Holding Judges To Account. Policy, Performance and Impartiality

Silje Synnøve Lyder Hermansen
Daniel Naurin

December 20, 2021

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 attention has been given to judicial accountability. Researchers have demonstrated 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) as well as by concerns about shoring up public support in power struggles with executives and legislators (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 accountability. How—if at all—can judges with de facto law-making powers

be held accountable? More important still, how do we ensure representation in a non-majoritarian institution for which independence is key?

We approach the question by studying the most significant mechanism for judicial accountability that acceptable in democratic rule-of-law systems – renewable terms. Setting limits on judges' terms in office, and letting an appointer decide whether they should be allowed back in, creates strong incentives for judges to respond to that appointer's desires. This is well understood by designers of judiciaries around the world. Although the varieties of ways in which states select judges is striking (Garoupa and Ginsburg, 2009), few have opted for renewable terms. Putting judges up for regular re-appointments has often been seen as one step too far from judicial independence (Feldt and Voigt 2003).

Renewable terms are nevertheless common in two contexts. 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 re-appointments. 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 other context where renewable terms are common, but where less research exists, is international courts (international courts). In 2014, 17 out of 24 active international courts had renewable terms (Squatrito 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 as collective bodies adjust their decisions in anticipation of political reactions (Clifford J. Carrubba, M. Gabel, and Hankla, 2008; Larsson and Naurin, 2016). However, we have little empirical knowledge on how renewable terms are used to exercise individual-level accountability.

This dearth of research is particularly striking considering the string of reforms recommended to several international judiciaries precisely to increase judges' independence and competence. Thus, the European Court of Human Rights has transitioned to a single non-renewable term among judges (2010) in an attempt to decrease members' political accountability. Similarly, judges at both the Inter-American Court for Human Rights and the African Court of Justice and Human Rights face term limits. In a similar vein, several courts have expanded their appointment procedures to include supranational selection panels with the intention to ensure judges' competence (e.g. Economic Community of West African States (2006), the Inter-American Court of Human Rights and the Court of Justice of the European Union (2010), International Criminal Court (2013)). These moves towards supranational checks on member states' nomination are often motivated by criticisms that appointmens are overly politicized and a mutual suspicion that governments are either unwilling or unable to select judges with the appropriate competence level (Pérez, 2015).

We make two general propositions relating to appointers' motivation and opportunity when considering incumbent judges for re-appointment. First, incumbent judges are evaluated both with respect to the direction of their preferences and their proven ability to influence case law. While extant research has focused mainly on the direction of judges' decision, we highlight that policy-oriented governments necessarily care about impact. If competence leads judges to positions of power, then even politically driven appointers will care about legal excellence.

Second, representation can be obtained either by selecting new representatives with the desired qualities or influencing the behavior of those already in office through supervision and sanctions. We argue that judicial selection is a combination of both. Judicial appointments aim to select judges for the future. However, information about past behavior may inform that decision. This is how accountability occurs (J. Ferejohn et al., 1999). Principal-agent

theory assumes that agents invest in the performance criteria that are visible, and shirk on the ones that are not (Holmstrom and Milgrom, 1991). We take the flip-side assumption that the principal acts on the information she has access to. Courts can be designed to provide or withhold information about judges behavior. The focus in the debate about international courts has been on the transparency of the court's organization. We add to this by arguing that the appointer also considers the quality of the signals it receives.

Extant research emphasizes the importance of the direction of judges policy choices. The strong anticipatory effect found in studies of U.S. state courts, where votes and separate opinions are public, is because it is easy for the appointer to monitor the direction of the case law that judges pursue. International courts are often designed with that effect in mind. Although there are examples of international courts with both renewable terms and public votes and opinions (such as the International Court of Justice and the International Tribunal for the Law of the Sea), renewable terms are commonly accompanied by restrictions in transparency (Dunoff and Pollack, 2017). Thus, international courts such as the Court of Justice of the EU (CJEU) and the WTO Appellate Body have renewable terms, but the information on votes and opinions is restriced. Governments consequently have the opportunity to regularly select judges they believe will represent their views. However, judges are not accountable insofar as their voting has no bearing on their future on the bench.

In contrast, accountability is absent from the debate on judges' legal skills and integrity. The narrow focus on selection panels in the current debate seems to indicate that such qualities can only be obtained through an initial screening. Yet students of political agency have repeatedly shown that term limits have important side effects. They increase representatives' corruption (Ferraz and Finan, 2011), decrease their economic performance (Besley and Case, 1995) and legislators exert overall less effort in their last term (Bailer and Ohmura, 2017). These side effects may be masked to the

occasional observer, however, since representatives in their last term are more experienced, and thus more efficient, than motivated newcomers (Alt, Bueno de Mesquita, and Rose, 2011).

To make our argument, we study the CJEU. We leverage two of its design features. On the one hand, the Court represents an ideal-type institution where appointments are political but judges' individual policy choices remain secret. On the other hand, the Court is exceptionally transparent with respect to judges' roles in the policy-making process. This is important information for influence-seeking appointers. However, the quality of those signals varies. Sometimes positions of influence are allocated through (peer-)selection and sometimes as a matter of procedure.

The CJEU is often described as an exceptionally independent international court. Yet the procedure for judicial selection ties governments closely to their judges. All governments nominate one judge each to the higher-level Court of Justice and two judges to the lower-level General Court. Until the establishment of a supranational advisory selection committee in 2010, governments were de facto free to appoint the judges of their choice. Judges' terms can furthermore be renewed indefinitely. Yet no research exists on how governments use these appointments to influence the Court (Kelemen, 2012, p. 50). To mend the gap, we analyze all 247 potential retention decisions by member states over the entire period of the Court's existence (1952-2019). Our analysis generates two key findings with general application for our understanding of judicial accountability. Although we focus on the international level, our theoretical insights apply also to domestic courts (Staton and Moore, 2011).

First, with respect to the direction of policy preferences, we find that sheltering judges from oversight by reducing transparency on votes and opinions does not protect them from politically motivated re-appointment choices. The lack of information reduces the opportunity for retrospective accountability, but the appointer may still engage in prospective selection based on

other (possibly less precise) information about the incumbent's policy preferences.

Second, we find that appointers act on available and high-quality information relating to the potential impact of individual judges. Positions of influence allocated selectively within the Court provides high-quality signals about judges potential future performance, and governments will factor this into their appointment choices. In contrast, positions of influence that are allocated as a matter of procedure are uninformative signals that governments disregard. While extant research has focused on secret voting, we demonstrate that renewable terms still generate retrospective accountability along other dimensions, as long as governments are able to discriminate between incumbents. One implication is that even if the appointer is primarily interested in seeing her policy preferences implemented, when the proper transparency rules are in place, renewable terms may steer towards an increase in the overall performance of the court.

Our findings contribute to ongoing debates about how to balance judicial independence and accountability, in particular by emphasizing the informational aspects of accountability. The value that governments put on performance provides an additional motivation for the increase in expert selection panels at the international level. These panels are not merely measures for mutual constraints and surveillance, but may also help governments select better judges in the absence of reliable information. However, courts can also provide information directly to governments. This requires (i) a certain autonomy in the court's internal organization and (ii) that the court uses that autonomy to selectively distribute influence.

The results also speak 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 findings indicate that states do not necessarily trade one against the other: Even when judicial accountability is

6

motivated by policy goals, selecting highly competent judges with a strong 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 makes legislative and constitutional commitments less credible. 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 usually thought of as roughly antonymous; more accountability means less independence (Cross 2008: 566). Institutional designers of judiciaries around the world are therefore 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 somewhat mixed results (Melton and Ginsburg, 2014; Hayo and Voigt, 2007). To the extent that rules on paper matter at all for real-life 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 formulated "the judicial trilemma". Out of three values that institutional designers may want to promote – indepen-

dence, accountability and transparency – they have to forego one. Judicial accountability, in this account, is understood specifically as judges having renewable terms, with states controlling their re-appointments (Dunoff and Pollack, 2017, p. 234). According to the logic of the trilemma, judicial independence and accountability may be compatible, but 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 take the court's case law, they will be independent, as they dont' have to fear sanctions based on the substance of their decisions.

The judicial trilemma, as a theoretical construct, is clearly helpful in pointing at how the interaction of 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. How – if at all – do states use the sanctioning mechanism that they have given themselves, if they are unable to follow what a judge has been doing on the bench?

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

What motivates the (re)appointment of judges?

Research on the selection of judges is sparse compared to the voluminous literature on judicial behavior and has mostly focused on the U.S.. Studies of

the Senate's confirmation of justices to the 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 the actual motivation and behavior of the (re)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 of 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 re-election 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 tend to 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 the appointments of judges to these courts. The principal-agent model expects that screening and nomination of candidates is infused with strategic political considerations concerning candidates' policy prefer-

9

ences (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 Organizations 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 (Posner and de Figueiredo, 2005). On the other hand, one important study of the European Court of Human Rights found large variations in judges' tendency to defer to their home states (Voeten, 2008, p. 417).

The trustee model emphasizes that politics unfold differently in courts compared to power struggles outside the court. Influence is obtained through legal arguments, so that impartiality and legal skills are central tools for trade. The model expects 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). 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. The main criticism of the model pertains to governments' ability to credibly commit to such ideals without institutional constraints. While they may wish to have an impartial judiciary, they also face incentives to shirk on that commitment by seeking influence over their own judge. Pertaining to the appointment of judges, proponents of the trustee model point out that the debate is irrelevant. 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 (Alter 2008, p. 46, see also Kelemen 2012). States are thus unlikely to bother using their (re)appointment powers to promote policy goals.

We conclude from these studies that two primary selection criteria are present in the minds of appointers; ideological or policy proximity and professional competence. However, we do not think that the two are necessarily contradictory. 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 representatives of any policy goals.

This means that we give professional competence as a selection criterion a different interpretation than the previous literature, which has focused on electoral interests or concerns for quality and impartiality. We also deviate from the trustee model's proposition that states are unwilling to make use of the influence they have – even if marginal for a single state – by appointing a judge that is willing and able to promote policy goals that proximate those of the appointer. Thus, 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. To the contrary, any appointer – whether at the domestic or international level – would prefer a candidate who is both ideologically committed to its agenda and possesses the authority to persuade other judges about the direction of the case law.

In the next section, we contextualize our argument in the case of the European Union, before specifying our theoretical propositions and hypotheses.

The Court of Justice of the European Union

Surprisingly, in spite of its elevated status as the motor of the much-debated judicialization of politics in Europe (Pollack, 2003; Kelemen, 2011; Schmidt, 2018), "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).

Yet the Court has served as a model for a slew of other international courts. Its institutional features determine who the appointers are and the information available to them. These rules structure how selection criteria are expressed.

Judicial appointments are done in a domestic political context

When the number of judges on an international court is smaller than the number of member states, the appoinment process involves political bargaining between coalitions of member states (Elsig and Pollack, 2014). In contrast, in full-representation courts like the CJEU, the selection of each judge results from the political choice of a single government. All EU member states nominate one judge each to the CJEU's highest formation (the "Court of Justice"), and two judges each to the lower formation (the "General Court"). All judges are subsequently appointed by consensus by the 27 governments. This latter stage is a mere formality, however (Dunoff and Pollack, 2017).

Member-state governments have cautiously held on to their prerogative. The General Court recently received more human resources to address its increasing case-load. In the lead-up to the re-haul in 2015, the Court and the European Commission suggested increasing the number of judges by 12 and supplement any further need with legal clerks. This suggestion was bluntly rejected by the member states, which instead preferred the more costly option to double the number of judges and thus keep their privilege intact (Commission of the European Union, 2021; *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* 2015).

Since 2010, the procedure also includes a merit selection committee, composed of high-ranking national judges and former CJEU judges, which gives

an opinion on the qualifications of the member states' nominees before they are appointed (Dumbrovsky, Petkova, and Van der Sluis, 2014). Selection panels can be seen as agents set up to ensure judges' legal skills because governments consider themselves unable to acquire this information on their own. Although a revision of the selection procedure to the CJEU had been in discussion for a long time, the consultative panel finally saw the light of day following the 2004 EU enlargement. However, it was motivated mainly by member states' wish to ensure the quality of the *other* governments' choice rather than by a frustration over their own appointments. As we will see, judicial performance had been a common selection criterion already since the outset.

The key role of single governments also determines how we conceptualize policy preferences. Judges have six-year renewable terms, and it is up to each government to decide whether to re-nominate their judge or not. The political interaction implied in an appointment therefore takes place domestically between successive governments. The most relevant policy dimension would therefore be the same as the one dividing alternative government formations in domestic politics.

One of very few articles that addresses the selection of CJEU judges directly, concludes 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). This makes sense, given the institutional position of the Court at the top of a supranational judicial hierarchy. The Court's core task is to ensure a reasonably uniform application of European laws across the Union. In case of a conflict between EU and national legislations, the Court has the competence to strike down domestic laws; effectively thwarting government policies.

However, the Court's case law also impede on the most salient policy dimension in domestic politics: the economic left-right dimension. The CJEU has been described as a driver of economic integration in Europe (Pollack,

2003; Conant, 2007; Alter, 2009). Its case law has often raised controversy. For example, while some observers have perceived the strengthening of individual citizenship rights as a progressive step towards embedded liberalism and social rights at the international level (Caporaso and Tarrow 2009), others have lamented the "Hayekian" bias inherent in the Court's case law (Scharpf 2010, Hopner and Schafer 2012, Schmidt 2018). Political divisions in the Court's surroundings spill over to the Court's own deliberations. Previous research has shown that when governments express disagreement among themselves, the Court's leadership distribute influence among judges following the economic left-right spectrum (Hermansen, 2020).

The information provided by the Court

According to the logic of the judicial trilemma, the CJEU has sacrificed transparency to the benefit of independence because judges' voting is secret (Dunoff and Pollack, 2017, p. 247). However, this only limits judges' *political* accountability. In contrast, the Court is exceptionally generous with information on the judges' positions of influence on the bench. This speaks to judges past *performance.* We emphasize here two organizational features that are potentially important for judicial accountability even in the absence of information on votes and opinions. First, division of labor within the Court puts single judges center-stage. Second, the allocation of these key positions is sometimes selective, while other times it is a question of procedure.

The division of labor within the Court means that some judges yield more influence than others. We consider two types of positions.

First, in contrast to some other international courts (e.g. the European Court of Human Rights and the Inter American Court of Human Rights) the CJEU publishes information on the identity of the opinion writer. This is the "judge-rapporteur"; the agenda setter in each case. The rapporteur is a rank-and-file judge delegated with the responsibility to manage the case on behalf of the other judges. The rapporteur proposes to the other judges whether a

14

case is important enough to include an opinion of an Advocate General, 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 judge rapporteur is in charge of collecting information and presenting the case to the other decision makers. The rapporteur pens the first version of the judgment – which forms the basis of the deliberations – as well as the final version. The responsibility applies even if the judge rapporteur was in the minority. This means that states cannot be sure that the preferences of the rapporteur are reflected in the decision, but having this role in many important cases, communicates nevertheless that the judge often find him/herself in a position of influence.

Second, there is a top-level and mid-level management on the Court. The top-level leaders – the presidents and vice-presidents – manage the Court's day-to-day life and preside over grand chamber cases. They distribute influence and monitor the judges. The President of the Court of Justice – assisted by their vice-president – assigns cases to individual judges, while the President of the General Court assigns cases to chambers. The middle-management 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 lower-level court, they also select the rapporteur. These positions are prestigious and often seen as stepping-stones to higher offices. However, these influential positions are sometimes allocated selectively, and sometimes not. First, the rapporteur is always appointed on a case-to-case basis, but there is a difference in how this appointment is done within the CJEU. Report allocation in the higher-level Court of Justice follows a logic of selection. The President of the Court makes his choice early in the process and the choice of the chamber follows from that initial decision. The President has few restrictions in his selection of the rapporteur. 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 rapporteur is appointed, the President leaves the monitoring to the checks and balances of the chamber deliberations (Justice of the European Union, 2012, art. 15). The rapporteur enjoys a large autonomy in their case management. 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 rapporteur of many important cases distinguishes an influential judge from their peers. Moreover, since these positions are selectively allocated, they are also an indication of a judge's future performance, since the leadership's allocation criteria remain consistent over time.

In contrast, report allocation in the lower-level General Court follows a logic of monitoring. Cases are allocated to chambers immediately after their filing. The rapporteur is then appointed by the chamber president among the more restricted number of judges at their disposal (art. 26 Court, 2015). In return for this expediency, the upper leadership is kept informed throughout the deliberation. Once the rapporteur has a final draft of the judgment, it is communicated to the vice president. Their assessment is then returned to the rapporteur with the rest of the affected judges in copy. In short, while the rapporteur plays a central role also in the General Court, they are treated as inter-changeable.

Second, there is a distinction in how leadership positions are distributed. The president has always been elected by his peers for a three year period, renewable. Since 2003 the presidents of the chamber 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)[1]. In contrast, in earlier times these positions were distributed following a rotational setup.

In sum, although the CJEU is secretive when it comes to the direction of

---

[1]While the presidents of the chambers of three judges are also formally elected, in practice, these are positions are allocated on a proposal from the President and rarely make the object of a competition between judges.

decisions promoted by individual judges, it sometimes reveals significant information to re-appointing governments about the performance and potential impact of incumbent judges. In the empirical analysis, we leverage that variation in the quality of information to tease out the effect of transparency on governments' re-appointments. In the next section we specify our theoretical assumptions and derive hypotheses with regards to how governments exercise accountability given these informational constraints and opportunities.

Should (s)he stay or should (s)he go?

We propose that re-appointment decisions are based on the presumed direction of incumbent judges' policy preferences and their predicted ability to have an impact on the court's decisions. In other words, the ideal candidate is both willing and able to pursue the appointer's policy agenda.

We assume that political actors with the power to select and de-select judges will choose candidates that they believe have similar policy preferences to their own. In cases where there is an open seat, for instance following the successive EU enlargements or due to a self-chosen retirement, the nomination of a new judge is determined by a screening process of potential candidates based on their previous record and behavior in other positions. Reversly, in cases where a judge seeks a renewed mandate, the appointer may also contrast potential new candidates with information about the past behavior of the current judge on the bench. If deliberations and votes are secret, the appointer has no information on the direction of the case law that the judge has promoted. However, other sources of information from outside of the bench are still available.

One shortcut to information about the preferences of the incumbent judge is the preferences of the previous appointer. If the initial appointer has similar preferences as the current – as is the case when no change in government has occurred, for example – the latter may assume that the incumbent judge is

likely to have preferences similar to its own. If the initial appointer performed a reliable screening process – and assuming that a new screening process would be relatively costly – the current appointer will choose to retain the judge. In fact, secret voting further increases the reliability of this choice: Since the initial appointer knows that they will not be able to monitor the decisions of the appointed judge, they have incentives to invest more seriously in the screening process.

One implication is that 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 1 *The probability that a judge is reappointed decreases with the distance in preferences between the appointing and the reappointing governments.*

This highlights the limitations of what a lack of transparency can perform. While secret voting contributes to judges' independence, it does not shield them from politicized appointments in case of a change in government. Debates about appointment procedures often conflates the concept of judicial independence with a lack of politicization. However, these two are not conceptually related. In fact, they may even be reversely connected. Judges on a court where votes are public can credibly signal a willingness to defend the new appointer's preferences. They may therefore off-set potential second thoughts that a different re-appointer may have. The re-appointment process thus appears as less politicized, since governments of different colors make the same choice. The judge is nonetheless less independent.

Beyond political preferences, we also theorize that a policy-oriented appointer considers whether the candidate is able to gain sway on the bench. A candidate that is unlikely to influence the discussions on the Court is less valuable from a policy perspective. Another way an incumbent candidate may increase her chances of re-appointment, therefore, is to show that she

18

acquires positions of influence. Judges can send a credible signal to the re-appointing government if two criteria are present. First, past performance must be an indicator of future performance. That is, positions of influence must be selective, not a matter of procedure. Second, information about these positions must be available. The CJEU is, in this respect, exceptionally transparent. Some positions also reflect the trust that fellow judges have in their colleague, and are thus an indicator of future influence.

Hypothesis 2a *The probability that a judge is replaced decreases with their past performance on the bench, in terms of achieving positions of influence on the court.*

Hypothesis 2b *The effect of the judge's performance only holds when positions are allocated following a logic of selection within the court.*

In this setup, governments select competent judges to the extent that the Court's leadership and fellow judges value legal skills. As long as the latter is true, there is no trade-off for governments between power-seeking strategies and competence. Politics may unfold differently within the Court's chamber, and legal merits and integrity may well be valuable assets that governments seek for purely utilitarian reasons. As such, we propose to bridge some of the differences between the judges-as-trustees and judges-as-agents debate.

Alternative explanations and robusness tests

Our theorization is parsimonious, but yields a set of testable predictions that distinguishes it from other, competing explanations for judicial selection.

Denying a high-performing judge a renewed mandate may generate reputational costs if the electorate interprets it as an act of politicization of the judiciary. Reputation could explain the effect of performance, but would predict a moderate to no effect of preferences. Furthermore, we believe this mechanism is weaker at the international level, where appointment processes

are less visible to the broader public, compared to domestic appointments (Elsig and Pollack, 2014). In the empirical part, we draw on the differences between the selection procedures within the Court to bolster our claim that the effect of performance stems from governments' expectation about future influence and not for reputational reasons. If reputation is a primary motivator, we would expect governments to retain high-performing judges even in the event that it doesn't speak to the judge's future performance.

We may further ask whether there is a trade-off between preferences and performance observable through an interaction effect between the two selection criteria. Governments frequently choose between retaining the experienced and potentially influential incumbent judge that the domestic party opponent put in office, or replacing her with a more ideologically aligned newcomer. Having a high-performing policy opponent on the bench with a strong ability to influence policy would be a worst-case scenario. We might therefore expect that such a judge would be removed. However, we have no reason to believe that a – relatively speaking – underperforming judge appointed by a party opponent is any more desirable. From the government's perspective, it would in any case be better to replace the incumbent with a skilled newcomer, assuming there is no shortage of competent candidates at the government's disposal. Building on our theory, we thus expect the two selection criteria to have a linear effect, without any interaction. We do, however, report the results of such a model in the Appendix.

Empirical strategy

Our analyses rely on an original data set listing all 421 appointment decisions made to the Court over its entire history (1952-2019). Information on names, appointment dates, the duration of mandates as well as the cause of exit is collected from the *Official Journal of the European Union*. Judges exit the Court for several reasons. While many appointments (52%) 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 (5%) or was promoted to other positions at the CJEU (5%).

We use this information to identify situations where the government had the opportunity to replace the sitting judge. That is, we consider only potential *re*appointments. 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' decision to replace a judge from the judges' own decision to leave the Court. We therefore discard all exit decisions for which the government's approval is not required, retaining only decisions 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' retention 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 247 (re)appointment decisions to the Court of Justice (61%) and the General Court (39%), while the placebo test is conducted on a similar-sized sample of 255 decisions.

Dependent variable: Replacement

Our dependent variable, *Replacement* is binary. It captures all member state decisions that may 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 (72%) decide to renew the mandate of judges

21

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Missing
Change of judge	0	0	0	0.28	1	1	0
Preference distance (economic issues)	0	0.08	0.24	0.35	0.47	2.52	11
Preference distance (integration issues)	0	0.04	0.13	0.2	0.24	1.69	11
Preference distance (general left-right issues)	0	3.8	10.99	13.18	20.41	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	-0.17	-0.24	0.3	5.28	0
Performance (selective leadership positions)	0	0	0	0.25	0.5	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	54.59	59.93	60.02	65.69	83.75	0
Tenure	1	3.7	6	7.22	9.47	21	0
Change in attendance	-57	-5	5	5.13	16	61	12

Table 1: Descriptive statistics

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%). This propensity has remained fairly stable over time.

Explanatory variables

In the absence of new information on preferences, successive governments will rely on past assessments of the judge. Thus, we expect that the probability of a change of judge increases with the distance between successive governments' preferences (H_1). 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 issues is a continuous measure of the distance between the appointing and the reappointing governments on economic issues. It is calculated as the absolute difference between the factor scores estimated from cabinet parties' electoral manifestos. The Court's most far-reaching and long-established case law has related to the common market, with profound implications for the state's role in economic governance. At the domestic level, it also corresponds to the main policy dimension dividing the government from the opposition. We therefore expect that governments

emphasize economic issues when selecting policy makers for the future. The bivariate statistics already give an indication to that effect. While the median distance between governments that prefer the same judge is 0.18, it increases to 0.38 when the incumbent is replaced. In the Appendix, we report two alternative operationalizations of preferences: Divisions along a general left-right axis and preferences on EU integration.

We further expect that the probability of a replacement is lower for high-performing judges (H_{2a}) that acquired their position through peer selection (H_{2b}) because governments see them as indications of future influence. While the CJEU keeps information on judges' votes secret, it is exceptionally transparent with regards to the role of individual judges in the decision making. We consider two important positions that judges may hold within the internal judicial hierarchy: The position as judge rapporteur and being part of the Court's leadership. For each position, we then distinguish when it was allocated by peer selection.

In each specific case, the judge-rapporteur is the single-most important member of the Court. This role is selective in the Court of Justice, while case management in the General Court builds on extensive monitoring, treating judges as interchangeable. Holding this role in significant cases is therefore a signal of trust in the Court of Justice, while merely an indication of past efforts in the General Court. We therefore analyze these two courts separately, expecting only an effect of judges' performance in the Court of Justice.

Case significance is measured in two alternative ways. Our first operationalization uses panel size to capture the court's own assessment. We then consider the level of attention that cases have attracted in the legal community.

Performance - cases of interest to the Court reports the proportion of cases in a judge's portfolio that were decided by a panel 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 panels of three

The distribution of PERFORMANCE
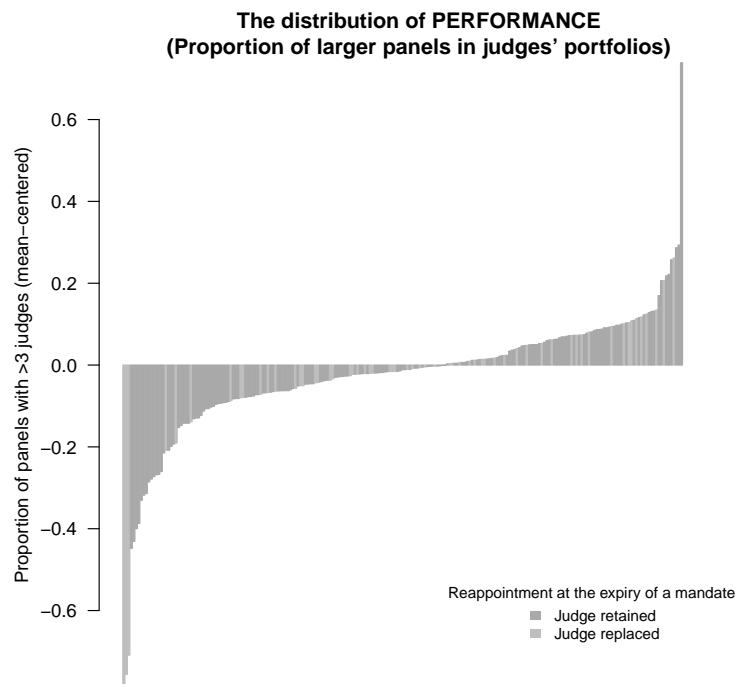**(Proportion of larger panels in judges' portfolios)**

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

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 portfolio includes disproportionately more cases of interest to the Court, while those below indicate 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 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 of 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.

Our third measure of performance captures leadership positions. Following the Court's reorganization in 2003, the status of the presidents in 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 in 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.[2] *Performance - selective leadership positions* is a binary variable that identifies the (Vice-)President and the chamber-of-five presidents post 2003. It is included in the 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 former 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 policy preferences and 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 seek reappointment. 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 would be 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 new appointment decision.

Thus, *Length of tenure* approximates the judge's career stage. The average tenure upon exit was 9.84 years, with the median judge sitting for 2 periods. Similarly, we control for a judge's *Age* at the time of the (re)appointment 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

---

[2]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.

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*). The variable controls for the possibility that the ruling party uses Court 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 (re)appointment (*(Vice-)President*), since these positions are clearly prestigous although the division of labor implies that they handle few cases 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).

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

The definition of the variables of interest, X, vary across models because we explore different operationalizations of the two hypothesized selection criteria; preferences and performance. All models also include relevant controls, Z.

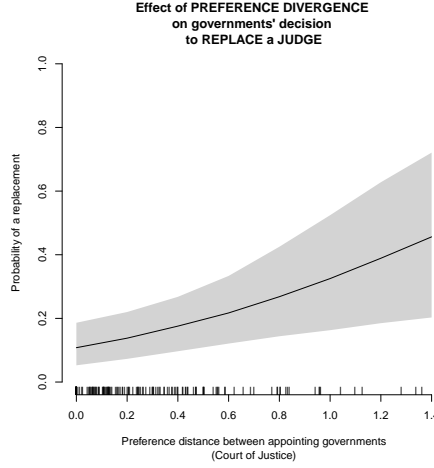The results are obtained from Bayesian models with MCMC estimation

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

(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 largely support our expectations.

The probability that a judge is replaced increases as the political distance between the appointing and reappointing governments grows (H_1). This holds for both courts. 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[3]. Such a shift would more than double the odds of a replacement.

We have argued that court cases overall revolve around economic questions. In the Appendix, we explore alternative operationalizations of government preferences. While divisions on a general left-right dimension yield similar but more moderate effects, we find no support for earlier suggestions that governments consider questions of European integration during the selection of judges (Kenney, 1998–1999). The results also put in perspective the Court's reputation as a run-away agent. The Court may well have been instrumental to European integration in a spill-over from economic questions to other policy areas (Burley and Mattli, 1993). However, the consistency of governments' attention to economic preferences during judicial appointments indicate their continuous focus on policy rather than the distribution of competence between governance levels.

We have argued that governments seek to influence the Court's decision making through the appointment of its members. However, selecting judges with the desired political compass, only takes them one step of the way. If governments seek political influence, they will look for cues that their judge can gain sway in office (H_{2a}). Yet we found no indication in the bivariate analysis that performance matters for reselection. This may give the impression to onlookers that governments only consider candidates' political orientation, while disregarding legal competence when they select judges. However, when controlling for judges' career stage and government preferences, we find a substantial effect of judges past performance. We explore three alternative measures of incumbent judges' past influence.

Handling high-impact cases in the Court of Justice, where such positions are selective, is clearly associated with judges' chances of being reappointed.

---

[3]There are 35 instances in the data where the prime minister's party has shifted from social democratic to conservative (or vice-versa) between appointments. The median distance in their estimated economic preferences is 0.51.

Dependent variable: 'Replacement'	Court of Justice	Court of Justice	General Court	General Court
Intercept	-0.34	-1.13	-2.25	-2.06
	(-1.9,1.15)	(-2.03,-0.25)	(-3.33,-1.32)	(-3.24,-0.97)
Preference distance (economic issues)	1.41	1.38	1.76	1.85
	(0.44,2.36)	(0.45,2.34)	(0.51,3.07)	(0.58,3.26)
Performance (cases of court interest)	-1.79		-0.07	
	(-3.39,-0.25)		(-0.67,0.56)	
Performance (cases of interest to legal community)		-1.02		-0.3
		(-1.89,-0.24)		(-1.27,0.52)
Change of PM party	0.02	0.03	0.81	0.85
	(-0.81,0.87)	(-0.81,0.87)	(-0.14,1.8)	(-0.13,1.8)
(Vice-)President	0.7	1.01	-3.39	-3.33
	(-0.72,2.18)	(-0.49,2.56)	(-6.99,-0.83)	(-7,-0.72)
Length of tenure	0.13	0.13	0.11	0.12
	(0.01,0.24)	(0.02,0.25)	(0,0.23)	(0,0.24)
Age	0.17	0.17	0.11	0.11
	(0.11,0.24)	(0.1,0.24)	(0.04,0.18)	(0.05,0.18)
Change in attendance	-0.02	-0.02	0.03	0.03
	(-0.04,0)	(-0.04,0)	(0,0.07)	(0,0.07)
Number of observations	150	150	97	97
Proportion of correct predictions	0.76	0.79	0.73	0.74
... correct positive predictions	0.71	0.79	0.77	0.73
... correct negative predictions	0.78	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.

This holds regardless of how we measure the importance of judges' portfolio.

Model 1 shows that the more large-panel cases a judge has obtained as rapporteur, the more likely they are to remain in office. All things equal, judges whose portfolio contains 10 percentage points more large-panel cases than the overall distribution in the Court, have a 86 % lower odds of replacement. Model 2 then considers the mean number of academic articles discussing cases where the judge has acted as rapporteur. Here too, we find support for our hypothesis. If we consider the difference between a typical under-performing (20th percentile) and a typical over-performing judge (80th percentile), the most influential judge has a 54% higher probability of retaining their seat.

Last, Table 3 and Figure 4 report the effect of being endorsed for a leadership position in either of the two courts. 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.

Our theory speaks to the agency that courts have even when terms are renewable and appointments are political. This hinges on the Court's auton-

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.

omy to organize its own work, the transparency of its decision making and governments' policy seeking.

First, our results have implications for courts' institutional design. Governments base their assessment of judges' potential influence on information provided by the Court. Under the right transparency rules, the Court can thereby signal to governments which judges it would like to keep in office. However, this only holds if the Court allocates influence selectively (H_{2b}). If positions are acquired as a matter of procedure – through geographical quotas or extensive monitoring – a new judge would stand the same chance of obtaining them in the future as the incumbent had in the past. Thus, the Court's signals would carry no information. Hence, in our study, the effect of past influence is indistinguishable from zero when we consider judges' case portfolio in the General Court where allocations treat judges as interchangable (models 3 and 4 in Table 2) or when leadership positions rotate between members (Table 3).

Second, conditional on positions being selective, high-performing judges are to some extent shielded against political (de)selection. Governments face
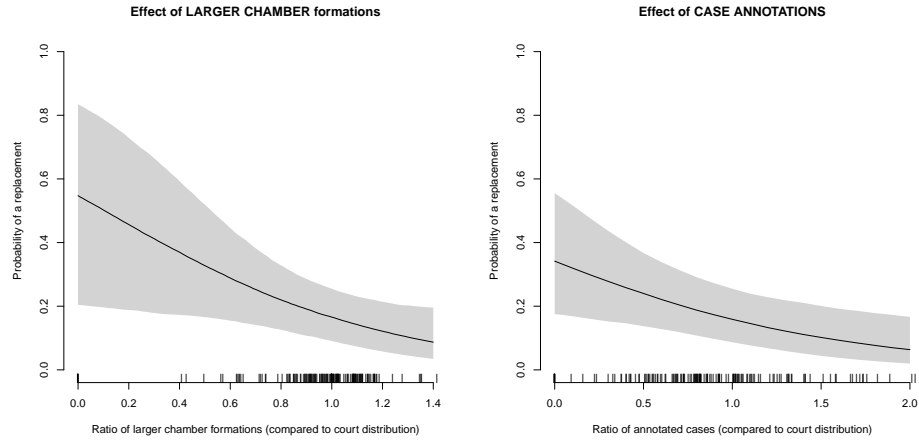
Figure 3: Predicted probability of replacement as a function of performance in office among judges in the Court of Justice.
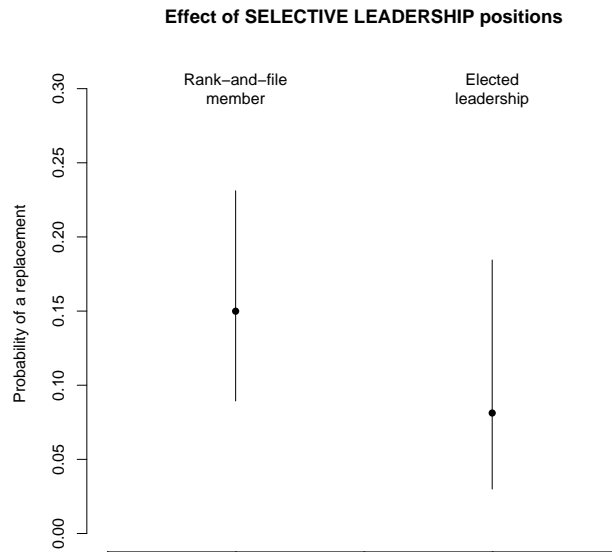

Figure 4: The effect of elected and non-elected leadership positions on governments' choice to replace their incumbent judge.

a tradeoff between preferences and performance when they consider replacing the sitting judge. Thus, if we consider the same scenario as before where there has a been a substantial change in government preferences (i.e. between a social democratic and conservative prime minister), a high-performing judge would still have about the same probability of being replaced (17%) as a low performer in a situation with no change in government preferences (19%).

More generally, governments have an incentive to appoint judges whose political orientation is palatable the opposition. The Court tends to reserve the most influential positions to experienced judges. Case allocation in the Court of Justice also allows judges to specialize, letting members gain a disproportionate influence over certain policy areas (Hermansen, 2020). 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. By replacing the sitting judge by an equally competent alternative, the government would therefore still see their investment nullified and would have to wait at least another 6 years before the new judge acquires a similar standing within the Court. An implication is that governments in principle stand to gain from replacing the incumbent only if they believe the new candidate's skills are substantially higher than their predecessor. As the preference distance increases between successive governments, they may be more willing to sustain the cost of changing judges. This may also explain why replacing an incumbent judge is a relatively rare event. Even a low performer facing an opposing government only has a predicted probability of 32 percentage points of losing their job.

Robustness of the results

The models provide a fair description of the data with an in-sample correct prediction rate of 76%. They are also robust to several alternative specifications, relating to variation over time and between member states. In particular, we find no consistent evidence that the introduction of the ad-

visory selection committee changed member states' emphasis on preferences or performance when they decide whether to replace the incumbent. Details about these additional tests are in the appendix.

We pin the effects of preferences and performance on governments' choice of judge rather than judges' choice 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 theorization, we only find an effect of preference distance and performance (in the Court of Justice) when governments' decisions were involved.

An intuitive expectation could be that governments see high-performing judges appointed by the opposition as a worst-case scenario such that performance is valued less (or negatively) when preference distance is high. As we have already argued, this is an awkward prediction to test empirically. Governments have always the opportunity to replace the sitting judge by someone who maximizes both criteria. We nevertheless report the results from a model that includes an interaction between preferences and performance in the Appendix. We find no trace of any such effect. In fact, our model suggests the opposite. Binomial logistic models are non-linear, meaning that the total effect of one variable is dependent on the level of the other variables in the data. If we compare the predicted difference between a high and a low performer when there has been no change in government, on the one hand, and a substantial shift in preferences, on the other, we find an increase of 10 and 16 percentage points, respectively. In other words, governments are to some extent more lenient to judges appointed by themselves, while they emphasize performance more when the incumbent was appointed

Dependent variable: 'Replacement'	Mandate expired	Resignation
Intercept	-0.34	-0.75
	(-1.9,1.15)	(-2.45,0.86)
Preference distance (economic issues)	1.41	-0.7
	(0.44,2.36)	(-2.09,0.58)
Performance (cases of interest of court interest)	-1.79	-0.82
	(-3.39,-0.25)	(-2.44,0.77)
Change of PM	0.02	-0.22
	(-0.81,0.87)	(-1.1,0.66)
(Vice-)President	0.7	0.25
	(-0.72,2.18)	(-1.3,1.71)
Length of tenure	0.13	0.19
	(0.01,0.24)	(0.07,0.31)
Age	0.17	0.13
	(0.11,0.24)	(0.06,0.21)
Change in attendance	-0.02	-0.03
	(-0.04,0)	(-0.05,-0.01)
Number of observations	150	139
Proportion of correct predictions	0.76	0.76
... correct positive predictions	0.71	0.74
... correct negative predictions	0.78	0.76

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.

by the opposition.

Discussion

We have argued that appointments and reappointments of judges to international is based on the information that nominating governments have regarding potential candidates' policy preferences and their ability to influence other judges. That information, in turn, depends on design features of these courts.

When dissenting votes and opinions are public, governments may update their beliefs about incumbent judges' policy preferences on a regular basis. If terms are renewable, this information may then be used to reward or sanction incumbent judges by allowing or denying them a new term in office. In courts

where deliberations and votes are secret, on the other hand, as is the case in the Court of Justice of the EU and (to some extent) the WTO Appellate Body, reappointing governments need to look elsewhere for information on the policy preferences of the incumbent judge. A probable source is the policy preferences of the initially appointing government, whose screening of candidates is likely to have placed a judge with similar preferences on the bench.

This means that judges are independent, insofar as they are not held accountable for the content of their decisions, but secrecy does not shield judges from politically motivated reappointments. Our empirical findings in the case of the CJEU indicate that EU governments choose to replace incumbent judges based on their presumed preferences regarding economic policy. Since no voting records are available from the Court, governments have no other information on the policy preferences of the judge than what they can infer from the political ideology of the government who made the previous selection. In the event of a shift in government, therefore, the secrecy of the Court's deliberations work in practice against opportunistic judge's chances of staying in office.

However, because governments are policy seekers, performance matters for reappointment insofar as it signals future potential influence. Again, the presence and quality of that information depend on the court's institutional design. Two conditions have to be present: (i) Past performance has to be indicative of future influence. That is, positions cannot be distributed merely based on judges' availability or through geographical quotas, for example, but rather the result of a deliberate choice by 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. (ii) The information has to be available. The pre-reform ECtHR serves as a good 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 judge rapporteur in a non-

random way, the name of that rapporteur is never published. Thus, while voting decisions were known, the relative influence of judges was not. In contrast, we have argued that the large variation in influence among judges of the CJEU is at times both informative and observable. Hence, performance – measured both through members' portfolio of cases and leadership positions – has a substantive effect on judges' likelihood of a reappointment to the CJEU.

Governments value not only that judges have similar policy preferences, but also that they have the ability to promote those preferences in competition for influence with judges from other states. Evidence that judges receive responsibility for important cases, or important positions within the court's internal organization, may signal confidence that the incumbent is respected by peers and in a position to influence others.

Our empirical findings in the case of the CJEU indicate that EU governments choose to replace incumbent judges based on their presumed preferences regarding economic policy. Since no voting records are available from the Court, governments have no other information on the policy preferences of the judge than what they can infer from the political ideology of the government who made the previous selection. In the event of a shift in government, therefore, the secrecy of the Court's deliberations in practice work against opportunistic judge's chances of staying in office.

While secrecy of judicial votes increases the judicial independence in terms of political accountability judges may still be held accountable for their performance on the bench. Our results indicate that the internal organization of the CJEU 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 their mandate expires.

References

Alt, James, Ethan Bueno de Mesquita, and Shanna Rose (2011). "Disentangling Accountability and Competence in Elections: Evidence from U.S. Term Limits". In: *The Journal of Politics* 73.1, pp. 171–186.

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.

Bailer, Stefanie and Tamaki Ohmura (2017). "Exploring, Maintaining, and Disengaging—The Three Phases of a Legislator's Life". In: *Legislative Studies Quarterly* 0.0, Online first.

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.

Besley, Timothy and Anne Case (1995). "Does Electoral Accountability Affect Economic Policy Choices? Evidence from Gubernatorial Term Limits". In: *The Quarterly Journal of Economics* 110.3, pp. 769–798.

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.

Burley, Anne-Marie and Walter Mattli (1993). "Europe Before the Court: A Political Theory of Legal Integration". In: *International Organization* 47.1, pp. 41–76.

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.*

Conant, Lisa (2007). "Review Article: The Politics of Legal Integration". In: *JCMS: Journal of Common Market Studies*. Annual Review 45.1, pp. 45–66.

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.

Ferraz, Claudio and Frederico Finan (2011). "Electoral Accountability and Corruption: Evidence from the Audits of Local Governments". In: *American Economic Review* 101.4, pp. 1274–1311.

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.

Holmstrom, Bengt and Paul Milgrom (1991). "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design". In: *Journal of Law, Economics & Organization* 7 (Special Issue), pp. 24–52.

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: `UPiqkRrOMcoC`.

— (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.

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.

Pérez, Aida Torres (2015). "Can Judicial Selection Secure Judicial Independence?: Constraining State Governments in Selecting International Judges". In: *Selecting Europe's Judges: A Critical Review of the Appointment Procedures to the European Courts*. Ed. by Michal Bobek. Oxford University Press, pp. 181–201.

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.

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 (2015).

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: `RlcUsmKA5iIC`.

Staton, Jeffrey K. and Will H. Moore (2011). "Judicial Power in Domestic and International Politics". In: *International Organization* 65.3, pp. 553–587.

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.

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 substantial 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 then reported in paranthesis. 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 caclulate an alternative variable where preferences are not weighted.

Last, the preference distance between successive government 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). See the next section for further information about the imputation of missing information.

- *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 can be classified in terms of their 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_ren|`) 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 (Volkens et al., 2017) and reports the absolute distance on the standard right-left scale provided by the Manifesto Project that differentiates the parties in the appointing and reappointing governments. 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.

- *Redestribution preferences* (`redestribution`) 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 continous measure of the mean number of legal annotations that cases garnered in which the judge acted as a rapporteur during his mandate. It is centered on the mean number of annotations garnered 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 - managing cases of member state interest (`Observations.mean – Observations.court.mean`) is a continous measure of the mean number of observations from member states that were filed in cases in which the judge acted as a rapporteur during his mandate. It is centered on the mean number of observations submitted in 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 from Eur-lex.

Performance - leadership positions (`Chamber5President`) reports whether the judge had precided over deliberations in panels with more than 5 participants during his 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 occured 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) 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 substracting 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 examplify 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:

$$Pr(Replacement_i = 1) \sim Bernouilli(\pi_i)$$
$$logit(\pi_i) = \alpha +$$
$$\beta_1 \times |Preferences\ gvt\ 1_i - Preferences\ gvt\ 2_i| +$$
$$\beta_2 \times (Annotations_i - Avg.\ Annotations_i) +$$
$$\beta_3 \times Change\ in\ Prime\ Minister_i +$$
$$\beta_4 \times (Vice-)President +$$
$$\beta_5 \times Tenure_i - mean(Tenure) +$$
$$\beta_6 \times Age_i - mean(Age) +$$
$$\beta_7 \times Change\ in\ Attendance_i$$

$$(2)$$

The regression parameters have relatively vague prior distributions:

$$\alpha \sim N(0, 10)$$
$$\beta_k \sim N(0, 10)$$
$$\gamma_k \sim N(0, 10)$$
$$\delta_k \sim N(0, 10)$$

$$(3)$$

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 10th 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[4]. They are therefore treated as randomly missing and imputed through priors informed by the

---

[4]Tenure (0 observation, 0%) and Legal annotations (0 observation, 0%

mean and standard deviation among the observed units:

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

$$(4)$$

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:

$$Preferences\ gvt\ 1_i \sim N(\mu 1_i, \tau_i)$$
$$\mu 1_i = \gamma 1_1 +$$
$$\gamma 1_1 \times Economic\ left-right1_i +$$
$$\gamma 1_2 \times Economic\ deregulation1_i$$

$$(5)$$

$$Preferences\ gvt\ 2_i \sim N(\mu 2_i, \tau_i)$$
$$\mu 2_i = \gamma 2_1 +$$
$$\gamma 2_1 \times Economic\ left-right2_i +$$
$$\gamma 2_2 \times Economic\ deregulation2_i$$

$$(6)$$

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 model 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)	-1.53
	(-3.36,1.3)
Economic left-right preferences (appointing gvt)	-0.13
	(-0.2,-0.05)
Economic deregulation (appointing gvt)	0.23
	(0.18,0.28)
Intercept (reappoiting gvt)	0.62
	(-1.11,3.65)
Economic left-right preferences (reappointing gvt)	-0.12
	(-0.21,0.01)
Economic deregulation (reappointing gvt)	0.18
	(0.12,0.24)
Number of observations	150

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	-0.34	-0.18
	(-1.9,1.15)	(-1.73,1.42)
Preference distance (economic issues)	1.41	1.44
	(0.44,2.36)	(0.46,2.4)
Performance (court interest)	-1.79	-1.97
	(-3.39,-0.25)	(-3.65,-0.38)
Change of PM	0.02	-0.01
	(-0.81,0.87)	(-0.86,0.84)
(Vice-)President	0.7	0.55
	(-0.72,2.18)	(-1.07,2.03)
Length of tenure	0.13	0.11
	(0.01,0.24)	(0,0.23)
Age	0.17	0.17
	(0.11,0.24)	(0.11,0.25)
Change in attendance	-0.02	-0.02
	(-0.04,0)	(-0.04,0)
Number of observations	150	145
Proportion of correct predictions	0.76	0.77
... correct positive predictions	0.71	0.75
... correct negative predictions	0.78	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 secondary issues less relevant to CJEU case law. This might nevertheless 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	-0.34	-0.38	-0.12
	(-1.9,1.15)	(-1.81,1.1)	(-1.64,1.36)
Preference distance (economic left-right)	1.41		
	(0.44,2.36)		
Preference distance (general left-right)		0.02	
		(-0.01,0.06)	
Preference distance (European integration)			0.01
			(-1.75,1.66)
Performance (cases of court interest)	-1.79	-1.56	-1.57
	(-3.39,-0.25)	(-3.13,-0.15)	(-3.12,-0.02)
Change of PM party	0.02	0.22	0.34
	(-0.81,0.87)	(-0.58,1.06)	(-0.45,1.15)
(Vice-)President	0.7	0.58	0.48
	(-0.72,2.18)	(-0.87,1.98)	(-1.02,1.93)
Length of tenure	0.13	0.11	0.13
	(0.01,0.24)	(0,0.23)	(0.02,0.24)
Age	0.17	0.16	0.16
	(0.11,0.24)	(0.1,0.23)	(0.1,0.23)
Change in attendance	-0.02	-0.02	-0.02
	(-0.04,0)	(-0.04,0)	(-0.04,0)
Number of observations	150	150	150
Proportion of correct predictions	0.76	0.77	0.77
... correct positive predictions	0.71	0.74	0.74
... correct negative predictions	0.78	0.79	0.78

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

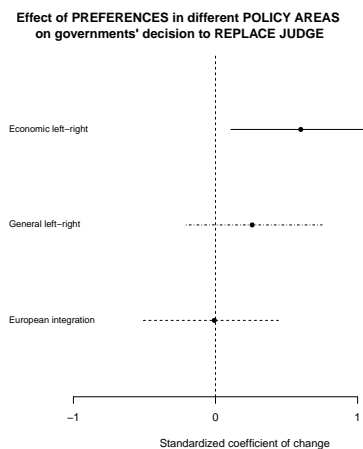Table 7: The models explore DIFFERENT OPERATIONALIZATIONS of PREFERENCES.

**Effect of PREFERENCES in different POLICY AREAS
on governments' decision to REPLACE JUDGE**

Figure 5: Different operationalizations of government 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; 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 (i.e. the "255 Panel").

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 (80%). The results will therefore be driven by selection in the last three decades.

To verify that the selection criteria have not changed substantially over
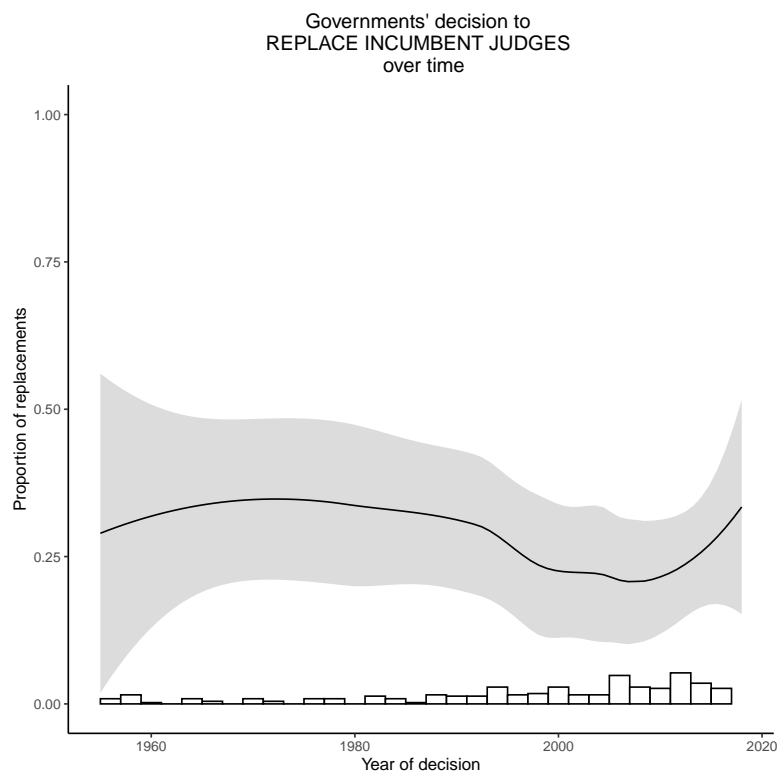
55

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).

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.

Did the introduction of an advisory panel change the criteria (i.e. "the 255 Panel")?

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

Dependent variable: 'Replacement'	1952-1980	1980-2000	20o0-2020
Intercept	-0.41	-0.27	-0.95
	(-2.79,1.93)	(-3.62,2.81)	(-3.2,1.22)
Preference distance (economic issues)	2.94	2.32	0.38
	(1.28,4.94)	(0.75,4.2)	(-1.38,2.06)
Performance (cases of court interest)	-1.34	-1.92	-2.58
	(-3.78,0.94)	(-5.06,1.31)	(-5,-0.31)
Change of PM party	-1.1	-0.37	2
	(-2.77,0.44)	(-2.32,1.41)	(0.55,3.59)
(Vice-)President	0.35	-0.84	0.42
	(-2.3,2.87)	(-3.91,2.03)	(-2.29,2.9)
Length of tenure	0.14	0.28	0.1
	(-0.04,0.35)	(0.03,0.63)	(-0.08,0.28)
Age	0.3	0.18	0.13
	(0.17,0.48)	(0.06,0.33)	(0.02,0.24)
Change in attendance	-0.05	-0.03	-0.04
	(-0.13,0.01)	(-0.08,0.02)	(-0.06,-0.01)
Number of observations	50	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.

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.

Is the there a tradeoff between competence and preferences?

Dependent variable: 'Replacement'	CJ	GC
Intercept	0.03	-3.28
	(-1.67,1.67)	(-5.06,-1.71)
Preference distance (economic issues)	1.19	1.53
	(0.19,2.2)	(-0.02,3.15)
255 Panel * preference distance	0.78	1.48
	(-1.76,3.36)	(-1.21,4.14)
255 Panel	-1.21	0.97
	(-3.89,1.49)	(-0.78,2.84)
Performance (court interest)	-1.88	0.57
	(-3.58,-0.18)	(-0.77,1.86)
255 Panel * performance	-0.48	-0.67
	(-3.26,2.17)	(-2.11,0.75)
Change of PM	0.26	0.86
	(-0.65,1.11)	(-0.13,1.89)
(Vice-)President	0.44	-3.68
	(-1.09,1.91)	(-7.11,-0.92)
Length of tenure	0.17	0.11
	(0.05,0.3)	(-0.01,0.23)
Age	0.18	0.1
	(0.12,0.26)	(0.03,0.18)
Change in attendance	-0.01	0.03
	(-0.03,0)	(-0.01,0.06)
Number of observations	150	97
Proportion of correct predictions	0.78	0.77
... correct positive predictions	0.71	0.77
... correct negative predictions	0.81	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.

Dependent variable: 'Replacement'	Court of Justice
Intercept	-2.19
	(-2.96,-1.46)
Preference distance (economic issues)	1.48
	(0.49,2.51)
Preference distance * Performance	-3.63
	(-11.87,4.75)
Performance (cases of court interest)	-1.66
	(-4.41,1.07)
Change of PM party	0.01
	(-0.86,0.87)
(Vice-)President	0.75
	(-0.84,2.29)
Length of tenure	0.13
	(0.02,0.25)
Age	0.18
	(0.11,0.25)
Change in attendance	-0.02
	(-0.04,0)
Number of observations	150
Proportion of correct predictions	0.77
... correct positive predictions	0.71
... correct negative predictions	0.8

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

Table 10: Is there an interaction effect between governments' preferences and judges' performance?