How the Chamber System at the CJEU Undermines the Consistency of the Court's Application of EU Law

Joshua C. Fjelstul*

The University of Geneva

Forthcoming at the Journal of Law and Courts

Abstract

The Court of Justice of the European Union (CJEU) uses a chamber system to more efficiently decide cases. To what extent, and under what conditions, does the CJEU's chamber system undermine the consistency of the Court's application of EU law? This paper contributes to the literature on the internal organization of collegial courts by presenting a computational formal model that predicts (a) that hearing cases in smaller chambers undermines the consistency of the Court's application of EU law and (b) that the magnitude of this effect is larger when judges' preferences are more heterogenous and smaller when plaintiffs strategically bring cases. Based on these findings, I use machine learning and empirical data on CJEU judgments in infringement cases to estimate the degree to which we should expect the chamber system to undermine the consistency of the CJEU's application of EU law in practice.

*The replication materials are available at the Journal's Dataverse archive. I thank Cliff Carrubba, Matt Gabel, Allison Östlund, Olof Larsson, Andreas Hofmann, Dimiter Toshkov, Urška Šadl, Michal Ovadek, Silje Synnøve Lyder Hermansen, Tommaso Pavone, Stein Arne Brekke, the anonymous reviewers, and the editor for their helpful comments and suggestions.

The Court of Justice of the European Union (CJEU) is one of the most active and productive international courts, currently deciding around 1,000 cases per year.[1] To reach this level of productivity, the CJEU has developed a complex chamber system in which most cases are decided in small chambers of 3 or 5 judges. There is very little theoretical or empirical research on the the internal organization of the Court (Kelemen 2012; Malecki 2012; Carrubba and Fjelstul 2021), but the chamber system plays a central role in the day-to-day functioning of one of the world's most influential courts.

While the chamber system is critical to the productivity of the CJEU, it may also have drawbacks. Hearing cases in small chambers creates the potential for inconsistency in the CJEU's application of European Union (EU) law, as smaller chambers are less representative of the whole Court and could choose to decide similar cases in different ways. The literature on collegial courts has examined how the structure of judicial bargaining can lead to inconsistency in how collegial courts decide similar cases (e.g., Kornhauser and Sager 1986; Kornhauser 1992; Landa and Lax 2009; Lax 2011), but this literature has not explored how the internal organization of collegial courts affects the consistency of their application of the law, leaving a gap in our theoretical understanding of how the institutional design of courts affects judicial decision-making.

To what extent, and under what conditions, does the CJEU's chamber system undermine the consistency of the Court's application of EU law? This paper contributes to the literature on the internal organization of collegial courts by presenting a computational formal model that predicts (a) that hearing cases in smaller chambers undermines the consistency of the Court's application of EU law and (b) that the magnitude of this effect is larger when judges' preferences are more heterogenous and smaller when plaintiffs strategically bring cases. Based on these findings, I use machine learning and empirical data on CJEU judgments to estimate the degree to which we should expect the chamber system to undermine the consistency of the CJEU's application of EU law in practice.

[1] Unlike many domestic courts, like the Supreme Court of the United States, the CJEU does not have docket control and has to adjudicate every case that it receives. If a case is inadmissible, the CJEU still has to issue an order dismissing the case.

Beyond the theoretical motivations of better understanding how the internal organization of collegial courts affects judicial decision-making, this question has normative implications. Suppose you are a litigant at the CJEU. If you lose your case, but see a similar litigant winning a similar case, and the only real difference was that the other case was heard by a different set of judges, you might feel aggrieved and see the judicial process as arbitrary or capricious. From a litigant's perspective, we could frame the question as: How much do your chances of winning a case depend on which judges you get?

Before I go any further, what do we mean by *the consistency of a court's application of the law*? The phrase *application of the law* refers to the process whereby a court uses legal rules to decide the disposition of a case based on the facts. The *disposition* refers to which party (i.e., the plaintiff or the defendant) wins. Judicial scholars use the term *consistency* in a variety of ways (Lax 2011). In the context of looking at the CJEU's chamber system, I conceptualize *consistency* as the degree to which the disposition of a case would change if it were decided by a different chamber of judges. Based on this definition, the CJEU's application of EU law is *consistent* only if the same party would win regardless of the composition of the chamber. As such, *consistency*, in this context, is fundamentally about the counterfactual — about comparing the dispositional rulings of counterfactual chambers in the same case. Even more precisely, we can conceptualize *consistency* as the variance of the distribution of the expected probability that the plaintiff (or defendant) wins a case across counterfactual chambers.

This paper proceeds as follows. First, I develop a computational formal model of judicial decision-making, based on a case-space model (Lax 2011), that captures the major institutional features of the CJEU, including the chamber system. Consistent with my conceptualization of *consistency*, the endogenous parameter of interest (i.e., the dependent variable of the theory) is the variance of the distribution of the probability that the plaintiff wins a case across counterfactual chambers, and the exogenous parameter of interest (i.e., the independent variable of the theory) is chamber size. I generate theoretical predictions by computationally deriving comparative statics.

Second, I conduct an empirical analysis to estimate the extent to which the chamber system undermines the consistency of the CJEU's application of EU law in practice, given the actual preferences of CJEU judges. In other words, I want to estimate how much your chances of winning your case depends on which judges you get. First, I use machine learning (random forests) to model the probability that the Commission wins infringement cases, which are cases in which the Commission sues an EU member state over an alleged violation of EU law, as a function of which judges are in the chamber. Infringement cases are a hard test because the Commission strategically brings cases it expects to win (Fjelstul and Carrubba 2018) — a condition under which my formal model predicts that the impact of the Court's chamber system on consistency will be mitigated. Second, I use the trained model to generate predicted probabilities for simulated counterfactual chambers. Third, I calculate the variance of the distribution of the predicted probabilities, which, connecting back to my research question, characterizes the extent to which the CJEU's chamber system undermines the consistency of the Court's application of EU law.

I find that the probability that the Commission wins varies dramatically across counterfactual chambers, suggesting that the chamber system meaningfully undermines the consistency of the CJEU's application of EU law. I conclude by highlighting the implications of my findings for other collegial courts that use chambers and for the contemporary policy debate about reforming the Supreme Court.

The Chamber System at the CJEU

The Court of Justice of the European Union (CJEU) consists of two constituent courts, the Court of Justice and the General Court. Both courts have a chamber system. The rules of the chamber system have evolved over time, but currently, nearly any case can be heard by a chamber. At the Court of Justice, there are currently 10 standing chambers that hear cases in panels of 3 or 5 judges. Judges are attached to one or more chambers and those assignments rotate regularly (Rules of Procedure, Article 11).[2] There is also a Grand Chamber, currently with 15 judges, that the Court generally reserves for important cases (Kelemen 2012). At the General Court, there are 8 standing chambers, each with 3 or 5 judges. The General Court used to use single judges in some cases. The General Court has only used a Grand Chamber for a handful of cases.

Figure 1 shows how the CJEU's use of chambers has evolved over time. At the Court of Justice, 5-judge chambers are now far more common than 3-judge chambers. The Full Court used to be very common but has now been completely replaced by 5-judge chambers (for ordinary cases) and the Grand Chamber (for important cases). At the General Court, 3-judge chambers are standard.

When a case is filed at the Court, the President of the Court (who is elected by all of the judges) appoints a Judge-Rapporteur (Rules of Procedure, Article 15), who is the judge that writes the judgment. After the written stage (e.g., the filing of briefs), the Judge-Rapporteur presents a preliminary report to the Plenum (i.e., all of the judges) at a general meeting of the Court. At this time, the Plenum decides (by majority vote) which chamber should hear the case (Rules of Procedure, Article 25). If the Plenum assigns the case to a 3-judge or 5-judge chamber, the chamber that hears the case must be one of the standing chambers that the Judge-Rapporteur is currently assigned to (Rules of Procedure, Article 28). Thus, by assigning a Judge-Rapporteur, the President has indirect influence over which judges hear the case. Judges often substitute for each other due to absences and scheduling conflicts, so there are many more chamber configurations in practice than formal chamber assignments would imply.

After oral arguments (if there is a hearing), the Judge-Rapporteur prepares a draft judgment and the chamber meets to deliberate. The chamber votes on the disposition using a simple majority decision-rule. At the CJEU, unlike at the Supreme Court, there is no majority opinion and concurrences and dissents are not permitted. All judges sign the

[2] There can be more judges attached to a standing chamber than actually hear the case (e.g., there could be 6 or 7 judges attached to a 5-judge chamber), in which case the judges rotate from case to case.

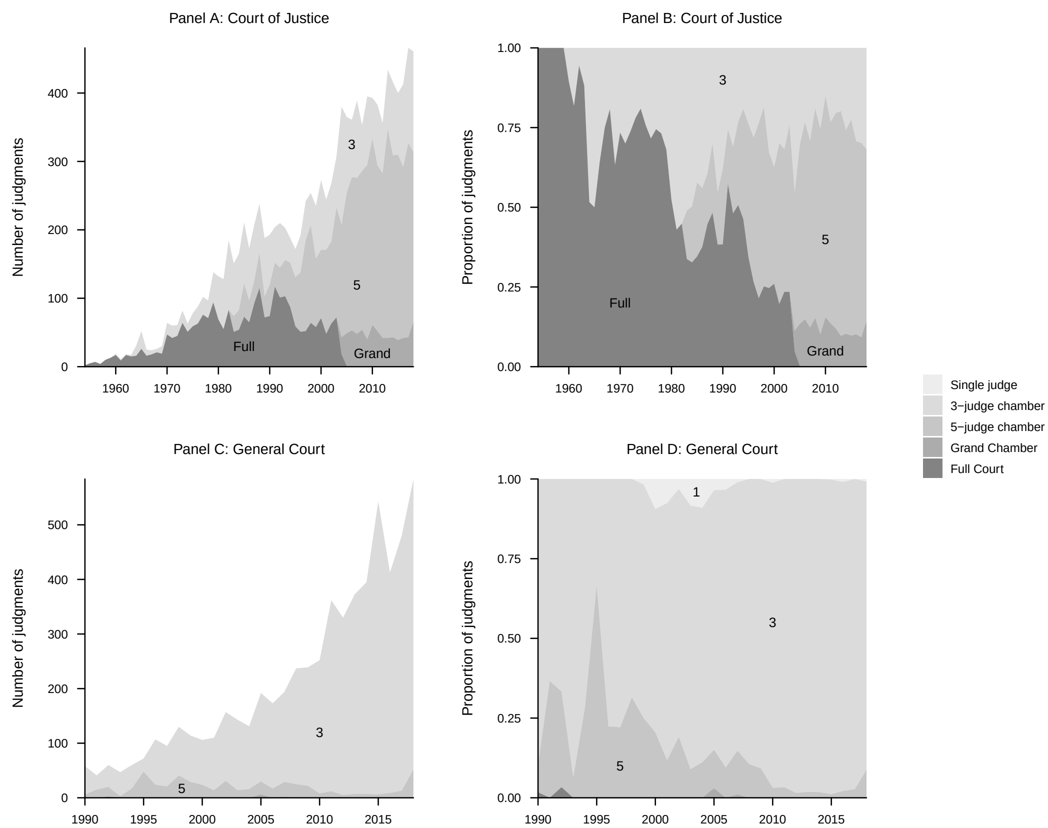

Figure 1. Chamber Sizes at the CJEU

Note: This figure shows how the use of formations has evolved over time at the CJEU. Panels A and B show the Court of Justice and Panels C and D show the General Court.

final judgement, whether they agree with it or not, and no information about voting is ever made public. Even retired judges do not comment on voting.

There is little research on the internal organization of the CJEU, much less the chamber system. The literature has generally focused on the external politics of the Court, including how threats of noncompliance with CJEU rulings (e.g., Carrubba, Gabel and Hankla 2012; Carrubba and Gabel 2015) and threats of judicial override by legislative actors (e.g., Larsson and Naurin 2016; Larsson et al. 2017) influence the Court's decision-making. Scholars are only just beginning to study the internal organization of the CJEU. For example, Frankenreiter (2018) looks at the role of Advocates General, who provide non-binding independent legal opinions to the Court. Cheruvu (2019) looks at how the Court's French-language mandate affects the efficiency of judicial decision-making. Hermansen (2020) finds that the President is more likely to assign Judge-Rapporteurs whose governments are moderate when case law is not yet well-developed in order to build legitimacy. While scholars are certainly interested in how the internal organization of the Court affects judicial decision-making, a lack of empirical data on case dispositions and chamber composition has deterred research (Carrubba and Fjelstul 2021).

The broader literature on collegial courts has explored how the structure of judicial bargaining at collegial courts (i.e., how preferences are aggregated) can affect the consistency of their application of the law across similar cases (e.g., Kornhauser and Sager 1986; Kornhauser 1992; Collins 2008; Landa and Lax 2009; Lax 2011). Several studies point to the imprecision of doctrine as a source of inconsistency (e.g., Staton and Vanberg 2008; Lax 2012; Fox and Vanberg 2014; Clark 2016). However, this literature often assumes that judicial decision-making occurs under institutional rules that are specific to the Supreme Court. Consequently, the literature has not explored how the internal organization of collegial courts — how institutions like the CJEU's chamber system — affects the consistency of their application of the law, leaving a gap in our understanding of how the institutional design of courts affects judicial decision-making.

Theory

In this section, I present a computational model of a collegial court with a chamber system, modeled on the CJEU. I use the model to generate theoretical expectations about how a decrease in chamber size affects the consistency of a court's application of the law (Siegel 2018). My computational model is similar to an analytical game-theoretical model. There is a steady-state, analogous to a game-theoretic equilibrium. As such, there are exogenous parameters (i.e., independent variables) and endogenous parameters (i.e., dependent variables). I derive comparative statics using a computational simulation to rigorously develop theoretical predictions (e.g., hypotheses).

In an analytical game-theoretic model, a comparative static is the derivative of an endogenous parameter with respect to an exogenous parameter, and the sign of the derivative is the hypothesized relationship between the two variables. In my computational model, a comparative static is still a change in an endogenous parameter with respect to an exogenous parameter, but I determine the sign using a computational simulation. To be clear, even though I use computational simulations to derive the comparative statics, this is still a purely theoretical exercise, not an empirical one (Siegel 2018). The objective is to rigorously derive theoretical predictions, rather than to test theory.

In my model, a collegial court hears a set of cases, and a set of counterfactual chambers decides the disposition in each case. For each counterfactual chamber, there is a probability that the plaintiff wins the case. In my analysis of the model's comparative statics, I derive how chamber size, which is the independent variable (i.e., an exogenous parameter), affects the variance of the distribution of the probability that the plaintiff wins across counterfactual chambers, which is the dependent variable (i.e., an endogenous parameter). Consistent with my conceptualization of *consistency*, this variance captures the consistency of the court's application of the law across counterfactual chambers. Our intuition should be that hearing cases in smaller chambers will increase the variance of this distribution, as smaller panels will be less representative of the whole court and different chambers could prefer to decide the same case in different ways. I confirm this intuition and then derive the conditions under which this effect will be larger or smaller.

My computational model is based on a case-space model, which is a widely-used class of models in the literature on collegial courts (e.g., Lax 2007; Lax and Cameron 2007; Landa and Lax 2008; Carrubba et al. 2012; Clark 2016; Ainsley, Carrubba and Vanberg 2021). In a case-space model, the facts of cases and the preferences of judges exist in a uni-dimensional case-space. The preferences of judges are represented as a legal rule, which is a cut-point that partitions the case space (Lax 2011). The position of the facts (a single point in the case-space) relative to a judge's cut-point determines how the judge votes, in terms of the disposition. If the facts of the case are located to one side of the cut-point, the judge will prefer to vote in favor of plaintiff, and if the facts are located to the other side of the cut-point, the judge will prefer to vote in favor of the defendant.

To make my discussion of the model more concrete, I use the running example of CJEU infringement cases, which will also be the focus of my empirical analysis. Infringement cases are cases at the Court of Justice in which the Commission (the plaintiff) sues an EU member state (the defendant) over an alleged violation of EU law under Article 258 of the Treaty on the Functioning of the European Union (TFEU). Infringement cases always have a clear disposition: either the Commission wins or the member state wins.

Computational Model

The Court is composed of $N_j = 27$ judges, which is the current number of judges at the Court of Justice. Let $J = \{1, 2, \ldots, N_j\}$ be the set of all judges. The Court hears a set of $N_c = 1000$ infringement cases. Let $C = \{1, 2, \ldots, N_c\}$ be the set of all cases. Each infringement case is heard by $N_i = 1000$ counterfactual chambers. Let $J_{ci} \subset J$ be the subset of judges in the i-th counterfactual chamber that hears case c. I randomly select each judge in J_{ci} from the set of all judges J without replacement.[3] Each counterfactual chamber J_{ci} is composed of s judges. The size of the chamber (s) is the exogenous parameter of interest in the model (i.e., the independent variable of the theory).

Each counterfactual chamber i hearing case c issues a dispositional ruling, $d_{ci} \in \{0, 1\}$. Let $d_{ci} = 1$ indicate that the chamber rules in favor of the plaintiff (i.e., the Commission) and let $d_{ci} = 0$ indicate that the chamber rules in favor of the defendant (i.e., the member state). Let $p_{ci} \equiv \Pr(d_{ci} = 1)$ be the probability that the Commission wins case c in counterfactual chamber i. As I discuss in a moment, this probability depends on how each judge chooses to vote. Across the $N_i = 1000$ counterfactual chambers that hear case c, this probability has a distribution. Let $V_c \equiv \mathrm{Var}(p_{ci})$ be the variance of that distribution. The

[3] I model the assignment of judges to cases as a random process, but this assumption does not drive the predictions of the model. If the Court strategically assigns judges to cases based on the facts of the case (i.e., if the probability of a judge being selected depends on the facts of the case), that would induce a correlation between the facts of the cases and the preferences of judges, but the model's predictions are the same.

average variance of that distribution across all $N_c = 1000$ cases, $\bar{V} = \frac{1}{N_c}\sum_{c=1}^{N_c} V_c$, is the endogenous parameter of interest (i.e., the dependent variable of the theory). It captures the consistency of the court's application of the law.

In the next section, I will calculate comparative statics computationally to determine how a decrease in chamber size affects this variance. I consider a variety of contemporary and historical chamber sizes used at the CJEU. I simulate a 1-judge chamber as a lower baseline. The Court of Justice has never used 1-judge chambers, but the General Court has. I also simulate a 27-judge chamber, which is the size of the full Court, as an upper baseline. The primary chamber sizes of interest are 3-judge and 5-judge chambers. These are by far the most common chamber sizes at the Court (see Figure 1). The Court of Justice primarily uses 5-judge chambers, but 3-judge chambers are also common. The General Court sometimes uses 5-judge chambers, but 3-judge chambers are standard. I simulate a 7-judge chamber to capture the Small Plenum, which used to be used by the Court of Justice. I simulate a 15-judge chamber to capture the Grand Chamber, which is used by the Court of Justice.[4]

Each case has facts f_c and each judge on the Court has a preferred legal rule θ_{cj} that captures how the judge would prefer to decide the case based on the facts. The legal rule is a cut-point in the case-space. The cut-point partitions the case-space and indicates the "correct" disposition, given the facts. If the facts fall to the right of the cut-point ($f_c > \theta_{cj}$), the "correct" disposition, according to the legal rule, is that the plaintiff wins. Otherwise ($f_c < \theta_{cj}$), the "correct" disposition is that the defendant wins.

Lower values of f_c indicate facts that favor the member state and higher values of f_c indicate facts that favor the Commission. In contrast, lower values of θ_{cj} indicate a judge who favors the Commission and higher values of θ_{cj} indicate a judge who favors the member state. If f_c is high, it is more likely that the facts will be located to the right of the judge's cut-point, $f_c > \theta_{cj}$, meaning that judge j will be more likely to agree with the Commission,

[4] The Grand Chamber has a quorum, and the quorum has changed over time, so empirically, we can see 9-judge, 11-judge, and 13-judge chambers.

and *vice versa*. Similarly, a judge with a low θ_{cj} has a predisposition to favor the Commission because it is more likely that the facts will be to the right of the judge's cut-point, $f_c > \theta_{cj}$, and *vice versa*.

I model the case facts and the judge cut-points as normally distributed random variables, $f_c \sim \text{Normal}(\mu_{f_c}, \sigma^2_{f_c})$ and $\theta_{cj} \sim \text{Normal}(\mu_{\theta_{cj}}, \sigma^2_{\theta_{cj}})$. I draw new cut-points for all judges on the court for each case, but the cut-points of all judges are always constant across counterfactual panels for the same case. If the mean of the cut-point distribution ($\mu_{\theta_{cj}}$) is small relative to the mean of the case facts distribution (μ_{f_c}), then judges will tend to favor the Commission, as case facts are more likely to be to the right of the judges' cut-points. Conversely, if the mean of the cut-points is large relative to the mean of the case facts, then judges will be more likely to rule for the member state, as the case facts are less likely to be to the right of the judges' cut-points.

The variance of the cut-point distribution captures the heterogeneity of judges' preferences. When the variance is small, judges are more homogeneous in their preferences, and when the variance is large, they are more heterogeneous. In the next section, I will derive comparative statics to determine how an increase in the heterogeneity of judges' preferences (i.e., an increase in $\sigma_{\theta_{cj}}$) conditions the effect of a decrease in chamber size (s) on the average variance of the probability that the Commission wins ($\bar{V}$). Intuitively, we should expect the effect to be mitigated as judges become more homogenous — it should not matter which judges are in the chamber if judges all apply the same legal rule.

At the CJEU, the facts of infringement cases tend to favor the Commission due to the strategic behavior of the Commission in bringing cases. Recent research finds that the CJEU is strategic and anticipates the likelihood that losing parties will ignore its rulings if the costs of compliance are too high (Carrubba and Gabel 2015) and the likelihood of that the EU's legislative institutions will override its rulings by amending EU law (Larsson and Naurin 2016; Fjelstul 2019). This induces the Commission to be strategic in the cases that it chooses to bring — only bringing cases with sufficiently pro-Commission facts (König and Mäder 2014; Fjelstul and Carrubba 2018; Fjelstul 2019).

This suggests that, for infringement cases, the mean of the case facts distribution should be located to the right of the mean of the cut-point distribution. In the next section, I will also derive comparative statics to determine how an increase in the predisposition of the facts to favor the plaintiff (i.e., an increase in μ_{f_c} relative to $\mu_{\theta_{cj}}$) conditions the effect of a decrease in chamber size (s) on the average variance of the probability that the Commission wins ($\bar{V}$). In this case, it is less clear what our intuition should be. However, if the facts of a case strongly favor the Commission, we might expect that the composition of the chamber will not matter as much as it would if a case were a close call.

Judges decide how to vote based on the facts. (Note that only judges in the counterfactual chamber, $j \in J_{ci}$, participate in voting.) This is the only choice variable in the model. Let $v_{cj} \in \{0, 1\}$ be the vote of judge j in case c, where $v_{cj} = 1$ indicates a vote for the Commission and $v_{cj} = 0$ indicates a vote for the member state. I assume that the steady-state behavior of judges is that they vote on the disposition according to their preferred legal rule, which indicates the "correct" disposition based on the facts. In other words, I assume that judges vote sincerely on the disposition. How a judge votes is therefore not influenced by which other judges are in the chamber.

This way of modeling voting on the disposition makes sense given the institutional rules of the CJEU. At the Supreme Court, where judges have to decide whether or not to join a majority opinion, judges could join majority opinions strategically (i.e., their decision to join the majority opinion could be insincere). Judges might even trade their vote on the disposition for a better legal rule, as only a majority opinion can establish precedent (e.g., Lax 2011; Carrubba et al. 2012). In that institutional setting, voting on the disposition could be strategic. But at the civil law CJEU, there are no majority opinions. All judges sign the judgment, and votes are never made public.

I assume that judges sometimes misinterpret the facts of the case when deciding how to vote. Let $\epsilon \sim \phi(\mu_\epsilon, \sigma_\epsilon^2)$ be a normally distributed error term. If the facts of the case, as interpreted by the judge, are to the right of a judge's cut-point ($f_c + \epsilon > \theta_{cj}$), the judge will vote for the Commission ($v_{cj} = 1$). Otherwise ($f_c + \epsilon < \theta_{cj}$), the judge will vote for the member state ($v_{cj} = 0$). The implication is that voting is probabilistic. The probability that a judge votes for the Commission is $p_{cj} = \Pr(\epsilon < f_c - \theta_{cj})$, which is equivalent to the cumulative density function (CDF) of the normal distribution evaluated at $f_c - \theta_{cj}$. We can think of these votes as Bernoulli distributed random variables, $v_{cj} \sim \text{Bernoulli}(p_{cj})$.

Next, I aggregate up from the votes of individual judges to the disposition of the court. The CJEU decides the disposition of a case using a simple majority decision-rule. Therefore, the probability that the Court sides with the Commission is the probability that a majority of the judges in the chamber vote for the Commission. Let $m_{ci} = \sum_{j \in J_{ci}} v_{cj}$ be the number of judges in counterfactual chamber i hearing case c who vote for the plaintiff. The threshold for a simple majority is $m^* = \frac{s+1}{2}$. Thus, the chamber will side with the Commission when $m_{ci} \geq m^*$ judges vote for the Commission.

The probability that a judge votes for the Commission (p_{cj}) varies by judge, so votes are non-identically distributed Bernoulli random variables. The number of judges that vote for the Commission (m_{ci}) is therefore the sum of non-identically Bernoulli distributed Bernoulli trials, making it a discrete random variable. The discrete probability distribution for a sum of non-identically distributed Bernoulli trials is the Poisson binomial distribution. Thus, the number of judges that vote for the Commission (m_{ci}) is a discrete random variable that follows a Poisson binomial distribution.

The probability that the Commission wins case c in counterfactual chamber i is the probability that the number of judges in chamber i who vote for the Commission reaches the threshold for a simple majority, $p_{ci} = \Pr(m_{ci} \geq m^*)$, which we can rewrite as $p_{ci} = 1 - \Pr(m_{ci} \leq m^* - 1)$. Let $G(\cdot)$ be the CDF of the Poisson binomial distribution. Then, the probability that the Commission wins is $p_{ci} = 1 - G(m^* - 1)$.[5] The disposition of the chamber's ruling, $d_{ci} = \{0, 1\}$, is a Bernoulli random variable, $d_{ci} \sim \text{Bernoulli}(p_{ci})$. As noted earlier, the probability that the Commission wins varies across counterfactual

[5] Note that the Poisson binomial distribution also depends on the probability that each judge in the chamber votes in favor of the plaintiff, p_{cj} for all $j \in J_{ci}$.

chambers, and the average variance of that distribution across all of the cases ($\bar{V}$) captures the consistency of the Court's application of the law.

Computational Comparative Statics

Next, I derive comparative statics conditional on the steady-state voting behavior of judges. First, I derive the effect of chamber size (s) on the average variance of the probability that the plaintiff (i.e., the Commission) wins across counterfactual chambers ($\bar{V}$). Then, I derive how the heterogeneity of judges' preferences ($\sigma_{\theta_{cj}}$) and the predisposition of case facts to favor the plaintiff (μ_{f_c} relative to $\mu_{\theta_{cj}}$) condition this effect. Note that the numerical values are completely arbitrary. What is important is the sign of the effect. We can also compare the relative magnitude of effects.

To calculate the effect of a decrease in chamber size on the average variance of the probability that the plaintiff wins, I simulate the model multiple times (each run has $N_c \times N_i$ simulated dispositions) and vary the size of the chamber. I calculate the average variance in each simulation, $\bar{V}(s)$. Then, I calculate how decreasing chamber size affects the average variance. For example, let $\bar{V}(s = 5)$ be the average variance of the probability that the plaintiff wins for 5-judge chambers, and let $\bar{V}(s = 3)$ be the average variance for 3-judge chambers. The effect of a decrease in chamber size from 5 judges to 3 judges on the average variance of the probability that the plaintiff wins is $\bar{V}(s = 3) - \bar{V}(s = 5)$.

Figure 2 visualizes the theoretical predictions of the model. Panel A shows the average variance (the y-axis) conditional on chamber size (the x-axis) according to the simulation. Panel B shows the effect of a decrease in chamber size (the y-axis) for several different decreases in chamber size (the x-axis). The model predicts that the average variance increases as chamber size decreases. The variance is highest when a single judge hears the case and decreases at a decelerating rate as the size of the chamber increases. This is an intuitive result. This is essentially an application of the law of large numbers to internal organization of the Court: the variance is higher when the sample size (i.e., the size of the chamber) is smaller. As shown in Figure 2, the effect of decreasing chamber size becomes larger when

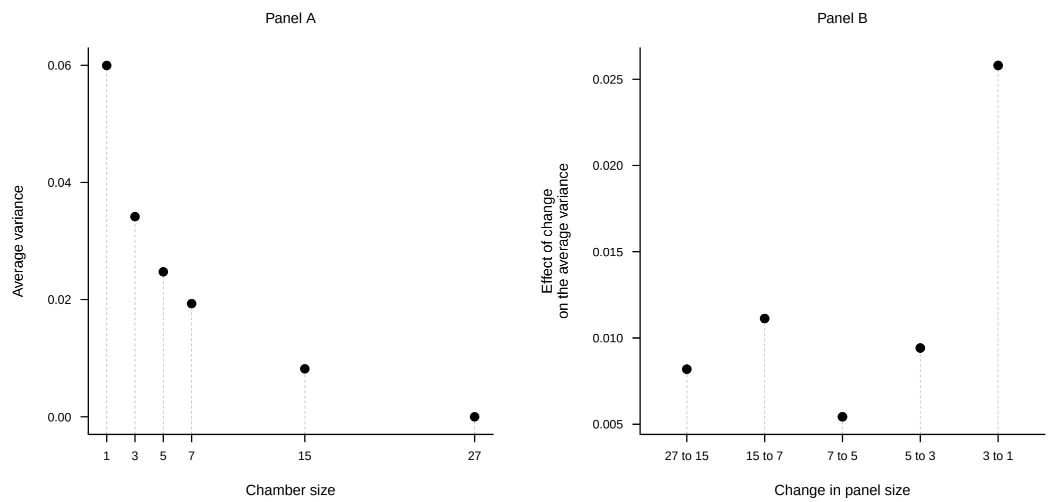

Figure 2. Comparative statics: The effect of chamber size

Note: This figure shows the predicted effect of chamber size (s) on the average variance of the probability that the plaintiff wins ($\bar{V}$). Panel A shows the average variance for various chamber sizes. Panel B shows the sign and relative magnitude of various changes in chamber size.

the starting chamber is smaller. The positive effect of decreasing chamber size by two judges (from 7 to 5, from 5 to 3, and from 3 to 1) on the variance increases non-linearly as the starting chamber size decreases.

Result 1. *Hearing cases in smaller chambers leads to more inconsistency in the application of the law.* A decrease in chamber size increases the variance of the probability that the plaintiff wins across counterfactual chambers.

The substantive interpretation of Result 1 is that a court's application of the law is more inconsistent when chambers are smaller. This is because the probability that the plaintiff wins across counterfactual chambers has a higher variance when chamber size is smaller. In other words, when chamber size is smaller, it matters more which judges happen to get assigned to each case. This result has important implications for the functioning of the CJEU. Many observers believe that the CJEU tends to reserve the Grand Chamber for salient, politically sensitive cases (e.g., Kelemen 2012). The model suggests that this practice ensures that the most important rulings are made by more representative chambers that are more likely to reflect the preferences of the majority of judges.

We can use these predictions to speculate about the relative impact of historical changes to the CJEU's use of chambers. For example, the Court eliminated the 7-judge Small Plenum in the early 1990s, as the number of judges expanded due to the enlargement of the EU. Cases that used to be heard by the Small Plenum are now heard either by a 5-judge chamber or the 15-judge Grand Chamber. Using 5-judge chambers instead of a 7-judge chamber is unlikely to have significantly worsened the consistency of the CJEU's application of EU law (going from 7 judges to 5 judges has the smallest effect in the simulation). The most important of the cases that would have been heard by the Small Plenum would now go to the 15-judge Grand Chamber, likely leading to a significant improvement in consistency (going from 7 judges to 15 judges has the second largest effect).

The model also suggests that the trend of hearing politically important cases in the Grand Chamber instead of the Full Court has likely had less of an impact on consistency than the Court's recent push to hear more cases in 3-judge chambers instead of 5-judge chambers. Going from the Full Court (27 judges) to the the Grand Chamber (15 judges) has the second smallest effect. Going from a 5-judge chamber to a 3-judge chamber actually has a larger effect. Moreover, a 5-judge chamber has a much higher starting variance than the Full Court. Thus, we should expect the CJEU's push to hear more cases in 3-judge chambers instead of 5-judge chambers to have a meaningful impact on consistency of the Court's application of EU law.

Next, I consider how other exogenous parameters in the model condition the effect of a decrease in chamber size on the average variance of the probability that the plaintiff wins. I start by calculating the effect of an increase in the heterogeneity of judges' preferences on the effect of a decrease in chamber size. In the model, the heterogeneity of judges' preferences is captured by the standard deviation of the cut-point distribution ($\sigma_{\theta_{cj}}$). I increase the standard deviation from $\sigma_{\theta_{cj}} = 1$ (low heterogeneity) to $\sigma_{\theta_{cj}} = 2$ (high heterogeneity) and re-run the simulation. Conditional on low heterogeneity, the effect of a decrease in chamber

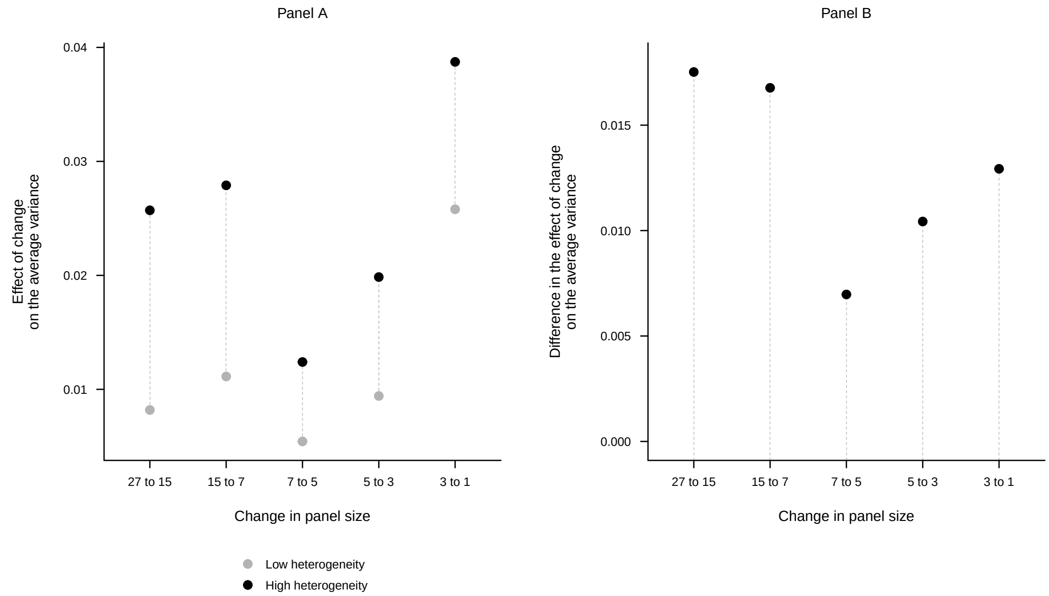

Figure 3. Comparative statics: The heterogeneity of judges' preferences

Note: This figure shows the predicted effect of chamber size (s) on the average variance of the probability that the plaintiff wins ($\bar{V}$), conditional on the heterogeneity of judges' preferences ($\sigma_{\theta_{cj}}$). Panel A shows the effect of various changes in chamber size when heterogeneity is low ($\sigma_{\theta_{cj}} = 1$) and high ($\sigma_{\theta_{cj}} = 2$). Panel B shows the change in each effect caused by increasing heterogeneity.

size from 5 to 3 judges is $\bar{V}(s = 3 \mid \sigma_{\theta_{cj}} = 1) - \bar{V}(s = 5 \mid \sigma_{\theta_{cj}} = 1)$, just as before. Conditional on high heterogeneity, the effect is $\bar{V}(s = 3 \mid \sigma_{\theta_{cj}} = 2) - \bar{V}(s = 5 \mid \sigma_{\theta_{cj}} = 2)$. The effect of an increase in heterogeneity on the effect of a decrease in chamber size is given by a difference-in-differences: $\big(\bar{V}(s = 3 \mid \sigma_{\theta_{cj}} = 2) - \bar{V}(s = 5 \mid \sigma_{\theta_{cj}} = 2)\big) - \big(\bar{V}(s = 3 \mid \sigma_{\theta_{cj}} = 1) - \bar{V}(s = 5 \mid \sigma_{\theta_{cj}} = 1)\big)$.

Figure 3 visualizes the theoretical predictions of the model. Panel A shows the effect on the average variance (the y-axis) of various changes in chamber size (the x-axis), conditional on less heterogeneous and more heterogeneous preferences (the color of the points). Panel B shows the change in the effect of decreasing chamber size (i.e., the vertical distance between the points in Panel A) caused by increasing the heterogeneity of judges' preferences (the y-axis) for various decreases in chamber size (the x-axis). The model predicts that decreasing chamber size increases the variance of the probability that the Commission wins to a greater degree when judges' preferences are more heterogeneous. The intuition is that, if all judges vote based on the same legal rule (i.e., the all have the same preferences), they should be interchangeable, and the dispositional outcome of a case should tend to be the same no matter which judges are on the panel.

Result 2. *Chamber systems matter more when the preferences of judges are more heterogeneous.* Increasing the heterogeneity of judges' preferences increases the positive effect of a decrease in chamber size on the variance of the probability that the plaintiff wins across counterfactual panels.

The substantive interpretation of Result 2 is we should only expect the CJEU's chamber system to undermine the consistency of the CJEU's application of EU law if CJEU judges have sufficiently different preferences. This is an important result because it points to what we need to investigate empirically in order to understand the real-world impact of the chamber system: we need to be able to characterize the heterogeneity of CJEU judges' preferences. However, as we will see, this is challenging given that the CJEU does not publish judges' votes. If all CJEU judges have similar preferences, there is no real danger of the chamber system undermining the Court's application of EU law. However, if different judges have markedly different preferences, we should be more concerned about the consistency of the Court's application of EU law.

Finally, I consider how the predisposition of case facts to favor the plaintiff over the defendant, due to plaintiffs strategically bringing cases they expect to win, conditions the effect of a decrease in chamber size. In the model, the mean of the case fact distribution (μ_{f_c}) relative to the mean of the cut-point distribution ($\mu_{\theta_{cj}}$) captures the predisposition of the case facts favor the plaintiff over the defendant. To calculate the effect of an increase in the predisposition of case facts to favor the plaintiff, I increase the mean of the case facts distribution from $\mu_{f_c} = 0$ (neutral case facts) to $\mu_{f_c} = 1$ (pro-plaintiff case facts) and re-run the simulation. (Equivalently, I could decrease the mean of the cut-point distribution.) This makes it more likely that the case facts will be to the right of the judges' cut-points, increasing the probability that the Commission will win. The effect of an increase in the

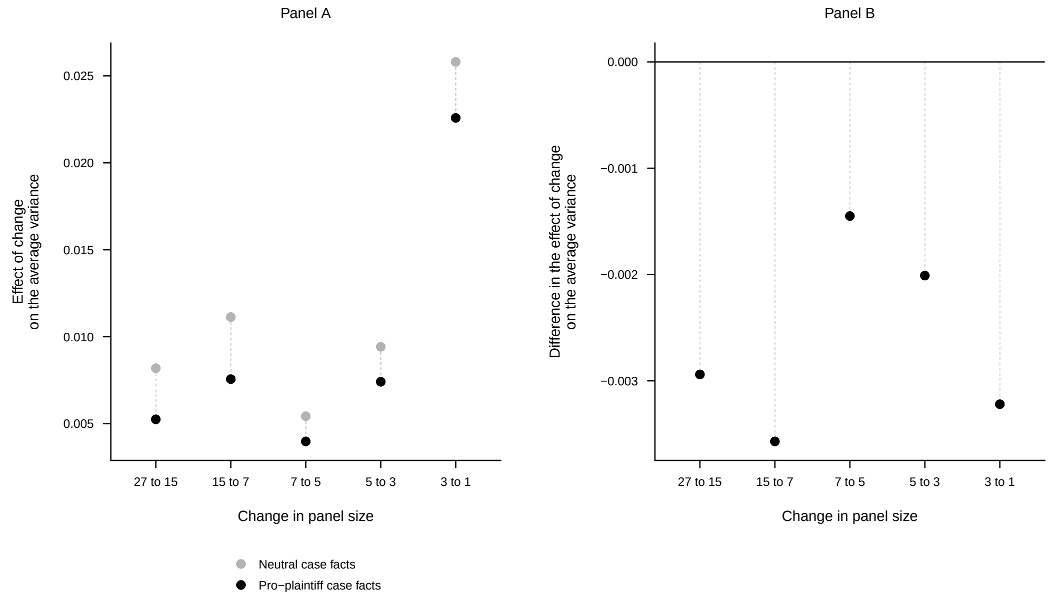

Figure 4. Comparative statics: The location of the case facts

Note: This figure shows the predicted effect of chamber size (s) on the average variance of the probability that the plaintiff wins ($\bar{V}$), conditional on the mean of the case facts distribution (μ_{f_c}). Panel A shows the effect of various changes in chamber size when case facts are neutral ($\mu_{f_c} = 0$) and when case facts favor the plaintiff ($\mu_{f_c} = 1$). Panel B shows the change in each effect caused by increasing the mean.

mean of the case facts distribution is: $\big(\bar{V}(s = 3 \mid \mu_{f_c} = 1) - \bar{V}(s = 5 \mid \mu_{f_c} = 1)\big) - \big(\bar{V}(s = 3 \mid \mu_{f_c} = 0) - \bar{V}(s = 5 \mid \mu_{f_c} = 0)\big)$.

As shown in Figure 4, which shows the effect of an increase in the mean of the case facts distribution (favoring the Commission), I find that increasing the predisposition of the case facts to favor the plaintiff over the defendant (e.g., to favor the Commission over the member state in infringement cases) attenuates the effect of a chamber system. Specifically, it decreases the magnitude of the positive effect of a decrease in chamber size on the average variance of the probability that the Commission wins.

The same would be true if the case facts favored the member state. When there is less overlap between the distribution of case facts and the distribution of judge cut-points, it is less likely that the composition of the panel will matter, in terms of the disposition. The intuition is that, when one party is strongly favored, which judges are on the panel will tend not to matter, as all judges will tend to reach the same decision on the disposition. But when the case is a close call, and could go either way, the preferences of the judges that do happen to get assigned to case could make the difference.

> **Result 3.** *Chamber systems matter less when case facts systematically favor some parties over others, given the preferences of judges.* Increasing the predisposition of case facts to favor the plaintiff (or the defendant) decreases the positive effect of a decrease in chamber size on the variance of the probability that the plaintiff wins across counterfactual panels.

The substantive interpretation of Result 3 is that, when plaintiffs strategically bring strong cases, the chamber system will have a smaller impact on the Court's consistency of the application of EU law. As such, to the extent that the Commission strategically brings infringement cases that it expects to win, that should mitigate the impact of the chamber system on the consistency of the CJEU's application of EU law.

Empirical Analysis

The objective of my empirical analysis is to estimate the extent to which the CJEU's chamber system undermines the consistency of the Courts's application of EU law in practice, given the real-world heterogeneity of CJEU judges' preferences. (Note that the predictions of my formal model are about the dispositional rulings of counterfactual chambers, so it is not possible to test the predictions of the model directly.) I focus my analysis on infringement cases, which are cases in which the Commission sues an EU member state over an alleged violation of EU law, because infringement cases are a hard test. According to Result 3, the impact of the chamber system is smaller when one party is strongly favored, and the Commission's win-rate in infringement cases is approximately 79 percent. I show that the chamber system can still undermine the consistency of the Court's application of EU law even when we should expect the effect to be relatively small. Another advantage of focusing on infringement cases is there is always a clear disposition: either the Commission wins or the member state wins.

Scholars have developed a variety of increasingly sophisticated methods for estimating judges' preferences (Martin and Quinn 2002; Bailey 2007; Clark and Lauderdale 2010; Lauderdale and Clark 2012, 2014, 2016; Bonica and Sen 2017). Some models only take into account votes, whereas others include additional information, such as citations (e.g., Clark and Lauderdale 2010), in order to estimate time-varying, issue-specific (Lauderdale and Clark 2012), and even case-specific positions (Lauderdale and Clark 2016). However, these methods are focused on the Supreme Court and require voting data, which is never made public at the CJEU. Malecki (2012) is the only study to develop a strategy for measuring the preferences of CJEU judges (on a pro-Commission dimension), but this method requires data on third-party briefs for identification, and the CJEU has not published this information for recent decades.

Given this limitation, I develop an empirical strategy that does not require measuring CJEU judges' preferences directly. First, I use machine learning, specifically random forests, to estimate the probability that the Commission wins as a function of the composition of the chamber. Second, I use the trained model to generate predicted probabilities for simulated counterfactual chambers. Third, I calculate the variance of the distribution of these predicted probabilities. Connecting back to my research question, the variance of this simulated distribution characterizes the extent to which the chamber system undermines the consistency of the CJEU's application of EU law in infringement cases.

Data

I use original data on the universe of infringement cases at the CJEU between 1952 and 2018. During this period, the Court published 2,313 judgments in infringement cases. I collect data on the disposition of the case and the composition of each chamber from InfoCuria, the CJEU's official online database. There are 867 unique chamber configurations across

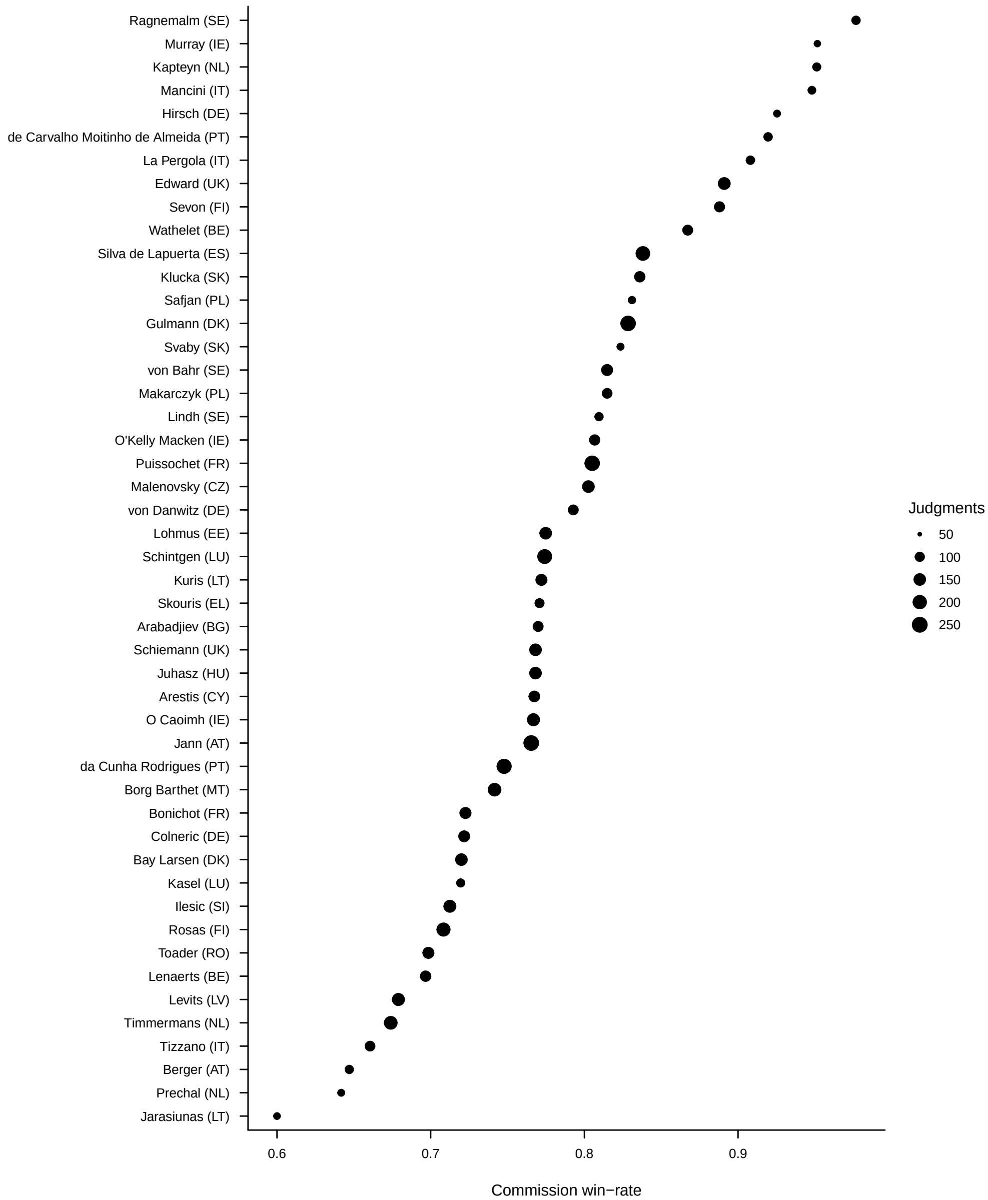

Figure 5. Commission win-rate by judge

Note: This figure shows win-rate of the Commission for all infringement cases that each judge participated in. Only judges that participated in at least 50 cases are included.

these 2,313 infringement cases. The median number of times a configuration appears is only 2, and the mean is only 2.7.

The dispositions are highly skewed, with the Commission winning 79 percent of cases.[6] This is due to strategic selection. The EU has a complex, multi-stage infringement procedure that the Commission uses to negotiate with member states (Fjelstul and Carrubba 2018). If the Commission and a member state cannot resolve the case by the end of the infringement procedure, the Commission can refer the case to the CJEU. As mentioned previously, recent research shows that the Commission strategically brings cases that it expects to win (König and Mäder 2014; Fjelstul and Carrubba 2018). Thus, the infringement cases that reach the CJEU will tend to have case facts that favor the Commission.

We can make a *rough* inference about the heterogeneity of judges' preferences in infringement cases by calculating the average win-rate of the Commission for each judge based on all cases that the judge has participated in. However, we have to be careful in interpreting this descriptive statistic. Figure 5 shows the Commission's win rate for all judges who have heard at least 50 infringement cases. The average win-rate of the Commission varies from 60 percent to over 95 percent. If judges are more or less randomly assigned to cases, then, given the number of unique chamber configurations in the data, this measure should provide a reasonable indication of the extent to which judges differ in their preferences. If we assume random assignment, we would infer that judges appear to be meaningfully heterogeneous, which suggests that which judges you get does matter.

However, it is plausible that the Court could strategically assign judges to cases to limit the influence of outliers and increase the probability of pro-EU outcomes (Kelemen 2012). Under the rules of procedure, the Court has considerable control over which judges get assigned to which cases. Recall that the President of the Court, who is elected, and therefore whose preferences reflect the preferences of the majority, assigns Judge-Rapporteurs. The Full Court then decides chamber assignments, with the constraint that the chamber that

[6] Sometimes, the Commission will be partially successful and partially unsuccessful. I code partially unsuccessful cases as unsuccessful. When the Court rules that a case is inadmissible, I code the Commission as unsuccessful.

hears the case is one of the standing chambers that the Judge-Rapporteur is currently assigned to. Thus, the Court has an opportunity to strategically assign cases.

If this is happening, the Commission's win-rate would be a biased indicator. But we can still learn something about the heterogeneity of judges' preferences from the Commission's win-rate. If the Court wants to maximize pro-EU dispositions, it could disproportionately assign pro-member state judges to cases with pro-Commission facts and pro-Commission judges to cases with pro-member state facts. Thus, pro-member state judges would be associated with higher Commission win-rates under strategic assignment than they would under random assignment, and pro-Commission judges would be associated with lower Commission win-rates. Thus, the Commission's win-rate, as a measure, should *underestimate* the heterogeneity of judges' preferences. So even though Commission win-rates do not account for the facts of cases, we can still infer that there is substantial variation in judges' preferences with respect to infringement cases.

In sum, the descriptive statistics indicate that there is meaningful variation in the preferences of judges, suggesting that the Commission's chances of winning do depend on which judges are in the chamber.

Method

I use random forests to estimate the relationship between chamber composition and the disposition and to predict cases dispositions for counterfactual chambers. This allows me to characterize the extent to which the Commission's chances of winning depends on which judges it gets. This is a supervised classification problem: the objective is to train an algorithm to predict a *class* for each observation based on a set of *features.* In this application, the class is the disposition of the ruling, which is binary. Either the Commission wins or the member state wins. The features are dummy variables that indicate the composition of the chamber. There is one dummy variable for each judge, which is coded 1 of the judge is in the chamber and 0 otherwise. Between 1952 and 2018, 94 judges have participated in at least one infringement case, so the number of features is 94.

A limitation of this analysis is that it does not account for the location of the case facts, which we cannot directly observe and measure. Including a measure of the location of the case facts as a predictor in the model might allow the model to more accurately learn the relationship between chamber composition and the disposition. However, since we cannot control for case facts, we need to be careful in interpreting the model's predictions. Note that, if the Court strategically assigns cases to judges based on case facts to increase the probability of its preferred outcomes (i.e., pro-Commission dispositions), then the location of the case facts would be correlated with chamber composition. Then, what the model is learning is how to predict case dispositions based on chamber composition, conditional on the (possibly strategic) assignment process that generated the empirical data, which means the model's predictions assume the same assignment process.

Random forests are a machine-learning algorithm for classification tasks based on decision trees that can learn complex non-linear relationships between features and classes, which is useful in this application. My formal model assumes the judges' voting decisions are independent, but in practice, a judge could be influenced by the other judges on the panel, and some judges might be more effective at persuading their colleagues than others. A random forest can learn these patterns, if they exist, which improves the model's ability to plausibly predict case dispositions for counterfactual chambers.

Random forests are more robust than decision trees. Individual decision trees tends to overfit the training data — in other words, they tend to memorize the training data instead of learning generalizable patterns. Because of this, they tend to have low bias but high variance. Random forests overcome this problem by combining the predictions of many decision trees, which reduces variance. Each tree votes on the predicted class, and the percent of votes for a class across all trees is the predicted probability. Random forests are random in two ways. First, they use bootstrap aggregation (or bagging): each decision tree is trained on a bootstrap sample of the training data (with replacement). Second, they use feature bagging: at each node in a decision tree, a random sample of features is used to split the data. This allows the algorithm to learn generalizable patterns and make more accurate out-of-sample predictions.

Random forests have one hyper-parameter that needs to be optimized: the number of features (i.e., independent variables) to sample at each branch of a decision tree. The optimal value depends on the data (i.e., there is no way to know *a priori* what an appropriate value is). Following standard practice, I use a grid search to choose a value for this parameter that optimizes the accuracy of the predictions. Empirically, however, the value of this hyper-parameter has little effect on the predictive accuracy of the trained model.[7]

The criteria that I use to evaluate trained models when optimizing this hyper-parameter is the out-of-bag (OOB) error. Unlike other machine-learning algorithms, there is no need to use k-fold cross-validation (where you check out-of-sample accuracy using withheld data). Since each tree uses a bootstrap sample of the training data, we can calculate an unbiased estimate of the error using the out-of-bag observations (i.e., the observations that are not in the bootstrap sample).

Findings

I train the model on the full sample of infringement cases (instead of a randomly selected training sample), as the predictive accuracy of the model can be evaluated using OOB error. I estimate 1,000 trees. One challenge is that the training data is unbalanced. Since the Commission wins 79 percent of the time, there are more pro-Commission dispositions than pro-member state dispositions in the training data. When the training data is unbalanced, any machine-learning algorithm will struggle to predict the less frequent class in a two-class classification problem. A naive model could guess the same class for every observation and be right 79 percent of the time, but that model has not learned anything. With unbalanced data, a better set of metrics to evaluate model performance is precision (the proportion of positive predictions that are correct) and recall (the proportion of actual positives that are

[7] Note that the number of decision trees does not need to be tuned, as the error rate is generally decreasing in the number of trees, with diminishing marginal returns. In this application, increasing the number of trees does not improve model fit.

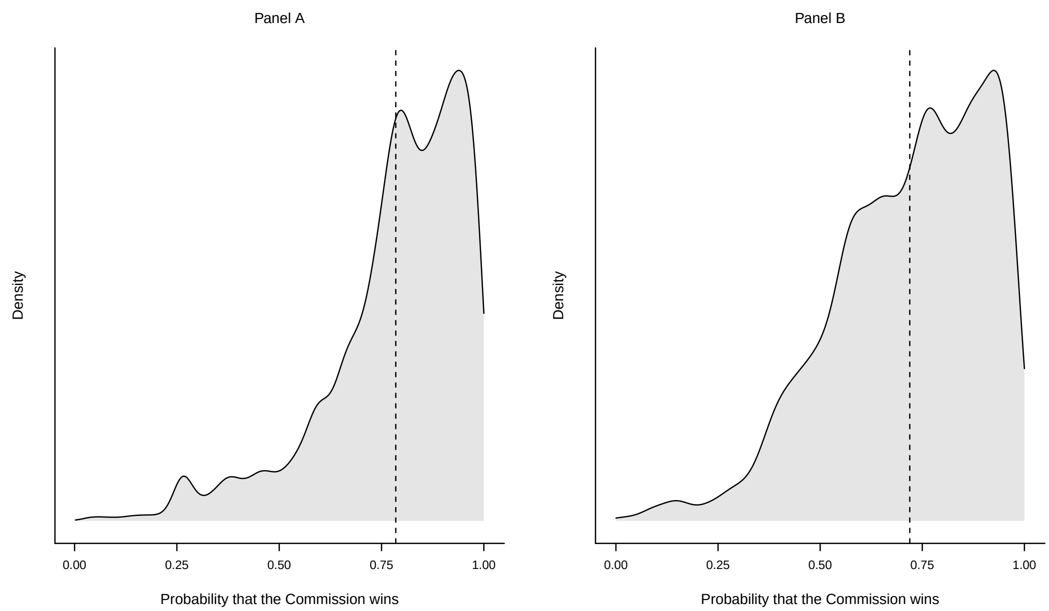

Figure 6. Distribution of predicted probabilities across counterfactual chambers

Note: This figure shows the distribution of the predicted probability that the Commission wins an infringement case across counterfactual chambers based on the trained random forest model. Panel A shows all counterfactual chambers and Panel B shows counterfactual chambers composed of current judges.

correctly predicted). To address the fact that the training data is unbalanced, I oversample cases in which the member state wins (with replacement) when training the model. This is a standard approach for small training datasets.

Based on my computational formal model, we have theoretical reason to believe that, if there is sufficient heterogeneity in judges' preferences, there should be a relationship between chamber composition and the disposition — a pattern for the algorithm to learn. However, there is no way to know *a priori* how predictive chamber composition will be. It could explain a lot of the variation in dispositions or only a little. We have no way of knowing what the theoretical upper bound on the performance metrics is. It just depends on how predictive the features are of the classes. Thus, there is no specific benchmark that we can aim for in terms of the performance metrics.

The trained model, with optimized hyper-parameters, has an overall accuracy of 81.8 percent. For dispositions that favor the member state (the smaller class), precision is 90.2

percent and recall is 77.2 percent. For dispositions that favor the Commission (the larger class), precision is 73.4 percent and recall is 88.2 percent. These performance metrics indicate that the model has learned, based on the training data, how to predict the disposition of an infringement case based only on the composition of the chamber.

I use the trained random forest model to calculate the predicted probability that the Commission wins for 10,000 simulated 3-judge or 5-judge chambers.[8] Consistent with the empirical data on chamber size in infringement cases, I allow each chamber size to be equally likely. In simulating the counterfactual chambers, I use two approaches. First, I randomly draw counterfactual chambers without any restrictions, which can produce configurations that could not have occurred historically (e.g., because the judges did not serve at the same time). Second, I randomly draw chambers composed only of current judges. Figure 6 shows the distribution of the predicted probabilities. Panel A shows the distribution for all counterfactual chambers and Panel B shows the distribution for counterfactual chambers composed of current judges. The distributions are similar in shape. I find that the predicted probability that the Commission wins varies considerably across counterfactual chambers. The interquartile range is 20 percentage points for distribution in Panel A and 28 percentage points for the distribution in Panel B.

These findings depend on the trained model plausibly predicting counterfactual dispositions without our being able to measure the location of the facts of the case, but this simulation suggests that the Commission's chances of winning an infringement case depend meaningfully on which judges are in the chamber.

Conclusion

The CJEU's chamber system greatly increases the Court's productivity, but the fact that all judges do not participate in all cases creates the possibility that the Court will not apply the law consistently. In this paper, I use a computational formal model to study the effect of

[8] The predicted probability for each observation is the proportion of the decision trees in the random forest that vote for the positive outcome (i.e., a Commission win).

hearing cases in smaller chambers on the consistency of the disposition across counterfactual chambers. My model shows how chamber systems undermine the consistency of the law and predicts that the impact of a chamber system will be greater when judges are heterogenous and smaller when selection bias in the filing of cases causes case facts to systematically favor the plaintiff over the defendant. Using machine learning, I show that the predicted probability that the Commission wins infringement cases depends meaningfully on which judges are in the chamber, suggesting that the chamber system does undermine the Court's application of EU law in practice.

My findings suggest that the Court faces a tradeoff between productivity and consistency: using smaller chambers increases the Court's productivity (by allowing the Court to process more cases simultaneously), thereby reducing its backlog of cases, but decreases the consistency of the Court's application of EU law. One way the Court could manage this tradeoff is by strategically sending clear-cut cases (where judges would likely agree how to decide the case, based on the facts), to the higher-variance 3-judge chambers and less clear-cut cases (where judges would likely disagree) to the lower-variance 5-judge chambers. Future research can explore the implications of using smaller chambers for the Court's productivity and assess strategies for managing this tradeoff.

These findings have implications for other collegial courts that use a chamber system. For example, the Bundesverfassungsgericht (the German Federal Constitutional Court) is divided into two senates, each with three chambers. Cases can be heard by a senate or a chamber. My model predicts that, if there is sufficient heterogeneity in the preferences of judges, it could matter which senate hears a case. However, the Bundesverfassungsgericht also has an interesting feature that should mitigate the effect of using 3-judge chambers: the dispositional ruling of a chamber must be unanimous, and if the chamber cannot reach a unanimous decision, the case is heard by the full senate. Future research on chamber systems should theorize how this kind of variation in institutional rules could shape judicial outcomes and develop new empirical strategies for testing the predictions.

Finally, my findings also have implications for the contemporary policy debate about reforming the Supreme Court. The Biden Administration has established the Presidential Commission on the Supreme Court of the United States to study possible institutional reforms to the Supreme Court, and one of the possible reforms that the Commission will consider is enlarging the court. A substantial expansion of the court could prompt the use of chambers. Other proposals that have been discussed in the media include rotating judges onto the Supreme Court from the Courts of Appeal. Given the polarization of American politics, which has also resulted in the ideological polarization of judges, my findings suggest that we should be concerned that any proposal that frequently changes the composition of the court or decreases the average number of judges that participate in a case could undermine the consistency of the Supreme Court's application of the law, which could lead to even more intense partisan fights over the federal judiciary.

References

Ainsley, Caitlin, Cliff Carrubba and Georg Vanberg. 2021. "The future matters: Judicial preferences over legal rules and decision-making on collegial courts." *Journal of Law and Courts* 9(1):1–15.

Bailey, Michael A. 2007. "Comparable Preference Estimates across Time and Institutions for the Court, Congress, and Presidency." *American Journal of Political Science* 51(3):433–448.

Bonica, Adam and Maya Sen. 2017. "A Common-Space Scaling of the American Judiciary and the Legal Profession." *Political Analysis* 25(1):114–121.

Carrubba, Cliff, Barry Friedman, Andrew Martin and Georg Vanberg. 2012. "Who controls the content of Supreme Court opinions?" *American Journal of Political Science* 56(2):400–412.

Carrubba, Clifford J, Matthew Gabel and Charles Hankla. 2012. "Understanding the Role of the European Court of Justice in European Integration." *American Political Science Review* 106(1):214–223.

Carrubba, Clifford J and Matthew J Gabel. 2015. *International Courts and the Performance of International Agreements*. New York: Cambridge University Press.

Carrubba, Clifford and Joshua Fjelstul. 2021. European International Courts: The CJEU and the ECtHR. In *High Courts in Global Perspective: Evidence, Methodologies, and Findings*, ed. Nuno Garoupa, Rebecca Gill and Lydia Tiede. University of Virginia Press.

Cheruvu, Sivaram. 2019. "How do institutional constraints affect judicial decision-making? The European Court of Justice's French language mandate." *European Union Politics* 20(4):562–583.

Clark, Tom and Benjamin Lauderdale. 2010. "Locating Supreme Court Opinions in Doctrine Space." *American Journal of Political Science* 54(4):871–890.

Clark, Tom S. 2016. "Scope and Precedent: Judicial Rule-Making under Uncertainty." *Journal of Theoretical Politics* 28(3):353–384.

Collins, Paul M. 2008. "The Consistency of Judicial Choice." *Journal of Politics* 70(3):861–873.

Fjelstul, Joshua C. 2019. "The evolution of European Union law: A new data set on the Acquis Communautaire." *European Union Politics* 20(4):670–691.

Fjelstul, Joshua C. and Clifford J. Carrubba. 2018. "The Politics of International Oversight: Strategic Monitoring and Legal Compliance in the European Union." *American Political Science Review* 112(3):429–445.

Fox, Justin and Georg Vanberg. 2014. "Narrow Versus Broad Judicial Decisions." *Journal of Theoretical Politics* 26(3):355–383.

Frankenreiter, Jens. 2018. "Are Advocates General Political? Policy Preferences of EU Member State Governments and the Voting Behavior of Members of the European Court of Justice." *Review of Law and Economics* 14(1).

Hermansen, Silje Synnøve Lyder. 2020. "Building legitimacy: strategic case allocations in the Court of Justice of the European Union." *Journal of European Public Policy* 27(8):1215–1235.

Kelemen, R Daniel. 2012. "The Political Foundations of Judicial Independence in the European Union." *Journal of European Public Policy* 19(1):43–58.

König, Thomas and Lars Mäder. 2014. "The Strategic Nature of Compliance: An Empirical Evaluation of Law Implementation in the Central Monitoring System of the European Union." *American Journal of Political Science* 58(1):246–263.

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

Kornhauser, Lewis A. and Lawrence G. Sager. 1986. "Unpacking the Court." *Yale Law Journal* 96:82–117.

Landa, Dimitri and Jeffrey R. Lax. 2008. "Disagreements on Collegial Courts: A Case-Space Approach." *Journal of Constitutional Law* 10(2):305–329.

Landa, Dimitri and Jeffrey R. Lax. 2009. "Legal Doctrine on Collegial Courts." *Journal of Politics* 71(3):946–963.

Larsson, Olof and Daniel Naurin. 2016. "Judicial Independence and Political Uncertainty: How the Risk of Override Affects the Court of Justice of the EU." *International Organization* 70(2):377–408.

Larsson, Olof, Daniel Naurin, Mattias Derlén and Johan Lindholm. 2017. "Speaking Law to Power: the Strategic use of Precedent of the Court of Justice of the European Union." *Comparative Political Studies* 50(7):879–907.

Lauderdale, Benjamin E and Tom S Clark. 2014. "Scaling politically meaningful dimensions using texts and votes." *American Journal of Political Science* 58(3):754–771.

Lauderdale, Benjamin E and Tom S Clark. 2016. "Estimating Vote-Specific Preferences from Roll-Call Data Using Conditional Autoregressive Priors." *Journal of Politics* 78(4):1153–1169.

Lauderdale, Benjamin and Tom Clark. 2012. "The Supreme Court's Many Median Justices." *American Political Science Review* 106(4):847–866.

Lax, Jeffrey. 2007. "Constructing Legal Rules on Appellate Courts." *American Political Science Review* 101(3):591–604.

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

Lax, Jeffrey and Charles Cameron. 2007. "Bargaining and Opinion Assignment on the US Supreme Court." *Journal of Law, Economics, and Organization* 23.

Lax, Jeffrey R. 2012. "Political Constraints on Legal Doctrine: How Hierarchy Shapes the Law." *Journal of Politics* 74(3):765–791.

Malecki, Michael. 2012. "Do ECJ judges all speak with the same voice? Evidence of divergent preferences from the judgments of chambers." *Journal of European Public Policy* 19(1):59–75.

Martin, Andrew D and Kevin M Quinn. 2002. "Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the U.S. Supreme Court, 1953-1999." *Political Analysis* 10(2):134–153.

Siegel, David A. 2018. "Analyzing computational models." *American Journal of Political Science* 62(3):745–759.

Staton, Jeffrey K. and Georg Vanberg. 2008. "The Value of Vagueness: Delegation, Defiance, and Judicial Opinions." *American Journal of Political Science* 52(3):504–519.