Sensitivity analysis for principal ignorability violation in estimating complier and noncomplier average causal effects

Trang Nguyen

joint work with Liz Stuart, Dan Scharfstein, Betsy Ogburn

arXiv:2303.05032 | trang.nguyen@jhu.edu

CMStatistics 2023-12-17

Outline

Background

- Principal causal effects
- Principal ignorability (PI) – one identification strategy
- Sensitivity analysis for PI violation

Mean-based sensitivity analysis: expanding options

- Allow different sensitivity parameters
- Accommodate different estimation methods

Distribution-based sensitivity analysis: a new thing

- Method limitations and information use
- A method using full information (ongoing work)

The noncompliance problem

People may

- not take their pill
- not attend the training they are assigned to
- volunteer less than they are asked to

The noncompliance problem

People may

- not take their pill
- not attend the training they are assigned to
- volunteer less than they are asked to

Z	treatment assigned
Y	outcome
$Y(z)$	potential outcomes
X	baseline covariates
S	treatment received

The noncompliance problem

People may

- not take their pill
- not attend the training they are assigned to
- volunteer less than they are asked to

Might be intersted in

- the effect of receiving treatment
- but
- those who received $\neq$ those who did not

Z	treatment assigned
Y	outcome
$Y(z)$	potential outcomes
X	baseline covariates
S	treatment received

Principal stratification

People may

not take their pill

not attend the training they are assigned to

volunteer less than they are asked to

Might be intersted in

the effect of receiving treatment

but the groups are not comparable

Principal stratification (Frangakis & Rubin 2002)

avoids this problem

by creating a new pre-treatment variable based on potential treatment received

and stratifying on it

Z	treatment assigned
Y	outcome
$Y(z)$	potential outcomes
X	baseline covariates
S	treatment received
C	principal stratum, defined based on potential values $S(1), S(0)$ of S

Principal causal effects:

$$E[Y(1) - Y(0) \mid C = c]$$

Principal stratification

People may

not take their pill

not attend the training they are assigned to

volunteer less than they are asked to

Might be intersted in

the effect of receiving treatment

but the groups are not comparable

Principal stratification (Frangakis & Rubin 2002)

avoids this problem

by creating a new pre-treatment variable based on potential treatment received

and stratifying on it

Two-sided noncompliance: 4 principal strata

$S(1)$	$S(0)$	C
1	1	always taker
1	0	complier
0	1	defier
0	0	never taker

Z	treatment assigned
Y	outcome
$Y(z)$	potential outcomes
X	baseline covariates
S	treatment received
C	principal stratum, defined based on potential values $S(1), S(0)$ of S

Principal causal effects:

$$E[Y(1) - Y(0) \mid C = c]$$

Principal stratification

People may

not take their pill

not attend the training they are assigned to

volunteer less than they are asked to

Might be intersted in

the effect of receiving treatment

but the groups are not comparable

Principal stratification (Frangakis & Rubin 2002)

avoids this problem

by creating a new pre-treatment variable based on potential treatment received

and stratifying on it

One-sided noncompliance: 2 principal strata

$S(1)$	$S(0)$	C
1	0	complier
0	0	noncomplier

i.e., $C = S(1)$

Z	treatment assigned
Y	outcome
$Y(z)$	potential outcomes
X	baseline covariates
S	treatment received
C	principal stratum, defined based on potential values $S(1), S(0)$ of S

Principal causal effects:

$$\mathrm{E}[Y(1) - Y(0) \mid C = c]$$

Examples

Our focus is one-sided noncompliance

target: (non)complier average causal effects (CACE and NACE)

JOBS II for unemployed workers

- Z: week-long training on job search and mental health
- S: attending training

Experience Corps for the elderly

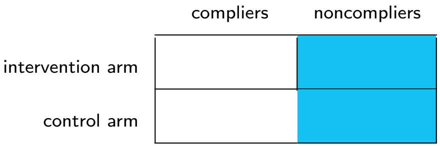
- Z: facilitated program for volunteering to help kids in school
- S: volunteering above a certain number of hours

Identification challenge: C is not observed under control

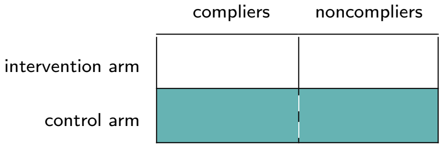
	compliers	noncompliers
intervention arm		
control arm		

Two major identification strategies

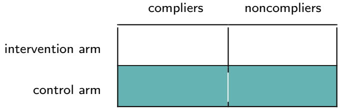
exclusion restriction/IV



principal ignorability



Principal ignorability (PI)



$$Y \perp\!\!\!\perp C \mid X, Z = 0$$

ie within covariate levels, compliers and noncompliers share the same outcome distribution under control

Principal ignorability (PI)

	compliers	noncompliers
intervention arm	$\mu_{11}(X)$	$\mu_{10}(X)$
control arm	$\mu_{01}(X)$	$\mu_{00}(X)$

$$Y \perp\!\!\!\perp C \mid X, Z = 0$$

or

$$\underbrace{\mathrm{E}[Y \mid X, Z = 0, C = 1]}_{\mu_{01}(X)} = \underbrace{\mathrm{E}[Y \mid X, Z = 0, C = 0]}_{\mu_{00}(X)}$$

Principal ignorability (PI)

	compliers	noncompliers
intervention arm	$\mu_{11}(X)$	$\mu_{10}(X)$
control arm	$\mu_{01}(X)$	$\mu_{00}(X)$

$$Y \perp\!\!\!\perp C \mid X, Z = 0$$

or

$$\underbrace{\mathrm{E}[Y \mid X, Z = 0, C = 1]}_{\mu_{01}(X)} = \underbrace{\mathrm{E}[Y \mid X, Z = 0, C = 0]}_{\mu_{00}(X)}$$

Combined with treatment assignment ignorability, PI identifies CACE, NACE:

$$\mathrm{E}[Y(1) - Y(0) \mid C = c] = \frac{\mathrm{E}\{[\mu_{1c}(X) - \kappa_0(X)]\pi_c(X)\}}{\mathrm{E}[\pi_c(X)]}$$

where

$$\kappa_0(X) := \mathrm{E}[Y \mid X, Z = 0] \quad \text{(mixture outcome mean)}$$
$$\pi_c(X) := \mathrm{P}(C = c \mid X, Z = 1) \quad \text{(principal score)}$$

Principal ignorability (PI)

	compliers	noncompliers
intervention arm	$\mu_{11}(X)$	$\mu_{10}(X)$
control arm	$\mu_{01}(X)$	$\mu_{00}(X)$

$$Y \perp\!\!\!\perp C \mid X, Z = 0$$

or

$$\underbrace{\mathrm{E}[Y \mid X, Z = 0, C = 1]}_{\mu_{01}(X)} = \underbrace{\mathrm{E}[Y \mid X, Z = 0, C = 0]}_{\mu_{00}(X)}$$

PI is untestable

Need sensitivity analyses

Prior sens analysis method that inspired this work

Ding and Lu (2017) use a mean ratio sensitivity parameter

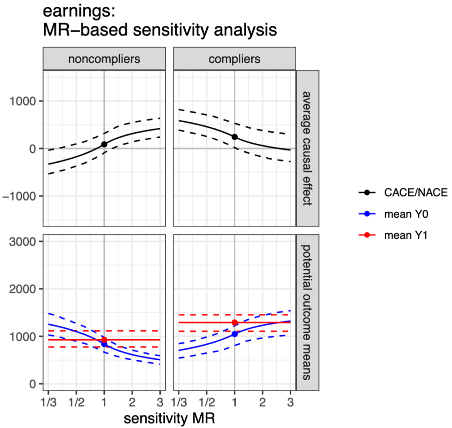
$$\frac{\mu_{01}(X)}{\mu_{00}(X)} = \rho$$

and modify a PI-based weighting estimator to incorporate ρ

See also Jiang, Yang and Ding (2022)

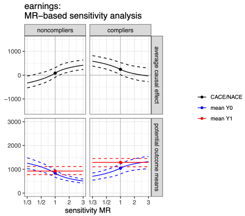
Example of MR-based sens analysis

earnings in JOBS II



Example of MR-based sens analysis

earnings in JOBS II



other outcomes for which MR param not ideal

- JOBS II: having a job (binary), depressive symptoms (bounded)
- Experience Corps: generativity (bounded)

Outline

Background

- Principal causal effects
- Principal ignorability (PI) – one identification strategy
- Sensitivity analysis for PI violation

Mean-based sensitivity analysis: expanding options

- Allow different sensitivity parameters
- Accommodate different estimation methods

Distribution-based sensitivity analysis: a new thing

- Method limitations and information use
- A method using full information (ongoing work)

A range of sens assumptions with different sens params

PI: $\mu_{01}(X) = \mu_{00}(X)$.

sens-MR: $$\frac{\mu_{01}(X)}{\mu_{00}(X)} = \rho,$$

sens-OR: $$\frac{\mu_{01}(X)/[1 - \mu_{01}(X)]}{\mu_{00}(X)/[1 - \mu_{00}(X)]} = \psi,$$

sens-GOR: $$\frac{[\mu_{01}(X) - l]/[h - \mu_{01}(X)]}{[\mu_{00}(X) - l]/[h - \mu_{00}(X)]} = \psi$$

where l and h are the outcome lower and upper bounds,

sens-SMD: $$\frac{\mu_{01}(X) - \mu_{00}(X)}{\sqrt{[\sigma^2_{01}(X) + \sigma^2_{00}(X)]/2}} = \eta$$

where $\sigma^2_{0c}(X) := \text{var}(Y \mid X, Z = 0, C = c)$,

for some range of ρ, ψ or η that is considered plausible.

Identification

sens-MR and sens-GOR result in point identification of CACE, NACE

because they help solve the mixture equation

$$\pi_1(X)\mu_{01}(X) + \pi_0(X)\mu_{00}(X) = \kappa_0(X).$$

Identification

sens-MR and sens-GOR result in point identification of CACE, NACE

because they help solve the mixture equation

$$\pi_1(X)\mu_{01}(X) + \pi_0(X)\mu_{00}(X) = \kappa_0(X).$$

sens-SMD obtains bounds for CACE, NACE

bounds are narrowed if also assume $1/k \leq \frac{\sigma_{01}^2(X)}{\sigma_{00}^2(X)} \leq k$ for some $k > 1$

and reduce to point identification if assume $\sigma_{01}^2(X) = \sigma_{00}^2(X)$ (aka sens-SMDe)

Identification

sens-MR and sens-GOR result in point identification of CACE, NACE

because they help solve the mixture equation

$$\pi_1(X)\mu_{01}(X) + \pi_0(X)\mu_{00}(X) = \kappa_0(X).$$

sens-SMD obtains bounds for CACE, NACE

bounds are narrowed if also assume $1/k \leq \frac{\sigma^2_{01}(X)}{\sigma^2_{00}(X)} \leq k$ for some $k > 1$

and reduce to point identification if assume $\sigma^2_{01}(X) = \sigma^2_{00}(X)$ (aka sens-SMDe)

in all cases, effect identification is via identification of $\mu_{0c}(X)$

by a function of sens param, $\pi_c(X)$, $\kappa_0(X)$ (and $\text{var}(Y \mid X, Z = 0)$ w/ sens-SMD)

Sens analysis techniques to go with 3 types of PI-based estimators

Sens analysis = a modification of main analysis

Sens analysis techniques to go with 3 types of PI-based estimators

Sens analysis = a modification of main analysis

- Type A ($\approx$ outcome regression estimators)
  - estimates $\kappa_0(X)$ to first estimate effects conditional on covariates and then aggregates them to estimate CACE/NACE, eg

$$\frac{\sum_{i=1}^n \hat{\pi}_c(X_i)[\hat{\mu}_{1c}(X_i) - \hat{\kappa}_0(X_i)]}{\sum_{i=1}^n \hat{\pi}_c(X_i)}, \quad \frac{\sum_{i=1}^n \frac{Z_i}{\hat{e}(X_i,Z_i)} \mathrm{I}(C_i = c)[Y_i - \hat{\kappa}_0(X_i)]}{\sum_{i=1}^n \frac{Z_i}{\hat{e}(X_i,Z_i)} \mathrm{I}(C_i = c)}$$

  - sens analysis technique: replace $\kappa_0(X)$ by the identification result of $\mu_{0c}(X)$ under the sens assumption

Sens analysis techniques to go with 3 types of PI-based estimators

Sens analysis = a modification of main analysis

- Type B (≈ influence function based estimators)
  - write the CACE/NACE as
    $$\frac{\nu_{1c} - \nu_{0c}^{\mathsf{PI}}}{\pi_c}$$
    where $\nu_{zc} := \mathsf{E}[\pi_c(X)\mu_{zc}(X)]$, $\nu_{0c}^{\mathsf{PI}} := \mathsf{E}[\pi_c(X)\kappa_0(X)]$, $\pi_c := \mathsf{E}[\pi_c(X)]$
  - a type B estimator can be expressed as combination of IF-based estimators of π_c, ν_{1c} and ν_{0c}^{PI}
    $$\frac{\hat{\nu}_{1c,\text{IF}} - \hat{\nu}_{0c,\text{IF}}^{\mathsf{PI}}}{\hat{\delta}_{c,\text{IF}}}$$
  - sens analysis technique: replace $\hat{\nu}_{0c,\text{IF}}^{\mathsf{PI}}$ with an IF-based estimator of ν_{0c} under the sens assumption

Sens analysis techniques to go with 3 types of PI-based estimators

Sens analysis = a modification of main analysis

- Type C (≈ other/weighting estimators)
  - an example is the pure weighting estimator

$$\frac{\sum_{i=1}^{n} \frac{Z_i}{\hat{e}(X_i, Z_i)} \mathrm{I}(C_i = c) Y_i}{\sum_{i=1}^{n} \frac{Z_i}{\hat{e}(X_i, Z_i)} \mathrm{I}(C_i = c)} - \frac{\sum_{i=1}^{n} \frac{1-Z_i}{\hat{e}(X_i, Z_i)} \hat{\pi}_c(X_i) Y_i}{\sum_{i=1}^{n} \frac{1-Z_i}{\hat{e}(X_i, Z_i)} \hat{\pi}_c(X_i)}$$

  - no general sens analysis technique

Sens analysis techniques to go with 3 types of PI-based estimators

Sens analysis = a modification of main analysis

- Type C ($\approx$ other/weighting estimators)
  - an example is the pure weighting estimator

$$\frac{\sum_{i=1}^{n} \frac{Z_i}{\hat{e}(X_i,Z_i)} \mathrm{I}(C_i = c) Y_i}{\sum_{i=1}^{n} \frac{Z_i}{\hat{e}(X_i,Z_i)} \mathrm{I}(C_i = c)} - \frac{\sum_{i=1}^{n} \frac{1-Z_i}{\hat{e}(X_i,Z_i)} \hat{\pi}_c(X_i) Y_i}{\sum_{i=1}^{n} \frac{1-Z_i}{\hat{e}(X_i,Z_i)} \hat{\pi}_c(X_i)}$$

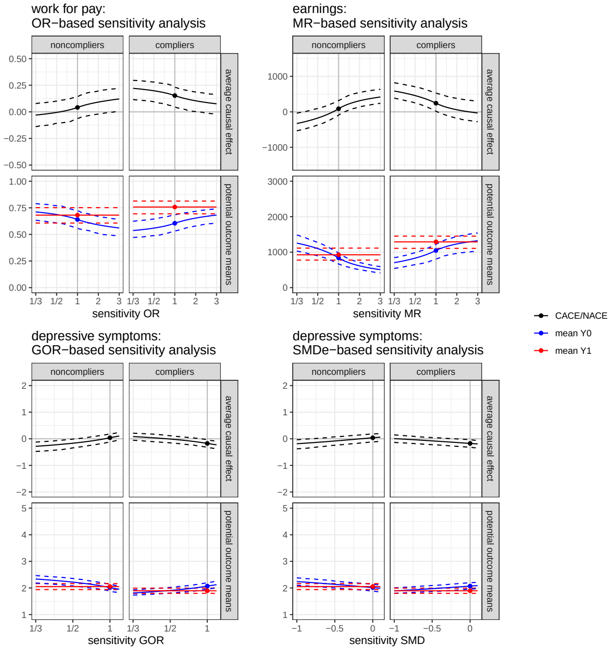
  - no general sens analysis technique
  - sens-MR: scale Y in control units by a simple function of ρ and $\pi_c(X)$

Sens analysis techniques to go with 3 types of PI-based estimators

Sens analysis = a modification of main analysis

- Type A (≈ outcome regression estimators)
  - estimates $\kappa_0(X)$ to first estimate effects conditional on covariates and then aggregates them to estimate CACE/NACE
  - sens analysis technique: replace $\kappa_0(X)$ by the identification result of $\mu_{0c}(X)$ under the sens assumption
- Type B (≈ influence function based estimators)
  - can be expressed as combination of IF-based estimators of π_c, ν_{1c} and ν_{0c}^{PI}
  - sens analysis technique: replace $\hat{\nu}_{0c,\text{IF}}^{\text{PI}}$ with an IF-based estimator of ν_{0c} under the sens assumption
- Type C (≈ other/weighting estimators)
  - no general sens analysis technique
  - sens-MR: scale Y in control units by a simple function of ρ and $\pi_c(X)$

JOBS II results



Some other things we noticed/figured out

- Partial loss of multiple robustness

  because $\mu_{0c}(X)$ is a function of $\pi_c(X)$ and $\kappa_0(X)$

- A pattern of finite-sample bias for the sens analysis where effect estimates are less extreme than should be

  because $\text{E}[Y_0 \mid C = c] = \dfrac{\text{E}[\pi_c(X)\mu_{0c}(X)]}{\text{E}[\pi_c(X)]}$ is a weighted average where the function being averaged depends on the weight

- If IF-based nonparametric estimation: Rate conditions on nuisance estimation for the estimator to be root-n consistent

Outline

Background
- Principal causal effects
- Principal ignorability (PI) – one identification strategy
- Sensitivity analysis for PI violation

Mean-based sensitivity analysis: expanding options
- Allow different sensitivity parameters
- Accommodate different estimation methods

Distribution-based sensitivity analysis: a new thing
- Method limitations and information use
- A method using full information (ongoing work)

A general limitation of above methods

is the risk of contradicting the observed data distribution

Assumption	Risk level	Info used from the observed mixture outcome distribution
sens-MR	greatest risk	mean $\kappa_0(X) = \mathrm{E}[Y \mid X, Z = 0]$ lower bound (0)
sens-GOR (for nonbinary Y)	less risk	mean both bounds
sens-SMD	less risk	mean variance $\mathrm{var}(Y \mid X, Z = 0)$
sens-OR (binary Y)	no risk	full distribution (as mean = probability)

For nonbinary Y, to avoid contradicting the observed data distribution,

sens analysis needs to be fully informed by $\mathrm{P}(Y \mid X, Z = 0)$

Constructing that sens analysis: foundation

PI: $Y \perp\!\!\!\perp C \mid X, Z = 0$

(Everything here conditions on $X, Z = 0$, so this will be implicit.)

Constructing that sens analysis: foundation

PI: $Y \perp\!\!\!\perp C \mid X, Z = 0$

(Everything here conditions on $X, Z = 0$, so this will be implicit.)

Want a sens assumption
that allows Y and C to be dependent AND helps identify $\mu_{0c}(X)$

Constructing that sens analysis: foundation

PI: $Y \perp\!\!\!\perp C \mid X, Z = 0$

(Everything here conditions on $X, Z = 0$,
so this will be implicit.)

Want a sens assumption
that allows Y and C to be dependent AND helps identify $\mu_{0c}(X)$

Principal scores

Recall/recast	$\pi_c(X) = \mathrm{P}(C = c \mid X, Z = 0)$	(outcome-agnostic)
Now define	$\tilde{\pi}_c(X, Y) = \mathrm{P}(C = c \mid X, Z = 0, Y)$	(outcome-specific)

Constructing that sens analysis: foundation

PI: $Y \perp\!\!\!\perp C \mid X, Z = 0$

(Everything here conditions on $X, Z = 0$,
so this will be implicit.)

Want a sens assumption
that allows Y and C to be dependent AND helps identify $\mu_{0c}(X)$

Principal scores

Recall/recast	$\pi_c(X) = \mathrm{P}(C = c \mid X, Z = 0)$	(outcome-agnostic)
Now define	$\tilde{\pi}_c(X, Y) = \mathrm{P}(C = c \mid X, Z = 0, Y)$	(outcome-specific)

$\tilde{\pi}_c(X, Y)$ is not identified but we know

- $\mathrm{E}[\tilde{\pi}_c(X, Y) \mid X, Z = 0] = \pi_c(X)$
- If Y and C are dependent, $\tilde{\pi}_c(X, Y)$ is a function of Y in addition to X

Constructing that sens analysis: foundation

PI: $Y \perp\!\!\!\perp C \mid X, Z = 0$

(Everything here conditions on $X, Z = 0$,
so this will be implicit.)

Want a sens assumption
that allows Y and C to be dependent AND helps identify $\mu_{0c}(X)$

Principal scores

Recall/recast	$\pi_c(X) = \text{P}(C = c \mid X, Z = 0)$	(outcome-agnostic)
Now define	$\tilde{\pi}_c(X, Y) = \text{P}(C = c \mid X, Z = 0, Y)$	(outcome-specific)

$\tilde{\pi}_c(X, Y)$ is not identified but we know

- $\text{E}[\tilde{\pi}_c(X, Y) \mid X, Z = 0] = \pi_c(X)$
- If Y and C are dependent, $\tilde{\pi}_c(X, Y)$ is a function of Y in addition to X

(Side note: Very different from the unobserved confounding problem b/c of the observed mixture.
Hence exponential tilting doesn't work, except for a binary outcome, in which case = sens-OR.)

Constructing that sens analysis: 2 steps

Use shorthand $\tilde{\pi}_1$ for $\tilde{\pi}_1(X, Y)$

Step 1: Assume a distribution for $\tilde{\pi}_1$ with mean $\pi_1(X)$ that allows $\tilde{\pi}_1$ to vary, indexed by a dispersion param

Step 2: Connect Y to $\tilde{\pi}_1$ to induce Y-C dependence

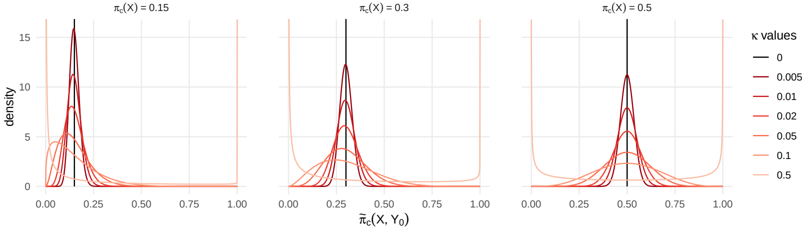
Constructing that sens analysis: 2 steps

Use shorthand $\tilde{\pi}_1$ for $\tilde{\pi}_1(X, Y)$

Step 1: Assume a distribution for $\tilde{\pi}_1$ with mean $\pi_1(X)$ that allows $\tilde{\pi}_1$ to vary, indexed by a dispersion param

- use the beta distribution (idea borrowed from Victor Veitch, 2020)

$$\tilde{\pi}_1 \mid X, Z = 0 \;\sim\; \text{Beta}\left(\pi_1(X)\frac{1-\kappa}{\kappa}, \pi_0(X)\frac{1-\kappa}{\kappa}\right)$$



Constructing that sens analysis: 2 steps

Use shorthand $\tilde{\pi}_1$ for $\tilde{\pi}_1(X, Y)$

Step 1: Assume a distribution for $\tilde{\pi}_1$ with mean $\pi_1(X)$ that allows $\tilde{\pi}_1$ to vary, indexed by a dispersion param

- use the beta distribution

$$\tilde{\pi}_1 \mid X, Z = 0 \ \sim \ \text{Beta}\left(\pi_1(X)\frac{1-\kappa}{\kappa}, \pi_0(X)\frac{1-\kappa}{\kappa}\right)$$

Step 2: Connect Y to $\tilde{\pi}_1$ to induce Y-C dependence

- use quantile-to-quantile mapping between distributions of Y and of $\tilde{\pi}_1$ (given $X, Z = 0$)
  - same-quantiles: positive Y-C association
  - opposite-quantiles: negative Y-C association

Constructing that sens analysis: 2 steps

Use shorthand $\tilde{\pi}_1$ for $\tilde{\pi}_1(X, Y)$

Step 1: Assume a distribution for $\tilde{\pi}_1$ with mean $\pi_1(X)$ that allows $\tilde{\pi}_1$ to vary, indexed by a dispersion param

- use the beta distribution

$$\tilde{\pi}_1 \mid X, Z = 0 \;\sim\; \text{Beta}\left(\pi_1(X)\frac{1-\kappa}{\kappa}, \pi_0(X)\frac{1-\kappa}{\kappa}\right)$$

Step 2: Connect Y to $\tilde{\pi}_1$ to induce Y-C dependence

- use quantile-to-quantile mapping between distributions of Y and of $\tilde{\pi}_1$ (given $X, Z = 0$)
  - same-quantiles: positive Y-C association
  - opposite-quantiles: negative Y-C association

The sens param: the "signed" κ

Identification and estimation

Identification:

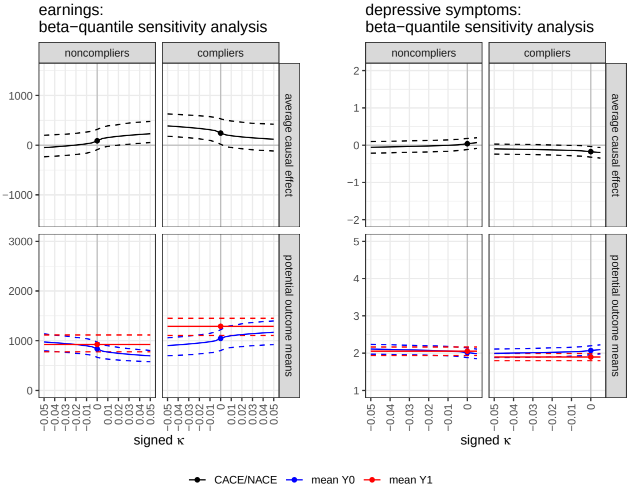
$$\mu_{0c}(X) = \frac{\mathrm{E}[\tilde{\pi}_c Y \mid X, Z = 0]}{\mathrm{E}[\tilde{\pi}_c \mid X, Z = 0]},$$

where $\tilde{\pi}_c$ and Y are quantile-to-quantile connected

Estimation:

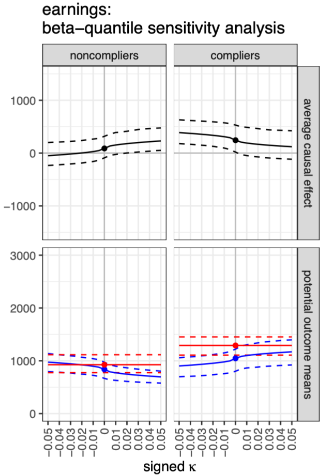
- estimate $\mathrm{P}(Y \mid X, Z = 0)$ so can compute quantiles
- then estimate $\mu_{0c}(X)$ using numerical integration

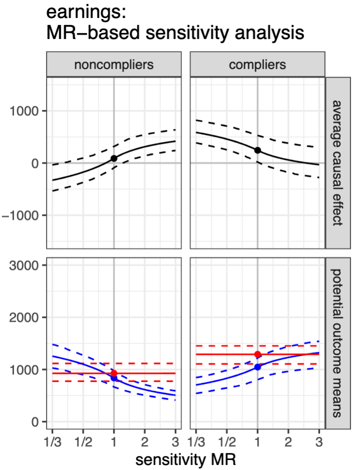
Example: JOBS II preliminary results



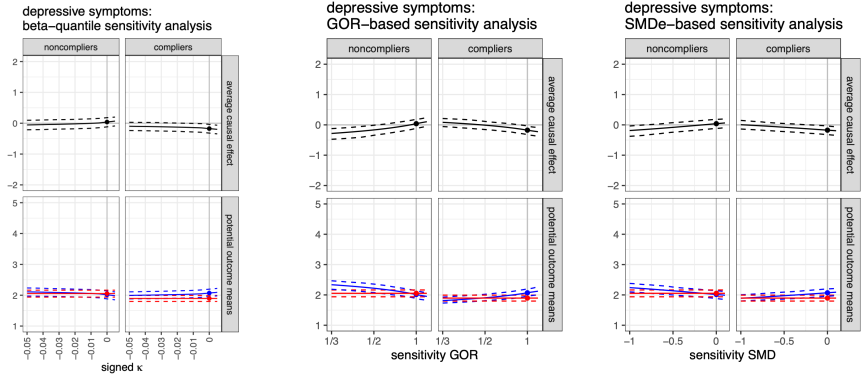
Results are less extreme than mean-based sens analysis

earnings: distribution-based (left), mean-based (right)





depressive symptoms: distribution-based (left), mean-based (middle and right)



Outline

Background
- Principal causal effects
- Principal ignorability (PI) – one identification strategy
- Sensitivity analysis for PI violation

Mean-based sensitivity analysis: expanding options
- Allow different sensitivity parameters
- Accommodate different estimation methods

Distribution-based sensitivity analysis: a new thing
- Method limitations and information use
- A method using full information (ongoing work)