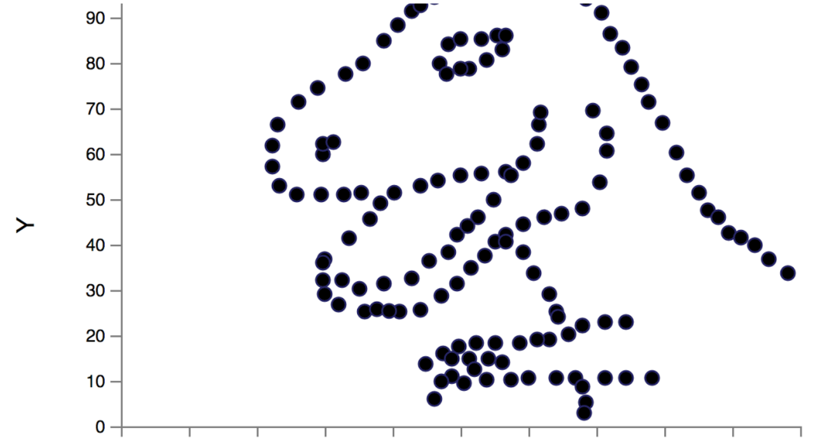
School of Criminology and Criminal Justice
Arizona State University

The Limit of Administrative Data?



Shi Yan

Symposium at the School of Criminology and Criminal Justice

March 29th, 2019

The Statement
(Baumer, 2013)

While all research areas much concede some ground to scientific rigor in light of data constraints,
it seems clear that the modal approach…
yields uncertain evidence about the presence (or absence) of racial disparities
because the data available for the task is likely to omit key attributes that vary in prevalence across racial groups and which are related to sentencing decisions.

Motivation: How far can we go?

- Problems: Missing information + linear assumptions
- How much can less restrictive parametric assumptions compensate the missing info?

Data: Computerized Criminal History (CCH) of New York State (N > 13 mil)

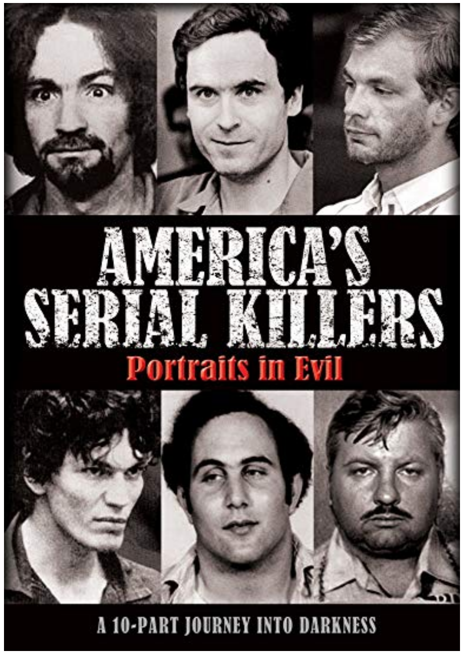
CCH	Conventional datasets
• Arrest sample	• Conviction sample
• Arrest events nested under person IDs	• No person IDs, only arrest IDs
• No prosecutor/judge IDs	• May or may not have prosecutor and judge info
• No sentencing guidelines and broad discretion	• Sentencing guidelines and discretion vary by state

Overview: Tasks

- Investigate criminal specialization
- Examine criminal escalation and the sentence

Study 1
How should we understand specialization?

The Assumption of Specialization



Theory & Policy: Same or Different Distributions?

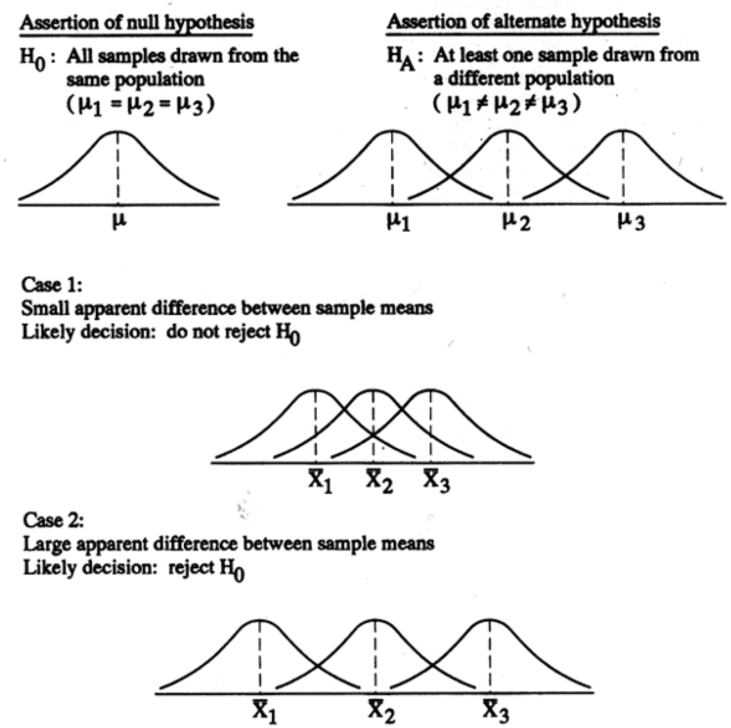


FIGURE 10.1
Null and Alternate Hypotheses in Analysis of Variance (ANOVA)

source: McGrew and Monroe (2000)

Center for Problem-Oriented Policing

Specialization Research as Part of the Criminal Careers Paradigm
(Piquero et al., 2003)

- Variation or stability
- Transition among offense types

The Measures
(Sullivan et al., 2009)

- Forward Specialization Coefficient
  - "Among those arrested for robbery, 12% also had robbery as their immediately subsequent arrest"
- Diversity Index
  - 0 to 1, where 0 means full specialization
- Latent class analysis
  - Groups: "Driving Specialists," "Drug Specialists" etc.

- Plus…
  - Repeated most recent crime type?
  - Total priors for the current crime type?

However…

(Yan, 2016)

	Diversity Index	Repeated most recent	Priors of current crime
Repeated most recent	0.155		
Priors of current crime	0.023	0.185	
Drug Gen.	-0.257	-0.023	0.259
High Invol. Gen.	-0.211	-0.031	0.175
Driving Spec.	0.228	0.098	-0.071
Property Spec.	0.114	-0.001	-0.038
Drug Spec.	0.094	0.054	-0.007
Violent Spec.	-0.045	-0.061	-0.119

To take a closer look…

Crime	Type
1	Drug possession
2	Drug sale
3	Drug sale
4	Drug possession

- DI = 0.5
- Class = Drug Spec.
- Recent crime? Depends

What Does Specialization Suggest?

- Similar nature
- Common cause
- Necessity for special prevention/intervention strategy?

- Classification-prediction?

Also Similar: A Sparse Matrix Problem

$$\begin{pmatrix} 1.0 & 0 & 5.0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3.0 & 0 & 0 & 0 & 0 & 11.0 & 0 \\ 0 & 0 & 0 & 0 & 9.0 & 0 & 0 & 0 \\ 0 & 0 & 6.0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 7.0 & 0 & 0 & 0 & 0 \\ 2.0 & 0 & 0 & 0 & 0 & 10.0 & 0 & 0 \\ 0 & 0 & 0 & 8.0 & 0 & 0 & 0 & 0 \\ 0 & 4.0 & 0 & 0 & 0 & 0 & 0 & 12.0 \end{pmatrix}$$

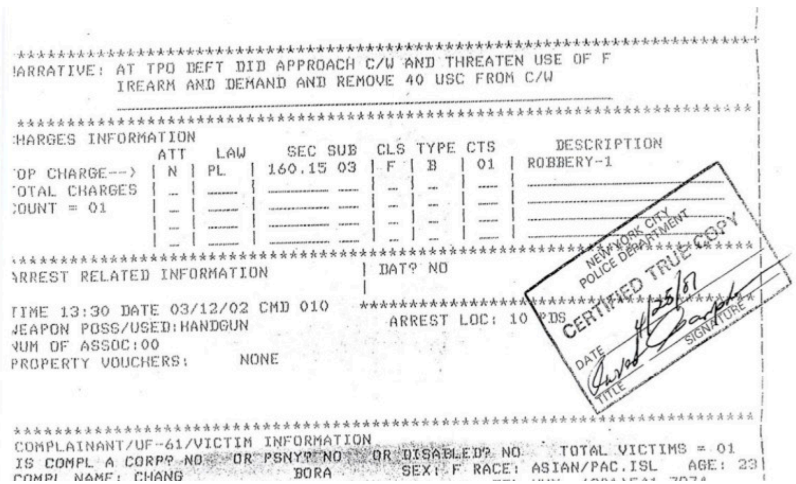
Results?

- On the one hand, algorithm targets sparse matrices
- On the other hand… too sparse?

Study 2
How close can criminal records predict the sentence?

(Spoiler: Not very close)

Overview



NARRATIVE: AT TPO DEFT DID APPROACH C/W AND THREATEN USE OF F
IREARM AND DEMAND AND REMOVE 40 USC FROM C/W

CHARGES INFORMATION

	ATT	LAW	SEC SUB	CLS	TYPE	CTS	DESCRIPTION
TOP CHARGE-->	N	PL	160.15 03	F	B	01	ROBBERY-1
TOTAL CHARGES							
COUNT = 01							

ARREST RELATED INFORMATION | DAT? NO

TIME 13:30 DATE 03/12/02 CMD 010 ARREST LOC: 10 PDS
WEAPON POSS/USED: HANDGUN
NUM OF ASSOC: 00
PROPERTY VOUCHERS: NONE

NEW YORK CITY POLICE DEPARTMENT CERTIFIED TRUE COPY
DATE 4/25/07 TITLE SIGNATURE

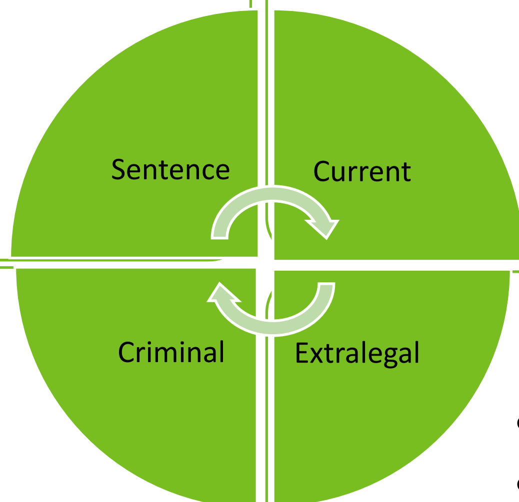
COMPLAINANT/UF-61/VICTIM INFORMATION TOTAL VICTIMS = 01
IS COMPL A CORP? NO OR PSNY? NO OR DISABLED? NO SEX: F RACE: ASIAN/PAC.ISL AGE: 23
COMPL NAME: CHANG BORA

- Criminal records are correlated with the sentence
  - But most existing research only control for number of priors
- What happens if I put the entire rap sheet in?

We’ve seen progress in all areas

- Alternative sanctions
- Pre-conviction outcomes
- Cumulative disadvantages

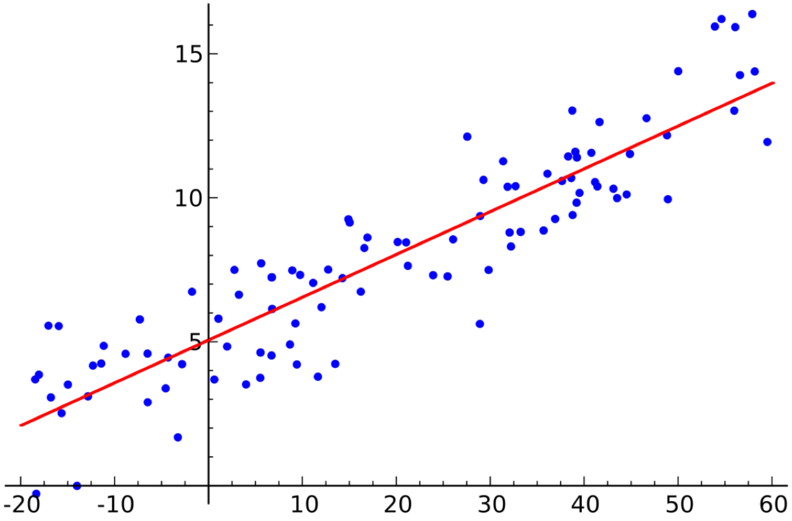
- Strength of evidence
- Aggravating and mitigating factors



- Criminal career properties
- Prior delinquency and school disciplines

- Physical appearance
- Immigration status
- Less-studied groups (Asians, Native Americans, etc.).

However, there is one thing remaining…



- Regression is useful
  - Examines relationship between variables and the sentence
  - Predicts the sentence given observed characteristics

Can we get better predictions?

(Abrams, 2016; Piehl & Bushway, 2007)

- Regression models assume underlying functional forms
  - Human decisions can be highly non-linear
- Prediction of the sentence can be useful when…
  - Seeking to reduce extralegal disparities
  - Estimating counterfactuals
  - Just trying to decipher the sentencing process

The present study compares three modeling approaches

- Among felony defendants, who get incarcerated?
  - In the full sample, 42.8% get incarceration
- Key outcome: Prediction accuracy
  - Naïve guess of no incarceration for all leads to 57.2% accuracy
  - Simple logistic regression
  - Classification tree
  - Random forest

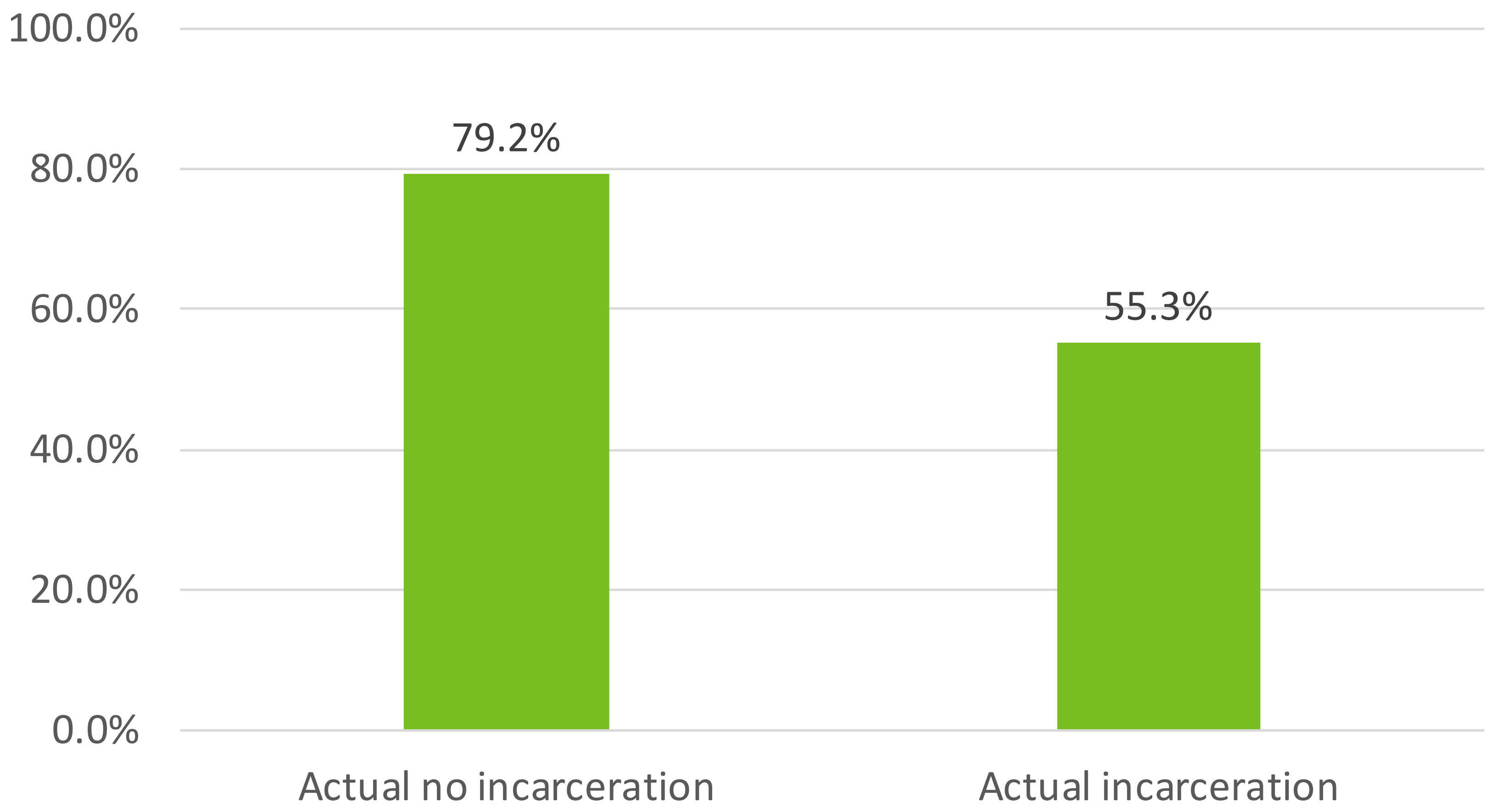
Data: New York State Computerized Criminal History

- All felony defendants between 2008 and 2012, who already had one or more prior convictions (n = 168,811)
  - To make sure we can connect them to their priors
- 70% random cases as training sample, 30% as testing sample
  - To prevent the model from overfitting the training data
  - “Hide” the testing sample first and train all models on training sample
  - Then examine performance on testing sample

Simple logistic model

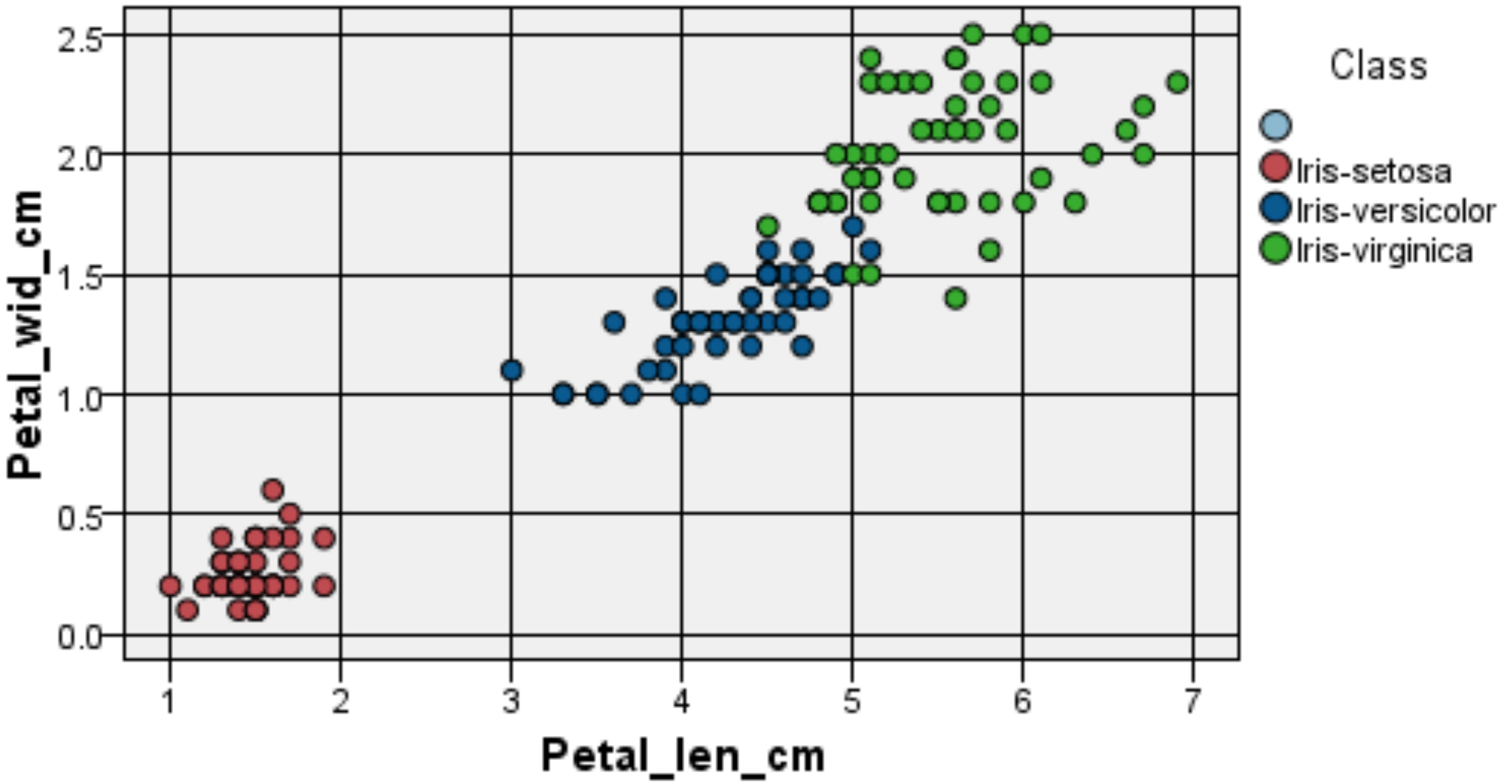
- DV: Incarceration (prison or jail)
  - Severity + type of current crime
  - Number of prior felony and misdemeanor convictions
  - Race, sex, age, county & year fixed effects
- Model findings consistent with literature
  - Strong predictors: Crime severity + number of priors
  - Small but significant racial, gender, ethnic disparities
- Predicts p(inc) for testing sample
  - prediction = 1 if p(inc) >= 0.5

Accuracy = 69.01%

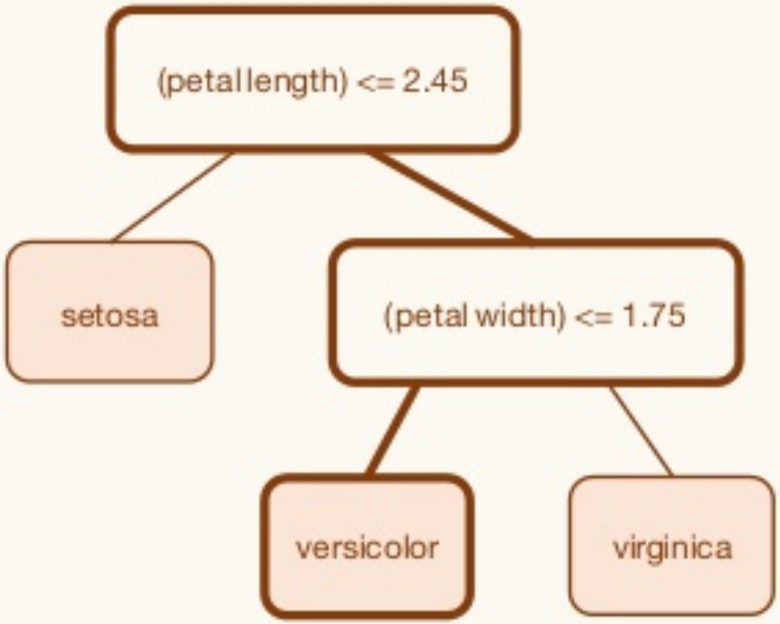


Decision tree, CART algorithm

(Breiman et al., 1984)



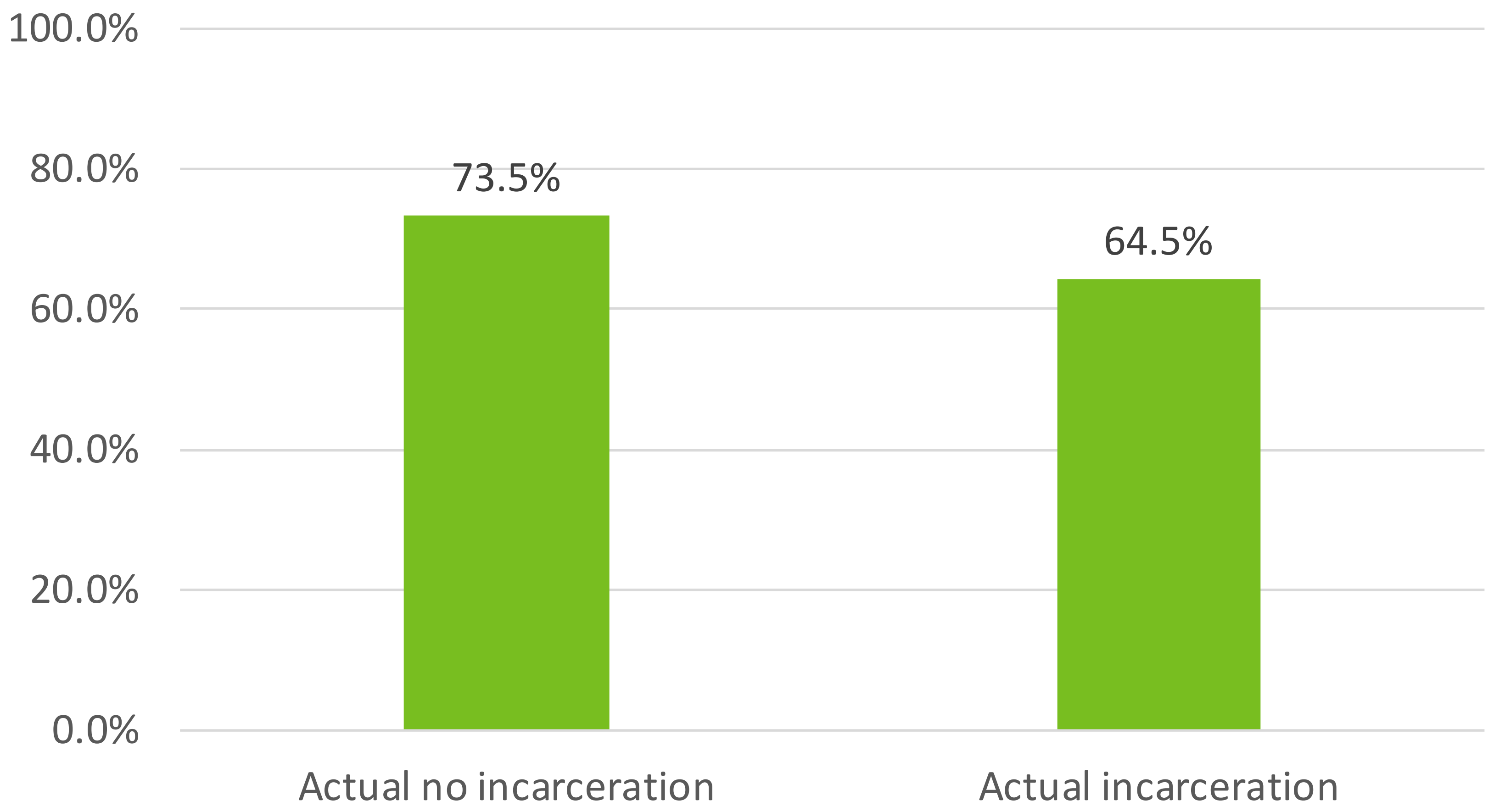
What is Decision Tree?



an iris whose
- petal length is 2.8
- petal width is 1.6

is predicted as versicolor.

Accuracy = 69.70%



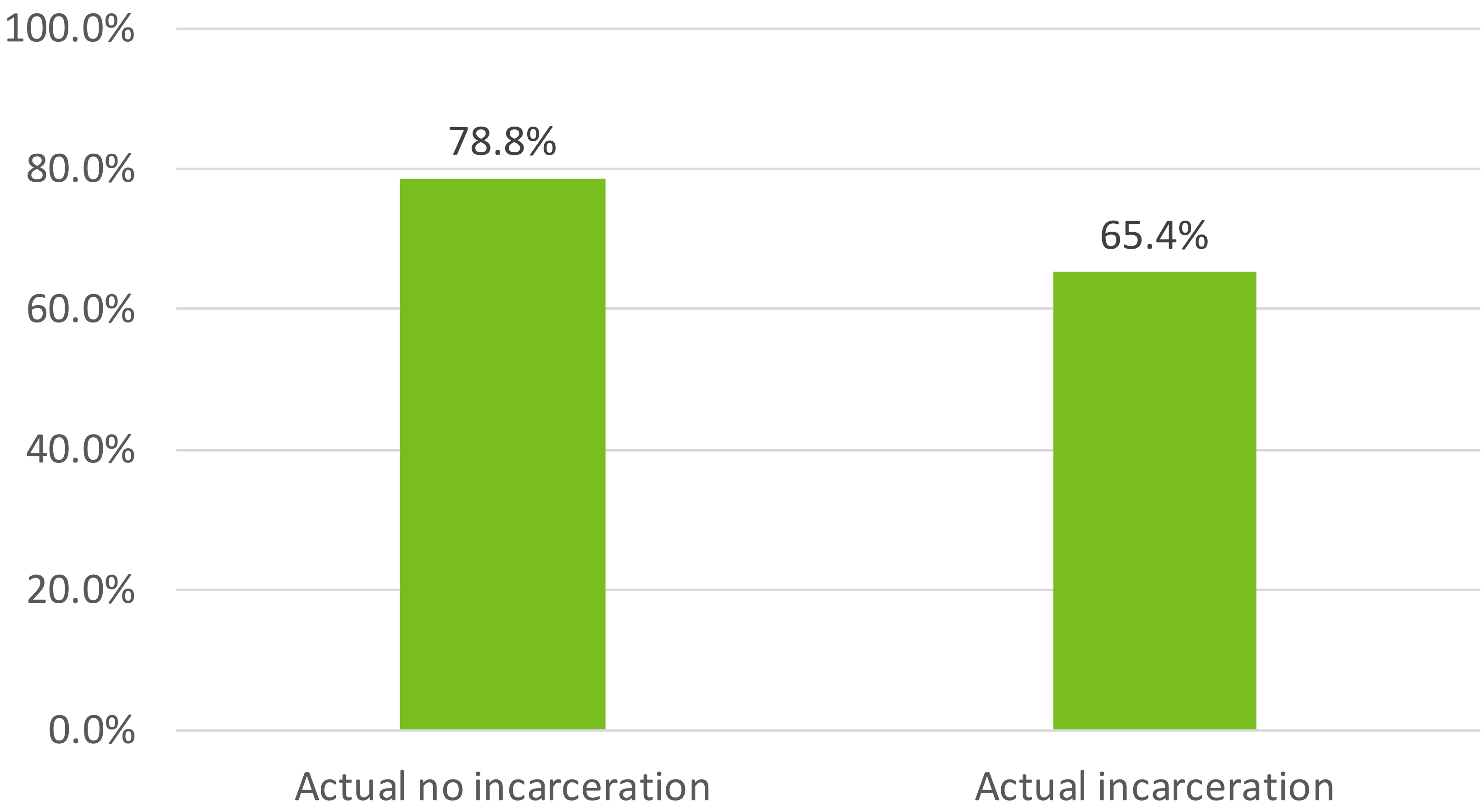
Random forest: Combination of trees

Random forest algorithm
(Breiman, 2001)

- Bootstrapped samples
- Random subset of predictors

- Takes minutes to resolve, not hours

Accuracy = 73.07%



Summary of findings

- Modeling non-linearity of criminal records helps somewhat…but only to a certain extent

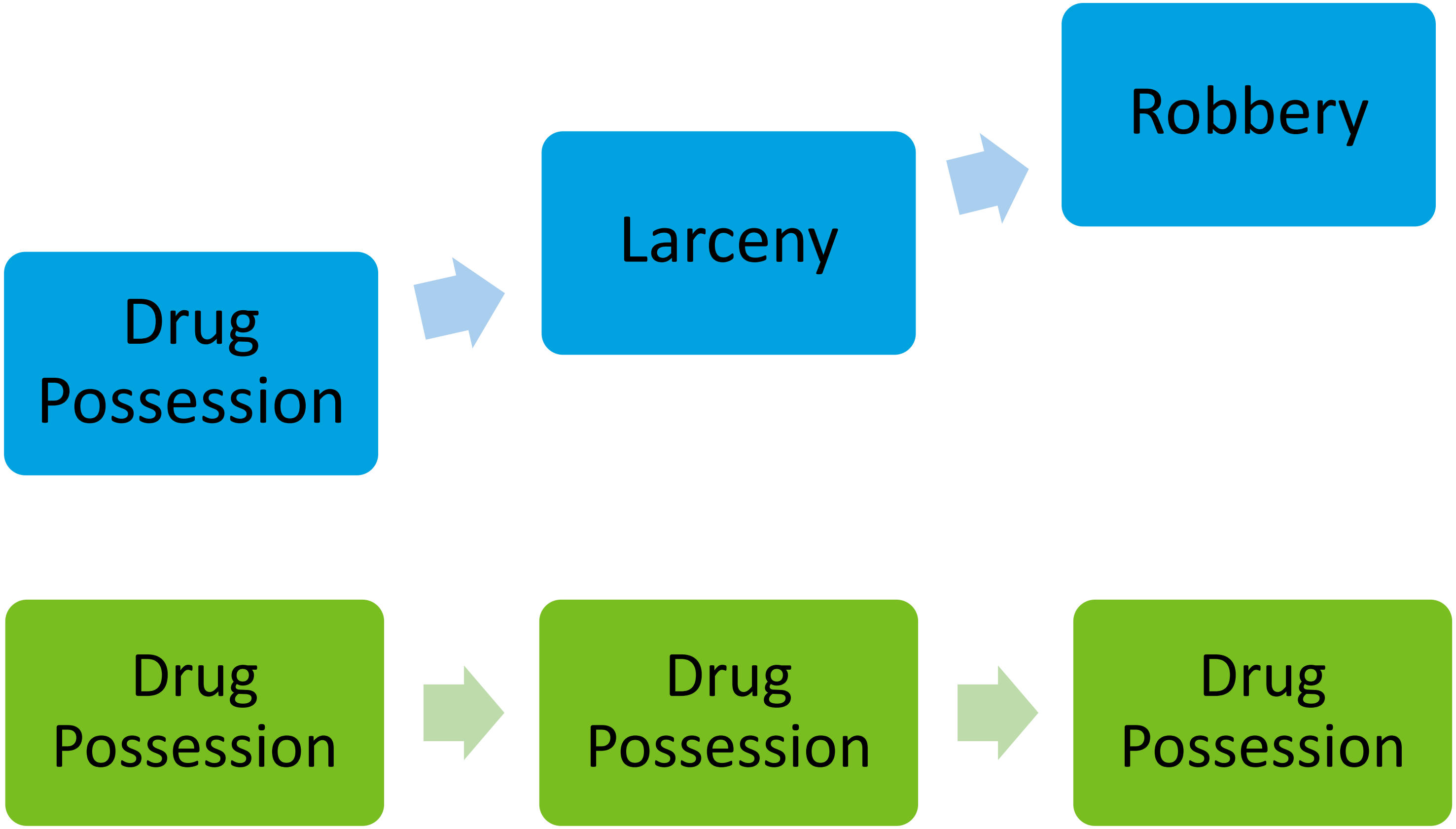
Study 3
Is Criminal Escalation Related to the Sentence?

(Spoiler: Yes)

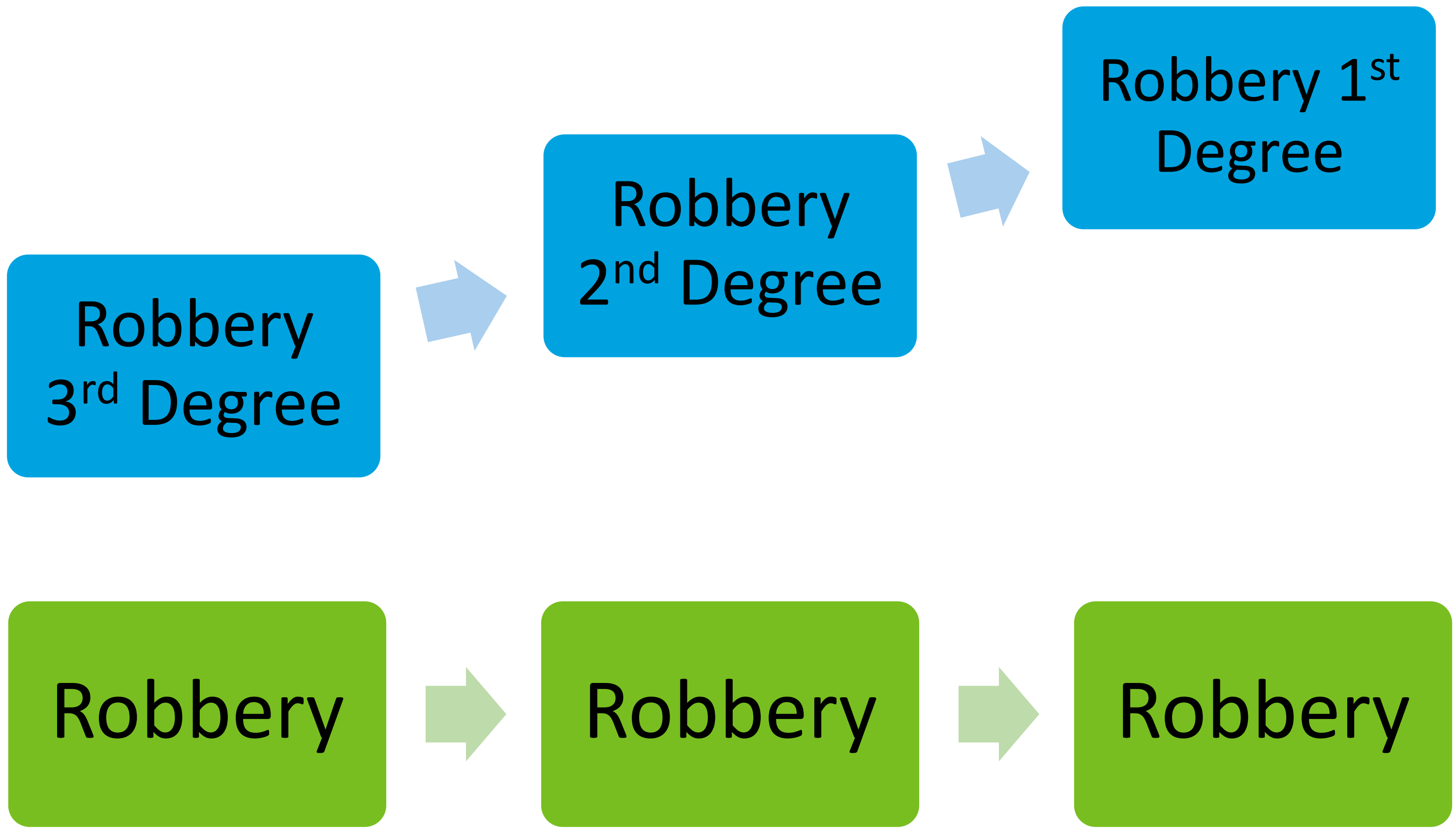
Research vs. Practice

ID	Prior Felonies	Prior Misdemeanors
87465	2	4
98475	4	6
11254	1	0
47586	1	5

Escalation Intertwined with Specialization
(Le Blanc, 2002)

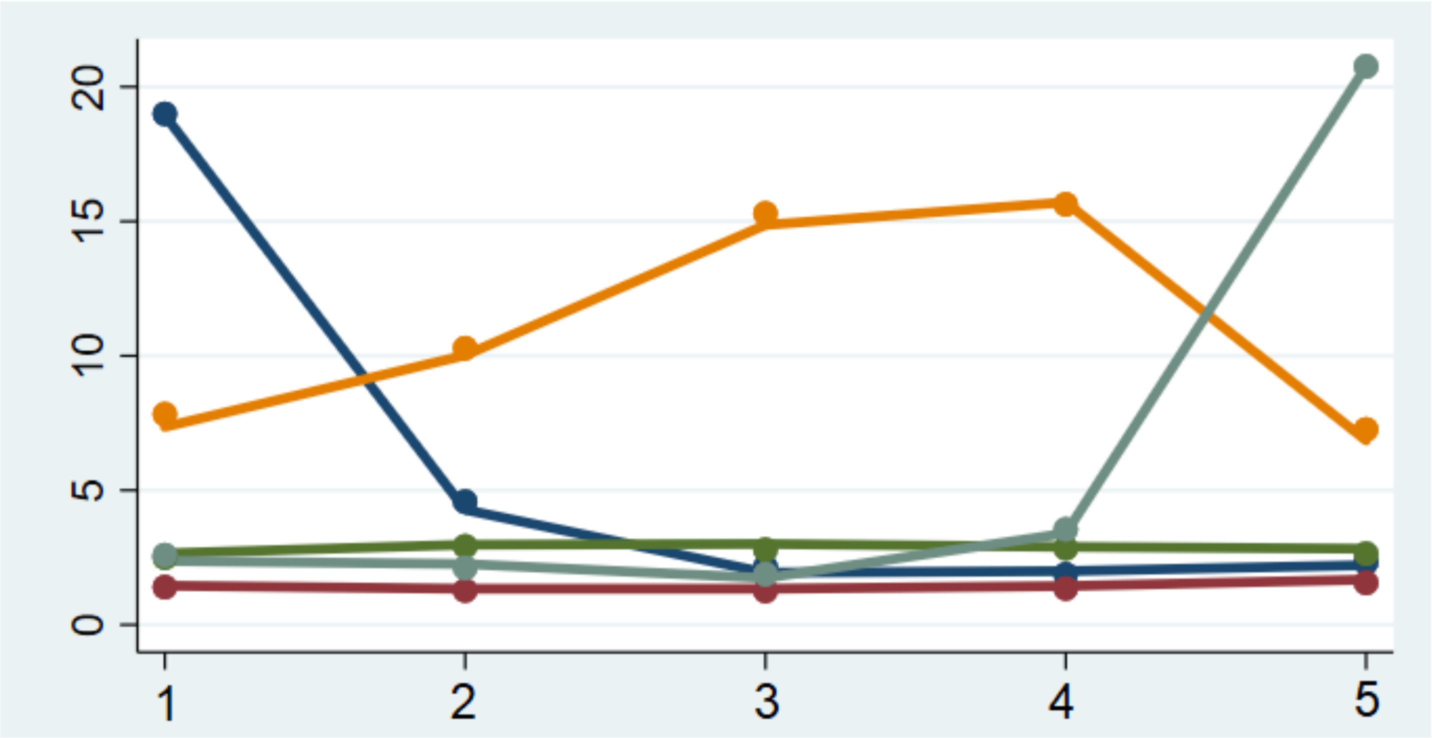


Escalation Independent from Specialization



Escalation: Group-based Trajectory Models (GBTM, Nagin, 2005)

- Identifies different longitudinal patterns within sample.
- Censored normal dependent variable up to cubic term.

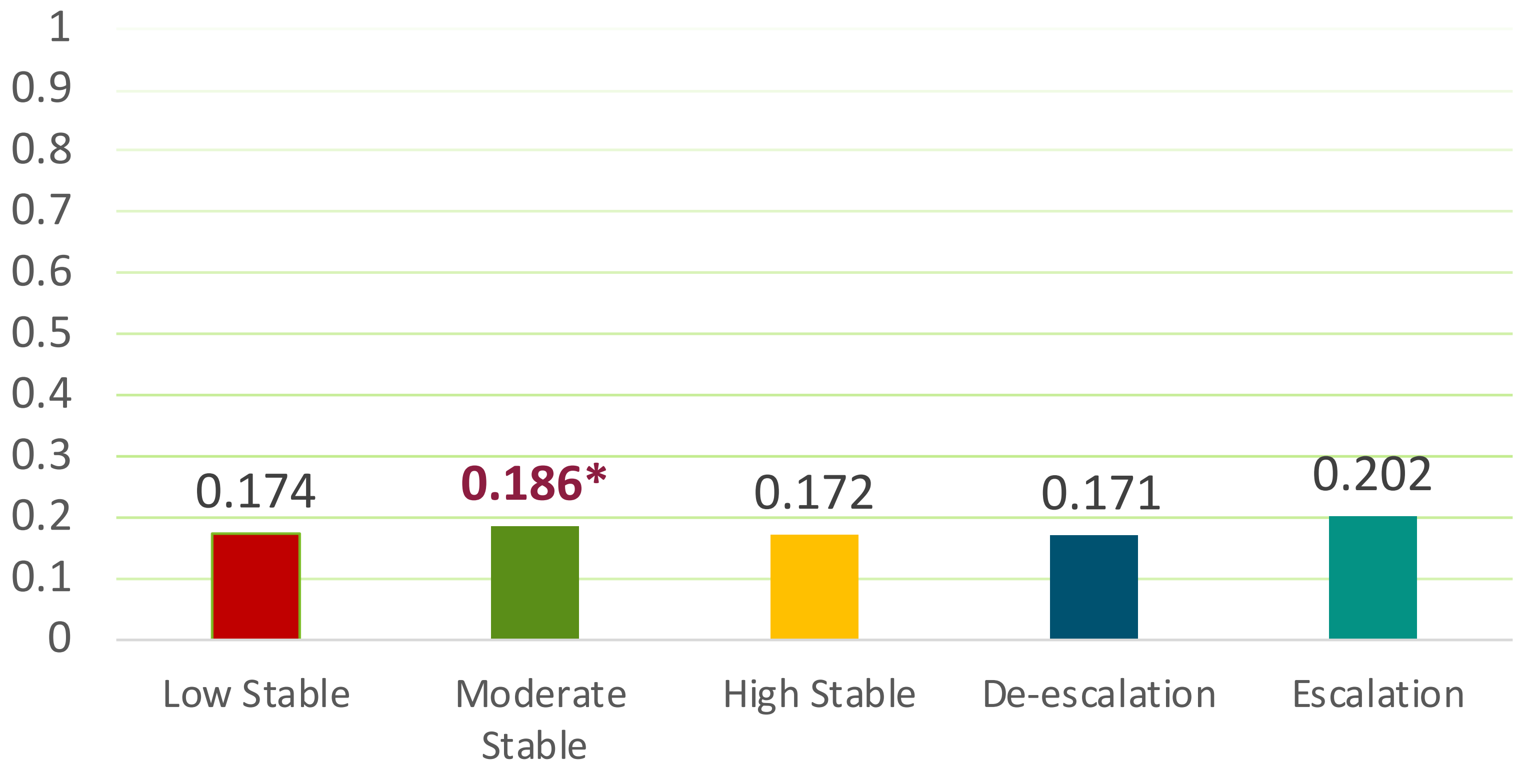


- Low Stable (68.0%)
- Moderate Stable (27.8%)
- High Stable (.6%)
- De-escalating (1.6%)
- Escalating (1.9%)

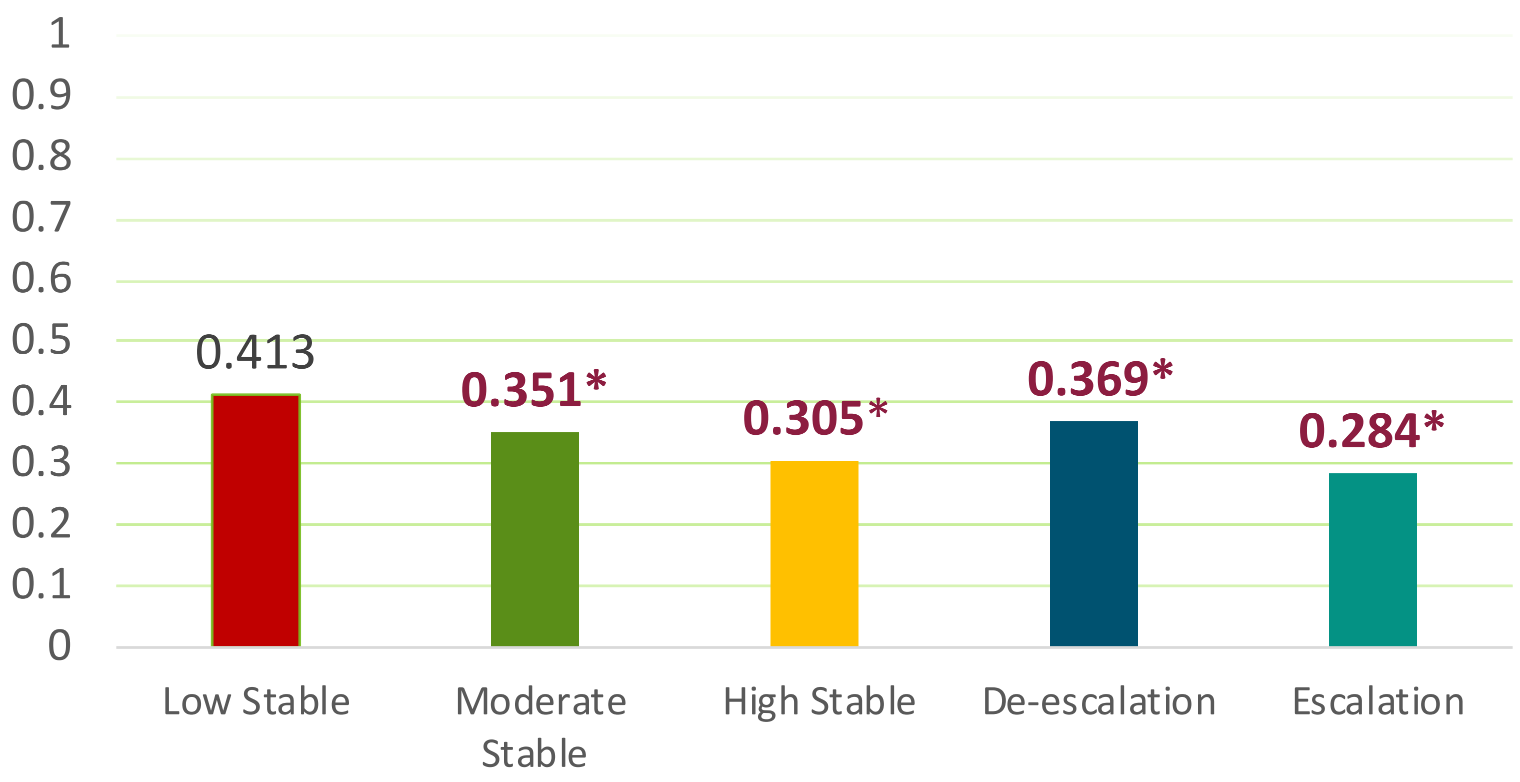
DV: Different Stages in Sentencing
(Cumulative disadvantage, Kutateladze et al., 2014)

- Dismissal
- Reduction to misdemeanor
- In/out decision to incarcerate

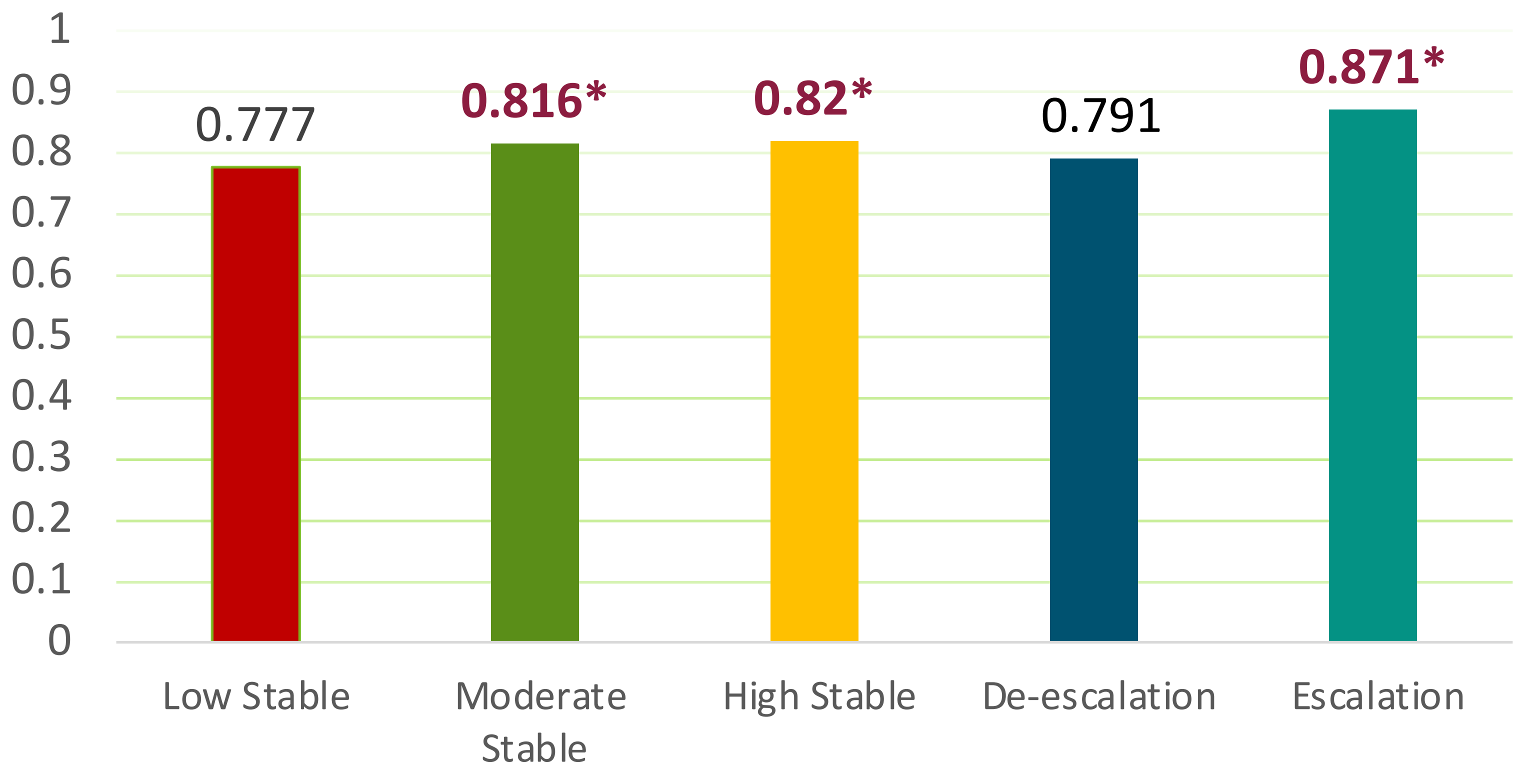
Dismissal
(1=Dismissed)



Reduction to Misdemeanor (1=Reduced)



In/Out Decision
(1=Incarcerated)



Summary of Findings

- Own and unique explanatory power on and above number of priors
- Not necessarily all about escalation—less favorable outcomes as long as defendants stand somewhere high

General Discussion

Why Should Criminal Records Matter?
(Hester et al., 2018; Roberts, 1997)

- Not self-evident: Levels of personal risk and diminished chances
- Considered holistically, especially when without guidelines—partially solved
- Cumulative process at multiple stages—not solved
  - Overall lack of exclusion restrictions

Additional Variables Are Still Necessary

- Clearly, flexibility is not everything (Study 2)
- Experimental designs and psychometric insights
- RCTs?

Thank you!

Shi Yan

shiyan@asu.edu