além do impacto, as suas causas ou outros possíveis mecanismos. Pode responder também a questões sobre o que mudar nas regras de participação de um programa, como nos processos do orçamento participativo. Nesse caso, se o uso de tecnologia poderia melhorar o processo ainda mais, por exemplo. Também avaliar qual é a interação com a capacidade do usuário, do cidadão que usa o sistema, que usa o processo. A avaliação de impacto pode responder a essas perguntas que outros estudos, com outras metodologias, não necessariamente conseguem responder.

A avaliação de impacto em governança não só é factível, mas é muito promitente. Por esse motivo existem alguns estudos sendo desenvolvidos atualmente. A evidência não precisa ser, necessariamente, relacionada a grandes perguntas. É importante enfocar em pequenas mudanças, porque essas podem ter um impacto muito grande.

Também é importante realçar que a questão da medição dos resultados é tão importante como a questão da experimentação. O DIME, e outras unidades do Banco Mundial (como o setor de Governança), podem apoiar nesse processo.

## REFERÊNCIAS

- COLONNELLI, Emanuele; TESO, Edoardo; PREM, Mounu. *Patronage in the Allocation of Public Sector Jobs.* 2017. Disponível em: <[https://scholar.harvard.edu/files/edoardoteso/files/edoardoteso\\_jmp.pdf](https://scholar.harvard.edu/files/edoardoteso/files/edoardoteso_jmp.pdf)>. Acesso em: 03 de maio de 2018.
- FUJIWARA, Thomas. Voting Technology, Political Responsiveness, and Infant Health: Evidence from Brazil. *Econometrica Journal of The Econometric Society*, Vol. 83, Nº 2, pp. 423-464. 2015.
- GERTLER, Paul J.; MARTÍNEZ, Sebastián; PREMAND, Patrick; RAWLINGS, Laura B.; VERMEERSCH, Christel M. J.. *Avaliação de Impacto na Prática*. Washington, D.C.: Banco Internacional para Reconstrução e Desenvolvimento/Banco Mundial e Banco Interamericano de Desenvolvimento (BID), 2018.
- GONÇALVES, Sónia. The Effects of Participatory Budgeting on Municipal Expenditures and Infant Mortality in Brazil. *World Development*, Volume 53, pp. 94-110. 2014.
- LEGOVINI, Arianna. *Development impact evaluation initiative: a World Bank-wide strategic approach to enhance developmental effectiveness*. Washington, D.C.: Banco Mundial, 2010. Disponível em: <<http://documents.worldbank.org/curated/pt/689141468161083406/Development-impact-evaluation-initiative-a-World-Bank-wide-strategic-approach-to-enhance-developmental-effectiveness>>. Acesso em: 02 de maio de 2018.
- RASUL, Imran; ROGGER, Daniel; WILLIAMS, Martin J.. Management and Bureaucratic Effectiveness: A Scientific Replication. *Under review: Economic Journal*. 2017. Disponível em: <[https://martinjwilliamsdotcom.files.wordpress.com/2017/06/css\\_replication\\_ej.pdf](https://martinjwilliamsdotcom.files.wordpress.com/2017/06/css_replication_ej.pdf)>. Acesso em: 04 de maio de 2018.

## POLÍTICAS DE COMPRAS GOVERNAMENTAIS: UMA PROPOSTA DE AVALIAÇÃO\*

Laura de Castro Zoratto<sup>1</sup>

**Resumo:** O texto apresenta o desenho de uma avaliação de impacto na área de compras governamentais, realizado a partir da parceria entre o Governo do Estado do Amazonas e o Programa *Development Impact Evaluation Initiative* (DIME) do Banco Mundial. A área de compras públicas representa parcela importante do orçamento público, sendo fundamental sua avaliação, tendo em vista que um sistema de compras deficiente tem impacto sobre todos os setores de provisão de bens e serviços públicos. Além do desenho da avaliação proposto, serão apresentados resultados de análises iniciais a partir dos dados já disponíveis.

**Palavras-chave:** Avaliação de impacto. Compras governamentais. Gasto público.

**Abstract:** The text presents the research design of an impact evaluation in the area of government procurement, based on a partnership between the State Government of Amazonas and the Development Impact Evaluation Initiative (DIME) of the World Bank. The public procurement area represents an important part of the public budget. So its evaluation is fundamental, once a deficient procurement system has an impact on all sectors of public goods and services provision. In addition to the research design for the proposed evaluation, initial results of the analysis will be presented from the data already available.

**Keywords:** Impact evaluation. Government procurement. Public expenditure.

---

\* Texto elaborado a partir da transcrição da fala e da apresentação em slides de palestras proferida no Workshop Avaliação de Impacto de Políticas Públicas, realizado nos dias 29 e 30 de novembro de 2017, no Auditório da Secretaria de Planejamento, Governança e Gestão (SPGG-RS), com o apoio do Banco Mundial.

<sup>1</sup>Economista Sênior do *Governance Global Practice for Latin America and the Caribbean* (GGP).

## INTRODUÇÃO

Este texto objetiva apresentar o desenho de uma avaliação de impacto na área de compras governamentais, realizado a partir da parceria entre o Governo do Estado do Amazonas e o Programa *Development Impact Evaluation Initiative* (DIME) do Banco Mundial<sup>2</sup>.

A importância do estudo se dá na medida em que a área de compras públicas é a atividade do governo mais sujeita a desperdícios, ao mesmo tempo em que representa, em média, 15% do PIB global (OECD, 2007). A intervenção apresentada ainda não foi implementada. Portanto, serão mostrados apenas resultados de análises de dados iniciais e o desenho da avaliação de impacto proposto.

## PARCERIA PARA SOLUÇÕES NA ÁREA DE COMPRAS PÚBLICAS

A parceria entre o governo estadual do Amazonas e o DIME-Banco Mundial começou através de um Projeto de Empréstimo para Políticas de Desenvolvimento (*Development Policy Loan Project – DPL*) em 2014. O Projeto contemplava várias reformas na área de compras públicas, desde a introdução de mais registros de preço e a padronização de contratos de serviço, até a criação de um banco de preços, o qual permitiu basear as compras governamentais nos preços unitários contidos nas Notas Fiscais Eletrônicas (NF-e) de bens comercializados no estado.

Essa parceria tem sido muito rica. Pesquisadores têm interesse em trabalhar com dados bons e boas perguntas. Por esse motivo, parcerias com governos são interessantes para que as perguntas de pesquisa não se tornem algo somente acadêmico, mas também relacionado com algum problema específico que se queira resolver. No Amazonas, em específico, a equipe e o sistema de compras são muito bons, sendo possível realizar muitas análises a partir dos dados disponíveis, que são muitos ricos.

Mas por que a importância do assunto compras públicas? Não é uma área finalística e geralmente é de pouco acesso aos cidadãos, que têm um contato mais direto com a escola ou os serviços de saúde, por exemplo. Mas a área de compras públicas está relacionada com a eficiência do serviço público em todos os setores. Envolve não só a compra do remédio ou do livro, mas também do equipamento, do computador, da viatura policial, etc. Em virtude disso, representa uma parcela muito importante do orçamento público, em qualquer governo. Consequentemente, um sistema de compras deficiente tem impacto sobre todos os setores de provisão de serviços do governo.

Estima-se que 15% do PIB global referem-se a compras de governo (OECD, 2007). Alguns países fazem até política industrial por meio de compras governamentais. Contudo, não existem muitas avaliações de impacto sobre esse setor. Adicionalmente, essa é uma atividade

---

<sup>2</sup>Equipe responsável pelo estudo: Michael Best (*Columbia University*), Alexandre Borges (Banco Mundial), Evan Kresch (*Oberlin College*), Joana Naritomi (*London School of Economics*), Laura Zoratto (Banco Mundial).

muito sujeita a desperdícios e corrupção. Logo, os ganhos possíveis com reformas nessa área são significativos, e é de interesse captar esses ganhos através de uma avaliação de impacto.

Apesar de o Governo do Amazonas já ter implantado muitas reformas, ainda se paga consideravelmente mais por um produto do que qualquer outro comprador do setor privado. Isso porque existem custos associados à venda para o governo: não só a questão do pagamento, que nem sempre é em dia, pois há outras prioridades, como o pagamento dos salários, mas também por uma questão de excesso de documentos exigidos pela legislação. Esses custos de transação reduzem a competitividade do processo de compra de uma maneira geral, fazendo com que menos fornecedores queiram participar, o que afeta negativamente o preço e a qualidade dos bens e serviços comprados. Esse retrato é brevemente explorado na seção seguinte.

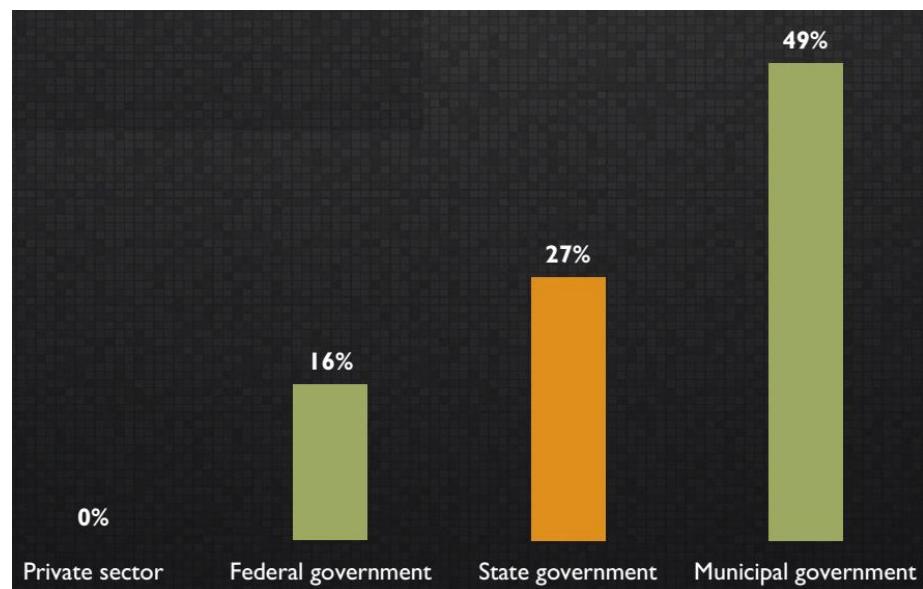
## DIAGNÓSTICO DO SETOR DE COMPRAS PÚBLICAS

Os dados administrativos do setor são muito bons. No caso das compras públicas, foram analisados 1.910.286 dados – 2010 e 2017, por produto, método de compra, agência destinatária. A parceria não obteve acesso a todo o banco de dados da nota fiscal eletrônica, pois isso envolve a questão de sigilo e confidencialidade, mas se pôde acessar o suficiente para poder identificar, por exemplo, os fornecedores potenciais. Essa base de dados contava com mais de 2 milhões de itens, em 139 mil notas emitidas. Foi realizada uma análise de rede para compor *clusters* de fornecedores, pelo fato de muitos fornecedores venderem os mesmos produtos. Trata-se de análise complexa, mas possível de ser fazer graças à riqueza de dados existentes.

Foram identificadas 4 mil firmas que vendem os produtos que o governo compra, mas não vendiam para o governo. Agora, com todos os dados de aproximadamente 18 mil firmas cadastradas no sistema de compras do governo, é possível conhecer toda a dinâmica de quem vende, para de vender, volta a vender, se é um produtor final, intermediário, etc. Vale dizer que esses dados já estavam sendo utilizados para criar indicadores de monitoramento dos resultados das licitações. Logo, já existia toda uma mentalidade de monitoramento, criado pela equipe de compras do Governo do Estado.

A partir desses dados, já se sabia, por exemplo, que o Governo do Amazonas pagava 27% a mais do que o setor privado pelo mesmo produto. Comparando o setor privado, Governo Federal, Governo do Amazonas e Manaus, porém, verificava-se que a capital do estado pagava muito mais do que os demais entes (Figura 1).

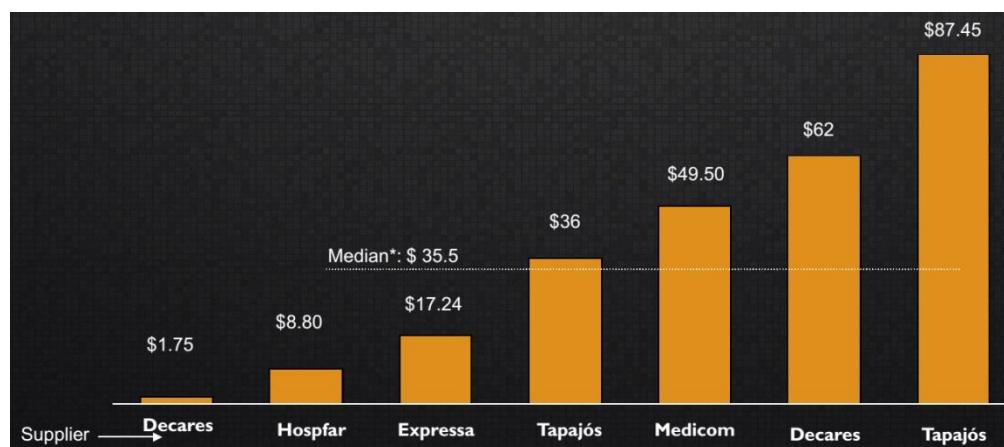
Figura 1 – Prêmio pago pelo governo com base no preço do setor privado – 2013



Fonte: TRIDAPALLI, 2013.

Comparando métodos diferentes para fazer a compra de um mesmo produto, no caso Claritromicina em pó 500 gramas, verificou-se, também, a variabilidade no preço unitário desse produto, vendido por diferentes fornecedores. Foi possível encontrar um mesmo fornecedor, Decares, vendendo o mesmo produto de acordo com um método por R\$ 1,75 para uma agência, e por R\$ 62,00 de acordo com outro método para outra agência (Figura 2). Verifica-se, então, uma variação muito grande nos preços, seja entre governos, seja dentro do próprio governo.

Figura 2 – Preços unitários pagos por Claritromicina em pó, 500 mg, em R\$ e Mediana para o conjunto de dados completo – 2015



Fonte: Análise do Banco Mundial a partir dos dados do e-Compras do Governo do Estado do Amazonas.

Com a finalidade de identificar os desafios enfrentados pelos produtores ou fornecedores ao vender para o governo, realizou-se também uma pesquisa com 200 fornecedores. Identificou-se que a demora em receber o pagamento é o principal desafio enfrentado por eles. O segundo ponto levantado foi o requerimento de documentos: 7,3% das firmas relataram ser esse o principal obstáculo (Figura 3).

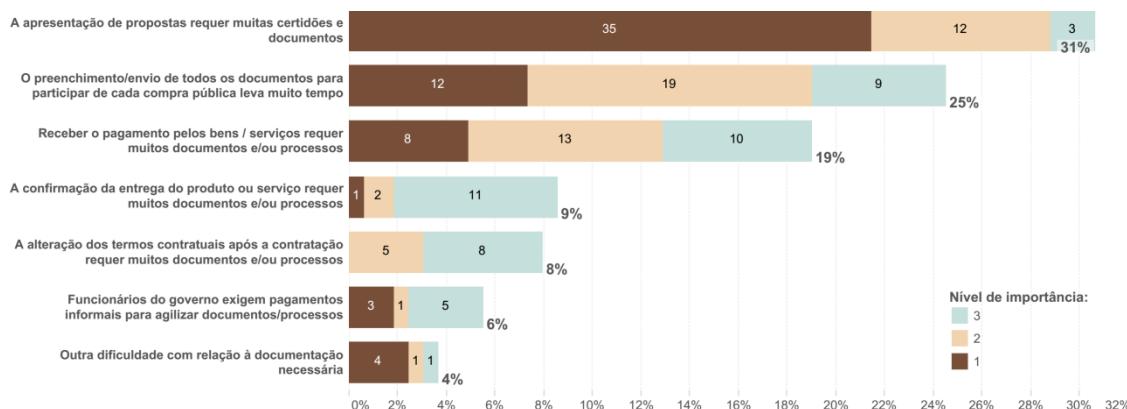
**Figura 3 – Principais barreiras enfrentadas pela amostra de empresas ao vender para o Governo do Estado do Amazonas – 2015**



Fonte: World Bank (2015).

Quanto a esse ponto específico, o Governo do Amazonas buscou compreender melhor essas dificuldades, para identificar os principais obstáculos no processo de venda para o governo (Figura 4). Verificou-se que, ao se registrar, o fornecedor precisa apresentar vários documentos, apresentá-los novamente quando deseja fazer uma proposta, outra vez ao entregar o produto ou serviço e uma vez mais para receber o pagamento correspondente. Se for preciso modificar o contrato de alguma forma, tanto por parte do governo quanto do fornecedor, a demora será muito maior, e serão necessários outros procedimentos. Todo o processo, então, depende da emissão e apresentação de várias certidões de cartório, que cobram um alto custo, sendo muitas vezes necessário que se contrate uma pessoa para lidar especificamente com isso dentro da empresa. Por esse motivo, muitas firmas não desejam nem mesmo iniciar seu registro.

**Figura 4 - Principais barreiras enfrentadas pela amostra de empresas para preencher a documentação exigida pelo Governo do Estado do Amazonas – 2015**



Fonte: World Bank (2015).

Diante disso, o Governo do Amazonas planejou várias reformas visando digitalizar praticamente todo o processo de compras, desde o registro do fornecedor até o pagamento. Isso vai além de simplesmente fazer a licitação e o pregão eletrônico, que já existem há bastante tempo. Trata-se não só de fazer todo o processo eletronicamente, mas também de criar um domicílio de fornecedores, batizado de Domicílio Eletrônico de Licitantes e Fornecedores (DLF-e). Por meio do DLF-e, a documentação passaria a ser apresentada uma única vez. O Governo armazenaria esses documentos e somente solicitaria novamente para o fornecedor caso uma certidão houvesse vencido, por exemplo.

Dessa forma, espera-se reduzir o custo de participação no pregão e, assim, atrair novos fornecedores. Material de divulgação já foi elaborado sobre esse novo processo e será enviado para empresas potenciais fornecedoras por e-mail, correio e ligações telefônicas. A base de dados da nota fiscal eletrônica tem sido útil para identificar também fornecedores de determinados produtos, mas que até então não vendem para o governo – inclusive a preços muito mais baratos.

A proposta de criação desse domicílio do fornecedor é que se pretende avaliar. A pergunta que a avaliação de impacto buscará responder é *como as barreiras à entrada afetam a competitividade do processo de compras do governo*. Os resultados a serem avaliados incluem o número de empresas que participam dos processos licitatórios, o preço final obtido, a qualidade e o tempo total de compra.

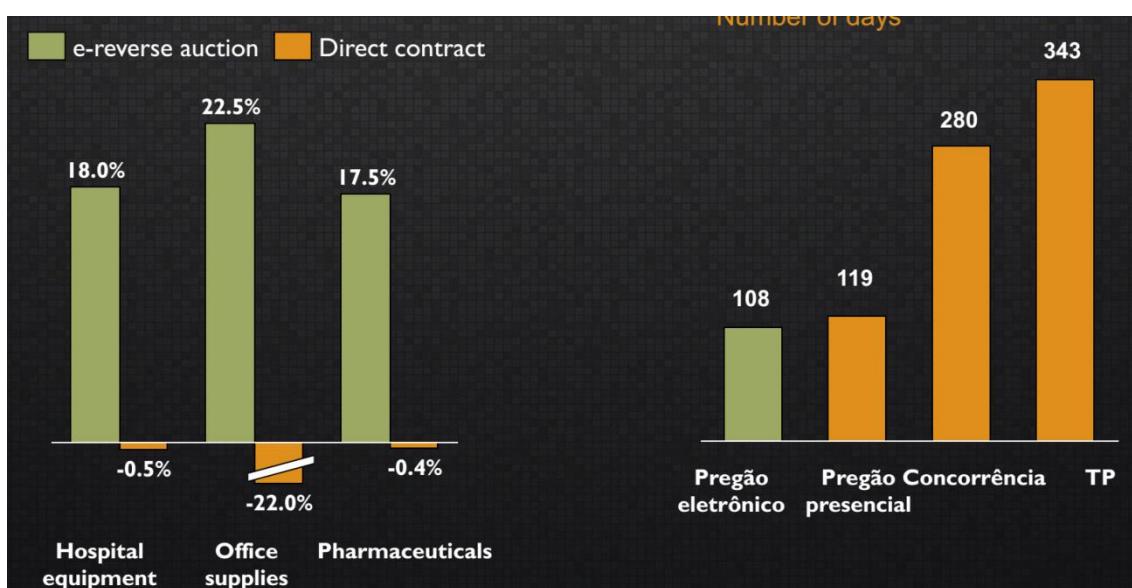
O DLF-e atualmente aguarda implementação. Embora a avaliação ainda não tenha sido realizada, uma análise de dados inicial já permite explorar várias evidências na área de compras públicas, apresentadas a seguir.

## ANÁLISE DE DADOS

Além do DLF-e, outras reformas implementadas pelo governo estadual do Amazonas incluem a compra em escala, a padronização de contrato de serviços, o uso da NF-e como preço de referência, etc. Nem todas essas reformas precisam de uma avaliação de impacto. A avaliação de impacto leva tempo e precisa de parceiros. O que o Governo do Estado do Amazonas tem feito é uma análise de custo-benefício e o monitoramento de indicadores de todas essas reformas, fazendo uso dos dados administrativos que possui.

Com base nos registros administrativos disponibilizados, a análise dos dados indica que as compras feitas por meio do pregão eletrônico são mais ágeis e mais econômicas para o governo, quando comparadas a outros métodos (Figura 5).

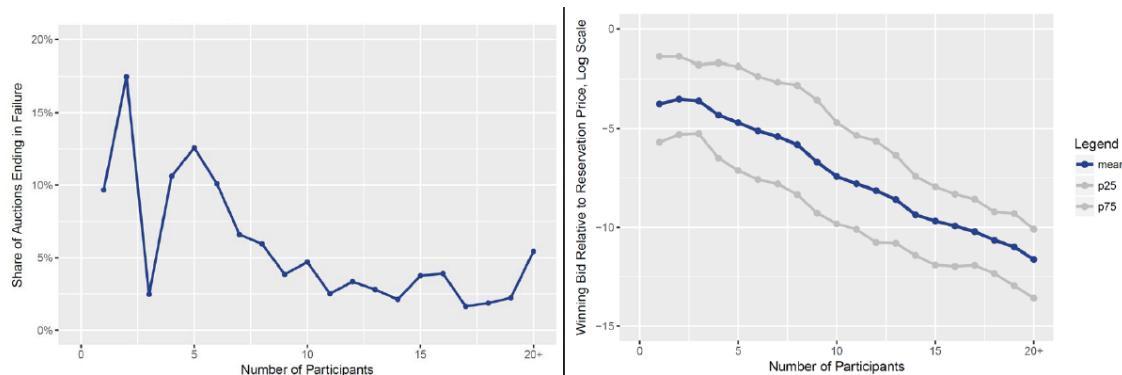
Figura 5 – Economia por método de licitação (à esquerda) e tempo médio de duração por processo em número de dias (à direita) – 2015



Fonte: Análise do Banco Mundial a partir dos dados do e-Compras do Governo do Estado do Amazonas.

A análise constatou, ainda, correlações entre o número de participantes no processo, o número de falhas e os preços dos leilões. Quanto mais participantes existem em uma licitação, menos falhas e fracassos ocorrem (Figura 6, à esquerda). Além disso, quanto mais participantes, menor o preço obtido, gerando mais economia (Figura 6, à direita). Logo, uma maior participação está correlacionada com uma maior eficiência.

Figura 6 – Distribuição das falhas em leilões por número de participantes (à esquerda) e relação entre os preços dos lances vencedores e o número de participantes (à direita) – 2016



Fonte: Elaboração própria, a partir de dados administrativos.

Diante disso, a avaliação parte da hipótese de que a implantação do Domicílio de Fornecedores atrairá mais fornecedores/participantes, podendo ter um impacto positivo sobre preços e reduzir o número de fracassos – hipótese a ser verificada atrás da avaliação de impacto.

Em resumo, a avaliação de impacto investigará como a desburocratização da participação nas compras públicas, via adoção de um domicílio de fornecedores que agilize os procedimentos (do ponto de vista do fornecedor), pode aumentar a competitividade do processo licitatório. O fato de os processos licitatórios se tornarem mais eficientes e gerarem um aumento na participação de concorrentes não é uma conclusão tão óbvia quanto parece. O processo todo é complexo. Com o aumento no número de participantes dos processos licitatórios, aquelas empresas que se especializaram em vender para governos podem ficar de fora do processo, caso a competição seja muito grande, por serem menos eficientes.

## CONSIDERAÇÕES FINAIS

Neste texto, apresentamos o desenho de uma avaliação na área de compras governamentais. A avaliação objetiva, pela primeira vez, analisar quais são os efeitos da desburocratização sobre o processo licitatório e a economia em geral. Com tanto, será possível analisar como essa intervenção afeta o número de participantes, o preço, o surgimento de novos fornecedores, a mudança de perfil do fornecedor, a qualidade e a eficiência do pregão.

Dessa forma, busca-se informar outros estados, outros países e outros governos. A cidade de Manaus, por exemplo, passou a adotar esse tipo de procedimento a partir da experiência do governo do estado do Amazonas. Apesar de a intervenção não ter sido implementada ainda, os dados administrativos já foram coletados e utilizados para algumas análises simples, tais como a identificação de fornecedores potenciais que o governo desejava obter, mas ainda não tinha identificado.

## REFERÊNCIAS

- LEGOVINI, Arianna. *Development impact evaluation initiative*: a World Bank-wide strategic approach to enhance developmental effectiveness. Washington, D.C.: Banco Mundial, 2010. Disponível em: <<http://documents.worldbank.org/curated/pt/689141468161083406/Development-impact-evaluation-initiative-a-World-Bank-wide-strategic-approach-to-enhance-developmental-effectiveness>>. Acesso em: 02 de maio de 2018.
- OECD. Organisation for Economic Co-operation and Development. *Integrity in Public Procurement, Good Practice from A to Z*. Paris: OECD Publishing, 2007.
- PIMENTA, C.; PESSOA, M. (Ed.) *Public financial management in Latin America: the key to efficiency and transparency*. Washington, D.C: BID, 2015.
- TRIDAPALLI, J. P. Gestão da Cadeia de Suprimentos no Setor Público. Apresentação. V Seminário de Avaliação de Políticas Públicas e Qualidade do Gasto. Fundação de Economia e Estatística, Porto Alegre, 13 a 14 de novembro de 2013. Disponível em: <<https://www.slideshare.net/feers/apresentao2-13112013-11h-juarez-paulo-tridapalli>>. Acesso em: 02 de maio de 2018.
- WORLD BANK. *Pesquisa Fornecedores*. [online]. 2015. Disponível em: <[http://dataviz.worldbank.org/views/Pesquisa\\_Fornecedores/Inicio?%3Aembed=y&%3AshowShareOptions=true&%3Adisplay\\_count=no#3](http://dataviz.worldbank.org/views/Pesquisa_Fornecedores/Inicio?%3Aembed=y&%3AshowShareOptions=true&%3Adisplay_count=no#3)>. Acesso em: 02 de maio de 2018.

## EXPERIÊNCIA DO CENTRO DE LIDERANÇA PÚBLICA: UMA AVALIAÇÃO\*

Fabio Hideki Ono<sup>1</sup>

**Resumo:** O texto apresenta brevemente a experiência do Centro de Liderança Pública (CLP). São apresentados projetos em que o CLP atua, destacando os resultados de uma avaliação de processos aplicada ao Master em Liderança e Gestão Pública (MLG), promovido pelo Centro. Serão compartilhadas, ainda, algumas experiências positivas de ex-alunos do MLG no tocante a aspectos de competências, rede, carreira, políticas públicas e impactos socioeconômicos.

**Palavras-chave:** Centro de Liderança Pública. Avaliação de processos. Master em Liderança e Gestão Pública.

**Abstract:** The text briefly introduces the experience of the Center for Public Leadership (CLP). It will be presented projects in which the CLP acts, highlighting the results of a process evaluation applied to the Master in Leadership and Public Management (MLG), promoted by the Center. It will also be shared some positive experiences of MLG ex-students regarding aspects like skills, network, career, public policies and socioeconomic impacts.

**Keywords:** Center for Public Leadership. Process evaluation. Master in Leadership and Public Management.

---

\* Texto elaborado a partir da transcrição da fala e da apresentação em slides de palestras proferida no Workshop Avaliação de Impacto de Políticas Públicas, realizado nos dias 29 e 30 de novembro de 2017, no Auditório da Secretaria de Planejamento, Governança e Gestão (SPGG-RS), com o apoio do Banco Mundial.

<sup>1</sup>Diretor de Desenvolvimento de Lideranças do Centro de Liderança Pública (CLP).

## INTRODUÇÃO

O tema de avaliação é importante para a gestão pública, especialmente no Brasil. Neste texto, será relatado um pouco da experiência do Centro de Liderança Pública (CLP), sendo interessante compartilhar um pouco da sua visão. Será trazida mais uma avaliação, não tanto de impacto, mas em nível de processo, de um programa do CLP.

## O CENTRO DE LIDERANÇA PÚBLICA (CLP) E A GOVERNANÇA PÚBLICA

O Centro de Liderança Pública é uma organização do terceiro setor de São Paulo, criada pelo cientista político Luiz Felipe D'Ávila. Em 2006, D'Ávila retornou de um curso de Mestrado na *Harvard Kennedy School*, em Harvard, e decidiu fazer algo com o conhecimento adquirido. Nesse sentido, decidiu compartilhar conhecimentos e experiências na formação e desenvolvimento de líderes públicos, fundando o CLP. A ideia do CLP é de que é preciso mudar atitudes e crenças dos líderes públicos para mudar o Brasil.

Para promover o desenvolvimento de lideranças, o CLP oferece cursos, palestras e *workshops*, com o intuito de capacitar os líderes públicos para enfrentar os problemas da sociedade e dos próprios governos. O primeiro curso da CLP, realizado em 2008, foi um curso para prefeitos, seguindo o modelo aprendido na *Kennedy School*. Foi realizada uma semana de curso para 18 prefeitos, um deles o então prefeito de Caxias do Sul, hoje governador do estado do Rio Grande do Sul. Não obstante, desde o início, o CLP sempre presou pela diversidade, inclusive de partidos.

Atualmente, o Centro atua em duas principais frentes: uma de mobilização e uma de desenvolvimento de lideranças, com o intuito de buscar a melhoria do Estado brasileiro. A missão do CLP é desenvolver e valorizar o Estado de direito, o Estado Democrático e a prestação de serviços públicos. Acredita-se que, para isso, é importante atuar numa frente de mobilização e também numa frente de desenvolvimento de lideranças. Nessa questão de mobilização, o CLP vem atuando, já há alguns anos, com o tema da competitividade. Pelo sexto ano consecutivo, em 2017, o Centro divulgou o *Ranking de Competitividade dos Estados*.

Além disso, a CLP tem tratado de temas que entende ser importantes para o Estado e para o futuro da Nação, tais como a reforma da previdência e a reforma política, entre outros. Recentemente, tem tratado também sobre o tema da avaliação de impacto em nível federal.

Governança é entendida como um conjunto de práticas de responsabilidade das lideranças, organizações necessárias para avaliar, direcionar, monitorar a atuação da gestão, com vistas ao desempenho das políticas públicas e à prestação de serviços. Sendo assim, a responsabilidade do CLP tem tudo a ver com a questão da governança. Significa formar pessoas para ter o mote de melhor avaliação, tomar melhores decisões e com vistas à eficiência das políticas públicas pelas quais elas são responsáveis.

## ALGUNS EXEMPLOS DA ATUAÇÃO DO CLP

Em 2017, foram mais de 150 prefeitos capacitados. No Rio Grande do Sul, mais de 30 prefeitos fizeram o curso, como, por exemplo, o prefeito de Bento Gonçalves. Há excelentes relatos dessas lideranças a partir do programa de formação.

O CLP também possui um programa em parceria com o SEBRAE do Rio de Janeiro, para a formação de servidores públicos dos municípios fluminenses. Do interior, são oito municípios, com vistas à formação e ao desenvolvimento de projetos. Dessa forma, busca-se atuar na capacitação e no auxílio diretamente aos servidores na execução de projetos junto com o SEBRAE do Rio de Janeiro.

Em Goiás, o CLP atua com o governo do estado na formação de um escritório de resultados, formando técnicos a partir de uma visão de estruturação de uma governança do estado, melhorando todos os processos decisórios – desde reuniões que o próprio governador tinha com secretários, escalando até os níveis da execução da operação das secretarias.

Outro programa interessante de se entrar em mais detalhes, que será trazido como exemplo de uma avaliação, não de impacto, mas uma avaliação de processos, é o programa de pós-graduação, chamado *Master* em Liderança e Gestão Pública (MLG). Todos os programas de formação estão amparados em algumas competências-chave. Foram trazidos para os programas do CLP os conceitos de liderança adaptativa e de uma visão sobre a resolução de problemas complexos, muito a partir de uma visão da *Kennedy School* de Harvard.

Considerando a complexidade, o CLP entende que é essencial para a tomada de decisão a consideração de aspectos políticos. Pode-se ter o melhor embasamento técnico, a melhor justificativa, mas se não forem consideradas questões políticas, fatalmente o gestor não será bem sucedido. Assim foi trazido para dentro de todos os nossos programas o aspecto político.

Esse lado político, chamado de comportamental, é bastante importante para a execução das ações. Já o lado mais técnico, um olhar sobre a resolução de problemas complexos, traz uma visão sobre modelos e ferramentas de gestão pública. Dentre essas ferramentas de gestão pública estão um olhar sobre técnicas de diagnóstico e a estruturação de um modelo racional para a intervenção.

## TEORIA DA MUDANÇA: EXPERIÊNCIAS DO MLG-CLP

A teoria da mudança foi aplicada primeiramente sobre os programas de formação. O desafio do CLP é como levar esses aspectos para uma capacitação mais ampla de servidores, olhando para o impacto que isso pode trazer, associando não só a formação, mas também o acesso à informação, o diagnóstico e a tomada de decisão e a implantação dos planos de ação.

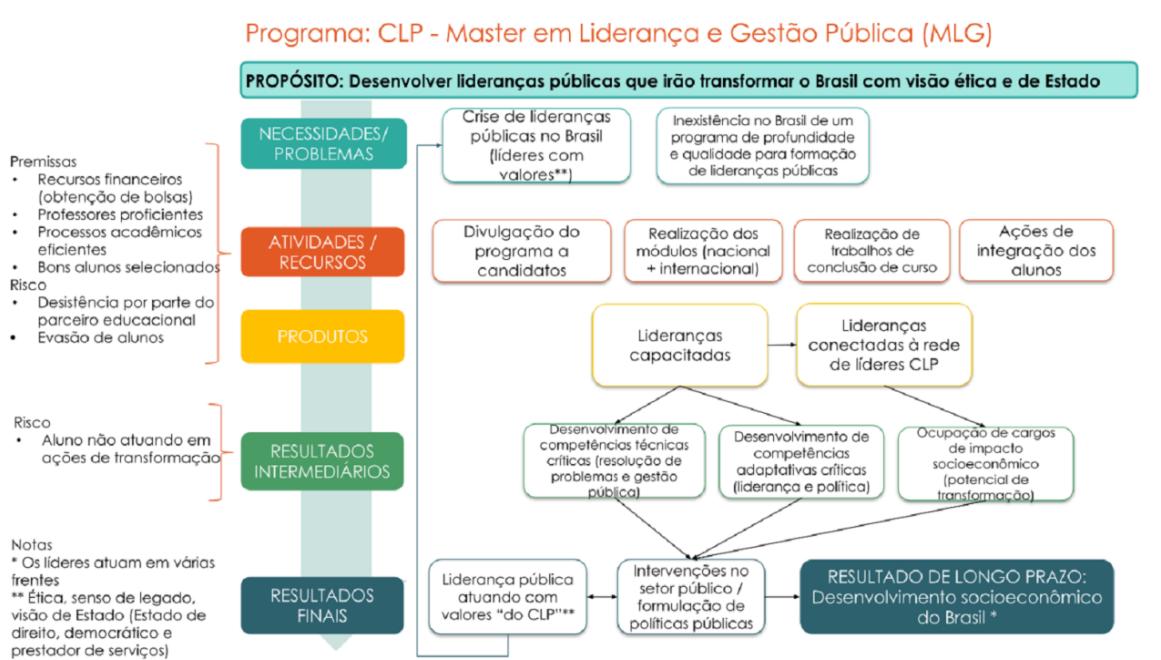
A teoria da mudança foi aplicada inicialmente ao *Master* em Liderança e Gestão Pública (MLG), a “cereja do bolo” do CLP, por ser o programa com mais profundidade. O MLG é

um programa de pós-graduação que atualmente está na quarta turma. São 35 pessoas por ano, em um programa de 17 meses de duração, com módulo internacional, cujos três últimos anos (na turma atual) foram na própria *Kennedy School*. Para a próxima turma, o módulo internacional será na Escola de Governo de Oxford. Contando com a participação de servidores de vários níveis de experiência, passaram pelo CLP mais de 21 estados.

O MLG busca levar uma oportunidade de formação para líderes públicos de alto potencial, com o propósito de desenvolver as lideranças públicas que irão transformar o Brasil, com a visão ética e de Estado. A partir dessa visão de propósito, o que se espera no final é que essas lideranças promovam intervenções no setor público e, finalmente, um desenvolvimento econômico de impacto socioeconômico.

Basicamente, a teoria da mudança do MLG envolve a capacitação de líderes e a conexão entre eles, formando uma rede. O potencial de atuação em rede é enorme, o que torna fundamental conectá-los ainda mais a essa rede. A Figura 1 sintetiza a teoria da mudança do programa.

Figura 1 – Teoria da mudança do MLG



Fonte: Elaboração própria.

Outros exemplos de uso da teoria da mudança foram aplicados em programas de formação do CLP junto a municípios brasileiros, tais como o Orçamento Participativo. Nesse sentido, citam-se trabalhos realizados em conjunto com o MGov e o COLAB. O CLP também está aplicando a teoria da mudança para ações no campo da institucionalização de políticas públicas e construção de capacidades de governo, bem como em programas de ensino a

distância visando qualificar a gestão das finanças municipais, que é um tema muito valorizado em relatos recentes de prefeitos e prefeitas.

## AVALIAÇÃO DO MLG: COMPETÊNCIAS, REDE, CARREIRA, POLÍTICAS PÚBLICAS E IMPACTOS

Foi realizada uma avaliação junto aos alunos e ex-alunos. Questionou-se primeiramente sobre a avaliação da competência de liderança adaptativa, perguntando como o MLG auxiliou nos processos de liderança. A liderança adaptativa entende que as mudanças de comportamento e cultura são fundamentais para a consolidação de transformações e são capazes de deixar um legado, sobretudo quando atrelado aos valores defendidos pelo líder. O MLG foi muito bem avaliado por alunos e ex-alunos. Ressaltam-se alguns relatos de alunos. Por exemplo, para uma aluna de Goiás (Turma 2014), muitas vezes elementos comportamentais são mais relevantes do que aspectos técnicos. Um aluno da Turma 2014, da Prefeitura do Rio de Janeiro, utiliza o referencial de *nudges* em suas ações, por exemplo, em iniciativas de mudança de comportamento de lideranças, através do programa Líderes Cariocas, por ele liderado.

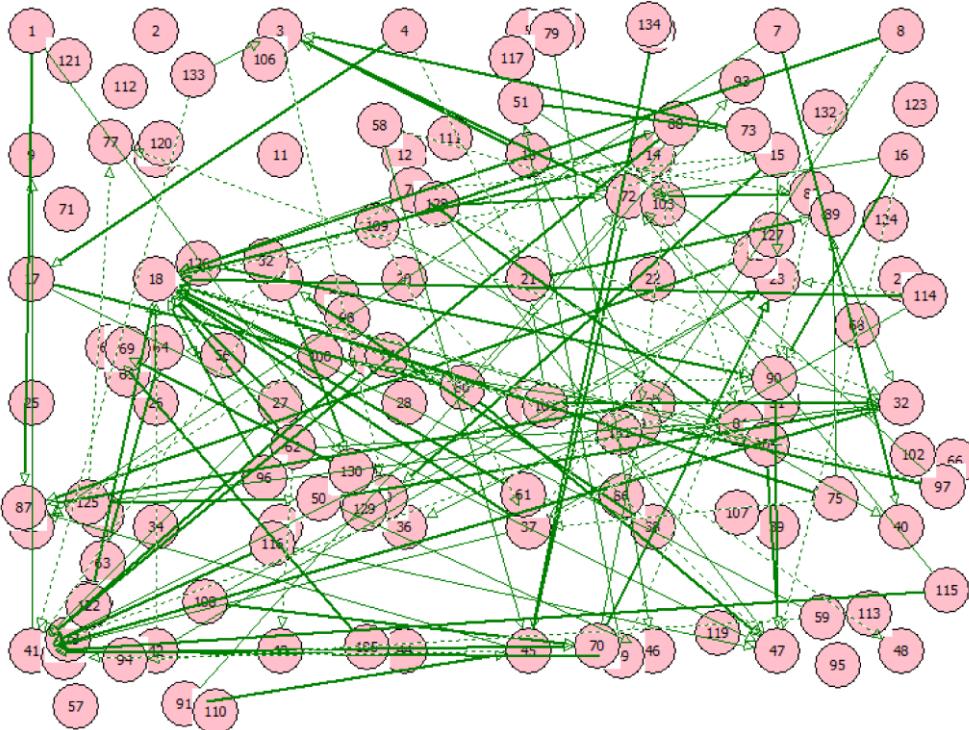
Sobre a competência política, o programa MLG foi relativamente bem avaliado. Por política, entende-se a habilidade de persuadir uns aos outros em objetivos comuns, envolvendo negociação na arte daquilo que é possível e zelando pela ética. Nesse ponto, um caso interessante é o de uma aluna (Turma 2015), que, quando iniciou o MLG, era Secretária de Educação de Ourinhos. Seu caso ganhou certo destaque no início de 2017. O Prefeito de Londrina abriu um concurso público para selecionar uma Secretaria de Educação junto com a Votor Brasil, e ela foi selecionada. Possui uma visão técnica, mas teve que fazer toda a negociação junto com os vereadores para aprovar suas próprias ideias, tendo afirmado que o MLG contribuiu para facilitar isso. Outro ex-aluno, Turma 2016, é o fundador de uma plataforma chamada Politize, que possui milhões de usuários e divulga a causa política.

Em relação à competência de gestão pública, o MLG buscou verificar sua contribuição para a aquisição conhecimento sobre a máquina pública, seus processos, marcos legais e dominar modelos e ferramentas são condicionantes da capacidade de transformação das lideranças. Aqui, relata-se o caso de um aluno (Turma 2014), então Secretário de Teresina, que, após o MLG, passou a aplicar algumas técnicas de gestão na melhoria de suas ações. Ainda, um aluno da Turma 2015, secretário municipal de Mesquita, que era o técnico da Prefeitura do Rio e trabalhou no programa de desburocratização na prefeitura, fazendo também um excelente trabalho que perdura em alguma medida até hoje.

Outro aspecto avaliado pelo CLP foi a promoção da rede. O CLP realizou duas consultas a seus ex-alunos para que eles apontassem três líderes com os quais tiveram mais contato e três líderes que mais lhes inspiravam por seus feitos. A partir disso, foi realizado um sociograma para entender bem essas conexões, com uma riqueza de detalhes (Figura 2). Foram identificados três principais influenciadores dessa rede: o Sócio Fundador da Muove Brasil (consultoria que promove o uso de dados para auxiliar os pequenos municípios, levando

informações e técnicas e a possibilidade de ferramentas de gestão); o ex-prefeito de Olímpia-SP e atualmente Coordenador do Programa Cidade Legal, da Secretaria de Estado de Habitação de São Paulo, que promoveu uma transformação no município de Olímpia, reconhecido como um dos principais polos turísticos de polo aquático; e o Secretário de Segurança do Estado Rio de Janeiro. Por meio do CLP, foram promovidas várias discussões sobre a segurança pública.

Figura 2 – Sociograma dos ex-alunos do CLP



Fonte: Elaboração própria.

Como o MLG conta com apoiadores, é importante prestar contas das ações e dos impactos promovidos pelos alunos. Os recursos das bolsas provêm da iniciativa privada e são destinados às bolsas para os servidores públicos participantes. O impacto intermediário foi captado pelo aspecto da evolução na carreira das pessoas desde que entraram no MLG, até onde estão atualmente. Constatou-se que 66% das pessoas tiveram algum tipo de incremento ao longo da carreira. Alguns exemplos podem ser citados, como o de uma aluna que era coordenadora do curso de Pedagogia e atualmente é Coordenadora Regional de Educação em Rondônia. Também o Secretário de Educação de Santa Catarina, hoje presidente do Conselho Nacional de Educação e membro do Conselho Estadual de Educação de SC. Ainda, existe o caso de uma aluna, que era inspetora do Conselho Regional de Engenharia e Agronomia (CREA) e atualmente é Diretora do Departamento Municipal de Obras e Serviços da Prefeitura de Salto Grande-SP.

Além disso, mensura-se o impacto intermediário do MLG também pelas intervenções promovidas pelos líderes. Nesse sentido, destaca-se o Projeto de alvarás, iniciativa de

desburocratização na prefeitura do Rio de Janeiro, conduzido por uma ex-aluna. Também o projeto de diminuição do desmatamento ilegal e de enfrentamento ao tráfico de animais no Pará, de uma aluna, policial federal. Dois alunos, por sua vez, lideraram o Programa Cidade Legal, de regularização fundiária, do Estado de São Paulo. Já uma outra aluna esteve à frente do Centro de Informações e Decisões Estratégicas em Saúde, o Conecta SUS, um programa exemplo do Estado de Goiás.

Finalmente, destacam-se alguns exemplos de programas liderados por ex-alunos do MLG e que tiveram impactos socioeconômicos relevantes. Como a implantação de processos digitais na Prefeitura de Santos, pelo então Secretário de Planejamento. A ação reduziu drasticamente o número de processos da prefeitura, bem como o tempo de abertura de empresas. Já um aluno de Teresina estruturou, durante o período do MLG, um programa de enfrentamento ao desemprego, que gerou mais de 15 mil postos de trabalho na cidade.

## CONSIDERAÇÕES FINAIS

Neste texto, foi discutido o tema de governança, a partir de uma perspectiva aplicada à realidade do Centro de Liderança Pública (CLP), notadamente as intervenções promovidas por lideranças públicas capacitadas. Além de alguns exemplos de aplicação da teoria da mudança para a estruturação de programas do CLP, foram apresentados alguns resultados da avaliação de processos aplicada ao Master em Liderança e Gestão Pública (MLG), destacando resultados positivos de ex-alunos em relação a aspectos de competências, rede, carreira, políticas públicas e impactos socioeconômicos. Vale frisar que se tratou de uma avaliação de processos, não de impacto. De toda forma, a experiência trouxe muito aprendizado para o CLP, provocando reavaliações quanto aos serviços oferecidos.

## REFERÊNCIAS

- GERTLER, Paul J. et al. *Impact Evaluation in Practice*. Washington, DC: Inter-American Development Bank and World Bank, 2016. Disponível em: <<https://openknowledge.worldbank.org/handle/10986/25030>>. Acesso em: 07 de maio de 2018.
- HEIFETZ, Ronald; LINSKY, Marty. *Leadership on the Line, With a New Preface: Staying Alive Through the Dangers of Change*. Boston-MA: Harvard Business Press, 2017.
- HEIFETZ, Ronald Abadian; GRASHOW, Alexander; LINSKY, Martin. *The practice of adaptive leadership: Tools and tactics for changing your organization and the world*. Boston-MA: Harvard Business Press, 2009.

## INICIATIVAS DE AVALIAÇÃO DE IMPACTO REALIZADAS POR ÓRGÃOS PÚBLICOS ESTADUAIS: RS, RN e CE\*

Antônio Paulo Cargnin<sup>1</sup>

Cristian Quijada Torres<sup>2</sup>

Jaílma Carvalho<sup>3</sup>

Jimmy Lima de Oliveira<sup>4</sup>

**Resumo:** O texto apresenta experiências de avaliação de impacto realizada por três órgãos estaduais, em colaboração com a equipe do Banco Mundial. São abordadas experiências da Secretaria do Planejamento, Governança e Gestão do Governo do Estado do Rio Grande do Sul (SPGG-RS), da Secretaria da Educação do Governo do Estado do Rio Grande do Norte (SEEC-RN) e do Instituto de Pesquisa e Estratégia Econômica do Ceará (IPECE). São pontuadas algumas questões enfrentadas ao longo do processo de avaliação, tais como: as barreiras para construir e incorporar a agenda de avaliação no setor público; a importância da sustentação dessa agenda por parte das autoridades; as estratégias de convencimento das burocracias responsáveis pelos programas avaliados, bem como as parcerias firmadas para viabilizar a coleta e análise dos dados.

**Palavras-chave:** Políticas públicas. Avaliação de impacto. Programas governamentais.

**Abstract:** The text presents experiences of impact evaluations carried out by three state agencies, in collaboration with the World Bank Group. It will be presented experiences of the Secretariat of Planning, Governance and Management of the Rio Grande do Sul State (SPGG-RS), the Secretariat of Education of the Rio Grande do Norte State (SEEC-RN) and the Research and Economic Strategy of Ceará (IPECE). Some issues are tackled from these evaluation experiences, such as: barriers to building and incorporating the evaluation agenda in the public sector; the importance of authorities support for this agenda; the strategies of convincing the bureaucracies responsible for the evaluated programs, as well as the partnerships established to enable the collection and analysis of the data.

**Keywords:** Public policy. Impact evaluation. Governmental programs.

---

\* Texto elaborado a partir da transcrição da fala e da apresentação em slides de palestras proferida no Workshop Avaliação de Impacto de Políticas Públicas, realizado nos dias 29 e 30 de novembro de 2017, no Auditório da Secretaria de Planejamento, Governança e Gestão (SPGG-RS), com o apoio do Banco Mundial.

<sup>1</sup> Diretor do Departamento de Planejamento Governamental da Secretaria do Planejamento, Governança e Gestão do Governo do Estado do Rio Grande do Sul.

<sup>2</sup> Especialista em setor privado e inovação do Banco Mundial.

<sup>3</sup> Gerente da unidade executora setorial do projeto Governo Cidadão da Secretaria de Educação do Governo do Estado do Rio Grande do Norte.

<sup>4</sup> Analista de Políticas Públicas do Instituto de Pesquisa e Estratégia Econômica do Ceará (IPECE).

## INTRODUÇÃO

Experiências de avaliação de impacto de programas governamentais podem ser realizadas seja por equipe externa, seja por equipe interna ao governo. Neste artigo, são apresentadas experiências de avaliações conduzidas por três órgãos estaduais diferentes, todos com intensa colaboração com a equipe do Banco Mundial.

Serão abordadas brevemente as experiências da Secretaria do Planejamento, Governança e Gestão do Estado do Rio Grande do Sul (SPGG-RS) e da Secretaria da Educação do Estado do Rio Grande do Norte (SEEC-RN) e, mais detidamente, a experiência do Instituto de Pesquisa e Estratégia Econômica do Ceará (IPECE). Os relatos foram apresentados em uma rodada de discussão e perguntas no *Workshop Avaliação de Impacto de Políticas Públicas*, cujo mediador foi o Especialista setorial do Banco Mundial, Cristian Quijada Torres.

Em cada caso, buscam-se abordar algumas questões enfrentadas pela equipe de avaliação ao longo do processo, tais como: as barreiras para construir e incorporar a agenda de avaliação; a importância da sustentação dessa agenda por parte das autoridades; as estratégias de convencimento das burocracias responsáveis pelos programas avaliados, bem como as parcerias firmadas e necessárias não só para estruturar a capacidade de avaliação, mas também para coletar e analisar os dados. A partir desses relatos, extraem-se algumas lições das experiências.

### A EXPERIÊNCIA DO ESTADO DO RIO GRANDE DO SUL

Especialmente no que tange ao planejamento do Estado e ao planejamento das secretarias, investir na questão de avaliação é trabalhar no sentido de qualificar a elaboração das políticas. Avaliar requer parar e repensar o desenho das políticas que são feitas e como sempre foram feitas, para que se possa então medir seus resultados. O período atual é o momento de investir nisso, pois marca o início do processo de elaboração do novo plano plurianual. Essa é a reflexão inicial.

No Estado do Rio Grande do Sul, a preocupação com a questão da avaliação começa basicamente a partir da agenda nacional e da reestruturação do processo de planejamento, especialmente no PPA 2004-2007. A partir daí, iniciam-se alguns trabalhos na Secretaria do Planejamento, feitos em conjunto com a Secretaria da Fazenda (SEFAZ) e a Fundação de Economia e Estatística (FEE), que passam a despertar o interesse na avaliação e no processo propriamente dito. Além disso, houve um interesse individual, especialmente por parte dos técnicos, quanto à necessidade e à importância desse tema.

Do ponto de vista da construção da temática na Secretaria, pode-se citar, inicialmente, a participação da diretora-adjunta do Departamento de Planejamento Governamental (DEPLAN-SPGG) em curso do Projeto Minerva em Nova York, no ano de 2006, do qual resultou um trabalho que até hoje é usado como referência sobre avaliação (CUNHA, 2006). Depois, os servidores da Secretaria fizeram alguns cursos, dentre os quais da Fundação Itaú Social.

Servidores também iniciaram estudos de pós-graduação nessa área, já tendo sido publicada dissertação de mestrado referente à avaliação de política pública de educação (XERXENEVSKY, 2012). Atualmente, está em curso tese de doutorado na área de avaliação em saúde (HOFFMANN, no prelo). Diante desse interesse individual, a Secretaria tem incentivado a busca de conhecimento para qualificar as avaliações realizadas.

Além da qualificação, um grande avanço foi a estruturação do Núcleo de Monitoramento e Avaliação (NUMA) no DEPLAN-SPGG, que reuniu diferentes servidores de distintos departamentos e que já vinham trabalhando com a temática. Afora isso, sabe-se que se trata de um processo de convencimento dos secretários de estado e gestores de que a avaliação é importante. Tem-se ainda um caminho a traçar, mas agora já existe um ambiente muito mais favorável para a avaliação de impacto no Estado do Rio Grande do Sul, tendo em vista melhorar a efetividade das políticas.

O primeiro estudo de avaliação de impacto elaborado pela Secretaria do Planejamento internamente, sem contratação externa, foi o do Programa de Extensão Produtiva e Inovação (PEPI). Nessa primeira experiência, a Secretaria de Desenvolvimento Econômico, que coordena o Programa, mostrou-se disponível e receptível à avaliação. Hoje, as demais secretarias também estão se engajando nessa proposta de avaliação de políticas públicas para próximos trabalhos.

Contudo, é claro que esse convencimento não se encerra aqui, porque, na verdade, ainda há a necessidade de deixar mais claro a diferença, por exemplo, de se acompanhar uma política, de se monitorar uma política e de se avaliar uma política. Todas as etapas e todos os processos são importantes e têm sua função, mas são coisas diferentes, que cumprem papéis diferentes no processo de elaboração de política pública. Mas a avaliação é uma etapa e deve ser percebida como uma etapa de qualquer processo. Se houver técnicos com o conhecimento e convencidos disso, fica muito mais fácil. Capacidade interna serve também como um cartão de visitas para poder mostrar para o gestor que o programa tem determinadas características e certos resultados – e que não é preciso contratar externamente para saber como está andando essa política.

Porém, o menos importante nesse momento é quem está fazendo avaliação. O que importa é fazer a avaliação. A credibilidade de uma avaliação de impacto não está no fato de ser feita externamente ou internamente. O fator determinante é o rigor metodológico, a capacidade de quem faz e a avaliação que é feita. Em algumas circunstâncias é melhor ter uma avaliação interna. Se uma secretaria tiver condições de fazer, acredita-se ser melhor, especialmente por facilitar o acesso às informações administrativas. Além disso, muitas vezes, a avaliação não é para ser exposta externamente, mas para corrigir uma política internamente, um programa ou suas ações – e para isso independe se é feita internamente ou externamente. Se houver capacidade técnica, qualidade metodológica, pessoas habilitadas, é possível fazer internamente. Se não, podemos procurar ajuda externamente. Isso não faz diferença, o importante é avaliar. Todo o caminho é válido, o que não pode acontecer é abdicar de avaliar política pública porque não é a melhor forma. Resta claro que a temática é complexa.

Diante disso, é possível fazer parcerias com universidades e outras instituições. Ademais, eventos para apresentação e discussão do tema são importantíssimos, na medida em

que permitem que as pessoas que trabalham nos órgãos do Estado, nas secretarias e os próprios gestores consigam saber da importância de fazer avaliação de impacto e saber também como trabalhar, no caso de alguém externamente que venha a fazer avaliação. Tratar do tema da avaliação em seminários, aperfeiçoamentos internos e em diálogo com as universidades é uma evolução.

Nesse processo, o gestor precisa entender de forma clara que o que está se fazendo não é no sentido punitivo ou de uma avaliação pessoal. Na verdade, é uma forma de mostrar para a sociedade que um programa desenvolvido por determinada secretaria pode ser direcionado e pode ser aperfeiçoado. Na maioria das vezes, um programa não é criado por uma gestão. É algo antigo, que tem história. É muito raro encontrar um exemplo de programa fundamental para a sociedade que tenha sido interrompido. Ele pode ser redirecionado, e a avaliação de impacto vai permitir isso, ao dar clareza para que o gestor tenha subsídios para decidir sobre melhorias na implementação, pois talvez, a forma como esteja sendo implementado, não seja a melhor.

Essa percepção depende de uma criação de cultura de avaliação e de um trabalho de despertar para a importância da avaliação e mostrar seus caminhos. Não se trata só de que um número maior de pessoas possa, cada vez mais, fazer a avaliação, mas, sim, de que mais pessoas entendam o que é e para que serve.

Atualmente, é muito comum ouvir nas falas dos gestores que é preciso saber se o programa está dando certo, se está tendo o resultado que deveria ter. Claro que ainda não é o ideal, não temos ambiente totalmente aberto à avaliação, mas as dificuldades desses vinte anos permitiram que os gestores já tivessem um pouco mais de conhecimento e de clareza sobre a importância da avaliação.

## A EXPERIÊNCIA DA SECRETARIA DE EDUCAÇÃO DO ESTADO DO RIO GRANDE DO NORTE

A experiência do Estado do Rio Grande do Norte com a temática da avaliação está inserida no contexto do projeto Governo Cidadão, um projeto multissetorial, que conta com o empréstimo do Banco Mundial para executar ações para o desenvolvimento da educação no Estado. Há uma intencionalidade e um recurso disponível para destinar às escolas, para que elas próprias desenvolvam ações inovadoras no intuito de modificar práticas pedagógicas relacionadas ao professor, à motivação docente, à relação professor-aluno, entre outras. A ideia da implementação existia; no entanto, não havia a estrutura formal do projeto.

No Estado do Rio Grande do Norte, não existe a cultura de formular políticas próprias. A exemplo do Brasil, como um todo, os estados federais e os municípios recebem políticas já desenhadas por terceiros. O primeiro desafio do Estado foi, portanto, justamente formular e implementar uma política própria, com o apoio do Banco Mundial. Nesse sentido, a experiência com avaliação de impacto do Estado coincide com a necessidade de formular uma política e implementá-la, considerando sua estrutura, seus processos e procedimentos, dentro

de uma lógica de gestão por resultados. Isso também foi algo novo que a avaliação traz. A gestão por resultados foi um ponto muito incentivado pelo Banco Mundial. A equipe do *Development Impact Evaluation* (DIME) fazia esse papel, às vezes questionador, mas muito oportuno, de instigar a pensar melhor os objetivos, as metas, o público de focalização. Isso gerou um conflito na “subjetividade” dos participantes do processo, no que se refere ao entendimento da avaliação relacionada ao projeto que se estava querendo implementar. Eram muitos participantes envolvidos e, num primeiro momento, foi necessário esclarecer o que se queria, o que era necessário refletir, para então conseguir escrever a proposta.

O processo de avaliação do Projeto de Inovação Pedagógica (PIP) ainda está em andamento. A discussão começou em 2014 e 2015, quando houve os procedimentos de coleta de dados e de pensar o desenho da cadeia de resultados. Verificou-se que a carência do Estado em termos de cultura de planejamento, inclusive do ensino, abriu espaço para que o grupo de orientadores que acompanha as escolas no projeto acabasse estruturando uma metodologia própria de acompanhamento da implementação do projeto, assim como contribuindo diretamente no fortalecimento do planejamento do professor – inclusive na articulação dos diversos programas que, com frequência, a escola recebe, mas entre os quais não sabe, às vezes, fazer as conexões. Hoje, os indicadores vêm sendo medidos, e foi instituído um *blog* para trocas de experiências, no qual se ouvem professores, gestores e alunos, que compartilham depoimentos, matérias específicas e vivências. Por meio desse *feedback*, o programa confere um protagonismo à escola. A satisfação é percebida nos depoimentos postados no *blog* e nas visitas às unidades educativas.

De toda forma, o Estado do Rio Grande do Norte ainda é principiante, tanto em termos de formulação de políticas públicas, quanto em avaliação. Essa é a primeira experiência de avaliação de impacto realizada. Assim que saírem os primeiros resultados, será possível uma nova conotação e um novo direcionamento. Isso porque, percebe-se a avaliação como um elemento de tomada de decisão, de continuidade ou não, mas também um elemento para repensar os desenhos daquilo que já é feito. Nisso a Secretaria de Educação está hoje muito focada.

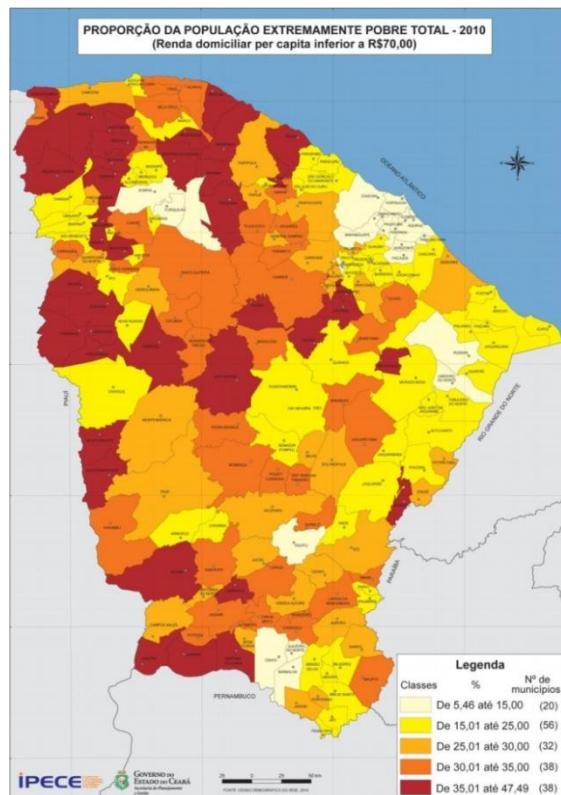
É importantíssimo que os gestores tenham ciência da concepção da avaliação e liderem esse processo. Seria mais fácil do que o movimento contrário, de baixo para cima. Para institucionalizar essa forma de trabalhar e promover essa conscientização, é preciso avaliar o que é feito para melhor fazer. Muito embora seja importante que tanto mais servidores se apropriem e entendam do processo de avaliação, mais importante ainda é compartilhar essa clareza, dando visibilidade à avaliação, sem medo de críticas, de expor para julgamento, a fim de se realmente conseguir melhorar o que está fazendo. Dessa forma, a melhor forma de conduzir esse processo é compartilhando essa responsabilidade internamente, juntamente com o corpo técnico, para caminhar todos na mesma direção.

## A EXPERIÊNCIA DO INSTITUTO DE PESQUISA E ESTRATÉGIA ECONÔMICA DO CEARÁ (IPECE): AVALIAÇÃO DE IMPACTO DO PADIN

O Instituto de Pesquisa e Estratégia Econômica do Ceará (IPECE) é uma autarquia ligada à Secretaria de Planejamento e Gestão do Ceará, que tem o papel de auxiliar as secretarias do Estado do Ceará no monitoramento e avaliação de políticas públicas. A seguir, relata-se experiência de avaliação experimental de programa de visitação domiciliar para crianças na primeira infância, sob a responsabilidade da Secretaria de Educação.

O Programa de Apoio ao Desenvolvimento Infantil (PADIN) é um programa do Governo do Estado do Ceará realizado em parceria com os municípios, que se caracteriza por visitas domiciliares, encontros coletivos e comunitários para apoiar e orientar as famílias no cuidado e na educação das crianças de 0 a 3 anos e 11 meses não atendidas nos Centros de Educação Infantil (CEI). O objetivo é a formação de competências familiares necessárias para estimular o desenvolvimento cognitivo e socioemocional por meio de brincadeiras e da interação com os adultos. O objetivo do Programa é a redução dos déficits de habilidades, que se originam antes do ingresso no sistema escolar, com o fim de tornar mais eficiente o processo de aprendizagem em idades posteriores. O PADIN pretende atender 80 famílias por município, nos 36 municípios com os maiores percentuais de pessoas em extrema pobreza, de acordo com os dados do Censo Demográfico de 2010 (Figura 1, em bordô).

Figura 1 – Proporção da população extremamente pobre no Estado do Ceará – 2010

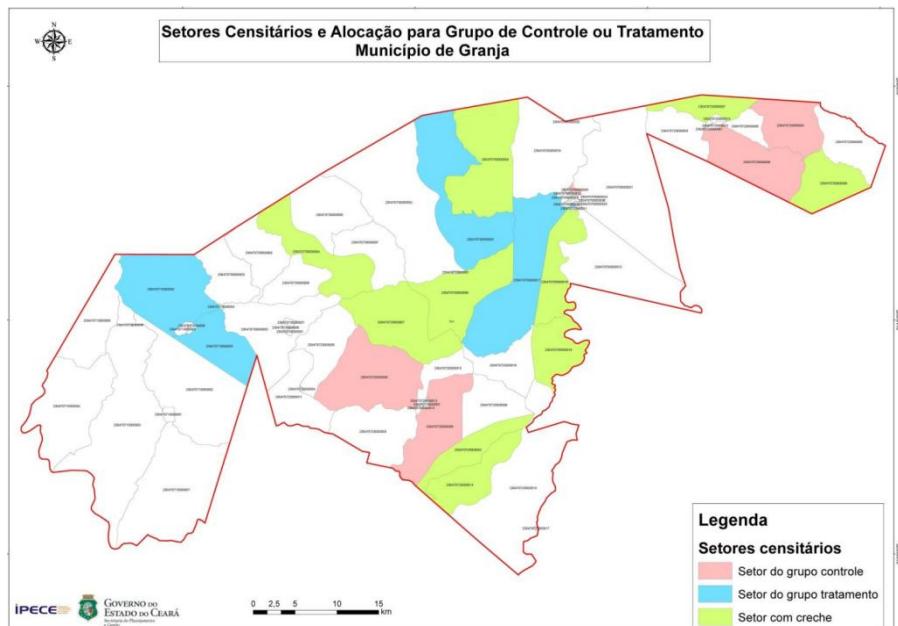


Fonte: Brasil, 2010.

O plano de avaliação do PADIN pretende: (a) avaliar o impacto do programa sobre o desenvolvimento infantil; (b) calcular quanto desse impacto decorre de efeitos diretos no aumento da qualificação dos pais; (c) estimar quanto do impacto decorre de efeitos indiretos do aumento da qualificação; (d) analisar o impacto do programa sobre a percepção dos pais a respeito da importância do investimento familiar no desenvolvimento infantil; (e) documentar o quanto do aumento da percepção se traduz em aumento dos investimentos familiares e (f) entender porque algumas famílias podem se beneficiar mais que outras do PADIN (efeitos heterogêneos).

É importante mencionar como foi feita a aleatorização no processo de planejamento da avaliação de impacto do Programa de Apoio ao Desenvolvimento Infantil. O fato de o programa ser implementado nos 36 municípios mais pobres do Estado do Ceará facilitou o convencimento do gestor, pois nenhum questionaria uma atuação a favor dos municípios mais pobres. A dificuldade, contudo, foi convencer o gestor quanto à área em que se iria atuar dentro do município. Foi preciso construir mapas de pobreza usando os setores censitários para mostrar que havia áreas mais pobres do que outras. Feito isso, cabia verificar como aleatorizar dentro desses cenários. Buscou-se, então, comparar como o Programa impactava crianças com idade de 0 a 3 anos que não eram atendidas em creches em contraposição às crianças atendidas por creches. A partir desse desenho da avaliação, a aleatorização foi feita. A Figura 2 apresenta as áreas de tratamento (azul), as áreas do grupo controle (rosa), bem como as áreas onde havia creches (verde).

Figura 2 – Distribuição dos setores de tratamento, controle e com creche: exemplo do município de Granja-CE.



Fonte: Brasil, 2010.

Feita a aleatorização, deve-se ter um cuidado quando se vai a campo, pois é importante analisar como é possível garantir que as famílias estejam dentro daquele setor estipulado. A empresa que coleta os dados pode chegar na localidade e confirmar por GPS, mas as casas estão espalhadas na região, havendo um limite claro entre setores censitários. Assim, uma dificuldade que só se veio a perceber na prática é que o limite dos setores muitas vezes é uma estrada, sendo que, no interior, muitas casas ficam em torno da estrada. Então você tem um setor censitário de um lado, que poderia estar na amostra, e um setor do outro lado, com famílias que partilham uma mesma realidade, mas que não estão na amostra. Portanto, a dificuldade não é fazer o sorteio necessário para a aleatorização. O difícil é convencer o gestor a atuar apenas em um lugar, se ele tem recurso para expandir.

Esse é o tipo de discussão que é feita quando da decisão da aleatorização. Há certa resistência em utilizar aleatorização pra escolher quem vai participar do programa, uma vez que os gestores tendem a assumir de antemão que o programa tem resultado e que, ao aleatorizar, se estaria negando o benefício para alguém. Nesse caso da avaliação de impacto do PADIN, a dificuldade era convencer o gestor da pasta estadual da necessidade de se chegar ao gestor municipal e explicar-lhe que ele não poderia escolher as áreas onde o programa iria atuar e onde não iria. Somado a isso, o gestor não poderia expandir o Programa para as outras áreas enquanto não houvesse o resultado da avaliação. No caso do Ceará, como há uma boa cooperação entre Secretaria Estadual de Educação e as secretarias municipais, por conta do programa de alfabetização na idade certa, isso foi mais ou menos tranquilo. De toda forma, levaram-se sete meses até convencer o gestor da pasta a utilizar a aleatorização.

Outra discussão por que se passou na avaliação do PADIN foi quanto aos instrumentos que seriam utilizados, o que exigiu igualmente um processo de convencimento. O Programa em questão consistia na estimulação ao desenvolvimento infantil. Questiona-se, pois, como se mede desenvolvimento infantil. No campo da Educação Infantil, essa não é uma discussão fácil, pois se entende que não é para haver avaliação de crianças nessa fase. Podem-se avaliar processos, métodos, mas não crianças. Sendo assim, a dificuldade foi convencer que não era uma avaliação das crianças, mas dos resultados pretendidos pelo programa. Então, houve a necessidade de discutir os tipos de instrumentos a utilizar.

Além dessas discussões, é preciso entrar na questão dos custos. A avaliação do PADIN não foi barata, por causa da quantidade de instrumentos empregados e pela complexidade do que se queria medir. Quando foi apresentado o custo de R\$ 3,6 milhões para coleta de dados de 4 mil famílias, em 22 municípios, surgiu a pergunta: "Se eu pegasse esse valor e aplicasse no programa, quantas crianças eu atenderia sem precisar ter que fazer avaliação?". Diante disso, coube uma réplica: "Quanto será gasto para ofertar um programa que não funciona ou que precisa de ajustes?" O que se pretende demonstrar é que avaliação, à primeira vista cara, na verdade economizará recursos, na medida em que será possível fazer os devidos ajustes antes de expandir o Programa ou se concluirá que tal programa não deverá ser expandido, por não produzir efeito. Isso sim é economia de recurso.

Na verdade, tem sido um aprendizado pensar nesse processo completo. Além do desenho e dos custos, tem que se pensar nos prazos. Por exemplo, muitas avaliações são feitas em acordo de cooperação com o Banco, que presta assessoria técnica. Para contratar, é

necessário realizar as licitações seguindo as regras do Banco. No Ceará, uma das dificuldades é cumprir os prazos dos processos, e muitas vezes acaba-se prejudicando o tempo da intervenção, porque isso demanda tempo.

Portanto, idealmente, o desenho, o custo e os prazos da avaliação têm que já estar discriminados desde o início do programa, na fase de planejamento. Se pensar na avaliação somente depois, perde-se o momento inicial que se tinha para coletar dados. É preciso pensar desde logo como será a coleta de dados. Será contratada uma empresa para isso? O que é preciso para garantir a qualidade dos dados necessários? Muitas vezes, implementa-se um programa sem nem ter se pensado nos indicadores com os quais se mediria o resultado. Além disso, se não se orçar antes, depois se aloca todo o orçamento para as intervenções, e faltarão recursos para se fazer a avaliação.

Então, muitas vezes, mesmo que haja a cultura da avaliação, ela é pensada e colocada como se fosse uma coisa à parte da intervenção, enquanto não é. No Estado, tem-se tentando fomentar essa cultura para todos os projetos que sejam apresentados para financiamento do fundo de combate à pobreza, os quais já devem vir com o desenho da avaliação e com a sua previsão de custo.

Essa experiência da avaliação do PADIN deixa evidente que a questão principal é que as avaliações sejam feitas de forma sistemática e regular. Não adianta fazer uma avaliação pontual de algum programa ou outro. O Estado emprega uma quantidade grande de recursos para fazer diversos programas, e nem todos são de fato avaliados. Muitos sequer são colocados em um formato passível de avaliação. Isso vem sendo trabalhado há certo tempo no Estado. O Ceará já tem uma experiência de longa data com a avaliação de alguns programas, tais como na área de agricultura, Projeto São José, em parceria com o Banco Mundial, e na área de educação. Mas é necessário um trabalho de convencimento das secretarias e do Governo para realização da avaliação.

Essa revolução por baixo está sendo feita há uns 10 anos no Ceará. Conseguiram-se algumas vitórias, mas facilita muito quando a determinação vem de cima. É muito tranquilo no Estado quando se tem que fazer uma avaliação de impacto de projeto que é financiado pelo Banco Mundial ou outra instituição, porque isso já vem no pacote, e essa negociação é feita “em cima”. No entanto, quando a proposta vem de baixo, dizendo que tem um programa que merece ser avaliado pelo seu formato, na maioria das vezes o processo de convencimento precisa ser maior, pois as resistências são muito maiores. Muitas vezes, cabe ao técnico sempre estar enfatizando a importância da avaliação<sup>5</sup>. Ou seja, quando isso vem de cima, a

---

<sup>5</sup>O IPECE é um órgão técnico que não executa as políticas públicas finalísticas. É importante que se tenha alguém ou algum órgão do Estado participando das avaliações, mas que não seja da secretaria executora. Há uma dificuldade por parte da secretaria executora quando é feita uma contratação direta da avaliação, mas não se tem alguém com conhecimento técnico para especificar o que é preciso na avaliação. Sem esse conhecimento mínimo, não será possível avaliar a qualidade do produto entregue. Houve uma situação em que já se estava no meio de um processo de avaliação contratado pela Secretaria de Assistência Social, e o consultor utilizou um modelo de regressão que não fazia muito sentido. Se não houvesse alguém com conhecimento de avaliação de impacto na Secretaria, o resultado da avaliação seria recebido sem questionamento. Além disso, deve-se prestar atenção a que, quando uma secretaria contrata diretamente alguém externo, existem outros interesses, os quais muitas vezes podem não ser responder às perguntas que estão colocadas. Sem contar que, quando está sendo vendido um serviço de consultoria, há o receio de se chegar ao final e verificar que o resultado foi um impacto zero. Nesse

pauta está posta. Quando vem por baixo, tem que tentar criar essa pauta. O risco de fazer por baixo é o risco de desenhar a avaliação, propor um método e, quando chegar na hora de implementar, o gestor não aceitar fazer a avaliação. Portanto, a ideia é a criação de um decreto governamental que institua o trabalho de avaliação de forma permanente, criando para tanto um núcleo de monitoramento e avaliação de políticas públicas no IPECE. Assim, busca-se conseguir de fato com que o processo de avaliação seja colocado como uma função que tem que ser executada sempre que se fizer uma intervenção.

## CONSIDERAÇÕES FINAIS

Investir na questão de avaliação é trabalhar no sentido de qualificar a elaboração das políticas públicas. As experiências com avaliação de impacto coincidem com a necessidade de formular uma política e implementá-la, considerando sua estrutura, seus processos e seus procedimentos, dentro de uma lógica de gestão por resultados.

Avaliar o impacto de um programa é verificar as mudanças diretamente atribuíveis ao programa. Para isso, é necessário separar o efeito do programa do efeito que outras variáveis podem ter sobre os resultados observados. A aleatorização é muito utilizada nas avaliações, uma vez que permite serem gerados grupos estatisticamente idênticos. No entanto, há certa resistência em utilizar aleatorização para escolher quem vai participar do programa, uma vez que os gestores tendem a assumir de antemão que o programa tem resultado e que, ao aleatorizar, se estaria negando o benefício para alguém. Quanto aos custos, é importante citar que, à primeira vista, a avaliação pode parecer cara. No entanto, avaliar gera economia de recursos na medida em que será possível fazer os devidos ajustes antes de expandir o programa.

Neste texto, foram apresentadas diferentes experiências de avaliação, levadas a cabo por órgãos estaduais. Não obstante, a credibilidade de uma avaliação de impacto não está no fato de ser feita externamente ou internamente. O fator determinante é o rigor metodológico utilizado no processo de avaliação. O menos importante nesse momento é quem está fazendo avaliação. O que importa é fazer a avaliação.

## REFERÊNCIAS

- BRASIL. IBGE. *Censo demográfico*. Rio de Janeiro: Instituto Brasileiro de Geografia e Estatística, 2010.
- CUNHA, C. G. *Avaliação de políticas públicas e programas governamentais: tendências recentes e experiências no Brasil*. Porto Alegre: Secretaria de Planejamento do Rio Grande do Sul, 2006.

---

cenário, dificilmente o consultor conseguiria outra licitação naquela secretaria. No entanto, a contratação externa é boa quando se traz alguém de fato especialista na área. A questão de credibilidade é muito importante.

HOFFMANN, J. *Avaliação de impacto dos Núcleos de Apoio à Atenção Básica (NAABs) no Rio Grande do Sul.* No prelo. Tese (Doutorado em Epidemiologia) – Programa de Pós-Graduação em Epidemiologia, Universidade Federal do Rio Grande do Sul, Porto Alegre.

XERXENEVSKY, L. *Programa mais educação:* avaliação do impacto da educação integral no desempenho de alunos no Rio Grande do Sul. 2012. Dissertação (Mestrado em Economia do Desenvolvimento) – Programa de Pós-Graduação em Economia do Desenvolvimento, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre.

## EXPERIÊNCIAS DE AVALIAÇÃO DE IMPACTO NO TERCEIRO SETOR\*

Caio Piza<sup>1</sup>

Thiago Nascimento<sup>2</sup>

Marta Aresta<sup>3</sup>

Janaína Soares<sup>4</sup>

**Resumo:** O texto apresenta experiências de avaliação de impacto em curso ou finalizadas, realizadas pela Associação de Educação Financeira do Brasil (AEF-Brasil), pelo Serviço Brasileiro de Apoio às Micro e Pequenas Empresas do Rio de Janeiro (SEBRAE-Rio) e pelo Banco Central do Brasil, em parceria com o Banco Mundial. São apresentadas as etapas para a construção das avaliações, bem como a decisão institucional quanto à realização de avaliações de impacto, dos resultados encontrados nos experimentos e de sua aplicação para captar recursos e/ou validar a agenda do projeto ou da ação propostos pela organização. Essas experiências expõem diferentes ganhos em termos de aprendizado quanto aos processos e procedimentos levados a cabo antes, durante ou depois da implementação de ações, projetos e intervenções.

**Palavras-chave:** Avaliação de impacto. Educação financeira. Cidadania financeira.

**Abstract:** The text presents current or completed impact assessment experiments which have been carried out by the Association of Financial Education in Brazil (AEF), the Brazilian Service of Support to Micro and Small Enterprises (SEBRAE) of Rio de Janeiro and the Central Bank of Brazil in partnership with the World Bank Group. The steps for the construction of the evaluation researches are presented, as well as the institutional decision to carry out impact evaluation in each sector. It will also be presented some of the researches results and their use in order to attract resources and/or validate the project or action agenda proposed by the organization. These experiences show different gains in terms of learning about the processes and procedures which are carried out before, during or after the implementation of public actions, projects and interventions.

**Keywords:** Impact evaluation. Financial education. Financial citizenship.

---

\* Texto elaborado a partir da transcrição da fala e da apresentação em slides de palestras proferida no Workshop Avaliação de Impacto de Políticas Públicas, realizado nos dias 29 e 30 de novembro de 2017, no Auditório da Secretaria de Planejamento, Governança e Gestão (SPGG-RS), com o apoio do Banco Mundial.

<sup>1</sup> Representante do Banco Mundial.

<sup>2</sup> Representante da Associação de Educação Financeira do Brasil.

<sup>3</sup> Representante do SEBRAE do Rio de Janeiro.

<sup>4</sup> Representante do Banco Central do Brasil.

## INTRODUÇÃO

O Banco Mundial apoia a realização de avaliações de impacto em vários setores. Neste texto, apresentamos experiências realizadas em parceria com o Banco Central do Brasil, o Serviço Brasileiro de Apoio às Micro e Pequenas Empresas do Rio de Janeiro (SEBRAE-Rio) e da Associação de Educação Financeira do Brasil (AEF-Brasil). Com isso, objetiva-se apresentar os passos para a construção das avaliações em cada caso. Busca-se discorrer acerca da decisão institucional quanto à realização de avaliações de impacto, dos resultados encontrados nos experimentos e de sua aplicação para captar recursos e/ou validar a agenda do projeto ou da ação propostos pela organização. Além disso, são pontuados alguns desafios e aprendizados advindos dessas experiências.

## AVALIANDO PROJETOS DE EDUCAÇÃO FINANCEIRA: A EXPERIÊNCIA DA AEF-BRASIL

A Associação de Educação Financeira do Brasil (AEF-Brasil) é uma organização da sociedade civil de interesse público (OSCIP), criada para operar a política de estado chamada Estratégia Nacional de Educação Financeira – ENEF, regulada por meio do Decreto Federal nº 7.397, de 22 de dezembro de 2010. A política institui ações de educação financeira pelo País e confere à AEF a responsabilidade de desenvolver, pilotar, avaliar e escalar os projetos de educação financeira por todo o Brasil.

A política é coordenada pelo Comitê Nacional de Educação Financeira, formado por treze membros: os quatro reguladores do mercado financeiro (Banco Central do Brasil, Comissão de Valores Mobiliários – CVM, Superintendência Nacional de Previdência Complementar – SUSEP e Superintendente da Superintendência de Seguros Privados – PREVIC), três ministérios (Fazenda, Educação, Justiça) e seis representantes da sociedade civil, sendo quatro entidades do mercado financeiro do setor privado (Federação Brasileira de Bancos – FEBRABAN, Bolsa de Valores de São Paulo – B3, Associação Brasileira das Entidades dos Mercados Financeiro e de Capitais – ANBIMA e Confederação Nacional das Seguradoras – CNSeg), o SEBRAE Nacional e o Conselho Nacional dos Secretários de Educação (CONSED).

Dessa forma, a ENEF conta com uma configuração moderna e inovadora, com um comitê formado por grandes representantes do governo e também grandes entidades do setor privado. As quatro entidades do setor financeiro privado (FEBRABAN, B3, ANBIMA e CNSeg) mantêm a Associação, por meio de um recurso fixo anual que custeia a equipe e a estrutura de espaço. Por sua vez, os projetos de educação financeira dependem de captação de recursos. Todos os recursos captados vão para ponta, ou seja, são aplicados no projeto de educação financeira, o que é muito positivo.

Dada a escala do programa, a ser desenvolvido em âmbito nacional, o Comitê Nacional de Educação Financeira entendeu ser necessário desenvolver uma avaliação de impacto para verificar a melhor forma de implementá-lo. Foram realizadas duas avaliações (BRUHN *et al.*, 2013). A primeira ocorreu de 2010 para 2011 e buscou avaliar um projeto-piloto de educação financeira no Ensino Médio. Esse piloto durou 18 meses e foi o maior piloto de educação

financeira no mundo, já avaliado em escolas. Envolveu 900 escolas, quase 20 mil alunos e mais de 2 mil professores. O material utilizado consistia em livros com conteúdo de educação financeira, com uma linha de comportamento e de competências alinhada ao currículo escolar. O material foi desenvolvido e coordenado pelo Ministério da Educação, com conteúdo transversal, capaz de conversar com todas as disciplinas. Não houve, portanto, criação de novas disciplinas, de modo que qualquer professor poderia ser um disseminador do programa. Cabe dizer que a educação financeira entrou na base nacional comum curricular, compondo um dentre os 11 temas transversais estabelecidos.

A avaliação desse programa-piloto de educação financeira chegou a resultados interessantes. Verificou-se que o programa afetou positivamente não só o letramento financeiro (mensurado pela capacidade de poupar, de fazer planejamentos, entre outras coisas), mas também melhorou a capacidade de leitura e interpretação de texto, bem como as habilidades matemáticas dos jovens participantes do programa.

Esse ponto é fundamental, tendo em vista o desafio de melhorar o Índice de Desenvolvimento da Educação Básica (IDEB) brasileiro – o equivalente ao Programa Internacional de Avaliação de Estudantes (PISA). O programa de educação financeira na escola melhorou a educação como um todo, e só um estudo como o realizado poderia ter identificado esses resultados, nessa profundidade.

Uma segunda experiência de avaliação foi realizada em 2015 com um projeto-piloto no Ensino Fundamental, no mesmo modelo que foi feito no Ensino Médio, com material específico desenvolvido por integrantes do Comitê Nacional e alinhado ao currículo dos nove anos do Ensino Fundamental. Foram pesquisadas 200 escolas em dois municípios com características bem diferentes, Joinville e Manaus. Embora fossem menos escolas, também foi composta uma amostra grande. O material foi testado nos anos 3º, 5º, 7º e 9º do Ensino Fundamental, por uma questão de custos e de metodologia. Isso porque a metodologia do terceiro ano é a mesma do primeiro e do segundo. A metodologia do quarto é a mesma do quinto, e assim sucessivamente.

Tendo em vista a necessidade de captar recursos para esse piloto acontecer, houve dificuldades na sua implementação, uma vez que ocorreu atraso no pagamento. A situação do investimento social privado tem mudado nos últimos cinco anos, e em 2015 foi um ponto crítico. Em virtude disso, o escopo do projeto precisou ser reduzido, e houve demora para iniciar a implementação. Algumas escolas demoraram para receber o material, sendo desconhecido o efeito desse atraso sobre o resultado final.

De toda forma, o resultado dessa segunda avaliação demonstrou que os anos 5º, 7º e 9º, especialmente os dois últimos, tiveram um ganho melhor. O terceiro ano, por sua vez, apresentou um ganho bem menor. Compreender essa diferença de resultados será tarefa para uma próxima avaliação. Espera-se responder se ou o tempo de implementação não foi suficiente para idades mais jovens, ou se deveria ter sido feita de outra forma (outro material mais adequado para a idade), ou ainda se alunos do terceiro ano são muito jovens e não têm maturidade suficiente para entender qualquer conceito de planejamento.

Portanto, um resultado importante da avaliação do programa diz respeito à qualidade da implementação. Isso porque não existe avaliação de impacto experimental em que não haja problemas com a implementação, tendo em vista que não se consegue antecipar todos os problemas que podem ocorrer. Isso influencia a decisão de escalar ou não a intervenção. Afora isso, deve-se ter cuidado, pois é possível testar um piloto, encontrar efeito e, depois, quando escalar, o efeito desaparece – não porque o programa não funciona, mas porque a forma como ele foi implementado no piloto é diferente. No piloto, consegue-se monitorar melhor as atividades. Geralmente há um coordenador de campo responsável por certificar-se de que as atividades estão acontecendo de fato. Quando se expande o projeto, porém, não se trata mais da mesma equipe em campo – e isso não é um problema do programa, é um problema de quem implementa. A implementação, com a qual as instituições envolvidas têm que estar comprometidas, é extremamente importante.

Os resultados da avaliação serviram também a um dos desafios da AEF, que é justamente o de captar recursos. A necessidade de buscar o setor privado para investir em um programa do tamanho da ENEF não é fácil. O setor privado que investe em educação de fato olha para esse tipo de avaliação, porque quer saber como funcionam os programas. O setor social tem que demonstrar o que é factível e escalável. Sendo assim, a realização das avaliações conta para a decisão dos investidores. Um desafio é trazer investimento social privado para fazer a avaliação em si. Em parte porque as grandes empresas parceiras não entendem ainda tanto sobre avaliação, embora reconheçam a força de uma avaliação com a marca do Banco Mundial. Contudo, tomar decisão de gastar “x” do orçamento em uma avaliação de impacto enfrenta alguma resistência.

Após haver captado os recursos para a avaliação e validado o material e seu impacto sobre o aluno, o novo desafio da AEF foi escalar o programa de educação financeira, o que depende única e exclusivamente de financiamento privado. São 170 mil escolas públicas no País, e a missão da AEF é chegar a todas. Somam-se 50 milhões de alunos em toda a Educação Básica. Como é um modelo analógico, questionou-se como imprimir livros para 50 milhões de alunos contando apenas com investimento social privado. Com esse modelo de financiamento, é pouco provável que se conseguisse<sup>5</sup>.

Diante desse impasse, ao longo de 2017, a AEF focou na construção de um projeto multimídia com produtos escaláveis que refletem o material testado. Foi feita uma campanha em um programa de televisão, em parceria com a TV Escola (braço de comunicação do Ministério da Educação). Também foi criada plataforma de educação a distância para a formação de professores, e encontra-se em fase de construção um *game*. Trata-se basicamente da transposição dos conteúdos dos livros para um formato de jogo educativo. Dessa forma, busca-se alcançar a meta de alunos, cobrando apenas custo de hospedagem e manutenção das plataformas. Contudo, não se avaliou ainda se o efeito sobre a aprendizagem

---

<sup>5</sup> O Tocantins é um *case* de sucesso, pois é o estado que tem maior adesão no País: 100% das escolas públicas aderiram à educação financeira. No Estado, a AEF focou todo o investimento social privado captado por meio de parcerias para formar professores, ao passo que o Estado se comprometeu a imprimir livros para todos os alunos do Ensino Fundamental e Ensino Médio. Evidentemente, o ideal seria que o mesmo acontecesse em todas as Unidades da Federação, mas não há uma fórmula.

do aluno é o mesmo que o material físico<sup>6</sup>. A expectativa é de que o jogo funcione melhor no Ensino Fundamental 2 e no início do Ensino Médio, pois é essa é a idade de quem joga mais.

## AVALIAÇÃO DE PROJETOS DE APOIO À GESTÃO: A EXPERIÊNCIA DO SEBRAE-RIO

O Banco Mundial vem apoiando o desenho e a implementação de duas avaliações de impacto junto ao Serviço Brasileiro de Apoio às Micro e Pequenas Empresas, uma no Paraná e outra no Rio de Janeiro. Nesta seção, relataremos a experiência como SEBRAE-Rio e, em específico, o processo de construção da avaliação do projeto do SEBRAE Na Sua Empresa.

O projeto do SEBRAE Na Sua Empresa é um projeto porta-a-porta, que visita 60 mil empresas já constituídas, com a finalidade de fazer diagnóstico de gestão das empresas e, a partir disso, repassar orientações de fácil aplicação para melhoria dessa gestão. O projeto sempre foi monitorado por meio de pesquisa quantitativa e qualitativa, por meio de *focus group*, e essa experiência já demonstrava uma predisposição para esse tipo de avaliação.

A avaliação de impacto é vista como um aprendizado para o projeto, bem como para a instituição. Isso porque, quando se avalia, a instituição tem que estar “de coração aberto”. Isto é, precisa ter desejo de aprendizado, de mudança e estar pronta para receber críticas construtivas, positivas ou de melhoria constante. Essa visão tem que existir para se conseguir absorver os ganhos da avaliação, caso contrário, talvez não seja o momento para a avaliação. Essa percepção já é um grande aprendizado.

A experiência da avaliação de impacto também é importante pelo desenvolvimento em conjunto no pré-campo. O Banco Mundial conhecia pouco do projeto “SEBRAE Na Sua Empresa”, ao passo que o SEBRAE-Rio conhecia pouco do Banco Mundial. Em face disso, houve a necessidade de alinhar objetivos, expor as limitações (inclusive orçamentárias) e as dificuldades, especialmente as relacionadas ao campo – ou seja, à visitação de 60 mil empresas. Isso não é fácil, pois exige uma equipe muito grande na rua, com mais de 120 agentes trabalhando. Essa imersão no projeto e na realidade do SEBRAE-Rio permitiu um alinhamento entre as equipes.

A saída a campo também trouxe bastante aprendizado. Dificuldades serviram para a melhoria de listagem, por exemplo. Logo, poder testar as iniciativas antes de ampliar para 60 mil casos mostrou-se fundamental. Pilotar em pequenas escalas e experimentar antes de se investir um recurso maior e apostar em uma meta maior. Essa visão foi disseminada com sucesso também entre os diretores, que aderiram à avaliação.

Realizado o levantamento, aguardam-se atualmente a análise e os resultados, que serão entregues pelo Banco Mundial, para então se chegar a proposições de melhorias para o

<sup>6</sup> Ao escalar via eletrônica, provavelmente o efeito não vai ser o mesmo daquele encontrado no piloto, ainda que tenha sido um piloto com uma amostra ampla. Na intervenção, tem-se a certeza de que a informação é entregue. Quando você coloca o material disponível *online*, por sua vez, o sujeito é mais passivo. Requer-se dele que esteja online, acesse e leia o conteúdo. A economia comportamental, por exemplo, aponta que as pessoas procrastinam – o que é um problema. Com a gestão é a mesma coisa. Diante do consultor, a empresa adota as práticas de gestão. Contudo, quando se pede para a empresa: “faça você mesmo”, é muito mais difícil.

projeto. Algumas melhorias necessárias já são perceptíveis, como no pós-visita, ações realizadas depois da visita presencial. O material entregue e o canal de e-mail e SMS não se mostraram os mais adequados. Tem-se estudado, por exemplo, a possibilidade de contato via WhatsApp.

A experiência da avaliação também contribuirá para o Sistema SEBRAE, pois ajudará não só a entender a importância de avaliação de impacto, por meio da identificação dos resultados, mas também para entender mais detalhadamente as características do nosso público. Por exemplo, se estamos lidando com empreendedores de fato ou com “donos de negócio”, pessoas que largariam o negócio diante de uma proposta de trabalho com carteira assinada. Essa compreensão poderá mostrar para o SEBRAE como direcionar seus recursos de uma forma mais adequada, seja para um profissional que demonstra características empreendedoras, seja para um “dono de negócio”. Dessa forma, continuaríamos atendendo a todos, mas utilizando estratégias diferentes e que gerem resultados para cada um dos públicos.

## **PROGRAMA DE CIDADANIA FINANCEIRA: A EXPERIÊNCIA DO BANCO CENTRAL DO BRASIL**

Assim como a AEF-Brasil, o Banco Central do Brasil (BC) tem um programa de cidadania financeira, com várias intervenções dentro dessa agenda já em andamento e reunidas sob o Departamento de Promoção da Cidadania Financeira (DEPEF). O DEPEF antigamente era chamado Departamento de Educação Financeira. Atualmente, emprega-se o termo cidadania financeira para englobar não só educação financeira, como inclusão financeira e proteção ao consumidor de serviços financeiros.

O Banco Central vem trabalhando para desenvolver uma agenda de avaliação de impacto de programas de cidadania financeira, com o objetivo de verificar se as ações promovidas estão na direção correta ou se é necessário redesená-las. Ou seja, verificar se os recursos estão sendo investidos corretamente e se o benefício para a população está sendo gerado. Um segundo motivo diz respeito ao regimento interno do Banco Central, segundo o qual cabe ao DEPEF avaliar o efeito das normas e das ações, sob a perspectiva do cidadão e das Micro e Pequenas Empresas.

O processo de desenvolvimento dessa agenda de avaliação de impacto iniciou no Banco com a consolidação e identificação de dados, que compõe uma base muito grande. Há dados de várias bases internas do BC com informações de crédito, de posse de contas de cartões, de pontos de atendimento, entre outros. Contudo, até então, esses dados eram olhados de forma macro. Sentia-se a necessidade de um olhar micro, do cidadão ou da pequena empresa. Em virtude disso, iniciou-se um trabalho de identificação de dados e de construção de indicadores, a partir do cruzamento com dados externos, tais como os da Receita (dados de cadastro), da Secretaria Especial da Micro e Pequena Empresa (SEMPE) e do Instituto Brasileiro de Geografia e Estatística (IBGE). Esses dados e indicadores foram reunidos em uma base única, a partir da qual se passou a monitorá-los. Cerca de 2 mil indicadores já

foram divulgados na página do Banco Central na internet<sup>7</sup> e no Portal de Dados Abertos do Governo Federal. São indicadores de pessoa física, pessoa jurídica, Microempreendedor Individual (MEI), Micro e Pequenas Empresas, dados referentes a crédito, a relacionamento com sistema financeiro, a pontos de atendimento, cooperativas de crédito, consórcio, entre outros.

Tem sido um processo árduo de construção. Porém esses dados dizem mais respeito à inclusão financeira, que trata do acesso e do uso. Para avaliar completamente e ter uma visão maior da educação financeira, é necessário complementar com pesquisas. Já foram realizadas, por exemplo, Uso e Qualidade dos Serviços Financeiros (2014) e aplicação do Toolkit da OCDE no Brasil, que traz dados interessantes sobre educação financeira e uma seção de inclusão financeira (2015). Essas e outras pesquisas e estudos podem ser encontrados na Série Cidadania Financeira do BC e fazem parte igualmente desse processo de construção da agenda de avaliação de impacto<sup>8</sup>.

Nesse sentido, cabe citar avaliação de impacto realizada em 2015, em parceria com a Fundação Capital e com a Universidade Federal da Bahia. A pesquisa avaliou um projeto de educação financeira para população em situação de pobreza e extrema pobreza. Para tanto, foram utilizados *tablets*, fornecidos pela Fundação Capital. Os equipamentos circularam em dez comunidades carentes piloto. No desenho inicial, os *tablets* teriam que circular ao mesmo tempo no grupo de controle e no grupo de tratamento. Contudo, durante o piloto, muita coisa aconteceu fora do previsto, nos levando a pensar que o Projeto estava sendo uma perda de tempo. Mas, pelo contrário, lições estavam sendo aprendidas. Entendeu-se que a metodologia desenhada era muito complexa para lidar com uma população que precisava de uma flexibilidade maior.

Vale citar ainda a avaliação do projeto voltado a atendentes dos PROCONs. O projeto visa fornecer informação sobre o funcionamento do Banco Central e sobre que tipo de informação deve ser encaminhada para o Banco, e como tratar o encaminhamento de outras reclamações. Adicionalmente, foram desenhadas e disponibilizadas aos atendentes dos PROCONs diversas fichas com informações que auxiliem o melhor encaminhamento das reclamações junto aos cidadãos. O projeto foi realizado em parceria com a MGov Brasil. Como resultados positivos, identificaram-se as principais necessidades que os atendentes do PROCONs tinham, bem como algumas questões de cidadãos que eram encaminhadas ao Banco Central pelo PROCON, mas que não diziam respeito ao Banco. Identificaram-se, ainda, outras questões a que os atendentes não sabiam responder e/ou dar procedimento.

Nesse sentido, o processo de avaliação é bem interessante pela aprendizagem e pela necessidade de aprender a lidar com as críticas.

---

<sup>7</sup> Disponível: <<http://www.bcb.gov.br/pt-br#!/home>> . Acesso em 23 abr. 2018.

<sup>8</sup> Disponível em: <<http://www.bcb.gov.br/pt-br#!/n/SERIECIDFINAN>>. Acesso em 23 abr. 2018.

## CONSIDERAÇÕES FINAIS

Não é necessário que as avaliações de impacto estejam concluídas para que se tenham ensinamentos a compartilhar. Neste texto, buscou-se mostrar que diferentes experiências, finalizadas ou em andamento, já dão conta de identificar ganhos em termos de aprendizado quanto aos processos e procedimentos levados a cabo antes, durante ou depois da implementação de ações, projetos e intervenções as mais variadas.

## REFERÊNCIAS

BRUHN, M.; LEÃO, L. S.; MARCHETII, R.; ZIA, B. The Impact of High School Financial Education: Experimental Evidence from Brazil. *Policy Research Working Paper* 6723. Washington D.C.: The World Bank, 2013. Disponível em:  
<<https://openknowledge.worldbank.org/bitstream/handle/10986/16940/WPS6723.pdf?sequence=1&isAllowed=y>>.

Número 10

Estudo

- Avaliação de Impacto de Políticas Públicas: Conceitos, Metodologias e Experiências