Supporting Information/Online Appendix for *Predicting Drift on Politically Insulated Institutions: A Study of Ideological Drift on the United States Supreme Court*

Data Collection

As we discuss in the main text, we collected and examined three sources of texts to calculate measures of cognitive (in)consistency and cognitive complexity: (1) every published article by a justice prior to her nomination; (2) every separate opinion authored by a justice while presiding as a circuit court judge or state supreme court justice; and (3) every speech delivered by a justice prior to her nomination. Obtaining each justice's published articles was straightforward. We searched Hein Online and Google Scholar for every article written by the justice prior to the date of his or her nomination. Obtaining each justice's pre-nomination opinions (where relevant) was also straightforward. We searched LEXIS for separate opinions written by the nominee when he or she was a circuit court judge, state supreme court justice, or other lower court judge. Obtaining the speeches delivered by each justice was much more difficult, as there are few published pre-nomination speeches available. Accordingly, we had to search through archival data in presidential libraries and the justices' private papers, taking digital images of each page in each speech.

Once we obtained all the articles, opinions, and speeches, we converted them into a text-readable format. Once again, converting articles and opinions was fairly straightforward, as some of them are downloadable in text formats. Where they were not, we had to transcribe each of them individually. And, once again, the justices' speeches were more difficult. We had to convert digital images into text-readable format. To so so, we employed two methods. First, some images were converted by manual transcription. The second method employed ABBY FineReader, optical character recognition software that converts digital images to text format. Once the software converted the speeches to a text-readable format, we examined each speech to check for conversion errors. The distribution of texts for each justice are displayed in Table 1.

Justice	**Separate Opinions**	**Speeches**	**Writings**	**Total**	**Regime Type**
Black	0	38	1	39	Constrained
Blackmun	20	18	2	40	Constrained
Brennan	20	1	0	21	Constrained
Breyer	46	7	13	66	Constrained
Burger	96	4	6	106	Constrained
Burton	0	50	1	51	Unconstrained
Clark	0	39	6	45	Constrained
Douglas	0	25	19	44	Constrained
Frankfurter	0	3	37	40	Constrained
Ginsburg	38	17	30	85	Semi-constrained
Harlan	2	2	1	5	Constrained
Jackson	0	43	20	63	
Kennedy	57	0	1	58	Constrained
Marshall	12	2	5	19	Constrained
Murphy	0	47	1	48	Unconstrained
OConnor	39*	0	1	40	Semi-constrained
Powell	0	5	24	29	Constrained
Reed	0	24	1	25	Constrained
Rehnquist	0	2	5	7	Constrained
Scalia	28	6	14	48	Constrained
Souter	22	0	0	22	Constrained
Stevens	50	2	3	55	Constrained
Stewart	13	2	0	15	Constrained
Thomas	2	2	3	7	Constrained
Warren	0	54	11	65	Constrained
Total	445	393	205	1,043	

Table 1: Distribution of texts and regime type for Justices during their pre-Court era. Note, all opinions are lower court separate opinions (dissents and concurrences) with the exception of O'Connor, which includes majority opinions. For the regimes in which presidents made nominations to the Court, the table reflects the filibuster as the veto pivot in the senate. Entry for Justice Jackson omitted, as the data do not clearly indicate in which regime his nomination took place. Our results remain stable regardless of which regime he is coded under.

Auxiliary Analyses Discussed in Text

In addition to the analyses in the paper, for a robustness check we re-estimated the models by first removing the *Cognitive Complexity* measure and then, estimating the models with the *Cognitive Inconsistency* measure removed. These results are shown in Table 2 and Table 3 and they are consistent with the results reported in the paper. *Cognitive Inconsistency* remains significant when *Cognitive Complexity* is removed. And, *Cognitive Complexity* does not reach statistical significance in the models where *Cognitive Inconsistency* is removed.

Table 2: Ordinary Least Squares Estimates of Ideological Drift: Only Inconsistency

	(1) St. Dev.	(2) Total Dist.	(3) Avg. Drift	(4) Range
Cognitive Inconsistency	0.135*	0.498**	0.016*	0.481**
	(0.074)	(0.185)	(0.007)	(0.181)
Semi-Constrained Regime	0.042	0.314	0.002	0.151
	(0.295)	(0.737)	(0.028)	(0.723)
Unconstrained Regime	-0.864**	-2.559**	-0.093**	-2.744**
	(0.289)	(0.722)	(0.027)	(0.709)
Court Composition	0.177*	0.619**	0.019*	0.496*
	(0.088)	(0.219)	(0.008)	(0.215)
Public Opinion	-1.504*	-4.235*	-0.159*	-4.025*
	(0.702)	(1.756)	(0.067)	(1.723)
Number of Utterances	-0.009*	-0.030**	-0.001*	-0.024*
	(0.005)	(0.011)	(0.000)	(0.011)
Judicial Experience	-0.007*	-0.030**	-0.001**	-0.020*
	(0.004)	(0.009)	(0.000)	(0.009)
Constant	1.276**	3.404**	0.130**	3.427**
	(0.416)	(1.040)	(0.039)	(1.021)
N	25	25	25	25
adj. R^2	0.363	0.564	0.528	0.501

Standard errors in parentheses

* $p < .05$, ** $p < .01$, one tailed

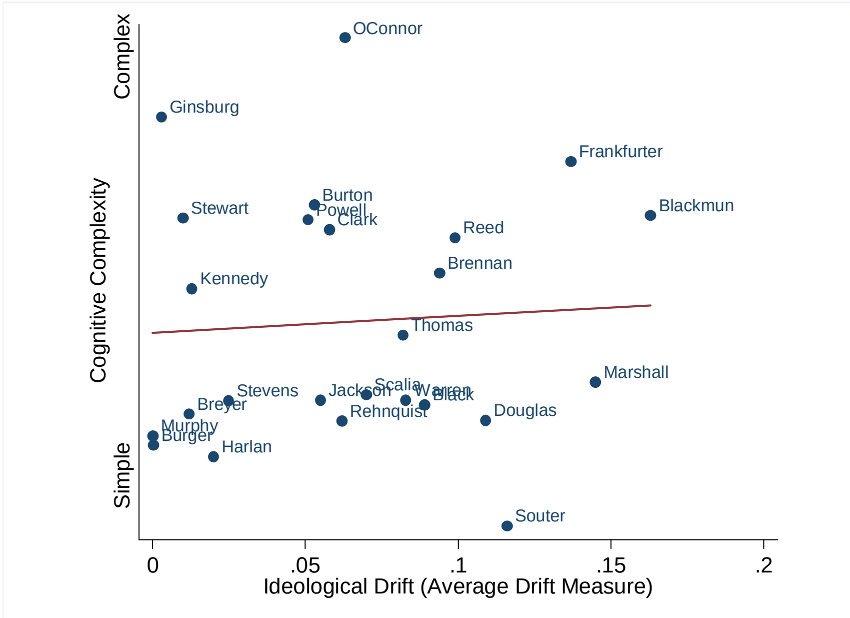
Table 3: Ordinary Least Squares Estimates of Ideological Drift: Only Complexity

	(1) St. Dev.	(2) Total Dist.	(3) Avg. Drift	(4) Range
Cognitive Complexity	-0.084	0.043	0.005	-0.190
	(0.062)	(0.178)	(0.006)	(0.168)
Semi-Constrained Regime	0.282	-0.305	-0.033	0.540
	(0.399)	(1.145)	(0.041)	(1.080)
Unconstrained Regime	-0.796**	-2.376**	-0.088**	-2.521**
	(0.299)	(0.858)	(0.031)	(0.810)
Court Composition	0.224*	0.666*	0.019*	0.624*
	(0.093)	(0.268)	(0.010)	(0.253)
Public Opinion	-1.732*	-5.485**	-0.203**	-4.962**
	(0.709)	(2.036)	(0.072)	(1.922)
Number of Utterances	-0.005	-0.015	-0.000	-0.011
	(0.004)	(0.012)	(0.000)	(0.011)
Judicial Experience	-0.006	-0.023*	-0.001*	-0.014
	(0.004)	(0.011)	(0.000)	(0.010)
Constant	1.825**	5.110**	0.181**	5.282**
	(0.343)	(0.985)	(0.035)	(0.929)
N	25	25	25	25
adj. R^2	0.311	0.379	0.410	0.343

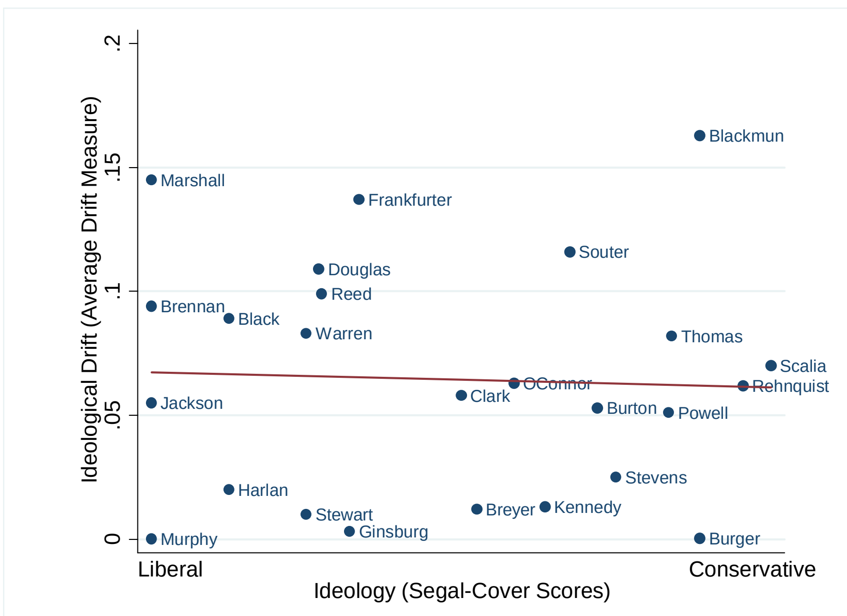
Standard errors in parentheses

* $p < .05$, ** $p < .01$, one tailed

In addition, as a further check on the robustness of this finding, we examined the correlation between *Cognitive Complexity* and ideological drift and it was not significant. To illustrate this, the top portion of Figure 1 contains a scatterplot of the justices based on ideological drift (the Average drift measure) and Cognitive Complexity, along with a "best-fit" line. As the figure shows, there is no relationship between the two. The bottom portion of Figure 1 represents the correlation between pre-nomination ideology and ideological drift, and it also shows no relationship. In short, this suggests that liberal nominees, once confirmed, are not more likely to drift than conservative nominees.



(1)



(2)

Figure 1: The top graph represents the correlation between cognitive complexity and drift on the court (using the average drift measure). The line represent the best linear fit of the data, and the correlation is not significant. Note, the correlations between cognitive complexity and the other two measures of drift are not statistically significant. The bottom graph represents the correlation between pre-nomination ideology and drift on the court (average drift measure). The best-fit line is not statistically significant. Note, the correlations between pre-nomination ideology and the other two measures of drift are not statistically significant.

Alternative Model Specification

We re-estimated our model excluding *Number of Utterances* and the number of years on the Court, which was a key part of our *Court Composition* measure. The p-values for our main independent variable (*Cognitive Inconsistency*) using each of our four dependent variables are: 0.04, 0.045, 0.105, and 0.03 (one-tailed).

Table 4: Ordinary Least Squares Estimates of Ideological Drift

	(1) Standard Deviation Drift	(2) Total Distance Drifted	(3) Average Drift	(4) Drift Range
Cognitive Inconsistency	0.013*	0.356*	0.093+	0.364*
	(0.007)	(0.198)	(0.071)	(0.182)
Cognitive Complexity	0.008+	0.161	-0.045	-0.082
	(0.006)	(0.181)	(0.065)	(0.167)
Semi-Constrained Regime	-0.059+	-1.296	-0.070	-0.330
	(0.037)	(1.100)	(0.395)	(1.014)
Unconstrained Regime	-0.095**	-2.663**	-0.865**	-2.758**
	(0.029)	(0.877)	(0.315)	(0.808)
Public Opinion	-0.154*	-4.038*	-1.292+	-3.500*
	(0.071)	(2.103)	(0.755)	(1.939)
Judicial Experience	-0.001**	-0.028**	-0.007*	-0.020*
	(0.000)	(0.010)	(0.004)	(0.009)
Constant	0.072*	1.549+	0.725*	1.893*
	(0.030)	(0.909)	(0.326)	(0.838)
N	25	25	25	25
adj. R^2	0.420	0.314	0.193	0.308

Standard errors in parentheses

+ $p < .1$, * $p < .05$, ** $p < .01$ one tailed

Political Regimes

We used the Poole and Rosenthal first dimension Common Space scores (Poole and Rosenthal 1997) to determine the revealed preferences of senators and presidents. The senator who constituted the filibuster pivot depended on the party of the president (Krehbiel 1998). For Democrat presidents after 1974, the filibuster pivot was the 60th most liberal senator. For Republican presidents during that time period, the filibuster pivot was the 40th most liberal senator. In 1975, the senate lowered the number of votes needed to invoke cloture from $\frac{2}{3}$ to $\frac{3}{5}$. Thus, for Democrat presidents prior to 1975, the filibuster pivot was the 67th senator while, for Republican presidents, it was the 34th senator.

To determine the location of the Court's policy, given the vacancy, we employed Judicial Common Space (JCS) scores (Epstein et al. 2007). Importantly, these scores are scaled into Poole and Rosenthal Common Space, which means that the justices' revealed preferences can be compared directly to the preferences of the president and senate. Further, the JCS score for each justice varies per term (because the JCS scores are, themselves, derived from the dynamic Martin-Quinn scores). As such, not only could we measure the justice's preferences on the same scale as the president and senate, we also gained a measure of the policy the Court would make in the absence of the nominee (and, by logic, the justice who retired). Thus, unlike Johnson and Roberts (2005), who were forced by contemporary measurement capabilities to rely on Segal-Cover scores—which are fixed pre-nomination estimations of nominee ideology (Segal and Cover 1989)—we were able to employ dynamic scores and, thereby, capture the Court's policy behavior in "real time."[1] Once we located the president, filibuster pivot, and Court in policy space, we were able to determine which among the three regimes the nomination took place.[2] We then created binary variables for all three categories

[1]This explains why our list of nominees per regime (found in the online supplement) differs somewhat from the Johnson and Roberts list (2005, Table 1a, 36). The Appendix provides a breakdown of the regimes in which the justices' nominations transpired.

[2]While the common space did not include scores for Presidents Truman and Roosevelt, we were still able

(unconstrained, semi-constrained, and fully constrained) that equalled 1 if the nomination took place in the respective regime, and 0 otherwise. (We used the fully constrained variable as our baseline category because it is the most common).

The Figure below presents the three regimes discussed in the manuscript. It reflects a one-dimensional policy space moving from most liberal (on the left) to most conservative (on the right). Each vertical tick represents the ideal point of the particular actor, the point in policy space that the actor prefers to all others. The model assumes that actors have continuous, single-peaked, symmetric preferences, which means that they equally dislike, say, a one-unit shift to the left of their ideal point the same as a one-unit shift to the right of their ideal point. We note that our results below are robust to either the median or filibuster actor being defined as the pivotal position. Following the logic of Johnson and Roberts (2005) and Krehbiel (1998), we model the pivotal actor in the Senate as the filibuster pivot rather than the Senate median.

Consider, first, the unconstrained regime, which is shown in Figure 2(a). The president can nominate someone to the Court at his ideal point without fear of obstruction from the Senate filibuster. The logic is simple. The Court, operating with a vacancy (i.e., $\frac{J_4+J_5}{2}$) is more liberal than the president who, in turn, is more liberal than the Senate filibuster[3] The policy that the Court would make with the vacancy is too liberal for both the president and the senate. Since the president is located between the Court and the senate, he is free to nominate someone at his ideal point. The result would be a gain for both the president and the senate. Because both actors are made better off than J, the nomination would

to determine the regimes in which they made their nominations. We used President Truman's common space scores from his time as a senator. And, we made the reasonable assumption that President Roosevelt's preferences were to the left of Truman's. (Our belief in the reasonableness of this strategy is buttressed not only by history, but also by the fact that FDR's DW-NOMINATE score is more liberal than Truman's.) The only justice in whose nomination regime we were not confident was Jackson, who could have been nominated in either a constrained or unconstrained regime. We are happy to report, nevertheless, that our results are robust to coding his nomination as either a constrained or unconstrained regime.

[3] The logic of all these figures is the same if we instead employ their mirrors, such that, in Figure 2(a), $S < P < J$.

succeed.

Consider, next, the semi-constrained regime, which is shown in Figure 2(b). In this regime, both the president and senate can arrive at a nominee who improves Court policy, but the president is limited in how far he can shift Court policy. While both the president and senate are more conservative than the Court, the senate now falls between the Court and the president. As such, the president will face obstruction if he nominates someone who moves policy "too far" to the right. More specifically, the nomination will fail if the nominee shifts Court policy to the right of the senate filibuster's indifference point vis-à-vis the status quo (i.e., existing Court policy). The filibuster's indifference point is the point in policy space that the filibuster pivot prefers to existing Court policy.[4] Any nominee who will shift Court policy between J and I_S will be acceptable to the senate filibuster since the policy would be an improvement over J. On the other hand, any nominee who will shift policy to a space outside $\overline{S, I_S}$ will be unacceptable and, thus, filibustered. Accordingly, a strategic president will nominate someone at the filibuster's indifference point. This is the point in policy space that makes the president better off and avoids a filibuster.

Finally, consider the fully constrained regime, which is shown in Figure 2(c). In this case, the president can only nominate someone who will generate the same types of policies as the current Court. Figure 2(c) shows that the senate is more liberal than the Court which, in turn, is more liberal than the president. If the president nominated anyone who would shift policy to the right of the existing Court, the senate filibuster pivot would be worse off and, therefore, would block the nomination. The only way the senate would be made better off is if the president nominated someone to the left of J which he, of course, has no incentive to do. Accordingly, gridlock prevails. The president must nominate someone who will simply continue the Court's current policies. The justices and their corresponding regimes can be seen in the Table 1.

[4]In Figure 2(b), $|S - J| = |S - I_S|$.

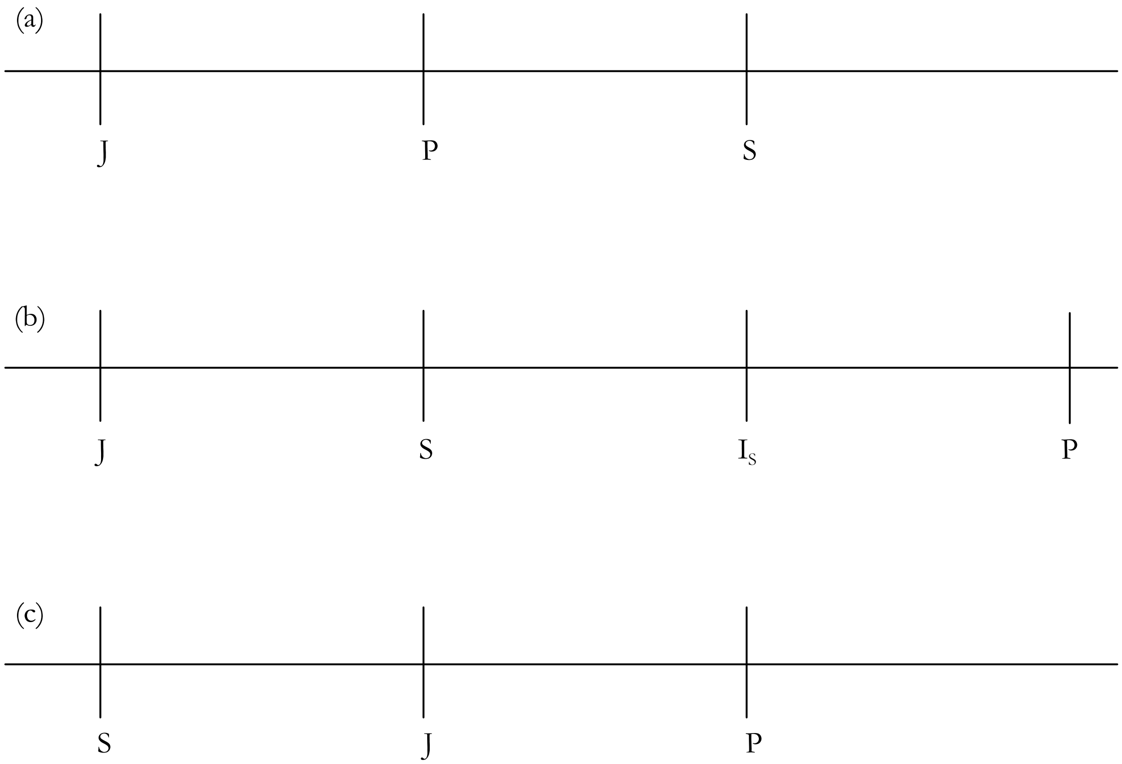


Figure 2: J=Policy the Court makes with the vacancy ($\frac{J_4+J_5}{2}$). P=President's ideal point. S=Senate Filibuster's ideal point. I_S=Senate Filibuster's indifference point vis-à-vis J. Figure (a) represents an unconstrained regime. Figure (b) represents a semi-constrained regime. Figure (c) represents a fully constrained regime.

LIWC Classifications

In the main text, we identified a host of measures employed by LIWC to estimate cognitive complexity. We employed ten such measures in this paper. Here, we explain each of these measures.

LIWC measures the *causation* dimension by searching for words like "because," "effect," and "hence" that refer to causal processes. This dimension taps into the degree that an individual sees relationships between different parts or components, and how changes in one may influence changes in another (e.g., thinking in terms of cause and effect). Increased use of causation words corresponds with more cognitive complexity. LIWC measures the *insight* dimension by searching for words such as “think,” “know,” and “consider.” This dimension captures the degree that individuals differ in how much one is able to discern a more in-depth understanding of a subject or its underlying nature. The *discrepancy* dimension examines words like “should,” “would,” and “could,” as it measures the degree to which an individual identifies discrepancies, differences, or inconsistencies between, for example, two situations or cases (e.g., the fact patterns of two searches and seizures). Higher scores along the discrepancy and insight dimensions correspond with increased levels of cognitive complexity.

The *inhibit* dimension searches for words like “block,” “stop,” and “constrain” as it measures the level of inhibition displayed by the decision maker. Inhibition is theorized to be how much restraint one expresses or to what degree a person displays how their actions are hindered. Increased amounts of inhibition in speech is associated with higher levels of cognitive complexity. The *tentative* dimension counts words like “maybe,” ”fairly” and “perhaps,” and measures the level of tentativeness each text or decision maker shows. Tentativeness is theorized to be how hesitant or unsure one is about something. Increased amounts of tentativeness in speech is associated with higher levels of cognitive complexity. The *certainty* dimension counts words like “always,” ”absolutely,” and ”clearly” and it is theorized to measure the degree of how confident one is about something. Generally, higher levels of certainty correspond with expressing or portraying issues less complex. Increased amounts of certainty in speech is associated with lower levels of cognitive complexity.

The *inclusiveness* dimension searches for words like “with” and “and.” It captures the degree to which one sees many connections or relations between ideas and concepts. Increased amounts of inclusiveness in speech is associated with higher levels of cognitive complexity. The *exclusiveness* dimension looks for words such as “but” and “except” and it is theorized to capture how distinct or separate one sees concepts and ideas. People use exclusion words to help make distinctions, especially when determining whether something belongs in a category or not. Increased amounts of exclusiveness in speech is associated with lower levels of cognitive complexity. The *negations* dimension examines words like “no” and “never,” and is theorized to measure how much an individual acknowledges the absence or opposite of something that is positive or affirmative. Increased amounts of negation in speech is associated with lower levels of cognitive complexity. The *six-letter* dimension seeks out the number of words in the text containing six or more letters. This is a commonly used linguistic measure of a person’s sophistication. Increased amounts of six-letter words in speech are associated with higher levels of cognitive complexity.

Note that we standardized all indicators by subtracting off the mean and dividing by the standard deviation indicated by a Z in the following formula: Cognitive complexity= $Zsixletter - Zcausation - Zinsight - Zdiscrepancy - Zinhibit - Ztentative - Zcertainty - Zinclusive - Zexclusive - Znegations$. This formula generates a cognitive complexity score for each document, and to generate our measure of cognitive inconsistency we simply take the standard deviation for each justices set of documents. This formula was based on a replication of Gruenfeld’s (1995) published measures of the justices, which used human coders.

To assess whether all ten categories represent one underlying concept, we subjected them to an exploratory factor analysis and it returned a one factor solution. The results of the exploratory factor analysis provides us with confidence that all ten indicators are part of the same underlying dimension that we theorize to be cognitive complexity. The alpha scale reliability coefficient for these 10 items is .68.

The Martin-Quinn Scores

One potential concern with our approach is that the Martin-Quinn scores may not represent preference change as much as they highlight decisional changes by justices. That is, the changing behavior we observe by justices might simply be a shift in how justices vote rather than a change in their underlying preferences. While it is possible that the drift is not preference change, but something else, we ultimately do not believe it changes the underlying story we seek to tell. Some justices have certain characteristics that lead them to change their voting behavior over time, and others do not. That "something," we argue, is cognitive inconsistency.

On a related note, one might argue that drift reflected in the Martin-Quinn scores is a function of the Court's changing docket and the way justices think about certain *issues*. For example, one study asks us to imagine an instance where Justice Kennedy, for example, rules for the government consistently in criminal procedure cases but against the government in free speech cases while Chief Justice Rehnquist votes for the government in both types of cases (Farnsworth 2007). If, in one term, the Court heard all criminal procedure cases, Kennedy would look quite conservative. If, in the next term, the Court heard a significant number of free speech cases, Kennedy would look less conservative. Did Kennedy drift? Perhaps not, argues the author.

We cannot accede to this argument, however. For starters, most justices have fairly consistent views across issue areas. That is, if they are conservative in, say, privacy cases, they are also generally conservative in criminal procedure cases. Second, justices themselves control the Court's docket (see, e.g., Black and Owens 2009). So, if docket shifts occur, they are likely to be a function of justices' changing preferences. Third, the types of docket shifts that would need to occur to support this example are too far-fetched to cast doubt on the methodology. Most drift appears uni-dimensional. What this means is that for a justice to look conservative in the beginning of her term and liberal later, her early tenure would have to have been marked by numerous cases in one issue area. Then, as time goes on, the Court would have to hear significantly fewer of those cases. Even a bare look at the descriptive data suggest that this did not occur. To be sure, the Court backed off of economics cases since the 1930s, but the docket-based shifts that would need to occur to support this argument are quite unlikely to have occurred. Simply put, while we recognize that Martin-Quinn scores are not perfect measure of justices' preferences (their creators never intended them to be), they do tap into revealed preferences and show that justices have changed over time.

Black, Ryan C. and Ryan J. Owens. 2009. "Agenda-Setting in the Supreme Court: The Collision of Policy and Jurisprudence." *Journal of Politics* 71(3):1062–1075.

Epstein, Lee, Andrew D. Martin, Jeffrey A. Segal and Chad Westerland. 2007. "The Judicial Common Space." *Journal of Law, Economics, & Organization* 23(2):303–325.

Farnsworth, Ward. 2007. "The Use and Limits of Martin-Quinn Scores to Assess Supreme Court Justices, with Special Attention to the Problem of Ideological Drift." *Northwestern University Law Review* 101(4):1891–1904.

Johnson, Timothy R. and Jason Roberts. 2005. "Pivotal Politics, Presidential Capital, and Supreme Court Nominations." *Congress and the Presidency* 32(1):31–48.

Krehbiel, Keith. 1998. *Pivotal Politics*. University of Chicago Press.

Poole, Keith and Howard Rosenthal. 1997. *Congress: A Political-Economic History of Roll Call Voting*. Oxford University Press.

Segal, Jeffrey A. and Albert D. Cover. 1989. "Ideological Values and the Votes of Supreme Court Justices." *American Political Science Review* 83(2):557–565.