

# Holding Judges To Account. Policy, Performance and Impartiality

Silje Synnøve Lyder Hermansen

Daniel Naurin

June 14, 2022

## Introduction

Democracy and rule of law require a judiciary that is both independent and accountable. But while the determinants and effects of judicial independence have been widely studied (Ginsburg and Melton, 2015; Garoupa and Ginsburg, 2009; Hayo and Voigt, 2007; La Porta et al., 2004; J. Ferejohn et al., 1999; Ramseyer, 1994; North and Weingast, 1989), much less scholarly attention has been given to judicial accountability. Researchers have identified a range of mechanisms by which courts as collective bodies can be constrained by other branches of government and by public opinion: Through legislative override, threats of non-compliance, court packing or other court-curbing means (Clark, 2009; Clifford J. Carrubba, M. Gabel, and Hankla, 2008; Vanberg, 2005; Epstein and Knight, 1997; J. A. Ferejohn and Weingast, 1992) and by concerns about shoring up public support and diffuse legitimacy (Gibson and Nelson, 2015; Staton, 2010; Clifford James Carrubba, 2009; Caldeira and Gibson, 1992). These mechanisms specify institutional checks and balances, but have little to say about individual judges' accountability. How – if at all – can judges with de facto law-making powers be held accountable?

Specifically, how can accountability be institutionalized so as to incentivize judicial performance without at the same time sacrificing independence?

We approach this question by studying the strongest mechanism for judicial accountability that has been found acceptable in democratic rule-of-law systems: renewable terms. Setting limits on judges' terms in office, and letting a (re)appointer decide whether they should be allowed to stay on the bench, creates strong incentives for judges to respond to the appointer's desires. This is well understood by court designers around the world. The wide variety of ways in which states select judges is striking (Garoupa and Ginsburg, 2009), but few have opted for renewable terms. Putting judges up for regular reappointments has often been seen as one step too far away from judicial independence (Feldt and Voigt 2003).

Renewable terms are nevertheless common in two contexts. First, almost all US state supreme courts rely on renewable terms (Geyh 2019:47). In some states, judges are re-appointed through more or less contested democratic elections. In other states, the governor or the state legislature decides on reappointment. Regardless of who the appointer is, research indicates that renewable terms have substantive effects on judicial behavior – judges anticipate accountability and adjust to the appointer's policy goals (Gray 2019, 2017, Shepherd 2009).

The second context where renewable terms are common, but where less research exists, is international courts. In 2014, 17 out of 24 active international courts had renewable terms (Squatrino 2018). There is no lack of evidence that governments sometimes select international judges on political grounds (Posner and de Figueiredo, 2005; Elsig and Pollack, 2014) and that international courts and judges sometimes adjust their decisions in anticipation of political reactions (Clifford J. Carrubba, M. Gabel, and Hankla, 2008; Larsson and Naurin, 2016; Stiansen, 2022). However, few studies have theorized or provided empirical evidence of how – and with what effects – renewable terms are used to hold judges to account at the international level.

In this article, we give a theoretical account of how the power to reappoint judges is exercised and what some of the implications are for judicial independence and performance. In brief, we argue that state governments who appoint international judges are concerned with reappointing candidates that they believe are likely to have an effective impact in the right direction, from their point of view, on the case law of the court. This requires attention both to the judge's preferences with regards to the direction of case law and to their potential ability to sway other judges on the bench. To make that call, the appointer seeks information on what judges are both *willing and able* to do if selected.

We demonstrate how the institutional design of courts that operate under renewable terms affect to what extent such information is available to the appointer. Previous research has already highlighted the significance of transparency of votes and individual opinions. The combination of renewable terms with public votes and opinions has been seen as incompatible with judicial independence, as judges will adjust their votes so as to please the appointer and secure reappointment (Dunoff and Pollack, 2017). As a consequence, most international courts that have renewable terms, including the WTO Appellate Body and the Court of Justice of the European Union, make decisions *per curiam* without publishing individual opinions. The purpose of secrecy with regards to votes and opinions is precisely to make it harder for the appointer to evaluate an incumbent judge's behavior on the bench based on the direction of case law they have promoted.

However, we argue that the appointer may still gain access to and make use of relevant information about judges' ability to have an impact on the court's decisions. This happens when the court's organization and decision-making procedures allows for credible information about the performance of judges, in particular with regards to reaching important positions of influence within the court.

Our argument suggests that under such conditions renewable terms may

be a powerful mechanism for judicial accountability, even in the absence of public votes and separate opinions. One important implication is that, with the right transparency rules, high performing judges have a better chance of being reappointed. This is likely to positively impact the overall performance of courts.

Furthermore, even though accountability with regards to the decisions made by judges is severely hampered by secrecy of votes and opinions, appointers may still resort to out-of-court information about judges' preferences. In particular, we propose that an appointer is likely to take into account the policy preferences of its predecessor, who first vetted and appointed the judge whose term is up for renewal. A second important implication of our argument, therefore, is that secrecy of votes and opinions does not necessarily protect judges from politically or ideologically motivated reappointment decisions.

We develop and test our theory with data from the Court of Justice of the European Union (CJEU). The CJEU (or the Court) is often described as an exceptionally independent and powerful international court (Pollack, 2003; Alter, 2009) and an institutional model for many other international courts (Alter, 2014). The CJEU only takes *per curiam* decisions and does not publish separate opinions. Yet the procedure for judicial selection ties judges closely to the member state governments. All member states of the European Union (EU) nominate one judge each to the Court of Justice (CJ), and two judges each to the General Court (GC), for six-year terms that can be renewed indefinitely.

We analyze all 248 reappointment decisions by member state governments since the Court started its activities in 1952 until 2020. Our research design leverages some of the specific institutional features of the CJEU that generate variation in the information available about the performance of judges.

Our findings show that judges that can credibly signal that they will have an impact on the Court's case law increase their chances of being reap-

pointed for a new term. They also show that that shifts in government during a judge's term significantly reduces their chances for reappointment. This is what we would expect from an appointer that is unable to track individual votes and opinions, but who knows the policy preferences of those who appointed the judge in the first place.

These findings indicate that although *per curiam* decisions reduce the opportunity for retrospective accountability with regards to policy preferences, the appointer may still engage in (i) both retrospective accountability and prospective selection with regards to the performance of incumbent judges and (ii) prospective selection based on other – less precise – information about the incumbent's preferences.

An additional important implication of our findings is that – to the extent that a court decides over its own procedures and can influence the information it provides about judges' performance – the court (or its leadership) has a tool for influencing the re-appointment of judges. Judges that for some reason lack the trust of their colleagues or leaders may be blocked from accessing the positions that signal high performance to the appointer.

Our study contributes to the long-standing question of how to balance judicial independence, accountability and performance, in particular by emphasizing the informational aspects of accountability and the institutional features governing transparency of judicial decision-making. It also speaks to the relationship between international courts and the states that created them. Scholars have debated whether international judges are best perceived as policy agents or expert trustees (Elsig and Pollack, 2014; Alter, 2008). Our argument implies that states do not necessarily trade one against the other: Even when judicial appointments are motivated by policy goals, selecting highly competent judges with a strong professional reputation may be the best strategy to have an impact on the court.

## Balancing judicial independence and accountability

Judicial independence is widely recognized as a prerequisite for well-functioning rule-of-law-based societies (J. Ferejohn et al., 1999; North and Weingast, 1989). Judges that are dependent upon those who appoint them cannot be fully trusted to take impartial decisions based on law and facts, which reduces the credibility of legislative and constitutional commitments. Democracy, on the other hand, requires that a judiciary with de facto law-making powers is to some extent responsive to society (Dahl, 1957). Judicial powers, like all public authority, should be accompanied by accountability. The two concepts are often thought of as antonymous; more accountability means less independence (Cross 2008: 566). Institutional designers of judiciaries around the world have been involved in an “eternal struggle” (Garoupa and Ginsburg, 2009, p. 105) for balance between independence and accountability.

Procedures for selection and retention of judges are particularly central to this balancing act. Comparative research on domestic high courts has studied to what extent de jure independence (constitutional or statutory provisions) leads to de facto independence (judicial decision making free from undue external influence), with mixed results (Melton and Ginsburg, 2014; Hayo and Voigt, 2007). To the extent that rules on paper matter at all for real independence, research indicates that the provisions relating to how judges enter and exit the court are the most important (Melton and Ginsburg, 2014, p. 190).

Dunoff and Pollack (2017) have proposed that institutional designers of international courts face “the judicial trilemma”. Out of three values that creators of international courts may want to promote – independence, accountability and transparency – they have to forego one. Accountability here is understood specifically as judges having renewable terms, with states controlling their reappointments (ibid., p. 234). According to the logic of the

trilemma, judicial independence and accountability may be compatible only if there is no transparency with regards to votes and opinions of individual judges. If states are unable to monitor in which direction individual judges prefer to take the court's case law, they will be independent as they do not have to fear sanctions based on the substance of their decisions.

As a theoretical construct, the judicial trilemma is helpful in pointing at how the interaction between rules relating to monitoring (transparency) and sanctioning (renewable terms) affect how judicial accountability may be exercised. However, it leaves unanswered the question of what type of accountability is left for the principal if votes and opinions are secret. How – if at all – do states use the sanctioning mechanism that they have given themselves, if they are unable to follow the direction of case law a judge has been promoting on the bench?

The answer, we believe, is that accountability may still be exercised, but based on other information that is available to the appointer. The absence of transparency about votes and opinions (indicating the direction of preferences) does not rule out information on the individual performance of judges (indicating potential impact of preferences). It also does not rule out prospective selection based on out-of-court information, where the appointer projects an incumbent judge's future voting behavior and compares it with a possible replacement judge. We specify this argument in more detail below. Before that, we review the existing literature with regards to the motives of actors with powers to select judges to international and domestic high courts.

## **Selecting judges to international and domestic high courts**

Research on the appointment of judges is sparse compared to the voluminous literature on judicial behavior and has mostly focused on the United States. Studies of the Senate's confirmation of justices to the U.S. Supreme Court have found that both political ideology and professional competence matter for how senators vote: "Ideologically proximate nominees will be attractive,

poorly qualified nominees unattractive, and nominees who are both ideologically distant and poorly qualified very unattractive” (Cameron et al 1990: 528). The underlying motivation is assumed to be electoral concerns, with senators being rewarded by their constituents for confirming competent justices with the preferred ideological leaning (Cameron et al 1990, Epstein et al 2006).

The literature on U.S. state courts has focused more on judges’ anticipation of possible sanctions than on the actual motivation and behavior of the appointer. We are not aware of any study that analyzes to what extent demonstrated ideological proximity and professional competence on the bench increases judges’ chances of survival when terms are renewable. Nevertheless, to the extent that judges are cognizant of the motives of the appointer, judicial behavior may also be informative as to the goals and strategies of those who have the power to exercise accountability. Several studies indicate that, at least in the minds of the judges, governors and state legislatures with reappointment powers value ideological (Gray 2019, 2017) and policy (Shepherd, 2009) proximity. There is also evidence that federal circuit court judges that are in the race for a nomination to the Supreme Court profile themselves by writing more dissenting opinions, and adjust their behavior to the ideology of the President (Black and Owens, 2016). On the other hand, one study failed to find significant effects of professional competence (measured as publishing and writing dissenting opinions and being frequently cited) on the President’s choice of promotion of district judges to circuit courts (Choi, Gulati, and Posner, 2015).

The literature on international courts contains two main contenders on how to understand judicial appointments. First, the principal-agent model proposes that screening and nomination of candidates is infused with strategic political considerations concerning candidates’ policy preferences (Pollack, 2003; Elsig and Pollack, 2014). Empirical support for that view has been found in a study of states’ appointments of judges to the World Trade



Organization’s Appellate Body. The authors concluded that the process, “far from representing a pure search for expertise, is deeply politicized” (Elsig and Pollack, 2014, p. 3). Empirical evidence that international judges follow the preferences of those who appoint them can be seen in a study of the International Court of Justice (ICJ) (Posner and de Figueiredo, 2005). On the other hand, one important study of the European Court of Human Rights (ECtHR) found large variations in judges’ tendency to defer to their home states (Voeten, 2008, p. 417).

Second, the trustee model emphasizes that decision-making unfolds differently in courts compared to political institutions. International judges are trustees rather than agents because they draw legitimacy from the rational-legal expertise that they hold and the normative ideal of impartial dispute resolution. Influence is obtained through legal arguments, which makes legal skills a key asset. The model expects appointing governments to focus less on political criteria, and more on finding the most competent candidates based on professional legal merits and personal reputation in the legal community (Alter, 2008; Stone Sweet and Brunell, 2013). Furthermore, even if states attempted to use their power to appoint international judges for political purposes, the decentralized character of the appointment process – where no single state or group of states can control the ideological composition of the court – makes it a rather futile exercise. States are therefore unlikely to bother using their appointment powers to promote policy goals (Alter 2008, p. 46, see also Kelemen 2012).

We conclude from these studies that two primary selection criteria are present in the minds of appointers at the national and international level; ideological or policy proximity and professional competence. However, in contrast to much of the literature on international courts, we do not assume that the two are necessarily contradictory. Selecting judges with the right political compass, from the appointer’s point of view, does not preclude taking legal competence and professional status into account. Instead, we propose

that selecting judges on the basis of competence is perfectly compatible with the interests of a policy-oriented appointer. The reason is that professional competence is related to the ability to obtain impact on the bench. Unqualified judges are less likely to influence other judges and therefore less useful promoters of any policy goals. We thereby give professional competence as a selection criterion an alternative interpretation compared to the previous literature, which has focused primarily on electoral interests or concerns for quality and impartiality.

We also challenge the proposition that states refrain from making use of the influence they may gain – even if marginal for a single state – from appointing a judge that is willing and able to promote their policy goals. To the contrary, any appointer – whether at the domestic or international level – should prefer a candidate who is both ideologically committed to its agenda and possesses the authority to persuade other judges.

## Hypotheses

We propose that the ideal candidate is both willing and able to pursue the appointer’s policy agenda. Appointment decisions are therefore based on both the presumed direction of judges’ policy preferences and their predicted ability to impact the court’s decisions.

In cases where there is an open seat to fill, for instance following a voluntary retirement, a reform of the court’s composition or the entrance of new member states to an international court, the appointment of a judge will be determined by a screening process of potential candidates by the appointer based on their previous record and behavior during their careers. In cases where a judge seeks a renewed mandate, on the other hand, the appointer may exercise judicial accountability in the sense of contrasting past behavior on the bench of the incumbent judge with the merits of potential new candidates. However, if decisions are taken *per curiam* and deliberations and votes are secret, as in most international courts, the appointer has no infor-

mation on the direction of the case law that the judge has promoted during the previous term.

We argue that this must not preclude retrospective accountability. Depending on the rules of procedure, the court may reveal relevant information to the appointer about the performance of the judge and their ability to influence decisions. This is the case when significant positions of influence within the court – such as leadership positions or assignments to important cases – are distributed selectively, rather than randomly or by predetermined procedure, and when there is transparency about who is (s)electd to these positions.

We suggest that receiving such selective positions within the court signals to the appointer that a judge is trusted and respected by other colleagues, or by the court’s leadership, and therefore in a position to influence decisions. Our first two hypotheses posit that these conditions affect the appointment decisions:

**Hypothesis 1a** *The probability that a judge is reappointed increases with their past performance on the bench, in terms of being trusted with selective positions of influence.*

**Hypothesis 1b** *The effect only holds when the court makes credible information about performance available to the appointer.*

Transparency with regards to performance does not help the appointer evaluate the direction of case law promoted by an incumbent judge. However, even when votes and opinions are secret, other sources of information on the preferences of the incumbent judge may still be available.

One important shortcut to such information is the preferences of the previous appointer. If the previous appointer has similar ideology as the current appointer – such as when no change in government has occurred during a judge’s term – and if the current appointer have reason to trust that the initial appointer performed a reliable screening process, the current

appointer may assume that the incumbent judge is likely to have preferences similar to its own. Secret voting, in fact, increases the reliability of this assumption: Since the initial appointer knows that they will not be able to monitor the voting decisions of the appointed judge, they have stronger incentives to invest seriously in the screening process.

Although this source of information is clearly less precise than being able to monitor a judge's votes and opinions on the bench – assuming that a new screening process would be relatively costly – the current appointer is likely to retain a judge selected by a predecessor with similar policy preferences. One implication is that under a system of secret voting and renewable terms a shift in government that increases the preference distance between the initial and the current appointer also decreases the chances that a judge is reappointed.

**Hypothesis 2** *The probability that a judge is reappointed decreases with the distance in preferences between the appointing and the reappointing governments.*

In the next section, we contextualize our argument and hypotheses in the case of the Court of Justice of the European Union.

## Judicial appointments at the Court of Justice of the European Union

In international courts, such as the ICJ and the WTO Appellate Body, when the number of judges is smaller than the number of member states the appointment process involves political bargaining between coalitions of states (Elsig and Pollack, 2014). In contrast, in full-representation courts like the ECtHR and the CJEU, the appointment of a judge results from the choice of a single government. All EU member states nominate one judge each to

the Court of Justice and two judges each to the General Court, the two formations of the CJEU. Since 2010, the appointment procedure also includes a merit selection committee (the Article 255 Committee), composed of high-ranking national judges and former CJEU judges, which gives an opinion on the qualifications of the member states' nominees (Dumbrovsky, Petkova, and Van der Sluis, 2014). All judges are subsequently appointed by consensus by the 27 governments, but this latter stage has so far been a mere formality (Dunoff and Pollack, 2017).

The member states of the EU have cautiously held on to their prerogative to appoint the judges of the CJEU. When the General Court was reformed in 2015 to address its increasing case-load, the Court and the European Commission suggested increasing the number of judges by 12 and supplement any further needs with legal clerks. This suggestion was rejected by the member states, who instead preferred the more costly option to double the number of judges (from 27 to 54) and thus keep intact every member states' right to an equal number of judges (Commission of the European Union, 2021; Ministers, 2015).

Individual member state governments are therefore the key actors in the selection of judges to the CJEU. Surprisingly, there has been little research on how those decisions are taken. Despite its elevated status as the motor of the much-debated judicialization of politics in Europe (Pollack, 2003; Alter, 2009; Kelemen, 2011; Schmidt, 2018), scholars have noted that "there is shockingly little written on the process through which ECJ [CJEU] justices are appointed" (Kelemen, 2012, p. 50; see also Kenney, 1998–1999, p. 104).

As outlined in the previous section, we suggest that the selection of judges is motivated by governments' interest in having an effective impact on the Court's case law in their preferred direction. While impact is a function of the professional competence and ability of a judge to persuade others on the bench, preferred direction is a question of policy or case law preferences. We now turn to the question of how to conceptualize and operationalize

potential impact and policy preferences in the case of the CJEU and how governments may find credible information about these characteristics of incumbent judges.

## **Information about performance**

Two features of the rules of procedure of courts are potentially important for performance-based accountability of incumbent judges. First, the division of labor within a court gives some judges a position to have more influence over the decisions than other judges. Second, the allocation of these positions may sometimes be selective, while other times it is random or a question of procedure. Selectively acquired positions may provide valuable information in the appointer's assessment of the potential impact of incumbent judges.

We consider two positions of influence that are especially important in the CJEU, the Reporting Judge and the Chamber President.

First, in contrast to some other international courts, such as the ECtHR and the Inter American Court of Human Rights, the CJEU publishes information on the identity of the Reporting Judge (or Judge Rapporteur) in a case. In the Court of Justice, the Reporting Judge is selected by the President of the Court. The assignment implies being a case manager, agenda setter and opinion writer for a specific case. The Reporting Judge writes up a preliminary report to the General Meeting of the Court, with proposals for important procedural steps in the case management. These procedural steps include whether a case is significant enough to include an opinion of an Advocate General (an advisor to the chamber), an oral hearing, and whether it requires a decision in the grand chamber (with 15 judges) or if a smaller chamber of five or three judges is enough. Most importantly, the Reporting Judge is in charge of collecting information and presenting the case to the other judges in the chamber. They pen the first version of the judgment – which forms the basis of the deliberations in the chamber – as well as the final judgment.

The responsibility to write up the judgement applies even if the Reporting Judge was in the minority. This means that member state governments cannot be sure that the preferences of the Reporting Judge are reflected in the decision. However, being selected for this role in many important cases signals that the judge often find themselves in a position of influence.

Second, the organizational hierarchy of the Court includes top-level and mid-level leadership positions. The top-level leaders – the Presidents and Vice-Presidents of the Court – manage the Court’s day-to-day life, preside over grand chamber cases and represent the Court to the outside world. They distribute influence and monitor the rank-and file judges. The President of the Court of Justice – assisted by the Vice-President – assigns cases to a Reporting Judge, while the President of the General Court assigns cases to chambers. The mid-level leadership includes the Presidents of Chambers of five judges. They preside over the deliberations and decide when the debate is over and it is time to call a vote. In the General Court, they also select the Reporting Judge. The mid-level positions as Chamber Presidents are prestigious and come with enhanced influence over the deliberations.

Importantly, these influential and prestigious positions are sometimes allocated selectively, and sometimes not. The Reporting Judge is always appointed on a case-by-case basis, but the assignment is done differently in the two formations of the Court. Report allocation in the Court of Justice follows a logic of selection. The President of the Court makes the choice early in the process and the composition of the chamber follows from that initial decision. The President has few restrictions in their selection of the Reporting Judge. While political considerations may play a role, the President generally promotes individual-level specialization; effectively allowing judges to capture a disproportional influence over certain issue areas (Hermansen, 2020). Once the Reporting Judge is appointed, the President leaves the monitoring of the case management to the checks and balances of the chamber deliberations (Justice of the European Union, 2012, art. 15). The Reporting Judge

enjoys a large autonomy in their case management in the Court of Justice. The President assigns the most important cases – which usually go to the grand chamber and which eventually attract more attention from the legal community – to judges that are perceived as trustworthy and competent. Being the Reporting Judge of many important cases therefore distinguishes an influential judge from their peers. Moreover, since these positions are selectively allocated, they also speak to a judge’s future performance, since the leadership’s allocation criteria remain consistent over time.

In contrast, report allocation in the General Court follows a logic of monitoring. Cases are allocated to chambers immediately after their filing. The Reporting Judge is then appointed by the Chamber President among the more restricted number of judges at their disposal (art. 26 Court, 2015). Furthermore, the top leadership is kept informed throughout the deliberation. Once the Reporting Judge has a final draft of the judgment, it is communicated to the Vice President. Their assessment is then returned to the Reporting Judge with the rest of the chamber’s judges in copy. In short, Reporting Judges in the General Court have less autonomy and influence compared to Reporting Judges in the Court of Justice.

There is also a difference over time in how leadership positions have been distributed. The President of the Court has always been elected by his peers for a three-year renewable period. Since 2003 the Presidents of the Chambers of five judges are also elected for a three-year term, once renewable (Justice of the European Union 2012, art. 12, Court 2015, art. 18)<sup>1</sup>. In contrast, before 2003 these positions were distributed following a rotational setup.

In sum, although the CJEU is secretive when it comes to the direction of decisions promoted by individual judges, it sometimes reveals significant

---

<sup>1</sup>The new higher status of Chamber Presidents applies only to presidents in the larger formations of five and not to chambers of three. The chambers of three (and their presidents) are in practice subsections of the larger chambers. While the presidents of the chambers of three judges are also formally elected, in practice, these positions are allocated on a proposal from the President of the Court and do not make the object of a competition between judges.



information to reappointing governments about the past performance and potential future influence of incumbent judges in terms of being selected for prestigious and influential positions. In the empirical analysis, we leverage this variation in the credibility of information to tease out the effect of transparency on governments' reappointments.

## Information about policy preferences

While the CJEU reveals information about the past performance of judges it is secretive when it comes to policy or case law preferences. For information about preferences, we argue that the current appointer may instead take cues from the preferences of the previous appointer when deciding on the reappointment of an incumbent judge. When conceptualizing the relevant policy preferences in the selection of judges to the CJEU, we assume that the dominant conflict dimension in European politics – the economic left-right dimension – is the most significant. Not only is the left-right dimension historically the central distinguishing feature of European national party systems, it is also highly relevant for EU politics and large parts of EU law. The role of the state in the economy, and individuals' rights towards the state, is at the center of the creation of the common European market.

The case law of the CJEU has often raised controversy along the left-right dimension. For example, while some observers have perceived the Court's strengthening of individual citizenship rights as a progressive step towards embedded liberalism and social rights at the international level(**CaporasoTarrow2009**), others have lamented the “Hayekian” bias inherent in the Court's case law (**HopnerSchafer2012**; Scharpf, 2010; Schmidt, 2018). Scholars have also shown that the economic left-right dimension structures divisions among member states in amicus briefs submitted to the Court (**LarssonNaurin2019**) and that the Court's leadership distributes influential positions among judges following the the same policy spectrum (Hermansen, 2020).

One of few articles that have addressed the selection of CJEU judges

directly, posits that “some anecdotal evidence suggests Member States sought candidates who were perhaps less of a Euro-enthusiast in a general sense than their predecessors” (Kenney, 1998–1999, p. 128). In our extended empirical analysis, we also take into account the possibility that governments select judges based on preferences relating to European integration and national sovereignty.

## Empirical strategy

Our empirical analyses rely on an original data set listing all 422 appointment decisions that were made to the Court since its inception (1952-2019). Information on names, appointment dates, the duration of mandates as well as the cause of exit was collected from the *Official Journal of the European Union*. Judges exit the Court for several reasons. While many appointments (51%) coincide with the end of a mandate, a large number of judges also resigned before the due expiry date (37%). Early terminations have also occurred because the judge deceased (4%) or was promoted to other positions at the CJEU (6%).

We use this information to identify situations where a government had the opportunity to replace a sitting judge. That is, we consider only potential *reappointments*. There are two reasons for this. First, we are interested in governments’ selection criteria, but have no data on the alternative candidates. For each decision we therefore measure the difference in preferences between the appointing and reappointing governments. In this way, we can test whether two governments from the same member state have different preferences over the same judge.

Second, we seek to isolate governments’ decisions to replace a judge from the judges’ own voluntary decision to leave the Court. We therefore discard all exit decisions over which member states did not have a say, retaining only exit decisions that were due to the expiration of a mandate. For the same

reason, the analysis also includes a set of controls designed to capture the judges’ career stage. We further verify the findings in a placebo test comparing judges’ resignations with governments’ reappointment decisions. We expect that the career-related controls have similar effects in both samples, while our explanatory variables – government preferences and judges’ performance – have no bearing on decisions where governments are not involved (the placebo).

We end up with a final data set of 248 reappointment decisions to the Court of Justice (61%) and the General Court (39%), while the placebo test is conducted on a similar-sized sample of 257 decisions.

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Missing
Change of judge	0	0	0	0.27	1	1	0
Preference distance (economic issues)	0	0.08	0.24	0.35	0.47	2.52	11
Preference distance (integration issues)	0	0.05	0.13	0.2	0.24	1.69	11
Preference distance (general left-right issues)	0	3.77	10.85	13.13	20.32	58.58	11
Performance (cases in larger panels)	-1	-0.07	-0.01	-0.03	0.05	0.74	0
Performance (cases of interest to the legal community)	-6.3	-0.92	-0.16	-0.23	0.3	5.28	0
Performance (selective leadership positions)	0	0	0	0.25	0.25	1	0
Non-selective leadership positions	0	0	0	0.28	1	1	0
Change of prime minister	0	0	1	0.51	1	1	0
Age	37.72	54.52	59.93	59.99	65.64	83.75	0
Tenure	1	3.67	6	7.2	9.44	21	0
Change in attendance	-57	-5	5	5.07	16	61	12

Table 1: Descriptive statistics

### Dependent variable: Replacement

Our dependent variable, *Replacement* is binary. It captures all member state decisions that could lead to a judge’s exit from the CJEU and flags those resulting in replacement. From the descriptive statistics in Table 1, we see that governments most often (73%) decide to renew the mandate of judges whose term has come to an end. In the placebo test, we rely on the same reference group – judges who are retained when their mandate expires – but consider instead voluntary exits (resignations). The univariate distribution remains very similar (81%). The propensity to receive a new mandate has remained fairly stable over time.

## Explanatory variables

In the absence of new information on policy preferences, we suggest that reappointing governments will rely on past assessments of a judge. Thus, we expect that the probability of replacement increases with the distance between successive governments' preferences ( $H_2$ ). We place governments in a single policy space using party manifestos covering the Court's entire 70-year history (Döring and Manow, 2018; Volkens et al., 2017; M. J. Gabel and Huber, 2000).

*Preference distance - economic left-right issues* is a continuous measure of the distance between the appointing and the reappointing governments on economic left-right issues. It is calculated as the absolute difference between the factor scores estimated from cabinet parties' electoral manifestos. The bivariate statistics already give an indication that the distance with regards to economic policy preferences matter for reappointment. While the median distance between governments that prefer the same judge is 0.17, it increases to 0.38 when the incumbent judge is replaced. In the Appendix, we report two alternative operationalizations of policy preferences: Divisions along a general left-right axis and preferences on EU integration.

We expect that the probability of a replacement is lower for high-performing judges ( $H_{1a}$ ) that acquired prestigious positions through peer selection ( $H_{1b}$ ) because governments take this information as indications of future influence. We consider two important positions that judges may hold within the internal judicial hierarchy: The position as Reporting Judge and the position as President of the Court or of a larger Chamber formation. For each position, we distinguish whether it was allocated by peer selection.

Case significance is measured in two alternative ways. The first operationalization uses chamber size, i.e. the number of judges deciding a case, to capture the court's own assessment of the significance of the case. We then consider the level of attention that cases have attracted in the legal community.

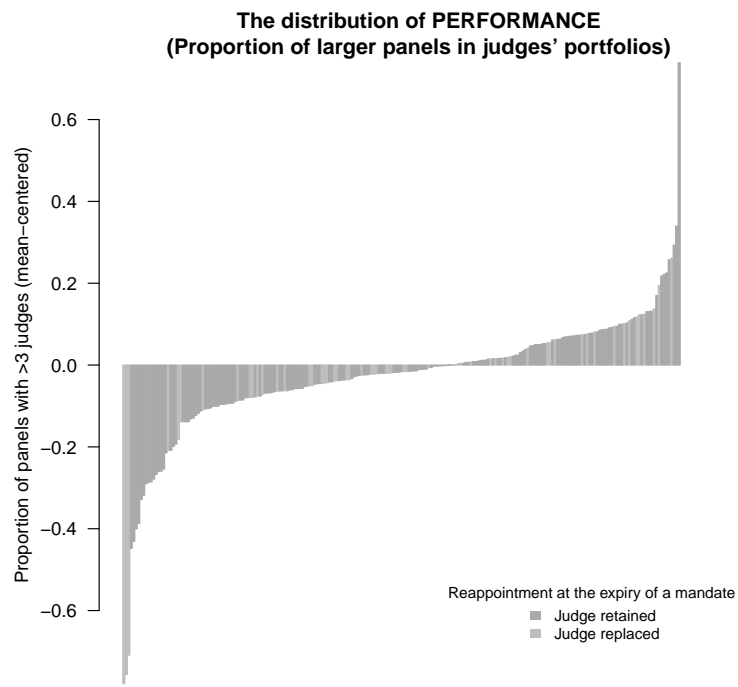


Figure 1: Variation in the proportion of salient cases that judges' handle compared to the average member on the Court.

*Performance – cases of interest to the Court* reports the proportion of cases in a judge’s portfolio as a Reporting Judge that were decided by a chamber of more than three judges. While the Court’s reliance on small chambers has increased over time, the cases of least consequence have always been delegated to chambers of three judges. Figure 1 illustrates a substantial variation in judges’ influence as measured in this way. Each bar represents a potential reappointment decision. Bars above the x-axis indicate judges whose portfolios include disproportionately more cases of interest to the Court, while those below show judges who handle disproportionately fewer such cases. Despite these substantial differences, the bivariate replacement rate among the high performers was marginally *higher* (29%) than the low performers (27%). One reason may be that senior members are both more likely to take up influential positions and more likely to retire. The multivariate analysis is designed to distinguish these effects.

*Performance – cases of interest to the legal community* reports the mean number of annotations (i.e. journal articles) that a judge’s portfolio of cases has attracted since their last appointment. The legal community regularly comments on judgments in academic journals. Information on such annotations is collected and reported by the CJEU itself. The annotations are thus reflective of the academic salience of cases that the judge has been entrusted with.

The general level of these two measures has changed over time and may vary according to the length of each judge’s term in office. Both variables are therefore reported as ratios to the type of cases handled by the Court during the judge’s term. In other words, we approximate the benchmarking that governments can reasonably do themselves when assessing the incumbent’s performance by comparing the performance of each judge with their peers.

In each specific case, the Reporting Judge is the single-most important member of the Court. However, this role comes with more autonomy and influence in the Court of Justice compared to the General Court. Having

this role in many significant cases is therefore a stronger signal of trust and potential impact in the Court of Justice. When considering judges' performance as the Reporting Judge, we analyze the two formations of the Court separately, expecting only an effect in the Court of Justice.

Our third measure of performance captures leadership positions. Following the Court's reorganization in 2003, the status of the Presidents of Chambers of five judges was substantially increased. While all Chamber Presidents prior to that point were appointed following a principle of rotation, they now became positions acquired in competitive elections (RoP of CJ, 1991Article 10(1); RoP of CJ, 2003Article 10(1)). The Presidents of Chambers of five are now elected to their offices by their fellow judges, which is a signal of trust and competence. Again, we argue that this conveys information about judges' future performance to the appointing governments. *Performance - selective leadership positions* is a binary variable that identifies the President and Vice President of the Court and the Presidents of Chambers of five judges post 2003. It is included in a separate analysis together with an indicator of all judges who have held such a position prior to the reform (labelled *non-elective leadership positions*). In total 51 of the 141 appointment decisions involved elected leaders. We expect that only the selective positions have a significant and negative effect on the odds of replacing a judge.

## Controls

In addition to presumed policy preferences and manifested individual performance, there are a number of other factors that might affect either judges' decisions to exit the CJEU or governments' assessments of the incumbent. The multivariate analyses account for these.

First and foremost, there are instances when judges might complete their term, but do not wish to continue. Few positions are more attractive for European judges than being a judge in the CJEU. Very few alumni have pursued high-ranking, work-intensive jobs upon their exit. Except for government de-

selection, therefore, the most likely reason for a judge to leave the Court is retirement. Judges often signal their intent to their respective governments in a letter of motivation. While we do not have access to these letters, we control for the judge's career stage and judicial behavior immediately prior to the reappointment decision.

Thus, *Length of tenure* approximates the judge's career stage. The average tenure upon exit was 9.97 years, with the median judge sitting for two six-year periods. Similarly, we control for a judge's *Age* at the time of the reappointment decision. The average age of judges exiting the court was 65 years. Thus, at the next reappointment decision, the judge would be 71 years; well beyond the retirement age in most member states. Both variables are mean-centered and we expect them to correlate positively with the decision to exit. We furthermore control for changes in a judge's investment in their mandate. A judge that plans on an exit may decrease their participation in Court activities. *Change in attendance* therefore measures the change in the number of deliberations in which a judge has partaken in the year immediately preceding the official exit decision (as compared to the previous year). In the multivariate analysis the variable is centered around the median judge, and we expect that when attendance decreases, the odds of replacing the judge increases.

All models also contain an indicator for whether there has been a change in the prime minister's party between appointments (*Change in PM*). This variable controls for the possibility that the ruling party uses appointments to distribute spoils to political friends. Models exploring judges' performance furthermore include an indicator of whether the judge held the position as President or Vice-President of the Court at the time of the reappointment (*(Vice-)President*), since these positions clearly come with influence and prestige although the division of labor implies that they handle few cases as a Reporting Judge themselves.

An incumbent judge can either stay in office or exit the Court. Given that



our dependent variable is binary, we use a binomial logistic model (Long, 1997, p. 34-84).

$$\begin{aligned} Pr(y_i = 1) &\sim \text{Bernoulli}(\pi_i) \\ \text{logit}(\pi_i) &= \alpha + \beta_k \times X_i + \beta_z \times Z_i \end{aligned} \tag{1}$$

The definition of the variables of interest,  $X$ , varies across models because we explore different operationalizations of the two hypothesized selection criteria; preferences and performance. However, all models include the same controls,  $Z$ .

The results are obtained from Bayesian models with MCMC estimation (Plummer, 2003). Some observations lack information on governments' preferences, judges' birth or entry dates. These are imputed in parallel to the estimation of the main model. Furthermore, the results reported in the main part of the article are estimated on pooled data. In the Appendix, we verify that these results hold also when we account for systematic variation that may be present in cross-sectional time-series. Overall, our results are robust to a range of alternative modelling strategies.

## Results

The results reported in Tables 2 and 3 support our expectations.

First, the probability that a judge is reappointed decreases as the political distance between the appointing and reappointing governments grows ( $H_2$ ). Table 2 shows that this effect holds for both formations of the Court. The effect is illustrated for the Court of Justice in Figure 2, and its size is non-trivial. Even a median change in a government's attitudes on economic policies would imply a 40% increase in the odds of a replacement. As a way to illustrate the effect, we may consider situations where a member state has shifted between a conservative and a social democratic prime minister<sup>2</sup>. Such

---

<sup>2</sup>There are 35 instances in the data where the prime minister's party has shifted from

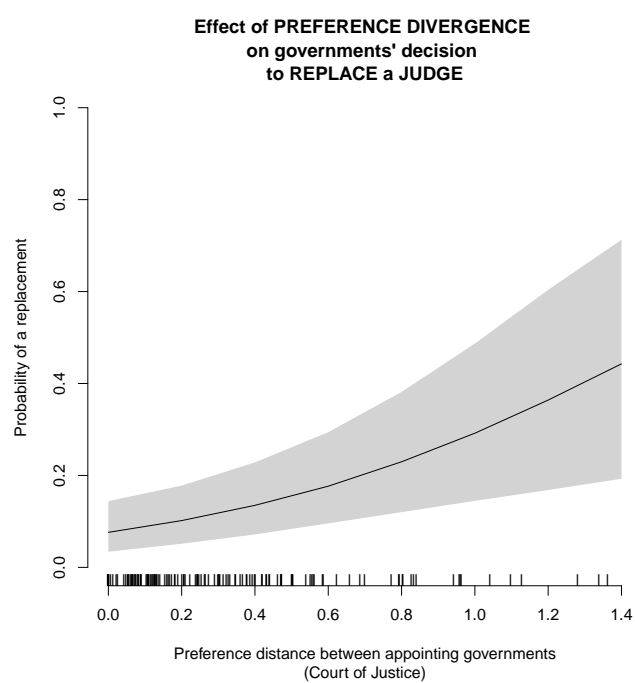


Figure 2: Predicted probability of replacement when there is a change in appointing government's economic preferences in the Court of Justice.

a shift would more than double the odds of a replacement.

In the Appendix, we explore alternative operationalizations of government preferences. Divisions on a general left-right dimension yield similar, but more moderate effects compared to economic left-right issues. However, we find no support for earlier suggestions that governments may prioritize preferences related to the distribution of competence between national and European governance levels in the selection of judges (Kenney, 1998–1999). This finding is consistent with more recent research on governments submissions of amicus briefs to the Court, which indicate that divisions generated from differences in welfare state models structure competition between member states in cases at the CJEU (LarssonNaurin2019).

Dependent variable: 'Replacement'	Court of Justice	Court of Justice	General Court	General Court
Intercept	-0.27 (-1.76,1.32)	-1.15 (-2.07,-0.27)	-2.25 (-3.33,-1.32)	-2.06 (-3.24,-0.97)
Preference distance (economic issues)	1.41 (0.44,2.35)	1.39 (0.45,2.37)	1.76 (0.51,3.07)	1.85 (0.58,3.26)
Performance (cases of court interest)	-1.88 (-3.56,-0.36)		-0.07 (-0.67,0.56)	
Performance (cases of interest to legal community)		-1.02 (-1.89,-0.23)		-0.3 (-1.27,0.52)
Change of PM party	0.05 (-0.83,0.88)	0.01 (-0.84,0.87)	0.81 (-0.14,1.8)	0.85 (-0.13,1.8)
(Vice-)President	0.67 (-0.81,2.12)	0.99 (-0.48,2.46)	-3.39 (-6.99,-0.83)	-3.33 (-7,-0.72)
Length of tenure	0.13 (0.02,0.24)	0.13 (0.02,0.25)	0.11 (0,0.23)	0.12 (0,0.24)
Age	0.17 (0.11,0.24)	0.17 (0.1,0.24)	0.11 (0.04,0.18)	0.11 (0.05,0.18)
Change in attendance	-0.02 (-0.04,0)	-0.02 (-0.04,0)	0.03 (0,0.07)	0.03 (0,0.07)
Number of observations	151	151	97	97
Proportion of correct predictions	0.77	0.79	0.73	0.74
... correct positive predictions	0.71	0.79	0.77	0.73
... correct negative predictions	0.79	0.79	0.72	0.75

Median effects with 90% symmetric posterior density interval in parenthesis.

Table 2: Replacement of judge as a function of CASE ALLOCATIONS. The models explore different operationalizations in the two courts.

We have argued that selecting judges with the right policy compass is only one part of the appointer’s calculus. The other part is about selecting judges that are likely to have an impact on the court’s decisions ( $H_{1a}$ ). We found no indication in the bivariate analysis that past performance, in terms social democratic to conservative (or vice-versa) between appointments. The median distance in their estimated economic preferences is 0.51.

Dependent variable: 'Replacement'	Both courts
Intercept	-2.04 (-2.71,-1.47)
Preference distance (economic issues)	0.91 (0,1.81)
Elected leadership	-0.7 (-1.5,0.02)
Non-elected leadership	0.3 (-0.96,1.47)
Change of PM party	0.83 (0.16,1.52)
Length of tenure	0.11 (0.03,0.19)
Age	0.09 (0.05,0.14)
Change in attendance	-0.01 (-0.03,0.01)
Number of observations	197
Proportion of correct predictions	0.68
... correct positive predictions	0.62
... correct negative predictions	0.7

Median effects with 90% symmetric posterior density interval in parenthesis.

Table 3: Replacement of judge as a function of LEADERSHIP positions. The models explore different operationalizations in the two courts.

of having been trusted with many important cases as the Reporting Judge, matters for reappointment. However, when controlling for judges' career stage and government preferences, there is a substantial effect of judges past performance.

First, handling high-impact cases as Reporting Judge in the Court of Justice, where such positions are selective and associated with significant influence over the case management process, is clearly linked to judges' chances of being reappointed. This relationship holds regardless of how we measure the importance of judges' portfolios.

Model 1 shows that the more large-chamber cases a judge has obtained as Reporting Judge, the more likely they are to remain in office. All things equal, judges whose portfolio contains 10 percentage points more large-chamber cases than the overall distribution in the Court, have a 87 % lower odds of replacement. Model 2 then considers the mean number of academic articles discussing cases where the judge has acted as Reporting Judge. Here too, we find support for our hypothesis. If we consider the difference between a typical under-performing (20<sup>th</sup> percentile) and a typical over-performing judge

(80<sup>th</sup> percentile), the most influential judge has a 54% higher probability of retaining their seat.

Second, table 3 and Figure 4 report the effect of being trusted with a leadership position, as President of the Court or of a larger Chamber, in either of the two formations of the Court. Once again, the effect is sizable, although less precise ( $p < 0.06$ ). The odds of replacing an elected leader is 50% lower than that of a rank-and-file judge, while the effect is indistinguishable from zero for non-elected leaders.

This means that, conditional on positions being selective, high-performing judges are to some extent shielded against politically motivated deselection. If we consider the same scenario as before, where there has been a substantial change in government preferences from a social democratic to a conservative prime minister, a judge who can signal high performance by achieving a selective leadership position would have about the same probability of being replaced (17%) as a low performer in a situation where no change in government took place (19%). Our results also indicate that the Court tends to reserve the most influential positions to experienced judges. Among the incumbent candidates in our data, most of the Court’s elected leaders – as well as those judges who handled a disproportionately high share of large-chamber cases – were in their second term.

### **Robustness tests**

The models in Tables 2 and 3 provide a fair description of the data, with an in-sample correct prediction rate of 77%. They are also robust to several alternative specifications, relating to variation over time and between member states. For example, we find no evidence that the introduction of the advisory merit selection committee in 2010 (the Article 255 committee) changed member states’ emphasis on preferences or performance when they decide whether to replace an incumbent judge. Details about these additional tests can be found in the appendix.

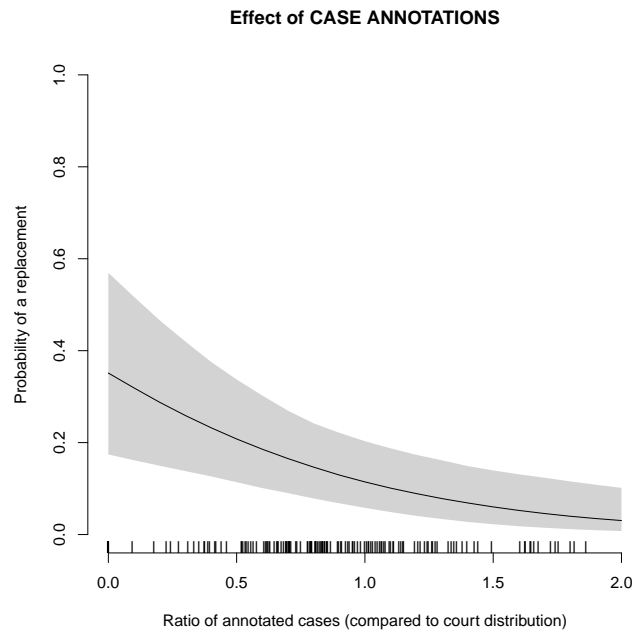
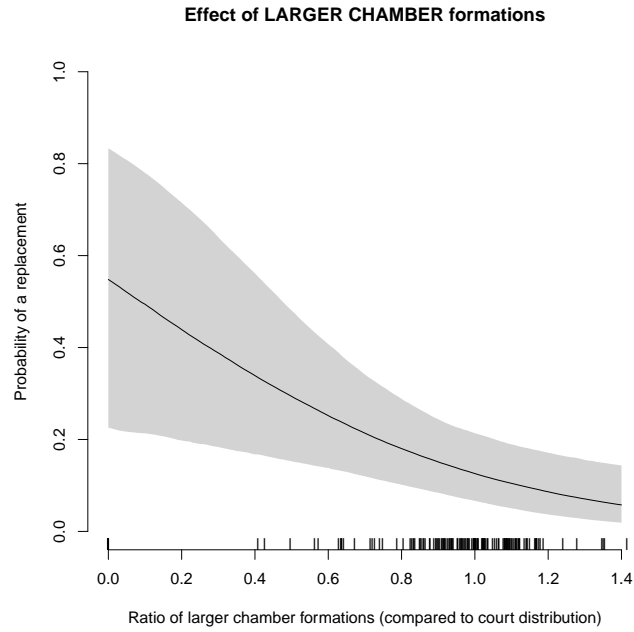


Figure 3: Predicted probability of replacement as a function of case portfolio among judges in the Court of Justice.



Figure 4: The effect of elected and non-elected leadership positions on governments' choice to replace their incumbent judge.

We also perform an additional test to further probe our assumption that replacements are driven by governments' decisions rather than judges' choices to exit the Court voluntarily. Table 4 reports results from a placebo test where we substitute those occasions where a judge was replaced at the end of a term for those instances when a judge exited from the court during a term. Career-related factors such as age, effort and seniority, should be related to both types of exits, as judges that retire voluntarily may do so both during and after a term. In contrast, governments can only replace a judge when the incumbent's mandate has expired. In line with our expectations, we only find an effect of preference distance and performance in the Court of Justice for the subset of the data where governments had the opportunity to make a decision to replace a judge.

Dependent variable: 'Replacement'	Mandate expired	Resignation
Intercept	-0.27 (-1.76,1.32)	-0.81 (-2.65,0.81)
Preference distance (economic issues)	1.41 (0.44,2.35)	-0.79 (-2.29,0.58)
Performance (cases of interest of court interest)	-1.88 (-3.56,-0.36)	-0.93 (-2.61,0.77)
Change of PM	0.05 (-0.83,0.88)	-0.17 (-1.12,0.72)
(Vice-)President	0.67 (-0.81,2.12)	0.3 (-1.24,1.83)
Length of tenure	0.13 (0.02,0.24)	0.2 (0.09,0.33)
Age	0.17 (0.11,0.24)	0.15 (0.07,0.24)
Change in attendance	-0.02 (-0.04,0)	-0.03 (-0.05,-0.01)
Number of observations	151	139
Proportion of correct predictions	0.77	0.77
... correct positive predictions	0.71	0.77
... correct negative predictions	0.79	0.77

Median effects with 90% symmetric posterior density interval in parenthesis.

Table 4: Placebo test on appointments to the Court of Justice: Reference level for both is instances where the term is expired and the judge is reappointed. Results from two logistic regressions.



## Discussion

We have explored the implications for judicial accountability, performance and independence of combining renewable terms with secrecy of votes and opinions, which is a particularly common institutional design feature of international courts. Our account of how the power to reappoint judges is exercised challenges previous conceptions of the appointer's choice as being either politically motivated or based on (genuine or reputational) concerns with professional competence. Instead, we have argued that a policy seeking appointer will look for a candidate that is both willing and able to make an impact on the court's case law in the direction preferred by the appointer.

The extent to which such information is available depends on the organization and transparency of the court. If the court publishes credible information about judges' performance on the bench, judges may promote their chances of being reappointed by increasing their efforts. From the perspective of the institutional design of courts, such transparency rules may therefore positively incentivize judicial performance. On the other hand, if positions with prestige and influence are being distributed by the court's leadership this will also give that leadership significant powers over judges' chances of reappointment.

We have argued that secrecy of votes and opinions do not necessarily protect judges from reappointment decisions motivated by policy preferences. When the appointer lacks the opportunity to update their beliefs about an incumbent judge's policy preferences from the voting records, they are likely to look for out-of-court information, such as the political ideology of the previous appointer. Under public voting, judges have a chance to respond to a government change and protect their reappointment chances by opportunistically catering to the preferences of the new government. Under secret voting, on the other hand, a government change will be an arbitrary factor over which neither opportunistic judges, nor judges with high integrity, have any influence. Only if the court publishes information about performance

will judges have a chance to mitigate the increased risk of replacement that follows from a change in government, by increasing their efforts and boosting their performance.

While secrecy of votes increases judicial independence in terms of being shielded from retrospective policy based accountability, judges may still be held accountable for their past performance on the bench. Our empirical study of the CJEU indicates that the internal organization of the Court provides valuable information to governments regarding the performance – and possible future influence – of incumbent judges. Those who are trusted with important positions within the court, by the President or their peers, are more likely to keep their job when the mandate expires, regardless of the policy concerns of the government. On the other hand, even high performing judges in the CJEU cannot completely counteract the effect of government change. In the case of the CJEU, we found the effect to be quite significant: A government change corresponding to a shift from a social democratic to a conservative prime minister (or vice versa) more than double the probability of an incumbent judge of being replaced.

For performance to matter under secret voting and renewable terms, two conditions that depend on the institutional design of the court must be present: First, past performance must be indicative of future influence. That is, positions cannot be distributed merely based on judges' availability or through pre-determined (e.g. geographical) quotas, but rather be the result of a deliberate choice by other members of the court. In other words, the court has to enjoy autonomy in how it organizes its work and actively use this autonomy when it distributes influence. Second, the information must be available to the appointer. The pre-reform ECtHR serves as an example of a court that only satisfies the first of the two criteria. Although being a court with renewable terms where each case is allocated to a reporting judge in a non-random way, the name of that judge was never published. Thus, while voting decisions were known, the relative influence of judges was not.

In contrast, the large variation in influence among judges of the CJEU is at times both informative and observable. Hence, performance – measured both through judges’ portfolios of cases and leadership positions – has a substantive effect on judges’ likelihood of a reappointment to the CJEU.

This article contributes to long-standing research on how to balance judicial independence, accountability and performance, by focusing on renewable terms as the strongest form of judicial accountability. We have pointed at several significant implications of institutional design choices that impact on the transparency of the performance and policy preferences of judges. Our findings speak to both domestic and international courts that combine *per curiam* decisions with renewable terms and where the appointer is likely to be motivated by policy goals.

## References

- Alter, Karen J. (2008). “Agents or Trustees? International Courts in Their Political Context”. In: *European Journal of International Relations* 14.1, pp. 33–63.
- (2009). *The European Court’s Political Power: Selected Essays*. Oxford University Press. 332 pp.
- (2014). *The New Terrain of International Law: Courts, Politics, Rights*. New Jersey: Princeton University Press. 477 pp.
- Bakker, Ryan et al. (2015). “Measuring Party Positions in Europe: The Chapel Hill Expert Survey Trend File, 1999–2010”. In: *Party Politics* 21.1, pp. 143–152.
- Black, Ryan C. and Ryan J. Owens (2016). “Courting the President: How Circuit Court Judges Alter Their Behavior for Promotion to the Supreme Court”. In: *American Journal of Political Science* 60.1, pp. 30–43.

- Caldeira, Gregory A. and James L. Gibson (1992). “The Etiology of Public Support for the Supreme Court”. In: *American Journal of Political Science* 36.3, p. 635. JSTOR: 2111585.
- Carrubba, Clifford J., Matthew Gabel, and Charles Hankla (2008). “Judicial Behavior under Political Constraints: Evidence from the European Court of Justice”. In: *American Political Science Review* 102.4, pp. 435–452.
- Carrubba, Clifford James (2009). “A Model of the Endogenous Development of Judicial Institutions in Federal and International Systems”. In: *The Journal of Politics* 71.1, pp. 55–69.
- Choi, Stephen J., Mitu Gulati, and Eric A. Posner (2015). “The Role of Competence in Promotions from the Lower Federal Courts”. In: *The Journal of Legal Studies* 44.S1, S107–S131.
- Clark, Tom S. (2009). “The Separation of Powers, Court Curbing, and Judicial Legitimacy”. In: *American Journal of Political Science* 53.4, pp. 971–989.
- Commission of the European Union (2021). *COMMISSION OPINION on the Requests for the Amendment of the Statute of the Court of Justice of the European Union, Presented by the Court*.
- Court, General (2015). *Rules of Procedure (General Court)*.
- Dahl, Robert A. (1957). “Decision-Making in a Democracy: The Supreme Court as a National Policy-Maker Policy Making in a Democracy: The Role of the United States Supreme Court: Role of the Supreme Court Symposium, No. 1”. In: *Journal of Public Law* 6.2, pp. 279–295.
- Döring, Holger and Philip Manow (2018). *Parliaments and Governments Database (ParlGov): Information on Parties, Elections and Cabinets in Modern Democracies*.
- Dumbrovsky, Tomas, Bilyana Petkova, and Marijn Van der Sluis (2014). “Judicial Appointments: The Article 255 TFEU Advisory Panel and Selection Procedures in the Member States”. In: *Common Market Law Review* 51.2, pp. 455–482.

- Dunoff, Jeffrey L. and Mark A. Pollack (2017). "The Judicial Trilemma". In: *American Journal of International Law* 111.2, pp. 225–276.
- Elsig, Manfred and Mark A. Pollack (2014). "Agents, Trustees, and International Courts: The Politics of Judicial Appointment at the World Trade Organization". In: *European Journal of International Relations* 20.2, pp. 391–415.
- Epstein, Lee and Jack Knight (1997). *The Choices Justices Make*. Washington: SAGE. 251 pp.
- Ferejohn, John et al. (1999). "Accountability and Authority: Toward Theory of Political". In: *Democracy, Accountability, and Representation*. Vol. 2. Cambridge: Cambridge University Press, p. 131.
- Ferejohn, John A. and Barry R. Weingast (1992). "A Positive Theory of Statutory Interpretation". In: *International Review of Law and Economics* 12.2, pp. 263–279.
- Gabel, Matthew J. and John D. Huber (2000). "Putting Parties in Their Place: Inferring Party Left-Right Ideological Positions from Party Manifestos Data". In: *American Journal of Political Science* 44.1, p. 94. JSTOR: 2669295.
- Garoupa, Nuno and Tom Ginsburg (2009). "Guarding the Guardians: Judicial Councils and Judicial Independence". In: *The American Journal of Comparative Law* 57.1, pp. 103–134.
- Gibson, James L. and Michael J. Nelson (2015). "Is the U.S. Supreme Court's Legitimacy Grounded in Performance Satisfaction and Ideology?" In: *American Journal of Political Science* 59.1, pp. 162–174.
- Gill, Jeff (2002). *Bayesian Methods - A Social and Behavioral Sciences Approach*. 1st ed. New York: Chapman and Hall/CRC. 480 pp.
- Ginsburg, Tom and James Melton (2015). "Does the Constitutional Amendment Rule Matter at All?: Amendment Cultures and the Challenges of Measuring Amendment Difficulty". In: *International Journal of Constitutional Law* 13.3, pp. 686–713.

- Hayo, Bernd and Stefan Voigt (2007). “Explaining de Facto Judicial Independence”. In: *International Review of Law and Economics* 27.3, pp. 269–290.
- Hermansen, Silje Synnøve Lyder (2020). “Building Legitimacy: Strategic Case Allocations in the Court of Justice of the European Union”. In: *Journal of European Public Policy* 27.8, pp. 1215–1235.
- Justice of the European Communities, Court of (1991). *Rules of Procedure of the Court of Justice of the European Communities of 19 June 1991*.
- Justice of the European Union, Court of (2003). *Amendments to the Rules of Procedure of the Court of Justice Following the Entry into Force of the Treaty of Nice*.
- (2012). *Rules of Procedure*.
- Kelemen, R. Daniel (2011). *Eurolegalism: The Transformation of Law and Regulation in the European Union*. Harvard University Press. 379 pp. Google Books: UPi~~q~~kRr0McoC.
- (2012). “The Political Foundations of Judicial Independence in the European Union”. In: *Journal of European Public Policy* 19.1, pp. 43–58.
- Kenney, Sally J. (1998–1999). “The Members of the Court of Justice of the European Communities”. In: *Columbia Journal of European Law* 5, pp. 101–134.
- La Porta, Rafael et al. (2004). “Judicial Checks and Balances”. In: *Journal of Political Economy* 112.2, p. 26.
- Larsson, Olof and Daniel Naurin (2016). “Judicial Independence and Political Uncertainty: How the Risk of Override Affects the Court of Justice of the EU”. In: *International Organization* 70.2, pp. 377–408.
- Long, J. Scott (1997). *Regression Models for Categorical and Limited Dependent Variables*. 1 edition. Thousand Oaks: SAGE Publications, Inc. 328 pp.

- Melton, James and Tom Ginsburg (2014). “Does De Jure Judicial Independence Really Matter?: A Reevaluation of Explanations for Judicial Independence”. In: *Journal of Law and Courts* 2.2.
- Ministers, Council of (2015). *Regulation (EU, Euratom) 2015/2422 of the European Parliament and of the Council of 16 December 2015 Amending Protocol No 3 on the Statute of the Court of Justice of the European Union*.
- North, Douglass C. and Barry R. Weingast (1989). “Constitutions and Commitment: The Evolution of Institutions Governing Public Choice in Seventeenth-Century England”. In: *The Journal of Economic History* 49.4, pp. 803–832.
- Plummer, Martyn (2003). *JAGS: A Program for Analysis of Bayesian Graphical Models Using Gibbs Sampling*.
- Polk, Jonathan et al. (2017). “Explaining the Salience of Anti-Elitism and Reducing Political Corruption for Political Parties in Europe with the 2014 Chapel Hill Expert Survey Data”. In: *Research & Politics* 4.1, p. 2053168016686915.
- Pollack, Mark A. (2003). *The Engines of European Integration: Delegation, Agency, and Agenda Setting in the EU*. USA: Oxford University Press. 511 pp.
- (2012–2013). “The New EU Legal History: What’s New, What’s Missing Toward a New History of EU Law”. In: *American University International Law Review* 28, pp. 1257–1310.
- Posner, Eric A. and Miguel F. P. de Figueiredo (2005). “Is the International Court of Justice Biased?” In: *The Journal of Legal Studies* 34.2, pp. 599–630.
- Ramseyer, J. Mark (1994). “The Puzzling (In)Dependence of Courts: A Comparative Approach”. In: *The Journal of Legal Studies* 23.2, pp. 721–747.

- Scharpf, F. W. (2010). “The Asymmetry of European Integration, or Why the EU Cannot Be a ‘Social Market Economy’”. In: *Socio-Economic Review* 8.2, pp. 211–250.
- Schmidt, Susanne K. (2018). *The European Court of Justice and the Policy Process: The Shadow of Case Law*. Oxford University Press.
- Shepherd, Joanna M. (2009). “Are Appointed Judges Strategic Too?” In: *Duke Law Journal* 58.7, pp. 1589–1626. JSTOR: 20684765.
- Staton, Jeffrey K. (2010). *Judicial Power and Strategic Communication in Mexico*. Cambridge University Press. 237 pp. Google Books: R1cUsmKA5iIC.
- Stegmueller, Daniel (2013). “How Many Countries for Multilevel Modeling? A Comparison of Frequentist and Bayesian Approaches”. In: *American Journal of Political Science* 57.3, pp. 748–761.
- Stiansen, Øyvind (2022). “(Non)Renewable Terms and Judicial Independence in the European Court of Human Rights”. In: *Journal of Politics* 1.
- Stone Sweet, Alec and Thomas L. Brunell (2013). “Trustee Courts and the Judicialization of International RegimesThe Politics of Majoritarian Activism in the European Convention on Human Rights, the European Union, and the World Trade Organization”. In: *Journal of Law and Courts* 1.1, pp. 61–88.
- Vanberg, Georg (2005). *The Politics of Constitutional Review in Germany*. New York: Cambridge University Press.
- Voeten, Erik (2008). “The Impartiality of International Judges: Evidence from the European Court of Human Rights”. In: *American Political Science Review* 102.4, pp. 417–433.
- Volkens, Andrea et al. (2017). *The Manifesto Data Collection. Manifesto Project (MRG/CMP/MARPOR)*.



# Appendix

## Description of variables

The manifest name of each variable, as used in the main text of this paper, is reported in bold. The variable name, as used in the data set, is reported in parentheses. The unit of observation is individual terms in office.

**Replacement** (`exit`) is a binary variable reporting whether a judge remains in office for another term (0) or exits the Court (1). We consider cases where the **ExitCause** is set to "Mandate expired". The information is collected from the announcement published in the *Official Journal of the European Union* that member states have appointed a new judge.

**Preference distance - economic issues** (`|FreeEconomy - FreeEconomy_ren|`)

is a continuous variable reporting the absolute difference between the economic preferences of a judge's appointing and reappointing governments respectively. The government's economic preferences are calculated in the following way:

In the first step, we identify the government in power (`DecisionCabinet` and `DecisionCabinetExit`) at the time of the appointment decision (`DecisionDate` and `DecisionDateExit`) using the "Cabinet" data provided by the ParlGov Project (Döring and Manow, 2018).

In the second step, we identify the parties in government using the "Parties" data (ibid.) and link these to the manifesto data provided by the Manifesto Project (Volkens et al., 2017).

In the third step, we estimate the economic preferences expressed in all party manifestoes using the vanilla method (M. J. Gabel and Huber, 2000). The indicators we rely on are questions related to the economic preferences of parties ("per401", "per402", "per403", "per404", "per405", "per406", "per409", "per410", "per412", "per413", "per414", "per415" and "per416").

- “per401”: Favourable mentions of the free market and free market capitalism as an economic model.
- “per402”: Favourable mentions of supply side oriented economic policies (assistance to businesses rather than consumers).
- “per403”: Support for policies designed to create a fair and open economic market.
- “per404”: Favourable mentions of long-standing economic planning by the government.
- “per405”: Favourable mentions of cooperation of government, employers, and trade unions simultaneously. The collaboration of employers and employee organisations in overall economic planning supervised by the state.
- “per406”: Favourable mentions of extending or maintaining the protection of internal markets (by the manifesto or other countries).
- “per409”: Favourable mentions of demand side oriented economic policies (assistance to consumers rather than businesses).
- “per410”: The paradigm of economic growth.
- “per412”: Support for direct government control of economy.
- “per413”: Favourable mentions of government ownership of industries, either partial or complete; calls for keeping nationalised industries in state hand or nationalising currently private industries. May also include favourable mentions of government ownership of land.
- “per414”: Need for economically healthy government policy making.

In the fourth step, we average the estimated preferences over all parties in government and link this to the appointing and reappointing gov-

ernments in the main data frame. In the main text these preferences are weighted by the party size of each government coalition partner in parliament. To check the robustness of these results we also calculate an alternative variable where preferences are not weighted.

Last, the preference distance between successive governments is calculated during the estimation of models as the absolute difference between two ideal-points.

When information is lacking, we impute the variable with two variables drawn from the Chapel Hill Expert Survey trend file (Bakker et al., 2015; Polk et al., 2017).

- *Economic left-right preferences* (`lrecon`) ranges from 0 (extreme left) to 10 (extreme right) and classifies parties in terms of their ideological stance on economic issues. Parties on the economic left want government to play an active role in the economy. Parties on the economic right emphasize a reduced economic role for government: privatization, lower taxes, less regulation, less government spending, and a leaner welfare state.
- *Economic deregulation* (`deregulation`) ranges from 0 (strongly favors deregulation) to 10 (strongly opposes deregulation) and classifies parties in terms of their position on deregulation.

#### **Preference distance - integration issues (`|Integration - Integration_renl|`)**

is a continuous variable calculated in a similar way as the previous measure. We rely on four indicators relating to international and European integration from the manifesto data: "per107", "per108", "per109" and "per110".

- "per107": Need for international co-operation.
- "per108": Favourable mentions of European Community/Union in general.

- “per109”: Negative references to international co-operation.
- “per110”: Negative references to the European Community/Union.

When information is lacking, we impute the variable with two variables drawn from the Chapel Hill Expert Survey trend file (Bakker et al., 2015; Polk et al., 2017).

- *European integration* (**position**) ranges from 0 (strongly opposes) to 7 (strongly favors) and reports the overall orientation of the party leadership towards European integration
- *European Parliament powers* (**eu\_ep**) ranges from 0 (strongly opposes) to 7 (strongly favors) and reports the overall orientation of the position of the party leadership on the powers of the European Parliament.

**Preference distance - general left-right issues** (`|rile - rile_ren|`) is drawn from the same data (Volgens et al., 2017) and reports the absolute distance on the standard right-left scale provided by the Manifesto Project. The rile score ranges from 0 to 100. When information is lacking, we impute the variable with two variables drawn from the Chapel Hill Expert Survey trend file (Bakker et al., 2015; Polk et al., 2017).

- *General left-right* (**lrgen**) ranges from 0 (extreme left) to 10 (extreme right) and reports the overall ideological stance of the party leadership.
- *Redistribution preferences* (**redistribution**) ranges from 0 (strongly favors redistribution) to 10 (strongly opposes redistribution) and classifies parties in terms of their position on redistribution of wealth from the rich to the poor.

**Performance - managing cases of legal interest** (`Annotations.mean - Annotations.court.mean`) is a continuous measure of the mean num-

ber of legal annotations that cases generated in which the judge acted as a Reporting Judge during their mandate. It is centered on the mean number of annotations generated by cases delivered by the Court in the same period. The variable thereby reports the difference between the actual performance of the judge and what the government could reasonably expect. Information is collected and provided by the CJEU itself.

**Performance - leadership positions** (`Chamber5President`) reports whether the judge had presided over deliberations in chambers with five or more judges during their mandate. The information is collected from the text of the judgments. In the analysis, the variable is contrasted with a subset of observations reporting all instances that have occurred after the reform in 2003, labelled (*Performance - selective leadership positions*).

**Length of tenure** (`Tenure`) is a continuous variable. It reports the cumulated sum of all the judge's mandates (`TermLength`).

**Age** (`AgeExit`) is a continuous variable reporting the difference between the judge's birth date (`Birth`) and the date of the member states' reappointment decision (`DecisionDateExit`). Information is collected from the judge's CVs, their profile at the Court's website ([www.curia.eu](http://www.curia.eu)) and the announcement of the next judge's appointment published in the *Official Journal of the European Union*.

**Change in attendance** (`Attendance_diff - Attendance_diff.median`) is a continuous variable. It reports the difference in the count number of deliberations a judge has attended in the year preceding the government's reappointment decision. For comparability, we normalize the measure by subtracting the median change in attendance in the same period.

## Description of the statistical model

All models are Bayesian, estimated using MCMC in JAGS (Plummer, 2003). The Bayesian approach has two compelling features: First, unobserved information can be estimated in parallel to the main regression, possibly by inserting information from other sources (Gill, 2002, p. 43). Second, Bayesian models also provide a more robust approach to multilevel non-linear regressions (such as our logistic regressions), in particular when the number of countries (or time periods) is low (Stegmueller, 2013). While the results reported in the article are produced by a pooled model, we verify the results in this appendix using random intercepts.

In the following, we exemplify with the estimation of the model using economic left-right preferences and performance measured as portfolio of cases of interest to the academic community. It is a binomial logistic regression:

$$\begin{aligned}
 Pr(Replacement_i = 1) &\sim Bernoulli(\pi_i) \\
 logit(\pi_i) &= \alpha + \\
 &\quad \beta_1 \times |Preferences_{gvt\ 1_i} - Preferences_{gvt\ 2_i}| + \\
 &\quad \beta_2 \times (Annotations_i - Avg. Annotations_i) + \\
 &\quad \beta_3 \times Change\ in\ Prime\ Minister_i + \\
 &\quad \beta_4 \times (Vice-)President + \\
 &\quad \beta_5 \times Tenure_i - mean(Tenure) + \\
 &\quad \beta_6 \times Age_i - mean(Age) + \\
 &\quad \beta_7 \times Change\ in\ Attendance_i
 \end{aligned} \tag{2}$$

The regression parameters have relatively vague prior distributions:

$$\begin{aligned}
\alpha &\sim N(0, 10) \\
\beta_k &\sim N(0, 10) \\
\gamma_k &\sim N(0, 10) \\
\delta_k &\sim N(0, 10)
\end{aligned} \tag{3}$$

We let the model run through 500 iterations in adaptation mode to set the Monte Carlo step size and another 5 000 iterations of burn-in before starting the sampling. We then sample every 10<sup>th</sup> iteration for the next 10 000 simulations. The final sample shows no signs of no-convergence.

### Imputation of missing information

Most variables contain no or few missing observations. They are therefore treated as randomly missing and imputed through priors informed by the mean and standard deviation among the observed units:

$$\begin{aligned}
Economic\ left - right_i &\sim N(\mu^{Economic\ left-right}, \tau^{Economic\ left-right}) \\
Economic\ deregulation_i &\sim N(\mu^{Economic\ deregulation}, \tau^{Economic\ deregulation}) \\
Annotations_i &\sim N(\mu^{Annotations}, \tau^{Annotations}) \\
Avg.\ Annotations_i &\sim N(\mu^{Avg.\ Annotations}, \tau^{Avg.\ Annotations}) \\
Tenure_i &\sim N(\mu^{Tenure}, \tau^{Tenure}) \\
Age_i &\sim N(\mu^{Age}, \tau^{Age})
\end{aligned} \tag{4}$$

Most of the missing observations come from our estimation of political preferences. Thus, we sometimes lack information on the party positions of the appointing (7 observations, 3%) and the reappointing governments on economic issues (9 observations, 4%). In these cases, we impute values relying on information provided by the Chapel Hill expert survey trend file (Bakker et al., 2015; Polk et al., 2017). While only available from 1998, this supplementary data source nevertheless covers a substantial part of the

missing observations (43% and 56% respectively). Missing information on government preferences is thus imputed through an ordinary linear model:

$$\begin{aligned}
\text{Preferences gvt } 1_i &\sim N(\mu 1_i, \tau_i) \\
\mu 1_i &= \gamma 1_1 + \\
&\quad \gamma 1_1 \times \text{Economic left} - \text{right} 1_i + \\
&\quad \gamma 1_2 \times \text{Economic deregulation} 1_i
\end{aligned} \tag{5}$$

$$\begin{aligned}
\text{Preferences gvt } 2_i &\sim N(\mu 2_i, \tau_i) \\
\mu 2_i &= \gamma 2_1 + \\
&\quad \gamma 2_1 \times \text{Economic left} - \text{right} 2_i + \\
&\quad \gamma 2_2 \times \text{Economic deregulation} 2_i
\end{aligned} \tag{6}$$

Results from these secondary regression models are reported in table 5, while the two columns in table 6 compare results from the main model when preferences are imputed to using listwise exclusion. In comparison to a listwise exclusion, the imputation allows us to estimate all remaining variables on the universe of cases. It also avoids making the assumption that missingness is independent of values on other covariates in the model. Results remain largely similar across models.

Predictors	Model 1
Intercept (appointing gvt)	-1.05 (-2.22,1.18)
Economic left-right preferences (appointing gvt)	-0.14 (-0.2,-0.06)
Economic deregulation (appointing gvt)	0.23 (0.18,0.28)
Intercept (reappointing gvt)	-0.33 (-4.91,2.54)
Economic left-right preferences (reappointing gvt)	-0.13 (-0.21,-0.03)
Economic deregulation (reappointing gvt)	0.18 (0.13,0.24)
Number of observations	151
Median effects with 90% symmetric posterior density interval in parenthesis.	

Table 5: Results from secondary models imputing values of preferences.



Predictors	w/imputations	listwise exclusion
Intercept	-0.27 (-1.76,1.32)	-0.2 (-1.74,1.33)
Preference distance (economic issues)	1.41 (0.44,2.35)	1.43 (0.49,2.42)
Performance (court interest)	-1.88 (-3.56,-0.36)	-1.97 (-3.6,-0.38)
Change of PM	0.05 (-0.83,0.88)	-0.02 (-0.87,0.84)
(Vice-)President	0.67 (-0.81,2.12)	0.57 (-1.08,2.08)
Length of tenure	0.13 (0.02,0.24)	0.11 (0,0.23)
Age	0.17 (0.11,0.24)	0.17 (0.1,0.25)
Change in attendance	-0.02 (-0.04,0)	-0.02 (-0.04,0)
Number of observations	151	146
Proportion of correct predictions	0.77	0.77
... correct positive predictions	0.71	0.75
... correct negative predictions	0.79	0.78

Median effects with 90% symmetric posterior density interval in parenthesis.

Table 6: Results from a logistic regression with imputations of preferences and with listwise exclusion respectively.

## Alternative operationalizations of political preferences

In this subsection, we verify whether the effects reported in the main article hold for different operationalizations of preferences. For brevity, we focus on the results relating to the Court of Justice.

We measure preferences along three different dimensions; economic left-right, general left-right and support for European integration. The results are reported in Table 7. Figure 5 furthermore illustrates the standardized coefficients for an explicit comparison of effects across preference dimensions.

The first model supports the proposition that a shift in economic preferences between the initial and the present appointer increases the chances of replacing a judge. The second model can be considered as a robustness check. It reports the effect of changes along the broader left-right cleavage present in European politics. As is apparent from the comparison in Figure

5, the direction is similar to that of economic left-right, although the effect is more moderate. This is unsurprising, as the measure includes a number of issues less relevant to CJEU case law. This might evolve in the future. The broader the scope of issues solved at the supranational level, the larger the overlap between the domestic debate and the Court’s output.

In contrast, the third model gives no support for a similar logic in questions of European integration. The previous academic debate has focused extensively on the cleavage between the institutional interests of the Court and the member states (Pollack, 2012–2013). Yet, these results indicate that governments prioritize substantive economic left-right issues rather than national sovereignty in their selection of judges to the CJEU.

Dependent variable: 'Replacement'	model 1a	model 1a	model 1a
Intercept	-0.27 (-1.76,1.32)	-0.4 (-1.9,1.21)	-0.21 (-1.67,1.35)
Preference distance (economic left-right)	1.41 (0.44,2.35)		
Preference distance (general left-right)		0.02 (-0.01,0.06)	
Preference distance (European integration)			-0.01 (-1.77,1.66)
Performance (cases of court interest)	-1.88 (-3.56,-0.36)	-1.57 (-3.28,-0.1)	-1.5 (-3.15,-0.07)
Change of PM party	0.05 (-0.83,0.88)	0.24 (-0.61,1.06)	0.34 (-0.43,1.14)
(Vice-)President	0.67 (-0.81,2.12)	0.61 (-0.81,2.01)	0.47 (-0.99,1.95)
Length of tenure	0.13 (0.02,0.24)	0.11 (0.01,0.23)	0.13 (0.02,0.24)
Age	0.17 (0.11,0.24)	0.16 (0.1,0.23)	0.16 (0.1,0.23)
Change in attendance	-0.02 (-0.04,0)	-0.02 (-0.04,0)	-0.02 (-0.04,0)
Number of observations	151	151	151
Proportion of correct predictions	0.77	0.78	0.77
... correct positive predictions	0.71	0.74	0.74
... correct negative predictions	0.79	0.8	0.78

Median effects with 90% symmetric posterior density interval in parenthesis.

Table 7: The models explore DIFFERENT OPERATIONALIZATIONS of PREFERENCES.

## Are the effects constant over time?

The results reported in the main article are estimated on a pooled data set. However, the dynamic of reappointments may vary over time and institutions;

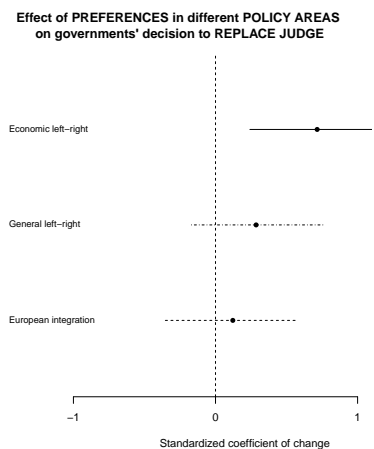


Figure 5: Different operationalizations of government preferences.

both national and supranational. We start by exploring different ways of treating the data as time-series cross-sectional, before we explore whether the selection criteria changed after the introduction of the expert advisory panel in 2010 (the Article 255 committee).

### Has there been a trend over time?

Figure 6 illustrates the proportion of replacement decisions following the expiration of mandate over time. Although we can distinguish a certain drop from the earliest period, the overall propensity to replace incumbents has remained low. Due to increases in the membership following successive EU enlargements, most of the observations were made in the period after 1990 (79%). The results will therefore be driven by selection in the last three decades.

To verify that the selection criteria have not changed substantially over time, we divide the data into three separate time periods and estimate separate models. As is clear from the results reported in Table 8, the effects remain essentially the same over the seven decades.

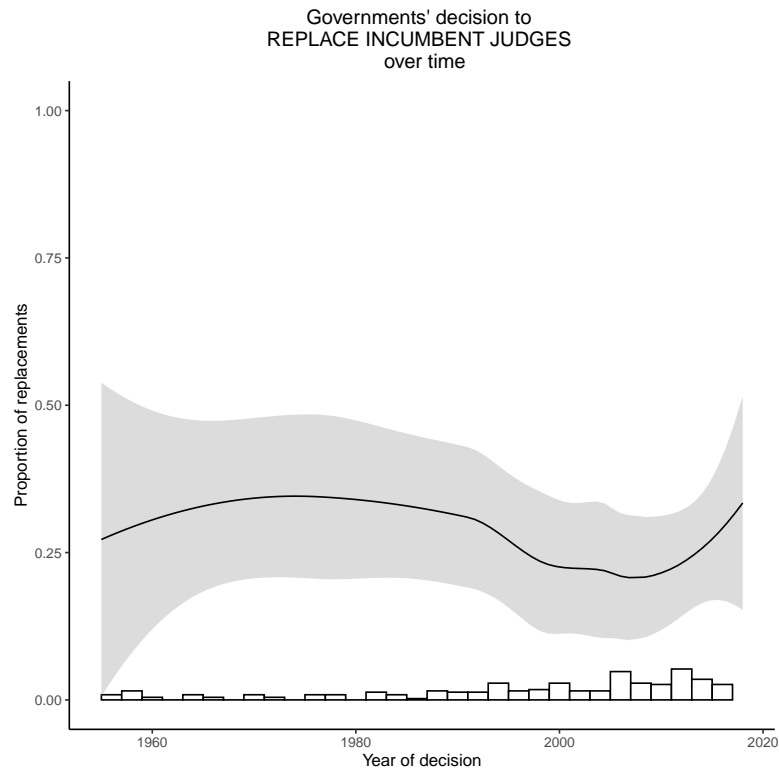


Figure 6: Governments' decision to replace an incumbent judge has remained relatively low (Locally estimated scatter plot smoothing overlaid by a histogram showing the number of observations over time).

Dependent variable: 'Replacement'	1952-1980	1980-2000	2000-2020
Intercept	-0.46 (-2.75,1.97)	-0.27 (-3.62,2.81)	-0.95 (-3.2,1.22)
Preference distance (economic issues)	2.94 (1.29,4.88)	2.32 (0.75,4.2)	0.38 (-1.38,2.06)
Performance (cases of court interest)	-1.43 (-3.89,0.92)	-1.92 (-5.06,1.31)	-2.58 (-5,-0.31)
Change of PM party	-1 (-2.73,0.49)	-0.37 (-2.32,1.41)	2 (0.55,3.59)
(Vice-)President	0.3 (-2.3,2.97)	-0.84 (-3.91,2.03)	0.42 (-2.29,2.9)
Length of tenure	0.14 (-0.04,0.34)	0.28 (0.03,0.63)	0.1 (-0.08,0.28)
Age	0.3 (0.16,0.48)	0.18 (0.06,0.33)	0.13 (0.02,0.24)
Change in attendance	-0.06 (-0.13,0.01)	-0.03 (-0.08,0.02)	-0.04 (-0.06,-0.01)
Number of observations	51	38	78
Proportion of correct predictions	0.86	0.82	0.81
... correct positive predictions	0.78	0.77	0.83
... correct negative predictions	0.91	0.84	0.8

Median effects with 90% symmetric posterior density interval in parenthesis.

Table 8: Replacement of judge in the Court of Justice where DIFFERENT TIME PERIODS are analyzed separately.

### Did the introduction of an advisory panel change the selection criteria?

In 2010, an advisory selection panel was set up at the supranational level to help member states screen nominees for the position as a judge. The panel has been impressively active and has advised to reject a number of candidacies at each round of nominations. The introduction of the panel is part of a broader movement towards opening the highly politicized process of judicial appointments to scrutiny, but has profiled itself mainly as an additional guarantee for the appointees' independence, legal expertise and professional experience. While its activism has caused scholars to warn against judicial self-government, it remains unclear to what extent it can constrain governments' retention choices (Dumbrovsky, Petkova, and Van der Sluis, 2014). The panel has been particularly critical towards first-time nominations as well as to appointments suggested for the General Court.

Here, we test whether the selection criteria identified in the main article have changed substantially following the panel's introduction through a set of interaction effects. Table 9 reports the change in the effect of governments'

preferences after 2010. As is evident from both the direction and the precision of these interactions, there is no consistent evidence of any alteration in governments' retention choices that affect the generality of our argument.

Dependent variable: 'Replacement'	CJ	GC
Intercept	0.06	-3.28
	(-1.62,1.76)	(-5.06,-1.71)
Preference distance (economic issues)	1.21	1.53
	(0.16,2.24)	(-0.02,3.15)
255 Panel * preference distance	0.68	1.48
	(-1.83,3.39)	(-1.21,4.14)
255 Panel	-1.34	0.97
	(-3.97,1.36)	(-0.78,2.84)
Performance (court interest)	-1.92	0.57
	(-3.7,-0.23)	(-0.77,1.86)
255 Panel * performance	-0.31	-0.67
	(-3.18,2.33)	(-2.11,0.75)
Change of PM	0.25	0.86
	(-0.63,1.12)	(-0.13,1.89)
(Vice-)President	0.44	-3.68
	(-1.07,1.91)	(-7.11,-0.92)
Length of tenure	0.17	0.11
	(0.06,0.3)	(-0.01,0.23)
Age	0.19	0.1
	(0.12,0.26)	(0.03,0.18)
Change in attendance	-0.01	0.03
	(-0.04,0.01)	(-0.01,0.06)
Number of observations	151	97
Proportion of correct predictions	0.77	0.77
... correct positive predictions	0.71	0.77
... correct negative predictions	0.8	0.77

Median effects with 90% symmetric posterior density interval in parenthesis.

Table 9: Selection criteria before and after the introduction of an ADVISORY SELECTION COMMITTEE in 2010.