

# A Model of the Endogenous Development of Judicial Institutions\*

Clifford J. Carrubba

Associate Professor of Political Science  
Emory University

Why do sovereign governments create judicial institutions and grant these institutions the power to rule their actions invalid? Once such a court is created, under what conditions is that court able to rule against these governments and get compliance with their rulings? Finally, how might the influence of the court change over time? This study presents a general theory of judicial institutions that provides a unified answer to these three questions. I argue that governments create judicial institutions to help solve collective action problems endemic to operating under a common regulatory regime. Once established, a court in its institutional “infancy” is capable of facilitating compliance with that regulatory regime’s rules but only consistent with the purpose for which the governments created the court. And finally, once the court has earned the trust of a government’s public, its ability to enforce the regulatory regime’s rules qualitatively expands. In demonstrating this last point, I derive how and why public’s can rationally come to believe that supporting a court, if there was a conflict between its government and the court, would be in that public’s interest.

---

\* I would like to thank Randy Calvert, Mike Giles, Mark Hallerberg, Eric Reinhardt, Craig Volden, Tom Walker, and Suzanne Werner for their helpful feedback on this project.

In 1787, the thirteen American States convened a Constitutional Convention with the goal of redressing the failings of the Articles of Confederation. In particular, the states had come to realize that the weak central institutions created under the Articles were incapable of ensuring desired cooperation among the states. For example, consider the issue of interstate commerce. While the Articles of Confederation established limits on the regulation of interstate trade, the national government proved incapable of enforcing those limits.<sup>1</sup> As a result, though state officials recognized the benefits they would receive from a system of free interstate trade, they found it hard to resist the short-term incentives to defect from this regime by imposing tax-based and other regulatory barriers.

To help facilitate cooperation with a newly strengthened national regulatory regime, the Constitutional Convention created the United States Supreme Court (USSC) and granted it the authority to rule upon possible violations of that regulatory regime's rules.<sup>2</sup> At the same time, however, the convention gave the court no ability to enforce its own rulings. As famously stated, the court had neither the "the purse nor the sword". Thus, the Founders constructed a court that could rule on these issues, but not a court that could directly enforce them. The unwillingness to concede any form of enforcement powers to the court was to a large degree a product of the fundamental tension of the Constitutional Convention; the Founders were trying to balance the desire to coordinate state activity with the demands to protect state sovereignty.

---

<sup>1</sup> Hamilton states, "the want of a power to regulate commerce is by all parties" seen as a "defect" of the Articles of Confederation (1788/1999:111).

<sup>2</sup> Of course, the convention also created a strengthened Congress and executive with an eye towards more effective national policy. And, the court was partly seen as a possible counter-balance to those institutions.

This inability to enforce its decisions proved problematic for the court for much of the 1800s. While the court did assert the right to judicial review over both state and federal law, the court tended to be cautious in its exertion of that authority. At the federal level, while the court asserted federal judicial review as early as 1803 in *Marbury v. Madison* (U.S. 1803), it was decades before the court actually used that power to annul a government action. At the state level, the court was much quicker to strike down state laws as unconstitutional (*Fletcher v. Peck*, U.S. 1810). However, even in legal areas as central to the federal enterprise as interstate commerce, evidence suggests that the court established doctrine that allowed states to burden trade contrary to the court's sincere preferences specifically because the court was worried about state compliance with its rulings (see author 2003). And, as Georgia's refusal to comply with the Court's ruling in *Worcester v. Georgia* (U.S. 1832) demonstrates, the threat of noncompliance by states was credible.

However, as of today, these concerns with state and federal compliance with court rulings have largely subsided. The perceived legitimacy of the court is high and stable,<sup>3</sup> and we observe the court ruling on highly contentious issues with no evident concern over whether its rulings will be followed (e.g. *McCreary County, Kentucky v. ACLU*, U.S. 2005; *Gonzales v. Raich*, U.S. 2005; *Bush v. Gore*, U.S. 2000). Thus, while the lack of enforcement powers once limited the Court's ability to rule on government action, this state of affairs seems to no longer hold. Something important has changed.

The USSC is far from the only court to have been formally empowered to rule over the legitimacy of government action, but denied any ability to directly enforce its decisions. We observe this tension across a wide array of national and international courts, including the

---

<sup>3</sup> But see Caldeira and Gibson (1992) for analysis of variation in public support for the USSC.

German Constitutional Court, the European Court of Justice, the World Trade Organization's arbitrating body, and the Russian Constitutional Court just to name a few. The degree to which these courts seem to have the authority to successfully constrain government behavior varies tremendously. On one extreme are courts, like the German Constitutional Court, that have more limited concerns with government compliance; in the middle are courts, like the WTO arbitrating body and the ECJ, that do not yet have the stature of established national courts, and on the other extreme are courts, like the Russian Constitutional Court, that appear cowed by their national governments.

These observations highlight a clear puzzle. In particular, how does a court that is incapable of enforcing its decisions develop into a court that can act *as if* its decisions are binding? And, more generally, what explains why some courts succeed and some courts fail in making this transition? To answer these questions, I present a formal model of the endogenous development of judicial influence. This model will characterize why a set of sovereign political actors might agree to a common regulatory regime, why they would choose to create a court to help facilitate enforcement of that regime's rules, how the court can (and cannot) facilitate compliance with the regime's rules in its institutional "infancy", and how the role of the court might (or might not) evolve over time into a "mature" court that can act (in general) as if its rulings are binding upon even sovereign political actors. Thus, this model will not only provide an explanation of court influence, but it will also endogenously explain the creation and evolution of the institution.

The first principle used in this model to explain judicial evolution builds directly off of recent insights in the literature on judicial influence. In general, models of judicial influence tend to assume that judicial rulings, at least at the time they are made, are binding. A decision might

be over-turned by subsequent legislation or constitutional amendment, but, absent one of those actions, the decision stands.<sup>4</sup> More recently, however, scholars have been asking why we can treat court rulings *as if* they are binding, when we know that these courts cannot enforce their decisions. An increasingly popular answer to this question is that the public acts as an indirect enforcement mechanism for the court.<sup>5</sup> For example, Vanberg (2005) argues that the German Constitutional Court does worry about whether the German federal government will comply with its rulings. However, the severity of the concern depends upon the German public's awareness of the case; the more aware the public, the less likely the German government will be to ignore an adverse ruling and, thus, the freer the court is to rule as it wishes. Similarly, in a study of the Mexican Supreme Court, Staton (2006) demonstrates that the Court actually tries to heighten public awareness in contentious cases in order to ensure that the Mexican government will comply with its decisions.

Of course, these works leave unexplained why we should expect the public to back the court. Two of the most important recent works on this question are Weingast (1997) and Stephenson (2004). Weingast (1997) argues that domestic publics support the rule of law (e.g. legal constraint on government action) to prevent being exploited by their sovereign. Stephenson (2004), in contrast, argues that the public supports a court when it believes its policy preferences are more closely aligned to those of the court than to those of their elected government. While

---

<sup>4</sup> This assumption holds for a wide class of separation of powers and principal-agent models of judicial influence (e.g. Ferejohn and Weingast 1992; Gely and Spiller 1990; Hansford and Damore 2000; Harvey and Friedman 2006; Rogers, 200; Shipan 1997, 2000).

<sup>5</sup> See Stephenson (2003), Landes and Posner (1975), Ramseyer (1994), Salzberger (1993), for a set of elite-based explanations of judicial independence.

these arguments are very different in some ways, both argue that public support for the court fundamentally is based upon self-interest.<sup>6</sup>

This study build off of these scholars' insights by assuming that the public can, if it wishes, act as an indirect enforcement mechanism for the court when the court is ruling on government behavior. The model then derives under what conditions judicial influence over government behavior will change as a result of endogenous changes in public support for the court.

Importantly, in answering this question, the model necessarily addresses the following subsidiary puzzles as well. First, what can the court do to maximize its chances of gaining influence? All a court can do is rule against a government or not. One might suppose that seeing court decisions flouted by that government would undermine its public support, while seeing its rulings obeyed would do the opposite. But why? How do those observations cause the public to believe that supporting the court is in its own interests? Second, why would sovereign governments allow such a transition to take place? Governments control the rules of the game under which the court operates. Why do not governments simply remove from the court the right to judicial review if they fear the court is becoming too powerful? And third, why would the public ever back the court over its own elected government? Presumably the government reflects majoritarian preferences. What would cause the public to believe that it would be better served having its directly elected agent curtailed by an unelected body?<sup>7</sup>

---

<sup>6</sup> See Ferejohn, Rosenbluth, and Shipan (2004) for a useful review of the courts literature.

<sup>7</sup> Note that both Weingast and Stephenson stop short of explaining why a public would transition from not supporting the rule of law, or the court, over their government, to doing so. In Weingast's work, whether the coordination problem is resolved or not depends upon which of

Before proceeding, two points are worth noting. First, this is not explicitly a study of the development of judicial review over federal law. Rather, it is focused on how courts gain influence over a collective of sovereign states. However, as some scholars have noted (e.g. Whittington 2005), in many ways, establishing review over state law has been a necessary, and substantial, precursor to establishing review over federal law, at least in the U.S. context. As a result, the model does have several interesting implications for the development of judicial review at the federal level. These implications are discussed in detail later. Second, this model departs from previous efforts to model the development of political institutions by fully endogenizing institutional change. It neither relies upon semi-exogenous parameters as done in Greif and Laitin (2004), nor upon a veil of ignorance story in which institutions are fixed before the game starts (e.g. Author 2001). Thus, this study contributes to the literature on institutional change, by deriving fully endogenous, in equilibrium changes in the institutional rules of the game.<sup>8</sup>

---

multiple equilibria is played. Thus, this model cannot explain how a society would transition from not respecting the rule of law to doing so in equilibrium. In Stephenson's model, whether the public backs the court is a function of a set of exogenous parameters of the model. Thus, the model does not endogenously explain why a public would transition from not supporting a court against its government to doing so. That is, while the model can be used to argue that the exogenous conditions have changed over time, it cannot be used to endogenously derive why those exogenous conditions have changed. This study picks up where these studies leave off.

<sup>8</sup> Also see work by Calvert (e.g. 1995) for an iterated game approach to explaining the purpose of social institutions.

The rest of the paper proceeds as follows. Part two presents the model. Part three derives equilibrium behavior. Part four present implications and conclusions.

### **A Model of the Evolution of Judicial Institutions**

The following model is designed to address the previously described empirical puzzle. The model consists of three main elements, two governments operating under a common regulatory regime, a court created to help facilitate compliance with that regulatory regime's rules, and a public – associated with each government – that cares about, but is also imperfectly informed about, how the regulatory regime affects its interests. Most of the assumptions used to generate the model directly follow from the empirical motivation. Where important additional assumptions are made further motivation is provided. I first provide an intuitive characterization of the model and then formalize it.

Suppose two governments agree to operate under a common regulatory regime. The purpose of this regime is to help the states overcome a collective action problem. In particular, the regulatory regime consists of a set of rules that the states recognize generally benefit them if everyone follows, but that they each have a short term incentive to violate at least occasionally. How strong the incentive is to violate the agreement will vary over time, and this incentive is not common knowledge.<sup>9</sup> If one state does violate the agreement, the other state does not receive the benefit it would otherwise accrue from participating in the regulatory regime. For example, suppose a set of states decide that it is in their common interest to establish a free trade zone.

---

<sup>9</sup> By not assuming common knowledge, I allow a government to have a better sense of its incentive to defect than the partner government. If the incentive to defect were common knowledge, the governments could perfectly self-adjudicate disputes (see Author 2005).

While a state government might have domestic reasons not to levy a tariff on imported goods, that government at least occasionally will have a political incentive to burden trade. How strong that incentive is can be only imperfectly known by the trading partner. Burdening trade denies the trading partners the benefit of participating in the free trade regime, because their exporting firms no longer have privileged access to the foreign market.

To help facilitate compliance with this regime, the governments create a court and allow some set of litigants standing before the court. Standing could be as restrictive as with the WTO, in which only governments can bring cases, or as permissive as the US, in which private litigants can bring cases and make appeals all the way up to the Supreme Court.<sup>10</sup> If a case is brought, the litigants make their arguments, relevant information about the case is revealed, and the court issues its ruling.<sup>11</sup> Thus, the court acts as a fire alarm and information clearinghouse. It acts as a fire alarm by providing a venue in which interested parties can bring cases and signal possible violations of the regulatory regime's rules.<sup>12</sup> It acts as an information clearinghouse by providing a venue in which the facts of the case can be vetted.

Importantly, the governments have not granted the court the power to enforce its decisions. Thus, if a government is ruled in violation of the regulatory regime's rules, that government decides whether it will choose to abide by the decision or not.

---

<sup>10</sup> See Author (2005) for further discussion of how to think about modeling litigants.

<sup>11</sup> Presumably, making arguments helps convey information. If not, no one would know more about the potential violation of the regulatory regime's rules than before the hearing, and adjudication would serve no purpose.

<sup>12</sup> Creating a fire alarm mechanism can be beneficial for governments, because it diffuses the costs of monitoring.

Associated with each government is a public. While members of the public are not necessarily very interested in, let alone aware of, any given case, they do care about how the regulatory regime as a whole affects their lives. Ideally, individuals would like the regulatory regime's rules followed when it serves their interests, but not otherwise. However, short of that, they would like the rules applied such that, in expectation, those rules benefit them as much as possible. How beneficial the regulatory regime is, on average, is unclear to the public.<sup>13</sup>

The preferences of the public are correlated with, but not perfectly reflected by, the preferences of their government.<sup>14</sup> Further, the public only gets to experience the outcome associated with an application of the regulatory regime's rules if the government actually chooses to comply. If the government does not comply, the public might have some beliefs over whether that application of the regulatory regime's rules would have been beneficial, but they cannot know with certainty. Thus, a public cannot always know that its government is acting in its best interests. Rather, if it observes noncompliance, it must make an assessment, and, if it finds the government's behavior wanting, sanction it.<sup>15</sup>

---

<sup>13</sup> The public's uncertainty over the benefit of the regulatory regime allows the public to update its beliefs about how much it likes that regime, and therefore the kinds of behavior it would like its government and the court to engage in with respect to that regime's rules.

<sup>14</sup> Think of the public and government as in a principal-agent relationship. The government is the public's agent, and while the public wants to put into office an agent that represents its preferences, it will always do so imperfectly.

<sup>15</sup> Beyond being realistic, the fact that the public is imperfectly informed is an important feature of the model. If completely informed, the public could perfectly sanction its government for not

### *Formalization*

The following model formalizes the intuitive game described above. Figure 1 illustrates the sequence of events, and table 1 provides a summary of the notation discussed below.<sup>16</sup> Each period starts with governments privately drawing a value of compliance,  $\chi_{t,i}$ . The subscript  $t$  is used to indicate the period of the game and the subscript  $i=1,2$  indicates the government. Note that, unless necessary for exposition, I suppress the government subscript to minimize notation. This cost consists of two components. One component,  $\beta_t$ , is the overall societal benefit of complying with the regulatory regime's rules in that period.  $\beta_t$  is an unbounded random variable, distributed with some mean,  $\mu_\beta$ , and variance,  $\sigma_\beta$ . This term is the common factor in the payoffs for a government and its public. The second component,  $c_t$ , is a government specific cost. For example, this cost term might entail foregone campaign contributions, or lost votes, from refusing to violate a particular application of the regime's rules.  $c_t$  is a non-negative, positively unbounded random variable, distributed with some mean,  $\mu_c$ , and variance,  $\sigma_c$ . Define  $\chi_t = \beta_t - c_t$ .

Once costs are drawn, each government chooses whether to comply with the regulatory regime's rules or not,  $X \in \{\text{comply}, \text{not comply}\}$ . If a government complies, it receives  $\chi_t$ , and the other government receives some benefit,  $b > 0$ . If a government does not comply, a potential litigant can choose to bring a case against that government,  $L \in \{\text{litigant}, \text{not litigant}\}$ . A litigant

---

behaving consistent with the public's preferences. Thus, the equilibrium would degenerate into the government acting as a perfect agent of the public.

<sup>16</sup> This model builds off of previous work (see Author 2003, 2005).

pays a cost,  $k_t$ , in time and resources for bringing cases.<sup>17</sup> Again,  $k_t$  is an unbounded random variable, distributed with some mean,  $\mu_k$ , and variance,  $\sigma_k$ .

If a case is brought, in the process of hearing the case  $\chi_t$  becomes common knowledge among the litigants, governments, and court. The court then issues a ruling,  $R \in \{\textit{for defendant, against defendant}\}$ , and, if the court rules *against defendant*, the court imposes a judgment,  $j_t$ , where  $j_t > \chi_t$ .

Because court decisions are not automatically enforced, when ruled against, the defending government must decide whether to acquiesce in the decision or not ( $A \in \{\textit{acquiesce, not acquiesce}\}$ ). If the government acquiesces, it pays the judgment,  $-j_t$ , and the litigant receives some benefit. For parsimony, assume the benefit equals the judgment,  $j_t$ .

After all cases are tried, with probability  $q$  a public observes the round of play.<sup>18</sup> If the public does not observe the round of play, the period ends, a discount factor,  $\delta \in (0,1)$ , is accrued, and the game repeats. If the public does observe the round, the public learns some, but not all, of what transpired. In particular, the public observes whether the government is brought to court,  $L$ , whether the government is ruled against,  $R$ , and whether the government complies with the adverse ruling,  $A$ .<sup>19</sup> The public also observes the net societal benefit of compliance with the regime's rules in that period,  $\beta_t$ , if and only if the government complies with the regime's

---

<sup>17</sup> If bringing cases were costless, litigants would always bring cases, and the equilibrium would unravel in a degenerate fashion (see Author (2005)).

<sup>18</sup> This feature characterizes the idea that publics are only probabilistically paying attention to the game being played between its government and the court.

<sup>19</sup> These are actions that might get reported in newspapers and other media outlets.

rules either initially or after an adverse court ruling. Allowing the public to only probabilistically observe a round of play, and restricting the set of information available to the public even if it does observe the round of play, allows me to demonstrate that the public can come to back the court and the court can thereby become a meaningful constraint on government behavior even when the public is relatively inattentive and poorly informed.

The public uses these observations to update its beliefs about states of the world. In particular, the public does not know the expected benefit of compliance with the regulatory regime,  $\mu_\beta$ , with certainty.<sup>20</sup> Rather, it has some belief,  $\hat{\mu}_\beta$ , over the true value.<sup>21</sup> Thus, as the public observes rounds of the game, it learns about how applications of the regulatory regime's rules affect it. Based upon what the public learns that period, it chooses whether to sanction its government,  $S \in \{\textit{sanction}, \textit{not sanction}\}$ , by imposing a cost,  $c_p$ , or not. The period then ends,  $\delta$  is accrued, and the game repeats.

At the end of each period, payoffs are realized. Government and litigant payoffs are already characterized above. The exact preferences of the court could be specified in a variety of ways. All that matters is that the court and the governments have a conflict of interests. Without some conflict, the court would have neither the incentive nor the ability to demonstrate its influence by getting compliance with the regulatory regime's rules independent of the interests of the governments. Thus, to generate this conflict, I assume that the court wants to maximize

---

<sup>20</sup> Assume the distribution of  $\mu_\beta$  is a nicely behaved, unbounded, unimodal function, such as the normal distribution.

<sup>21</sup> Without this uncertainty, the equilibrium would degenerate into one in which the public would know immediately whether backing the government or the court was in its interests. Thus, the central question of this study could not be addressed.

total instances of compliance with the regulatory regime's rules, where either *ex ante* compliance with the regime's rules or *ex post* compliance with the court's ruling counts as an instance of compliance. Because governments will not want the regulatory regime's rules followed when  $\chi_t$  is sufficiently large, this specification ensures that the court wants compliance in situations that the governments do not.<sup>22</sup> The public's payoff is a function of whether the two governments comply with the regime's rules or not. It receives  $b$ , if the other government complies and  $\beta_t$  if its government complies.<sup>23</sup>

## Analysis

---

<sup>22</sup> One equally well could assume that the court has its own agenda and that it wants the regulatory regime's rules followed in some instances, but not others. However, as long as the court's preferences and governments' preferences are not perfectly correlated, such a model should generate equivalent results. Thus, the preference function assumed here is a more parsimonious way of generating the necessary conflict.

<sup>23</sup> The aggregate public's preferences are rooted in the idea that each member of the public cares whether they benefit from the regulatory regime. On any given application of the legal regime's rules, each individual is affected by its application to a greater or lesser extent. As such, one could sum up across individuals to get a measure of the net societal benefit (or cost) of that application. Operating from behind a veil of ignorance, the majority of the public would want the law followed whenever its impact is net beneficial. Thus, while the model assumes a unitary public, this assumption is built upon a particular notion of how individual preferences sum up to the "public's" preferences.

The formal model characterized above allows me to address the following strategic scenario. Governments face a complex collective action problem. They would like to sustain cooperation with a common regulatory regime's rules; however, doing so is problematic. To help facilitate compliance, the governments create a court that can act as a fire alarm and an information clearinghouse. Unfortunately for the government, it cannot perfectly control the preferences of the court. Further, its domestic public, while not necessarily overly attentive or knowledgeable, both cares about what its government is doing and has at least somewhat different tastes from its government. Thus, if the public decides its interests would be better served by the court's application of the regulatory regime's rules, this situation could become problematic for the governments in its own right.

I derive equilibrium behavior in four parts. I first motivate why the governments would create a court without enforcement powers. Here I demonstrate that governments can use the court to help them sustain a desirable pattern of "selective compliance" with the regulatory regime.<sup>24</sup> Note that rather than fully re-derive this result here, for parsimony I rely upon previously derived results from Author (2005). Given this role for the court, I next characterize equilibrium behavior when the public would not sanction its government for ignoring an adverse court decision, and for when the public would sanction its government. These parts of the equilibrium characterize a court in its institutional "infancy" and after it has "matured", respectively. Here, I demonstrate that while the court can be no more than a faithful agent of the governments without public support, the court is actually able to push its own agenda, to a great

---

<sup>24</sup> Specifically, I characterize the punishment strategy the governments would agree upon that would take maximal advantage of the creation of a court. This provides the government's off-equilibrium path behavior for the rest of the on-equilibrium path analysis.

extent independent of the preferences of both the governments and the public, once it does have public support. Last, I derive what must happen for a public to decide to support the court when in a conflict with its government. Here the critical issue is the degree to which the public believes its government is acting as a faithful agent.

### *Why Create a Court?*

Governments face an important strategic tension in the stage game (see table 2). If both governments cooperate, they receive  $b + \chi_t$ ; if only one cooperates it receives  $\chi_t$  and the other government receives  $b$ ; and if neither cooperates they both receive zero. When  $\chi_t$  is positive, it is a dominant strategy for a government to comply. It receives a benefit from complying unilaterally and an additional benefit if the other government complies as well. When  $\chi_t < 0$ , but  $b + \chi_t > 0$ , cooperation in the stage game is *mutually beneficial*. If both governments cooperate, both governments are better off. However, each government is best off if they defect and the other government cooperates. Thus, in this situation, the stage game is a prisoner's dilemma. They would like to sustain cooperation, but without iteration and credible threats of punishment, they will both defect. Finally, when  $b + \chi_t$  is negative, cooperation in the stage game is *mutually costly*. If both governments cooperate in this situation, they are both worse off. Thus, in this case, the governments would not want to attempt to sustain cooperation.

Ideally, the governments would like to sustain cooperation as long as it is at least mutually beneficial. However, the informational asymmetries make doing so problematic. Because governments only get to observe if a defection occurs, they can not know whether they would be punishing a government for defecting when cooperation would have been mutually beneficial or not. By creating a court that can act as a fire alarm and information clearinghouse,

the governments can use the court to facilitate their desired bargain. In particular, such a court allows the governments to target their punishment strategy by conditioning it upon the realized costs of compliance. As demonstrated in Author (2005), assuming the governments engaged in pre-play negotiation, they would agree upon the following punishment strategy; if a government defects, is taken to court, litigation reveals that the net present value of compliance would have been at least mutually beneficial (i.e.  $b + \chi_t + \delta CV_{\beta,t} > \delta CV_{\sim\beta,t}$ , where  $CV_{i,t}$  for  $i = \beta, \sim\beta$  is the continuation value from playing out the rest of the game conditional upon that period's moves) and the government refuses to comply with an adverse ruling, the governments punish. Otherwise, the governments do not punish. Punishment entails the punishing government defecting while the other government cooperates for some number of periods. Once the violating government has been punished for a sufficient number of periods, it is forgiven and the two governments return to cooperation.<sup>25</sup> While governments are not always best off under this regime, this institutional arrangement and punishment strategy can be an optimal, feasible arrangement for the governments. Thus, while we should not always expect governments to choose to create a court without enforcement powers to help facilitate compliance with the regime's rules, this provides an explanation for why governments would do so.<sup>26</sup>

---

<sup>25</sup> More precisely, punishment is a renegotiation-proof  $T$  period punishment strategy, in which the government being punished must comply for a total of  $T$  periods, while the other government gets to defect whenever unilateral cooperation is costly ( $\chi_t < 0$ ). Punishment is for a total of  $T$  periods, rather than for  $T$  periods in a row, because governments probabilistically will draw costs over which they would refuse to cooperate.

<sup>26</sup> See Author (2005) for a more detailed discussion.

Having identified why governments would create a court without enforcement powers, and thereby the class of punishment strategies the governments would agree upon in pre-play negotiations, we can now identify on-equilibrium path behavior (i.e. observable behavior) by the governments, the court, and the publics. This behavior is broken into three sets of results, behavior when the public does not have confidence about the average benefits of the regulatory regime ( $\hat{\mu}_{\beta,t} < \mu_{\beta}^*$ ), behavior when it does ( $\hat{\mu}_{\beta,t} \geq \mu_{\beta}^*$ ), and transitioning behavior between these two phases. Note that the cut-point  $\mu_{\beta}^*$  is a critical value, because it defines the conditions under which the public is indifferent over punishing or not punishing its government for noncompliance with an adverse court ruling. Why this cut-point is critical is discussed in the third set of results.

### *A Court in Its Institutional “Infancy”*

Assume the public’s assessment of the average benefit of compliance with the regulatory regime’s rules is less than some cut-point,  $\hat{\mu}_{\beta,t} < \mu_{\beta}^*$ . Equilibrium behavior when this condition holds, labeled the “selective compliance” (SC) phase is as follows.

#### “Selective Compliance” Equilibrium Behavior:

$$X = \begin{cases} \text{comply} & \text{if } \chi_t \geq \chi_t^* \\ \text{not comply} & \text{otherwise} \end{cases}$$

$$L = \begin{cases} \text{litigate} & \text{if } k_t \leq k_t^* \\ \text{not litigate} & \text{otherwise} \end{cases}$$

$$R = \begin{cases} \text{against defendant} & \text{if } \beta_t \geq \beta_t^*, \text{ and } \chi_t \geq \delta q(CV_{-\beta} - CV_{\beta}) - b \\ \text{for defendant} & \text{otherwise} \end{cases}$$

$$j_t' \leq \begin{cases} j_t^* & \text{if } R = \textit{against defendant} \\ 0 & \text{otherwise} \end{cases}$$

$$A = \begin{cases} \textit{acquiesce} & \text{if } \chi_t \geq \delta q (CV_{-\beta} - CV_{\beta}) b, \text{ and } j_t' \leq j_t^* \\ \textit{not acquiesce} & \text{otherwise} \end{cases}$$

$$S = \begin{cases} \textit{sanction} & \text{if } \hat{\mu}_{\beta,t} \geq \mu_{\beta}^* \text{ and } A = \textit{not acquiesce} \\ \textit{not sanction} & \text{otherwise} \end{cases}$$

where,  $\chi_t^* = p_k j_t^* + \delta q (1 - p_k) (CV_{\beta} - CV_{-k, -\beta})$ ,

$$k_t^* = j_t^* p \{ \beta_t \geq \beta_t^* \cup b + \chi_t + \delta q CV_{\beta} \geq \delta q CV_{-\beta} \mid c_t - \beta_t > c_t^* \},$$

$\beta_t^*$  solves  $1 + \delta^{court} CV_{I=1}^{court} = \delta^{court} CV_{I=0}^{court}$ ,  $j_t^* = \delta (q CV_{\beta} - q CV_{pp, -\beta})$ , and

$CV_i$  are continuation values defined in the appendix.

In equilibrium, governments probabilistically comply with the regime's rules, litigants probabilistically bring cases against governments, the court probabilistically rules against the governments, the court only imposes judgments that governments will accept, and governments comply with all adverse rulings. To reason through the logic of this equilibrium behavior, a series of propositions are presented, starting with the end of the tree and moving back up it.

*Proposition 1<sub>SC</sub>*: Governments obey adverse rulings if and only if cooperation is not too costly ( $\chi_t \geq \delta q (CV_{-\beta} - CV_{\beta}) b$ ), and the judgment the court imposes is not too large ( $j_t' \leq j_t^*$ ).

If cooperation is too costly, large enough to make compliance mutually costly, the governments will not punish each other for noncompliance with an adverse ruling, and therefore the litigating government has no incentive to obey the court. Further, if the imposed judgment is too large, larger than the cost of being punished by the other government for noncompliance, the

litigating government is better off disobeying the court ruling. Thus, governments only obey adverse court rulings if cooperation with the regulatory regime's rules would have been mutually beneficial and if the imposed judgment is not too large.

*Proposition 2<sub>SC</sub>*: The court rules against a government if and only if cooperation is not too costly ( $\chi_t \geq \delta q(CV_{-\beta} - CV_{\beta})b$ ) and the societal benefits of cooperation are sufficiently large ( $\beta_t \geq \beta_t^*$ ), and the court only imposes judgments that governments will pay ( $j_t' \leq j_t^*$ ).

The court knows that governments will only cooperate with an adverse ruling if cooperation with the regulatory regime's rules would have been mutually beneficial, and if the judgment imposed is not too large. Further, as will be demonstrated subsequently, the court knows that the public is more likely to punish its government for noncompliance with an adverse court ruling in the future if it does not observe noncompliance with court rulings today and if it does not observe applications of the regulatory regime's rules that are too costly (i.e.  $\beta_t$  sufficiently large). Thus, the court only rules against governments when governments will comply with those rulings, and when the societal benefits of cooperation are sufficiently large, and the court only imposes judgments that will be paid.

*Proposition 3<sub>SC</sub>*: Litigants only bring cases when the cost of bringing a case is not too large ( $k_t \leq k_t^*$ ).

Litigants know that they will only win cases probabilistically. Thus, litigants will only bring cases when the cost of bringing a case does not exceed the expected value of that litigation.

*Proposition 4<sub>SC</sub>*: Governments only comply with a regulatory regime's rules ex ante when compliance is not too costly ( $\chi_t \geq \chi_t^*$ ).

Governments know that cases are only brought probabilistically. Further, governments know that they will win any cases in which ex ante compliance with the regulatory regime's rules would have been mutually costly. Thus, governments only will comply with the regulatory regime's rules ex ante when doing so is sufficiently beneficial. In particular, governments will always defect when compliance would have been mutually costly, and even will defect sometimes when it would have been mutually beneficial, though not too beneficial, because they know the case might not be brought against them.

*Proposition 5<sub>SC</sub>*: The public does not punish its government, because ( $\hat{\mu}_{\beta,t} < \mu_{\beta}^*$ ).

Finally, the public does not believe that the average benefit of compliance with the regulatory regime is sufficiently large to merit punishing its government for noncompliance with an adverse court ruling. Further, because the government never voluntarily complies with the regulatory regime when it is not mutually beneficial for the governments, and therefore also for the governments' publics, the public also will not punish its government for the compliance in which the government does engage.

In sum, when the court does not have the backing of the public – i.e. the public will not punish its government for noncompliance with an adverse court ruling – the court's ability to enforce the regulatory regime's rules is limited. It can facilitate compliance with those rules when doing so is mutually beneficial for the governments involved, but it cannot get compliance with those rules independent of the interests of the governments. Further, the court has to be wary of public opinion. The court actually might choose not to apply the regulatory regime's

rules even when governments would have complied, because it does not want the public experiencing especially bad outcomes associated with that application of those rules. Thus, while the court can help governments achieve a more stable and efficient regulatory regime than would exist without the court, the court's influence is limited.<sup>27</sup>

### *An Institutionally "Mature" Court*

Now, assume the public's assessment of the average benefit of compliance with the regulatory regime's rules is greater than some cut-point,  $\hat{\mu}_{\beta,t} \geq \mu_{\beta}^*$ . Equilibrium behavior when this condition holds, labeled the "universal compliance" (*UC*) phase, is indicated below. Note that this equilibrium requires assuming a cost of public sanctioning,  $c_p$ , large enough that governments will comply at least occasionally when compliance would be mutually costly. Further, this equilibrium also assumes that governments would not prefer the non-coordinated equilibrium (only comply when doing so is unilaterally beneficial) to the equilibrium behavior characterized below. If these assumptions are not made, public sanctioning cannot alter government behavior and transitioning to a "mature" court is impossible.

### "Universal Compliance" Equilibrium Behavior:

$$X = \begin{cases} \text{comply} & \text{if } \chi_t \geq \chi_t^* \\ \text{not comply} & \text{otherwise} \end{cases}$$

$$L = \begin{cases} \text{litigate} & \text{if } k_t \leq k_t^* \\ \text{not litigate} & \text{otherwise} \end{cases}$$

---

<sup>27</sup> These findings are consistent with previous work (author 2001, 2005).

$$R = \begin{cases} \textit{against defendant} & \text{if } \beta_t \geq \beta_t^*, \text{ and } \chi_t \geq -qc_p \\ \textit{for defendant} & \text{otherwise} \end{cases}$$

$$j_t^* \leq \begin{cases} c_p & \text{if } R = \textit{against defendant} \\ 0 & \text{otherwise} \end{cases}$$

$$A = \begin{cases} \textit{acquiesce} & \text{if } j_t \leq qc_p \\ \textit{not acquiesce} & \text{otherwise} \end{cases}$$

$$S = \begin{cases} \textit{sanction} & \text{if } \hat{\mu}_{\beta,t} \geq \mu_{\beta}^* \text{ and } A = \textit{not acquiesce} \\ \textit{not sanction} & \text{otherwise} \end{cases}$$

where,  $\chi_t^* = p_k j_t^* + \delta q(1 - p_k)(CV_{\beta} - CV_{\sim k, \sim \beta})$ ,

$$k_t^* = j_t^* p \left\{ b + \chi_t + \delta q CV_{\beta} \geq -c_p + \delta q CV_{\sim \beta} \cup \beta_t \geq \beta_t^* \mid c_t - \beta_t > c_t^* \right\},$$

$\beta_t^*$  solves  $1 + \delta^{court} CV_{l=1}^{court} = \delta^{court} CV_{l=0}^{court}$ ,  $j_t^* \leq \delta(q CV_{\beta} - q CV_{pp, \sim \beta})$ , and

$CV_i$  are continuation values defined in the appendix.

Superficially, behavior in this stage of the equilibrium appears quite similar to the behavior derived in the stage of the equilibrium discussed above. Again, governments probabilistically comply with the regime's rules, litigants probabilistically bring cases against governments, the court probabilistically rules against the governments, the court only imposes judgments that governments will accept, and governments comply with all adverse rulings. The important difference is that the public will punish its government for noncompliance with an adverse court ruling. As a result, the conditions under which governments comply with the regulatory regime's rules and court decisions qualitatively changes. The following propositions characterize behavior in this stage.

*Proposition 1<sub>UC</sub>*: Governments obey adverse rulings as long as paying the judgment is less costly than the expected cost of ignoring a court ruling ( $qc_p$ ).

Without the threat of public sanctions, governments only comply with adverse court rulings if doing so is mutually beneficial. However, once the threat of public sanctions becomes credible, governments no longer base their decision on whether other governments will punish them, but rather over the risk that their public will punish them. If the expected cost of public sanctioning is sufficiently severe, either because the probability of being caught or the cost the public would impose is sufficiently large, the government would prefer to comply with the adverse judgment.

*Proposition 2<sub>UC</sub>*: The court rules against a government if and only if cooperation is not too costly ( $\chi_t \geq -qc_p$ ) and the societal benefits of cooperation are sufficiently large ( $\beta_t \geq \beta_t^*$ ).

Recognizing that the government's decision to comply with a ruling or not now depends upon possible public backlash, the court is now freer to rule against the government. While the court still does not want to have the regulatory regime's rules applied when the net societal benefit too small, the court can at least occasionally rule against governments and get compliance even when compliance would be mutually costly. The more severe the potential public backlash, the more latitude the court has.

*Proposition 3<sub>UC</sub>*: Litigants only bring cases when the cost of bringing a case is not too large ( $k_t \leq k_t^*$ ).

Litigant decision-making is basically unaffected. While the probability with which the litigant draws a cost sufficiently small to litigate will change, the litigant is still focused upon balancing the expected payoff from bringing a case with the cost of doing so.

*Proposition 4<sub>UC</sub>*: Governments only comply with a regulatory regime's rules ex ante when compliance is not too costly ( $\chi_t \geq \chi_t^*$ ).

Similarly, the governments' decision to comply ex ante with the regulatory regime's rules is not qualitatively different from before. The probability with which the government will comply ex ante will change (presumably increase since it knows it will lose cases over a wider range of costs than before), the government is still deciding on a risky gamble: defect and risk getting caught and having to pay a putative fine, or simply comply in the first place.

*Proposition 5<sub>UC</sub>*: The public punishes its government if it observes noncompliance with an adverse court ruling, because ( $\hat{\mu}_{\beta,t} \geq \mu_{\beta}^*$ ).

Finally, this time the public does believe that the average benefit of compliance with the regulatory regime is sufficiently large to merit punishing its government for noncompliance with an adverse court ruling.

In sum, when the court *does* have the backing of the public – i.e. the public will punish its government for noncompliance with an adverse court ruling – the court's ability to enforce the regulatory regime's rules is qualitatively expanded. Now it can facilitate compliance with the regulatory regime's rules even when doing so is mutually costly for the governments involved. Further, the more likely the public is to notice a conflict between its government and the court, and the more sensitive to public sanctioning the government is, the greater the latitude available

to the court. Thus, while there are still limits to what the court can do, the court cannot go beyond what the public can enforce and the court remains sensitive to public reactions to applications of the regulatory regime's rules, the court is now in a position to push its own agenda (even contrary to public tastes on particular cases).<sup>28</sup>

### *Transitioning from a Court in its "Infancy" to a "Mature" Court*

The above discussion characterizes observable behavior in the two states of the equilibrium. However, the discussion omitted both the logic behind both the public's behavior and beliefs. And, by implication, the discussion omitted any explanation for why transitioning can happen in equilibrium.

To start, suppose there were no public. In this version of the model, the court could never arise out of its "infancy". Rather, the selective compliance phase of the equilibrium would be a steady state; all of the continuation values in the characterized equilibrium would cancel out, and as a result the critical cut-points,  $\chi^*$ ,  $k^*$ ,  $j^*$ , would be time invariant (i.e. constants).<sup>29</sup>

What happens when the public is introduced? The governments and the court start anticipating the possibility of transitioning from the selective compliance to the universal compliance phase. Their assessment of the likelihood of transitioning depends upon the public's beliefs over the average value of compliance with the regulatory regime's rules,  $\hat{\mu}_{\beta,t}$ . The following propositions present the public's moves and beliefs. Proof is provided in the appendix.

---

<sup>28</sup> These findings are broadly consistent with work by Vanberg (2005), Staton (2006), and Stephenson (2004).

<sup>29</sup> While modeled slightly differently, this equilibrium would match the one derived in author (2005).  $\beta^*$  would no longer exist.

*Proposition 1<sub>T</sub>*: The public will punish its government for noncompliance with adverse court rulings when  $EU_p(\textit{sanction} \mid \hat{\mu}_{\beta,t}) \geq EU_p(\textit{not sanction} \mid \hat{\mu}_{\beta,t})$ , where  $EU_p(\textit{sanction} \mid \hat{\mu}_{\beta,t}) - EU_p(\textit{not sanction} \mid \hat{\mu}_{\beta,t})$  is increasing in  $\hat{\mu}_{\beta,t}$ .

The public does not observe  $\beta_t$  when its government is noncompliant. Therefore, the public must make an expected utility calculation; is the expected value of the cases in which the government did not comply, but would have under threat of public sanctioning, positive? If  $EU_p(\textit{sanction} \mid \hat{\mu}_{\beta,t})$  is greater than  $EU_p(\textit{not sanction} \mid \hat{\mu}_{\beta,t})$ , the public punishes its government for noncompliance with an adverse court ruling and the equilibrium enters the universal compliance phase, otherwise the public leaves the government alone. Whether  $EU_p(\textit{sanction} \mid \hat{\mu}_{\beta,t}) \geq EU_p(\textit{not sanction} \mid \hat{\mu}_{\beta,t})$  holds or not depends upon  $\hat{\mu}_{\beta,t}$ . Recall that the government's payoff from complying with the regulatory regime's rules,  $\chi_t$ , is a function of two components, the net social benefit of compliance,  $\beta_t$ , and the government's private cost of compliance,  $c_t$ . If the public believes that  $\beta_t$  is sufficiently large on average – i.e.  $\hat{\mu}_{\beta,t}$  is sufficiently large – the public will logically conclude that the government's decision to not comply is a product of a bad draw on  $c_t$ . Thus, the larger  $\hat{\mu}_{\beta,t}$ , the more likely the public is to believe that on average governmental noncompliance is only benefiting the government, and therefore  $EU_p(\textit{sanction} \mid \hat{\mu}_{\beta,t}) \geq EU_p(\textit{not sanction} \mid \hat{\mu}_{\beta,t})$  is more likely to hold.<sup>30</sup>

---

<sup>30</sup> Note that sanctioning or not sanctioning its government in any period is a best reply for the public in this model. Thus, I identify a strategy that maximizes public's expected utility. Note that, as with government punishment strategies, public sanctioning will only hinge upon

*Proposition 2<sub>T</sub>*: If the public observes  $\beta_t$ , the public's posterior belief over the expected value of compliance with the regulatory regime's rules is  $\hat{\mu}_{\beta,t+1} = \frac{\lambda\hat{\mu}_{\beta,t} + r\beta_t}{\lambda + r}$ , while if the public does not observe  $\beta_t$ , the public's posterior belief over the expected value of compliance with the regulatory regime's rules is defined as  $\mu'_\beta = \int_{-\infty}^{\bar{\beta}_t} \left( \frac{\lambda\mu_\beta + r\beta'}{\lambda + r} \right) \frac{\phi(\beta_t = \beta' | \mu_\beta)}{\Phi(\bar{\beta}_t | \mu_\beta)} \partial\beta'$ . Beliefs will update in the direction of  $\beta_t$ ,  $E[\beta_t | A = \text{not acquiesce}]$ ,  $E[\beta_t | R = \text{for defendant}]$ , or  $E[\beta_t | L = \text{not litigate}]$  depending upon the signal received.

If the government complies with the regulatory regime's rules, the public gets to observe  $\beta_t$ . As such, the public does not have to make any uncertain inferences over the societal benefit from compliance in that period. If the draw in that period is larger than the prior  $\hat{\mu}_{\beta,t}$ , the public increases its assessment, and if the draw is smaller, the public decreases its assessment.

If the government does not comply with the regulatory regime's rules in that period, the public does not get to observe  $\beta_t$ . In this case, the public must make its best inference over what  $\beta_t$  would have been if the government had complied. The exact inference depends upon the reason the public did not observe  $\beta_t$ . Independent of exactly why  $\beta_t$  is unobserved, if the expected value is larger than the public's prior over  $\hat{\mu}_{\beta,t}$ , the public's posterior belief will be larger, while if the expected value is smaller, the posterior belief will be smaller. Importantly, for all three expectations, the upper bound of possible  $\beta_t$ 's will be truncated (at  $\bar{\beta}_t$ ), because the government noncompliance with adverse court rulings. Otherwise, equilibrium behavior can unravel.

public knows that it will always observe large draws of  $\beta_t$ . Thus, unless  $\hat{\mu}_{\beta,t}$  is very small, the public will tend to decrease its assessment of  $\hat{\mu}_{\beta,t}$  upon observing noncompliance with the regulatory regime's rules.

In sum, the more the public observes governments complying with the regulatory regime's rules, and the better the outcomes, the more the public likes the regulatory regime. Conversely, the more the public observes noncompliance with the regulatory regime, the more likely the public is to downgrade its assessment of the regime's benefits.<sup>31</sup> If the public observes enough good outcomes, the public will decide that violations of the regulatory regime's rules are not in its interests, and it will punish its government for noncompliance with adverse court decisions.

## **Implications and Conclusions**

Why do sovereign governments create judicial institutions and grant them the authority to rule their acts invalid? Once created, to what degree can a court in its institutional "infancy" enforce government compliance with a regulatory regime's rules? And, how can such a court evolve over time into a "mature" court that can act (in general) as if its rulings are binding upon even sovereign actors?

---

<sup>31</sup> Of course, if the public's assessment of the net benefits are sufficiently skewed, the public will actually increase its assessment of the net benefits unless it observes a bad  $\beta_t$ . However, the closer the public is to supporting the court, the larger  $\hat{\mu}_{\beta,t}$  and therefore the less likely this condition is to hold.

The model presented here provides a unified answer to these questions. Governments have an incentive to rationally create a court that can act as a fire alarm and information clearinghouse to facilitate selective compliance with a common regulatory regime. Once created, an institutionally “immature” court – one without public backing – will fulfill the role set out by the governments admirably; the court can facilitate compliance with the regulatory regime’s rules, but it cannot impose its own preferences over those of the governments. If the same court “matures” and gains public backing, the story changes dramatically. While the court is not totally free to rule as it wishes, up to certain limits it is capable of getting government compliance with the regulatory regime’s rules even if the governments would prefer otherwise. A public can come to back a court if it observes governments complying with the regulatory regime’s rules frequently enough and the impact of that compliance is sufficiently beneficial.

A number of additional, substantively interesting implications can be derived from this equilibrium solution. The first set of implications identifies in what ways a court can influence the likelihood of successfully transitioning into a “mature” court. First, the Court does not always want compliance with the regulatory regime’s rules. As indicated in proposition 2<sub>T</sub>, if the net societal benefit of the regime’s rules is sufficiently small, e.g. when  $\beta_t < \hat{\mu}_{\beta,t}$ , compliance actually can decrease the probability of transitioning by decreasing the public’s assessment of the benefits of the regime. Thus, if the probability of transitioning is sufficiently high, and the draw of  $\beta_t$  is sufficiently bad ( $\beta_t < \beta_t^*$ ), the Court would rather forgo that instance of compliance in order to prevent damage to its long-term chances of transitioning. Second, the court has a weakly dominant strategy to not rule against governments when governments would ignore its decisions. The public knows that noncompliance with an adverse court ruling only happens when governments find compliance mutually costly, whereas the court might not rule against the

government either because the government would not have complied or because  $\beta_t$  is sufficiently bad ( $\beta_t < \beta_t^*$ ). Since there may exist ranges of  $\beta_t$  over which the Court would not rule against the government, but the government would comply if the court had ruled against it, the public's expectation of  $\beta_t$  is always the same or larger if the court avoids a confrontation with the government. Note that not ruling against the government is more likely to dominate when transitioning is more likely.

The second set of implications relate to public support for the court. Most importantly, in this model public support for the court is instrumental. That is, publics back a court only if doing so protects policy interests, not because publics intrinsically value the supremacy of rule of law or judicial institutions. Thus, diffuse support is directly related to expected benefits from court decisions. Additionally, courts can never take public support for granted. While a popular court can make individual decisions that go contrary to the public will, the court has to be careful not to make decisions that are too far out of line with public opinion, or too many unpopular decisions in a row. There is no guarantee that courts cannot lose public support as easily as they gain it.

The third set of implications relates to government behavior. Governments want a subservient, not an independent, court. Thus, given the governments perfectly well know what would lead to transitioning, and governments can engage in behavior to prevent it from happening (by not complying with the regulatory regime's rules), why is transitioning possible in equilibrium?

Suppose a government draws a  $\beta_t$  such that, if the public observes that  $\beta_t$ , the public will start backing the court in the future. Further, suppose the government also draws a  $c_t$  such that compliance is mutually beneficial. The result is that the government is faced with a risky lottery.

If the government goes to court and does not comply with the court ruling, it will be punished by the other government for sure. However, if the government does cooperate with the adverse ruling, only with probability  $q$  does the public observe the court case and the equilibrium transition. Since the judgment the court imposes,  $j_t'$ , is less costly to the government than being punished, as long as  $q$  is sufficiently small, the government will take the gamble.

Of course, the government only wants to take the gamble because it expects to be punished for noncompliance. So, why not wait to observe  $\beta_t$  in the court case, and decline to punish each other when transition is possible? While this behavior might appear desirable, the punishing government cannot credibly commit to not punishing; punishing results in more compliance by the other government and less undesirable compliance ( $\chi_t < 0$ ) by you. Further, once a government enters the universal compliance phase, that government just complies more often. Thus, once at this point, the governments suffer from a coordination problem. While they would both prefer not to transition to universal compliance, each is actually perfectly happy if the other does.

Finally, why not pre-negotiate their way out of this coordination problem by agreeing to punishment strategies in which neither government punishes if cooperation in that period could lead to transitioning? Even if governments were sufficiently foresighted to anticipate this possibility in pre-game negotiations, doing so would not prevent the possibility of transitioning in equilibrium. Because transitioning only happens with probability  $q$ , for sufficiently large  $\beta_t$ , the government prefers to unilaterally comply and reap the short-term benefit of getting  $\beta_t$ , even at the risk of transitioning. Thus, no matter how sophisticated and foresighted the governments, governments rationally could choose to create regulatory regimes over which they might eventually lose control.

The fourth set of implications involve why some courts fail to make this transition and others succeed. Courts that are unlikely to ever make the transition are ones in which the judges are less foresighted (i.e. smaller  $\delta^{court}$ ), the average societal benefit of the regulatory regime is relatively small on average (i.e. smaller  $\beta_i$ ), and government preferences are more closely aligned to their publics (i.e. smaller  $E[c_i]$ ). Conversely, the more foresighted the judges, the greater the societal benefits on average, and the larger the gap between public and government preferences, the more likely the transition and the more stable the court will be upon transitioning. Thus, transition is more likely when courts are more willing to trade-off getting compliance today for getting more influence tomorrow, when the public finds that the benefits of the regulatory regime are substantial, and when the public has good reason to be suspicious of the motives of an unchecked government.

Finally, note that this study also has implications for the development of judicial review of federal law. Assuming the standing government is opposed to an annulment of federal law, that decision is equivalent to ruling contrary to a majority position of the states.<sup>32</sup> Thus, we would expect the court to be able to exercise this authority exactly when it has the capacity to do so over the states individually. By implication, the development of federal judicial review may be incidental to the development of review over state actions.

In sum, this model is relevant to a wide variety of judicial institutions. It suggests why we see courts created and granted authority (or not) over government decision-making at both the national and international levels. Further, it suggests why some of these courts mature into strong and independent courts and some do not. Future work can take advantage of this wide empirical

---

<sup>32</sup> See Whittington (2005) for examples when the judiciary can exercise judicial review consistent with federal government preferences.

applicability for both testing and refinement of the theory. If supported, this work then would provide the foundation for a general theory of judicial authority over sovereign bodies.

## Appendix

This appendix provides an existence proof for the characterized equilibrium. I work backwards through the model, first defining the on-equilibrium path and punishment path continuation values for the governments, then equilibrium behavior backwards through the stage game. Note that, because of the complexity of the model, there is no closed form solution to the game.

### Selective Compliance Phase:

Each period, governments comply with the regulatory regime's rules when  $\chi_t > \chi_t^*$ , litigants bring cases when  $k_t < k_t^*$ , the court rules against governments when  $\beta_t > \beta_t^*$  and compliance is mutually beneficial,  $b + \chi_t + \delta q CV_{\beta} \geq \delta q CV_{\sim\beta}$ , and the court imposes a judgment small enough that the governments will pay,  $j_t < j_t^*$ . The public observes the round of play with probability  $q$ . If the public observes  $\beta_t$ , the subsequent government continuation value is  $CV_{\beta_t}$ , while if the public does not observe  $\beta_t$ , the subsequent government continuation value is  $CV_{\sim\beta_t}$ . This equilibrium behavior yields the government continuation value  $CV$  indicated below. Notice that the equilibrium approaches a steady-state as  $q \rightarrow 0$ , because the continuation values approach steady-state values. Note that for the proofs, the  $CV$ 's period subscript is suppressed to minimize notational clutter.

$$\begin{aligned}
 CV = & p_{\chi} E[\chi_t + \delta(q CV_{\beta} + (1-q)CV) | \chi_t \geq \chi_t^*] + (1-p_{\chi}) p_k p_{ad} (-j_t^* + \delta(q CV_{\beta} + (1-q)CV)) \\
 & + (1-p_{\chi}) p_k ((1-p_{ad}) (\delta(q CV_{\sim\beta} + (1-q)CV)) + p_{\beta} (\delta(q CV_{\sim\beta} + (1-q)CV))) \\
 & + (1-p_{\chi}) (1-p_k) (\delta(q CV_{\sim k, \sim\beta} + (1-q)CV)) + p_{\chi, og} b,
 \end{aligned}$$

where  $p_{ad} = p\{\beta_t \geq \beta_t^* \cup b + \chi_t + \delta q CV_\beta \geq \delta q CV_{\sim\beta} \mid \chi_t < \chi_t^*\}$ ,  $p_k = p\{k_t \leq k_t^*\}$ ,  $p_\chi = p\{\chi_t \geq \chi_t^*\}$ ,  $p_{\chi,og} = p\{\chi_{og,t} \geq \chi_{og,t}^*\}$ , and  $p_\beta = p\{\beta_t < \beta_t^*\}$ .

If a government enters the punishment path, it will be punished for  $T$  total periods in expectation. During that time, the government cooperates each period  $\chi_t$  is not so costly that the government would prefer delaying the return to cooperation by a period, and pays the expected value of  $\chi_t$  given it chose to cooperate. Simultaneously, the other government cooperates whenever doing so is unilaterally beneficial. Once punishment is over, the governments return to on-equilibrium path behavior. This calculation yields the punishment path continuation value,

$$CV_\tau^{pp} = \frac{(\delta - \delta^\tau) p\{\chi_t + \delta CV_{\tau-1}^{pp} > \delta CV_\tau^{pp}\} E[\chi_t \mid \chi_t + \delta CV_{\tau-1}^{pp} > \delta CV_\tau^{pp}]}{1 - \delta} + \frac{(\delta - \delta^\tau) p\{\chi_{og,t} + \delta(q CV_\beta + (1-q)CV) > \delta(q CV_{\sim\beta} + (1-q)CV)\} b}{1 - \delta} + \delta^\tau CV_{\hat{\mu}_{\beta,t}}$$

where  $\tau$  indicates the number of periods remaining in the punishment phase in expectation, and the subscript  $\hat{\mu}_{\beta,t}$  indicates the public's beliefs after the punishment phase. The lower boundary on the punishment path payoff equals the payoff from the non-cooperative equilibrium in which both governments only comply when it is unilaterally beneficial,

$$\underline{CV}^{pp} = \frac{\delta}{1 - \delta} (p\{\chi_t + \delta q CV_\beta > \delta q CV_{\sim\beta}\} E[\chi_t \mid \chi_t + \delta q CV_\beta > \delta q CV_{\sim\beta}] + p_{\chi,og} b).$$

Governments pay adverse judgments as long as the cost of that judgment plus the expected continuation value from remaining on-equilibrium path is greater than or equal to the expected continuation value of going down the punishment path,

$$-j_t^* + \delta(q CV_\beta + (1-q)CV) \geq \delta(q CV_{\sim\beta}^{pp} + (1-q)CV^{pp}). \text{ Solving for } j_t^* \text{ yields}$$

$$j_t^* \leq \delta((q CV_\beta + (1-q)CV) - (q CV_{\sim\beta}^{pp} + (1-q)CV^{pp})). \text{ In the limit, as } q \rightarrow 0, j_t^* \leq \delta(CV_{SC} - CV_{SC}^{pp}). \text{ The}$$

steady-state on-equilibrium continuation value is

$$CV_{sc} = p_\chi E[\chi_t + \delta(qCV_{sc} + (1-q)CV_{sc}) | \chi_t \geq \chi_t^*] + (1-p_\chi)p_k p_{ad}(-j_{sc}^*) + (1-p_\chi)b + \delta CV_{sc}. \text{ Solving yields}$$

$$CV_{sc} = \frac{p_\chi E[\chi_t + \delta(qCV_{sc} + (1-q)CV_{sc}) | \chi_t \geq \chi_t^*] + (1-p_\chi)p_k p_{ad}(-j_{sc}^*) + (1-p_\chi)b}{1-\delta}. \text{ The steady-state}$$

punishment path continuation value is

$$CV_{sc}^{pp} = \frac{(\delta - \delta^T) \left( p \left\{ \chi_t + \delta CV_{sc}^{pp(\tau-1)} > \delta CV_{sc}^{pp(\tau)} \right\} E[\chi_t | \chi_t + \delta CV_{sc}^{pp(\tau-1)} > \delta CV_{sc}^{pp(\tau)}] + p \left\{ \chi_{og,t} > 0 \right\} b \right)}{1-\delta} + \delta^T CV_{sc}. \text{ On}$$

punishment path, the opposing government will only cooperate when it is unilaterally beneficial. Thus, for arbitrarily large  $b, j_t^*$  can get arbitrarily large.

The court can get compliance with an adverse ruling when compliance would be mutually beneficial for the governments,  $b + \chi_t + \delta q CV_\beta \geq \delta q CV_{-\beta}$ . At  $\beta_t^*$ , the court is indifferent over ruling against the government and having the government comply versus not ruling against, because the additional instance of compliance plus the continuation value given  $\beta_t$  might be observed equals the continuation value associated with not getting compliance this period. That is,  $\beta_t^*$  solves

$$1 + \delta^{court} CV_{I=1}^{court} = \delta^{court} CV_{I=0}^{court}, \text{ where the index } I=1 \text{ if } \beta_t \text{ could be observed, and } 0 \text{ otherwise. For } \beta_t > \beta_t^*$$

the court prefers ruling against the government and getting compliance and for  $\beta_t < \beta_t^*$  the court does not.

In the limit, as  $q \rightarrow 0$ ,  $\beta_t^* \rightarrow -\infty$ . Thus,  $\beta_t > \beta_t^*$  exists for  $q$  sufficiently small. While  $\beta_t < \beta_t^*$  does not

necessarily exist, it can. Suppose  $\hat{\mu}_{\beta,t}$  and  $b$  are arbitrarily large. Then, the equilibrium is firmly in the

universal compliance state, but the government will comply with are arbitrarily small  $\beta_t$  such that  $\hat{\mu}_{\beta,t}$

could be driven arbitrarily small and the equilibrium could be driven arbitrarily firmly into the selective

compliance state. As  $\delta^{court} \rightarrow 1$ , the difference between these steady states approaches

infinity,  $\frac{\delta^{court} (CV_{I=1,SC}^{court} - CV_{I=0,UC}^{court})}{1 - \delta^{court}} \rightarrow \infty$ . Thus,  $\beta_t < \beta_t^*$  can exist as well.

Litigants are indifferent over bringing cases when the probability of winning the case, times the value of winning the case, minus the cost of bringing the case, are equal to the payoff from not bringing a case,  $p_{ad} j_t^* - k_t^* = 0$ . Solving yields  $k_t^* = j_t^* p_{ad}$ . For  $k_t > k_t^*$  the litigant does not bring cases and for  $k_t < k_t^*$  the litigant does. Since  $j_t^*$  is some positive number,  $p_{ad}$  is a probability, and the distribution of  $k_t$  is unbounded,  $k_t^*$  exists for  $q$  sufficiently small.

The governments are indifferent over initially complying or not when payoff associated with complying equals the expected value of risking going to court and losing,

$\chi_t^* + \delta(qCV_\beta + (1-q)CV) = p_k(-j_t^* + \delta(qCV_\beta + (1-q)CV)) + (1-p_k)\delta(qCV_{\sim k, \sim \beta} + (1-q)CV)$ . Solving yields

$\chi_t^* = -p_k j_t^* - \delta q(1-p_k)(CV_\beta - CV_{\sim k, \sim \beta})$ . For  $\chi_t > \chi_t^*$  the court does not defect and for  $\chi_t < \chi_t^*$  the court

does. In the limit, as  $q \rightarrow 0$ ,  $\chi_t^* = p_k(-j_t^*)$ . Since  $j_t^*$  is some positive number,  $p_k$  is a probability, and the distribution of  $\chi_t$  is unbounded,  $\chi_t^*$  exists for  $q$  sufficiently small.

### Universal Compliance Phase:

The only difference in this phase is that the public sanctions its government for noncompliance with adverse court ruling by imposing  $-c_p$ . Thus, whereas governments would only comply with an adverse ruling when  $b + \chi_t + \delta CV_\beta > \delta CV_{\sim \beta}$  holds, now they comply when  $b + \chi_t + \delta CV_\beta > -c_p + \delta CV_{\sim \beta}$  holds. By implication two equations from the previous solution change, the probability that the court rules against a government, and the maximal judgment that can be imposed. They are

$p_{ad,UC} = p\{\beta_t \geq \beta_t^* \cup b + \chi_t + \delta q CV_\beta \geq -c_p + \delta q CV_{\sim \beta} \mid \chi_t < \chi_t^*\}$  and

$j_{t,UC}^* \leq \delta(qCV_\beta + (1-q)CV) + c_p - \delta(qCV_{\sim\beta} + (1-q)CV)$ , where UC indicates that the equilibrium is in the universal compliance phase. No other equations defined in the selective compliance phase change.

### Public Behavior and Beliefs:

When there is ex ante or ex post compliance, and the public observes the round of play, the public gets to observe  $\beta_t$ . Assuming  $\beta_t$  and  $\mu_\beta$  are drawn from normal distributions, the posterior mean is a weighted average of the prior over  $\mu_\beta$  and the draw of  $\beta_t$ ,  $\hat{\mu}_{\beta,t+1} = \frac{\lambda\hat{\mu}_{\beta,t} + r\beta_t}{\lambda + r}$  where  $\lambda$  and  $r$  are the precision of  $\mu_\beta$  and  $\beta_t$  respectively (DeGroot 2004; 166-167). If the public observes the round of play, but the government does not comply, the public does not observe  $\beta_t$ . Instead, the public observes no case if  $\chi_t \leq \chi_t^*$  and  $k_t \geq k_t^*$  hold, or a ruling for the government if

$\chi_t \leq \chi_t^*, k_t < k_t^*$ , and  $b + \chi_t + \delta q CV_\beta < \delta q CV_{\sim\beta}$  or  $\beta_t \leq \beta_t^*$  hold. These constraints provide an upper boundary on  $\beta_t$  when  $\frac{\partial(\beta_t - \chi_t^*)}{\partial\beta_t} = 1 - \chi_t^{*'} > 0$  and  $\frac{\partial(\beta_t + \delta q CV_\beta)}{\partial\beta_t} = 1 + \delta q CV_\beta' > 0$  hold. These two

conditions hold for  $q > 0$ . As such, if  $\beta_t$  is unknown, the public updates based upon the probability

distribution of  $\beta_t$  as follows:  $\hat{\mu}_{\beta,t+1} = \int_{-\infty}^{\bar{\beta}_t} \left( \frac{\lambda\hat{\mu}_{\beta,t} + r\beta'}{\lambda + r} \right) \frac{\phi(\beta_t = \beta' | \hat{\mu}_{\beta,t})}{\Phi(\bar{\beta}_t | \hat{\mu}_{\beta,t})} d\beta'$ , where  $\Phi(\bullet)$  is a CDF.

Note if the court chose to rule against a government when that government would not comply, this event would happen when  $\chi_t \leq \chi_t^*, k_t < k_t^*$ , and  $b + \chi_t + \delta q CV_\beta < \delta q CV_{\sim\beta}$  held. Thus, because there would be one less condition under which the public would not observe  $\beta_t$  than when the court chooses to not rule against its government, the public would discount  $\beta_t$  less when  $\beta_t^* > \delta q (CV_{\sim\beta} - CV_\beta) - b - c_t$  holds.

The public will want to punish when the expected value of the  $\beta_t$ 's over which the government now complies, but would not have if the public did not punish, is positive. Assume  $c_p$  and the true mean,  $\mu_\beta$ , are sufficiently large such that  $p_k \rightarrow 1$  if the public is anticipated to punish (as these get large, the probability of losing a case or not getting compliance drops to zero). The first term covers the times when the government was not taken to court, but would have complied if the public had sanctioned and the second term covers the times when the government was taken to court, would not have complied if the public did not sanction, but would have complied if the public did sanction.

$$EU_p [\text{sanction} | \hat{\mu}_{\beta,t}] - EU_p [\text{not sanction} | \hat{\mu}_{\beta,t}] = p_\chi (1 - p_k) p_{ad,UC} \omega + p_\chi p_k p_\psi \psi,$$

where  $\omega = E[\beta_t | \chi_t \leq \chi_t^* \cup k_t > k_t^* \cup \beta_t \geq \beta_t^* \cup b + \chi_t + \delta q CV_\beta \geq q(-c_p + \delta CV_{c_p, \sim \beta}) + (1-q)\delta CV]$ ,  
 $p_\psi = p\{\beta_t \geq \beta_t^* \cup b + \chi_t + \delta q CV_\beta < \delta CV_{\sim \beta} \cup q(-c_p + \delta CV_{c_p, \sim \beta}) + (1-q)\delta CV < b + \chi_t + \delta q CV_\beta\}$ , and  
 $\psi = E[\beta_t | \chi_t \leq \chi_t^* \cup k_t \leq k_t^* \cup \beta_t \geq \beta_t^* \cup q(-c_p + \delta CV_{c_p, \sim \beta}) + (1-q)\delta CV < b + \chi_t + \delta q CV_\beta \cup b + \chi_t + \delta q CV_\beta < \delta q CV_{\sim \beta}]$ .

There are two upper boundary constraints on  $\beta_t$ ,  $\chi_t = \beta_t - c_t \leq \chi_t^*$  and

$b + \beta_t - c_t + \delta q CV_\beta < \delta q CV_{\sim \beta}$ . If  $E[c_t + \chi_t^*] > 0$  and  $E[c_t - b + \delta q(CV_{\sim \beta} - CV_\beta)] > 0$  hold, the conditional upper boundary on  $\beta_t$  can be positive. First, consider  $E[c_t + \chi_t^*] > 0$ . As  $E[c_t]$  gets large,  $\chi_t^*$  gets smaller. We know this because  $j_t^*$  decreases as  $E[c_t]$  gets large (all continuation values approach zero as  $E[c_t]$  gets large), and  $\chi_t^*$  is increasing in  $j_t^*$  (for  $q > 0$ ,  $c_t^*$  only relies upon  $j_t^*$  and  $p_K$ , and the probability term is upwardly bounded by one, while  $j_t^*$  will approach zero, thus  $E[c_t]$  will grow arbitrarily large relative to  $\chi_t^*$ ). Now consider  $E[c_t - b + \delta q(CV_{\sim \beta} - CV_\beta)] > 0$ . In general, we expect  $CV_{\sim \beta} > CV_\beta$  to hold. However, even if not, as  $E[c_t]$  increases,  $\delta q(CV_{\sim \beta} - CV_\beta) \rightarrow 0$ . Thus, both  $\omega$  and  $\psi$  can be positive as long as  $E[c_t]$  is sufficiently large. As long as enough of the probability density is packed between zero and these upper boundaries (i.e. a large enough  $\mu_\beta$  with a small enough variance)

then  $EU_p[\textit{sanction} | \hat{\mu}_{\beta,t}] - EU_p[\textit{not sanction} | \hat{\mu}_{\beta,t}] > 0$  can hold. The strategy is sequentially rational because the public is always indifferent between punishing or not punishing in any given round.

## References

- Calvert, Randall L. 1995. "Rational Actors, Equilibrium, and Social Institutions." In Jack Knight and Itai Sened, *Explaining Social Institutions*. University of Michigan Press.
- Ferejohn, John, and Barry Weingast. 1992. "A Positive Theory of Statutory Interpretation." *International Journal of Law and Economics*. 12:263.
- Ferejohn, John, Frances Rosenbluth, and Charles Shipan. 2004. *Comparative Judicial Politics*. Manuscript.
- Gely, Rafael and Pablo T. Spiller. 1990. "A Rational Choice Theory of Supreme Court Statutory Decisions with Applications to the State Farm and Grove City Cases." *Journal of Law, Economics, & Organization*. 6(2): 263-300.
- Hansford, Thomas G. and David F. Damore. 2000. "Congressional Preferences, Perceptions of Threat, and Supreme Court Decision Making." *American Politics Quarterly*. 28(4): 490-510.
- Harvey, Anna and Barry Friedman. 2006. "The Limits of Judicial Independence: The Supreme Court's Constitutional Rulings, 1987-2000." *Legislative Studies Quarterly*. Forthcoming.
- Landes, William M. and Richard A. Posner. 1975. "The Independent Judiciary in an Interest-Group Perspective." *Journal of Law and Economics*. 18(3): 875-901.
- Ramseyer, J. Mark. 1994. "The Puzzling (In)Dependence of Courts: A Comparative Approach." *The Journal of Legal Studies*. 23(2): 721-47.
- Rasmusen, Eric. 1994. "Judicial Legitimacy as a Repeated Game." *Journal of Law, Economics, & Organization*. 10(1): 63-83.

- Rogers, James R. 2001. "Information and Judicial Review: A Signaling Game of Legislative-Judicial Interaction." *American Journal of Political Science*. 45(1): 84-99.
- Rosenberg, Gerald N. 1992. "Judicial Independence and the Reality of Political Power." *The Review of Politics*. 54(3): 369-98.
- 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(4): 349-79.
- Shipan, Charles. 1997. *Designing Judicial Review: Interest Groups, Congress, and Communications Policy*. Ann Arbor: University of Michigan Press.
- Shipan, Charles. 2000. "The Legislative Design of Judicial Review: A Formal Analysis." *Journal of Theoretical Politics*. 12(3): 269-304.
- Staton, Jeffery K. 2006. "Constitutional Review and Selective Promotion of Case Results." *American Journal of Political Science*. 50(1): 98-112.
- Stephenson, Matthew C. 2003. "'When the Devil Turns...': The Political Foundations of Independent Judicial Review." *Journal of Legal Studies*. 32(1): 59-89.
- Stephenson, Matthew C. 2004. "Court of Public Opinion: Government Accountability and Judicial Power." *Journal of Law, Economics, & Organization*. 20(2): 379-99.
- Vanberg, Georg S. 2005. *The Politics of Constitutional Review in Germany*. 2005. Cambridge: Cambridge University Press.
- Weingast, Barry R. 1997. "The Political Foundations of Democracy and the Rule of Law." *The American Political Science Review*. 91(2): 245-63.
- Whittington, Keith E. 2005. "'Interpose Your Friendly Hand': Political Supports for the Exercise of Judicial Review by the United States Supreme Court." *The American Political Science Review*. 99(4): 683-96.

**Table 1**

<b>Notation</b>	<b>Description</b>
$\chi_t = \beta_t - c_t$	government's total cost of compliance in period $t$
$\beta_t \sim (\mu_\beta, \sigma_\beta)$	net societal benefit in period $t$
$c_t \sim (\mu_c, \sigma_c)$	government's private cost of compliance in period $t$
$b$	societal benefit from the other government complying
$k_t \sim (\mu_k, \sigma_k)$	litigant's cost of bringing a case in period $t$
$c_p$	public sanctioning cost
$\delta$	discount rate
$\hat{\mu}_{\beta,t}$	public's belief over the expected net societal benefit in period $t$
$j_t$	court's judgment in period $t$

**Table 2**

		<b>Government 2</b>	
		<b>Cooperate</b>	<b>Defect</b>
<b>Government 1</b>	<b>Cooperate</b>	$b + \chi_t$ $b + \chi_t$	$\chi_t$ $b$
	<b>Defect</b>	$b$ $\chi_t$	$0$ $0$

**Figure 1: The Sequence of Events**

