Game Theory and Judicial Behaviour

Arthur Dyevre, *Max Planck Institute for International and Comparative Law*
Game Theory and Judicial Behaviour

Arthur Dyevre
Max Planck Institute for International and Comparative Public Law
Im Neuenheimer Feld 535
69120 Heidelberg
Germany
arthurdyevre@hotmail.com

Abstract

The present paper surveys applications of game theory in the positive literature on judicial decision-making. As we shall see, scholars have tried to model various aspects of the strategic environment in which judges are thought to operate. Early applications of game theory to judicial behaviour concentrated on the interactions between courts and the other branches of power. More recently, scholars have tried to address the internal deliberative process of judicial institutions, to predict how judges sitting on the same court may engage in bargaining with each others to ensure that the final outcome is as close as possible to their ideal point. Meanwhile, others have sought to model interactions among judges sitting on different courts through the doctrine of precedent and the hierarchical structure of the judiciary. The paper concludes with a brief consideration of issues neglected in the existing literature and some suggestions for future research.
The use of game theory to model judicial behaviour is motivated by the understanding that judging has a strategic dimension. Judicial decisions are not made in a vacuum. Judges are aware that their decisions may bring about reactions. Legislators may respond to their rulings by passing override legislation, by stripping the court of its jurisdiction, by cutting its budget or, worse, by starting an impeachment procedure. Similarly, unless they themselves sit at the top of the judicial pyramid, judges are aware that their judgments may be reversed by the appellate or supreme court. So, in order to advance their interpretation of the law or any other policy goal, they know they need to anticipate the social responses elicited by their own behaviour.

As in other fields, game-theoretic models of judicial decision-making often greatly simplify reality and cannot claim to rest on incontrovertible assumptions. Judges are supposed to be rational agents acting in purposive, forward-looking fashion with transitive and cardinally ordered preferences. More importantly perhaps, they are supposed to engage in complex calculations – carefully weighing costs and benefits while computing their probability and discount factor – before making any choice. Still, even though no single model may ever be able to capture all the complexities and intricacies of judging, game theoretic models provide a useful consistency-check for the development of more sophisticated theories of the judicial
process. By making all assumptions explicit and forcing us to work out their exact implications, game theory brings discipline to our intuitions and helps us in identifying the exact set of conditions under which a particular outcome will obtain.

The present paper surveys applications of game theory in the positive literature on judicial decision-making.\(^1\) As we shall see, scholars have tried

\(^1\) This is not to say that models which have a more explicitly or directly normative purpose are not interesting. They certainly are. Many of the models presented in the law and economics literature serve to back what are in fact policy recommendations addressed to the courts. Other normative applications, most notably in institutional economics and political economy, focus on the broader issue of institutional choice. Maskin and Tirole (2004) for example compare various policy-making regimes and their impact on public welfare depending on whether elected officials are ready to pander to the electorate to gain re-election, the probability that judicial preferences are congruent with those of the population, the policy expertise of the average citizen, the cost of acquiring information, etc. They conclude that the most important decisions should be taken by elected officials while technical issues are better allocated to judges. They also suggest that the judicial regime does better than both direct and representative democracy when the majority’s preferences are very likely to inflict large negative externalities on a minority. Arguably models of this kind can be used as positive models to make predictions about real world institutions (e.g. one implication of the Maskin and Tirole’s model would be that minority rights are better protected, other things being equal, in countries with a strong judiciary). However, these models rest on broad assumptions about judicial motives and, because they focus on institutional choice, they do little to explain or predict variations in judicial outcomes. The same goes for the aforementioned law and economics models. They may offer guidance as to how judges ought to decide cases, but they are of very little avail when it comes to make
to model various aspects of the strategic environment in which judges operate. Most of the early applications of game theory to judicial behaviour concentrated on the interactions between the courts and the other branches of power. More recently, scholars have tried to address the internal deliberative process of judicial institutions, to predict how judges sitting on the same court may engage in bargaining with each others to ensure that the final outcome is as close as possible to their ideal point. Meanwhile, others have sought to model interactions between judges sitting on different courts through the doctrine of precedent and the hierarchical structure of the judiciary. I'll conclude with a brief consideration of issues that have been neglected by the existing literature and some suggestions for future research.

**Courts and the Separation of Powers**

The emergence of game theoretic models of judicial decision-making coincides with the expansion of the rational choice variant of the neo-institutional paradigm in American political science. Drawing on the work of modern Congressional scholars who had tried to model relations among legislative committees or between legislators and administrative agencies (Shepsle and Weingast 1987), Marks (1988), Gelly and Spiller (1990) and Eskridge (1991a, b) considered the potential for Congress to overturn the decisions of the Supreme Court.

predictions about how judges will decide cases in fact – the focus of the models reviewed in this paper.
Judicial Decision-Making and Political Fragmentation

Figure 1.1 depicts the extended form of the game laid out by Eskridge (1991), admittedly one of the most accessible and influential among the early studies of the judicial process using insights from game theory.

Figure 1.1. Extended Form of the Separation-of-Powers Game

At the initial stage of the sequence of play, the Court interprets a federal statute. Next the relevant congressional committee has to decide how to react to the ruling. The committee may choose to do nothing, in which case the ruling is left untouched. But it may also decide to refer to Congress a bill overturning the ruling. If it does, Congress will have to choose whether to adopt or to reject this attempt to override the Court. If Congress adopts the bill (or a modified version thereof), the President will have to decide whether to veto the bill or to sign it into law. Then, if the President puts his veto, Congress will have to decide by a two-third majority whether to
override it. With the strategic space thus specified, identifying the set of conditions under which Supreme Court decisions can be reversed is relatively straightforward if we assume that each player will choose the course of action that achieves the outcome which is as close as possible to her ideological ideal point in the knowledge the other players will do the same.

As in Figure 1.2, the players’ preferences can be arrayed on a single-peaked, one dimensional policy space. Here liberal-conservative. The letters stand for the ideal points of the different actors. J denotes the preferred position of the Supreme Court; M is the preferred position of the median member of Congress\(^2\); V denotes the “veto median”, the point at which one third of the legislators are on one side of the policy outcome and two-thirds on the other (that is on the side of C); P is the ideal point of the President; and C represents the most preferred position of the key committees in Congress that decide whether to propose a bill to their respective houses, while C(M)

\(^2\) Following the Median Voter Theorem, in an institution where decisions are taken by majority vote, what matters most, ultimately, is the position of the median voter, because she will determine which side the majority falls on.
denotes the committees’ indifference point in relation with M (they have no preference for a policy at M over a policy at C(M) and vice-versa). In the situation depicted in Figure 1.2, Eskridge’s Separation-of-Powers Game predicts the Supreme Court will be able to vote its preferred position into its decision. The congressional committees will not want to send a bill to the floor because the ideal point of Congress (M) is not closer to the their ideal point (C) than a Court decision at (J). Note, moreover, that (J) coincides with (P) and (V). This means that even if the committees preferred (M) to a Court decision on (J) and sent a bill to the floor, the President would put his veto and Congress would lack the two-thirds majority required to override it. The outcome x = J satisfies the definition of a Nash equilibrium: no player will get better off by unilaterally deviating from it.

In Eskridge’s model, this equilibrium holds as long as the Supreme Court’s ideal point lies within the interval [M, C(M)], i.e. as long as the congressional committees’ preferred position is not closer to (M) than it is to (J). By contrast, when the Court’s ideal point lies outside this interval, an outcome on its ideal point will not be in equilibrium. Consider the situation represented in Figure 1.3. Were the Court to issue a ruling at its ideal point or anywhere to the right of C(M), the committees would want to send a bill to the floor of Congress because an outcome at (M) would make them better off. So the Court’s best strategy in that situation is to behave strategically and issue a ruling at (C(M)) so that committees have nothing to gain from sending a bill to Congress. A ruling at (C(M)) is in equilibrium.
**Figure 1.3. Constrained Court**

We can easily transpose Eskridge’s framework to other institutional contexts. In Figure 2.1 we see it applied to the ECJ under the codecision procedure.

**Figure 2.1. The ECJ in the Codecision Game**

Under the codecision procedure, overriding an ECJ decision interpreting a directive requires a Commission proposal, the approval of the European Parliament (EP) as well as that of a qualified majority in the Council – the body representing national governments. From Figure 2.1 it is easy to see that the ECJ needs to be aligned with just one of the three players in the
codecision game to prevent the enactment of override legislation. Similar to
the US Supreme Court, the model predicts that ideological fragmentation
among the actors of the legislative process will result in equilibrium
outcomes favourable to the ECJ. Figure 2.2 illustrates one such outcome.

\[ \text{Equilibrium Result, } x = \text{Com} = J \]

**Figure 2.2. ECJ with Commission Support**

Here the judges are in position to issue a ruling on their ideal point (J)
because the Commission (Com) has no interest in proposing an override
directive. What is more, even if the Commission were to make a proposal,
the Parliament (EP) and the Council’s qualified majority (QM) would not be
able to agree on an override bill as long as a Court decision lies somewhere
in the space between QM and EP. This is because every outcome in the
interval [QM, EP] is Pareto-optimum from the viewpoint of the Parliament
and the Council’s qualified majority. Any change to the outcome would
necessarily make one of these two players worse off.

The policy debate about the definition of working hours in labour law
illustrates how this kind of model can help make sense of the EU law-
making process. In its **SIMAP**\(^3\) and **Jaeger**\(^4\) rulings, the ECJ had held that

---

\(^3\) Case C-303/98, 3 October 2000, *Sindicato de Medicos de Asistencia Pública (SiMAP) v. Conselleria de Sanidad y Consumo de la Generalidad Valenciana.*

under the 1993 Working Time Directive on-call duties should count as working hours for the purpose of work and rest calculation when employees are required to be present on site. The impact on the health care sector in Member States where medical staff and junior doctors were traditionally required to be resident on site while on call was tremendous. So much so that when the Commission made a proposal for a new working time directive the Council insisted on having the definition of working hours revised to exclude on-call duties. This was clearly an attempt to overrule the ECJ jurisprudence. The Council and national governments, however, could not bring the European Parliament round to support their override attempt. In April 2009, after years of negotiations, MEPs rejected the redefinition of working hours and, with it, the new directive. According to the Parliament, on-call time must remain working time in accordance with the jurisprudence of the ECJ. In other words, the Court’s case-law was left undisturbed because MEPs preferred it to the proposal backed by the Council.

---


7 Note that quantitative studies have shown that the position of the Commission is a strong predictor of Court of Justice decisions, especially in infringements proceedings (Stone Sweet 2004). Interpreted as evidence of ideological convergence (rather than as a judicial...
Having said this, a model like the one built by William Eskridge has obvious limitations. First, it focuses on how legislators and the executive branch may react \textit{ex post} to judicial pronouncements and as such it says nothing of how courts may influence the legislative process \textit{ex ante}, notably at the constitutional level through the power of judicial review. Second, it assumes that the players have perfect information about each others’ preferences. It is not a plausible assumption because if it were true Supreme Court decisions would never be overruled by Congress or only when the preferences of legislators and committee members change. An implication that even Eskridge himself shows to be empirically incorrect (Eskridge 1991).

\textit{Imperfect Information, Public Support and Legislative-Judicial Relations}

Describing the interactions between a court and a legislature as a two-player game, the model put forward by Vanberg (2005) goes some way to address these limitations. Incorporating public support for the court and transparency (the public’s awareness that the policy issue under consideration is being dealt with by the Court and the legislature) as parameters of legislative and judicial behaviour, it also demonstrates how stratagem designed to encourage the Commission to bring more cases), this fact would suggest that many attempts to override the Court are nipped in the bud, as it were, because the Commission is able to prevent override proposals from being made in the first place through its agenda-setting monopoly. For students of EU judicial politics, alas, this means that documenting the size of the phenomenon will be difficult, if not impossible, because override attempts leave few apparent traces in the legislative process.
public opinion can strengthen judicial power even in strongly majoritarian systems where political fragmentation is low.

Figure 3.1. Extended Form of Vanberg (2005)'s Court-Legislature Game

The game (Figure 3.1) starts with Nature choosing an environment (public support or no public support for the Court)\(^9\) and a Court type (convergent or not convergent with the Legislature). Then the Legislature must decide whether to legislate. If it legislates, the Court will decide whether to veto the new legislation. If the bill is vetoed by the Court, the last move belongs

\(^9\) For the sake of simplicity I leave out the transparency parameter. Since transparency matters only when there is public support for the Court and public support only when there is transparency, this makes very little difference.
to the Legislature, which will decide whether to evade or to obey the judicial veto. In the model, the Legislature’s utility function has three components: (1) legislators want to implement their policy preferences and receive $\alpha$ whenever they achieve to do so; but (2) legislating is costly and legislators have to pay cost $\epsilon$ every time they choose to legislate (with $\alpha > \epsilon > 0$); (3) moreover, attempts to evade judicial pronouncements may result in a public backlash if the public backs the Court, damaging the Legislature’s political capital with cost $\beta$ (with $\beta > 0$). As to the Court’s utility function, it has two elements: (1) the court wants the law to reflect its preferred policy, so it gains $A$ whenever its preferred policy is implemented ($A > 0$); (2) the court wants to avoid non-compliance on the part of the legislature, so it pays institutional cost $I$ when the legislature evades its decision ($I > 0$). Obviously, the players' payoffs depend on whether the public backs the Court and on whether the Court is convergent or not. However, neither the judges, nor the members of the legislature know for sure whether the public will back the Court in the event of a showdown. Similarly, at the initial drafting stage, when they have to decide whether to legislate, the legislators do not know with certainty whether the Court shares their policy preferences.

From these assumptions we can derive six perfect Bayesian equilibria corresponding to positions of varying judicial strength vis-à-vis the legislature. In Figure 3.2 the six equilibria are plotted against the parameters of public support and ideological divergence (i.e. the probability that the Court will not be convergent).
Proceeding by backward induction from the last stage of the game, we can see, for example, that if the expected payoff of evading a judicial veto is less than the cost of legislating \((1-q)(\alpha - \varepsilon) + q(-\beta - \varepsilon) < -\varepsilon\)\(^{10}\), then the Legislature will decide not to evade the veto (\(\neg E\)). At the previous stage, the non-convergent Court, knowing the Legislature will not want to evade its ruling, will not hesitate to veto the Legislature’s bill. With no need to worry about its institutional standing, the non-convergent Court is strictly better off using its veto \((A > 0)\). Knowing that, the Legislature’s decision to legislate at the drafting stage will depend on whether legislators believe the Court is likely to be convergent. If the expected payoff from the Court being

---

\(^{10}\) This formula simply captures the fact that the legislature has to weigh the probability of a successful evasion (i.e. \((1-q)(\alpha - \varepsilon)\)) as well as that of an unsuccessful evasion \((p(-\beta - \varepsilon))\) against the cost of compliance \((-\varepsilon)\).
convergent, minus the expected payoff from the Court being non-convergent, is less than the payoff of not legislating (that is if: \( p(\alpha - \varepsilon) - (1 - p)\varepsilon < 0 \), then the legislature will refrain from legislating. If it is greater, then the Legislature will legislate, but if the Court turns out to be non-convergent and invalidates its bill then the Legislature will comply. The former situation corresponds to the Legislative Self-Censorship Equilibrium \( A \). The latter to the Judicial Supremacy Equilibrium. Other equilibria can be derived in like fashion.

Vanberg's model has many interesting implications. One is that high rates of judicial annulments are most likely to be found where a powerful court faces legislators who often wrongly believe that the judges are on their side. Where legislators expect – rightly or wrongly – the court to be divergent, the model predicts that, other things being equal, they will prefer to refrain from legislating. The studies of Christine Landfried on the German legislative process (Landfried, 1984) and Alec Stone on French judicial politics (Stone, 1992) are consistent with this prediction. They show that the fear of judicial annulment may induce self-censorship on the part of legislative majorities. German politicians call the phenomenon “Karlsruhe Astrologie”. Trying to guess how the judges will respond to their policy initiatives, the legislators prefer to water down their bills or abandon them altogether rather than endure a judicial veto (von Beyme, 1997: 311). Meanwhile, the model suggests that courts are likely to be the weakest where they enjoy little public support and legislators feel safe to evade
rulings they dislike, as would seem to be the case with the Russian Constitutional Court (Epstein et al. 2001).

The comparative statics enables us to work out how changes in the model’s parameters can lead to a change in the prevailing equilibrium.
Table 1. Comparative Statics of Vanberg’s Model

<table>
<thead>
<tr>
<th></th>
<th>$A$ increases (favours the Court)</th>
<th>$I$ increases (favours the Legislature)</th>
<th>$\alpha$ increases (favours the Legislature)</th>
<th>$\beta$ increases (favours the Court)</th>
<th>$p$ increases (favours the Legislature)</th>
<th>$q$ increases (favours the Court)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Legislative Self-Censorship Equilibrium</td>
<td>No change</td>
<td>Judicial Self-Censorship</td>
<td>Judicial Supremacy</td>
<td>Contentious</td>
<td>Judicial Supremacy</td>
<td>Contentious</td>
</tr>
<tr>
<td>Judicial Supremacy Equilibrium</td>
<td>No change</td>
<td>No change</td>
<td>Contentious</td>
<td>No change</td>
<td>No change</td>
<td>No change</td>
</tr>
<tr>
<td>Contentious Equilibrium</td>
<td>No change</td>
<td>Judicial Self-Censorship</td>
<td>No change</td>
<td>Judicial Supremacy</td>
<td>Legislative Self-Censorship</td>
<td>No change</td>
</tr>
<tr>
<td>Judicial Self-Censorship Equilibrium</td>
<td>Contentious</td>
<td>Judicial Supremacy</td>
<td>Legislative Self-Censorship</td>
<td>No change</td>
<td>Contentious</td>
<td>Legislative Self-Censorship</td>
</tr>
</tbody>
</table>

18
Vanberg’s model provides a powerful framework to compare judicial institutions across time, countries and policy areas (Dyevre 2010).

The Internal Deliberative Process of Judicial Bodies: The Collegial Game

The models we have looked at so far treat courts as unitary actors. Yet most courts are collegial bodies and judges sitting on the same court do not necessarily share the same preferences. So how about using game theory to explain the behaviour of individual judges?

At first blush, if we assume that the court’s external institutional environment gives the judges sufficient leeway to vote their preferences into their decisions, it seems that beyond the preferences of the individual judges there is little else we need to know to predict the outcome of the collegial game. If decisions are taken by simple majority, then they should reflect the position of the median judge.

Accordingly, in a nine-judge tribunal like the US Supreme Court, the justice in the court’s ideological middle should be the swing-vote (judge J5 in Figure 4).

However, although it may work for the decision on the merits, this account of the judicial deliberation process seems inadequate when it comes
to explaining the content of judicial opinions. Judicial opinions generally go beyond merely stating the reasons for the decision on the merits. They also specify the position of the court on the broader policy issue raised by the case. Now the rules (formal or informal) governing the judicial process do not automatically assign opinion writing to the median judge. The power to assign opinion is often wielded by the chief justice or the president of the court. So scholars have tried to model the influence of this institutional feature on the deliberative dynamic.

Figure 5. The Opinion Writer as Agenda-Setter

Figure 5 depicts the ideal points of the members of a five-judge court in a two-dimensional policy space. The circles around the judges' ideal points define their respective Preferred-to Set, the set of outcomes each judge prefers to the Status Quo (SQ). The hatched areas where at least three circles overlap define the Win-Sets of the Status-Quo, the sets of outcomes that a majority of judges prefer to SQ. Here we straightforwardly see that $J_3$ is the swing vote and that there are two possible winning coalitions: $\{J_1, J_2,$
Now let us suppose that $J_5$ wields the power to assign opinion writing. As every judge will seek to get an outcome as close as possible to her ideal point, assigning the opinion to $J_1$ would produce an outcome somewhere in the more rightward Win-Set, such as $A$. But $J_5$ would be better off if she strategically assigned the opinion either to $J_4$ or to herself. This would result in an outcome somewhere in the Win-Set to the left, such as $B$.

While spatial voting analysis suggests that the opinion-assigner can indirectly influence the content of the final opinion, its exact content, however, will devolve on whether the assignee is able to present her draft opinion as a take-it-or-leave-it proposal. If she is, then she will be able to pick the point closest to her preferred position in the Win-Set (Hammond et al. 2005). If she is not – that is, if the other judges can amend her proposal – the final outcome should be somewhat closer to the median judge on the relevant issue dimension. There is no closed rule prohibiting counter-offers in the US Supreme Court deliberation protocol. So some scholars have considered models that make both the assignee and the assigner irrelevant. In these models, the opinion of the court, just as the decision on the merits, corresponds to the point closest to the median judge in the Win-Set of the status-quo (Hammond et al. 2005: 161-2; Jacobi 2009; Spiller and Spitzer 1995). Yet this conclusion does not square with the findings of extant empirical studies, which indicate that, despite the absence of a formal gag rule, the author’s identity does have some effect on the content of opinions (see e.g. Clark and Lauderdale 2009). Lax and Cameron (2007), Carrubba et
al. (2007) and Cameron and Kornhauser (2010) address this problem by adding an opinion-writing cost to the judges’ utility function. Because crafting high-quality opinions is costly, a judge may prefer to join an opinion he would otherwise challenge.

Of the more recent attempts to model adjudication on collegial courts, Cameron and Kornhauser (2010) is admittedly the most developed and refined. Unlike pre-existing models, which generally conflate decision-making over case dispositions and decision-making over policies (opinions),¹¹ they carefully distinguish dispute resolution and policy-making, which are conceptualised as both distinct and interrelated.¹² The sequence of play in their adjudication game is as follows:

¹¹ Carrubba et al. (2008) was the notable exception.

¹² The content of an opinion is defined as a rule or function mapping cases into dispositions. A rule has a cut-point \( x \) in policy space \( X \) establishing two equivalence classes with respect to dispositions. (Dispositions are thus conceived as dichotomous, i.e. “for Claimant” or “for Defendant”.) For any case \( \hat{x} \), the rule yields the “correct” disposition:

\[
 r(\hat{x}, x) = \begin{cases} 
 0 & \text{if } \hat{x} < x \\
 1 & \text{if } \hat{x} \geq x 
\end{cases}
\]

Where 0 indicates one disposition and 1 the other. Arguably, many legal rules take this form. Speed limits are an obvious example. They employ a cut-point to mark off the class of acceptable driving behaviour from the class of reprehensible driving behaviour potentially facing legal sanctions. For instance, if 60 mph is the cut-point, a driver exceeding the limit will normally fall in the class of reprehensible behaviour while a driver below the limit will normally fall in the class of accepted driving behaviour. Of course, a lower or higher cut-


1. A case $\hat{x}$ arrives.
2. A judge $j$ is designated, who writes an opinion $x_j$.
3. Acting simultaneously, the non-writing judges first choose whether to join the opinion.
4. Then they vote on the disposition of the case.

In accordance with prevailing jurisprudential norms, the pair of actions $\{3,4\}$ must obey the endorsement-consistency constraint: if a judge decides to join opinion $x_j$, she must also vote for the case disposition entailed by $x_j$. In other words, if the rule spelled out in the opinion entails that the defendant should win, a judge cannot join the opinion and then cast a vote for the plaintiff.

Figure 6.1 illustrates the corresponding game form associated with the sequence of play for a three member court. Justice 1 is the opinion writer. Opinions to one side of the case $\hat{x}$ entail disposition 1 while opinions on the other side entail disposition 2. After Justice 1 has delivered her opinion, Justice 2 decides whether to join and then casts a dispositional vote. Simultaneously (information sets are shown with dashed lines) Justice 3 does likewise. Due to the endorsement-consistency constraint, some portions of the game tree are unreachable. They appear in grey in the figure. The opinion author is assumed to join her own opinion. Outcomes for each permissible combination of moves are shown at the terminal nodes.

_____________________________

point entails a different partition of the case space. Someone driving at 50 mph will be “safe” with a cut-point at 60 but not with a cut-point at 45.
Figure 6.1. Game Form of Cameron and Kornhauser (2010) for a Three Member Court

Non-authors receive payoffs based on their dispositional vote, join decision, the opinion’s content, and the case location. Each non-writing judge as an ideal rule $\vec{x}_i$ on policy space $X$, which in turn determines her preferred...
disposition for every case on case space $\hat{X}$. When a non-writing judge joins an opinion $x_j$, she suffers a loss in proportion to the distance between $x_j$ and $\overline{x}_i$. The larger the distance, the larger the incurred loss. If she declines to join the opinion, she does not suffer this loss but she must pay the effort cost $k$ required to write a concurring or dissenting opinion. Finally, if her dispositional vote is not in accord with her preferred disposition of the case, she incurs a dispositional loss $\gamma$ ($\gamma \leq 0$).\(^{13}\) Although the writing judge cares about case disposition and opinion content too, her utility function differs in three respects. First, she cares about the number of joins received by her opinion. She wants her opinion to be as authoritative as possible. Accordingly, provided her opinion is compatible with the majority-winning disposition, she receives $\beta$ for each additional join. Second, given that she must write an opinion anyway, she does not face any relevant writing cost. However, she cares whether a majority of the judges are in dissent and she suffers a large loss $\kappa$ when she fails to author a majority-disposition compatible opinion. Cameron and Kornhauser require $\kappa$ to be large enough so that the most attractive majority-disposition incompatible opinion is always worse for the author than the least attractive majority-disposition安置

\(^{13}\) Accordingly, the utility function of judge $i$ over dispositional vote $d_i$ ($d_i \in D = \{0,1\}$) and join decision $s_i$ ($s_i \in S = \{0,1\}$) given opinion $x_j$ and case $\hat{x}$ is defined as:

$$u_i(d_i, s_i, x_j, \hat{x}) = s_i v(x_j, \overline{x}_i) - (1 - s_i)k - \gamma(d_i, \hat{x}, \overline{x}_i)$$

Where $I(d_i, \hat{x}, \overline{x}_i)$ is the indicator function defined as:

$$I(d_i, \hat{x}, \overline{x}_i) = \begin{cases} 1 \quad \text{if} \quad d \neq r(\hat{x}, \overline{x}_i) \\ 0 \quad \text{otherwise} \end{cases}$$
compatible opinion (Cameron and Kornhauser 2010: 17). In short, the writing judge is always better off penning an opinion with majority support.\textsuperscript{14}

Equilibrium strategy profiles are derived from these specifications, using backward induction. Starting from the last stage of the game, the dispositional vote, we easily see that a non-writing judge should vote for the disposition reflecting her ideal point \( \bar{x}_j \), unless she joins the opinion \( x_j \), in which case she must vote for the disposition required by the opinion. At the join decision stage, the non-author must weigh the payoff associated with a join against the payoff associated with a decision not to join. Each judge has a join region around her ideal-point. Her join decision calculus can thus be represented by the function in Figure 6.2.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{join_region.png}
\caption{Non-Writing Judge Join Region}
\end{figure}

\textsuperscript{14} The opinion writer’s utility function takes the following form:

\[
u_j(d_j, s_j = 1, x, \hat{x}) = \begin{cases} 
\beta n(x_j) + \nu(x_j, \bar{x}_j) - \mathcal{H}(d_j, \hat{x}, \bar{x}_j) & \text{if } \tilde{d} = r(\hat{x}, x) \\
\nu(x_j, \bar{x}_j) - \mathcal{H}(d_j, \hat{x}, \bar{x}_j) - \kappa & \text{otherwise}
\end{cases}
\]
The horizontal axis represents both the case and the policy space, while the vertical axis indicates the probability to join (1 = join, 0 = not join). As we can see, given her ideal rule \( \bar{x}_i \) and case \( \hat{x} \), judge \( i \)'s decision to join depends on the location of opinion \( x_j \). If \( x_j \) is sufficiently close to her ideal point, she will join, if not, she won't. An important consideration in the judge's decision calculus is whether the opinion and the judge's ideal cut-point lie on opposite side of the case. To the extent that a judge cares about cases disposition (\( \gamma < 0 \)), opposite-side opinions are less attractive to her than same-side opinions. A judge will join an opposite-side opinion only if the sum of the resulting policy loss and dispositional loss is less than the cost of writing a separate opinion. For same-side opinions, by contrast, only the size of the policy loss and the size of the writing cost are relevant.

The opinion author must choose her opinion from the set of majority-disposition compatible opinions so as to maximise the net gains from joins less the loss of departing from her most-preferred rule and any dispositional loss. For any parameter values, the model predicts the location of the opinion, the size of the dispositional majority, as well as the number of joins, dissents and concurrences.

The baseline example provided by Cameron and Kornhauser helps illustrate the basic logic and rich comparative statics of the model. We assume that the case at hand does not present the nine judges with a dispositional value (\( \gamma = 0 \)), that writing cost \( k = 0.225 \) and that the court is non-polarized, so that judge 1 has an ideal point at 0.1 on the unit interval [0,1], judge 2 at 0.2, judge 3 at 0.3, and so on. Suppose we have a case \( \hat{x} = \)
0.55 and Judge 9 is the author. Figure 6.3 shows the number of dissents (dashed line) and joins (solid line) corresponding to the location of her opinion. Opinions to the right of the vertical at \( x_j = 0.725 \) are not majority-disposition compatible (a majority of the court will be in dissent). Consequently, Judge 9 must locate her opinion to the left of that point.

**Figure 6.3. Dissent Function and Aggregate Join Function**

The location she will choose within the majority-disposition compatibility set will depend on how much she values authoritativeness. If the value of additional join \( \beta > 0.1 \), she will want to move further away from her ideal point to write an opinion at \( x_j = 0.625 \) that will receive five joins in addition to her own (recall that the author is assumed to join her own opinion). If the value she puts on joins is comparatively lower, \( \beta \leq 0.1 \), then \( x_j = 0.725 \) will be the equilibrium opinion.

Another interesting implication of the model is that the relation between opinion content and opinion assignment is non-monotonic:
assigning an opinion to a more extreme judge may result in a more moderate opinion.

Figure 6.4. Non-monotonicity of Opinion Location and Author Ideal Point

The example in Figure 6.4 assumes an extreme case location ($\hat{x} = 0.0$ or $\hat{x} = 1.0$), a join value $\beta = 0.06$, and, as in the previous example, a non polarized court, a writing cost $k = 0.225$, and a dispositional value of $\gamma = 0$. The diagram shows the effect of change in opinion assignment on opinion content holding these parameter values constant. It highlights at least two things. First, the decision to assign an opinion to a particular judge is potentially very consequential. Here each judge authors a different opinion. Second, opinion location is not monotonic in ideal point of the opinion

---

15 Extreme case location will result in a unanimous dispositional vote.
author. As we can see, Judges 3 through 7 write opinions at their ideal point. Because they occupy the court’s middle ground, they need not deviate from their preferred rule to garner joins. In contrast, Judges 2 and 8 do. They are respectively better off writing an opinion slightly to the right and slightly to the left of their ideal point. What is more, Judges 1 and 9 write opinions that are even more moderate. Because they do not have to worry about losing the join of a more extreme judge, they can move towards the centre to seek additional joins.

Judges and Future Judges: Modelling the Effect of Precedents on Judicial Decision-Making

Especially in Common Law countries, where courts are supposed to follow the doctrine of *stare decisis*, but also in Civil Law jurisdictions where past judicial rulings are regularly invoked as constituting an authoritative guide for present decisions, precedents are believed to affect judicial behaviour significantly. Arguably, this means that judges sometimes forego what would otherwise be their favourite outcome because a past decision suggests a different conclusion. Rasmusen (1994) seeks to uncover the conditions under which judges will do so by modelling interactions among past, present and future judges as an infinitely repeated game. In his model, the players’ payoffs reflect the cost of following past precedents and the gain derived from having their own precedents followed. Infinite repetition makes strategies possible that would not be sustainable in finite games. In a finite game, the expectation that judges will follow precedent may rapidly unravel.
as judges face the temptation to free-ride on their colleagues. On the other hand, when the game is repeated an infinite number of times punishing defectors becomes possible and this alone may suffice to deter defection and ensure cooperation. Rasmusen examines the following six strategy profiles from the standpoint of judge \( j \) in an infinite sequence of judges:

1. **Strategy**: Violate every precedent. **Outcome**: Every judge violates every precedent (Judicial Breakdown Equilibrium).

2. **Strategy**: Obey every precedent unless some previous judge has violated a precedent. In that case violate every precedent. **Outcome**: Every judge obeys every precedent (Punishment Breakdown Equilibrium).

3. **Strategy**: Obey the precedent of judge \( j − i \), for \( i = 1, \ldots, n \), if she “retained legitimacy”, and violate it otherwise. Judge \( j − i \) is considered to have “retained legitimacy” if she herself followed this strategy. **Outcome**: Every judge follows every precedent (Specific Punishment Equilibrium).

4. **Strategy**: Obey the precedent of judge \( j − i \), for \( i = 1, \ldots, n − 1 \), if she “retained legitimacy”, and violate it otherwise. Violate the precedent of judge \( j − n \). Judge \( j − i \) is considered to have “retained legitimacy” if she herself followed this strategy or obeyed all \( n \) precedents. **Outcome**: Every judge obeys \( n − 1 \) precedents and violates one precedent (Lax Specific Punishment Equilibrium).

5. **Strategy**: Obey the precedent of every judge. **Outcome** (if all judges follow the strategy): Every judge follows every precedent.
(6) **Strategy**: Obey the precedent of every judge unless the immediately preceding judge violated any precedent. In that case, violate every precedent. **Outcome** (if all judges follow the strategy):

Every judge obeys every precedent.

Unlike the other four, (5) and (6) are not equilibrium strategies. (5) is not a stable equilibrium for the simple, aforementioned reason: if obeying precedent is costly, then \( j \) is better off deviating from the strategy. (6) is not a stable strategy profile because punishment is not credible. Consider the situation of judge \( j \) after judge \( j - 1 \) has violated a precedent. If she follows the strategy, then judge \( j + 1 \) will also violate all precedents if she in turn follows the strategy. However, as long as the gain from having her precedent followed is greater than the cost of obeying precedents, judge \( j \) will be better off obeying all precedents and foregoing punishment so that \( j + 1 \) and all subsequent judges uphold her precedent. In other words, the strategy is not stable because punishment costs the punisher too much. On the face of things, strategy (2) appears similar to (6). But in fact it is more effective because no single judge has the power to stop the destructive consequence of her predecessor’s deviation. To continue the previous example, after judge \( j - 1 \) has violated a precedent, \( j \) has no interest in foregoing punishment, since \( j + 1 \) will violate precedent anyway.

As in other infinitely repeated games, the existence of multiple Nash equilibria underscores the importance of initial expectations. When the gain from having one’s favourite holding followed outweighs the cost of obeying past precedents, the Punishment Breakdown, the Specific Punishment and
the Lax Punishment Equilibria are all Pareto superior to the Judicial Breakdown Equilibrium. Even when this condition is met, however, nothing in the model’s specifications guarantees that the Pareto superior equilibria will prevail. The question as to which equilibrium will be played out cannot be answered within the model. This is where the notion of “focal point” (Schilling 1960: 57) becomes useful to explain which equilibrium will ultimately emerge in the real world.

**Lower Courts and Higher Courts: Acquiescence and Defection in the Judicial Hierarchy**

Supreme courts are supposed to ensure law’s uniform application throughout the judicial system. Given their limited resources, however, they cannot review each and every lower court decision. From the perspective of the judges sitting on the tribunals at the bottom of the judicial heap, this limitation creates potential opportunities to deviate from the high court’s doctrines (judicial drift).

Building on Cameron et al. (2000) and with a view to explain the US Supreme Court’s certiorari practice, Lax (2003) develops a game in which a lower court $L$ decides whether to comply with the doctrine of high court $H$, which in turn must decide whether to review the lower court’s ruling (Figure 7.1).
Figure 7.1. Compliance in the Judicial Hierarchy

Each court has a preferred doctrine – e.g. concerning the degree to which evidence obtained by coercive means should be admitted in criminal proceedings. Unlike other scholars who prefer to model judicial preferences using cut-points, Lax uses indifference points. A court wants cases on the left side of its indifference point to be adjudicated one way (e.g. evidence should be admitted) and cases on its right side another (e.g. evidence should be excluded) and is indifferent for cases on its indifference point. More importantly though, the loss a court incurs when a case is “incorrectly decided” from the viewpoint of its doctrinal preferences is not a constant (as in Cameron et al. 2000 and Cameron and Kornhauser 2010). Rather, it is commensurate to the distance between the case and the court’s indifference point. This specification captures the idea that some cases are more compelling than others. Cases that are far away from the indifference point
are more painful to lose and more important to win.\textsuperscript{16} That cases are adjudicated according their policy preferences is not the courts’ sole concern, however. The lower court $L$ faces a reversal cost $c > 0$ when the high court reviews its decision and reverses it, while the high court $H$ faces an auditing cost $k > 0$.

![Diagram](image)

\textbf{Figure 7.2. Certworthy and Uncertworthy Cases}

In Figure 7.2 the courts’ indifference points are respectively denoted as $x_L$ (lower court) and $x_H$ (high court).\textsuperscript{17} To take again the example of admissible

\begin{align*}
\text{CR} & \quad 0 \quad \text{UR} \quad 1 \\
& x_L \quad x_H - k \quad x_H \quad x_H + k
\end{align*}

\textsuperscript{16} Note that the model does not systematically distinguish between policy and case disposition in the way Cameron and Kornhauser (2010)’s does. More to the point: it does not have separate payoffs for policy (preferred rule) and dispositional outcomes. In fact, what matters for the courts is that the case disposition corresponds to what their preferred doctrine requires irrespective of the doctrine actually invoked to justify that disposition. In the example depicted in Figure 7.2, for $x < x_L$, the doctrines respectively favoured by $L$ and $H$, different though they are ($x_L \neq x_H$), nonetheless entail the same case disposition (= admit evidence in the search and seizure example). Now this means that from both $H$ and $L$’s perspective whether that outcome is justified by one doctrine or the other is indifferent.

\textsuperscript{17} Assuming $x_L < x_H$ (which can be readily interpreted as meaning that $L$ is more liberal than $H$), $L$’s utility function can be defined as:
evidence in criminal proceedings (a favourite in the literature), the case space can be viewed as representing the degree of coerciveness by which evidence has been obtained, from outright torture (case \( x = 1 \)) to voluntary relinquishment (case \( x = 0 \)). Both courts would exclude evidence to the right of their indifference point and admit evidence to their left. Yet \( H \) is more on the side of prosecution and \( L \) more on the side of the accused. Therefore, all cases falling between \( x_L \) and \( x_H \) are contentious: \( L \) wants to exclude them while \( H \) wants to admit them. The basic version of the model assumes perfect information. The courts know each others’ preferences. Crucially, the

\[
\begin{align*}
\epsilon

\begin{cases}
  x \leq x_L, l = c, h = g & \text{or } \neg g \\
  x < x < x_H, l = \neg c, h = \neg g \\
  x \geq x_H, l = c, h = g & \text{or } \neg g \\

  x \leq x_L, l = \neg c, h = \neg g \\
  x > x_H, l = c, h = g & \text{or } \neg g \\
  x > x_H, l = \neg c, h = \neg g \\
  -\epsilon & \text{if } x_L < x < x_H, l = \neg c, h = g
\end{cases}
\end{align*}
\]

For \( H \):

\[
\begin{align*}
\epsilon

\begin{cases}
  x \leq x_H, l = c, h = \neg g \\
  x > x_H, l = c, h = \neg g \\
  x = c & \text{or } \neg c, h = g \\
  x < x_H, l = \neg c, h = \neg g \\
  x > x_H, l = \neg c, h = \neg g
\end{cases}
\end{align*}
\]

Where \( h \) and \( l \) stand for the behaviour of respectively the higher and lower court; \( c \) stands for the lower court’s decision to comply, \( \neg c \) its decision not to comply, while \( g \) stands for the higher court’s decision to grant and \( \neg g \) not to grant certiorari.
lower court knows the size and location of the high court’s Uncertworthy Region (UR). This region corresponds to the set of cases where the gain derived from correcting a contrary decision is less or equal to the cost of auditing. Formally, if \( x_H - k < x < x_H + k \), \( H \) is better off letting the lower court’s decision stand. \( L \) may thus safely deviate from \( H \)'s doctrine when contentious cases fall within UR. On the other hand, when contentious cases fall in the Certworthy Region (CR) \( L \) is better off complying. Consequently, in equilibrium \( H \) never reviews any lower court decision.

Lax shows that procedural rules may work to reduce the size of UR. On the US Supreme Court, the decision to review a case (to grant certiorari) is subject to the famous Rule of Four: it takes the votes of four Justices to get a case heard by the Court. As illustrated in Figure 7.3, a Rule of Four results in a smaller UR than a Rule of Five.

**Figure 7.3. Rule of Four and Rule of Five**

This example suggests that the Rule of Four actually operates to increase the power of the median Justice. Because cases closer to the median are worth reviewing for at least four Justices, lower courts have less latitude to deviate from the median Justice’s preferred doctrine.
In a repeated version of the game and assuming imperfect information, conditions can be identified under which the shortfall in Supreme Court control may be totally eliminated. From Figure 7.3 we can see that complete compliance could be ensured if Justice$_4$ were thought to be slightly more liberal or Justice$_6$ slightly more conservative. In a repeated game – the normal configuration for interactions between higher and lower courts – judges may thus have an incentive to appear more extreme than they really are. Provided the gain from ensuring total compliance in future cases is sufficiently high, the cert-pivots (Justice$_4$ and Justice$_6$ in Figure 7.3) will want to grant certiorari to a case falling within their UR to signal “toughness”. Lax considers a refinement of his basic model in which the high court interacts with lower courts over two periods of play. In the first period, a case $x_1$ is adjudicated by the lower court $L_1$. After that the high court decides whether to grant certiorari and, in the event it does, whether to reverse or to affirm $L_1$’s decision. In the second period, the lower court $L_2$, which has the same indifference point as $L_1$, must decide $n$ cases at $x_2^{18}$ having observed $H$’s behaviour in reaction to $L_1$’s decision in the first period. Finally, $H$ must decide whether to review each of these cases. If certiorari is granted, then the case is decided as per the median Justice’s indifference point. The relevant cert-pivot, i.e. the conservative one for liberal lower courts and the liberal one for conservative lower courts, may be of two types:

\[^{18}\text{Picked from the uniform distribution on the case space } X \text{ (unit interval from 0 to 1). This means that the } n \text{ cases may or may not fall within the conflict region.}\]
either moderate (his/her UR overlaps with the median’s own UR) or extreme (no overlap with median’s UR).

Now, $L_1$’s strategy in the first period and for cases falling in the region of conflict will depend on the initial probability that the relevant cert-pivot is of the moderate or extreme type. A conservative cert-pivot of the extreme type will always grant certiorari to reverse decisions that are more liberal than the median Justice. A conservative cert-pivot of the moderate type, by contrast, would normally not review a liberal decision falling in the conflict region, except if the expected gain from compliance in the second period outweighs the net cost of review in the first.\(^{19}\) This is because granting certiorari may help persuade $L_2$ that the relevant cert-pivot is of the extreme type, which in turn would make compliance more likely in the second period. Going back to the choice faced by $L_1$ in the first period, we realise that $L_1$ must also deal with the risk that the moderate type uses its case for a strategic reversal. Other things being equal, the higher the initial probability that the cert-pivot is of the extreme type, the stronger the incentive for the moderate type to mimic the behaviour of the extreme type. A high initial probability of an extreme type may thus induce $L_1$ to comply pre-emptively.

Of course, pre-emptive compliance by $L_1$ rules out the possibility for the cert-pivot to reverse the decision to signal extremeness. However, in this situation affirmation can be used as an alternative signal. Affirmance is a

\(^{19}\)The argument for a liberal cert-pivot and a conservative lower court decision is, of course, symmetrical.
costlier signal, since, unlike reversal, it does not have the added benefit of
correct a lower court mistake. Paradoxical though it may seem, it
constitutes for that very reason a more convincing signal of extremeness for
the lower courts. If the prospect of greater compliance in the second period
is attractive enough, both the moderate and the extreme type will want to
review and affirm at least some cases of compliance in the first period.

In the second period, $L_2$ updates the probability that the relevant cert-pivot
is of the extreme type after observing his/her behaviour in the first period.\(^\text{20}\)
If non-compliance and a failure to grant certiorari are observed, then the
moderate type is revealed and $L_2$ can safely decline compliance with the
high court’s doctrine in the conflict region. In other circumstances – non-
compliance followed by reversal, compliance followed by affirmance, and
compliance followed by denial of certiorari – the extent of $L_2$’s compliance
varies according to the location of $x_2$ and the initial probability that the cert-
pivot is of the extreme type (Lax 2003: 77-81).

\(^{20}\) $L_2$ does so applying Bayes’ Theorem. Bayes’ Theorem states how to modify the probability
of an hypothesis so as to take into account new evidence:

$$P(H|E) = \frac{P(E|H)P(H)}{P(E)}$$

Where $P(H|E)$ is the posterior probability of the hypothesis $H$ given new evidence
$E; P(E|H)$ is the conditional probability of observing $E$ if $H$ were true; $P(H)$ is the prior
probability that $H$ is true that was inferred before the observation of $E$; and $P(E)$ is the
probability of observing $E$ under all possible hypotheses (marginal probability of $E$).
Conclusion: Beyond the American Judiciary

As can be seen from the foregoing survey of the literature, the application of game theory to judicial institutions has been a largely American affair. The truth is this literature is almost entirely the product of American scholars. And it comes as little surprise that most of models have been developed with a view to explain features of the American judiciary. Rare are those that have been developed with another context in mind. Vanberg (2005) and Carrubba (2005 and 2009) are the exception, not the rule.

Thus much work remains to be done to develop models addressing the peculiarities and idiosyncrasies of judicial institutions outside the United States. In Europe, the rules governing the courts’ internal deliberation process often differ from those found in the United States: in abstract review proceedings constitutional courts can decide policy matters without simultaneously having to adjudicate a specific dispute; the invalidation of governmental acts is sometimes subject to a supermajority requirement; judges do not always have the right to file separate opinions, etc. These differences seriously limit the relevance of models geared to the US judicial process.

Another important feature of judicial behaviour outside the United States is that interactions between courts may take place outside a hierarchical framework. The European Union is a case in point. The European Court of Justice has a complex relationship with the courts operating in its Member States. While ordinary and lower instance courts have displayed an impressive degree of willingness to cooperate with the
ECJ, constitutional courts have proved more recalcitrant. The German Federal Constitutional Court in particular has made repeated threats of defiance. The relationship between this court and the ECJ has many characteristics of a Hawk-Dove game. Each player wants to expand or maintain his turf at the other’s expense. But both face a tremendous cost in the case of outright confrontation, which would bring about a European-scale constitutional crisis (Dyevre 2011). In any case, all these are issues that may benefit from formalization using game theory’s rich insights.
References


