

# Como Economistas Descobrem Relações de Causa e Efeito no Mundo Real

Correlação, causalidade e o problema do contrafactual

---

Beatriz Gietner

O Puzzle Central

Uma Breve História do Problema

Por Que a Correlação Falha

A Estrutura Contrafactual

Quatro Formas de Construir um Contrafactual

Ensaio Aleatorizados (RCT)

Diferenças-em-Diferenças (DiD)

Regressão Descontínua (RDD)

Variáveis Instrumentais (IV)

Predição, Causalidade e Transporte

Efeitos, Mecanismos e Interpretação

Como Ler Afirmações Causais

Leituras Complementares

# O Puzzle Central

---

# O Puzzle Central: Correlação é Fácil

## Pergunta:

*Pessoas com diploma universitário ganham salários mais altos.  
Isso significa que a universidade **causa** salários mais altos?*

**Resposta curta:** talvez.

**O problema:** pessoas que vão para a universidade podem já ser diferentes antes da universidade.

**A pergunta causal não é:** quem ganha mais?

**A pergunta causal é:** quanto essa mesma pessoa ganharia se não tivesse ido para a universidade?

# O Que Observamos vs. O Que Queremos Saber

O que observamos	O que queremos saber
Salário de quem foi para a universidade	Salário dessas mesmas pessoas se não tivessem ido
Salário de quem não foi para a universidade	Salário dessas mesmas pessoas se tivessem ido

## O problema fundamental:

- Para cada pessoa, observamos apenas um caminho.
- Não observamos a mesma pessoa em dois mundos diferentes.
- O segundo mundo é o **contrafactual**.

# O Problema da História Que Não Vemos

Para saber se uma política causou um resultado, gostaríamos de observar dois mundos:

- o mesmo país **com** a política;
- o mesmo país **sem** a política.

Mas o mundo só revela uma dessas histórias.

A inferência causal é a tentativa de construir uma comparação plausível para a história que não observamos.

Esta é a dificuldade central por trás de quase todo artigo empírico em economia.

**Pergunta:** Alunos de escolas privadas têm notas mais altas nos exames.

Podemos concluir que escolas privadas causam melhor desempenho?

1. Sim, porque os alunos têm notas mais altas.
2. Não, porque famílias que escolhem escolas privadas podem ser diferentes.
3. Só se compararmos alunos com características observáveis parecidas.
4. Só se houver uma fonte convincente de variação exógena.

**Melhor resposta:** 4. Controlar características observáveis ajuda, mas não resolve automaticamente seleção não observada.

# Uma Breve História do Problema

---

# Uma Pergunta Antiga

David Hume<sup>1</sup> fez uma pergunta simples:

*Quando dizemos que X causa Y, o que observamos de fato?*

Observamos sequências:

- X acontece;
- depois Y acontece.

Mas a ligação causal não aparece diretamente nos dados.

Por isso precisamos de teoria, desenho de pesquisa e pressupostos.

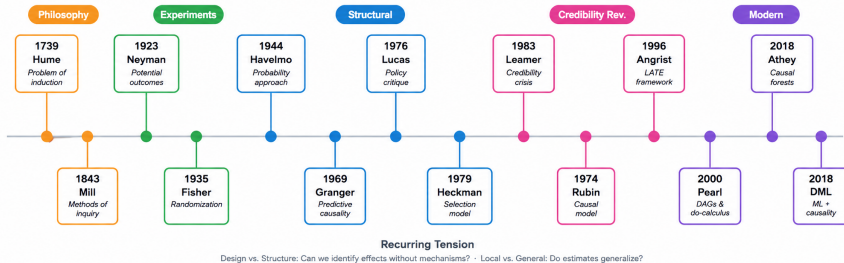
---

<sup>1</sup>A Treatise of Human Nature (1739); An Enquiry Concerning Human Understanding (1748).

# Causalidade: Uma Pergunta Antiga

## A Timeline of Causal Inference in Economics

*From Hume's skepticism to machine learning*



A economia moderna desenvolveu ferramentas para tornar perguntas causais empiricamente tratáveis.

Referência: linha do tempo de Carlos Chavez, ON CAUSALITY.

## A virada pragmática na economia:

*Em vez de perguntar “o que é causalidade em abstrato?”, economistas geralmente perguntam: “que comparação identificaria o efeito causal?”*

Essa mudança é útil porque leva a discussão para perguntas concretas:

- Quem recebe o tratamento?
- Quem é o grupo de comparação?
- Qual é o resultado?
- O que torna a atribuição do tratamento plausivelmente independente dos resultados potenciais?

Os dados, sozinhos, não respondem a essas perguntas. O desenho de pesquisa responde.

# Por Que a Correlação Falha

---

# Correlação e Causalidade: Dois Erros (Parte 1)

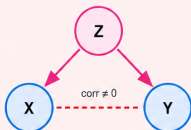
## The Two Directions of the Correlation-Causation Fallacy

*Both mistakes are common in applied research*

### Part 1: Correlation $\neq$ Causation (the famous one)

#### A. Confounding

(common cause)



Ice cream  $\leftrightarrow$  drowning  
(Z = summer heat)

#### B. Reverse Causation

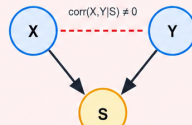
(wrong direction)



Police  $\leftrightarrow$  crime rates  
(crime  $\rightarrow$  police deployment)

#### C. Selection Bias

(collider)



Talent  $\leftrightarrow$  beauty (actors)  
(S = became actor)

A frase famosa é correta: **correlação não implica causalidade**. A figura mostra três razões clássicas.

Figura adaptada de "The Two Directions of the Correlation-Causation Fallacy", de Carlos Chavez, ON CAUSALITY.

# Por Que Correlação Não Implica Causalidade

## 1. Confounding / variável omitida

Vendas de sorvete e afogamentos podem estar positivamente correlacionados. Mas sorvete não causa afogamentos. A variável omitida é o calor: dias quentes aumentam tanto o consumo de sorvete quanto o número de pessoas nadando.

## 2. Causalidade reversa

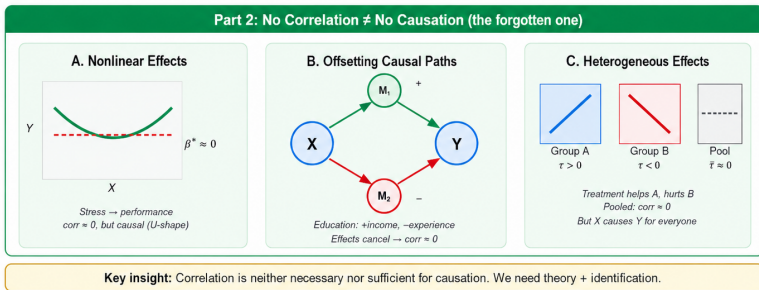
Podemos observar mais policiais em lugares com mais crime. Isso não significa que policiais causam crime. A direção pode ser inversa: áreas com mais crime recebem mais policiamento.

## 3. Viés de seleção

Se olharmos apenas para pessoas que se tornaram atores, talento e beleza podem parecer correlacionados. Mas a amostra já foi selecionada: para entrar no grupo observado, a pessoa provavelmente precisou ter pelo menos uma dessas características.

O problema comum: os dados observados misturam causalidade com seleção, variáveis omitidas ou direção inversa.

# Correlação e Causalidade: Dois Erros (Parte 2)



A frase também é **incompleta**: ausência de correlação não implica ausência de causalidade.

Figura adaptada de “The Two Directions of the Correlation-Causation Fallacy”, de Carlos Chavez, ON CAUSALITY.

# Por Que Pode Haver Causalidade Sem Correlação?

## 1. Efeitos não lineares

Estresse pode melhorar desempenho até certo ponto, mas prejudicar desempenho quando fica alto demais. Se olhamos apenas para uma relação linear média, a correlação pode parecer próxima de zero.

## 2. Caminhos causais que se anulam

Treinamento profissional pode aumentar habilidades, mas também reduzir horas trabalhadas durante o treinamento. Um canal aumenta salários futuros; outro reduz renda no curto prazo. Dependendo do período observado, os efeitos podem se compensar.

## 3. Efeitos heterogêneos

Um programa pode ajudar alguns alunos e prejudicar outros. Se juntamos todos em uma média, os efeitos podem se cancelar, mesmo que o programa tenha efeito causal para cada grupo.

A ausência de uma correlação simples não prova ausência de efeito causal.

# Erro 1: Correlação Sem Causalidade

Existem três problemas clássicos:

1. **Variável omitida:** uma terceira variável causa X e Y.
2. **Causalidade reversa:** Y causa X, não X causa Y.
3. **Seleção:** pessoas ou unidades entram no tratamento por razões ligadas ao resultado.

A mesma correlação pode ser compatível com várias histórias causais.

Por isso, um coeficiente de regressão não é automaticamente um efeito causal.

# Problema 1: Variável Omitida

**Correlação:** pessoas com mais educação ganham mais.

**Possíveis variáveis omitidas:**

- contexto familiar;
- habilidade prévia;
- motivação;
- qualidade da escola;
- redes de contato.

O **problema**: a diferença salarial pode refletir educação, mas também diferenças que já existiam antes da educação.

## Problema 2: Causalidade Reversa

**Suponha que observamos:**

*Quando o preço é mais alto, empresas vendem mais.*

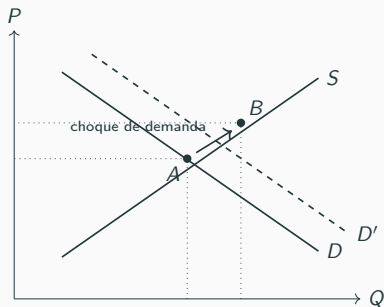
Isso significa que a curva de demanda é inclinada para cima?

**Não necessariamente.** Talvez estejamos observando choques de demanda.

- Demanda aumenta  $\Rightarrow$  preço sobe e quantidade sobe.
- Oferta cai  $\Rightarrow$  preço sobe e quantidade cai.

Mesma correlação. Histórias causais diferentes.

# Oferta, Demanda e Identificação



A curva de demanda se desloca de  $D$  para  $D'$ .

O equilíbrio vai de  $A$  para  $B$ .

**Dados observados:**

preço e quantidade aumentam juntos.

**Problema de identificação:**

isso não significa que a demanda inclina para cima.

Podemos estar vendo deslocamentos de demanda, não movimento ao longo de uma curva fixa.

Para estimar a demanda, precisamos de variação que desloque a oferta.

## Problema 3: Viés de Seleção

**Observação:** pessoas que vão ao hospital têm maior probabilidade de morrer do que pessoas que não vão ao hospital.

**Conclusão ruim:** hospitais causam morte.

**Explicação melhor:** pessoas mais doentes têm maior probabilidade de ir ao hospital.

**Problema de seleção:** o tratamento não é aleatório.

## Erro 2: Causalidade Sem Correlação Simples

O erro menos óbvio é pensar que ausência de correlação significa ausência de causalidade.

Isso também pode falhar.

- **Efeitos não lineares:** uma relação em  $U$  pode ter correlação linear próxima de zero.
- **Canais que se anulam:** um efeito positivo e outro negativo podem cancelar-se na média.
- **Efeitos heterogêneos:** o tratamento ajuda alguns e prejudica outros.

Correlação não é necessária nem suficiente para causalidade.

## Exemplo: Efeitos Heterogêneos

Imagine um programa de tutoria.

- Para alunos com dificuldades moderadas, a tutoria aumenta notas.
- Para alunos muito avançados, a tutoria pode substituir estudo independente e não ajudar.
- Para alguns alunos, o efeito pode ser próximo de zero.

A média pode esconder efeitos muito diferentes.

A pergunta causal muitas vezes não é apenas “qual é o efeito médio?”, mas “para quem?”

**A correlação muitas vezes compara pessoas que fizeram escolhas diferentes.**

**A causalidade faz outra pergunta:**

*O que teria acontecido ao grupo tratado se ele não tivesse recebido o tratamento?*

**O desafio:** encontrar um grupo de comparação que represente esse contrafactual.

# A Estrutura Contrafactual

---

# O Contrafactual Ausente

Para cada pessoa, podemos imaginar dois resultados potenciais:

- $Y_i(1)$ : resultado da pessoa  $i$  se ela receber o tratamento;
- $Y_i(0)$ : resultado da pessoa  $i$  se ela não receber o tratamento.

**O efeito causal para a pessoa  $i$  seria:**

$$Y_i(1) - Y_i(0)$$

O problema: para cada pessoa, observamos apenas um desses dois resultados.

Se a pessoa recebe o tratamento, observamos  $Y_i(1)$ , mas não observamos  $Y_i(0)$ . Se a pessoa não recebe o tratamento, observamos  $Y_i(0)$ , mas não observamos  $Y_i(1)$ .

Geralmente, não conseguimos estimar o efeito causal individual.

Por isso, costumamos perguntar sobre efeitos médios.

**Efeito médio do tratamento na população:**

$$ATE = E[Y_i(1) - Y_i(0)]$$

**Efeito médio do tratamento entre os tratados:**

$$ATT = E[Y_i(1) - Y_i(0) \mid D_i = 1]$$

**ATE:** efeito médio do tratamento para toda a população.

**ATT:** efeito médio do tratamento para quem recebeu o tratamento.

$D_i = 1$ : a pessoa  $i$  recebeu o tratamento.

## Exemplo: Educação e Salários

Suponha que o tratamento seja frequentar a universidade.

- $Y_i(1)$ : salário da pessoa  $i$  se ela frequentar a universidade;
- $Y_i(0)$ : salário da pessoa  $i$  se ela não frequentar a universidade.

**O efeito causal da universidade para essa pessoa seria:**

$$Y_i(1) - Y_i(0)$$

Mas nunca observamos a mesma pessoa nos dois cenários ao mesmo tempo.

A dificuldade da inferência causal é encontrar uma comparação convincente para o resultado que não observamos.

# Estratégia de Identificação

Uma estratégia de identificação explica por que a comparação é convincente.

## **A pergunta central:**

*De onde vem a variação no tratamento?*

- atribuição aleatória?
- mudança de política?
- ponto de corte para elegibilidade?
- regra institucional?
- acidente histórico?

Sem uma fonte clara de variação, a interpretação causal fica frágil.

# A Metáfora do Transporte

Uma forma útil de pensar sobre inferência causal é esta:

*Todo método tenta mover informação de onde a observamos para onde precisamos dela.*

## Exemplo:

Queremos saber o efeito de frequentar a universidade sobre salários.

Para uma pessoa que frequentou a universidade, observamos:

$$Y_i(1)$$

Mas queremos saber também:

$$Y_i(0)$$

Esse é o salário que ela teria recebido se não tivesse frequentado a universidade.

# De Onde Vem o Contrafactual?

Como não observamos o contrafactual diretamente, tentamos procurá-lo em outro lugar.

Podemos usar informação de:

- outra pessoa parecida;
- outra escola parecida;
- outra região antes e depois de uma política;
- outro período de tempo;
- pessoas logo acima e logo abaixo de um ponto de corte.

**A pergunta de identificação é:**

*Essa comparação representa bem o que teria acontecido sem o tratamento?*

A parte difícil é defender que a comparação é convincente.

# Quatro Formas de Construir um Contrafactual

---

# Quatro Formas de Construir um Contrafactual

Método	De onde vem o contrafactual
RCT	um grupo de controle sorteado
DiD	outro grupo antes e depois de uma política
RDD	unidades logo abaixo ou acima de um ponto de corte
IV	variação deslocada por um fator externo

**Mesmo objetivo:** encontrar uma comparação que indique o que teria acontecido sem o tratamento.

Estas estratégias não são mutuamente exclusivas. Em alguns estudos, elas podem ser combinadas.

O método depende da pergunta, do contexto institucional, dos dados disponíveis e da fonte de variação no tratamento.

# RCT: Quando Criamos a Comparação

**Ideia:** atribuir o tratamento aleatoriamente.

- O grupo de tratamento recebe o programa.
- O grupo de controle não recebe o programa.
- A aleatorização torna os grupos comparáveis, em média.

**Por que funciona:** a situação de tratamento não é escolhida pelos indivíduos.

Isso importa porque pessoas que escolhem receber um tratamento podem ser diferentes das que não escolhem. A aleatorização reduz esse problema: antes do tratamento, os dois grupos tendem a ser parecidos tanto em características observáveis quanto em características não observáveis.

# Exemplo: Sorteio de Bolsas

Imagine que uma universidade tem 1.000 candidatos elegíveis para 300 bolsas.

Se as 300 bolsas são atribuídas por sorteio:

- antes do sorteio, cada candidato elegível tem a mesma chance de receber a bolsa;
- por isso, receber a bolsa não depende de renda, habilidade, motivação ou contexto familiar;
- os grupos sorteado e não sorteado tendem a ser comparáveis em média;
- em amostras maiores, grandes diferenças prévias entre os grupos se tornam menos prováveis;
- qualquer diferença posterior nos resultados pode ser interpretada como efeito causal da bolsa, em média.

O sorteio cria uma fonte de variação no tratamento que não depende das escolhas ou características dos indivíduos.

# Por Que o Sorteio Ajuda?

O sorteio não garante que os grupos serão idênticos em uma amostra finita.

Mas ele muda algo fundamental:

*receber a bolsa não depende das escolhas ou características dos candidatos.*

Em outras palavras, a bolsa não é recebida porque o estudante é mais rico, mais motivado, mais preparado ou tem uma família com mais informação.

**A variação no tratamento vem do sorteio ao invés da seleção dos próprios indivíduos.**

## O Que a Aleatorização Garante?

Em um RCT, a independência vem do desenho do estudo.

A atribuição ao tratamento é independente dos resultados potenciais:

$$D_i \perp (Y_i(1), Y_i(0))$$

Isto significa que receber a bolsa não está relacionado ao que o estudante teria alcançado com ou sem a bolsa.

Também esperamos equilíbrio nas características pré-tratamento:

$$D_i \perp X_i$$

Aqui,  $X_i$  representa características medidas antes do tratamento, como renda familiar, notas anteriores ou contexto escolar.

# Aleatorização e Amostras Finitas

A aleatorização não elimina todas as diferenças entre os grupos em uma amostra específica.

Pode acontecer, por acaso, que o grupo sorteado tenha mais alunos com notas altas, ou mais alunos de famílias ricas.

O ponto é que essas diferenças surgem por acaso, não porque os indivíduos escolheram o tratamento.

Com amostras maiores, a lei dos grandes números torna grandes desequilíbrios menos prováveis.

Por isso, em um RCT, diferenças médias posteriores entre tratamento e controle podem ser interpretadas como efeito causal médio do tratamento, desde que o sorteio tenha sido implementado corretamente.

RCTs são poderosos, mas nem sempre são possíveis.

- Não podemos atribuir aleatoriamente todos os anos de educação de uma pessoa.
- Não podemos atribuir recessões aleatoriamente.
- Alguns experimentos seriam antiéticos.
- Alguns experimentos são caros demais.
- Alguns experimentos estudam contextos muito específicos.

Por isso, economistas procuram experimentos naturais.

# Diferenças-em-Diferenças (DiD)

**Pergunta:** qual é o efeito de uma política?

**Exemplo:**

- A Região A aumenta o salário mínimo.
- A Região B não muda o salário mínimo.
- Comparamos como o número de trabalhadores empregados mudou nas duas regiões, antes e depois da política.

**Ideia:** o grupo de controle ajuda a estimar a tendência que o grupo tratado teria seguido sem a política.

$$\text{DiD} = (\text{Depois} - \text{Antes})_{\text{Tratado}} - (\text{Depois} - \text{Antes})_{\text{Controle}}$$

## **Interpretação:**

O grupo de controle estima o que teria acontecido com o grupo tratado na ausência da política.

O DiD desconta mudanças comuns ao longo do tempo que afetam os dois grupos.

# Uma Tabela Simples de DiD

Exemplo: número de trabalhadores empregados.

	Antes	Depois	Mudança
Região tratada	50	60	10
Região de controle	40	45	5
Diferenças-em-diferenças			5

**Estimativa ingênua antes-depois:** 10.

**Mudança no grupo de controle:** 5.

**Estimativa DiD:**  $10 - 5 = 5$ .

A mudança no grupo de controle é usada como aproximação para o que teria acontecido com a região tratada na ausência da política.

O efeito estimado da política é a mudança adicional observada na região tratada.

## Tendências paralelas:

*Na ausência do tratamento, os grupos tratado e de controle teriam seguido trajetórias semelhantes.*

Esse pressuposto não é diretamente testável.

Mas podemos perguntar:

- os grupos já tinham trajetórias parecidas antes da política?
- o grupo de controle é uma comparação plausível?
- aconteceu algo diferente no grupo tratado ao mesmo tempo?

Se o grupo tratado já estava mudando mais rapidamente antes da política, o DiD pode confundir uma tendência anterior com o efeito do tratamento.

Outra condição importante é que os indivíduos não reajam ao tratamento antes dele acontecer.

### **Exemplo:**

Se empresas sabem que o salário mínimo vai subir no próximo ano, elas podem ajustar contratações antes da mudança oficial.

Nesse caso, o período “antes” já foi afetado pela política.

O problema: o DiD compararia um antes que já não é realmente anterior ao tratamento.

## DiD: Outras Ameaças à Interpretação

Mesmo quando a ideia básica do DiD é clara, a interpretação causal ainda pode falhar.

- **Spillovers:** a política no grupo tratado afeta o grupo de controle.
- **Políticas simultâneas:** outra mudança acontece ao mesmo tempo que o tratamento.
- **Mudança na composição:** os grupos mudam ao longo do tempo.
- **Diferenças de exposição:** nem todos os tratados são afetados pela política da mesma forma.

A pergunta prática é sempre a mesma: o grupo de controle representa bem o que teria acontecido com o grupo tratado sem a política?

# Regressão Descontínua (RDD)

**Ideia:** o tratamento é atribuído com base em um ponto de corte.

**Exemplo:**

- Estudantes com nota  $\geq 70$  são elegíveis para uma bolsa.
- Estudantes com nota  $< 70$  não são elegíveis.

Comparamos estudantes logo acima e logo abaixo de 70.

**A intuição:** perto do ponto de corte, os estudantes tendem a ser muito parecidos, mas têm situação de tratamento diferente.

Existem dois casos principais.

## Sharp RDD:

- cruzar o ponto de corte determina completamente o tratamento;
- exemplo: todos com nota  $\geq 70$  recebem a bolsa, todos abaixo de 70 não recebem.

## Fuzzy RDD:

- cruzar o ponto de corte aumenta a probabilidade de receber o tratamento;
- exemplo: estudantes com nota  $\geq 70$  ficam elegíveis, mas alguns não usam a bolsa.

Nos dois casos, a variação importante vem da descontinuidade criada pelo ponto de corte.

# Por Que o RDD Pode Ser Convicente

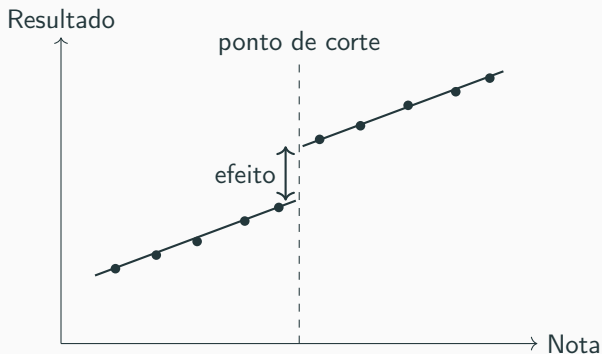
Estudantes com notas 69,8 e 70,2 provavelmente são muito parecidos.

Mas um fica logo abaixo do ponto de corte e o outro fica logo acima.

O ponto de corte cria variação local quase aleatória.

A estimativa é convincente perto do ponto de corte. Ela não informa automaticamente o efeito para estudantes longe do ponto de corte.

# Intuição Visual do RDD



O salto descontínuo no ponto de corte é interpretado como efeito causal se os resultados potenciais seriam contínuos ao redor do limiar na ausência do tratamento. Também precisamos que as unidades não consigam manipular precisamente sua posição em torno do ponto de corte.

## Variáveis Instrumentais (IV)

**Pergunta:** qual é o efeito de mais educação sobre salários?

**Problema:** pessoas não escolhem educação aleatoriamente.

Pessoas que estudam mais podem já ser diferentes antes da decisão educacional.

- podem ter maior habilidade;
- podem vir de famílias com mais recursos;
- podem ter mais informação sobre oportunidades educacionais;
- podem ter frequentado escolas melhores.

Por isso, comparar salários de pessoas com mais e menos educação pode misturar o efeito da educação com diferenças pré-existentes.

# Ideia das Variáveis Instrumentais

**Ideia:** encontrar uma fonte de variação em educação que não venha da escolha dos indivíduos.

**Exemplo:** distância até a universidade mais próxima.

- morar mais perto pode aumentar os anos de educação;
- mas, para ser um bom instrumento, a distância não pode afetar salários por outros canais.

O instrumento tenta separar a variação em educação que é quase externa da variação que vem de seleção individual.

# As Condições para um Instrumento Válido

Para que  $Z$  seja um instrumento válido para o tratamento  $D$ , precisamos de três ideias.

- **Relevância:**  $Z$  afeta  $D$ .
- **Independência:**  $Z$  não está correlacionado com fatores não observados que afetam  $Y$ .
- **Restrição de exclusão:**  $Z$  afeta  $Y$  apenas por meio de  $D$ .

Em geral, a restrição de exclusão é a parte mais difícil de defender.

Um instrumento isola apenas a parte do tratamento que foi deslocada por uma fonte externa de variação.

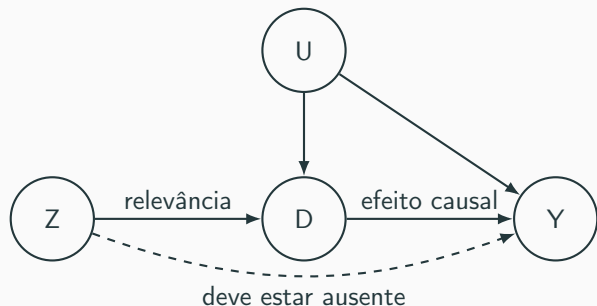
**Perguntamos:**

*Entre as pessoas cuja educação mudou por causa do instrumento, qual foi o efeito sobre os salários?*

IV geralmente estima um efeito local.

A estimativa vale para as pessoas cuja situação de tratamento mudou por causa do instrumento. Ela não representa necessariamente o efeito para toda a população.

## IV em Diagrama



$U$  representa fatores não observados, como habilidade, motivação ou contexto familiar. Um instrumento válido desloca o tratamento, não está associado a  $U$  e não afeta o resultado diretamente. Esta última condição é a restrição de exclusão.

# Predição, Causalidade e Transporte

---

# Prever Não É o Mesmo Que Causar

A mesma variável pode servir para duas perguntas muito diferentes.

## Pergunta de previsão

*Dado o que sei sobre X, consigo prever Y?*

- Vendas de sorvete ajudam a prever afogamentos.
- Mas vender mais sorvete não causa afogamentos.
- O calor aumenta os dois.

## Pergunta causal

*Se eu mudar X, o que acontece com Y?*

- Mais escolaridade aumenta salários?
- Salário mínimo altera emprego?
- Bolsas aumentam permanência na universidade?

Previsão é sobre adivinhar melhor. Causalidade é sobre o efeito de uma mudança.

# Predição vs. Inferência Causal

## **Pergunta de predição:**

*Dado o que sei sobre X, consigo prever Y?*

## **Pergunta causal:**

*Se eu intervenho e mudo X, o que acontece com Y?*

Essas são perguntas diferentes.

- Predição pode funcionar com correlações.
- Inferência causal precisa de um contrafactual defensável.

# De Onde Vem o Contrafactual?

Método	Contrafactual vem de...	Risco principal
RCT	grupo de controle sorteado	atrito, não adesão, validade externa
DiD	outro grupo ao longo do tempo	tendências não paralelas
RDD	pessoas perto do ponto de corte	manipulação do ponto de corte
IV	variação externa no tratamento	restrição de exclusão inválida

Todo método causal faz uma afirmação de transporte.

Ele move informação de um grupo, tempo, lugar ou margem para o contrafactual de que precisamos.

# Transporte Entre Contextos

Um resultado pode ser internamente convincente e ainda ter validade externa limitada.

**Exemplo:** um programa de tutoria funciona em uma cidade.

Podemos concluir que ele vai funcionar:

- em outra cidade?
- em escala nacional?
- com professores diferentes?
- com uma população estudantil diferente?

A estimativa identifica o que aconteceu no contexto do estudo. Aplicá-la em outro lugar exige outro argumento.

Quando um artigo empírico termina com uma recomendação de política, ele muitas vezes faz duas afirmações:

1. **Afirmação interna:** o efeito estimado é causal no contexto do estudo.
2. **Afirmação de transporte:** esse efeito é informativo para o contexto de política que nos interessa.

A primeira afirmação é sobre identificação.

A segunda afirmação é sobre validade externa, escala, contexto e comportamento.

**Ambas precisam ser defendidas.**

# Efeitos, Mecanismos e Interpretação

---

# E o Mecanismo?

Até aqui, a pergunta principal foi:

*X causa Y?*

Mas muitas vezes também queremos saber:

*Por que X causa Y? Por qual canal?*

**Exemplo:** educação pode aumentar salários porque:

- aumenta habilidades produtivas;
- funciona como sinal para empregadores;
- amplia redes de contato;
- muda o tipo de ocupação acessível ao trabalhador.

Identificar um efeito causal não identifica automaticamente o mecanismo.

## Como Ler Afirmações Causais

---

# Todo Método Tem uma Ameaça

Método	Por que funciona	Ameaça principal
RCT	aleatorização	atrito / não adesão
DiD	comparação de tendências	tendências não paralelas
RDD	ponto de corte	manipulação ao redor do ponto de corte
IV	variação externa	restrição de exclusão

Não existe método mágico. Existem pressupostos mais ou menos convincentes em cada contexto.

# Como Economistas Defendem Essas Estratégias?

Na prática, pressupostos de identificação raramente são “provados” diretamente.

Eles são defendidos combinando:

- **conhecimento institucional:** por que a regra, o sorteio, o ponto de corte ou a política gerou variação plausivelmente externa?
- **testes de balanceamento:** os grupos eram parecidos antes do tratamento?
- **tendências pré-tratamento:** no DiD, os grupos já se moviam de forma semelhante antes da política?
- **testes placebo:** aparece “efeito” onde não deveria haver efeito?
- **robustez:** os resultados mudam muito quando alteramos especificações razoáveis?

Esses exercícios não provam a causalidade, mas ajudam a avaliar se a história causal é convincente.

# Como Ler uma Afirmação Causal

Quando alguém diz “X causa Y”, pergunte:

1. Qual é o tratamento?
2. Qual é o resultado?
3. Qual é o contrafactual?
4. Quem é comparado com quem?
5. O que cria variação no tratamento?
6. Qual é a principal ameaça à identificação?
7. Existe evidência sobre o mecanismo, ou apenas sobre o efeito total?
8. A estimativa responde à pergunta de política em questão?

## **Manchete:**

*Estudo mostra que pessoas que bebem café vivem mais.*

## **O que um economista deveria perguntar?**

- Quem bebe café?
- Comparado a quem?
- Qual é o contrafactual?
- O consumo de café foi atribuído aleatoriamente?
- Renda, saúde, tipo de trabalho ou estilo de vida poderiam explicar o resultado?
- Qual é a estratégia de identificação?
- O estudo identifica um efeito causal ou apenas uma associação?
- Há evidência sobre o mecanismo?

**Economistas não perguntam apenas se X e Y se movem juntos.**

Eles perguntam:

*Que comparação revelaria o que teria acontecido sem o tratamento?*

**Inferência causal é a busca por um contrafactual convincente.**

# Leituras Complementares

---

- Angrist, J. D. and Pischke, J.-S. (2015). *Mastering 'Metrics*. Princeton University Press.
- Angrist, J. D. and Pischke, J.-S. (2009). *Mostly Harmless Econometrics*. Princeton University Press.
- Cunningham, S. (2021). *Causal Inference: The Mixtape*. Yale University Press.
- Holland, P. W. (1986). "Statistics and Causal Inference". *Journal of the American Statistical Association*.
- Imbens, G. W. and Rubin, D. B. (2015). *Causal Inference for Statistics, Social, and Biomedical Sciences*. Cambridge University Press.

## Exemplos Aplicados: RCTs e Sorteios

- Finkelstein, A. et al. (2012). “The Oregon Health Insurance Experiment: Evidence from the First Year”. *Quarterly Journal of Economics*.
  - Exemplo de sorteio para estudar o efeito de acesso a seguro de saúde.
- Angrist, J. D., Pischke, D., Imbens, G. W., and Krueger, R. F. (2002). “Vouchers for Private Schooling in Colombia: Evidence from a Randomized Natural Experiment”. *American Economic Review*.
  - Exemplo de sorteio de vouchers escolares.
- Schultz, T. P. (2004). “School Subsidies for the Poor: Evaluating the Mexican Progresa Poverty Program”. *Journal of Development Economics*.
  - Exemplo de rollout aleatorizado em comunidades, com comparação antes/depois.
- Miguel, E. and Kremer, M. (2004). “Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities”. *Econometrica*.
  - Útil para discutir experimentos, efeitos indiretos e spillovers.

## Exemplos Aplicados: Diferenças-em-Diferenças

- Card, D. and Krueger, A. B. (1994). “Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania”. *American Economic Review*.
  - Exemplo clássico de comparação antes/depois entre uma região tratada e uma região de controle.
- Bertrand, M., Duflo, E. and Mullainathan, S. (2004). “How Much Should We Trust Differences-in-Differences Estimates?”. *Quarterly Journal of Economics*.
  - Texto importante sobre inferência e erros-padrão em DiD.
- Callaway, B. and Sant’Anna, P. H. C. (2021). “Difference-in-Differences with Multiple Time Periods”. *Journal of Econometrics*.
  - Referência moderna para DiD com vários períodos e adoção escalonada.

## Exemplos Aplicados: RDD

- Lee, D. S. (2008). “Randomized Experiments from Non-Random Selection in U.S. House Elections”. *Journal of Econometrics*.
  - Usa eleições decididas por margem estreita como desenho de regressão descontínua.
- Angrist, J. D. and Pischke, J. K. (2009). “Using Maimonides’ Rule to Estimate the Effect of Class Size on Scholastic Achievement”. *Quarterly Journal of Economics*.
  - Exemplo IV/RDD-style usando uma regra institucional de tamanho máximo de turma.
- Black, S. E. (1999). “Do Better Schools Matter? Parental Valuation of Elementary Education”. *Quarterly Journal of Economics*.
  - Exemplo de comparação em torno de fronteiras escolares.

- Imbens, G. W. and Lemieux, T. (2008). “Regression Discontinuity Designs: A Guide to Practice”. *Journal of Econometrics*.
  - Guia metodológico para aplicar RDD.
- Lee, D. S. and Lemieux, T. (2010). “Regression Discontinuity Designs in Economics”. *Journal of Economic Literature*.
  - Revisão acessível e muito usada sobre RDD em economia.

## Exemplos Aplicados: Variáveis Instrumentais

- Angrist, J. D. and Krueger, A. B. (1991). “Does Compulsory School Attendance Affect Schooling and Earnings?”. *Quarterly Journal of Economics*.
  - Usa trimestre de nascimento como instrumento para escolaridade.
- Card, D. (1995). “Using Geographic Variation in College Proximity to Estimate the Return to Schooling”. In *Aspects of Labour Market Behaviour*.
  - Usa proximidade a universidades como fonte de variação em escolaridade.
- Imbens, G. W. and Angrist, J. D. (1994). “Identification and Estimation of Local Average Treatment Effects”. *Econometrica*.
  - Referência central para interpretar IV como efeito local.

- Chavez, C. (2026). "On Causality". Manuscrito / material de aula.
  - Texto útil para discutir correlação, causalidade, contrafactuais e a história da inferência causal.
- Chavez, C. (2026). "The Transportation Problem in Causality". Manuscrito / material de aula.
  - Útil para pensar em como evidência é transportada de um contexto para outro.
- Deaton, A. and Cartwright, N. (2018). "Understanding and Misunderstanding Randomized Controlled Trials". *Social Science & Medicine*.
  - Discussão crítica sobre RCTs, validade externa e interpretação de evidência.