Challenges in Causality: Sensitivity Analysis and Automation

Carlos Cinelli, University of Washington, Seattle

Challenges in Causality[1]

Carlos Cinelli, Avi Feller Guido Imbens, Edward Kennedy, Sara Magliacane, Jose Zubizarreta,

Benchmarks, Evaluation, and Validation

Complex Experiments and Modern Experimental Design

Interference and complex systems

Heterogeneous Effects & Policy Learning

Reliable and Scalable Causal Discovery

Aggregation and Synthesis of Causal Knowledge

Optimality

New Identification Strategies

Sensitivity Analyses

Automation

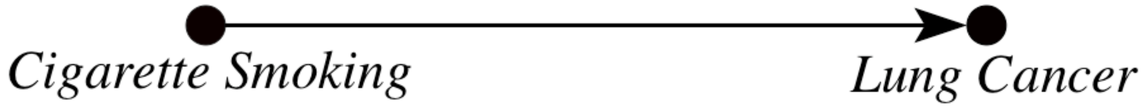
Motivating Example

Motivating example: smoking and cancer

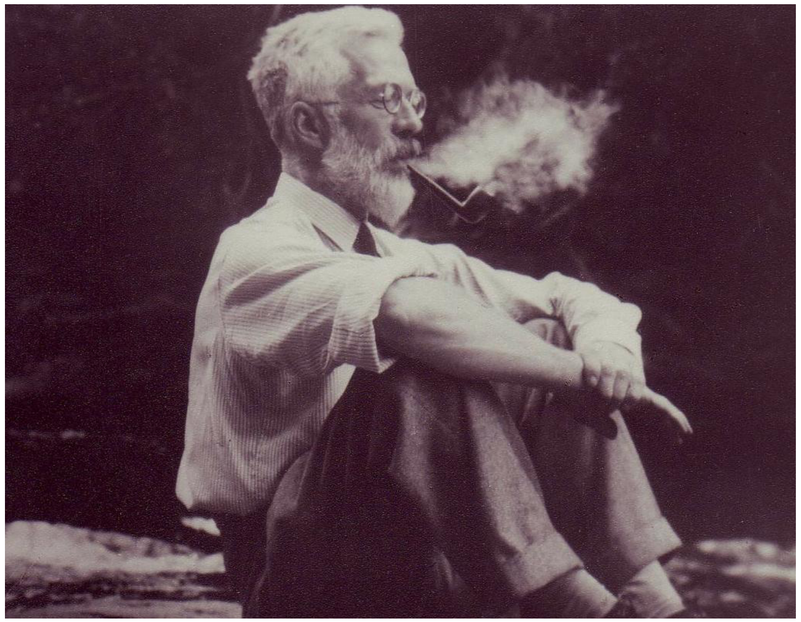
Let's start with a motivating example: the debate on cigarette smoking and lung cancer (50's/60's).

Strong association: smokers had 9 times the risk of nonsmokers to develop lung cancer (eg. Dorn, 1959).

Causal?



Not everyone agreed with this claim.



"For my part, I think it is more likely that a ***common cause*** supplies the explanation… The obvious common cause to think of is the ***genotype***"

- Ronald Fisher (1958)

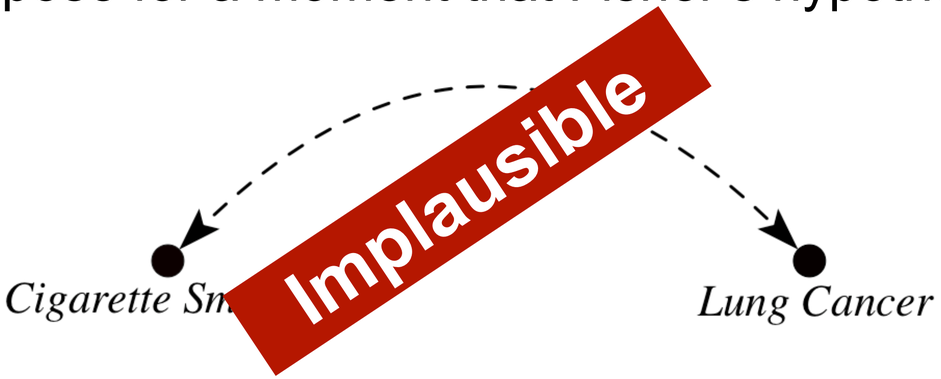
Observational data alone cannot distinguish both models.

No matter how big the data.

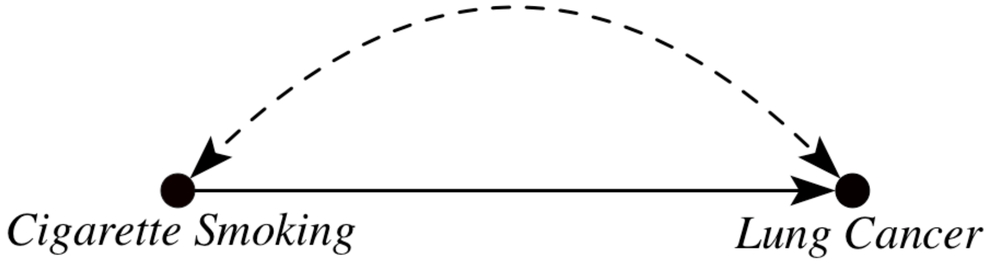
No matter how deep your NN.

Motivating example: smoking and cancer

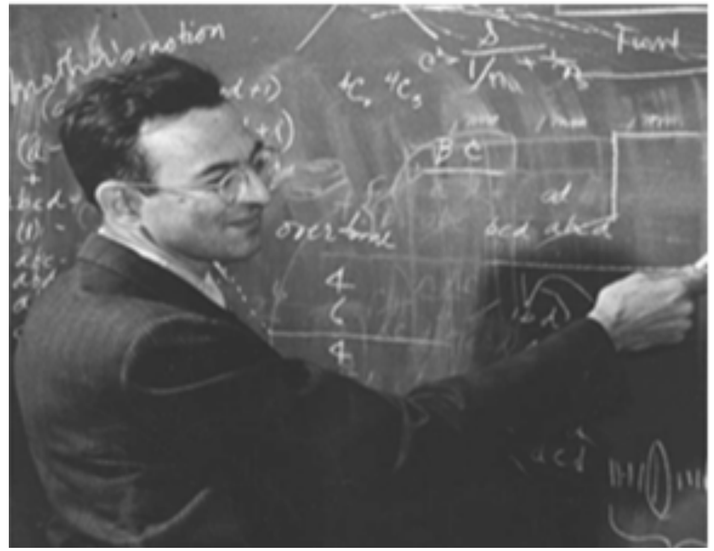
Let's suppose for a moment that Fisher's hypothesis were true.



How strong would unobserved confounding need to be to explain all the observed association?



Sensitivity analysis + plausibility judgments = there must be a causal path between cigarette smoking and lung cancer.



"…if cigarette smokers have 9 times the risk of nonsmokers for developing lung cancer, ***and this is not because cigarette smoke is a causal agent,*** …, then ***the proportion of hormone-X-producers among cigarette smokers must be at least 9 times greater than that of nonsmokers***"

- Cornfield et al (1959)

Why do we need sensitivity analyis (and automation)?

Most of traditional causal inference still relies on strong ***exact assumptions*** such as the ***absence of unobserved confounders, or the absence of certain direct effects.***

And the truth is that hardly anyone believes that those assumptions hold *exactly.*

We need tools that ***make it easy*** to ***routinely discuss*** the ***sensitivity*** of our estimates when our assumptions are called into question, as in the smoking and cancer debate.

Moreover, we need ***"automated Cornfields"*** – derivations such as those performed by Cornfield should not have to be done "by hand," for each new different question, model or assumption. They should be automatized.

Sensitivity Analyses

Sensitivity today: it is getting better

Although often praised, sensitivity analysis is still rarely practiced. However, we do see a recent uptake in many disciplines. What is changing? Here's a partial list of challenges that are gradually being addressed:

- **Challenge**: strong parametric and distributional assumptions about the unobserved confounders;
  - **Recent methods** impose no extraneous parametric assumptions on confounders.
- **Challenge:** lack of simple, interpretable sensitivity measures users can readily apply and routinely report;
  - **Recent methods** derive simple measures to summarize the robustness of an estimate to systematic biases, such as the the E-Value (Vanderweele and Ding) or the robustness value (Cinelli and Hazlett).
- **Challenge:** difficulty in connecting formal results to a cogent argument about which confounders are plausible;
  - **Recent methods** provide more interpretable sensitivity parameters (e.g. maximum explanatory power of confounders) and formal benchmarking exercises, comparing observed with unobserved confounders.
- **Challenge:** methods restricted to binary treatments, binary outcomes, or a specific estimands.
  - **Recent methods** are fully non-parametric, cover a broad class of estimands, and allow for flexible estimation with machine learning.

Omitted Variable Bias Approach to Sensitivity Analysis

Making Sense of Sensitivity: Extending Omitted Variable Bias

Carlos Cinelli * Chad Hazlett†

An Omitted Variable Bias Framework for Sensitivity Analysis of Instrumental Variables

Carlos Cinelli* Chad Hazlett†

LONG STORY SHORT:
OMITTED VARIABLE BIAS IN CAUSAL MACHINE LEARNING

VICTOR CHERNOZHUKOV†, CARLOS CINELLI*, WHITNEY NEWEY‡, AMIT SHARMA‖, AND VASILIS SYRGKANIS§

Long Story Short: Omitted Variable Bias in Causal ML

We provide a ***general nonparametric theory of omitted variable bias*** for a ***broad class of causal parameters:***

- average potential outcomes;

- average treatment effects (e.g, ATE/ATT/ATU);

- average causal derivatives (e.g. continous treatments).

- average effects from transporting covariates;

- average effects from distributional changes.

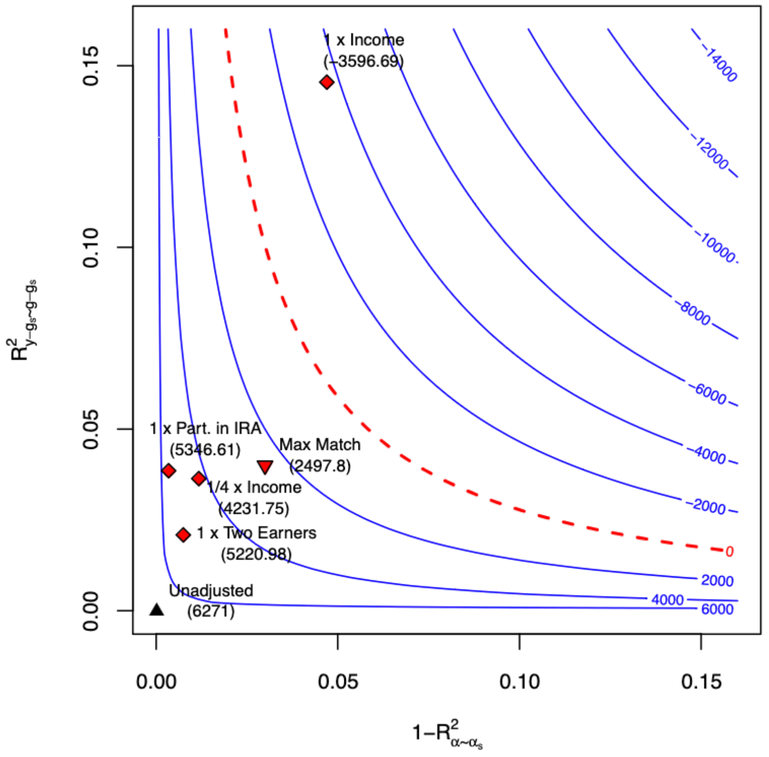
Bounds depend on simple plausibility judgments on the ***maximum explanatory power*** of latent variables (R2's).

We provide ***robustness statistics*** for routine reporting.

We derive ***relative bounds*** on the strength of confounder if they were ***as strong as observed covariates.***

Flexible statistical inference using ***debiased machine learning.***

	Results Under Conditional Ignorability			Robustness Values
Model	Short Estimate	Std. Error	Confidence Bounds	$RV_{\theta=0,\ a=0.05}$
Partially Linear	9,002	1,394	[6,271; 11,733]	5.4%
Nonparametric	7,949	1,245	[5,509; 10,388]	4.5%



(A) Lower limit confidence bound ($|\rho| = 1$).

Open problems

Goal: make sensitivity analyisis routine and standard practice.

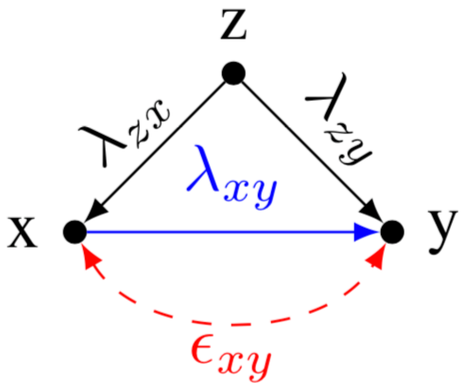
- Theory, methods and easy to use software to perform sensitivity analysis for all common study designs.
  - Eg: IV, DID, RDD, Synthetic Controls.
- Theory, methods and easy to use software for handling all common types of biases, simulatenously:
  - Selection Bias;
  - Missing Data;
  - Measurement error;
  - Cross-Population Bias;
  - Sensitivity to functional constraints (e.g. monotonicity);
- All the above should be acompanied with theory for modern estimation and inference.

While some of these have been solved for specific cases (e.g., specific estimands), we still lack general, easy to use, broadly applicable results.

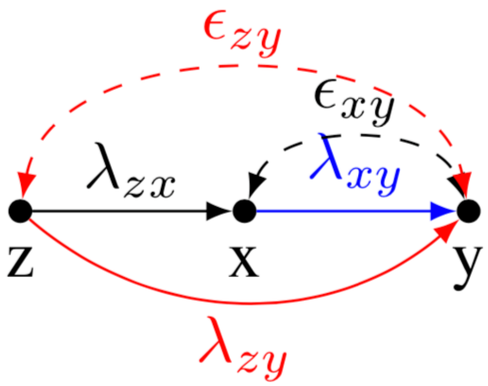
Automation (for sensitivity)

Current sensitivity analysis literature

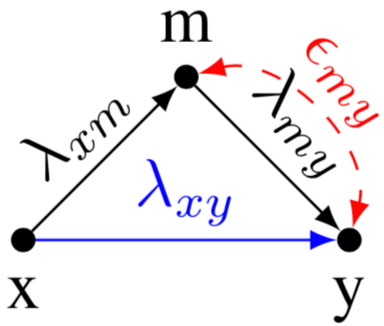
Limited to ***specific model structures, solved on a case-by-case basis;***



(a) Backdoor



(b) IV



(c) Mediation

Sensitivity analyses results for canonical models, as we have seen, are very useful.

But moving forward we need to address the essence of the problem in a more general way.

This calls for a flexible, systematic approach to incorporate credible and realistic constraints on causal models. Derivations of partial identification results or sensitivity curves should be performed automatically.

Algorithmic Tools for Sensitivity Analysis

Efficient Identification in Linear Structural Causal Models with Auxiliary Cutsets

Daniel Kumor [1] **Carlos Cinelli** [2] **Elias Bareinboim** [3]

Exploiting equality constraints in causal inference

Chi Zhang[1] Carlos Cinelli[2] Bryant Chen[3] Judea Pearl[1]

Sensitivity Analysis of Linear Structural Causal Models

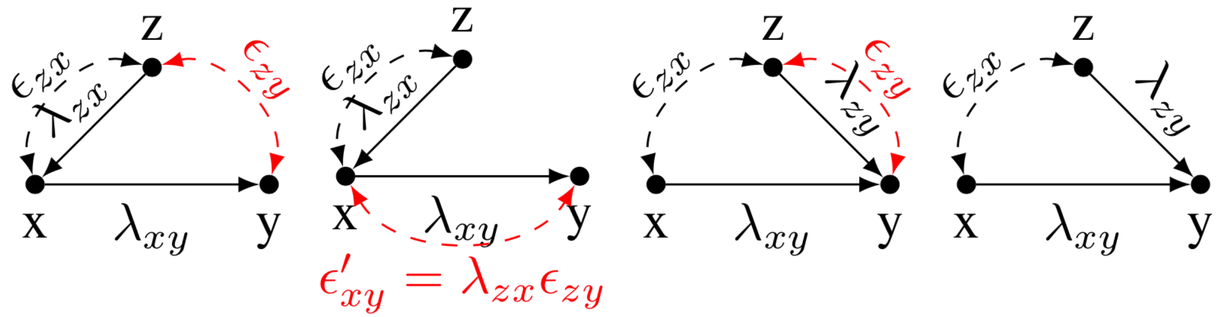
Carlos Cinelli [1] **Daniel Kumor** [2] **Bryant Chen** [3] **Judea Pearl** [1] **Elias Bareinboim** [2]

Example: linear structural equation models

Goal: systematic approach to sensitivity analysis for ***arbitrary*** linear structural equation models (SEM)

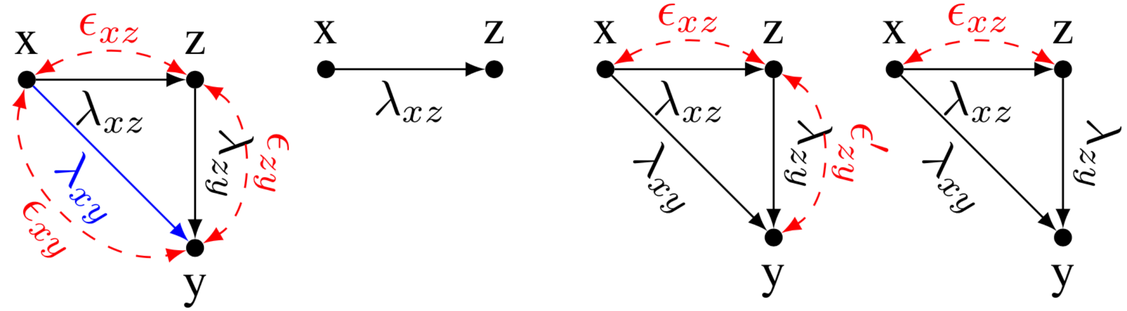
1. Formalize sensitivity analysis as **identification with non-zero constraints;**

2. Devise a novel **graphical procedure** (PushForward) to incorporate numerical constraints on bidirected edges;



(a) G_A (b) G'_A (c) G_B (d) G'_B

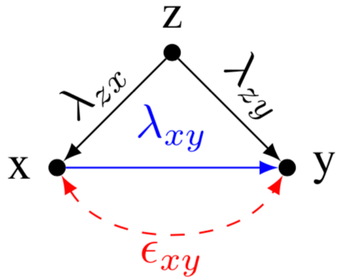
3. Develop an **efficient graph-based identification algorithm:**



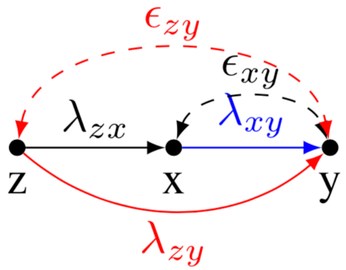
(a) G (b) G'_z (c) G'_y (d) G''_y

Example: linear structural equation models

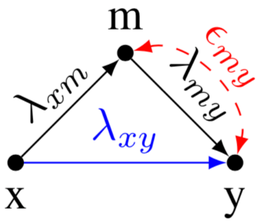
The algorithm captures all canonical cases.



(a) Backdoor ✔

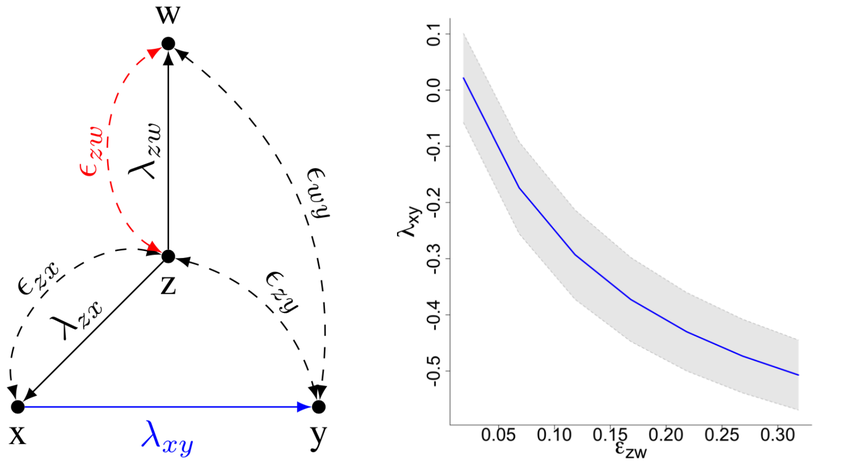


(b) IV ✔



(c) Mediation ✔

These are a **small subset** of all possible models.



(a) λ_{xy} is ϵ_{zw}-identifiable (b) Sensitivity of λ_{xy} in terms of ϵ_{zw}

Example: here, you can use bounds on the strength of confounding between *Z* and *W* to bound the causal effect of *X* and *Y.*

An algorithmic approach frees the researcher to model what they know, and to choose a sensitivity parameter according to the available expert knowledge.

Open problems

Goal: solve causal inference problems at scale. Automatize (partial) identification.

- We still do not have general algorithms for partial identification and bounds in arbitrary non-parametric DAGs;
- In certain cases, we actually do have potentially complete solutions to partial identification and sensitivity using tools from computer algebra or optimization:
  - Linear SEMs, use Groebner Basis; Discrete systems, use polynomial programming. These algorithms are inneficient. Perhaps, we still should try to scale them up?
- Many constraints not implemented by (partial) identification algorithms:
  - Inequality constraints; shape constraints.
- Knowledge Representation and Elicitation:
  - It remains difficult to elicit and represent causal knowledge. Graphical models have been extremely helpful, but many open problems remain: e.g. variable importance?
- Software for automatic identification and estimation is still a major bottleneck.

Thank you!

carloscinelli.com