# Sensitivity analysis via stochastic programming

Qingyuan Zhao

Statistical Laboratory, University of Cambridge

December 20, 2022 @ ICSDS

# Sensitivity analysis

Sensitivity analysis is widely used in any area that uses mathematical models.

## The broader concept (Saltelli et al. 2004)

- "The study of how **the uncertainty in the output** of a mathematical model or system can be apportioned to different sources of **uncertainty in its inputs**".
- Model inputs may be any factor that "can be changed in a model prior to its execution", including **"structural and epistemic sources of uncertainty"**.

## In observational studies

- Making causal inference from observational studies require **untestable assumptions**.
- So a typical question is:

  *How do the qualitative and/or quantitative conclusions if the identification assumptions (e.g. no unmeasured confounding) are violated?*

# Sensitivity analysis

Sensitivity analysis is widely used in any area that uses mathematical models.

## The broader concept (Saltelli et al. 2004)

- "The study of how **the uncertainty in the output** of a mathematical model or system can be apportioned to different sources of **uncertainty in its inputs**".
- Model inputs may be any factor that "can be changed in a model prior to its execution", including **"structural and epistemic sources of uncertainty"**.

## In observational studies

- Making causal inference from observational studies require **untestable assumptions**.
- So a typical question is:

  *How do the qualitative and/or quantitative conclusions if the identification assumptions (e.g. no unmeasured confounding) are violated?*

# Get ready for a slippery ride...

# Three components of sensitivity analysis

1. **Model augmentation:** Extend the model used by primary analysis to allow for unmeasured confounding or other forms of violations.
2. **Statistical inference:** Vary the sensitivity parameter, estimate the causal effect, and control suitable statistical errors.
3. **Interpretation of the results:** Probe different "directions" of unmeasured confounding and "make sense" of the results.

Each step can be challenging!

# Example: Child soldiering

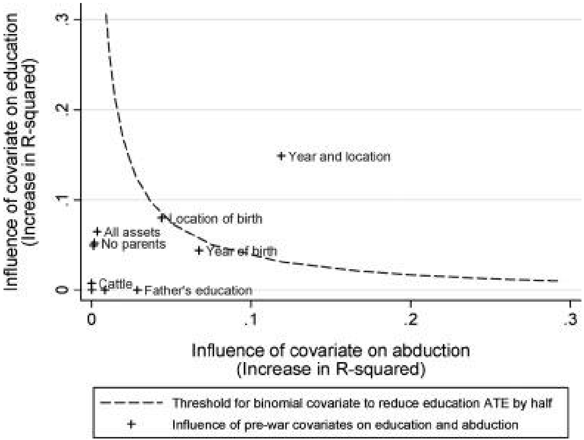
- About $60,000$ to $80,000$ youths were abducted in Uganda by a rebel force in 1995-2004.
- Question: What is the impact of child soldiering on the years of education?
- Blattman and Annan (2010) controlled for a variety of covariates $\boldsymbol{X}$ (age, household size, parental education, etc.) but were concerned about a unmeasured confounder $U$ (ability to hide from the rebel).
- They used the following model proposed by Imbens (2003):

$$
\begin{aligned}
&A \perp\!\!\!\perp Y(a) \mid \boldsymbol{X}, U, \text{ for } a = 0, 1, \\
&U \mid \boldsymbol{X} \sim \text{Bernoulli}(0.5), \\
&A \mid \boldsymbol{X}, U \sim \text{Bernoulli}(\text{expit}(\boldsymbol{\kappa}^T \boldsymbol{X} + \lambda U)), \\
&Y(a) \mid \boldsymbol{X}, U \sim \text{N}(\beta a + \boldsymbol{\nu}^T \boldsymbol{X} + \delta U, \sigma^2) \text{ for } a = 0, 1,
\end{aligned}
$$

- $A$ is the exposure and $Y(a)$ is the potential outcome.
- $(\lambda, \delta)$ are sensitivity parameters; $\lambda = \delta = 0$ corresponds to a primary analysis assuming no unmeasured confounding.

# Main results of Blattman and Annan (2010)

- Their primary analysis found that the ATE is -0.76 (s.e. 0.17).
- Sensitivity analysis was summarized by a single calibration plot:



# Some issues with the last analysis

Recall the model:

$$
\begin{aligned}
&A \perp\!\!\!\perp Y(a) \mid \boldsymbol{X}, U, \text{ for } a = 0, 1, \\
&U \mid \boldsymbol{X} \sim \text{Bernoulli}(0.5), \\
&A \mid \boldsymbol{X}, U \sim \text{Bernoulli}(\text{expit}(\boldsymbol{\kappa}^T \boldsymbol{X} + \lambda U)), \\
&Y(a) \mid \boldsymbol{X}, U \sim \text{N}(\beta a + \boldsymbol{\nu}^T \boldsymbol{X} + \delta U, \sigma^2) \text{ for } a = 0, 1,
\end{aligned}
$$

- Model augmentation: $(\lambda, \delta)$ are identifiable, so incoherent to vary them.
- Interpretation: In the calibration plot, partial $R^2$ for observed and unobserved confounders are not directly comparable because they use different reference models.
- Statistical inference: Cinelli and Hazlett (2020) proposed a nice solution to the last problem, but their heuristic sensitivity interval achieves no confidence guarantees.

# Rest of the talk

1. A general formulation of sensitivity analysis as a stochastic program.
2. Example 1: Linear models with no closed-form solutions.
3. Example 2: Nonparametric models with closed-form solutions.

# Sensitivity analysis as a stochastic program

- Consider an i.i.d. sample $(V_i, U_i) \sim \mathsf{P}_{V,U}$ but only $V_i$ is observed, $i = 1, \ldots, n$.
- In the last example, $V = (A, \boldsymbol{X}, Y)$ and $U = (Y(0), Y(1))$.
- Interested in a real-valued "causal parameter/functional" $\beta = \beta(\mathsf{P}_{V,U}) = \beta(\theta, \psi)$ where $\theta = \theta(\mathsf{P}_V)$ is identifiable and $\psi = \psi(\mathsf{P}_{V,U})$ is not.

## Sensitivity analysis

- A **sensitivity model** constrains the non-identifiable parameter: $\psi \in \Psi(\theta)$.
- This leads to a **partially identified region** for $\beta$, which can be bounded by solving

$$\begin{aligned} \text{minimize/maximize} \quad & \beta(\theta(\mathsf{P}_V), \psi) \\ \text{subject to} \quad & \psi \in \Psi(\theta(\mathsf{P}_V)). \end{aligned}$$

## Challenges

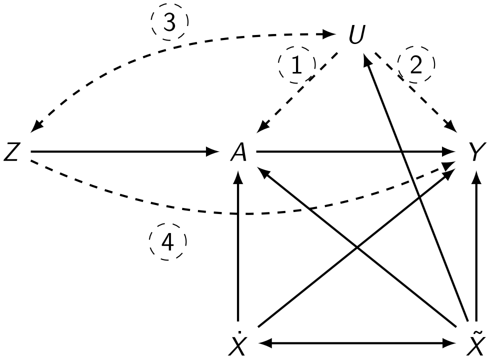
1. How to specify interpretable $\Psi$?
2. How to solve this optimization problem with a finite sample?
3. How to make inference for $\beta$ and its partially identified region?

# Example 1: Linear models with no closed-form solutions

Tobias Freidling and Qingyuan Zhao (2022). *Sensitivity Analysis With the $R^2$-calculus*. arXiv: 2301.00040 [stat.ME]

- Key idea: Specify an interpretable sensitivity model with the help of causal graphical models and the $R^2$-calculus.

# Causal graph




- Regression analysis assumes no (1) and (2).
- Instrumental variable analysis assumes no (3) and (4).

# Proposed method

- Objective function (Hosman, Hansen, and Holland 2010; Cinelli and Hazlett 2020)

$$\beta_{Y\sim A|X,Z,U} = \beta_{Y\sim A|X,Z} - R_{Y\sim U|X,Z,A}\, f_{A\sim U|X,Z} \frac{\sigma_{Y\sim X+Z+A}}{\sigma_{A\sim X+Z}}.$$

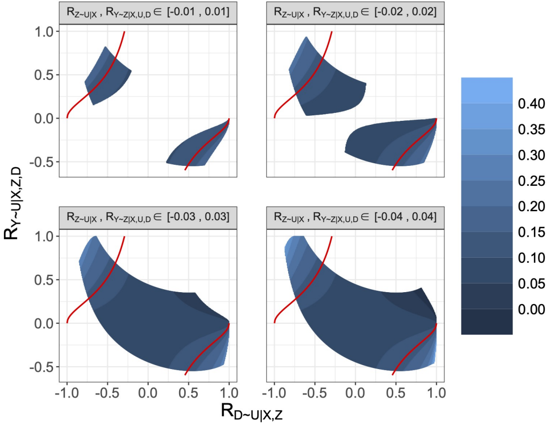
- Sensitivity models are formed by a mix-and-match of following constraints.

$$\begin{array}{lll}
① \; U \to A & 1. & R_{A\sim U|X,Z} \in [B^l_{UA}, B^u_{UA}] \\
 & 2. & R^2_{A\sim U|\tilde{X},\dot{X}_I,Z} \le b_{UA} R^2_{A\sim \dot{X}_J|\tilde{X},\dot{X}_I,Z} \\
② \; U \to Y & 1. & R_{Y\sim U|X,Z,A} \in [B^l_{UY}, B^u_{UY}] \\
 & 2. & R^2_{Y\sim U|\tilde{X},\dot{X}_I,Z} \le b_{UY} R^2_{Y\sim \dot{X}_J|\tilde{X},\dot{X}_I,Z} \\
 & 3. & R^2_{Y\sim U|\tilde{X},\dot{X}_I,Z,A} \le b_{UY} R^2_{Y\sim \dot{X}_J|\tilde{X},\dot{X}_I,Z,A} \\
③ \; U \leftrightarrow Z & 1. & R_{Z\sim U|X} \in [B^l_{UZ}, B^u_{UZ}] \\
 & 2. & R^2_{Z\sim U|\tilde{X},\dot{X}_{-j}} \le b_{UZ} R^2_{Z\sim \dot{X}_j|\tilde{X},\dot{X}_{-j}} \\
④ \; Z \to Y & 1. & R_{Y\sim Z|X,U,A} \in [B^l_{ZY}, B^u_{ZY}] \\
 & 2. & R^2_{Y\sim Z|X,U,A} \le b_{ZY} R^2_{Y\sim \dot{X}_j|\tilde{X},\dot{X}_{-j},Z,U,A}
\end{array}$$

- $R$ partial correlation, $f = R/\sqrt{1-R^2}$, $\beta$ least-squares coefficient, $\sigma$ residual variance.

# Stochastic optimization is not simple

- The optimization problem is generally nonlinear and non-convex (example below). A grid-search algorithm tailored for this problem was developed in the paper.




- Asymptotic theory (for bootstrap) requires further development.

# Example 2: Nonparametric models with closed-form solutions.

Yao Zhang and Qingyuan Zhao (2022). "Sharp Bounds and Semiparametric Inference in $L^\infty$- and $L^2$-sensitivity Analysis for Observational Studies". In: arXiv: 2211.04697 [stat.ME]


- Key idea: Identify variational optimization problems that admit closed-form solutions.

# Background: Marginal sensitivity model ($L^\infty$)

- Denote $\pi(X) = \mathrm{P}(A = 1 \mid X)$, $\pi(X, U) = \mathrm{P}(A = 1 \mid X, U)$, and $h(X, U) = \pi(X)/\pi(X, U)$.
- Tan (2006) proposed the following sensitivity model for unmeasured confounders:

$$\Gamma^{-1} \leq \frac{\pi(X)/\{1 - \pi(X)\}}{\pi(X, U)/\{1 - \pi(X, U)\}} \leq \Gamma, \quad \Gamma \geq 1.$$

- Zhao, Small, and Bhattacharya (2019) rediscovered it and called it "marginal sensitivity model", in relation to the celebrated Rosenbaum sensitivity model.
- For simplicity, let $U = Y(1)$. Bounds for $\beta(X) = \mathrm{E}\{Y(1) \mid X\}$ can be obtained by solving

$$\begin{aligned} \text{minimize/maximize} \quad & \mathrm{E}\{h(X, Y)Y \mid X, A = 1\} \\ \text{subject to} \quad & (1 - \Gamma^{-1})\pi(X) + \Gamma^{-1} \leq h(X, Y) \leq (1 - \Gamma)\pi(X) + \Gamma, \\ & \mathrm{E}\{h(X, Y) \mid X, A = 1\} = 1. \end{aligned}$$

- ZSB ignored the equality constraint and proposed a percentile bootstrap method.
- Dorn and Guo (2022) gave a closed-form solution to this linear stochastic program.

# Marginal sensitivity analysis ($L^2$)

- Key idea: One can also constrain the degree of unmeasured confounding by $\mathrm{Var}(h(X, U))$.
- This leads to a quadratic program:

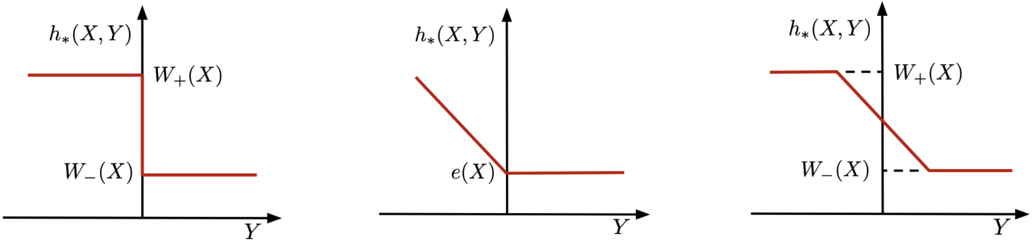
$$\begin{aligned} \text{minimize/maximize} \quad & \mathrm{E}\{h(X, Y)Y \mid X, A = 1\} \\ \text{subject to} \quad & \mathrm{E}\{h(X, Y)^2 \mid X, A = 1\} \leq \Gamma, \\ & \mathrm{E}\{h(X, Y) \mid X, A = 1\} = 1, \\ & h(X, Y) \geq \pi(X). \end{aligned}$$

- Closed-form solutions are found in the paper by considering the Lagrangian problem.
- The paper also derived the influence function for the bounds of the $L^\infty$- and $L^2$-sensitivity problems and discussed how to estimate them.[1]
- Calibration using observed confounders is much easier with the $L^2$ model than the $L^\infty$ model.

---

[1] An alternative form of the influence function for the $L^\infty$ problem was obtained by Dorn, Guo, and Kallus (2021).

# Schematic for the solution




(a) $L^\infty$ constraint (linear program).

(b) $L^2$ constraint (quadratic program).

(c) $L^\infty$ and $L^2$ constraints (quadratic program).

# Discussion

- Skipped literature review. Indeed massive literatures on sensitivity analysis and stochastic optimization, and several recent articles have made the connection more clear.
- A wide spectrum of problems between the two examples in this talk:
  1. Linear models with no closed-form solutions.
  2. Nonparametric models with closed-form solutions.
- In some sense this is a "wide open" field.

# References I


- Blattman, Christopher and Jeannie Annan (2010). "The Consequences of Child Soldiering". In: *The Review of Economics and Statistics* 92.4, pp. 882–898.
- Cinelli, Carlos and Chad Hazlett (2020). "Making sense of sensitivity: Extending omitted variable bias". In: *Journal of the Royal Statistical Society Series B: Statistical Methodology* 82.1, pp. 39–67.
- Dorn, Jacob and Kevin Guo (2022). "Sharp Sensitivity Analysis for Inverse Propensity Weighting Via Quantile Balancing". In: *Journal of the American Statistical Association*, in press. DOI: 10.1080/01621459.2022.2069572.
- Dorn, Jacob, Kevin Guo, and Nathan Kallus (2021). "Doubly-Valid/doubly-Sharp Sensitivity Analysis for Causal Inference With Unmeasured Confounding". In: arXiv: 2112.11449 [stat.ME].
- Freidling, Tobias and Qingyuan Zhao (2022). *Sensitivity Analysis With the $R^2$-calculus*. arXiv: 2301.00040 [stat.ME].
- Hosman, Carrie A., Ben B. Hansen, and Paul W. Holland (2010). "The sensitivity of linear regression coefficients' confidence limits to the omission of a confounder". In: *The Annals of Applied Statistics* 4.2. ISSN: 1932-6157. (Visited on 11/29/2021).

# References II


- Imbens, G. W. (2003). "Sensitivity to exogeneity assumptions in program evaluation". In: *American Economic Review* 93, pp. 126–132.
- Saltelli, Andrea et al. (2004). *Sensitivity analysis in practice: A guide to assessing scientific models*. John Wiley & Sons, Ltd.
- Tan, Zhiqiang (2006). "A distributional approach for causal inference using propensity scores". In: *Journal of the American Statistical Association* 101.476, pp. 1619–1637.
- Zhang, Yao and Qingyuan Zhao (2022). "Sharp Bounds and Semiparametric Inference in $L^\infty$- and $L^2$-sensitivity Analysis for Observational Studies". In: arXiv: `2211.04697 [stat.ME]`.
- Zhao, Qingyuan, Dylan S. Small, and Bhaswar B. Bhattacharya (2019). "Sensitivity Analysis for Inverse Probability Weighting Estimators Via the Percentile Bootstrap". In: *Journal of the Royal Statistical Society Series B: Statistical Methodology* 81.4, pp. 735–761. DOI: `10.1111/rssb.12327`.