Safe Policy Learning through Extrapolation

Application to Pre-trial Risk Assessment

Eli Ben-Michael

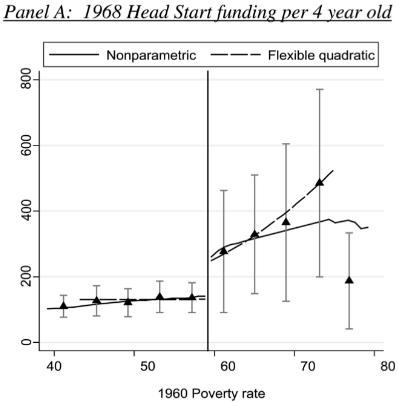
Harvard University

(joint work with Kosuke Imai, Jim Greiner, and Zhichao Jiang)

Berkeley Machine Learning and Science Forum
December 2021

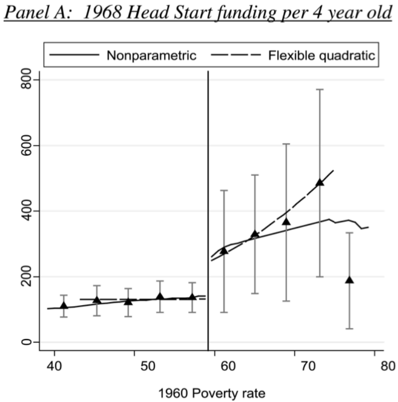
Rule-based policies – i.e. algorithms – show up everywhere

Rule-based policies – i.e. algorithms – show up everywhere

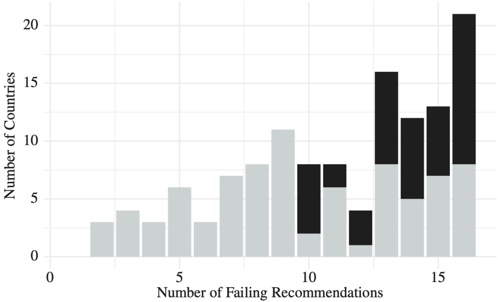


[Ludwig and Miller, 2007]

Rule-based policies – i.e. algorithms – show up everywhere

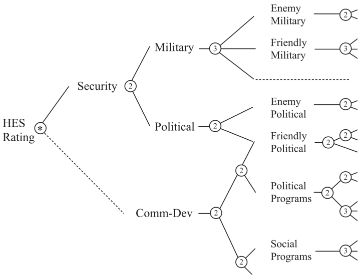


[Ludwig and Miller, 2007]



[Morse, 2019]

Rule-based policies – i.e. algorithms – show up everywhere

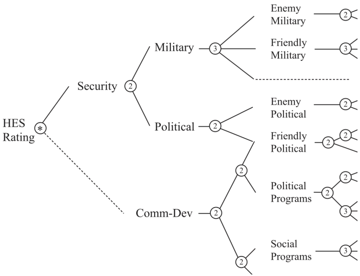


[Dell and Querubin, 2018]

Rule-based policies – i.e. algorithms – show up everywhere

Algorithmic policies are in many organizational levels

- Guiding high-level policies
- Aiding human decision makers with discretion



[Dell and Querubin, 2018]

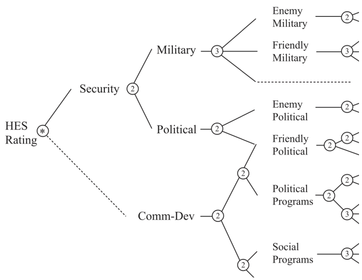
Rule-based policies – i.e. algorithms – show up everywhere

Algorithmic policies are in many organizational levels

- Guiding high-level policies
- Aiding human decision makers with discretion

Often based on known, **deterministic** rules

- Transparency and interpretability gives accountability



[Dell and Querubin, 2018]

Rule-based policies – i.e. algorithms – show up everywhere

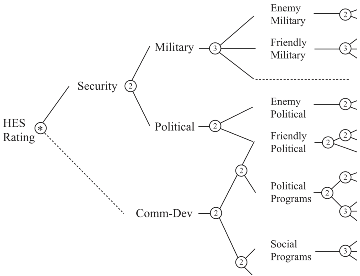
Algorithmic policies are in many organizational levels

- Guiding high-level policies
- Aiding human decision makers with discretion

Often based on known, **deterministic** rules

- Transparency and interpretability gives accountability

We study the **effects** of policies all the time



[Dell and Querubin, 2018]

Rule-based policies – i.e. algorithms – show up everywhere

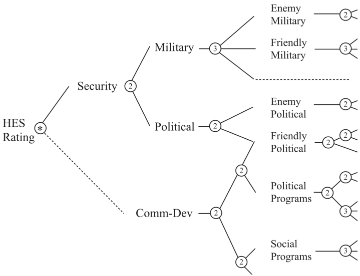
Algorithmic policies are in many organizational levels

- Guiding high-level policies
- Aiding human decision makers with discretion

Often based on known, **deterministic** rules

- Transparency and interpretability gives accountability

We study the **effects** of policies all the time



[Dell and Querubin, 2018]

But how can we **improve** the underlying rules and algorithms?

Goals and contributions

We know how to evaluate local effects of deterministic rules (e.g. RDD)

- To learn new rules we need to extrapolate from the existing status quo rule
- In some special cases we can extrapolate uniquely [Angrist and Rokkanen, 2015; Cattaneo et al., 2020]
- In general there are many ways to extrapolate from one dataset, so need to be careful! [King and Zeng, 2006]

Goals and contributions

We know how to evaluate local effects of deterministic rules (e.g. RDD)

- To learn new rules we need to extrapolate from the existing status quo rule
- In some special cases we can extrapolate uniquely [Angrist and Rokkanen, 2015; Cattaneo et al., 2020]
- In general there are many ways to extrapolate from one dataset, so need to be careful! [King and Zeng, 2006]

This paper: a safe extrapolation approach to learning rule-based policies

- Characterize all of the ways to extrapolate under assumptions on the model
- Then find the **best policy in the worst case**
- Guaranteed to be at least as good as the status quo in terms of average utility
- Incorporates uncertainty from both extrapolation and noise

Goals and contributions

We know how to evaluate local effects of deterministic rules (e.g. RDD)

- To learn new rules we need to extrapolate from the existing status quo rule
- In some special cases we can extrapolate uniquely [Angrist and Rokkanen, 2015; Cattaneo et al., 2020]
- In general there are many ways to extrapolate from one dataset, so need to be careful! [King and Zeng, 2006]

This paper: a safe extrapolation approach to learning rule-based policies

- Characterize all of the ways to extrapolate under assumptions on the model
- Then find the **best policy in the worst case**
- Guaranteed to be at least as good as the status quo in terms of average utility
- Incorporates uncertainty from both extrapolation and noise

Apply this methodology to pre-trial risk assessment algorithms

- Robust algorithms learn to classify fewer arrestees as risky

Outline

1. Background on pre-trial risk assessment algorithms
2. Methodological framework for learning new rule-based policies
3. Results when applied to pre-trial risk assessment

Pre-Trial Risk Assessment

First appearance hearings

- Judge decides pre-trial release conditions
- Cash bail? How much? Monitoring?
- Short, many in one day

First appearance hearings

- Judge decides pre-trial release conditions
- Cash bail? How much? Monitoring?
- Short, many in one day

Judges balance between

- Risk of new crime or failing to appear
- Costs of pre-trial detention to arrestee, community

First appearance hearings

- Judge decides pre-trial release conditions
- Cash bail? How much? Monitoring?
- Short, many in one day

Judges balance between

- Risk of new crime or failing to appear
- Costs of pre-trial detention to arrestee, community

Assessment scores designed to help judges

The PSA-DMF System

Public Safety Assessment (PSA) classifies 3 risks

1. Failure To Appear in court (FTA)
2. New Criminal Activity (NCA)
3. **New Violent Criminal Activity (NVCA)**

The PSA-DMF System

Public Safety Assessment (PSA) classifies 3 risks

1. Failure To Appear in court (FTA)
2. New Criminal Activity (NCA)
3. **New Violent Criminal Activity (NVCA)**

Decision Making Framework (DMF) combines scores for bail recommendation

- Signature bond vs. cash bail

The PSA-DMF System

Public Safety Assessment (PSA) classifies 3 risks

1. Failure To Appear in court (FTA)
2. New Criminal Activity (NCA)
3. **New Violent Criminal Activity (NVCA)**

Decision Making Framework (DMF) combines scores for bail recommendation

- Signature bond vs. cash bail

Millions of people across 19 states live in jurisdictions that use the PSA-DMF system

The PSA-DMF System

Public Safety Assessment (PSA) classifies 3 risks

1. Failure To Appear in court (FTA)
2. New Criminal Activity (NCA)
3. **New Violent Criminal Activity (NVCA)**

Decision Making Framework (DMF) combines scores for bail recommendation

- Signature bond vs. cash bail

Millions of people across 19 states live in jurisdictions that use the PSA-DMF system

Two goals: minimize pre-trial detention **and** NVCAs

- Ideally, no detentions and no NVCAs
- The amount of weight we apply to these two goals will be important
- This is different from predicting NVCAs well

Name: [redacted] **Spillman Name Number:** [redacted]
DOB: [redacted] **Gender:** Male
Arrest Date: 03/25/2017 **PSA Completion Date:** 03/27/2017

New Violent Criminal Activity Flag

No

New Criminal Activity Scale

1	2	3	4	5	6

Failure to Appear Scale

1	2	3	4	5	6

Charge(s):

961.41(1)(D)(1) MFC DELIVER HEROIN <3 GMS F 3

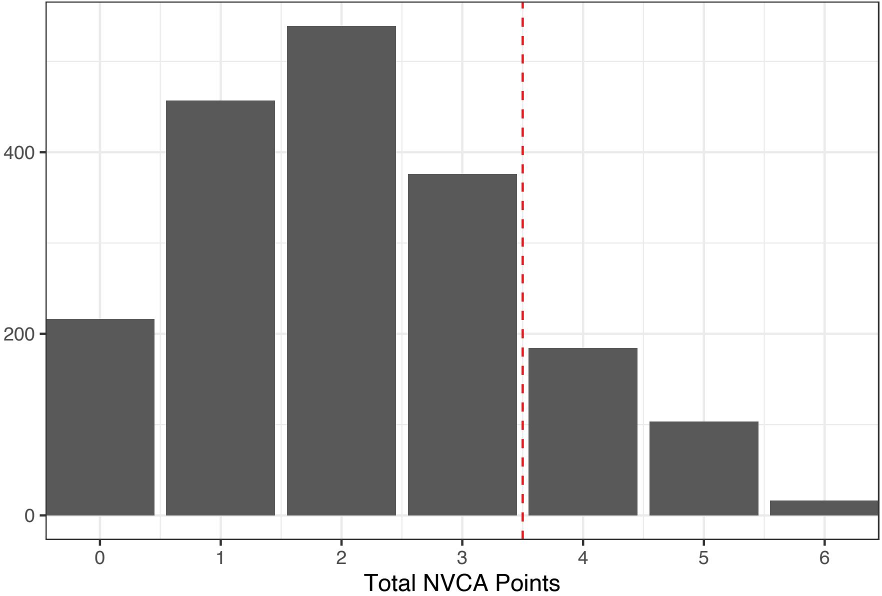
Risk Factors:	**Responses:**
1. Age at Current Arrest	23 or Older
2. Current Violent Offense	No
a. Current Violent Offense & 20 Years Old or Younger	No
3. Pending Charge at the Time of the Offense	No
4. Prior Misdemeanor Conviction	Yes
5. Prior Felony Conviction	Yes
a. Prior Conviction	Yes
6. Prior Violent Conviction	2
7. Prior Failure to Appear Pretrial in Past 2 Years	0
8. Prior Failure to Appear Pretrial Older than 2 Years	Yes
9. Prior Sentence to Incarceration	Yes

Recommendations:

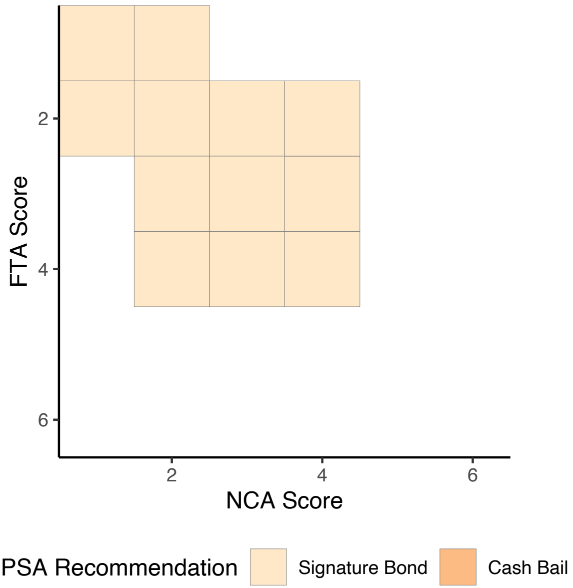
Release Recommendation - Signature bond
Conditions - Report to and comply with pretrial supervision

New Violent Criminal Arrest: Points		
PSA FACTOR	RESPONSE	POINTS
Current violent offense	No	0
	Yes	2
Current violent offense and 20 years old or younger	No	0
	Yes	1
Pending charge at the time of arrest	No	0
	Yes	1
Prior conviction (misdemeanor or felony)	No	0
	Yes	1
Prior violent conviction	No	0
	Yes, 1 or 2	1
	Yes, 3 or more	2

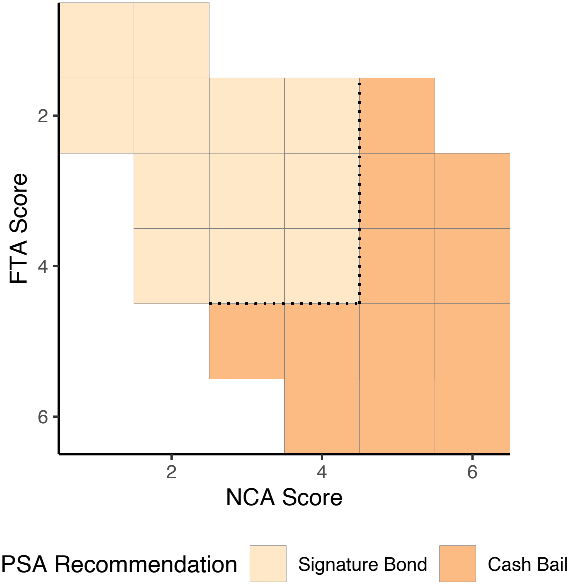
We want to change the NVCA threshold **and** how points are assigned



Also want to change the boundary for the cash bail recommendation



Also want to change the boundary for the cash bail recommendation



A randomized control trial evaluating pre-trial risk assessment

Ideal randomized experiment: randomize the algorithm's output

- Randomly flag arrestees as NVCA risk and randomly recommend cash bail
- This is totally unethical!

A randomized control trial evaluating pre-trial risk assessment

Ideal randomized experiment: randomize the algorithm's output

- Randomly flag arrestees as NVCA risk and randomly recommend cash bail
- This is totally unethical!

Instead, use a unique RCT developed to evaluate the algorithms

[Greiner et al., 2020; Imai et al., 2020]

- 1891 arrests in Dane County, WI 2017-2019, 2-year-follow-up for half the sample
- Randomly make scores and recommendations available to judges

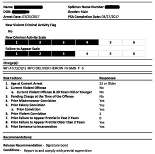
A randomized control trial evaluating pre-trial risk assessment

Ideal randomized experiment: randomize the algorithm's output

- Randomly flag arrestees as NVCA risk and randomly recommend cash bail
- This is totally unethical!

Instead, use a unique RCT developed to evaluate the algorithms
[Greiner et al., 2020; Imai et al., 2020]

- 1891 arrests in Dane County, WI 2017-2019, 2-year-follow-up for half the sample
- Randomly make scores and recommendations available to judges



Name: [redacted] **Spillman Name Number:** [redacted]
DOB: [redacted] **Gender:** Male
Arrest Date: 03/25/2017 **PSA Completion Date:** 03/27/2017

New Violent Criminal Activity Flag

No

New Criminal Activity Scale

1	2	3	4	5	6

Failure to Appear Scale

1	2	3	4	5	6

Charge(s):
961.41(1)(D)(1) MFC DELIVER HEROIN <3 GMS F 3

Risk Factors:	Responses:
1. Age at Current Arrest	23 or Older
2. Current Violent Offense	No
a. Current Violent Offense & 20 Years Old or Younger	No
3. Pending Charge at the Time of the Offense	No
4. Prior Misdemeanor Conviction	Yes
5. Prior Felony Conviction	Yes
a. Prior Conviction	Yes
6. Prior Violent Conviction	2
7. Prior Failure to Appear Pretrial in Past 2 Years	0
8. Prior Failure to Appear Pretrial Older than 2 Years	Yes
9. Prior Sentence to Incarceration	Yes

Recommendations:

Release Recommendation - Signature bond
Conditions - Report to and comply with pretrial supervision

A randomized control trial evaluating pre-trial risk assessment

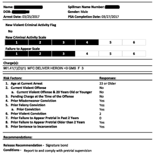
Ideal randomized experiment: randomize the algorithm's output

- Randomly flag arrestees as NVCA risk and randomly recommend cash bail
- This is totally unethical!

Instead, use a unique RCT developed to evaluate the algorithms

[Greiner et al., 2020; Imai et al., 2020]

- 1891 arrests in Dane County, WI 2017-2019, 2-year-follow-up for half the sample
- Randomly make scores and recommendations available to judges



Name: [redacted] Spillman Name Number: [redacted]
DOB: [redacted] Gender: Male
Arrest Date: 03/25/2017 PSA Completion Date: 03/27/2017

New Violent Criminal Activity Flag

No

New Criminal Activity Scale

1	2	3	4	5	6

Failure to Appear Scale

1	2	3	4	5	6

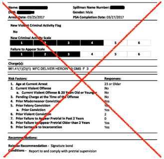
Charge(s):

961.41(1)(D)(1) MFC DELIVER HEROIN <3 GMS F 3

Risk Factors:	Responses:
1. Age at Current Arrest	23 or Older
2. Current Violent Offense	No
a. Current Violent Offense & 20 Years Old or Younger	No
3. Pending Charge at the Time of the Offense	No
4. Prior Misdemeanor Conviction	Yes
5. Prior Felony Conviction	Yes
a. Prior Conviction	Yes
6. Prior Violent Conviction	2
7. Prior Failure to Appear Pretrial in Past 2 Years	0
8. Prior Failure to Appear Pretrial Older than 2 Years	Yes
9. Prior Sentence to Incarceration	Yes

Recommendations:

Release Recommendation - Signature bond
Conditions - Report to and comply with pretrial supervision



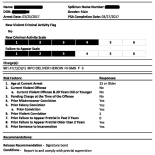
A randomized control trial evaluating pre-trial risk assessment

Ideal randomized experiment: randomize the algorithm's output

- Randomly flag arrestees as NVCA risk and randomly recommend cash bail
- This is totally unethical!

Instead, use a unique RCT developed to evaluate the algorithms
[Greiner et al., 2020; Imai et al., 2020]

- 1891 arrests in Dane County, WI   2017-2019, 2-year-follow-up for half the sample
- Randomly make scores and recommendations available to judges



Name: [redacted] Spillman Name Number: [redacted]
DOB: [redacted] Gender: Male
Arrest Date: 03/25/2017 PSA Completion Date: 03/27/2017

New Violent Criminal Activity Flag

No

New Criminal Activity Scale

1	2	3	4	5	6

Failure to Appear Scale

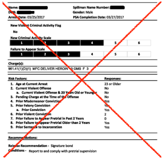
1	2	3	4	5	6

Charge(s):
961.41(1)(D)(1) MFC DELIVER HEROIN <3 GMS F 3

Risk Factors:	Responses:
1. Age at Current Arrest	23 or Older
2. Current Violent Offense	No
a. Current Violent Offense & 20 Years Old or Younger	No
3. Pending Charge at the Time of the Offense	No
4. Prior Misdemeanor Conviction	Yes
5. Prior Felony Conviction	Yes
a. Prior Conviction	Yes
6. Prior Violent Conviction	2
7. Prior Failure to Appear Pretrial in Past 2 Years	0
8. Prior Failure to Appear Pretrial Older than 2 Years	Yes
9. Prior Sentence to Incarceration	Yes

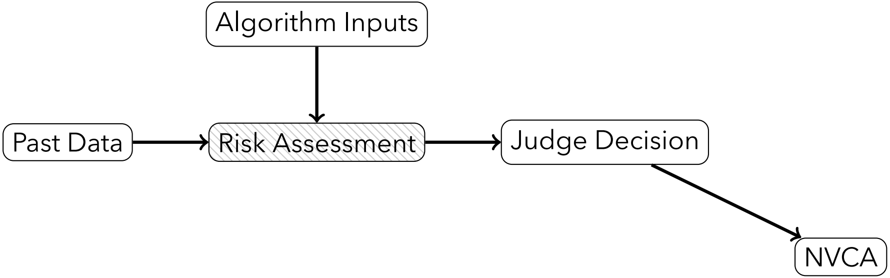
Recommendations:

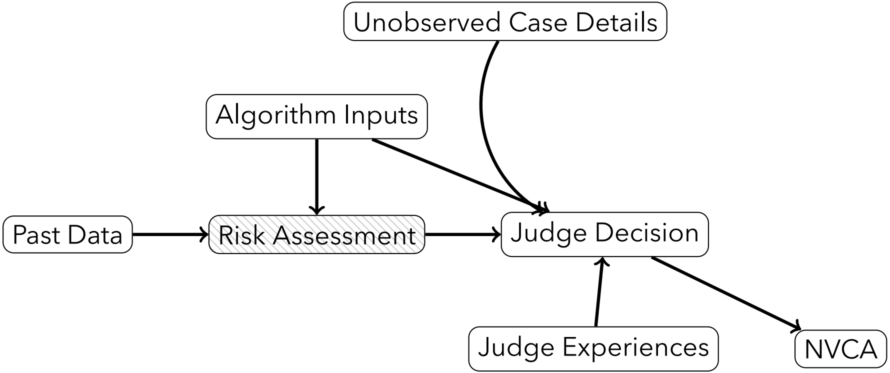
Release Recommendation - Signature bond
Conditions - Report to and comply with pretrial supervision

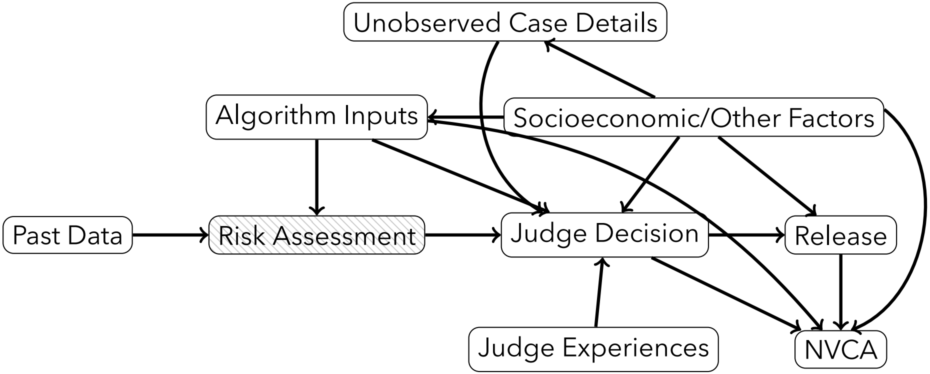


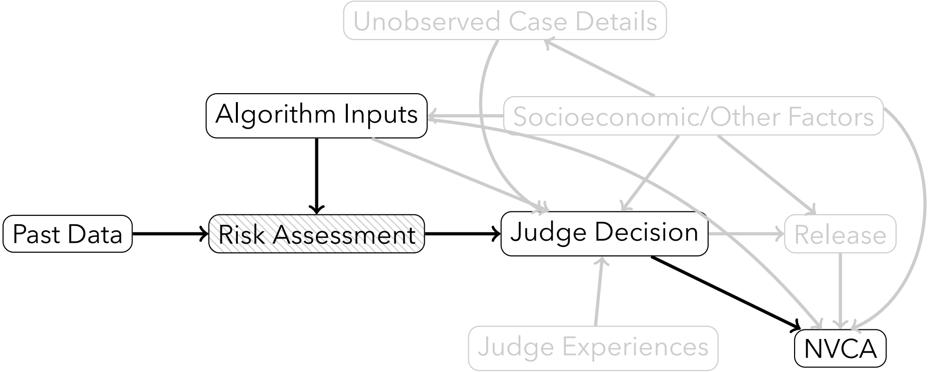
Name: [redacted] Spillman Name Number: [redacted]
DOB: [redacted] Gender: Male
Arrest Date: 03/25/2017 PSA Completion Date: 03/27/2017
New Violent Criminal Activity Flag
No
New Criminal Activity Scale
1 2 3 4 5 6
Failure to Appear Scale
1 2 3 4 5 6
Charge(s):
961.41(1)(D)(1) MFC DELIVER HEROIN <3 GMS F 3
Risk Factors: Responses:
1. Age at Current Arrest 23 or Older
2. Current Violent Offense No
a. Current Violent Offense & 20 Years Old or Younger No
3. Pending Charge at the Time of the Offense No
4. Prior Misdemeanor Conviction Yes
5. Prior Felony Conviction Yes
a. Prior Conviction Yes
6. Prior Violent Conviction 2
7. Prior Failure to Appear Pretrial in Past 2 Years 0
8. Prior Failure to Appear Pretrial Older than 2 Years Yes
9. Prior Sentence to Incarceration Yes
Recommendations:
Release Recommendation - Signature bond
Conditions - Report to and comply with pretrial supervision

We'll use this data to **learn a better algorithm**, rather than evaluate the existing one

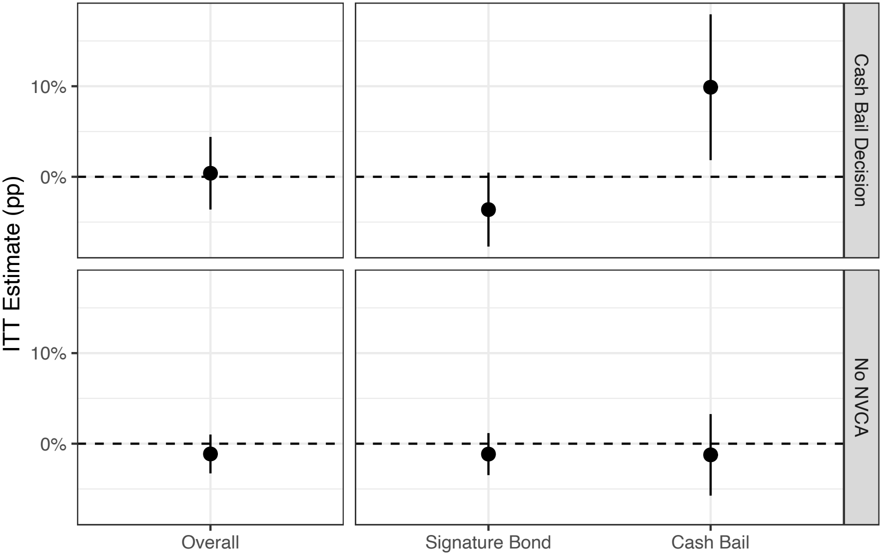

Algorithm Inputs
Past Data
Risk Assessment
Judge Decision
NVCA


Unobserved Case Details
Algorithm Inputs
Past Data
Risk Assessment
Judge Decision
Judge Experiences
NVCA

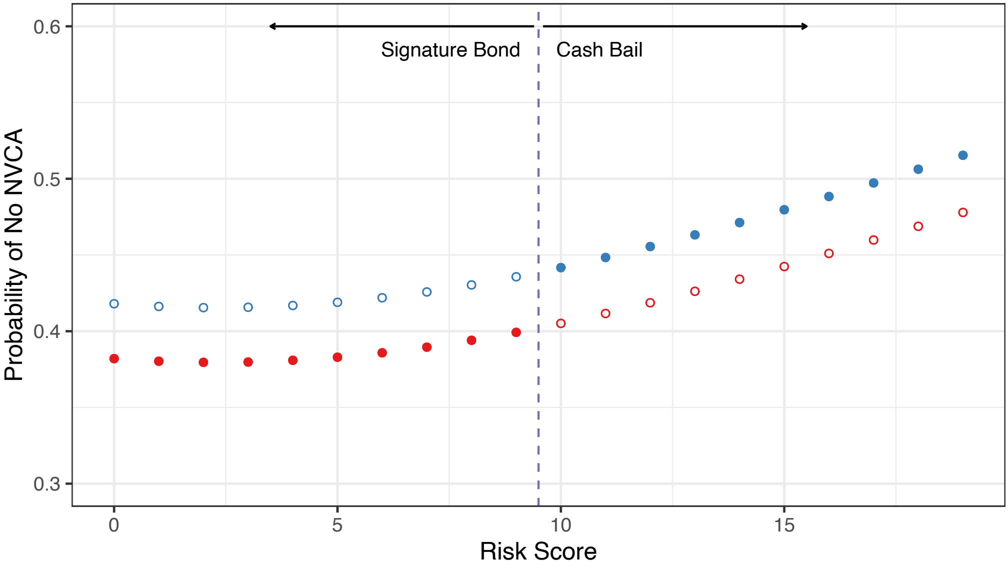

Unobserved Case Details
Algorithm Inputs
Socioeconomic/Other Factors
Past Data
Risk Assessment
Judge Decision
Release
Judge Experiences
NVCA

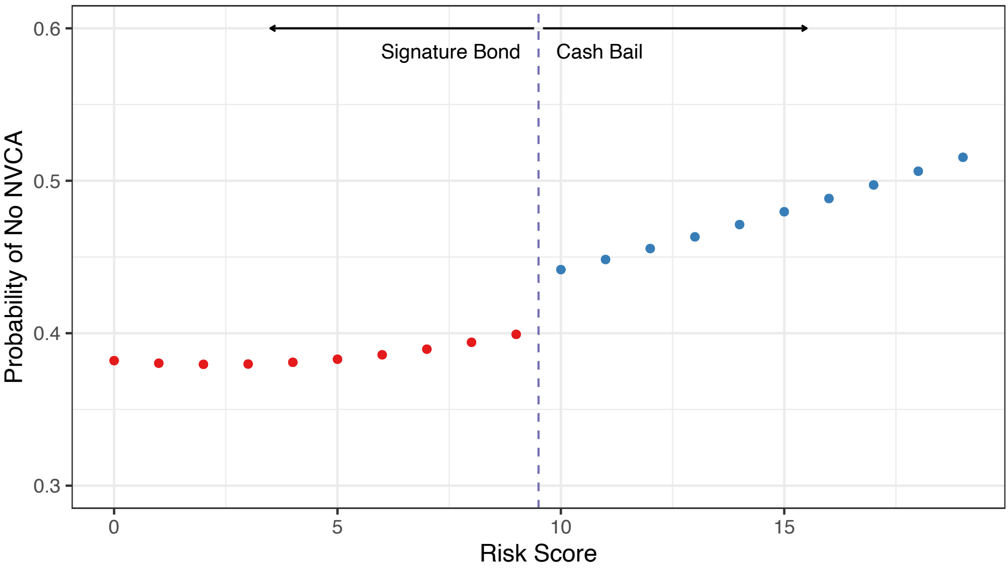

Unobserved Case Details
Algorithm Inputs
Socioeconomic/Other Factors
Past Data
Risk Assessment
Judge Decision
Release
Judge Experiences
NVCA

Suggestive but inconclusive evidence that PSA content has effects

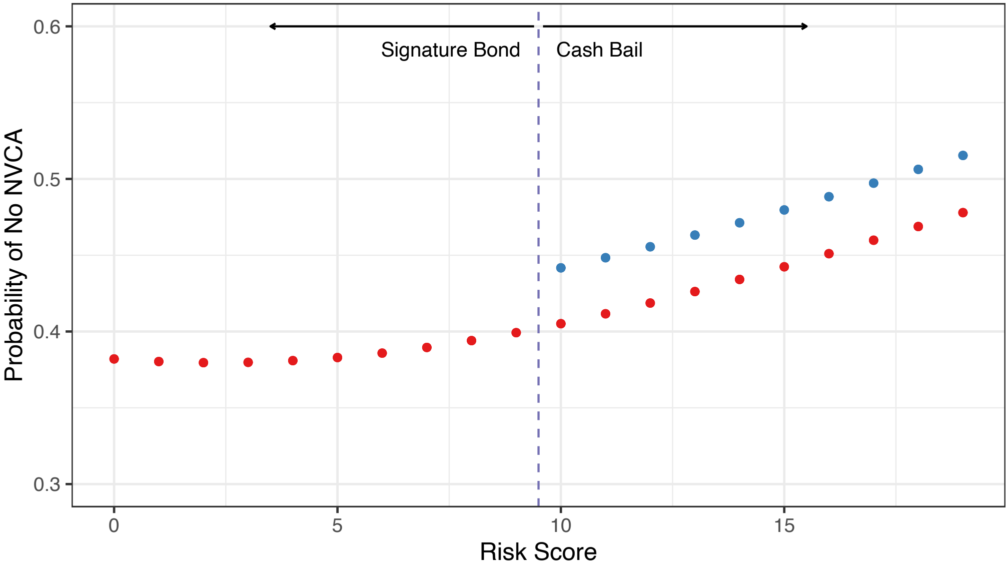


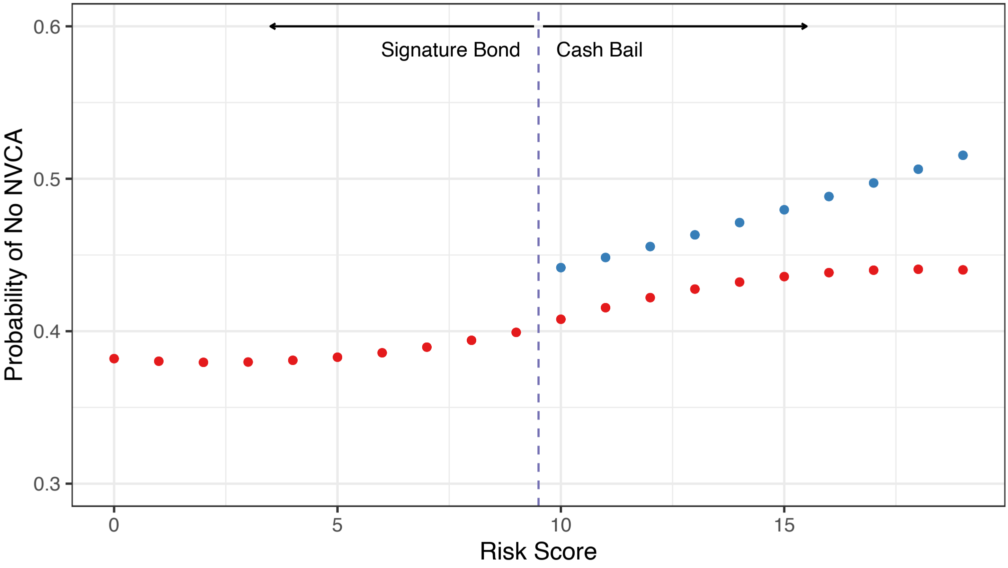
Safe Policy Learning

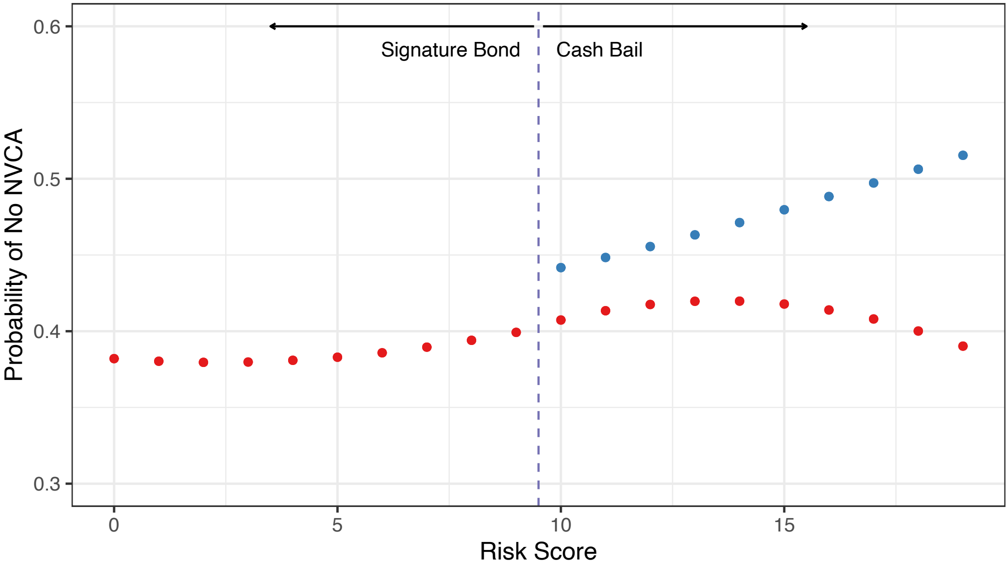

Signature Bond
Cash Bail
Probability of No NVCA
0.6
0.5
0.4
0.3
0
5
10
15
Risk Score

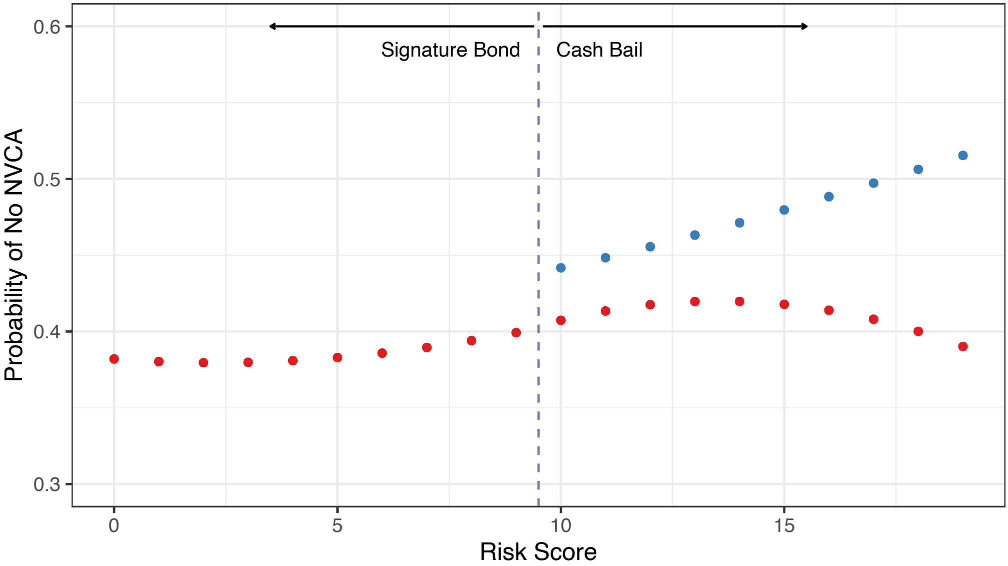


Signature Bond
Cash Bail
Probability of No NVCA
Risk Score
0.3
0.4
0.5
0.6
0
5
10
15

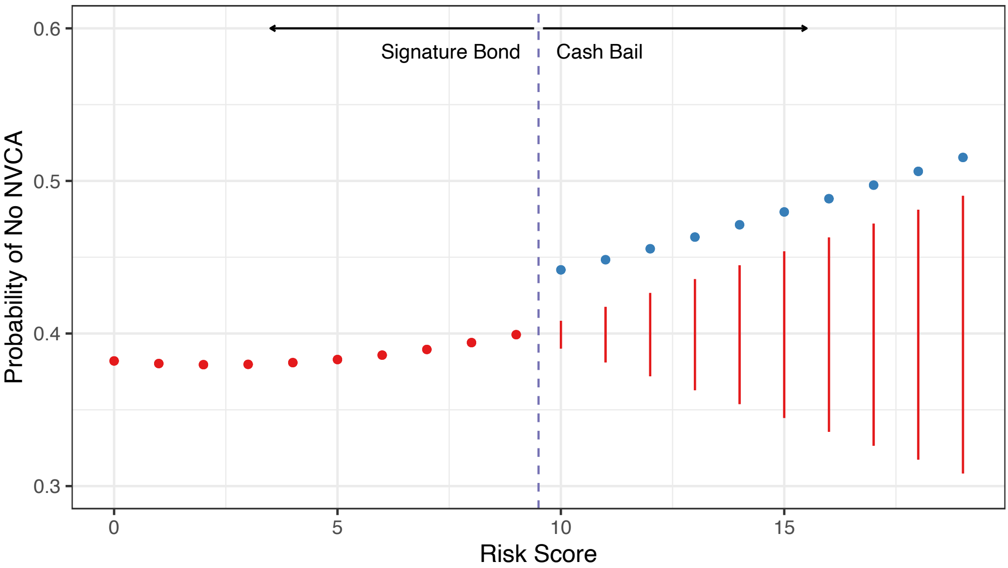

Signature Bond
Cash Bail
0.6
0.5
0.4
0.3
Probability of No NVCA
0
5
10
15
Risk Score


Signature Bond
Cash Bail
Probability of No NVCA
0.6
0.5
0.4
0.3
0
5
10
15
Risk Score

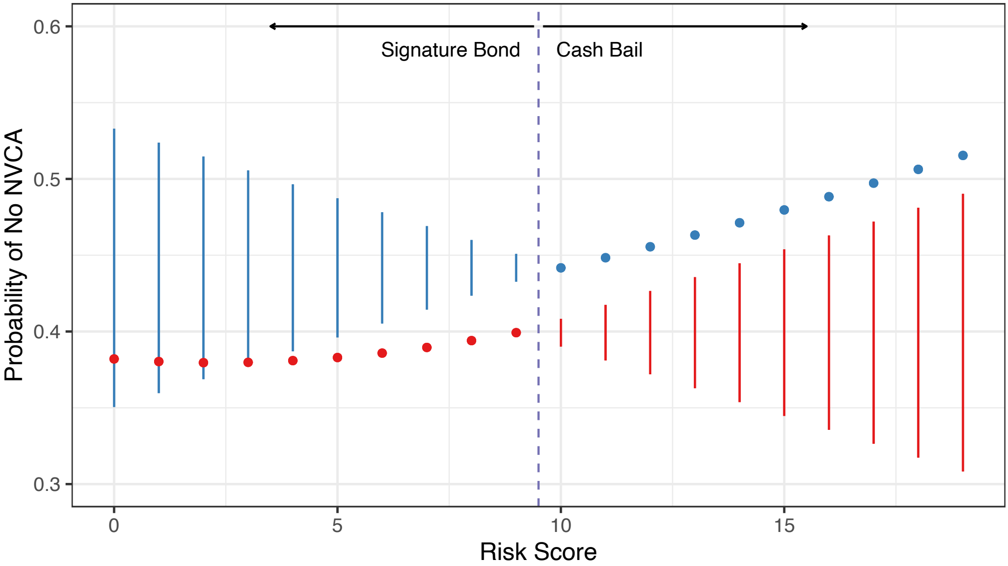

Signature Bond
Cash Bail
Probability of No NVCA
Risk Score
0.6
0.5
0.4
0.3
0
5
10
15



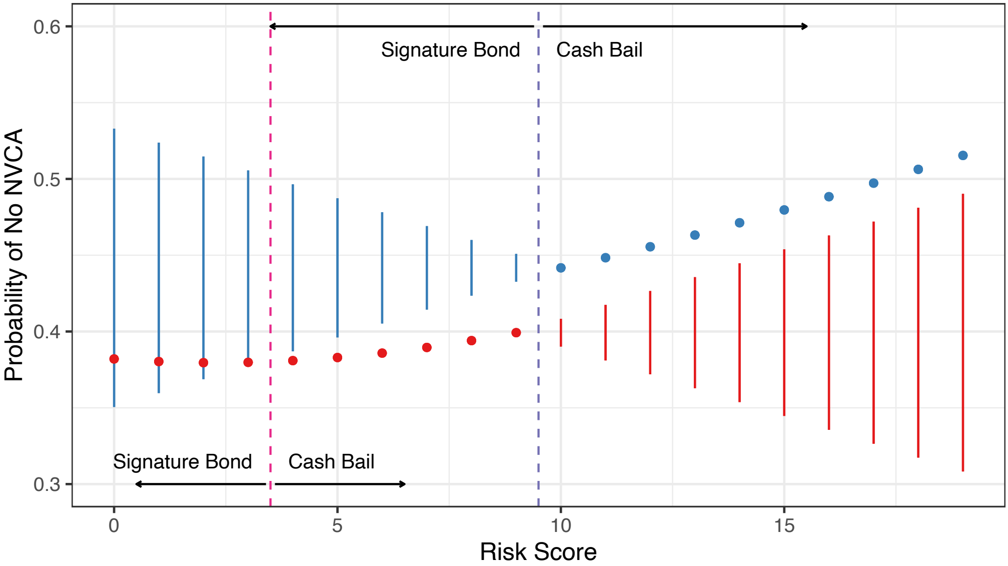
$$|f(a) - f(b)| \leq \lambda |a - b|$$



$$|f(a) - f(b)| \leq \lambda |a - b|$$



$$|f(a) - f(b)| \leq \lambda |a - b|$$



$$|f(a) - f(b)| \leq \lambda |a - b|$$

Setup

For each individual i, observe

- Covariates $X_i \in \mathcal{X}$ e.g. pre-computed risk scores or criminal history
- Action taken $A_i \in \mathcal{A} = \{0, 1\}$ e.g. trigger NVCA flag or recommend cash bail
- Binary outcome $Y_i \in \{0, 1\}$ no NVCA occurring

Deterministic status quo policy $\tilde{\pi}$, where $A_i = \tilde{\pi}(X_i)$

Setup

For each individual i, observe

- Covariates $X_i \in \mathcal{X}$ e.g. pre-computed risk scores or criminal history
- Action taken $A_i \in \mathcal{A} = \{0, 1\}$ e.g. trigger NVCA flag or recommend cash bail
- Binary outcome $Y_i \in \{0, 1\}$ no NVCA occurring

Deterministic status quo policy $\tilde{\pi}$, where $A_i = \tilde{\pi}(X_i)$

Don't observe potential outcome under action a, $Y(a)$

- Conditional expectation $m(a, x) = \mathbb{E}[Y(a) \mid X = x]$

Setup

For each individual i, observe

- Covariates $X_i \in \mathcal{X}$ e.g. pre-computed risk scores or criminal history
- Action taken $A_i \in \mathcal{A} = \{0, 1\}$ e.g. trigger NVCA flag or recommend cash bail
- Binary outcome $Y_i \in \{0, 1\}$ no NVCA occurring

Deterministic status quo policy $\tilde{\pi}$, where $A_i = \tilde{\pi}(X_i)$

Don't observe potential outcome under action a, $Y(a)$

- Conditional expectation $m(a, x) = \mathbb{E}[Y(a) \mid X = x]$

Observed outcomes are

$$Y_i = \begin{cases} Y_i(0), & \tilde{\pi}(X_i) = 0 \\ Y_i(1), & \tilde{\pi}(X_i) = 1 \end{cases}$$

	$\tilde{\pi}(X_i) = 0$	$\tilde{\pi}(X_i) = 1$
$Y_i(0)$	Y_i	?
$Y_i(1)$	?	Y_i

To find the optimal algorithmic policy, we need counterfactual information

Our goal: Find a policy with high **expected utility** (value/welfare)

$$V(\pi, m) = \mathbb{E}\left[\text{benefit} \times m(\pi(X), X) - \text{cost} \times \pi(X)\right]$$

- This is **not a prediction problem**, it involves consequences of actions
- Costs and benefits determine the objective
- *In the paper*: include Judge's decisions into utility

To find the optimal algorithmic policy, we need counterfactual information

Our goal: Find a policy with high **expected utility** (value/welfare)

$$V(\pi, m) = \mathbb{E}\left[\text{benefit} \times m(\pi(X), X) - \text{cost} \times \pi(X)\right]$$

- This is **not a prediction problem**, it involves consequences of actions
- Costs and benefits determine the objective
- *In the paper*: include Judge's decisions into utility

But how do we impute the counterfactuals?

To find the optimal algorithmic policy, we need counterfactual information

Our goal: Find a policy with high **expected utility** (value/welfare)

$$V(\pi, m) = \mathbb{E}\left[\text{benefit} \times m(\pi(X), X) - \text{cost} \times \pi(X)\right]$$

- This is **not a prediction problem**, it involves consequences of actions
- Costs and benefits determine the objective
- *In the paper*: include Judge's decisions into utility

But how do we impute the counterfactuals?

Existing work uses **stochastic** policies for identification
[e.g. Qian and Murphy, 2011; Zhao et al., 2012; Kitagawa and Tetenov, 2018; Athey and Wager, 2021]

- Inverse probability weighting or model-based imputation

To find the optimal algorithmic policy, we need counterfactual information

Our goal: Find a policy with high **expected utility** (value/welfare)

$$V(\pi, m) = \mathbb{E}\left[\text{benefit} \times m(\pi(X), X) - \text{cost} \times \pi(X)\right]$$

- This is **not a prediction problem**, it involves consequences of actions
- Costs and benefits determine the objective
- *In the paper*: include Judge's decisions into utility

But how do we impute the counterfactuals?

Existing work uses **stochastic** policies for identification
[e.g. Qian and Murphy, 2011; Zhao et al., 2012; Kitagawa and Tetenov, 2018; Athey and Wager, 2021]

- Inverse probability weighting or model-based imputation

Deterministic policies $\longrightarrow$ many ways to extrapolate and impute the counterfactual

Rather than choose one particular imputation, optimize for the worst case

We partially identify the model $m \in \mathcal{M}$, then find the best policy in the worst case

Rather than choose one particular imputation, optimize for the worst case

We partially identify the model $m \in \mathcal{M}$, then find the best policy in the worst case

$$\pi^{\text{inf}} \in \underset{\pi \in \Pi}{\operatorname{argmax}} \min_{m \in \mathcal{M}} V(\pi, m)$$

- A robust optimization approach [Bertsimas et al., 2011; Kallus and Zhou, 2021; Pu and Zhang, 2021]
- *In the paper*: with an RCT use effect relative to no policy instead of outcomes

Rather than choose one particular imputation, optimize for the worst case

We partially identify the model $m \in \mathcal{M}$, then find the best policy in the worst case

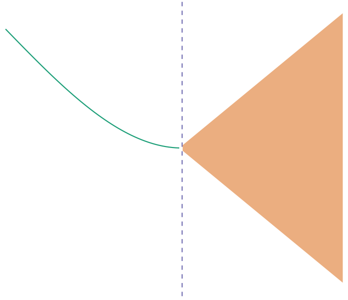
$$\pi^{\text{inf}} \in \operatorname*{argmax}_{\pi \in \Pi} \min_{m \in \mathcal{M}} V(\pi, m)$$

- A robust optimization approach [Bertsimas et al., 2011; Kallus and Zhou, 2021; Pu and Zhang, 2021]
- *In the paper*: with an RCT use effect relative to no policy instead of outcomes

Many model assumptions result in point-wise bounds

$$B_\ell(a,x) \leq m(a,x) \leq B_u(a,x)$$

- Lipschitz functions, additive models, linear models
- Similar assumptions on outcomes as RD, but globally



Easy to compute!

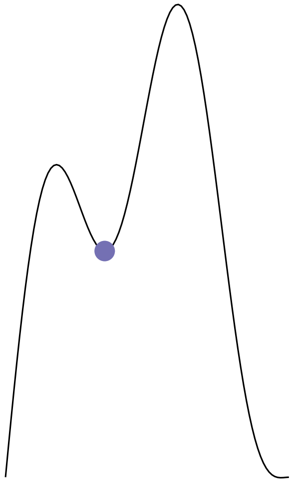
- Plug in the worst-case bound [Pu and Zhang, 2021]

$$\min_{m \in \mathcal{M}} V(\pi, m) = V(\pi, B_\ell)$$

This is a safe policy that is at least as good as the status quo on average

The value of π^{inf} is at least as high as the status quo

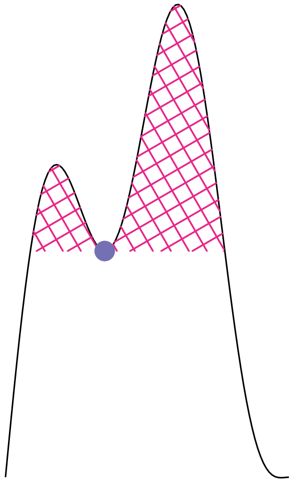
$$V(\tilde{\pi}) - V\left(\pi^{\text{inf}}\right) \leq 0$$



This is a safe policy that is at least as good as the status quo on average

The value of π^{inf} is at least as high as the status quo

$$V(\tilde{\pi}) - V\left(\pi^{\text{inf}}\right) \leq 0$$



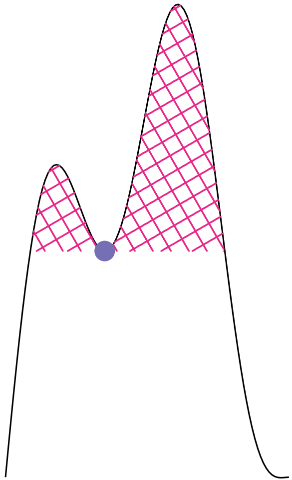
This is a safe policy that is at least as good as the status quo on average

The value of π^{inf} is at least as high as the status quo

$$V(\tilde{\pi}) - V\left(\pi^{\text{inf}}\right) \leq 0$$

Allows policy makers to know things won't get worse

- Too much uncertainty $\longrightarrow$ Fall back on status quo
- Conservative, "pessimistic" principle [Cui, 2021]



This is a safe policy that is at least as good as the status quo on average

The value of π^{inf} is at least as high as the status quo

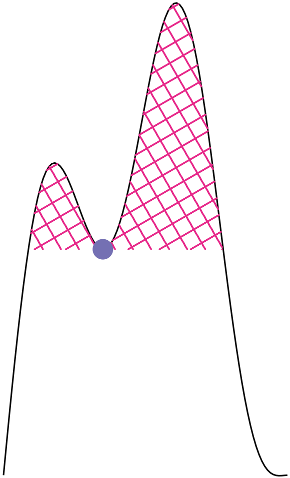
$$V(\tilde{\pi}) - V\left(\pi^{\text{inf}}\right) \leq 0$$

Allows policy makers to know things won't get worse

- Too much uncertainty $\longrightarrow$ Fall back on status quo
- Conservative, "pessimistic" principle [Cui, 2021]

Many other possible objectives in this framework [Manski, 2005]

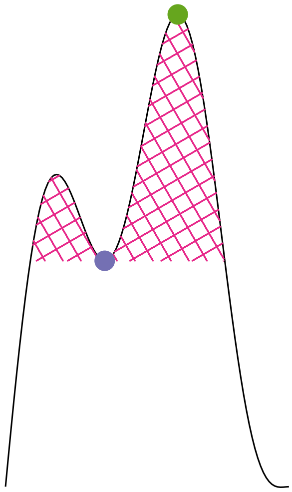
- Possibly also ensure safety for subgroups individually



The safe policy is sub-optimal, but we can bound how much

Compare to the best possible policy

$$\pi^* \in \mathrm{argmax}_{\pi \in \Pi} V(\pi)$$



The safe policy is sub-optimal, but we can bound how much

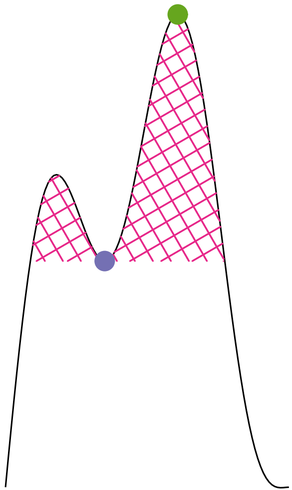
Compare to the best possible policy

$$\pi^* \in \mathrm{argmax}_{\pi \in \Pi} V(\pi)$$

Optimality gap controlled by size of $\mathcal{M}$

$$V(\pi^*) - V\left(\pi^{\text{inf}}\right) \leq u\mathbb{E}\left[\max_{a \in \mathcal{A}} B_u(a, X) - B_\ell(a, X)\right]$$

- Tighter partial identification $\longrightarrow$ better policy
- If we can extrapolate uniquely, π^{inf} is also optimal



The safe policy is sub-optimal, but we can bound how much

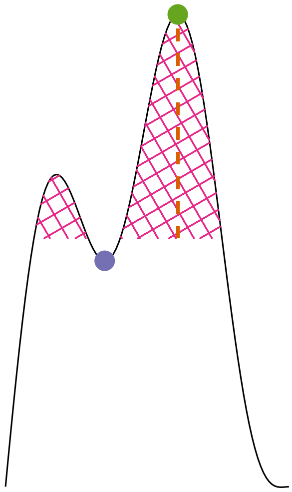
Compare to the best possible policy

$$\pi^* \in \mathrm{argmax}_{\pi \in \Pi} V(\pi)$$

Optimality gap controlled by size of $\mathcal{M}$

$$V(\pi^*) - V\left(\pi^{\mathrm{inf}}\right) \leq u\mathbb{E}\left[\max_{a \in \mathcal{A}} B_u(a, X) - B_\ell(a, X)\right]$$

- Tighter partial identification $\longrightarrow$ better policy
- If we can extrapolate uniquely, π^{inf} is also optimal



The safe policy is sub-optimal, but we can bound how much

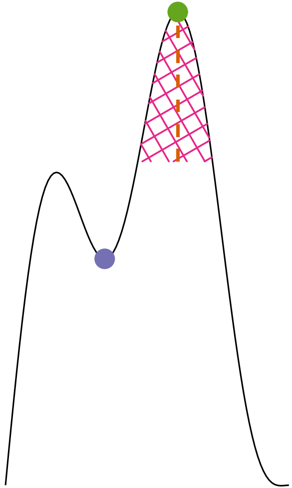
Compare to the best possible policy

$$\pi^* \in \operatorname{argmax}_{\pi \in \Pi} V(\pi)$$

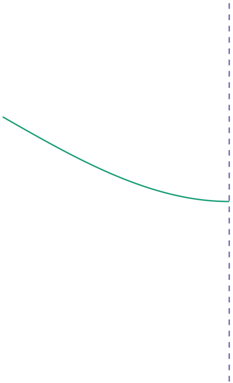
Optimality gap controlled by size of $\mathcal{M}$

$$V(\pi^*) - V\left(\pi^{\text{inf}}\right) \le u\mathbb{E}\left[\max_{a \in \mathcal{A}} B_u(a, X) - B_\ell(a, X)\right]$$

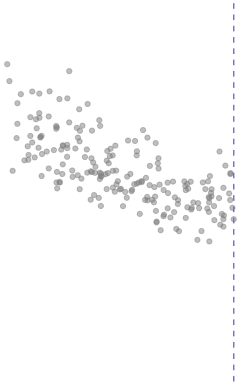
- Tighter partial identification $\longrightarrow$ better policy
- If we can extrapolate uniquely, π^{inf} is also optimal



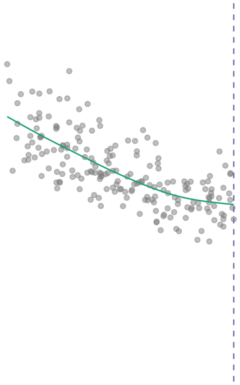
To find the safe policy empirically from data, we need to account for noise



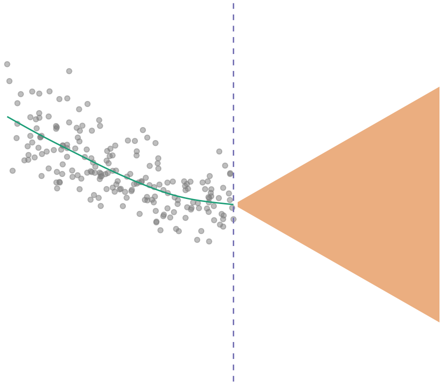
To find the safe policy empirically from data, we need to account for noise



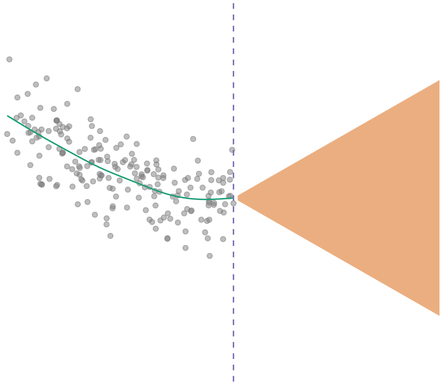
To find the safe policy empirically from data, we need to account for noise



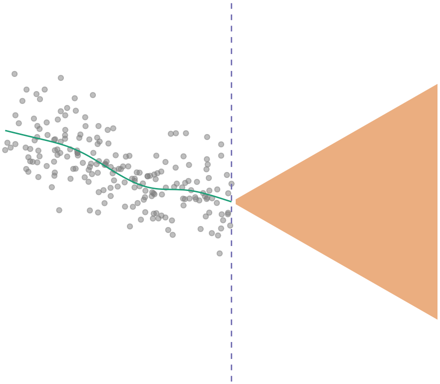
To find the safe policy empirically from data, we need to account for noise



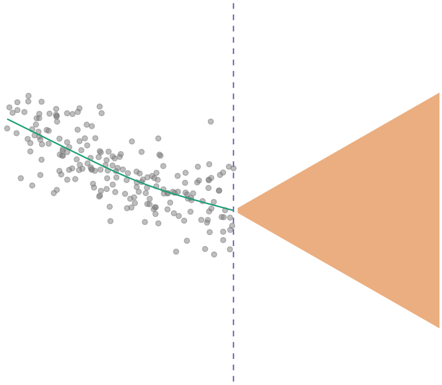
To find the safe policy empirically from data, we need to account for noise



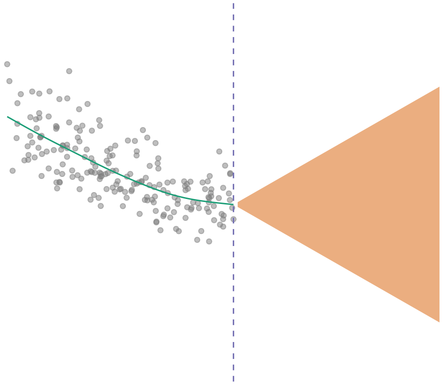
To find the safe policy empirically from data, we need to account for noise



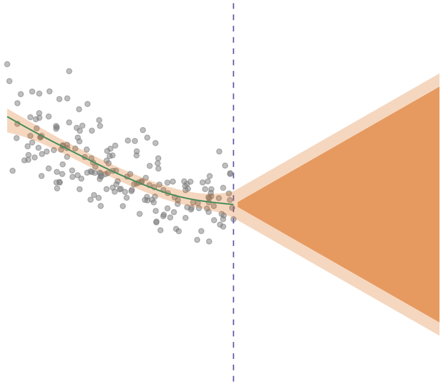
To find the safe policy empirically from data, we need to account for noise



To find the safe policy empirically from data, we need to account for noise



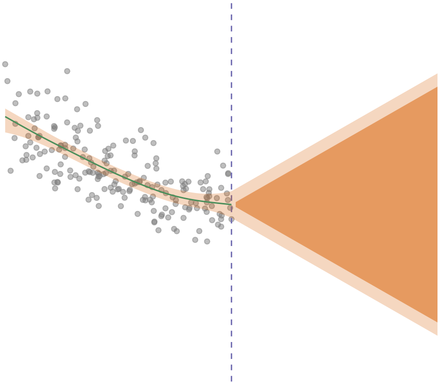
To find the safe policy empirically from data, we need to account for noise



Construct a $1 - \alpha$ confidence set $\widehat{\mathcal{M}}_n(\alpha)$

$$P\left(\mathcal{M} \in \widehat{\mathcal{M}}_n(\alpha)\right) \geq 1 - \alpha$$

To find the safe policy empirically from data, we need to account for noise



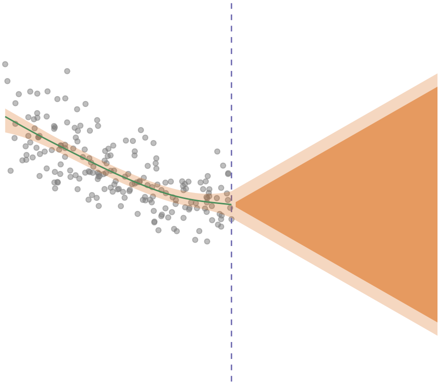
Construct a $1-\alpha$ confidence set $\widehat{\mathcal{M}}_n(\alpha)$

$$P\left(\mathcal{M} \in \widehat{\mathcal{M}}_n(\alpha)\right) \geq 1-\alpha$$

Empirical welfare maximization problem with imputed counterfactuals $\widehat{\Upsilon}_i(a)$

$$\hat{\pi} \in \underset{\pi \in \Pi}{\text{argmax}} \frac{1}{n} \sum_{i=1}^{n} \text{benefit} \times \widehat{\Upsilon}_i(\pi(X_i)) - \text{cost} \times \pi(X_i)$$

To find the safe policy empirically from data, we need to account for noise



Construct a $1-\alpha$ confidence set $\widehat{\mathcal{M}}_n(\alpha)$

$$P\left(\mathcal{M} \in \widehat{\mathcal{M}}_n(\alpha)\right) \geq 1-\alpha$$

Empirical welfare maximization problem with imputed counterfactuals $\widehat{\Upsilon}_i(a)$

$$\hat{\pi} \in \underset{\pi \in \Pi}{\operatorname{argmax}} \frac{1}{n} \sum_{i=1}^{n} \text{benefit} \times \widehat{\Upsilon}_i(\pi(X_i)) - \text{cost} \times \pi(X_i)$$

Gives a statistical safety guarantee

- Approximately holds with prob. $\geq 1-\alpha$
- Tradeoff between level α and tighter bounds

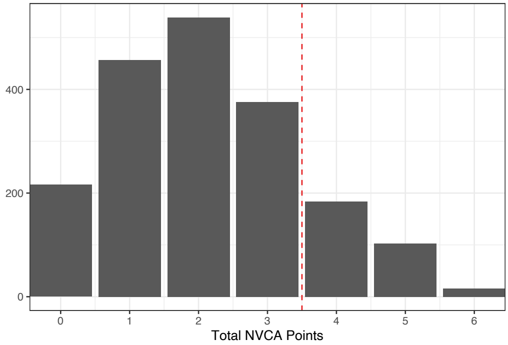
Learning a new PSA-DMF system

Learning a new NVCA Flag: choosing the algorithm

Construct a new NVCA flag using the same risk factors and structure

Learning a new NVCA Flag: choosing the algorithm

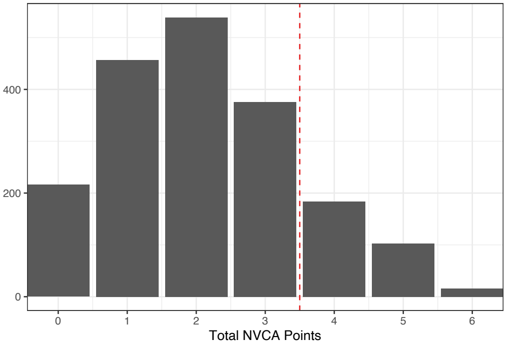
- Construct a new NVCA flag using the same risk factors and structure
  - Change the threshold but keep the number of points assigned to each factor fixed



Learning a new NVCA Flag: choosing the algorithm

Construct a new NVCA flag using the same risk factors and structure

- Change the threshold but keep the number of points assigned to each factor fixed
- Change the number of points assigned to each factor but keep the threshold fixed



New Violent Criminal Arrest: Points		
PSA FACTOR	RESPONSE	POINTS
Current violent offense	No	0
	Yes	2
Current violent offense and 20 years old or younger	No	0
	Yes	1
Pending charge at the time of arrest	No	0
	Yes	1
Prior conviction (misdemeanor or felony)	No	0
	Yes	1
Prior violent conviction	No	0
	Yes, 1 or 2	1
	Yes, 3 or more	2

How do we weigh the costs of flagging arrestees vs an NVCA?

Define utility based on triggering the flag and whether NVCA occurs

- Monetary cost of triggering the flag is zero
- But fiscal costs on jurisdiction and social and economic costs on individual and community
- Presumption of innocence, so limit pre-trial detention

How do we weigh the costs of flagging arrestees vs an NVCA?

Define utility based on triggering the flag and whether NVCA occurs

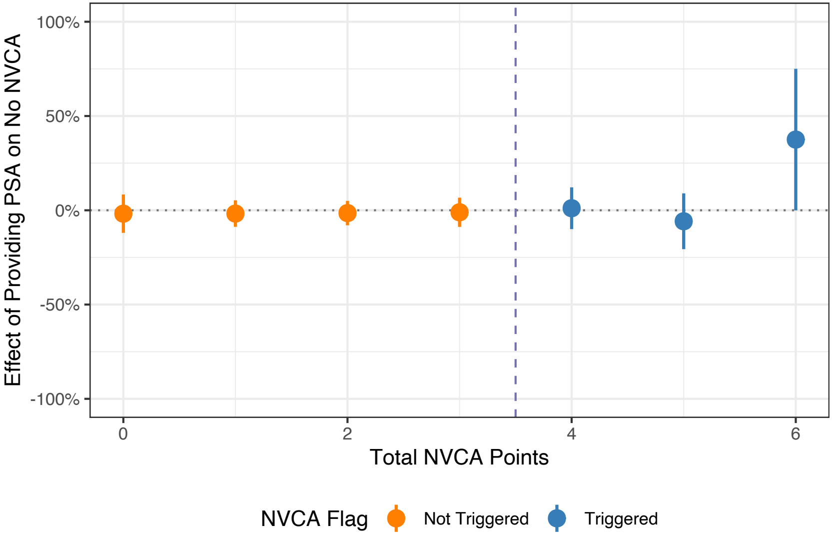
- Monetary cost of triggering the flag is zero
- But fiscal costs on jurisdiction and social and economic costs on individual and community
- Presumption of innocence, so limit pre-trial detention

Use a single parameterization:

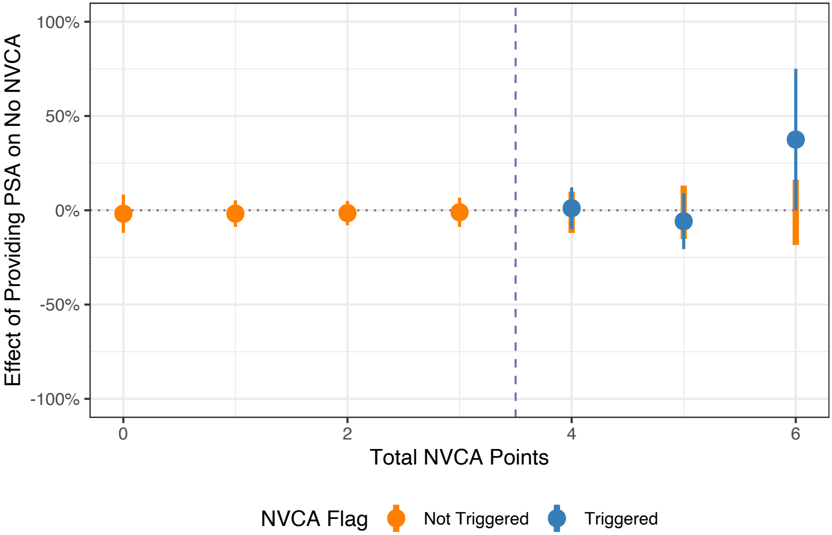
$$\text{NVCA Cost} \times Y(a) - \pi(a)$$

- Pin fiscal and societal costs to be 1
- Cost of an NVCA starts at 1 and grows

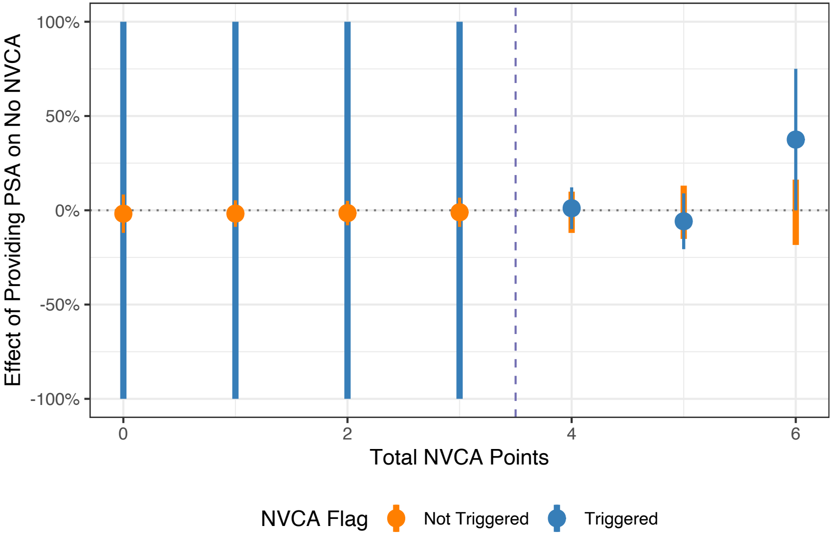
Robust approach raises the NVCA risk threshold



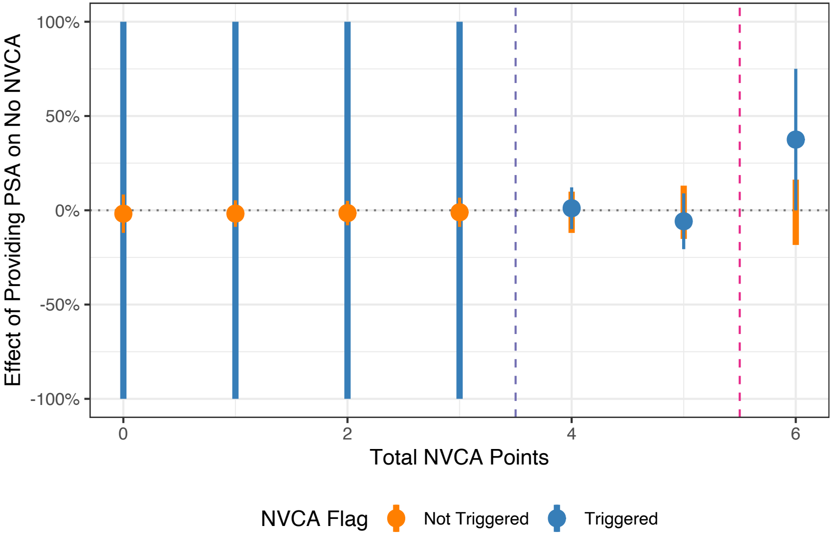
Robust approach raises the NVCA risk threshold



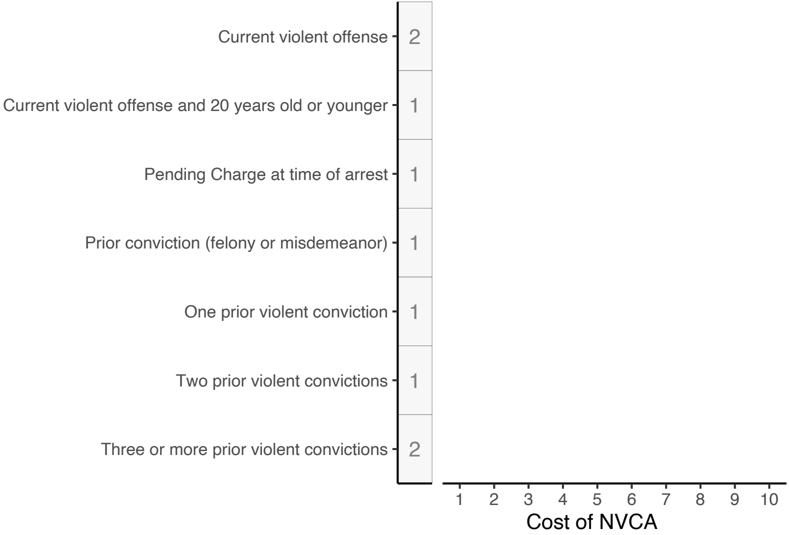
Robust approach raises the NVCA risk threshold



Robust approach raises the NVCA risk threshold

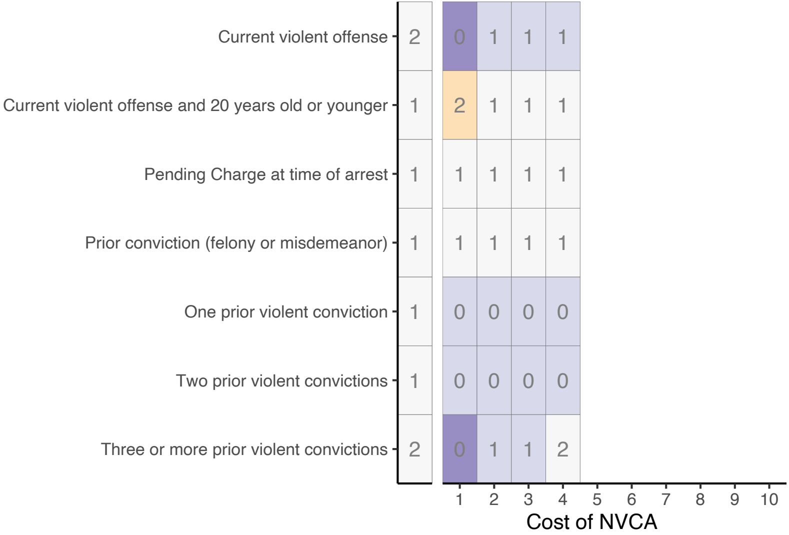


Robust approach places less weight on violent convictions and triggers flag less



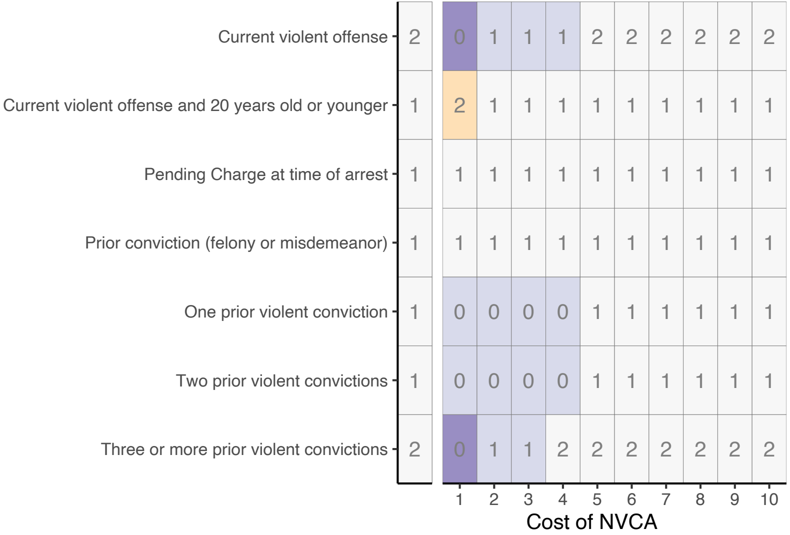
$$\text{Effect} = f_1(\text{current violent offense}) + f_2(\text{prior violent convictions}) + \ldots$$

Robust approach places less weight on violent convictions and triggers flag less



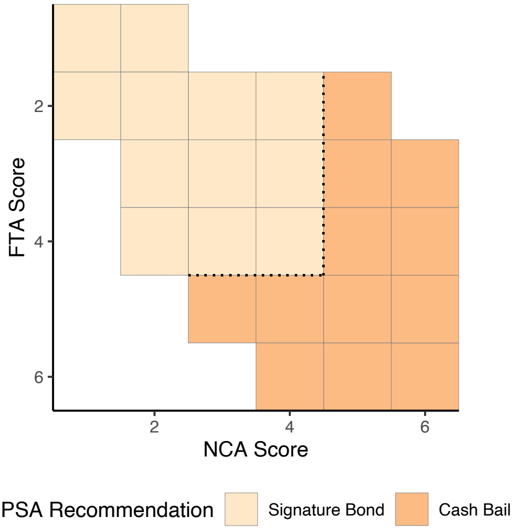
$$\text{Effect} = f_1(\text{current violent offense}) + f_2(\text{prior violent convictions}) + \ldots$$

Robust approach places less weight on violent convictions and triggers flag less



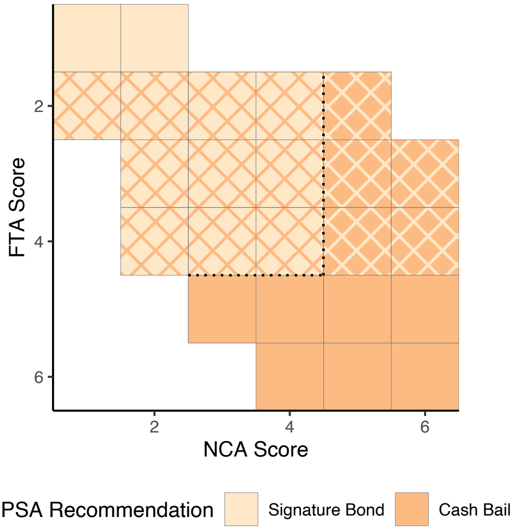
$$\text{Effect} = f_1(\text{current violent offense}) + f_2(\text{prior violent convictions}) + \ldots$$

With an additive model we can identify a slice of the DMF matrix...



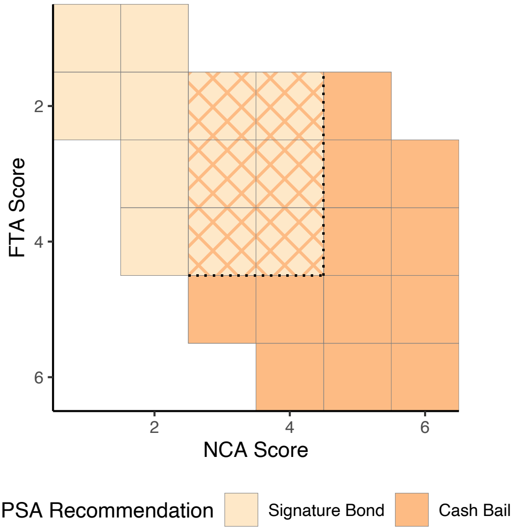
$$\text{Effect} = f_1\,(\text{FTA Score}) + f_2\,(\text{NCA Score})$$

With an additive model we can identify a slice of the DMF matrix...



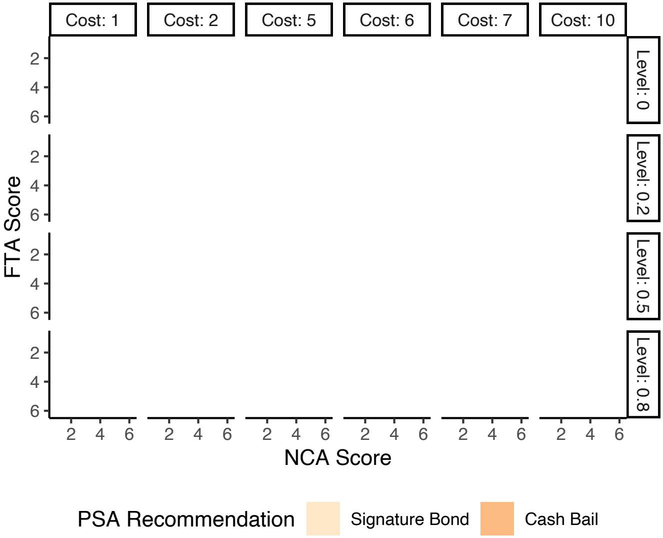
$$\text{Effect} = f_1\,(\text{FTA Score}) + f_2\,(\text{NCA Score})$$

With an additive model we can identify a slice of the DMF matrix...

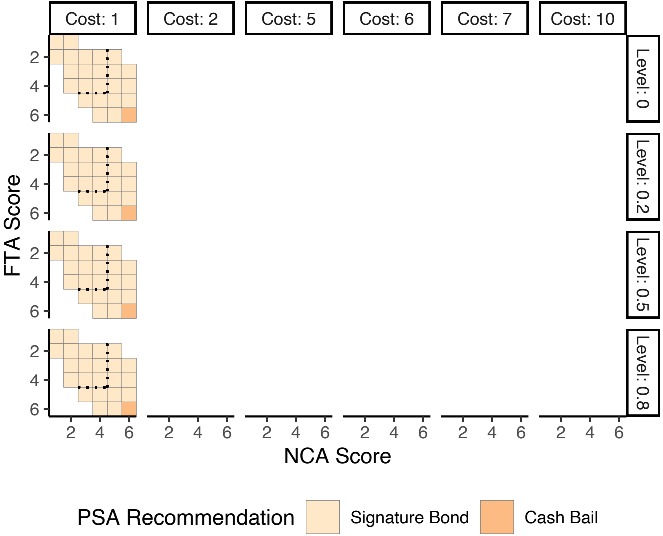


$$\text{Effect} = f_1\,(\text{FTA Score}) + f_2\,(\text{NCA Score})$$

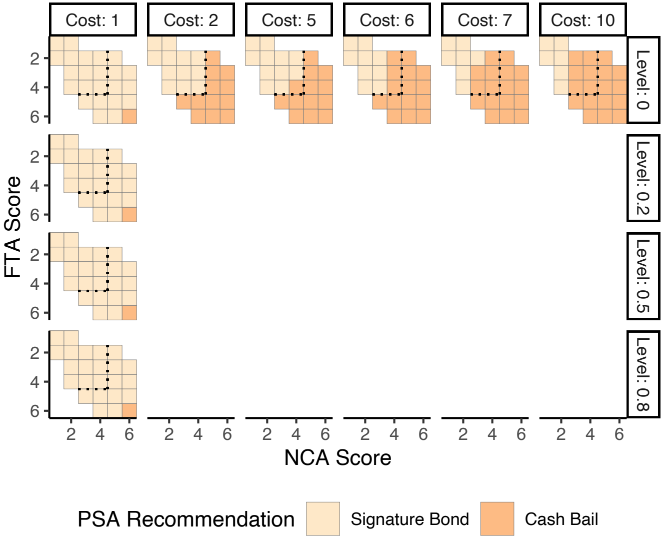
...but we can't reliably learn a new algorithm when accounting for statistical noise



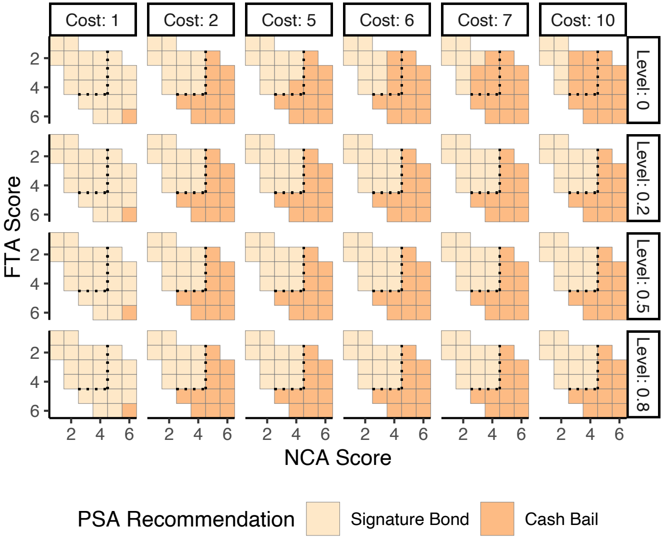
…but we can't reliably learn a new algorithm when accounting for statistical noise



...but we can't reliably learn a new algorithm when accounting for statistical noise



…but we can't reliably learn a new algorithm when accounting for statistical noise



Recap: Safe policy learning through extrapolation

Deterministic rule-based and algorithmic policies are everywhere

- Generate a lot of data! But deterministic nature means we have to extrapolate

Recap: Safe policy learning through extrapolation

Deterministic rule-based and algorithmic policies are everywhere

- Generate a lot of data! But deterministic nature means we have to extrapolate

This paper: Extrapolate in a safe way with robust optimization to learn a new algorithm

- Characterize all of the ways to extrapolate and find the **best policy in the worst case**
- Gives a statistical safety guarantee: at least as good as the status quo
- Some evidence we can improve the PSA, but noisy. Need more data!

Recap: Safe policy learning through extrapolation

Deterministic rule-based and algorithmic policies are everywhere

- Generate a lot of data! But deterministic nature means we have to extrapolate

This paper: Extrapolate in a safe way with robust optimization to learn a new algorithm

- Characterize all of the ways to extrapolate and find the **best policy in the worst case**
- Gives a statistical safety guarantee: at least as good as the status quo
- Some evidence we can improve the PSA, but noisy. Need more data!

Many more questions on designing algorithms to assist human decision makers

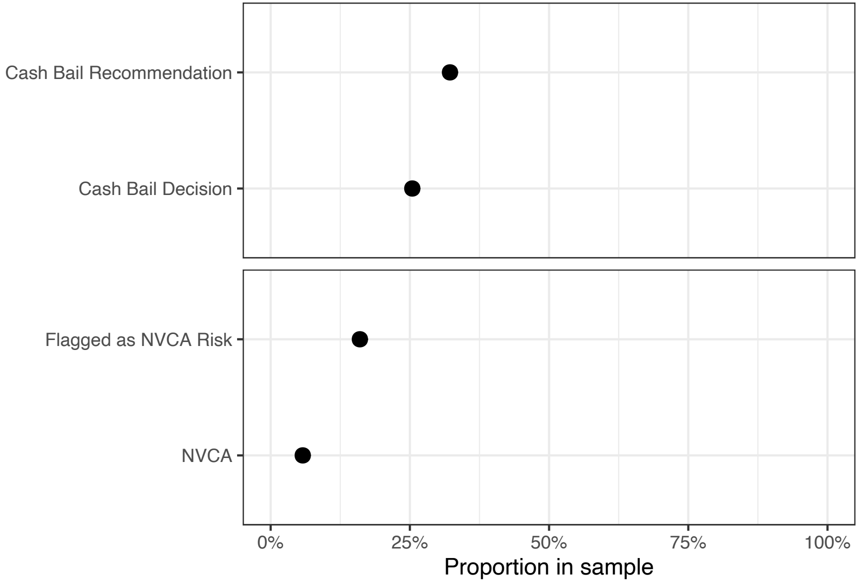
- Asymmetric utility functions lead to unidentifiable objectives
- Incorporating fairness and alternative notions of "safety"
- Optimizing for long term outcomes when we only can measure short term outcomes
- Learning policies when human decisions mediate future outcomes and decisions

Thank you!

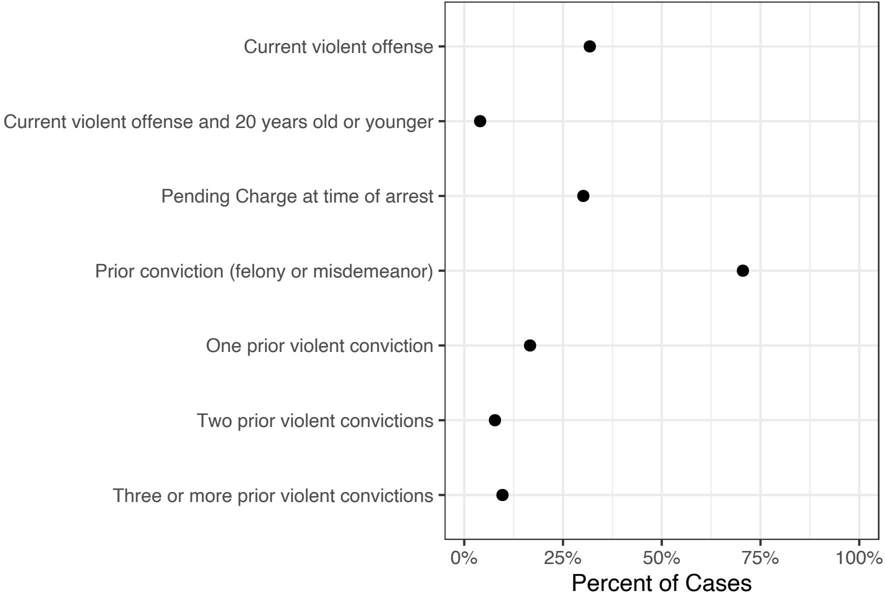
ebenmichael.github.io

Appendix

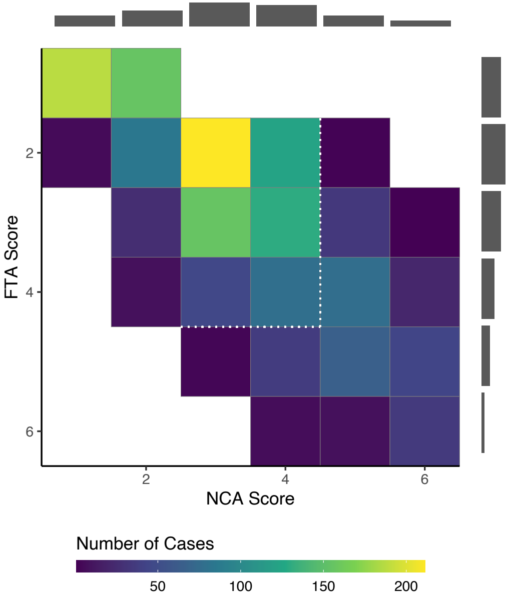
Cash Bail and NVCAs are less common



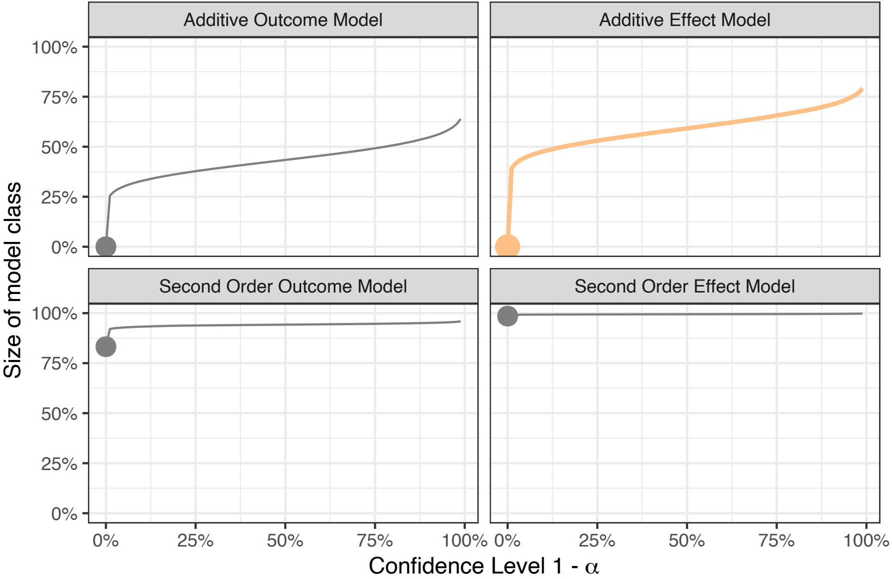
Frequency of attributes entering the NVCA Flag



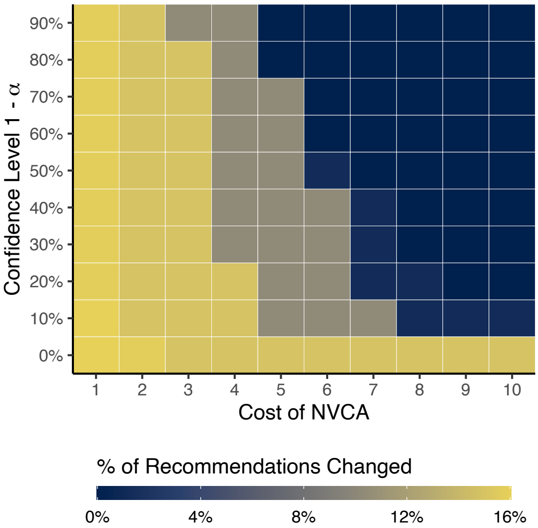
FTA and NCA scores move in unison



Additive treatment effects are at the sweetspot of robustness and optimality



Robust approach changes scores to trigger NVCA flag less often



$$\text{Effect} = f_1(\text{current violent offense}) + f_2(\text{prior violent convictions}) + \ldots$$

Incorporating experiments and human decisions

Incorporating experiments evaluating a deterministic policy

- In our study, judges randomly receive the "null policy" $\emptyset$, no access to PSA

Allows us to work with **treatment effects** instead of outcomes

$$\tau(a, x) = \mathbb{E}[Y(a) - Y(\emptyset) \mid X = x]$$

- Treatment effects are often considered to be simpler than baseline outcomes [Künzel et al., 2019; Hahn et al., 2020; Nie and Wager, 2021]

Incorporating judge decisions from algorithmic recommendations

- Define utility based on potential decision $D(a)$ and potential outcome $Y(D(a))$

$$\text{benefit} \times Y(D(a)) - \text{cost} \times D(a)$$

Value includes two unidentified components, outcomes **and decisions**

- Need to find the worst case potential decision and outcome for cost and benefit

Statistical properties

Value is probably, approximately at least as high as baseline

$$V(\tilde{\pi}) - V(\hat{\pi}) \lesssim \text{Complexity}(\Pi) \qquad \text{with probability at least } \gtrsim 1 - \alpha$$

- Conservative approach gives a statistical safety guarantee with level α
- If policy class Π is complex, need more samples to avoid overfitting

Empirical optimality gap controlled by size of $\widehat{\mathcal{M}}_n(\alpha)$ and complexity of Π

$$V(\pi^*) - V(\hat{\pi}) \lesssim \frac{u}{n} \sum_{i=1}^{n} \max_{a \in \mathcal{A}} \widehat{B}_{\alpha u}(a, X_i) - \widehat{B}_{\alpha \ell}(a, X_i) + \text{Complexity}(\Pi)$$

with probability at least $\gtrsim 1 - \alpha$

- Tradeoff between safety and optimality

References I

Angrist, J. D. and Rokkanen, M. (2015). Wanna Get Away? Regression Discontinuity Estimation of Exam School Effects Away From the Cutoff. *Journal of the American Statistical Association*, 110(512):1331–1344.

Athey, S. and Wager, S. (2021). Policy Learning With Observational Data. *Econometrica*, 89(1):133–161.

Bertsimas, D., Brown, D. B., and Caramanis, C. (2011). Theory and applications of robust optimization. *SIAM Review*, 53(3):464–501.

Cattaneo, M. D., Keele, L., Titiunik, R., and Vazquez-Bare, G. (2020). Extrapolating Treatment Effects in Multi-Cutoff Regression Discontinuity Designs. *Journal of the American Statistical Association*, 0(0):1–48.

Cui, Y. (2021). Individualized decision making under partial identification: three perspectives, two optimality results, and one paradox. *Harvard Data Science Review*. Just accepted.

Dell, M. and Querubin, P. (2018). Nation building through foreign intervention: evidence from discontinuities in military strategies. *Quarterly Journal of Economics*, 133(2):701–764.

Greiner, D. J., Halen, R., Stubenberg, M., and Chistopher L. Griffen, J. (2020). Randomized control trial evaluation of the implementation of the psa-dmf system in dane county. Technical report, Access to Justice Lab, Harvard Law School.

References II

Hahn, P. R., Murray, J. S., and Carvalho, C. M. (2020). Bayesian Regression Tree Models for Causal Inference: Regularization, Confounding, and Heterogeneous Effects. *Bayesian Analysis*, pages 1–33.

Imai, K., Jiang, Z., Greiner, D. J., Halen, R., and Shin, S. (2020). Experimental Evaluation of Computer-Assisted Human Decision-Making: Application to Pretrial Risk Assessment Instrument (with discussion). *Journal of the Royal Statistical Society, Series A (Statistics in Society)*, page Forthcoming. arxiv preprint `https://arxiv.org/pdf/2012.02845.pdf`.

Kallus, N. and Zhou, A. (2021). Minimax-optimal policy learning under unobserved confounding. *Management Science*, 67(5):2870–2890.

King, G. and Zeng, L. (2006). The dangers of extreme counterfactuals. *Political Analysis*, 14(2):131–159.

Kitagawa, T. and Tetenov, A. (2018). Who Should Be Treated? Empirical Welfare Maximization Methods for Treatment Choice. *Econometrica*, 86(2):591–616.

Künzel, S. R., Sekhon, J. S., Bickel, P. J., and Yu, B. (2019). Metalearners for estimating heterogeneous treatment effects using machine learning. *Proceedings of the National Academy of Sciences of the United States of America*, 116(10):4156–4165.

Ludwig, J. and Miller, D. L. (2007). Does head start improve children's life chances? Evidence from a regression discontinuity design. *Quarterly Journal of Economics*, 122(1):159–208.

References III

Manski, C. F. (2005). *Social Choice with Partial Knowledge of Treatment Response*. Princeton University Press.

Morse, J. C. (2019). *Blacklists, Market Enforcement, and the Global Regime to Combat Terrorist Financing*, volume 73.

Nie, X. and Wager, S. (2021). Quasi-oracle estimation of heterogeneous treatment effects. *Biometrika*, 108(2):299–319.

Pu, H. and Zhang, B. (2021). Estimating optimal treatment rules with an instrumental variable: A partial identification learning approach. *Journal of the Royal Statistical Society Series B*, pages 1–28.

Qian, M. and Murphy, S. A. (2011). Performance guarantees for individualized treatment rules. *The Annals of Statistics*, 39(2):1180–1210.

Zhao, Y., Zeng, D., Rush, A. J., and Kosorok, M. R. (2012). Estimating individualized treatment rules using outcome weighted learning. *Journal of the American Statistical Association*, 107(499):1106–1118.