


Artigo original

Potencial custo-efetividade: Proposta alternativa para avaliação de projetos sociais

Cost-effectiveness potential: An alternative proposal for evaluating social projects

Matheus Gomes Mendonça Ferreira¹ , Nadine Mattos Depin¹ , Ernesto Javier Fernández Tovar^{1,2} , Olga de Laforet Belohlavek¹ , Mateus Pinto da Luz Ligocki¹ , Bruna Larissa Seibel^{1,2} , Thais Lopes Nicolau^{1*} 

¹Incentiv.me, Florianópolis, SC, Brasil

²Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPQ), Brasília, DF, Brasil

COMO CITAR: Ferreira, Matheus Gomes Mendonça, Depin, Nadine Mattos, Fernández Tovar, Ernesto Javier, Belohlavek, Olga de Laforet, Ligocki, Mateus Pinto da Luz, Seibel, Bruna Larissa, & Nicolau, Thais Lopes (2024). Potencial custo-efetividade: Proposta alternativa para avaliação de projetos sociais. *Revista Brasileira de Avaliação*, 13(1), e130524. <https://doi.org/10.4322/rbaval202412005>

Resumo

A avaliação de projetos sociais é um tema pertinente à responsabilidade social, parte da agenda ESG. Mensurações efetivas das práticas sociais são também relevantes para o combate ao *socialwashing*. A avaliação de impacto é o método mais recomendado, pois indica causalidade entre ações e seus efeitos. Porém, sua usabilidade é limitada, especialmente no setor privado. Métodos simplificados, como a avaliação de resultados, podem ser imprecisos, ampliando riscos de vieses ou investimentos financeiros altos em ações pouco eficazes. Este trabalho propõe um método alternativo de análise dos projetos sociais, utilizando métricas comuns às avaliações de resultados, combinadas a informações sobre o contexto beneficiado. Para isso, uma empresa do setor privado analisou seu banco de dados de projetos sociais patrocinados por leis de incentivo fiscal, desenvolvendo o potencial custo-efetividade, uma forma de estimar o efeito ótimo para o investimento realizado, e ranquear os projetos pelo potencial efeito.

Palavras-chave: Avaliação de impacto. Avaliação de resultados. Projetos sociais. Responsabilidade social. Custo-efetividade.

Abstract

The evaluation of social projects is a relevant issue for social responsibility as part of the ESG agenda. Effective measurement of social practices is also relevant to combating socialwashing. Impact assessment is the most recommended method, as it shows the causality and effects of actions. However, its applicability is limited, especially in the private sector. Simplified methods, such as outcome evaluation, can be inaccurate, increasing the risk of bias or high financial investment in ineffective interventions. This paper proposes an alternative method for analyzing social projects, using metrics common to outcome evaluations and information about the context in which they are benefiting. To this end, a private sector company analyzed its database of social projects supported by tax incentive legislation, developed potential cost-effectiveness, a method for estimating the optimal impact for the investment made, and ranked projects according to their potential impact.

Keywords: Impact evaluation. Outcome evaluation. Social projects. Social responsibility. Cost-effectiveness.

Matheus Gomes Mendonça Ferreira, branco, analista de dados e métricas, Incentiv via edital Fapesc.

Nadine Mattos Depin, branca, coordenadora de produto, Incentiv.

Ernesto Javier Fernández Tovar, negro, pesquisador bolsista CNPq, Incentiv.

Olga de Laforet Belohlavek, branca, assistente administrativa, Incentiv via edital Fapesc.

Mateus Pinto da Luz Ligocki Vieira, branco, analista de dados, Incentiv.

Bruna Larissa Seibel, branca, pesquisadora bolsista CNPq, Incentiv.

Thais Lopes Nicolau, branca, COO, Incentiv.

A RBAVAL apoia os esforços relativos à visibilidade dos autores negros na produção científica. Assim, nossas publicações solicitam a autodeclaração de cor/etnia dos autores dos textos para tornar visível tal informação nos artigos.

Recebido: Junho 27, 2023

Aceito: Março 11, 2024

***Autor correspondente:**

Thais Lopes Nicolau

E-mail: thais@incentiv.me



Este é um artigo publicado em acesso aberto (Open Access) sob a licença Creative Commons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.



Introdução

O interesse sobre avaliação de impacto tem extrapolado a esfera acadêmica. Empresas e outras organizações (Institutos, Fundações, etc.) têm buscado cada vez mais implementar essas análises com diversas finalidades, entre elas, mensurar o impacto de projetos sociais que fazem parte do seu portfólio de responsabilidade social. Tal movimento ocorre junto à crescente preocupação com a agenda de responsabilidade socioambiental, orientadas pela abordagem do *environmental, social, and corporate governance* (ESG) (Li et al., 2021).

Diante da relevância dessa agenda, há a necessidade de avaliar se as ações sociais e/ou ambientais dessas empresas impactam de fato a sociedade ou se enquadram como medidas chamadas de *greenwashing*, *socialwashing* ou *pinkwashing* (Caporale et al., 2022)¹. Nesses casos, as empresas assumem que há impactos ou exageram na magnitude das suas ações. Portanto, entre outras funções, as pesquisas de avaliação de impacto servem para avaliar possíveis desvios de finalidade da agenda ESG.

Apesar disso, mesmo com intenções orientadas para essa agenda, a medição envolvendo ações de sustentabilidade tem limitações (Li et al., 2021). Muitas empresas e organizações não conseguem mensurar o impacto de projetos do seu portfólio de responsabilidade social devido aos custos operacionais, à falta de conhecimento científico sobre o tema e à profissionalização do setor. Ainda há dificuldade, por parte de algumas empresas, em adotarem essa estratégia analítica como a mais adequada. Tradicionalmente, pesquisas são aplicadas para avaliar os efeitos de programas sociais e políticas públicas de larga escala, promovidas pelo Estado. Como os projetos das empresas ocorrem em menor escala e com menor disponibilidade de recursos, as avaliações de impacto nem sempre são consideradas, ou mesmo realizadas tardiamente.

Diante dessas barreiras, uma estratégia comum utilizada pelas partes interessadas é a mensuração de resultados. Em uma avaliação de resultados, é comum analisar métricas como número de beneficiários, ações criadas, custos das ações, tempo do projeto, pessoas empregadas, entre outras. Dessa forma, o escopo da avaliação deixa de ser o impacto do projeto na vida do beneficiário, estabelecendo um comparativo pré e pós ação, e passa a ser uma análise sobre os *outputs*, ou os resultados diretos.

Com este artigo, pretende-se contribuir para o campo de avaliação de projetos sociais, contemplando a dimensão de responsabilidade social do ESG, especificamente de empresas como *stakeholders* centrais. Optou-se por esse escopo devido às barreiras apontadas. Poucos são os trabalhos dedicados a sistematizar análises de projetos sociais promovidos ou apoiados por empresas no Brasil. Embora não existam dados consolidados sobre a participação do setor empresarial nesses projetos, a ausência de bancos de dados disponíveis e a predominância do Estado como agente promotor de iniciativas sociais sugerem que a participação do setor privado ainda é limitada.

Dessa forma, o artigo apresenta um método alternativo de monitoramento e avaliação de projetos sociais, indo além da avaliação de resultado, contornando as dificuldades de análise da avaliação de impacto. Trata-se, portanto, de uma abordagem intermediária. A hipótese sustentada é de que as avaliações de projetos sociais alcançam outro patamar quando é dado o tratamento devido às métricas tradicionais e quando são mobilizadas outras fontes de informações, relacionados à qualidade de vida das pessoas beneficiárias.

O artigo está estruturado de forma que na primeira seção é apresentado o debate sobre avaliação de projetos sociais e métricas utilizadas para mensurar resultados e impactos de projetos sociais. Na segunda seção, são destacados a metodologia, os dados utilizados e os resultados. Por fim, apontam-se algumas limitações do método de estudo.

Avaliando projetos sociais

Atinente às avaliações de projetos, existem aquelas que se ocupam do momento anterior à implementação (*ex ante*) e aquelas que ocorrem durante e após a implementação destes (*ex*

¹ Segundo os autores, *greenwashing* ocorre quando as empresas exageram suas credenciais ambientais; *socialwashing* ocorre quando as empresas superestimam o impacto na área de direitos humanos e; *pinkwashing* ocorre quando as empresas criam a falsa impressão de serem inclusivas das comunidades LGBTQIA+.



post). Neste trabalho, o escopo é sobre o segundo tipo. Esta escolha se justifica pelo fato de que, geralmente, os projetos sociais apoiados pelas empresas já estão em andamento, ou seja, vêm sendo conduzidos por parceiros. Dessa forma, as empresas não costumam participar da construção dos respectivos modelos lógicos.

Sobre os tipos de avaliação *ex post*, destacam-se as avaliações de impacto e de resultados. A primeira é considerada “padrão ouro” da ciência, uma vez que, por meio dela, é possível inferir causalidade, ou seja, se um projeto causou o que pretendia e qual a magnitude desse efeito. Com relação à segunda, trata-se de uma análise exploratória de indicadores associados aos resultados esperados, que também são utilizados nas avaliações de impacto.

Nos subtópicos a seguir, são apresentadas as principais características dos dois modelos e quais as variáveis-chave que os analistas devem mensurar. Por fim, avaliam-se os *trade-offs* dessas análises, considerando suas virtudes e custos.

Avaliação de impacto

A avaliação de impacto é uma das mais importantes no campo de políticas públicas e avaliação de projetos sociais. Por meio dela, é possível atribuir (ou não) às ações das empresas (ou dos projetos apoiados por elas) as mudanças ocorridas em uma população assistida. Portanto, é a ferramenta desejável para que empresas decidam se continuam investindo em um projeto (caso haja impacto), interrompam-no ou realizem modificações para alcançar um impacto que antes não existia.

Na agenda ESG, há uma preocupação para que ações de empresas não sejam classificadas como *socialwashing*, *pinkwashing* ou *greenwashing*. Estes termos indicam alguma superestimativa de avaliação quanto aos efeitos das ações patrocinadas, geralmente relacionados à imagem das marcas e consequente valorização do produto ou serviço pelo público. Porém, é importante destacar que a avaliação de impacto não é necessariamente o único indicativo de que uma empresa não esteja praticando esses fenômenos. Pode haver empresas que mensurem impacto e pratiquem *socialwashing*, e empresas que não avaliem impacto e não o pratiquem. Por isso, para saber se as empresas realmente estão impactando positivamente a sociedade, provocando a inclusão de minorias, trabalhando em prol dos direitos humanos e do meio ambiente, uma avaliação de impacto é recomendada, mas não é o único fator a ser considerado.

Apesar de sua relevância, avaliações de impacto nem sempre são viáveis, devido aos custos financeiros, operacionais e aos esforços intelectuais envolvidos. Um dos principais desafios para medir o impacto social é definir as estratégias de identificação para inferência causal. Existem várias abordagens que orientam sobre os tipos de desenhos e suas aplicações (Angrist & Pischke, 2014; Cunningham, 2021). O ponto de partida são os experimentos, conhecidos como *Randomized Controlled Trial* (RCT). Porém, além dos RCTs, existem outras abordagens alternativas para a avaliação de impacto, mais simples ou intermediárias. Um guia de avaliação de impacto do INSPER destaca várias dessas estratégias, que vão desde a avaliação de resultados até a avaliação de impacto de “padrão ouro” (Insper Metricis, 2022). Portanto, a impossibilidade de se obter um grupo de tratamento e um de controle não impede o avanço da pesquisa sobre desenhos *quasi-experimentais*. Nesses desenhos, a estratégia de identificação requer mais pressupostos do que nas RCTs, tornando sua aplicação ainda mais exigente.

Importante destacar que existem diferentes maneiras de avaliar o impacto social, e esta diversidade está ligada ao fato de que os empreendimentos variam em tamanho, capacidade, atividades, foco e objetivos. Assim, não existe um modelo único que se adeque a todas as organizações e, portanto, as iniciativas têm o desafio de encontrar uma metodologia apropriada para a sua realidade e para seus diferentes interesses de impacto (Grieco et al., 2015).

A título de exemplo, considere a diferença entre uma avaliação de impacto de projetos culturais, que pretendem levar atrações artísticas para locais remotos, e projetos da área de saneamento e saúde pública, que almejam reduzir os riscos à saúde de uma população. Enquanto o objetivo dos projetos de saúde é melhorar a qualidade de vida das pessoas por meio da redução de doenças, podendo ter seu impacto mensurado por meio de indicadores



como longevidade, taxa de contaminação, entre outras, os projetos da área de cultura esbarram nos desafios de mensurar resultados que podem ser subjetivos, como satisfação e felicidade das pessoas.

Apesar dessas diferenças, tratar de impacto social é lidar com uma variável de performance relacionada às ações de uma entidade, como um projeto social (Rawhouser et al., 2019). Nesse sentido, Sandri et al. (2020) ressaltam que qualquer atividade realizada por uma entidade gera algum tipo de efeito na população em que está inserida, alterando aspectos situacionais de determinado grupo socioeconômico. Importante ressaltar que, para se conseguir medir o impacto social gerado por uma organização, é antes necessário ter definido este conceito. Esta pode ser uma tarefa árdua, pois, conforme observado por Rawhouser et al. (2019), em um levantamento teórico sobre medição de impacto social, dependendo do contexto em que este tema é abordado, os pesquisadores tendem a utilizar diferentes nomenclaturas. Especificamente sobre a avaliação do impacto social, Rodrigues et al. (2021) ressaltam que, considerando o objetivo de identificar se as ações realizadas resultaram nas alterações desejadas para o contexto social, ela se diferencia de outras análises em razão de ter como característica estabelecer uma relação de causalidade entre as ações propostas e os efeitos gerados.

A fim de facilitar o estabelecimento dessa relação causal, mas expandindo para a mensuração e monitoramento, Sugahara & Rodrigues (2019) apontam a utilização da teoria da mudança como ferramenta para entender o contexto em que determinada atividade está inserida e quais as alterações que são pretendidas pelas ações da entidade. A literatura sobre efeitos causais é rica com relação às possibilidades de desenhos de pesquisas que encontrem as condições necessárias e suficientes para inferir causalidade. Nessa esteira, algumas limitações, como a capacidade técnica das organizações, são apontadas como desafios para uma avaliação de impacto adequada. Se as definições conceituais de impacto social já são extensas, as formas de avaliação dos resultados de projetos sociais também são.

Rodrigues et al. (2021) argumentam que, apesar de serem consideradas relevantes, as ferramentas para avaliação de projetos sociais no Brasil são ainda bastante desconhecidas e inutilizadas. Nesse sentido, Bengo et al. (2016) comentam sobre a limitação da capacidade das organizações sociais, que sofrem por não possuírem conhecimentos mais amplos de métricas e indicadores de avaliação de impacto.

Tão importante quanto pensar os melhores desenhos de pesquisa para inferência causal é pensar sobre as métricas utilizadas. Um desenho de pesquisa que consiga garantir estratégia de identificação causal não resolve um problema de natureza primária: a qualidade das métricas. Métricas ruins podem fazer com que bons desenhos de pesquisa cheguem a resultados equivocados.

Imagine que um grupo de pesquisadores esteja avaliando o impacto que a internet (de telefonia móvel) tem sobre a economia de algumas localidades. Como métrica para a variável de tratamento foi utilizada a presença da torre de transmissão, que permite o uso dessa tecnologia ($D=1$). Porém, essa métrica nada informa se essas torres estão funcionando ou se o sinal de telefonia móvel está chegando com qualidade para os usuários. Pode ser que algumas localidades recebam as torres, mas estas não estão funcionando (o que deveria ser considerado como $D=0$, mas está como $D=1$). Portanto, ao considerar a “existência das torres” na localidade como uma métrica para avaliar o impacto da internet móvel, os pesquisadores estariam violando alguns supostos necessários para garantir uma estratégia de identificação que permite inferir causalidade. Uma solução para esse problema seria substituir a métrica “ter torre de internet” pela métrica “volume de dados” (3G por exemplo).

Tal preocupação com as métricas se faz presente em alguns trabalhos, que investigam os impactos gerados por organizações sociais, em especial no terceiro setor (Kardec et al., 2002; Melo, 2008; Milani Filho, 2009). Para as autoras, as métricas (pensadas também por meio de indicadores) são guias que permitem medir não apenas a eficácia das ações tomadas, mas os vieses entre o programado e o realizado. Em suma, sem indicadores não seria possível avaliar o desempenho das organizações e se suas ações realmente impactaram a sociedade e de que forma impactaram.



Mesmo quando essas barreiras são superadas, existem desafios inerentes à temporalidade do efeito causal dos projetos. Nem sempre uma intervenção social gera mudanças logo após a sua implementação. Estas podem ocorrer ao longo da vida das pessoas. O exemplo acima, do efeito da telefonia móvel sobre o desenvolvimento econômico de uma localidade, é típico de se esperar um efeito tardio. Após a instalação das antenas, os usuários precisam “aprender” a manusear a nova tecnologia. Apenas com o tempo é que eles conseguirão transformá-la em valor monetário. Sendo assim, como garantir que, no decorrer da vida das pessoas, aquela ação que ocorreu há muito tempo pode ser considerada como a causa do resultado obtido? Levando em conta esse aspecto, consideramos que nem sempre uma avaliação de impacto é a melhor alternativa. Diante dessa inviabilidade, avaliações de resultados são alternativas importantes, mesmo que não substituam as de impacto.

Além disso, Hadad & Gauca (2014) consideram que a dificuldade em calcular o impacto social está relacionada aos conceitos de problemas e inovação sociais, uma vez que para fazer esta mensuração é necessário identificar os desafios que determinado grupo enfrenta e, por meio da inovação social, buscar formas de alterar o contexto e gerar resultados. As autoras ponderam sobre a dificuldade de determinar quais ações geram um maior impacto e comparar o resultado obtido por diferentes atividades e, buscando uma forma de medir o impacto social que se adeque aos diferentes objetivos propostos por diferentes entidades, concluíram que tal hipótese não é possível.

Apesar dessa impossibilidade apontada pelas autoras, existe um indicador que auxilia na comparação entre diferentes projetos sociais: avaliação de custo-efetividade (ACE). Matematicamente ele pode ser representado pela fórmula:

$$\omega = \frac{\tau - \tau_0}{|\rho - \rho_0|} \quad (1)$$

onde:

1. τ representa o impacto produzido pelo projeto;
2. τ_0 o impacto esperado;

$$3. \rho = \frac{\text{R\$captado}}{\text{Beneficiários}} \cdot \text{Tempo};$$

$$4. \rho_0 = \frac{\text{R\$esperado}}{\text{Beneficiários esperados}} \cdot \text{Tempo estimado};$$

5. $\tau - \tau_0$ é a efetividade da política;
6. $\rho - \rho_0$ é o resultado financeiro.

Projetos que apresentam um custo-efetividade (ω) ≥ 1 são considerados como projetos melhores que aqueles que apresentam o valor menor que 1. Isso porque, para cada unidade investida, espera-se que haja um (1) ponto de resultado efetivo. A vantagem de se utilizar o custo-efetividade (CE) é que ele permite comparar projetos que tenham objetivos distintos, uma vez que padroniza as unidades de análise. O desafio é que este indicador, além de exigir uma análise de impacto prévia, exige que os projetos tenham desenvolvido uma teoria da mudança capaz de estimar o impacto, os recursos, beneficiários e tempo estimados. Estimativas muito ruins dessas métricas podem levar a conclusões equivocadas. No próximo subtópico é explorado o que são avaliações de resultados e como elas se aproximam de atividades de monitoramento de projetos. A proposta, como será apresentada na próxima seção, apresenta uma estratégia que combina análise de custo-efetividade com dados



contextuais, que informam sobre a qualidade de vida dos beneficiários dos projetos. Com isso, pretende-se superar essa impossibilidade apontada por Hadad & Gauca (2014).

Avaliação de resultados

Diferente das avaliações de impacto, a de resultado não almeja inferir causalidade. Seu objetivo é uma análise exploratória dos indicadores de resultados de um projeto social. Isso não significa que uma avaliação de resultado é necessariamente anterior a de impacto. É possível realizar avaliação de impacto sem realizar avaliação de resultado. Por outro lado, análises de impacto tendem a ser mais robustas quando temos em mãos uma boa avaliação de resultados.

Tal modelo trata das mudanças de curto prazo, mas não apenas nos projetos em si, levando em conta também as mudanças que ocorrem no nível dos beneficiários. Por também envolver variáveis ligadas estritamente aos processos dos projetos, ela se aproxima de atividades de monitoramento.

A diferença consiste em que o monitoramento se restringe aos indicadores diretamente relacionados à execução da política pública, enquanto a avaliação de resultados é mais ampla, permitindo analisar a população beneficiária sob as mais diferentes dimensões. (Brasil, 2018, p. 241).

Em uma avaliação de resultado, pesquisadores podem discutir o cumprimento das metas e mapear os processos e insumos que favoreceram ou dificultaram o projeto. Outra atividade chave é o levantamento do nível de satisfação dos usuários da política, que pode ser direcionada para aprimorar os processos e resultados de um projeto.

Um exemplo de avaliação de resultado que ajudou a compreender melhor o impacto do Projeto Bolsa Família (PBF) no Brasil foi um estudo etnográfico coordenado pela Secretaria de Avaliação e Gestão da Informação do Ministério do Desenvolvimento Social (SAGI/MDS) com famílias indígenas (ver Brasil, 2018). Enquanto uma avaliação de impacto poderia informar que o PBF causou uma melhora na vida dessas famílias, a avaliação de resultado informou que:

Diversos indígenas se expõem para conseguir participar do programa: dificuldades para sacar os benefícios devido às longas distâncias das comunidades em relação aos centros urbanos; inadequação das condicionalidades previstas para a realidade dos grupos pesquisados; problemas com documentação civil e preconceitos sofridos por esses povos; comerciantes que prejudicam as famílias indígenas com retenção de cartões e endividamentos, entre outros. (Brasil, 2018, p. 246).

Esse tipo de pesquisa revela que a presença de resultados positivos em avaliação de impacto não implica em resultados positivos também na avaliação de resultados. Por isso a importância de ambas serem conduzidas para uma melhor compreensão dos projetos e programas sociais é enfatizada pela literatura.

Apesar dessa importância, como informado anteriormente, avaliações de resultados são estudos exploratórios, que não permitem inferir causalidade. Essa limitação impede que *stakeholders* avaliem se um projeto impacta positiva ou negativamente a população assistida, ou se tal resultado se deu por outras variáveis associadas no processo.

Considerando que empresas com portfólio de responsabilidade social utilizam recursos próprios ou via leis de incentivo fiscais para apoiar projetos, esse tipo de informação sobre o real impacto dos projetos não pode ser alcançado por meio de uma avaliação de resultados. No tópico seguinte, apresentamos um modelo alternativo, que se coloca entre as avaliações de resultado e impacto, e visa contribuir para essa agenda de pesquisa e avaliação.

Método: Potencial Custo-Efetividade (PCE)

Tendo em vista as barreiras de uma avaliação de impacto e a limitação de uma avaliação de resultados, este estudo propõe uma terceira abordagem: a análise do potencial custo-efetividade (PCE) de um projeto. Potencial porque utilizamos apenas o denominador do custo-efetividade,



uma vez que dificilmente temos informações sobre o efeito de um programa (τ). Por outro lado, é mais comum o acesso às métricas de monitoramento e de avaliação de resultado, como valor captado, número de beneficiários e tempo de duração do projeto (métrica-chave para calcular o denominador do custo-efetividade). Sabendo que um projeto com custo-efetividade (CE) ótimo precisa apresentar $\omega \geq 1$, o valor do numerador ($\tau - \tau_0$) precisa ser, no mínimo, igual ao resultado do denominador ($\rho - \rho_0$). Como o resultado do denominador é acessível, saberemos o quanto de impacto o projeto precisa ter para que alcance um nível satisfatório de CE.

Caso o analista ou o responsável pelo projeto não possuam nenhuma estimativa confiável dos valores para τ_0 e ρ_0 , basta igualá-los a zero (0), resultando em $\omega = \frac{\tau - 0}{\rho - 0}$. Importante destacar que, para inferir valores esperados (τ_0 e ρ_0), é preciso um esforço de revisão da literatura sobre a área do projeto e seus impactos catalogados e de um modelo lógico robusto. Caso contrário, estimativas supervalorizadas de τ_0 e ρ_0 podem levar a resultados de custo-efetividade negativos (quando se estima $\tau_0 > \tau$) ou supervalorizados (quando se estima concomitantemente $\tau_0 < \tau$ e $\rho_0 < \rho$).

Encontrado o denominador do custo-efetividade (DCE), sabemos o quanto um projeto precisa impactar para obter um resultado custo-efetividade ótimo. Projetos com valor do DCE mais baixos tendem a entregar um custo-efetividade (CE) maior. Porém, apenas essa informação não é suficiente. Um projeto que tenha um DCE maior do que outro ($DCE_{projeto X} > DCE_{projeto Z}$), ainda sim pode apresentar um CE melhor, dado que o impacto daquele projeto pode ser bem maior do que do outro projeto ($\tau_{projeto Z} > \tau_{projeto X}$). Para tornar esse cálculo mais robusto, sugerimos ponderar os valores de DCE pelo Índice de Vulnerabilidade Social (IVS) do Instituto de Pesquisa Econômica Aplicada (IPEA), criando assim o Potencial Custo-Efetividade de um projeto.

O IVS é um índice que mede a vulnerabilidade social no Brasil. Seus dados podem ser desagregados até o nível dos municípios e por variáveis como sexo, cor e condição do domicílio. O índice é construído com base em vários indicadores do Atlas do Desenvolvimento Humano (ADH), retratando três dimensões da vida: i) Infraestrutura Urbana; ii) Capital Humano e; iii) Renda e Trabalho. Cada uma delas representa um subíndice do IVS. Cada subíndice é construído com base na média ponderada dos indicadores. Ao todo, são levados em conta dezesseis indicadores. Após o cálculo dos subíndices, o IVS é calculado com base na média aritmética dos subíndices. Ao final, gera-se valores que variam entre 0 e 1, onde 1 é o grau máximo de vulnerabilidade social. Esses valores ainda são categorizados, pelo IPEA (Costa & Marguti, 2015), na escala apresentada na Tabela 1.

Tabela 1. Categorias de vulnerabilidade do IVS.

IVS	Intervalo
Muito baixa	0-0,200
Baixa	0,201-0,300
Média	0,301-0,400
Alta	0,401-0,500
Muito alta	0,501-1

Fonte: Costa & Marguti (2015).

O IVS será utilizado como um ponderador do DCE, gerando assim o Potencial Custo-Efetividade (PCE). Quanto menor o PCE, maior a chance de o custo efetividade ser maior ou igual a 1. A tese sustentada neste trabalho é que, projetos sociais que ocorrem em regiões com maior vulnerabilidade social, em comparação com aqueles que ocorrem em regiões com IVS menor, tem mais chance de apresentarem projetos com um impacto (τ) maior. A Tabela 2 auxilia na explicação desse argumento ao simular dois projetos.

**Tabela 2.** Simulação do Potencial Custo-Efetividade (PCE).

Variáveis	Projeto 1	Projeto 2
DCE	2	4
IVS	0,233	0,837
IVS tratado*	$1 - 0,233 = 0,767$	$1 - 0,837 = 0,163$
Categoria IVS	Baixa	Muito Alta
PCE	$2 \times 0,767 = 1,534$	$4 \times 0,163 = 0,652$
Impacto (τ)	Desconhecido	Desconhecido

*Nota: Subtraí-se 1% do valor do IVS para inverter a escala do índice. Assim, valores menores representam maior vulnerabilidade (menor qualidade de vida), ao passo que valores maiores representam menor vulnerabilidade (maior qualidade de vida). Fonte: Autoria própria (2023).

Considere, a título de exemplo, que ambos os projetos têm por objetivo melhorar as notas de matemática de alunos do ensino básico. Cada projeto é desenvolvido em regiões distintas. A região do projeto 2 é mais vulnerável do que a do projeto 1. Dessa forma, antes mesmo dos projetos acontecerem já é possível assumir que a probabilidade da média das notas de matemática dos alunos atendidos pelo projeto 2 é menor do que a do projeto 1. Estudos apontam que condições sociais impactam o rendimento acadêmico dos alunos (Fernandes et al., 2018; Américo & Lacruz, 2017; Palermo et al., 2014).

O que nós consideramos chave nesta análise é que o potencial de crescimento da nota de matemática no projeto 2 é maior do que no projeto 1. Por esse potencial de crescimento ser maior, o peso atribuído ao projeto 2 deve contribuir mais para que este alcance um custo-efetividade ótimo (igual ou maior a 1).

Olhando apenas para o DCE, teríamos a informação de que o custo-efetividade poderia ser maior para o projeto 1. Por outro lado, quando inserida a variável IVS na análise, passamos a ter outra compreensão sobre o projeto 2, apostando neste como o projeto que teria um resultado de custo-efetividade melhor. Porém, esse tipo de entendimento só é válido quando comparamos categorias de vulnerabilidade distintas (quando comparamos, por exemplo, regiões com vulnerabilidade muito alta e média).

Questionamos, portanto, como esse peso deve ser calculado quando comparamos projetos que ocorrem em regiões que estão na mesma categoria de vulnerabilidade, mas que se diferenciam em termos do valor do índice. Considere, por exemplo, dois projetos que ocorrem em uma região de IVS "Muito Alta", mas que possuem índices distintos: 0,600 e 0,800. Seguindo a lógica anterior, o projeto que ocorre na região com vulnerabilidade 0,999 teria um peso maior. Para essa situação, é preciso um outro tipo de leitura acerca do PCE.

Argumentamos que, dentro de cada categoria de Vulnerabilidade Social, o potencial impacto de um projeto é maior entre aqueles menos expostos a vulnerabilidades. Por exemplo: Entre as pessoas que residem em regiões de IVS muito alto, quanto menor for o IVS maior a chance de o projeto obter um impacto ótimo. Isso decorre do fato de que quanto menor a vulnerabilidade, mais recursos os beneficiários terão à sua disposição e que contribuirão para o sucesso dos projetos. A Tabela 3 reproduz o exemplo anterior, mostrando os resultados antes e depois da aplicação do peso.

Essa abordagem não se propõe substituir as demais. Pelo contrário, almeja contribuir com uma melhor análise dos dados quando se tem em mãos apenas uma avaliação de resultados. Para testar essa abordagem em dados não simulados, optamos por aplicá-la em um banco de dados que contém informações sobre projetos sociais que fazem parte do portfólio de responsabilidade social de algumas empresas. Com isso, será possível ver o desempenho do nosso modelo em dados reais.

**Tabela 3.** Simulação do PCE para comparação de projetos na mesma categoria do IVS.

Variáveis	Projeto 1	Projeto 2
DCE	5	4
IVS	0,600	0,800
Categoria IVS	Muito Alta	Muito Alta
PCE	$5 \times 0,600 = 3,000$	$4 \times 0,800 = 3,200$
Impacto (τ)	Desconhecido	Desconhecido

Fonte: Autoria própria (2023).

Resultados e discussão

O banco de dados que utilizamos neste trabalho pertence a uma empresa líder na gestão de recursos para projetos com impacto social. Por ser responsável em conectar projetos sociais a companhias que tenham interesse em patrociná-los por leis de incentivo fiscal, a empresa preza pela excelência quanto à mensuração dos efeitos das iniciativas. Ela possui informações sobre projetos sociais que estão cadastrados na plataforma da empresa e que já receberam aporte financeiro e estão em fase de implementação. Os projetos estão distribuídos da seguinte forma entre os estados: 18 projetos em Santa Catarina; 4 em São Paulo; 2 em Minas Gerais; 1 projeto em cada um dos estados restantes (Rio de Janeiro, Paraná, Pernambuco, Rio Grande do Sul, Rio Grande do Norte, Ceará, Bahia, Espírito Santo). Como variáveis do banco, temos: i) categorias de impacto; ii) quantidade de ações realizadas pelo projeto, iii) pessoas empregadas pelos projetos; iv) total de pessoas beneficiadas pelo projeto e; v) valor captado pelo projeto. O banco não traz informações sobre os valores esperados das métricas (ρ_0). A Tabela 4 apresenta as informações sobre esses dados.

Tabela 4. Estatísticas descritivas das variáveis tratadas

Métricas	Média	Desvio Padrão	Min.	Max.
Beneficiários	723,07	1319,55	7	5435
Ações	4880,53	17771,47	1	149
R\$ captado	1.106.503,00	1.988.490,00	4.500,00	80.000.926,00
Tempo	9,533	7,65	1	24

Fonte: Autoria própria (2023).

Com relação aos dados do Atlas da Vulnerabilidade Social, optamos por utilizar o índice mais recente, que é de 2021. Porém, os dados de 2021 só fornecem informações do IVS para o nível dos estados, mas não por município. O último ano em que se tem dados disponíveis no nível dos municípios é 2015. Essa ausência de informação ocorre por causa da ausência do censo demográfico de 2020 no Brasil. Alguns indicadores são coletados pelo censo do IBGE. Mesmo assim, o ideal seria ter informações ainda mais desagregadas, no nível das localidades onde os projetos são desenvolvidos (bairros ou regiões censitárias). Utilizar informações no nível do estado pode mascarar a vulnerabilidade real das pessoas beneficiárias dos projetos, uma vez que dentro de um estado pode haver grande variação no IVS². A Figura 1 apresenta o IVS dos estados onde os projetos analisados ocorrem.

² Um estado com IVS 0,5 pode possuir regiões com IVS variando entre 0,1 e 0,9, por exemplo. Imagine que um projeto está sendo desenvolvido em uma região cujo IVS é 0,9, mas a única informação que temos é que naquele estado, no último ano, o IVS é 0,5.

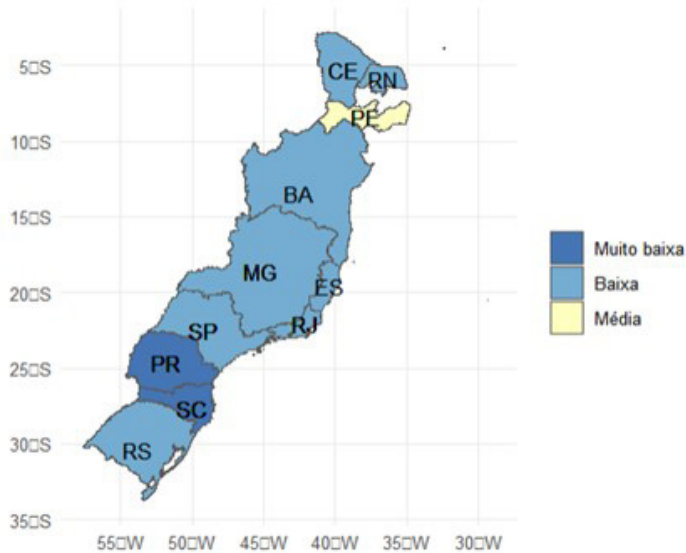


Figura 1. IVS dos estados que receberam os projetos.

Fonte: IPEA (2021).

Com relação aos resultados da aplicação, uma forma de verificar o impacto dos pesos sobre o denominador do custo-efetividade é analisar se algum projeto mudou de quartil. Para isso, subtraímos o quartil da variável DCE da variável PCE. O resultado foi que apenas 2 projetos mudaram suas posições nos intervalos interquartis. Isso mostra que apenas esses dois projetos sofreram uma alteração como aquela da Tabela 2. Esse resultado pode ser explicado pela pouca quantidade de casos no banco de dados. Conforme há um aumento no número de casos, espera-se maior impacto da atribuição dos pesos.

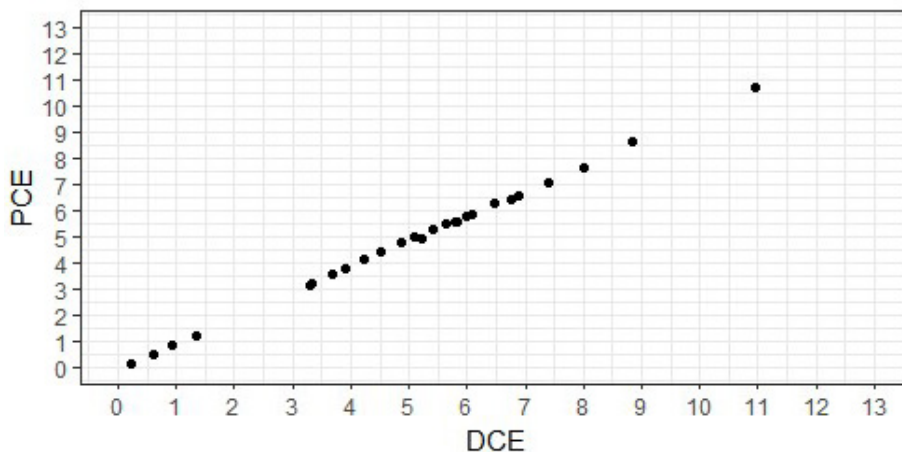


Figura 2. Gráfico de dispersão (PCE × DCE).

Analisando a mudança em cada um dos projetos, encontramos que a aplicação dos pesos apresentou alterações diversas nos valores do denominador do custo-efetividade (DCE). Em caso de utilizar dados ainda mais desagregados, no nível dos bairros ou até mesmo dos beneficiários, haverá maior variação do IVS, gerando outros resultados. Isso ocorre porque a maioria dos projetos estão localizados em Santa Catarina e acabam por receber o mesmo peso. Analisando o gráfico apresentado na Figura 2, é possível observar uma proximidade dos valores dos eixos PCE e DCE. Utilizamos o logaritmo natural dos indicadores por causa de outliers. A maior diferença entre DCE e PCE foi de 0,3975 e a menor, 0,1370. Por outro lado, quando calculamos a porcentagem da redução, encontramos que tais mudanças variaram entre 12,8% (valor mínimo) a 32,8% (valor máximo) dos valores de DCE.



Por fim, analisando apenas o PCE, encontramos que 25% dos resultados estão abaixo do valor de 39,82. Isso indica que, para 75% dos casos, o resultado da avaliação de impacto precisa produzir mudanças realmente significativas para produzir um custo-efetividade no nível ótimo.

Como não tivemos acesso a outro banco de dados desse tipo, que contenham mais casos, optamos por simular dados a partir do IVS municipal do Atlas. Para isso juntamos as informações sobre o IVS dos municípios brasileiros e simulamos dados para variáveis “recurso captado”, “beneficiários” e “tempo do projeto”. Após essa simulação, reproduzimos a análise apresentada na Figura 3.

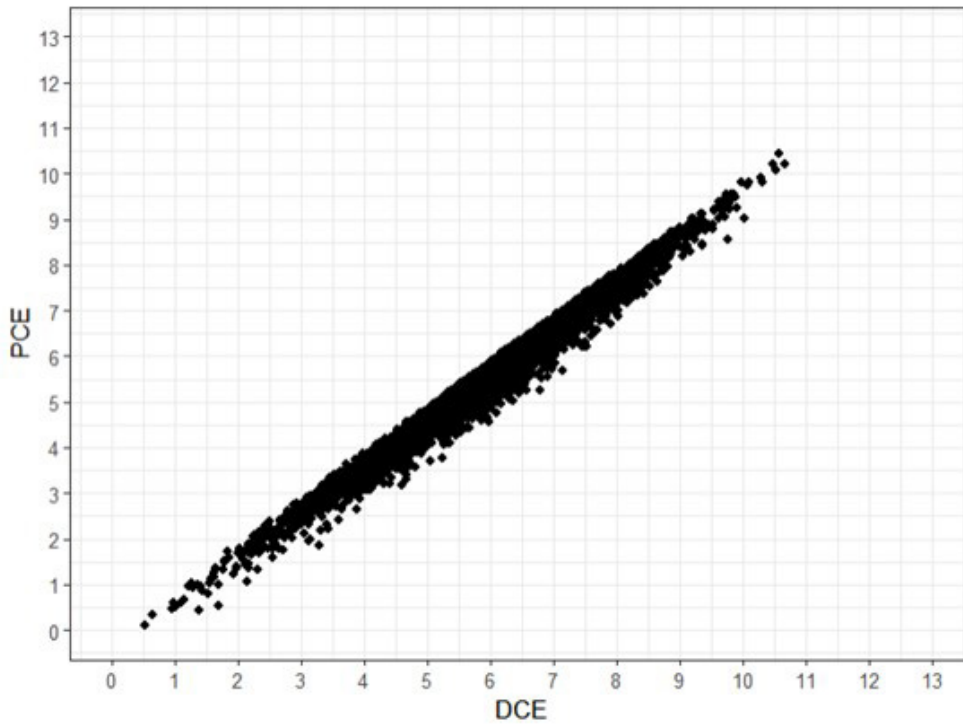


Figura 3. Gráfico de dispersão (PCE × DCE) com dados simulados.

Com os dados obtidos por meio da simulação é possível observar maior impacto da atribuição dos pesos. Esse fenômeno pode ser visto por meio do desvio da linearidade dos pontos. Ao analisar as mudanças provocadas pela atribuição dos pesos, encontramos que o valor do DCE apresentou reduções que variaram entre 9% e 78,4%. Tal resultado corrobora nossa expectativa de que o aumento do número de casos do banco original e a consideração de níveis mais desagregados mudariam a diferença entre DCE e PCE.

Considerações finais

Este trabalho buscou apresentar uma nova proposta de análise de métricas padrões das análises de resultados de projetos sociais que auxilie na melhor compreensão dos seus potenciais impactos. O objetivo era produzir uma abordagem que aproximasse as avaliações de resultados das avaliações de impacto, por meio do indicador de custo-efetividade.

O argumento chave para essa abordagem é que, na ausência de informações sobre o impacto do programa, considerando informações sobre a vulnerabilidade das populações atendidas pelos projetos, é possível ter informações mais qualificadas que auxiliem nos processos de tomada de decisão por parte dos *stakeholders*.

Apesar das limitações dos dados, o resultado mostrou que a aplicação dos pesos afetou de forma significativa os valores do DCE, alcançando uma redução de mais de um terço.



Esperamos que, com o aumento do banco de dados e com a disponibilização do IVS por municípios atualizada para 2022, essa abordagem entregue resultados ainda mais robustos em termos da avaliação de quais projetos apresentam os melhores PCE.

Atinente ao modelo, uma limitação ocorre quando comparamos projetos que pertencem a categorias de vulnerabilidade distintas, mas cujo índices são muito próximos (0,499 e 0,501, por exemplo). Nesse caso, a lógica de implementação do modelo proposta não é tão eficiente, uma vez que não é esperado que haja muita diferença no controle de recursos entre os beneficiários dos projetos nesses contextos e que a margem de impacto não se diferencia tanto. Uma alternativa seria criar um *cut-off* para aplicação do modelo de ponderação com base na diferença do índice de vulnerabilidade. Um exemplo disso seria: quando as categorias de vulnerabilidade forem distintas, a aplicação da ponderação deve ocorrer somente quando a vulnerabilidade de um grupo representar X% da vulnerabilidade do outro grupo de comparação. Estabelecer essa porcentagem do *cut-off* é um desafio que os próximos estudos podem se debruçar.

A proposta de uma medida fidedigna que considere o efeito das ações sociais relativo ao investimento aportado pode ser uma estratégia viável para a área de responsabilidade social do setor privado. Uma vez que a agenda ESG vem ganhando espaço nas discussões sobre o uso adequado de recursos e o desenvolvimento equitativo da sociedade, tem-se percebido maior interesse das empresas em financiarem projetos sociais que impactem seu contexto de inserção. O potencial custo-efetividade pode se apresentar como uma solução adequada para incrementar estudos dos reais efeitos sociais promovidos pelas ações, superando lacunas deixadas pela avaliação de resultados, visto que vai além de apenas apresentar dados finais de uma ação. Adotar medidas apropriadas de responsabilidade social previne situações como *socialwashing*, aspecto que pode prejudicar a cadeia de desenvolvimento em ESG na medida em que mina a confiança de clientes e da sociedade frente a essas iniciativas.

Fonte de financiamento

Financiamentos FAPESC No 01/2021 [Protocolo 2021TR1280] e Financiamento CNPq/SEMPI/MCTI nº 021/2021 [Processo 424657/2021-0].

Conflito de interesse

Não há.

Referências

- Américo, Bruno Luiz, & Lacruz, Adonai José. (2017). Contexto e desempenho escolar: Análise das notas na Prova Brasil das escolas capixabas por meio de regressão linear múltipla. *Revista de Administração Pública*, 51(5), 854-878. <http://doi.org/10.1590/0034-7612160483>
- Angrist, Joshua, & Pischke, Jorn-Steffen. (2014). *Mastering metrics: The path from cause to effect*. New Jersey: Princeton University Press.
- Bengo, Irene, Arena, Marika, Azzone, Giovanni, & Calderini, Mario. (2016). Indicators and metrics for social business: A review of current approaches. *Journal of Social Entrepreneurship*, 7(1), 1-24. <http://doi.org/10.1080/19420676.2015.1049286>
- Brasil (2018). *Estudos Etnográficos sobre o Programa Bolsa Família entre Povos Indígenas*. Brasil: Ministério do Desenvolvimento Social.
- Caporale, Guglielmo Maria, Gil-Alana, Luis, Plastun, Alex, & Makarenko, Inna. (2022). Persistence in ESG and conventional stock market indices. *Journal of Economics and Finance*, 46(4), 678-703. <http://doi.org/10.1007/s12197-022-09580-0>
- Costa, Marco Aurélio, & Marguti, Bárbara Oliveira. (2015). *Atlas da vulnerabilidade social nos municípios brasileiros*. Brasília: IPEA.
- Cunningham, Scott. (2021). *Causal inference: The mixtape*. New Haven: Yale University Press.
- Fernandes, L., Leme, Vanessa Barbosa R, Elias, Luciana Carla S., & Soares, Adriana B. (2018). Preditores do desempenho escolar ao final do ensino fundamental: Histórico de reprovação, habilidades sociais e apoio social. *Temas em Psicologia*, 26(1), 215-228. <http://doi.org/10.9788/TP2018.1-09Pt>



- Grieco, Cecilia, Michelini, Laura, & Iasevoli, Gennaro. (2015). Measuring value creation in social enterprises: A cluster analysis of social impact assessment models. *Nonprofit and Voluntary Sector Quarterly*, 44(6), 1173-1193. <http://doi.org/10.1177/0899764014555986>
- Hadad, Shahrazad, & Gauca, Oana. (2014). Social impact measurement in social entrepreneurial organizations. *Management & Marketing*, 9(2), 119-136.
- Insper Metricis. (2022). Guia de avaliação de impacto socioambiental para utilização em projetos e investimentos de impacto: guia geral com foco em monitoramento e verificação de adicionalidade. São Paulo: Insper.
- IPEA (2021). Atlas da Vulnerabilidade Social. Recuperado em 20 de março em <http://ivs.ipea.gov.br/index.php/pt/>.
- Kardec, Allan Kardec Afonso, Arcuri, Rogério, & Cabral, Nelson. (2002). Gestão estratégica e avaliação do desempenho. Rio de Janeiro: Qualitymark Editora Ltda.
- Li, Ting-Ting, Wang, Kai, Sueyoshi, Toshiyuki, & Wang, Derek. (2021). ESG: Research progress and future prospects. *Sustainability*, 13(21), 11663. <http://doi.org/10.3390/su132111663>
- Melo, Anamélia Cavalcanti Carvalho. (2008). *Indicadores de desempenho como instrumento de gestão das entidades do terceiro setor: Um estudo das organizações da sociedade civil de interesse público do Estado da Paraíba* (Dissertação de mestrado). Programa Multiinstitucional e Inter-Regional de Pós-graduação em Ciências Contábeis, Universidade de Brasília/UFPA/UFPE/UFRN, João Pessoa.
- Milani Filho, Marco Antonio Figueiredo. (2009). Resultado econômico em organizações do Terceiro Setor: Um estudo exploratório sobre a avaliação de desempenho. *Revista Contabilidade e Controladoria*, 1(1), 35-44. <http://doi.org/10.5380/rcc.v1i1.14710>
- Palermo, Gabrielle, Silva, Denise Britz do Nascimento, & Novellino, Maria Salet Ferreira. (2014). Fatores associados ao desempenho escolar: Uma análise da proficiência em matemática dos alunos do 5º ano do ensino fundamental da rede municipal do Rio de Janeiro. *Revista Brasileira de Estudos de População*, 31(2), 367-394. <http://doi.org/10.1590/S0102-30982014000200007>
- Rawhouser, Hans, Cummings, Michael, & Newbert, Scott. (2019). Social impact measurement: Current approaches and future directions for social entrepreneurship research. *Entrepreneurship Theory and Practice*, 43(1), 82-115. <http://doi.org/10.1177/1042258717727718>
- Rodrigues, Patrícia Peres, Sugahara, Cibele Roberta, Branchi, Bruna Angela, & Ferreira, Denise Helena Lombardo. (2021). Teoria da mudança e metodologias de avaliação de projetos sociais nas organizações. *Revista de Empreendedorismo, Negócios e Inovação*, 6(1), 55-74. <http://doi.org/10.36942/reni.v6i1.332>
- Sandri, Emanuel, Kumasaka, Júlia Mitsue Vieira Cruz, Cruz, June Alisson Westarb, & Cruzara, Giovanni. (2020). Avaliação do impacto social: Um levantamento bibliométrico. *Teoria e Prática em Administração*, 11(1), 106-121. <http://doi.org/10.22478/ufpb.2238-104X.2021v11n1.52611>
- Sugahara, Cibele Roberta, & Rodrigues, Patrícia Peres. (2019). Avaliação de impacto de negócios sociais e teoria da mudança. *Revista Nacional de Gerenciamento de Cidades*, 7(46), 119-130. <http://doi.org/10.17271/2318847274620192080>