

# Inference & Causality

Week 5 Session 9 • November 17, 2025

Lecturer: Narges Chinichian

IU University of Applied Sciences, Berlin



Unit 5

# Fallacies in Causal Reasoning

Mediation Fallacy, Collider Bias

## Recap from Unit 4

Backdoor paths help us remove confounding.

Frontdoor paths help with unobserved confounding.

**But adjusting blindly can create new bias.** This leads us to today's fallacies.

# Overview of Today

1	Mediation fallacy
2	Collider bias

# Why This Unit Matters

Causal inference is not only about **what to adjust for**.

More important is **what NOT to adjust for**.

Many failed studies result from conditioning on the wrong variable.

This unit gives you "anti-patterns" to avoid.

# Mediation Fallacy: Intuition

A **mediator** lies on the causal path:

$$X \rightarrow M \rightarrow Y$$

If you adjust for M, you **block part of the effect** of X on Y.

# Mediation Fallacy Example 1

Education → Income → Health

What happens if you control for income?



# Historical Example: Education → Income → Health

Early epidemiological studies often underestimated the profound impact of education on health outcomes. This occurred when researchers inadvertently controlled for income, which acts as a mediator rather than a confounder in the causal pathway. By adjusting for income, these studies effectively blocked the indirect pathway through which education influences health, leading to a biased and diminished estimation of education's true effect.

The causal structure posits: **Education → Income → Health**. Adjusting for income in this model removes a critical component of education's beneficial influence.

## Link & Phelan (1995)

Demonstrated how education provides access to crucial socioeconomic resources, influencing health outcomes indirectly through these advantages.

## Mirowsky (2017)

Provided extensive evidence that income functions as a key mediator, translating educational attainment into tangible health advantages.

## Cutler & Lleras-Muney (2006)

Their work highlighted that conditioning on income significantly diminishes, or "collapses," the observed effect of education on health, revealing the methodological error.

## Glymour, Avendano & Kawachi (2014)

Explicitly cautioned against adjusting for income when examining the education-health gradient, emphasizing its role as a principal mediator that should not be controlled for.



# Mediation Fallacy Example 2

Colonialism → Institutions → Development

What happens if you control for institutions?

Acemoglu & Robinson [2024 Nobel Economics Prize winners for their contribution in comparative studies of prosperity between nations] in their book "Why Nations Fail: The Origins of Power, Prosperity, and Poverty" (2012) argue that nearly every difference in long-run development can be traced to whether a society has **inclusive** or **extractive** institutions.



# Colonialism → Institutions → Development

## The Mediation Fallacy in Development Economics

Institution Type	Inclusive Institutions (Promote development)	Extractive Institutions (Create stagnation)	Historical Examples
Political	Broad political participation; competitive elections; checks and balances	Power concentrated in elite; autocracy; colonial governors with unchecked authority	UK after 1688; modern Botswana vs. North Korea; colonial Latin America
Economic	Secure property rights; open markets; independent courts; widespread access to finance	Forced labor (mita, encomienda); extractive taxes; monopolies; restricted entry	Japanese Meiji reforms; U.S. frontier; Belgian Congo rubber regime
Legal	Rule of law; independent judiciary; equal application of laws	Dual legal systems (colonizers vs. natives); arbitrary confiscation; corruption	Modern OECD states vs. colonial India; Tsarist Russia
Social	Universal education; gender equality; social mobility	Caste hierarchies; racial segregation; dynastic political control	Finland's education reforms vs. apartheid South Africa

# Interpretation: Colonialism → Institutions → Development

**Colonialism as a Foundation:** Colonial powers established specific types of institutions (e.g., extractive for resource exploitation or inclusive for settler self-governance) that profoundly shaped the administrative, economic, and legal frameworks of colonized regions.

**Long-Run Development Outcomes:** The path-dependent nature of these institutions means they largely dictate long-run development trajectories. It is these institutional structures, rather than inherent geographical or cultural factors, that are considered the primary drivers of subsequent economic success or stagnation.

**Institutional Impact on Incentives:** These foundational institutions, once established, determine the incentives for economic participation, investment, and political engagement. Extractive institutions disincentivize broad participation and innovation, while inclusive institutions foster them.

**The Mediation Fallacy:** Adjusting for institutions (M) in studies examining the link between colonialism (X) and development (Y) constitutes a **mediation fallacy**. Since colonialism's primary effect on development operates **through** the institutions it established, controlling for institutions in a regression analysis effectively "blocks" this main causal pathway, leading to a significant underestimation of colonialism's total impact on long-term development.

# When Does Mediation Fallacy Matter?

When estimating the total causal effect

When M is a "mechanism variable"

When adjusting was recommended  
"because it improves prediction" →  
remember: prediction  $\neq$  causation

## Transition: From Mediator to Collider

Mediators block causal paths when conditioned on.

But **colliders open non-causal paths** when conditioned on.

**Opposite effect but equally dangerous!**

## Collider Bias: Definition

A **collider** is a variable that is influenced by two other variables:

$$X \rightarrow Z \leftarrow Y$$

- ❏ Conditioning on Z makes X and Y **dependent**, even if originally independent.

Do you remember the hospitalization, air pollution, and smoking example?

# Collider Bias: Intuition

When you restrict data to a subgroup ( $Z = \text{anything}$ ):



the subgroup selection creates artificial correlations



spurious relationships emerge



causal interpretation breaks

# Classic Collider Example



Among people who got the job, skill and connections appear negatively correlated.

But in the population they are independent.



**Smoking → Hospital Admission ← Air Pollution**

Among hospital patients, smoking and air pollution appear correlated.

→ But only because you conditioned on "being in the hospital."



# Collider Bias Example 1: WWII Bomber Armor Problem

Allies had planes coming back from missions with bullet holes on them.

What strategy is the best to reinforce these planes?



# Collider Bias Example 1:

## WWII Bomber Armor Problem

- **The Problem**

Only surviving planes were in the data: planes hit in critical areas didn't return.

- **The Insight**

Armor should go where there are NO holes (because planes hit there didn't survive to be counted).

- **The Collider**

"Returned to base" (survival) was influenced by hit location AND random factors.

## Collider Bias Example 2: Success stories

To reach your dream, you have to leave everything behind... or have you?



## Collider Bias Example 2: Success Stories

- **The Popular Narrative:** "Quit your job. Take the risk. Follow your passion. That's how I made it."
- **Success as Collider:** An outcome influenced by risk-taking AND luck, timing, resources, and networks.
- **The Problem:** We only hear from people who succeeded, creating selection bias.
- **The Bias:** Conditioning on success creates misleading advice about the necessity of extreme risk-taking.

# Why Colliders Are Dangerous

1

We tend to control for "everything we observe"

2

Selection bias often hides unnoticed

3

Machine learning pipelines often introduce it unintentionally (feature selection)

## Caution: When Collider = Selection Variable

### Examples:

- "admitted to university"
- "passed the clinical exam"
- "people who responded to a survey"
- "people who got the job"

# Summary

1

Mediators block paths (don't adjust for them when estimating total effect)

2

Colliders open paths (don't adjust for them unless required for identification)

3

Proper adjustment requires understanding DAG structure, not intuition

# Group Activity: Find Your Own Examples

## 1 Work in Groups

Form small groups to discuss and brainstorm together.

## 3 Task 2: Collider Bias

Identify one real-world example of **collider bias**. Consider scenarios where conditioning on a collider (e.g., selection criteria, common outcome) opens a spurious path between two variables.

## 5 Present Your Findings

Each group will present their examples to the class, briefly explaining the context and why it illustrates mediation fallacy or collider bias.

## 2 Task 1: Mediation Fallacy

Identify one real-world example of the **mediation fallacy**. Think about situations where you might incorrectly adjust for a mediator, leading to a biased estimate of a total effect.

## 4 Time Allocation

You will have **10-15 minutes** for discussion and example identification within your groups.

## 6 Be Creative!

Draw examples from your own fields of study, interests, or everyday life. The more relevant and imaginative, the better!



# Next: Hands-On Python Practice



## Explore Python Notebooks

Two dedicated notebooks, available on GitHub, for today's session.

## Mediation Fallacy Notebook

Dive into code examples that illustrate the mediation fallacy (the health example).

## Collider Bias Notebook

Understand collider bias through interactive simulations (the success story example).

# Official Uni Evaluation Form:

Valid until 25.11.2025

