



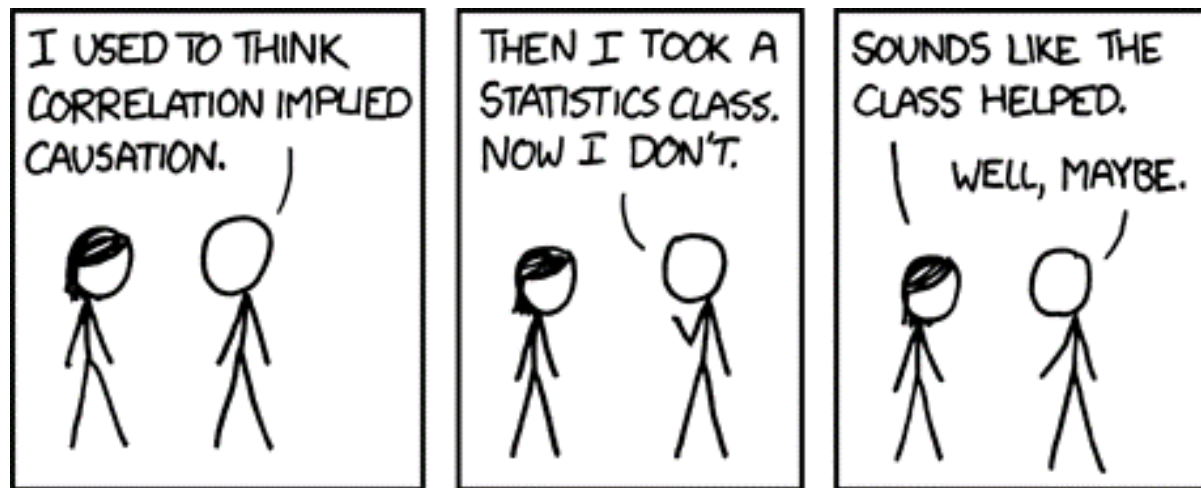
DCP098

Fundamentos para Avaliação Quantitativa de Políticas Públicas

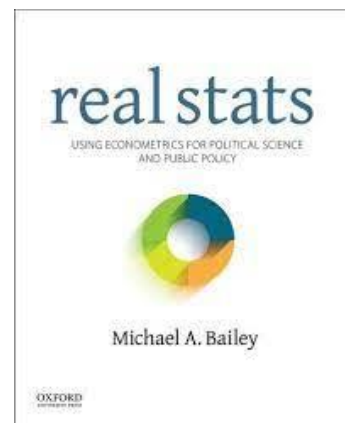
Espuriedade, endogeneidade e viés (vs. exogeneidade, acurácia e validade interna)

Aula 04
11 de abril de 2022

Ana Paula Karruz



Fonte: Bailey (2016: 3). Original em: <https://xkcd.com/552/>



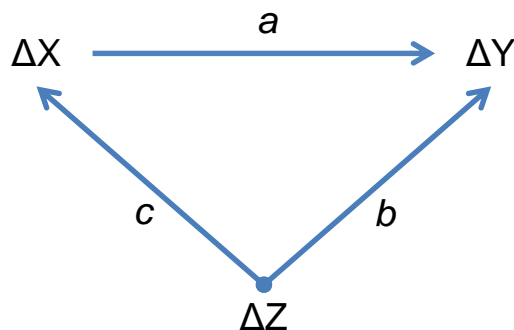
OK. We get it. Correlação/ associação não implica causalidade.

Mas, então, o que implica causalidade? Resposta: Atendimento aos três requisitos para inferência causal: ANTERIORIDADE, ASSOCIAÇÃO E NÃO ESPURIEDADE.

Se conseguirmos identificar alguma variação da suposta causa (X) que seja independente de variáveis omitidas (i.e., variação exógena; neste caso, independente de Z) e ainda assim essa suposta causa se correlacionar/ associar com o efeito em questão, provavelmente essa correlação/ associação é causal (digo “provavelmente” pois a correlação/ associação ainda poderia ser fortuita).

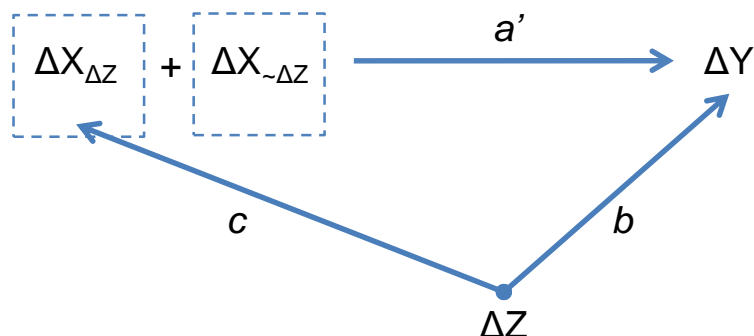
Zooming in na associação espúria

a é associação espúria; a' pode não ser espúria



Associação a é espúria (i.e., não é uma medida válida da associação causal entre X e Y) porque reporta uma combinação de dois efeitos: efeito de X sobre Y e efeito de Z sobre Y.

- Para que a fosse não espúria, para cada potencial fonte de espuriedade Z, as associações b e c não poderiam existir concomitante (ou seja, pelo menos uma delas teria de ser nula)
- **Atribuição aleatória ao tratamento** (desenho experimental) e **análise por estrato de Z eliminam variações de Z**; se Z não varia, não se associa com X ou Y na amostra
- Análise por estrato de Z é uma solução inferior à atribuição aleatória ao tratamento; estratificação deve ocorrer para todo potencial Z, e pode render estratos com poucas observações



- Se associação a' for não nula para todas as potenciais fontes de espuriedade Z, então a proposição de que variações em X causam variações em Y é válida.
- **O controle estatístico** (i.e., regressão múltipla em que se estima o efeito de X sobre Y controlando-se por Z) e o uso de **variável instrumental** são formas de se **identificar a associação a'**
- Porém, ambas são soluções inferiores à atribuição aleatória ao tratamento

Nem todos os fatores incidentes sobre Y são conhecidos ou podem ser observados, e alguns podem causar viés

Fatores desconhecidos ou desconsiderados possivelmente associam-se ao tratamento

Viés ou “inacurácia” é a tendência a estimar efeitos sistematicamente afastados do efeito “verdadeiro”. Uma fonte frequente de viés é a existência de um condicionante da variável dependente (de resultado) que esteja omitido do modelo (lógico e/ou estatístico) e que se correlacione/ associe com a exposição ao tratamento. Quando isso acontece, dizemos que a exposição ao tratamento é endógena.

Endogeneidade gera viés/ Busca por exogeneidade

Na prática, devemos **começar cada análise avaliando a endogeneidade**. Primeiro, desvie o olhar do modelo por um momento e **liste todas as coisas que podem determinar a variável dependente**. Em segundo lugar, pergunte se **alguma coisa na lista se correlaciona [associa] com a variável independente no modelo** e explique por que isso acontece. É isso. Faça isso e estaremos a caminho de identificar a endogeneidade.

Bailey (2016, p. 19)

Nosso desafio central é **evitar a endogeneidade e, assim, alcançar a exogeneidade**. Se tivermos sucesso, podemos estar mais confiantes de que **fomos além da correlação [associação] e estamos mais perto de entender se X causa Y** – nosso objetivo fundamental. Este processo não é automático nem fácil. Muitas vezes não conseguimos encontrar variações puramente exógenas, então teremos que pensar em quão perto podemos chegar. No entanto, a conclusão é esta: se pudermos encontrar variação exógena em X, podemos usar dados para fazer uma inferência razoável sobre o que acontecerá com a variável Y se mudarmos a variável X.

Bailey (2016, p. 16-17)

Nem todos os fatores incidentes sobre Y são conhecidos ou podem ser observados, e alguns podem causar viés

Fatores desconhecidos ou desconsiderados possivelmente associam-se ao tratamento

Estimação de efeitos das variáveis incluídas associadas com omitidas **confundirá a influência das duas**, resultando em viés – i.e., tendência a **superestimar ou subestimar** a importância da variável incluída na produção da variável dependente.

Diante da **atribuição aleatória ao tratamento** (típica de desenhos experimentais), presume-se que o nível da variável omitida é semelhante nos grupos tratado e não tratado, portando neutralizando eventual associação entre variáveis incluídas e omitidas.

Desenhos **quase experimentais** não contam com atribuição aleatória ao tratamento, mas aplicam outras estratégias para aprofundar a comparabilidade entre grupos em dados observacionais (i.e., não experimentais).

Voltaremos a este ponto

Randomized experiments as a gold standard

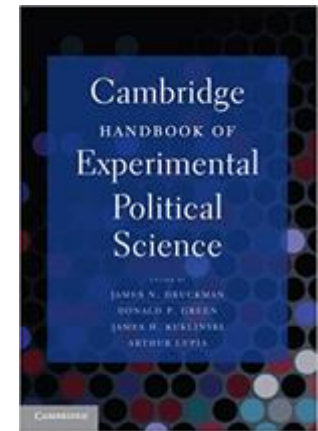
A good way of having exogenous variation is to create it

- Com base na **pergunta** de pesquisa, identifique uma **população** relevante e obtenha uma **amostra (aleatória ou não)**
- Use algum **procedimento aleatório para atribuir as unidades observadas a um de dois grupos**:
 - Grupo de **tratamento**: recebe a intervenção (e.g., política pública)
 - Grupo de **controle**: não recebe a intervenção

Amostragem vs. atribuição aleatória ao tratamento

*Um objetivo comum na amostragem aleatória é selecionar participantes de uma população mais ampla de modo a oferecer a cada potencial participante a mesma probabilidade de ser recrutado para o estudo. A atribuição aleatória é diferente. [...] O requisito principal é que um procedimento aleatório, como um sorteio com cara ou coroa, determine se os participantes do estudo receberão um certo tratamento. **Assim como um experimento não requer amostragem aleatória, um estudo com amostra aleatória nem sempre é um experimento.***

(Druckman et al., 2011, p. 17-18; tradução nossa)



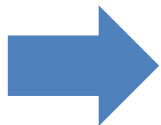
Randomized experiments as a gold standard

A good way of having exogenous variation is to create it

- **Após o tratamento, compare** o comportamento dos grupos de tratamento e controle no tocante ao **resultado de interesse**
- Se o grupo de tratamento **diferir substancialmente** do grupo de controle, **acreditamos que o tratamento teve um efeito**; caso contrário, estamos inclinados a acreditar que o tratamento não teve efeito
- Sabe-se com certeza que **nenhum dos condicionantes sistemáticos de Y afetaram a atribuição** ao tratamento
- A atribuição aleatória ao tratamento enseja uma **presunção de comparabilidade** entre grupos
- Em outras palavras, não há razão para acreditar que possíveis variáveis omitidas estejam associadas com o grau de exposição ao tratamento (já que este foi determinado aleatoriamente)
- Isso vale para as **variáveis (condicionantes) observáveis e não observáveis**. O grande valor da atribuição aleatória ao tratamento é a presunção de comparabilidade entre grupos nas dimensões não observáveis

Desenho experimental em avaliação de políticas públicas

Indicação:



- Mensurar o efeito (resultado ou impacto) do programa sobre seus beneficiários ou comunidade atendida
- Dito de outra forma: Investigar a estrutura e intensidade de causalidade entre uma variável-efeito dependente e seus fatores determinantes

Características:

- Inspirado no formato clássico do experimento nas **ciências naturais**
- Para o desenho **experimental** de pesquisa, a **atribuição ao tratamento** (i.e., definição de quem vai se beneficiar do programa) **deve ser aleatória**

Sobre o **desenho experimental** e a **comparabilidade presumida entre grupos formados aleatoriamente** a partir de uma **mesma população**:

*Como os **dois grupos** foram expostos às mesmas condições ‘de laboratório’ e são **muito parecidos** (pois são amostras de uma mesma população), a **eventual diferença** entre eles seria consequência do fato de um deles ter tido **acesso ao programa** (enquanto o outro não). Se a eventual diferença das medidas é significativa em termos estatísticos, estaria comprovada a tese de que o programa teria algum impacto sobre seus beneficiários.*

Jannuzzi (2016, 97)

EXEMPLO: EXPERIMENTO DE CAMPO



Attanazio et al. (2017, p. ii)

Resumo traduzido do relatório “Impact of free availability of childcare on labour supply and child development in Brazil”, datado de 2017 e disponível em: <http://www.3ieimpact.org/evidence-hub/publications/impact-evaluations/impact-free-availability-public-childcare-labour>.

Resumo

Neste relatório, examinamos os impactos do acesso a creche gratuita em período integral sobre o desenvolvimento da criança, a oferta de trabalho dos membros do domicílio e os ambientes domésticos. Usamos dados da cidade do Rio de Janeiro, Brasil, primariamente de domicílios localizados em favelas.

O programa de creches públicas do Rio de Janeiro é um programa integrado de desenvolvimento da primeira infância para crianças de 0-3 anos vivendo em bairros de baixa renda. O programa consiste de uma variedade de intervenções baseadas nas creches, incluindo oferta de creche em período integral, serviços de saúde, alimentação e provisão de brinquedos pedagógicos e materiais escolares para as crianças.

Espera-se que o acesso a essas creches bem equipadas e adequadamente administradas, ao propiciar um ambiente acolhedor e estimulante para crianças carentes, impulsione a acumulação de capital humano pelos pobres. Porque os pais são encorajados a participar da vida de seus filhos na creche, e até recebem da equipe da creche informação sobre como cuidar das crianças, também poderíamos esperar uma mudança nos comportamentos parentais. Adicionalmente, é esperado que mães e outros membros do domicílio encarregados de cuidar das crianças passem a dispor do tempo para buscar emprego e aumentar sua renda, melhorando assim o bem-estar dos domicílios. Este seria outro mecanismo através do qual acesso à creche poderia levar a um melhor desenvolvimento da criança.

[Continua]

EXEMPLO: EXPERIMENTO DE CAMPO



Attanazio et al. (2017, p. ii)

[Continuação]

Este estudo identifica a relação causal entre acesso a creche e resultados relacionados ao desenvolvimento infantil e participação maternal no mercado de trabalho, tirando proveito de um sorteio realizado pelo governo municipal do Rio de Janeiro em 2007. Todo ano, o governo do Rio oferece aproximadamente 10.000 novas vagas em creches para crianças de 0-3 anos. In 2007, como em anos anteriores, a demanda por essas vagas foi largamente superior à oferta. Mais de 25.000 famílias inscreveram-se para 10.000 novas vagas. Para garantir igualdade de oportunidade, um sorteio foi usado para atribuir as vagas aos interessados elegíveis (aproximadamente 24.000 dos 25.000). Novas crianças beneficiárias começaram a receber serviços em fevereiro de 2008.

Construímos um rico banco de dados que inclui várias medições das crianças e um questionário muito detalhado sobre o domicílio.

Encontramos impactos bastante fortes da frequência à creche pública sobre a altura e o peso das crianças, diversos anos após terem deixado as creches. Também, achamos evidência de que ter acesso à creche produz impactos substantivos e duradouros na renda domiciliar, gastos com bens duráveis e investimos nas crianças – tanto em tempo como em bens.

Particularmente, há um aumento da oferta de trabalho e renda dos avós (principalmente das avós) residindo no mesmo domicílio que uma criança frequentando a creche. Notadamente, esses impactos estão presentes mesmo quatro anos e meio após a aleatorização inicial, num momento em que pouquíssimas crianças de nossa amostra ainda frequentam creches.

Autoria: Orazio Attanasio (Institute for Fiscal Studies); Ricardo Paes de Barros (Insper); Pedro Carneiro (University College London); David Evans (World Bank); Lycia Lima (Fundação Getulio Vargas, São Paulo); Pedro Olinto (World Bank); Norbert Schady (Inter-American Development Bank).

RETOMANDO: Delineamentos experimentais nem sempre são aplicáveis em casos concretos de políticas públicas

*Os grupos de tratamento e comparação acabam sendo definidos em função de **atrasos** ou do **calendário de implantação** de um programa social ou, ainda, pela **cobertura deficiente** de público-alvo deste. Jannuzzi (2016, 97)*

São desafios dos desenhos experimentais:

- Vencer problemas éticos na escolha e justificação de **quem será beneficiário** e quem vai ficar de fora de um programa (supostamente benéfico)
- Evitar a **evasão (atrito)** dos casos analisados
- Garantir que os **efeitos medidos sejam apenas do programa** em foco, em meio a um contexto crescente de intervenções sociais concomitantes
- Escolher a **melhor medida** para captar a dimensão supostamente impactada
 - O que se supõe como dimensão impactada guarda, pelo desenho lógico do programa, vinculação estreita com as ações e atividades deste?
 - O impacto deve ser medido sobre os beneficiários e/ou na comunidade a que pertencem?
- Avaliar em que medida as **circunstâncias “artificiais”** que rodeiam a avaliação do programa (especialmente a atribuição aleatória ao tratamento) conspiram **contra a generalização** dos resultados (**validade externa** do desenho)



Validities, observational data

A **validade interna** refere-se a se a inferência é enviesada; **validade externa** refere-se a se uma inferência se aplica de forma mais geral. Um **experimento bem executado será internamente válido**, o que significa que os resultados nos levarão, em média, a fazer as inferências corretas sobre o tratamento e o resultado no contexto do experimento.

[...]

Mesmo com validade interna, no entanto, um **experimento pode não ser válido externamente**, porque a **relação** causal entre o tratamento e o resultado **pode diferir** em outros contextos.

[...]

Consequentemente, mesmo que os experimentos ofereçam uma abordagem conceitualmente clara para derrotar a endogeneidade, eles **nem sempre podem oferecer a palavra final** na pesquisa em economia, políticas públicas e política. **Portanto, a maioria dos estudiosos na maioria dos campos precisa lidar com dados não experimentais.** Os estudos observacionais usam dados que foram **gerados por processos não experimentais**.

[...]

A endogeneidade será um problema crônico, mas não estamos totalmente indefesos em a luta contra isso.

Exemplos de dados observacionais?

Desenho típico: quase experimento com algum tipo de análise multivariada – isso é bom o suficiente?

Bailey (2016, p. 36; tradução nossa)



Lista de desejos para a nossa avaliação

- **Ausência de espuriedade** nas associações observadas entre exposição à política (X) e dimensões de resultado (Y): permite que confiemos que a associação entre X e Y não é fortuita, nem pode ser explicada por uma terceira variável, omitida da análise
- **Variação exógena** na exposição ao tratamento (X): **afasta espuriedades**
 - Variação exógena significa que X é independente de fatores influentes sobre Y não incluídos na análise
 - Com variação exógena de X, evitamos que X esteja correlacionado com o termo de erro (ε)
 - Do contrário, se X for endógeno, teremos **viés** (pois no processo de estimação da equação de regressão populacional, a parte de ε que se correlaciona com X vai compor o β_{hat})
- **Acurácia** (i.e., **ausência de viés ou pouco viés**): é pré-requisito para um desenho de pesquisa **internamente válido**, que oferece estimativas próximas ao verdadeiro efeito da política
 - Do contrário, se houver viés β_{hat} expressa não apenas o efeito de X sobre Y, mas uma combinação desse efeito com o efeito das variáveis omitidas correlacionadas com X

Um pequeno glossário

Em ordem de aparição dos conceitos nesta disciplina, mais ou menos 😊

Conceito	Enfoque	Referência	Descrição
Espuriedade	Estatística	Wikipedia (https://en.wikipedia.org/wiki/Spurious_relationship)	"In statistics, a spurious relationship or spurious correlation is a mathematical relationship in which two or more events or variables are associated but not causally related, due to either coincidence or the presence of a certain third, unseen factor (referred to as a 'common response variable', 'confounding factor', or 'lurking variable')."
		Siegel (2016, "Practical Business Statistics", https://www.sciencedirect.com/topics/mathematics/spurious-correlation)	"One possible basis for correlation without causation is that there is some hidden, unobserved, third factor that makes one of the variables seem to cause the other when, in fact, each is being caused by the missing variable. The term spurious correlation refers to a high correlation that is actually due to some third factor. For example, you might find a high correlation between hiring new managers and building new facilities. Are the newly hired managers 'causing' new plant investment? Or does the act of constructing new buildings 'cause' new managers to be hired? Probably there is a third factor, namely, high long-term demand for the firm's products, which is causing both."

Um pequeno glossário

Em ordem de aparição dos conceitos nesta disciplina, mais ou menos ☺

Conceito	Enfoque	Referência	Descrição
Endogeneidade	Equação de regressão	Bailey (2016, p. 14-15)	<p>"An independent variable is endogenous if changes in it are related to factors in the error term. The prefix 'endo' refers to something internal, and endogenous variables are 'in the model' in the sense that they are related to other things that also determine Y.</p> <p>In the donuts example, donut consumption is likely endogenous because how many donuts a person eats is not independent of other factors that influence weight gain. Factors that cause weight gain (such as eating Ben and Jerry's ice cream) might be associated with donut eating; in other words, factors that influence the dependent variable Y might also be associated with the independent variable X, muddying the connection between correlation and causation. If we can't be sure that our variation in X is not associated with factors that influence Y, we worry about wrongly attributing to X the causal effect of some other variable. We might wrongly conclude donuts cause weight gain when really donut eaters are more likely to eat tubs of Ben and Jerry's, which is the real culprit.</p> <p>In all these examples, something in the error term that really causes weight gain is related to donut consumption. When this situation arises, we risk spuriously attributed to donut consumption the causal effect of some other factor. Remember, anything not measured in the model is in the error term and here, at least, we have a wildly simple model in which only donut consumption is measured. So Ben and Jerry's, genetics, and everything else are in the error term."</p>
Exogeneidade	Equação de regressão	Bailey (2016, p. 16)	<p>"The opposite of endogeneity is exogeneity. An independent variable is exogenous if changes in it are not related to factors in the error term. The prefix 'exo' refers to something external, and exogenous variables are 'outside the model' in the sense that their values are unrelated to other things that also determine Y. For example, if we use an experiment to randomly set the value of X, then changes in X are not associated by factors that also determine Y. This gives us a clean view of the relationship between X and Y, unmuddled by associations between X and other factors that affect Y."</p>

Um pequeno glossário

Em ordem de aparição dos conceitos nesta disciplina, mais ou menos ☺

Conceito	Enfoque	Referência	Descrição
Viés	Avaliação de políticas públicas	Slides da disciplina	Viés ou “inacurácia” é a tendência a estimar efeitos sistematicamente afastados do efeito “verdadeiro”. Uma fonte frequente de viés é a existência de um condicionante da variável dependente (de resultado) que esteja omitido do modelo (lógico e/ou estatístico) e que se correlacione/ associe com a exposição ao tratamento. Quando isso acontece, dizemos que a exposição ao tratamento é endógena.
		Gertler et al. (2018, p. 60)	"Quando o grupo de comparação não fornece o verdadeiro contrafactual, o impacto estimado do programa será inválido. Em termos estatísticos, a estimativa do impacto será enviesada." "Intuitivamente, um grupo de comparação inválido é aquele que difere do grupo de tratamento de alguma forma que não apenas a ausência do tratamento. Essas diferenças adicionais podem fazer com que a estimativa do impacto seja inválida ou, em termos estatísticos, enviesada: a avaliação do impacto não estimará o verdadeiro impacto do programa. Em vez disso, ela estimará o efeito do programa combinado com essas outras diferenças."
	Equação de regressão	Bailey (2016, p. 87)	Viés significa que a distribuição teórica de β hats não se encontra centrada no verdadeiro β . Assim, há uma tendência sistemática a sobrestimar β (se o centro da distribuição teórica de β hats estiver à direita do verdadeiro β) ou a subestimar β (se o centro da distribuição teórica de β hats estiver à esquerda do verdadeiro β).

Um pequeno glossário

Em ordem de aparição dos conceitos nesta disciplina, mais ou menos ☺

Conceito	Enfoque	Referência	Descrição
Validade interna	Methodologia de pesquisa em geral	Shadish, Cook e Campbell (2002, "Experimental and quasi-experimental designs for generalized causal inference", p. 53)	"We use the term internal validity to refer to inferences about whether observed covariation between A and B reflects a causal relationship from A to B in the form in which the variables were manipulated and measured. To support such an inference, the researcher must show that A preceded B in time, that A covariates with B [...] and that no other explanations for the relationship are plausible."
	Avaliação de políticas públicas	Gertler et al. (2018, p. 81-82)	"Validade interna significa que o impacto estimado do programa é livre de todos os outros fatores potenciais externos que afetam os resultados (fatores confundidores), ou seja, que o grupo de comparação fornece uma estimativa válida do contrafactual, de modo que podemos estimar o verdadeiro impacto do programa."
Validade externa	Methodologia de pesquisa em geral	Shadish, Cook e Campbell (2002, "Experimental and quasi-experimental designs for generalized causal inference", p. 83)	"External validity concerns inferences about the extent to which a causal relationship holds over variations in persons, settings, treatments, and outcomes."
	Avaliação de políticas públicas	Gertler et al. (2018, p. 82)	"Validade externa significa que a amostra de avaliação representa corretamente a população de unidades elegíveis. Dessa maneira, os resultados da avaliação podem ser generalizados para essa população."

Um pequeno glossário

Em ordem de aparição dos conceitos nesta disciplina, mais ou menos 😊

Conceito	Enfoque	Referência	Descrição
Acurácia	Avaliação de políticas públicas	Slides da disciplina	Ausência de viés ou pouco viés; alta validade interna. Implica que a relação identificada não é espúria e que o desenho da pesquisa possibilita mitigar eventuais endogeneidades.
	Medição de variáveis	Babbie (2004, "The practice of social research", p. 140)	"Don't confuse precision with accuracy, however. Describing someone as 'born in New England' is less precise than 'born in Stowe, Vermont--but suppose the person in question was actually born in Boston. The less-precise description, in this instance, is more accurate, a better reflection of the real world."



DCP098

Fundamentos para Avaliação Quantitativa de Políticas Públicas

Espuriedade, endogeneidade e viés (vs. exogeneidade, acurácia e validade interna)

Aula 04
11 de abril de 2022

Ana Paula Karruz