Incentive Effects of Recall Elections:
Evidence from Criminal Sentencing in California Courts

Sanford C. Gordon and Sidak Yntiso

Wilf Family Department of Politics, New York University

April 8, 2022

*We promise as judges to rule on the facts and on the law, not on public opinion... When public opinion affects a juror's decision or a judge's decision, it corrupts the rule of law. This recall, if successful, will make it harder for judges to keep that promise ...**The judicial recall, if successful, will be a silent force, a silent corrupting force. A force that will enter the minds of judges as they contemplate difficult decisions.***

- Judge Aaron Perksy, May 18 2018

Introduction

Does threat of recall affect sentencing behavior of trial judges?

Introduction

Does threat of recall affect sentencing behavior of trial judges?

Immediate significance: recall of CA Judge Aaron Persky.

Introduction

Does threat of recall affect sentencing behavior of trial judges?

Immediate significance: recall of CA Judge Aaron Persky.

Broader significance: electoral incentives of public officials

Introduction

Does threat of recall affect sentencing behavior of trial judges?

Immediate significance: recall of CA Judge Aaron Persky.

Broader significance: electoral incentives of public officials

Challenges

- recall campaigns are not randomly assigned;
- recall threat priced in.

What we're up to

Examine the effect of two potential shocks to salience of recall threat brought about by Persky recall initiative.

What we're up to

Examine the effect of two potential shocks to salience of recall threat brought about by Persky recall initiative.

Evaluate claims that the recall effort

- Increased judicial punitiveness
- Disproportionately burdened minority defendants

What we're up to

Examine the effect of two potential shocks to salience of recall threat brought about by Persky recall initiative.

Evaluate claims that the recall effort

- Increased judicial punitiveness
- Disproportionately burdened minority defendants

Findings

- Recall *announcement* led to an instantaneous 33% increase in average sentence length
- No evidence that effect mitigated racial disparities.
- Estimates of cumulative effects

Background: related research

Empirical research on electoral incentives generally (e.g., Besley and Case 1995; Ferraz and Finan 2011; Alt, Bueno de Mesquita, and Rose 2011)

Electoral incentives and sentencing behavior (e.g., Huber and Gordon 2004; Gordon and Huber 2007; Lim 2011; Park 2017)

Racial bias in criminal trials (e.g., Anwar, Bayer, Hjalmarsson 2012; Alesina and La Ferrara 2014, Abrams, Bertrand, Mullainathan 2012; Park 2017)

Background: recall elections

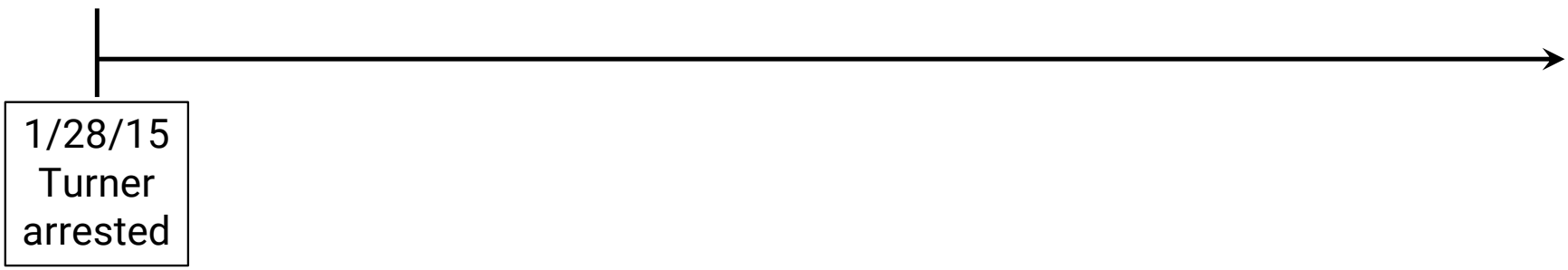
39 US states have some form of recall election; eight authorize recall of judges

Extant empirical research focuses on voter behavior in recall elections (e.g., Ho and Imai 2012; Segura and Fraga; Masket 2011; Shaw and McKenzie 2005)

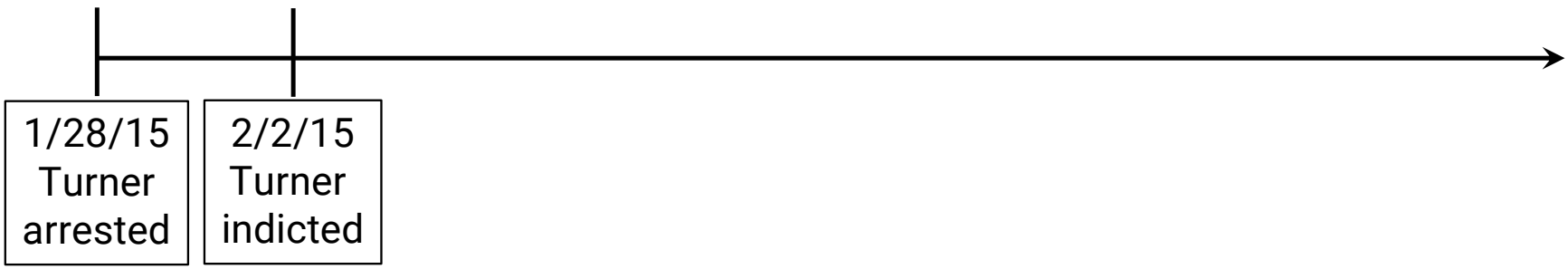
Recall elections in California

- Adopted in 1911 (Article 2, §§13-19 of CA Constitution)
- State: 165 attempts since 1913; 9 qualified; 5 successful
- Local: 333 qualified since 1995; 244 successful
  - One Superior Court judge: Aaron Persky (2018)

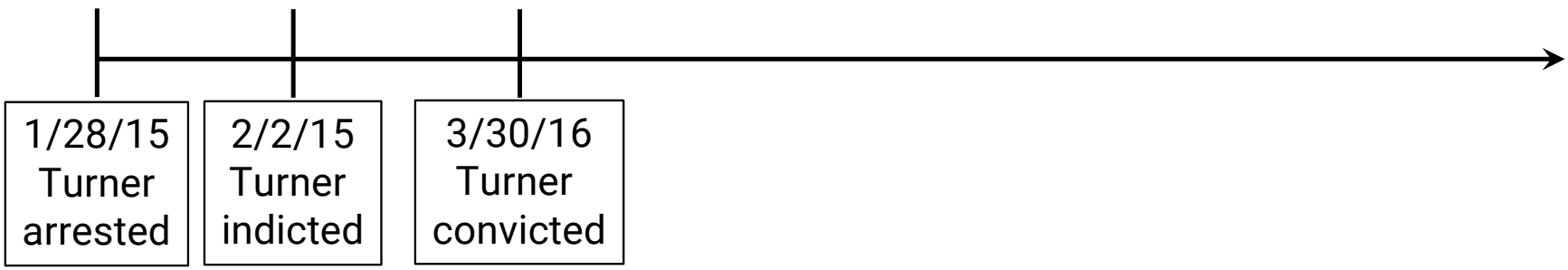
Timeline: the recall of Judge Persky



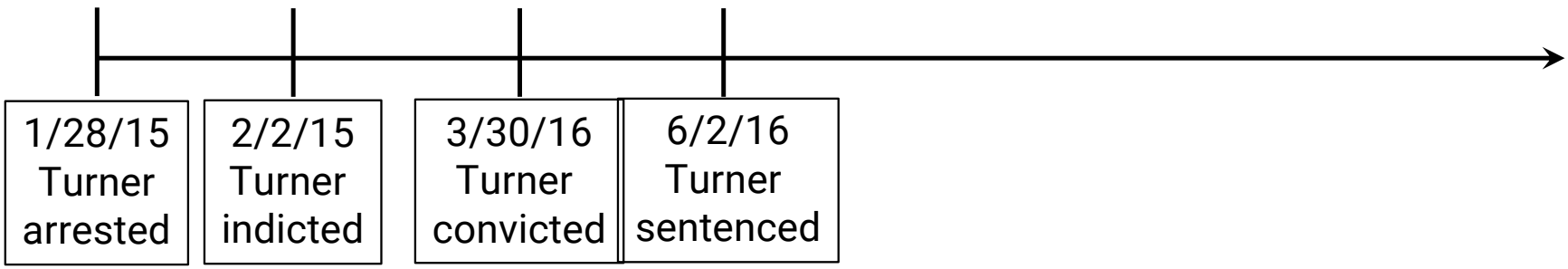
Timeline: the recall of Judge Persky



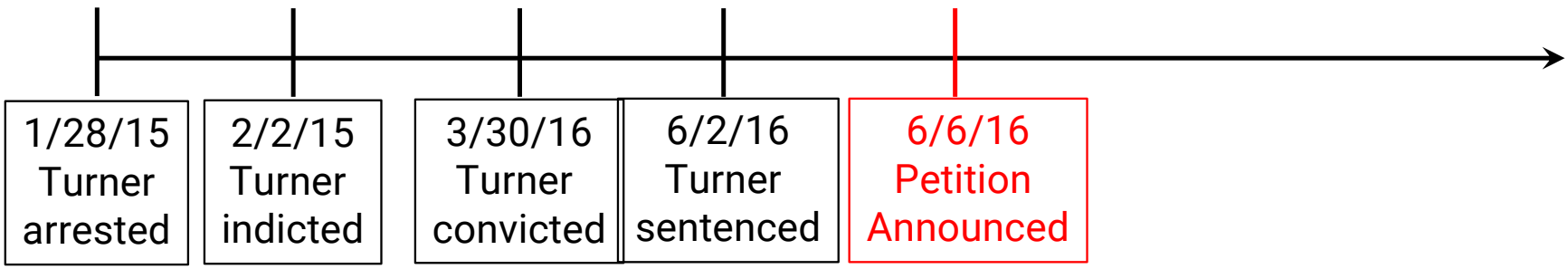
Timeline: the recall of Judge Persky



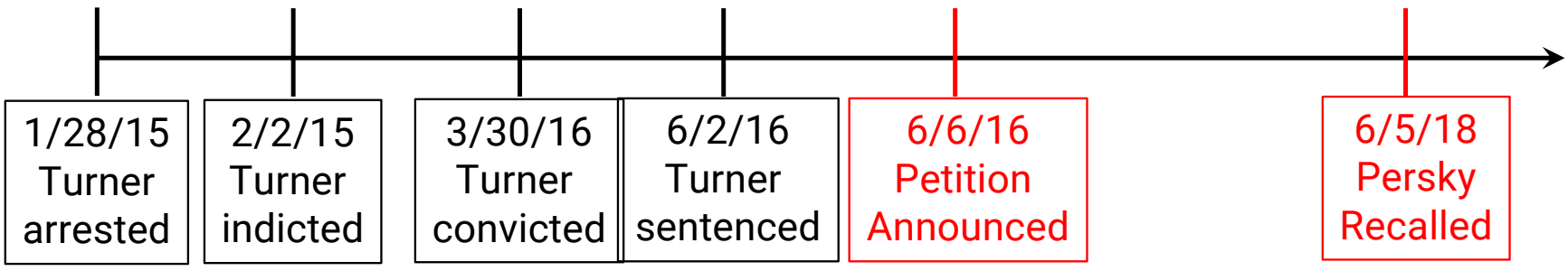
Timeline: the recall of Judge Persky



Timeline: the recall of Judge Persky



Timeline: the recall of Judge Persky



Empirical approach

1. Immediate effect of recall salience
   Let t_k be calendar date of critical event k. We estimate:

$$y_{ijt} = \beta_0 + \beta_1 \mathbb{I}(t > t_k) + \beta_2 f(t - t_k) + \cdots$$

   where $y_{ijt} \equiv \min\{s_i/\overline{s}_i, 1\}$ is normalized sentence of conviction i by judge j at time t; $f(\cdot)$ is smooth function

Empirical approach

1. Immediate effect of recall salience
   Let t_k be calendar date of critical event k. We estimate:

$$y_{ijt} = \beta_0 + \beta_1 \mathbb{I}(t > t_k) + \beta_2 f(t - t_k) + \cdots$$

   where $y_{ijt} \equiv \min\{s_i/\overline{s}_i, 1\}$ is normalized sentence of conviction i by judge j at time t; $f(\cdot)$ is smooth function
   - evaluate whether recall affected prosecutorial behavior;
   - conduct placebo tests for contemporaneous shocks.

Empirical approach

1. Immediate effect of recall salience
   Let t_k be calendar date of critical event k. We estimate:

$$y_{ijt} = \beta_0 + \beta_1 \mathbb{I}(t > t_k) + \beta_2 f(t - t_k) + \cdots$$

   where $y_{ijt} \equiv \min\{s_i/\overline{s}_i, 1\}$ is normalized sentence of conviction i by judge j at time t; $f(\cdot)$ is smooth function
   - evaluate whether recall affected prosecutorial behavior;
   - conduct placebo tests for contemporaneous shocks.
2. Heterogeneous effects by severity, race

Empirical approach

1. Immediate effect of recall salience
   Let t_k be calendar date of critical event k. We estimate:

$$y_{ijt} = \beta_0 + \beta_1 \mathbb{I}(t > t_k) + \beta_2 f(t - t_k) + \cdots$$

   where $y_{ijt} \equiv \min\{s_i/\overline{s}_i, 1\}$ is normalized sentence of conviction i by judge j at time t; $f(\cdot)$ is smooth function
   - evaluate whether recall affected prosecutorial behavior;
   - conduct placebo tests for contemporaneous shocks.
2. Heterogeneous effects by severity, race

Empirical approach

1. Immediate effect of recall salience
   Let t_k be calendar date of critical event k. We estimate:

$$y_{ijt} = \beta_0 + \beta_1 \mathbb{I}(t > t_k) + \beta_2 f(t - t_k) + \cdots$$

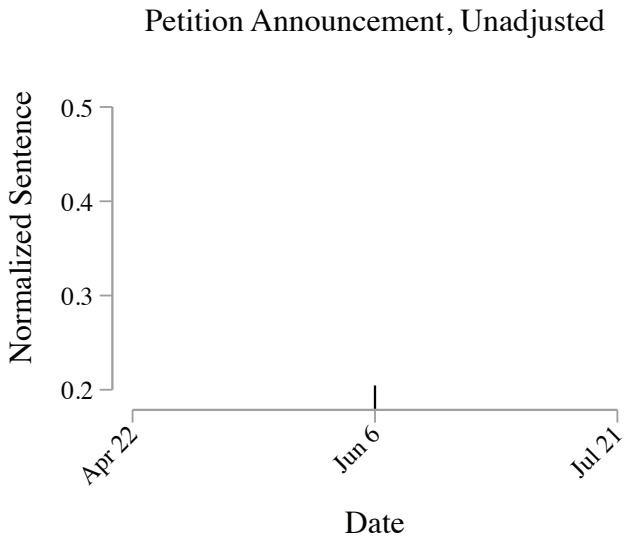
   where $y_{ijt} \equiv \min\{s_i/\overline{s}_i, 1\}$ is normalized sentence of conviction i by judge j at time t; $f(\cdot)$ is smooth function
   - evaluate whether recall affected prosecutorial behavior;
   - conduct placebo tests for contemporaneous shocks.
2. Heterogeneous effects by severity, race
3. Cumulative effects of recall shock

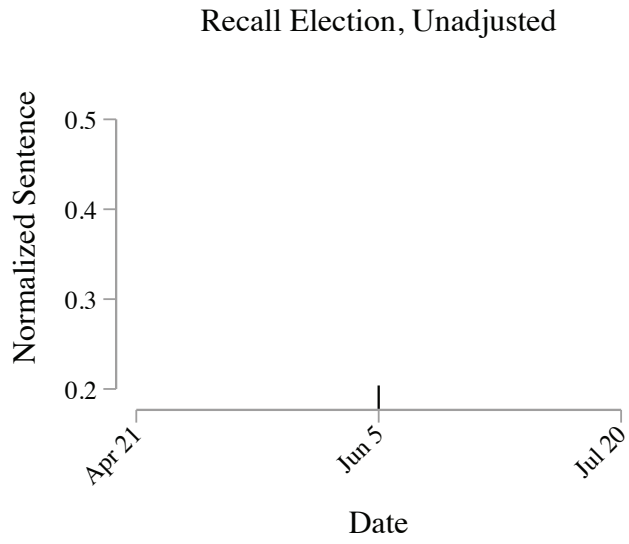
Data: sentencing and defendants

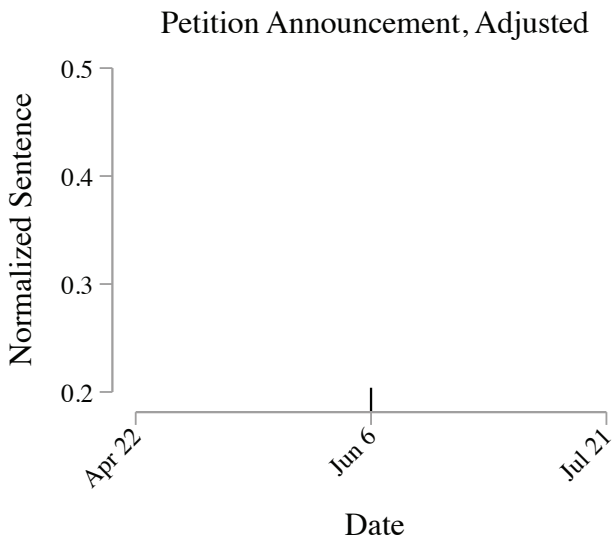
- No centralized repository of California sentencing data
- Crawled 494,840 CA superior court cases from six counties (Fresno, Napa, Santa Barbara, Sacramento, San Bernardino, Santa Cruz)
  - Sacramento data missing judge identifier
- Crawled 201,066 arrest records to get defendant characteristics (e.g., race)
- Extracted 22,139 criminal felony charges involving 19,845 cases (Jan 2015 to Dec 2018, inclusive)
  - Successfully matched with 12,844 arrest records

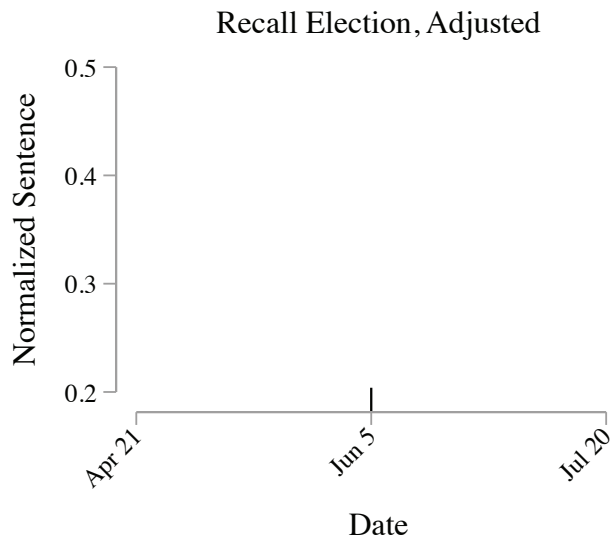
Immediate effect of recall salience

Graphical Analysis

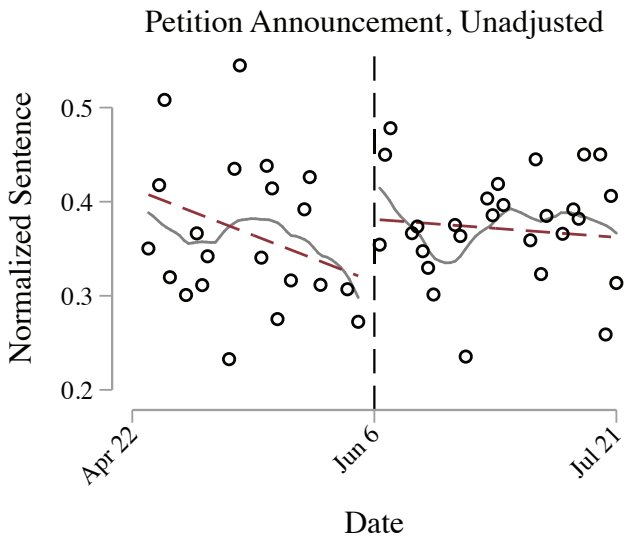


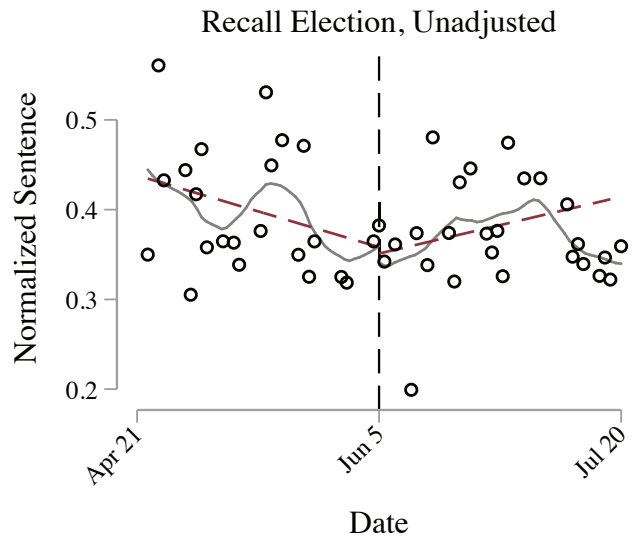


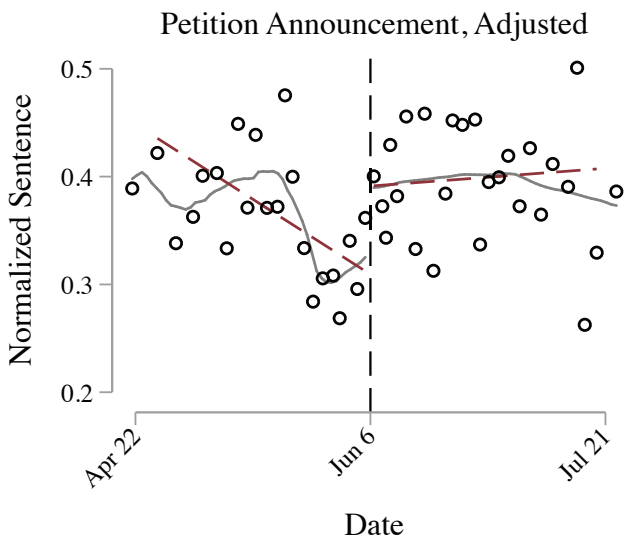


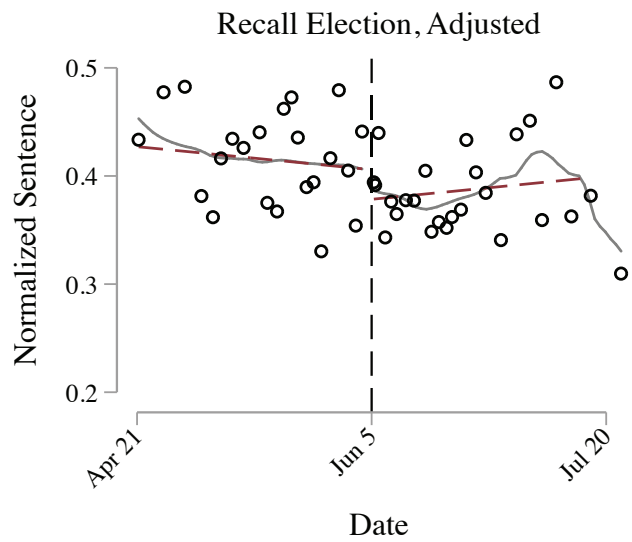


Graphical Analysis

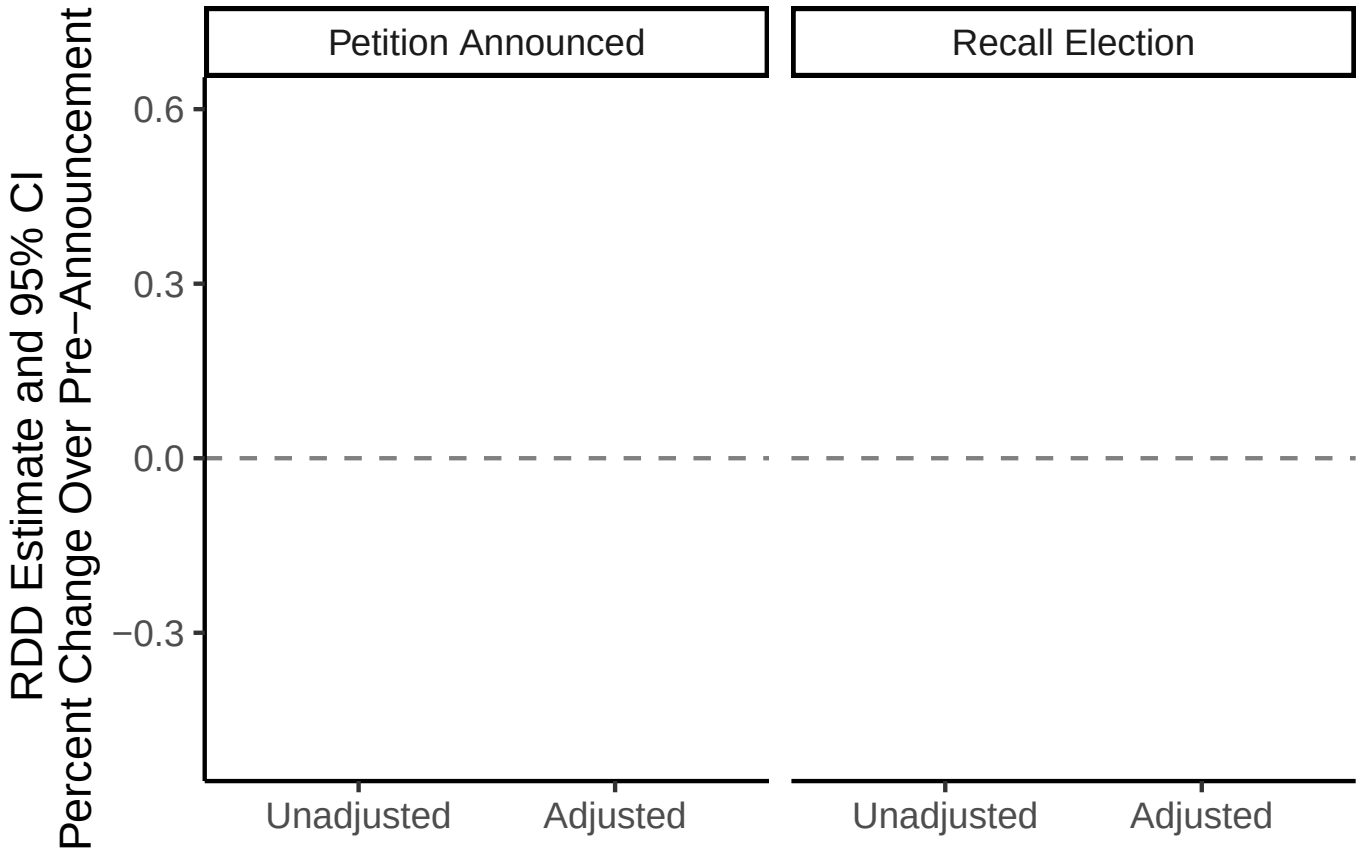




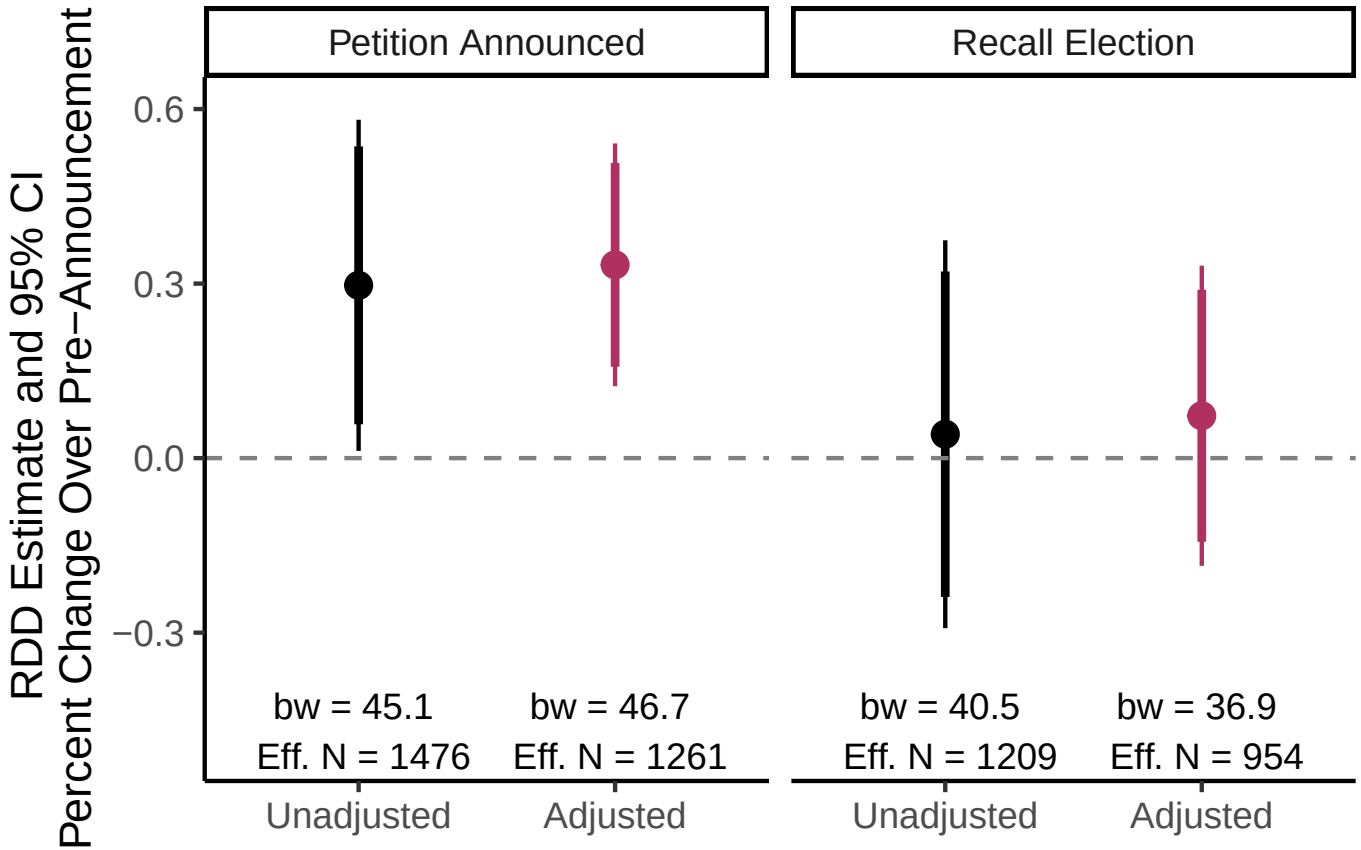




RDit Estimates



RDit Estimates



- Implies instantaneous 33% increase in average sentence length on day petition is announced

Robustness

Prosecutors or Judges?

1. Prosecutors may file more severe charges.
   - Substitute statutory maximum for the top count at *arraignment* – not conviction
   - Restrict sample to cases arraigned before the critical events. Estimates

Prosecutors or Judges?

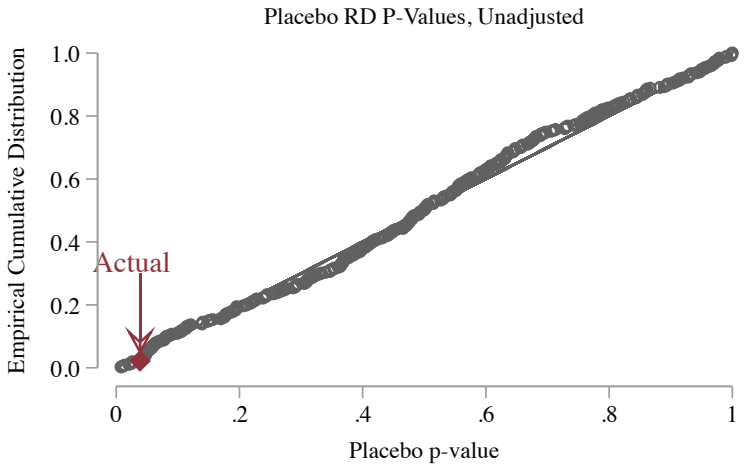
1. Prosecutors may file more severe charges.
   - Substitute statutory maximum for the top count at *arraignment* – not conviction
   - Restrict sample to cases arraigned before the critical events. Estimates

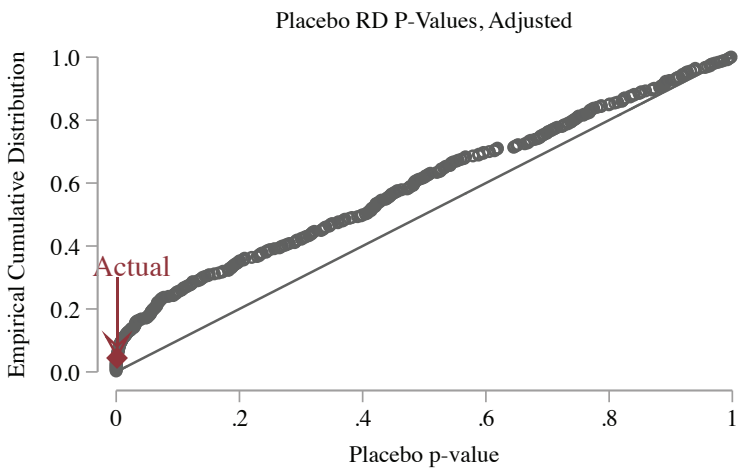
2. Prosecutors might seek fewer charge reductions. Estimates

Prosecutors or Judges?

1. Prosecutors may file more severe charges.
   - Substitute statutory maximum for the top count at *arraignment* – not conviction
   - Restrict sample to cases arraigned before the critical events. Estimates
2. Prosecutors might seek fewer charge reductions. Estimates
3. Prosecutors might otherwise alter the composition of convictions Estimates

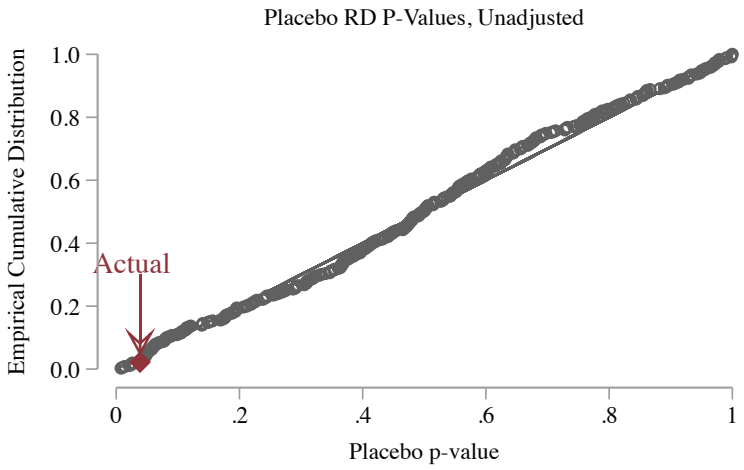
Placebo tests: main effect

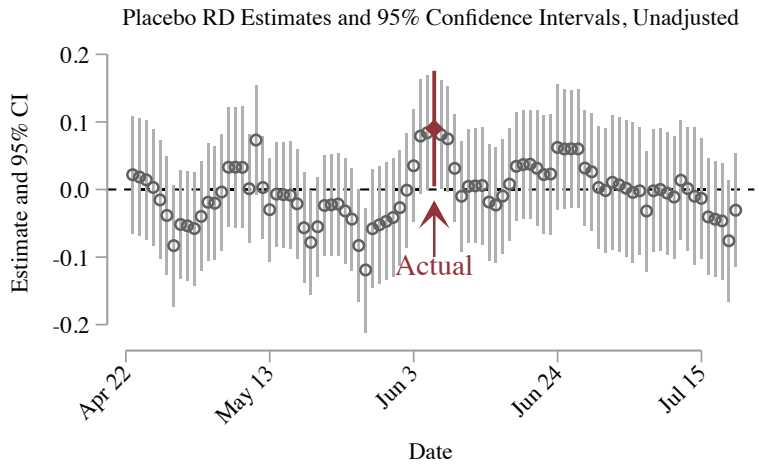


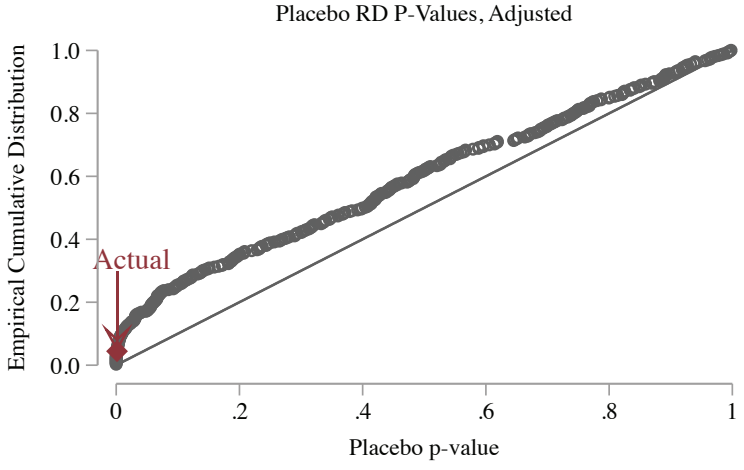


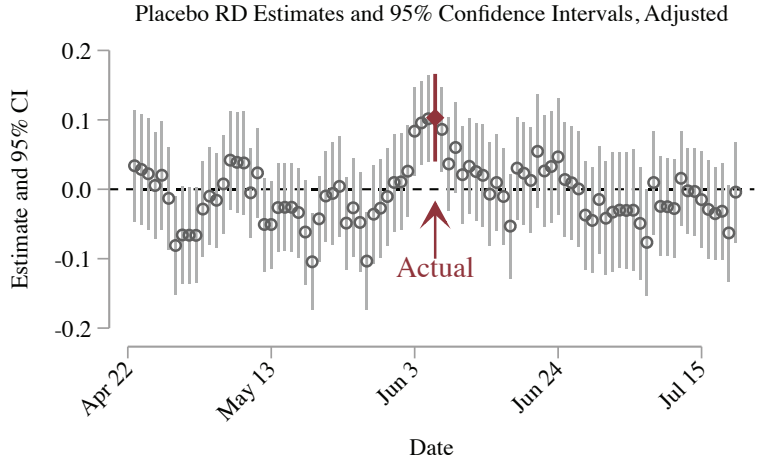
- June 6 p-value lower than 98.6% of placebos

Placebo tests: main effect









- Only larger estimates are August 24/25
- RD using June 2 cutpoint small, statistically insignificant

Effect heterogeneity: RDiT estimates

Effect driven by violent crimes

	Sex Crimes		Other Violent Crimes		Nonviolent Crimes	
RD estimate	-0.033	-0.003	0.248	0.188	0.077	0.106
	(0.17)	(0.048)	(0.098)	(0.056)	(0.048)	(0.035)
Left-side intercept	0.29	0.448	0.211	0.2	0.318	0.322
	(0.133)	(0)	(0.046)	(0.031)	(0.028)	(0.022)
Bandwidth	65.5	33.9	43.8	38.1	46.5	51.6
Adjusted	N	Y	N	Y	N	Y
Effective observations	85	29	207	151	1,232	1,131

Estimates employ triangular kernel. Standard errors clustered at the county-charge level/judge-charge level

No evidence for direct racial burden hypothesis

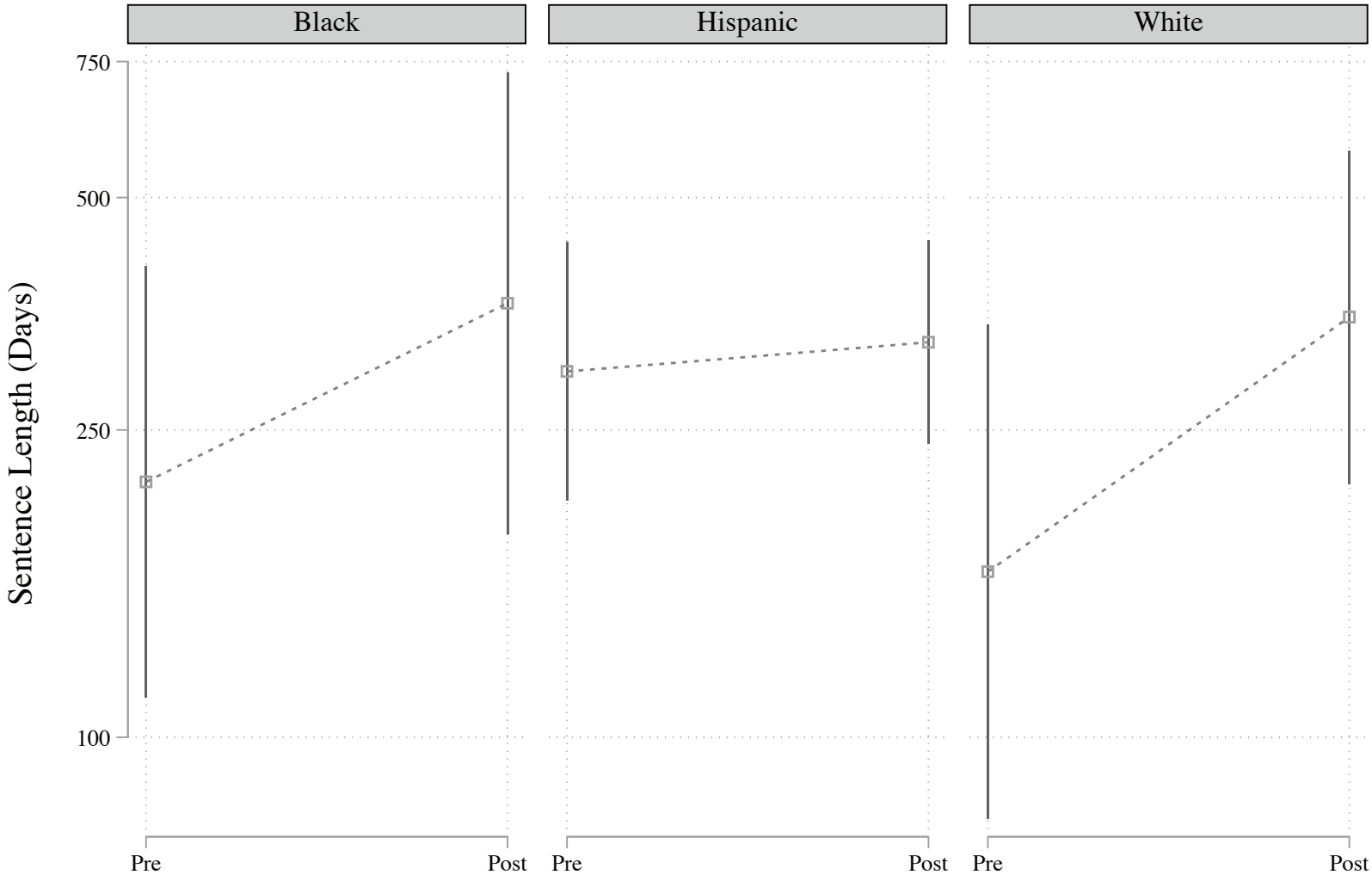
No evidence for direct racial burden hypothesis

	Petition Announced					
	Blacks		Hispanics		Whites	
RD estimate	0.136	0.24	0.098	0.062	0.206	0.545
	(0.12)	(0.074)	(0.059)	(0.045)	(0.096)	(0.072)
Left-side intercept	0.355	0.37	0.302	0.31	0.247	0.225
	(0.066)	(0.055)	(0.036)	(0.026)	(0.064)	(0.05)
Bandwidth	64.2	41.6	65.9	54.1	56	26.6
Adjusted	N	Y	N	Y	N	Y
Effective observations	304	136	689	516	326	126

Estimates employ triangular kernel. Standard errors clustered at the county-charge level/judge-charge level

Assessing evidence of indirect racial burden

Assessing evidence of indirect racial burden



Cumulative Effects

Substantive Impact

RD estimates are LATEs: causally identified, but of questionable substantive relevance

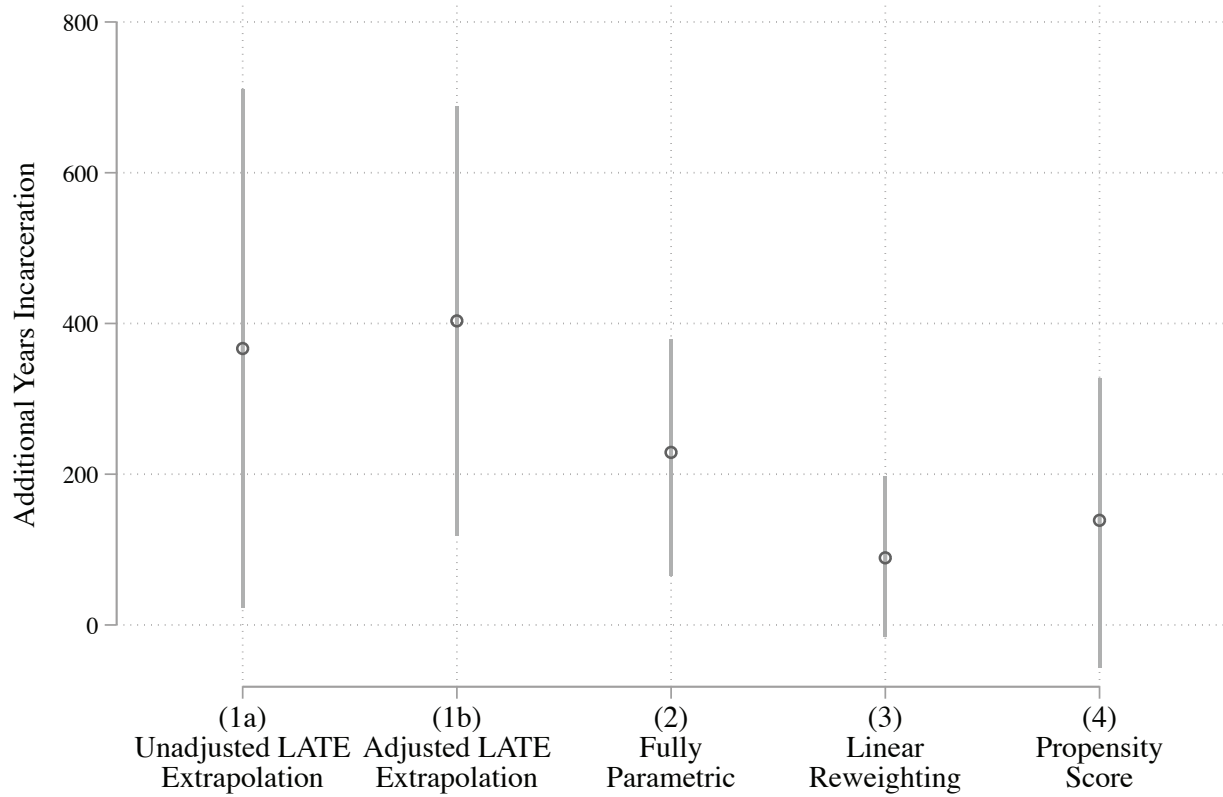
Assessing cumulative effects requires stronger assumptions.

Rather than extrapolate over a longer period, we restrict attention to 45 day window.

Three different extrapolation approaches:

- Assume LATE is the ATE (unadjusted and adjusted)
- Fully parametric
- Linear reweighting and propensity score estimators (see Angrist and Rokkanen 2015).

Cumulative Effects



Using the most conservative 88 year estimate, total cost to five counties is $6.25 million. Total effect statewide is 733 years or $52.1 million.

Conclusion

Conclusion

Judges highly attuned to political environment and alter behavior accordingly

- Announcement of recall petition creates “new normal” in which recall now applies to lower court judges
- Other important events have no effect, because either anticipated or don’t reflect comparable shift in norms

Conclusion

Judges highly attuned to political environment and alter behavior accordingly

- Announcement of recall petition creates “new normal” in which recall now applies to lower court judges
- Other important events have no effect, because either anticipated or don’t reflect comparable shift in norms

Indirect effects of direct democracy

- Persky recall aimed at sanctioning judge for sentence of affluent white defendant for felony sexual assault
- But once contemplated, impossible to control circumstances under which recall can be used to sanction lenient sentences

Appendix

Replication of Main Analysis Using Non-Normalized Sentence Length as Outcome

	Petition Announced		Recall Election	
RD estimate	172.941	194.651	-31.927	-41.616
	(82.756)	(68.852)	(147.211)	(67.915)
Left-side intercept	401.934	393.664	602.326	600.891
	(40.154)	(31.04)	(127.296)	(57.545)
Bandwidth	55.7	55.4	48.6	38
Judge fixed effects	N	Y	N	Y
Statute fixed effects	N	Y	N	Y
Effective observations	1777	1466	1461	966

Replication of Main Analysis Restricted to Cases with One Count

	Petition Announced		Recall Election	
RD estimate	0.085	0.104	0.019	0.036
	(0.042)	(0.035)	(0.056)	(0.048)
Left-side intercept	0.291	0.297	0.321	0.332
	(0.029)	(0.025)	(0.046)	(0.040)
Bandwidth	47.509	45.807	40.836	37.490
Adjusted	N	Y	N	Y
Effective observations	1421	1140	1108	872

Effect on Sentence Normed to Top Arraignment Count

	Petition Announced		Recall Election	
RD estimate	0.144	0.158	0.044	0.03
	(0.056)	(0.043)	(0.059)	(0.043)
Left-side intercept	0.282	0.278	0.316	0.303
	(0.027)	(0.022)	(0.047)	(0.035)
Bandwidth	46.7	48.2	39.1	34.8
Adjusted	N	Y	N	Y
Effective observations	1326	1180	1104	820

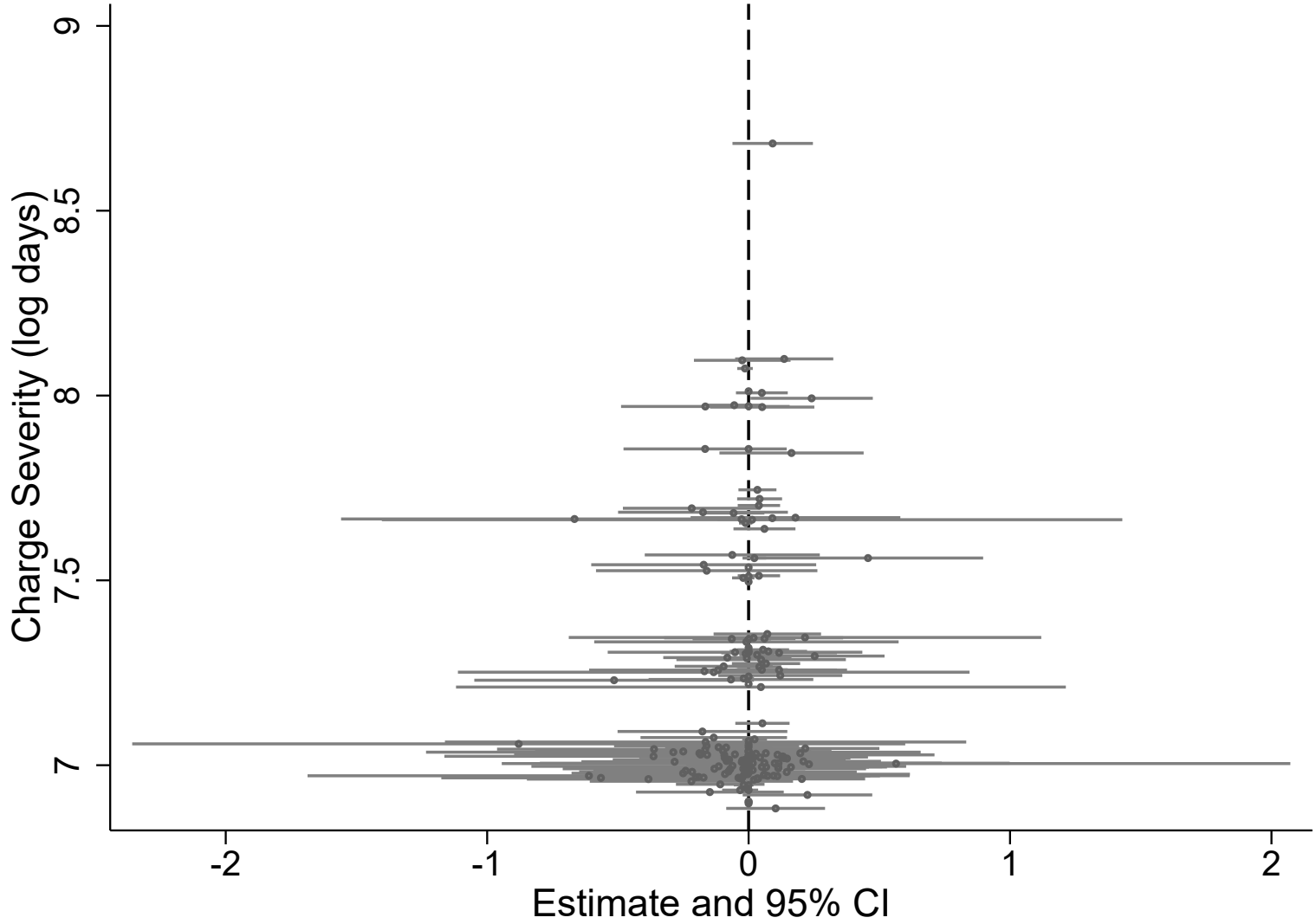
Back to Presentation

Effect on Charge Reductions

	Petition Announced		Recall Election	
	B. Charge Reduction			
RD estimate	-0.022	-0.03	-0.017	-0.022
	(0.027)	(0.019)	(0.023)	(0.022)
Left-side intercept	0.104	0.097	0.064	0.072
	(0.022)	(0.015)	(0.018)	(0.018)
Bandwidth	45.6	45.4	45	39
Adjusted	N	Y	N	Y
Effective observations	1268	1088	1209	915

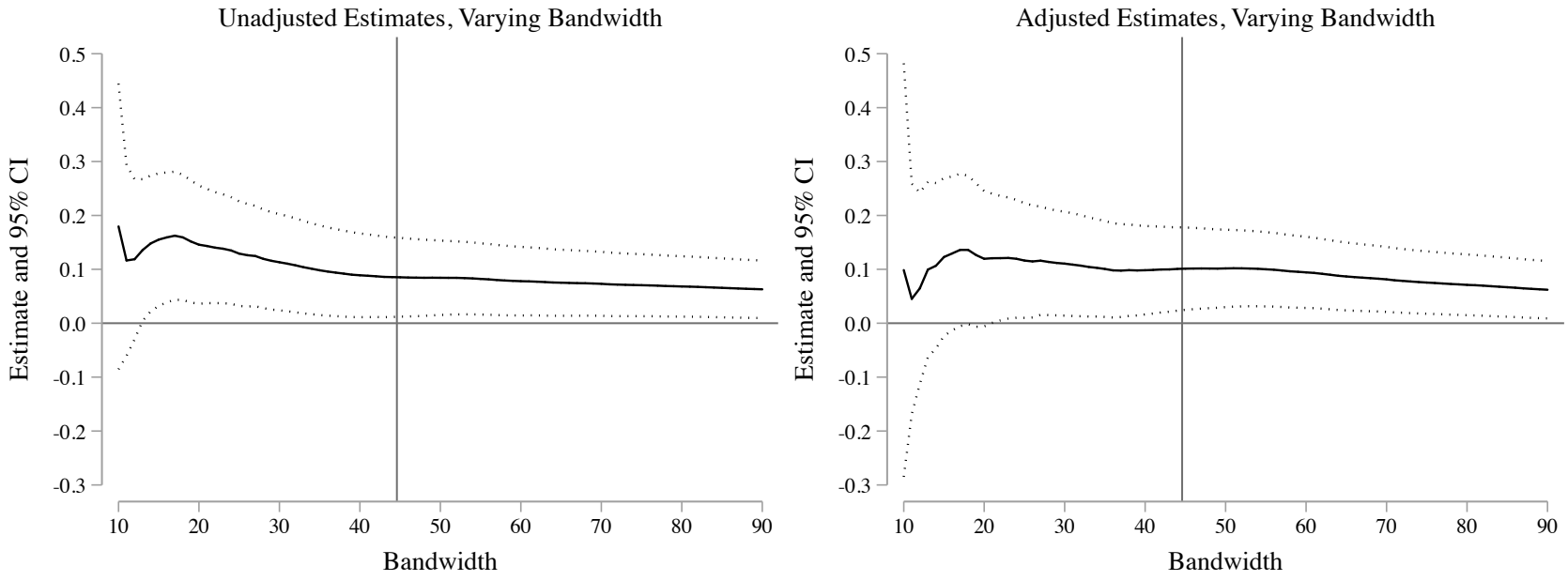
Back to Presentation

Charge-FE RD Estimates and 95% Confidence Intervals



Each grey circle (and grey line) represents the RD estimate (and 95% confidence interval) associated with a unique crime's daily count.

RD Estimates Varying Bandwidth



As in the main analysis, estimates employ triangular kernel, with standard errors clustered at the judge-charge level. The solid line denotes the MSE-optimal bandwidth.

Effect of Petition Announcement in Washington State: Placebo Test