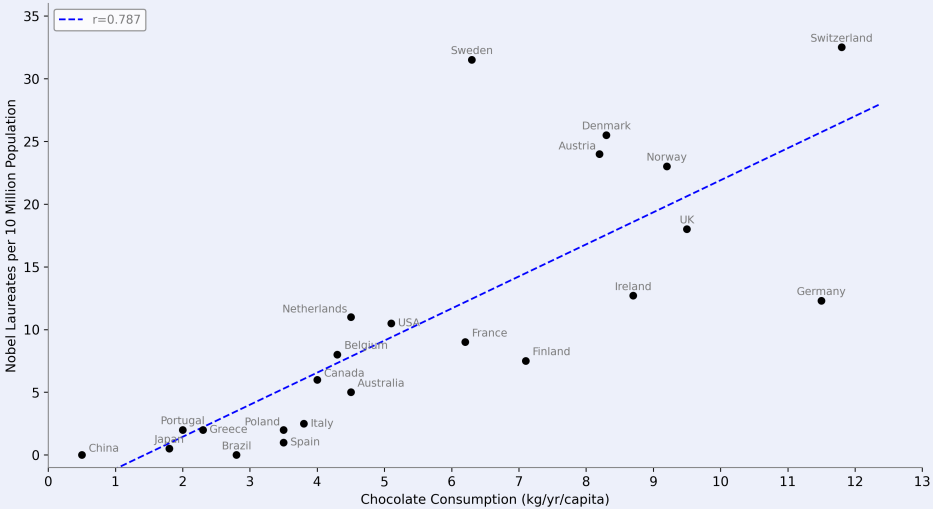
Revisiting the Big Three

- Sufficiency: All common causes are **measured**. (No doubts about parents!)
- Faithfulness: Graph reflects **all** data independencies. (WYSIWYG!)
- Markov: Data reflects **all** independencies in graph. (Graph predicts data!)

Is Chocolate the Secret Ingredient for Genius?

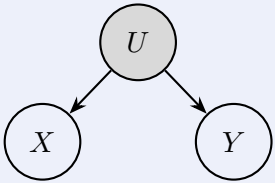


Messerli (2012): Chocolate Consumption, Cognitive Function, and Nobel Laureates

Three Causes of Scientific Sadness

Hidden Confounding

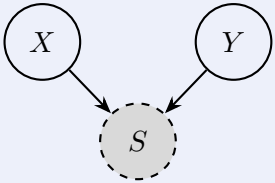
(Sufficiency Sabotage)



Illusory causation

Selection Bias

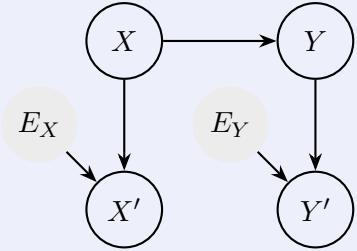
(Faithfulness Fail)



Sample $\neq$ Population

Measurement Error

(Markov Mayhem)

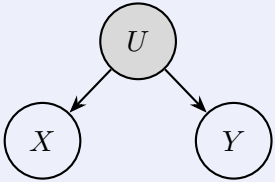


Fuzzy data, fuzzy truths

Three Causes of Scientific Sadness

Hidden Confounding

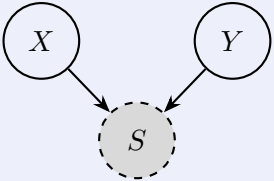
(Sufficiency Sabotage)



Illusory causation

Selection Bias

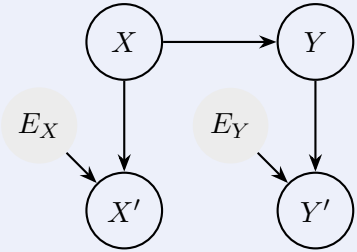
(Faithfulness Fail)



Sample $\neq$ Population

Measurement Error

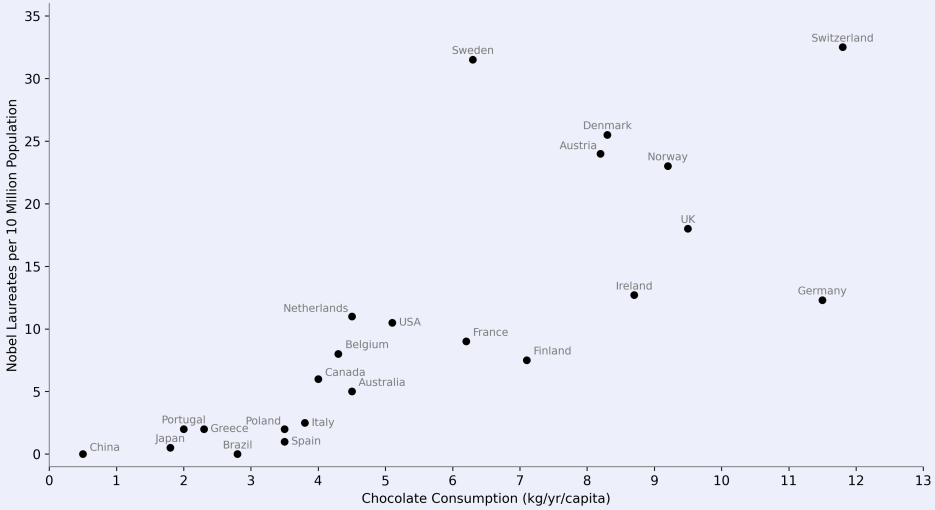
(Markov Mayhem)



Fuzzy data, fuzzy truths

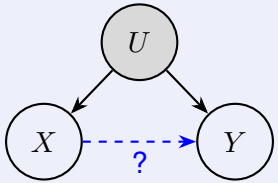
Hidden Confounding

Is Wealth the Hidden Hand?

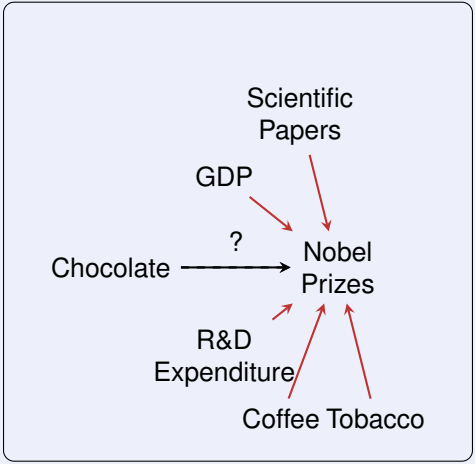


Hidden Confounding: The Unseen Puppeteer

- Unmeasured common cause (U).
- Violated **Causal Sufficiency**.
- "Correlation is not causation".
- Spurious associations found.
- Real associations masked.
- **Wrong** conclusions, **bad** policies.



The Obvious Fix: Just Control For Confounders

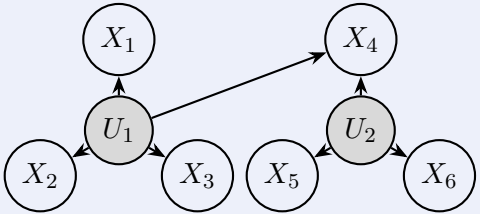


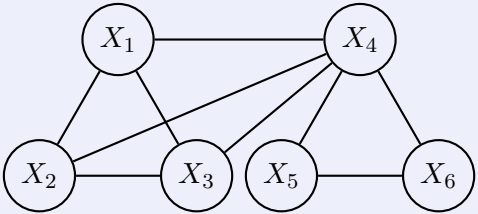
Simply Controlling Does Not Cut It

"Nobel laureates per capita is regressed on chocolate consumption per capita, as well as the number of published scientific papers, R&D expenditures per capita and GDP per capita. In addition, coffee and tobacco consumption per capita are also included. In contrast to the expectation, it is found that a positive correlation between chocolate consumption per capita and the stock of Nobel laureates per capita persists." (emphasis added)

Prinz (2020): Chocolate consumption and Nobel laureates

Graphical Patterns of Confounding





Parametric Patterns of Confounding

If X is generated by

$$X = b^\top U + \epsilon\,,$$

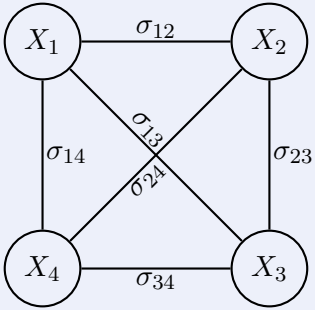
then the **six** covariances

$$\sigma_{ij} = b_i b_j\,,$$

lie on a **four**-dimensional manifold:

$$\sigma_{ij}\sigma_{uv} = \sigma_{iu}\sigma_{jv} = \sigma_{iv}\sigma_{ju} = b_i b_j b_u b_v\,.$$

I.e., the covariance matrix Σ is rank one.



Some Methods for Confounding

- **Compression Discovers Confounding**
  Causal Discovery with Hidden Confounders using the Algorithmic Markov Condition
- **Linear Non-Gaussians, ICA, and Confounding**
  A Versatile Causal Discovery Framework to Allow Causally-Related Hidden Variables
- **Hierarchical Confounding in Linear Non-Gaussians**
  Identification of Linear Non-Gaussian Latent Hierarchical Structure
- **Confounding in Nonlinear Models**
  Nonlinear Causal Discovery with Latent Confounders

RCTs: Not All That Glitters is a Gold Standard

Wait, but why didn't we just do a Randomized Controlled Trial (RCT?)

- **Feasible** RCT? Unlikely.
- **Sample**? Oompa-Loompas or representative humans?
- **Treatment**? What kind of chocolate? Administered how?

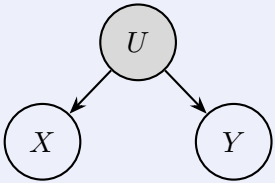
Precise estimate for one kind of chocolate, in one group, in one context.[1]

[1] In behavioral sciences: replication crisis. In CS: "Works on my machine."

Three Causes of Scientific Sadness

Hidden Confounding

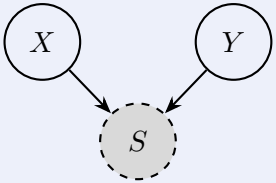
(Sufficiency Sabotage)



Illusory causation

Selection Bias

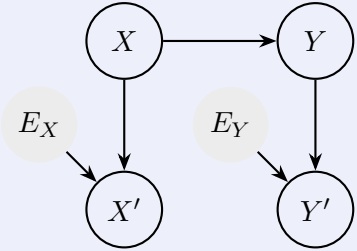
(Faithfulness Fail)



Sample $\neq$ Population

Measurement Error

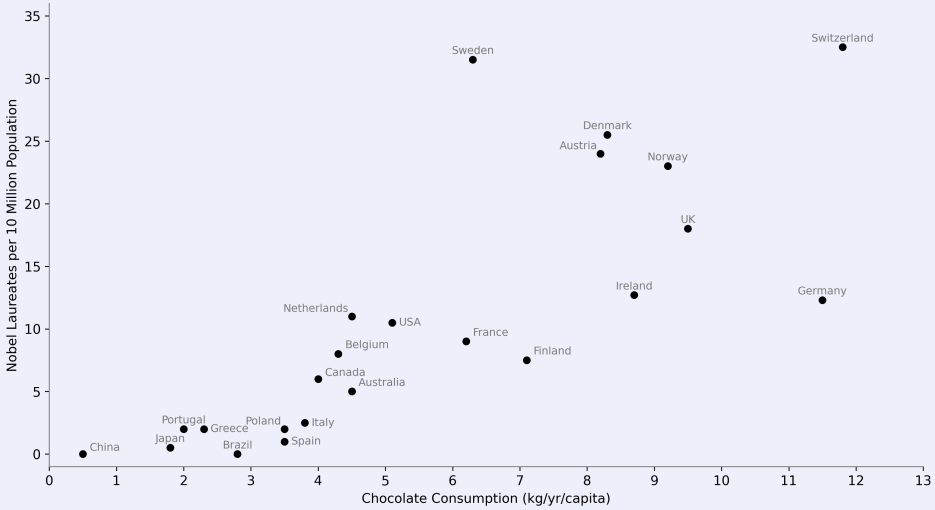
(Markov Mayhem)



Fuzzy data, fuzzy truths

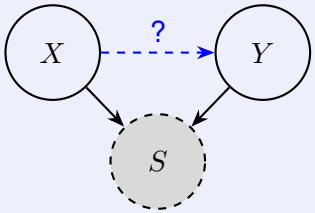
Selection Bias

How Did We Choose These Countries?

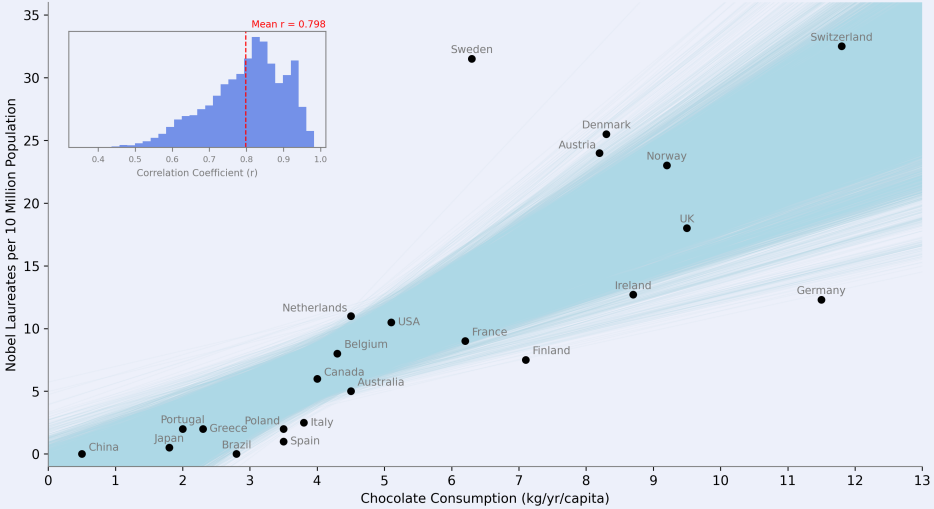


Selection Bias: The VIP Room Problem

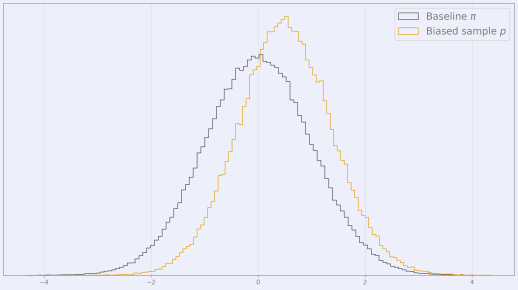
- Conditioning on collider S
- Violated **Faithfulness**
- "Absence of evidence is not evidence of absence".
- Mistaking sample quirks for causal laws.
- Spurious associations found.
- Real associations masked.
- Famous example: **Berkson's Paradox**.

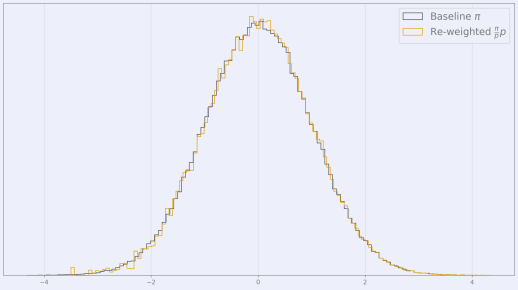


What if We Subsample Some Countries?



Propensity Weights

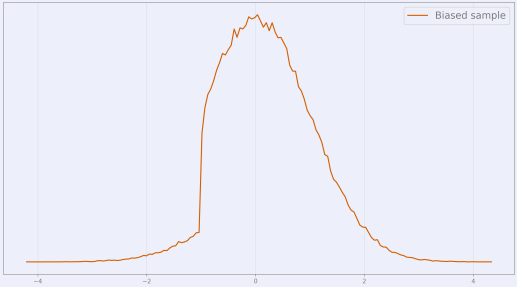


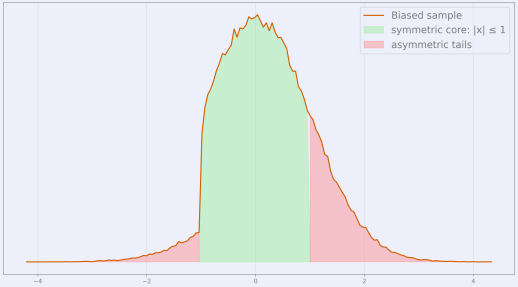


If we know the true distribution, we can use **propensity weights** to adjust models

$$\int h(x)\,\pi(x)\,dx \;=\; \int h(x)\,\frac{\pi(x)}{p(x)}\,p(x)\,dx$$

Boulevard of Broken Symmetries





If we don't know the distribution, we can still notice broken symmetries to adjust

$$\pi \approx p_{\text{sym}} = p(x) \cdot 1_{\{x>-1\}} + p(-x)1_{\{x\leq -1\}}$$

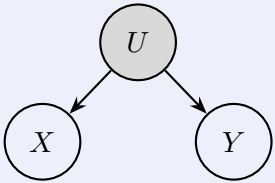
Selection Bias Methods

- **Kernel Mean Matching**
  Correcting Sample Selection Bias by Unlabeled Data
- **Importance Reweighting**
  Covariate Shift Adaptation by Importance Weighted Cross Validation
- **Selection Bias as Class of Distributions**
  A Review on Theory and Applications of Distributionally Robust Optimization
- **Bounding Causal Estimates**
  Nonparametric Bounds on Treatment Effects
- **Parametric and non-Parametric Single-Sample Detection**
  Identifying Selection Bias from Observational Data

Three Causes of Scientific Sadness

Hidden Confounding

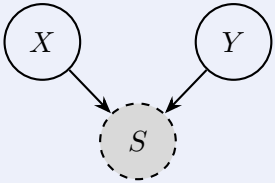
(Sufficiency Sabotage)



Illusory causation

Selection Bias

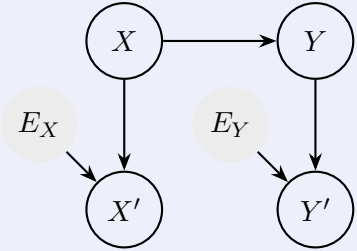
(Faithfulness Fail)



Sample $\neq$ Population

Measurement Error

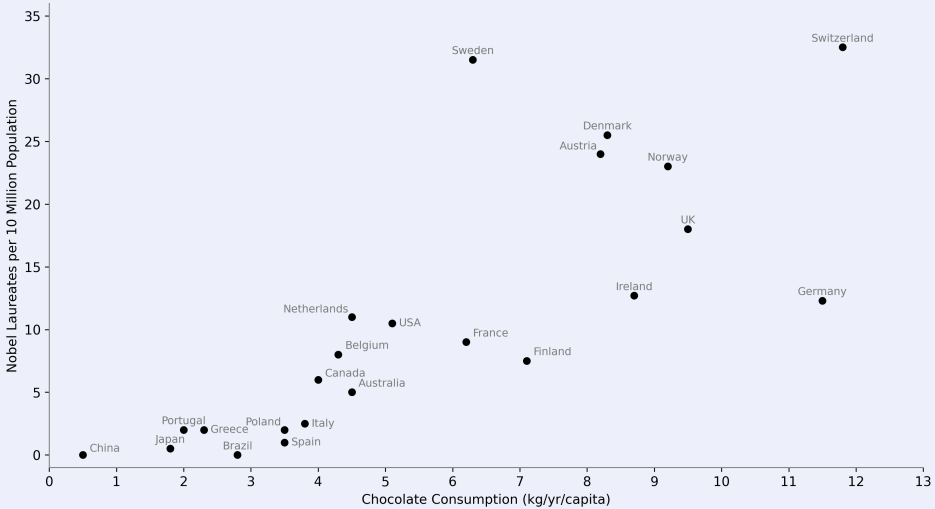
(Markov Mayhem)



Fuzzy data, fuzzy truths

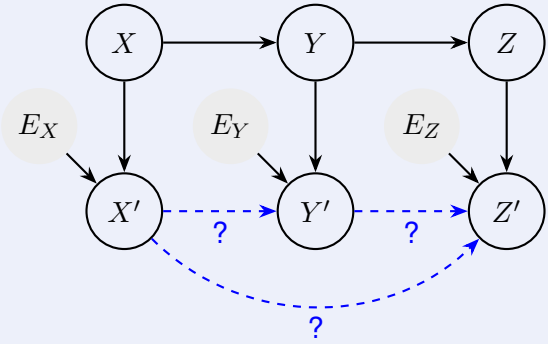
Measurement Error

How Much Can We Trust Those Measurements?

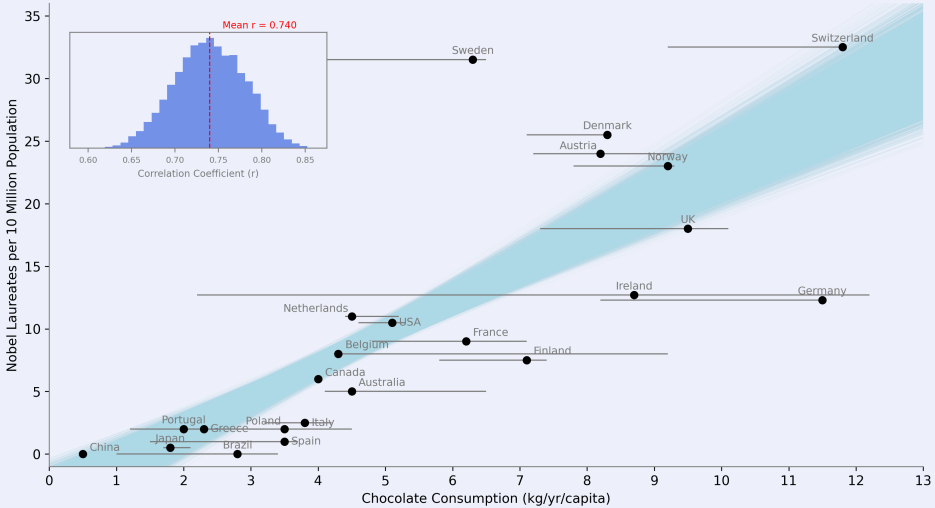


Measurement Error: The Funhouse Mirror

- Observe noisy $X' = f(X, E_X)$.
- Violated **Faithfulness** and **Markov Condition**.
- "The map is not the territory".
- Spurious links found.
- True effects masked.



Measurement Errors and Regression



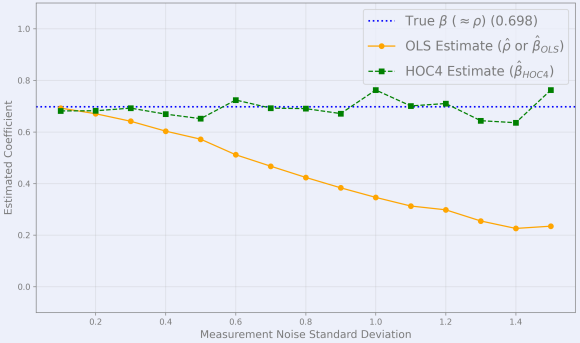
Higher Order Cumulants to the Rescue

If measurement error is Gaussian,

$$\hat{\beta}_{HOC4} = \frac{(n+1)m_{31} - 3(n-1)m_{20}m_{11}}{(n+1)m_{40} - 3(n-1)m_{20}^2}$$

$$m_{pq} = \frac{1}{n}\sum_{k=1}^{n}(x'_k - \bar{x}')^p(y'_k - \bar{y}')^q$$

is still consistent for the true β.



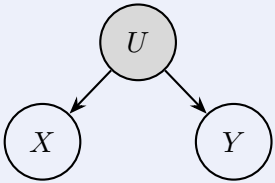
Measurement Error Methods

- **Adding Even More Noise**
  Simulation-Extrapolation Estimation in Parametric Measurement Error Models
- **Correcting Scores for Measurement Error**
  Conditional Scores and Optimal Scores for Generalized Linear Measurement-Error Models
- **Changing the Objective: Total Least Squares**
  Overview of total least squares methods
- **Independence Tests for Linear Non-Gaussian Models**
  Independence Testing-Based Approach to Causal Discovery under Measurement Error and Linear Non-Gaussian Models
- **Regression Calibration and Higher Order Cumulants**
  A comparison of regression calibration, moment reconstruction and imputation for adjusting for covariate measurement error in regression

Three Causes of Scientific Sadness

Hidden Confounding

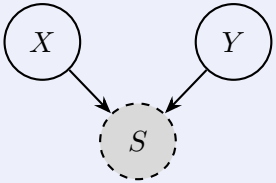
(Sufficiency Sabotage)



Illusory causation

Selection Bias

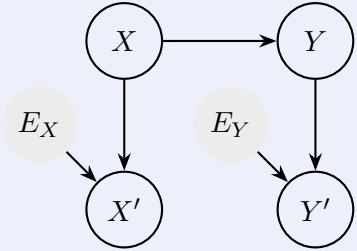
(Faithfulness Fail)



Sample $\neq$ Population

Measurement Error

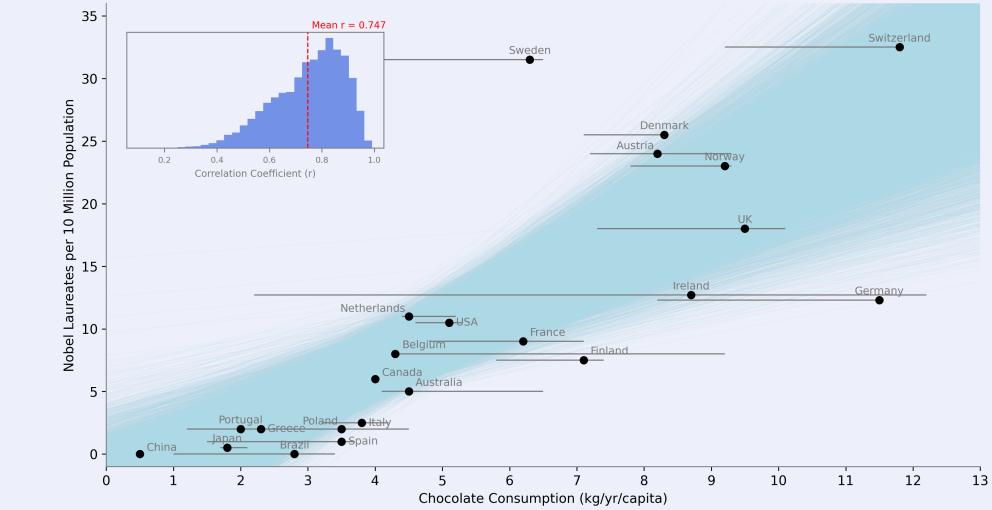
(Markov Mayhem)



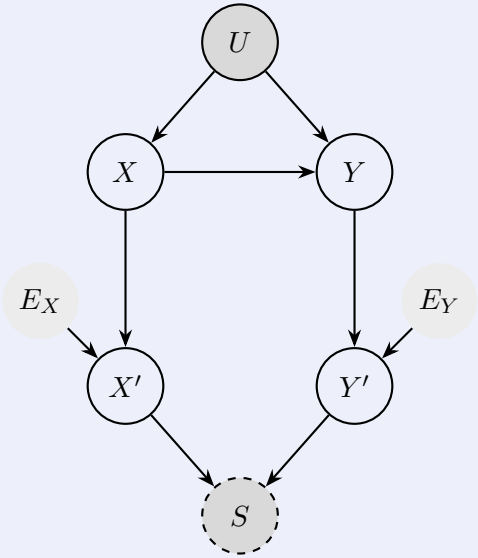
Fuzzy data, fuzzy truths

Cursed Combinations

Selection Meets Measurement Error



Lasciate ogne speranza, voi ch'intrate



What can we do when multiple biases are simultaneously present?

- Denial: Just ignore them.
- Anger: Blame the data.
- Bargain: Make a lot of assumptions.
- Depression: Abandon all hope.[a]
- Acceptance: Get better data.

[a] Most research on causality is here.

Towards Acceptance

Why did we run into problems in the first place?

- RCTs: costly, narrow, unethical, potential selection.
- Observational data: cheap, confounded, biased, unreliable.

What would it take to get better data?

- Naturally existing shifts emulate randomized interventions.
- Leverage existing heterogeneity: no new trials.
- No selection through admission criteria.
- Invariance filters spurious correlations.