The Force of the Law*

JBrandon Duck-Mayr

September 12, 2021

Abstract

(How much) does "the law" affect judges decisions? Though among the most fundamental of questions in judicial politics, the effect of the law is difficult to identify. A number of past studies have taken clever approaches to tackling this difficult question; however, they offer mixed findings at best, and are sometimes critiqued on methdological grounds. I provide new evidence by taking a new perspective on the law: I conceptualize the law as the mapping from cases to outcomes implied by all past court decisions. I use that perspective to develop a strategy that exploits attributes of Gaussian process classification to isolate the effect of the law on judges' votes from the effect of judges' own preferences. I apply this strategy to data on First Amendment Free Exercise cases at the U.S. Supreme Court. I find the justices exhibit varying levels of deference to the implications of past decisions, with some justices showing substantial effects and others essentially unaffected. In addition to providing the best available evidence of the constraining effect of law, I provide a measure of the legal status quo, an important theoretical concept in judicial politics, and highlight past studies' vulnerability to misspecification bias, a problem my modeling strategy also solves.

Word Count: 10,997

*I would like to thank Jim Spriggs and Jacob Montgomery for their helpful comments.

1 Introduction

A fundamental question at the heart of judicial politics is whether, how, and to what extent "law" impacts judges' decisions. Judges often profess to simply apply existing law to the cases before them, untainted by their own preferences, exemplified by Chief Justice Roberts' (in)famous statement, "Judges are like umpires. Umpires don't make the rules, they apply them… I will decide every case based on the record, according to the rule of law… and I will remember that it's my job to call balls and strikes, and not to pitch or bat." Nevertheless, judges are typically portrayed as conservative or as liberal just as other political actors are, and with good reason: A long line of political science research finds ideology to be a strong predictor of US Supreme Court justices' decisions (e.g. Epstein and Knight 1998; Segal and Spaeth 2002) and media coverage may highlight political or negative aspects of the Court's decisions (Denison, Wedeking, and Zilis 2020; Johnson and Socker 2012).

Although the public need not view judges as "legal automatons" to view the courts as legitimate (Gibson and Caldeira 2011), to the extent judges are viewed as mere "politicians in robes", the courts' legitimacy may suffer. More fundamentally, understanding the way legal constraint interacts with political ideology in judicial decision making is crucial for generating useful theories of judicial behavior and accurate inferences in quantitative studies of judicial politics. Given the stakes for understanding the extent to which the law constrains judicial decision making, it is no surprise the literature offers a variety of clever ways to test for the law's influence, such as examining the behavior of dissenting justices in future related cases (Segal and Spaeth 1996), studying the use of precedent (Hansford and Spriggs 2006; Hinkle 2015), considering legal forces particular to agenda-setting decisions (Black and Owens 2009), interacting legal and ideological variables to find conditional effects (Bartels 2009, 2011), and isolating ideological effects by considering positions taken by political actors outside the Court (Bailey and Maltzman 2008).

While the political science literature is unambiguous that ideology plays a large role at the Court, the evidence on what constraining effect, if any, the law exerts on the justices' decisions is mixed. Segal and Spaeth (2002) posited Supreme Court justices as more or less unconstrained by law, deciding cases as they like, but other scholars advocate a more nuanced view, with effects of ideology as well as constraint of ideological behavior by legal factors (e.g. Bartels 2009). Even some work that concludes that "law matters" finds that legal *constraint* is not among the chief legal effects; for example, Hansford and Spriggs (2006) conclude that law (that is, precedent) can act as both an opportunity (to interpret existing precedent positively or negatively in line with their own ideological preferences) in addition to a constraint—but the weight of their presented evidence is on the side of law as opportunity rather than constraint.

Perhaps because constraining effects of law are difficult to uncover in Supreme Court justices' ultimate decisions on the merits, work on this topic often focuses on slightly different effects or settings. For example, Bailey and Maltzman (2008) estimate justices' willingness to explicitly overrule precedent, controlling for justices' preferences by using positions taken on Supreme Court cases by members of Congress and the president.[1] Black and Owens (2009) show factors such as circuit conflicts predict agenda-setting decisions at the Supreme Court, controlling for whether the justices would ideologically prefer to grant or deny cert given the status quo from the lower court decision. Hinkle (2015) finds Circuit Court judges are more likely to cite precedent that is binding in their jurisdiction, utilizing random panel assignment at the U.S. Courts of Appeals to identify the effect.

Moreover, even when empirical studies uncover constraining effects of law, methodological critiques cloud the findings. For example, Richards and Kritzer (2002) find evidence of "jurisprudential regimes": When a (typically landmark) decision dictates an important change to how certain facts are treated under the law, their work suggests justices seem to vote differently afterward, suggesting these precedents exert some binding force on the justices' choices. However, Lax and Rader (2010a) show that work in this area failed to account for crucial forms of error correlation, such as by term, resulting in overconfident statistical tests.

[1]Bailey and Maltzman (2008) also look at two other legal effects: Whether the justices exhibit "judicial restraint"—i.e., defers to Congress—and whether they adhere to a strict interpretation of the First Amendment.

Some work, such as Bartels (2009), has used mutli-level modeling to account for this issue, though we may still worry about "parametric assumptions of error distributions" since we "must assume what sort of clustering can exist [and] assume that other forms... do not exist" (Lax and Rader 2010b, 289). A particularly difficult problem for inference in this area is a type of time-varying confounding: as a justice serves on the Court, their preferences influence the law, so finding an impact of "law" on the justice's decisions may just be picking up on agreement with their own past selves. Some work, such as Segal and Spaeth (1996), focuses on the actions of dissenters in future related cases for this reason, finding little constraining effect of the law.

I provide a novel examination of this issue by using Gaussian process (GP) classification, a machine learning technique, to measure the law and its effect on justices' decisions. I conceptualize the law as the implication in a particular case from all the cases that came before it; given the outcome the Court has assigned to similar cases in the past, what should we think about this case if we did not let anything else influence our thinking? I operationalize this as the predicted outcome in the present case given its facts from a GP classification model trained only on the cases that came before it. This approach aligns more closely than past empirical approaches with the process typically identified as legal reasoning: determining how the present case, as a whole, relates to the Court's past decisions (Levi 1949). It provides a single variable whose effect in justice-level models will give the impact of the law on their decisions. This contrasts with approaches in traditional case fact studies, such as Richards and Kritzer (2002) and Bartels (2009), who look at the effect of a particular (set of) legally relevant case fact(s), where I instead look at the effect on justices' individual decisions of the implied legal outcome given all the legally relevant case facts. Importantly, I show how to control for justices' own contribution to the current state of the law to properly identify the impact of the current state of precedent on justices' individual decisions, independent of their own preferences, addressing the time-varying confounding issue. I find ample effects of ideological and policy preferences, but also that law provides a significant and substantial effect on Supreme Court decision making, for some justices more than others.

This study provides several main contributions. First, I provide a fresh theoretical perspective on legal constraint by conceptualizing the law as the implied outcome in each case given the cases that came before it. Second, I overcome methodological issues with past studies to provide the best available evidence that the law exerts a constraining force on the actions of (some) justices. Third, I provide a new measure of the law, or of the legal status quo. This new measure can help reinvigorate studies of legal constraint and modeling law at the Court, as the volume of work on this important issue has diminished in part because some think the justices are largely unconstrained, but also due to difficult methodological issues (see, e.g., Klein 2017). A measure of the legal status quo is also an important quantity in a number of contexts judicial politics researchers face (see, e.g., Black and Owens 2009) just as a status quo is important in many areas of political science more broadly (e.g. Krehbiel 1998). I also provide suggestive evidence differentiating between reasons *why* the law matters, indicating some justices have a preference for following the law rather than seeing it as a constraint due to (for example) legitimacy needs, which few studies have attempted. Finally, I highlight a widespread mismatch between theory and methods in judicial politics and show how to apply a method better suited to studying decision making in the setting of adjudication. By restricting attention to linear models, past work was susceptible to misspecification bias (see Kastellec 2010), while the modeling strategy taken here avoids that issue while more closely matching the concept of legal reasoning. In so doing, I offer the most compelling evidence to date of how law constraints Supreme Court Justices' decisions.

2 Measuring the Law

Part of the difficulty in studying this issue is one of measurement; how do you measure the law in a way to facilitate analyzing its constraint on justices' decisions? A number of clever approaches to measuring the law have been utilized. For example, the law can be conceptualized as a collection of precedents; Hansford and Spriggs (2006) use this idea, and analyze citations to those precedents to assess legal change. Bailey and Maltzman (2008) measure justices' willingness to overrule past precedent as a measure of legal constraint. Bartels (2009) interacts Martin-Quinn scores (Martin and Quinn 2002)[2] with legally relevant case facts to determine if facts constrain the effect of ideology, and in particular to determine whether some types of cases provide more constraint on ideology than others.

However, lack of a measure of what outcome the law implies in each case has led some to feel "we have reached a point of rapidly diminishing returns in our study of [the] issue" of legal constraint, though with perhaps some hope that advances in text analysis methods may provide better measurement tools (Klein 2017). I provide a different approach: Joining Justice Oliver Wendell Holmes in claiming, "the prophecies of what the courts will do . . . are what I mean by the law" (Holmes 1897), I propose estimating a *predictive* mapping between case characteristics and legal outcomes as a measure of "the law." That is, for each case, we place ourselves in the shoes of a decision maker at the time the case was decided, and say "the law" in that case is what our best predicted outcome in the case would be if we examined *only* the past decisions of the Court at the time.

This conception matches a number of theoretical perspectives on the law, judicially-created policy, and judicial decision making. Consider the attitudinal model, which posits that judges make decisions based on the characteristics of the cases presented to them and their own sincere attitudes and values. Holmes' legal realist approach comports well with this approach. "Case space" models (Lax 2011) can be considered a formalization of the attitudinal model; judges' policy preferences in these models are a partition of a case space, where each dimension corresponds to a case fact, into outcomes. Lax (2007) shows how to represent judicial policy-making using such models. My predictive approach can be viewed as a Bayesian update after each case as to the mapping from the case space to outcomes, or an online update about the justices' attitudes.

[2]Martin-Quinn scores provide ideological estimates for Supreme Court justices somewhat analogous to (e.g.) NOMINATE scores (Poole and Rosenthal 1985) for members of Congress.

This approach can also comport with a legal model of judicial decision making; Levi (1949) explains that in legal analysis, factually similar cases should receive similar outcomes. So this is exactly how law students learn and how lawyers typically argue cases, and how judges would make decisions under the legal model: We take all past cases of the Court, compare the present case to those cases, and assign it the outcome implied by the cases most similar to the present case. Callander and Clark (2017) show how to extend case space models so that the case space maps onto a continuous latent legal outcome rather than onto discrete dispositions; under an assumption that the high court has perfect knowledge of the "correct" mapping from the case space to legal outcomes, lower courts update their knowledge of that mapping after every observed decision of the high court by comparing similarity of the cases they are presented with the past observed decisions of the high court.

Thus, I use the following approach to measure the law: Let $X \subset \mathbb{R}^n$ be the n case factors relevant in the area of law at issue and $\mathbf{X}$ be a finite number of observed cases from the space X; for each case i, we estimate the mapping between the characteristics $\overline{\mathbf{X}}_{i-1}$ of all cases observed *prior* to i and outcomes $\overline{\mathbf{y}}_{i-1}$ in those cases, so that "the law" is the outcome predicted from that model given the characteristics of the instant case, $\mathbf{x}_i$.

There is a long history in the study of judicial politics of estimating the relationship between case characteristics and Court decisions. For example, in the context of Fourth Amendment challenges to police searches and seizures, Segal (1984) shows the relationship between case facts such as the existence of probable cause and the Court's decisions. However, scholars have stopped short of using such models to develop measures of the law. In addition to this innovation, I also use GP classification for estimating this mapping rather than a generalized linear model, as in, e.g., Segal (1984). I use this model because it flows naturally from my conception of law here: The GP classifier is a nonparametric model that compares the test case to all training cases on the basis of all combinations of predictors. That is, a judge hearing a case is not so concerned with any particular fact in isolation. Instead, they care about how the facts as a collective bundle affect the legal outcome; the effect of one fact almost always depends on the presence or absence of other facts. Therefore a linear model unreasonably constrains the class of legal rules we can recover from observed cases (see Kastellec 2010), resulting in misspecification bias. GP classification allows us to uncover any smooth function mapping case characteristics to legal outcomes, as explained in Section 4.2, and provides a convenient way to separate out the effects of this predictor and judges' own preferences as explained in Section 4.3.

3 Modeling Judges' Decisions

I start from the assumptions of an attitudinal model or the case space model (Kornhauser 1992)[3]. For any issue area, or type of case (e.g. cases about police searches and seizures), a judge has a number of factors relevant to their preferred outcome. For example, in a police search case, a judge's preferred outcome (i.e., whether they prefer to rule that the search was constitutional or unconstitutional) may be influenced by whether the search was of a home or not, or if the search was incident to arrest; these case factors may interact in potentially complex ways to determine the judge's sincerely preferred outcome. However, we also must take into account that judges may not follow their own preferred dispositions precisely; they may feel constrained by "the law". For example, suppose a judge would always prefer to rule that searches of a home require a warrant to be lawful. However, they have also observed the Supreme Court's ruling in *Maryland v. Buie* (1990)[4] that officers may sweep the area during an in-home arrest to uncover hidden persons who could pose a danger to those at the scene. This judge may—or may not—then feel constrained to rule differently from their preferred disposition in a closely related case.

To formalize this idea, let ϕ_j represent judge j's sincere policy preferences; in other words, ϕ_j is a function that maps X to outcomes. Now, Supreme Court justices are likely less concerned with the outcome in any particular case they hear than broader issues of how

[3]See Lax (2011) for an overview of case space models in political science.

[4]494 U.S. 325.

the law is interpreted, so it may seem awkward to define judicial preferences in this way. However, think about the implications of a case outcome in light of the discussion in the previous section; every time the Supreme Court puts out a new decision, it clarifies their interpretation of the mapping between X and y. That is, the justices' utility from deciding a case one way or another may come in part from preferring a particular party wins the case, but it also provides a vehicle for updating observers on judicial policy. Let λ represent the Court's mapping from X to y, or the association between cases in X and the outcomes that would result from majority voting over outcomes among the justices on the Court. λ_i will indicate the outcome implied in case i from the prior cases $\overline{D}_{i-1} = \left(\overline{\mathbf{X}}_{i-1}, \overline{\mathbf{y}}_{i-1}\right)$.

For example, let's revisit the *Buie* case mentioned above. An armed robber was described as wearing a red track suit. The police obtained a warrant for Buie, and executed it at his home. An officer coaxed Buie out of the basement and arrested him. Another officer then "entered the basement 'in case there was someone else' down there", at which point he found a red track suit in the basement. Buie argued the red track suit should be excluded as evidence. There are a number of prior decisions of the Court in $\overline{D}_{i-1}$ that influence our estimate of λ_i, or the legal outcome. In *Terry v. Ohio* (1968)[5], the Court ruled officers can "stop and frisk" an individual without violating the Fourth Amendment if they have reasonable suspicion the person is armed and thus presents a threat to the officers. In *Michigan v. Tyler* (1978)[6], the Court held government actors may seize evidence of a crime if it is in "plain view" from a place they were allowed to be in. Thus our estimate of λ_i may well be that the law implies a ruling that the track suit was legally seized, since the officer suspected "someone else" (who could pose a danger to the officers at the scene) was "down there" in the basement where he found the track suit in "plain view". The sincere preferences ϕ_j of judge j may be to rule it was a constitutional seizure or an unconstitutional one.

[5] 392 U.S. 1.

[6] 436 U.S. 499. The plain view doctrine relied on in *Tyler* was first articulated by the Court in *Coolidge v. New Hampshire*, 403 U.S. 443 (1971), though in that case the Court described it only to state the facts of *Coolidge* did not meet the requirements of the plain view doctrine. In *Horton v. California*, 496 U.S. 128 (1990), decided the same term as *Buie*, the Court held that does not even matter if the discovery of the evidence "in plain view" was inadvertent or not.

It will be useful to consider λ to be a mapping from cases to continuous outcomes rather than dichotomous ones. Judges must assign a dichotomous outcome to each case (government action is constitutional or unconstitutional; or the plaintiff wins or defendant wins, etc.), but some cases are a "closer call" than others. Then with some function σ mapping $\mathbb{R}$ to $[0, 1]$, which I will take to be the logistic function,[7] $\sigma(\lambda_i)$ gives the probability case i should receive a positive outcome. Cases with large positive λ_i are cases where we are quite certain the case should receive a positive outcome, while cases with large negative λ_i are cases where we have high confidence the case should receive a negative outcome; cases with λ_i close to zero are "close calls".

Approaches to model or account for judicial preferences in models of case adjudication differ. Case space models typically concern themselves only with ϕ_j, and how judges act strategically to enact their preferences within particular institutional arrangements or strategic situations. Bartels (2009) instead focused on controlling for legal effects via λ and capturing judicial preferences via Martin-Quinn scores. Here I focus on accounting for the fuller picture of judicial preferences ϕ_j and utilizing λ as the predictor of interest to capture the effect of the law on judicial preferences. That is, the judges' sincere preferences in these cases are not wholly captured by a unidimensional measure such as Martin-Quinn scores. The Martin-Quinn measure is useful as an overall measure of ideology, but issue-specific preferences are best described with reference to the case space defined by the relevant facts in such cases. However, the law, which we must measure or operationalize in some way to assess legal constraint, is also best described as a mapping between this space and outcomes.

Separating out these effects—the pull toward a particular vote from the judge's individually truly preferred outcome given the characteristics of the present case and the pull from the outcome implied by the Court's past decisions given the characteristics of the present case—is non-trivial. We cannot observe the justices' sincere policy preferences ϕ_j, and both ϕ_j and λ are themselves functions of the same case factors X.

[7] I will assume the use of the logistic function throughout the paper, but a number of other functions such as the normal cumulative distribution function could be used.

However, there are some things we can say even with such a general theoretical setup. First, consider our quantity of interest is the average marginal effect of the law, or how much, on average, λ influences a justice's decision; considering all the cases this justice has heard, how much on average would the value of f_j change with an increase in λ (i.e. with an increase in the probability that *the law* says the outcome should go a particular way)? This average marginal effect is by definition

$$\gamma_j = \frac{1}{N} \sum_{i=1}^{N} \frac{\partial}{\partial \lambda_i} f_j \left(\mathbf{x}_i \right). \tag{1}$$

Next, if the law exerts no constraining effect on judges' decisions, then $f_j = \phi_j$. This implies that typically we should see an average effect of 0 if judges' behavior is unconstrained, or an average positive effect if the law is constraining their behavior.

Could λ have an average *negative* effect on f_j? Let's consider a basic sequential game where the justices are voting on cases one by one (as we see in reality). Absent legal constraint, under a wide variety of assumptions about preferences, it is a weakly dominant strategy for justices to always vote precisely in line with their preferences.[8] That is, no matter how the other justices are voting, if justice j votes how they would sincerely prefer, they can only make the outcome they prefer more likely rather than less, and further, their preferred outcome occurring in that case can only make the law get closer to their preferred policy rather than farther away.

In other words, in no circumstances will such a vote cause λ to deviate further from ϕ_j than it would given a contrary vote, and in some circumstances—when j would be a pivotal vote in a majority coalition—it will cause λ to come more into line with ϕ_j. So when j might be pivotal, if j is not constrained by law, they are strictly better off by voting their preferences. However, when there is no chance for j to be pivotal, they are indifferent between votes except perhaps for expressive reasons. If we assume some noise or imperfection in a

[8]See Proposition 5 in Lax (2007).

judge's actions[9], this can result in a negative relationship between λ and f_j since then her expressive motivations break her indifference more often in more extreme cases (in terms of the difference between λ and ϕ_j).

It is also important to note that in the applied (empirical) setting, where we cannot directly observe ϕ_j, and may not even be certain about all its inputs, omitted variable bias could bias our *estimate* of γ—if some factor Z in truth is an input to both λ and ϕ_j, but is not included in the statistical model, we may see a spurious (positive or negative) effect. However, if this relationship is only seen in non-pivotal votes, it provides at least suggestive evidence that an expressive mechanism is at play rather than omitted variable bias.

We may also consider different reasons *why* the law would enter into judges' decisions. From one perspective, judges may allow the law to influence their decisions contrary to their personal policy preferences because they also intrinsically care about the law itself. That is, they have both individual policy preferences, but also some preference for deciding cases consistent with the Court's past decisions (and these things may sometimes conflict). However, we may also think that judges only allow the law to constrain their decisions because they do not want to be *seen* as going against the law, perhaps for institutional legitimacy reasons. Bartels (2009) and Bartels (2011) suggest certain legal factors constrain ideological voting by judges more than others. In particular, Bartels (2009) finds that ideology influences Supreme Court justices' votes to a greater degree in cases involving "intermediate scrutiny" than in cases involving either a "rational basis" or "strict scrutiny" test. In the latter two cases, legal doctrine strongly suggests the correct legal outcome should likely be a ruling that government action is constitutional and unconstitutional respectively, while there is much more variation in the implied legal outcome for intermediate scrutiny cases. In the context of the model of judicial decision making considered here, when the absolute value of λ is higher, the implied legal outcome is more certain; large positive values of λ indicate a high probability the present case should, legally speaking, receive a positive outcome while

[9]Or perhaps we may assume some imperfect knowledge of their own preferences; sometimes judges might be confronted with a "hard case".

large negative values indicate a high probability the present case should receive a negative outcome). When the absolute value of λ is close to zero, there is a roughly 0.5 probability that either outcome is legally correct. So if judges have some intrinsic concern for following the law, the effect of λ should be more or less constant over the range of λ values, whereas if the mechanism argued for in Bartels (2009) drives law's constraining effect, then the effect of λ should be high when the absolute value of λ is high, but lower when λ is closer to zero.

In sum,

1. If $\gamma_j > 0$, it implies the law exerts some positive constraining effect on j's decision making.

2. If $\gamma_j = 0$, then j's decisions are not constrained by the law and they act only according to their own preferences ϕ_j.

3. If $\gamma_j < 0$, we should expect the negative relationship between λ and f_j to exist only in cases where j is unlikely to be a pivotal justice.

4. If γ_j is increasing in $|\lambda|$, it implies j follows the law only when the legal outcome is more certain (offering less ideological discretion), whereas if γ_j is constant across $|\lambda|$, j may have some intrinsic concern for following the law.

4 Data and Methods

To assess legal constraint at the Supreme Court, I will utilize both court- and justice-level data on case dispositions. A court-level model provides predicted values to serve the role of λ in Section 3; at both levels I use Gaussian process (GP) classification, a method described further in Section 4.2.

4.1 Data

I use data on First Amendment Free Expression cases at the U.S. Supreme Court. This setting has been frequently studied, both in the context of studying legal change (Richards and Kritzer 2002; Bartels and O'Geen 2015) as well as studying legal constraint (Bartels 2009) as here. Richards and Kritzer (2002) original coded the facts of all Free Expression cases heard by the Supreme Court in the 1953 to 1998 terms, and Bartels and O'Geen (2015) backdated and updated this data to include cases from the 1946 to 2004 terms.[10] I updated this data to include cases from the 1946 to 2019 terms (adding the cases from the 15 most recent terms). These free expression cases are coded for: whether the court-level outcome was liberal (i.e., pro-expression) or conservative (i.e., anti-expression), measures of the facts of each case (discussed further below), and the median Martin-Quinn score (Martin and Quinn 2002) for the Court at the time of that decision.[11] I use the justice-level Supreme Court Database (Spaeth et al. 2020) and the justice-level Martin-Quinn score data (Martin and Quinn 2020) to expand the data to the justice level. The relevant case facts in Free Expression cases are: the *Category* of restriction, or whether the restriction on speech is content-based, content-neutral, or of a less protected category of speech; the *Actor*, or who is restricting speech, such as the federal government, a state government, or a private actor; what the *Restriction* is, such as a criminal sanction or a loss of employment; and the *Identity* of the speaker, such as whether they are a politician, an alleged communist, or a racial minority (Bartels and O'Geen 2015). After merging records, there are 677 court-level outcomes to analyze and 5,922 justice-level votes to analyze.[12]

[10]The data for Bartels and O'Geen (2015) are published as Bartels and O'Geen (2014).

[11]As Martin and Quinn (2005) explain, despite the tautological issue of "votes explaining votes", use of Martin-Quinn scores as explanatory variables is appropriate when the votes in question are only from one issue area (see also Bartels 2009).

[12]There are 6,076 justice-decision combinations, but sometimes justices do not participate in a particular case (such as for ethical reasons, or because arguments were given before they joined the Court), leaving 5,922 observations remaining after accounting for missingness in the outcome.

4.2 Gaussian process classification

Gaussian process (GP) models are a class of flexible machine learning models (Rasmussen and Williams 2006) that political scientists have recently begun to utilize (Carlson 2021; Duck-Mayr, Garnett, and Montgomery 2020; Gill 2021).[13] Duck-Mayr (2021b) provides a fuller introduction to the method for political science, but as the method is fairly new to political science, I provide a brief introduction here.

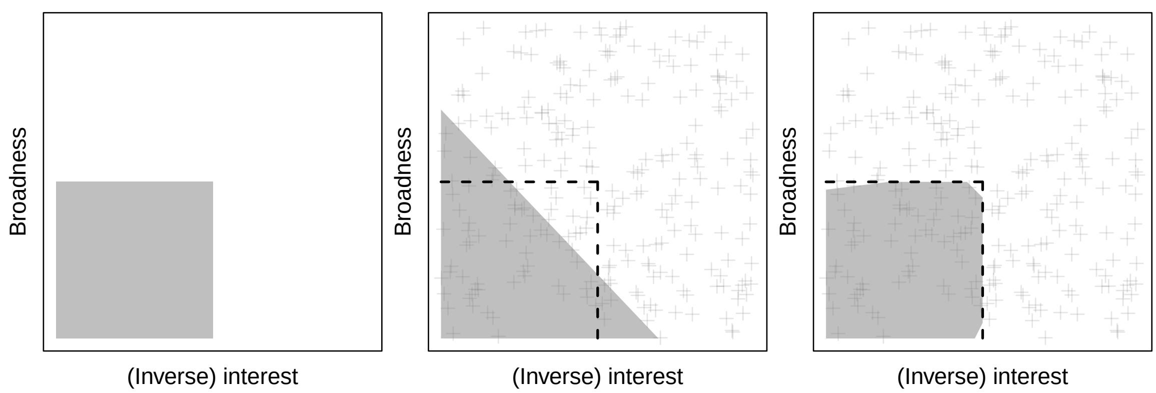
GP models are used to learn the mapping from predictors to outcomes when its functional form is unknown. Think back to my discussion in Section 3 of what the law might imply in the *Buie* case. I did not have to specify some linear relationship between individual relevant facts and outcomes, or even tell you anything about the shape of the relationship between X and y; I told you about the cases that were close to *Buie* in X and what their outcomes were, which gave you some idea of what the legal outcome would be in *Buie*. That is essentially what GP models do: They learn about what the outcome should be in our test cases by assuming it will be similar to the outcome in the cases that are closest in the covariate space. That GP models can accommodate a wide variety of "shapes" is of particular interest in the context of judicial politics; Kastellec (2010) rightly points out that linear models often used to relate case factors to case outcomes are too inflexible to accommodate many common types of legal rules.

For example, consider a common type of legal rule, called a "conjunctive rule", where all elements of the rule must be met for a positive outcome, such as a law facing strict scrutiny analysis. For a law to pass constitutional muster under a strict scrutiny test, it must both serve a compelling governmental interest and be narrowly tailored to achieve that interest. Such a rule is depicted in Figure 1. The x-axis captures whether the governmental interest is "compelling"; the axis is reverse coded, such that low values indicate a compelling interest and high values indicate an interest that is less than compelling. The y-axis is for the broadness

[13] They share some similarities to the kernel-regulated least squares model introduced by Hainmueller and Hazlett (2014), but are more versatile, accommodating categorical outcomes, as well as a variety of kernels, and providing a Bayesian approach (see Cheng et al. 2019).

of the regulation at issue; low values indicate a narrowly tailored regulation and high values a regulation that is not narrowly tailored.

Taking each dimensions as being a continuum from 0 to 1, with 0.5 being the threshold on each dimension for meeting that element of the strict scrutiny test, I simulated 250 cases, drawing the case factors independently from a standard uniform distribution for each case. I assigned them outcomes according to the rule that a "constitutional" outcome should be given if and only if both broadness and inverse interest were less than 0.5. I then trained both a logit model (the standard approach in prior literature) and a GP classifier on these cases. I then determined the set of cases each model would assign a "constitutional" outcome to, and compared these estimated rules to the true strict scrutiny test; the results are depicted in Figure 1.



(a) True strict scrutiny test (b) Logit estimate (c) GP classification estimate

Figure 1: Comparing estimation of legal rules between linear models and GP classification. In panel (a), the true strict scrutiny test is depicted with the set of cases receiving a "constitutional" outcome indicated with gray shading. In panels (b) and (c), estimates of the rule from a logit model and GP classifier respectively are depicted with gray shading. In those panels the true rule is indicated with a dashed line, and the simulated cases the models were trained on are indicated with light gray crosses.

The GP classifier is able to develop a much more accurate estimate of rules such as the conjunctive rule in Figure 1.[14] This is accomplished by making less restrictive assumptions. GP classifiers typically only make some (usually fairly minimal) assumptions about the

[14]To be more specific, using a linear model is equivalent to the very restrictive assumption that judges' preferred rules must be a hyperplane in the case space.

covariance between outcomes conditional on the predictors. Then Bayes rule combined with some linear algebra allows us to derive a posterior distribution. More specifically, the model uses a logistic likelihood as in a typical logit model,

$$\Pr(\mathbf{y} \mid \mathbf{X}) = \prod_i \sigma\left(y_i f\left(\mathbf{x}_i\right)\right), \tag{2}$$

(where σ is the logistic function). However, rather than assuming a linear (or any particular) form for f, we simply put a prior distribution on the latent outcomes,

$$f(\mathbf{X}) \sim \mathcal{N}\left(\mathbf{0}, K\left(\mathbf{X}, \mathbf{X}\right)\right), \tag{3}$$

where K is a matrix-valued function whose i, j entry gives the prior covariance between latent outcomes for observations i and j.[15] An overwhelmingly popular choice for K, which I use in this article, is the squared exponential covariance function,

$$K\left(\mathbf{x}_i, \mathbf{x}_j, \sigma_f, \boldsymbol{\ell}\right) = \sigma_f^2 \exp\left(-0.5 \sum_d \frac{\left(x_{id} - x_{jd}\right)^2}{\ell_d^2}\right), \tag{4}$$

where σ_f is a scaling factor for the covariance matrix and $\boldsymbol{\ell}$ is a vector of length scales. Essentially, the assumption of the model here is simply that $\mathbf{X}$ observations that are closer together in the covariate space will be more likely to have latent outcomes that are close together; $\boldsymbol{\ell}$ basically tells us what "close" means on each dimension. From this simple assumption, the posterior over f—in other words, an update of our belief about what sort of relationship there is between X and y after observing some data—can actually be quite easily approximated with a Taylor expansion (for details, see Duck-Mayr 2021b).

This posterior, to return back to the notation of our specific context, is referred to above as λ from the court-level model. Duck-Mayr (2021b) also derives average marginal effects for GP models, which I use to calculate γ, the ultimate quantity of interest.

[15] Technically this prior distribution need not have a zero mean either; see Duck-Mayr (2021b) for more details.

4.3 Identification

I am interested in determining the extent to which the decisions of the justices on the United States Supreme Court are influence by the law, or the latent legal outcome implied by the Court's past decisions. As discussed above, this conception of the law clearly implies a measure of the law: the prediction in case i given its characteristics $\mathbf{x}_i$ from a model trained only on the Court's cases $1, \ldots, i-1$, denoted by $\boldsymbol{\lambda}$, or λ_i in case i. However, I am more specifically interested in the effect of the law on justices' decisions, *independent of* the justices' own preferences ϕ over outcomes. Case characteristics can influence justice j's decisions not only through the law, but also through their own preferred mapping from case characteristics to outcomes. However, we cannot directly observe ϕ_j, so we must include in our model of justice j's decisions not only $\boldsymbol{\lambda}$, but also $\mathbf{X}$ itself. But, moreover, we must account for the fact that justice j's decisions also have the potential to dynamically impact our variable of interest, λ. This issue becomes clearer by considering the directed acyclic graph for the influences on justices' decisions.

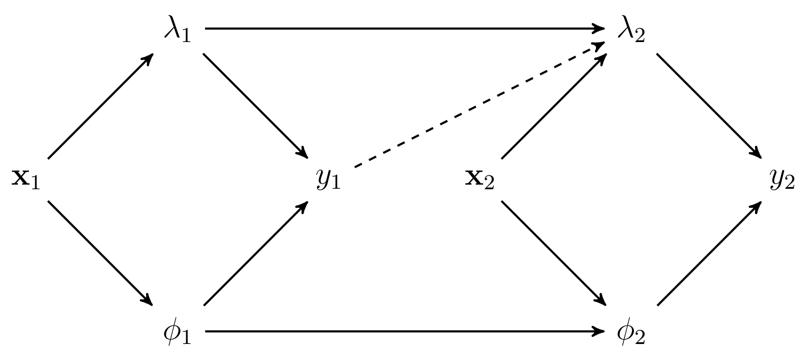


Figure 2: Directed acyclic graph for the influence on a justice's decisions. $\mathbf{x}_i$ denotes the facts of case i, y_i denotes the justice's decision in case i, λ_i denotes the predicted legal outcome in case i, and ϕ_i represents the justice's policy preferences in case i.

We want to estimate the influence of the law on justices decisions; in Figure 2, this is represented by the arrows from λ_i, the predicted legal outcome in case i to y_i, the justice's decision in case i. However, this relationship can be confounded by the justice's own policy preferences that get "baked in" to the law when they are a part of the majority coalition in a case; in Figure 2, this is represented by the dashed arrow from y_i to λ_{i+1}. This line is dashed to illustrate that it is a *potential* pathway;[16] the justice's preferences only become part of the law if they are in the majority coalition. If we do not block these dashed pathways, *when they are active*, our estimate of the effect of λ may be biased[17] by the justice's mere agreement with their own past self rather than acquiescence to the law. So, consider the sequence of events. In case 1, controlling for $\mathbf{x}_1$ is sufficient to identify the effect of λ_1; then λ_1 has no path to y_1 through $\mathbf{x}_1$ and ϕ_1. However, in case 2, *if* the justice was part of case 1's majority coalition, controlling for $\mathbf{x}_2$ is no longer sufficient; a path then runs from λ_2 to y_2 through $y_1, \phi_1,$ and ϕ_2.

This is a form of time-varying confounding, a particularly difficult issue for inference. Political science has recently begun taking cues from biostatistics to deal with various forms of time-varying confounding, such as using structural nested mean models (Acharya, Blackwell, and Sen 2016), or marginal structural models (MSM) (Robins, Hernán, and Brumback 2000; Torres 2020) to estimate controlled direct effects of past treatments when there are intervening treatments, or to estimate cumulative effects of dynamic treatments with MSM (Blackwell 2013). However, while these frameworks account for dynamic treatments, they do not account for dynamics where past outcomes can affect current treatments.[18]

Sequential conditional mean models (Liang and Zeger 1986) are sometimes used with longitudinal data. These models are closer to our setting, where past outcomes, treatments, and confounders can all have an effect on present treatment. However, the approach does not directly extend to our setting, as the past outcomes' effects on current treatments are actually a function of present covariates, rather than an unconditional effect.

In our case, we can take advantage of the nature of $\boldsymbol{\lambda}$ as the prediction from a GP model

[16] Technically in a DAG all pathways are considered potential pathways, but here we observe whether the pathway is active or not, which is what we emphasize.

[17] In the Bayesian setting, we are not strictly speaking concerned with the frequentist conception of bias. However, we are concerned with the correct interpretation of the parameters we estimate. Without conditioning on the proper set of variables, the average marginal effect of λ_i we estimate will not represent solely the constraining effect of the law, but may also include an effect of agreement with your own past self, whose views have been partially enshrined in the law.

[18] Additionally, the MSM framework is used for discrete treatment values rather than our continuous λ.

to give us a direct way to control for the effect of past values of $\mathbf{y}$ that have impacted $\boldsymbol{\lambda}$. For case i,

$$\lambda_i = \sum_{t=1}^{i-1} k\left(\mathbf{x}_i, \mathbf{x}_t\right) \alpha_t, \tag{5}$$

$$\boldsymbol{\alpha} = \nabla \log p\left(\overline{D}_{i-1} \mid \overline{\boldsymbol{\lambda}}_{i-1}\right), \tag{6}$$

where $\overline{D}_{i-1}$ and $\overline{\boldsymbol{\lambda}}_{i-1}$ indicate the decisions of the Court and the predicted legal outcomes in cases before case i. Letting $\mathcal{Y}_i$ be the set of indices of the justice's $\mathbf{y}$ decisions that have been in the majority up to case i, then

$$\rho_i \triangleq \sum_{t \in \mathcal{Y}_i} k\left(\mathbf{x}_i, \mathbf{x}_t\right) \alpha_t \tag{7}$$

gives the total impact justice j's preferences have had on λ_i. In other words, λ_i is a weighted sum, and the elements of that sum attributable to justice j's chosen outcomes are directly identifiable, so we can control for them using that portion of the weighted sum. Thus, controlling for ρ_i blocks the backdoor path from λ_i to y_i through $\overline{\mathbf{y}}_{i-1}$ and $\overline{\boldsymbol{\phi}}_{i-1}$, allowing us to obtain a correct estimate of the effect of λ on y, *assuming* all other confounders on present decisions are accounted for.

4.4 Model specification

For the Court-level model, I use as predictors each of the case factors identified by prior literature as relevant: the *Category* of restriction, the *Actor* imposing the restriction, the *Restriction* itself, and the *Identity* of the speaker. I also use the *Term* of the Court as a predictor, to allow doctrine to fluctuate over time. Finally, I include the *Median Martin-Quinn Score* on the Court to capture any remaining ideological nature of these cases outside the dimensions of preference given by the case factors. To ensure sufficient training data in the

Court-level model, and sufficient observations in the justice-level models, I use data before the appointment of Rehnquist as purely training data and focus on the final natural court in the Rehnquist court—consisting of Chief Justice Rehnquist and Justices Breyer, Ginsburg, Kennedy, O'Connor, Scalia, Souter, Stevens, and Thomas—for analysis.[19]

I cannot optimize the prior in the Court-level model as a model-selection step or ρ would be infected by other information in the prior optimization step, so I specify at the outset a somewhat agnostic prior for the Court-level model. I use a scale factor of approximately 2.2, or the 90% quantile of the logistic distribution. This corresponds to an assumption that on average, cases that make it to the Supreme Court will have a probability of between 0.2 and 0.8 of having a liberal outcome; that is, we keep the model flexible enough to predict extreme outcomes but assume that most cases that make it to the Supreme Court are not "easy cases". For the length scales, I use a length scale of one for each category of every case factor so that we are not *a priori* imposing a relative importance of case factors but letting the model learn which case factors are more important over time as we add data. Finally, for the *Term* and *Median Martin-Quinn Score* variables, I use the inter quartile range for the length scale, which is about 24 and 0.11 respectively. This allows the model to understand these variables are on a much larger and smaller scale respectively than the others.

In the justice-level models, I use each of the case factors as predictors to capture the justice's own preferences, as well as the justice's individual *Martin-Quinn Score* to capture any residual ideological influence on their votes. Again, I include the *Term* to allow justices' preferences to vary over time. Finally, I include λ, or the predicted outcomes from the Court-level model, to represent the law, and ρ to control for the justice's own contribution to λ. As is standard for GP classification, I first engage in a model-selection step to set the prior's hyperparameters, then fit the justice-level models to their individual voting records and calculate the average marginal effect γ as described in Duck-Mayr (2021b). All analyses were conducted in `R` (R Core Team 2021) utilizing the R extension package `gpmss` (Duck-Mayr

[19]For completeness I also conduct analysis on the other justices as well; the results in these other models are substantively similar to those presented in the main paper and are reported in Appendix B.

2021a) for GP classification-specific functionality.

5 Results

The effect of λ is listed for each justice in Table 1 with 95% credible intervals.[20] For interpretability, I calculated the effect on the probability scale, and as a discrete difference (also called "first differences") rather than marginal effect of instantaneous change. As choosing any particular points at which to calculate the discrete difference is somewhat arbitrary, I report two different choices of difference points: A change between the mean of λ minus one standard deviation and the mean plus one standard deviation (-1.05 and 1.27, respectively), and a change between the minimum and maximum values of λ (-2.42 and 3.09 respectively); both are common choices for reporting first differences. The mean plus and minus a standard deviation also offers a nice substantive interpretation: We are comparing a point at which we are fairly certain the law implies a liberal ruling ($\lambda = 1.27$ corresponds to a 78% probability the ruling should be liberal according to the Court's past decisions) and a point at which we are fairly sure the law implies a conservative ruling ($\lambda = -1.05$ corresponds to a 26% probability the ruling should be liberal according to the law). So the effect tells us how much more likely a justice is to vote liberally in a case if the Court's past decisions imply a 78% probability the "correct legal outcome" in the case is liberal versus a 26% probability it is liberal. This effect is averaged over all observations in the sample. For example, the average marginal effect of λ for Justice Kennedy is 0.13 with a 95% credible interval of [0.06, 0.22]. This means that for every set of case facts Kennedy was actually presented, on average Kennedy's probability of voting liberally would increase by 13% if past cases implied a probability of 78% the correct legal outcome is liberal versus if past cases implied a probability of 26% the correct outcome is liberal. For the mean plus and minus a standard deviation, I also depict the average marginal effect estimates in Figure 3 with 90% and 95% credible intervals.

[20] Average marginal effects for all predictors are given in Appendix A

Table 1: Average marginal effect of λ on judges' decisions. For each justice, I report the estimate and 95% confidence interval for the difference in the probability the justice will vote liberally between two different values of λ, averaged over all observations in the sample.

	Mean of $\lambda \pm$ sd of λ	Range of λ
Breyer	0.07	0.18
	[0.01, 0.13]	[0.12, 0.25]
Ginsburg	-0.14	-0.29
	[-0.23, -0.05]	[-0.39, -0.18]
Kennedy	0.13	0.31
	[0.06, 0.22]	[0.22, 0.39]
O'Connor	0.08	0.21
	[0.04, 0.13]	[0.16, 0.25]
Rehnquist	0.05	0.12
	[0.00, 0.09]	[0.07, 0.16]
Scalia	0.10	0.22
	[0.02, 0.17]	[0.13, 0.31]
Souter	0.04	0.11
	[-0.12, 0.21]	[-0.07, 0.27]
Stevens	0.06	0.16
	[0.02, 0.11]	[0.11, 0.20]
Thomas	0.02	0.06
	[-0.06, 0.10]	[-0.04, 0.15]

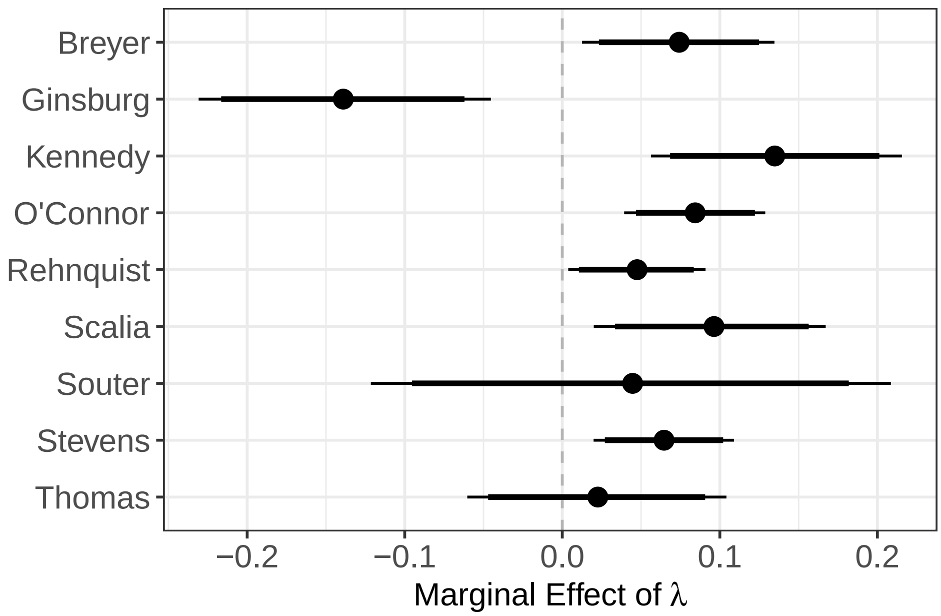


Figure 3: Average marginal effect of λ on judges' decisions. Estimates are depicted with circles, 90% credible intervals are depicted with thick line segments, and 95% credible intervals are depicted with thinner line segments. Effects are on the probability scale and calculated for a concrete difference; they reflect the difference in the probability the justice will vote liberally when the Court's past decisions give a 0.78 probability the outcome should be liberal vs. a 0.26 probability the outcome should be liberal (corresponding to latent legal outcomes of 1.27 and -1.05 respectively—the mean of λ plus and minus one standard deviation).

We see that for most justices, the law exerts a reliable influence on their decision making; the 95% credible interval for Justices Breyer, Kennedy, O'Connor, Scalia, Stevens and Chief Justice Rehnquist contains only positive values at both difference levels. However, for other justices, the credible intervals bound zero, indicating the law does not have a reliable effect on their decisions. Justices Souter and Thomas fall into this group. In some ways, this affirms past results in the literature; some justices act relatively unconstrained as argued in Segal and Spaeth (2002). However, contrast the closest analogous attitudinalist finding: Segal and Spaeth (1996) studied dissenters to landmark cases and their subsequent votes in related cases. They found the justices *most* deferential to the precedent they disagreed with still voted in line with their preferences (and against the legal outcome) two thirds of the time, and for a supermajority of justices, that occurred over 90% of the time. By contrast, here we see reliable evidence that the law on average affects the decisions of a supermajority of the Court![21]

Similarly, when we consider studies that do find some constraining effect of law, there are important differences to the results here. First, and perhaps most importantly, past work in this area has often been clouded by methodological issues, casting doubt on even the mixed evidence offered in favor of legal constraint (see Lax and Rader 2010a). This study uses an approach that not only handles error correlation over (e.g.) term,[22] but crucially takes into account the "state dependence" problem, or the problem that justices' own preferences from the past get incorporated into the law for the future. Without devising a control for the justices' own influence on the current legal status quo, any constraining effect of law found could be a spurious effect. An important limitation in this approach is it requires us to identify the case factors relevant to judges' preferences and court outcomes. However, it allows us to control for this issue when we believe we have identified these factors, and there is strong evidence in several contexts of the case factors we must include (see Segal 1984;

[21] That is, a majority of the natural court studied.

[22] As Carlson (2021) explains, GP models are a natural answer to the common methodological problem in political science of violating an assumption of conditional independence.

Richards and Kritzer 2002). By addressing the state dependence issue in contexts where we can identify these factors, I offer the best available evidence to date of the constraining effect of law.[23]

Next, there are contextual differences. For example, Bailey and Maltzman (2008) focus on justices' willingness to explicitly overrule precedent; while important to assess, this is decidedly a narrower focus than the present study, which looks for the average influence the law has on justices' decisions. Similar contextual differences separate the present study from past work (e.g. Black and Owens 2009).

There are two crucial differences to Bartels (2009). First, the approach here allows for law to influence justices differently; I allow Justice Breyer to consider the current state of the law more important in his decision making than Justice Scalia does. In contrast, Bartels (2009) focuses on an effect that is homogeneous across justices by construction. Moreover, Bartels (2009) treats the case factors as relevant *legal* factors, but uses Martin-Quinn scores *only* to capture the justices' preferences, while we would expect from theoretical approaches like the attitudinal model or the case space model that we should in fact also treat justices' issue-specific preferences as defined with reference to the case factors. The modeling approach here allows us to capture both justice-level preferences and legal implications with reference to the case factors.

The average marginal effect of law for Justice Ginsburg is reliably negative. This means that as the Court's past decisions indicate a higher probability that the correct legal outcome is liberal, Justice Ginsburg becomes more likely to vote conservative. While it is possible there is some unmeasured aspect of these cases that cause Justice Ginsburg's preferred outcomes to differ from the law, which could result in the negative marginal effect we see, it is also possible the mechanism discussed in Section 3 is at play: Ginsburg may be affected by the

[23] A potential additional consideration in the Free Expression context is the ideological nature of the expression in question. Epstein, Parker, and Segal (2018) show justices may provide more protection to speech they agree with (for example, a conservative justice ruling to protect religious expression) than speech they disagree with (for example, a liberal justice refusing protection to commercial speech). I include the variables from Epstein, Parker, and Segal (2018) in the justice-level models in Appendix C, which gives substantively similar results.

law in cases where her vote is not pivotal. As shown in Figure 4, that is indeed the case. The negative correlation between Ginsburg's latent outcomes and the latent legal outcome is 0.19 [-0.12, 0.47] when Ginsburg might be pivotal (i.e., the correlation is not reliably negative), but is -0.32 [-0.50, -0.12] when Ginsburg is likely not pivotal.[24] This provides suggestive though not conclusive evidence that the mechanism from Section 3 is driving the result for Ginsburg rather than omitted variable bias.

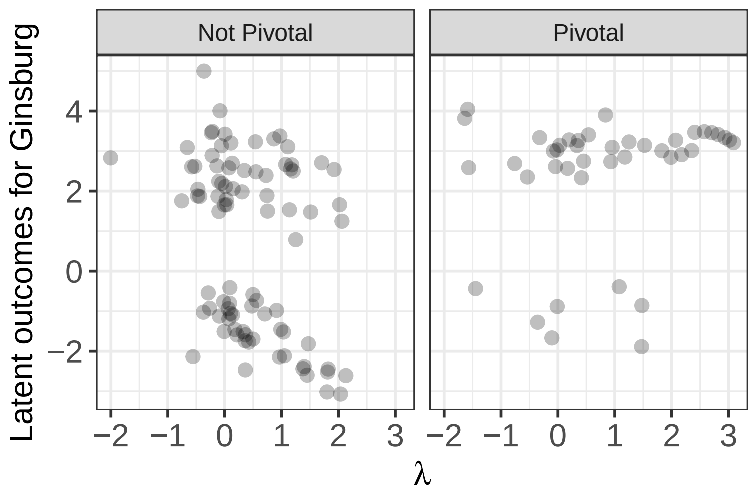


Figure 4: Comparing predicted outcomes for Justice Ginsburg against λ. Values are depicted with gray circles. The right panel labeled "Pivotal" contains observations from all cases where Ginsburg was part of a majority coalition of six or fewer justices—the situations where Ginsburg may have been pivotal. The left panel labeled "Not Pivotal" contains observations from all remaining cases.

Is the constraining effect of law here driven by situations in which doctrine allows little ideological discretion, while judges can follow their policy preferences more in less certain cases? Bartels (2009) presents evidence that ideological constraint is higher in cases where more "certain" legal tests such as the rational basis or strict scrutiny tests should apply than in cases where a more fluid test such as intermediate scrutiny should apply. In the Free Expression context, this means we should see a higher marginal effect of law in "Content

[24] Notice that I use a pivotality threshold of six here, whereas strictly speaking a justice is only pivotal if the majority coalition is five justices. Using the five justice threshold gives similar results: 0.44 [0.06, 0.71] and -0.32 [-0.48, -0.14] respectively. However, I use the higher threshold to add observations to the pivotal group, to avoid concern that the pattern was driven by too few observations in the pivotal group. This also serves to cover cases where Ginsburg is in truth pivotal but her presence in the coalition causes another justice to join, or where Ginsburg has some uncertainty about whether she would be pivotal.

based" restriction and "Less protected" forms of expression cases than in "Content neutral" restriction cases. Figure 5 shows the marginal effect of λ averaged over each case that each justice heard from each of those three categories.

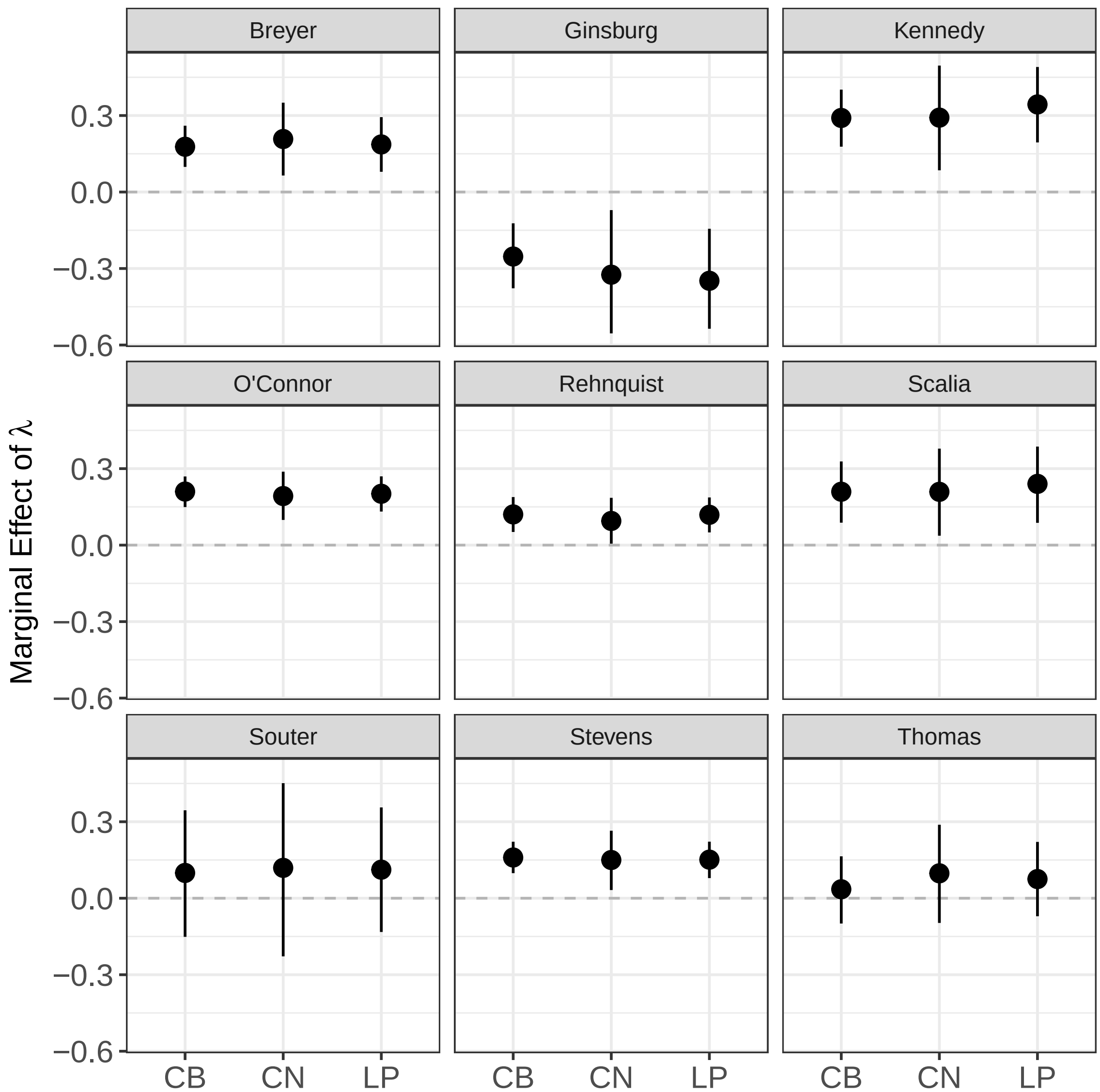


Figure 5: Average marginal effect of λ by category. The effect displayed in this figure uses the difference between minimum and maximum values of λ. Content based is abbreviated as CB, Content neutral as CN, and Less protected as LP.

While we do not see a difference in magnitude of the effect of law by category as Bartels (2009) found, we see an effect in certainty; the credible intervals for the "Content based" and "Less protected" categories are somewhat tighter than that for the "Content neutral" category.

However, this is likely simply caused by the fact that the "Content neutral" category has fewer cases in it, as shown in Table 2.

Table 2: Number of observations in each category for each justice.

Justice	Less protected	Content neutral	Content based
Breyer	35	19	71
Ginsburg	36	19	75
Kennedy	53	25	107
O'Connor	73	27	125
Rehnquist	145	42	188
Scalia	58	26	114
Souter	37	19	70
Stevens	120	36	171
Thomas	42	21	80

So we can consider the more general formulation from Section 3: If justices only feel the constraining effect of law when legal outcomes are more certain (thus giving them less ideological discretion), the pointwise partial derivatives, giving the effect of law in each case, should be increasing in $|\lambda|$. However, as shown in Figure 6, this is not the case, providing suggestive evidence that for justices constrained by the law, they feel some *internal* constraint of the law, or in other words, a preference for following the law that can conflict with—and perhaps in some cases override—their sincere policy preferences.

6 Conclusion

I asked at the outset, "(How much) does the law affect judges decisions?" The answer is the law does affect judges decisions, substantially… at least some of them. Several studies approach the question of legal influence on judicial decision making using contexts other than votes on the merits and find the legal effects in decisions such as agenda-setting decisions (Black and Owens 2009), citation choices (Hansford and Spriggs 2006; Hinkle 2015), or willingness to explicitly overrule precedent (Bailey and Maltzman 2008). Bartels (2009) studies justices' votes on the merits and finds a constraining effect of law, but with an analysis

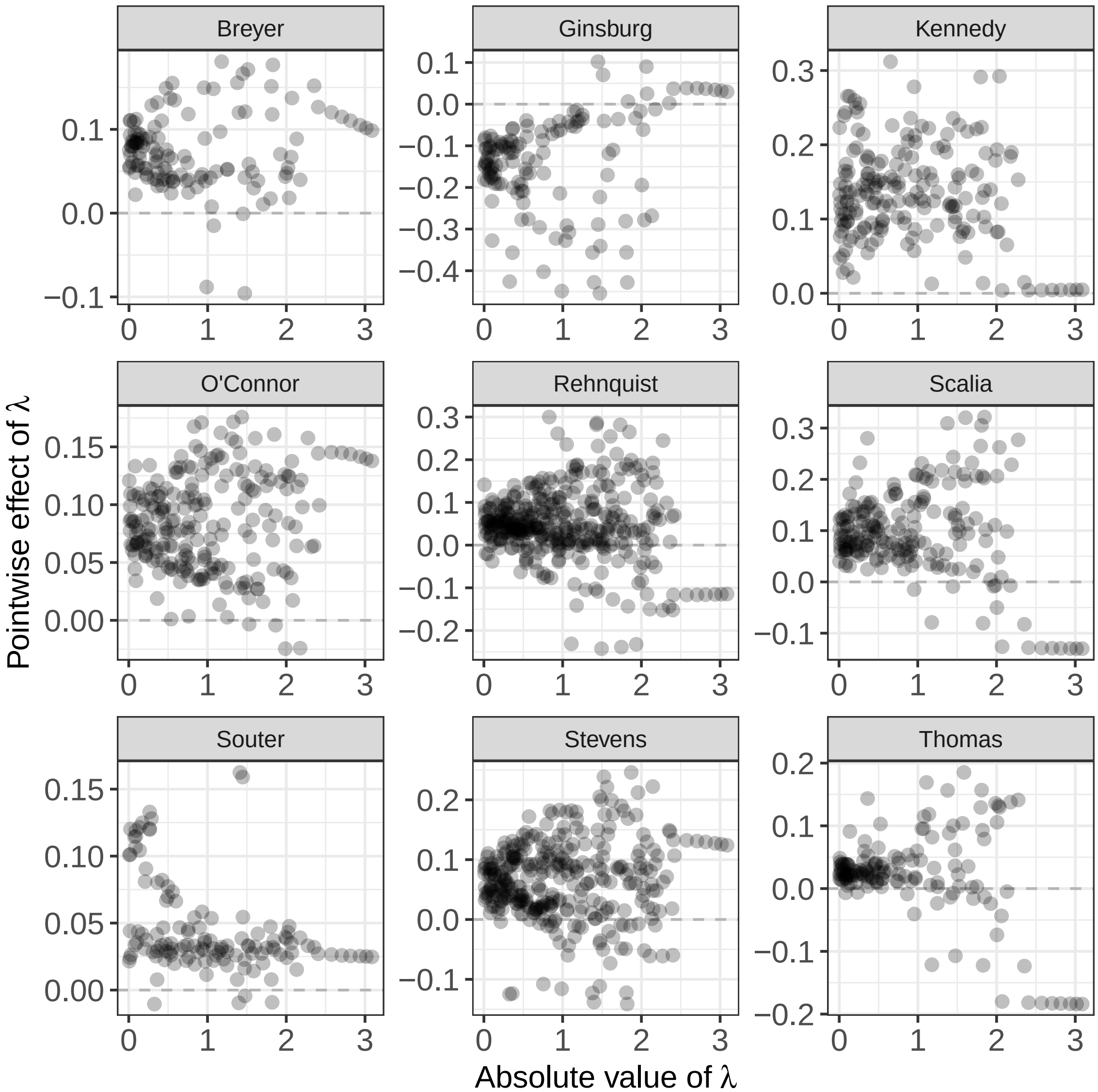


Figure 6: Pointwise effect of λ by the absolute value of λ for each justice.

that pools the justices together rather than an analysis that allows some justices to be unfazed by the force of law.

I start from a general theoretical approach based on case space models (Lax 2011) and extended it to consider the law as a potential explicit influence on judges' preferred outcomes. Decision makers in this model can consult the set of cases previously decided by the Court, comporting with the typical model of legal reasoning (Levi 1949): we consider that like cases ought to be treated alike, and determine the outcome reached in the most closely analogous cases to the one we are deciding. This framework suggests a measure of the law: The predicted outcome in each case given a model trained only on the cases that came before it. I use a model to generate these predictions, GP classification (Rasmussen and Williams 2006), that accommodates realistic forms of Court doctrine better than some more restrictive approaches to modeling doctrine taken in the past (see Kastellec 2010). Importantly, this approach readily allows for a way to not only directly test the average influence the law exerts on justices' decisions, but crucially while controlling for the justices' own impact on the law—an issue of time-varying confounding that would otherwise pose a danger to inference. With the effect of law thus carefully identified, I apply the method to the natural court beginning with the appointment of Stephen Breyer and ending with the end of the Rehnquist Court. I show several justices exhibit reliable constraint from the law in their decision making, though others do not.

This study makes several major contributions. First, I provide a fresh theoretical perspective on legal constraint by conceptualizing the law as the implied outcome in each case given the cases that came before it. Second, I present credible evidence of a substantial constraining effect of law for some justices by addressing the danger to inference presented by the justices' own votes affecting the legal status quo whose effect we are trying to measure. Third, developing this measure of the legal status quo in itself is a contribution; status quo policy is an important consideration in a wide variety of political contexts. While, for example, Black and Owens (2009) simply uses the median Martin-Quinn score of the lower court panel to measure the legal status quo, the measure I develop provides the legal outcome implied by the Court's past decisions, a quantity that is a closer theoretical match to the concept of the legal status quo. I also provide suggestive evidence differentiating between reasons why the law matters, indicating some justices have a preference for following the law rather than seeing it as a constraint due to (for example) legitimacy needs, which few studies have attempted. Finally, I highlight common mismatches between methods and theory in studies of judicial politics: judges' preferences within an issue area are best conceptualized as multidimensional, and with respect to case facts; and empirical models seeking to capture either the law or judges' individual preferences with respect to case facts should be flexible enough to accommodate any shape rather than imposing the strict assumption of linearity as in past studies. Thus, I provide the most convincing evidence to date of a substantively important effect on judicial decision making as well as provide methodological tools and measurement strategies useful for future research in judicial politics.

References

Acharya, Avidit, Matthew Blackwell, and Maya Sen. 2016. "Explaining Causal Findings Without Bias: Detecting and Assessing Direct Effects." *American Political Science Review* 110 (3): 512–529.

Bailey, Michael A., and Forrest Maltzman. 2008. "Does Legal Doctrine Matter? Unpacking Law and Policy Preferences on the U.S. Supreme Court." *American Political Science Review* 102 (3): 369–384.

Bartels, Brandon L. 2009. "The Constraining Capacity of Legal Doctrine on the U.S. Supreme Court." *American Political Science Review* 103 (3): 474–495.

———. 2011. "Choices in Context: How Case-Level Factors Influence the Magnitude of Ideological Voting on the U.S. Supreme Court." *American Politics Research* 39 (1): 142–175.

Bartels, Brandon L., and Andrew J. O'Geen. 2014. *Replication data for: The Nature of Legal Change on the U.S. Supreme Court: Jurisprudential Regimes Theory and Its Alternatives.* Version V3. Harvard Dataverse. https://doi.org/10.7910/DVN/26522.

———. 2015. "The Nature of Legal Change on the U.S. Supreme Court: Jurisprudential Regimes Theory and Its Alternatives." *American Journal of Political Science* 59 (4): 880–895.

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.

Blackwell, Matthew. 2013. "A Framework for Dynamic Causal Inference in Political Science." *American Journal of Political Science* 57 (2): 504–520.

Callander, Steven, and Tom S. Clark. 2017. "Precedent and Doctrine in a Complicated World." *American Political Science Review* 111 (1): 184–203.

Carlson, David. 2021. "Modeling Without Conditional Independence: Gaussian Process Regression for Time-Series Cross-Sectional Analyses." Under review, https://mysite.ku.edu.tr/dcarlson/research/.

Cheng, Lu, Siddharth Ramchandran, Tommi Vatanen, Niina Lietzén, Riitta Lahesmaa, Aki Vehtari, and Harri Lähdesmäki. 2019. "An Additive Gaussian Process Regression Model for Interpretable Non-Parametric Analysis of Longitudinal Data." *Nature Communications* 10 (1): 1–11.

Denison, Alexander, Justin Wedeking, and Michael A. Zilis. 2020. "Negative Media Coverage of the Supreme Court: The Interactive Role of Opinion Language, Coalition Size, and Ideological Signals." *Social Science Quarterly* 101 (1): 121–143.

Duck-Mayr, JBrandon. 2021a. *gpmss: Gaussian Process Models for Social Science.* R package version 0.1.0.

———. 2021b. "Inference in Gaussian Process Models for Political Science." Working paper, http://jbduckmayr.com/research.

Duck-Mayr, JBrandon, Roman Garnett, and Jacob Montgomery. 2020. "GPIRT: A Gaussian Process Model for Item Response Theory." In *Proceedings of the 36th Conference on Uncertainty in Artificial Intelligence (UAI),* edited by Jonas Peters and David Sontag, 124:520–529. Proceedings of Machine Learning Research. PMLR. http://proceedings.mlr.press/v124/duck-mayr20a.html.

Epstein, Lee, and Jack Knight. 1998. *The Choices Justices Make.* CQ Press.

Epstein, Lee, Christopher M. Parker, and Jeffrey A. Segal. 2018. "Do Justices Defend the Speech They Hate? An Analysis of In-Group Bias on the U.S. Supreme Court." *Journal of Law and Courts* 6 (2): 237–262.

Gibson, James L., and Gregory A. Caldeira. 2011. "Has Legal Realism Damaged the Legitimacy of the U.S. Supreme Court?" *Law & Society Review* 45 (1): 195–219.

Gill, Jeff. 2021. "Measuring Constituency Ideology Using Bayesian Universal Kriging." *State Politics & Policy Quarterly* 21 (1): 80–107.

Hainmueller, Jens, and Chad Hazlett. 2014. "Kernel Regularized Least Squares: Reducing Misspecification Bias With a Flexible and Interpretable Machine Learning Approach." *Political Analysis* 22 (2): 143–168.

Hansford, Thomas G., and James F. Spriggs II. 2006. *The Politics of Precedent.* Princteon University Press.

Hinkle, Rachael K. 2015. "Legal Constraint in the US Courts of Appeals." *Journal of Politics* 77 (3): 721–735.

Holmes, Oliver Wendell, Jr. 1897. "The Path of the Law." *Harvard Law Review* 10:457–478.

Johnson, Tyler, and Erica Socker. 2012. "Actions, Factions, and Interactions: Newsworthy Influences on Supreme Court Coverage." *Social Science Quarterly* 93, no. 2 (434–463 2012).

Kastellec, Jonathan P. 2010. "The Statistical Analysis of Judicial Decisions and Legal Rules With Classification Trees." *Journal of Empirical Legal Studies* 7 (2): 202–230.

Klein, David. 2017. "Law in Judicial Decision-Making." Chap. 12 in *The Oxford Handbook of U.S. Judicial Behavior,* edited by Lee Epstein and Stefanie A. Lindquist. Oxford University Press.

Kornhauser, Lewis A. 1992. "Modeling Collegial Courts II. Legal Doctrine." *Journal of Law, Economics, & Organization* 8 (3): 441–470.

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

Lax, Jeffrey R. 2007. "Constructing Legal Rules on Appellate Courts." *American Political Science Review.*

Lax, Jeffrey R. 2011. "The New Judicial Politics of Legal Doctrine." *Annual Review of Political Science* 14:131–157.

Lax, Jeffrey R., and Kelly T. Rader. 2010a. "Legal Constraints on Supreme Court Decision Making: Do Jurisprudential Regimes Exist?" *Journal of Politics* 72 (2): 273–284.

———. 2010b. "The Three Prongs of a Jurisprudential Regimes Test: A Response to Kritzer and Richards." *Journal of Politics* 72 (2): 289–291.

Levi, Edward. 1949. *An Introduction to Legal Reasoning.* University of Chicago Press.

Liang, Kung-Yee, and Scott L. Zeger. 1986. "Longitudinal Data Analysis Using Generalized Linear Models." *Biometrika* 73 (1): 13–22.

Martin, Andrew D., and Kevin M. Quinn. 2002. "Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the U.S. Supreme Court." *Political Analysis* 10 (2): 134–153.

———. 2005. "Can Ideal Point Estimates Be Used as Explanatory Variables?" Working paper, October 8, 2005. https://mqscores.lsa.umich.edu.

———. 2020. *Martin-Quinn Scores.* Version 2019 Term Release, July 12, 2020. https://mqscores.lsa.umich.edu/measures.php.

Poole, Keith T., and Howard Rosenthal. 1985. "A Spatial Model for Legislative Roll Call Analysis." *American Journal of Political Science* 29 (2): 357–384.

R Core Team. 2021. *R: A Language and Environment for Statistical Computing.* Version 4.1.0. R Foundation for Statistical Computing. https://www.R-project.org/.

Rasmussen, Carl Edward, and Christopher K. I. Williams. 2006. *Gaussian Processes for Machine Learning.* MIT Press.

Richards, Mark J., and Herbert M. Kritzer. 2002. "Jurisprudential Regimes in Supreme Court Decision Making." *American Political Science Review* 96 (2): 305–320.

Robins, James M., Miguel Ángel Hernán, and Babette Brumback. 2000. "Marginal Structural Models and Causal Inference in Epidemiology." *Epidemiology* 11 (5): 550–560.

Segal, Jeffrey A. 1984. "Predicting Supreme Court Cases Probabilistically: The Search and Seizure Cases, 1962–1981." *American Political Science Review* 78 (4): 891–900.

Segal, Jeffrey A., and Harold J. Spaeth. 1996. "The Influence of Stare Decisis on the Votes of United States Supreme Court Justices." *American Journal of Political Science* 40 (4): 971–1003.

———. 2002. *The Supreme Court and the Attitudinal Model Revisited.* Cambridge University Press.

Spaeth, Harold J., Lee Epstein, Andrew D. Martin, Jeffrey A. Segal, Theodore J. Ruger, and Sara C. Benesh. 2020. *2020 Supreme Court Database.* Version 2020 Release 01. http://supremecourtdatabase.org.

Torres, Michelle. 2020. "Estimating Controlled Direct Effects Through Marginal Structural Models." *Political Science Research and Methods* 8 (3): 391–408.

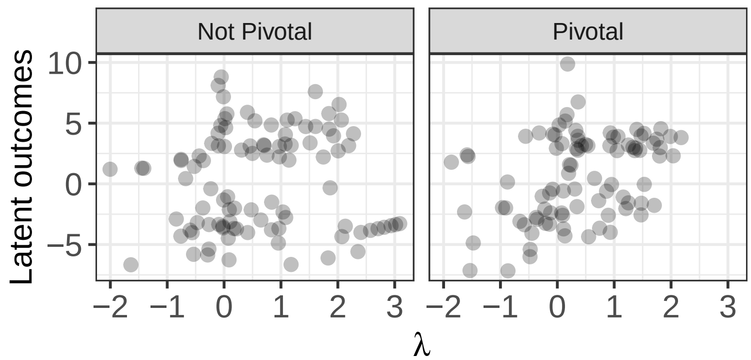
Appendix

A Full results of main model

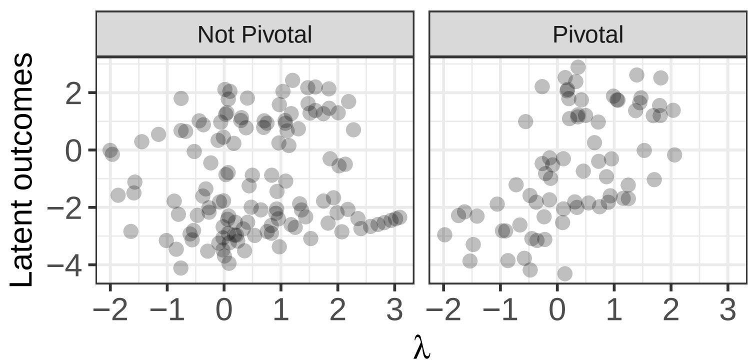
The main text of the paper presents results of the average marginal effect of λ, or the law. For those interested in the average marginal effect of other predictors, I provide the average marginal effect of every variable in the models in Table A.1. The average marginal effect here is on the link, or latent, scale rather than on the probability scale; that is, it provides $\partial x / \partial f$ rather than $\partial x / \partial \sigma(f)$. Of note is that the results for λ are consonant with the probability scale results in the main paper, with the exception that Justice Souter has a reliable positive average marginal effect of λ on the latent scale even though the result is not reliable at a 95% level on the probability scale. Looking at the marginal effects of other predictors, we see that most justices are more likely to vote liberally in cases with content-based restrictions, which we may interpret as in line with the prior literature (see Bartels 2009; Bartels and O'Geen 2015; Richards and Kritzer 2002). Other case factors appear to have more nuanced effects, and may be highly conditional on the values of other variables. A reader may also want to see breakdowns like Figure 4 between cases where a justice is pivotal or not pivotal for justices other than Ginsburg; I provide such figures for each justice analyzed in the main paper in Figure A.1.

Table A.1: Average marginal effects for all predictors in the justice-level models. For each justice, I report the estimate and 95% confidence interval for the average marginal effect on the link scale. For the continuous variables ‘MQ Score’, ‘Term’, ‘λ’, and ‘ρ’, this is the instantaneous rate of change. For the categorical variables ‘Category’, ‘Government’, ‘Action’, and ‘Speaker Identity’, this is a discrete difference against a reference category (listed in the table).

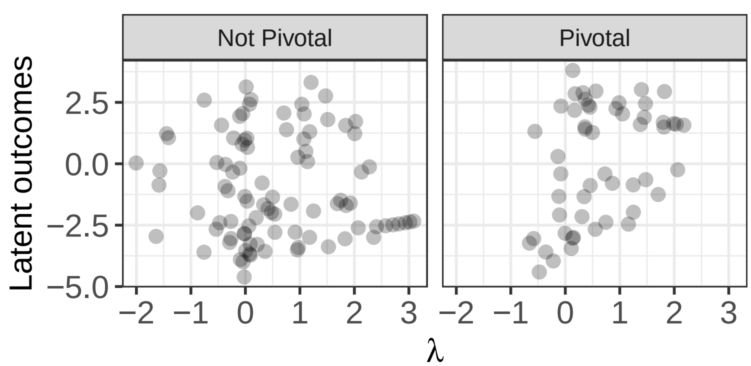
	Breyer	Ginsburg	Kennedy	O’Connor	Rehnquist	Scalia	Souter	Stevens	Thomas
λ	0.20	-0.54	0.76	0.18	0.16	0.34	1.09	0.15	0.12
	[0.00, 0.40]	[-0.91, -0.17]	[0.57, 0.94]	[0.07, 0.29]	[0.03, 0.29]	[0.14, 0.53]	[0.13, 2.05]	[0.03, 0.27]	[-0.06, 0.31]
ρ	0.38	0.20	-0.64	0.02	-0.09	-0.36	-1.14	0.13	-0.22
	[0.19, 0.58]	[-0.25, 0.65]	[-0.85, -0.43]	[-0.09, 0.12]	[-0.21, 0.03]	[-0.55, -0.16]	[-2.47, 0.18]	[0.01, 0.25]	[-0.40, -0.04]
Term	0.00	-0.07	-0.04	-0.00	0.01	0.03	-0.05	-0.00	0.04
	[-0.25, 0.25]	[-0.66, 0.52]	[-2.42, 2.33]	[-0.16, 0.16]	[-0.23, 0.24]	[-0.46, 0.51]	[-22.82, 22.71]	[-0.16, 0.15]	[-0.87, 0.94]
MQ Score	0.12	0.22	-2.75	-0.47	0.02	-0.15	0.45	0.08	0.09
	[-0.15, 0.39]	[-0.57, 1.00]	[-4.23, -1.28]	[-0.63, -0.32]	[-0.21, 0.24]	[-0.71, 0.41]	[-8.34, 9.23]	[-0.09, 0.26]	[-0.60, 0.77]
Category (baseline is “Less protected”)									
Content based	1.49	1.67	4.08	0.76	0.85	1.29	-1.08	-0.12	1.88
	[1.06, 1.91]	[0.65, 2.68]	[2.57, 5.60]	[0.51, 1.01]	[0.52, 1.18]	[0.56, 2.01]	[-14.24, 12.07]	[-0.39, 0.16]	[1.05, 2.71]
Content neutral	1.14	1.10	-0.84	0.08	-0.19	-0.56	-1.03	-0.17	0.08
	[0.70, 1.59]	[-0.02, 2.22]	[-2.61, 0.94]	[-0.18, 0.34]	[-0.54, 0.17]	[-1.40, 0.27]	[-14.40, 12.35]	[-0.45, 0.12]	[-0.79, 0.96]
Government (baseline is “Other”)									
Education	-0.15	-0.06	-2.19	-0.04	0.54	-0.27	1.17	-0.01	-0.26
	[-0.62, 0.32]	[-1.33, 1.21]	[-4.07, -0.31]	[-0.32, 0.24]	[0.14, 0.94]	[-1.14, 0.61]	[-14.26, 16.59]	[-0.32, 0.30]	[-1.28, 0.77]
Federal	-0.10	-0.49	-1.24	0.76	0.77	1.09	-2.78	-0.48	2.18
	[-0.55, 0.36]	[-1.64, 0.65]	[-2.92, 0.43]	[0.49, 1.03]	[0.41, 1.14]	[0.29, 1.90]	[-17.40, 11.84]	[-0.78, -0.19]	[1.23, 3.12]
Local	0.01	-0.18	-0.56	0.29	-0.39	-0.40	-0.56	0.10	-0.62
	[-0.46, 0.48]	[-1.44, 1.08]	[-2.34, 1.22]	[0.01, 0.57]	[-0.78, 0.00]	[-1.27, 0.46]	[-16.76, 15.64]	[-0.21, 0.41]	[-1.60, 0.36]
Private	0.12	0.04	-1.75	-0.09	-0.23	0.28	-1.56	-0.34	0.04
	[-0.36, 0.59]	[-1.23, 1.31]	[-3.61, 0.11]	[-0.36, 0.19]	[-0.63, 0.16]	[-0.60, 1.15]	[-17.61, 14.48]	[-0.66, -0.03]	[-0.99, 1.06]
State	-0.60	-0.83	1.38	0.29	0.27	0.66	-4.83	-0.14	1.25
	[-1.06, -0.14]	[-2.01, 0.35]	[-0.34, 3.10]	[0.02, 0.57]	[-0.11, 0.64]	[-0.17, 1.48]	[-19.26, 9.60]	[-0.45, 0.16]	[0.28, 2.22]
Action (baseline is “Civil suit”)									
Criminal	0.11	-0.54	1.21	0.15	0.00	0.46	-0.43	1.05	0.87
	[-0.35, 0.58]	[-1.89, 0.81]	[-0.66, 3.09]	[-0.12, 0.43]	[-0.37, 0.38]	[-0.41, 1.33]	[-14.45, 13.59]	[0.75, 1.35]	[-0.13, 1.86]
Deny benefit	-0.33	-0.16	-1.09	-0.47	-0.74	-1.85	-4.68	0.12	-0.03
	[-0.80, 0.13]	[-1.57, 1.24]	[-3.13, 0.96]	[-0.75, -0.20]	[-1.13, -0.36]	[-2.79, -0.90]	[-19.72, 10.35]	[-0.19, 0.44]	[-1.06, 1.00]
Deny expression	-1.18	-1.61	1.65	0.23	0.24	0.51	-5.56	0.47	2.29
	[-1.63, -0.74]	[-2.79, -0.44]	[-0.19, 3.49]	[-0.04, 0.49]	[-0.13, 0.60]	[-0.33, 1.35]	[-19.75, 8.62]	[0.17, 0.77]	[1.38, 3.19]
Disciplinary	-0.66	-1.02	0.76	-0.14	-0.28	-0.54	-3.98	0.93	1.30
	[-1.13, -0.19]	[-2.45, 0.41]	[-1.30, 2.81]	[-0.41, 0.13]	[-0.67, 0.11]	[-1.50, 0.42]	[-19.02, 11.07]	[0.62, 1.24]	[0.28, 2.33]
Lose employment	-0.47	-1.14	-1.18	-0.60	-0.45	-1.27	-1.96	0.76	-0.38
	[-0.94, -0.00]	[-2.56, 0.27]	[-3.29, 0.94]	[-0.87, -0.33]	[-0.84, -0.06]	[-2.22, -0.31]	[-16.53, 12.62]	[0.44, 1.07]	[-1.40, 0.64]
Regulation	-0.64	-1.06	-0.69	-0.96	-0.04	-0.94	-1.57	0.33	0.33
	[-1.10, -0.19]	[-2.41, 0.28]	[-2.69, 1.31]	[-1.24, -0.69]	[-0.43, 0.35]	[-1.88, 0.01]	[-15.59, 12.45]	[0.02, 0.65]	[-0.67, 1.32]
Identity (baseline is “Other”)									
Alleged communist				0.60	0.66			-0.08	
				[0.34, 0.87]	[0.29, 1.02]			[-0.37, 0.21]	
Broadcast media	-0.34	0.39	-3.15	1.01	1.05	0.71	-1.33	-0.11	0.70
	[-0.79, 0.11]	[-0.74, 1.52]	[-4.90, -1.41]	[0.75, 1.28]	[0.69, 1.41]	[-0.10, 1.53]	[-17.31, 14.66]	[-0.40, 0.18]	[-0.20, 1.61]
Business	-1.24	-0.45	2.92	0.76	0.91	2.00	0.91	-0.23	2.35
	[-1.67, -0.80]	[-1.52, 0.62]	[1.37, 4.46]	[0.50, 1.02]	[0.56, 1.26]	[1.25, 2.75]	[-11.98, 13.80]	[-0.51, 0.06]	[1.49, 3.20]
Military protester		0.60	-0.77	0.70	0.47	0.87	2.68	0.11	1.10
		[-0.54, 1.74]	[-2.54, 1.00]	[0.43, 0.97]	[0.11, 0.84]	[0.05, 1.70]	[-12.66, 18.02]	[-0.18, 0.40]	[0.19, 2.01]
Politician	-0.25	0.75	-3.68	1.04	0.77	-0.07	3.95	0.03	-1.04
	[-0.69, 0.19]	[-0.35, 1.84]	[-5.34, -2.01]	[0.77, 1.30]	[0.41, 1.14]	[-0.87, 0.73]	[-10.93, 18.83]	[-0.27, 0.32]	[-1.93, -0.14]
Print media			-2.35	0.72	0.31	0.50	0.98	-0.13	
			[-4.06, -0.63]	[0.45, 0.98]	[-0.05, 0.67]	[-0.30, 1.31]	[-14.84, 16.79]	[-0.42, 0.16]	
Racial minority			-1.30	0.68	0.75	0.16	0.90	0.23	-0.07
			[-3.05, 0.45]	[0.41, 0.94]	[0.38, 1.11]	[-0.66, 0.98]	[-14.87, 16.68]	[-0.07, 0.52]	[-0.98, 0.83]
Religious	-0.31	1.15	-1.52	0.96	0.71	0.89	6.16	0.36	0.48
	[-0.75, 0.14]	[0.02, 2.28]	[-3.22, 0.19]	[0.70, 1.23]	[0.35, 1.07]	[0.08, 1.70]	[-7.88, 20.21]	[0.07, 0.65]	[-0.40, 1.37]



(a) Kennedy

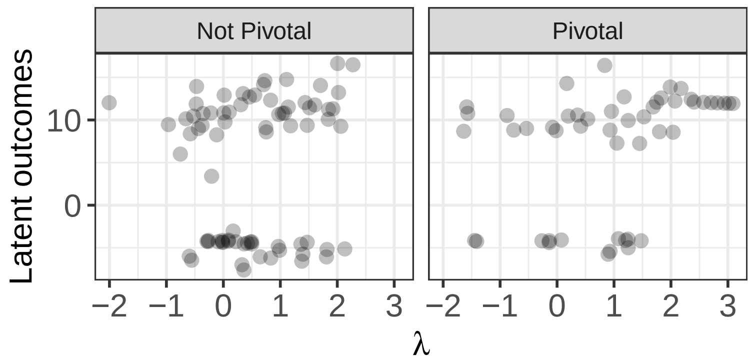


(b) Scalia

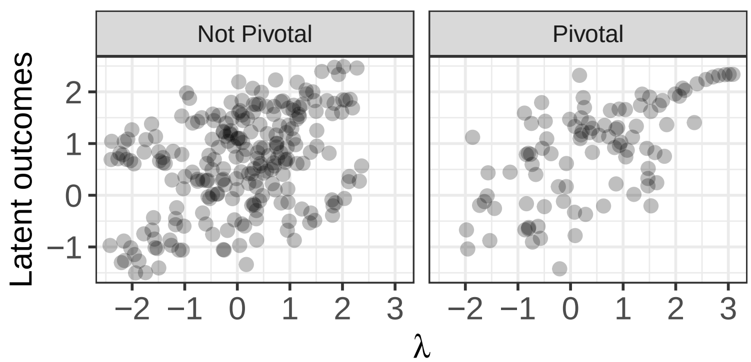


(c) Thomas

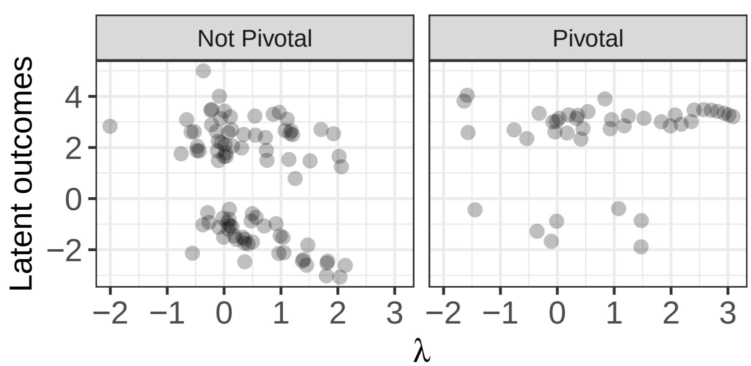
Figure A.1: Comparing predicted outcomes for each justice against λ.



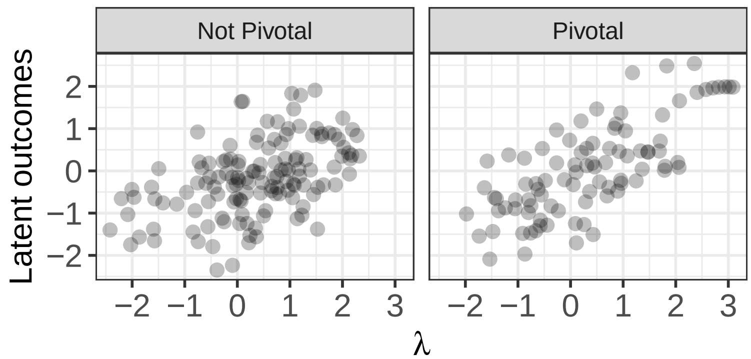
(d) Souter



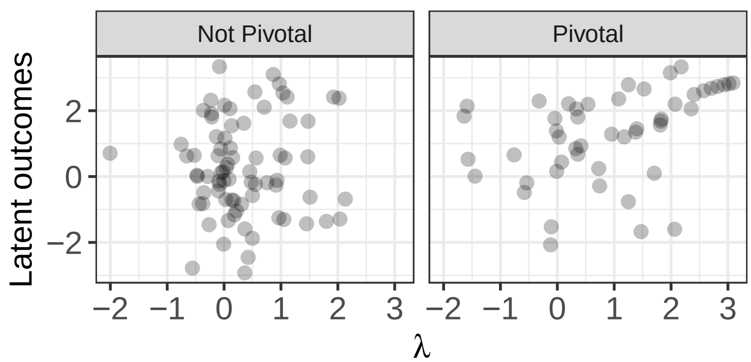
(e) Stevens



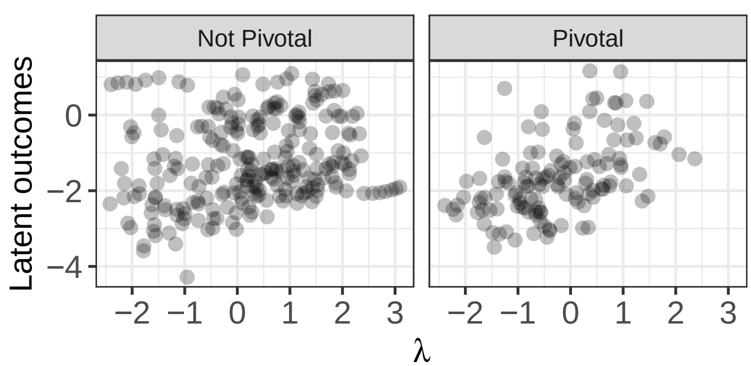
(f) Ginsburg



(g) O'Connor



(h) Breyer



(i) Rehnquist

B Marginal effect of law for additional justices

In the main paper I focus on results for the last natural court of the Rehnquist court to ensure that

1. a sufficient number of observations are used as purely training data for the court-level model to ensure the quality of the λ measure used in the justice-level models;
2. we are using every case decided by each justice we're analyzing; and
3. we have a sufficient number of observations for each justice we're analyzing to ensure the quality of estimates.

In this appendix I provide supplemental results for additional justices. However, I do not analyze the Trump appointees, who each have observations in the single digits in this data. Additionally, even the Bush and Obama appointees all have less than 50 observations, so we may want to see more data from them as well (all justices whose results are presented in the main paper have over 100 observations), since GP classification can be a somewhat data-hungry procedure (see Duck-Mayr, Garnett, and Montgomery 2020). The number of observations in the data for each justice is given in Table B.1.

I present the average marginal effect of λ for the justices who decided any cases outside of the training data that were not presented in the main paper (except for the Trump appointees) in Table B.2. These results are substantively similar to the results for the natural court studied in the main paper, with a majority of justices exhibiting a reliably positive average marginal effect of law, while others are more or less unconstrained.

Table B.1: Number of complete observations for each justice.

Justice	N
Kavanaugh	6
Gorsuch	8
Kagan	25
Sotomayor	31
Alito	46
Roberts	49
Douglas	67
Breyer	125
Souter	126
Ginsburg	130
Thomas	143
Stewart	156
Kennedy	185
Scalia	198
Powell	218
Burger	219
O'Connor	225
Brennan	262
Marshall	279
Blackmun	302
White	302
Stevens	327
Rehnquist	375

Table B.2: Average marginal effect of λ on judges' decisions. For each justice, I report the estimate and 95% confidence interval for the difference in the probability the justice will vote liberally between two different values of λ, averaged over all observations in the sample.

	Mean of $\lambda \pm$ sd of λ	Range of λ
Alito	-0.03	-0.08
	[-0.15, 0.07]	[-0.21, 0.04]
Blackmun	0.12	0.29
	[0.10, 0.15]	[0.26, 0.31]
Brennan	0.02	0.04
	[-0.08, 0.12]	[-0.06, 0.15]
Burger	0.11	0.27
	[0.09, 0.14]	[0.25, 0.29]
Douglas	0.11	0.25
	[-0.05, 0.29]	[0.08, 0.44]
Kagan	-0.50	-0.94
	[-0.55, -0.45]	[-0.95, -0.92]
Marshall	0.02	0.05
	[-0.10, 0.13]	[-0.07, 0.16]
Powell	0.26	0.58
	[0.23, 0.30]	[0.55, 0.61]
Roberts	-0.28	-0.57
	[-0.29, -0.26]	[-0.59, -0.56]
Sotomayor	0.27	0.57
	[0.25, 0.29]	[0.55, 0.58]
Stewart	0.11	0.26
	[0.07, 0.16]	[0.21, 0.30]
White	0.13	0.30
	[0.08, 0.18]	[0.25, 0.35]

C Robustness check: In-group bias

Epstein, Parker, and Segal (2018) find that justices may offer greater protection to speech they agree with. We may worry this feeds into ϕ and thus may bias the γ estimate unless we control for whether the speech at issue is liberal or conservative. I re-run the analysis with a variable coding whether the speech at issue in a case is liberal (for example, obscenity), conservative (for example, commercial or religious speech), or neutral (such as where, for example, campaign spending of both Republicans and Democrats is implicated), as well as other variables from Epstein, Parker, and Segal (2018) such as whether the law at issue is conservative (e.g. anti-obscenity laws), liberal (e.g. a law criminalizing depiction of animal cruelty), or neutral, whether the challenge is an "as applied" or "facial" challenge, and the type of expression (spoken, written, other expression, or association) at issue. As the Epstein, Parker, and Segal (2018) data covers the 1953-2014 terms, I update and backdate that data, as well as filling in some missingness from cases that were included in the Richards and Kritzer (2002) data but not in the Epstein, Parker, and Segal (2018) data.

The results are largely substantively similar. The main differences to note are that our estimate of the average marginal effect of law for Kennedy is now essentially zero and not reliably positive or negative, and the result for Rehnquist is weakened. However, a majority of the Court still exhibits a reliable effect of the law on their decision making.

Table C.1: Average marginal effect of λ on judges' decisions. For each justice, I report the estimate and 95% confidence interval for the difference in the probability the justice will vote liberally between two different values of λ, averaged over all observations in the sample.

	Mean of $\lambda \pm$ sd of λ	Range of λ
Breyer	0.04	0.10
	[-0.01, 0.10]	[0.05, 0.16]
Ginsburg	-0.15	-0.34
	[-0.22, -0.09]	[-0.41, -0.27]
Kennedy	0.00	0.00
	[-0.10, 0.11]	[-0.11, 0.12]
O'Connor	0.05	0.12
	[0.01, 0.08]	[0.08, 0.16]
Rehnquist	0.02	0.02
	[-0.03, 0.06]	[-0.02, 0.07]
Scalia	0.04	0.10
	[-0.03, 0.11]	[0.01, 0.18]
Souter	0.12	0.28
	[0.04, 0.20]	[0.19, 0.37]
Stevens	0.05	0.12
	[-0.00, 0.10]	[0.07, 0.18]
Thomas	-0.08	-0.18
	[-0.16, 0.01]	[-0.27, -0.09]