Online Appendix to

"Rules Versus Norms: How Formal and Informal Institutions Shape Judicial Sentencing Cycles"

by Christian Dippel and Michael Poyker

Online Appendix A Data Description

Online Appendix A.1 Sentencing Data

Sentencing data was collected separately from each state. 15 states were willing to share their data with us for free or at reasonable cost: Alabama, Arkansas, Georgia, Kentucky, Maryland, Minnesota, Mississippi, Nevada, North Carolina, Oregon, Pennsylvania, Tennessee, Texas, Virginia, and Washington.

We contacted each state with the following initial data request:
The data we are looking for has a court case (or 'sentencing event') as the unit of observation. In some states the data is organized by charge (with several charges making up the case or sentencing event) and that is equally fine. The key data that we need are:
1. date, month and year of sentencing,
2. type of crime,
3. length of sentencing,
4. type of sentencing (low-security, high security, etc),
5. defendant's sex,
6. defendant's race,
7. court identifier
8. name of judge or judge identifier number,
9. type of court that convicted (trial, appeal, etc),
10. in what prison the person was sent
We do not seek any information that identifies defendants.
Sincerely, XXX

There were 10 states that (*i*) shared their sentencing data in digitized form and (*ii*) their data included the judge identifiers needed to estimate judge political cycles.[32] The following reports for each state the office responsible for storing the data, as well as relevant contacts at the time we requested the data between late 2016 and late 2018. Some states had considerably longer processing times than others. These were typically do either to backlogs of data-technicians or to having to go get our request vetted and signed off on by other individuals.

1. Alabama
    - Initial contact with the Sentencing Commission at http://sentencingcommission.alacourt.gov/
    - After emailing sentencing.commission@alacourt.gov, Bennet Wright processed our request.
    - Time between data application and delivery: 16 months.
2. Colorado
    - Initial contact with the Colorado Court Services Division, at https://www.courts.state.co.us/Administration/Division
    - Jessica Zender, the Court Programs Analyst at the Court Services Division processed our request.

[32] We also obtained sentencing data from Arkansas, Maryland, Mississippi, Nevada, Oregon, and Texas, but these states' data does not include judge identifiers

- Time between data application and delivery: 1 month.

3. Georgia
   - Initial contact with Department of Corrections at http://www.dcor.state.ga.us/Divisions/ExecutiveOperations/OPS/OpenRecords.
   - After emailing open.records@gdc.ga.gov it was recommended we go through their 'Media Inquiries' under +1-478-992-5247, where Jamila Coleman coordinated our request with their data technicians.
   - Time between data application and delivery: 3 months.
4. Kentucky
   - We spoke on the phone to Cathy Schiflett at the Kentucky Courts Research and Statistics Department.
   - She guided us to https://courts.ky.gov/Pages/default.aspx, where we had to select 'Statistical Reports' and then submit our data request.
   - Daniel Sturtevant handled our request.
   - Time between data application and delivery: 9 months.
5. Minnesota
   - Initial contact with the Minnesota Sentencing Guidelines Commission at http://mn.gov/sentencing-guidelines/contact/contact-us.jsp
     Email address: sentencing.guidelines@state.mn.us
   - Kathleen Madland was the Research Analyst who processed our request
   - Time between data application and delivery: 2 months
6. North Carolina
   - Initial contact though http://www.ncdoj.gov/Top-Issues/Public-Integrity/Open-Government/Understanding-Public-Records.aspx
   - Then we were put in touch with the North Carolina Administrative Office of the Courts, where our data request was processed by the 'Remote Public Access' data technicians
   - Time between data application and delivery: 3 months
7. Pennsylvania
   - In Pennsylvania, sentencing data can be requested from the Sentencing Commission at http://pcs.la.psu.edu/data/request-and-obtain-data-reports-and-data-sets/sentencing/data-sets
   - Leigh Tinik processed our request
   - Time between data application and delivery: 1 month
8. Tennessee
   - Initial contact with Tennessee's Department of Corrections at https://www.tn.gov/correction/article/tdoc-prison-directory

- Tanya Washington, the DOC's Director of Decision Support: Research & Planning, processed our request
- Time between data application and delivery: 6 months

9. Virginia
   - Initial contact was through a web-form of the Virginia Criminal Sentencing Commission at http://www.vcsc.virginia.gov/
   - After being initially denied on the grounds that FOIA requests could only be processed for Virginia residents, we called +1-804-225-4398, and were eventually approved after speaking to the director Meredith Farrar-Owens.
   - Time between data application and delivery: 3 months
10. Washington
    - Initial contact with the Department of Corrections at http://www.doc.wa.gov/aboutdoc/publicdisclosure.asp, where Duc Luu processed our request
    - We use essentially the same data as Berdejó and Yuchtman (2013)
    - Time between data application and delivery: 2 weeks

Online Appendix A.2 Judicial Biography Data

All data about judge electoral cycles was taken from the ballotpedia.org. The site contain information about the judges of each circuit court for each state.[33] The individual page of each judge contain data for age and gender of a judge, the dates when she was appointed/elected, date of retirement (if already retired), name of a governor by whom she was appointed (if appointed), and whom the judge replaced.

To collect the data research assistants started with the contemporary judges, collected their data and proceeded with their predecessor judges. This procedure resulted in collecting information for approximately 80% of the judges mentioned in the sentencing data. For the states where the name of a judge was known we searched those judges individually on the sites of their courts and added them to the dataset.

Ten of the states in this paper include judge names or identifiers in the sentencing data: Alabama, Colorado, Georgia, Kentucky, Minnesota, North Carolina, Pennsylvania, Tennessee, Virginia, and Washington. We then code up judge biographies, including when they are up for reelection from Where judges are identified by name, merging the judge biographies is straightforward. Where only judge identifiers are given, these identifiers still almost always include a variant of the judges' initials. As well we observe entry and exit dates and which circuit a judge id is identified with.

[33] Or courts of the similar level.

Online Appendix B Robustness Checks

Online Appendix Table 1 reports on a version of Table 3 where we alternatively cluster by calendar-quarter, replacing $4 + t$ fixed effects with $4 \times t$ fixed effects (for t years). We find similar results.

Table Online Appendix Table 1: Baseline with S.E. Clustered by Calendar-Quarter

	Dependent variable: Sentence (months)									
	WA	GA	KY	MN	NC	AL	TN	PA	CO	VA
Panel A: ~ year x quarter s.e.										
Proximity to election	4.318	1.013	0.346	2.595	2.989***	4.375	3.822	15.214*	-0.500	-7.739
	(7.8991)	(3.2327)	(9.0335)	(2.8742)	(0.9593)	(9.5358)	(5.9321)	(7.7291)	(11.6138)	(8.8728)
R-squared	0.527	0.737	0.420	0.386	0.423	0.119	0.656	0.389	0.465	0.323
Observations	13,124	2,433	4,150	11,888	34,906	10,701	20,515	12,866	6,395	2,412

Notes: *(a)* Each panel reports on results of a one specification, run for each state separately across columns. *(b)* All regressions include defendant's race, gender, age, age squared, and an indicator for recidivism. All regressions also include the case's severity, and the number of charges in each case. Finally, all regressions include judge fixed effects and year as well as three quarter-of-year fixed effects. *(c)* This table replicates the baseline specification from Panel A of Table 3 but clusters standard errors by calendar quarter (i.e., quarter-year). *** p<0.01, ** p<0.05, * p<0.1

Online Appendix Table 2: Baseline with Days of the Cycle

	Dependent variable: Sentence (months)									
	WA	GA	KY	MN	NC	AL	TN	PA	CO	VA
Panel A: ~ x days of cycle										
Proximity to election	0.0030**	0.0007	0.0001	0.0012	0.0010**	0.0020	0.0013	0.0042	-0.0002	-0.0027
	(0.0008)	(0.0035)	(0.0031)	(0.0009)	(0.0003)	(0.0061)	(0.0021)	(0.0025)	(0.0070)	(0.0019)
R-squared	0.527	0.737	0.420	0.386	0.423	0.119	0.656	0.389	0.465	0.323
Observations	13,124	2,433	4,150	11,888	34,906	10,701	20,515	12,866	6,395	2,412

Notes: *(a)* Each panel reports on results of a one specification, run for each state separately across columns. *(b)* All regressions include defendant's race, gender, age, age squared, and an indicator for recidivism. All regressions also include the case's severity, and the number of charges in each case. Finally, all regressions include judge fixed effects and year as well as three quarter-of-year fixed effects. *(c)* This Table replaces the linear regressor from expression (1) with a "number of days since the election"; for example, Washington's electoral cycle is 4 years, i.e., 1461 days, so that the coefficient goes from 0 to 1461 instead of from 0 to 1. *(d)* Standard errors are two-way clustered by calendar-year and quarter-of-year in all panels. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Online Appendix Table 3: Robustness to State-Specific Data Issues

	Dependent variable: Sentence (months)									
	WA	GA	KY	MN	NC	AL	TN	PA	CO	VA
Panel A: redefine recidvism as dummy										
Proximity to election	4.318**	0.899	-	2.485	3.174**	-	4.653	14.880	-	-
	(1.3219)	(5.0656)	-	(2.0686)	(0.7896)	-	(5.9734)	(9.1835)	-	-
R-squared	0.527	0.736		0.385	0.418		0.654	0.386		
Observations	13,124	2,434		11,888	34,915		20,517	12,866		
Panel B: w/o recidivism variable										
Proximity to election	4.550**	0.906	-	3.252	3.039**	-	3.964	15.189	-	-
	(1.1617)	(4.9945)	-	(1.9956)	(0.9334)	-	(5.7643)	(9.7623)	-	-
R-squared	0.509	0.736		0.340	0.403		0.645	0.370		
Observations	13,124	2,434		11,888	34,915		20,517	12,866		
Panel C: add sent. guidelines										
	4.062**	-	-	1.410	-	1.132	-	32.454*	-	-
	(0.9546)	-	-	(1.7554)	-	(12.9622)	-	(12.9622)	-	-
R-squared	0.810			0.713		0.803		0.510		
Observations	13,124			11,888		10,701		12,866		

Notes: *(a)* Each panel reports on results of a one specification, run for each state separately across columns. *(b)* All regressions include defendant's race, gender, age, age squared, and an indicator for recidivism. All regressions also include the case's severity, and the number of charges in each case. Finally, all regressions include judge fixed effects and year as well as three quarter-of-year fixed effects. *(c)* This Table's specifications are based on Panel A of Table 3. Panels A–B re-code recidivism as a dummy, or omit it. (The number of observations goes up in some states because fewer observations are absorbed by recidivism-category fixed effects.) Panel C add sentencing guidlines as a control variable. *(d)* Standard errors are two-way clustered by calendar-year and quarter-of-year. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

To confirm that the broad patterns in the sentencing data are consistent in all states, Online Appendix Table 4 reports the coefficients on defendant characteristics (race, gender and recidivism) that went unreported Table 3. All of these patterns have the expected signs, match previous research, and are sign consistent with each other: judges in all states pass shorter sentences for women, judges in all but one state pass longer sentences for black defendants, and judges in all states pass longer sentences for recidivists. The coefficient on recidivism is the most variable across the states, which reflects the fact that —unlike the race and gender dummies— the recidivism dummy can cloud substantial variation in the degree of recidivism. This is discussed in footnote 20.

Online Appendix Table 4: Effect of Defendant Characteristics on Sentence Length

	Dependent variable: Sentence (months)									
	WA	GA	KY	MN	NC	AL	TN	PA	CO	VA
Female	-4.982***	-7.566***	-3.277***	-5.513***	-3.748***	-15.466***	-13.180***	-1.750***	-2.846	-
	[0.0007]	[0.0000]	[0.0063]	[0.0001]	[0.0001]	[0.0001]	[0.0001]	[0.0004]	[0.2406]	-
Black	2.291***	2.250**	2.251*	3.101***	1.552***	6.390***	5.322***	1.961***	-0.058	-
	[0.0030]	[0.0138]	[0.0582]	[0.0005]	[0.0018]	[0.0014]	[0.0035]	[0.0019]	[0.9709]	-
Recidivist, (0 or 1)	11.946***	69.223***	-	29.499***	10.267***	-	140.035***	10.779***	-	-
	[0.0000]	[0.0005]	-	[0.0000]	[0.0000]	-	[0.0000]	[0.0001]	-	-
R-squared	0.564	0.254	0.802	0.542	0.420	0.125	0.477	0.360	0.492	
Observations	139,900	100,413	81,442	122,616	251,907	94,071	215,539	463,236	53,683	

Notes: *(a)* Each panel reports on results of one specification, run for each state separately. All cases included. *(b)* We use dummy for recidivism instead of the scaled variable for the sake of data representation; however, estimates for the proximity to election do not change if we use scaled recidivism. *(c)* We report p-values in square brackets. *** p<0.01, ** p<0.05, * p<0.1

Online Appendix Table 5: Pooled Panel with Interactions - Data Quality

	I	II	III	IV
	Dependent variable: Sentence (months)			
Proximity to election	2.624	7.379*	3.717	3.632
	(2.9788)	(3.5866)	(3.0241)	(2.0652)
Proximity to election				
x D(Observed recidivism)	1.556			
	(2.4891)			
x D(Concurently sentence)		-5.496		
		(3.1115)		
x D(Ordinal severity categories)			0.048	
			(3.6533)	
x D(# of crime categories)				0.0003
				(0.0031)
R-squared	0.445	0.445	0.445	0.445
Observations	116,082	116,082	116,082	116,082

Notes: *(a)* This table re-runs Column I of Table 7 with interactions related to state-level data quality. *(b)* Standard errors are multi-way-clustered by quarter-year and state level. *** p<0.01, ** p<0.05, * p<0.1

Online Appendix Table 6: Baseline with Alternative Fixed Effects

	Dependent variable: Sentence (months)									
	WA	GA	KY	MN	NC	AL	TN	PA	CO	VA
Panel A: ~ county x quarter FEs										
Proximity to election	4.8208***	0.6746	1.4409	2.5621	3.0358**	7.1997	3.7154	15.3821	-1.5959	-5.3227
	(0.6155)	(3.6948)	(11.4120)	(2.1447)	(0.7857)	(12.9804)	(6.1508)	(9.4199)	(14.5865)	(5.8507)
R-squared	0.532	0.761	0.451	0.390	0.429	0.127	0.661	0.395	0.470	0.369
Observations	13,124	2,433	4,150	11,888	34,906	10,701	20,515	12,866	6,395	2,412
Panel B: ~ county x year x FEs										
Proximity to election	7.2910**	-1.0076	11.0339	2.9153	3.3364**	0.8342	2.5250	15.7413	-1.2419	9.1324
	(1.3671)	(3.4039)	(21.8589)	(3.0298)	(0.8374)	(12.6809)	(8.1264)	(8.9478)	(13.8308)	(9.6950)
R-squared	0.541	0.758	0.536	0.404	0.437	0.148	0.676	0.425	0.476	0.401
Observations	13,124	2,433	4,150	11,888	34,906	10,701	20,515	12,866	6,395	2,412

Notes: *(a)* Each panel reports on results of a one specification, run for each state separately across columns. *(b)* All regressions include defendant's race, gender, age, age squared, and an indicator for recidivism. All regressions also include the case's severity, and the number of charges in each case. Finally, all regressions include judge fixed effects and year as well as three quarter-of-year fixed effects. *(c)* This Table replicates the baseline specification from Panel A of Table 3 but uses alternative fixed effects. Panel A adds county-quarter fixed effects. Panel B adds county-year fixed effects. *(d)* Standard errors are two-way clustered by calendar-year and quarter-of-year in all panels. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

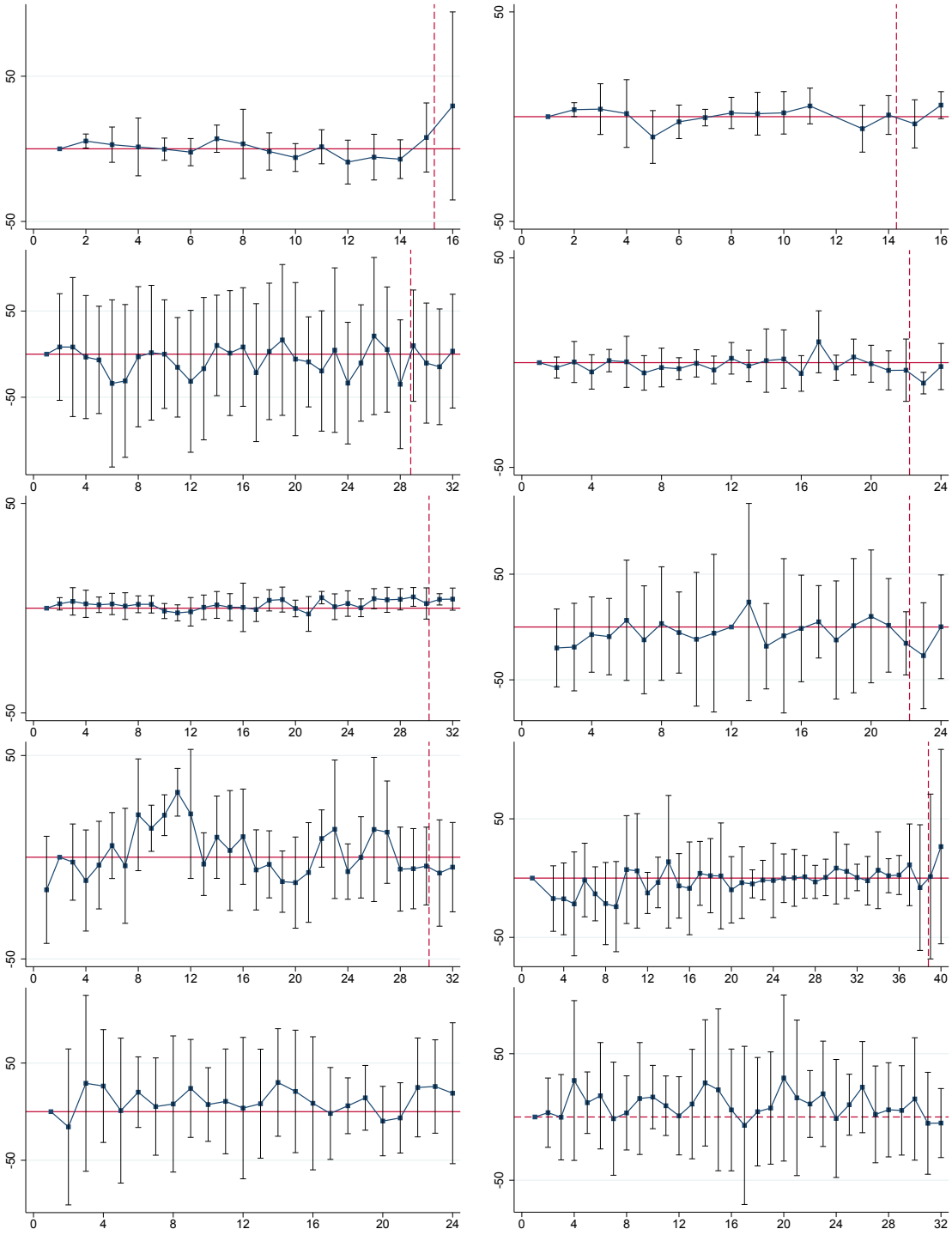
Online Appendix Table 7: Quarters and Filing Date

	Dependent variable: Sentence (months)									
	WA	GA	KY	MN	NC	AL	TN	PA	CO	VA
Panel A: Include Cases after Filing-Date										
Proximity to election	8.822**	1.629	-2.250	-0.574	2.844**	3.270	6.166	13.988	-	-
	(2.7494)	(1.9602)	(7.5145)	(1.5761)	(0.7428)	(11.3719)	(3.2564)	(7.6693)	-	-
R-squared	0.517	0.571	0.420	0.386	0.421	0.105	0.641	0.427		
Observations	13,774	7,664	4,250	15,532	35,290	14,045	24,866	17,781		
Panel B: Quarters (Ordinal-Scale)										
Proximity to election	0.283*	0.119	0.006	0.056	0.098**	0.165	0.187	0.320	0.101	-0.261
	(0.0962)	(0.1769)	(0.2941)	(0.0599)	(0.0262)	(0.5103)	(0.2136)	(0.2148)	(0.6328)	(0.1823)
R-squared	0.527	0.737	0.420	0.386	0.423	0.119	0.656	0.389	0.465	0.323
Observations	13,124	2,433	4,150	11,888	34,906	10,701	20,515	12,866	6,395	2,412

Notes: *(a)* Each panel reports on results of a one specification, run for each state separately across columns. *(b)* All regressions include defendant's race, gender, age, age squared, and an indicator for recidivism. All regressions also include the case's severity, and the number of charges in each case. Finally, all regressions include judge fixed effects and year as well as three quarter-of-year fixed effects. *(c)* Panel A adds cases after the filing date. Panel B replaces the linear regressor from expression (1) with a "count of quarters"; for example, Washington's electoral cycle is 4 years, i.e., 16 quarters, so that the coefficient in Panel B is roughly $1/16$ the baseline coefficient ($0.28 \approx 4.32/16$). *(d)* Standard errors are two-way clustered by calendar-year and quarter-of-year in all panels. *** p<0.01, ** p<0.05, * p<0.1

As discussed in footnote 21, Online Appendix Figure 1 adds confidence bands to Figure 2.

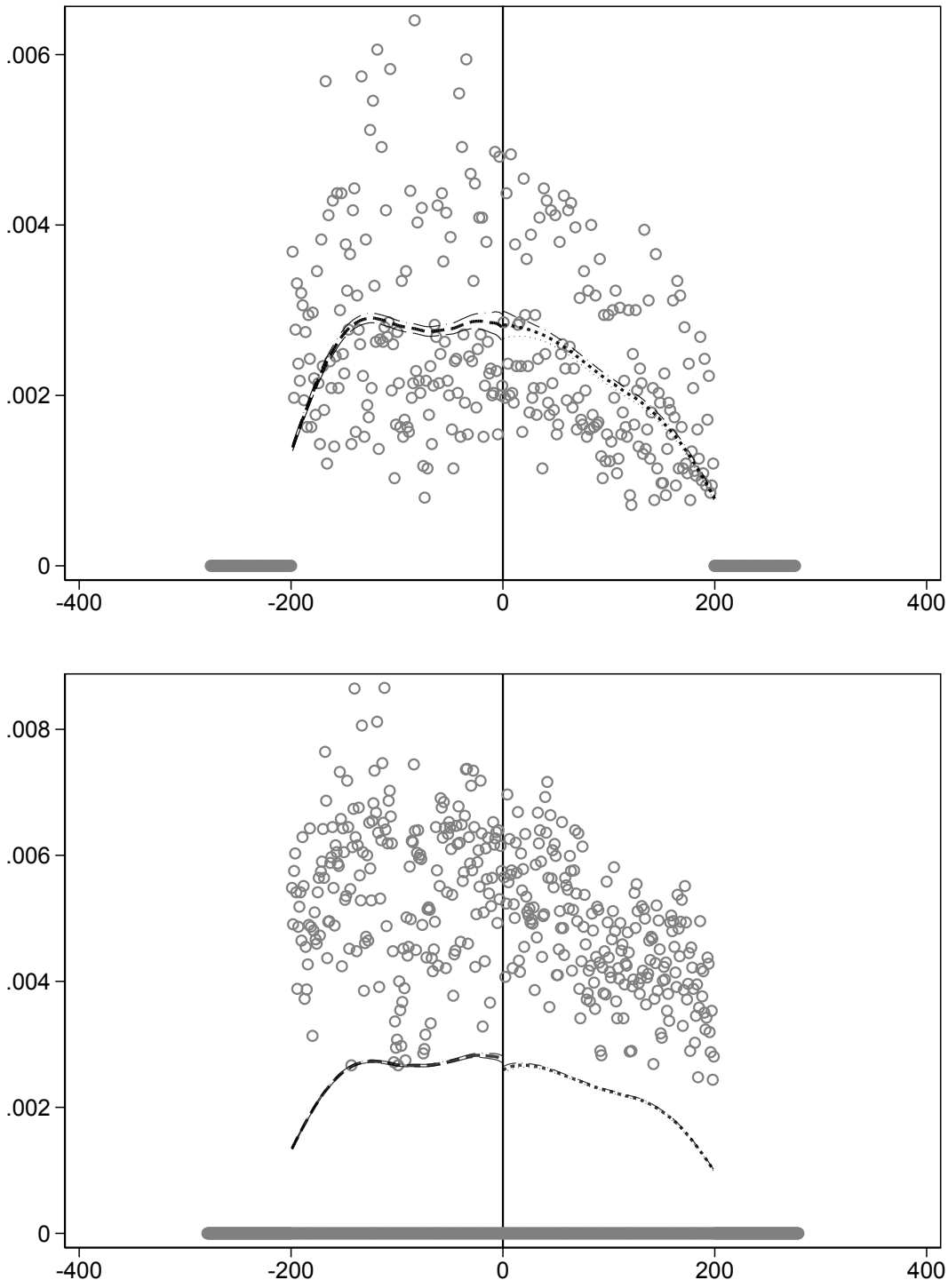
Online Appendix Figure 1: Quarterly Indicators with Confidence Bands



Notes: The 0 vertical line is the date of the general election. The second vertical line left of 0 marks the filing date. (In Georgia, there were 3 different election dates within our data.) This figure reports on the estimated quarterly dummies when cases after the filing date are omitted.

One concern with our omission of cases between the filing date and the general election date is that judges may , instead of levying harsher sentences in the lead-up to the filing date, postpone contentious or visible cases until after the filing date to avoid a challenger running against them on the basis of a contentious ruling. If this was the case we would expect some bunching of cases after the filing date, and we would expect this to be concentrated in the severe cases. Online Appendix Figure 2 presents the results of a McCrary (2008) test to test for this. There is no evidence of bunching either side of the filing date for severe-crime cases (top-panel). The associated test shows a log difference in height of 0.015, with a standard error of 0.042, giving rise to a t-statistic of 0.359, i.e., the hypothesis of no bunching is not rejected. But there is some evidence for bunching of non-severe cases before the filing date (bottom-panel). The associated test shows a log difference in height of -0.061, with a standard error of 0.015, giving rise to a t-statistic of -4.11. If anything, this suggests that judges may try to get smaller cases dealt with before the filing date in case they need to devote some of their time after the filing date to the campaign trail.

Online Appendix Figure 2: McCrary Tests



Notes: (*a*) This figure shows the McCrary Test for bunching of a running variable (McCrary, 2008). In our case, that running variable is days within an election cycle, centered around the filing date. The sample is cases that fall within six month either side of a filing date and inside the same electoral cycle. (*b*) The top-panel displays the test for 25,000 severe-crime cases. The bottom-panel displays the test for 202,000 non-severe cases. (Because the number of observations in the bottom panel is very large, the scatter has to use coarser bins than the smoothing function so that it lies everywhere above the smoothed function.) (*c*) The associated test in the top-panel shows a log difference in height of 0.015, with a standard error of 0.042, giving rise to a t-statistic of 0.359, i.e. the hypothesis of no bunching is not rejected. The associated test in the bottom-panel shows a log difference in height of $-.061$, with a standard error of 0.015, giving rise to a t-statistic of -4.11.

In Table 7, we set the measures of electoral competition to zero in Virginia, where judges are always appointed. To test that the inclusion of Virginia in the pooled sample does not drive our results, we re-run all specification without Virginia in Online Appendix Table 8.

Online Appendix Table 8: Pooled Panel with Interactions - Drop VA

	I	II	III	IV	V	VI	VII	VIII	IX	X
Panel: no VA	Dependent variable: Sentence (months)									
Proximity to election	4.048*	1.340	3.034	-5.074	-3.639**	-4.740*	-2.923**	-3.283**	-6.441	-9.711
	(2.2169)	(1.8064)	(4.8508)	(4.0759)	(1.3879)	(2.6640)	(1.2674)	(1.2882)	(3.8229)	(5.8370)
Proximity to election										
x Partisan Election		7.568*	5.874				6.042	4.687		
		(3.4288)	(3.4034)				(3.3923)	(3.1340)		
x Nonpartisan Election			-1.811							
			(4.0586)							
x Cycle-Length				1.312*					0.589	0.837
				(0.6157)					(0.6594)	(0.5930)
x Prob. electoral challenge					0.279***		0.173***		0.236***	
					(0.0731)		(0.0373)		(0.0711)	
x # donors per judge-race						0.303**		0.196***		0.281**
						(0.1166)		(0.0368)		(0.1055)
R-squared	0.440	0.440	0.440	0.440	0.440	0.260	0.440	0.260	0.440	0.260
Observations	131,028	131,028	131,028	131,028	131,028	101,583	131,028	101,583	131,028	101,583

Notes: *(a)* This table re-runs Table 7, omitting Virginia because judges are always appointed there. *(b)* Standard errors are multi-way-clustered by quarter-year and state level. *** p<0.01, ** p<0.05, * p<0.1

Online Appendix Table 9 shows that the results look similar when we exclude Colorado and Tennessee, where the Bonica (2016) measure is unavailable.

For Online Appendix Table 10, we constructed the principal component of the two measures of revealed competition. Alternatively, we also took the average (for columns 2 and 4). In all specifications, the average of these measures interacts significantly in shaping the strength of electoral cycles.

Online Appendix Table 9: Pooled Panel with Interactions - Drop CO and TN

	I	II	III	IV	V	VI	VII	VIII	IX	X
Panel: no CO and TN	Dependent variable: Sentence (months)									
Proximity to election	3.405**	1.457	-4.946*	-3.294	-1.640	-4.740*	-1.323	-3.283**	-4.331	-9.711
	(1.3784)	(1.0282)	(2.1973)	(4.6413)	(1.7459)	(2.2640)	(1.5925)	(1.2882)	(4.8613)	(5.8370)
Proximity to election										
x Partisan Election		6.448	12.852**				5.276	4.687		
		(3.6063)	(4.5709)				(3.5375)	(3.1340)		
x Nonpartisan Election			6.750**							
			(2.6038)							
x Cycle-Length				1.010					0.557	0.837
				(0.7912)					(0.7056)	(0.5930)
x Prob. electoral challenge					0.199*		0.123		0.159*	
					(0.0968)		(0.0746)		(0.0712)	
x # donors per judge-race						0.303**		0.196***		0.281**
						(0.1166)		(0.0368)		(0.1055)
R-squared	0.261	0.261	0.261	0.261	0.261	0.261	0.261	0.261	0.261	0.261
Observations	104,311	104,311	104,311	104,311	104,311	104,311	104,311	104,311	104,311	104,311

Notes: *(a)* This table re-runs Table 7, omitting Virginia, Colorado and Tennessee because we do not observe Bonica's average number of donors per race in those three states. *(b)* We report p-values in square brackets. Standard errors are multi-way-clustered by quarter-year and state level. *** p<0.01, ** p<0.05, * p<0.1

Online Appendix Table 10: Principle Component and/or Average of Revealed Competition

	I	II	III	IV
	Dependent variable: Sentence (months)			
Sample	All		No CO and TN	
Proximity to election	3.787*	3.922*	3.582**	3.782**
	(1.9911)	(1.9926)	(1.2992)	(1.3742)
Proximity to election				
x PCA	3.076**		2.371*	
	(0.9775)		(1.0166)	
x avg of normalized		4.362***		3.280*
		(1.2367)		(1.4089)
R-squared	0.440	0.440	0.261	0.261
Observations	133,756	133,756	104,311	104,311

Notes: Standard errors are two-way clustered by state-year and state-specific quarter-year, thus essentially just stacking the number of clusters of the individual regressions in Section 3. *** p<0.01, ** p<0.05, * p<0.1