

Apêndice A

Variável instrumental

Ivan Filipe de Almeida Lopes Fernandes

SciELO Books / SciELO Livros / SciELO Libros

FERNANDES, I.F.A.L. Apêndice A - Variável instrumental. In: *A democracia reduz a desigualdade econômica? Um estudo sobre as possibilidades de construção de uma sociedade mais igual por meio da democracia* [online]. São Bernardo do Campo, SP: Editora UFABC, 2017, pp. 265-269. ISBN: 978-85-68576-79-3. <https://doi.org/10.7476/9788568576793>.



All the contents of this work, except where otherwise noted, is licensed under a [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Todo o conteúdo deste trabalho, exceto quando houver ressalva, é publicado sob a licença [Creative Commons Atribuição 4.0](https://creativecommons.org/licenses/by/4.0/).

Todo el contenido de esta obra, excepto donde se indique lo contrario, está bajo licencia de la licencia [Creative Commons Reconocimiento 4.0](https://creativecommons.org/licenses/by/4.0/).

Apêndice A - Variável instrumental

O principal objetivo de uma pesquisa que busca avaliar os efeitos de uma variável independente em uma variável dependente é garantir a validade da relação causal proposta. No plano ideal, a melhor maneira de garantir que a relação entre duas variáveis seja causal seria trabalhar com contrafactuais, ou seja, observar a presença ou ausência da variável de interesse, dadas condições ambientais exatamente iguais. Uma forma clássica de garantia contrafactual é o uso de um experimento randômico. Contudo em estudos observacionais, aqueles nos quais não é possível o uso da técnica experimental, é necessário a utilização de métodos que permitam a superação de problemas de autosseleção¹²¹.

Para descrever essa questão de maneira precisa, definimos o tratamento como uma variável binária; $D_i = \{0, 1\}$. O resultado de interesse é definido por Y_i . A questão a ser pesquisada é se a variação em D_i afeta o valor de Y_i . Assim, para qualquer indivíduo i existem dois resultados potenciais:

¹²¹ Angrist e Pischke (2008) ilustram essa questão discutindo se o atendimento hospitalar melhora ou piora a saúde de seus pacientes. Obviamente, a mera comparação entre os que foram ao hospital e aqueles que não foram não é suficiente, pois obviamente, quem busca um hospital possui uma condição de saúde mais frágil.

$$\text{resultado potencial} = \begin{cases} Y_{0i}, & \text{se } D_i = 0 \\ Y_{1i}, & \text{se } D_i = 1 \end{cases} \quad (\text{Eq. A.1})$$

Caso fosse possível observar os dois resultados potenciais, o efeito de D_i sobre Y_i seria facilmente calculado subtraindo-se $Y_{1i} - Y_{0i}$ para cada indivíduo e assim estimaríamos o efeito causal calculando a diferença média encontrada na amostra. Contudo, inevitavelmente só podemos ver um dos dois resultados potenciais. Outra forma possível é atribuir os tratamentos $D_i = 0$ e $D_i = 1$ de forma aleatória na amostra observada, de maneira a controlar para possíveis fatores que interfiram na relação (fatores de confusão). A atribuição aleatória do tratamento resolveria os problemas de seleção, pois torna D_i independente dos resultados potenciais. O experimento assegura, portanto, que a relação causal de interesse seja independente dos resultados potenciais. (ANGRIST; PISCHKE, 2008).

Em estudos observacionais o ideal experimental é muitas vezes impraticável. O estudo de regimes políticos comparado é exemplo típico desses casos. Por razões óbvias, não é possível atribuir o tipo de regime de dado país aleatoriamente. As características de cada um dos indivíduos (nesse caso os países) da população de países se autoseleciona a adotar certo tipo de regime. Assim, é necessário controlarmos a análise de consequências da democracia para as variáveis conectadas ao próprio processo de democratização.

A primeira forma de aproximar os estudos observacionais do ideal experimental reside na utilização de controles e assumir a suposição da seleção em observáveis. Isto é, assumir a suposição de que as variáveis que interferem na relação causal são conhecidas e observáveis e, portanto, podem ser controladas (ANGRIST; PISCHKE, 2008).

Outra ferramenta mais sofisticada para permitir a inferência causal em um cenário de pesquisa observacional é o uso de uma variável instrumental (IV), que permite solucionar o problema da autosseleção. Em nosso tema de estudo, o problema da autosseleção é claro, uma vez que a literatura sobre a redemocratização discute os efeitos que a própria desigualdade econômica possui sobre a possibilidade de democratização e consolidação democrática (BOIX, 2003; ACEMOGLU; ROBINSON, 2006; HOULE, 2009).

Suponha as seguintes equações sobre a relação entre democracia (D) e desigualdade (Y):

$$Y_i \equiv f_i(s) \tag{Eq. A.2}$$

$$f_i(s) = \alpha_1 + \beta * D_i + \delta_1 * X_i + \varepsilon_i \tag{Eq. A.3}$$

onde D_i é o indicador de democracia, X_i é o vetor de variáveis de controle e ε_i é o termo de erro. O problema crucial dessa equação é que o termo de erro é correlacionado com D_i ; caso não fosse, estaríamos assumindo que não existe qualquer relação entre desigualdade e democracia em um país, o que parece pouco plausível:

$$Cov(D_i, \varepsilon_i) \neq 0. \tag{Eq. A.4}$$

Com isso uma estimação pelo tradicional método de MQO (mínimos quadrados ordinários) gerarão resultados inconsistentes e enviesados, pois a equação não atende às Hipóteses de Gauss-Markov. O método de IV fornece uma solução a esse problema. Para isso é necessário escolher uma variável instrumental Z_i que não esteja na equação A.3 e que satisfaça três condições. (WOOLDRIDGE, 2010):

Primeiro, Z_i deve ser não correlacionado com o termo de erro ε_i .

$$Cov(Z_i, \varepsilon_i) = 0. \quad (\text{Eq. A.5})$$

Isto é Z_i deve ser exógeno à equação A.3, não sendo, portanto, um determinante da desigualdade econômica dentro de um determinado país.

Segundo, Z_i deve ser parcialmente correlacionado com D_i , uma vez controlada todas as outras variáveis de controle da equação A.3.

$$Cov(Z_i, D_i | X_i) \neq 0. \quad (\text{Eq. A.6})$$

A forma mais simples de identificar essa relação é estimando a equação:

$$D_i = \alpha_1 + \rho * Z_i + \delta_2 * X_i + \xi_i ; \text{ onde } \rho \neq 0 \quad (\text{Eq. A.7})$$

Quando Z_i satisfaz as condições A.5 e A.6 ela torna-se um instrumento em potencial da equação A.3. O segundo passo na análise é inserir a Equação A.7 na Equação Principal A.3:

$$f_i(s) = \alpha_3 + \lambda * Z_i + \delta_3 * X_i + \varepsilon_i \quad (\text{Eq. A.8})$$

rearranjando, assim, o termo de erro na forma reduzida:

$$u_i = \varepsilon_i + \beta * \xi_i ; \quad (\text{Eq. A.9})$$

e o coeficiente do efeito do tratamento D_i :

$$\lambda = \beta * \rho \quad (\text{Eq. A.10})$$

A última e final condição, denominada como restrição de exclusão, assume que a relação entre Z_i e Y_i é única e exclusivamente transmitida por D_i . Desta forma, todo e qualquer efeito de Z_i estimado na equação A.8 é de fato o efeito de D_i em Y_i purificado do problema de causalidade reversa, ou qualquer outro problema de endogeneidade no termo de erro. Assim, segundo os pressupostos assumidos u_i é não correlacionado com todas as variáveis explanatórias de A.8 e uma regressão de MQO pode assim estimar os parâmetros reduzidos de forma consistente¹²² (WOOLDRIDGE, 2010).

Conforme indicado no Capítulo 3, a variável selecionada como instrumento da relação entre democracia e desigualdade é o próprio processo de difusão de democracia no mundo, que afeta a desigualdade econômica de um dado país, única e exclusivamente pelos efeitos que possui sobre o seu regime político.

¹²² Do ponto de vista estritamente econométrico a hipótese de restrição de exclusão é decorrência lógica da condição 1. Apenas a apresentamos para sublinhar a importância da hipótese de que um regime político democrático, isto é, a competição política eleitoral, seja o único meio de transmissão dos efeitos da difusão de democracia sobre a desigualdade econômica.