How to place non-majoritarian institutions and political actors in a common policy space: Spatial modeling of court–executive interactions

Benjamin G. Engst
University of Konstanz
Department of Politics and Public Administration
benjamin.engst@uni-konstanz.de

David M. Grundmanns
University of Mannheim
Collaborative Research Center 884
david.grundmanns@mzes.uni-mannheim.de

Thomas Gschwend
University of Mannheim
School of Social Sciences
gschwend@uni-mannheim.de

Abstract. How can we estimate positions of non-majoritarian institutions in a common policy space? To answer this question, we take highest courts as examples of powerful non-majoritarian institutions and develop a new scaling approach to estimate their position in a common policy space with other political actors. In contrast to previous research, our approach neither relies on individual votes of justices nor assumes that justices "inherit" positions from political actors who nominated them. Instead, for each court decision, we use the positions of political actors expressed in written statements as well as the courts' decision outcome to estimate comparable policy positions. In two applications, we position the German Federal Constitutional Court with different German governments and the European Court of Justice with different European governments in common policy spaces and validate them. Finally, we show how our common policy scores can be used to study court–executive relations and inter-institutional interactions.

We thank Linda Schill and Vanessa Müden for valuable research assistance and the Collaborative Research Center 884 at the University of Mannheim, funded by the German Research Foundation, for generous research opportunities. We are also grateful for valuable copy-editing by Hannah Laumann. The article has been developed as part of the project "Measuring a common space and the dynamics of reform positions" (C4), conducted at the aforementioned research center.

Political scientists often use spatial models to simplify complex political behavior (Romer and Rosenthal, 1978; Clinton, Jackman and Rivers, 2004), e.g., when studying power struggles. To assess the implications of these models empirically, we must assume that the positions of all actors are comparable in a common space. This assumption is true when comparing the positions of the same type of actor, such as parties or legislators. Unfortunately, power struggles often involve a variety of different actors, and we frequently lack a reliable and valid measurement strategy to position all involved actors within the same policy space. Non-majoritarian institutions, such as central banks, regulatory agencies, or constitutional courts, serve as prime examples of such actors. Non-majoritarian actors play a role in policymaking, but they differ from typical political actors. How can we identify the positions of non-majoritarian actors in a common policy space with political actors?

To answer this question, we use highest courts as examples of particularly powerful non-majoritarian institutions and develop a measurement strategy to place courts in a common policy space with political actors. In doing so, we contribute to prior research by estimating—rather than assuming (Hönnige, 2009; Brouard and Hönnige, 2017)—comparable scores for various actors. Our measurement strategy can be applied to different contexts and actors for which existing methods fail to construct a common policy space. Specifically, we develop valid and reliable common policy scores on the national level for the German Federal Constitutional Court (GFCC) and various German governments as well as on the supranational level for the European Court of Justice (ECJ) and various European governments.

Existing approaches to constructing a common space—widely used in studies on court–executive relations in the US—scale judicial decisions based on individual judicial votes and relate them to the known political ideologies (Martin and Quinn, 2002; Epstein, Martin, Quinn and Segal, 2007; Epstein, Martin, Segal and Westerland, 2007; Hanretty, 2012*a*,*b*, 2014). This allows for placing a highest court as a collective actor in a

common space with political actors. Nevertheless, this approach has at least three major shortcomings.

First, publishing individual votes by members of non-majoritarian institutions is not common practice in all countries. Powerful courts such as the German *Bundesverfassungsgericht*, the Austrian *Verfassungsgerichtshof*, or the Italian *Corte costituzionale* either are not allowed to publish votes or rarely do so (Raffaelli 2012, 30, Kelemen 2013, 1345). Moreover, even if non-majoritarian institutions report votes, they seldom publish dissent but follow a norm of consensus. This makes it difficult to apply standard scaling approaches using roll call votes (Baerg and Lowe, 2020). Second, not all justices are selected by political actors or through a politicized process, making it challenging to relate the views expressed in individual judicial votes to prior information about the justices' political views. Third, even if we can infer certain doctrinal leanings from court decisions, it is not valid to place them directly in a common policy space because the legal language used in court decisions is not necessarily comparable to the language used in party manifestos, which are often used to estimate the positions of political actors.

We address these shortcomings by using features of court decisions that are not based on dissenting votes or positions justices have "inherited" from the political actors who nominated them. Instead, we employ a strategy used in legislative research whereby scholars compute a matrix of roll call votes to position legislators in a common policy space (e.g., Clinton, Jackman and Rivers, 2004). In our context, we consider the outcome of a court decision (or, in the ECJ context, a ruling on a "question") as a collective "vote" by the justices on whether a referral to the court is constitutional or not. To construct a vote matrix, we add the "votes" by political actors who submit briefs (or, in the ECJ context, "observations") expressing their opinion on the constitutionality of the same referral. We then scale the positions of the courts relative to the positions of the political actors who submitted the briefs. To establish a common policy space, we use the manifesto scores of the political actors as bridging observations. The bridging observations enable us to map

judicial positions from a legal case space (cf. Clark and Lauderdale, 2012; Fischman, 2019; Arnold, Engst and Gschwend, 2023) onto a common policy space.

In what follows, we briefly review the literature on estimating comparable positions to study inter-institutional interactions. We then introduce our scaling approach and present two applications, one to estimate common policy scores for the GFCC and different German governments and another to measure common policy scores for the ECJ and different European governments. In both cases, we show how existing research can benefit from our scaling approach. We conclude by summarizing the advantages of our scaling approach.

Our new strategy to develop common policy scores has significant implications for the assessment of court–executive relations. Scholars no longer need to limit their analyses to the few countries in which dissenting votes are published. Instead, our measurement strategy is more general and extends to systems in which political actors can file briefs that are referenced in court decisions. Importantly, our approach can be adapted to examine other non-majoritarian institutions that make collective decisions with policy implications, such as central banks or regulatory agencies.

Comparable positions to study inter-institutional interactions

To study the interactions between political actors in a spatial model empirically, scholars need to conceptualize a common space and then determine the positions of political actors within it. The aim of this study is to estimate the positions of courts in a common policy space with political actors. Courts are the best-studied non-majoritarian institutions for which scholars have developed strategies to measure their positions (on the positions of central banks see, for example, Baerg and Lowe, 2020). In this section, we argue that the existing strategies are not necessarily applicable in different contexts without relying on strong assumptions and that we therefore require new methods.

The most extensively studied non-majoritarian institution is the US Supreme Court.

Martin and Quinn (2002) use individual judicial votes to estimate its ideological leaning (see also Epstein, Martin, Quinn and Segal, 2007; Epstein, Martin, Segal and Westerland, 2007) with models based on traditional item response theory (IRT). Hanretty (2014) adopts a similar strategy to study the Bulgarian Constitutional Court. He uses published dissenting opinions to scale the court, but these are not commonly available for non-majoritarian institutions. Instead, justices often follow a norm of consensus (Epstein, Segal and Spaeth, 2001). Unanimous decisions are generally not informative for scaling, as Hanretty (2012*b*, 706) points out when estimating the positions of British Law Lords, who often publish consensual decisions.[1] Research on legislative politics has shown that scaling models work well in situations with published individual votes or roll call votes (e.g., Clinton, Jackman and Rivers, 2004). However, it is more difficult to apply them in contexts in which there are fewer decisions with identifiable votes or in which there is a predominant norm of consensus when justices, legislators, or central bankers cast their votes (Baerg and Lowe, 2020; Bonica and Sen, 2021; Epstein, Segal and Spaeth, 2001).

Alternative approaches derive justices' individual positions not from court decisions but from external sources. Common methods used by scholars to infer justices' positions from their preferences are coding newspaper editorials on congressional hearings when they were nominated (Segal and Cover, 1989), identifying justices' campaign contributions prior to their nomination (Bonica and Sen, 2017), or using process tracing of expert judgments, historical accounts, and "all other existing and relevant documentary materials" (Epstein, Knight and Shvetsova, 2001, 140). These approaches require in-depth ("thick") knowledge of each political environment, which makes them labor-intensive and less

[1] Hanretty (2012*a*) faces similar challenges when estimating the positions of Spanish and Portuguese justices, as there are few dissenting opinions and many unanimous decisions. Similarly, Engst et al. (2017) are only able to identify 20 decisions by the GFCC over a period of about eleven years to analyze networks among justices.

reliable. Additionally, such positions, once estimated, will not be updated during a justice's tenure, for example, because they no longer make campaign contributions. The validity of the inferred positions is also threatened when outside sources do not adequately reflect true judicial preferences. For example, justices strategically conceal their sincere preferences during congressional hearings (Segal and Cover, 1989, 560-561).

Approaches that draw on individual judicial votes to estimate policy positions are not generally applicable because publishing votes is not common practice across all judicial systems (Raffaelli, 2012; Kelemen, 2013). Consequently, scholars who analyze European domestic courts, for example, which seldom publish votes, have to make strong assumptions instead. They often employ what we call the "party label approach" to identify the political leaning of courts. According to this approach, a position is assigned to a court based on the ideological views of the mean or median justice. These views are derived from the policy position of the actor nominating this justice (Hönnige, 2009; Carrubba et al., 2012; Brouard and Hönnige, 2017). The approach therefore assumes that justices "inherit" fixed political positions from their nominating actors. However, research on the US Supreme Court (Martin and Quinn, 2002; Epstein, Martin, Quinn and Segal, 2007) shows that justices change positions during their tenure. Additionally, not all justices are selected by a scalable actor, i.e., an actor with an identifiable position in a common policy space.

In sum, existing scholarship scales justices primarily on the basis of individual judicial votes. This approach is not suitable for cross-country comparisons. Assigning justices' policy positions based on their nominating actors is also an ineffective way to overcome this limitation. Furthermore, the in-depth analysis of judicial behavior is labor-intensive and cannot be applied to larger settings.

We propose scaling courts directly rather than on the basis of individual judicial votes. To this end, we draw on two pieces of information that are typically available in published court decisions to construct a matrix of "votes." The first is the dichotomous outcome of

a decision as a collective "vote" by all justices, indicating whether a highest court rules a referral to be constitutional or unconstitutional. The second are the positions taken by scalable political actors in written briefs submitted prior to the court's decision, in which they express their opinions on the constitutionality of the referral. We operationalize these as the actors' "votes." In appendix A, we demonstrate that briefs by political actors are a cross-country comparative feature of court decisions and more prevalent than individual judicial votes.

We are not the first to use briefs to estimate the policy positions of courts. Fischman (2015) uses briefs submitted to the US Supreme Court by two particular interest groups to estimate its position in a particular policy space using multidimensional scaling. Our approach, in contrast, draws on behavioral theory to derive IRT scaling models. In the following sections, we explain in more detail how we determine the positions of highest courts as non-majoritarian institutions in a common policy space.

The vote matrix approach to estimate common policy scores

How can we place highest courts and political actors in a common policy space? Positioning political actors in a common policy space is straightforward. Political actors are partisan, and the literature has established common policy scales for parties, often based on their manifestos (e.g., Lowe et al., 2011; König, Marbach and Osnabrügge, 2013; Lehmann et al., 2022). Understanding the relationship between highest courts and political actors requires comparable policy positions. Therefore, we need to locate courts in the same common policy space, too. To do so, we adapt a scaling approach that allows us to map different types of actors in a common unidimensional space using appropriate bridging observations (e.g., Bailey and Chang, 2001; Bailey, 2007; Jessee, 2009, 2016).

Critics argue that unidimensional scales inaccurately reflect the preference formation of actors mapped onto them (e.g., Fischman and Jacobi, 2015; Fischman, 2019). Empirically, there are both findings that support unidimensional scales (Poole and Rosenthal, 1991,

2007) and those that support multidimensional scales (e.g., Roberts, Smith and Haptonstahl, 2016). Given these mixed results, we follow the literature in prioritizing parsimony over complexity (similar Jessee, 2016, 1110). For the first application, the most prominent way to structure the political landscape in Germany is by using a common left–right scale, with progressive and liberal actors positioned to the left of conservative actors (see, e.g., Matthieß, 2020; Lehmann et al., 2022). For the second application, we locate both the ECJ and the respective national governments on a scale of European integration (König, Marbach and Osnabrügge, 2013).

To position a highest court and political actors in a common policy space, we leverage two pieces of information from court decisions. First, political actors file briefs using legal language to express their preferred outcome of a referral. For instance, a minister of the federal government (A) presents a brief to the court arguing whether a referral is constitutional or not. Similarly, a state government (B) presents a different argument. Second, the court decides on the constitutionality of the referral. If we consistently observe that actor A agrees with the court more often than actor B, then the former should be placed closer to the court in a common space than the latter. This allows us to estimate the court's position relative to the positions of political actors filing the briefs. We then anchor the positions in a common policy space based on the established manifesto scores of the political parties.

To scale the court, we adopt the unidimensional spatial voting model, which is equivalent to the two-parameter IRT model (Clinton, Jackman and Rivers, 2004; Jessee, 2016). In particular, we estimate a standard two-parameter IRT model with a probit link:

$$P(Y_{ij} = 1) = \Phi[\beta_j(x_i - \gamma_j)], \tag{1}$$

where β_j is the discrimination parameter for decision j, indicating the strength and direction of the relationship between actor i's ideal point x_i and the likelihood that the

actor agrees to uphold it, i.e., $P(Y_{ij} = 1)$, while γ_j is the location of the cut point of decision j, i.e., the point that discriminates between "yes" and "no" votes in the common space.

We have three sets of information: First, we have the dichotomous outcome of a court decision, which indicates whether constitutional provisions are violated or not. A decision outcome represents a collective "vote" of the court, not to be confused with individual judicial votes, which may not always be available.[2] Second, we have briefs submitted by various political actors taking a position ("votes") on whether constitutional provisions are violated or not. These briefs are reported per decision. Therefore, we compare "votes" of the court and various political actors on a decision level. Third, in some decisions, a referral is concerned with a federal law, and we have information about the political actor who introduced this law. It is reasonable to assume that this actor argues that the law does not violate constitutional provisions. Hence, we have a reported position ("vote") for this actor, too.

We collect this information in a vote matrix. Each column of this matrix represents a single decision, and the rows indicate the "votes" on the decision's outcome. The first row is the collective "vote" by the justices of the court. The court (or a respective panel of justices) can "vote" whether constitutional provisions or provisions of international law are violated (= 1) or not violated (= 0). Additionally, we add a row to our vote matrix for each political actor submitting a brief or observation that argues whether a provision is violated (= 1) or not (= 0). All this information is pooled to estimate a common policy space, assuming that the particular policy space underlying the preferences of each actor is structured in the same way. Specifically, we "glue" the different spaces together by

[2] Sometimes, highest courts combine related referrals into one decision, and although there are nuances across different referrals, the overall outcome of the court decision has implications for all referrals.

assuming that β_j and γ_j is the same in each actor's utility function. This assumption is justified at least to the extent that the court and the political actors respond to the exact same legal question and present their opinion in legal language.

Scaling the German Federal Constitutional Court

In this section, we present the German Federal Constitutional Court (GFCC) as a suitable case to implement our scaling approach and outline our estimation procedure.

The GFCC is a powerful constitutional court in Europe and a typical Kelsenian court, ruling exclusively on constitutional matters (Kelsen 2008 [1931]; Vanberg 2005; Engst 2021, Ch. 3). Although separate opinions are allowed, individual judicial votes are very rarely published (Wittig, 2016), making scaling based on individual votes unsuitable. Therefore, the GFCC is an ideal case to illustrate our scaling approach.

The GFCC consists of two Senates with a panel of eight justices each, serving a non-renewable term of 12 years. Half are elected by Germany's first legislative chamber (*Bundestag*) and the other half by the second chamber (*Bundesrat*). To be elected, justices require a two-thirds majority and must meet a number of professional requirements. The legislative chambers take turns electing the chief justice and the deputy chief justice, who each preside over a Senate (§2 to §9 Act on the Constitutional Court [AoCC]). To manage the GFCC's heavy caseload, each Senate has chambers (panels of three justices) that decide unanimously on individual complaints and concrete reviews initiated by ordinary courts. The chambers cannot declare a referred law unconstitutional. Thus, controversial issues are resolved by the Senates. This is why we only use Senate decisions for our scaling approach.

Different procedures allow different plaintiffs to refer to the GFCC (see Art. 93 German Constitution). Constitutional complaints are the most common procedures (about 96 percent of all referrals since 1951). They permit individuals affected by a public act or law to refer to the GFCC after exhausting all legal remedies. The second most common

procedure is the law review, either initiated by ordinary courts as concrete reviews or initiated by the federal government, a state government, or one-quarter of the members of the *Bundestag* as abstract reviews. Other procedures include disputes between high state organs, disputes between federal states, or electoral complaints.[3] Both Senates frequently combine multiple proceedings of similar nature into one decision.

To illustrate our scaling approach, we assess all 584 main decisions published between Germany's 12th (beginning December 1990) and 16th legislative term (ending October 2009).[4] We extract information about Senate decisions from the Constitutional Court Database (CCDB), which includes decision characteristics, justice details, and information about the political context (Engst et al. 2020; see also appendix B). We supplement the data with information about briefs extracted from the decision texts.

The assessed period saw five federal governments of three ideological blocs governing Germany. The conservative Christian Democrats (CDU/CSU) were in coalition with the smaller Liberal Party (FDP) from 1990 to 1994 and from 1994 to 1998. The Social Democrats (SPD) were in coalition with the smaller Green Party from 1998 to 2002 and from 2002 to 2005. Finally, the CDU/CSU governed in a grand coalition with the SPD from 2005 to 2009. Thus, we estimate the GFCC's positions during center–right, center–left,

---

[3] See the GFCC's annual report at `https://www.bundesverfassungsgericht.de/SharedDocs/Pressemitteilungen/EN/2022/bvg22-014.html` (last access 08/04/2022).

[4] We focus on decisions the GFCC labels as main decisions (*Hauptentscheidung*). They regulate substantive matters. Other decisions are, e.g., provisional orders temporarily regulating referrals but eventually requiring main decisions, requests to exclude justices from partaking in main decisions, or the reimbursement of expenses following main decisions. Information on the type of decision is included in the introductory part of decisions.

and centrist grand coalition governments. Additionally, we ensure institutional stability on the GFCC's side, since Germany's reunification was formally completed in 1990.

Consider a law referred to the GFCC for review. The justices apply two steps to make a collective ruling. First, they assess its admissibility. This is a technical assessment of whether a referral meets the requirements for substantive review. Second, the justices decide on the merits and substantively assess the law's constitutionality. We use the rulings on the admissibility and on the merits to code the decision outcome. If the outcome is that a referral is (partially) admissible *and* (partially) justified, then it violates constitutional provisions. Thus, the justices "vote" (collectively) against the law's constitutionality. In all other instances, the law stands and the justices of the respective Senate "vote" (collectively) for its constitutionality. Out of the 584 decisions made, 298 "votes" find a constitutional violation, while 284 "votes" find no violation. Two decisions are excluded because it is not possible to code the decision outcome.[5]

Our proposed scaling approach is also based on data regarding how scalable political actors would "vote" on the constitutionality of the same issue. We obtain this information from their submitted briefs, which are subsequently summarized in a court's decision and written in a similar legal language. We assume that political actors do not strategically oppose their sincere interests. In legislative politics, actors may act strategically and trade votes or merge bills when negotiating policy outcomes. However, in the judicial sphere rules of procedure are different. First, courts may make decisions in small panels before political actors have the opportunity to file briefs. Second, courts have full authority to merge cases leaving no room for package-deals among political actors. Third,

[5] Scholars often only account for the ruling on the merits to estimate a court's position (e.g., Martin and Quinn, 2002; Lauderdale and Clark, 2014). We combine admissibility and merits rulings to reflect the legal procedures in Germany. In appendix D we use merits rulings only and show that our substantive findings remain unchanged.

courts have full authority over the final decision. Thus, political actors cannot credibly commit to specific outcomes. Consequently, political actors lack incentives to strategically misrepresent their positions.

The procedural regulations to submit briefs are not conclusive (Kühne, 2015, 319). In abstract or concrete reviews, briefs may be filed by both legislative chambers, the federal government, state governments, or parliaments (§77 & §82 I, II AoCC). In conflicts between state or federal organs, all affected organs may join the proceeding (§65 I & 69 AoCC). The GFCC also invites political actors whose acts are addressed in constitutional complaints to file briefs (§94 AoCC). Since all these proceedings constitute the majority of the GFCC's caseload, political actors can submit briefs for most referrals.

However, political actors are not obligated to file briefs, and the absence of provisions empowering them to do so does not necessarily imply that they cannot express their opinions (cf. Kühne, 2015, 319). Nevertheless, it is reasonable to assume that political actors invest resources and file briefs in decisions that are politically relevant. These decisions are particularly suitable for our purpose of scaling courts in political terms. Finally, justices receive briefs before they make a decision, which means they may update their judgment based on them. However, a causal explanation for why the GFCC takes a particular position is not required for the application of our scaling approach, as we only aim to locate the GFCC as a collective actor in a common policy space.

We extracted briefs from the full decision texts using regular expressions in *R*. We then drew 100 briefs submitted by political actors in Senate decisions published between 1973 and 2010. Two coders with a background in public law and political science classified the briefs in a double-blind process. They evaluated whether a brief argues that any referral violates constitutional provisions or not. The intercoder reliability was 96 percent. Following the training, the coders classified 695 unique briefs fielded by political actors in 421 out of the 582 decisions included in our data. Afterwards, one of the authors reviewed all briefs. In total, 301 briefs were filed by the federal government, 349 by different state

governments, 13 by the *Bundestag*, 13 by the *Bundesrat*, and 19 by political parties. We excluded 60 neutral briefs, as they were uninformative for scaling.

Finally, our data includes the date on which federal laws referred to the GFCC were initially passed in the *Bundestag* and the actor who presented the law. For federal laws passed after 1972 and referred to the court, we assume that the political actor who presented a law "votes" that the law is constitutional.[6] This "vote" is added to our vote matrix similar to briefs filed by political actors. The federal government introduced 162 of the laws referred to the GFCC, members of the *Bundestag* 73, and members of the *Bundesrat* 6. Incorporating the additional information into our vote matrix increases the "votes" by political actors to 863 in 446 of our 582 decisions. [7]

Usually, scaling approaches require non-unanimous votes (cf. Martin and Quinn, 2007). Since we use informative priors for all political actors, scaling the GFCC also works when we have unanimous votes. Even unanimous votes carry information when utilizing prior information for (some) scalable actors: we learn that the GFCC and all political actors "vote" the same and, hence, are on the same side of this decision. Unanimous votes occur in 204 main decisions. 498 votes occur in 242 main decisions in which at least two actors oppose each other. The latter set of decision makes up 42 percent of all decisions. They are most informative, as we can estimate the cut point and discrimination parameter more accurately, and prior assumptions are less influential. In total we use about 77 percent of the 582 main decisions published by the GFCC between 1990 and 2009 to scale

---

[6] Information on federal laws is taken from the dataset summarizing Germany's legislative proceedings (initially published by Burkhart 2008 and extended by Stecker 2016), which is available from 1972 onward.

[7] In 13 cases, the political actors who introduced a law under review also filed a brief. The respective briefs always spoke in favor of the constitutionality of the law, and we thus considered the actors' "vote" only once.

the GFCC. This is a major improvement considering that only about 7 percent of the decisions in our sample include separate opinions by individual justices. Thus, there is insufficient information to scale the GFCC based on individual votes. Moreover, in a European comparison, the highest courts in Austria, Belgium, France, Italy, Luxembourg, and Malta are not even allowed to publish individual votes, while briefs by political actors are available in all these countries (see also appendix A).

In sum, our scaling approach using outcomes of decisions and briefs by political actors can benefit research on courts that do not publish individual judicial votes. Briefs are commonly available while individual votes are not, as a comparison of highest courts in Europe shows (see appendix A). In the next section, we apply our scaling approach to the GFCC.

Common policy scores for the GFCC

To estimate our model from equation 1, we compute a vote matrix for the GFCC. Each column represents a single decision, and the rows indicate the "votes" on the decision's outcome by the GFCC and various political actors. We use the vote matrix and `R STAN` to estimate our model. Remember that our aim is to map the position of the GFCC in a common policy space with political actors. To achieve this, we use prior information about the position of the political actors who submitted the briefs. In particular, we assign political actors posterior values from the Manifesto Common Space Scores (MCSS, König, Marbach and Osnabrügge 2013) for German parties. We take into account that briefs are often filed by collective actors, such as governments. Prior information about governments is generated via a linear combination of the posterior MCSS distribution of each governing party, weighted by the party's share of government portfolios. Similarly, we estimate the position of the *Bundestag* and *Bundesrat* as seat-weighted linear combination of the posterior MCSS distribution of each parliamentary party.

Finally, we use Markov Chain Monte Carlo (MCMC) to simulate the posterior distribu-

tion of a Senate's ideal point. We base our estimates on 160,000 iterations, with the first 80,000 iterations omitted as warmup. Although we have strong prior information about the position of the political actors, our assumption about a Senate's ideal point is weakly informative. We assume that each Senate is a non-extreme actor within the space covered by the parties in the *Bundestag*. This is a necessary assumption to estimate ideal points using an IRT model. In legislative research, extreme members of parliament are placed on either end of a space to identify the scale on which other members are situated. This is of little concern when scaling non-majoritarian institutions, such as courts. Justices are unlikely to represent extreme positions, as they are often elected to highest courts via supermajorities, which requires broader consensus among legislators.

To express this as prior information, we assign the GFCC a—weakly informative—prior ideal point drawn from a normal distribution with a standard deviation of two. Thus, the GFCC's ideal point is assumed to be anywhere within the ideological space encompassing all ideal points of the parties in the *Bundestag*. We calculate the respective mean across the positions of the parties that nominate each justice to assign a prior mean to each Senate.[8] Weakly informative priors are also set for the cut point and the discrimination parameter, drawn from a standard normal distribution.

In the next section we present the estimated positions and explain how we extend existing research on the GFCC using those positions. Afterwards, we illustrate the general applicability of our approach by outlining how it can be used to estimate common policy scores for the European Court of Justice.

---

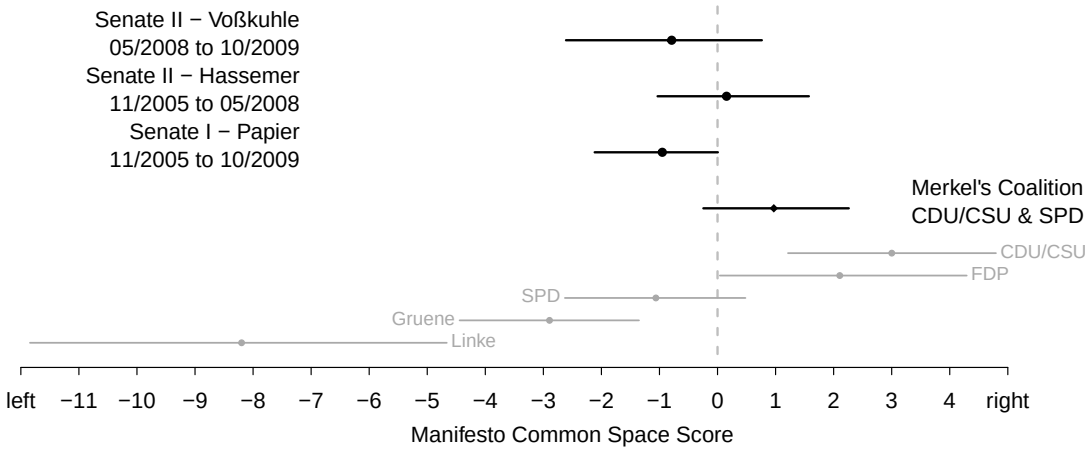
[8] Leveraging external information as prior information is common practice when estimating ideal points (e.g., Martin and Quinn, 2002). However, we also estimate ideal points using weakly informative priors by assigning justices a prior mean of zero. This does not change our substantive findings, as demonstrated in appendix D.

Validity of the common policy scores for the GFCC

In this section we present the common policy scores for the GFCC. First, we discuss their face validity. Second, we assess how the scores benefit existing research in judicial politics that predicts when justices will hold oral hearings (see Vanberg, 2001; Krehbiel, 2016).

In essence, we estimate a position for each Senate decision in our vote matrix. However, as the number of actors who "vote" per decision is low, the credible intervals are large. Thus, we aggregate the positions. First, we aggregate by the tenure of German chancellors. The three chancellors included in our data—Kohl, Schroeder, and Merkel—led coalition governments representing different ideological blocs. Second, within the chancellors' terms, we aggregate by Senate chairs in the GFCC. The first aggregation accounts for changes in the political environment and the second for changes in the judicial environment. Moreover, aggregating at the level of justices who chair the Senates within a chancellor's term allows us to assess the face validity of our scores. The German media frequently publish expert opinions on the Senates' behavior over the term of chairing justices. These opinions allow us to derive expectations about the Senates' positions, and we can compare these expectations to our estimates.

Figure 1: Positions of the Senates of the GFCC in a common policy space with Merkel's governing coalition



To assess face validity, we focus on ideal point estimates for the Senates during

Chancellor Merkel's term, presented in figure 1. Positions for the other terms are available in appendix C. The First Senate, chaired by Justice Papier, was often in conflict with Merkel's grand coalition. When Papier retired, a major German newspaper even titled an article "The Troublemaker" (Müller, 2010), highlighting that the First Senate frequently clashed with the governing branches. The GFCC made several high-profile decisions that addressed major reforms and led to political tensions with the "omnipotent justices" (Wefing, 2010, own translation). The decisions included increasing welfare spending and strengthening individual rights. Shortly into Merkel's second term, the justices directed the government to raise unemployment benefits for those affected by the labor market reform Hartz IV (BVerfGE 125, 175). Additionally, the justices rejected the Aviation Security Law, which allowed for the downing of hijacked planes (BVerfGE 115, 118) and limited the authorities' powers to secretly investigate private computer systems (BVerfGE 120, 274). The complaints voiced by political actors suggest that the First Senate tended to be more progressive (left-leaning), on average, than the government.

The Second Senate acted as a referee in the conflict between the political branches, though it was partially at odds with Merkel's coalition. Justice Hassemer chaired the Second Senate from 2005 to 2008, followed by Voßkuhle from 2008 to 2009. During their tenures, the Second Senate invalidated regulations concerning the unequal distribution of parliamentary seats (BVerfGE 121, 266) and tax policies disadvantaging smaller local parties (BVerfGE 121, 108). The Senate also rejected the use of voting machines (BVerfGE 123, 39; Kneip 2015, 291). While the former decision caused judicial–political conflict, the latter did not. In conflicts with Merkel's government, the Second Senate often sided with the legislature (BVerfGE 124, 78; BVerfGE 124, 161) and was often characterized as a "referee" (Kneip, 2015, 294). An alliance among conservative justices formed under Hassemer, but Voßkuhle broke it early on (Janisch, 2020). Initially, Voßkuhle was described as a strong yet quietly acting justice (Wefing, 2010). This changed after the Second Senate's decision on Europe's Lisbon Treaty (BVerfGE 123, 267), which drew widespread criticism

(Müller 2009; Fischer 2009; Kneip 2015, 286-287). Thus, we expect that the Second Senate drifted away from the position of Merkel's coalition over time as the conservative alliance dissolved and the Lisbon decision was made. Given its role as a referee, the distance between the Second Senate and Merkel's government should not be as significant as that between the First Senate and Merkel's government. We therefore expect the First Senate to be, on average, further from the government than the Second Senate.

Figure 1 shows the positions of both Senates (round, black estimates) and the position of Merkel's government (diamond, black estimate) along with the corresponding 95 percent credible interval. The manifesto scores of the parties in the first chamber are shown in gray. What can we infer from figure 1? First, compared to the party system, neither Senate takes an extreme position. This is plausible considering that the major German parties must agree on judicial nominees, which should result in a centrist court. Second, both Senates behave as expected when compared to the government. With parliamentary parties having a say in selecting the justices, the Senates are somewhat more progressive than the moderately conservative government. The First Senate tends to the left of the government, consistent with its characterization as a troublemaker (Müller, 2010). Moreover, the Second Senate is, on average, closer to the government than the First Senate, and after 2008, the Second Senate shifted slightly (although insignificantly) to the left. These positions align with our expectations.

The descriptive assessment suggests that the positions of the Senates are plausible. We would expect the GFCC to generally take a centrist position, and the positions reflect the nuances in the GFCC's behavior as suggested by observers. In the next section, we use our estimated positions to study why courts hold oral hearings.

Applying the common policy scores from the GFCC

In this section, we use our scores to assess an important question in court–executive relations: under what conditions do courts hold oral hearings (Johnson, Wahlbeck and

Spriggs, 2006; Krehbiel, 2016; Vanberg, 2005)? We also compare our scores to the current best practice to measure positions of courts when individual votes are not available.

Courts with constitutional review powers limit the political power of the legislative majority and the executive. However, courts cannot enforce decisions themselves and have to rely on the other branches of government to comply with the decisions. An important strategy to encourage compliance is to raise the government's cost of noncompliance through public hearings. Hearings increase the public awareness of a decision, as the media is likely to report them. One can expect that as media attention increases, so will the government's fear of negative public reaction as long as courts enjoy a high level of public support (Krehbiel, 2016; Vanberg, 2001, 2005). Krehbiel (2016) shows that the GFCC is more likely to hold oral hearings when political compliance is at risk, as the government defends the constitutionality of a referred law in a brief.

In addition, scholars argue that oral hearings serve not only to raise awareness and address policy compliance risks but also to gather information (Johnson, Wahlbeck and Spriggs, 2006, 99). However, the credibility of information gathered at such hearings depends on the source. Studies suggests that information provided by a government in briefs and hearings is more credible to the court the closer the government and the court are ideologically (Bailey, Kamoie and Maltzman, 2005; Johnson, Wahlbeck and Spriggs, 2006). Conversely, a court may perceive information from a more distant government as less credible (Bailey, Kamoie and Maltzman, 2005; Segal and Spaeth, 1993, 313). Thus, when courts utilize hearings as a means of gathering information, they may do so regardless of the perceived risk of government noncompliance, as the incentives to enforce compliance and to gather information are different.

To identify the functions of oral hearings, we start with replicating Krehbiel's (2016, Model 1) logistic regression model as a baseline. Thus, we also take the occurrence ($= 1$) or non-occurrence ($= 0$) of oral hearings as the dependent variable. We then extend his analysis of the GFCC by introducing a new covariate: the ideological distance between

the GFCC and the government in a given decision. This variable allows us to examine the hypothesis that *the closer the government and the GFCC are ideologically, the more likely the GFCC is to schedule oral hearings* (H1). We test this hypothesis while acknowledging Krehbiel's (2016) finding that the GFCC is more likely to schedule oral hearings when the government defends the constitutionality of the law under review in a brief.

To operationalize the distance between the government and the GFCC, we use two strategies: First, our vote matrix approach assigns the GFCC the estimated position of the respective Senate based on our model from equation 1. Second, using the party label approach, we assign each justice the MCS Score of the party nominating them. We then use the mean justice position of the respective senate members as a measure for the GFCC's position (Hönnige, 2009). Next, we calculate the coalition government's position as a convex combination of the respective party positions drawn from the posterior distribution of the MCSS (König, Marbach and Osnabrügge, 2013), weighted by each party's portfolio share. Finally, we measure the distance between the government and the GFCC by taking the absolute difference between the two positions, as they are measured on the same scale. A negative regression coefficient of the distance indicates that the GFCC is more likely to schedule oral hearings the closer the government is ideologically.

Table 1 summarizes three models. Model 1 confirms the subsample stability of Krehbiel's (2016, Model 1) findings. In decisions without government briefs, the probability of a hearing is systematically lower. This finding is consistent with his hypothesis that hearings are used to raise public awareness. In Model 2, we introduce the absolute distance between the GFCC and the government, measured using the party label approach. The estimate is positive (against our expectation) but not significant at conventional levels, indicating that the ideological distance between the GFCC and the government does not systematically increase the likelihood of oral hearings. However, this conclusion appears premature when calculating the distance based on our vote matrix approach, introduced in model 3. The respective estimate of the distance is systematically negative,

Table 1: Logit analysis to predict under what conditions public oral hearings occur

	Model 1	Model 2	Model 3
	Baseline Krehbiel (2016)	Party label approach	Vote matrix approach
Distance GFCC - govt.		0.11	-0.54^{***}
		(0.19)	(0.18)
No government brief (=1)	-1.18^{**}	-1.2^{**}	-1.06^{**}
	(0.49)	(0.5)	(0.49)
Controls	✓	✓	✓
N	313	313	313
logLik	−163.53	−163.37	−158.75

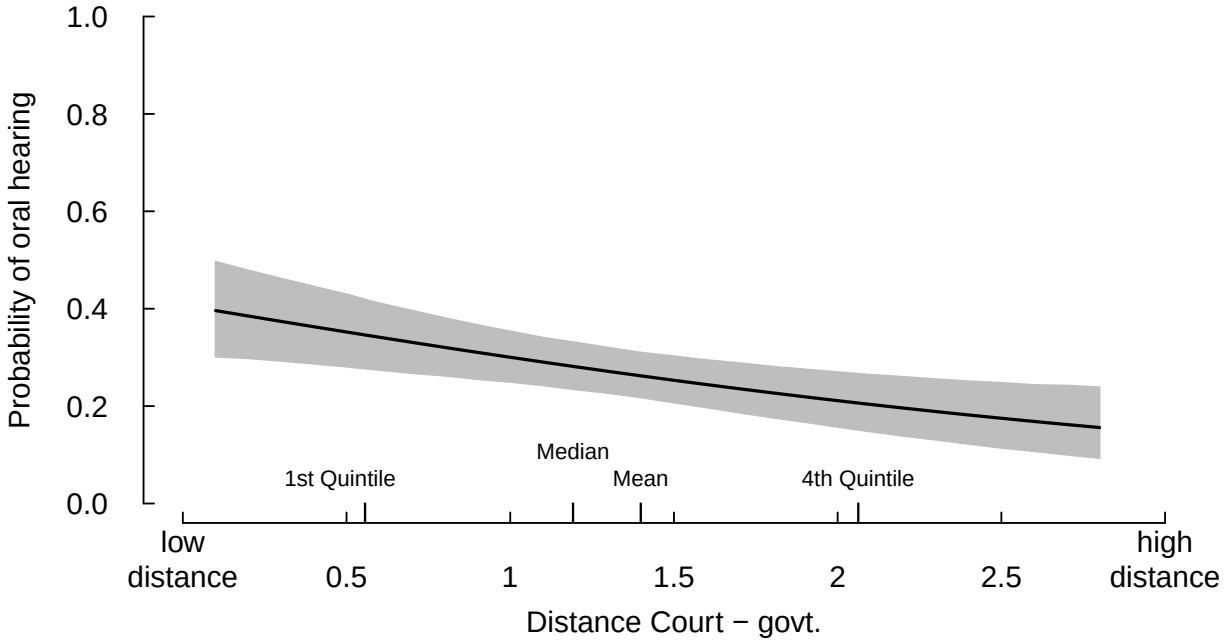
Note: 84 observed oral hearings in 313 decisions. Standard errors in parenthesis. Constant and controls reported in table E.1 in appendix E. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. The baseline model is a replication of model 1 in Krehbiel (2016), re-estimated on a subsample of 313 decisions. Model 2 and 3 include distance measures based on the party label approach and our vote matrix approach respectively.

indicating that the smaller the distance between the GFCC and the federal government, the more likely it is that oral hearings will be held.[9] This holds true while also supporting Krehbiel's (2016) substantive conclusions (with similar estimated coefficient sizes). Thus, we find evidence that the GFCC schedules oral hearings strategically to raise public awareness when political compliance is at risk, but also to gather credible information.

Figure 2 shows the effect of the absolute distance between the GFCC and the government in a common policy space, as calculated using our vote matrix approach, on the probability that oral hearings are held. As the distance increases, the probability decreases. In the first quintile of our distance measure, the probability is about 35 percent, while it decreases to about 20 percent in the fourth quintile. This suggests that the objective of holding oral hearings is not just to raise awareness when political compliance is at risk. Instead, the effect is consistent with the notion that courts also use oral hearings to gather credible information. The GFCC is more likely to schedule oral hearings when the government aligns closely with it in a common policy space, regardless of whether the government submits a brief or not.

[9] Appendix E shows that our results are robust when accounting for measurement error.

Figure 2: Predicted probabilities of public oral hearings based on the distance between the GFCC and the government in a common policy space



Note: The black line indicates predicted probabilities of a hearing. The gray area indicates the corresponding 95 percent credible intervals using an observed value approach based on the results from model 3.

Our analysis extends Krehbiel's (2016, 999) findings on the function of oral hearings at the GFCC. Our scaling approach enables us to operationalize the distance in a common policy space between courts and political actors without relying on strong assumptions—such as those required by the party label approach (e.g., Hönnige, 2009; Brouard and Hönnige, 2017)—and without requiring individual judicial votes, which are often unavailable. With our vote matrix approach, we provide new evidence on the strategic role of oral hearings as a means of gathering information. In the following section, we apply our scaling approach to the European Court of Justice as a significant international court, further demonstrating the generalizability of our approach.

Scaling the European Court of Justice

In this section, we estimate the ideal points of the European Court of Justice (ECJ) and confirm the common finding that the ECJ favors European integration (Carruba, Gabel

and Hankla, 2012; Larsson and Naurin, 2016). Then, we use the scores to extend the research by Larsson and Naurin (2016) on the role of compliance in ECJ decision-making.

The ECJ is the highest court in the European Union (EU), with each EU member state providing one judge. The ECJ ensures that the EU member states abide by EU law and that EU law is applied uniformly across all member states. The justices are assisted by eleven advocates general, who prepare non-binding, advisory opinions on questions referred to the ECJ concerning new points of law (Carrubba, Gabel and Hankla, 2008, 447-449). The ECJ is an excellent case to illustrate the generalizability of our scaling approach. First, it is an international court whose decisions have major implications for all EU residents. Second, since the ECJ does not publish individual judicial votes, our scaling approach is helpful in estimating the ECJ's position.

The ECJ is ascribed to strongly support European integration in decisions referred to as preliminary references (Carruba, Gabel and Hankla, 2012; Larsson and Naurin, 2016). Preliminary references are initiated by national courts seeking guidance from the ECJ on the interpretation of EU law (Carruba, Gabel and Hankla 2012, 217; Larsson and Naurin 2016, 391-392). If the ECJ indeed supports more European integration, it should position itself accordingly in a policy space ranging from more national sovereignty to more European integration. Our scaling approach allows us to assess this claim.

To implement our approach, we require three pieces of information. First, we need the ECJ's collective "vote" on a referred question. Questions are functional equivalent to referrals at the GFCC. Larsson and Naurin (2016) provide data that classifies the opinions as favoring more European integration (pro-integration), preserving national sovereignty (anti-integration), or being ambivalent. Second, to anchor the judicial opinions in a common space with EU member states, we require opinions of different EU member states on the referred question. These opinions are summarized in the *Report for the Hearing* by the reporting justice on a decision. Larsson and Naurin (2016) also classify these opinions

as favoring European integration, preserving national sovereignty, or being ambivalent.[10] In addition, the opinion of the European Commission on a referred question is classified accordingly. The opinions of the EU member states, the European Commission, and the ECJ allow us to compute the vote matrix required for our scaling approach. Each column in our data represents a referred question, and the rows indicate the "votes" on the question by the ECJ, the EU member states, and the European Commission.

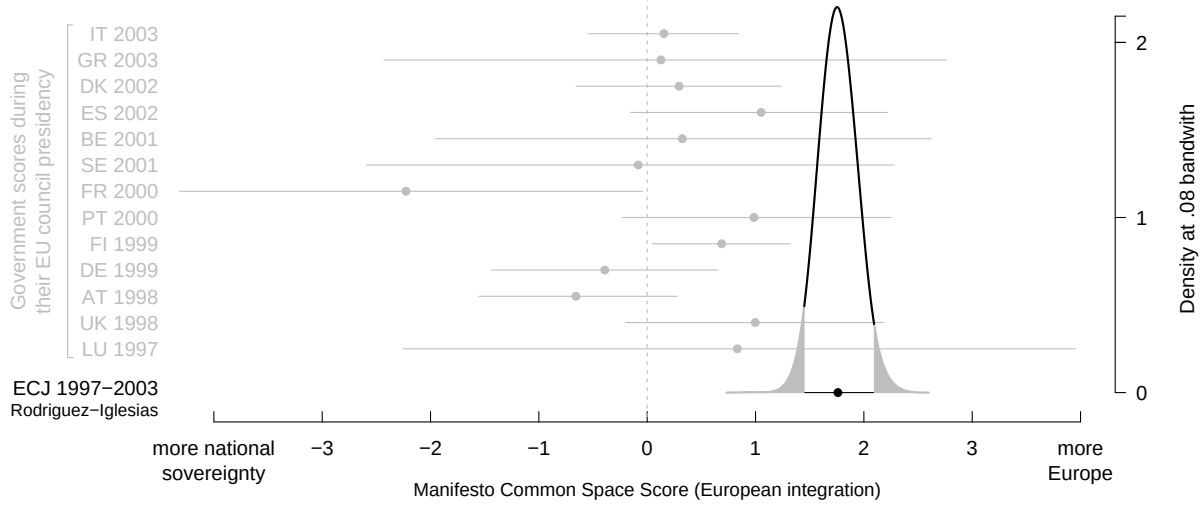
Third, we need the manifesto scores of the EU member states on European integration as prior information to anchor our estimated positions in a common policy space. These member state positions are represented by the respective member state government, calculated as a convex combination of the positions of each government party drawn from the posterior distribution of the MCSS on European integration (König, Marbach and Osnabrügge, 2013) and weighted by each party's share of government portfolios. In the next section, we apply our scaling approach to the ECJ.

Common policy scores for the ECJ

The common policy scores for the ECJ, similar to the GFCC scores, are estimated per presidency using the IRT model from equation 1. Our sample, using Larsson and Naurin's (2016) data, includes 1,241 questions that meet two conditions. First, the ECJ did not make an ambivalent decision and, second, at least one actor—the European Commission or a member state—filed a "vote" opposing the ECJ. The sample consists of one third of all questions included in the original data. We can identify MCS Scores for 65 of 108 governments that filed 2,155 "votes", which constitutes 74 percent of all government "votes" in our sample. We use weakly informative zero priors for the remaining 43 governments, three EU Commissions, and two ECJ presidencies. Higher MCS scores for European integration indicate a preference for European integration and lower scores

[10] Please refer to Larsson and Naurin (2016, 392-393) for details on their coding procedure.

Figure 3: Position of the ECJ in a common policy space with EU member state governments



indicate a preference for national sovereignty. As all "votes" in our data are coded based on whether the actors support more European integration (= 1) or not (= 0), we can constrain the discrimination parameter to be non-negative.

The density curve in figure 3 illustrates the distribution of the resulting common policy score estimate for the ECJ from 1997 until the end of Justice Rodriguez-Iglesias' presidency in 2003. In figure F.1 of appendix F, we show equivalent results for the presidency of Justice Skouris (until 2007). We find that the ECJ consistently favors European integration (black estimate) over national sovereignty. We can also directly compare our estimates to the MCS Scores on European integration of the EU member states. Figure 3 includes the positions of the EU member state governments at the time of their EU Council presidency (gray estimates). In line with the literature (Carruba, Gabel and Hankla, 2012; Larsson and Naurin, 2016; Krehbiel, 2021), we find that the ECJ is on average more supportive of European integration than the governments. This supports the validity of our scores.

Applying the common policy scores from the ECJ

In this section, we use the estimated ECJ positions to assess the ECJ's decision-making. Larsson and Naurin (2016, 382-386) argue that the justices rely on signals by national

governments to assess their willingness to comply with rulings. These signals are included in written and oral "observations" submitted by the governments. Observations are functional equivalent to briefs. To estimate the effect of government observations, the observations should be weighted by a respective government's vote share on the European Council. More votes on the council make a signal stronger, as a government has a higher impact on European policymaking (Larsson and Naurin, 2016, 389-390). These considerations imply that *the more observations favor more national sovereignty, the more likely the ECJ will rule against more European integration* (H1).

This perspective, however, does not account for the weight that the ECJ places on the policy under review. Our scaling approach allows us to incorporate this perspective, extending the study by Larsson and Naurin (2016). Figure 3 supports the assumption that the ECJ significantly favors more European integration over more national sovereignty. Thus, we expect the ECJ to rule in favor of integration, in particular when a decision would substantially increase integration. In spatial terms, the ECJ wants to draw a policy under review close to its ideal point. This should influence its rulings when national governments file opinions supporting more national sovereignty and signal noncompliance with opposing rulings. The assumption is that with increasing distance of a policy, the concerns about that policy outweigh the ECJ's concerns about noncompliance. Our scaling approach allows us to assess this assumption. We hypothesize that *the ECJ is more likely to rule in favor of more European integration if the distance between the ECJ to a policy under review increases and the member states file observations in favor of more national sovereignty* (H2). This implies that the effect of member state observations favoring more national sovereignty decreases as the ECJ's distance to the policy increases.

To test H1 and H2, we utilize the data by Larsson and Naurin (2016) and require three variables. First, our dependent variable is an ordinal measure indicating whether the ECJ ruled in favor of more national sovereignty (=0), more European integration (=2), or whether the ruling was ambivalent (=1) (Larsson and Naurin, 2016, 393). Second,

Table 2: Ordered logit analysis to predict under what conditions the ECJ rules in favor of more Europe

	Model 1	Model 2
	Baseline Larsson/Naurin (2016)	with distance measure
Distance		-0.06
		(0.05)
MS anti	-2.05^{***}	-3.96^{***}
	(0.69)	(1.18)
Distance × MS anti		1.51^{**}
		(0.71)
Controls	✓	✓
N	3845	3835
logLik	-2865.51	-2854.21

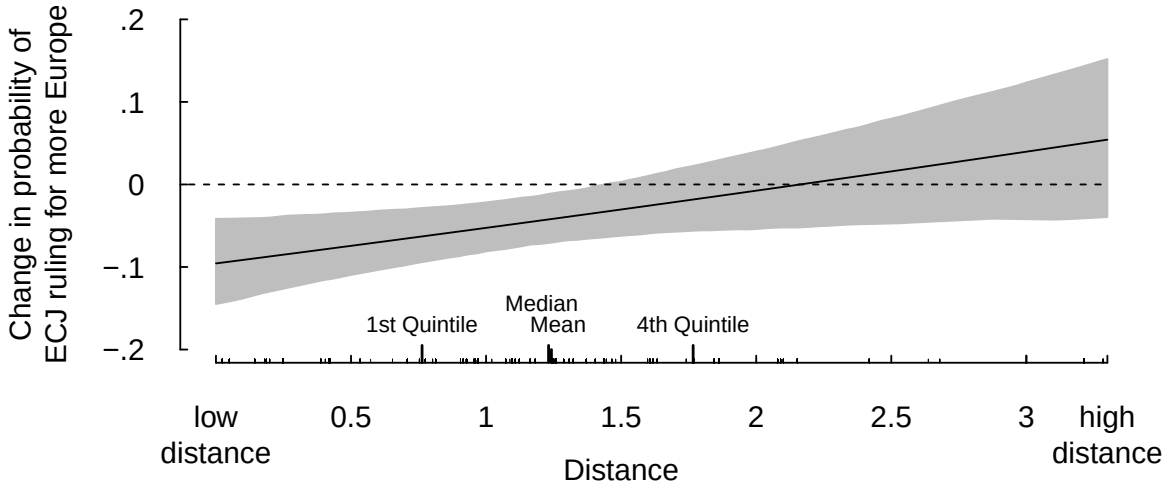
Note: Results from ordered logistic regressions. Standard errors in parenthesis; observations clustered by case. Cut points and controls reported in table H.1 in appendix H. *p<0.1; **p<0.05; ***p<0.01

Larsson and Naurin (2016, 392-393) categorize member state observations (weighted by Council votes) similar to ECJ rulings as favoring more national sovereignty, favoring more European integration, or being ambivalent. `MS anti` indicates the number of member state observations in favor of more national sovereignty as the share of all member states weighted by their Council votes. Finally, we approximate the `distance` between the ECJ and the policy under review. To do so, we calculate the absolute distance between the ECJ and the MCS Score on European integration of the national government that initiated the referred policy. The assumption is that the position of this national government best reflects the policy. We also use the control variables of Larsson and Naurin (2016).

Table 2 summarizes two ordered logistic regression models. The baseline model replicates model 3 in Larsson and Naurin (2016, 399) to predict ECJ rulings. The second model adds our distance measure as well as the interaction term between the observations of member states favoring more national sovereignty and our distance measure.[11] Both models confirm our expectations. Member state observations favoring more national

[11] For ten questions, we cannot estimate the distance between the ECJ and the member state that initiated the policy under review. The questions are excluded in model 2.

Figure 4: Difference in the probability of the ECJ ruling for more European integration depending on the share of member state observations favoring more national sovereignty and distance



Note: The black line indicates the mean of the differences in the predicted probability that the ECJ rules for more European integration when `MS anti` increases from 0 to 20 percent. The grey area indicates the corresponding 95 percent credible intervals using an observed value approach based on the results from model 2.

sovereignty decrease the likelihood of the ECJ ruling in favor of more integration (H1). The significant interaction term in model 2 illustrates that, as the distance increases, the negative effect of observations favoring more national sovereignty decreases. Thus, the ECJ pays less attention to the member state observations the larger the distance (H2).

Figure 4 illustrates the first differences in the ECJ's likelihood of ruling in favor of more European integration as the number of member state observations favoring more national sovereignty increases from zero to 20 percent. This threshold represents a "realistic shift" (Larsson and Naurin, 2016, 400). According to the first differences, the ECJ is significantly less likely to favor more European integration when the distance is small but becomes indifferent to ruling in favor of more integration when the distance increases. This implies that the influence of member state observations favoring more national sovereignty decreases as the distance increases. Why does the ECJ no longer take into account member states' signals of potential noncompliance as the distance increases? Two explanations seem plausible. First, when member states do not comply, this can

encourage additional litigation, providing the ECJ with further opportunities to promote European integration (Stone Sweet and Brunell, 2012). Second, the ECJ may choose to test the reactions to its decisions and adopt a more cautious approach in subsequent rulings (Carruba, Gabel and Hankla, 2012, 217). While we do not study these mechanisms here, our aim is to show that our scaling approach may encourage further research.

Conclusion

How can we identify the positions of non-majoritarian actors in a common policy space? To comprehensively assess spatial models of inter-institutional interactions, it is often necessary to compare the positions of political actors with the positions of non-majoritarian actors, such as courts. In this study, we have presented a scaling approach that allows us to map different highest courts in a common policy space with political actors.

In contrast to previous research, our strategy relies neither on individual judicial votes (e.g., Martin and Quinn, 2002; Epstein, Martin, Segal and Westerland, 2007; Hanretty, 2012*a*) nor on assumptions that justices have "inherited" positions from political actors who nominated them (e.g., Hönnige, 2009; Carrubba et al., 2012; Brouard and Hönnige, 2017). Instead, we leverage two features of published decisions that travel comparatively. These are, first, briefs that contain opinions of scalable political actors and, second, the courts' positions as implied by the collective decision outcome. This information allows us to compute a vote matrix similar to roll call data in legislative studies (e.g., Clinton, Jackman and Rivers, 2004; Bailey, 2007) and scale the actors in a common policy space.

We applied our approach to develop common policy scores for the German Federal Constitutional Court and the European Court of Justice. Both applications have produced new and valid position of the courts. This implies that scholars can use our method in contexts in which courts do not publish individual votes but political actors file briefs. Scholars no longer need to limit their analyses to the small number of countries in which individual votes are published to apply standard scaling approaches. Moreover,

our approach can be adapted to examine other non-majoritarian institutions that make collective decisions with policy implications, such as central banks or various regulatory agencies. As long as we have information about the outcome of decisions made by non-majoritarian actors and about how political actors position themselves regarding these outcomes (e.g., through analyzing speeches, statements such as briefs, or press releases), scholars can use our vote matrix approach to map *any* non-majoritarian actor in a common policy space with political actors.

We presented several validity checks demonstrating that the estimated scores align with our expectations from the literature. Additionally, we used the scores for the GFCC and the ECJ to establish the distance between these courts and various governments in common policy spaces. These new measures of ideological alignment between different institutions enable us to gain a better understanding of court–executive interactions. First, we have shown that the GFCC employs oral hearings not only to raise public awareness when political compliance is at risk but also to gather credible information. This extends Krehbiel's (2016) influential study. Second, we have shown that the ECJ favors European integration over national sovereignty. The ECJ is willing to ignore national signals of noncompliance with its rulings if a decision could strongly enhance European integration in its interest. This extends a prominent study by Larsson and Naurin (2016).

What is the advantage of our approach over traditional approaches to the comparative study of courts, such as the party label approach? Because we also rely on information included in briefs, our approach leverages more information from court decisions than traditional studies in comparative judicial politics currently do. It speaks to a more recent literature that takes judicial output more seriously (e.g., Lauderdale and Clark, 2014; Frankenreiter, 2017; Arnold, Engst and Gschwend, 2023). This helps to study the interaction of courts with political branches in a broader and interdisciplinary way.

References

Arnold, Christian, Benjamin G. Engst and Thomas Gschwend. 2023. "Scaling Court Decisions with Citation Networks." *Journal of Law and Courts* 11(1):25–44.

Baerg, Nicole and Will Lowe. 2020. "A textual Taylor rule." *Political Science Research and Methods* 8(1):106–22.

Bailey, Michael A. 2007. "Comparable Preference Estimates across Time and Institutions for the Court, Congress, and Presidency." *American Journal of Political Science* 51(3):433–48.

Bailey, Michael A., Brian Kamoie and Forrest Maltzman. 2005. "Signals from the Tenth Justice." *American Journal of Political Science* 49(1):72–85.

Bailey, Michael A. and Kelly H. Chang. 2001. "Comparing Presidents, Senators, and Justices." *Journal of Law, Economics & Organization* 17(2):477–506.

Bonica, Adam and Maya Sen. 2017. "A Common-Space Scaling of the American Judiciary and Legal Profession." *Political Analysis* 25:114–21.

Bonica, Adam and Maya Sen. 2021. "Estimating Judicial Ideology." *Journal of Economic Perspectives* 35(1):97–118.

Brouard, Sylvain and Christoph Hönnige. 2017. "Constitutional courts as veto players." *European Journal of Political Research* 56(3):529–52.

Burkhart, Simone. 2008. "Deutsche Bundesgesetzgebung 1972 - 2005." *GESIS Datenarchiv* ZA4569(1.0.0):Köln.

Carruba, Clifford J., Matthew Gabel and Charles Hankla. 2012. "Understanding the role of the European Court of Justice in European integration." *American Political Science Review* 106(1):214–23.

Carrubba, Cliff, Barry Friedman, Andrew D. Martin and Georg Vanberg. 2012. "Who Controls the Content of Supreme Court Opinions?" *American Journal of Political Science* 56(2):400–12.

Carrubba, Clifford J., Matthew Gabel and Charles Hankla. 2008. "Judicial behavior under political constraints." *American Political Science Review* 102(4):435–52.

Clark, Tom S. and Benjamin E. Lauderdale. 2012. "The Genealogy of Law." *Political Analysis* 20(3):329–50.

Clinton, Joshua, Simon Jackman and Douglas Rivers. 2004. "The Statistical Analysis of Roll Call Data." *American Political Science Review* 98(2):355–70.

Engst, Benjamin G. 2021. *The Two Faces of Judicial Power*. Cham: Palgrave Macmillan.

Engst, Benjamin G., Thomas Gschwend, Christoph Hönnige and Caroline E. Wittig. 2020. "The Constitutional Court Database." *Working Paper* .

Engst, Benjamin G., Thomas Gschwend, Nils Schaks, Sebastian Sternberg and Caroline E. Wittig. 2017. "Zum Einfluss der Parteinähe auf das Abstimmungsverhalten der Bundesverfassungsrichter – eine quantitative Untersuchung." *Juristenzeitung* 72(17):816–26.

Epstein, Lee, Andrew D. Martin, Jeffrey A. Segal and Chad Westerland. 2007. "The Judicial Common Space." *Journal of Law, Economics, & Organization* 23(2):303–25.

Epstein, Lee, Andrew D. Martin, Kevin M. Quinn and Jeffrey A. Segal. 2007. "Ideological Drift among Supreme Court Justices." *Northwestern University Law Review* 101(4):1483–541.

Epstein, Lee, Jack Knight and Olga Shvetsova. 2001. "The Role of Constitutional Courts in the Establishment and Maintenance of Democratic Systems of Government." *Law & Society Review* 35(1):117–64.

Epstein, Lee, Jeffrey A. Segal and Harold J. Spaeth. 2001. "The Norm of Consensus on the U.S. Supreme Court." *American Journal of Political Science* 45(2):362–77.

Fischer, Joschka. 2009. "Ein nationaler Riegel." *ZEIT Online* (Jul.,17).

Fischman, Joshua B. 2015. "Do the Justices Vote Like Policy Makers?" *Journal of Legal Studies* 44(S1):269–93.

Fischman, Joshua B. 2019. "Politics and Authority in the U.S. Supreme Court." *Cornell Law Review* 104:1513–92.

Fischman, Joshua B. and Tonja Jacobi. 2015. "The Second Dimension of the Supreme Court." *William and Mary Law Review* 57(5):1671–715.

Frankenreiter, Jens. 2017. "The Politics of Citations at the ECJ-Policy Preferences of E.U. Member State Governments and the Citation Behavior of Judges at the European Court of Justice." *Journal of Empirical Legal Studies* 14(4):813–857.

Hanretty, Chris. 2012*a*. "Dissent in Iberia." *European Journal of Political Research* 51(5):671–92.

Hanretty, Chris. 2012*b*. "The Decisions and Ideal Points of British Law Lords." *British Journal of Political Science* 43(3):1–14.

Hanretty, Chris. 2014. "The Bulgarian Constitutional Court as an Additional Legislative Chamber." *East European Politics & Societies* 28(3):540–58.

Hönnige, Christoph. 2009. "The Electoral Connection." *West European Politics* 32(5):963–84.

Janisch, Wolfgang. 2020. "Der Prozess." *Süddeutsche Zeitung* (May,5):3.

Jessee, Stephen. 2016. "(How) Can We Estimate the Ideology of Citizens and Political Elites on the Same Scale?" *American Journal of Political Science* 60(4):1108–24.

Jessee, Stephen A. 2009. "Spatial Voting in the 2004 Presidential Election." *American Political Science Review* 103(01):59–81.

Johnson, Timothy R., Paul J. Wahlbeck and James F. Spriggs. 2006. "The Influence of Oral Arguments on the U.S. Supreme Court." *American Political Science Review* 100(1):99–113.

Kelemen, Katalin. 2013. "Dissenting Opinions in Constitutional Courts." *German Law Journal* 14(8):1345–71.

Kelsen, Hans. 2008. *[1931] Wer soll der Hüter der Verfassung sein?* Tübingen: Mohr Siebeck.

Kneip, Sascha. 2015. Konfliktlagen des Bundesverfassungsgerichts mit den Regierungen Schröder und Merkel, 1998 - 2013. In *Handbuch: Bundesverfassungsgericht im Politischen System*, ed. Robert Chr van Ooyen and Martin H. W. Möllers. 2nd ed. Wiesbaden: Springer Fachmedien Wiesbaden pp. 281–95.

König, Thomas, Moritz Marbach and Moritz Osnabrügge. 2013. "Estimating Party Positions across Countries and Time." *Political Analysis* 21(4):468–91.

Krehbiel, Jay N. 2016. "The Politics of Judicial Procedures." *American Journal of Political Science* 60(4):990–1005.

Krehbiel, Jay N. 2021. "Public Awareness and the Behavior of Unpopular Courts." *British Journal of Political Science* 51(4):1601–19.

Kühne, Ulrich. 2015. *Amicus Curiae*. Tübingen: Mohr Siebeck.

Larsson, Olof and Daniel Naurin. 2016. "Judicial Independence and Political Uncertainty." *International Organization* 70(2):377–408.

Lauderdale, Benjamin E. and Tom S. Clark. 2014. "Scaling Politically Meaningful Dimensions Using Texts and Votes." *American Journal of Political Science* 58(3):754–71.

Lehmann, Pola, Tobias Burst, Theres Matthieß, Sven Regel, Andrea Volkens, Bernhard Weßels and Lisa Zehnter. 2022. *The Manifesto Data Collection. Manifesto Project (MRG / CMP / MARPOR). Version 2022a.* Berlin: Berlin Social Science Center.

Lowe, Will, Kenneth Benoit, Slava Mikhaylov and Michael Laver. 2011. "Scaling Policy Preferences from Coded Political Texts." *Legislative Studies Quarterly* 36(1):123–55.

Martin, A. D. and K. M. Quinn. 2007. "Assessing Preference Change on the US Supreme Court." *Journal of Law, Economics, and Organization* 23(2):365–85.

Martin, Andrew D. and Kevin M. Quinn. 2002. "Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the U.S. Supreme Court, 1953-1999." *Political Analysis* 10(2):134–53.

Matthieß, Theres. 2020. "Retrospective pledge voting." *European Journal of Political Research* 59(4):774–796.

Müller, Reinhard. 2009. "Wenn Karlsruhe weh tut." *Frankfurter Allgemeine Zeitung* (Nov.,13):10.

Müller, Reinhard. 2010. "Der Störenfried." *Frankfurter Allgemeine Sonntagszeitung* (Feb.,28):10.

Poole, Keith T. and Howard Rosenthal. 1991. "Patterns of Congressional Voting." *American Journal of Political Science* 35(1):228–278.

Poole, Keith T. and Howard Rosenthal. 2007. *Ideology and Congress: A Political Economic History of Roll Call Voting*. 2nd ed. New Brunswick: Transaction Publishers.

Raffaelli, Rosa. 2012. "Dissenting opinions in the Supreme Courts of the Member States." *European Parliament: Directorate General for Internal Policies* (PE 462.470).

Roberts, Jason M., Steven S. Smith and Stephen R. Haptonstahl. 2016. "The Dimensionality of Congressional Voting Reconsidered." *American Politics Research* 44(5):794–815.

Romer, Thomas and Howard Rosenthal. 1978. "Political Resource Allocation, Controlled Agendas, and the Status Quo." *Public Choice* 33(4):27–43.

Segal, Jeffrey A. and Albert D. Cover. 1989. "Ideological Values and the Votes of U.S. Supreme Court Justices." *American Political Science Review* 83(2):557–65.

Segal, Jeffrey A. and Howard J. Spaeth. 1993. *The Supreme Court and the Attitudinal Model*. New York: Cambridge University Press.

Stecker, Christian. 2016. "The effects of federalism reform on the legislative process in Germany." *Regional and Federal Studies* 26(5):603–624.

Stone Sweet, Alec and Thomas Brunell. 2012. "The European Court of Justice, state noncompliance, and the politics of override." *American Political Science Review* 106(1):204–13.

Vanberg, Georg. 2001. "Legislative-Judicial Relations." *American Journal of Political Science* 45(2):346–61.

Vanberg, Georg. 2005. *The Politics of Constitutional Review in Germany*. Cambridge: Cambridge University Press.

Wefing, Heinrich. 2010. "Der Mann, der schweigt." *Die Zeit* (Mar,18):6.

Wittig, Caroline E. 2016. *The Occurrence of Separate Opinions at the Federal Constitutional Court*. Berlin: Logos Verlag Berlin GmbH.

Online Appendix

Table of Contents

A Cross-European comparison of the publication of judicial votes

In this appendix, we aim to validate our claim that individual judicial votes are not widely available across countries, whereas briefs of political actors are. To do so, we randomly select ten decisions made by each constitutional court in the European Union in 2018 and count how often we can identify individual judicial votes for each justice in this sample. We then compare this number to the number of identifiable briefs of political actors in the same set of decisions. The results from these ten decisions are summarized in table A.1.

Table A.1: Judicial votes and briefs by political actors in EU member states with constitutional courts

Country	*Individual judicial votes, identifiable for each and every justice in N of 10 decisions*	*Brief by political actor or political actor is a participant with a clear opinion on the decision outcome in N of 10 decisions*
Austria	not allowed	6
Belgium	not allowed	7
Bulgaria	2	10
Croatia	0	7
Czech Republic	0	6
France	not allowed	10
Germany	0	6
Hungary	0	3
Italy	not allowed	10
Latvia	0	10
Lithuania	0	7
Luxembourg	not allowed	8
Malta	not allowed	5
Poland	0	8
Portugal	0	2
Romania	0	10
Slovakia	0	6
Slovenia	2	6
Spain	0	9

The count is based on information from ten decisions published in 2018 by the constitutional courts of the countries listed in table A.2. A value of zero in the second column indicates that there is insufficient information available to identify how each justice voted. The count also excludes unanimous votes, as they do not provide enough information to scale courts based on individual votes. Decisions that include separate opinions or unanimous votes are marked in table A.2. All countries except Austria, Belgium, France, Italy, Luxembourg, and Malta allow justices to make separate opinions publicly available (according to Wittig 2016, 153-154; Raffaelli 2012).

The second column of table A.1 shows that courts rarely publish individual votes, with some countries even not allowing their publication. As a result, common scaling approaches based on individual votes of justices (e.g., Hanretty, 2012*b*,*a*, 2014) are not sufficient to estimate positions of courts. Our scaling approach, in contrast, can be applied in all countries because, first, decisions are published in all democracies and, second, scalable political actors frequently submit briefs, as shown in the last column of table A.1. These findings validate our claim and show that our approach can be extended to other countries in order to locate constitutional courts in a common policy space with political actors.

Table A.2 summarizes the ten decisions we reviewed per country. When decisions were not available in German or English, we used Google Translate to translate the sampled decisions into English. The decisions were then reviewed to determine whether political actors submitted briefs with clear "votes" on cases. In addition, we assessed the availability of individual judicial votes.

Table A.2: List of decisions used in table A.1

Country			
Austria		**Belgium**	
Database			
`https://www.ris.bka.gv.at /Vfgh/`		`https://www.const-court.be /nl/judgments?year=2018`	
Special search parameter			
Erkenntnisse			
Decisions			
No. 1 on case 3/2017	No. 17 on case 9/2017	154/2018	130/2018
No. 14 on case 12/2017	No. 12 on case 1/2018	126/2018	122/2018
No. 2 on case 8/2017	No. 6 on case 10/2017	118/2018	85/2018
No. 7 on case 7/2017	No. 8 on case 13/2017	80/2018	74/2018
No. 11 on case 8/2018	No. 13 on case 3/2018	70/2018	40/2018
Country			
Bulgaria		**Croatia**	
Database			
`http://www.constcourt.bg /Search/Search.aspx`		`https://sljeme.usud.hr /usud/praksaw.nsf/`	
Special search parameter			
decisions		U-I, U-II, U-III	
Decisions			
UA3/2018	V 11/2018-8	UI-1574/2016, UI-1244/2017*	UI-439/2015
E547/2018*	E547/2018*	UI-3019/2018, UI-3337/2018*	U-II-2392/2014
G248/2017 ua*	G241/201*	U-II-872/2017	U-II-343/2015
G136/2017 ua*	V 97/2017-11	U-III-4868/2017†	U-III-272/2017†
KR1/2018 ua*	E1416/2018	U-III-432/2017*	U-III-871/2017†
Country			
Czech Republic		**France**	
Database			
`https://nalus.usoud.cz /Search/Search.aspx`		`https://www.conseil-constitutionnel.fr /les-decisions/annee/2018/type/dc`	
Special search parameter			
Excluding resolutions and opinions of *plenna*		*DC* and *QPC*	
Decisions			
Pl.US.28.18.1*	Pl.US.37.18.1	2017-757 DC	2018-763 DC
Pl.US.27.16.1*	Pl.US.4.18.1*	2018-761 DC	2018-770 DC
1.US.1099.18.1*	1.US.178.15.1	2018-777 DC	2017-689 QPC
2.US.482.18.1	3.US.647.15.1	2017-690 QPC	2018-701 QPC
2.US.1685.17.1	Pl.US.7.17.1*	2018-711 QPC	2018-699 QPC

Country			
Germany		**Hungary**	
Database			
www.bundesverfassungsgericht.de		https://alkotmanybirosag.hu /ugykereso	
Special search parameter			
Senatsentscheidungen		*decisions*, excluding orders, keyword: "észrevételeket"	
Decisions			
BVerfGE 150, 244*	BVerfGE 150, 309	II / 01483/2017*	IV / 00344/2016
BverfGE 149, 407	BVerfGE 149, 382	IV / 00987/2015	III / 02007/2017
BVerfGE 148, 11	BVerfGE 150, 1	IV / 01801/2016*	IV / 01096/2018
BVerfGE 149, 293	BVerfGE 149, 346	IV / 00339/2017*	IV / 01096/2018
BVerfGE 149, 222	BVerfGE 148, 296	IV / 01087/2017*	IV / 02111/2017

Country			
Italy		**Latvia**	
Database			
https://www.cortecostituzionale.it /actionPronuncia.do		https://www.satv.tiesa.gov .lv/cases/	
Special search parameter			
Excluding orders; constitutional legitimacy judgement in main way		*judgments*, no decision to close proceedings	
Decisions			
2018: 249	2018: 247	2018-12-01*	2018-10-0103*
2018: 238	2018: 210	2018-25-01*	2018-23-03
2018: 198	2018: 183	2018-21-01*	2018-17-03*
2018: 147	2018: 138	2018-15-01*	2011-01-11
2018: 107	2018: 68	2018-08-03	2018-07-05

Country			
Lithuania		**Luxembourg**	
Database			
https://www.lrkt.lt/lt/teismo-aktai /nutarimai-isvados-ir-sprendimai /138/y2018?exception=on		https://justice.public.lu/fr /jurisprudence/cour-constitutionnelle.html ?r=f%2Faem_first_released%2F2018&	
Special search parameter			
Decisions			
KT1-N1/2018	KT2-S1/2018	132/18	133/18
KT5-S2/2018	KT6-N4/2018	134/18	135/18
KT9-N5/2018	KT10-N6/2018	136/18	137/18
KT13-N8/2018	KT14-N9/2018	138/18	139/18
KT15-N10/2018	KT18-S8/2018	140/18	141/18

Country			
Malta		**Poland**	
Database			
`https://ecourts.gov.mt /onlineservices/Judgements/`		`http://otkzu.trybunal.gov.pl/`	
Special search parameter			
		Only decisions of type "A"	
Decisions			
42/2014/1 LSO	171/14 JPG	K 17/14 - Judgement†	SK 20/15†
30/16 JPG	1118/09 JPG	SK 25/15†	P 14/17†
95/17 SM	98/2016 MCH	SK 6/17†	U 2/17†
38/2016 JPG	38/17 LSO	SK 24/17†	SK 28/17†
23/13 JRM	5/15 JZM	SK 22/17†	K 39/15†

Country			
Portugal		**Romania**	
Database			
`http://w3.tribunalconstitucional.pt/ /AcordaosV22/`		`http://ccrsearch.ccr.ro/ccrSearch/ MainSearch/SearchForm.aspx`	
Special search parameter			
Decisions of type "Plenário"		"Act Solution A"	
Decisions			
377/2018*	79/2018	2129E/2018*	1018AI/2018
319/2018	379/2018*	439AI/2018	462AI/2018
372/2018	1083/2018	2267D/2017	2151D/2016†
428/2018	122/2018	3158D/2016	78D/2016*
160/2018	558/2018	2186D/2016	612AI/2018†

Country			
Slovakia		**Slovenia**	
Database			
`https://www.ustavnysud.sk /vyhladavanie-rozhodnuti /#!DmsSearchView`		`https://www.us-rs.si/odlocitve/`	
Special search parameter			
Decision-type *findings*; negotiation type *hearing*		*decisions*	
Decisions			
1.US.629.2017.1	3.US.557.2017.1	U-I-43/16†	Up-217/14*
2.US.559.2017.1	PL.US.11.2016.1*	U-I-80/16, U-I-166/16, U-I-173/16*	UI-79/16†
3.US.122.2018.2	PL.US.12.2016.1*	U-I-157/16, Up-729/16, Up-55/17†	U-I-50/16*†
PL.US.10.2016.1*	2.US.152.2018.2	U-I-38/16†	U-I-14/15†
PL.US.8.2017.1*	PL.US.9.2016.1*	Up-616/15†	UI-80/17†

Country	
Spain	
Database	
`https://hj.tribunalconstitucional.es/`	
Special search parameter	
plenary session	
Decisions	
48/2018*	13/2018
11/2018	89/2018
120/2018*	135/2018
128/2018*	119/2018 - Judgement
104/2018*	90/2018

* indicates decisions including separate opinion(s)
but no information on how the justices
who did not file separate opinions voted;
† indicates decisions with unanimous votes
that do not allow for discriminating between justices
using scaling approaches based on votes.

B The Constitutional Court Database (CCDB)

The Constitutional Court Database (CCDB) summarizes information on 2,006 senate decisions made by the German Federal Constitutional Court (GFCC) between 1972 and 2010. The 2,006 decisions contain information on 3,284 referrals to the court by 4,088 plaintiffs. In addition, the database includes sociodemographic information on the justices serving on the GFCC. The CCDB includes 6,790 different referrals to the court, mostly (1) laws, (2) decisions by ordinary or federal courts, or (3) administrative acts. Federal laws passed by the *Bundestag* after 1972 and until 2005 are included in the established data set on the German *Bundesgesetzgebung*, compiled by Burkhart (2008). Data compiled by Stecker (2016) allow this data set to be extended to include laws until the end of 2009. The CCDB provides identifiers to link both data sets. The most recent version of the database is available for download from `www.ccdb.eu` (last accessed 04/06/2023).

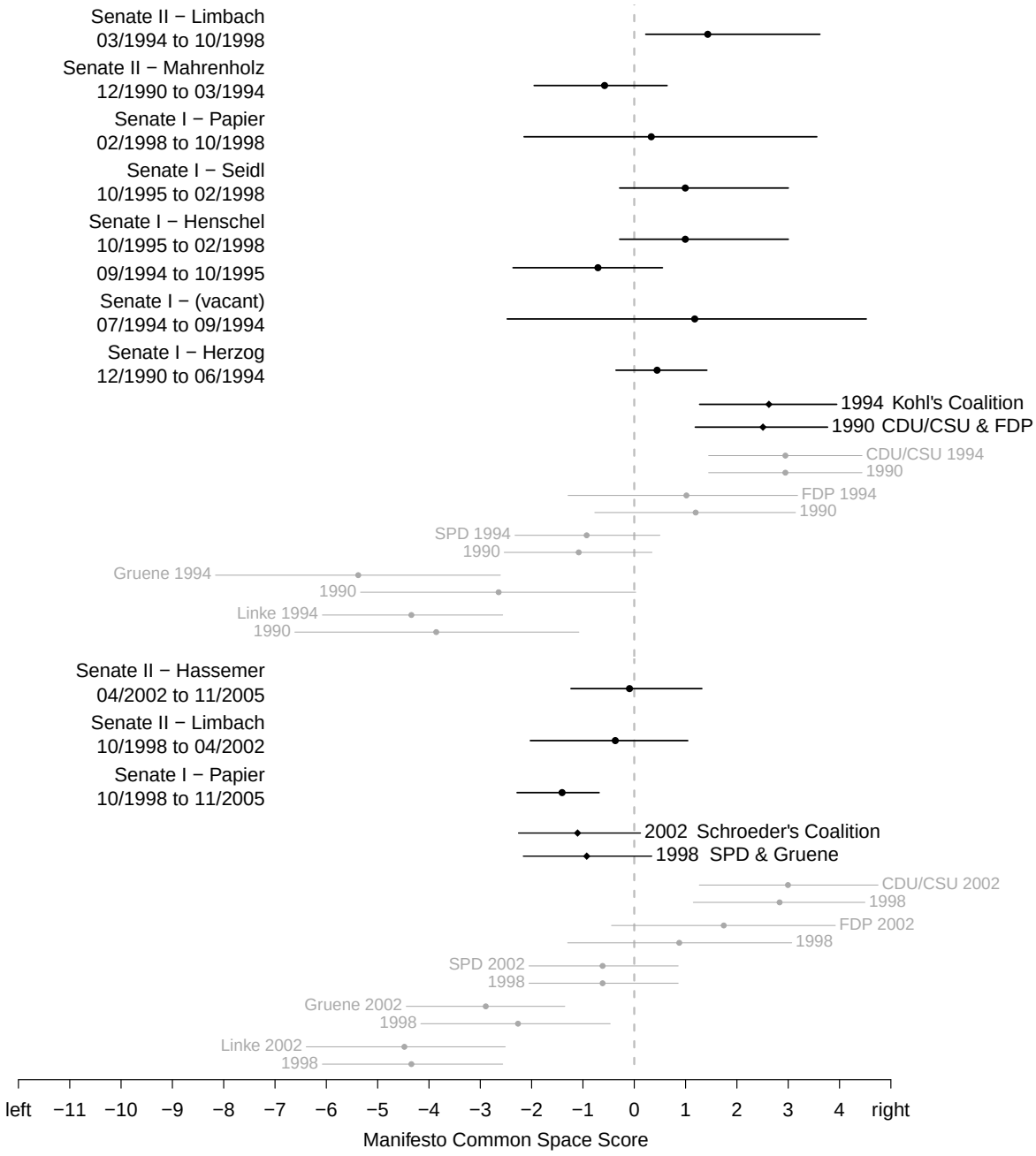
The article uses data from the former version of the CCDB. The codebook and the materials necessary to replicate all findings from this article are available in a comprehensive repository at [upon publication at Harvard Dataverse] and can also be replicated with the current version of the CCDB (CCDB V22.01-Mar). Please refer to the online codebook to learn more about the architecture of the CCDB. To generate the article's basic data we proceeded as follows: First, we used the variables `casesType` and `casesDate` to identify the 584 main decisions made between the 12^{th} and the 16^{th} legislative term. Second, we merged the variables `proceedingsAdm` (ruling on admissibility) and `proceedingsMer` (ruling on the merits) with the data to compute the judicial outcome of 582 decisions, as described in the main text. Third, we used the identifiers `gestaBurkhart2008` and `gestaStecker2020`, included in the database, to merge the data sets by Burkhart (2008) and Stecker (2016) with our data. Both data sets allow us to identify the political actor who originally proposed a law that was submitted to the GFCC for review.

C Ideal points estimated for the GFCC

In the main text, we presented the ideal point estimates of the Senates during Chancellor Merkel's term to discuss the face validity of our estimated common policy scores (see figure 1). In this appendix, we provide a corresponding figure for the terms of Chancellors Kohl and Schröder.

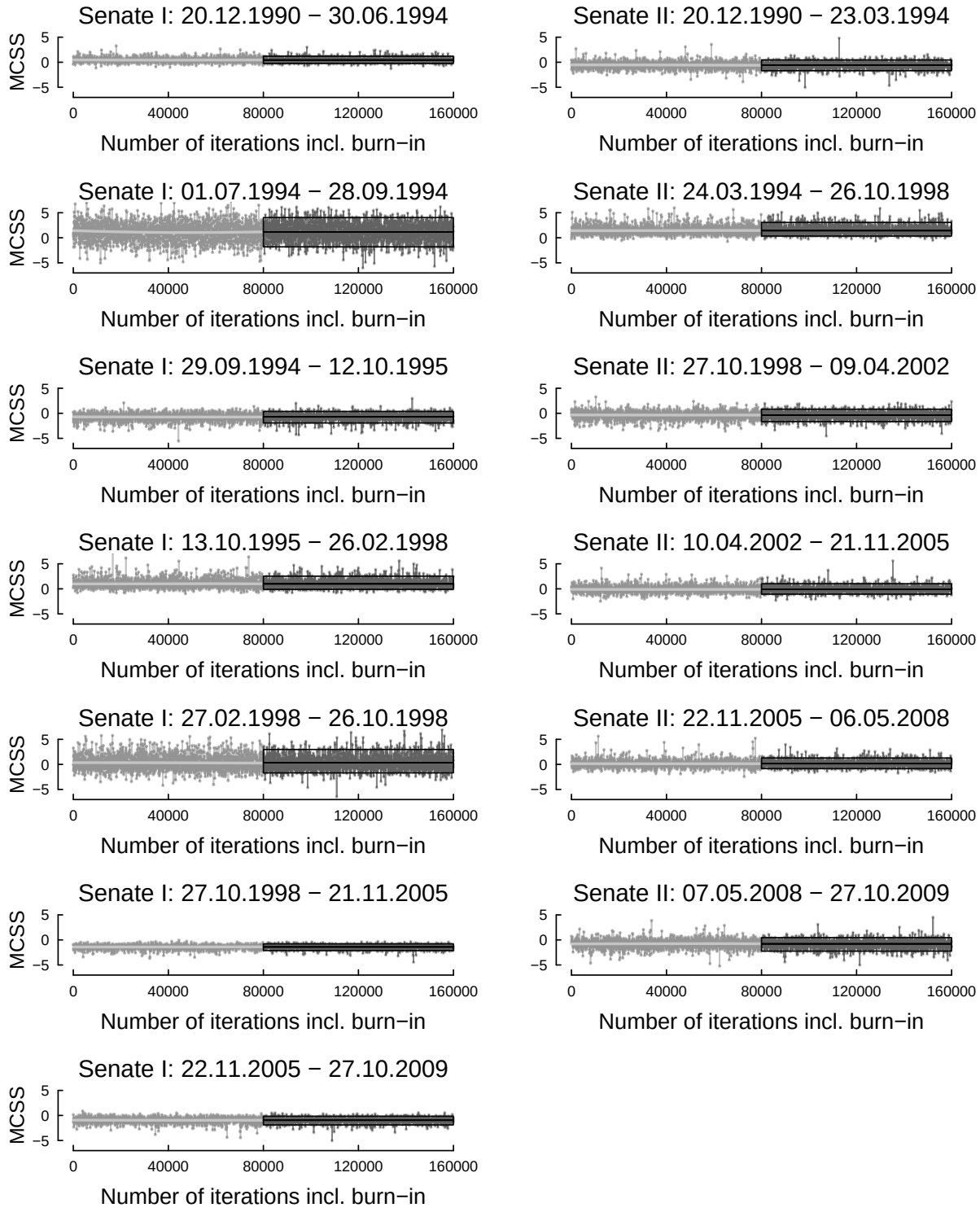
Figure C.1 is similar to figure 1 in the main text but displays the governing periods of Chancellors Kohl and Schröder. The figure presents the estimated positions for all Senates (dot, black estimates) and the positions of the governments (diamond, black estimates), along with their 95 percent credible intervals. The MCS Scores of the parties present in the *Bundestag* (first chamber) are depicted in gray. Since each chancellor served two terms, we included estimates for both terms, indicated by the years in which the chancellors took office. Chancellor Kohl took office in 1990 and 1994, while Chancellor Schröder took office in 1998 and 2002.

Figure C.1: Ideal point estimates of the Senates in comparison with the manifesto scores for governments and parties in the *Bundestag* during Chancellor Kohl's (1990–1998) and Chancellor Schröder's terms (1998–2005).



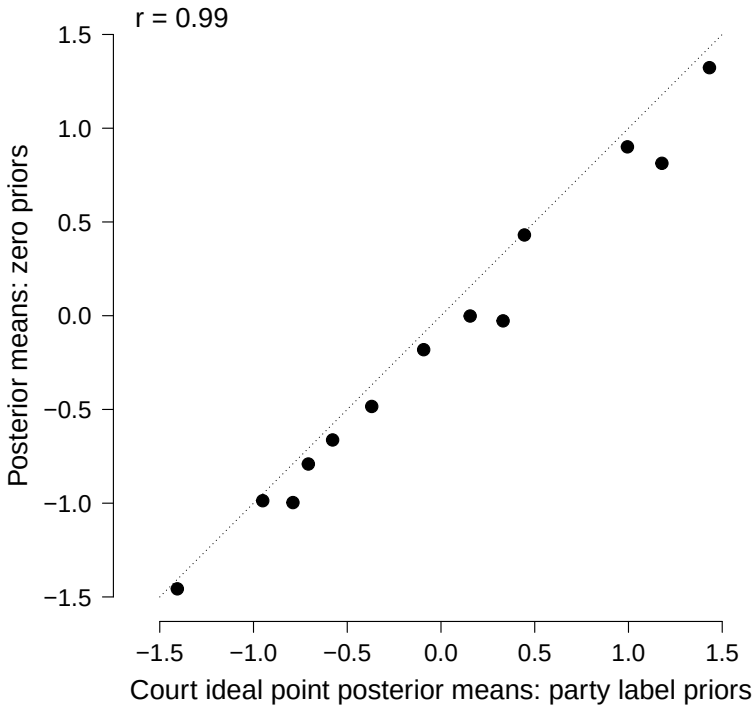
D Robustness of the ideal points estimated for the GFCC

Figure D.1: GFCC ideal point trace plots by Senate obtained with party label priors



In the main text, the ideal points are estimated using prior values defined by applying the party label approach. Figure D.1 illustrates the posterior chains for each Senate included in our analysis. The gray portion of the chains represents 80,000 warmup iterations, which are excluded from the results. The black portion of the chains includes posterior values from the 80,000 iterations following the warmup. We applied a thinning factor of 40 in the analyses, meaning that we retained only every 40^{th} posterior value to minimize autocorrelation, as the draws in Bayesian ideal point estimation are not independent of each other. The resulting autocorrelation functions indicate no issues with autocorrelation for any of the estimated ideal points (cf. Clinton, Jackman and Rivers, 2004). In total, we obtained 2,000 posterior values to inform each estimated ideal point.

Figure D.2: Correlation between ideal points estimated with party label priors and zero priors

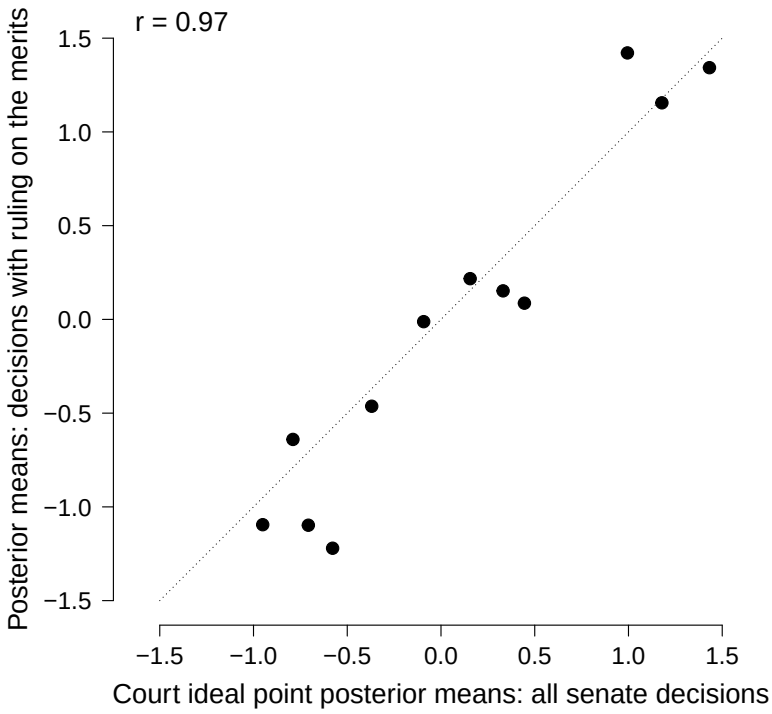


Note: The black dots indicate the mean ideal point estimates for the GFCC obtained from the baseline model with party label priors (x-axis) and with zero priors (y-axis). The gray diagonal illustrates the position of perfectly correlated GFCC ideal point estimates (r = 1). Estimates are obtained with the following prior values: $sd_{\gamma_j} = sd_{\beta_j} = 1$, $sd_{x_i} = 2$.

To evaluate the robustness of our findings, we compare the estimates using the party label priors to another set of plausible priors. Specifically, we set the prior mean for each Senate chair to zero, which represents a centrist court. This is plausible because courts are typically moderate actors, particularly as justices are elected by supermajorities in parliament and are unlikely to represent extreme partisan views. In figure D.2 we compare the estimates using party label priors (x-axis) to those resulting from zero priors (y-axis). The resulting estimates suggest that the ideal points are robust to changes in priors. The estimates from both sets of priors exhibit a high bivariate correlation of 0.99. On average, the party label approach establishes more conservative priors on the common policy scale (higher values) than zero priors, and the slightly more conservative Senates compared to the zero prior posteriors reflect this tendency. The first differences between the posteriors in the estimated ideal points are not statistically significant on a 95 percent confidence level.

We also conducted a re-estimation of the Senates' positions using only the rulings on the merits. As explained in the main text, we account for the data-generating process by linking rulings on admissibility with rulings on the merits for 582 decisions. When using only the 487 rulings on the merits, we find that the resulting positions are not substantively different from the positions reported in the main text. Figure D.3 depicts the ideal point estimates for each Senate using all decisions (x-axis) versus using decision outcomes on rulings on the merits only (y-axis). The estimates exhibit a high bivariate correlation of 0.97. There is no systematic difference between the positions derived from the two approaches.

Figure D.3: Correlation between ideal points estimated for all decisions and rulings on the merits only



Note: The black dots indicate the mean ideal point estimates for the GFCC obtained from the baseline model for all decisions (x-axis) and rulings on the merits only (y-axis). The gray diagonal illustrates the position of perfectly correlated GFCC ideal point estimates (r = 1). Estimates are obtained with the following prior values: $sd_{\gamma_j} = sd_{\beta_j} = 1$, $sd_{x_i} = 2$.

E Measurement error when predicting oral hearings

In this appendix, we present all estimates from the regression models in table E.1, including all control variables. Additionally, we demonstrate the robustness of the effect of the spatial distance between the GFCC and government on the occurrence of oral hearings. The effects reported in the main text remain robust when accounting for estimation uncertainty in the Senates' positions resulting from our vote matrix approach. In table E.1, we present the complete specifications of models 1 to 3 introduced in table 1 in the main text, including the estimates of the control variables. In model 4, we account for estimation uncertainty in the Senates' positions to assess the robustness of our distance measure shown in model 3. To address measurement error in our distance measure, we randomly draw one value from the posterior distribution of each Senate's ideal point and calculate the distance to the respective government in each of the 313 decisions. We then estimate a regression and save the results. This process is repeated 1,000 times for different draws from the posterior distribution of the Senate's position, resulting in different distance estimates for each draw. To reproduce the results reported in model 3 in table E.1, we combine the 1,000 estimated logit results. First, we calculate the mean of all estimated coefficients for each variable as our overall point estimate. Second, we compute the respective standard error, accounting for the variance within and across all regressions. To do so, we follow established strategies for combining coefficients in multiple imputation (see King et al., 2001, eq. 3). The results are presented in model 4 in table E.1.

We observe that, compared to model 3, the effect of the spatial distance in model 4 in table E.1 remains robust even when considering the estimation uncertainty in the Senates' positions. Although the coefficient's size decreases when accounting for measurement error and, to a lesser extent, the standard error as well, the coefficient remains negative and statistically significant. The estimated coefficients of the other covariates in the

baseline model (Krehbiel, 2016, Model 1) as well as the model fit remain almost identical when comparing models 3 and 4. Hence, we can conclude that the results obtained with the distance measure based on our vote matrix approach are robust, even when accounting for estimation uncertainty in the positions of the Senates.

Table E.1: Logit analysis to predict under what conditions public oral hearings occur

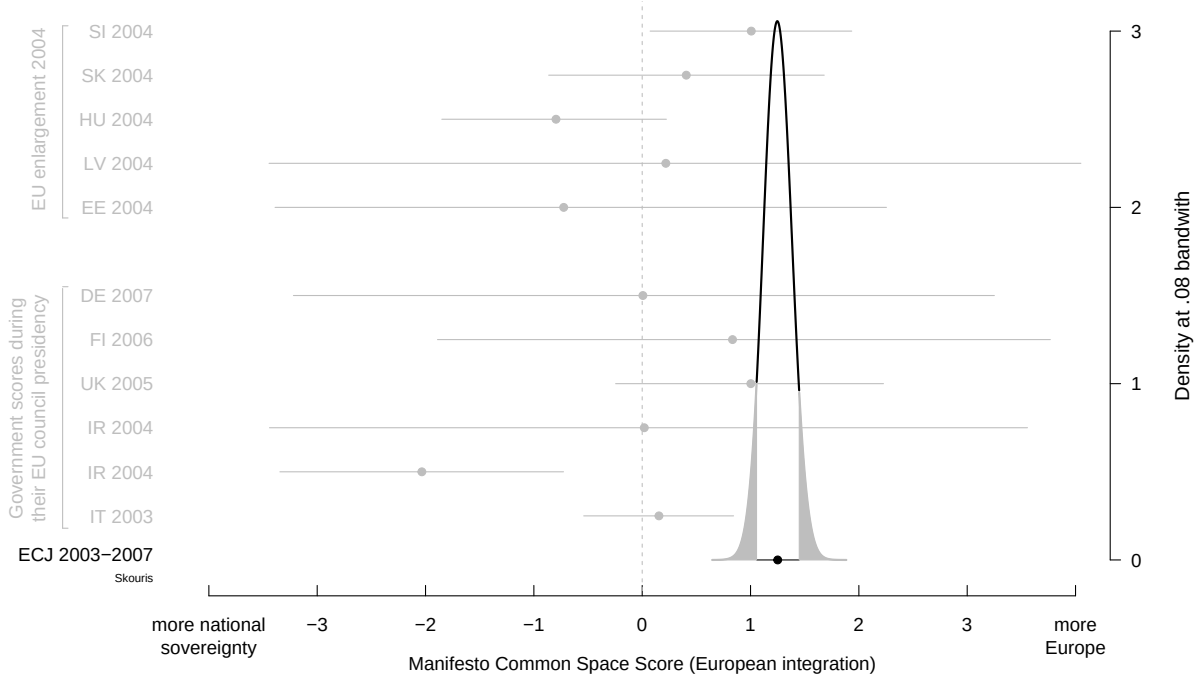
	Model 1	Model 2	Model 3	Model 4
	Baseline	Party label	Vote matrix	
	Krehbiel (2016)	Point estimate	Point estimate	Measurem. error
Distance GFCC - govt.		0.11	-0.54^{***}	-0.42^{**}
		(0.19)	(0.18)	(0.16)
No government brief (=1)	-1.18^{**}	-1.2^{**}	-1.06^{**}	-1.08^{**}
	(0.49)	(0.5)	(0.49)	(0.49)
Second Senate (=1)	0.48	0.49	0.42	0.46
	(0.3)	(0.3)	(0.31)	(0.31)
Total briefs	0.16^{***}	0.16^{***}	0.16^{***}	0.16^{***}
	(0.05)	(0.05)	(0.05)	(0.05)
Court brief	0.47^{*}	0.48^{*}	0.5^{*}	0.49^{*}
	(0.28)	(0.28)	(0.28)	(0.28)
Federal law	0.76^{**}	0.78^{**}	0.65^{*}	0.67^{*}
	(0.37)	(0.37)	(0.37)	(0.37)
Complainant support	0.3	0.31	0.21	0.23
	(0.33)	(0.34)	(0.34)	(0.34)
Constant	-2.7^{***}	-2.92^{***}	-1.9^{***}	-2.06^{***}
	(0.48)	(0.61)	(0.54)	(0.53)
N	313	313	313	313
logLik	-163.53	-163.37	-158.75	-159.77
AIC	341.07	342.73	333.5	335.55

Note: 84 observed oral hearings in 313 decisions. Standard errors in parenthesis. * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$. The baseline model is a replication of model 1 in Krehbiel (2016), re-estimated on a subsample of 313 decisions. Models 2 and 3 include distance measures based on the party label approach and our vote matrix approach respectively. To estimate model 4, we first draw one value from the posterior distribution of each senates' ideal point. We then calculate the distance of the Senate to the corresponding government in each of the 313 decisions to re-estimate the regression presented in model 3. This procedure is repeated to combine the coefficient parameters and their respective standard errors across 1,000 regressions (cf. King et al., 2001).

F Ideal points estimated for the ECJ

In the main text, we present the ideal points estimated during the presidency of Justice Rodriguez-Iglesias to evaluate the face validity of our estimated common policy scores for the ECJ (see figure 3). In this appendix, we provide the corresponding figure for the years 2003–2007, when Justice Skouris was the president.

Figure F.1: ECJ ideal point trace plot by presidency

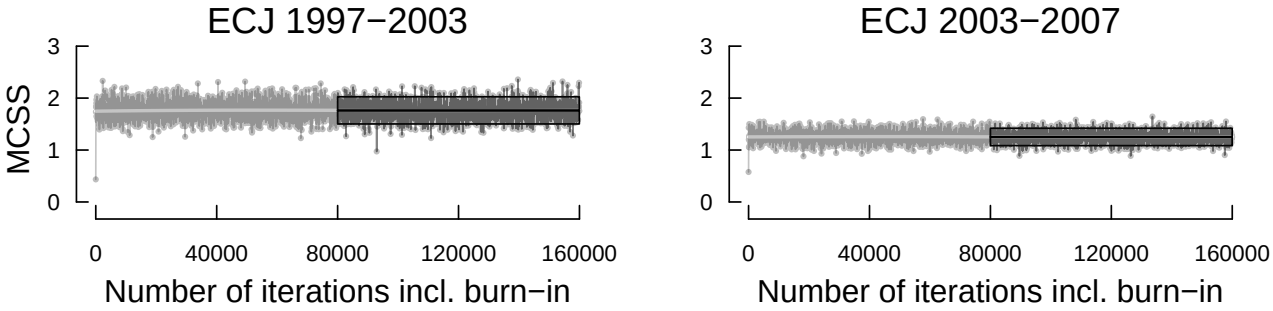


The density curve in figure F.1 illustrates the distribution of the posterior estimates for the ECJ's ideal points from the start of Justice Skouris' presidency until 2007 on the Manifesto Common Space (MCS) Score for European integration (König, Marbach and Osnabrügge, 2013). Consistent with our findings for the presidency of Justice Rodriguez-Iglesias, the ECJ significantly favors more European integration (black estimates) over more national sovereignty. Moreover, we observe that, on average, the ECJ favors more European integration compared to the position of all national governments when presiding over the European Council (gray estimates). Thus, in general, the ECJ emerges as an institution strongly in favor of 'more Europe' integration, also during Justice Skouris' term. These

findings align with our expectations from the literature (Carruba, Gabel and Hankla, 2012; Larsson and Naurin, 2016) and provide further support for the validity of our estimated positions.

G Robustness of the ideal points estimated for the ECJ

Figure G.1: ECJ ideal point trace plots by presidency
Estimates for prior values: $sd_{\gamma_j} = sd_{\beta_j} = 1$; $sd_{x_i} = 2$.

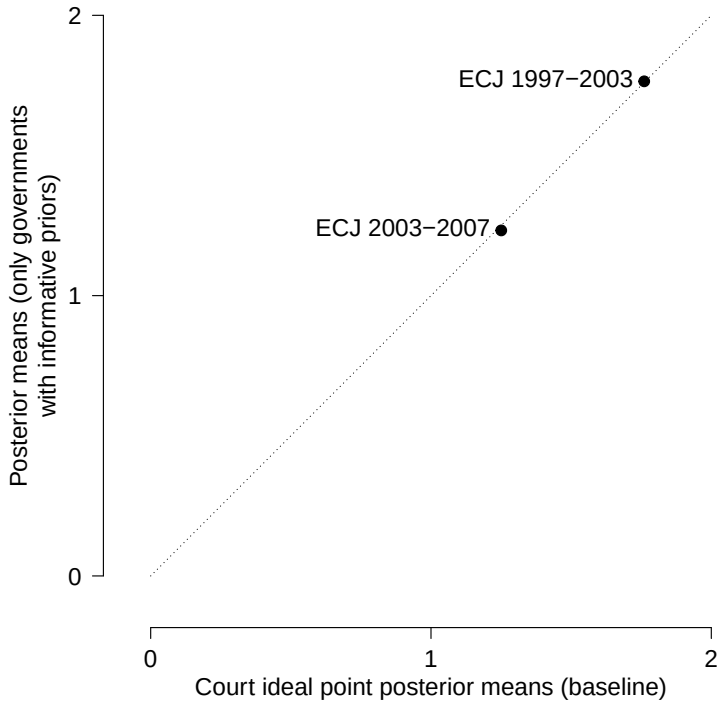


In the main text, the ECJ's ideal points are estimated using all recorded "votes" by the EU member state governments and the EU Commission in 1,241 questions in which at least one vote opposes the ECJ's ruling. Figure G.1 depicts the posterior chains for the two presidencies of the ECJ included in our analysis. The gray portion of the chains represents 80,000 warmup iterations, which are excluded from the results. The black portion represents posterior values from the 80,000 iterations after the warmup. We applied a thinning factor of 40 for the analyses, omitting 39 posterior values and including the 40th in our results. In total, we obtained 2,000 posterior values to inform each estimated ideal point.

In this appendix, we test the robustness of our ideal point estimates to including different sets of actors. To do so, we omit all "votes" by actors who are assigned weakly informative priors $\mathcal{N}(x_i, sd_{x_i})$, namely the EU Commission and some governments. The robustness check illustrates that the ECJ's estimated ideal points remain unchanged when excluding the EU Commission from the estimation. By estimating ideal points for the EU Commission, we show that our approach can be used to estimate the position for

any set of actors without a substantial impact on the estimated positions of the court. The requirement is that any actor has "voted" a sufficient number of times to estimate meaningful positions. For reference, the EU Commission casts 1,143 "votes" in our sample of 1,241 decisions. It is difficult to specify a specific number of "votes" necessary to estimate meaningful positions, as a virtually limitless number of combinations of "votes" (and prior information) may be observed.

Figure G.2: Correlation between ideal points estimated for votes by all actors and votes by actors with informative priors only



Note: The black dots indicate the mean ideal point estimates for the ECJ obtained from the baseline model including all "votes" (x-axis) and a model excluding votes by the EU Commission and governments with weakly informative priors (y-axis). The gray diagonal illustrates the position of perfectly correlated estimates (r = 1). Estimates are obtained with the following prior values: $sd_{\gamma_j} = sd_{\beta_j} = 1$, $sd_{x_i} = 2$.

In figure G.2, we compare the estimates presented in the main text including EU Commission "votes" (x-axis) to the estimates obtained when using only "votes" by governments with informative priors (y-axis); i.e., we use a subset of the vote matrix in which we omit all rows of "votes" by actors who are assigned weakly informative priors. Specifically, we

exclude 26 percent of government "votes" and 1,143 EU Commission "votes." We observe no meaningful difference in the resulting positions compared to the positions reported in the main text. Including or omitting one actor without informative prior does not impact the ideal point estimates of another actor. The results suggest that researchers who choose to adapt the approach presented in the main text can limit their data collection on specific sets of actors of interest to their analysis.

H Effects of member state observations on ECJ rulings

In the main text, we test the hypothesis that *the ECJ is more likely to rule in favor of more European integration if the distance between the ECJ to a policy under review increases and the member states file observations in favor of more national sovereignty* (H2). Our findings indicate that the impact of member state observations in favor of more national sovereignty diminishes as the ECJ's distance to the policy referred for review expands. Thus, when the distance to the policy under review is substantial, the ECJ does not take into consideration adverse member state observations (see figure 4 in the main text).

What should we expect when EU member states file observations favoring more European integration? In this instance, the ECJ has member states joining "its team." Therefore, we assume that the effect of member state observations favoring more European integration should not vary according to the distance between the ECJ and the policy under review. Instead, the pro-European ECJ receives support from the member states, enabling the ECJ to rule sincerely.

In table H.1, we present the complete specifications of models 1 to 3, as introduced in table 2 in the main text, including the estimates of the control variables. Model 4 incorporates an interaction term between member state observations favoring more European integration (`MS pro`) and our distance measure. Similar to `MS anti`, the variable `MS pro` indicates the number of member state observations filed in favor of more European integration, as the share of all member states weighted by their EU Council votes. This allows us to assess our claim that the ECJ can sincerely rule in favor of more European integration when it is supported by the member states, regardless of its distance to the policy under review.

The estimates presented in table H.1 reveal two key findings. First, the interaction term `Distance` $\times$ `MS pro` demonstrates that the impact of `MS pro` on the likelihood of the ECJ ruling in favor of more European integration does not vary with the distance between the ECJ and the member state government whose policy is under review. This

suggests that the ECJ rules sincerely when it receives support from the member states. Second, we observe that the effects reported in the main text remain robust even when including `Distance × MS pro`. The coefficients of all other covariates in the baseline model (Larsson and Naurin, 2016) are almost identical when comparing models 2 and 3. These findings support the robustness of our assumed effect that the ECJ gives less weight to member state observations favoring more national sovereignty as the distance to the member state government whose policy is under review increases.

Table H.1: Ordered logit analysis to predict under what conditions the ECJ rules in favor of more European integration depending on the distance in a common policy space between the ECJ and the national government which initiated a referred policy

	Model 1	Model 2	Model 3
	Baseline	with distance measure	
	Larsson/Naurin (2016)	Main Model	Robustness Check
Distance		−0.06	−0.05
		(0.05)	(0.05)
MS anti	.05***	−3.96***	−3.98***
	(0.69)	(1.18)	(1.18)
Distance × MS anti		1.51**	1.53**
		(0.71)	(0.71)
MS pro	4.24***	4.25***	5.22**
	(1.18)	(1.19)	(2.15)
Distance × MS pro			−0.81
			(1.48)
Chamber size	0.10	0.07	0.07
	(0.20)	(0.20)	(0.20)
_cut1	−1.33***	−1.41***	−1.40***
	(0.08)	(0.10)	(0.10)
_cut2	1.41***	1.33***	1.35***
	(0.08)	(0.10)	(0.10)
AG anti	−1.69***	−1.71***	−1.71***
	(0.17)	(0.17)	(0.17)
AG pro	1.93***	1.94***	1.95***
	(0.16)	(0.16)	(0.16)
Com anti	−0.87***	−0.86***	−0.87***
	(0.14)	(0.14)	(0.14)
Com pro	1.37***	1.36***	1.36***
	(0.15)	(0.15)	(0.15)
N	3845	3835	3835
logLik	−2865.51	−2854.21	−2853.91
AIC	5749.02	5730.42	5731.82

Note: Results from ordered logistic regressions. Standard errors in parenthesis; observations clustered by case. *p<0.1; **p<0.05; ***p<0.01

References

Burkhart, Simone. 2008. "Deutsche Bundesgesetzgebung 1972 - 2005." *GESIS Datenarchiv* ZA4569(1.0.0):Köln.

Carruba, Clifford J., Matthew Gabel and Charles Hankla. 2012. "Understanding the role of the European Court of Justice in European integration." *American Political Science Review* 106(1):214–23.

Clinton, Joshua, Simon Jackman and Douglas Rivers. 2004. "The Statistical Analysis of Roll Call Data." *American Political Science Review* 98(2):355–70.

Hanretty, Chris. 2012*a*. "Dissent in Iberia." *European Journal of Political Research* 51(5):671–92.

Hanretty, Chris. 2012*b*. "The Decisions and Ideal Points of British Law Lords." *British Journal of Political Science* 43(3):1–14.

Hanretty, Chris. 2014. "The Bulgarian Constitutional Court as an Additional Legislative Chamber." *East European Politics & Societies* 28(3):540–58.

King, Gary, James Honaker, Anne Joseph and Kenneth Scheve. 2001. "Analyzing Incomplete Political Science Data." *American Political Science Review* 95(1):49–69.

König, Thomas, Moritz Marbach and Moritz Osnabrügge. 2013. "Estimating Party Positions across Countries and Time." *Political Analysis* 21(4):468–91.

Krehbiel, Jay N. 2016. "The Politics of Judicial Procedures." *American Journal of Political Science* 60(4):990–1005.

Larsson, Olof and Daniel Naurin. 2016. "Judicial Independence and Political Uncertainty." *International Organization* 70(2):377–408.

Raffaelli, Rosa. 2012. "Dissenting opinions in the Supreme Courts of the Member States." *European Parliament: Directorate General for Internal Policies* (PE 462.470).

Stecker, Christian. 2016. "The effects of federalism reform on the legislative process in Germany." *Regional and Federal Studies* 26(5):603–624.

Wittig, Caroline E. 2016. *The Occurrence of Separate Opinions at the Federal Constitutional Court*. Berlin: Logos Verlag Berlin GmbH.