Nonparametric instrumental variable analysis of surgical care for gallstone diseases

In collaboration with
Edward Kennedy (CMU) and Luke Keele (UPenn)

Kenta Takatsu

Motivation

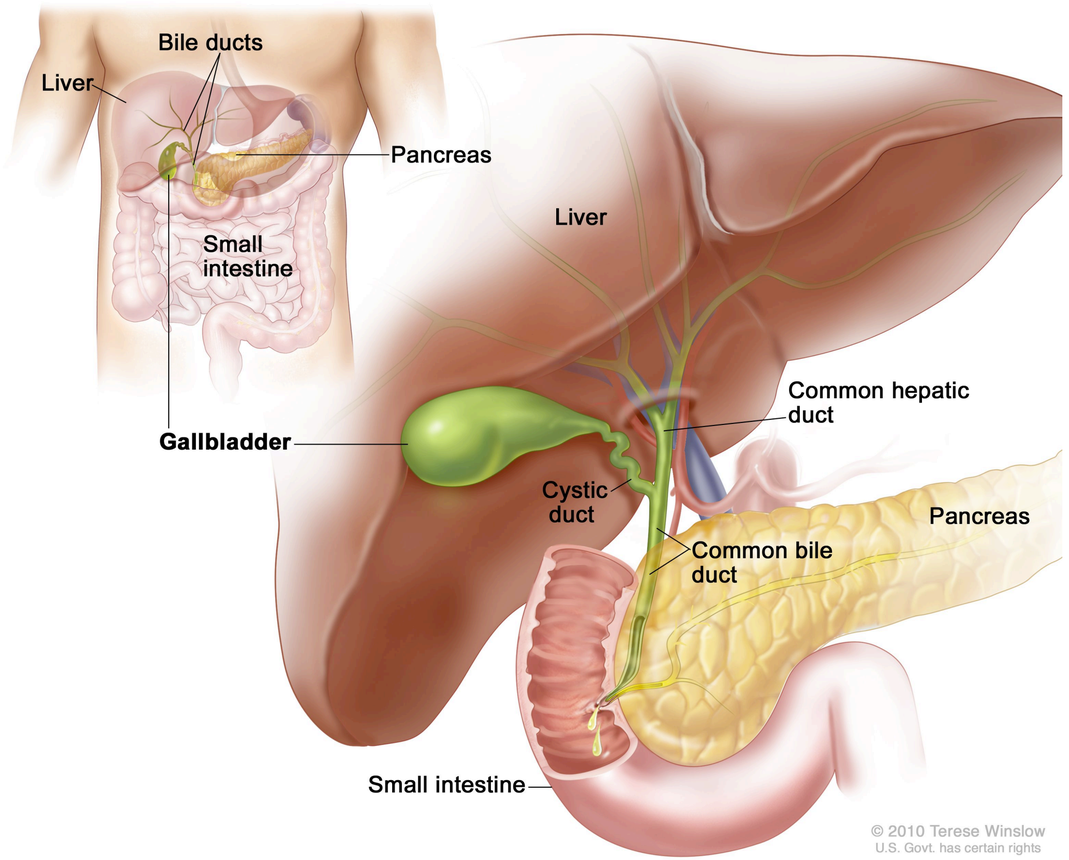


Fig: Anatomy of the gallbladder. Adapted from National Cancer Institute (https://www.cancer.gov/)

In 2012 - 2013, there were **56078** emergency cases associated with gallstone-related diseases in Florida.

In 2012 - 2013, there were **56078** emergency cases associated with gallstone-related diseases in Florida.

47633 patients receive operative treatment (e.g., surgical removal of gallbladder).

In 2012 - 2013, there were **56078** emergency cases associated with gallstone-related diseases in Florida.

47633 patients receive operative treatment (e.g., surgical removal of gallbladder).

However, surgery can lead to additional complications, which may delay recovery time.

In 2012 - 2013, there were **56078** emergency cases associated with gallstone-related diseases in Florida.

47633 patients receive operative treatment (e.g., surgical removal of gallbladder).

However, surgery can lead to additional complications, which may delay recovery time.

Q1: Is operative treatment more effective in reducing length-of-stay?

In 2012 - 2013, there were **56078** emergency cases associated with gallstone-related diseases in Florida.

47633 patients receive operative treatment (e.g., surgical removal of gallbladder).

However, surgery can lead to additional complications, which may delay recovery time.

Q1: Is operative treatment more effective in reducing length-of-stay?

A. Yes. Surgery seems effective on *average.*

In 2012 - 2013, there were **56078** emergency cases associated with gallstone-related diseases in Florida.

47633 patients receive operative treatment (e.g., surgical removal of gallbladder).

However, surgery can lead to additional complications, which may delay recovery time.

Q1: Is operative treatment more effective in reducing length-of-stay?

A. Yes. Surgery seems effective on *average.*

Q2: Should everyone receive surgery?

In 2012 - 2013, there were **56078** emergency cases associated with gallstone-related diseases in Florida.

47633 patients receive operative treatment (e.g., surgical removal of gallbladder).

However, surgery can lead to additional complications, which may delay recovery time.

Q1: Is operative treatment more effective in reducing length-of-stay?

A. Yes. Surgery seems effective on *average.*

Q2: Should everyone receive surgery?

A. No. Surgery may *not* be as effective for certain patients.

Observational data includes confounding

Observational data includes confounding

$\approx 8\,\%$ of the treatment arm (surgery) showed prolonged length-of-stay.

$\approx 24\,\%$ of the control arm showed prolonged length-of-stay.

Observational data includes confounding

$\approx 8\,\%$ of the treatment arm (surgery) showed prolonged length-of-stay.

$\approx 24\,\%$ of the control arm showed prolonged length-of-stay.

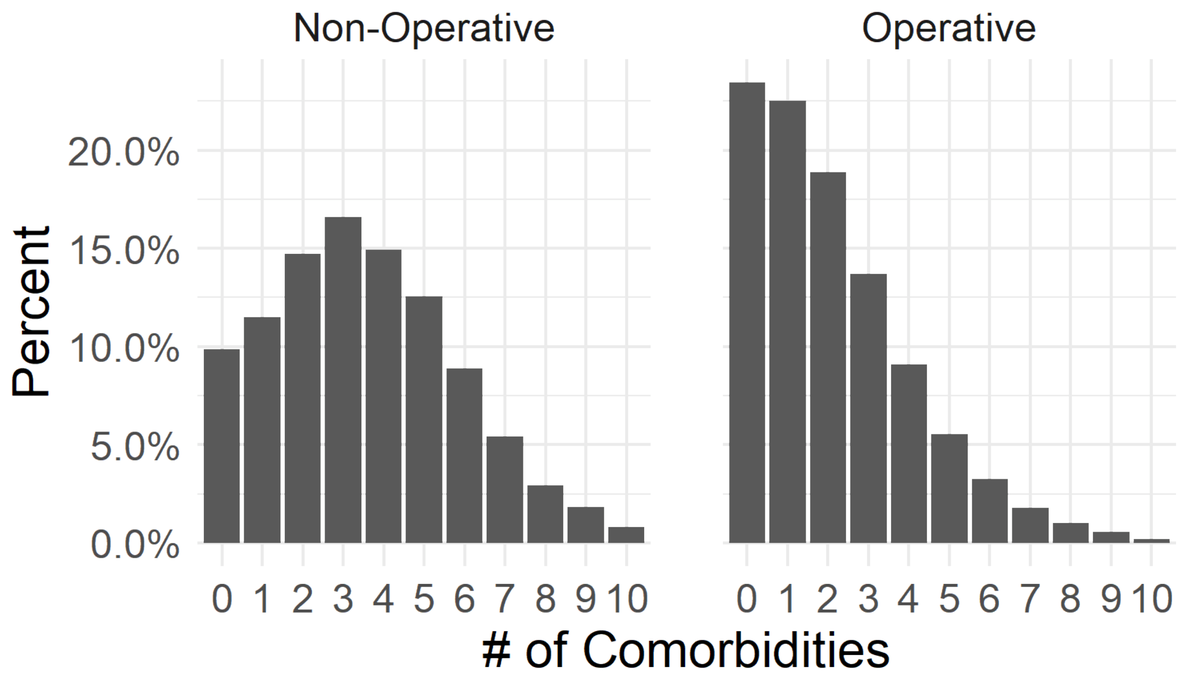


Fig: Empirical distribution of # of comorbidities b/w treatment and control arms.

Observational data includes confounding

$\approx 8\,\%$ of the treatment arm (surgery) showed prolonged length-of-stay.

$\approx 24\,\%$ of the control arm showed prolonged length-of-stay.

There can be **unmeasured confounding.**

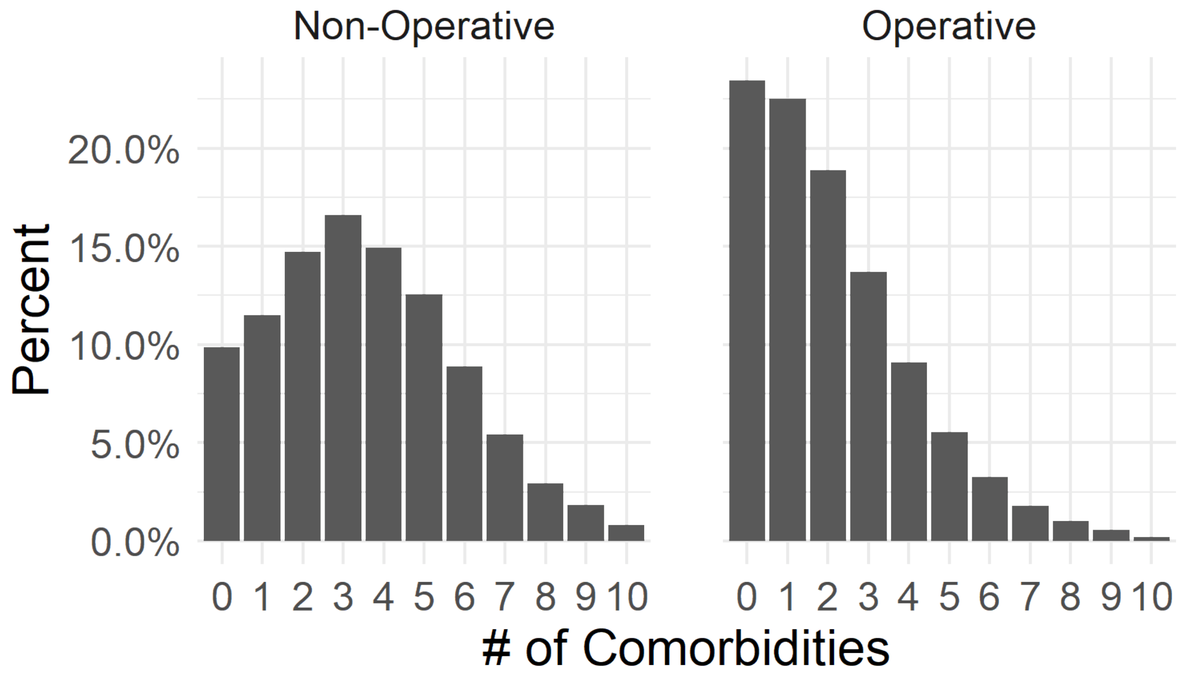


Fig: Empirical distribution of # of comorbidities b/w treatment and control arms.

Observational data includes confounding

What can we do?

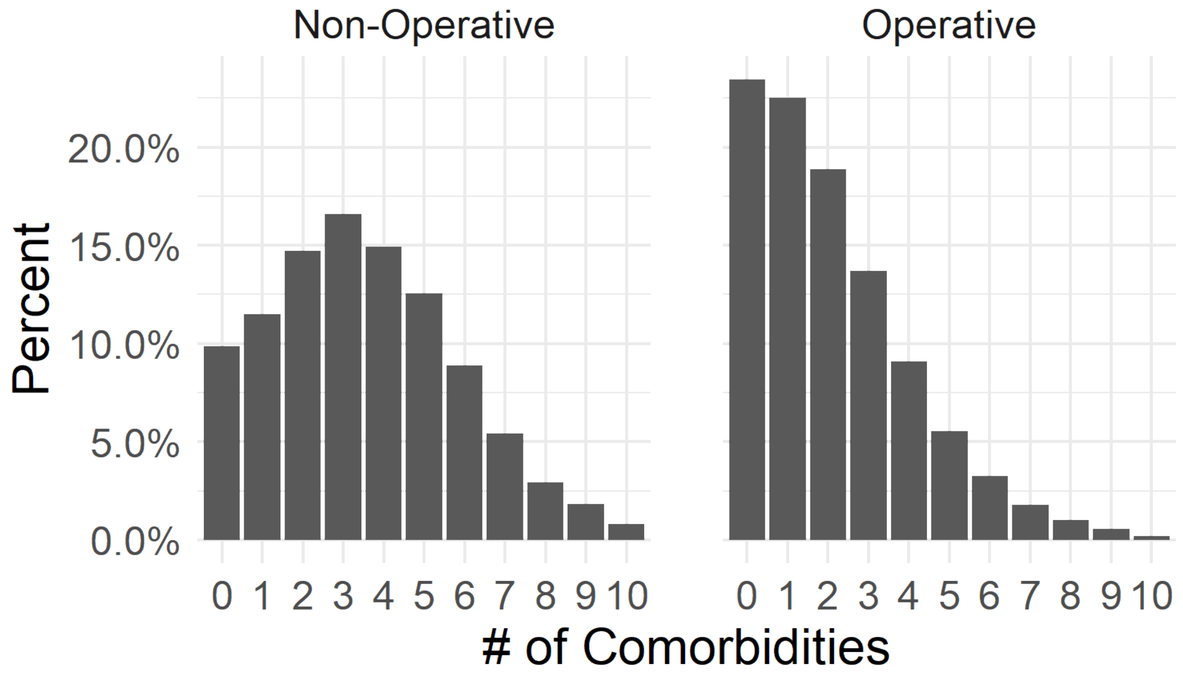


Fig: Empirical distribution of # of comorbidities b/w treatment and control arms.

Observational data includes confounding

What can we do?

If you have a special variable called an ***instrumental variable*** (IV), you can still estimate “certain” treatment effect.

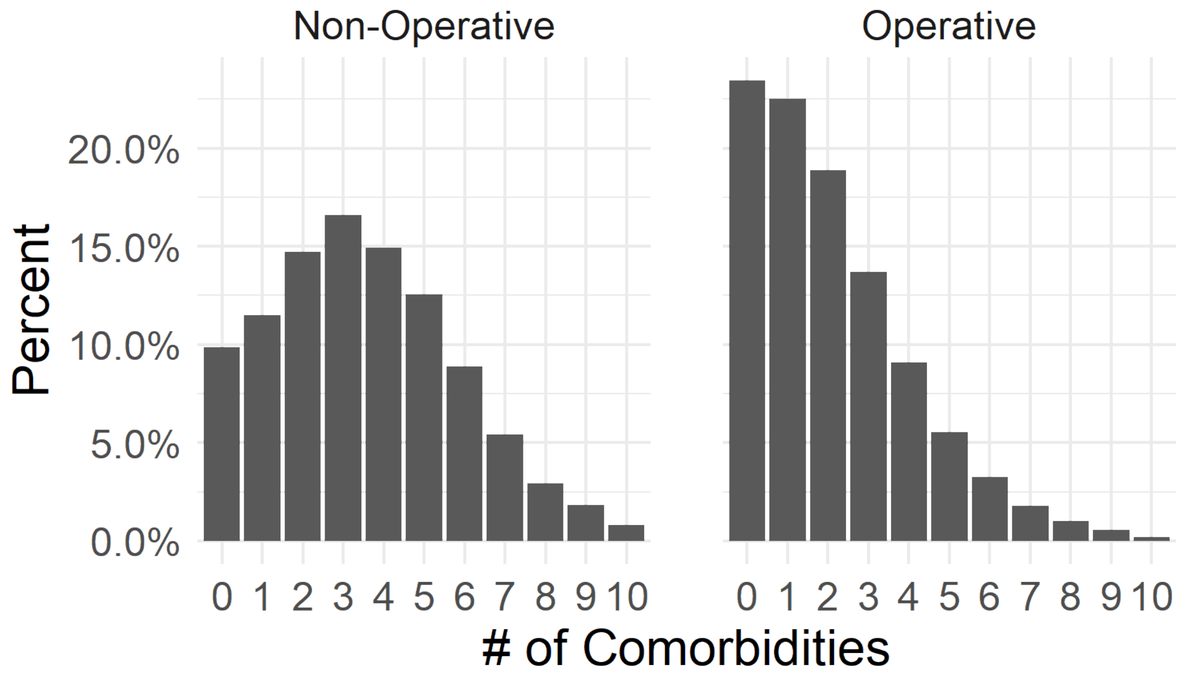
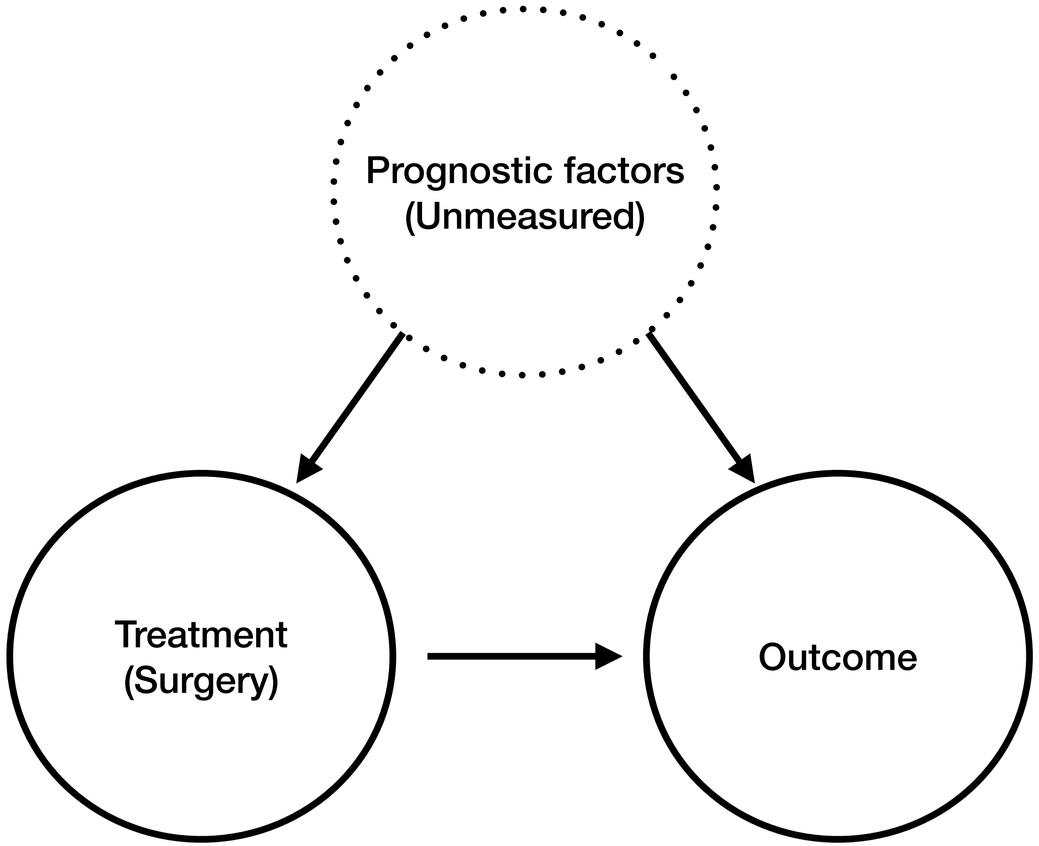
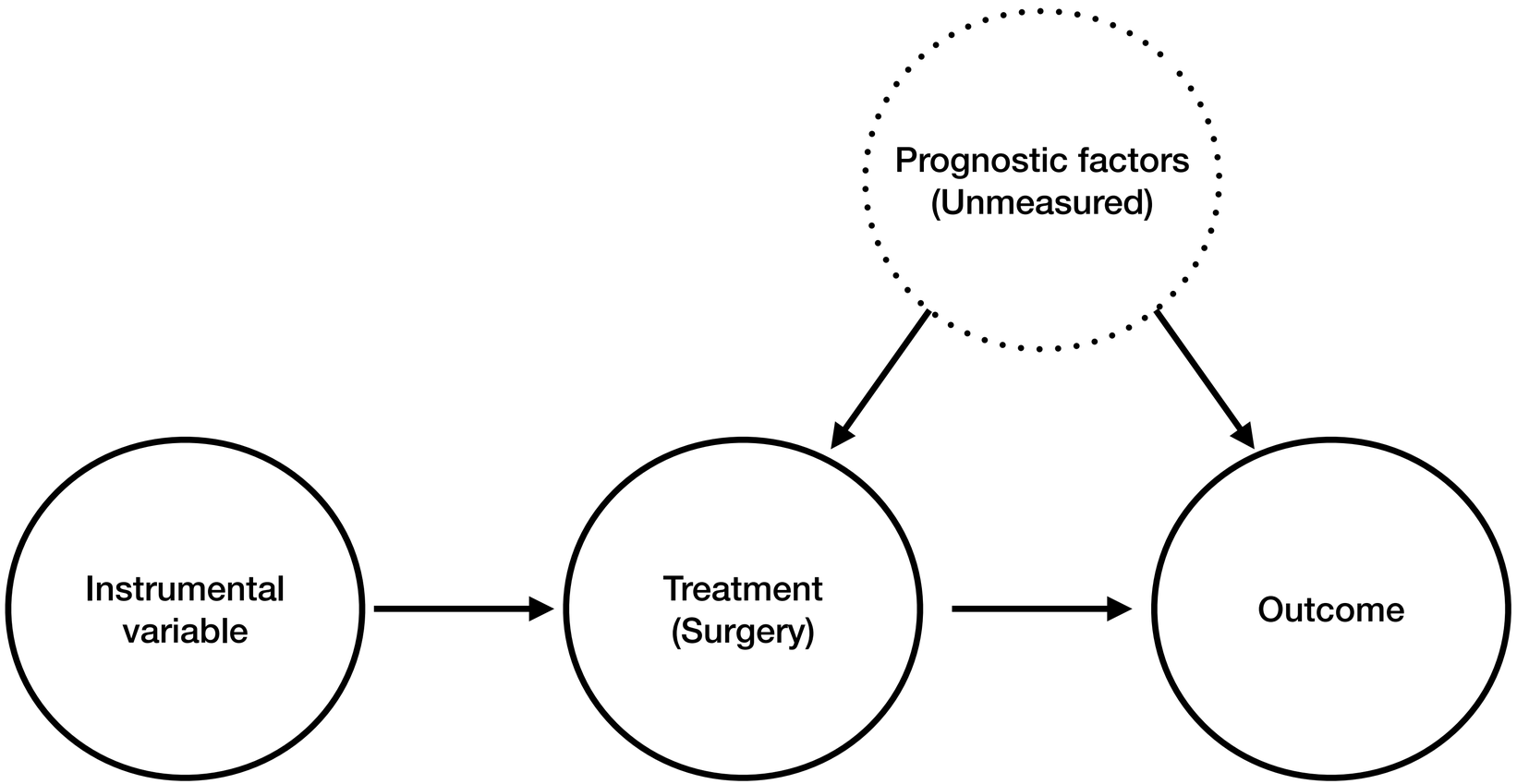


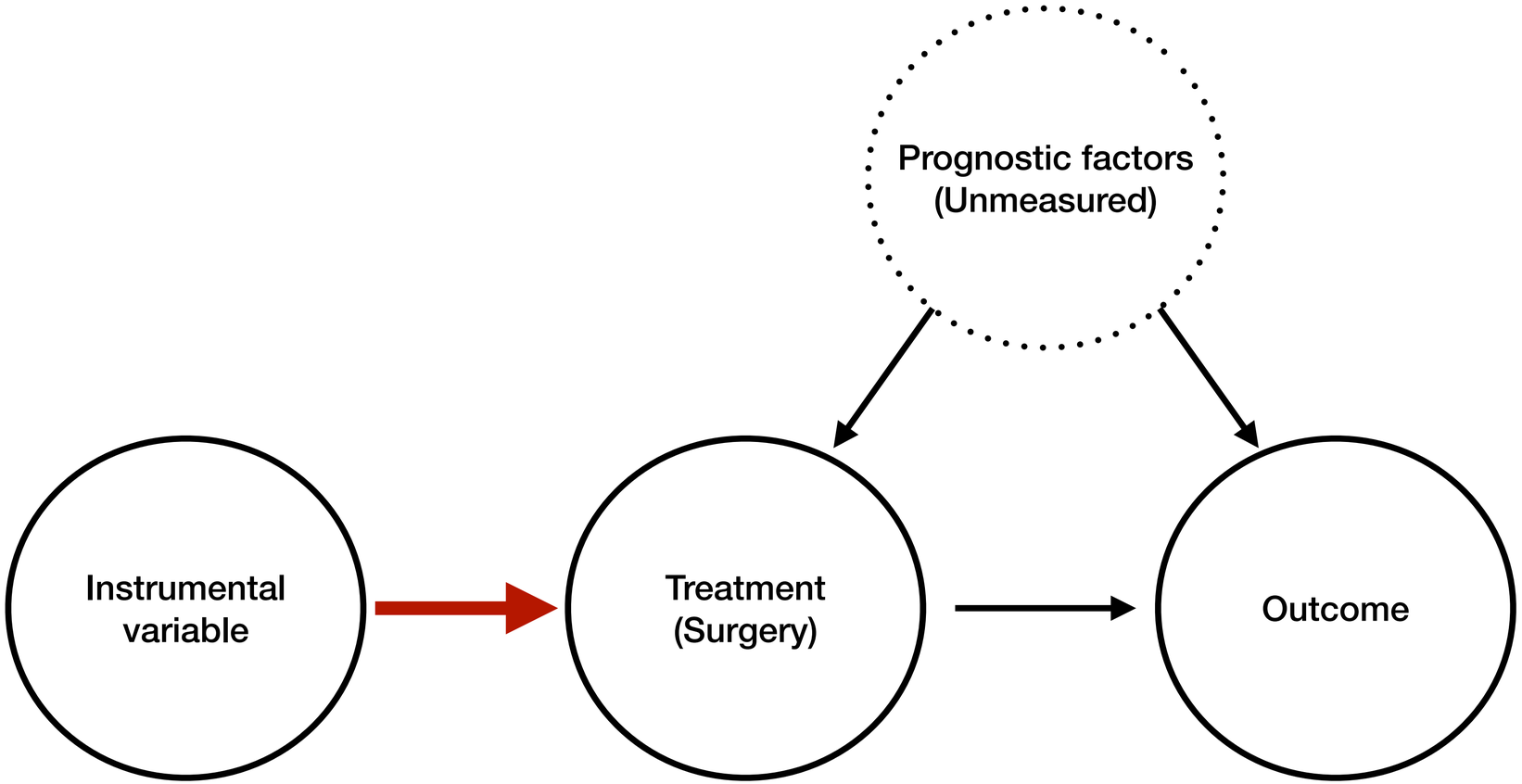
Fig: Empirical distribution of # of comorbidities b/w treatment and control arms.


Prognostic factors
(Unmeasured)
Treatment
(Surgery)
Outcome


Prognostic factors
(Unmeasured)
Instrumental
variable
Treatment
(Surgery)
Outcome

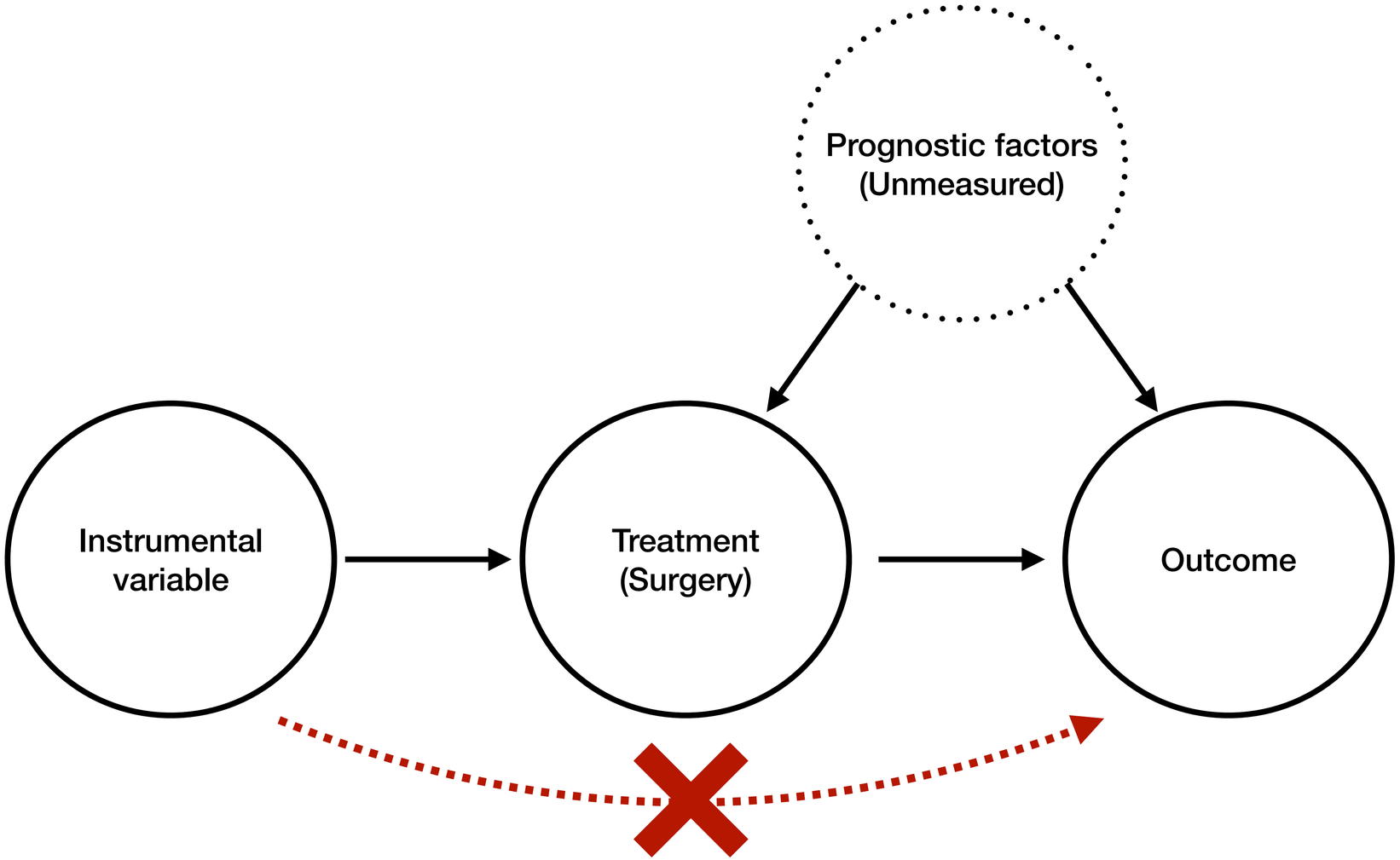
Relevance

An instrument must be associated with treatment



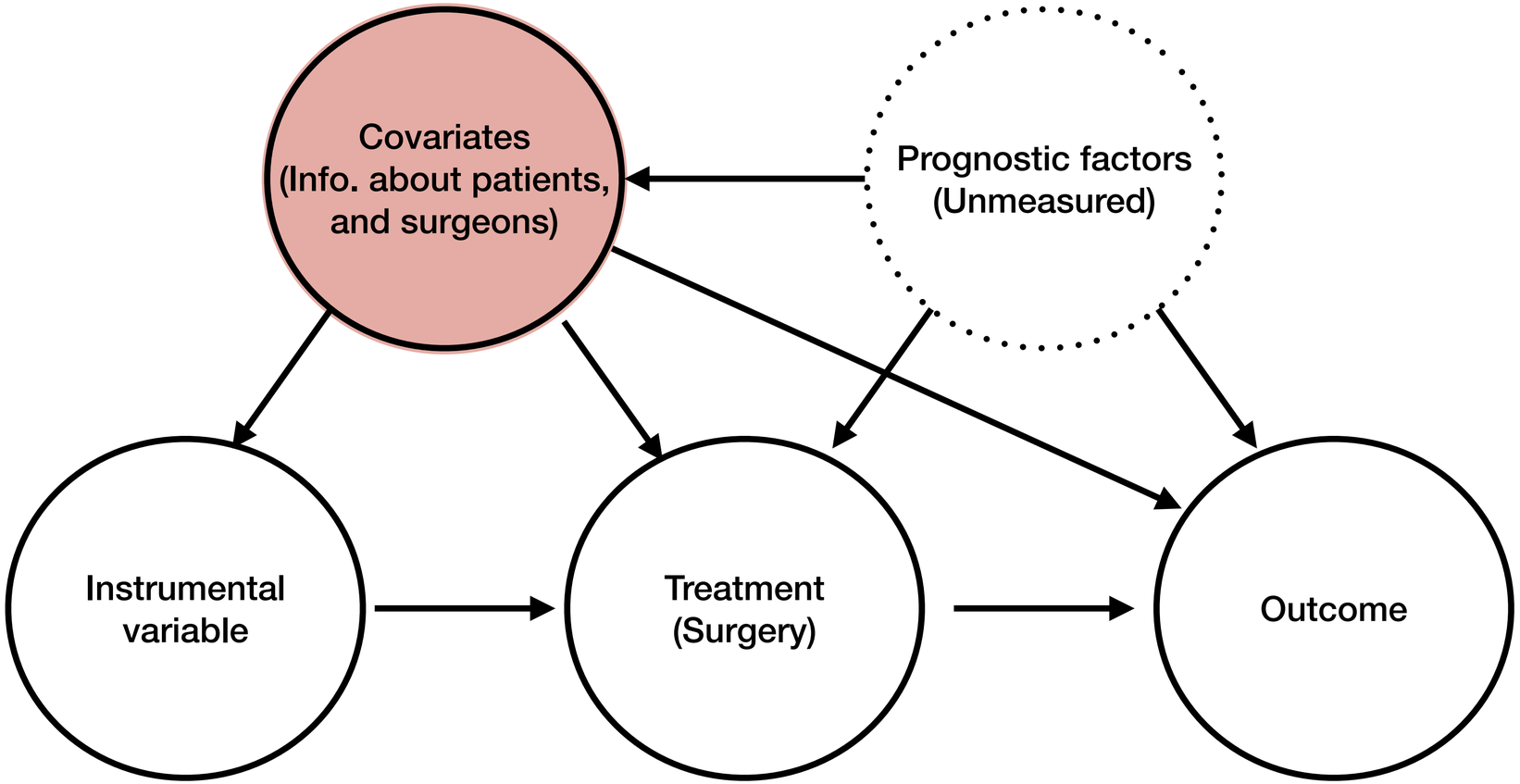
Exclusion Restriction

An instrument must not affect outcomes directly



Unconfounded IV

An instrument must itself be unconfounded



We use surgeons' preference as an IV

We use surgeons' preference as an IV

Relevance:

Patients are more likely to receive treatment surgeons prefer.

We use surgeons' preference as an IV

Relevance:

Patients are more likely to receive treatment surgeons prefer.

Exclusion restriction:

Preference may not directly affect the outcomes.

We use surgeons' preference as an IV

Relevance:

Patients are more likely to receive treatment surgeons prefer.

Exclusion restriction:

Preference may not directly affect the outcomes.

Unconfounded IV:

Patients for emergency care may not choose their surgeons (i.e., randomized).

We use surgeons' preference as an IV

Relevance:

Patients are more likely to receive treatment surgeons prefer.

Exclusion restriction:

Preference may not directly affect the outcomes.

Unconfounded IV:

Patients for emergency care may not choose their surgeons (i.e., randomized).

For each surgeon, compute # of operations/# of patients on a separate data. (Brookhart (2007) and Keele et al. (2018)).

$O := (Y, A, Z, W)$ and observe $\{O_i\}_{i=1}^{n} \overset{iid}{\sim} P_0$

$Y \in \{0,1\}$ denotes the outcomes. 1:="Adverse" outcomes.

$A \in \{0,1\}$ denotes the treatment. 1:=Surgery.

$Z \in \{0,1\}$ denotes the IV. 1:= # of operations/# of patients is above median.

$W \in \mathscr{W} \subseteq \mathbb{R}^d$ where d=226.

Ex) # of comorbidities, an indicator for sepsis, and age.

$O := (Y, A, Z, W)$ and observe $\{O_i\}_{i=1}^{n} \overset{iid}{\sim} P_0$

$Y \in \{0,1\}$ denotes the outcomes. 1:="Adverse" outcomes.

$A \in \{0,1\}$ denotes the treatment. 1:=Surgery.

$Z \in \{0,1\}$ denotes the IV. 1:= # of operations/# of patients is above median.

$W \in \mathscr{W} \subseteq \mathbb{R}^d$ where d=226.

Ex) # of comorbidities, an indicator for sepsis, and age.

Average treatment effect (ATE): $E[Y(1) - Y(0)]$

$O := (Y, A, Z, W)$ and observe $\{O_i\}_{i=1}^{n} \overset{iid}{\sim} P_0$

$Y \in \{0,1\}$ denotes the outcomes. 1:="Adverse" outcomes.

$A \in \{0,1\}$ denotes the treatment. 1:=Surgery.

$Z \in \{0,1\}$ denotes the IV. 1:= # of operations/# of patients is above median.

$W \in \mathscr{W} \subseteq \mathbb{R}^d$ where d=226.

Ex) # of comorbidities, an indicator for sepsis, and age.

Average treatment effect (ATE): $E[Y(1) - Y(0)]$

Avg. outcome *if* everyone *received* surgery

$O := (Y, A, Z, W)$ and observe $\{O_i\}_{i=1}^{n} \overset{iid}{\sim} P_0$

$Y \in \{0,1\}$ denotes the outcomes. 1:="Adverse" outcomes.

$A \in \{0,1\}$ denotes the treatment. 1:=Surgery.

$Z \in \{0,1\}$ denotes the IV. 1:= # of operations/# of patients is above median.

$W \in \mathscr{W} \subseteq \mathbb{R}^d$ where d=226.

Ex) # of comorbidities, an indicator for sepsis, and age.

Average treatment effect (ATE): $E[Y(1) - Y(0)]$

Local average treatment effect (LATE): $E[Y(1) - Y(0) \mid \text{Complier}]$

$O := (Y, A, Z, W)$ and observe $\{O_i\}_{i=1}^{n} \overset{iid}{\sim} P_0$

$Y \in \{0,1\}$ denotes the outcomes. 1:="Adverse" outcomes.

$A \in \{0,1\}$ denotes the treatment. 1:=Surgery.

$Z \in \{0,1\}$ denotes the IV. 1:= # of operations/# of patients is above median.

$W \in \mathscr{W} \subseteq \mathbb{R}^d$ where d=226.

Ex) # of comorbidities, an indicator for sepsis, and age.

Average treatment effect (ATE): $E[Y(1) - Y(0)]$

Local average treatment effect (LATE): $E[Y(1) - Y(0) \mid \text{Complier}]$

Recall $A \in \{0,1\}$ and $Z \in \{0,1\}$.

Recall $A \in \{0,1\}$ and $Z \in \{0,1\}$.

1. Always-takers
2. Never-takers
3. Compliers

   Patients who follow surgeon's preference

4. Defiers

   Patients who reject surgeon's preference

Recall $A \in \{0,1\}$ and $Z \in \{0,1\}$.

1. Always-takers
2. Never-takers
3. Compliers

   Patients who follow surgeon's preference

~~4. Defiers~~

Patients who reject surgeon's preference

Monotonicity

LATE := $E[Y(1) - Y(0)$ | Complier]

Monotonicity := No Defiers with prob. 1

LATE := $E[Y(1) - Y(0)$ | Complier]

Monotonicity := No Defiers with prob. 1

Both LATE and monotonicity have been controversial
(Imbens (2014), Swanson and Hernan (2014))

LATE := $E[Y(1) - Y(0)$ | Complier]

Monotonicity := No Defiers with prob. 1

Both LATE and monotonicity have been controversial
(Imbens (2014), Swanson and Hernan (2014))

1. We do not generally know who compliers are.

LATE := $E[Y(1) - Y(0)$ | Complier]

Monotonicity := No Defiers with prob. 1

Both LATE and monotonicity have been controversial
(Imbens (2014), Swanson and Hernan (2014))

1. We do not generally know who compliers are.
2. It might be unreasonable to assume no defiers in our context.

LATE := $E[Y(1) - Y(0)$ | Complier$]$

Monotonicity := No Defiers with prob. 1

Both LATE and monotonicity have been controversial
(Imbens (2014), Swanson and Hernan (2014))

1. We do not generally know who compliers are.
2. It might be unreasonable to assume no defiers in our context.

Remedy for 1

We can estimate covariate density for each patient type:
$P(V = v \mid$ Complier$)$, $P(V = v \mid$ Always-taker$)$, and
$P(V = v \mid$ Never-taker$)$.

LATE := $E[Y(1) - Y(0) \mid \text{Complier}]$

Monotonicity := No Defiers with prob. 1

Both LATE and monotonicity have been controversial
(Imbens (2014), Swanson and Hernan (2014))

1. We do not generally know who compliers are.
2. It might be unreasonable to assume no defiers in our context.

Remedy for 1

We can estimate covariate density for each patient type:
$P(V = v \mid \text{Complier})$, $P(V = v \mid \text{Always-taker})$, and
$P(V = v \mid \text{Never-taker})$.

Remedy for 2

We can study the robustness of our estimates when there are defiers.

$$\mathbb{P}(I\{\text{\# of Comorb. = v}\} \mid \text{Patient type})$$

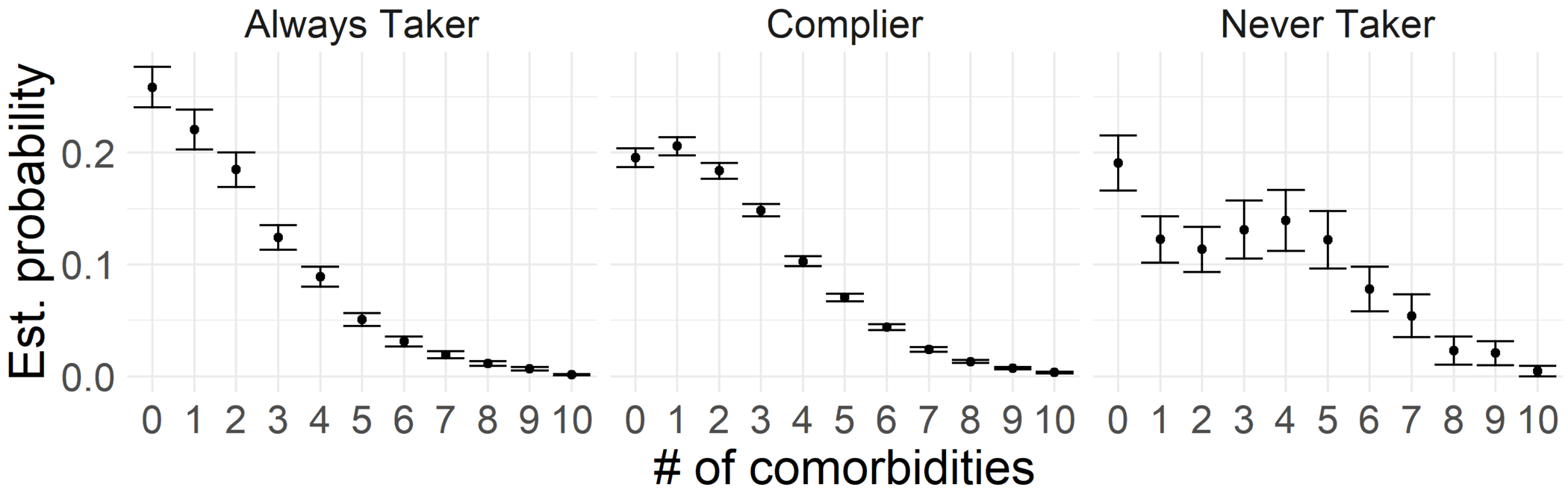


Fig: Estimated conditional probability of # of comorbidities for each patient type. Vertical bars are pointwise 95% CIs.

Nonparametric estimation and inference

$$E[Y(1) - Y(0) \mid \text{Complier}] = \psi_0$$

$$E[Y(1) - Y(0) \mid \text{Complier}] = \psi_0$$

$$= \frac{E[E[Y \mid Z = 1, W]] - E[E[Y \mid Z = 0, W]]}{E[E[A \mid Z = 1, W]] - E[E[A \mid Z = 0, W]]}$$

By Imbens and Angrist (1994)
(Valid IV, no-defiers, Positivity)

$$E[Y(1) - Y(0) \mid \text{Complier}] = \psi_0$$

$$= \frac{E[E[Y \mid Z = 1, W]] - E[E[Y \mid Z = 0, W]]}{E[E[A \mid Z = 1, W]] - E[E[A \mid Z = 0, W]]}$$

By Imbens and Angrist (1994)
(Valid IV, no-defiers, Positivity)

$$\approx \frac{n^{-1} \sum_{i=1}^{n} \widehat{\mu_n}(1, W_i) - n^{-1} \sum_{i=1}^{n} \widehat{\mu_n}(0, W_i)}{n^{-1} \sum_{i=1}^{n} \widehat{\lambda_n}(1, W_i) - n^{-1} \sum_{i=1}^{n} \widehat{\lambda_n}(0, W_i)}$$

$$E[Y(1) - Y(0) \mid \text{Complier}] = \psi_0$$

$$= \frac{E[E[Y \mid Z=1,W]] - E[E[Y \mid Z=0,W]]}{E[E[A \mid Z=1,W]] - E[E[A \mid Z=0,W]]}$$

By Imbens and Angrist (1994)
(Valid IV, no-defiers, Positivity)

$$\approx \frac{n^{-1}\sum_{i=1}^{n} \widehat{\mu_n}(1,W_i) - n^{-1}\sum_{i=1}^{n} \widehat{\mu_n}(0,W_i)}{n^{-1}\sum_{i=1}^{n} \widehat{\lambda_n}(1,W_i) - n^{-1}\sum_{i=1}^{n} \widehat{\lambda_n}(0,W_i)}$$

This estimator is suboptimal.

There *is* a root-n consistent estimator.

We can correct the first-order bias by *estimating* influence functions.

We can correct the first-order bias by *estimating* influence functions.

We have to estimate:

$$\mu_0(z, w) := E[Y \mid Z = z, W = w]$$

$$\lambda_0(z, w) := E[A \mid Z = z, W = w]$$

$$\pi_0(w) := P(Z = 1 \mid W = w)$$

We can correct the first-order bias by *estimating* influence functions.

We have to estimate:

$$\mu_0(z, w) := E[Y \mid Z = z, W = w]$$
$$\lambda_0(z, w) := E[A \mid Z = z, W = w]$$
$$\pi_0(w) := P(Z = 1 \mid W = w)$$

We use sample-splitting and construct estimators $\widehat{\mu_n}$, $\widehat{\lambda_n}$, and $\widehat{\pi_n}$ using machine learning.

We can correct the first-order bias by *estimating* influence functions.

We have to estimate:

$$\mu_0(z, w) := E[Y \mid Z = z, W = w]$$
$$\lambda_0(z, w) := E[A \mid Z = z, W = w]$$
$$\pi_0(w) := P(Z = 1 \mid W = w)$$

We use sample-splitting and construct estimators $\widehat{\mu_n}$, $\widehat{\lambda_n}$, and $\widehat{\pi_n}$ using machine learning.

An estimator of LATE ψ_0

$$\widehat{\psi_n} := \frac{n^{-1}\sum_{i=1}^{n} \phi_1(O_i; \widehat{\mu_n}, \widehat{\pi_n})}{n^{-1}\sum_{i=1}^{n} \phi_2(O_i; \widehat{\lambda_n}, \widehat{\pi_n})}$$

We can correct the first-order bias by *estimating* influence functions.

We have to estimate:

$$\mu_0(z,w) := E[Y \mid Z=z, W=w]$$
$$\lambda_0(z,w) := E[A \mid Z=z, W=w]$$
$$\pi_0(w) := P(Z=1 \mid W=w)$$

We use sample-splitting and construct estimators $\widehat{\mu_n}$, $\widehat{\lambda_n}$, and $\widehat{\pi_n}$ using machine learning.

An estimator of LATE ψ_0

$$\widehat{\psi_n} := \frac{n^{-1}\sum_{i=1}^{n}\phi_1(O_i;\widehat{\mu_n},\widehat{\pi_n})}{n^{-1}\sum_{i=1}^{n}\phi_2(O_i;\widehat{\lambda_n},\widehat{\pi_n})}$$

Asymptotic normality

$$n^{1/2}\left(\widehat{\psi_n}-\psi_0\right) \xrightarrow{d} N\left(0,\sigma^2(\mu_0,\lambda_0,\pi_0)\right)$$

We can correct the first-order bias by *estimating* influence functions.

We have to estimate:

$$\mu_0(z,w) := E[Y \mid Z=z, W=w]$$
$$\lambda_0(z,w) := E[A \mid Z=z, W=w]$$
$$\pi_0(w) := P(Z=1 \mid W=w)$$

We use sample-splitting and construct estimators $\widehat{\mu_n}$, $\widehat{\lambda_n}$, and $\widehat{\pi_n}$ using machine learning.

An estimator of LATE ψ_0

$$\widehat{\psi_n} := \frac{n^{-1}\sum_{i=1}^{n}\phi_1(O_i;\widehat{\mu_n},\widehat{\pi_n})}{n^{-1}\sum_{i=1}^{n}\phi_2(O_i;\widehat{\lambda_n},\widehat{\pi_n})}$$

Asymptotic normality

$$n^{1/2}\left(\widehat{\psi_n}-\psi_0\right) \xrightarrow{d} N\left(0,\sigma^2(\mu_0,\lambda_0,\pi_0)\right)$$

$$\Longrightarrow \left[\widehat{\psi_n} \pm 1.96\, n^{-1/2}\hat{\sigma}\right]$$

LATE can answer if surgery is effective for compliers on average.

LATE can answer if surgery is effective for compliers on average.

LATE *cannot* answer the following:

Does surgery become more effective if a patient is young?
How does the efficacy vary as the number of comorbidities increases?
“I am septic but have no comorbidities. Should I receive surgery?”

LATE can answer if surgery is effective for compliers on average.

LATE *cannot* answer the following:

Does surgery become more effective if a patient is young?
How does the efficacy vary as the number of comorbidities increases?
“I am septic but have no comorbidities. Should I receive surgery?”

$$\psi_0(v) = E[Y(1) - Y(0) \mid \text{Complier}, V = v]$$

An estimator of LATE

$$\widehat{\psi_n} := \frac{n^{-1}\sum_{i=1}^{n}\phi_1(O_i;\widehat{\mu_n},\widehat{\pi_n})}{n^{-1}\sum_{i=1}^{n}\phi_2(O_i;\widehat{\lambda_n},\widehat{\pi_n})}$$

An estimator of Cond. LATE

$$\widehat{\psi_n}(v) := \frac{\widehat{E}\left[\phi_1(O;\widehat{\mu_n},\widehat{\pi_n}) \mid V = v\right]}{\widehat{E}\left[\phi_2(O;\widehat{\mu_n},\widehat{\lambda_n}) \mid V = v\right]}$$

An estimator of LATE

$$\widehat{\psi_n} := \frac{n^{-1}\sum_{i=1}^{n}\phi_1(O_i;\widehat{\mu_n},\widehat{\pi_n})}{n^{-1}\sum_{i=1}^{n}\phi_2(O_i;\widehat{\lambda_n},\widehat{\pi_n})}$$

An estimator of Cond. LATE

$$\widehat{\psi_n}(v) := \frac{\widehat{E}\left[\phi_1(O;\widehat{\mu_n},\widehat{\pi_n}) \mid V = v\right]}{\widehat{E}\left[\phi_2(O;\widehat{\mu_n},\widehat{\lambda_n}) \mid V = v\right]}$$

An estimator of LATE

$$\widehat{\psi_n} := \frac{n^{-1}\sum_{i=1}^{n}\phi_1(O_i;\widehat{\mu_n},\widehat{\pi_n})}{n^{-1}\sum_{i=1}^{n}\phi_2(O_i;\widehat{\lambda_n},\widehat{\pi_n})}$$

An estimator of Cond. LATE

$$\widehat{\psi_n}(v) := \frac{\widehat{E}\left[\phi_1(O;\widehat{\mu_n},\widehat{\pi_n}) \mid V=v\right]}{\widehat{E}\left[\phi_2(O;\widehat{\mu_n},\widehat{\lambda_n}) \mid V=v\right]}$$

Inference of $\psi_0(v)$ is generally challenging.
We use bootstrap to construct CIs.

Empirical results

Surgery is effective on average if you are a complier

We first estimate LATE.

- Lower the better (i.e., surgery reduces the rate of “adverse” outcomes).

Unadjusted estimator ignores confounding.

TSLS is a parametric method based on linear regression.

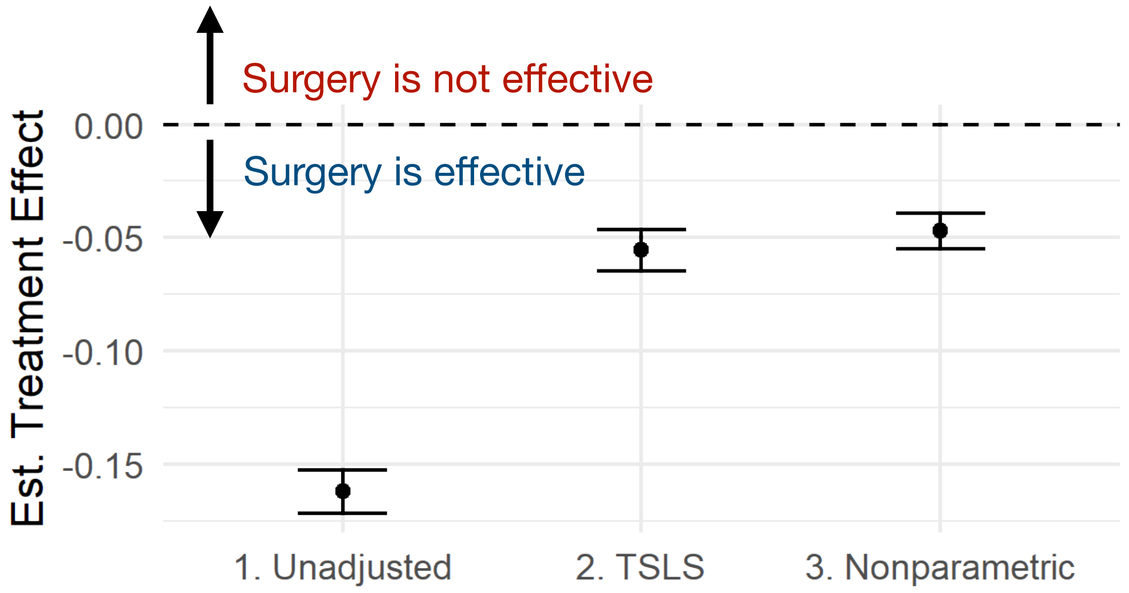


Fig: The point estimates of LATE and 95% CIs from three estimators.

Surgery *may not* be effective for most people

We estimate $E[Y(1) - Y(0) \mid \text{Complier}, W = w]$ where W is all covariates.

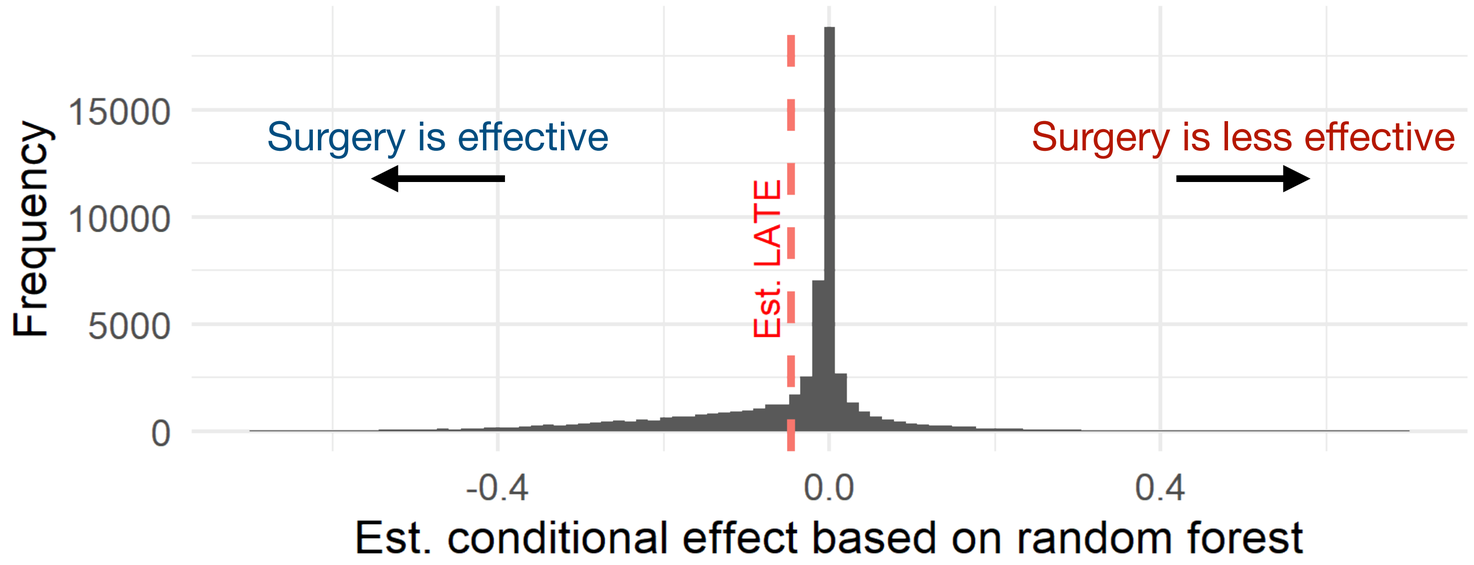


Fig: The distribution of the estimated cond. LATE on all available covariates.

How does the efficacy vary as a function of covariates?

1. We estimate $E[Y(1) - Y(0) \mid \text{Complier}, V = v]$ for V including # of comorbidities, an indicator for sepsis, and age.

2. For the regression model, we use a generalized additive model.

3. We use bootstrap samples to construct 95% confidence sets.

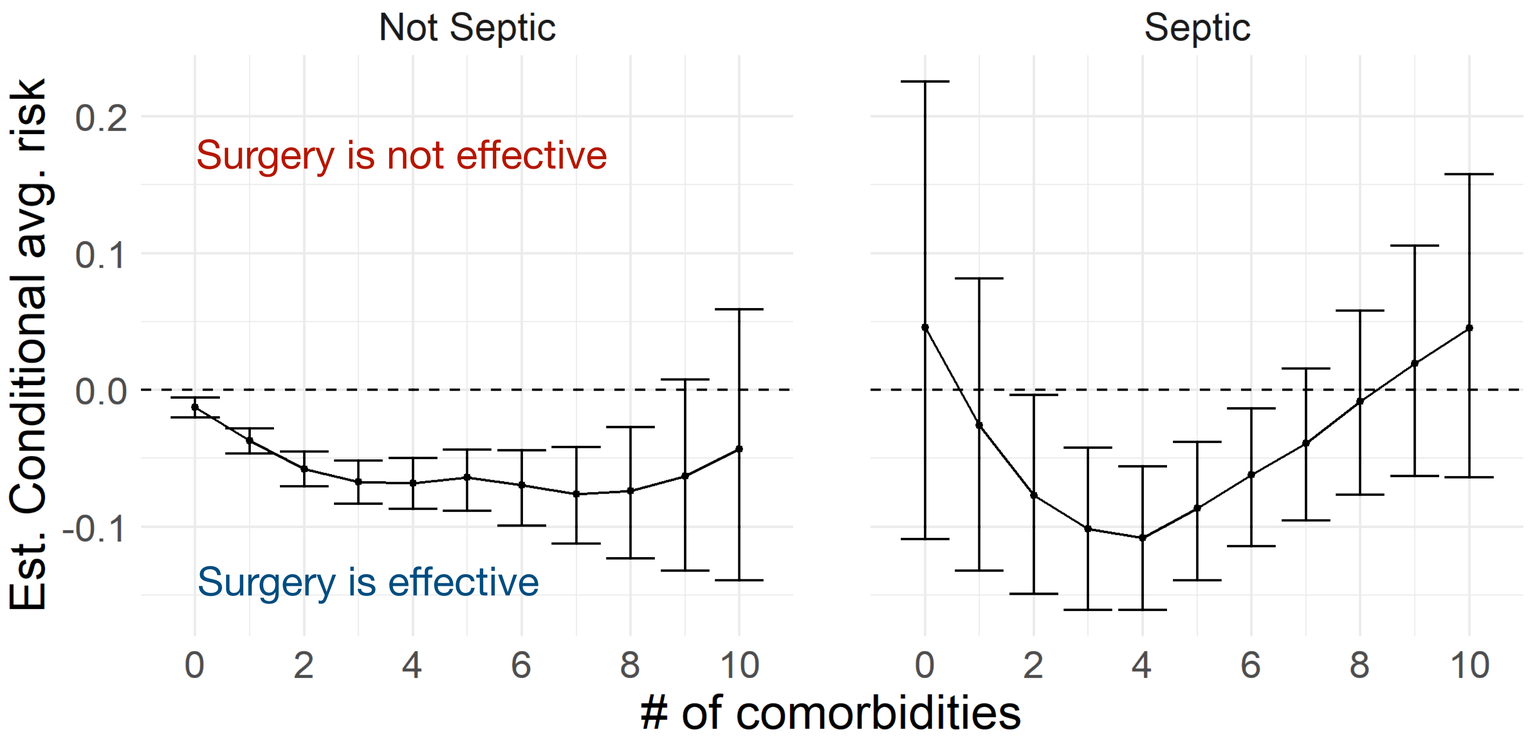


Fig: Estimated cond. LATE and bootstrap CIs as a function of comorbidities and sepsis.

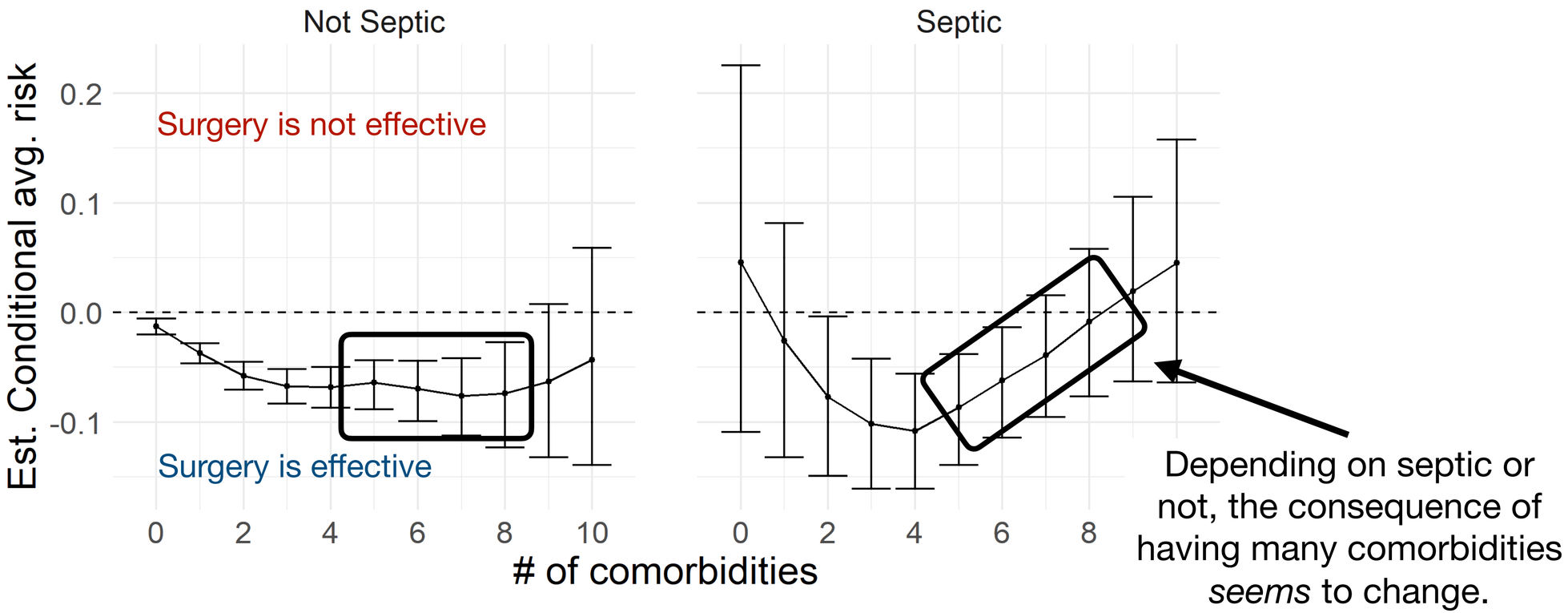


Fig: Estimated cond. LATE and bootstrap CIs as a function of comorbidities and sepsis.

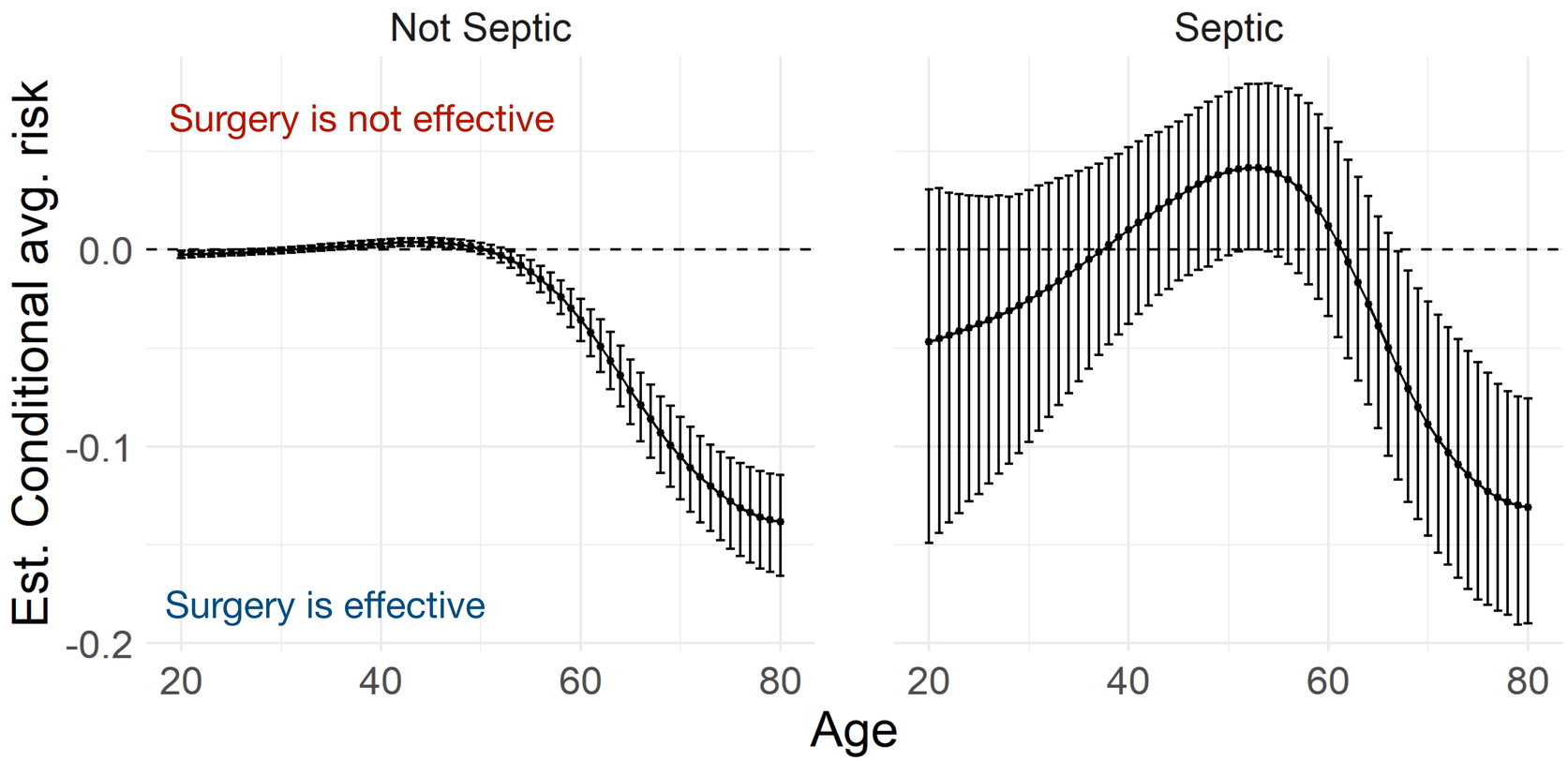
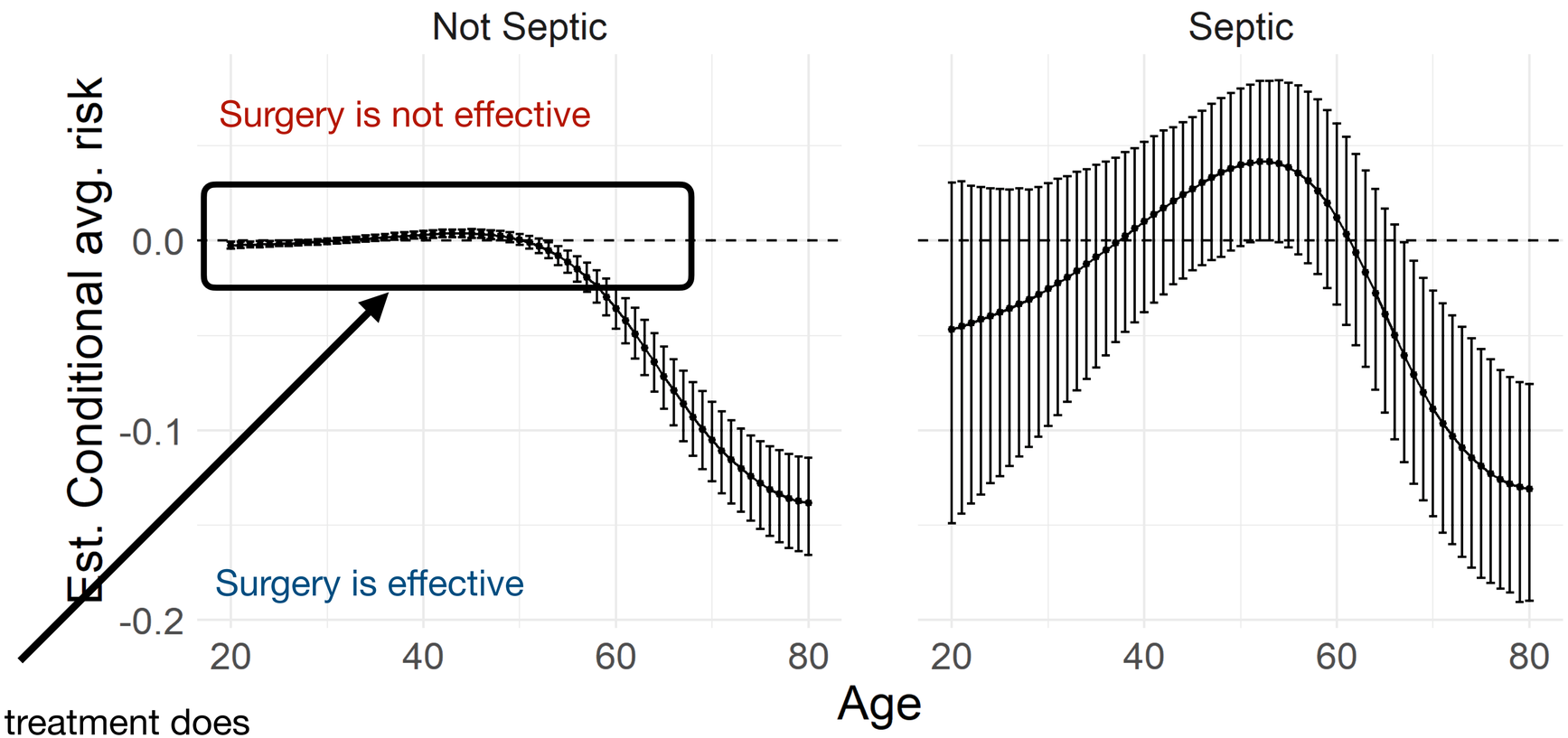


Fig: Estimated cond. LATE and bootstrap CIs as a function of age and sepsis.



The choice of treatment does not really matter for non-septic and young patients

Fig: Estimated cond. LATE and bootstrap CIs as a function of age and sepsis.

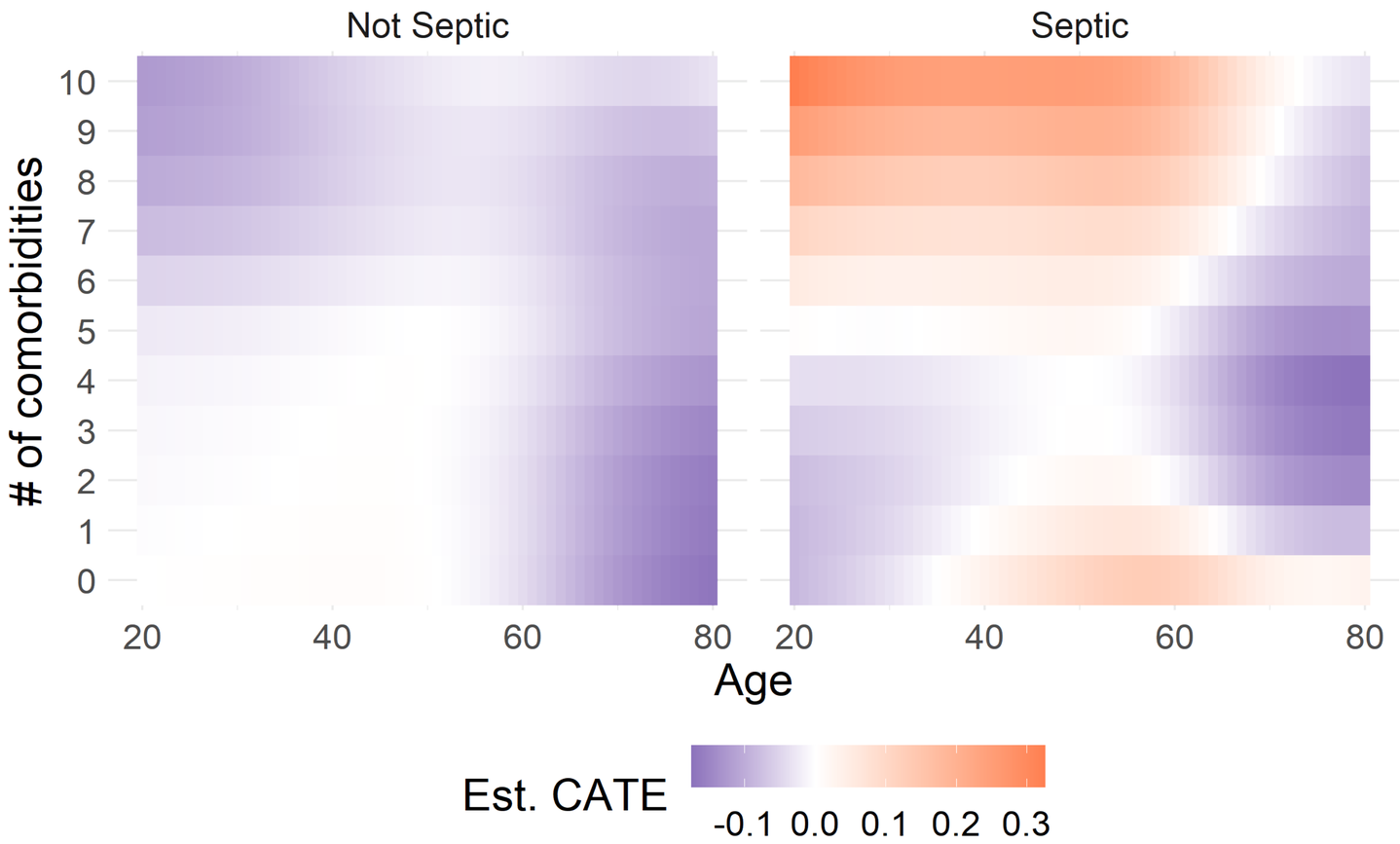
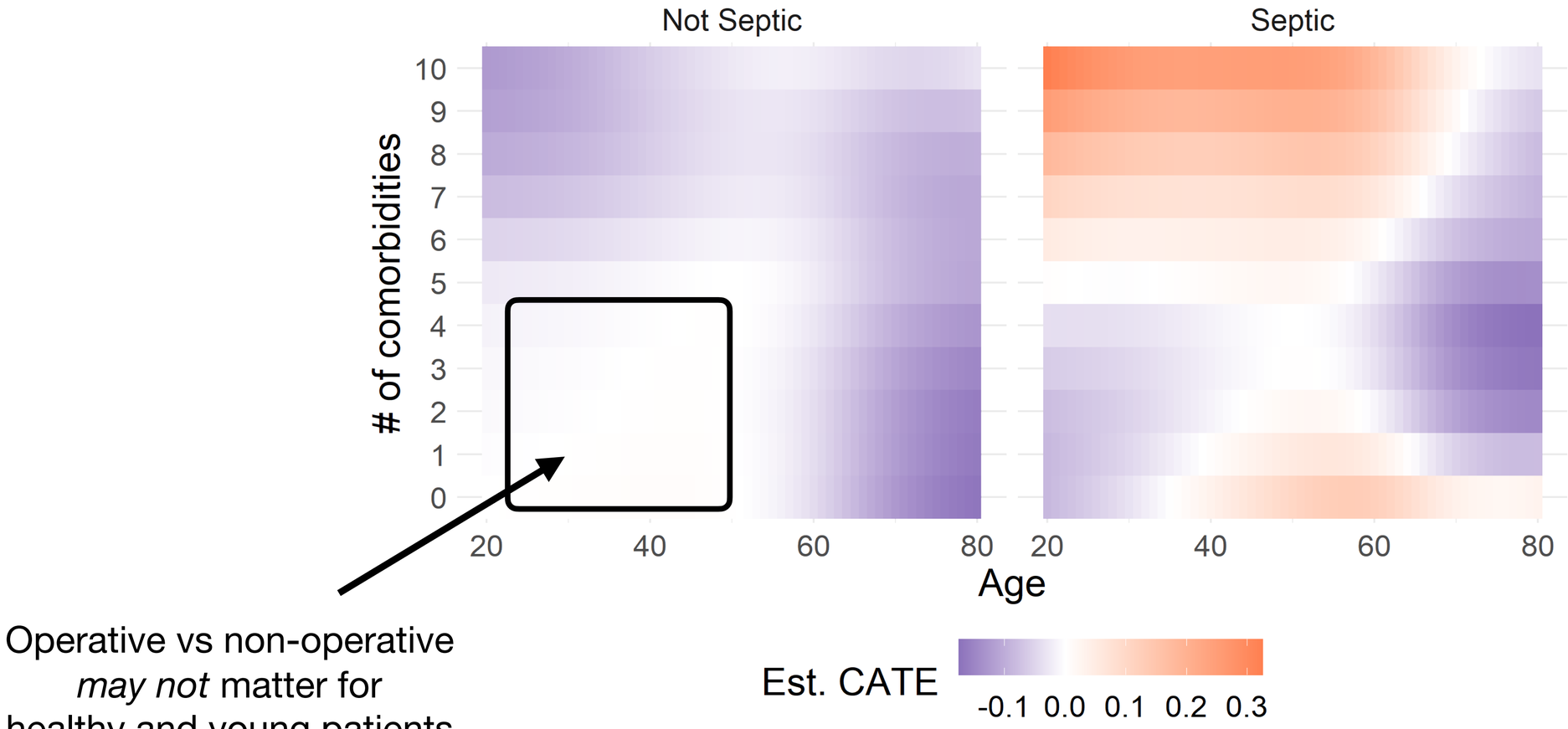


Fig: Heatmap of cond. LATE as a function of age, comorbidities and sepsis.



Operative vs non-operative *may not* matter for healthy and young patients

Fig: Heatmap of cond. LATE as a function of age, comorbidities and sepsis.

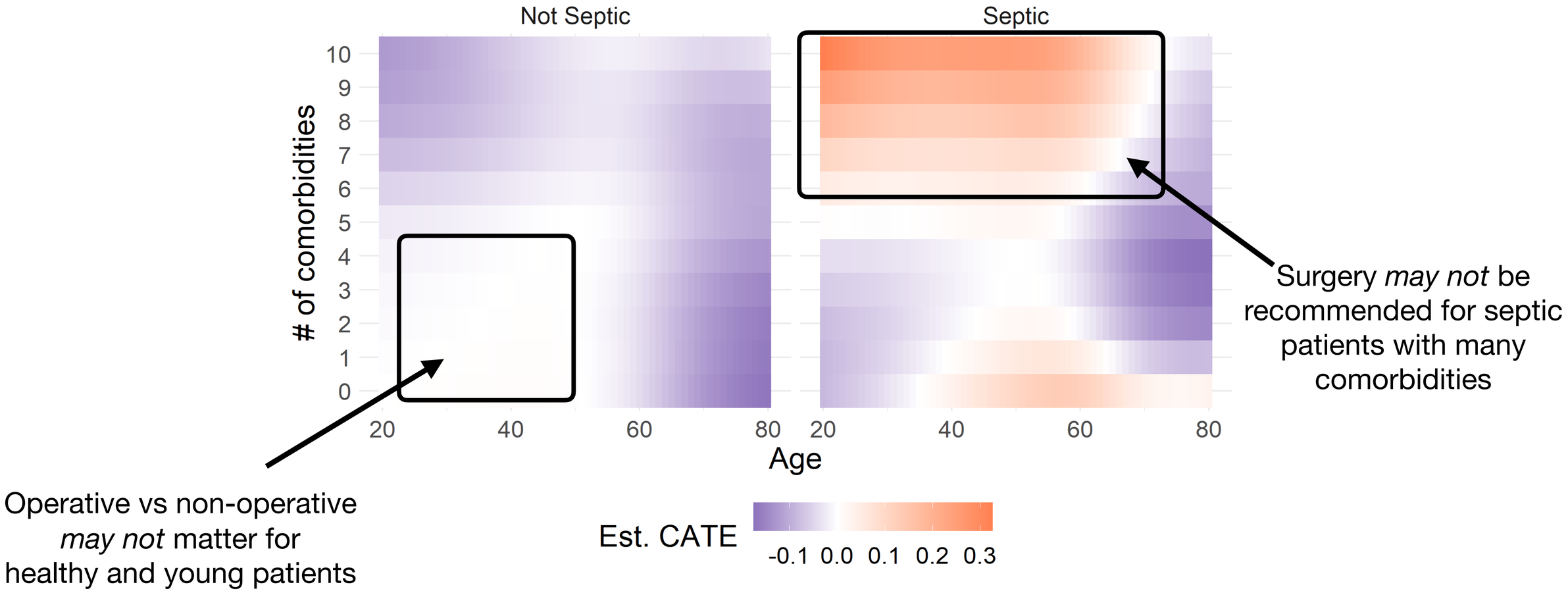


Fig: Heatmap of cond. LATE as a function of age, comorbidities and sepsis.

Sensitivity analysis

What if our data contained defiers?

When there are defiers, LATE can take any values in the following interval:

$$E[Y(1) - Y(0) \mid \text{Complier}] \in \left[\psi_0 - \frac{\delta_1 \delta_2}{\delta_3}, \psi_0 + \frac{\delta_1 \delta_2}{\delta_3}\right]$$

What if our data contained defiers?

When there are defiers, LATE can take any values in the following interval:

$$E[Y(1) - Y(0) \mid \text{Complier}] \in \left[\psi_0 - \frac{\delta_1\delta_2}{\delta_3}, \psi_0 + \frac{\delta_1\delta_2}{\delta_3}\right]$$

$\delta_1 := P(\text{Defier})$

$\delta_2 := E[Y(1) - Y(0) \mid \text{Defier}] - E[Y(1) - Y(0) \mid \text{Complier}]$

$\delta_3 := P(\text{Complier}) - P(\text{Defier})$

What if our data contained defiers?

When there are defiers, LATE can take any values in the following interval:

$$E[Y(1) - Y(0) \mid \text{Complier}] \in \left[\psi_0 - \frac{\delta_1 \delta_2}{\delta_3}, \psi_0 + \frac{\delta_1 \delta_2}{\delta_3}\right]$$

$\delta_1 := P(\text{Defier})$

$\delta_2 := E[Y(1) - Y(0) \mid \text{Defier}] - E[Y(1) - Y(0) \mid \text{Complier}]$

$\delta_3 := P(\text{Complier}) - P(\text{Defier})$

Angrist, et al (1996)

What if our data contained defiers?

Recall LATE $\leq \psi_0 + \frac{\delta_1 \delta_2}{\delta_3}$

and $\widehat{\psi_n} = -0.05$

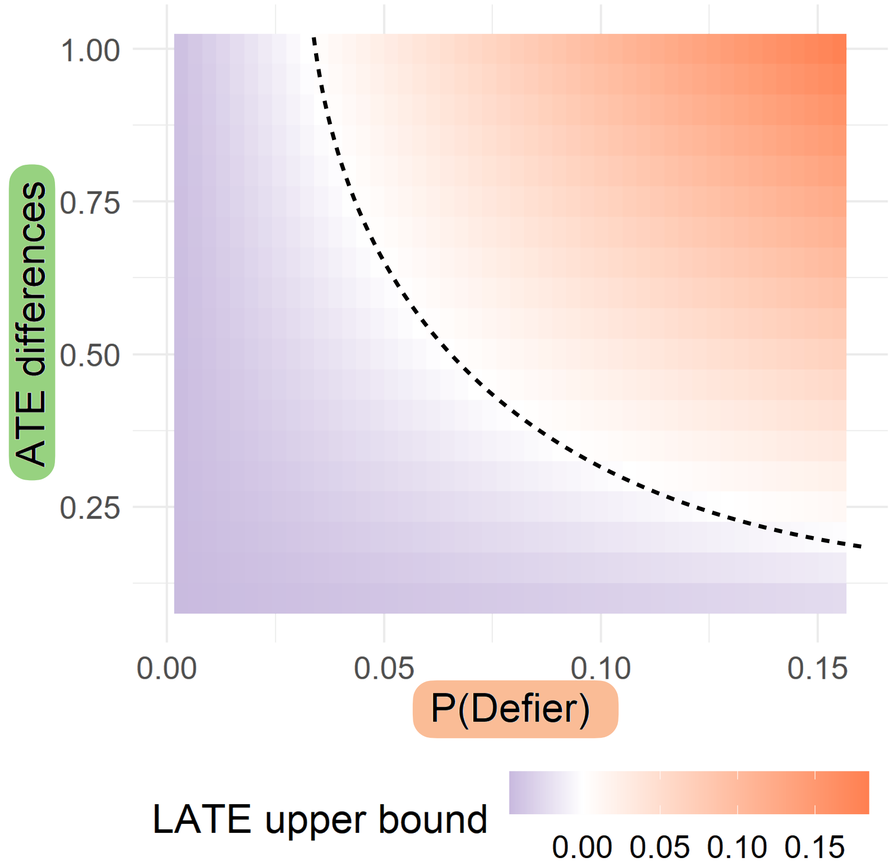
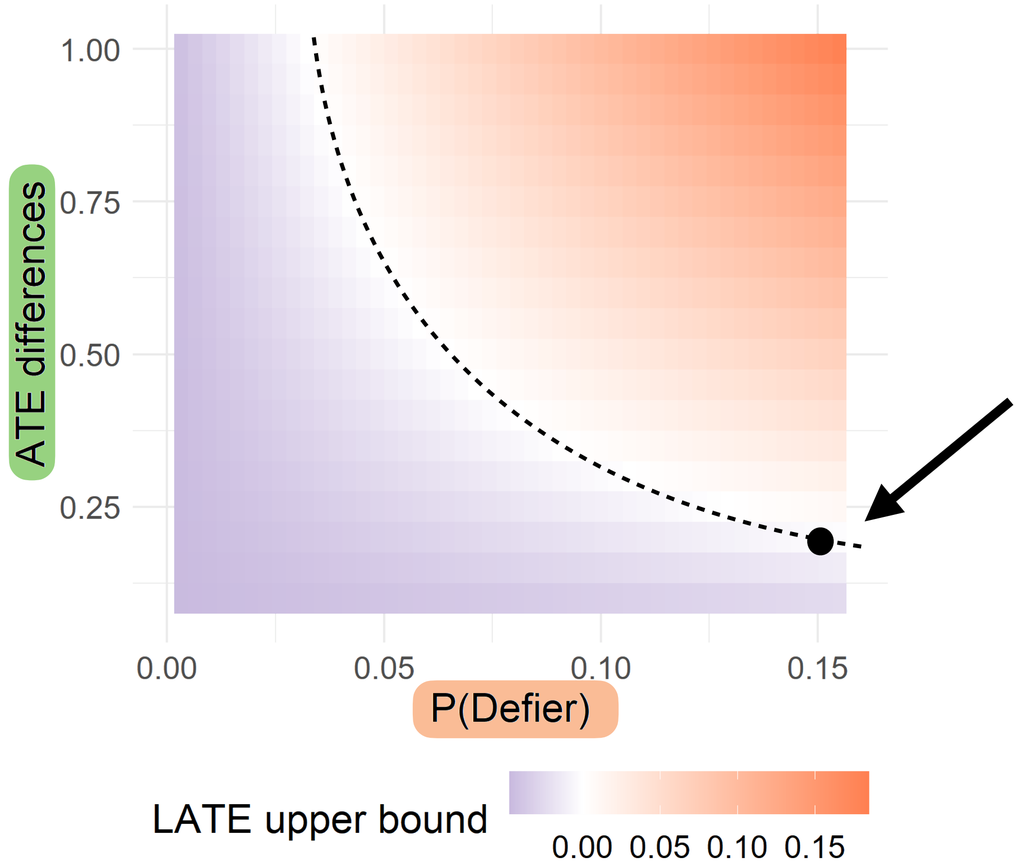


Fig: Heatmap of LATE upper bound as a function of two sensitivity parameters.

What if our data contained defiers?

Recall LATE $\leq \psi_0 + \frac{\delta_1 \delta_2}{\delta_3}$

and $\widehat{\psi_n} = -0.05$



15% of the studied population is defiers.

Surgery is 25% risker for defiers than compliers.

Fig: Heatmap of LATE upper bound as a function of two sensitivity parameters.

Conclusion

Morals of the story

We *can* estimate treatment effect under unmeasured confouding using an IV.

- Although it is an effect for compliers only, we can investigate their characteristics.
- We can conduct the sensitivity analysis against the no-defiers assumption.

We *should* look at conditional LATE.

- The conclusion from LATE can be misleading and may not be applicable to most people.

Thank you.

Appendix

A.1 Computing surgeon's "preference"

1. For each surgeon, we split his or her patient population in half.

2. Using one half of the data, we calculate the proportion of times a surgeon operates.

3. Surgeons were removed from our study if they did not perform at least 5 operations per year.

4. The resulting variable is binarized at median.

A.2 Definition of adverse outcomes

1. Prolonged length of stay is an indicator that equals one when the hospital and operation-specific length of stay is greater than the 75th percentile (5790 cases)

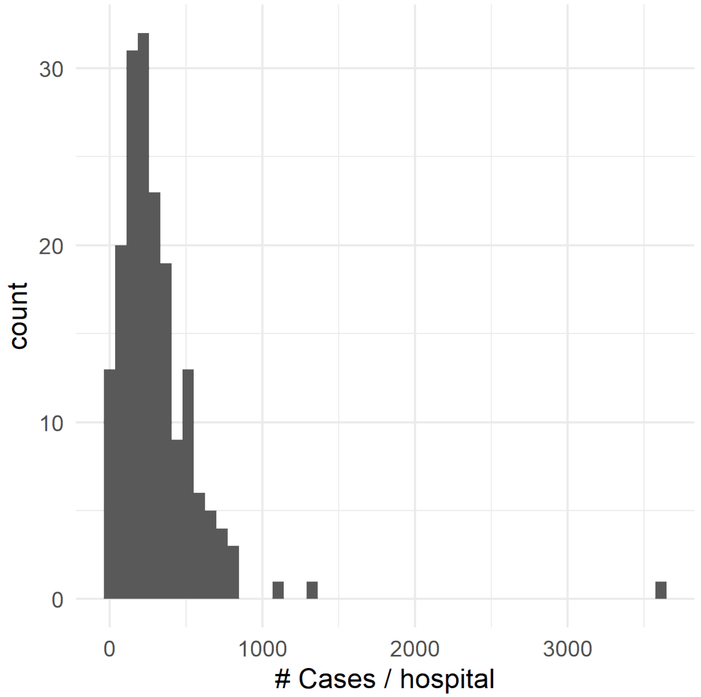
2. Include mortality as an adverse outcomes (332 cases)

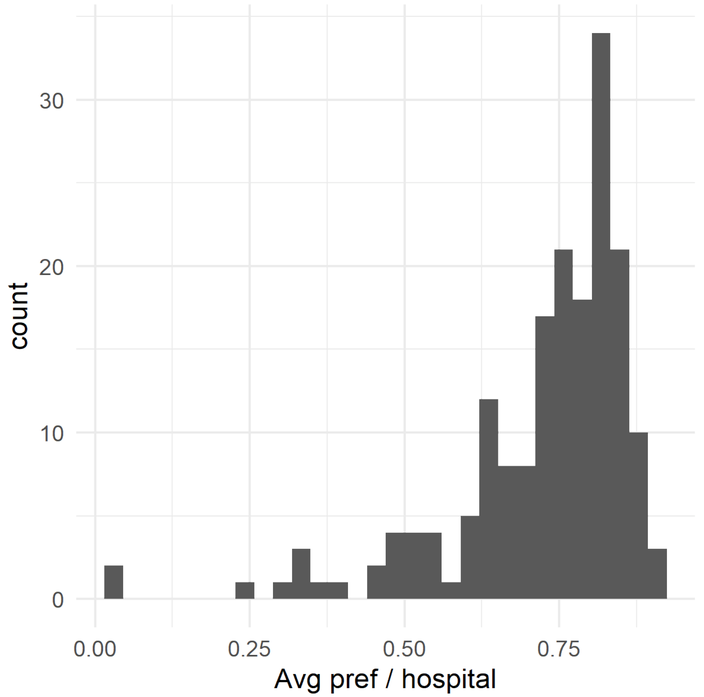
3. Together we have 5971 cases of adverse outcomes. (i.e., Prolonged LOS or mortality)

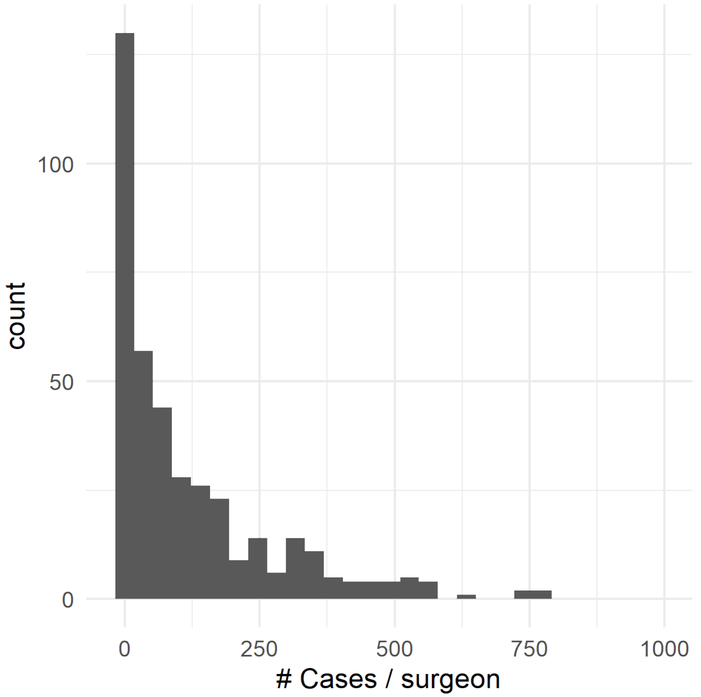
A.3 More description of the data

1. 181 unique hospitals and 397 unique surgeons.
2. IV strength varies between hospitals (approx 0.2~0.9)
3. Avg. preference per hospital varies (approx 0.03~0.90).
4. Covariates include 31 comorbidities based on Elixhauser indices, types of medical insurance, types of ethnicity (White, Black, Hispanic, and others), gender, the presence of sepsis, and disabilities. In addition to these binary variables, we also have the total number of comorbidities (count), the age of patients (continuous), and the surgeon's years of experience (continuous).

A.4 Cases per hospitals or surgeons







A.5 Definition of IVs

1. Relevance: $\mathbb{P}(A(1) = A(0)) \neq 1$
2. Exclusion restriction: $Y(z, a) = Y(a)$
3. Unconfounded IV: $Z \perp (A(z), Y(z)) \mid W$
4. Monotonicity: $\mathbb{P}(A(1) < A(0)) = 0$

A.6 Identification of LATE.

1. A valid IV (relevance, exclusion restriction, unconfounded IV)
2. **Monotonicity** (i.e., no defiers)
3. $0 < P(Z = 1 \mid W) < 1$ with prob. 1

$$E[Y(1) - Y(0) \mid \text{Complier}] = \frac{E[E[Y \mid Z = 1, W]] - E[E[Y \mid Z = 0, W]]}{E[E[A \mid Z = 1, W]] - E[E[A \mid Z = 0, W]]}$$

A.6 Identification of LATE.

1. A valid IV (relevance, exclusion restriction, unconfounded IV)
2. **Monotonicity** (i.e., no defiers)
3. $0 < P(Z = 1 \mid W) < 1$ with prob. 1

Imbens and Angrist (1994)

$$E[Y(1) - Y(0) \mid \text{Complier}] = \frac{E[E[Y \mid Z = 1, W]] - E[E[Y \mid Z = 0, W]]}{E[E[A \mid Z = 1, W]] - E[E[A \mid Z = 0, W]]}$$

A.7 Identification of cond. LATE.

Let $V \subseteq W$ (the subset of covariates).

1. Valid IV* and 2. Monotonicity (i.e., no defiers).
   *Relevance needs to be strengthen.

3. $0 < P(Z = 1 \mid W) < 1$ with prob. 1.

Abadie (2003)

$$E_0[Y(1) - Y(0) \mid \text{Complier}, V = v]$$
$$= \frac{E_0[E_0[Y \mid Z = 1, W] \mid V = v] - E_0[E_0[Y \mid Z = 0, W] \mid V = v]}{E_0[E_0[A \mid Z = 1, W] \mid V = v] - E_0[E_0[A \mid Z = 0, W] \mid V = v]}$$

A.8 A nonparametric estimator of LATE.

$$\widehat{\psi_n} := \frac{n^{-1}\sum_{i=1}^{n}\phi_{n,1}(O_i;\widehat{\mu_n},\widehat{\pi_n})}{n^{-1}\sum_{i=1}^{n}\phi_{n,2}(O_i;\widehat{\lambda_n},\widehat{\pi_n})}$$

Where

$$\phi_{n,1}(O_i;\widehat{\mu_n},\widehat{\pi_n}) := \left\{\frac{Z_i}{\widehat{\pi_n}(W_i)} - \frac{1-Z_i}{1-\widehat{\pi_n}(W_i)}\right\}\left\{Y_i - \widehat{\mu_n}(Z_i,W_i)\right\} + \widehat{\mu_n}(1,W_i) - \widehat{\mu_n}(0,W_i)$$

$$\phi_{n,2}(O_i;\widehat{\lambda_n},\widehat{\pi_n}) := \left\{\frac{Z_i}{\widehat{\pi_n}(W_i)} - \frac{1-Z_i}{1-\widehat{\pi_n}(W_i)}\right\}\left\{A_i - \widehat{\lambda_n}(Z_i,W_i)\right\} + \widehat{\lambda_n}(1,W_i) - \widehat{\lambda_n}(0,W_i)$$

A.9 Delta method for influence functions.

We can combine multiple asymptotic linear estimators as follows:

$$
\begin{aligned}
& h\left(\widehat{\psi}_{n,1}, \widehat{\psi}_{n,2}\right) - h\left(\psi_{0,1}, \psi_{0,2}\right) \\
& \quad = \frac{1}{n} \sum_{i=1}^{n} \nabla h\left(\psi_{0,1}, \psi_{0,2}\right)^T \left[\phi^*_{0,1}(O_i), \phi^*_{0,2}(O_i)\right] + o_p(n^{-1/2}) \\
& \quad := \widetilde{\phi}^*_0(O_i)
\end{aligned}
$$

This is known as Delta method for influence functions.

We heavily use this property for $h(u, v) = u/v$.

A.10 Influence function for covariate profile

$$\mathbb{P}(I(V=v) \mid A(1) > A(0)) = \frac{E_0\left[I(V=v)\left\{E_0[A \mid Z=1,W] - E_0[A \mid Z=0,W]\right\}\right]}{E_0[E_0[A \mid Z=1,W] - E_0[A \mid Z=0,W]]}$$

$$\mathbb{P}(I(V=v) \mid A(1) = A(0) = 1) = \frac{E_0[I(V=v)E_0[A \mid Z=0,W]]}{E_0[E_0[A \mid Z=0,W]]}$$

$$\mathbb{P}(I(V=v) \mid A(1) = A(0) = 0) = \frac{E_0[I(V=v)E_0[A \mid Z=1,W]]}{E_0[E_0[A \mid Z=1,W]]}$$

A.11 Influence function for LATE.

$$\phi_1 := O \mapsto \left\{ \frac{Z}{\pi_0(W)} - \frac{1-Z}{1-\pi_0(W)} \right\} \left\{ Y - \mu_0(Z,W) \right\} + \mu_0(1,W) - \mu_0(0,W)$$

$$\phi_2 := O \mapsto \left\{ \frac{Z}{\pi_0(W)} - \frac{1-Z}{1-\pi_0(W)} \right\} \left\{ A - \lambda_0(Z,W) \right\} + \lambda_0(1,W) - \lambda_0(0,W)$$

$$\widetilde{\phi_0^*}(O;\mu_0,\lambda_0,\pi_0) := \frac{1}{\mathbb{E}_0[\phi_2(O)]} \left(\phi_1(O) - \psi_0 \phi_2(O) \right)$$

A simple consequence of Delta method

A.12 Nonparametric estimator for covariate profile

$$\phi_2 := O \mapsto \left\{ \frac{Z}{\pi_0(W)} - \frac{1-Z}{1-\pi_0(W)} \right\} \{A - \lambda_0(Z, W)\} + \lambda_0(1, W) - \lambda_0(0, W)$$

$$\frac{E_0\left[I(V=v)\{E_0[A \mid Z=1, W] - E_0[A \mid Z=0, W]\}\right]}{E_0[E_0[A \mid Z=1, W] - E_0[A \mid Z=0, W]]} = \frac{E_0 I(V=v)\phi_2(O)}{E_0\phi_2(O)}$$

A.13 Nonparametric estimator for covariate profile

$$\phi_2^{(0)} := O \mapsto \frac{1-Z}{1-\pi_0(W)} \left\{ A - \lambda_0(Z,W) \right\} + \lambda_0(0,W)$$

$$\frac{E_0[I(V=v)E_0[A \mid Z=0,W]]}{E_0[E_0[A \mid Z=0,W]]} = \frac{E_0 I(V=v)\phi_2^{(0)}(O)}{E_0 \phi_2^{(0)}(O)}$$

A.14 Nonparametric estimator for covariate profile

$$\phi_2^{(1)} := O \mapsto \frac{Z}{\pi_0(W)} \left\{ A - \lambda_0(Z, W) \right\} + \lambda_0(1, W)$$

$$\frac{E_0[I(V = v) E_0[A \mid Z = 1, W]]}{E_0[E_0[A \mid Z = 1, W]]} = \frac{E_0 I(V = v) \phi_2^{(1)}(O)}{E_0 \phi_2^{(1)}(O)}$$

A.15 Profiling with continuous RVs

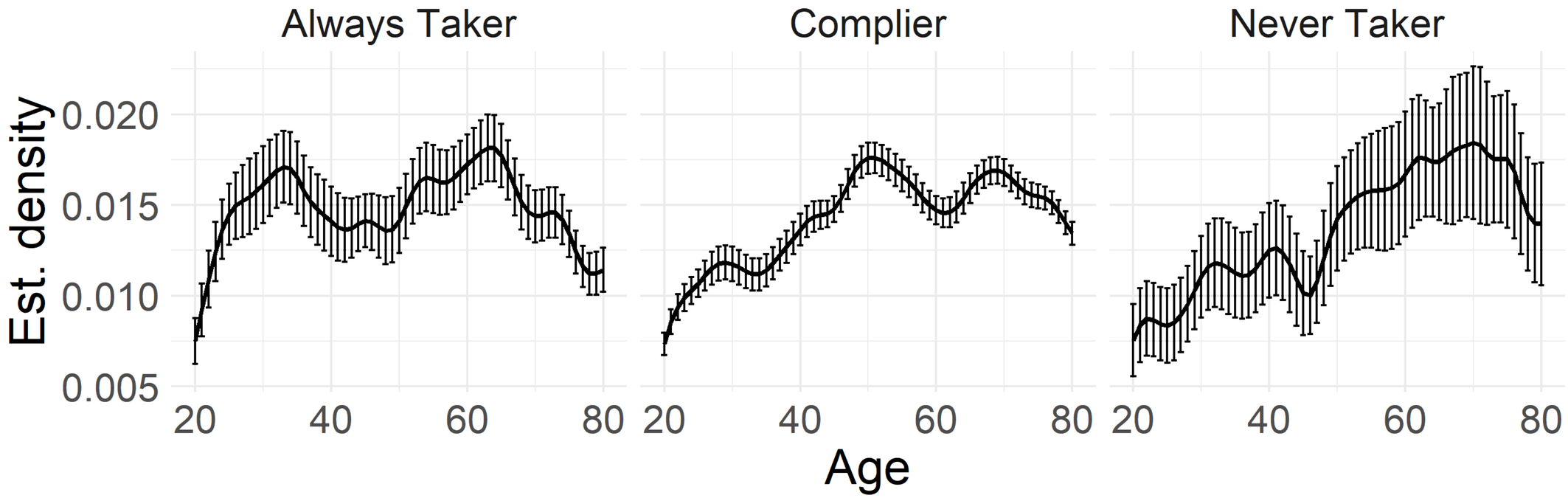


Fig: Estimated conditional density of age for each patient type. Vertical bars indicate pointwise 95% CIs.

A.16 An algorithm for LATE

Step 1: Use sample-splitting to construct machine learning estimators: $\widehat{\mu_n}$, $\widehat{\lambda_n}$, $\widehat{\pi_n}$.

Step 2: Plug-in to the (uncentered) influence functions: $\{\phi_{n,1}(O_i; \widehat{\mu_n}, \widehat{\pi_n})\}_{i=1}^n$ and $\{\phi_{n,2}(O_i; \widehat{\lambda_n}, \widehat{\pi_n})\}_{i=1}^n$.

Step 3: Return $\widehat{\psi_n} := \dfrac{n^{-1}\sum_{i=1}^n \phi_{n,1}(O_i; \widehat{\mu_n}, \widehat{\pi_n})}{n^{-1}\sum_{i=1}^n \phi_{n,2}(O_i; \widehat{\lambda_n}, \widehat{\pi_n})}$.

Step 4: 95%-CI is given by $\left[\widehat{\psi_n} \pm 1.96\sqrt{Var\,\widetilde{\phi}_n^*/n}\right]$ where $\widetilde{\phi}_n^*$ is an estimate of the influence function.

A.17 An algorithm for cond. LATE

Step 1: Use sample-splitting and construct machine learning estimators: $\widehat{\mu}_n$, $\widehat{\lambda}_n$, $\widehat{\pi}_n$.

Step 2: Plug-in to the (uncentered) influence functions: $\{\phi_{n,1}(O_i; \widehat{\mu}_n, \widehat{\pi}_n)\}_{i=1}^n$ and $\{\phi_{n,2}(O_i; \widehat{\lambda}_n, \widehat{\pi}_n)\}_{i=1}^n$.

Step 3: Regress $\{\phi_{n,1}(O_i; \widehat{\mu}_n, \widehat{\pi}_n)\}_{i=1}^n$ and $\{\phi_{n,2}(O_i; \widehat{\lambda}_n, \widehat{\pi}_n)\}_{i=1}^n$ on V using (nonparametric) regression.

Step 4: Return $\widehat{\psi}_n(v)$ as the estimates of $\dfrac{\widehat{E}_0[\phi_{n,1}(O) \mid V = v]}{\widehat{E}_0[\phi_{n,2}(O) \mid V = v]}$.

A.18 Properties of LATE estimator

A.18 Properties of LATE estimator

1. Our estimator is root-n consistent.

A.18 Properties of LATE estimator

1. Our estimator is root-n consistent.

$$n^{1/2}\left(\widehat{\psi_n} - \psi_0\right) \xrightarrow{d} N\left(0, Var\widetilde{\phi_0^*}(O; \mu_0, \lambda_0, \pi_0)\right)$$

A.18 Properties of LATE estimator

1. Our estimator is root-n consistent.

$$n^{1/2}\left(\widehat{\psi_n} - \psi_0\right) \xrightarrow{d} N\left(0, Var\,\widetilde{\phi_0^*}(O;\mu_0,\lambda_0,\pi_0)\right) \implies \left[\widehat{\psi_n} \pm 1.96\sqrt{Var\,\widetilde{\phi}_n^*/n}\right]$$

A.18 Properties of LATE estimator

1. Our estimator is root-n consistent.

$$n^{1/2}\left(\widehat{\psi_n} - \psi_0\right) \xrightarrow{d} N\left(0, Var\,\widetilde{\phi_0^*}(O;\mu_0,\lambda_0,\pi_0)\right) \implies \left[\widehat{\psi_n} \pm 1.96\sqrt{Var\,\widetilde{\phi}_n^*/n}\right]$$

2. It possesses *double-robustness.*

A.18 Properties of LATE estimator

1. Our estimator is root-n consistent.

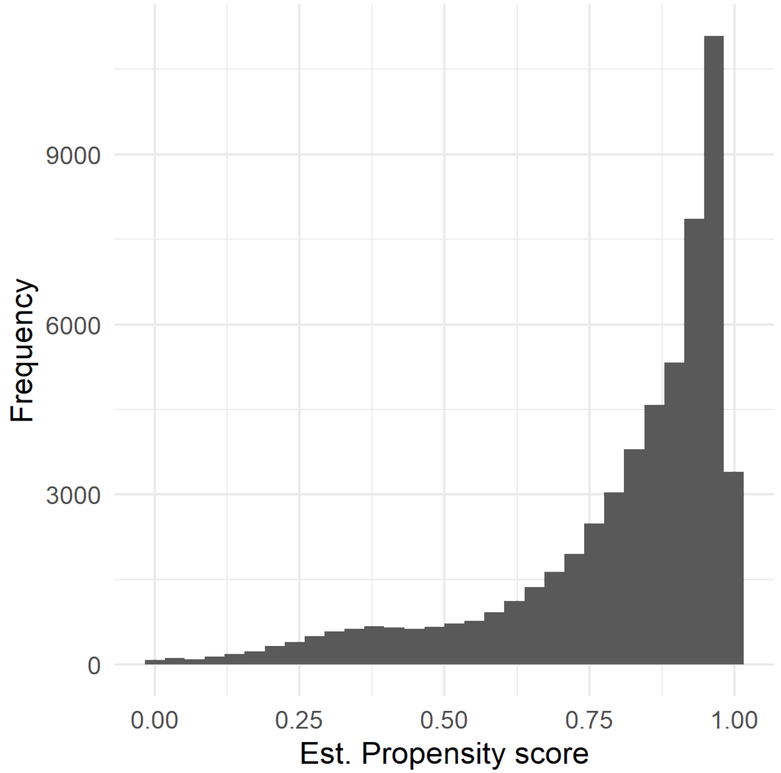
$$n^{1/2}\left(\widehat{\psi_n} - \psi_0\right) \xrightarrow{d} N\left(0, Var\,\widetilde{\phi_0^*}(O;\mu_0,\lambda_0,\pi_0)\right) \implies \left[\widehat{\psi_n} \pm 1.96\sqrt{Var\,\widetilde{\phi}_n^*/n}\right]$$

2. It possesses *double-robustness.*

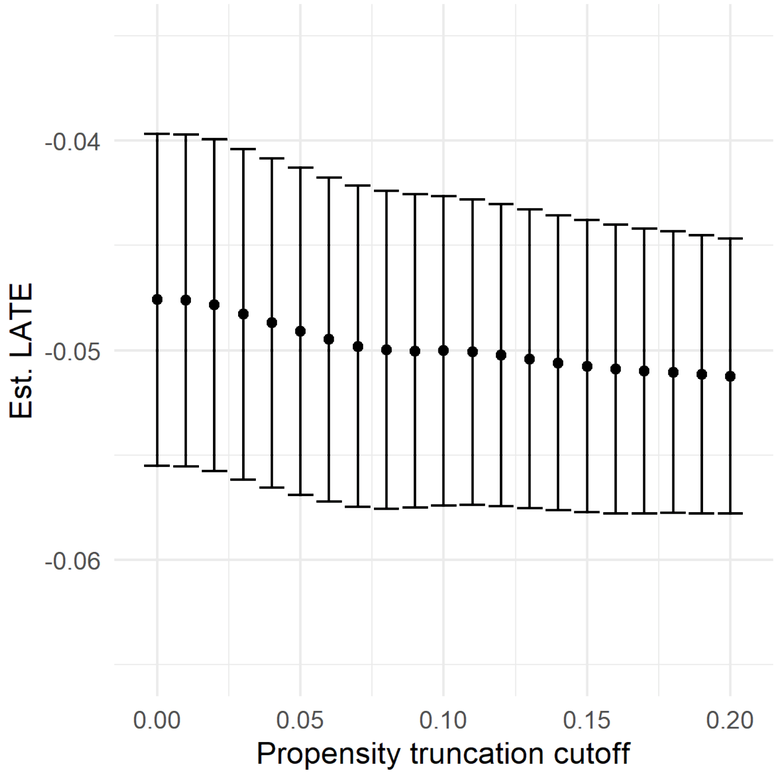
$\widehat{\psi_n}$ is root-n consistent if

$$\|\widehat{\pi_n} - \pi_0\|_2 \left(\|\widehat{\lambda_n} - \lambda_0\|_2 + \|\widehat{\mu_n} - \mu_0\|_2\right) = o_P(n^{-1/2})\,.$$

A.19 Positivity violation



Distribution of est. propensity scores



Est. LATE at different truncation values of propensity

A.20 F-test for relevance

A.20 F-test for relevance

1. Regress A on Z and W
2. Regress A on constant and W
3. Perform F-test on the nested model

A.22 Exclusion restriction

$Y(0,a) \neq Y(1,a)$ where $Y(z,a)$ is POs for both IV and trt.

$$E[Y(1) - Y(0) \mid \text{Complier}] \in \left[\psi_0 - \frac{\delta_1\delta_2}{\delta_3}, \psi_0 + \frac{\delta_1\delta_2}{\delta_3}\right]$$

A.22 Exclusion restriction

$Y(0,a) \neq Y(1,a)$ where $Y(z,a)$ is POs for both IV and trt.

$$E[Y(1) - Y(0) \mid \text{Complier}] \in \left[\psi_0 - \frac{\delta_1\delta_2}{\delta_3}, \psi_0 + \frac{\delta_1\delta_2}{\delta_3}\right]$$

$\delta_1 := 1 - P(\text{Complier})$

$\delta_2 := E[Y(1,a) - Y(0,a) \mid \text{Always taker} \cup \text{Never taker}]$

$\delta_3 := P(\text{Complier})$