# Judicial Review and Democratic Failure *

Justin Fox[†] Matthew C. Stephenson[‡]

February 6, 2010

**Abstract**

We use an agency model to analyze the impact of judicial review on democratic performance. We find that judicial review may increase "democratic failure" by rescuing elected officials from the consequences of ill-advised policies, but may also decrease democratic failure by alerting voters to unjustified government action. We further find that judges will defer to the decision of elected leaders unless the level of democratic failure is sufficiently high. We then show how judicial review affects voter welfare, both through its effect on policy choice and through its effect on the efficacy of the electoral process in selecting leaders. We also analyze how the desirability of judicial review is affected by characteristics of the leaders and the judges. Our welfare analysis establishes general conditions under which judicial review serves majoritarian interests—and thereby arguably increases the "democratic" character of political outcomes, despite the non-democratic nature of judicial review itself.

*We are grateful to Gabriella Blum, Rachel Brewster, Glenn Cohen, Barry Friedman, Jim Greiner, Leslie Johns, Dimitri Landa, Adriaan Lanni, Gilat Levy, Andrew Martin, Judkins Mathews, Nolan McCarty, Minor Myers, Ben Roin, Frances Rosenbluth, Ben Sachs, Ken Shepsle, Kenneth Shotts, Jed Shugerman, Jeffrey Staton, Bill Stuntz, Adrian Vermeule, Abby Wood, Kathy Zeiler, Aaron Zelinksy and seminar participants at Notre Dame, MIT, Harvard Law School, and Yale University for helpful comments and conversations.

†Assistant Professor, Department of Political Science, Yale University, ISPS, P.O. Box 208209, New Haven, CT 06520. Email: justin.fox@yale.edu

‡Assistant Professor, Harvard Law School, Griswold 509, Cambridge, MA 02138. Email: mstephen@law.harvard.edu

What is the appropriate role for judicial review in a democracy? When should independent judges be allowed to strike down the decisions of elected legislatures or executives? This question is of enormous practical and theoretical interest. It has been a central focus—some might say an obsession—of American constitutional theory (Friedman 2002; Tribe 2000), and has assumed increasing salience internationally as the power and influence of courts around the world has grown (Hirschl 2004; Tate and Vallinder 1997). Many have defended judicial review as a way to reduce or correct predictable, systematic failures in legislative and executive decision-making—thereby reducing the divergence between actual policy choices and those that would prevail in an ideally-functioning representative democracy. Appropriately designed judicial review, on this view, can be justified on democratic grounds, even if judicial review is not itself a democratic institution. Critics, however, have argued that judicial review tends to exacerbate rather than ameliorate democratic failures, and that the costs of judicial review typically exceed whatever benefits it may have.

This paper uses a formal model to analyze the impact of judicial review on democratic failure. The analysis elucidates the scope and limits of many of the positive theoretical claims that appear in debates about judicial review, and generates several additional implications. The paper is organized as follows. In Part I, we situate our contribution by presenting a brief overview of the contemporary debate about judicial review. In Part II, we develop a simple political agency model, without judicial review. Analysis of this baseline model allows us to isolate a particular type of democratic failure: the incentive that elections create for less-capable incumbents to take bold action, in order to appear more competent than they really are. In Part III, we modify the baseline model by introducing judicial review, and we investigate how this institutional change affects the incidence and impact of the sort of democratic failure we identified in Part II. We show that judicial review may have two effects on democratic failure: *First*, judicial review may rescue elected officials from the consequences of ill-advised policies, and this "bailout effect" tends to increase democratic failure; *second*, judicial approval or disapproval of a policy may affect public opinion of the government that enacted it, and this "legitimation effect" usually tends to decrease democratic failure. We next show that a rational judge's review strategy depends on the level of democratic failure. If democratic failure is sufficiently rare, the judge would rationally defer to the elected leader, while if democratic failure on some issue is sufficiently high, the judge would flatly prohibit government action in that area; judges rely on their own judgment only for "intermediate" levels of democratic failure. We then combine these partial-equilibrium analyses to assess the net impact of judicial review. Part IV turns to normative considerations, focusing on how judicial review affects voter welfare both through its effect on policy choice, and through its effect on the efficacy of the electoral process in selecting good leaders. We also consider how the desirability of judicial review is affected by characteristics of the leaders and the judges. Our results here are sometimes surprising. For example, increasing judicial competence can sometimes make judicial review less socially desirable. More generally, our welfare analysis establishes conditions under which judicial review serves majoritarian interests—and thereby arguably increases the "democratic" character of political outcomes, despite the non-democratic nature of judicial review itself. A brief conclusion summarizes our main findings and suggests directions for future research. A technical appendix contains proofs of all formal propositions in the text.

## I. Judicial Review: The Contemporary Debate

Although no explicit text in the U.S. Constitution grants the judiciary the power to strike down legislative or executive acts as unconstitutional, the power of judicial review was established early on, most famously by Chief Justice Marshall's opinion in *Marbury v. Madison* [5 U.S. (1 Cranch) 837 (1803)].[1] Yet concerns about the legitimacy of judicial review have persisted from the earliest days of the Republic. These concerns intensified as federal courts began to wield their power more aggressively in the twentieth century. The political valence of the debate has changed over time (Chemerinsky 2004; Friedman 2002): In the 1920s and 1930s, progressives castigated the conservative Supreme Court for blocking progressive reforms and economic regulation (including key pieces of President Roosevelt's New Deal), but beginning in the 1950s and 1960s, the political tenor changed, with conservative politicians complaining about liberal "judicial activism," particularly in the areas of civil rights and criminal procedure. More recently, as the federal bench has become increasingly conservative, some progressives have rediscovered their skepticism of judicial review, arguing for greater judicial modesty, or in some cases wholesale reconsideration of the practice.

The debate about judicial review is not all about opportunistic politics, however. Many scholars and jurists have tried to take a longer and broader view. The scholars who did the most to frame

[1]Some revisionist historians, however, have questioned whether *Marbury* was really as significant in establishing strong judicial review as the conventional account would have it (Clinton 1989; Klarman 2001).

the terms of the modern debate about judicial review are probably James Bradley Thayer and Alexander Bickel. Thayer (1893) was deeply skeptical about the legitimacy and wisdom of the strong-form judicial review that Chief Justice Marshall's *Marbury* opinion has come to symbolize. Thayer argued that courts should strike down a legislative act only when the legislation in question is so clearly unconstitutional that no reasonable person could disagree. Otherwise, Thayer argued, the courts would be assuming legislative responsibilities. Writing almost 70 years later, Bickel (1962, pp. 16-17) nicely summed up what he famously described as the counter-majoritarian difficulty with judicial review: "[W]hen the Supreme Court declares unconstitutional a legislative act or the action of an elected executive, it thwarts the will of representatives of the actual people of the here and now; it exercises control, not in behalf of the prevailing majority, but against it." In a polity that is otherwise organized as a representative democracy—and in a political culture that usually celebrates the virtues of majority rule—judicial review may therefore seem like a "deviant institution" (Bickel 1962, p. 18).[2]

Neo-Marshallian defenders of judicial review have advanced a variety of responses to the charge of counter-majoritarianism. One is to deny that majoritarianism is the only, or even a particularly important, normative consideration in assessing judicial review, and to assert moreover that the counter-majoritarian nature of judicial review is something to be celebrated rather than condemned.[3] A second response (in considerable tension with the first) denies that judicial review is in fact significantly counter-majoritarian in practice.[4] A third type of response—the one that we focus on in this paper—seeks to turn the tables on the critics by pointing out the divergence between ideally-functioning democratic institutions and the real-world institutions we actually have. Accord-

[2]Unlike Thayer, Bickel thought that there was room for courts legitimately to exercise their judgment in close cases, even against prevailing political majorities, but he thought courts should do so only occasionally, when such interventions were truly necessary.

[3]The most influential version of this argument asserts that judicial freedom from majoritarian constraints allows courts to serve as "forums of principle" that can promote important non-majoritarian values (Dworkin 1985). Other scholars, influenced by social choice theory, have suggested that "majoritarianism" is often an incoherent normative baseline (Riker 1982; Shepsle 1992), and welfare economists remind us that the median voter's ideal policy may diverge (sometimes substantially) from the policy that would maximize aggregate social welfare (Stiglitz 2000).

[4]This line of argument emphasizes that federal judges, though not directly elected, are appointed by elected officials and are embedded in a complex web of political relationships that ensure a reasonable degree of responsiveness to majoritarian preferences (Friedman 2009). Proponents of this view point to evidence suggesting that judicial decisions tend to track public opinion reasonably closely, at least when considered across a large number of cases over time (Barnum 1993; Dahl 1957; McGuire and Stimson 2004). There is also some evidence that even when courts strike down legislation, they often do so with the implicit blessing of current electoral majorities (Powe 2000; Tushnet 2006; Whittington 2005). A related but distinct argument as to why judicial review is not truly counter-majoritarian emphasizes that judicial rulings are rarely the last word on contested policy questions, but rather provoke subsequent debate among elected representatives and the general public (Friedman 1993; Seidman 2001).

ing to this line of argument, the elected branches are prone to systematic and predictable forms of "democratic failure"—defined here as a divergence between actual outcomes and the outcomes that would obtain if elected officials were perfect agents of the citizens—which appropriately-designed judicial review can correct or reduce. Thus judicial review, even if non-democratic when considered in isolation, can enhance rather than undermine the democratic performance of the political system overall.[5]

Neo-Marshallians have identified a wide variety of democratic failures that might justify judicial intervention. *First*, incumbent officials might use their power to entrench themselves, perhaps by interfering with rights of political speech or organization, or by manipulating electoral rules. Judicial review might be justified as a means for policing these anti-democratic "lock-ups" of the political process (Ely 1980; Issacharoff and Pildes 1998). *Second*, irrational prejudice might lead to the systematic exclusion of certain "discrete and insular minorities" from the ordinary process of legislative bargaining and coalition formation, which might in turn justify enhanced judicial scrutiny of government actions that burden those minorities (Ely 1980; Rogers 1999).[6] *Third*, parochial special interest groups might "capture" the government decision-making process, manipulating it to deliver rents at the expense of more diffuse majorities; judicial review may be a way to guard against this danger (Chemerinsky 1989; Sunstein 1984, 1985). *Fourth*, the government may exhibit a systematic bias in favor of too much legislation (either in general or in a particular domain), and the addition of a judicial veto that makes legislation more difficult or costly to enact can therefore enhance majority welfare (Cross 2000; Fallon 2008; Rogers and Vanberg 2007). *Fifth*, short-term electoral pressures may induce political leaders to advocate policies that are temporarily popular, or that tend to burnish the incumbent leaders' reputation, even when the leaders know that these policies are not in the long-term interest of the citizenry (Eisgruber 2001; Maskin and Tirole 2004). Counter-majoritarian judicial review might alleviate this problem by weeding out some of these misguided policies.

Scholars have reasonably questioned the existence and severity of these and other forms of democratic failure (e.g. Ackerman 1985; Persily 2002; Tushnet 1999). Even more importantly, neo-

[5]This is a form of institutional "second-best" argument (Coram 1996; Vermeule 2003).

[6]The "discrete and insular minorities" terminology, as well as this approach to judicial review, grows out of a footnote in the Supreme Court's 1938 decision in *United States v. Carolene Products Co.* [304 U.S. 144], which suggested that it might be appropriate for courts to review legislation that burdens discrete and insular minorities, such as religious or racial minorities, more aggressively than ordinary legislation.

Thayerian critics of judicial review have pointed out that the existence of a democratic failure does not necessarily justify judicial review, any more than the existence of a market failure necessarily justifies government regulation: Even if the alleged democratic failures exist, judicial review might make things worse, not better (Elhauge 1991; Vermeule 2006; Waldron 2006). There are three principal reasons for this. The first concerns *judicial bias*: Judges are more likely than elected politicians to have the "wrong" policy preferences, and may use their review power in ways that are adverse to most voters' interests. The second concern emphasizes *judicial incompetence*: Even if judges are well-motivated, courts lack the capacity to evaluate the moral and empirical questions at stake in any hard case. This means that whatever benefits may flow from correct judicial decisions are likely to be outweighed by the costs of erroneous decisions. Third, following a concern raised by Thayer, some critics warn of a kind of *judicial overhang* (Tushnet 1999) or "moral hazard" (Rogers 2009; Vermeule 2006): Judicial review may cause legislatures to pay less attention to certain issues (such as the impact of a proposed statute on constitutionally-protected values), because the legislators rely on the courts to address such concerns. If coupled with judicial bias or incompetence, this judicial overhang could make democratic failure problems worse.

So, we confront a difficult question: When will judicial review serve majoritarian interests, and when will judicial review harm those interests? We develop a political economy framework for addressing this question.[7] The scale and complexity of the topic mean that our contribution is necessarily limited. Perhaps most importantly, instead of considering the myriad forms of possible democratic failure, we focus on one particular type: the incentive that elections create for an incumbent politician to undertake bold policy gambles in order to appear competent, even when the politician suspects that the gamble is unwise. Also, although we explicitly incorporate the concern about judicial incompetence, and endogenously derive results related to judicial overhang,

[7]Somewhat surprisingly, despite the vast legal literature on judicial review and the counter-majoritarian difficulty, relatively little political economy work engages this problem directly. Most democratic accountability models focus on the relationship between voters and one or more *elected* agents (e.g. Austen-Smith and Banks 1989; Persson, Roland and Tabellini 1997), while most models of judicial review employ a separation-of-powers framework in which the accountability relationship with voters is suppressed (e.g. Ferejohn and Shipan 1990; Tsebelis 2002). Some important work compares decision-making by an elected official to decision-making by an unelected judge (Komesar 1994; Maskin and Tirole 2004), but the question whether one would want a judge *or* a politician to have authority over a given decision is distinct from the question whether one would want a judge to *review* a politician's decision. Similarly, the literature considering why elected officials ever *comply* with the decisions of independent courts (e.g. Rogers 2001; Stephenson 2003) is focused on a different question than the one we address here. Some prior political economy work does focus on how judicial review might affect outcomes in a political agency framework (e.g. Posner 2008; Rogers 1999, 2009; Rogers and Vanberg 2007); our work is thematically closest to, and builds on, this strand in the literature.

we do not incorporate the concern about judicial bias. Thus the agency problems in our analysis arise from sources other than preference divergence on the policy issue. We recognize that these limitations mean that our analysis omits several of the most important issues in debates over judicial review. Limiting our analysis in this way, however, allows us to focus on other important issues without the complexity and loss of clarity inherent in a "model of everything."

## II. Democratic Failure without Judicial Review

### The Baseline Model

In the baseline model, there is a single Voter and a single elected Leader. The Leader must select one of two policies, $a \in \{n, x\}$, where $a = n$ denotes a "normal" policy and $a = x$ denotes an "extraordinary" policy. The policy that the Voter would prefer depends on the underlying state of the world, $\omega \in \{n, x\}$, where $\omega = n$ denotes the "normal" state—that is, the state in which the Voter would prefer the normal policy—and $\omega = x$ denotes an "extraordinary" state, in which the Voter would prefer the extraordinary policy. For simplicity, we assume the Voter receives a policy payoff of 1 if the policy matches the state ($a = \omega$), and receives a policy payoff of 0 otherwise. The prior probability of the normal state is $p > \frac{1}{2}$. This means (without loss of generality) that the Voter would prefer the normal policy *ex ante*.

The Leader's information about the state depends on his type, $t \in \{l, h\}$, which we will refer to as his "ability" or "competence." If the Leader is low-ability (or "incompetent") ($t = l$), then his information is no better than the Voter's; such a Leader knows only the prior probability of the normal state ($p$). The high-ability (or "competent") Leader ($t = h$), by contrast, learns the true state with certainty. The Leader knows his own ability, but the Voter knows only the prior probability $q \in (0, 1)$ that the Leader is competent. Note that because the low-ability Leader knows only the prior $p$ that $\omega = n$ when choosing policy, and $p > \frac{1}{2}$, it follows that the low-ability Leader maximizes the Voter's policy payoff by selecting $a = n$. In contrast, since the high-ability Leader knows the state when choosing policy, he maximizes the Voter's policy payoff by matching policy to the state.

After the Leader selects the policy $a$, the Voter updates her assessment of the probability that the Leader is competent. Denote the Voter's posterior estimate of this probability as $\hat{q}(a)$. This posterior belief can be thought of as the Leader's *reputation*, with higher or lower values of $\hat{q}$ indicating "better" or "worse" reputations, respectively. There is then an election in which the incumbent Leader faces some challenger.[8] All else equal, the Voter's expected utility in future periods (which we do not model explicitly) is an increasing function of the competence of the winning candidate, so the Voter is more likely to reelect an incumbent with a good reputation. Thus, we assume that the probability the Leader wins the election is given by the function $F(\hat{q})$, where $F$ is continuous and strictly increasing in $\hat{q}$.[9] We will refer to $F(\hat{q})$ as the Leader's *electoral strength*. The better the Leader's reputation, the greater his electoral strength.

The Leader shares the Voter's policy preferences; all else equal, the Leader prefers that the policy match the state ($a = \omega$), receiving a policy payoff of 1 when $a = \omega$ and zero otherwise. The Leader, however, also receives a private benefit from holding office (e.g., ego rents, perks, or the ability to influence other policy issues). This private benefit, rather than a difference in policy preferences, creates the agency problem between the Voter and the Leader in our model.[10] Because the Leader cares about both policy and holding office, we write his total utility as $\alpha u + (1 - \alpha)$ if he wins reelection and as $\alpha u$ otherwise, where $\alpha \in [0, 1)$ is the weight that the Leader attaches to policy and $u \in \{0, 1\}$ is the Leader's policy utility.

To summarize, the baseline model is as follows:

1. Nature determines the state of the world $\omega$ and the Leader's underlying ability $t$.

2. *Policymaking Phase*: The Leader, knowing his ability, implements a policy $a \in \{n, x\}$.

3. *Election Phase*: The Voter, knowing the Leader's policy choice, draws an inference about the Leader's ability. An election is then held, in which the Leader's probability of reelection is an increasing function of the probability the Voter assigns to the Leader being of high ability.

It may be useful to consider a stylized example that illustrates our model in a concrete setting. Suppose the Leader is the mayor of a small city who is considering what to do about an

[8]We assume that the Voter does not acquire any additional information about the Leader's competence prior to the election. The Voter does not, for example, observe whether the chosen policy was successful.

[9]One natural interpretation is that the challenger's perceived quality, realized after the incumbent's choice of $a$, is a random variable with cumulative distribution function $F$.

[10]We assume that the Leader's interest in retaining office is independent of his ability. We recognize, however, that this assumption may be in tension with our assumption that the Leader shares the Voter's policy preferences: If the Leader cares about good policy, then one could argue that a low-ability Leader ought to have a weaker interest in reelection than a high-ability Leader. While we acknowledge this possibility, we think it is substantively reasonable to suppose that even a politician who tries to advance the public welfare while in office cares just as intensely about reelection, no matter how competent the challenger appears.

economically-depressed downtown area, and the Voter is the median voter in the local electorate. A private corporation has approached the mayor and expressed interest in acquiring some of the small residential parcels in the depressed area and using them to put up a new commercial facility. The firm requests that the city use its "eminent domain" power to seize the property in question, pay the legally-required compensation to the former owners, and transfer the property to the private firm so it can build its facility.

The median voter believes it is usually a bad idea for the government to use its eminent domain power to transfer property between private parties. That said, the median voter accepts that there are circumstances in which such takings would be appropriate, for example if holdout problems prevent efficient land assembly, or if there are significant positive externalities associated with the proposed new use of the property, or significant negative externalities associated with the current use of the property. But the median voter views these situations as rare. In this example, then, the "normal action" ($a = n$) would be allowing the private real estate market to allocate the property without government intervention, while the "extraordinary action" ($a = x$) would be for the city to use its eminent domain power to transfer the property from the current owners to the firm.

The mayor shares the median voter's views of when using eminent domain to effect private property transfers is appropriate. The mayor might, however, have better information as to whether the market failures that would justify this sort of intervention are actually present. If the mayor (or his administration) is highly capable ($t = h$), he will be able to correctly identify those extraordinary situations in which a government-mandated property transfer is in the public interest ($\omega = x$), and will be able to confidently ascertain when such transfers are unjustified ($\omega = n$). A low-ability mayor ($t = l$), however, will not be able to distinguish these cases. The voters start out with some belief about the probability the mayor is competent ($q$), but do not know for sure. The mayor's decision whether to invoke the eminent domain power will affect the voters' subsequent assessment of the mayor's competence ($\hat{q}$), which in turn affects the mayor's ability to win reelection ($F(\hat{q})$).

### Equilibrium

Our solution concept is Perfect Bayesian Equilibrium. We assume that the high-ability Leader always proposes the policy that matches the state.[11] Hence, to solve for an equilibrium, we solve

[11] This simplifies the exposition. We verify in the appendix that matching policy to the state is consistent with equilibrium behavior for the high-ability Leader.

for the low-ability Leader's equilibrium strategy and the Voter's equilibrium beliefs, where the low-ability Leader's strategy is a probability, $\pi$, of selecting the extraordinary policy, $a = x$. A strategy-belief pair is an *equilibrium* if (1) the low-ability Leader's policy choice (as prescribed by his strategy) maximizes his expected payoff given the Voter's beliefs, and (2) for each policy choice $a$, the Voter's posterior belief that the Leader is high-ability, $\hat{q}(a)$, is derived via Bayes' Rule when possible.

We now characterize equilibrium behavior in the game as follows:

**Proposition 1** *(a) In the absence of judicial review, the equilibrium probability that the low-ability Leader proposes the extraordinary action, denoted $\pi^*_{norev}$, has a uniquely defined value in the $[0, 1-p]$ interval.*

*(b) If $\frac{\alpha}{1-\alpha}(2p-1) \geq F(1) - F\left(\frac{pq}{pq+1-q}\right)$, then $\pi^*_{norev} = 0$; otherwise, $\pi^*_{norev} \in (0, 1-p]$.*

*(c) $\pi^*_{norev}$ is weakly decreasing in $\alpha$, equaling $1-p$ when $\alpha = 0$.*[12]

We can interpret $\pi^*_{norev}$ as a measure of "democratic failure." In an ideally-functioning democracy, in which the Leader always acts as the Voter's faithful agent, a high-ability Leader would match the action to the state, while a low-ability Leader would always select the normal action. In the equilibrium described in Proposition 1, the high-ability Leader behaves as he should, but the low-ability Leader may not. When $\pi^*_{norev} > 0$, the low-ability Leader's electoral interests lead him sometimes to select the extraordinary action despite the absence of sufficient evidence that doing so is in the public interest—indeed, despite the fact that the Leader himself expects a lower policy payoff from the extraordinary action than from the normal action.[13] Thus, $\pi^*_{norev}$ measures the

[12]It is straightforward to show that the effect of $q$ on $\pi^*_{norev}$ is ambiguous.

[13]This baseline result is a variant on Levy's (2004) analysis of "anti-herding" behavior by agents concerned with their reputation for competence (see also Avery and Chevalier 1999; Prendergast and Stole 1996; Trueman 1994). This anti-herding behavior is closely related to another form of democratic failure, usually characterized as "pandering," in which a less-competent leader's reputational interest causes him to select the policy the voters believe *ex ante* is more likely to be correct, even though the leader's own information suggests otherwise (Canes-Wrone, Herron and Shotts 2001; Prat 2005; Prendergast 1993). The principal difference between these two classes of models concerns the quality of the less-competent leader's information. In pandering models, even a less-competent leader gets a sufficiently reliable private signal that, but for electoral incentives, he would prefer to follow his signal. Furthermore, in these models the conditional probability that the low-ability leader gets the wrong signal, given the state, is equal (or nearly equal) for all states. These assumptions together mean that less-competent leaders disproportionately prefer the unpopular policy in the absence of electoral incentives. By contrast, in anti-herding models, the less-competent leader does not have sufficiently strong private information to alter his prior beliefs about which policy is more likely to be correct. Therefore, instead of being disproportionately likely to prefer the unpopular policy, less-competent leaders are (absent electoral incentives) disproportionately likely to prefer the popular policy. As should be clear from the comparison, the basic dynamic underlying both types of models is essentially the same: When

degree to which the Leader's private interest in reelection distorts his equilibrium behavior away from the behavior one would observe in an ideally-functioning representative democracy.[14]

We can illustrate the intuition for this result using our eminent domain example. Suppose the mayor always acts as the median voter's faithful agent. If so, a competent mayor would use the government's takings power to effect a property transfer if, but only if, doing so is justified by extraordinary circumstances (such as significant holdout problems or externalities), and an incompetent mayor would never engage in these sorts of transfers, because he lacks sufficient information to overcome the presumption that such transfers are undesirable. So, if the mayor always acts as a faithful agent, only a competent mayor would ever use the eminent domain power to transfer property to a private firm. That, in turn, means that voters who observe the mayor take this extraordinary action can infer with certainty that the mayor is competent. As a consequence, the electoral benefits to the mayor of using the eminent domain power this way are very high. Indeed, they might be so high that it is not optimal for a low-ability mayor always to abstain from this use of the takings power: If the low-ability mayor deviates by executing the taking, his expected policy payoff will decrease (because he thinks this intervention is unlikely to be justified by market conditions), but his expected electoral payoff will increase (because the voters will incorrectly infer that he is competent); if the latter effect is stronger, the incompetent mayor cannot always act as a faithful agent in equilibrium. Instead, the incompetent mayor will use a mixed strategy, usually allowing the private real estate market to operate without government involvement, but sometimes (with probability $\pi^*_{norev} \in (0, 1-p])$ exercising the eminent domain power to transfer property, despite the absence of sufficient evidence that such extraordinary action is justified by extraordinary circumstances.

This sort of democratic failure hurts the Voter in two ways. First, and most obviously, absent additional information the extraordinary action has a negative expected policy payoff. Second, the distortion in the low-ability Leader's behavior can make the electoral mechanism a less efficient means for the Voter to select a competent Leader, because the incumbent Leader's action provides

less-competent leaders disproportionately prefer one (popular or unpopular) policy, they have an electoral incentive to choose the other (unpopular or popular) policy, lest voters infer from the leader's choice that he is not competent. We therefore expect that the analysis we develop in this paper would be similar in a model where democratic failure took the form of pandering, though we defer consideration of that case to future research.

[14]Notice that the probability the final policy does not match the state is equal to $(1-q)[p\pi^*_{norev} + (1-p)(1-\pi^*_{norev})]$, and since $p > \frac{1}{2}$, this probability is strictly increasing in $\pi^*_{norev}$.

less information to the Voter about the Leader's ability.[15] One solution to the former problem might be to eliminate the electoral constraint—say, by allowing the incumbent Leader to stay in office indefinitely, or by imposing a term limit—or to keep the decision secret until after the election (Fox 2007; Prat 2005). Such approaches, however, would further undermine the capacity of the electoral system to improve the average competence of Leaders over time, and could have other adverse consequences as well. For these reasons, institutional designers might contemplate other mechanisms that preserve the system of selecting the Leader via competitive elections, but reduce or compensate for the democratic failure induced by the low-ability Leader's electoral incentives. Judicial review may be one such mechanism.

## III. Democratic Failure with Judicial Review

### The Modified Game

Now that we have isolated a particular form of democratic failure and understood its causes, we can analyze the impact of judicial review by modifying the baseline model as follows: After the Leader selects action $a$, this action is reviewed by a Judge before it is implemented. The Judge issues a decision $d \in \{uphold, strike\}$; if the Judge upholds the Leader's proposal ($d = uphold$), the Leader's proposed action is implemented, but if the Judge strikes down the Leader's proposal ($d = strike$), the Judge imposes the alternative policy as the final outcome (for instance, if $a = x$ and $d = strike$, then the final policy is $n$). Our analysis incorporates, in stylized fashion, three characteristics of judicial review that are often cited as important distinctions between judicial review and other forms of oversight or institutional control:

*First*, the Judge in our model has the power to strike down the Leader's action, but cannot implement a more refined incentive scheme that offers variable payments or penalties to the Leader that depend on the Leader's proposal. This is consistent with conventional understandings of the nature and limits of judicial power, and captures one of the ways that judicial review of government action is thought to differ from, for example, legislative oversight of a bureaucratic agency, or a firm's supervision of its employees.[16]

[15] To illustrate with an extreme case, if $\pi^*_{norev}$ were equal to $1 - p$, then low-ability and high-ability Leaders would select $a = x$ with equal probabilities, and the Voter would learn nothing at all about the Leader's type from the Leader's policy choice.

[16] A strain in legal scholarship has challenged this assumption, arguing that judicial review can sometimes raise

*Second*, we assume the Judge is insulated both from popular elections and direct interference by the Leader, and that the Judge does not consider how her decisions will affect the election. The Judge in our model simply tries to match the policy to the state. These assumptions are obviously simplifications. There is considerable evidence that real judges are sensitive to public opinion, care about their reputations, and are mindful of the political and electoral repercussions of their decisions (Friedman 2009; Gely and Spiller 1992; Schauer 2000). Nonetheless, these considerations probably matter much less to life-tenured, politically insulated judges than they do to elected politicians. We capture this *relative* difference, in stylized form, by stipulating that the Judge cares *only* about getting the correct answer in the case before her.

*Third*, we incorporate the neo-Thayerian concern about (relative) judicial incompetence by assuming that although the Judge's analysis of a proposed government action conveys some decision-relevant information (i.e., information about the true state), the Judge's information is worse than a competent Leader's, and is never by itself strong enough to overcome the Judge's prior belief that the state is normal. We model this by assuming that after the Leader proposes action $a$, the Judge gets a private signal $s \in \{n, x\}$; the probability that this signal is accurate (i.e., $s = \omega$) is $\gamma \in [\frac{1}{2}, p)$.[17] The $\gamma$ parameter, which is common knowledge, is a measure of the Judge's "ability." The assumption that $\gamma < p$ guarantees that the Judge's signal is never sufficient, on its own, to overcome the prior presumption that the normal action is correct.[18] Also, the Judge, like the Voter, does not know the Leader's true type, but knows only the prior probability, $q$, that the Leader is competent.

We further assume, less realistically, that the Judge's policy preferences are aligned with the

the costs of enacting certain government policies without necessarily prohibiting them (Stephenson 2008; Young 2000). As we show in the working paper version of this article [INSERT CITE], if the Judge in our model were able to implement such an approach, then judicial review could completely eliminate the democratic failure problem. That said, there are a number of potential obstacles—including the problems of credible commitment and accurate calibration—that may limit the applicability of such a strategy. It is thus reasonable to analyze settings where the Judge must make a simple yes-or-no decision whether to uphold the Leader's proposal.

[17]We assume that the Judge's signal is her private information. The working paper version of this article considers how the results change if the Judge can publicly disclose her signal, $s$, in addition to issuing her decision, $d$; we show that if the Judge can do this, judicial review is more likely to benefit the Voter [INSERT CITE]. That said, we think it is reasonable to presume that $s$ remains private information. Although judges do write opinions that may include non-binding observations ("dicta") that do not determine the ultimate holding, actual case dispositions are likely to be more salient for most voters. Also, a Judge who defers to the Leader's proposal might not have an incentive to disclose her contrary signal, as doing so may be mildly costly and does not affect the case outcome. Furthermore, if the Judge must pay some small cost to acquire the signal—if, for instance, she must do some research to assess the merits of the case—she may not bother to do so, especially if she does not expect her ultimate decision to be influenced by the realization of $s$.

[18]Formally, $\gamma < p$ implies that $Pr(\omega = n | s = x) > \frac{1}{2}$. This means that a Judge whose only information beyond her prior ($p$) about the state is a signal of $s = x$ still believes that $\omega = n$ is the the more likely state.

Voter's: The Judge receives a payoff of 1 if the policy matches the state, and a payoff of 0 otherwise. As we noted in Part II, the assumptions that both the Leader and the Judge share the Voter's policy preferences admittedly rule out considerations that are central in other analyses of judicial review. Our model does not speak directly to arguments in favor of judicial review that emphasize the potential for legislative bias, nor to arguments against judicial review that emphasize judicial bias. We acknowledge these limitations, but we are principally interested in costs and benefits of judicial review that arise from other sources. By assuming that the Judge and the Leader, like the Voter, prefer that the policy match the state all else equal, we guarantee that the effects we identify are not due to differences in policy preferences.

To summarize, the modified game is as follows:

1. Nature determines the state of the world $\omega$ and the Leader's underlying ability $t$.

2. *Policymaking Phase*: The Leader, knowing his ability, proposes a policy $a \in \{n, x\}$.

3. *Judicial Review Phase*: The Judge, knowing the Leader's policy proposal $a$, issues a decision $d$, either upholding or striking down the Leader's proposal. If the Judge upholds the proposal, it is implemented; otherwise, the alternative policy is implemented.

4. *Election Phase*: The Voter, knowing the Leader's policy choice and the Judge's decision, draws an inference about the Leader's ability. An election is then held, in which the Leader's probability of reelection is an increasing function of the probability the Voter assigns to the Leader being of high ability.

Before proceeding to the equilibrium analysis, we note that even in the Judge's presence, the Voter's policy payoff continues to be maximized when the high-ability Leader matches policy to the state and the low-ability Leader selects the normal policy. That the low-ability Leader should never select the extraordinary action in our setting with judicial review follows from our assumptions that the low-ability Leader knows only the prior that $\omega = n$ when proposing policy and the Judge's signal of the state is not strong enough on its own to support implementing the extraordinary action.[19]

[19]In other words, our model does not consider scenarios in which it is socially desirable for the low-ability Leader to delegate policymaking authority to the Judge, as would be the case if the low-ability Leader always proposed the extraordinary action, thus letting the Judge determine policy based upon her private signal.

### Equilibrium

As before, we assume the high-ability Leader always proposes the policy that matches the state ($a = \omega$). We also assume that there is no judicial review if the Leader proposes the normal action. That is, only extraordinary proposals are "justiciable." In our takings example, this assumption would mean that if the mayor decides to exercise the eminent domain power to transfer private property, a judge could potentially strike the taking down (for instance, on the grounds that the taking is not for a "public use" within the meaning of the Fifth Amendment of the U.S. Constitution), but if the mayor decides *not* to act, there is no mechanism by which a court could compel the mayor to use the eminent domain power to transfer the property.[20]

To solve for an equilibrium under these assumptions, we must solve for the probability that the low-ability Leader selects the extraordinary action, and for the behavior of the Judge upon observing that the Leader proposed such action. As before, we denote the low-ability Leader's strategy by the probability $\pi$ that he chooses the extraordinary action. We write the Judge's strategy as $(\sigma_n, \sigma_x)$, where $\sigma_n$ is the probability that the Judge upholds the extraordinary action even though her signal favors the normal action ($s = n$), while $\sigma_x$ is the probability that the Judge upholds the extraordinary action when her signal favors the extraordinary action ($s = x$).

We must also specify beliefs. In contrast to the baseline model, the Voter's beliefs about the Leader's type may be affected by the Judge's decision as well as the Leader's action. Therefore, the Voter's posterior estimate of the probability that the Leader is competent is $\hat{q}(a, d)$, rather than simply $\hat{q}(a)$. In addition to specifying beliefs for the Voter, we must also specify beliefs for the Judge. In particular, in deciding whether to uphold the extraordinary action, the Judge must update her prior that the state is normal based upon her signal of the state ($s$) and the fact that the extraordinary action has been proposed. We denote this posterior by $\hat{p}(s)$.

The strategies of the low-ability Leader and the Judge, together with the beliefs of the Judge and the Voter, constitute an *equilibrium* if: (1) the low-ability Leader's policy choice (as prescribed by his strategy) maximizes his expected payoff given the Judge's strategy and the Voter's posterior beliefs about the Leader's competence; (2) for each private signal $s \in \{n, x\}$, the Judge's ruling (as prescribed by her strategy) maximizes her expected payoff given her posterior beliefs about the

[20]This assumption is often empirically plausible (as in our takings example) and simplifies the exposition. We show in the appendix that this assumption is benign, for even if $a = n$ were justiciable, the Judge would always uphold the normal policy provided that the high-ability Leader follows his signal.

state; and (3) beliefs are derived via Bayes' Rule whenever possible.[21]

Our equilibrium analysis proceeds in three stages. *First*, we characterize the equilibrium behavior of the Leader, taking the Judge's strategy as exogenous. This partial-equilibrium analysis allows us to assess claims about the effect of aggressive judicial review on the probability of democratic failure. *Second*, we characterize the equilibrium behavior of the Judge, taking the Leader's strategy as exogenous. This partial-equilibrium analysis clarifies how the rate of democratic failure affects the aggressiveness with which the Judge reviews the Leader's proposals. *Third*, we combine these partial-equilibrium analyses to characterize the equilibria of the judicial review game.

**The Effect of the Judge's Review Strategy on the Level of Democratic Failure**

One of the important questions in debates over judicial review, as discussed in Part I, is whether judicial review tends to increase or decrease democratic failure. While our model does not permit a comprehensive answer to that question, it enables us to ask how judicial review affects the likelihood of the particular type of democratic failure we isolated in Part II. We now consider that question, taking the Judge's strategy as exogenous.

Most extant discussions of the effect of judicial review on democratic failure presume that the reviewing court makes some independent decision, based on the court's own analysis, about whether to uphold the proposed government action. In our model, this would imply a situation in which the Judge follows her signal, upholding the extraordinary policy if $s = x$ but striking it down if $s = n$ (i.e., $\sigma_n = 0$, $\sigma_x = 1$). We label this form of judicial review as *active* review. There are two other pure strategies the Judge might employ, however. First, the Judge might adopt a *passive* approach, upholding the Leader's proposal regardless of the Judge's own signal (i.e., $\sigma_n = \sigma_x = 1$). Second, the Judge might employ a *strict* review strategy, striking down the extraordinary action in all cases (i.e., $\sigma_n = \sigma_x = 0$).[22] We refer to any equilibrium in which the Judge employs an active review strategy as an *active equilibrium*. We define *passive equilibrium* and *strict equilibrium* analogously. The low-ability Leader's equilibrium strategy, conditional on the Judge adopting the

[21] Since the high-ability Leader always matches policy to the state, $x$ is always chosen with positive probability, so $\hat{p}$ can be completely specified via Bayes' Rule. In contrast, it is not always possible to completely specify $\hat{q}$ via Bayes' Rule. For example, if the Judge's strategy calls for her always to uphold the Leader's proposal of $a = x$, then $\hat{q}(x, \textit{strike})$ cannot be derived via Bayes' Rule. That said, the low-ability Leader's incentives are fully determined by on-path beliefs, so the specification of off-path beliefs is inconsequential.

[22] A fourth pure strategy, in which the Judge always does the opposite of her signal, is implausible and never occurs in equilibrium. We defer consideration of mixed strategies until the next subsection.

active strategy, is denoted $\pi^*_{act}$, while $\pi^*_{pass}$ and $\pi^*_{strict}$ denote the Leader's equilibrium strategy given passive and strict review, respectively.[23]

We can now address the question whether judicial review increases or decreases democratic failure, taking the Judge's review strategy as exogenous:

**Proposition 2** *(a)* Passive Review*: In a passive equilibrium, the level of democratic failure is the same as in the baseline model without judicial review (i.e., $\pi^*_{pass} = \pi^*_{norev}$).*

*(b)* Strict Review*: In a strict equilibrium, democratic failure is weakly greater than in the no review case (i.e., $\pi^*_{strict} = 1 - p \geq \pi^*_{norev}$).*

*(c)* Active Review*: In an active equilibrium, the level of democratic failure, $\pi^*_{act}$, is uniquely defined, but the ordering of $\pi^*_{act}$ and $\pi^*_{norev}$ is ambiguous: Active judicial review may increase or decrease democratic failure.*[24]

Part (a) of Proposition 2 is unsurprising: If the Judge is passive, the model with judicial review is functionally equivalent to the baseline model with no review. Part (b) is also straightforward: If the Judge uses a strict review strategy, the final outcome is always the normal policy; because that outcome is a foregone conclusion, the low-ability Leader's policy choice is driven solely by electoral considerations. Thus, in a strict equilibrium, the low-ability Leader selects $a = x$ with the exact same probability as the high-ability Leader $(1-p)$. Part (c) of Proposition 2—concerning the effect of active judicial review on a low-ability Leader's behavior—is more complicated, and also more relevant to debates over the effect of judicial review on democratic failure. The reason that active review may increase or decrease the level of democratic failure is that active review affects the low-ability Leader's incentives through two quite different channels, and these two effects may cut in opposite directions. One of these effects concerns the impact of active judicial review on the likely *policy consequences* if the low-ability Leader proposes the extraordinary action. The other effect concerns the impact of active judicial review on the likely *reputational consequences* for the low-ability Leader who proposes the extraordinary action. Let us consider each in turn.

First, the low-ability Leader's incentive to propose the extraordinary action depends in part on the expected policy consequences of making such a proposal. Because the extraordinary action

[23] In the appendix (Lemma 3), we show that the probabilities $\pi^*_{act}$, $\pi^*_{pass}$, and $\pi^*_{strict}$ are uniquely defined.

[24] It is straightforward to show that $\pi^*_{act}$ is weakly decreasing in $\alpha$, whereas the effects of $\gamma$ and $q$ on $\pi^*_{act}$ are ambiguous.

is probably a bad idea, a low-ability Leader is more likely to propose such action if there is some probability the Judge will strike it down. This would be true even if the Judge struck down proposals at random; the fact that the Judge's signal is somewhat informative strengthens the effect. This "bailout effect" means that active judicial review tends to increase democratic failure all else equal (though it will also correct at least some instances of such failure). This observation is broadly consistent with the Thayerian concern about "judicial overhang"—the fear that judicial review will make elected leaders more reckless, because they can rely on the judiciary to screen out objectionable policies. It is also consistent with the related claim that elected officials are sometimes pleased—perhaps even relieved—when the courts strike down some policy measure that the enacting officials viewed as ill-advised, but felt pressured by electoral interests to propose (Hirschl 2000; Salzberger 1993).

Active review also has a second effect, however, that neo-Thayerians (and others) tend to overlook: As long as the Judge's signal is somewhat informative, active review communicates useful information to the Voter about the true state, which the Voter can use to update her assessment of the Leader's competence.[25] Call this the "legitimation effect" of active judicial review.[26] This effect can dampen the low-ability Leader's electoral incentive to propose the extraordinary action.[27] This is because, holding the low-ability Leader's strategy fixed, the introduction of a Judge who employs an active review strategy reduces the low-ability Leader's expected reputation from proposing $a = x$. To see why active review has this reputational effect, begin by noting that when the Judge overrules the Leader, the Leader's reputation suffers; since the low-ability Leader is more likely to be overruled than the high-ability Leader, the former's expected reputation from proposing $a = x$ is less than the latter's. This fact, taken together with the Martingale property of Bayesian

[25]This is similar to the means by which an informative media's reporting about the state of the world can affect public perceptions of the competence of elected officials, as discussed in Ashworth and Shotts (2008). A crucial difference, however, is that the media in their model *only* communicate a signal about the state, whereas in our model the Judge's decision *simultaneously* communicates information about the state *and* affects the policy outcome directly. This difference reflects an important substantive difference between the media and the judiciary as constraints on elected officials.

[26]While there is little empirical evidence directly on the question of how judicial validation or invalidation of a policy affects public opinion of the officials who enacted it, there is evidence that judicial rulings may affect public evaluations of the policies themselves (Clawson, Kegler and Waltenburg 2001; Hoekstra and Segal 1996). The significance of this effect, however, has been questioned (Baas and Thomas 1984; Marshall 1989). If such an effect does sometimes exist, it is suggestive evidence for the sort of legitimation effect we derive here.

[27]If we assume that the electoral strength function $F$ is concave, then the legitimation effect *will* dampen the low-ability Leader's electoral incentive to propose $a = x$. In contrast, when $F$ is sufficiently convex, the addition of the Judge can increase the low-ability Leader's electoral incentive to propose $a = x$. See the working paper version of this article [INSERT CITE] for further discussion of this subtlety.

posteriors,[28] implies that the low-ability Leader's expected reputation from proposing $a = x$ is lower in the presence of active judicial review than in the case without judicial review.

Thus, active judical review affects the low-ability Leader's incentive to propose $a = x$ via two channels that often work at cross purposes. On the one hand, active review reduces the policy costs to the low-ability Leader form proposing the extraordinary action (the bailout effect). This makes selecting $a = x$ more attractive to the low-ability Leader. On the other hand, active judical review can reduce the electoral benefit of proposing $a = x$ (the legitimation effect), especially when the electoral hit the Leader takes from being overruled is large relative to the electoral boost he receives from being upheld. Thus, depending upon the relative magnitude of the bailout and legitimation effects, active judicial review can increase or decrease democratic failure, relative to the no review baseline.

We conclude this subsection by noting that our discussion so far has considered only the impact of judicial review on the *incidence* of democratic failure, not its potential to *correct* democratic failure (as when the Judge strikes down a low-ability Leader's incorrect decision) nor its potential to obstruct desirable government initiatives (as when the Judge incorrectly strikes down an extraordinary proposal).[29] We take up these factors, in conjunction with the effect of review on the frequency of democratic failure, in Part IV.

**The Effect of the Level of Democratic Failure on the Judge's Review Strategy**

The preceding subsection characterized the behavior of the Leader, taking the Judge's review strategy as fixed. But the Judge in our model is also a strategic actor. We now investigate the conditions that give rise to different forms of judicial behavior, temporarily treating the level of democratic failure as exogenous. Doing so allows us to characterize the Judge's optimal review strategy as follows:

[28] The Martingale property of Bayesian posteriors implies that the Voter's posterior about the Leader's ability upon observing $x$ proposed, $\hat{q}(x)$, must equal the Voter's expected posterior upon observing $x$ proposed *and* the Judge's decision $d$, $\hat{q}(x)E_d[\hat{q}(x,d)|t = h] + (1 - \hat{q}(x))E_d[\hat{q}(x,d)|t = l]$.

[29] Formally, under active review the probability that the Leader incorrectly *proposes* extraordinary action in the normal state is $p(1 - q)\pi^*_{act}$; the probability that the extraordinary policy is incorrectly *implemented* is $(1 - \gamma)p(1 - q)\pi^*_{act}$; and the probability that the active Judge incorrectly *strikes down* an extraordinary policy in the extraordinary state is $(1 - \gamma)(1 - p)(q + (1 - q)\pi^*_{act})$.

**Proposition 3** *Define the following two threshold values:*

$$\underline{T} \equiv \left( \frac{q(1-p)}{1-q} \right) \frac{(1-\gamma)}{p-(1-\gamma)}$$

*and*

$$\overline{T} \equiv \left( \frac{q(1-p)}{1-q} \right) \frac{\gamma}{p-\gamma},$$

*where* $\underline{T} < \overline{T}$.[30] *In an equilibrium in which the low-ability Leader selects* $a = x$ *with probability* $\pi^*$, *the Judge's strategy takes one of five forms—passive, strict, active, semi-active, or semi-strict—as follows:*

*(a) If* $\pi^* < \underline{T}$, *the Judge adopts the passive review strategy, upholding proposal* $a = x$ *even if the Judge's signal is* $s = n$ $(\sigma_n^* = \sigma_x^* = 1)$.

*(b) If* $\pi^* > \overline{T}$, *the Judge adopts the strict review strategy, striking down proposal* $a = x$ *even if the Judge's signal is* $s = x$ $(\sigma_n^* = \sigma_x^* = 0)$.

*(c) If* $\pi^* \in (\underline{T}, \overline{T})$, *then the Judge adopts the active review strategy, upholding proposal* $a = x$ *if* $s = x$, *but striking it down if* $s = n$ $(\sigma_n^* = 0; \sigma_x^* = 1)$.

*(d) If* $\pi^* = \underline{T}$, *the Judge adopts either the passive strategy, the active strategy, or a semi-active strategy in which the Judge always upholds proposal* $a = x$ *if* $s = x$ $(\sigma_x^* = 1)$ *and upholds it with probability* $\sigma_n^* \in (0, 1)$ *if* $s = n$.

*(e) If* $\pi^* = \overline{T}$, *the judge adopts either the active strategy, the strict strategy, or a semi-strict strategy in which the Judge always strikes down proposal* $a = x$ *if* $s = n$ $(\sigma_n^* = 0)$ *and upholds it with probability* $\sigma_x^* \in (0, 1)$ *if* $s = x$.

The threshold values $\underline{T}$ and $\overline{T}$ partition the $[0, 1]$ interval into (at most) three regions (depicted graphically in Figure 1): the $[0, \min\{\underline{T}, 1\}]$ interval; if $\underline{T} < 1$, the $[\underline{T}, \min\{\overline{T}, 1\}]$ interval; and, if $\overline{T} < 1$, the $[\overline{T}, 1]$ interval. The postulated rate of democratic failure $\pi^*$ will fall into one of these regions, or will fall exactly on one of the boundary points ($\underline{T}$ or $\overline{T}$). Proposition 3 establishes that the Judge's optimal review strategy depends on the region into which $\pi^*$ falls: If $\pi^*$ is sufficiently low (below $\underline{T}$), the Judge prefers passive review; if $\pi^*$ is sufficiently high (above $\overline{T}$), the Judge

[30] That $\underline{T} < \overline{T}$ follows from the fact that $\gamma > \frac{1}{2}$.

prefers strict review; if $\pi^*$ falls in the middle range (between $\underline{T}$ and $\overline{T}$), the Judge prefers active review. Thus, $\underline{T}$ is the value of $\pi^*$ at which the Judge is indifferent between passive and active review, while $\overline{T}$ is the value of $\pi^*$ at which the Judge is indifferent between active and strict review. If $\pi^*$ falls exactly on one of these indifference points, the Judge may adopt a mixed strategy as described in parts (d) and (e) of Proposition 3.

**Figure 1 about here.**

In interpreting Proposition 3, it is helpful to recall that the Judge (who cares only about making the correct policy choice) has three useful pieces of information at her disposal when she must issue her decision. *First*, she knows the prior probability of the normal state, $p > \frac{1}{2}$. *Second*, she knows that the Leader proposed the extraordinary policy, which implies either that the Leader was competent and observed the extraordinary state, or that the Leader was incompetent but proposed the extraordinary action anyway (which he does with probability $\pi^*$). *Third*, the Judge observes her own signal, $s$. The Judge uses her signal and the Leader's action to form an updated belief, $\hat{p}$, about the probability that the state is normal. If $\hat{p} > \frac{1}{2}$, the Judge prefers to strike down the extraordinary action; if $\hat{p} < \frac{1}{2}$, the Judge prefers to uphold the extraordinary action; if $\hat{p} = \frac{1}{2}$, the Judge is indifferent.

The fact the Leader proposed the extraordinary action tends to increase the Judge's estimate of the probability that such action is indeed correct, but the strength of the inference the Judge can draw from the Leader's extraordinary proposal is inversely proportional to the level of democratic failure, $\pi^*$. If $\pi^*$ is low enough—if it is less than $\underline{T}$—then the Judge can draw a sufficiently strong inference from the Leader's proposal of $a = x$ that even if the Judge's own signal is $s = n$, the Judge's posterior $\hat{p}$ is still less than $\frac{1}{2}$. In other words, if the low-ability Leader's behavior is sufficiently close to that of a faithful agent, the information contained in the Leader's proposal is greater than the information contained in the prior and the Judge's signal combined. If $\pi^*$ falls into this range, then, the judge rationally defers to the Leader's decision.

If, on the other hand, $\pi^*$ is sufficiently high—greater than $\overline{T}$—the inferences that the Judge can draw from the Leader's proposal of $a = x$, even when combined with a judicial signal of $s = x$, are too weak to overcome the Judge's prior belief that the normal action is correct. Recall that the Judge's signal is never strong enough to overcome the prior, so the Judge would not be willing to uphold an extraordinary proposal without some additional piece of reliable information. This additional information would have to be the information contained in the fact that the Leader made the extraordinary proposal. But if the likelihood of democratic failure is high, the Leader's proposal contains very little information because low-ability Leaders frequently propose extraordinary action. So, for a sufficiently high $\pi^*$, the Judge always strikes down the extraordinary proposal.

Perhaps most interesting is the intermediate case, in which $\pi^*$ lies between $\underline{T}$ and $\overline{T}$. In this case, the information contained in the prior (which favors the normal state) and the information contained in the Leader's extraordinary proposal (which favors the extraordinary state) offset sufficiently that the Judge's own signal becomes decisive. Thus, the Judge uses her signal only when it is rational for her to do so—when her signal, though weak, is determinative in light of the fact that the other pieces of information at her disposal cut in opposite directions.

Although we defer a comprehensive consideration of the welfare effects of judicial review until Part IV, the preceding observations imply one welfare implication that is sufficiently important (both substantively and for developing intuition) that we state it here:

**Remark 1 : Rational Judicial Activism** *The Judge strikes down the extraordinary action only when doing so does not harm her expected policy payoff, which is identical to the Voter's expected policy payoff. Thus, if democratic failure ($\pi^*$) is fixed, judicial review weakly improves Voter welfare relative to the baseline no review case.*

This result is notable in light of a familiar neo-Thayerian objection to judicial review that emphasizes the relative incompetence of judges in evaluating the uncertain empirical and moral questions at stake in hard cases. Translated into the language of our model, this criticism implies that $\gamma$ is low relative to $q$. One might think that for sufficiently low values of $\gamma$, introducing judicial review would harm Voter welfare, because ignorant Judges would strike down policies adopted by more competent elected Leaders. Our analysis, however, demonstrates that low judicial ability is not a sufficient condition for judicial review to have adverse effects on Voter welfare. The reason is that in our model the Judge, like Socrates, is aware of the limits of her knowledge, and she rationally discounts the inferences she draws from her own signal accordingly. Consequently, when the Judge's ability is sufficiently low, she defers to the Leader, making judicial review irrelevant but not harmful.[31]

[31] Although the "judicial incompetence" criticism usually emphasizes that judges lack *information or expertise* (i.e., that their signals are inaccurate), some have hypothesized that judges may also exhibit systematically *irrational*

The preceding discussion shows how the Judge's optimal review strategy depends on the relationship of $\pi^*$ to $\underline{T}$ and $\overline{T}$. We may also be interested in comparative statics on $\underline{T}$ and $\overline{T}$, as changes in these thresholds can affect the Judge's review strategy, even holding $\pi^*$ constant. These comparative statics are as follows:

**Remark 2 : Impact of Leader and Judge Characteristics on Judicial Strategy**

*(a) The lower threshold $\underline{T}$ is decreasing in $\gamma$, while the upper threshold $\overline{T}$ is increasing in $\gamma$. Thus, as the ability of the Judge ($\gamma$) increases, the $[\underline{T}, \overline{T}]$ interval expands.*

*(b) $\underline{T}$ and $\overline{T}$ are independent of the weight the Leader attaches to policy ($\alpha$).*

*(c) $\underline{T}$ and $\overline{T}$ are both increasing in $q$. Thus, as the probability of a competent Leader ($q$) increases, the $[\underline{T}, \overline{T}]$ interval shifts upward; the effect of an increase in $q$ on the size of this interval is ambiguous.*

The intuition for part (a) is straightforward: The more accurate the Judge's signal, the greater the Judge's payoff from relying upon it. This increases the attractiveness of the active strategy relative to the passive strategy and the strict strategy. Part (b) follows from the fact that the Judge's incentives are not directly affected by parameters of the Leader's payoff function. To see the intuition for part (c), note that if we hold $\pi^*$ fixed and increase the probability $q$ of a competent Leader, the Judge can draw a stronger inference about the state from the Leader's proposal. This makes the passive strategy more attractive relative to the active strategy (i.e., $\underline{T}$ shifts up), and makes the active strategy more attractive relative to the strict strategy (i.e., $\overline{T}$ shifts up).

behavior—for example, overconfidence in their own signals. The most prominent such hypothesis maintains that judges will be swayed excessively by the facts and circumstances of the individual cases before them (Schauer 2006; Vermeule 2009). In our setting, this sort of consideration would not matter, given that both the Judge and the Leader are evaluating a specific action rather than making a general rule. That said, we acknowledge the possibility that judicial incompetence may lead to worse outcomes in the presence of this or other forms of judicial irrationality, even if the judge is unbiased. Thus, the arguments of some neo-Thayerians might be interpreted as a call for real-world judges to behave more like the fully-rational Judge in our model. The extant literature, however, does not always sharply distinguish arguments about why judges are less *informed* than other political actors from arguments about why judges are systematically less *rational* than other actors. Many (though not all) of the structural and institutional arguments offered to support the former claim do not suffice to establish the latter claim. As for the specific argument that judges will exhibit irrational behavior because of the undue salience they attach to the specific cases before them, this is a possibility, but it is not clear how strong the effect actually is relative to potentially countervailing considerations, such as the additional information contained in a particular case or a series of cases (Rogers 2001; Sherwin 2006), and the analogous tendency of elected officials to overreact to salient events with ill-considered legislation or regulations (Kuran and Sunstein 1999; McGinnis and Mulaney 2008).

### Equilibria of Judicial Review Game

In this section, we move from our partial-equilibrium analyses in which either the Leader or the Judge's behavior was taken to be exogenous to a general-equilibrium analysis in which both the Leader and the Judge's behavior is taken to be endogenous. Our objective is to characterize the outcome of the strategic interaction between the Leader and the Judge as a function of the model's underlying parameters. The subsequent proposition illuminates how changes in the competence of the Leader ($q$) and the Judge ($\gamma$) affect the degree of judicial deference to the Leader – i.e., whether equilibria are passive, strict, or active.[32]

**Proposition 4** *$\pi^*_{act}$ and $\pi^*_{pass}$ are defined as in Proposition 2, and $\underline{T}$ and $\overline{T}$ are defined as in Proposition 3.*

*(a) A passive equilibrium exists if and only if $\pi^*_{pass} \leq \underline{T}$. In terms of the model's parameters, a sufficient condition for a passive equilibrium to exist is $q \geq \frac{p+\gamma-1}{p}$. This condition is not necessary, however, unless $\alpha = 0$.*

*(b) A strict equilibrium exists if and only if $(1-p) \geq \overline{T}$. In terms of the model's parameters, a necessary and sufficient condition for the existence of a strict equilibrium is $q \leq \frac{p-\gamma}{p}$.*

*(c) An active equilibrium exists if and only if $\pi^*_{act} \in [\underline{T}, \overline{T}]$. In terms of the model's parameters, a necessary condition for an active equilibrium to exist is that $q \leq \frac{p+\gamma-1}{p\gamma}$.*

*(d) Equilibria always exist; parameter values exist in which more than one judicial strategy is consistent with equilibrium behavior; parameter values also exist in which all equilibria involve the Judge employing a mixed strategy (i.e., equilibria are either semi-active or semi-strict).*

This proposition tell us that equilibria are passive when the probability of a competent Leader ($q$) is sufficiently large (part a), whereas equilibria are strict when the probability of a competent Leader is sufficiently small (part b). The basic intuition is straightforward: If most Leaders are high ability, the selection of the extraordinary proposal is sufficiently informative of the true underlying

[32] We note that our characterization of the conditions under which passive and active equilibria exist is incomplete. This is because the closed-form solution for the low-ability Leader's equilibrium strategy in non-strict equilibria tends to be an analytically intractable expression. This analytical intractability arises because the low-ability Leader's net expected payoff from proposing $a = x$ is a non-linear function of his strategy (see appendix for details). This non-linearity remains even if we take $F$, the electoral strength function, to be linear.

state, so it is optimal for the Judge to defer to the Leader. In contrast, when the Leader is likely to be incompetent, the Judge will be considerably more skeptical of the appropriateness of the extraordinary action when it is proposed. Hence, in such circumstances, the Judge will overrule the Leader regardless of her signal of the state. Notice that part (c) of Proposition 4 leaves open the possibility of active equilibria. However, for such equilibria to exist, the probability of a competent Leader must not be too large.

To complement our analytic results, we provide a graphical characterization of the conditions under which passive, strict, and active equilibria exist. In Figure 2, we fix $F(\hat{q}) = \hat{q}$, $p = 0.8$, and $\alpha = 0.4$, while varying both the Judge's ability ($\gamma$) and the probability of a competent Leader ($q$). We then use computational methods to solve for the model's equilibria over a discrete grid in $\gamma$-$q$ space. For instance, consider the left-most array in Figure 2. On the horizontal axis we vary $\gamma$ from 0.5 to 0.8, while on the the vertical axis we vary $q$ from 0 to 1. The black region of the array tells us those values of $q$ and $\gamma$ for which passive equilibria exist. The center and right arrays do the same for active and strict equilibria, respectively. Consistent with Proposition 4 and the intuitions form our partial equilibrium analysis (Remark 2), if we fix $\gamma$, a large $q$ facilitates the existence of passive equilibria, while small $q$ gives rise to strict equilibria. Likewise, and again consistent with the intuitions from Remark 2, if we fix $q$, increasing $\gamma$ tends to facilitate the existence of an active equilibrium.

**Figure 2 about here.**

Consideration of the general equilibrium case also reveals some potentially important features of the model that are not apparent in the partial-equilibrium analyses. First, the analysis suggests a possible explanation for the observation—often framed as a criticism—that judicial behavior in many domains is "unpredictable" (e.g. Tushnet 1979; Rose-Ackerman 1988; Pildes and Niemi 1993). In our model, equilibrium judicial behavior may be unpredictable for two reasons. First, and more straightforwardly, when the Judge uses an active review strategy her decision will depend on her signal, which is not known in advance. Second, our general-equilibrium analysis indicates that a deeper, and potentially more interesting, form of unpredictability may also arise: For some parameter values, there is no equilibrium in which the Judge plays a pure strategy (active, passive, or strict). In such cases, one cannot be certain how the Judge will respond to a given signal, because equilibrium behavior involves the Judge using a mixed strategy (semi-active or semi-strict) in which she randomizes her response to one of the two possible signals. This sort of uncertainty may be thought of as uncertainty with respect to the "standard of review" that the court will apply—a form of uncertainty often thought to be particularly problematic. Our model, however, suggests that this form of unpredictability may be the consequence of instrumentally rational and forward-looking judicial behavior, rather than judicial carelessness or neglect of long-term consequences.

Our equilibrium results also imply a different kind of unpredictability: For some parameter values, multiple equilibria may exist. For example, sometimes both active and passive review may be sustainable in equilibrium. (In particular, when $\pi^*_{pass} < \underline{T} < \pi^*_{act}$, both passive and active equilibria will exist.[33]) Such multiplicity can occur when active review exacerbates democratic failure relative to passive review. Whether the active or passive equilibrium is played when both exist depends on the equilibrium selection rule, which lies outside the scope of the model, and which may be thought of as an aspect—perhaps a manipulable aspect—of "legal culture."

## IV. Welfare Consequences of Judicial Review

We now consider the effect of judicial review on Voter welfare. In our setup, judicial review may affect Voter welfare through two distinct channels: *First*, judicial review affects the probability that the policy matches the state, which affects the Voter's *current* policy payoff. *Second*, judicial review affects the Voter's information regarding the incumbent Leader's type, which in turn affects the efficacy of elections as a device for selecting competent leaders. Through this channel, judicial review may affect the Voter's *future* policy payoffs. We consider each effect in turn, remaining agnostic as to their relative importance.

### Policy Effects

Consider first the effect of judicial review on the probability of a correct policy decision in the current period. The following proposition identifies how the form of judicial review (passive or non-passive) affects the Voter's current policy payoff.

**Proposition 5** *(a) If judicial review induces a passive equilibrium, judicial review has no effect on the Voter's expected current policy payoff.*

[33] That this possibility can arise can be gleaned by inspecting Figure 2 and observing that the intersection between the set of parameters for which passive equilibria exist and the set of parameters for which active equilibria exist is non-empty.

(b) *If judicial review induces a non-passive equilibrium in which democratic failure is lower than (equal to) that in the no review case, then judicial review improves (weakly improves) the Voter's expected current policy payoff.*

(c) *If judicial review induces a non-passive equilibrium in which democratic failure is greater than in the no review case, then judicial review may increase or decrease the Voter's expected current policy payoff.*

Part (a) of Proposition 5 is straightforward: In a passive equilibrium, behavior is identical to the no review case, so judicial review has no effect. To understand parts (b) and (c), it is helpful to recall, first, that non-passive review never decreases the Voter's current policy payoff if the level of democratic failure is held constant (see Remark 1), and, second, that if judicial review induces a non-passive equilibrium, democratic failure may increase or decrease. Thus, non-passive judicial review has two distinct effects on the Voter's expected policy payoff: *First*, non-passive judicial review will change the *probability* that the low-ability Leader will erroneously propose the extraordinary action in the normal state; *second*, non-passive judicial review will *correct* some of these erroneous decisions by striking them down.[34] If judicial review *decreases* democratic failure, both of these effects cut in the same direction, leading to an unambiguous increase in the Voter's expected policy payoff (part (b)). If, however, judicial review *increases* democratic failure, these effects cut in opposite directions, and, as a result, the effect of review on the Voter's current policy payoff is unclear (part (c)).[35] Figure 3 extends the example used in Figure 2, contrasting the Voter's current policy payoff with and without judicial review. Note in particular the right-most panel. The black region of this panel indicates the set of $q$ and $\gamma$ values for which the public's current policy payoff is strictly lower with review. For such parameters, the level of democratic failure without review is not low enough to induce the Judge to employ a passive strategy, yet non-passive review exacerbates democratic failure to such an extent that the Voter fails to benefit from the Judge's expertise.

[34] Of course, the Judge will also erroneously strike down some justified extraordinary proposals, but in equilibrium the expected benefits of correct reversals outweigh the expected costs of erroneous reversals (see Remark 1).

[35] For example, suppose judicial review induces an active equilibrium in which $\pi^*_{act} > \pi^*_{norev}$. Whether judicial review increases or decreases the Voter's current policy payoff in this instance turns on the question of whether judicial review creates more instances of democratic failure than it corrects, or corrects more than it creates. Alternatively, consider the case in which judicial review induces a strict equilibrium in which $\pi^*_{strict} > \pi^*_{norev}$. In a strict equilibrium, the final policy is always $n$. The cost of strict review is that the Voter no longer benefits from the high-ability Leader's expertise. The benefit of strict review is that the Voter no longer suffers from the low-ability Leader erroneously proposing $a = x$ when $\omega = n$. Which effect dominates depends on the relative magnitude of these costs and benefits.

**Figure 3 about here.**

Proposition 5, together with Figure 3, provides some support and some challenges both for neo-Marshallian defenders of judicial review and for neo-Thayerian critics. On the one hand, it is possible in our model for judicial review to worsen Voter welfare through a combination of judicial incompetence and judicial overhang, as critics have warned. Indeed, this adverse effect may occur even though the Judge in our model shares the Voter's state-contingent policy preferences. So, judicial bias is not a necessary condition for judicial review to have adverse effects on democratic performance: Judicial review may cause a large increase in democratic failure (most likely through a strong bailout effect) that overwhelms whatever benefits may accrue from judicial correction of erroneous decisions. This suggests a potentially serious problem with judicial review that neo-Marshallian defenders of the institution must take seriously.

On the other hand, our analysis also highlights important limits to this "judicial overhang" concern. Increased democratic failure is not an inevitable consequence of judicial review—indeed, judicial review can *reduce* democratic failure, turning the judicial overhang argument on its head. Moreover, even if judicial review does increase the frequency of democratic failure, it may correct enough democratic failures that the net impact on the Voter's expected current policy payoff is positive. The analysis also highlights the fact that relative judicial incompetence—a favorite theme of many critics of judicial review—is not a sufficient condition for judicial review to have adverse effects on the Voter's expected policy payoff, because of the endogeneity of judicial activism (see Remark 1).

We conclude our analysis of judicial review's effects on the Voter's current policy payoff with a proposition that indicates how variation in the Leader's electoral ambition (inversely measured by $\alpha$), as well as variation in the ability of the Leader ($q$) and the Judge ($\gamma$), jointly affect the desirability of judicial review.

**Proposition 6** *(a) Suppose that* $q \in \left(\frac{p-\gamma}{p}, \frac{p+\gamma-1}{p}\right)$ *and that the electoral strength function* $F$ *is concave. Then judicial review strictly increases the Voter's current policy payoff provided* $\alpha$ *is sufficiently small.*[36]

*(b) Suppose that* $q \leq \frac{p-\gamma}{p}$*. Then there exists an equilibrium with review (i.e., the strict equilib-*

[36]In addition, one can show that if $F$ is concave and $q > \frac{p+\gamma-1}{p}$, then the unique equilibrium is passive and review has no effect on the Voter's current policy payoff, regardless of $\alpha$.

*rium) under which the Voter's current policy payoff is strictly less than that with no review provided that $\alpha$ is sufficiently large.*

Part (a) establishes that when the Leader's electoral ambition is relatively large, judicial review tends to benefit the Voter, whereas part (b) establishes that when the Leader's electoral ambitions is relatively small, judicial review can harm the Voter.

The intuition for part (a) is as follows: The concavity of $F$ ensures that legitimation effect of review dampens the low-ability Leader's electoral incentive to propose $a = x$. Thus, so long as the bailout effect of review is sufficiently small—as will be the case when the weight the Leader attaches to policy is small—active judicial review will decrease democratic failure and thus increase the Voter's current policy payoff relative to the case of no review. The requirement that the prior on the Leader's competence $q$ is neither too small nor too large—i.e., $q \in \left(\frac{p-\gamma}{p}, \frac{p+\gamma-1}{p}\right)$—ensures that in any equilibrium the Judge both upholds and overrules the Leader with positive probability, so the legitimation effect of review is indeed operative.

The intuition for part (b) is a bit more straightforward: When the weight attached to policy is sufficiently large, the low-ability Leader's behavior approaches that of a faithful agent. Hence, there is little to be gained from judicial review. However, the supposition that $q \leq \frac{p-\gamma}{p}$ ensures the existence of strict equilibria, which, if induced by review, result in the Voter losing the benefit of the high-ability Leader's expertise. When $\alpha$ is high, this loss will outweigh the benefit that results from the fact that the low-ability Leader can no longer implement the extraordinary action. Interestingly, part (b) points to the possibility that review can be harmful even when it is likely that the Leader is incompetent ($q$ low). Notice that this is in fact the case in Figure 3, where review is harmful for low values of $q$ but not for high values of $q$.

### Electoral Selection Effects

It is also important to consider how judicial review affects the efficacy of the electoral process in selecting competent Leaders, which affects the Voter's long-term welfare. Because we do not model future periods explicitly, we cannot provide a fully micro-founded view of how judicial review affects future welfare. That said, our analysis permits the following observations:

First, and most straightforwardly, if judicial review induces a passive equilibrium, judicial review has no effect on Voter learning. If judicial review induces a strict equilibrium, then Voter learning is actually worse than in the no review case: If there is no judicial review, the Voter learns nothing from the Judge, but can at least learn something from the Leader's action (provided $\alpha > 0$); under strict review there is no helpful judicial signal *and* the competent and incompetent Leaders choose the extraordinary action with identical probabilities, making the Leader's action totally uninformative. If judicial review induces an equilibrium in which the Judge sometimes follows her signal (an active, semi-active, or semi-strict equilibrium), the result is more complicated because such judicial review can affect Voter learning via two distinct channels: *First*, when the Judge follows her signal, this conveys additional information to the Voter; *second*, by changing the level of democratic failure, judicial review affects the Voter's ability to draw inferences directly from the Leader's action. If review has a minimal effect on the low-ability Leader's behavior, then judicial review increases Voter learning; however, if judicial review has a substantial affect on the Leader's behavior, then Voter learning may be better or worse than in the no review case.

It is important to emphasize that it is not necessarily the case that if judicial review improves the Voter's expected current policy payoff, it also improves Voter learning (or vice versa). Although these effects will often cut in the same direction—favorable or unfavorable to judicial review—it is possible to construct scenarios in which they cut in opposite directions.[37] Therefore, institutional designers interested in assessing the value of judicial review as an institution must be attentive *both* to how judicial review affects policy outcomes *and* how it affects the efficacy of the electoral process in selecting high-quality political leaders. The former consideration is familiar and common in the literature. The latter consideration, however, is often overlooked.

## Additional Implications

The fact that the Judge in our model is rational and cares about getting the policy right, coupled with the fact for certain parameters judicial review leads to worse expected policy outcomes, has intriguing implications for our understanding of "justiciability" doctrines (standing, political question, etc.). There are several plausible explanations for why judges would devise doctrines to limit their ability to hear certain cases, including the interests in reducing workload, avoiding contro-

[37]For instance, a scenario in which judicial review enhances the Voter's current policy payoff yet hinders Voter learning is the following: Fix $\alpha = 0.1$, $q = 0.2$, $p = 0.8$, $F(\hat{q}) = \hat{q}$, and $\gamma = 0.6$. Then without review, $\pi^*_{norev} \approx 0.145$. With review, the unique equilibrium is strict, where $\pi^*_{strict} = 0.2$. While Voter learning is lower with review, the Voter's current policy payoff is higher. Specifically, the Voter's current policy payoff with review is 0.8, whereas without review it is approximately 0.77.

versy, pursuing substantive objectives surreptitiously, and controlling lower courts (e.g. Bickel 1962; Pierce 1999; Ho and Ross 2009). Our analysis suggests another possibility. Consider a situation in which judicial review would worsen the Voter's current expected policy payoff. To fix ideas, suppose that introducing judicial review induces an active equilibrium in which judicial screening does not offset the increased frequency with which the low-ability Leader proposes extraordinary action. Because the Judge shares the Voter's policy preferences, it follows that the Judge is also worse off than she would be without judicial review. This does not mean that the Judge is behaving irrationally in using the active strategy—given a high level of democratic failure, the Judge is better off following her signal. But if the Judge could credibly commit to uphold the Leader's proposal in all cases, the level of democratic failure would be lower, and the Judge would be better off from an *ex ante* perspective. Whether such a commitment could ever be credible is a legitimate question, but one might plausibly interpret certain justiciability doctrines as judicial attempts to achieve such commitment.

Our model also helps illuminate debates concerning the value of judicial competence. A natural conjecture is that if judicial review exists, the Voter would always be better off with a more competent Judge. Our model shows that this is not necessarily the case. As $\gamma$ increases, non-passive review becomes more attractive to the Judge relative to passive review. However, when the Judge employs a non-passive strategy, the policy cost to the low-ability Leader from proposing $a = x$ is lower, and, as a result, the low-ability Leader's incentive to propose $a = x$ is higher. Thus, increasing the Judge's competence $\gamma$ can increase the rate at which the low-ability Leader erroneously proposes $a = x$ when $\omega = n$. If this effect is large enough, then increasing $\gamma$ can reduce the Voter's current policy payoff despite the fact that the Judge has better information upon which to base her rulings.[38] This observation has implications for a variety of reform proposals that are intended to improve the quality of judicial review by, for example, facilitating judicial specialization (Jordan 1981; Meador 1981), selecting more judges with decision-relevant expertise (Vermeule 2007), facilitating judicial acquisition of outside expert advice (Schauer 2001), and reorienting legal scholarship to provide more information about the pragmatic consequences of judicial decisions (Posner 1998). If reforms along these lines increase $\gamma$, the impact on the quality of final policy decisions might well be desirable

[38] An example in which increasing $\gamma$ reduces the Voter's current policy payoff is the following: Fix $\alpha = 0.2$, $q = 0.5$, $p = 0.8$, and $F(\hat{q}) = \hat{q}$. When $\gamma = 0.501$, the unique equilibrium is passive, $\pi^*_{pass} \approx 0.1195$, and the Voter's current policy payoff is approximately 0.8642. In contrast, when $\gamma = 0.701$, the unique equilibrium is active, $\pi^*_{act} \approx 0.1517$, and the Voter's current policy payoff is approximately 0.8626.

in many cases, but if these reforms induce a formerly deferential judge to become active, they may increase democratic failure substantially, lowering expected Voter welfare overall.

## Conclusion

This paper used a formal political agency model to investigate how judicial review affects the incidence and impact of one form of democratic failure: the electoral incentive for low-ability leaders to undertake bold but ill-advised policy initiatives in order to project a false image of competence. Our analysis, though limited in scope, generates a number of potentially useful insights about the relationship between judicial review and democratic performance. For instance, we have shown that judicial review may increase or decrease the rate of democratic failure, depending on the relative strength of the "bailout" and "legitimation" effects of judicial review, and that rational (and unbiased) judges will only engage in active review if the level of democratic failure is high enough but not too high. Our analysis also produced a number of sometimes surprising comparative statics predictions, including the finding that increasing judicial ability may sometimes decrease rather than increase the desirability of judicial review. More generally, we have derived conditions under which judicial review may improve or worsen voter welfare through a direct impact on policy outcomes and through an indirect impact on electoral sorting. The analysis also predicts, and provides a rationalist explanation for, a number of phenomena often associated with real-world judicial systems, including judicial deference, legal unpredictability, and judicially-created limits on the courts' own jurisdiction.

Beyond these contributions, the larger objective of this paper is to help lay the groundwork for a research program that would rigorously analyze the effect of judicial review on democratic performance in a political agency framework. Subsequent research could build on our preliminary analysis in a number of ways. First, while we have focused narrowly on one particular form of democratic failure, the same general approach might be applied to other forms of democratic failure as well. Also, while our analysis made the simplifying assumption that the Judge shares the Voter's policy preferences, future work should relax that assumption, and could also incorporate a richer characterization of judicial institutions—taking into account, for example, the hierarchical structure of the courts, strategic litigation and judicial agenda-setting, collective decision-making by multi-member courts, the complexity of legal doctrine, and problems of endogenous compliance with judicial decisions. Research along these and other lines would further enrich our understanding of the costs and benefits of judicial review. In addition, the framework we have developed here, when combined with the emerging literature on separation-of-powers between elected branches of government (e.g Fox and Van Weelden 2008; Persson, Roland and Tabellini 1997; Vlaicu 2008), may facilitate direct comparisons between judicial and political oversight of government decision-making—a crucial issue in contemporary legal and policy debates, which the extant political economy literature has not fully engaged.

## References


Ackerman, Bruce A. 1985. "Beyond Carolene Products." *Harvard Law Review* 98:713–746.

Ashworth, Scott and Kenneth W. Shotts. 2008. "Does Informative Media Commentary Reduce Politicians' Incentives To Pander?" Unpublished Paper.

Austen-Smith, David and Jeffrey Banks. 1989. Electoral Accountability and Incumbency. In *Models of Strategic Choice in Politics*, ed. Peter C. Ordeshook. Ann Arbor: University of Michigan Press.

Avery, Christopher N. and Judith A. Chevalier. 1999. "Herding over the Career." *Economic Letters* 63(3):327–333.

Baas, Larry R. and Dan Thomas. 1984. "The Supreme Court and Policy Legitimation: Experimental Tests." *American Politics Quarterly* 12:335–360.

Barnum, David G. 1993. *The Supreme Court and American Democracy.* New York: St. Martin's Press.

Bickel, Alexander M. 1962. *The Least Dangerous Branch: The Supreme Court at the Bar of Politics.* New Haven: Yale University Press.

Canes-Wrone, Brandice, Michael C. Herron and Kenneth W. Shotts. 2001. "Leadership and Pandering: An Equilibrium Theory of Executive Policymaking." *American Journal of Political Science* 45(3):532–550.

Chemerinsky, Erwin. 1989. "The Supreme Court, 1988 Term–Forward: The Vanishing Constitution." *Harvard Law Review* 103:43–104.

Chemerinsky, Erwin. 2004. “In Defense of Judicial Review: The Perils of Popular Constitutionalism.” *University of Illinois Law Review* 2004:673–690.

Clawson, Rosalee A., Elizabeth R. Kegler and Eric N. Waltenburg. 2001. “The Legitimacy-Conferring Authority of the U.S. Supreme Court: An Experimental Design.” *American Politics Research* 29(6):566–591.

Clinton, Robert Lowry. 1989. *Marbury v. Madison and Judicial Review.* Lawrence: University Press of Kansas.

Coram, Bruce Talbot. 1996. Second Best Theories and the Implications for Institutional Design. In *The Theory of Institutional Design*, ed. Robert E. Goodin. Cambridge: Cambridge University Press.

Cross, Frank B. 2000. “Institutions and Enforcement of the Bill of Rights.” *Cornell Law Review* 85:1529–1608.

Dahl, Robert. 1957. “Decision-Making in a Democracy: The Supreme Court as a National Policy-Maker.” *Journal of Public Law* 6:279–295.

Dworkin, Ronald. 1985. *A Matter of Principle.* Cambridge: Harvard University Press.

Eisgruber, Christopher L. 2001. *Constitutional Self-Government.* Cambridge: Harvard University Press.

Elhauge, Einer R. 1991. “Does Interest Group Theory Justify More Instrusive Judicial Review?” *Yale Law Journal* 101:31–110.

Ely, John Hart. 1980. *Democracy and Distrust: A Theory of Judicial Review.* Cambridge: Harvard University Press.

Fallon, Richard H. 2008. “The Core of an Uneasy Case for Judicial Review.” *Harvard Law Review* 121:1694–1736.

Ferejohn, John and Charles R. Shipan. 1990. “Congressional Influence on Bureaucracy.” *Journal of Law, Economics, and Organization* 6:1–20.

Fox, Justin. 2007. “Government Transparency and Policymaking.” *Public Choice* 131:23–44.

Fox, Justin and Richard Van Weelden. 2008. "Partisanship and the Effectiveness of Oversight." Unpublished Paper.

Friedman, Barry. 1993. "Dialogue and Judicial Review." *Michigan Law Review* 91:577–682.

Friedman, Barry. 2002. "The Birth of an Academic Obsession: The History of the Countermajoritarian Difficulty, Part Five." *Yale Law Journal* 112:153–259.

Friedman, Barry. 2009. *The Will of the People: How Public Opinion Has Influenced the Supreme Court and Shaped the Meaning of the Constitution*. Farrar, Straus and Giroux.

Gely, Rafael and Pablo T. Spiller. 1992. "The Political Economy of Supreme Court Constitutional Decisions: The Case of Roosevelt's Court-Packing Plan." *International Review of Law and Economics* 12:45–67.

Hirschl, Ran. 2000. "The Political Origins of Judicial Empowerment Through Constitutionalization: Lessons from Four Constitutional Revolutions." *Law and Social Inquiry* 25(1):91–152.

Hirschl, Ran. 2004. *Towards Juristocracy: The Origins and Consequences of the New Constitutionalism*. Cambridge: Harvard University Press.

Ho, Daniel E. and Erica L. Ross. 2009. "Did Liberal Justices Invent the Standing Doctrine?: An Empirical Study of the Evolution of Standing, 1921-2006." *Stanford Law Review* . Forthcoming.

Hoekstra, Valerie J. and Jeffrey A. Segal. 1996. "The Shepherding of Local Public Opinion: The Supreme Court and Lamb's Chapel." *Journal of Politics* 58(4):1079–1102.

Issacharoff, Samuel and Richard H. Pildes. 1998. "Politics as Markets: Partisan Lockups of the Democratic Process." *Stanford Law Review* 50:643–717.

Jordan, Ellen R. 1981. "Specialized Courts: A Choice." *Northwestern University Law Review* 76(5):745–785.

Klarman, Michael J. 2001. "How Great Were the 'Great' Marshall Court Decisions?" *Virginia Law Review* 87:1111–1184.

Komesar, Neil K. 1994. *Imperfect Alternatives: Choosing Institutions in Law, Economics, and Public Policy*. Chicago: University of Chicago Press.

Kuran, Timur and Cass R. Sunstein. 1999. "Availability Cascades and Risk Regulation." *Stanford Law Review* 51:683–768.

Levy, Gilat. 2004. "Anti-Herding and Strategic Consultation." *European Economic Review* 48(3):503–525.

Marshall, Thomas R. 1989. *Public Opinion and the Supreme Court.* Boston: Unwin Hyman.

Maskin, Eric and Jean Tirole. 2004. "The Politician and the Judge: Accountability in Government." *American Economic Review* 94:1034–1054.

McGinnis, John O. and Charles W. Mulaney. 2008. "Judging Facts Like Law." *Constitutional Commentary* 25:69–130.

McGuire, Kevin T. and James A. Stimson. 2004. "The Least Dangerous Branch Revisited: New Evidence on Supreme Court Responsiveness to Public Preferences." *Journal of Politics* 66(4):1018–1035.

Meador, Daniel J. 1981. "Appellate Subject Matter Organization: The German Design from an American Perspective." *Hastings International and Comparative Law Review* 5(1):27–72.

Persily, Nathaniel. 2002. "In Defense of Foxes Guarding Henhouses: The Case for Judicial Acquiescence to Incumbent-Protecting Gerrymanders." *Harvard Law Review* 116:649–683.

Persson, Torsten, Gerard Roland and Guido Tabellini. 1997. "Separation of Powers and Political Accountability." *Quarterly Journal of Economics* 112:310–327.

Pierce, Richard J., Jr. 1999. "Is Standing Law or Politics?" *North Carolina Law Review* 77:1741–1789.

Posner, Eric A. 2008. "Does Political Bias in the Judiciary Matter?: Implications of Judicial Bias Studies for Legal and Constitutional Reform." *University of Chicago Law Review* 75:853–883.

Posner, Richard A. 1998. "Against Constitutional Theory." *New York University Law Review* 73:1–22.

Powe, Lucas A. 2000. *The Warren Court and American Politics.* Belknap Press.

Prat, Andrea. 2005. "The Wrong Kind of Transparency." *American Economic Review* 95(3):862–877.

Prendergast, Canice. 1993. "A Theory of 'Yes Men'." *American Economic Review* 83(4):757–770.

Prendergast, Canice and Lars Stole. 1996. "Impetuous Youngsters and Jaded Old-Timers: Acquiring a Reputation for Learning." *Journal of Political Economy* 104:1105–1134.

Riker, Willam H. 1982. *Liberalism Against Populism: A Primer.* W.H. Freeman and Co.

Rogers, James R. 1999. "Legislative Incentives and Two-Tiered Judicial Review: A Game Theoretic Reading of Carolene Products Footnote Four." *American Journal of Political Science* 43(4):1096–1121.

Rogers, James R. 2001. "Information and Judicial Review: A Signaling Game of Legislative-Judicial Interaction." *American Journal of Political Science* 45(1):84–99.

Rogers, James R. 2009. "Casting the Gimlet Eye on Judicial Review: Can Judicial Review Be Democratically Debilitating?" Unpublished paper.

Rogers, James R. and Georg Vanberg. 2007. "Resurrecting Lochner: A Defense of Unprincipled Judicial Activism." *Journal of Law, Economics, and Organization* 23(2):442–468.

Salzberger, Eli M. 1993. "A Positive Analysis of the Doctrine of Separation of Powers, or: Why Do We Have an Independent Judiciary?" *International Review of Law and Economics* 13:349–379.

Schauer, Frederick. 2000. "Incentives, Reputation, and the Inglorious Determinants of Judicial Behavior." *University of Cincinnatti Law Review* 68:615–636.

Schauer, Frederick. 2001. "The Dilemma of Ignorance: PGA Tour, Inc. v. Casey Martin." *Supreme Court Review* 2001:267–297.

Schauer, Frederick. 2006. "Do Cases Make Bad Law?" *University of Chicago Law Review* 73:883–918.

Seidman, Louis Michael. 2001. *Our Unsettled Constitution: A New Defense of Constitutionalism and Judicial Review.* New Haven: Yale University Press.

Shepsle, Kenneth A. 1992. "Congress Is a 'They,' Not an 'It': Legislative Intent as Oxymoron." *International Review of Law and Economics* 12(2):239–256.

Sherwin, Emily. 2006. "Judges as Rulemakers." *University of Chicago Law Review* 73:919–931.

Stephenson, Matthew C. 2003. "'When the Devil Turns...': The Political Foundations of Independent Judicial Review." *Journal of Legal Studies* 32:50–89.

Stephenson, Matthew C. 2008. "The Price of Public Action: Constitutional Doctrine and the Judicial Manipulation of Legislative Enactment Costs." *Yale Law Journal* 118(1):2–62.

Stiglitz, Joseph E. 2000. *Economics of the Public Sector, 3rd Edition*. W.W. Norton and Co.

Sunstein, Cass R. 1984. "Naked Preferences and the Constitution." *Columbia Law Review* 84:1689–1732.

Sunstein, Cass R. 1985. "Interest Groups in American Public Law." *Stanford Law Review* 38:29–87.

Tate, C. Neal and Torbjorn Vallinder. 1997. *The Global Expansion of Judicial Power*. New York: New York University Press.

Thayer, James Bradley. 1893. "The Origin and Scope of the American Doctrine of Constitutional Law." *Harvard Law Review* 7(3):129–156.

Tribe, Laurence H. 2000. *American Constitutional Law, 3rd Edition*. West Publishing Co.

Trueman, Brett. 1994. "Analyst Forecasts and Herding Behavior." *Review of Financial Studies* 7:97–124.

Tsebelis, George. 2002. *Veto Players: How Political Institutions Work*. Princeton: Princeton University Press.

Tushnet, Mark. 1999. *Taking the Constitution Away from the Courts*. Princeton: Princeton University Press.

Tushnet, Mark. 2006. "Political Power and Judicial Power: Some Observations on Their Relation." *Fordham Law Review* 75:755–768.

Vermeule, Adrian. 2003. "Hume's Second-Best Constitutionalism." *University of Chicago Law Review* 70:421–436.

Vermeule, Adrian. 2006. *Judging under Uncertainty: An Institutional Theory of Legal Interpretation.* Cambridge: Harvard University Press.

Vermeule, Adrian. 2007. "Should We Have Lay Justices?" *Stanford Law Review* 59:1569–1611.

Vermeule, Adrian. 2009. *Law and the Limits of Reason.* New York: Oxford University Press.

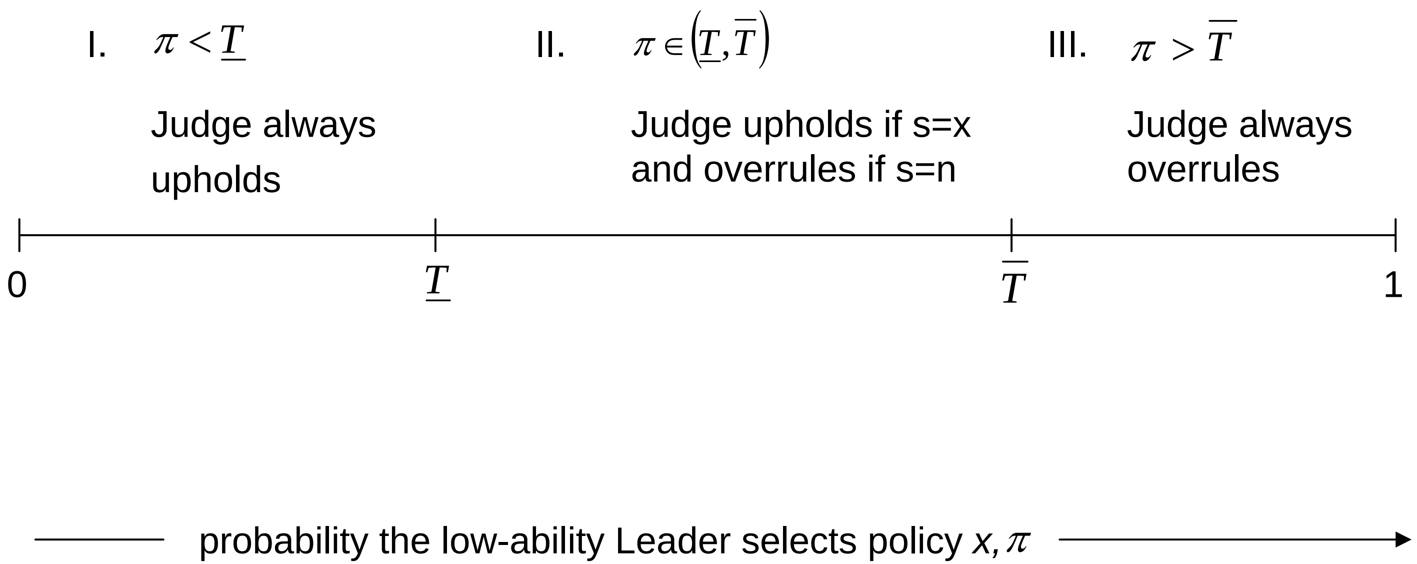
Vlaicu, Razvan. 2008. "Executive Performance under Direct and Hierarchical Accountability Structures: Theory and Evidence." Unpublished Paper.

Waldron, Jeremy. 2006. "The Core of the Case Against Judicial Review." *Yale Law Journal* 115:1346–1406.

Whittington, Keith E. 2005. "'Interpose Your Friendly Hand': Political Supports for the Exercise of Judicial Review by the United States Supreme Court." *American Political Science Review* 99(4):583–596.

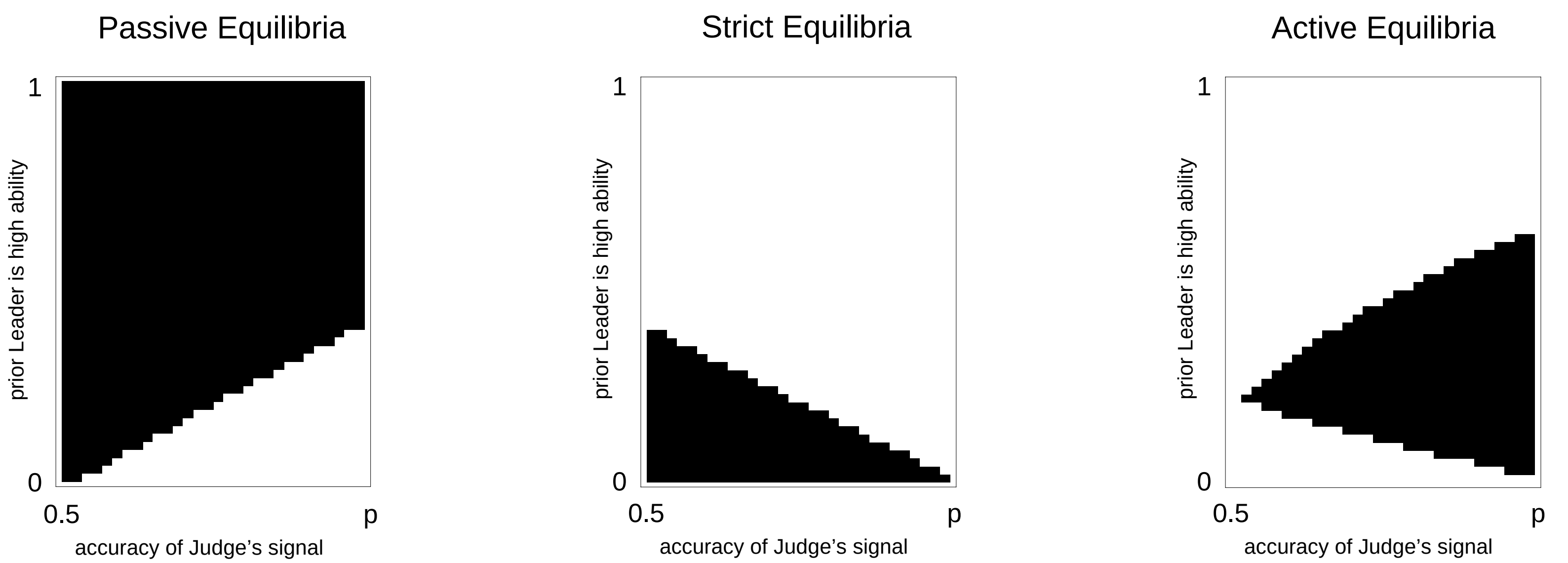
Young, Ernest A. 2000. "Constitutional Avoidance, Resistance Norms, and the Preservation of Judicial Review." *Texas Law Review* 78:1549–1614.

**Figure 1: Judge's best response as a function of the low-ability Leader's strategy**



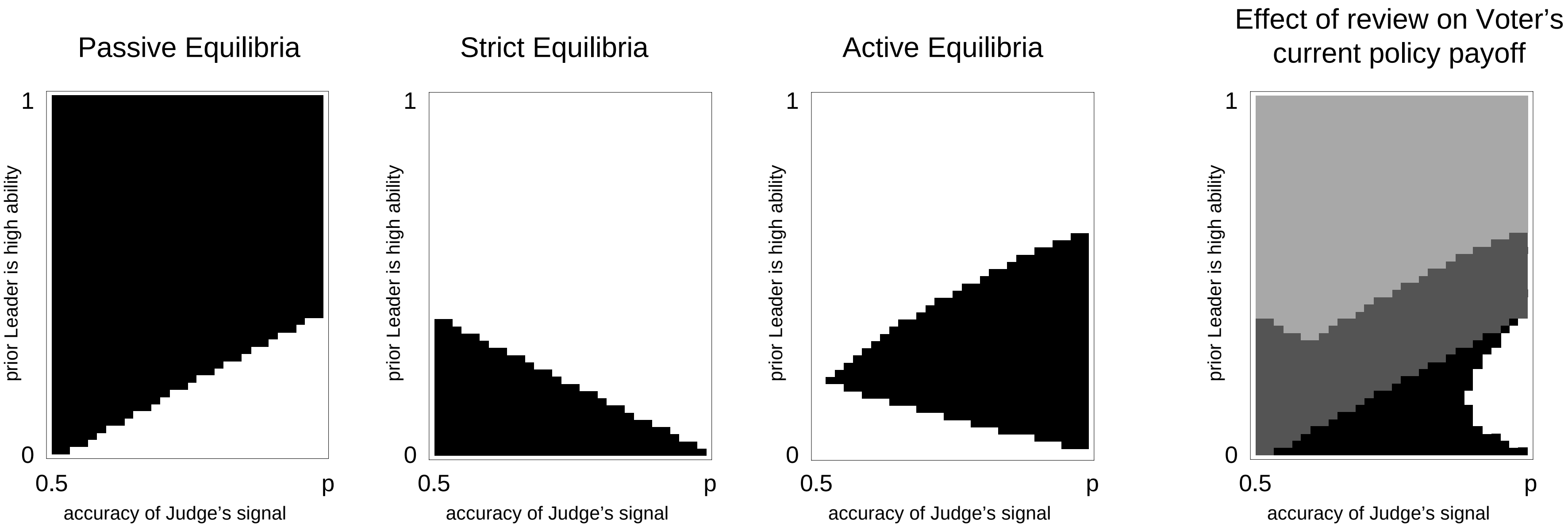

*This figure provides a graphical illustration of Proposition 3. The figure indicates that as the probability the low-ability Leader proposes the extraordinary action increases, the Judge's willingness to strike down the extraordinary action also increases. More specifically, when democratic failure is low (region I), the Judge's best response is to employ a passive strategy; when democratic failure is moderate (region II), the Judge's best response is to use an active strategy; and when democratic failure is high (region III), the Judge's best response is to use a strict strategy.

# Figure 2: The Judge's equilibrium strategy as a function of the accuracy of her signal and the prior belief about the Leader's ability




*Simulations fix $F(\hat{q})=\hat{q}$, $p$=0.8, and $\alpha$=0.4. In each panel, we vary the accuracy of the Judge's signal $\gamma$ along the horizontal axis and the prior $q$ the Leader is high ability along the vertical axis, where $\gamma$ ranges from 0.5 to 0.8 and q ranges from 0 to 1. The black region in a given panel indicates those values of $q$ and $\gamma$ for which an equilibrium of the specified form exists.

# Figure 3: Effect of judicial review on the Voter's current policy payoff




*Simulations fix F($\hat{q}$)=$\hat{q}$, *p*=0.8, and *α*=0.4. In each panel, we vary the accuracy of the Judge's signal *γ* along the horizontal axis and the prior the Leader is high ability *q* along the vertical axis. The right-most panel indicates whether the Voter is better off in terms of her current policy payoff with review than without review:

- •The Voter is strictly worse off with review in the black region of this panel.
- •The Voter is weakly worse off with review in the dark grey region (i.e., in this region, there exists a passive equilibrium that gives the Voter a payoff identical to that without review *and* a non-passive equilibrium which gives the Voter a payoff strictly lower than that without review).
- •The Voter's current policy payoff is the same with and without review in the light grey region.
- •The Voter is strictly better off with review in the white region.

# Technical Appendix

We begin by reiterating two of the assumptions made in the main text that will be useful to keep in mind when proving the manuscript's propositions. First, we assumed that the high-ability Leader matches policy to the state. Second, we assumed that policy $a = n$ is not justiciable. This is equivalent to assuming that when $a = n$ is proposed, the Judge always upholds the Leader. We conclude the appendix by proving that both behavioral postulates—i.e., the high-ability Leader matches policy to the state and the Judge upholds when $a = n$—are consistent with equilibrium behavior.

## Preliminaries

### The Voter's Beliefs About the Leader's Ability

Suppose the low-ability Leader proposes $a = x$ with probability $\pi$. And suppose the Judge's strategy is given by the double $\sigma = (\sigma_n, \sigma_x)$: When $s = n$, the Judge upholds the Leader with probability $\sigma_n$, and when $s = x$, the Judge upholds the Leader with probability $\sigma_x$.

Absent review, the Leader's reputation from action $a$ will be denoted by $\hat{q}(a; \pi)$. That is, $\hat{q}(a; \pi)$ denotes the Voter's posterior that the Leader is of high ability when policy $a$ is proposed. By Bayes' Rule, we have

$$\hat{q}(a; \pi) = \frac{Pr(a|t = h)Pr(t = h)}{Pr(a|t = h)Pr(t = h) + Pr(a|t = l)Pr(t = l)}.$$

Thus,

$$\begin{aligned} \hat{q}(x; \pi) &\equiv \frac{(1-p)q}{(1-p)q + \pi(1-q)} \\ \hat{q}(n; \pi) &\equiv \frac{pq}{pq + (1-\pi)(1-q)}. \end{aligned}$$

With judicial review, the Leader's reputation when he proposes policy $a$ and the Judge issues ruling $d$ will be denoted by $\hat{q}(a, d; \pi, \sigma)$. Since we assume the Judge always chooses $d = uphold$ when $a = n$, we have that $\hat{q}(n, uphold; \pi, \sigma) = \hat{q}(n; \pi)$. All that remains is to specify $\hat{q}$ when $a = x$ and this proposal is subject to judicial review. If the Judge issues ruling $d$ with positive probability when $a = x$, then, by Bayes' Rule, we have

$$\hat{q}(x, d; \pi, \sigma) = \frac{Pr(d|x, t = h)\hat{q}(x; \pi)}{Pr(d|x, t = h)\hat{q}(x; \pi) + Pr(d|x, t = l)(1 - \hat{q}(x; \pi))}. \qquad (1)$$

Write $\lambda(t, \sigma)$ for the probability proposal $a = x$ is upheld when the Leader's type is $t$ and the Judge uses strategy $\sigma$: $\lambda(t, \sigma) = Pr(uphold|x, t)$, where $Pr(uphold|x, t) = \sum_\omega Pr(\omega|x, t)[\sigma_x Pr(s = x|\omega) + \sigma_n Pr(s = n|\omega)]$. We thus have

$$\begin{aligned} \lambda(h, \sigma) &\equiv \sigma_x \gamma + \sigma_n (1 - \gamma) \\ \lambda(l, \sigma) &\equiv \sigma_x [(1 - p)\gamma + p(1 - \gamma)] + \sigma_n [(1 - p)(1 - \gamma) + p\gamma]. \end{aligned}$$

Using the convention that $\hat{q}(x, d; \pi, \sigma) = q$ when ruling $d$ is off path, and substituting both $\hat{q}(x; \pi)$ and $\lambda(t, \sigma)$ into (1) when ruling $d$ is on path, we have that

$$\hat{q}(x, uphold; \pi, \sigma) \equiv \begin{cases} \frac{\lambda(h,\sigma)(1-p)q}{\lambda(h,\sigma)(1-p)q + \lambda(l,\sigma)\pi(1-q)}, & \text{if } \sigma_n > 0 \text{ or } \sigma_x > 0 \\ q, & \text{otherwise} \end{cases},$$

and

$$\hat{q}(x, strike; \pi, \sigma) \equiv \begin{cases} \frac{[1-\lambda(h,\sigma)](1-p)q}{[1-\lambda(h,\sigma)](1-p)q + [1-\lambda(l,\sigma)]\pi(1-q)}, & \text{if } \sigma_n < 1 \text{ or } \sigma_x < 1 \\ q, & \text{otherwise} \end{cases}.$$

### The Leader's Expected Payoff

Write $u(a, d; \omega)$ for the (common) policy payoff that results when action $a$ is proposed, the Judge's ruling is $d$, and the state is $\omega$, where

$$u(a, d; \omega) = \begin{cases} 1, & \text{if } \{a = \omega;\ d = uphold\} \text{ or } \{a \neq \omega;\ d = strike\} \\ 0, & \text{otherwise} \end{cases}.$$

In what follows, we focus on the low-ability Leader's incentives (since we have assumed the high-ability Leader matches policy to the state). The low-ability Leader's incentive to propose a particular policy depends upon his expected policy and electoral payoffs from doing so.

Write $E_{\omega,d}[u(a, d; \omega); \sigma]$ for the expected policy payoff to the low-ability Leader from proposing policy $a$ when the Judge's strategy is $\sigma$. Now notice that

$$E_{\omega,d}[u(a,d;\omega);\sigma] = \sum_{\omega,d} Pr(d|a,\omega)Pr(\omega)u(a,d;\omega).$$

Thus, we have that $E_{\omega,d}[u(x,d;\omega);\sigma] = p \cdot [(1-\sigma_n)\gamma + (1-\sigma_x)(1-\gamma)] + (1-p) \cdot [\sigma_n(1-\gamma) + \sigma_x\gamma]$; and, since the Judge is assumed to always uphold the Leader when $a = n$, it follows that $E_{\omega,d}[u(n,d;\omega);\sigma] = p$.

Write $E_d[F(\hat{q}(a,d;\pi,\sigma))]$ for the low-ability Leader's expected probability of reelection when he proposes policy policy $a$, his strategy is $\pi$, the Judge's strategy is $\sigma$, and beliefs about the Leader's ability are derived via Bayes' Rule whenever possible. Since

$$\begin{aligned} E_d[F(\hat{q}(a,d;\pi,\sigma))] = Pr(uphold|a,t=l)F(\hat{q}(a,uphold;\pi,\sigma)) + \\ Pr(strike|a,t=l)F(\hat{q}(a,strike;\pi,\sigma)), \end{aligned}$$

we have that $E_d[F(\hat{q}(x,d;\pi,\sigma))] = \lambda(l,\sigma)F(\hat{q}(x,uphold;\pi,\sigma)) + (1-\lambda(l,\sigma))F(\hat{q}(x,strike;\pi,\sigma))$; and, since the Judge is assumed to always uphold the Leader when $a = n$, it follows that $E_d[F(\hat{q}(n,d;\pi,\sigma))] = F(\hat{q}(n;\pi))$.

We will write $\Delta^p(\sigma) \equiv E_{\omega,d}[u(x,d;\omega);\sigma] - p$ for the low-ability Leader's net policy payoff from choosing $a = x$ when the Judge's strategy is $\sigma$, and we will write $\Delta^e(\pi,\sigma) \equiv E_d[F(\hat{q}(x,d;\pi,\sigma))] - F(\hat{q}(n;\pi))$ for the low-ability Leader's net electoral payoff from choosing $a = x$ given his strategy is $\pi$, the Judge's strategy is $\sigma$, and beliefs about the Leader's ability are derived via Bayes' Rule whenever possible. Finally, we denote by $\Delta(\pi,\sigma) \equiv \alpha\Delta^p(\sigma) + (1-\alpha)\Delta^e(\pi,\sigma)$ the net benefit to the low-ability Leader from choosing $a = x$. (Recall that $\alpha$ is the weight that the Leader attaches to policy considerations.) If this net benefit is positive, then the low-ability Leader maximizes his expected payoff by choosing $a = x$; if it is equal to zero, he is indifferent between $a = x$ and $a = n$; and if it is negative, he maximizes his expected payoff by selecting $a = n$.

## The Judge's Beliefs about the State and Expected Payoff

Write $\hat{p}(s;\pi)$ for the probability the Judge assigns to $\omega = n$ when $a = x$, her signal is $s$, and the strategy of the low-ability Leader is $\pi$. By Bayes' Rule, we have

$$\hat{p}(s;\pi) = \frac{Pr(s|\omega = n)Pr(a = x|\omega = n)Pr(\omega = n)}{Pr(s|\omega = n)Pr(a = x|\omega = n)Pr(\omega = n) + Pr(s|\omega = x)Pr(a = x|\omega = x)Pr(\omega = x)}.$$

Thus,

$$\begin{aligned}
\hat{p}(x;\pi) &\equiv \frac{(1-\gamma)(1-q)\pi p}{(1-\gamma)(1-q)\pi p + \gamma(q+(1-q)\pi)(1-p)} \\
\hat{p}(n;\pi) &\equiv \frac{\gamma(1-q)\pi p}{\gamma(1-q)\pi p + (1-\gamma)(q+(1-q)\pi)(1-p)}.
\end{aligned}$$

Write $E_\omega[u(x,d;\omega);s,\pi]$ for the Judge's expected policy payoff from decision $d$ when $a = x$, her signal is $s$, and the strategy of the low-ability Leader is $\pi$. Since $E_\omega[u(x,d;\omega);s,\pi] = \hat{p}(s;\pi)u(x,d;\omega = n) + (1-\hat{p}(s;\pi))u(x,d;\omega = x)$, by the definition of $u(a,d;\omega)$, we have $E_\omega[u(x,strike;\omega);s,\pi] = \hat{p}(s;\pi)$ and $E_\omega[u(x,uphold;\omega);s,\pi] = 1-\hat{p}(s;\pi)$.

## The Voter's Policy Payoff

Write $E_{a,d,\omega}[u(a,d;\omega);(\pi,\sigma),t]$ for the Voter's expected current policy payoff when the Leader's ability level is $t$ and strategy profile $(\pi,\sigma)$ is played. Since $E_{a,d,\omega}[u(a,d;\omega);(\pi,\sigma),t] = \sum_{a,d,\omega} u(a,d;\omega)Pr(d|a,\omega)Pr(a|t,\omega)Pr(\omega)$,

$$E_{a,d,\omega}[u(a,d;\omega);(\pi,\sigma),h] = p + (1-p)(\sigma_x\gamma + \sigma_n(1-\gamma))$$

and

$$\begin{aligned}
E_{a,d,\omega}[u(a,d;\omega);(\pi,\sigma),l] &= p\left((1-\pi) + \pi((1-\sigma_x)(1-\gamma) + (1-\sigma_n)\gamma)\right) \\
&\quad + (1-p)\left(\pi(\sigma_x\gamma + \sigma_n(1-\gamma))\right).
\end{aligned}$$

In what follows, we denote the Voter's expected current policy payoff from strategy profile $(\pi, \sigma)$ as $V(\pi, \sigma)$:

$$V(\pi, \sigma) \equiv qE_{a,d,\omega}[u(a, d; \omega); (\pi, \sigma), h] + (1 - q)E_{a,d,\omega}[u(a, d; \omega); (\pi, \sigma), l].$$

## Proofs

We begin by proving three lemmas that will be invoked in proving the main text's results concerning equilibrium behavior with and without judicial review.

**Lemma 1** *(a) Suppose the Judge employs a passive strategy. Then the low-ability's Leader's net electoral payoff from the extraordinary action is positive (negative) if and only if the probability with which he proposes the extraordinary action is less (greater) than $1 - p$: $\Delta^e(\pi, (1, 1)) \gtreqless 0$ if and only if $\pi \lesseqgtr 1 - p$. (b) If the probability with which the Judge upholds the extraordinary action is positive, then the low-ability Leader's net policy payoff from the extraordinary action is negative; in contrast, when the Judge always strikes down the extraordinary action, the low-ability Leader is indifferent policy-wise between $a = n$ and $a = x$: If $\sigma_n > 0$ or $\sigma_x > 0$, then $\Delta^p(\sigma_n, \sigma_x) < 0$; also, $\Delta^p(0, 0) = 0$.*

Proof of part (a). $\Delta^e(\pi, (1, 1)) = F(\hat{q}(x; \pi)) - F(\hat{q}(n; \pi))$. Thus, $\Delta^e(\pi, (1, 1)) \gtreqless 0$ if and only if $F(\hat{q}(x; \pi)) \gtreqless F(\hat{q}(n; \pi))$. Since $F$ is increasing in $\hat{q}$, $F(\hat{q}(x; \pi)) \gtreqless F(\hat{q}(n; \pi))$ if and only if $\hat{q}(x; \pi) \gtreqless \hat{q}(n; \pi)$. Finally, one can show (via algebra) that $\hat{q}(x; \pi) \gtreqless \hat{q}(n; \pi)$ if and only if $\pi \lesseqgtr 1 - p$.

Proof of part (b). By definition, $\Delta^p(\sigma) = E_{\omega,d}[u(x, d; \omega); \sigma] - p$, where

$$E_{\omega,d}[u(x, d; \omega); \sigma] = p \cdot [(1 - \sigma_n)\gamma + (1 - \sigma_x)(1 - \gamma)] + (1 - p) \cdot [\sigma_n(1 - \gamma) + \sigma_x\gamma]$$

First, notice that $E_{\omega,d}[u(x, d; \omega); \sigma] = p$ when $\sigma_n = \sigma_x = 0$; therefore, $\Delta^p(0, 0) = 0$. Consequently, if we can show that $E_{\omega,d}[u(x, d; \omega); \sigma]$ is decreasing in both $\sigma_n$ and $\sigma_x$, we will have established $\Delta^p(\sigma_n, \sigma_x) < 0$ whenever $\sigma_n > 0$ or $\sigma_x > 0$. To see that $E_{\omega,d}[u(x, d; \omega); \sigma]$ is decreasing in $\sigma_n$ and $\sigma_x$, notice that $\frac{\partial E_{\omega,d}[u(x,d;\omega);\sigma]}{\partial \sigma_n} = 1 - p - \gamma$ and that $\frac{\partial E_{\omega,d}[u(x,d;\omega);\sigma]}{\partial \sigma_x} = -p + \gamma$. By assumption, $\gamma > \frac{1}{2}$ and $p > \frac{1}{2}$. Thus, $\frac{\partial E_{\omega,d}[u(x,d;\omega);\sigma]}{\partial \sigma_n} < 0$. Also, by assumption, $p > \gamma$. Thus, $\frac{\partial E_{\omega,d}[u(x,d;\omega);\sigma]}{\partial \sigma_x} < 0$. ■

**Lemma 2** *(a) The low-ability Leader's net expected payoff from the extraordinary action, $\Delta(\pi, \sigma)$,* *is decreasing in* $\pi$*, where* $\Delta(1,\sigma) < 0$*. (b) When* $\Delta(0,\sigma) > 0$*, there exists a unique solution to* $\Delta(\pi,\sigma) = 0$ *in* $\pi$*, say* $\hat{\pi}$*, where* $\hat{\pi} \in (0,1)$*.*

Proof of part (a). We first show that $\Delta(\pi,\sigma)$ is decreasing in $\pi$. Taking the derivative of the low-ability Leader's net benefit from the extraordinary action with respect to $\pi$, we have

$$\frac{\partial \Delta(\pi,\sigma)}{\partial \pi} = \lambda(l,\sigma)\frac{\partial F(\hat{q}(x,uphold;\pi,\sigma))}{\partial \hat{q}}\frac{\partial \hat{q}(x,uphold;\pi,\sigma)}{\partial \pi} + (1-\lambda(l,\sigma))\frac{\partial F(\hat{q}(x,strike;\pi,\sigma))}{\partial \hat{q}}\frac{\partial \hat{q}(x,strike;\pi,\sigma)}{\partial \pi} - \frac{\partial F(\hat{q}(n;\pi))}{\partial \hat{q}}\frac{\partial \hat{q}(n;\pi)}{\partial \pi}.$$

By assumption, the electoral strength function $F$ is increasing in $\hat{q}$; by inspection, $\hat{q}(x,d;\pi,\sigma)$ is weakly decreasing in $\pi$; and, by inspection, $\hat{q}(n;\pi,\sigma)$ is increasing in $\pi$. Thus, $\Delta(\pi,\sigma)$ is decreasing in $\pi$.

We now show that $\Delta(1,\sigma) < 0$. Since $\Delta^p(\sigma) \leq 0$ (Lemma 1(b)) and $\alpha < 1$ (by assumption), to prove that $\Delta(1,\sigma) < 0$, it is sufficient to show that $\Delta^e(1,\sigma) < 0$. If $\pi = 1$, then $\hat{q}(n;1) = 1$, so the low-ability Leader's probability of reelection from proposing $a = n$ is $F(1)$. Thus, to show that $\Delta^e(1,\sigma) < 0$, we must show that the low-ability Leader's expected probability of reelection from proposing $a = x$ is less than $F(1)$. By inspection of $\hat{q}(x,d;1,\sigma)$, for any ruling $d$, $\hat{q}(x,d;1,\sigma) < 1$. This fact, taken together with the fact that $F$ is increasing in the Leader's reputation $\hat{q}$, implies that for any ruling $d$, $F(\hat{q}(x,d;1,\sigma)) < F(1)$. Hence, the low-ability Leader's expected probability of reelection from proposing $a = x$ is less than $F(1)$.

Proof of part (b). Suppose that $\Delta(0,\sigma) > 0$. We know from part (a) that when $\pi = 1$, $\Delta(1,\sigma) < 0$. Since $\Delta(0,\sigma) > 0$ and $\Delta(1,\sigma) < 0$, any solution to $\Delta(\pi,\sigma) = 0$ in $\pi$ lies in $(0,1)$. Moreover, since $\Delta$ is continuous in $\pi$, $\Delta(0,\sigma) > 0$, and $\Delta(1,\sigma) < 0$, the Intermediate Value Theorem ensures that there exists $\hat{\pi} \in (0,1)$ such that $\Delta(\hat{\pi},\sigma) = 0$. That the solution to $\Delta(\pi,\sigma) = 0$ in $\pi$ is unique follows from the fact that $\Delta$ is decreasing in $\pi$ (see part (a)). ■

To state the next lemma, we define a function $\psi$ that maps the Judge's strategy into a probability that the low-ability Leader selects $a = x$.

$$\psi(\sigma) \equiv \begin{cases} 0, & \text{if } \Delta(0,\sigma) \leq 0 \\ \hat{\pi} \in (0,1), & \text{otherwise, where } \hat{\pi} \text{ is the unique solution to } \Delta(\pi,\sigma) = 0 \text{ in } \pi \end{cases}.$$

**Lemma 3** *In any equilibrium* $(\pi^*, \sigma^*, \hat{q}, \hat{p})$*,* $\pi^* = \psi(\sigma^*)$*.*

Proof. Let $(\pi^*, \sigma^*, \hat{q}, \hat{p})$ denote an equilibrium. Then the low-ability Leader's net benefit from proposing $a = x$ is given by $\Delta(\pi^*, \sigma^*)$. We need to show that $\pi^* = \psi(\sigma^*)$.

Begin by supposing that $\Delta(0, \sigma^*) \leq 0$. And, by way of contradiction, suppose that $\pi^* \neq \psi(\sigma^*)$—i.e., $\pi^* > 0$. $\pi^* > 0$ implies that the low-ability Leader's net benefit from selecting $a = x$ is non-negative. Hence, $\Delta(\pi^*, \sigma^*) \geq 0$. To derive a contradiction, notice the following: By supposition, $\Delta(0, \sigma^*) \leq 0$, and, by part (a) of Lemma 2, $\Delta(\pi, \sigma^*)$ is decreasing in $\pi$. Together, these observations imply that $\Delta(\pi^*, \sigma^*) < 0$, a contradiction.

Now suppose that $\Delta(0, \sigma^*) > 0$. Notice that $\pi^*$ cannot equal 0 or 1. If $\pi^* = 0$, then $\Delta(0, \sigma^*) \leq 0$, which yields a contradiction. If $\pi^* = 1$, then $\Delta(1, \sigma^*) \geq 0$. However, by part (a) of Lemma 2, we know that $\Delta(1, \sigma^*) < 0$, which yields a contradiction. It follows that $\pi^* \in (0, 1)$, which implies that the low-ability Leader's net benefit from $a = x$ equals zero: $\Delta(\pi^*, \sigma^*) = 0$. Since the solution to $\Delta(\pi, \sigma^*) = 0$ in $\pi$ is unique, it follows that $\pi^* = \psi(\sigma^*)$. ■

**Proof of Proposition 1.**

Proof of part (a). The game without judicial review is strategically equivalent to the game with judicial review provided that the Judge employs a passive strategy (i.e., $\sigma_n = \sigma_x = 1$). Thus, by Lemma 3, the low-ability Leader's equilibrium strategy, $\pi^*_{norev}$, is uniquely defined, where $\pi^*_{norev} = \psi(1, 1)$.

All that remains to verify is that $\pi^*_{norev} \leq 1 - p$. By way of contradiction, suppose that $\pi^*_{norev} > 1 - p$. Because $\pi^*_{norev}$ is positive, the low-ability Leader's net benefit form proposing $a = x$ is weakly positive: $\Delta(\pi^*_{norev}, (1, 1)) \geq 0$. However, since $\pi^*_{norev} > 1 - p$, part (a) of Lemma 1 implies that $\Delta^e(\pi^*_{norev}, (1, 1)) < 0$. This fact, combined with the fact that $\Delta^p(1, 1) < 0$ (Lemma 1(b)), implies that $\Delta(\pi^*_{norev}, (1, 1)) < 0$, a contradiction.

Proof of part (b). Since, by definition,

$$\Delta(0, (1, 1)) = \alpha(1 - 2p) + (1 - \alpha)\left(F(1) - F\left(\frac{pq}{pq + (1 - q)}\right)\right),$$

it follows that $\Delta(0, (1, 1)) \leq 0$ if and only if

$$\frac{\alpha}{1 - \alpha}(2p - 1) \geq F(1) - F\left(\frac{pq}{pq + 1 - q}\right).$$

Thus, by Lemma 3, whenever the preceding inequality holds, $\pi^*_{norev} = 0$; in contrast, when the preceding inequality fails to hold, Lemma 3, taken together with part (a) of this proposition, implies that $\pi^*_{norev} \in (0, 1-p]$.

Proof of part (c). Define $\bar{\alpha} \equiv \frac{F(1)-F\left(\frac{pq}{pq+1-q}\right)}{2p-1+F(1)-F\left(\frac{pq}{pq+1-q}\right)}$. Part (b) of this proposition implies that $\pi^*_{norev} \in (0, 1-p]$ if $\alpha < \bar{\alpha}$ and $\pi^*_{norev} = 0$ if $\alpha \geq \bar{\alpha}$. Thus, to show that $\pi^*_{norev}$ is weakly decreasing in $\alpha$, it is sufficient to show that $\pi^*_{norev}$ is decreasing on $[0, \bar{\alpha})$.

So, suppose $\alpha \in [0, \bar{\alpha})$. Since $\pi^*_{norev}$ is a non-degenerate probability on this interval of $\alpha$, it follows that $\Delta(\pi^*_{norev}, (1,1)) = 0$. This fact, taken together with the fact that $\Delta^p(1,1) < 0$ (Lemma 1(b)), implies that $\Delta^e(\pi^*_{norev}, (1,1)) \geq 0$. Consequently, $\frac{\partial \Delta(\pi^*_{norev},(1,1))}{\partial \alpha} = \Delta^p(1,1) - \Delta^e(\pi^*_{norev}, (1,1)) < 0$. The fact that $\Delta$ is differentiable in $\alpha$, taken together with the fact that $\frac{\partial \Delta}{\partial \pi} < 0$ (Lemma 2(a)), allows us to apply the Implicit Function Theorem to sign the effect of a change in $\alpha$ on $\pi^*_{norev}$. Applying the Implicit Function Theorem, it follows that $\frac{\partial \pi^*_{norev}}{\partial \alpha} = \frac{-\frac{\Delta(\pi^*_{norev},(1,1))}{\partial \alpha}}{\frac{\partial \Delta(\pi^*_{norev},(1,1))}{\partial \pi}}$. Since the numerator is positive and the denominator is negative, $\frac{\partial \pi^*_{norev}}{\partial \alpha} < 0$.

We now show that when $\alpha = 0$, $\pi^*_{norev} = 1-p$: Suppose that $\alpha = 0$. Then it follows that $\Delta(\pi, (1,1)) = \Delta^e(\pi, (1,1))$. Notice that $\Delta^e(0, (1,1)) > 0$ (Lemma 1(a)). Consequently, by Lemma 3, $\pi^*_{norev} \in (0,1)$, where $\pi^*_{norev}$ is the unique solution to $\Delta^e(\pi, (1,1)) = 0$ in $\pi$. Hence, by part (a) of Lemma 1, $\pi^*_{norev} = 1-p$. ■

**Proof of Proposition 2**.

Proof of part (a). Result is immediate.

Proof of part (b). Consider an equilibrium in which the Judge uses a strict strategy. Notice that $\Delta^e(\pi, (0,0)) = \Delta^e(\pi, (1,1))$ and that $\Delta^p(0,0) = 0$ (the latter equality follows from Lemma 1(b)). Thus, the low-ability Leader's net payoff from $a = x$ when the Judge uses a strict strategy is $\Delta(\pi, (0,0)) = (1-\alpha)\Delta^e(\pi, (1,1))$. Since $\Delta^e(0, (1,1)) > 0$ (Lemma 1(a)), it follows from Lemma 3 that $\pi^*_{strict}$ is the unique solution to $\Delta(\pi, (0,0)) = 0$ in $\pi$. Since $\Delta(\pi, (0,0)) = 0$ if and only if $\Delta^e(\pi, (1,1)) = 0$, it follows from part (a) of Lemma 1 that $\pi^*_{strict} = 1-p$.

All that remains to establish is that $\pi^*_{strict} \geq \pi^*_{norev}$. This follows from the fact that $\pi^*_{norev} \leq 1-p$ (Proposition 1(a)).

Proof of part (c). That the low-ability Leader's equilibrium strategy is uniquely defined when the Judge uses an active strategy follows from Lemma 3. That the ordering of $\pi^*_{act}$ and $\pi^*_{norev}$ is ambiguous can be seen from the following pair of examples. Consider a parameterization of our model in which $\alpha = 0$, $q = 0.5$, $p = 0.8$, $F(\hat{q}) = \hat{q}$, and $\gamma = 0.781$. Then, $\pi^*_{norev} = 0.2$ and

$\pi^*_{act} \approx 0.1326$. Next, consider a parameterization in which $\alpha = 0.2$, $q = 0.5$, $p = 0.8$, $F(\hat{q}) = \hat{q}$, and $\gamma = 0.781$. Then, $\pi^*_{norev} \approx 0.1195$ and $\pi^*_{act} \approx 0.13$.

**Proof of Proposition 3.**

Consider an equilibrium in which the low-ability Leader proposes $a = x$ with probability $\pi^*$. In such an equilibrium, the probability that the Judge assigns to $\omega = n$ when her signal of the state is $s$ and $a = x$ is $\hat{p}(s; \pi^*)$. Thus, the Judge's expected payoff from upholding $a = x$ is $1 - \hat{p}(s; \pi^*)$, whereas her expected payoff from overruling $a = x$ is $\hat{p}(s; \pi^*)$. As a result, when $\hat{p}(s; \pi^*) < \frac{1}{2}$, the Judge maximizes her expected payoff by upholding the Leader; in contrast, when $\hat{p}(s; \pi^*) > \frac{1}{2}$, the Judge maximizes her expected policy payoff by overruling the Leader; finally, when $\hat{p}(s; \pi^*) = \frac{1}{2}$, the Judge is indifferent between upholding and overruling the Leader. Now notice that $\hat{p}(n; \pi^*) \lesseqqgtr \frac{1}{2}$ iff $\pi^* \lesseqqgtr \underline{T}$, whereas $\hat{p}(x; \pi^*) \lesseqqgtr \frac{1}{2}$ iff $\pi^* \lesseqqgtr \overline{T}$. Parts (a) through (e) of this proposition follow from the preceding observations taken together with the fact that $\underline{T} < \overline{T}$. ■

**Proof of Proposition 4.**

Proof of part (a). That a passive equilibrium exists if and only if $\pi^*_{pass} \leq \underline{T}$ is a consequence of Proposition 3.

Part (a) of Proposition 2, together with part (a) of Proposition 1, implies that $\pi^*_{pass} \leq 1 - p$. Hence, a sufficient condition for the existence of passive equilibria is that $1 - p \leq \underline{T} = \left(\frac{q(1-p)}{1-q}\right)\frac{(1-\gamma)}{p-(1-\gamma)}$. This inequality holds if and only if $q \geq \frac{p+\gamma-1}{p}$. Therefore, passive equilibria exist if $q \geq \frac{p+\gamma-1}{p}$.

Now suppose that $\alpha = 0$. Part (a) of Proposition 2, taken together with Part (c) of Proposition 1, implies that $\pi^*_{pass} = 1 - p$. Thus, if passive equilibria exist, $1 - p \leq \underline{T}$. Hence, when $\alpha = 0$, a necessary condition for passive equilibria to exist is that $q \geq \frac{p+\gamma-1}{p}$.

Proof of part (b). That strict equilibria exist if and only if $\pi^*_{strict} \geq \overline{T}$ is a consequence of Proposition 3.

By part (b) of Proposition 2, $\pi^*_{strict} = 1 - p$. Hence, strict equilibria exist if and only if $1 - p \geq \overline{T} = \left(\frac{q(1-p)}{1-q}\right)\frac{\gamma}{p-\gamma}$. The preceding inequality holds if and only if $q \leq \frac{p-\gamma}{p}$. Hence, strict equilibria exist if and only if $q \leq \frac{p-\gamma}{p}$.

Proof of part (c). That active equilibria exist if and only if $\pi^*_{act} \in [\underline{T}, \overline{T}]$ is a consequence of Proposition 3.

Since $\pi^*_{act} \leq 1$, a necessary condition for the existence of an active equilibrium is that $1 \geq \underline{T}$.

This inequality holds if and only if $q \leq \frac{p+\gamma-1}{p\gamma}$. Hence, a necessary condition for the existence of active equilibria is that $q \leq \frac{p+\gamma-1}{p\gamma}$.

Proof of part (d).

*Existence.* That an equilibrium always exists is a consequence of the following argument: By part (a), if $\pi^*_{pass} \leq \underline{T}$, a passive equilibrium exists; by part (b), if $\pi^*_{strict} \geq \overline{T}$, a strict equilibrium exists; and, by part (c), if $\pi^*_{act} \in [\underline{T}, \overline{T}]$, an active equilibrium exists. So, consider the remaining possibility: a situation in which neither an active nor a passive nor a strict equilibrium exists. In other words, $\pi^*_{pass} > \underline{T}$ and $\pi^*_{act} \notin [\underline{T}, \overline{T}]$ and $\pi^*_{strict} < \overline{T}$. Provided that these conditions hold, we will show a semi-active equilibrium exists when $\pi^*_{act} < \underline{T}$ and a semi-strict equilibrium exists when $\pi^*_{act} > \overline{T}$.

Suppose $\pi^*_{act} < \underline{T}$ and $\pi^*_{pass} > \underline{T}$. As $0 \leq \pi^*_{act} < \underline{T} < \pi^*_{pass} < 1$, we have that $\Delta(\pi^*_{act}, (0,1)) \leq 0$ and $\Delta(\pi^*_{pass}, (1,1)) = 0$. These facts, taken together with the fact that $\Delta$ is decreasing in $\pi$ (Lemma 2(a)), imply that $\Delta(\underline{T}, (0,1)) < 0$ and $\Delta(\underline{T}, (1,1)) > 0$. This implication, taken together with the fact that $\Delta$ is continuous in $\sigma_n$, ensures that the Intermediate Value Theorem applies. Thus, the equation $\Delta(\underline{T}, (\sigma_n, 1)) = 0$ has a solution in $\sigma_n$ on $(0,1)$. Denoting this solution by $\bar{\sigma}_n$, it is easily verified that the strategy profile in which $\pi^* = \underline{T}$, $\sigma^*_n = \bar{\sigma}_n$, and $\sigma^*_x = 1$ (together with beliefs $\hat{q}(\cdot,\cdot;\pi^*,\sigma^*)$ and $\hat{p}(\cdot;\pi^*)$) constitutes an equilibrium.

Now suppose that $\pi^*_{act} > \overline{T}$ and $\pi^*_{strict} < \overline{T}$. As $0 < \pi^*_{strict} < \overline{T} < \pi^*_{act} < 1$, we have that $\Delta(\pi^*_{srict}, (0,0)) = 0$ and $\Delta(\pi^*_{act}, (0,1)) = 0$. These facts, taken together with the fact that $\Delta$ is decreasing in $\pi$, imply that $\Delta(\overline{T}, (0,0)) < 0$ and $\Delta(\overline{T}, (0,1)) > 0$. This implication, taken together with the fact that $\Delta$ is continuous in $\sigma_x$, ensures that the Intermediate Value Theorem applies. Thus, the equation $\Delta(\overline{T}, (0, \sigma_x)) = 0$ has a solution in $\sigma_x$ on $(0,1)$. Denoting this solution by $\bar{\sigma}_x$, it is easily verified that the strategy profile in which $\pi^* = \overline{T}$, $\sigma^*_n = 0$, and $\sigma^*_x = \bar{\sigma}_x$ (together with beliefs $\hat{q}(\cdot,\cdot;\pi^*,\sigma^*)$ and $\hat{p}(\cdot;\pi^*)$) constitutes an equilibrium.

*Multiplicity.* We now provide an example in which more than one judicial strategy is consistent with equilibrium behavior. Suppose that $\alpha = 0.2$, $q = 0.5$, $p = 0.8$, $F(\hat{q}) = \hat{q}$, and $\gamma = 0.651$. Then $\pi^*_{pass} \approx 0.119$, $\pi^*_{act} \approx .159$, $\underline{T} \approx 0.155$, $\overline{T} \approx 0.874$. Since $\pi^*_{pass} \leq \underline{T}$, a passive equilibrium exist. And since $\pi^*_{act} \in [\underline{T}, \overline{T}]$, an active equilibrium also exists.

*Example in which all equilibria involve the Judge using a mixed strategy.* We now provide an example in which no equilibrium exists in which the Judge uses a pure strategy. Suppose that $\alpha = 0$,

$q = 0.5$, $p = 0.8$, $F(\hat{q}) = \hat{q}$, and $\gamma = 0.601$. Then $\pi^*_{pass} = 0.2$, $\pi^*_{act} \approx 0.192$, $\pi^*_{strict} = 0.2$, $\underline{T} \approx 0.199$, $\overline{T} \approx 0.604$. Since $\pi^*_{pass} > \underline{T}$, there does not exist a passive equilibrium. Since $\pi^*_{strict} < \overline{T}$, there does not exist a strict equilibrium. And since $\pi^*_{act} < \underline{T}$, there does not exist an active equilibrium. The preceding observations, taken together with Proposition 3, imply that the Judge's equilibrium strategy must involve mixing. (For the specified parameters, one can further show that a semi-active equilibrium exists but a semi-strict equilibrium does not.)■

**Proof of Proposition 5.**

Proof of part (a). Suppose judicial review induces a passive equilibrium. Then the low-ability Leader proposes $a = x$ with probability $\pi^*_{pass}$, which is equivalent to $\pi^*_{norev}$. Hence, the Voter's current policy payoff with review is equivalent to that with no review.

Proof of part (b). Suppose judicial review induces a non-passive equilibrium in which the low-ability Leader's equilibrium strategy is $\pi^*$ and the Judge's equilibrium strategy is $(\sigma^*_n, \sigma^*_x)$.

Begin by noticing that the Voter's equilibrium current payoff with review and democratic failure level $\pi^*$, $V(\pi^*, (\sigma^*_n, \sigma^*_x))$, is at least as great as that without review and democratic failure level $\pi^*$, $V(\pi^*, (1, 1))$. This is because the Judge shares the Voter's policy preferences and overrules the Leader only when doing so would weakly improve the Voter's expected payoff. Thus, when $\pi^* = \pi^*_{norev}$, $V(\pi^*, (\sigma^*_n, \sigma^*_x)) \geq V(\pi^*_{norev}, (1, 1))$: The Voter's current policy payoff with review is at least as great as that without review. Now consider the case in which $\pi^* < \pi^*_{norev}$. Notice that

$$\frac{\partial V(\pi, (1, 1))}{\partial \pi} = -(1 - q)(2p - 1)$$

is negative, since $p > \frac{1}{2}$ and $q \in (0, 1)$. The fact that $V(\pi, (1, 1))$ is decreasing in $\pi$, taken together with the fact that $\pi^* < \pi^*_{norev}$, implies that $V(\pi^*, (1, 1)) > V(\pi^*_{norev}, (1, 1))$. This fact, taken together with the fact that $V(\pi^*, (\sigma^*_n, \sigma^*_x)) \geq V(\pi^*, (1, 1))$, implies that $V(\pi^*, (\sigma^*_n, \sigma^*_x)) > V(\pi^*_{norev}, (1, 1))$. Consequently, when non-passive review lowers democratic failure, the Voter's equilibrium current policy payoff with review is greater than that without review.

Proof of part (c). We now provide two examples in which judical review exacerbates democratic failure, one in which judicial review strictly decreases the Voter's current policy payoff and one in which review strictly increases the Voter's current policy payoff. Consider a parameterization of our model in which $\alpha = 0.2$, $q = 0.5$, $p = 0.8$, $F(\hat{q}) = \hat{q}$, and $\gamma = 0.701$. Then, $\pi^*_{norev} \approx 0.1195$. With review, the unique equilibrium is active, where $\pi^*_{act} \approx 0.1517$. Thus, the Voter's current policy payoff *without review* is $V(0.1195, (1,1)) \approx 0.8642$, whereas the Voter's current policy payoff *with review* is $V(0.1517, (0,1)) \approx 0.8626$. Consequently, in this example, the introduction of judicial review strictly decreases the Voter's current policy payoff. Now consider a parameterization identical to the preceding one with the exception of the accuracy of the Judge's signal $\gamma$: continue to fix $\alpha = 0.2$, $q = 0.5$, $p = 0.8$, and $F(\hat{q}) = \hat{q}$, but now set $\gamma = 0.781$. (Thus, relative to the preceding example, the accuracy of the Judge's signal is now higher.) Then, $\pi^*_{norev} \approx 0.1195$. With review, the unique equilibrium is active, where $\pi^*_{act} \approx 0.13$. Thus, the Voter's current policy payoff *without review* is $V(0.1195, (1,1)) \approx 0.8642$, whereas the Voter's current policy payoff *with review* is $V(0.1517, (0,1)) \approx 0.8769$. Thus, in this example, the introduction of judicial review strictly increases the Voter's current policy payoff despite exacerbating democratic failure. ■

We now prove three lemmas. Lemma 4 (parts (a) and (b)), Lemma 5, and Lemma 6 will be invoked in proving Proposition 6. Lemma 4 (parts (c) and (d)) will be invoked in our subsequent discussion of the high-ability Leader's incentives in the robustness section.

**Lemma 4** *Consider an equilibrium* $(\pi^*, \sigma^*, \hat{q}, \hat{p})$ *in which the Judge's strategy is either semi-active, active, or semi-strict.*

*(a) The probability the high-ability Leader is upheld upon proposing* $a = x$ *is greater than that of the low-ability Leader:* $\lambda(h, \sigma^*) > \lambda(l, \sigma^*)$.

*(b) The Leader's reputation upon being upheld is greater than that when overruled:* $\hat{q}(x, uphold; \pi^*, \sigma^*) > \hat{q}(x, strike; \pi^*, \sigma^*)$.

*(c) The low-ability Leader's net electoral payoff from proposing* $a = x$ *is less than the high-ability Leader's net electoral payoff from proposing* $a = x$ *when* $\omega = x$:

$$\lambda(l, \sigma^*)F(\hat{q}(x, uphold; \pi^*, \sigma^*)) + (1 - \lambda(l, \sigma^*))F(\hat{q}(x, strike; \pi^*, \sigma^*)) - F(\hat{q}(n; \pi^*)) <$$
$$\lambda(h, \sigma^*)F(\hat{q}(x, uphold; \pi^*, \sigma^*)) + (1 - \lambda(h, \sigma^*))F(\hat{q}(x, strike; \pi^*, \sigma^*)) - F(\hat{q}(n; \pi^*)). \quad (2)$$

*(d) The low-ability Leader's net electoral payoff from proposing* $a = x$ *is greater than the high-*

*ability Leader's net electoral payoff from proposing $a = x$ when $\omega = n$:*

$$\lambda(l,\sigma^*)F(\hat{q}(x, uphold; \pi^*, \sigma^*)) + (1 - \lambda(l,\sigma^*))F(\hat{q}(x, strike; \pi^*, \sigma^*)) - F(\hat{q}(n; \pi^*)) >$$
$$z(\sigma^*)F(\hat{q}(x, uphold; \pi^*, \sigma^*)) + (1 - z(\sigma^*))F(\hat{q}(x, strike; \pi^*, \sigma^*)) - F(\hat{q}(n; \pi^*)), \quad (3)$$

*where $z(\sigma^*) \equiv \sigma_x^*(1-\gamma) + \sigma_n^*\gamma$ is the probability the high-ability Leader is upheld when he proposes $a = x$, $\omega = n$, and the Judge's strategy is $\sigma^*$.*

*Proof:*

Proof of part (a). Consider an equilibrium in which the Judge's strategy is either semi-active, active, or semi-strict and denote the Judge's strategy by $(\sigma_n^*, \sigma_x^*)$. If the Judge's strategy is semi-active, then $\sigma_n^* \in (0, 1)$ and $\sigma_x^* = 1$. If the Judge's strategy is active, then $\sigma_n^* = 0$ and $\sigma_x^* = 1$. And if the Judge's strategy is semi-strict, then $\sigma_n^* = 0$ and $\sigma_x^* \in (0, 1)$. Consequently, in either case, $\sigma_x^* > \sigma_n^*$. Thus, it follows (via algebra) that $\lambda(h,\sigma^*) > \lambda(l,\sigma^*)$.

Proof of part (b). Consider an equilibrium in which the Judge's strategy is either semi-active, active, or semi-strict. Thus, the Leader is both upheld and overruled with positive probability, which implies that in each instance the Voter's posterior about the Leader's ability can be derived via Bayes' Rule. Now notice that $\hat{q}(x, uphold; \pi^*, \sigma^*) > \hat{q}(x, strike; \pi^*, \sigma^*)$ if and only if $\lambda(h,\sigma^*) > \lambda(l,\sigma^*)$. Thus, by part (a), our desired conclusion follows.

Proof of part (c). That inequality (2) holds follows from parts (a) and (b) of this lemma, taken together with the fact that $F$ is increasing in $\hat{q}$.

Proof of part (d). Consider an equilibrium in which the Judge's strategy is either semi-active, active, or semi-strict and denote the Judge's strategy by $(\sigma_n^*, \sigma_x^*)$. Then $\sigma_x^* > \sigma_n^*$. It thus follows (via algebra) that $\lambda(l,\sigma^*) > z(\sigma^*)$. This fact, taken together with the fact that $\hat{q}(x, uphold; \pi^*, \sigma^*) > \hat{q}(x, strike; \pi^*, \sigma^*)$ (see part (b) of this lemma) and the fact that $F$ is increasing in $\hat{q}$, implies that inequality (3) holds. ■

**Lemma 5** *Consider an equilibrium in which the Judge's strategy is either semi-active, active, or semi-strict. In addition, suppose that $\alpha = 0$ and that the electoral strength function $F$ is concave. The equilibrium probability with which low-ability Leader selects the extraordinary action, $\pi^*$, is less than that with no review, $\pi^*_{norev}$.*

Proof. Suppose $\alpha = 0$ and the electoral strength function $F$ is concave. Also, suppose that judicial review induces a semi-active, active, or semi-strict equilibrium. Finally, denote the Judge's strategy in this equilibrium by $(\sigma_n^*, \sigma_x^*)$.

Because $\alpha = 0$, it follows that $\Delta(\pi, \sigma) = \Delta^e(\pi, \sigma)$. Also, notice that when $\pi = 0$, it follows that $\Delta^e(0, \sigma) = 1 - \frac{pq}{pq+(1-q)} > 0$. Since $\Delta^e(0, \sigma) > 0$, it follows from Lemma 3 that in any equilibrium in which the Judge uses strategy $\sigma$, the low-ability Leader's equilibrium strategy is the unique solution to $\Delta^e(\pi, \sigma) = 0$ in $\pi$ on (0,1). Thus, in the absence of review, the low-ability Leader's equilibrium strategy, $\pi^*_{norev}$, is the solution to $\Delta^e(\pi, (1, 1)) = 0$ in $\pi$. And with review, the low-ability Leader's equilibrium strategy, $\pi^*$, is the solution to $\Delta^e(\pi, (\sigma_n^*, \sigma_x^*)) = 0$ in $\pi$.

We need to show that $\pi^*_{norev} > \pi^*$. To do so, it is sufficient to show that for all $\pi \in (0, 1)$, $\Delta^e(\pi, (1, 1)) > \Delta^e(\pi, (\sigma_n^*, \sigma_x^*))$. This is equivalent to showing that for all $\pi \in (0, 1)$,

$$F(\hat{q}(x; \pi)) > \lambda(l, \sigma^*) F(\hat{q}(x, uphold; \pi, \sigma^*)) + (1 - \lambda(l, \sigma^*)) F(\hat{q}(x, strike; \pi, \sigma^*)) \tag{4}$$

We now turn to showing that inequality (4) in fact holds. Write $Pr(uphold; \pi, \sigma^*)$ for the equilibrium probability the Leader is upheld conditional upon $a = x$ having been proposed. Hence, $Pr(uphold; \pi, \sigma^*) = \hat{q}(x; \pi)\lambda(h, \sigma^*) + (1 - \hat{q}(x; \pi))\lambda(l, \sigma^*)$. Now notice that by the Martingale property of Bayesian posteriors,

$$\hat{q}(x; \pi) = Pr(uphold; \pi, \sigma^*)\hat{q}(x, uphold; \pi, \sigma^*) + (1 - Pr(uphold; \pi, \sigma^*))\hat{q}(x, strike; \pi, \sigma^*).$$

Since $\lambda(h, \sigma^*) > \lambda(l, \sigma^*)$ (Lemma 4(a)), $Pr(uphold; \pi, \sigma^*) > \lambda(l, \sigma^*)$. This fact, taken together with the fact that $\hat{q}(x, uphold; \pi, \sigma^*) > \hat{q}(x, strike; \pi, \sigma^*)$ (Lemma 4(b)), implies that

$$\hat{q}(x; \pi) = Pr(uphold; \pi, \sigma^*)\hat{q}(x, uphold; \pi, \sigma^*) + (1 - Pr(uphold; \pi, \sigma^*))\hat{q}(x, strike; \pi, \sigma^*) > \lambda(l, \sigma^*)\hat{q}(x, uphold; \pi, \sigma^*) + (1 - \lambda(l, \sigma^*))\hat{q}(x, strike; \pi, \sigma^*).$$

This fact, taken together with the fact that $F$ is increasing in $\hat{q}$, implies that

$$F(\hat{q}(x; \pi)) > F\big(\lambda(l, \sigma^*)\hat{q}(x, uphold; \pi, \sigma^*) + (1 - \lambda(l, \sigma^*))\hat{q}(x, strike; \pi, \sigma^*)\big).$$

The concavity of $F$ implies that

$$F\big(\lambda(l,\sigma^*)\hat{q}(x,uphold;\pi,\sigma^*)+(1-\lambda(l,\sigma^*))\hat{q}(x,strike;\pi,\sigma^*)\big)>$$
$$\lambda(l,\sigma^*)F(\hat{q}(x,uphold;\pi,\sigma^*))+(1-\lambda(l,\sigma^*))F(\hat{q}(x,strike;\pi,\sigma^*)).$$

The preceding two inequalities, taken together, imply that $F(\hat{q}(x;\pi)) > \lambda(l,\sigma^*)F(\hat{q}(x,uphold;\pi,\sigma^*)) +(1-\lambda(l,\sigma^*))F(\hat{q}(x,strike;\pi,\sigma^*))$. Consequently, for all $\pi \in (0,1)$, inequality (4) holds. ■

**Lemma 6** *Define $\pi^*(\alpha)$ to be the probability with which the low-ability Leader selects $a = x$ in an equilibrium in which the Judge employs strategy $\sigma^*$ and the weight attached to policy is $\alpha$. $\pi^*(\alpha)$ is continuous in $\alpha$ on an open neighborhood of 0.*

Proof. Suppose an equilibrium exist in which the Judge uses strategy $\sigma^*$. In addition, suppose that $\alpha = 0$. Because $\alpha = 0$, it follows that $\Delta(\pi,\sigma^*) = \Delta^e(\pi,\sigma^*)$. Also, notice that when $\pi = 0$, $\Delta^e(0,\sigma^*) = 1 - \frac{pq}{pq+(1-q)} > 0$. Since $\Delta(0,\sigma^*) > 0$, it follows from Lemma 3 that in any equilibrium in which the Judge uses strategy $\sigma^*$, the low-ability Leader's equilibrium strategy, $\pi^*(0)$, is the unique solution to $\Delta(\pi,\sigma^*) = 0$ in $\pi$ on (0,1). The fact that $\Delta(\pi^*(0),\sigma^*) = 0$, taken together with the fact that $\Delta$ is differentiable in $\alpha$ and the fact that $\frac{\partial\Delta}{\partial\pi} \neq 0$ (Lemma 2(a)), enables us to invoke the Implicit Function Theorem in order to conclude that $\pi^*(\alpha)$ is continuous in $\alpha$ on an open neighborhood of 0. ■

**Proof of Proposition 6.**

Proof of part (a). Suppose $q \in \left(\frac{p-\gamma}{p}, \frac{p+\gamma-1}{p}\right)$ and that $F$ is concave. That $q > \frac{p-\gamma}{p}$ implies strict equilibria never exist (Proposition 4(b)). Hence, with review, equilibria must be passive, semi-active, active, or semi-strict.

We need to show that the Voter's current policy payoff with review is strictly greater than that without review when $\alpha$ is sufficiently small. A sufficient condition for review to strictly increase the Voter's current policy payoff is that review strictly decrease democratic failure (Proposition 5(b)). Thus, we first show that when $\alpha = 0$, the probability with which the low-ability Leader selects $a = x$ is strictly less than that with no review. We then show that this remains true provided $\alpha$ is sufficiently small.

So, suppose that $\alpha = 0$. This supposition, taken together with our assumption that $q < \frac{p+\gamma-1}{p}$, implies that passive equilibria do not exist (Proposition 4(a)). Hence, equilibria are either semi-active or active or semi-strict. Therefore, the equilibrium probability with which the low-ability Leader selects $a = x$ with review when $\alpha = 0$, $\pi^*(0)$, is an element of the set $\{\underline{T}, \pi^*_{act}(0), \overline{T}\}$. Further, the concavity of $F$, taken together with the fact that with review equilibria are either semi-active or active or semi-strict, allows us to apply Lemma 5. Hence, the equilibrium probability with which the low-ability Leader selects $a = x$ with review, $\pi^*(0)$, is less than that without review, $\pi^*_{norev}(0)$. Consequently,

$$\max\{\underline{T}, \pi^*_{act}(0), \overline{T}\} < \pi^*_{norev}(0).$$

Since $\underline{T}$ and $\overline{T}$ are independent of $\alpha$ (Remark 2), and $\pi^*_{norev}(\alpha)$ and $\pi^*_{act}(\alpha)$ are both continuous in $\alpha$ on an open neighborhood of 0 (Lemma 6), it follows that for $\alpha$ sufficiently small

$$\max\{\underline{T}, \pi^*_{act}(\alpha), \overline{T}\} < \pi^*_{norev}(\alpha). \tag{5}$$

This implies that the equilibrium probability with which the low-ability Leader selects $a = x$ with review, $\pi^*(\alpha)$, is less than that with no review, $\pi^*_{norev}(\alpha)$, provided $\alpha$ is sufficiently small. To see why, suppose that inequality (5) holds. Since $\pi^*_{norev}(\alpha) = \pi^*_{pass}(\alpha)$ (Proposition 2(a)), it follows that $\pi^*_{pass}(\alpha) > \underline{T}$. Thus, by part (a) of Proposition 4, passive equilibria do not exist. Consequently, the equilibrium probability with which the low-ability Leader selects $a = x$, $\pi^*(\alpha)$, is an element of $\{\underline{T}, \pi^*_{act}(\alpha), \overline{T}\}$. This fact, taken together with inequality (5), implies that $\pi^*(\alpha) < \pi^*_{norev}(\alpha)$.

Proof of part (b). Suppose $q \leq \frac{p-\gamma}{p}$. Thus, by part (b) of Proposition 4, a strict equilibrium exists. In addition, suppose judicial review induces the strict equilibrium. As a result, the policy that results from the interaction between the Judge and the Leader is always the normal policy. Hence, with review, the Voter's current policy payoff equals $p$, the prior probability that $\omega = n$. Now consider the Voter's current policy payoff without review. This payoff equals $V(\pi^*_{norev}(\alpha), (1, 1)) = q + (1-q)(p(1-\pi^*_{norev}(\alpha)) + (1-p)\pi^*_{norev}(\alpha))$, where $\pi^*_{norev}(\alpha)$ is the equilibrium probability with which the low-ability Leader select the extraordinary action when the weight he attaches to policy is $\alpha$ and there is no review. Consequently, the Voter's current policy payoff with review is less than that with no review provided $\pi^*_{norev}(\alpha) < \frac{q(1-p)}{(2p-1)(1-q)}$. Notice that an immediate implication of part (b) of Proposition 1 is $\pi^*_{norev}(\alpha) = 0$ if $\alpha \geq \bar{\alpha} \equiv \frac{F(1)-F\left(\frac{pq}{pq+1-q}\right)}{2p-1+F(1)-F\left(\frac{pq}{pq+1-q}\right)}$. Thus, for all $\alpha > \bar{\alpha}$, the Voter's current policy payoff with review is strictly less than that with no review. ■

## Robustness

In the main text, we assumed policy $a = n$ is not justiciable. This is equivalent to assuming that the Judge always upholds the Leader when $a = n$ is proposed. In addition, we assumed that the high-ability Leader matches policy to the state (i.e., selects $a = \omega$). We now establish the following two claims:

(a) Given that the high-ability Leader matches policy to the state, upholding when $a = n$ is consistent with equilibrium behavior for the Judge. (Hence, our assumption that only $a = x$ is justiciable is benign.)

(b) Matching policy to the state is consistent with equilibrium behavior for the high-ability Leader.

We begin with claim (a). Consider an equilibrium in which the high-ability Leader matches policy to the state and the low-ability Leader proposes the extraordinary action with probability $\pi^*$. We need to show that upholding the Leader when $a = n$ is optimal for the Judge regardless of her signal of the state. That is, we need to show that for any signal $s \in \{n, x\}$, the Judge's posterior that $\omega = n$ when $a = n$ is greater than 1/2. By Bayes' Rule,

$$Pr(\omega = n|s, a = n)) = \frac{Pr(s|\omega = n)(q + (1-q)(1-\pi^*))p}{Pr(s|\omega = n)(q + (1-q)(1-\pi^*))p + Pr(s|\omega = x)(1-q)(1-\pi^*)(1-p)}.$$

Since $Pr(\omega = n|s = n, a = n) > Pr(\omega = n|s = x, a = n)$, it is sufficient to show that $Pr(\omega = n|s = x, a = n) > \frac{1}{2}$. Since $p > \gamma$ (by assumption),

$$Pr(\omega = n|s = x, a = n) > \frac{1}{2} \Leftrightarrow \pi^* < \frac{1}{(1-q)} \frac{(p-\gamma) + q\gamma(1-p)}{(p-\gamma)}.$$

Finally, notice that $\pi^* < \frac{1}{(1-q)} \frac{(p-\gamma)+q\gamma(1-p)}{(p-\gamma)}$, as the left hand side is a probability and the right hand side is strictly greater than 1.

We now turn to claim (b). In our baseline model, we took it as given that the high-ability Leader matched policy to the state and solved for the equilibrium of the resulting game between the low-ability Leader, Judge and Voter. It turns out that in an equilibrium of this resulting game, matching policy to the state is in fact incentive compatible for the high-ability Leader. To develop intuition as to why this is so, consider the case in which the equilibrium to the game between the low-ability Leader, Judge and Voter is active. In such an equilibrium, $\sigma_n^* = 0$, $\sigma_x^* = 1$, and $\pi^* \in (0,1)$. Since $\pi^*$ is a non-degenerate probability, the low-ability Leader's net payoff from proposing $a = x$, $\Delta(\pi^*, \sigma^*) = \alpha\Delta^p(\sigma^*) + (1-\alpha)\Delta^e(\pi^*, \sigma^*) = 0$. Now consider the incentives of the high-ability Leader. Unlike the low-ability Leader, he knows the state of the world when proposing policy. This has two implications:

1. The net policy gain to the high-ability Leader from proposing $a = x$ when $\omega = x$ ($\omega = n$) is greater (less) than $\Delta^p(\sigma^*)$.

2. The net electoral gain to the high-ability Leader form proposing $a = x$ when $\omega = x$ ($\omega = n$) is greater (less) than $\Delta^e(\pi^*, \sigma^*)$ (Lemma 4, parts (c) and (d)).

The preceding two implications, taken together with the fact $\Delta(\pi^*, \sigma^*) = 0$, imply that the high-ability Leader has a strict incentive to propose $a = x$ ($a = n$) when $\omega = x$ ($\omega = n$).

While we know that with or without review, there always exists an equilibrium in which the high-ability Leader matches policy to the state, for completeness, we note that for certain parameterizations of our model, equilibria exist in which the high-ability Leader fails to match policy to the state. Examples include the following:

- With or without review, when $\alpha$ is sufficiently small, there exists an equilibrium in which both the high- and low-ability Leader always propose $a = n$. This equilibrium is supported by having the Voter believe that any Leader who makes the off-path proposal of $a = x$ is of low ability. (Such equilibria, however, do not survive refinements that are in the spirit of universal divinity.)

- Suppose that there exists an active equilibrium with review, where the high-ability Leader matches policy to the state and the low-ability Leader selects $a = x$ with probability $\pi^*_{act}$. If $a = n$ is justiciable, then there also exists a payoff-equivalent equilibrium in which the high-ability Leader mismatches policy to the state (i.e., he selects $a = n$ when $\omega = x$ and selects $a = x$ when $\omega = n$), the low-ability Leader selects $a = x$ with probability $(1 - \pi^*_{act})$, the Judge always strikes down the proposal of $a = x$ regardless of her signal, and when $a = n$ is proposed, she strikes it down if and only if her signal $s = x$.