Who reaches the Bench?
Evaluation of Judicial Nominees for Constitutional Courts[*]

BENJAMIN G. ENGST[†]
University of Mannheim

THOMAS GSCHWEND[‡]
University of Mannheim

SEBASTIAN STERNBERG[§]
University of Mannheim

How do citizens evaluate judicial nominees for highest courts? Previous research solely based on the US Supreme Court points at two dimensions: judiciousness and the nominee's political leaning. While these dimensions are not likely to be independent from one another the identification strategies previously applied are not suitable to untangle the independent effect of each dimension. We employ a discrete-choice experiment using panel data from a random sample of German citizens to estimate the weights the public places on each dimension. We find that the public has a strong preference for political independent nominees. Moreover, our results clarify the conditions under which a judicial nominee's perceived lack of political independence can be compensated by a higher degree of judiciousness. Finally, we put our findings about the public's perception of judicial nominee's in a comparative perspective.

Abstract: 138 words
Main text: 9,996 words (incl. all references)
Online supplementary materials: 2,408 words

1 Introduction

How do citizens evaluate potential judicial nominees for the highest courts? While the public does typically not select those judges, most of their decisions have important implications for the public at large. Although there seems to be no standard model democracies follow when selecting judges to the highest court, the comparative judicial politics literature shows that political elites overwhelmingly control the selection of judges (Hönnige, 2007, 2008; Venice Commission, 1997). How they play this out within recent constitutional crises, for instance in Hungary and Poland, reminds us that the selection of judges to the highest courts belongs to the fundamental challenges faced by newly established democracies.

Despite the fact that the public does not select nominees for highest courts directly, it is crucially important to understand the public's perception of those nominees. The comparative literature established above and beyond the best studied case, the US Supreme Court, that constitutional courts are influential political actors (e.g., Brouard and Hönnige, 2017; Hanretty, 2012, 2013; Hönnige, 2011; Krehbiel, 2016; Staton, 2006, 2010; Vanberg, 2001, 2015). They make decisions that influence the policy-making process and, therefore, affect the life of every citizen. Political elites as principals might have the power to stuff constitutional courts as agents. However, given that political elites also depend on the public themselves through popular elections, elites have a vital interest in selecting judicial nominees whose decisions are widely accepted by the public, and particularly when the court makes a decision in the political interest of the political elites rather than the public. Choosing acceptable nominees will facilitate public acceptance of the court's decisions and contributes to the overall legitimacy of the judiciary. Without understanding which nominees the public finds acceptable, this additional reservoir of legitimacy can neither be analyzed nor tapped.

What kind of judicial nominees does the public prefer? Previous research points to especially two factors, judiciousness and the nominee's political leaning. The former addresses a nominee's professional legal qualities and the later the nominee's ideological views. However, conventional approaches to measure the importance of both factors for citizens' preferences might give mislead-

ing results because they do not account for the fact that the relevant attributes that determine those factors come in bundles. At any given time a judicial nominee incorporates both features, legal qualities from a professional education and career as well as personal, ideological views. Using traditional surveys to asses perceptions of each dimension is challenging because in practice it is likely that a citizen would evaluate judiciousness differently, depending on the nominee's ideology, or vice versa. As the most recent confirmation hearings of Judge Kavanaugh in the US remind us, it is heroic to assume that the public is able to asses the judiciousness independently from the nominee's perceived political leaning. Instead both dimensions should be rather conceptualized as non-separable (Stoetzer and Zittlau, 2015).

We make three contributions. First, we provide a research design, a discrete-choice experiment (Louviere, Hensher and Swait, 2000), that allows us to elicit citizens' preferences and to provide information about the way in which individuals or segments of the public value different attributes. Our approach differs from previous studies using traditional surveys as we infer the public's values placed on a nominee's judiciousness and political leaning directly from an active choice between judicial nominees rather than by separately asking respondents about their perception of the nominee's judiciousness and her political leaning. We also differ from so-called 'conjoint experiments' (Rogowski and Stone, 2017; Sen, 2017; Hainmueller, Hopkins and Yamamoto, 2013; Hainmueller and Hopkins, 2015) that we infer the public's preferences from choice-based behavior based on random utility theory rather than ad-hoc from how respondents rank or rate a nominee.

While we leverage randomly generated profiles of judicial nominees to estimate the weights the public places on each dimension and also identify the type of nominee the public prefers most, our second contribution is that we take a closer look at the inherent trade-off between judiciousness and political leaning in citizens preferences. We provide evidence that there is a "price" in terms of deviation from the ideal of political independence that citizens are "willing to pay" in order to get a nominee on the bench with high judicial credentials.

Our third contribution is that we are able to provide evidence from a non-US case that helps to put the US findings on the perception of Supreme Court nominees in a comparative perspective.

To do so, we administered our discrete-choice experiment to a random sample of German citizens and leverage a panel-design to test the reliability of the survey responses. We repeatedly presented respondents with pairs of randomly generated profiles of judicial nominees for the German Federal Constitutional Court (GFCC) that depend on personal characteristics, their level of judiciousness, characteristics of who nominated them, and nominees' political leanings.

Our study indicates that the two different sources of the public's perceptions about judicial nominees we know from studying the case of the US Supreme Court – judiciousness and political leaning – seem to be prerequisites for the evaluation of judicial nominees in other democratic regimes as well. Hence, while we use our approach to map the perception of judicial nominees in one country, we provide a blueprint to study public perception comparatively. On the one hand, we find that political independence is more important than judiciousness if respondents have to choose between respective judicial nominees. On the other hand, when looking at the inherent trade-off between judiciousness and political leaning of a nominee our results imply that judiciousness can in fact compensate to some degree for the lack of political independence of a nominee. These findings have important implications for other institutions and their perceived reservoir of legitimacy that, like highest courts, are not directly accountable to the public.

2   How the Public Evaluates Judicial Nominees

Existing research on how the public views judicial nominees is mainly driven by two approaches. On the one hand, scholars consider a judge's professional characteristics and legal skills as decisive to the public. This is summarized under the term *judiciousness* (Gibson and Caldeira, 2009*b*, 140). On the other hand, the political and ideological values of a nominee are regarded as decisive which we define as the judge's *political leaning*. While both are distinct dimensions, we argue in the following sections that citizens have non-separable preferences (Lacy, 2001; Stoetzer and Zittlau, 2015) regarding both dimensions. Afterwards, we propose a discrete-choice experiment to assess the public's evaluation of judicial nominees on the two dimensions jointly.

2.1 Competing Views of How the Public Evaluates Judicial Nominees

Scholars have identified two dimensions used by the public to evaluate judicial nominees – a nominee's *judiciousness* and her *political leaning*. Judiciousness describes the "satisfaction of legalistic expectations citizens hold of judges" (Gibson and Caldeira, 2009*b*, 141). These expectations are, for instance, influenced by a nominee's judicial qualification, a prestigious educational background or previous work experience. The judiciousness hypothesis implies that the public places particular weight on a nominee's legal qualifications holding other attributes of a nominee constant. Support for this perspective comes from Gibson and Caldeira (2009*a*,*b*). They assess the public's perception of the Conservative Supreme Court nominee Samuel Alito. Their analysis shows that legal qualification outplays political leaning when respondents who evaluate Alito hold strong feelings of legitimacy towards the Court.

The second dimension the public uses to evaluate judicial nominees is their political leaning. The political leaning hypothesis implies that the public places a particular weight on a nominees ideological views. Responds prefer nominees who are politically close to them compared to nominees who are politically distant. Several studies generate findings in favor of this perspective. Gimpel and Wolpert (1996) investigate controversial US Supreme Court nominees. They find that presidential approval, party identification and ideological leaning are highly related to the approval of a judicial nominee. Moreover, Bartels and Johnston (2012) find that the more a respondent perceives the Court in political terms, the stronger her preferences in favor of a political appointment process. The authors conclude that "much of the mass public actually *prefers* that justices be chosen on the basis of political factors" (Bartels and Johnston, 2012, 112, emphasis in original). The Court is seen as an additional political battleground.

The research designs of current approaches do not allow to assess the relative importance of each dimension – a nominee's judiciousness and political leaning – for the evaluation of a judicial nominee. Instead, it seems that respondent's preferences vary depending on the dimension they are asked to evaluate. However, it seems plausible that a mixture of attributes from both dimensions is decisive to determine who the public's most preferred judicial nominee is. Subsequently, the

separate assessment of the two dimensions imposes an identification problem. This is why we argue in the next section that the two dimensions need to be assessed conjointly.

2.2   A Joint Perspective on the Evaluation of Judicial Nominees

The preceding section outlines two dimensions used by the public to evaluate judicial nominees. The analyses of those dimensions previously performed by scholars imply that the public places an emphasis either on a nominee's judiciousness *or* on a nominee's political leaning (Gibson and Caldeira, 2009*b*, 140). However, the approach of the existing literature of treating these two dimensions independently is valid only if these dimensions are separable in citizens' evaluation. We argue that this is unlikely: a citizen's evaluation of the judiciousness of a nominee is not independent of her ideology, and vice versa. Instead, both dimensions are inherently linked and *non-separable*. Subsequently, it is necessary to reassess the existing findings accounting for the non-separability of the two dimensions, since treating them independently as it is currently done might lead to invalid conclusions.

Non-separability is a concept used to study voting-behavior in mass-elections. The concept implies that a voter's evaluation of a certain platform on one policy dimension is conditional upon the voter's position on the second policy dimension (Stoetzer and Zittlau 2015, 415; see also Lacy 2001). We argue that the concept also applies to the evaluation of judicial nominees. Judicial nominees have candidate profiles that are composed of attributes characterizing their judiciousness *and* their political leaning. Attributes of either dimensions are always simultaneously present. As in the real world, a judicial nominee consists of a bundle of attributes addressing the nominee's judicial credentials and her perceived ideological leaning at the same time.

This is why citizens are forced to simultaneously value the attributes on both dimensions when evaluating a judicial nominee. The public faces a trade-off and will sometimes be forced to pay a "price" in terms of deviation from the ideal of judiciousness in order to receive a nominee with a certain political leaning or vice versa. In order to understand this trade-off we need to ask the public to choose between (at least) two judicial nominees with certain attributes summarizing

5

a nominees judiciousness and political leaning. The characteristics of those attributes need to randomly vary across the two nominees in order to ensure separability of the judiciousness *and* the ideological dimension.

To implement such a design we employ a discrete-choice experiment (DCE) (Hainmueller, Hopkins and Yamamoto, 2013; Hainmueller and Hopkins, 2015; Louviere, Hensher and Swait, 2000) which allows us to assess the evaluation of a nominee's judiciousness and a nominee's political leaning conjointly. Finally, we randomly generate all other attributes of a nominee's profile. From observing their choice between two nominees we can infer the relative importance respondents places on the different attributes. We force respondents to engage in a trade-off and the design allows to treat the two dimensions as separable. Afterwards, we are able to estimate the weight placed on each dimension independently. Through the experimental design we obtain a clearer identification strategy to estimate such weights than by using statistical control.

We are not the first to use an experimental designs in the context of the evaluation of judicial nominees. However, we are the first to look at the inherent trade-off between a nominee's judiciousness and her political leaning. So far, conjoint experiments have been used to assess the public's perception of judicial nominees in the context of the US Supreme Court (Rogowski and Stone, 2017; Sen, 2017). However, Sen (2017) randomly withholds partisan information from respondents and Rogowski and Stone (2017) randomly prime responds using statements by the President or the Senate. In addition, both studies ask respondents to only rate certain dimensions of nominees and not to actively select a preferred judge. We employ our DCE as a preference elicitation technique, that can provide information about the way in which individuals or segments of the public value different attributes when evaluating judicial nominees. We infer these values directly from observing individual choice behavior using random utility theory rather than by asking respondents about their preferences directly or inferring this from ratings of nominees as it is done in the existing conjoint experiments (Rogowski and Stone, 2017; Sen, 2017). We intentionally do not use a priming strategy but ask responds to actively choose between two judicial nominees. This way, we are able to identify the characteristics that drive this behavior using standard random

6

utility theory (Louviere, Hensher and Swait, 2000).

In sum, we jointly account for two theoretical views that drive the evaluation of judicial nominees – (1) judiciousness and (2) political leanings. A research design that requires respondents to take a joint assessment of the two dimensions describes the actual evaluation of judicial nominees more accurately than non-experimental designs chosen by previous studies. Moreover, existing studies who apply appropriate research designs have not been used to study the inherent trade-off between judiciousness and the nominees political leanings that drive the public's evaluation of them. Finally, we administer our experiment in a context outside the typically used US American political system. This has three major implications which we discuss in the next section.

2.3 A Comparative Perspective on the Evaluation of Judicial Nominees

The studies discussed here assess the public's evaluation of judicial nominees in the context of the US. We employ our experiment within the context of the German Federal Constitutional Court (GFCC). This way, we make three contributions:

First, the GFCC is the archetype of a Kelsenian constitutional court and influenced the institutional design of many other courts, including their judicial appointment procedures. Hence, our findings of a typical case are likely to be generalizable with regard to constitutional courts. Moreover, by studying a constitutional court instead of the US Supreme Court we add the other common institutional design of highest courts (see Epstein, Knight and Shvetsova, 2001, 7) to the study of the public's perception of judicial nominees.

Second, the political leaning of judicial nominees in Germany is less obvious to the public than in the US. This is due to the selection process. Both parliamentary houses – the *Bundestag* and the *Bundesrat* – elect half of the sixteen judges at the GFCC. The Upper House (*Bundesrat*) confirms nominees with a two-thirds majority. The *Bundestag* has a selection committee that nominates candidates which are to be confirmed by a two-third majority vote of the entire parliament. However, the actual selection for both Senates (8-judge panel) of the court is based on inter-party agreements. The requirement of a two-thirds majority implies that the two major parties in Ger-

many, the Christian Democrats (CDU/CSU) and the Social Democrats (SPD) need to coordinate. They alternately nominate judges for both Senates. The respective smaller coalition parties (Greens and FDP) are allowed to nominate one candidate each in consent with the bigger parties from time to time (Brouard and Hönnige, 2017). Nevertheless, political actors are constraint in their actual choice. Eligible candidates for the court need to be at least 40 years old and have obtained a particular law degree (be qualified to hold the office of a judge). Three of the eight judges in a Senate of the GFCC must be former judges from a federal court. Subsequently, nominees are often judges, lawyers, politicians, or university professors. The selection process allows political actors to suggest nominees with a political leaning but the additional requirements are in place to limit ideological selections.

This process is not an artifact, but true for most of the European countries[1]. Moreover, in systems where the opposition can have considerable influence on the selection of judicial nominees and higher professional barriers are in place, the public cannot anticipate the partisan leaning of a future judicial nominee with certainty. This seems different in an institutional context where one (political) actor with an openly defined political leaning selects a judicial nominee. For example, the US President nominates a candidate subject to a hearing by the Senate. Everyone is aware that a Republican President will nominate candidates who are Conservative and a Democratic President will nominate candidates who are Liberal. In other words, the election mechanism in the US places a higher obvious emphasis on a nominees political leaning compared to the election mechanism in many European countries. If the judicial election system influences the public's evaluation of judicial nominees, then we might find that the political leaning of a judge is of lesser relevance in Germany than in the US. Moreover, as the judicial selection process in Germany is more diffuse, and the political ideology of the institutions involved are less obvious than in the US, we also regard it as more plausible that a nominee who is selected by a confirming institution with an obvious political leaning will be less preferred than a judicial nominee selected by institutions with diffuse political position. Our experiment allows us to disentangle the role of the political

---

[1]Hönnige (2007, 112-115) finds that under certain conditions the opposition can have considerable influence on the selection of judges in 11 out of 15 European countries with constitutional courts.

leaning in the selection of judicial nominees in Germany and to compare these findings to the existing findings on the US Supreme Court.

Finally, it is plausible to assume that the lens through which a judicial nominee's political leaning is reviewed is based on the same short-cuts people use to maneuver the party system. The US Supreme Court is embedded in a political system driven by political ideologies originating from a two party system with an electorate becoming increasingly polarized (Abramowitz and Webster, 2016). We employ our study in the setting of a parliamentary democracy with a multi-party system. Subsequently, even if the two major US parties allow for diversification through different inner-party movements, the German party system allows for direct discrimination through different parties. If viewed through the lens of the German public, the variety of parties may affect the variety on the bench while in the US context the bench is more polarized between two parties. Our research design allows as to disentangle the degree to which people compare the political leaning of a judicial nominee to their own political ideology. Subsequently, we are able to show whether the choices a multi-party system offers lead to higher variety in the composition of the bench.

In sum, we widen the perspective by studying perceptions of judicial nominees in a different institutional setting, the German Federal Constitutional Court. Differences of the judicial selection process and the party system may lead to different findings than the ones generated by studies on the US Supreme Court. In the following section we summarize the implementation of our DCE and present the results of our study.

3   Experimental Design and Analysis

We implemented our discrete-choice experiment as part of wave 26 of an online panel. Consequently, we ask respondents to evaluate profiles of potential nominees for the Federal Constitutional Court in Germany. The panel collects information on political attitudes and preferences of respondents through bimonthly longitudinal online panel surveys. Although administered online,

all surveys are based on a random probability sample of face-to-face recruited households from the German population, which were provided with access to Internet and special computers if necessary. Wave 26 includes $N = 2,749$ registered participants and is representative of both the online and offline population aged $16 - 75$ in Germany.

3.1 The Judicial Nominee Experiment

We implement our *Judicial Nominee Experiment* using altogether six screens. Each respondent saw six pairs of profiles of judicial nominees that were presented side-by-side, with each pair of profiles on a separate screen. We describe profiles of our hypothetical would-be judicial nominees along seven attribute categories, from which we build our independent variables.[2] Each of the seven attribute categories can take on multiple values. Attribute categories include (1) current occupation (politician, law professor, lawyer, judge at federal or regional court, prosecutor), (2) political leaning (none, CDU, SPD, FDP, Left party, Greens, AfD), (3) confirming institution (President, Upper house (*Bundesrat*), *Bundestag* with and without public hearings, Constitutional Court Judge Selection Committee of the *Bundestag*, the government, non-partisan expert committee) (4) age (35, 40, 45, 50, 55, 60 or 65 years old), (5) origin (East or West German, each with or without migrational background), (6) gender (female, male), and (7) marital status (same-sex marriage, divorced, married, single, widowed). Given the findings of the previous literature regarding judiciousness and political leaning we include both dimension into the judicial profiles. We assume that citizens can infer judiciousness from career cues and political leaning from ideological cues. Thus, we operationalize the career cues with the current occupation of the respective judicial nominee and ideological cues with the (non-) existence of partisan leanings.[3] Theoretically, we thus can generate $(6 \cdot 7 \cdot 7 \cdot 7 \cdot 4 \cdot 2 \cdot 5 =) 82,320$ different judicial nominee profiles. In practice, each respondent gets to see a random subset from this universe of judicial nominee profiles.

In order to minimize the potential impact of the order in which these attributes are presented,

---

[2]We provide respondents with profiles that are richer than required by law because citizens are *not* likely to know the legal requirements for becoming a judge at the GFCC. However, even if we restrict our analysis to only realistic

Figure 1: Example Screen of Nominee Profiles Used in the Experiment

The German Federal Constitutional Court is the only Court in Germany that is allowed to declare laws unconstitutional. Laws that have been rejected by the Federal Constitutional Court cannot be implemented afterwards.

Suppose there are the following candidates for nomination to the German Federal Constitutional Court. Please read the description of the potential judges carefully. Then, please indicate which of the two candidates you would personally prefer. Note that there is no wrong answer, it is solely about your personal choice.

	Candidate 1	Candidate 2
Current Occupation	Politician	Law Professor
Origin	West German	East German
Gender	male	female
Age	35	45
Partisan leaning	Close to SPD	Close to CDU
Confirming Institution	President	Government
Marital Status	Single	Divorced

If you had to choose between Candidate 1 and Candidate 2, which one of the two would you prefer?

Candidate 1   Candidate 2
◎   ◎

Note: This table illustrates a random screen (translated by us) of our discrete-choice experiment. Each respondent saw six of such screens. While the order of all seven attributes was randomized between respondents, it did not change from screen to screen once determined for each respondent. Nevertheless, the values for each attribute that make up each nominee profile have been randomized across screens and respondents. Figure 7 in the Appendix provides an actual screen shot.

so-called profile order effects (Hainmueller, Hopkins and Yamamoto, 2013), we randomized the order of the attributes between respondents. To ease the cognitive burden of respondents we fixed a randomly chosen order of the attributes across the six screens for each respondent. We present an example screen in figure 1. After each screen, respondents were asked to choose the preferred nominee. Thus, each respondent ideally made up to six decisions between 2×6 randomly generated profiles of judicial nominees. The choice outcomes of these decisions serve as our dependent variable.

For example, we might be interested whether respondents generally tend to choose female over male nominees. Such a gender effect might differ depending on how old the nominees are or whether the nominees are leaning towards a particular party. Following the strategy proposed by Hainmueller, Hopkins and Yamamoto (2013), we estimate *average marginal component effects*

---

profiles we get similar results as we show in table 1 of the appendix.

[3]Thus, the systematic component of our statistical model does include both theoretically relevant dimensions but also accounts for a richer description of the judicial profiles as well.

(AMCEs). The AMCE represents the average difference in the probability of being the preferred judicial nominee for the GFCC when comparing two different attribute values. Thus, the AMCE represents a quantity that summarizes the overall effect of gender, i.e. being male rather than female, despite heterogeneity in effect sizes across other attributes of the judicial nominees, including age or political leaning, and so forth.

The AMCE of female on the choice probability can be understood as the result of the following hypothetical calculation: (1) take a female judicial nominee with a fixed but arbitrary set of attribute values and compute the probability that she is chosen over an opposing judicial nominee with another specific set of attributes. (2) Now take the attributes of this female nominee and hypothetically change her gender into male. Thus, we have constructed a hypothetical twin brother of our female nominee. (3) Next, compute the probability that this hypothetical twin brother is chosen over the same opponent judicial nominee, and take the difference between the probabilities for the female and the male judicial nominee. Then, (4) successively compute the same difference between a female and a male nominee for all other different sets of the nominee's and opponent's attributes (other than gender). Finally, (5) take the weighted average of these differences over all possible combinations of the attributes according to their joint distribution. The resulting AMCEs provide an overall measure of how much female nominees are preferred over male nominees for the GFCC.

To sum up, the key advantage of our design is that the profiles of the would-be judicial nominees are fully randomized. This allows us to disentangle in particular attributes that cue the perceived judiciousness as well as political leaning of the nominees. Moreover, we can evaluate the relative importance of effects that might be otherwise correlated, because in reality those attributes are bundled together. Note that since the unit of analysis when analyzing a respondent's decision is the respective judicial nominee profile, we examine data for $32,988$ different judicial nominees – each of our $2,749$ respondents rated up to six pairings, with two nominees profiles per pairing. To obtain accurate uncertainty assessments, we cluster the standard errors by respondent because the observed choice outcomes are obviously not independent across the choices by the
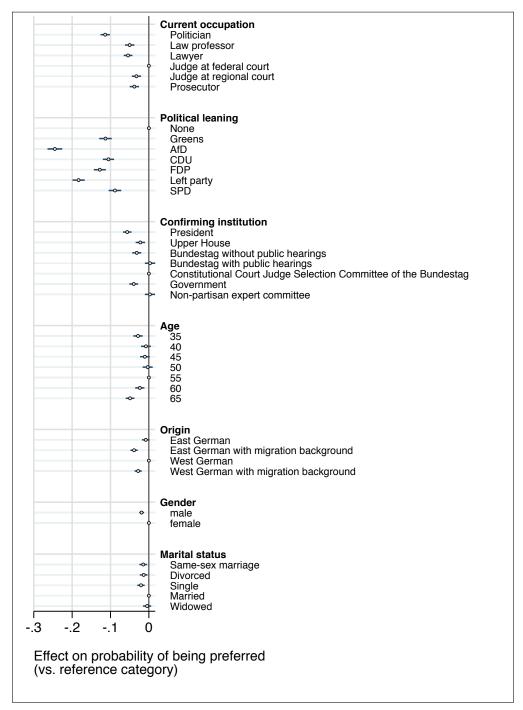
12

same respondent.

3.2 Results of the Baseline Model

We estimate a conditional (fixed-effects) logit model in order to predict the probability that certain judicial nominees are selected as GFCC judge by the respondents based on the attribute values that make up their profile. They are included in the systematic component of our model as a set of indicator variables for each attribute level (omitting the reference categories). The systematic component of our model therefore consists of 31 ($= 5 + 6 + 6 + 6 + 3 + 1 + 4$) indicator variables[4] and the estimated raw coefficients are presented in model (1) of table 1 in the appendix. Given that conditional logit models are neither additive nor linear, the estimated raw coefficients are less informative. Figure 2 provides therefore an overview of the estimated AMCEs and their corresponding 95% confidence intervals of the randomly assigned attribute values of judicial nominee profiles on the probability of being preferred to sit on the bench of the Federal Constitutional Court. All quantities of interest are presented on the vertical axis. The points represent the AMCE point estimates while the bars represent their uncertainty. The points without horizontal bars denote the attribute value that is the reference category for each attribute.

Relevant attributes in our experiment – as it is in reality – come in bundles. Judicial nominees have certain judicial credentials, their perceived degree of judiciousness, despite that fact that they are leaning potentially towards a particular party. Because of random assignment of the attributes we can identify the effect of the relevant dimensions. With respect to the judiciousness dimension, we find that judges who are currently employed at a federal court seem to be the most wanted nominees, and politicians are the least liked ones. For instance, politicians are about 11 ($\pm.6$) percentage points less likely to be chosen over current judges at a federal court. Next to current federal judges, respondents seem to prefer judges at the regional court and prosecutors (the

---

[4]We choose the following reference categories: 'Judge at federal court' (*Current Occupation*), 'None' (*Political Leaning*), 'Constitutional Court Judge Selection Committee of the Bundestag' (*Confirming Institution*), '55' (*Age*), 'West German' (*Origin*), 'Female' (*Gender*), and 'Married' (*Marital Status*).

Figure 2: Average Marginal Component Effects of Profile Attributes on being selected as preferred Judicial Nominee

Note: This figure shows estimates of the effects of the randomly assigned attribute values of judicial nominee profiles on the probability of being chosen. AMCE estimates and their 95% confidence intervals are presented on the vertical axis. The points represent the AMCE point estimates and the bars represent their uncertainty. The points without horizontal bars denote the attribute value that is the reference category for each attribute. All estimates are derived from a benchmark conditional logit model with clustered standard errors.

difference between both AMCEs is not significant at conventional levels). They are merely slightly less preferred (about 3 ±.6 percentage points). Between the occupations that signal high and low in judicial credentials there is finally the group of law professors and lawyers (whose AMCEs do not differ systematically as well). They are about 5 (±.6) percentage points less likely to be chosen over current judges at a federal court. Hence, we conclude that the public prefers professional judges and the more the higher they worked their way up within the judicial hierarchy. Such judicial nominees score seemingly higher on judiciousness than individuals who have merely obtained a law degree to work in the private sector, in academia or as politicians.

With respect to political signals, we find that, on average, political leanings of judicial nominees have a strong negative influence on the probability of being chosen compared to a politically independent nominee who does not lean towards any party. For instance, judicial nominees who lean towards extreme right (AfD) or left (Left party) parties are on average 25 (±1) and 18 (±.8) percentage points less likely to be chosen over nominees without any partisan leaning. These are the strongest effects we find across all AMCEs we estimate. The effect of leaning towards a mainstream party (like all the remaining parties) relative to a politically independent judicial nominee is around 10 percentage points. Thus, people seem to prefer judges without political leanings, i.e. nominees that are perceived to be politically independent. If nominees are not perceived as politically independent, citizens seem to prefer rather nominees that lean towards mainstream parties with moderate ideological positions.

The nature of the confirming institution can signal additional legitimacy and therefore might raise the judicial credibility of a nominee. With regard to the confirming institution we find that, on average, nominees seem to be less preferred if they are nominated by partisan actors such as the President or the Government. Instead citizens seem to like the transparency of confirmation hearings or confirmations by non-partisan expert committee, or by supposedly non-partisan or at least ideologically representative selection committees of the parliament. This confirms our expectation that German citizens dislike nominees selected by partisan actors alone, which we explain with the fact that the judicial selection process in Germany is ideological diffuse and the

15

leaning of the involved institutions is not immediately evident.

Moreover, the estimated AMCEs are less strong for various socio-demographic attributes compared to the ones that supposedly cue judiciousness and political independence such as current occupation, political leaning. Would-be judges should rather be females and should be neither to young nor too old, without migration background, and better not single. In a way, the preferred socio-demographic characteristics of judicial nominees seem to be similar to the average of the respondents themselves.

To sum up, how do citizens evaluate potential nominees for the highest courts? Our results indicate that both factors the literature studying the US judicial system identified as important seem to also matter in a comparative perspective when analyzing the public's evaluation of judicial nominees for the German Federal Constitutional Court. Consequently, nominees with many judicial credentials, i.e. nominees that score high on judiciousness and, at the same time, seem to be politically independent are most preferred by the public. In reality, though, they do not exist.

What is more relevant, therefore, is how citizen mange the trade-off when facing a choice between a nominee that seemingly has the qualifications to be a good judge but leans toward a political party and another nominee that lacks the qualities to be a good judge but is perceived to be politically independent. What is more important to the public, perceived judiciousness or perceived political independence? We explicitly assess this trade-off in the following section.

3.3 Disentangling the Trade-Off between Judiciousness and Political Independence

The analysis of our baseline model above makes transparent that attributes related to judiciousness and political leaning determine which kind of judge will be preferred in the eyes of the public. We basically find that nominees who seem to be qualified and politically independent are most likely to be chosen. In this section, we further explore the substantive meaning of these results in order to understand the implicit trade-off respondent's make when evaluating a judicial nominee.

How important is a nominee's degree of judiciousness compared to her perceived level of

political independence? To make this inherent trade-off transparent we, first, fix one profile to generate an ideal nominee who is politically independent and ranks as high as possible in terms of judiciousness. Accordingly, this nominee is currently a judge at a federal court, does not lean towards any party and for all other attributes posses the baseline categories. Given the results from the baseline model in figure 2, this baseline nominee will be on average preferred against every other conceivable nominee.

Second, we systematically choose profiles of other judicial nominees that vary on both dimensions and subsequently compare the model's predictions of who will be chosen in a pairwise comparisons with our baseline nominee. In order to operationalize the different judge types, we create five twin sisters of the baseline nominee that differ from her only in one or two attributes. The first twin sister has exactly the same attribute values as our baseline judge but ranks lower in judiciousness given her current occupation as lawyer. Furthermore, we create two more twin sisters, who also have a high level of judiciousness. The only difference between the baseline nominee is that one twin sister leans towards an extremist party (AfD) and the other twin leans towards a mainstream party (SPD). Lastly, we create two more twin sisters, who both are lawyers and therefore rank lower in judiciousness. They nevertheless differ because one leans towards an extremist party (AfD) and the other twin leans towards a mainstream party (SPD).

Figure 3 provides the simulated predicted probabilities together with their respective 95% confidence intervals for those five twins as judicial nominees when competing against our baseline nominee, who ranks high on judiciousness and does not lean towards any party.[5] Note that the respective estimation uncertainty of the predicted probabilities for each of the five pairings is always small enough such that all first-differences between any two values in figure 3 are systematic and, therefore, not due to chance alone. Our simulations make transparent that on average the public seems to value political independence more than judiciousness. The politically independent nominee who ranks low in judiciousness would loose against her twin sister, the baseline nominee, with

---

[5]We use a parametric bootstrap approach and calculate the respective choice probabilities based on 1000 draws from a multivariate normal distribution with a mean equal to the estimated coefficients and a variance that equals the estimated variance–covariance matrix of our baseline conditional (fixed-effects) logit model.
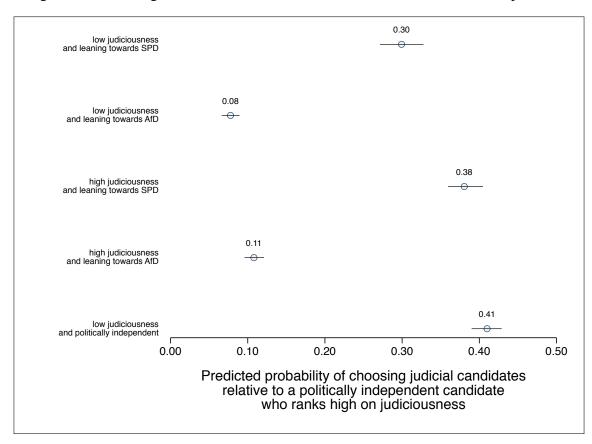
Figure 3: Evaluating the Trade-off between Judiciousness and Political Independence

Note: This figure shows the simulated predicted probabilities of five different nominees when competing against a baseline nominee that ranks high on judiciousness and does not lean towards any party. All simulations are derived from estimated coefficients of a baseline conditional logit model with clustered standard errors.

on average $41 : 59$, i.e., there is merely a 18 percentage point difference. This is the effect size that can be attributed to the difference in predicted probability of ranking high (judge at federal court) or low (lawyer) in terms of judiciousness given our simulated scenarios.

Next we estimate the size of the effect that can be attributed to the difference of nominees leaning towards a certain party compared to nominees that do not lean towards any party given our simulated scenarios. We therefore concentrate on the pairwise comparisons of the nominees who, like our baseline nominee, rank also high on judiciousness but lean towards an extremist party (AfD) and towards a mainstream party (SPD), respectively. Both of them will lose a pairwise comparison against our baseline nominee with $11 : 89$, which is a 78 percentage point difference in the case of the twin nominee that leans towards an extremist party and $38 : 62$, respectively, which

implies a 24 percentage point difference in the case of the nominee that leans towards a mainstream party. Even the smaller effect of both differences that represent the effect of being perceived as political or not is larger than the comparable effect of judiciousness (18 vs. 24 percentage points). Hence, we conclude that the public seems to value relative differences on the political dimension more than on the judiciousness dimension in our simulated scenarios.

Moreover, our simulations show that the picture is even more fine-grained when looking more closely at the political dimension. Our results indicate that the public seems to make a major difference between leaning towards a mainstream party and an extremist party. Irrespective of their respective degree of judiciousness, nominees leaning towards an extremist party are overall clearly less preferred ($8 : 92$ and $11 : 89$, respectively) than nominees leaning towards mainstream parties ($30 : 70$ and $38 : 62$, respectively). In fact, judicial credentials of nominees do not make much of a difference any more. The nominee leaning towards an extremist party with high judicial credentials is less likely to get chosen in a pairwise comparison with the baseline nominee than the twin nominee leaning towards a mainstream party with low judicial credentials.

What did we learn about the evaluation of judicial nominees? In the previous section we saw that the public prefers nominees who rank high on judiciousness and do not lean towards any party or if they do than rather lean towards a mainstream than towards an extremist party. Overall we find in this section that the public seems to prefer nominees that lean towards mainstream rather than extremist parties if nominees are not political independent. Judicial credentials seem to play less of a role. Even nominees with low credentials that lean towards a mainstream party (SPD) are preferred over nominees with high credentials that lean towards an extremist party (AfD). To sum up, the public values nuances on the political dimension more than on the judiciousness dimension.

3.4 Test of Model Assumptions and Robustness Checks

In this section we report four different robustness and diagnostic tests. Our first diagnostic test involves checking whether there are any carryover effects (Hainmueller, Hopkins and Yamamoto, 2013). The assumption of no carryover effects implies that respondents would choose the same

19

judicial nominee regardless of what particular profiles they get to see on a screen. Hence, the AMCEs should not depend on the data of particular screens. We assess the plausibility of this assumption by estimating the AMCEs for the two profiles respondents see on their very first screen. Obviously, we expect the resulting AMCEs to be less precise because we merely use data of one rather than six screens. In the Appendix (figure 5) we provide an overview of the estimation results. We find the estimated AMCEs to be very similar to the ones presented in figure 2 with the full data. Thus, the results would not be different even if we would rely on data from the first screen only.

Our second diagnostic test explores the effect of potential individual heterogeneity. Previous research mainly focuses on one particular respondent characteristic that potentially introduces heterogeneity in their decision-making process, namely knowledge about the court (e.g., Hoekstra, 2000; Sen, 2017). If decisions of respondents systematically differ depending on how much they know about the court, we would need to take this into account. Our respective analysis in the Appendix (figure 6), however, shows that respondents value the respective characteristics of judicial nominees similarly, irrespective of their knowledge about the court.

Our third robustness check involves a balance test to explore our randomization procedure. Although we fully randomized the judicial nominee attributes within each profile, we assess whether it actually produces well balanced experimental groups in our sample. In the context of our experiment, we are mostly interested in whether the profile attributes are actually balanced. We therefore conduct a multivariate balance test by regressing a particular respondent characteristic on the 31 indicator variables for all profile attributes. There should be no effect of those indicator variables on predicting the outcome. We use our knowledge scale (which we developed for the previous diagnostic test) as an outcome measure and estimate a ordered logit model in order to predict the respective scores of this knowledge scale (ranging from 0–2). Using a likelihood-ratio test we find that the indicator variables of the nominee attributes are jointly insignificant. The respective χ^2-value with 31 degrees-of-freedom is 29.47. Thus, the nominee attributes of the judicial nominees are jointly balanced.

Finally, our last robustness test leverages the panel design of our survey instrument. Owing

20

to this design, we have the possibility to check for the stability and, hence, the reliability of the individual decisions respondents make in our discrete-choice experiment across panel waves. We implemented the same experiment with the same screens, showing the very same judicial nominees in the next wave of the internet panel (wave 27) fielded two months later. 90% of all respondents (2486 out of 2749 respondents) of wave 26 have been re-interviewed in wave 27. We find that $11,204$ out of $14,916$ decisions lead to the same outcome for those 2486 respondents we have data across both waves. Thus, more than 3 out of 4 decisions when presented with the very same nominee attributes as one month before end up to be the same. We find this to be remarkable and strong evidence that the repeated decisions made by respondents are not done randomly (which would result in the same decision in 50% of all decisions).

In the next section, we expand our baseline model by adding another covariate to analyze more systematically the relationship between judiciousness and political independence. In particular, we investigate how much deviation from political independence citizens are "willing to pay" for a nominee who has high judicial credentials?

3.5 Can Judiciousness Compensate for Lack of Political Independence?

In the previous sections, we have seen that citizens generally prefer a politically independent nominee even if she has low judicial credentials over a nominee who has high judicial credentials but leans towards a party. This does not necessarily hold unconditionally. There might be a "price" in terms of deviation from the ideal of political independence that citizens are "willing to pay" by choosing a nominee who ranks high on judiciousness rather than a politically independent nominee who has a low judicial credentials.

Our discrete-choice experiment is ideally suited to entertain such an interpretation even if both assumed evaluative dimensions are potentially non-separable. Given that we randomize all attributes, the estimated AMCEs allow us to estimate such "willingness-to-pay" coefficients, i.e. the degree of perceived lack of political independence citizens are still willing to accept in order to receive a nominee who ranks high on judiciousness. We suggest that deviating from the ideal

of political independence can be measured using the perceived ideological distance between a judicial nominee and a respondent. Hence, we expand our baseline model and introduce ideological distance as another independent variable into our model[6]. We code ideological distance as the perceived absolute ideological distance between the perceived position of the party towards which a judicial nominee is leaning and the respondent's self-placement using a common left-right 1–11 scale. The resulting distance ranges between 0 and 10. The mean ideological distance to the judicial nominees which was assigned to the respondents is 2.3, with a standard deviation of 2.2.[7]

We use the estimates of the extended baseline model to simulate the respective choice probabilities when choosing between a nominee with high judicial credentials who lacks political independence and a nominee with low judicial credentials who is politically independent. We conceptualize the lack of political independence as ideological distance and let it systematically vary across its entire range between 0, where the respondent's ideological self-placement and the placement of the respective party coincide, and 10, where they are perceived to be maximally apart. By doing so, we can simulate how much ideological distance citizens are willing to accept before they start choosing the low judiciousness nominee who is politically independent. All other profile attributes are fixed at the values of the respective reference categories.[8]
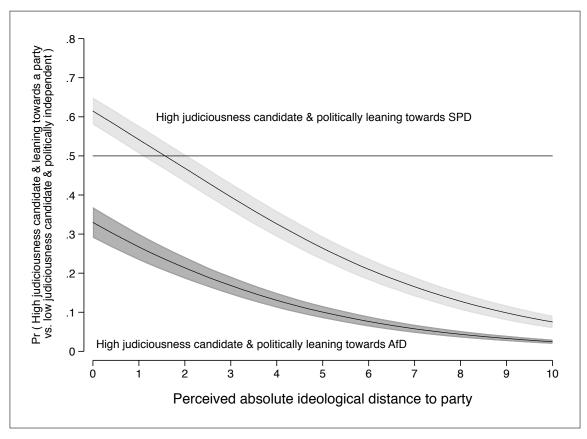
Figure 4 provides an overview about the respective choice probabilities between a nominee who ranks high on judiciousness but lacks political independence and a nominee who ranks low on judiciousness but is politically independent. As before, judges at a federal court rank high and lawyers rank low on judiciousness. In order to compare the 'willingness to pay' for a high judiciousness but partisan nominee with mainstream or extreme political views, we replicate all simulations for nominees who rank high on judiciousness and lean either towards the SPD or

---

[6]The estimation results of this extended baseline model are reported in table 1 of the appendix. They are similar to the baseline model in figure 2, with the exception of the AMCEs reported for variables that are related to political leaning.

[7]We set the ideological position of a politically independent judicial nominee to be on the same position where the respondent places herself on that scale. This implies that the ideological distance variable has no contribution in the model's systematic component when simulating the choice probabilities for such respondents.

[8]We use again a parametric bootstrap approach and calculate the respective choice probabilities based on 1000 draws from a multivariate normal distribution with a mean equal to the estimated coefficients and a variance that equals the estimated variance–covariance matrix of our conditional (fixed-effect) logit model that includes ideological distance.

Figure 4: Willingness to Deviate from Political Independence to get high Judicial Professionalism Candidates

Note: This figure shows the estimated effect of perceived ideological distance on the predicted probability to prefer a high judiciousness (but politically dependent) nominee vs. a low judiciousness nominee who is politically independent. The area in which the curve intersects with the .5 reference line indicates the estimated willingness to pay the price to have a nominee with a high judicial credentials on the bench despite being a partisan. All predicted probabilities are derived from a benchmark conditional logit model with clustered standard errors controlling also for ideological distance.

the AfD. How ideological distant can a high judiciousness nominee be before she will be chosen instead of a politically independent nominee with low judicial credentials?

The results indicate that citizens are willing to pay for deviations from the ideal of political independence in terms of accepting a partisan nominee as long as the ideological distance is not too far and the nominees' political views are perceived to be mainstream (SPD) rather than extreme (AfD). In fact, the predicted probabilities of the nominee with high judicial credentials leaning towards the AfD are always below the .5 reference line, i.e., she is never preferred over her twin sister nominee with low judicial credentials but who is politically independent. If citizens

perceive a nominee to hold extreme political views, no advantage on the judiciousness dimension can compensate for that. Citizens are not willing to accept judicial nominees with extreme views, no matter how high she ranks on judiciousness. This holds true even for respondents who place themselves in the same position on a left–right scale as the AfD-leaning nominee (lower curve), i.e. where the perceived ideological distance is 0.

Looking at the upper curve of a nominee with a high judicial credentials leaning towards the SPD shows that her predicted probabilities are sometimes above the .5 reference line, i.e. she is preferred over her low judiciousness twin sister who is politically independent as long as respondents do not see a large ideological difference between the party the nominee is leaning to and her own ideological self-placement. As long as this distance is 0 or 1 units on the 11-point ideological distance scale, respondents seem to be willing to pay this deviation from the ideal of political independence in order to get a high judiciousness nominee. If this perceived distance is greater than 2 units, the "price" seems to be too high for the average respondent to pay in order to get a partisan nominee with high judicial credentials on the bench.

To sum up, we find that judiciousness can compensate for the lack of political independence if the perceived ideological distance is not too large and if it facilitates the choice of a partisan nominee with rather mainstream political views. As long as the nominee's perceived ideological position is not reasonably similar to the respondent's own position on the same scale, though, citizens prefer a politically independent nominee with low judicial credentials over an high judiciousness but partisan nominee.

4   Conclusion

This study seeks to explain how citizens evaluate nominees for highest courts. When evaluating a nominee, the public seems to value both, judiciousness as well as her political leaning. However, no expert knowledge is necessary to do that. Attributes of the nominees' profile function as diagnostic cues that are readily available and help citizens to infer the nominees' level of judiciousness and

their degree of political leaning.

In order to identify the public's preferences and in particular to disentangle the relative importance of both dimensions, we administer a discrete-choice experiment to a random sample of German citizens. We repeatedly present respondents pairs of profiles of judicial nominees that systematically vary across attributes indicating among other things their judicial credentials and their political leaning. Evaluating the public's repeated choices of such profiles allows us to identify the relative importance of both dimensions and, consequently, find out which profiles gain the most public support.

Our results indicate that the public's ideal typical nominee combines both, political independence and a high level of judiciousness. Thus, the public prefers the robe and not a policy-maker wearing it. However, we also observe some interesting trade-offs when such an ideal typical nominee is not available. The results suggest that political independence is more important than judiciousness if respondents had to choose between respective judicial nominees. The public seems to value relative differences in political leanings more than relative differences in judiciousness given our simulated scenarios. However, no matter the degree of judiciousness, judicial nominees leaning towards an extremist party are clearly less preferred than nominees leaning toward mainstream parties.

Furthermore, our research design allows us to estimate to what degree citizens are willing to accept the lack of political independence of a nominee if getting a nominee on the bench with higher judicial credentials. We find that judiciousness can compensate for the lack of political independence. Citizens are "willing to pay a price" in form of some acceptable deviation from the ideal of political independence for a nominee that has high judicial credentials. As long as the perceived ideological distance is not too large and the partisan nominee leans towards a mainstream party, citizens are willing to compensate missing political independence with a higher level of judiciousness.

The hypothetical judicial profiles in our discrete-choice experiment describe potential nominees for the bench of the GFCC. This court is the archetype of a Kelsenian constitutional court,

which influenced the institutional design of highest courts in many other democracies. The findings based on a typical case are likely to travel also to other cases with similar highest courts and, therefore, contributes to the comparative literature. In addition, our discrete-choice experiment provides a blueprint to replicate such experiments in equivalent populations of citizens of other democracies as well.

We can compare our findings to studies on the US Supreme Court, given that we assess the public's perception of judicial nominees outside the US for the first time. One of our expectations was that the politicization of the existing selection procedures influences the public's perception of judicial nominees. German respondents are used to multiple political actors selecting judicial nominees. From an outside perspective, the selection process will include various political views and disperses an ideological concentration. On the contrary, in the US context the President with a particular ideological leaning is the sole actor nominating candidates for the Supreme Court and the Senate as another political actor confirming them. Subsequently, we expected that German respondents prefer institutions without a clear political leaning to select nominees, while US respondents seem to prefer a more politicized selection (Sen, 2017; Bartels and Johnston, 2012). Indeed, we find that political independence of the confirming institution to be vital to German citizens. Independent institutions are preferred over political institutions.

Moreover, we argued that the variety of ideological choices that the German multi-party system offers might lead to a higher variation in the composition of the bench. Consistent with studies on the US Supreme Court (Sen, 2017, 389), we find that German respondents use similar political shortcuts as respondents in the US when evaluating judicial nominees. Respondents compare their own political leaning to the political leaning of a nominee. Nevertheless, nominees who lean towards an extremist party are always less preferred than nominees who lean towards mainstream parties even by respondents who lean towards extremist parties. Thus, even if judicial nominees lean towards different parties of a multi-party system this does not imply that we will observe a higher variety of ideological judges.

Finally, the judicial literature finds inconsistent results about the impact of citizens' knowl-

edge about the court on the evaluation of nominees in the US. While Sen (2017) and Rogowski and Stone (2017) find a moderating effect of knowledge about the court, we find that German respondents who are knowledgeable about the court and those who do not evaluate judicial nominees similarly mirroring results of Hoekstra and LaRowe (2013) who employ an experimental design to predict support of US Senate confirmation of judicial nominees to federal courts.

More research is necessary to derive expectations about the relationship between political independence and judiciousness in democracies that, for instance, might not apply similar eligibility requirements for judges at the highest court than the case we studied here. An interesting "most-different" design would be to compare the results from this study, for instance, with a similar study using a French population of citizens. Nominees for the *Conseil constitutionnel*, the constitutional court in France, are neither required to have formal legal qualifications nor are they appointed in a particular non-partisan way (Venice Commission, 1997).

Finally, our findings have important implications for understanding the transformation process of the current democratic system of governance. In this transformation process newly created institutions are mushrooming around the globe and we observe a tendency that elected officials delegate more and more power to non-elected actors in order to make public policy. Regulatory bodies staffed with specialists, central banks or highest courts, are prime examples of such institutions that consists of non-elected actors that are not directly accountable to the public. Because of their role in the chain of delegation within the policy-making process such institutions act as *agents* for elected officials in their role as *principals*. This transformation of the democratic system of governance obviously undermines the traditional understanding of democratic accountability and, thus, has implications for how legitimate the public perceives such system of governance including the relevant political actors. Given that elected officials as principals are also themselves agents of the public, the public functions as a "second-order" principal. The implication of our results based on this observation is that the selection process to staff such institutions becomes an important but yet not well understood mechanism to legitimize the policy-making process. Our findings imply that if elected officials consistently staff such institutions with nominees the public does not prefer,

decisions of such institutions are less likely to be seen as legitimate. Conversely, even if institutions are not directly accountable to the public, the public perception of the appointment process might become instrumental in building-up a reservoir of legitimacy within the policy-making process using the public as an indirect but ultimate resource.

References

Abramowitz, Alan I. and Steven Webster. 2016. "The rise of negative partisanship and the nationalization of U.S. elections in the 21st century." *Electoral Studies* 41:12–22.

Bartels, Brandon L. and Christopher D. Johnston. 2012. "Political justice? Perceptions of politicization and public preferences toward the supreme court appointment process." *Public Opinion Quarterly* 76(1):105–116.

Brouard, Sylvain and Christoph Hönnige. 2017. "Constitutional courts as veto players: Lessons from the United States, France and Germany." *European Journal of Political Research* 56(3):529–552.

Epstein, Lee, Jack Knight and Olga Shvetsova. 2001. "Comparing Judicial Selection Systems." *William and Mary Bill of Rights Journal* 10:7–36.

Gibson, James L. and Gregory A. Caldeira. 2009*a*. *Citizens, Courts, and Confirmations: Posivity Theory and the Judgements of the American People*. Princeton University Press.

Gibson, James L. and Gregory A. Caldeira. 2009*b*. "Confirmation politics and the legitimacy of the U.S. Supreme Court: Institutional loyalty, positivity bias, and the Alito nomination." *American Journal of Political Science* 53(1):139–155.

Gimpel, James G. and Robin M. Wolpert. 1996. "Opinion-Holding and Public Attitudes toward Controversial Supreme Court Nominees." *Political Research Quarterly* 49(1):163–176.

Hainmueller, Jens and Daniel J. Hopkins. 2015. "The Hidden American Immigration Consensus:

A Conjoint Analysis of Attitudes toward Immigrants." *American Journal of Political Science* 59(3):529–548.

Hainmueller, Jens, Daniel J. Hopkins and Teppei Yamamoto. 2013. "Causal Inference in Conjoint Analysis: Understanding Multidimensional Choices via Stated Preference Experiments." *Political Analysis* 22(1):1–30.

Hanretty, Chris. 2012. "Dissent in Iberia: The ideal points of justices on the Spanish and Portuguese Constitutional Tribunals." *European Journal of Political Research* 51(5):671–692.

Hanretty, Chris. 2013. "The Decisions and Ideal Points of British Law Lords." *British Journal of Political Science* 43(3):703–716.

Hoekstra, Valerie J. 2000. "The Supreme Court and Local Public Opinion." *American Political Science Review* 94(1):89–100.

Hoekstra, Valerie and Nicholas LaRowe. 2013. "Judging Nominees: An Experimental Test of the Impact of Qualifications and Divisiveness on Public Support for Nominees to the Federal Courts." *Justice System Journal* 34(1):38–61.

Hönnige, Christoph. 2007. *Verfassungsgericht, Regierung und Opposition: Die vergleichende Analyse eines Spannungsdreiecks*. Wiesbaden: VS Verlag für Sozialwissenschaften.

Hönnige, Christoph. 2008. "Verfassungsgerichte in den EU-Staaten: Wahlverfahren, Kompetenzen und Organisationsprinzipien." *Zeitschrift für Staats- und Europawissenschaften* 6(3):524–553.

Hönnige, Christoph. 2011. "Beyond Judicialization: Why We Need More Comparative Research About Constitutional Courts." *European Political Science* 10(3):346–358.

Krehbiel, Jay N. 2016. "The Politics of Judicial Procedures: The Role of Public Oral Hearings in the German Constitutional Court." *American Journal of Political Science* 60(4):990–1005.

Lacy, Dean. 2001. "Preferences of Nonseparable in Survey Responses." *American Journal of Political Science* 45(2):239–258.

Louviere, Jordan J., David A. Hensher and Joffre D. Swait. 2000. *Stated Choice Methods: Analysis and Applications*. Cambridge: Cambridge University Press.

Rogowski, Jon C. and Andrew R. Stone. 2017. "How Politicized Judicial Nominations Affect Attitudes Toward the Courts.".

Sen, Maya. 2017. "How Political Signals Affect Public Support for Judicial Nominations." *Political Research Quarterly* 70(2):374–393.

Staton, Jeffrey K. 2006. "Constitutional Review and the Selective Promotion of Case Results." *American Journal of Political Science* 50(1):98–112.

Staton, Jeffrey K. 2010. *Judicial power and strategic communication in Mexico*. Cambridge University Press.

Stoetzer, Lukas F. and Steffen Zittlau. 2015. "Multidimensional spatial voting with non-separable preferences." *Political Analysis* 23(3):415–428.

Vanberg, Georg. 2001. "Legislative-Judicial Relations: A Game- Theoretic Approach to Constitutional Review." *American Journal of Political Science* 45(2):346–361.

Vanberg, Georg. 2015. "Constitutional Courts in Comparative Perspective: A Theoretical Assessment." *Annual Review of Political Science* 18(January):1–19.

Venice Commission. 1997. "The Composition of Constitutional Courts." *European Commission for Democracy Through Law* CDL-STD([1997]020).

Online Appendix: Assumption of No Carryover Effects

One diagnostic test involves checking whether there are any carryover effects (Hainmueller, Hopkins and Yamamoto, 2013). We asses the plausibility of assuming no carryover effects by estimating the AMCEs for the two profiles respondents see on their first screen only. Figure 5 provides an overview of the resulting AMCEs. The estimated AMCEs are very similar to the ones presented in figure 2. Thus, the results would not be different even if we rely on data from the first screen only. The respective uncertainty of our estimates, of course, are larger because the sample size is obviously smaller.
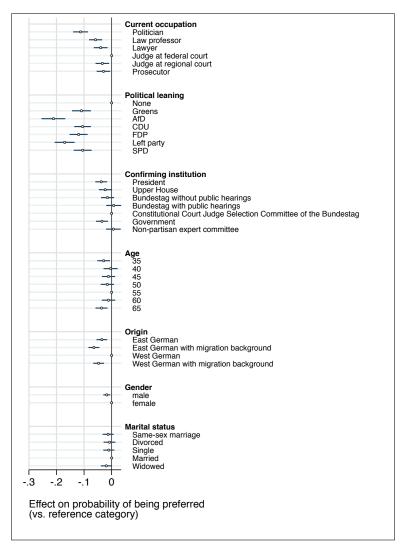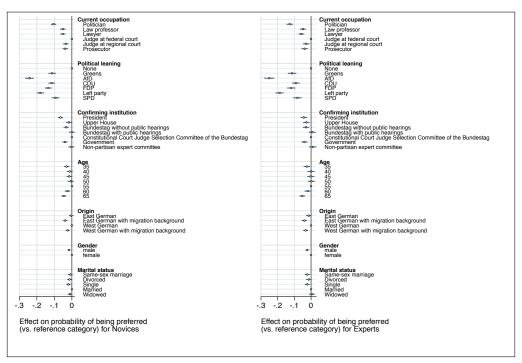
Figure 5: Assumption of No Carryover Effects

Note: This figure shows estimated AMCEs of the randomly assigned attribute values of judicial nominee profiles on the probability of being selected to the Federal Constitutional Court similar to figure 2. The difference is merely that we only use the data from the first screen. The results are virtually the same with slightly larger standard errors, of course, because the sample size is obviously reduced. There seems to be no spillover effects across the six screens.

Online Appendix: Individual Heterogeneity - Knowledge about the Court

Figure 6 shows the effects of the randomly assigned judicial nominee attributes on the probability of being selected conditional on respondents' knowledge about the court. We have two questions in which respondents need to identify the correct name of a judge currently sitting on the bench. Respondents who do not correctly identify any of the two individuals (Susanne Baer, Judge of the first Senate and Chief Justice Andreas Voßkuhle) comprise the group of *Novices* (roughly 62% of all respondents in wave 26). The remaining group are *Experts*. In the left panel of figure 6, we see estimates for respondents with low levels of knowledge (*Novices*). The right panel shows estimates for respondents with high knowledge about the court (*Experts*). Estimates are based on the same conditional logit estimator with clustered standard errors; bars represent 95% confidence intervals. Again, as in figure 2 in the main body of the paper, the points without horizontal bars denote the attribute value that is the reference category for each attribute. We find that the patterns of characteristics are generally similar for all respondents, irrespective of their particular knowledge about the court.

Figure 6: Average Marginal Component Effects of Profile Attributes on the Probability of being selected by Knowledge about the Court.
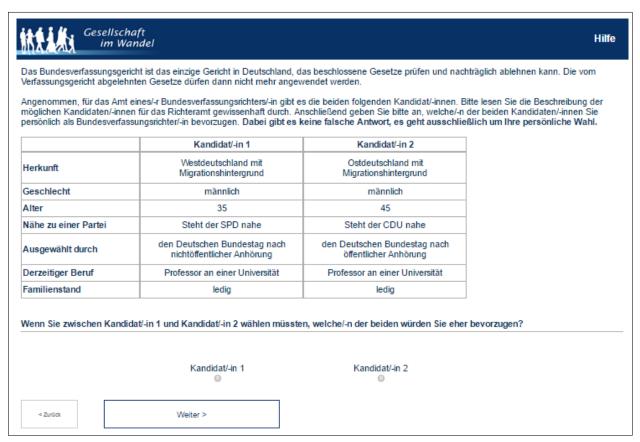

Note: This figure shows estimated AMCEs of the randomly assigned attribute values of judicial nominee profiles on the probability of being selected to the Federal Constitutional Court similar to figure 2. The difference is merely that we divided up the sample into *Novices* and *Experts*, based on two knowledge items that ask about particular judges on the court. The results are virtually the same across both samples, which indicates that both types of respondents evaluate judicial nominees in similar ways.

Online Appendix: Original Screenshot (in German)

Figure 7 shows an original screenshot of our discrete-choice experiment as implemented in wave 26 of the online panel.

Figure 7: Actual Screenshot (in German) of one Pairwise Comparison Used in the Experiment

	Kandidat/-in 1	Kandidat/-in 2
Herkunft	Westdeutschland mit Migrationshintergrund	Ostdeutschland mit Migrationshintergrund
Geschlecht	männlich	männlich
Alter	35	45
Nähe zu einer Partei	Steht der SPD nahe	Steht der CDU nahe
Ausgewählt durch	den Deutschen Bundestag nach nichtöffentlicher Anhörung	den Deutschen Bundestag nach öffentlicher Anhörung
Derzeitiger Beruf	Professor an einer Universität	Professor an einer Universität
Familienstand	ledig	ledig

Note: The following screenshot is an exemplar of an actual screen (in German) respondents saw at each stage of our discrete choice experiment.

Online Appendix: Baseline Estimation Results and Robustness Test

Below we document the estimated coefficients of two conditional logit models. The first model documents the raw coefficient estimates of the baseline model. These are the coefficients from which we calculate the AMCEs and other quantities-of-interest that are presented in the main body of the paper in figures 2 and 3.

The second model provides a further robustness test using a particular subsample. By law nominees to the constitutional court in Germany have to be at least 40 years old. Moreover, neither the president nor the government or any expert panel is nominating candidates. We thus deleted from the analysis any decision respondents had to make involving attributes of a candidate that were merely hypothetical. While the number of decisions and also respondents decreases considerably (from the full sample, i.e. model (1), of $N = 32,988$ to and subsample, i.e. model (2), with $N = 4,484$), the results stay essentially the same when comparing the estimates from the subsample in the second model to the ones from the baseline model. We therefore conclude that presenting respondents hypothetical attributes did not substantively change the results.

Table 1: Baseline Conditional Logit Estimation Results and Robustness Test

	(1) Full Sample		(2) Subsample	
	Coef.	Std.Err.	Coef.	Std.Err.
CURRENT OCCUPATION:				
Politician	-0.89***	0.05	-0.78***	0.12
Law professor	-0.33***	0.04	-0.29**	0.12
Lawyer	-0.36***	0.04	-0.30***	0.11
Judge at regional court	-0.21***	0.04	-0.21*	0.11
Prosecutor	-0.25***	0.04	-0.20*	0.11
POLITICAL LEANING:				
Greens	-0.65***	0.05	-0.57***	0.13
AfD	-2.11***	0.07	-2.05***	0.15
CDU	-0.59***	0.05	-0.33***	0.11
FDP	-0.75***	0.05	-0.71***	0.13
Left party	-1.23***	0.05	-1.03***	0.14
SPD	-0.49***	0.05	-0.36***	0.13
CONFIRMING INSTITUTION:				
President	-0.43***	0.05		
Upper House	-0.16***	0.05	-0.15*	0.08
Bundestag without public hearings	-0.23***	0.05	-0.21***	0.08
Bundestag with public hearings	0.02	0.05		
Government	-0.29***	0.05		
Non-partisan expert committee	0.02	0.05		
AGE:				
35 years	-0.20***	0.05		
40 years	-0.05	0.05	-0.18	0.11
45 years	-0.07	0.05	-0.01	0.11
50 years	-0.02	0.05	-0.04	0.12
60 years	-0.17***	0.05	-0.16	0.11
65 years	-0.37***	0.05	-0.39***	0.11
ORIGIN:				
East German	-0.06	0.04	-0.07	0.09
East German with migration background	-0.28***	0.04	-0.39***	0.10
West German with migration background	-0.20***	0.03	-0.28***	0.09
GENDER:				
Male	-0.14***	0.02	-0.04	0.07
MARITAL STATUS:				
Same-sex marriage	-0.11***	0.04	0.08	0.10
Divorced	-0.10***	0.04	-0.06	0.10
Single	-0.15***	0.04	-0.07	0.10
Widowed	-0.03	0.04	-0.09	0.10
Observations	32988		4484	
Log-Likelihood	-9975		-1364	
No. of choices	16494		2242	
No. of Respondents	2749		1603	

Reference categories omitted; * p < 0.10; ** p < 0.05; *** p < 0.01.

Online Appendix: Extended Baseline Model Including Ideological Distance

In model (1) of the following table 2 we present the raw estimates of the extended baseline model controlling additionally for the perceived absolute ideological distance of the voter to the respective party the nominee is leaning to. First, note that these are the estimated coefficients we use to simulate various quantities-of-interest that we present in figure 4 of the paper.
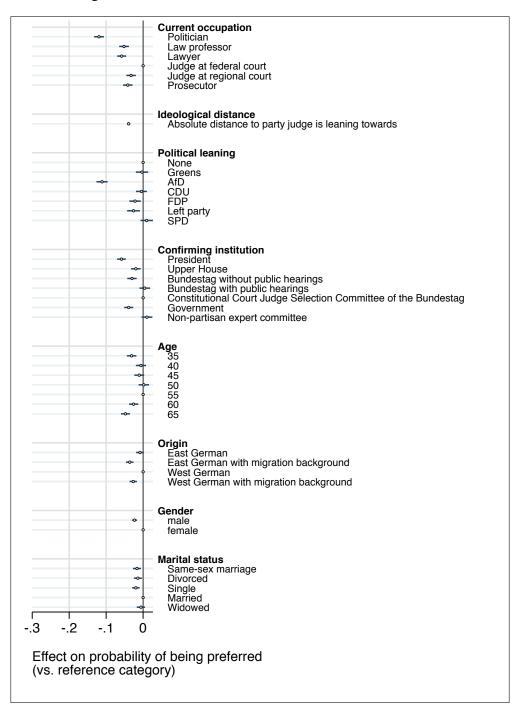
Second, given that conditional logit models such as the one we present in table 2 are non-linear and non-additive we also calculate the AMCEs based on model (1) and present them in figure 8 to ease interpretation of the results. Figure 8 provides an overview of the estimated AMCEs for our extended conditional logit model including perceived absolute ideological distance. The results are very similar to the baseline model we presented previously in figure 2, with the exception of the AMCEs for attributes that are related to political leaning. Controlling for the perceived ideological distance, the difference between parties essentially disappears. The major exception seems to be the AfD, for which we can identify a huge valence disadvantage of more than 10 percentage points. Additionally, nominees leaning towards the FDP and the Left party encounter small but systematic valence disadvantages as well. Moreover, the AMCE of the perceived ideological distance is very precisely estimated, indicating that the effect is real and not due to chance. The interpretation of this coefficient is straightforward. On average, citizens prefer judicial nominees who lean towards a party that is closer to their ideological position if no politically independent nominee is present.

Table 2: Conditional Logit Estimation Results

	(1) Using absolute distance		(2) Using squared distance	
	Coef.	Std.Err.	Coef.	Std.Err.
CURRENT OCCUPATION:				
Politician	-0.99***	0.05	-0.98***	0.05
Law professor	-0.36***	0.05	-0.36***	0.05
Lawyer	-0.41***	0.05	-0.40***	0.05
Judge at regional court	-0.22***	0.05	-0.21***	0.05
Prosecutor	-0.28***	0.05	-0.28***	0.05
IDEOLOGICAL DISTANCE:				
Ideological distance to party judge is leaning to	-0.30***	0.01	-0.44***	0.02
POLITICAL LEANING:				
Greens	-0.02	0.06	-0.32***	0.06
AfD	-1.11***	0.08	-1.43***	0.08
CDU	-0.03	0.05	-0.32***	0.05
FDP	-0.16***	0.06	-0.47***	0.06
Left party	-0.20***	0.07	-0.53***	0.06
SPD	0.07	0.06	-0.23***	0.06
CONFIRMING INSTITUTION:				
President	-0.47***	0.05	-0.47***	0.05
Upper House	-0.14***	0.05	-0.14***	0.05
Bundestag without public hearings	-0.23***	0.05	-0.22***	0.05
Bundestag with public hearings	0.03	0.05	0.03	0.05
Government	-0.30***	0.05	-0.30***	0.05
Non-partisan expert committee	0.07	0.05	0.07	0.05
AGE:				
35 years	-0.24***	0.05	-0.25***	0.05
40 years	-0.04	0.05	-0.05	0.05
45 years	-0.08	0.05	-0.08	0.05
50 years	0.01	0.05	0.01	0.05
60 years	-0.19***	0.05	-0.20***	0.05
65 years	-0.38***	0.05	-0.38***	0.05
ORIGIN:				
East German	-0.06	0.04	-0.06	0.04
East German with migration background	-0.28***	0.04	-0.27***	0.04
West German with migration background	-0.20***	0.04	-0.20***	0.04
GENDER:				
Male	-0.18***	0.03	-0.18***	0.03
MARITAL STATUS:				
Same-sex marriage	-0.13***	0.04	-0.12***	0.04
Divorced	-0.11**	0.04	-0.10**	0.04
Single	-0.15***	0.04	-0.14***	0.04
Widowed	-0.04	0.04	-0.04	0.04
Observations	28612		28612	
Log-Likelihood	-8054		-8085	
No. of choices	14306		14306	
No. of Respondents	2466		2466	

Reference categories omitted; * p < 0.10; ** p < 0.05; *** p < 0.01.

Figure 8: Average Marginal Component Effect of Perceived Ideological Distance within an extended Conditional Logit Model
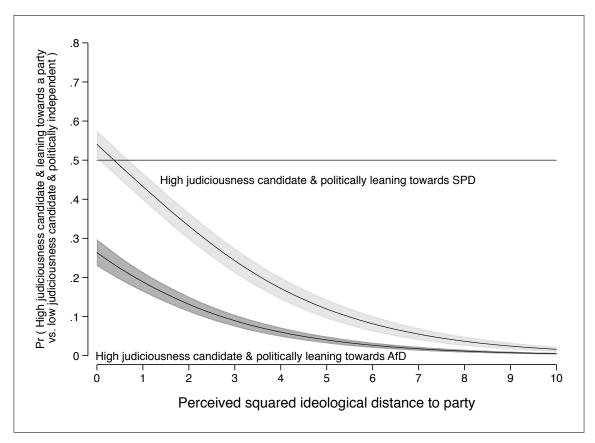
Online Appendix: Robustness of findings regarding the Trade-off between Judiciousness and Political Independence

How ideological distant can a high judiciousness nominee perceived to be before she will be chosen instead of a politically independent nominee with low judicial credentials? In order to test the robustness of the findings in section 3.5 we replicate our analysis and include the squared distance rather than the absolute ideological distance into our model, re-estimate it, present the raw coefficients as model (2) in table 2 of the appendix and rerun our simulations. Figure 9 provides an overview about the respective choice probabilities between a nominee who ranks high on judiciousness but lacks political independence and a nominee who ranks low on judiciousness but is politically independent using the squared ideological distance as a new measure.

Figure 9: Willingness to Deviate from Political Independence to get high Judicial Professionalism Candidates

Note: This figure shows the estimated effect of perceived ideological distance on the predicted probability to prefer a high judiciousness (but politically dependent) nominee vs. a low judiciousness nominee who is politically independent. The area in which the curve intersects with the .5 reference line indicates the estimated willingness to pay the price to have a nominee with a high judicial credentials on the bench despite being a partisan. All predicted probabilities are derived from a benchmark conditional logit model with clustered standard errors controlling also for squared ideological distance.

As can be seen, figure 9 is substantively similar to graph 4 in section 3.5 of the paper. The findings are robust to using a different operationalization of the ideological distance measure although, the estimated 'willingness to pay' for a high judiciousness but partisan nominee with mainstream views is somewhat lower using the squared distance compared to the absolute distance.