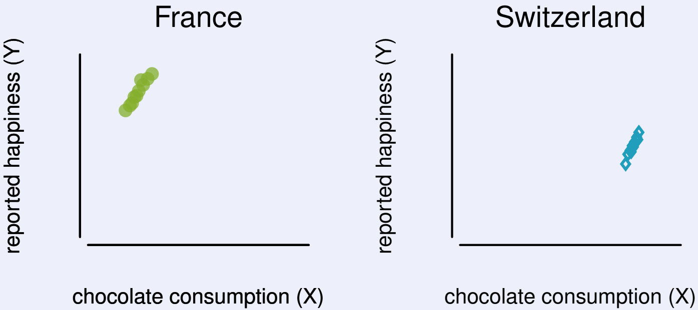
Cocoa Comfort across Contexts

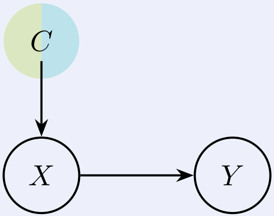


So far we assumed observations from a fixed context. But what if we obtain data from heterogeneous data sources?

Heterogeneous Data Sources

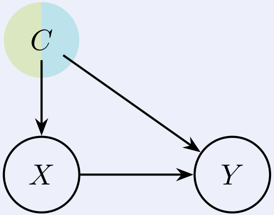
Roadmap: Tour du Chocolat

Invariance



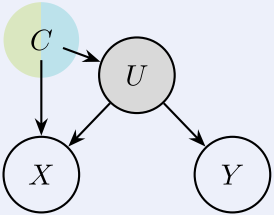
New worlds,
same rules

Independent Change



dependent samples,
independent mechanisms

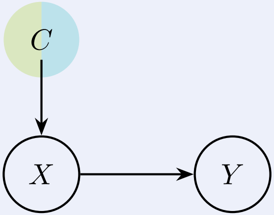
Contextual Confounding



dependent mechanisms?
... confounding!

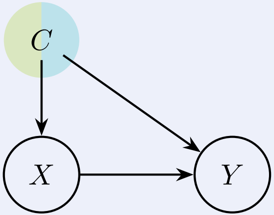
Roadmap: Tour du Chocolat

Invariance



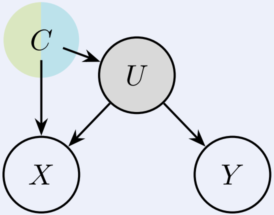
New worlds,
same rules

Independent Change



dependent samples,
independent mechanisms

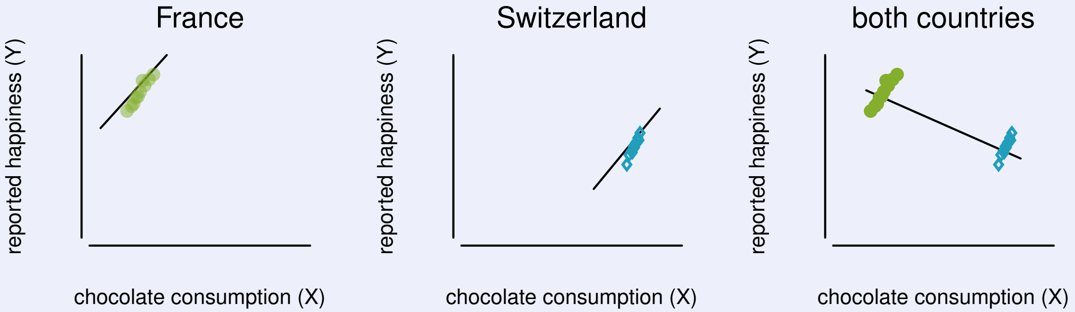
Contextual Confounding



dependent mechanisms?
... confounding!

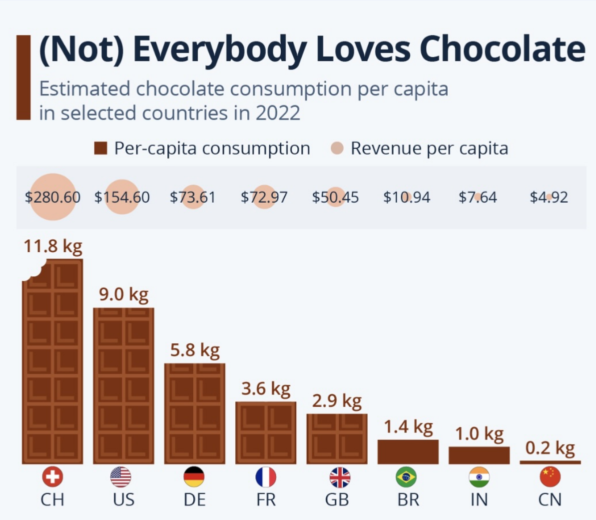
Causal Invariance

Sweet Swiss Paradox



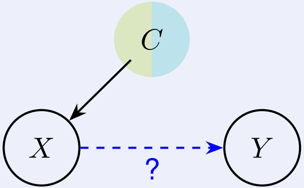
- **France, Switzerland**: More chocolate $\rightarrow$ improved happiness.
- Both countries: More chocolate $\rightarrow$ less happiness (?!)

Sweet Swiss Paradox, Explained



Multiple Contexts: A Puppeteer in Plain Sight

- Measured context C (countries, studies, or subpopulations).
- Here affects cause X (variation of chocolate per capita).
- Pooling data naively: averaging apples (●) and oranges (●).



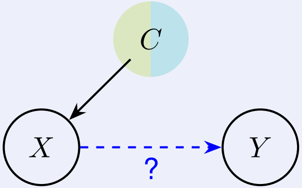
Causal Invariance

- Pooling contexts c: causal links may appear, vanish, or even reverse!
- But in each context c:

$$Y^{(c)} = \beta X^{(c)} + \varepsilon$$

- **Causal Invariance:**

$$\beta_{\textit{France}} \approx \beta_{\textit{Switzerland}}$$

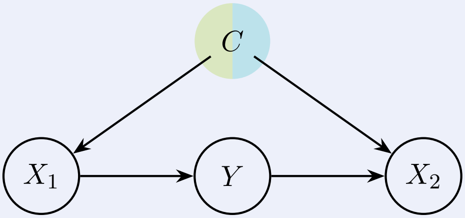


What Does Invariance Mean?

We want to find $\Phi(X)$ such that:

- $Y \perp\!\!\!\perp C \mid \Phi(X)$, and
- $\Phi(X)$ is informative about Y
- This is an optimization problem

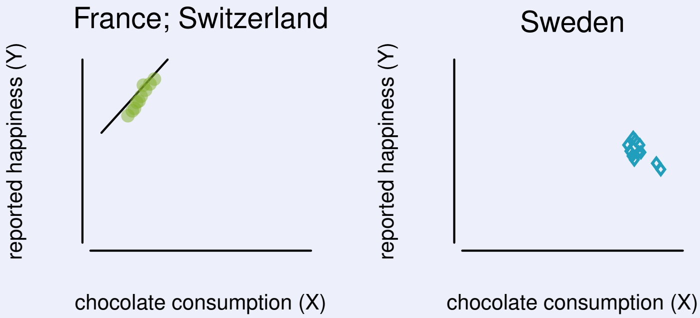
$$\min_{\Phi} \left(I(\Phi(X); Y) - \alpha I(C; Y \mid \Phi(X)) \right)$$



Invariance-Based Methods

- **Causal predictors are invariant**
  Causal Inference Using Invariant Prediction
- **Invariance as conditional independence of C**
  CD-NOD: Causal Discovery from Heterogeneous/Nonstationary Data
- **Invariance = optimality across environments**
  Invariant Risk Minimization
- **Generalization = Invariance**
  Invariant Models for Causal Transfer Learning

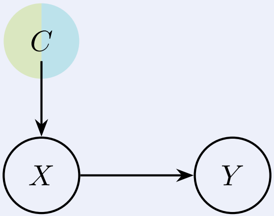
When Invariance Breaks ...



- So far, we assumed the effect of chocolate on happiness is stable.
- But what if $P(Y \mid X)$ changes across environments?

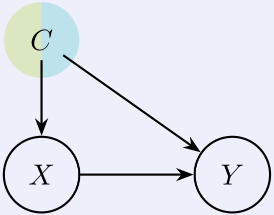
Three Layers of Distribution Shifts

Invariance



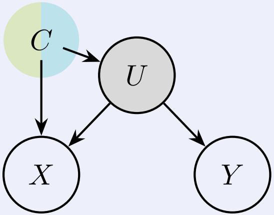
New worlds,
same rules

Independent Change



dependent samples,
independent mechanisms

Contextual Confounding

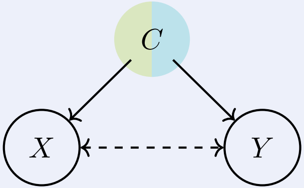


dependent mechanisms?
... confounding!

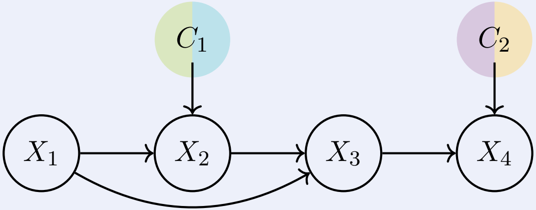
Independent Change

Beyond Invariance

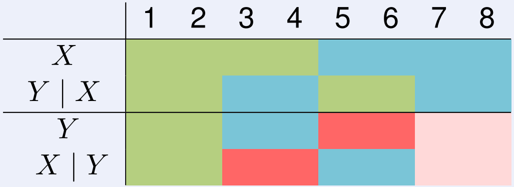
- Context C.
- Affecting both predictors X and target Y.
- **Modularity**: causal mechanisms change independently of one another.

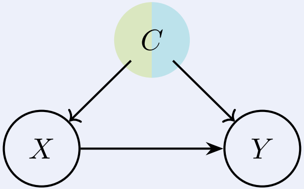


Independent Change: Nature's API Design



Independent Shifts = Sparse Shifts





The true causal graph entails fewer distinct mechanisms!

Independent Mechanism Shift Methods

- **Fewer Mechanism Shifts are Better**
  Causal Discovery in Heterogeneous Environments Under the Sparse Mechanism Shift Hypothesis
- **Compressing Multiple Mechanisms to Obtain Causal Graphs**
  Learning Causal Models under Independent Changes
- **Exogenous Variables Can Enforce Invariance**
  Anchor Regression: Heterogeneous Data Meet Causality
- **Neural Networks Can Leverage Independent Mechanism Shifts**
  Learning Robust Models Using The Principle of Independent Causal Mechanisms

Recall: Not all that Glitters is a Gold Standard

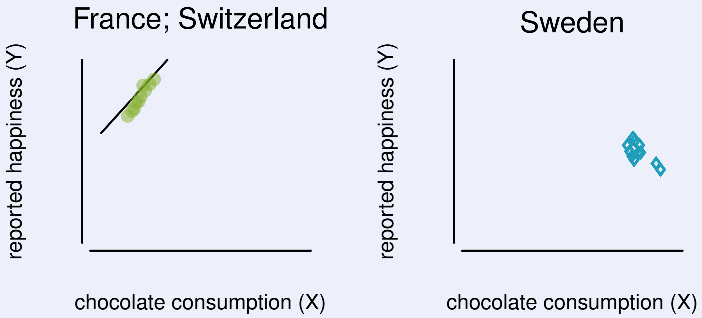
"Despite my efforts to keep the study groups blinded, some of the participants changed groups mid-study. Indeed, those in the control group (who received no extra chocolate) decided to start raiding the chocolate of those in the interventional arms (dark chocolate or milk chocolate). Furthermore, participants in the dark and milk chocolate groups traded chocolate based on their individual preferences." *(emphasis added)*

Chen 2007, A clinical trial gone awry: the Chocolate Happiness Undergoing More Pleasantness (CHUMP) study

The RCT We Didn’t Run

- In an ideal world, we randomize chocolate consumption (... forcing people to eat more chocolate? Keeping people from raiding chocolate?)
- In a nonideal world, we observe different countries with different habits. (... some countries naturally consume more chocolate than others!)
- We didn’t run an RCT — but the world kind of did.

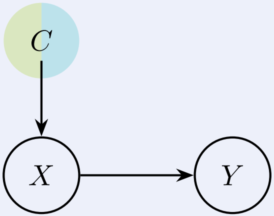
What else can we do with Nature's RCT?



- What if the changes of $P(X), P(Y \mid X)$ across environments are dependent?
- **Enter: Hidden Confounders.**

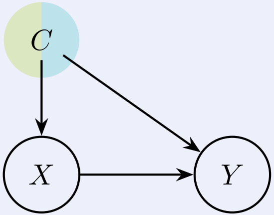
Three Layers of Distribution Shifts

Invariance



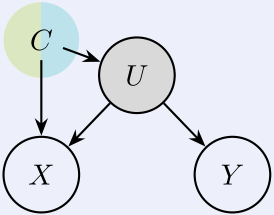
New worlds,
same rules

Independent Change



dependent samples,
independent mechanisms

Contextual Confounding



dependent mechanisms?
... confounding!

Contextual Confounders

Confounders in Disguise

- The context didn't just shift chocolate consumption — it also shifted something else:

$$\text{Wealth} \rightarrow \begin{cases} \text{more chocolate} \\ \text{higher happiness} \end{cases}$$

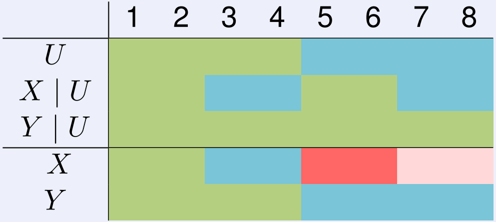
- Now changes in $P(X)$ and $P(Y \mid X)$ are no longer independent.

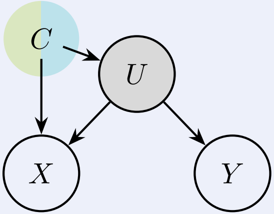
Confounders Revealed: Violations of ICM

- If both $P(X)$ and $P(Y \mid X)$ shift across environments...
- ...and they shift in coordinated ways...
- ... something is linking them behind the scenes.
- **Confounding implies correlated shifts.**

$$I_C\left(P^{(C)}(X),\ P^{(C)}(Y \mid X)\right) \neq 0 \quad \Rightarrow \text{ suspect a hidden confounder.}$$

Unmeasured Confounding Causal Correlated Mechanism Shifts





Context Confounding Compatible Approaches

- **Leveraging Invariance in Linear Models with Latent Confounders**
  Causal Dantzig: fast inference in linear structural equation models with hidden variables under additive interventions🔗
- **Running FCI on Data Including Environment Variables**
  Joint Causal Inference from Multiple Contexts🔗
- **Detecting Mechanism Shift Correlations for Confounder Detection**
  Identifying Confounding from Causal Mechanism Shifts🔗