

PESQUISA QUANTITATIVA EM CIÊNCIA POLÍTICA: DESENHOS OBSERVACIONAIS X EXPERIMENTOS NATURAIS

QUANTITATIVE RESEARCH IN POLITICAL SCIENCE: OBSERVATIONAL DESIGNS VS. NATURAL EXPERIMENTS

Gustavo Batista Araujo[†]

Resumo: Existem diversas perguntas interessantes a serem respondidas em todos os campos do conhecimento, e é assim também na Ciência Política. A questão é que a maior parte das situações empíricas em ciências sociais costuma dificultar a realização de experimentos, obrigando-nos a trabalhar com dados observacionais. O problema com esse tipo de dado é que, devido à ausência de controle por parte do pesquisador, podemos ter diversas dificuldades para estimar a relação entre as variáveis de interesse de maneira não-enviesada, como viés de seleção, omissão de variável relevante ou a dificuldade de se estabelecer o contrafactual adequado. Para evitar esses problemas existem diversas alternativas aos pesquisadores: procurar variáveis instrumentais, estimar um modelo de seleção em observáveis ou de *matching*, ou procurar por experimentos naturais. Em particular, este artigo irá se focar nos experimentos naturais e em como sua utilização vem impactando o campo da Ciência Política.

Palavras-chave: Ciência Política. Pesquisa quantitativa em Ciência Política. Desenhos Observacionais.

Abstract: There are several interesting questions to be answered in all fields of knowledge, and that is also true for Political Science. The question is that most empirical situations in the social sciences are often difficult to frame in an experimental way, requiring us to work with observational data. The trouble with this type of data is that, due to the lack of control by the researcher, we have difficulties to estimate the unbiased relationship between the variables of interest. Problems such as selection bias, relevant variable omission or difficulty in establishing the proper counterfactual are always a possibility. So as to avoid these problems, there are several alternatives to researchers: to look for instrumental variables, to estimate a model of selection-on-observables or matching, or to search for natural experiments. In particular, this article will focus on natural experiments and how their use is impacting the field of Political Science.

Keywords: Political Science. Quantitative research in Political Science. Observational Designs.

[†] Professor da FECAP e da Universidade Anhembi-Morumbi e pesquisador do CAENI/USP.

1 Introdução

Existem inúmeras perguntas interessantes e que merecem ser investigadas em todos os campos do conhecimento, e isso, claro, acontece também no campo da Ciência Política. Existem, ainda, múltiplas maneiras de procurar as respostas a essas perguntas, cada uma com seus benefícios, mas também suas desvantagens. Podemos, por exemplo, nos apoiar em métodos qualitativos ou quantitativos, e dentro desses campos metodológicos há ainda uma diversidade de escolhas a fazer. Neste artigo procuramos examinar algumas dessas escolhas dentro de um desenho de pesquisa quantitativo. Começaremos analisando pesquisas que tentaram mostrar o impacto de tomar empréstimos do FMI sobre o desempenho econômico dos países. Usaremos esse caso para observar a diferença entre desenhos observacionais e experimentais. Em seguida, procuraremos mostrar por que a abordagem experimental é desejável, em comparação à observacional. Por fim, analisaremos as possibilidades do desenho de pesquisa experimental no âmbito da Ciência Política e faremos o fechamento do artigo na última seção.

2 Motivação

Qual é o efeito de um empréstimo do FMI sobre o desempenho econômico dos países receptores? Essa foi uma das principais questões de interesse do livro *The IMF and Economic Development*, de James R. Vreeland (2003). A questão é relevante porque o FMI foi, em especial a partir de 1990, uma fonte importante de socorro para países com dificuldades em seu balanço de pagamentos. Mas a instituição colocava condicionalidades à concessão do empréstimo, geralmente relacionadas a reformas ortodoxas sobre a política econômica. Tais reformas, no entanto, foram acusadas de causar mais malefícios do que benefícios aos países receptores do empréstimo. Dessa forma, como poderíamos averiguar o efeito dos empréstimos sobre o desempenho econômico posterior dos países receptores?

Poderíamos pensar em desenhos de pesquisa baseados em metodologia qualitativa para contribuir com o entendimento dessa questão. Mas, neste artigo, o destaque será em desenhos de pesquisa focados em metodologia quantitativa. Assim, uma primeira – e ingênua – sugestão dentro do campo quantitativo poderia ser comparar o desempenho econômico de países que tomaram empréstimos com o desempenho de países que não tomaram empréstimos do FMI, imaginando que a diferença média no desempenho econômico desses países seria explicada pelo fato de alguns realizarem empréstimos e outros não.

Mas, para que uma estratégia desse tipo pudesse mostrar realmente o impacto dos empréstimos sobre o desempenho econômico dos países seria necessário que fosse a única característica diferente entre o grupo dos que realizaram empréstimos em relação ao dos que não realizaram empréstimos. Isso, obviamente, não é verdade. Os países não decidem aleatoriamente recorrer a empréstimos do FMI: sua situação econômica influencia fortemente essa decisão.

Dessa maneira, é quase natural que países que recorrem aos empréstimos do Fundo tenham um desempenho econômico pior do que o de países que não necessitam recorrer a esses empréstimos. Concluir, a partir dessas observações, que os empréstimos do FMI (através de

suas condicionalidades) prejudicam o desempenho econômico dos países receptores seria como considerar que a aspirina causa a dor de cabeça. As pessoas não tomam aspirinas aleatoriamente, tomam porque tem dor de cabeça. Deixar de levar em conta este fato nos atrapalha na tarefa de verificar o real efeito da aspirina ou dos empréstimos do FMI (VREELAND, 2003, p. 4).

Uma segunda sugestão poderia ser deixar de lado a comparação de países que receberam empréstimos com países que não receberam empréstimos, focando assim na comparação de diferentes momentos dos mesmos países: seu desempenho econômico antes de receberem um empréstimo do FMI e depois de receberem. A ideia aqui seria de que a possível diferença de desempenho econômico entre os momentos (antes e depois de receber o empréstimo) poderia ser explicada justamente pelo fato de o país ter recorrido ao FMI.

Para que essa estratégia mostrasse realmente o efeito dos empréstimos sobre o desempenho econômico dos países seria necessário que fossem o único fator que tivesse sido alterado entre os diferentes períodos de tempo, o que é uma hipótese muito forte na maioria das aplicações reais. No caso dos empréstimos do FMI e o desempenho posterior da economia, precisaríamos controlar uma série de fatores cujos efeitos poderiam confundir-se com o dos empréstimos (caso não fossem controlados), como: alterações no governo ou nas instituições domésticas, alterações em políticas domésticas (não relacionadas às condicionalidades do empréstimo), alterações no cenário econômico local e/ou global, entre outros.

Outra possibilidade, esta muito utilizada na literatura recente, volta à comparação do desempenho econômico entre países que tomaram empréstimos e países que não tomaram empréstimos do FMI. Mas, desta vez, há a tentativa de controlar os fatores que influenciam um país a recorrer a esses empréstimos. O que se faz é uma tentativa de corrigir o problema da possível diferença sistemática entre países receptores e não receptores através do cálculo da probabilidade da busca pelos empréstimos em função de seus prováveis determinantes observáveis, incluindo posteriormente essa probabilidade na análise (método de seleção em observáveis¹ – *selection on observables*).

O problema dessa abordagem é que ela não leva em consideração que podem existir fatores não observáveis afetando tanto a busca pelos empréstimos quanto o desempenho econômico dos países. Por exemplo, vamos supor que existam dois tipos de países: aqueles que, frente a uma crise no balanço de pagamentos, têm muita motivação para debelá-la e aqueles que têm pouca motivação. Os países que têm muita motivação procuram ativamente maneiras de resolver ou minorar o problema, inclusive recorrendo a empréstimos do FMI. Já os países pouco motivados não tomam tantas atitudes para resolver ou minorar seu problema.

Pode ser extremamente difícil observar a “motivação” dos países frente a uma crise no balanço de pagamentos, mas é possível que a motivação afete tanto a probabilidade de que os países recorram a empréstimos quanto o desempenho econômico desses países. Vamos supor que países altamente motivados se recuperem melhor de uma crise de balanço de pagamentos do que países com baixa motivação, mesmo na ausência de empréstimos do FMI. Se não levarmos em conta esse fator não observado, corremos o risco de atribuir aos empréstimos um efeito que não é completamente seu, sobrevalorizando seu impacto, neste exemplo (VREELAND, 2003, p. 6).

¹ Para saber mais sobre este método, confira, por exemplo, Dale e Krueger (2002, p. 1491-1527).

Podem existir outros fatores não observados que influenciam tanto a busca por empréstimos quanto o desempenho econômico dos países. Se qualquer desses fatores não for considerado, não será possível observar o real efeito dos empréstimos do FMI sobre o desempenho econômico dos países, pois iremos confundir esse efeito com o dos fatores não observados.

Esse é apenas um dos problemas das pesquisas observacionais. Sem ter muito controle sobre o processo gerador dos dados, o pesquisador tem dificuldades para isolar o efeito que está analisando, pois é possível que existam variáveis relacionadas tanto à variável dependente de interesse quanto à variável independente analisada e que ficaram de fora da pesquisa (muitas vezes por serem não observáveis). Mas as pesquisas observacionais não apresentam apenas este problema. Muitas vezes nos deparamos com outras dificuldades, como a de viés de seleção ou a de se estabelecer o contrafactual adequado.

É por esses motivos que os desenhos de pesquisa experimentais são preferíveis aos observacionais. Iremos agora discorrer um pouco sobre a diferença entre desenhos experimentais e observacionais, sustentando que os primeiros são preferíveis aos segundos.

3 Desenhos de Pesquisa Experimentais x Observacionais

Para captar o real impacto dos empréstimos contraídos junto ao FMI sobre o desempenho econômico dos países, idealmente gostaríamos de observar, para as mesmas unidades de análise, num mesmo instante do tempo, dois resultados potenciais distintos: qual teria sido o desempenho econômico do país sem que tivesse recorrido ao empréstimo do FMI e seu contrafactual: qual teria sido seu desempenho econômico caso tivesse recorrido ao auxílio do Fundo.

Se fosse possível observar ambas as situações para os mesmos indivíduos, num mesmo instante do tempo, seria fácil calcular o efeito individual desse “tratamento” (poderíamos olhar para os empréstimos do FMI como se fossem um “remédio” que os países tomam – e gostaríamos de estimar seu efeito): $\tau_i = y_{1i} - y_{0i}$, onde τ_i indica o efeito do tratamento para o indivíduo i , y_{1i} indica o desempenho econômico do país i se este tivesse recorrido ao empréstimo e y_{0i} indica o desempenho econômico do país i se este não tivesse recorrido ao empréstimo.

Da mesma forma seria relativamente fácil calcular o efeito médio do tratamento (*average treatment effect – ATE*), calculado como a média dos efeitos individuais dos empréstimos:

$$\bar{\tau} = \frac{\sum_{i=1}^n y_{1i}}{n} - \frac{\sum_{i=1}^n y_{0i}}{n}.$$

Este seria o experimento ideal porque manteria inalterados todos os outros fatores que afetam o desempenho econômico das unidades de análise, alterando apenas a variável de interesse – no caso, contrair ou não um empréstimo com o FMI (o “tratamento”). Dessa forma, qualquer diferença de desempenho econômico poderia ser atribuída ao único fator que varia sistematicamente entre os grupos de tratamento e controle: ter ou não contraído um empréstimo com o Fundo.

No entanto, infelizmente não é possível observar ambos os resultados potenciais para as mesmas unidades de análise em um mesmo instante do tempo: ou observamos o desempenho econômico de um país depois de ele ter recorrido a um empréstimo ou observamos seu desempenho

econômico sem que tenha recorrido ao auxílio do FMI (o contrafactual nunca será diretamente observável para as mesmas unidades de análise, em um mesmo instante do tempo).

Assim, se quisermos estimar o efeito do “tratamento” em questão, precisaremos utilizar e comparar os resultados potenciais de indivíduos diferentes. Para alguns países conseguiremos observar, em um determinado momento no tempo, seu desempenho econômico sem que tenham recorrido a empréstimo do FMI (mas não ao contrafactual: seu desempenho econômico, neste mesmo momento, tendo contraído o empréstimo). Para outros, conseguiremos observar, em um determinado momento no tempo, seu desempenho econômico tendo eles recorrido ao socorro do Fundo (mas novamente não ao contrafactual: seu desempenho econômico, neste mesmo momento, sem que tenham contraído o empréstimo).

O problema é que, ao compararmos resultados potenciais para indivíduos diferentes, nada garante que o único fator a variar sistematicamente entre os grupos é o tratamento. Este é o problema dos estudos observacionais: será que conseguimos estabelecer o contrafactual adequado (será que os grupos de controle e tratamento são comparáveis)? Será que há viés de seleção tornando os grupos sistematicamente diferentes? Será que estamos controlando o impacto de todas as variáveis que podem causar confusão na hora de verificar o efeito do tratamento?

Essas questões estão relacionadas entre si, são muitas vezes difíceis de serem respondidas e têm a ver com a própria natureza dos estudos observacionais, assim como com as dificuldades de se realizar estudos experimentais em ciências sociais. Os pressupostos, então, em desenhos de pesquisa observacionais são em geral muito fortes.

Para que estudos observacionais nos deem medidas não enviesadas da relação de interesse, precisamos que os grupos de tratamento e controle sejam comparáveis. Isto é, não pode haver nenhuma característica sistematicamente diferente entre os grupos, caso contrário não conseguiríamos distinguir se o efeito observado deveu-se ao tratamento (administrado a apenas um dos grupos) ou à outra característica sistematicamente diferente entre eles. Dessa forma, se houver viés de seleção – a seleção para os grupos de tratamento e controle for influenciada ou determinada por alguma característica das unidades de análise – os grupos não serão diretamente comparáveis e muito provavelmente iremos estimar a relação de interesse com viés.

A maneira mais simples de explicar esta ideia é pensarmos que os dois grupos, de tratamento e controle, deveriam apresentar resultado semelhante na ausência do tratamento. Voltando ao exemplo dos empréstimos do FMI, o desempenho econômico médio dos países na ausência de empréstimos do Fundo deveria ser parecido. No entanto, se os dois grupos forem sistematicamente diferentes no quesito “motivação” (ou em qualquer outra característica, observável ou não – e se esta característica estiver relacionada de alguma forma ao desempenho econômico do país), então os desempenhos econômicos na ausência dos empréstimos seriam diferentes. Assim, se não levarmos em conta essa diferença de pontos de partida entre os países, poderíamos confundir-la com o efeito do tratamento, obtendo uma medida enviesada do impacto dos empréstimos do FMI.

O problema é que muitas vezes, em estudos observacionais, as variáveis que introduzem diferenças sistemáticas entre os grupos de tratamento e controle são não observáveis. Se fossem observáveis, bastaria incluí-las em nossos modelos para que obtivéssemos estimativas não

enviesadas da relação de interesse (desde que nenhuma outra variável que causa diferenças sistemáticas entre os grupos de tratamento e controle permanecesse fora do modelo – problema conhecido como “omissão de variável relevante”). Ou seja, os pressupostos que fazemos em desenhos observacionais podem ser muito exigentes e de difícil verificação.

Nesse sentido, desenhos experimentais seriam desejáveis, pois permitiriam ao analista controle muito maior sobre a situação analisada. O pesquisador poderia controlar não só a seleção dos indivíduos para os grupos de tratamento e controle, garantindo que não houvesse característica sistemática que os diferenciasse, como também poderia controlar outros fatores que potencialmente poderiam afetar a observação da relação de interesse.

Em geral, os desenhos experimentais realizam a seleção amostral ou a alocação das observações para os grupos de tratamento e controle através de um processo de randomização. A seleção aleatória garantiria que as diferenças em amostras suficientemente grandes entre o grupo dos tratados e dos não tratados seriam não sistemáticas, tanto nas características observáveis quanto nas não observáveis (como decorrência da lei dos grandes números).

Dessa forma, os desenhos experimentais, ao promoverem a alocação aleatória do tratamento, produzem grupos comparáveis no que diz respeito tanto a características observáveis quanto a não observáveis. O problema é que muitas vezes não conseguimos controlar experimentalmente a situação de interesse.

O exemplo a que estamos nos referindo ao longo deste texto ilustra essa situação. Em uma primeira análise, não parece que seria fácil reproduzir experimentalmente a situação de um país com problemas no balanço de pagamentos e que teria a opção de buscar um empréstimo com o FMI. Nem conseguiríamos aleatorizar a concessão de empréstimos pelo fundo. Assim, o que nos resta é trabalhar com os dados observacionais, tendo a consciência de que é possível que haja fatores não observados afetando sistematicamente a relação de interesse.

4 Diminuindo os problemas de desenhos observacionais: a busca por experimentos naturais

No exemplo da relação entre os empréstimos do FMI e o desempenho econômico dos países, pode ser muito difícil encontrar um experimento ou algo que simule uma situação experimental. Dessa forma, é necessário trabalhar com um desenho de pesquisa baseado em dados observacionais.

Como vimos, dados observacionais podem levar a conclusões precipitadas por parte de analistas que não levam em conta a possibilidade de existirem variáveis não observadas afetando a relação de interesse. No caso dos empréstimos do FMI, pode ser que apenas países “motivados” busquem contrair empréstimos com o fundo. Assim, possíveis diferenças na performance econômica de países que contraíram empréstimos em relação aos que não contraíram não poderiam ser completamente atribuídas aos empréstimos, pois poderiam estar relacionadas em algum grau à “motivação”, característica não observada.

Se a característica não observada for observável, basta incorporá-la na análise para resolver o problema de seleção. Vreeland (2003) propõe, para o caso dos empréstimos, uma

estimação em dois estágios: no primeiro procuram-se as características que influenciam a escolha dos países em buscar empréstimos com o FMI. Assim, pode-se calcular a probabilidade de que qualquer país escolha contrair um empréstimo com o Fundo. Então, basta usar essa probabilidade no segundo estágio, como uma nova variável no modelo de regressão linear. Assim encontraríamos o efeito dos empréstimos, mantendo constante a probabilidade de que os países recorram a eles.

Ou seja, seria como se estivéssemos comparando países mais comparáveis, países que tivessem a mesma probabilidade de contraírem empréstimos – usaríamos países com probabilidades semelhantes como contrafactuais uns dos outros (método de seleção em observáveis ou métodos de *matching*). O problema com essa abordagem, como já visto, é que muitas vezes a seleção depende de alguma variável não observável. Se essa variável ficar de fora do modelo, pode ser que a relação verificada ainda seja enviesada.

Vreeland (2003) sugere ainda um teste para que se verifique se deve haver alguma variável importante não observável que ficou de fora do modelo de seleção: a partir das variáveis observadas seria possível prever quais países recorreriam a empréstimos junto ao FMI; seria possível, também, comparar essas previsões aos resultados reais, e as diferenças entre previsões e valores reais seriam os “termos de erro” do modelo. Esses “termos de erro” conteriam a parte da história não observada, mas também um componente aleatório. Se os erros se distribuíssem de maneira aleatória entre os países que recorreram a empréstimos e os que não recorreram, então não haveria variável que impactasse tanto a busca por empréstimos quanto o desempenho econômico dos países e que tivesse ficado de fora do modelo (por ser não observável). No entanto, se os erros não se distribuíssem de maneira aleatória, poderia haver relação entre os erros da regressão de seleção e os erros da regressão de desempenho econômico, o que poderia indicar que haveria um fator não observado influenciando tanto a seleção quanto o desempenho, caso em que nossas estimativas seriam enviesadas.

A questão é que em desenhos observacionais quase sempre corremos o risco de haver alguma variável não observável enviesando a estimativa da nossa relação de interesse. O que fazer, então? Podemos buscar a utilização de um desenho de pesquisa experimental. Mas os experimentos podem ser muito difíceis de realizar em alguns campos, como nas ciências sociais aplicadas. Nesse caso, a melhor alternativa talvez seja a procura por experimentos naturais (ou quase experimentos).

Experimentos naturais seriam experimentos não controlados pelo pesquisador. Existem situações em que a natureza age como se fosse um pesquisador experimental, separando de maneira aproximadamente aleatória unidades de análise para os grupos de controle e tratamento (ROBINSON; MCNULTY; KRASNO, 2009, p. 346). Dessa forma, embora o processo de obtenção dos dados fuja ao controle do pesquisador (e, por isso, pareça não poder ser chamado de experimental), ainda assim produz dados que são semelhantes aos que teriam sido produzidos por um pesquisador com controle sobre o processo gerador e que o realizasse por randomização.

Tais situações não são abundantes, no entanto. Pelo contrário. Encontrá-las exige tempo, paciência e até um pouco de sorte. Apesar do nome – experimentos naturais –, não se derivam apenas de “atos” da natureza, mas de maneira mais geral referem-se a intervenções exógenas que separam as unidades de análise em dois grupos distintos de maneira que aparente uma

alocação aleatória (ROBINSON; MCNULTY; KRASNO, 2009, p. 346).

O importante aqui é que a seleção simule de maneira muito próxima uma alocação randômica. Se isso não ocorrer, estamos novamente no âmbito de um desenho observacional em que pode haver uma característica não observada afetando tanto a seleção quanto o resultado de interesse (o que enviesaria nossas estimativas). Em experimentos naturais o pesquisador não controla o processo gerador da seleção e, por isso, não pode garantir com certeza que se trata de um processo aleatório. Cabe a ele argumentar e tentar mostrar que alguma situação criou um processo de seleção que fez as vezes de uma alocação randômica.

A melhor maneira de mostrar isso é procurando observar o balanceamento em diversas covariadas entre os grupos de tratamento e controle. Se o processo pode ser considerado aproximadamente aleatório, então não devem existir diferenças significativas entre os grupos em qualquer covariada de interesse (em amostras suficientemente grandes). Isso é possível testar empiricamente (através de testes de comparação de médias ou de regressões lineares). No entanto, é necessário haver também balanceamento entre variáveis não observáveis, o que não é possível testar empiricamente. O que se faz, então, é testar o balanceamento para as variáveis observáveis e, se ele se verificar, considerar que deve valer também para as variáveis não observáveis – o que ainda é uma hipótese forte, mesmo comparando-se com as hipóteses que temos que fazer nos desenhos observacionais. Porém, se o mecanismo de seleção for convincente como instrumento de seleção randômico, deposita-se mais confiança no desenho de experimento natural do que em um desenho observacional (ROBINSON; MCNULTY; KRASNO, 2009, p. 348).

Embora o desenho de experimento natural ainda não seja amplamente difundido na Ciência Política, já existem diversos exemplos de sua utilização nesse campo. Krasno e Green (2008), por exemplo, utilizam a construção arbitrária das fronteiras estaduais nos EUA para observar a relação entre propagandas eleitorais na televisão e o comparecimento às urnas dos eleitores. As fronteiras, desenhadas arbitrariamente muito tempo atrás, muito provavelmente não se relacionam a muitas das variáveis de interesse para os pesquisadores. No entanto, criam diferenças em leis, instituições, cultura e outras variáveis potencialmente interessantes para cientistas políticos. Dessa forma, é possível comparar localidades bem próximas, separadas por uma fronteira estadual, pois se imagina que sejam suficientemente semelhantes, exceto por estarem em distintos lados de uma linha imaginária que cria diferenças no tocante a uma determinada lei, por exemplo.

No caso de Krasno e Green (2008), os autores mostram que existem localidades em estados eleitoralmente não competitivos que são contíguas a outras localidades em estados competitivos. Como os mercados locais de mídia não são limitados pelas demarcações fronteiriças (as ondas da transmissão da televisão cruzam as fronteiras), é como se as localidades em estados não competitivos próximas da fronteira com estados competitivos fossem escolhidas aleatoriamente para receber campanha mais intensa, cujo objetivo era impactar as localidades contíguas do estado competitivo. Se compararmos as localidades do estado não competitivo que estiveram imersas em intensa competição (através de propagandas na televisão) com localidades no mesmo estado que não receberam tanta atenção dos partidos, poder-se-ia estimar o efeito da propaganda eleitoral sobre o comparecimento às urnas dos eleitores. Os resultados

encontrados pelos autores contrariam muitos dos estudos observacionais anteriores, concluindo que as propagandas eleitorais não têm grande influência sobre o comparecimento às urnas dos eleitores.

Outros exemplos da utilização de experimentos naturais baseados na geografia são os trabalhos de Miguel (2004), Posner (2004), Huber e Arceneaux (2007), Gerber, Kessler e Meredith (2011), entre outros.

Mas a geografia não é a única forma de intervenção exógena que é capaz de simular uma alocação aleatória. Outro exemplo usado na literatura é o de eleições competitivas. Eleições realizam a separação em dois grupos: os eleitos e os não eleitos. Os eleitos poderão dispor dos benefícios que exercer o cargo em disputa oferece; já os não eleitos não poderão dispor dos mesmos benefícios. Alguns autores procuram usar este fato para estimar a vantagem eleitoral de se tornar um incumbente ao vencer uma eleição. Mas não é possível realizar uma comparação direta entre qualquer eleito e qualquer não eleito, pois não são bons *contrafactuais* uns dos outros. Eleitos e não eleitos provavelmente diferem em características importantes, observáveis e também não observáveis.

Entre as características não observáveis, os candidatos devem apresentar, em média, diferenças em sua “qualidade”. Essa qualidade do candidato deve afetar tanto sua chance de eleição quanto a posterior chance de reeleição. Assim, candidatos com grande “qualidade” teriam boas chances de serem reeleitos mesmo que não houvesse vantagem eleitoral de ser o incumbente. Mas se não levarmos em consideração a “qualidade” do candidato ao estimarmos o efeito eleitoral de ser incumbente, corremos o risco de confundir o efeito de um com o efeito do outro.

Uma maneira quase experimental de analisar a situação seria considerar apenas candidatos separados por uma pequena margem de votos em uma eleição. Candidatos com votações parecidas devem ser parecidos na grande maioria das variáveis que influenciam sua quantidade total de votos, sejam elas observáveis ou não observáveis. Se considerarmos ainda que os candidatos não devem conseguir controlar precisamente o total de votos que irão obter, então candidatos com votações suficientemente próximas poderiam ser suficientemente parecidos e a definição de qual dos dois ficaria na frente se deveria unicamente a fatores aleatórios. Dessa forma, em eleições muito apertadas é como se ocorresse uma alocação aleatória da vitória eleitoral (como se fatores aleatórios definissem a disputa entre candidatos muito parecidos).

Este é o desenho de regressão descontínua. Nele o pesquisador procura se aproveitar de que há um ponto de corte exogenamente definido sobre uma variável contínua (no caso, a votação do candidato), que separa eleitos de não eleitos. É isso exatamente o que faz Lee (2008) em artigo que examina diversas eleições nos EUA para verificar o efeito de se ser o incumbente na probabilidade de eleição do candidato. A conclusão a que chega é que ser o incumbente realmente traz benefícios eleitorais na eleição subsequente.

A análise de Lee (2008) ajuda ainda a vermos um outro aspecto positivo desses desenhos de pesquisa quase experimentais: como já mencionado, é possível testar alguns de seus pressupostos – em contraposição à maior parte dos estudos observacionais, onde só nos resta torcer.

Caughey e Sekhon (2011) testam a hipótese de que as eleições americanas de 1942 a 2008 decididas por pequenas margens produziram seleções que simulariam uma alocação aleatória e

concluíram que não é possível afirmar com grande confiança que isso tenha acontecido. Outros estudos já mostraram que diversas outras eleições em inúmeras localidades obtiveram sucesso em promover uma seleção que se aproximou de uma atribuição randômica. São exemplos desses estudos Eggers, Fowler, Hainmueller, Hall e Snyder Jr. (2015), de la Cuesta e Imai (2016), Boas, Hidalgo e Richardson (2014), Araujo (2012), entre outros.

Um sorteio aleatório não controlado pelo pesquisador também pode fazer o papel de intervenção exógena que gera uma alocação randômica. Angrist, Chen e Song (2011) utilizaram o sorteio para o alistamento obrigatório durante a Guerra do Vietnã, nos EUA, para analisar o efeito do alistamento sobre o salário dos indivíduos depois de retornarem à vida civil. Como é possível que alguma variável não observável afete tanto a escolha individual com relação ao alistamento militar quanto o salário dos indivíduos depois de voltarem à vida civil, então uma estimativa que não levasse em conta essa variável forneceria uma estimativa enviesada do efeito do alistamento sobre o salário futuro dos indivíduos.

A exigência de grande número de soldados durante a Guerra do Vietnã fez com que o alistamento, até então facultativo, passasse a ser compulsório. Os novos recrutas seriam selecionados através de um sorteio (uma loteria). Dessa forma, os indivíduos não teriam a possibilidade de se autoselecionar para o tratamento: a loteria iria dividir o grupo dos que se alistariam do grupo dos que não se alistariam. E, se o sorteio não houver sido fraudado de alguma forma, não há por que acreditar que os grupos seriam diferentes em alguma característica que pudesse afetar também o salário dos indivíduos depois de servirem ao exército.

Utilizando tal desenho de pesquisa os autores concluem que o alistamento militar tem um impacto negativo no salário dos indivíduos depois que retornam à vida civil, ao menos no curto prazo. Outros autores também utilizaram a loteria para o alistamento militar nos EUA durante a Guerra do Vietnã para estudar os efeitos de se passar determinada quantidade de tempo no exército sobre a saúde dos indivíduos (ANGRIST; CHEN; FRANDBSEN, 2010) ou sobre a participação no mercado de trabalho e em programas de transferências governamentais (AUTOR; DUGGAN, 2008), por exemplo.

As fronteiras, eleições e loterias são apenas alguns dos exemplos de situações em que intervenções exógenas criam uma separação aproximadamente aleatória das unidades de análise em grupos de tratamento e controle. Não são os únicos exemplos e possibilidades, contudo, e ainda há muitos experimentos naturais a serem explorados nas ciências sociais. Encontrar essas situações, no entanto, exige grande perspicácia e paciência por parte dos pesquisadores.

Até onde conheço, não existem ainda aplicações de experimentos naturais para determinar o impacto causal dos empréstimos do FMI sobre o desempenho econômico posterior dos países. É possível até que não se possa encontrar um experimento natural que permitisse a análise de tal relação. Mas, pelo que se viu até aqui, é importante e desejável que se continue procurando.

5 Conclusão

Cientistas políticos, assim como analistas em outros campos das Ciências Sociais, enfrentam diversos desafios ao empreender pesquisas que têm a intenção de estabelecer relações

causais entre variáveis. Muitas vezes os pesquisadores não têm condições de realizar experimentos controlados e precisam se basear em dados cujo processo gerador não dominam.

O desenho de pesquisa que utilizam, no entanto, não precisa necessariamente se render ao fato de que precisará usar dados observacionais. Talvez um desenho observacional seja a única alternativa, mas há maneiras de conseguir estimativas mais adequadas em um desenho desse tipo através de métodos de seleção em observáveis ou de *matching*.

No entanto, se for possível, o pesquisador deve procurar por situações onde uma intervenção exógena cria condições de estimar uma relação causal sem viés, ao separar as unidades de análise em grupos distintos de tratamento e controle, como se houvesse sido realizado um processo de randomização. Ao encontrar esses experimentos naturais (ou quase experimentos) o pesquisador poderá ficar muito mais confiante de ter encontrado uma relação causal do que ficaria em um desenho meramente observacional.

Este artigo teve como principal objetivo discutir a importância dos desenhos de pesquisa em Ciências Sociais, enfatizando os benefícios de desenhos experimentais e quase experimentais. Esperamos com isso ter contribuído, mesmo que de maneira marginal, para a disseminação desses importantes instrumentos no campo das ciências sociais aplicadas – e em especial no campo da Ciência Política.

Referências

ANGRIST, J. D.; CHEN, S. H. Schooling and the Vietnam-Era GI Bill: evidence from the Draft Lottery. *American Economic Journal: Applied Economics*. v. 3, n. 2, p. 96-119, 2011.

ANGRIST, J. D.; CHEN, S. H.; FRANDBSEN, B. Did Vietnam Veterans Get Sicker in the 1990s? The complicated effects of military service on self-reported health. *Journal of Public Economics*, n. 94, p. 824-837, 2010.

ANGRIST, J. D., CHEN, S. H.; SONG, J. Long Term Consequences of Vietnam-Era Conscriptio: new estimates using social security data. *Papers and Proceedings*, n. 94, p. 824-837, 2011.

ANGRIST, J. D.; PISCHKE, J. S. *Mostly Harmless Econometrics: na empiricist's companion*. Princeton, NJ: Princeton University Press, 2009.

ARAUJO, G. B. *Contribuições de campanha influenciam decisões públicas? O caso dos contratos públicos federais e das emendas ao orçamento no Brasil*. Tese (Doutorado em Ciência Política). Universidade de São Paulo, 2012.

AUTOR, D. H.; DUGGAN, M. G. The Effect of Transfer Income on Labor Force Participation and Enrollment in Federal Benefits Programs: evidence from the veterans disability compensation program. *NBER Papers on Retirement Research Center Project, Paper NB08-07*. 2008.

BOAS, T.; HIDALGO, D.; RICHARDSON, N. The Spoils of Victory: campaign donations and government contracts in Brazil. *Journal of Politics*, n. 76, v. 2, p. 415-429, 2014.

CAUGHEY, D.; SEKHON, J. S. Elections and the Regression Discontinuity Design: Lessons from Close U. S. House Races, 1942-2008. *Political Analysis*, n. 19, p. 385-408, 2011.

de la CUESTA, B.; IMAI, K. Misunderstandings about the Regression Discontinuity Design in the Study of Close Elections. *Annual Review of Political Science*, v. 19, p. 375-396, 2016.

DALE, S. B.; KRUEGER, A. B. Estimating the Payoff to Attending a More Selective College: an application of selection on observables and unobservables. *The Quarterly Journal of Economics*, n. 117, v. 4, p. 1491-1527, 2002.

DUNNING, T. Improving Causal Inference: strenghts and limitations of natural experiments. *Political Research Quarterly*, n. 61, p. 282-293, 2008.

EGGERS, A. et al. On the Validity of the Regression Discontinuity Design for Estimating Electoral Effects: new evidence from over 40.000 close races. *American Journal of Political Science*, n. 59, v. 1, p. 259-274, 2015.

GERBER, A. S.; KESSLER, D. P.; MEREDITH, M. The Persuasive Effects of Direct Mail: a regression discontinuity based approach. *Journal of Politics*, n. 73, v. 1, p. 140-155, 2011.

HUBER, G. A.; ARCENEUX, K. Identifying the Persuasive Effects of Presidential Advertising. *American Journal of Political Science*, n. 51, v. 4, p. 957-977, 2007.

KRASNO, J. S.; GREEN, D. P. Do Televised Presidential Ads Increase Voter Turnout? Evidence from a natural experimete. *Journal of Politics*, n. 70, v. 1, p. 245-261, 2008.

LEE, D. S. Randomized Experiments from Non-Random Selection in U.S. House Elections. *Journal of Econometrics*, n. 142, p. 675-697, 2008.

MIGUEL, E. Tribe or Nation? Nation building and public goods in Kenya versus Tanzania. *World Politics*, n. 56, v. 3, p. 327-362, 2004.

POSNER, D. N. The Political Salience of Cultural Difference: why chewas and tumbukas are allies in Zambia and adversaries in Malawi. *The American Political Science Review*, n. 98, v. 4, p. 529-545, 2004.

ROBINSON, G.; MCNULTY, J. E.; KRASNO, J. S. Observing the Counterfactual? The search for political experiments in nature. *Political Analysis*, n. 17, p. 341-357, 2009.

RUBIN, D. B. Estimating Causal Effects of Treatments in Randomized and Non-Randomized Studies. *Journal of Educational Psychology*, n. 66, p. 688-701, 1974.

VREELAND, J. R. *The IMF and Economic Development*. Cambridge, UK: Cambridge University Press, 2003.