Does Efficiency Trump
Legality? The Case of the
German Constitutional Court

Christoph Engel

MAX PLANCK SOCIETY

Does Efficiency Trump Legality?
The Case of the German Constitutional Court

Christoph Engel

October 2017

Does Efficiency Trump Legality?

The Case of the German Constitutional Court

Christoph Engel

Abstract

The US Supreme Court has the power of certiorari. It may pick its fights. As a beneficial side effect, the court may allocate its resources, in particular the time and energy the justices spend on a case, to worthy causes. In economic parlance, this discretion makes the court more efficient. Efficiency comes at a political cost, though. This discretion also gives the court political power. It may direct its verdict to causes that are politically most relevant, or it may put an issue on the political agenda. Officially German constitutional law does not have certiorari. The Constitutional Court must decide each and every case that is brought. Yet over time the court has crafted a whole arsenal of more subtle measures for managing the case load. This paper shows that it uses these tools to engage in its version of allocating resources to cases. It investigates whether the ensuing efficiency gain comes at the cost of biasing the court's jurisprudence.

The paper exploits a new comprehensive data set. It consists of all (mostly only electronically) published cases the court has heard in 2011. While the data is rich, in many technical ways it is demanding. The paper uses a factor analysis to create a latent variable: to which degree has the court taken an individual case seriously? It then investigates whether observed indicators for bias explain this latent variable. Since the paper essentially investigates a single (independent) case, in statistical terms the findings are to be interpreted with caution. The paper can only aim at finding smoking guns.

I. Research Question

Supreme Courts are judicial and political bodies at a time. A visible expression of this dual role is the power of certiorari, as enjoyed by the US Supreme Court since the Judiciary Act of 1891 (26 Stat. 826). The court may pick its fights. As a beneficial side effect, the court may allocate its resources, in particular the time and energy the justices spend on a case, to worthy causes. In economic parlance, this discretion makes the court more efficient. Efficiency comes at a political cost, though. This discretion also gives the court political power. It may direct its verdict to causes on which it wants to have a political impact, or it may put an issue on the political agenda.

Officially German constitutional law does not grant certiorari. In principle, the Constitutional Court must decide each and every case that is brought. This obligation is qualified for constitutional complaints that are directly brought by citizens. They must be accepted by the court. But officially the court does not have discretion in such cases either. It has to accept the case, "in so far as it has general constitutional significance" or "if the complainant would suffer a particularly severe disadvantage if the Court refused to decide on the complaint" (§ 93a BVerfGG). Both prescriptions bind the court. As a matter of fact, the court hears a lot of cases. In 2011,[1] 6066 new cases have been brought. During the year, the court has issued 5733 verdicts.[2] Yet a closer look reveals that the court has crafted a whole arsenal of more subtle measures for managing the caseload. Only 34 rulings have been given by one of the two senates, all the remaining only by chambers of three justices. Only 76 complaints decided by the chambers have been successful. Only 467 of the dismissals have come with written reasons.[3]

From 1998 on, the court has made many of its decisions electronically available.[4] Unfortunately, this collection is incomplete. For instance for 2011, it only lists 255 decisions. The remaining decisions are not publicly available. Quite likely most of these cases have been summarily dismissed. But this is only a supposition. Strictly speaking there is an insurmountable selection problem. As this paper shows, it is nonetheless revealing to study this incomplete evidence with quantitative methods. For multiple reasons that transcend selection and are discussed below, this evidence is inconclusive. Causal effects cannot be established. But the available information suffices to develop an index that measures how seriously the court has taken a case. This index shows pronounced variance. Apparently the court does not devote its energy randomly or uniformly, but picks cases that it considers worth the effort. While the court does not have the power of certiorari, it has found indirect ways for allocating resources to cases.

---

1   I use 2011 since I have coded a dataset of all decisions made during the year that are electronically available on the website of the court.

2   http://www.bundesverfassungsgericht.de/DE/Verfahren/Jahresstatistiken/Archiv/2011/gb2011_pdf/ A-III.pdf?__blob=publicationFile&v=2.

3   Ibid.

4   http://www.bundesverfassungsgericht.de/SiteGlobals/Forms/Suche/Entscheidungensuche_Formular.html ?language_=de. The website also reports select decisions from earlier years.

This finding triggers the obvious follow up question: does the ensuing efficiency gain come at the cost of biasing the court's jurisprudence? In the political and in the academic discourse in the US, one concern dominates: that the Supreme Court is partisan and favors one of the political parties, or their cherished policies. In Germany, Constitutional Court justices are nominated by the political parties, which makes it possible to investigate the equivalent concern. But there are multiple other bench characteristics which might as well explain that the court, acting as a senate or chamber, takes a case more seriously. In this exercise, I find almost no signs of bias. Again, the finding has to be read with caution. I am not in a position to identify causal effects. But I at least do not find a smoking gun.

The remainder of this paper is organized as follows: section II relates the paper to the literature. Section III introduces the German case. Section IV explains how to generate the dependent variable: how seriously does the court take a case? Section V introduces bench characteristics and investigates whether they explain that the court has taken a case more seriously. Section VI concludes.

II.   Related Literature

The quantitative literature on the German Constitutional Court is still sparse. Georg Vanberg was interested in factors correlated with the decision of the court to strike down a statute. He uses all published senate decisions from 1983 to 1995 directly concerned with a statute. He finds that the court has been more likely to strike down a statute if it has held an oral hearing, and in particular if one of the parties heard claims the statute to be unconstitutional. He interprets both indicators as proxies for public scrutiny in case government were to neglect the ruling. He further finds that statutes are more likely to be struck down if the current government (after a change of majority in the elections) also claims unconstitutionality. By contrast, in policy areas that are more complex, like economic regulation or social insurance, statutes are less likely to be invalidated by the court (Vanberg 2004). I have another research question, another dependent variable, use different data, and have a different econometric approach. In one dimension, our results can be seen as complementary. While Georg Vanberg finds that the court is politically active, he too does not find that the court is active in support of one political orientation. The activity is a function of the relationship between the court on the one hand, and government and parliament on the other hand.

Shikano and Koch use dissenting votes of the justices from one of the two senates between 1970 and 2009 to identify their ideological positions. Justices selected by the Social Democrats, compared with the majority of the senate, tend to express more left-wing positions in their dissenting votes. Yet the dissenting opinions of three of these justices are fairly centrist, and one outlier even exhibits a pronounced right-wing position in his dissenting votes. Overall, the political positions expressed in the dissenting votes of justices selected by the Chris-

tian Democrats, including the one justice selected by the Liberals, are more centrist. Yet about half of them exhibit identifiable right-wing positions (Shikano and Koch 2010).

There is a rich quantitative literature on the jurisprudence of the US Supreme Court, and on lower US courts. Some of this literature studies the exercise of judicial discretion. In the period between 1891 and 1925, docket control was still incomplete for the US Supreme Court, which led to similar doctrinal and procedural bypasses as the ones today observed in the German Constitutional Court (for background see Eisenberg 1974, but see Freeborn and Hartmann 2010: on sensitivity of judges to the Congressional decision to reduce discretion to reduce criminal sentences). US State Supreme Courts are much less likely to reverse cases if their jurisdiction is mandatory, indicating that they use discretion to pick the cases in which they want to intervene (Eisenberg and Miller 2009).

It has been shown that, in recent years, in its decisions to grant certiorari, the Supreme Court has been sensitive to pressure from Congress (Harvey and Friedman 2009). If the Supreme Court has requested the Solicitor General to submit a brief, this is a good predictor of the court later granting certiorari (Thompson and Wachtell 2008). The Rhode Island Supreme Court's use of its discretion in granting sentence reduction could best be explained by judge effects, indicating that the exercise of discretion is influenced by idiosyncratic predilections (Braslow and Cheit 2011). Yet what looks like a bias in exercising discretion may, at closer scrutiny, result from a selection effect. While foreign plaintiffs have been more likely to win cases in US courts, this has been shown to result from successful efforts of plaintiffs to avoid US jurisdiction if odds were low they would win (Clermont and Eisenberg 2007).

Many papers have tested for an impact of judges' political orientation on their judiciary decisions (Schubert 1965, Segal and Spaeth 1993, Segal and Spaeth 2002, Sunstein, Schkade et al. 2004, Carroll and Tiede 2011). In the US Supreme Court, taking justices' ideological position with respect to civic rights into account, 60% of their pertinent decisions could be correctly predicted while, without this information, only 27% of the decisions were predicted (Segal and Spaeth 2002). In the Rehnquist court, ideological orientation predicted decisions, even if that meant for conservative justices to be activist (Solberg and Lindquist 2006). The ideology effect was particularly pronounced in judges who stood for reelection (Hall 1992, Huber and Gordon 2004, Shepherd 2009). But the ideology effect is not uncontested. Exploiting random assignment of cases for identification, one study has not found a significant effect of ideological orientation (Ashenfelter, Eisenberg et al. 1995). There also is extended debate over how to measure (pre-nomination) political orientation (Cameron and Park 2009). In the case of the US Forest Service, only published decisions were significantly explained by political orientation (Keele, Malmsheimer et al. 2009).

An equally rich literature investigates the effect of judge gender on the decisions they make, again with very mixed outcomes (see the survey in Boyd, Epstein et al. 2010). In terms of judicial quality and esteem, there was practically no difference, neither in state nor in federal

judges, despite the fact that female judges on average had received less prestigious training (Choi, Gulati et al. 2011).

It has been shown that the professional background of Justices on the US Supreme Court has considerably varied over time, but that in recent year judicial experience has been the dominant selection criterion. It has been argued that this has likely deprived the court of judicial talent, and that the criterion risks exacerbating other biases, like race or gender, that already impact on the likelihood of becoming a judge (Epstein, Knight et al. 2003).

III. The German Constitutional Court

German courts are specialized by subject matter. There are separate branches of the judiciary for criminal and private law disputes, for disputes between government and the citizens, for social security, for labor law and for tax law. Each of these branches has a specialized Supreme Court. In principle, for each dispute there can be both appeal and revision. On appeal, the case is heard again by a superior court. On revision, the facts are taken for granted. Only the violation of procedural rules, and the misconstruction of substantive law, are at issue. Appeal is in the hands of intermediate courts, while revision usually goes to the respective Supreme Court. Germany has an additional court that only decides about constitutionality. All other courts are entitled to apply the constitution, in particular using it for the reinterpretation of ordinary law, but they may not declare a statute anti-constitutional. This is the prerogative of the Constitutional Court.

There are multiple ways of bringing cases to the Constitutional Court. The bulk of cases results from constitutional complaints by individuals. In 2011, of 255 publicly available decisions 231, or 90.59 %, originated in constitutional complaints. A slightly different procedure results from complaints by individuals that provisions regarding federal elections have been violated. In 2011 the court has heard 9 such cases. If courts deem a statutory provision anti-constitutional, they may refer the case to the Constitutional Court. In 2011, the Constitutional Court has received 6 such referrals. Finally there are multiple procedures for disputes between federal authorities, or between the federation and the *Laender*.[5] In 2011, these procedures have led to 9 cases.

The court is composed of two senates with eight justices each. Justices are appointed by the President of the Federal Republic, but half of them are selected by the *Bundestag* and half of them by the *Bundesrat*, i.e. by the Federation and the *Laender*, always with a majority of two thirds. In practice, this has led to a right of initiative for half of the justices for the Christian Democrats, and for the other half for the Social Democrats. At the last occasion the Christian Democrats have let their then coalition partner, the Liberals, pick one justice, as have the So-

---

cial Democrats for their former coalition partner, the Greens. Justices are appointed for the duration of 12 years, and may not be reappointed.

Beginning in January 1998, the Constitutional Court has made its decisions publicly available on its website.[6] The data for this paper covers all 255 decisions made public in 2011. I have coded the data such that both the dependent variable and explanatory variables are available in numeric form. I will explain coding as I introduce the respective variables. The main reason for limiting the data to one year is pragmatic; coding the entire year has already been quite labor intense. But the limited time span also means that I do not have reason to be concerned about changes in the court's reactions to explanatory variables over time.

IV.   How Seriously Does the Court Take a Case?

The German Constitutional Court is obliged to decide all cases that are brought. Nonetheless, the court does not devote the same amount of effort to all cases. In this more subtle way it allocates scarce judicial resources to cases, and eases the caseload. While it seems plain that the court makes such choices, they are difficult to trace. In this section, I introduce the multiple observable indicators. I next show that they are highly correlated. This observation invites the aggregation of information from individual indicators. Using exploratory factor analysis, it turns out that a single factor best captures the information. Loadings from the individual indicators have an intuitive interpretation.

The first indicator comes from the decision itself. Doctrinally, whether a reference to the Constitutional Court is admissible or not is a question of law, not of discretion. Yet with tongue in cheek, observers of court practice have said: if the Court wants to hear a case, it'll find reasons why it is admissible; if the Court does not want to hear the case, it'll find ways to declare it inadmissible. There is one procedure that supports this view. In 2011, the Court has received six references from lower courts for preliminary ruling. The Court has declared all of them inadmissible, usually with lengthy reasons that do not only cover procedural issues, but substance. The obvious explanation stems from the fact that, in this procedure, chambers may only declare the reference inadmissible, and do not have jurisdiction to decide in substance.[7] If chambers did not want to burden the senate with cases they deem less relevant, declaring them inadmissible is the way out.

For sure, for some cases, admissibility is patently out of the question. The fact that the court declares the case inadmissible need therefore not follow from discretion but may only indicate that the court follows the procedural rules given to it by the legislator. Yet there is enough room at the margin to make the fact that the case is declared (in)admissible a potential indicator of how seriously the court has taken that particular case. Over all procedures, the court has declared 72 cases, or 28.24 % of all cases, inadmissible, or has confined to its decision to pre-

---

6   http://www.bundesverfassungsgericht.de/entscheidungen.html.
7   § 81a Bundesverfassungsgerichtsgesetz.

liminary procedural issues. Since the court has avoided to decide on the merits of the case, one might argue that it has taken the case less seriously.

The second indicator is the body that has taken the decision. Constitutional complaints are screened by chambers of three justices. The chamber may refer the case to the senate. But it may also unanimously decide that the case is inadmissible, admissible but unwarranted, or even admissible and warranted. As mentioned, if lower courts referred the case to the Constitutional Court, the chamber may only screen for admissibility. Over all procedures, in 2011 219 decisions, or 85.88 %, have been taken by a chamber. Propelling the case to the senate is certainly a way of taking a case more seriously.

The third indicator is a procedural choice. Chambers sit *in camera* by design, while senates may choose to have an oral hearing with the parties, and if the Court deems fit also with other stakeholders and experts. Arguably if there is a hearing, this indicates that the Court takes the case particularly seriously. Of the 36 senate decisions, this happened in four cases.

The fourth indicator is also only available for decisions taken by a senate. The only decision chambers may take if they are not unanimous is to refer the case to the senate. These intermediary decisions are not made publicly available. Senates are free to decide by majority. If they do, the size of the majority is reported. In 2011, 2 cases have been decided by a majority of 5, one case by a majority of 6, and 4 cases by a majority of 7. To make the size of the majority comparable across court composition, I translate these numbers into a fraction of the respective bench.[8] If the court decides by majority, this implies dissent that was so powerful that it could not be removed by debate or by striving for a compromise. Consequently, majority decisions indicate that the court had to struggle with the case more than with others, which one may interpret as a signal that the court has taken the case particularly seriously.

The fifth indicator results from the way how the court deals with the case. In principle, it is for the complainant to inform the Constitutional Court about the facts of the underlying conflict and, if necessary, to provide background information. Yet the court is entitled to give other stakeholders, and experts, the opportunity to submit the equivalent of an *amicus brief*.[9] However, unlike the law in the US, in Germany outsiders may only address the court if the court solicits their view. If it does, this is reported in the final decision. In 2011, the court has exercised this option in 75 cases, or in 29.41 % of all cases. Asking for *amicus briefs* indicates that the court either deems its decision particularly relevant for one stakeholder, or that it wants to further clarify the issue. Both suggest that the Court takes this case more seriously than others.

The sixth indicator is a straightforward measure of the energy the court devotes to a case: the length of the decision. In the electronic version of its decisions, the court numbers paragraphs,

---

8   Occasionally, senates decide with less than 8 justices, e.g. since the successor to a justice who has left the court has not yet been appointed. In 2011, one decision was taken by a senate of 6, and 3 were taken by a senate of 7. All these four decisions were unanimous.

9   § 27a Bundesverfassungsgerichtsgesetz.

the mean being 25.44 paragraphs. But variance is pronounced, with many decisions being fairly short. Unsurprisingly, the very lengthy decisions are all taken by senates. But chamber decisions may be remarkably extended as well, the longest of them having 76 paragraphs.

A seventh indicator is taken from the way how the court writes. It sometimes formulates a few sentences akin to black letter rules. These sentences precede the official decision and are mainly directed to the legal community, telling them how the court sees the current decision contribute to the evolution of constitutional law, but they are of course also noted by the wider public. The court only does so if the decision has been taken by a senate. In 2011, 15 of 36 senate decisions, or 41.67 % of them, have such rules. Arguably, black letter rules indicate that the Court takes the case seriously since it envisions that the decision will have ramifications for future cases.

An eighth indicator results from the way how the court communicates its decision. Whenever the Court formulated black letter rules, it also issued a press release. But press releases are also made available in 33 more cases. 27 cases decided by chambers are accompanied by a press release. All in all, they are issued in 18.82 % of all cases. A press release indicates that, in the Court's perception, the wider public has an interest in this particular case, or in the reasons for deciding it.

A more ambiguous ninth indicator for the effort the court devotes to the case is the time it takes for deciding. The indicator is weak, since the court might also procrastinate on cases it deems less relevant or interesting. The *rubrum* informs about the year when the case has reached the court. 94 cases have been decided in the same year of 2011, 89 are from 2010, 30 are from 2009, 25 are from 2008, 8 are from 2007, 3 are from 2006, and 6 are from 2005.

The tenth indicator exploits a feature of German constitutional doctrine. In its jurisprudence, the German Constitutional Court has shaped the German constitution in a way similar to the *Lochner*[10] era in US constitutional law. Any intervention into freedom or property, and any public act that might treat comparable cases differently, come under the scrutiny of the Constitution. If one of the specific guarantees of freedom or equality is applicable, it provides the constitutional standard. But if not, the general clauses of freedom (Art. 2 I Basic Law) or of equality (Art. 3 I Basic Law) may be invoked.[11] The fact that any exercise of sovereign powers comes under constitutional scrutiny is, however, not tantamount to constitutional verdict. In principle, any interference with a fundamental right may be justified if only it turns out proportional, given the aim pursued by government. One should therefore expect that all cases, if they are not declared inadmissible, have a section on proportionality. Actually, this is not true. The court applies the proportionality test only in 110, or in 43.14 %, of all cases. That practice makes discussion of the proportionality test a good indicator of how seriously

---

10   *Lochner vs. New York*, 198 U.S. 45 (1905); it is often said that the era came to an end with *West Coast Hotel Co. vs. Parrish*, 300 U.S. 379 (1937).
11   Leading case German Constitutional Court Jun 6, 1989, BVerfGE 80, 137 – Reiten im Walde.

the court took the case. If the test is applied, this means that the court discusses the merits of the case.[12]

As a first analytic step, Table 1 shows that these 10 indicators are highly correlated. Correlations are weakest, both in terms of the size of the correlation and of significance levels, between declaring a case inadmissible and the remaining indicators. Duration is not significantly correlated with having an oral hearing, and with not deciding unanimously. Applying the proportionality test is not significantly correlated with the decision being taken by a chamber. All remaining correlations are significant, and mostly strong.

	proc	chamb	oral	maj	brief	length	black	press	durat	prop
proc	1	-.021	-.009	.066	-.309***	-.161*	-.046	-.079	-.188**	-.265***
chamb		1	-.311***	.368***	-.233***	-.438***	-.617***	-.410***	-.202**	-.011
oral			1	-.581***	.196**	.491***	.505***	.262***	.096	.145*
maj				1	-.231***	-.426***	-.543***	-.309***	-.068	-.146*
brief					1	.520***	.387***	.284***	.242***	.289***
length						1	.761***	.525***	.308***	.361***
black							1	.519***	.229***	.186**
press								1	.322***	.290***
durat									1	.187**
prop										1

Table 1
Correlation of Indicators for Taking Cases Seriously
proc: court has not taken a decision on the merits; chamb: decision has been taken by a chamber (not a senate);
oral: (senate) has held an oral hearing; maj: (senate) has decided by majority; brief: court has solicited briefs;
length: number of paragraphs, in the official electronic version; black: (senate) has formulated black letter rules;
press: court has made a press release; durat: years that the case takes before decision; prop: proportionality test is applied
standardized correlations, N = 255
*** $p < .001$, ** $p < .01$, * $p < .05$

While these correlations suggest that these 10 indicators have something in common, it is still unclear what constitutes the common denominator, and how much and in which way the individual indicators contribute to the underlying construct. In the presentation of the indicators, I have already explained why one has reason to believe that all of them measure different dimensions of the court taking a case seriously. That invites a statistical procedure that is able to construct this underlying latent variable, and to investigate whether these indicators are manifestations of the court taking a case more or less seriously. The appropriate procedure to do this is (exploratory) factor analysis.

Arguably, the 15% cases not resulting from constitutional complaints are less comparable. To be on the safe side, I therefore confine the statistical analysis to the 231 decisions on constitutional complaints.[13] The procedure finds 10 separate factors. But only the first factor has an Eigenvalue > 1,[14] and therefore above the explanatory power of individual observed variables. This implies that the 10 indicators can be aggregated to a unique latent variable.

---

12   Except if it is one of the rare cases where fundamental freedoms are irrelevant for legality. Since for estimation I confine the dataset to cases resulting from constitutional complaints, for my statistical models this is not an issue.
13   Results look similar if I keep these cases in. The additional estimates are available upon request.
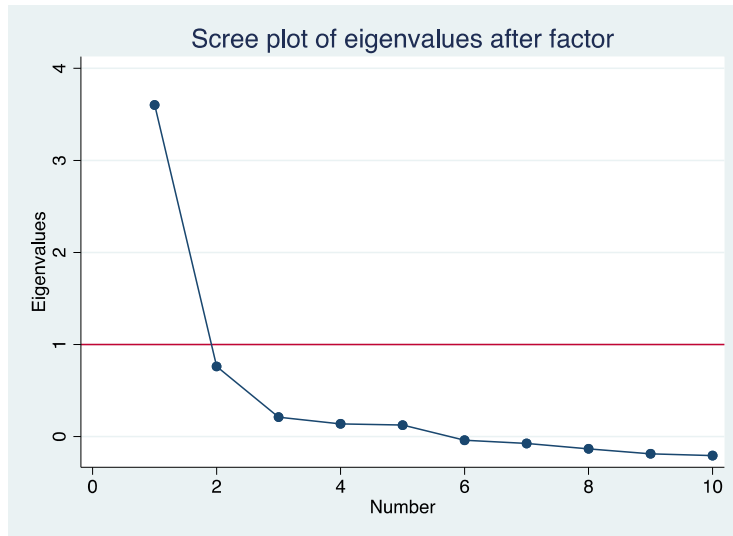14   The Eigenvalue of the first factor is 3.602.

Figure 1
Factor Analysis Finds Unique Measure

Table 2 shows that the first factor has an intuitive interpretation. Three indicators have a negative loading: if they are high, the court has taken the case less seriously. Results suggest that a decision by a chamber, and a decision only on procedural grounds, indicate that the court has taken the case less seriously. The variable "size of majority" is 1 except if there was a majority vote. Hence it must be interpreted in the reverse: if the court decides by majority, it has taken the case more seriously. This also holds if it has applied the proportionality test, if it has taken longer to decide, if it has solicited briefs, if it has held an oral hearing, if it has made a press release, and if it has written a longer verdict. The (positive or negative) size of the coefficient can be interpreted as a measure for the relative importance of the respective indicator. Hence a decision by a chamber is the strongest indicator that the court has not taken the case seriously, and black letter rules are the strongest indicator that the court has taken the case seriously.

	loading
chamber	-0.774
size of majority	-0.539
no decision on the merits	-0.215
proportionality test	0.376
duration	0.400
briefs solicited	0.491
oral hearing	0.504
press release	0.601
length	0.845
black letter rules	0.891

Table 2
Relative Weight of Indicators that Court Takes a Case Seriously
factor loadings for first factor

Figure 2 collects the seriousness scores of individual cases that are recovered from this statistical model. The distribution is heavily right-skewed. The court has taken the mass of all constitutional complaints not very seriously on which is has decided in 2011. Most of them do even have a negative score, indicating that the court has allocated very few resources to them.
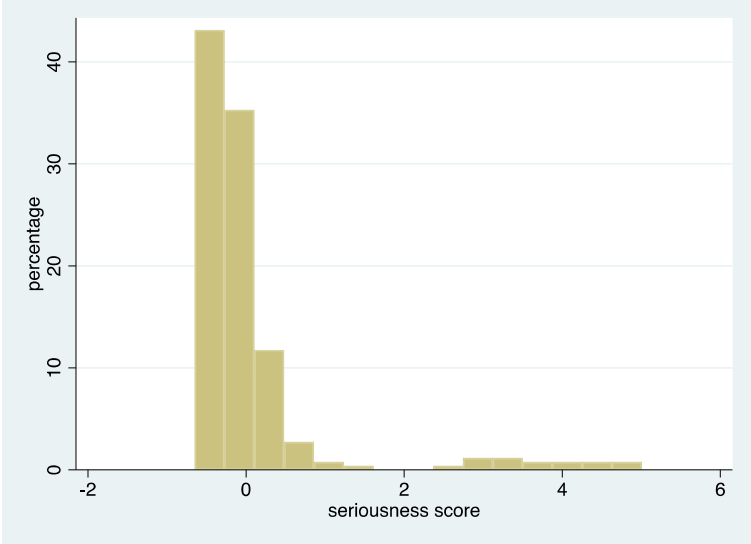

Figure 2
Seriousness Scores
predicted values from factor analysis (loadings reported in Table 2)

V.   Has the Court Been Biased?

As pointed out in the introduction, the fact that the court discriminates between cases when deciding how much effort to devote to one of them is not per se normatively problematic. To the contrary: if the court allocates its limited resources to the cases that deserve it, it does a better job. Or in the language used in the introduction: such discrimination would be efficient. But efficiency might come at the price of bias. In this section I want to explore whether there is reason for suspicion.

The international literature on supreme courts has been most interested in interpreting them as political actors or political arenas, and testing how the political procedure for selecting judges, their measurable political attitudes, demographic characteristics of individual judges or panels impact on case outcomes. The following section contributes the German case to this literature. It is limited in one way, but fairly information rich in another. The limitation results from the fact that (the rare case where justices write a dissenting opinion notwithstanding) votes of individual justices are not reported. More importantly, the large majority of all decisions are taken unanimously in the first place. Therefore the data does not allow to identify choices of individual judges. Yet the fact that more than 85% of all decisions are taken by chambers, and

that chamber composition varies widely, provides a rich dataset for measuring the effect of various indicators of bench composition on the degree by which the court takes the concrete case seriously.

The court always reports which justices have been on the bench. Justices differ in seven observed dimensions. Consequently, in each of these dimensions, the composition of the bench may be characterized. In so doing I always follow the same procedure. I first record or generate the respective measure for each justice. For the bench in question, I then calculate the average. Take the example of gender. If two members of a chamber of three have been female, my indicator for the fraction of females on that bench is at .66. Note that this procedure is not vulnerable to the identification problems known as ecological inference (King, Rosen et al. 2004). For I do not aim at inferring from the behavior of the bench how seriously female justices take cases. All I investigate is in which way entire benches behave differently if the fraction of females on the bench goes up.

Table 3 summarizes in which ways the 18 justices who have been on the court in 2011 differ in observable ways. During the year, three justices have left the court (Bryde, Hohmann-Dennhardt, Mellinghoff), and two justices have joined the court (Baer, Britz). The third open position has only been filled in 2012. The first column lists the number of months a justice has been on the court in January 2011. For decisions taken later in the year I add the pertinent number of months. The next column reports the age of the justice, calculated as 2011 minus the year of birth. The following two columns are dummies that are 1 if the justice is a woman or a law professor, respectively. Note that more than half of the justices are professors of law. They keep their university positions, with the right to return once they have completed their term at the Constitutional Court.

As mentioned, effectively the political parties have fixed slots at the court. The next column lists the party that has nominated the respective justice. Individual justices have heard differently many cases throughout the year. Those who have been on the court for the entire year have heard between 39 and 87 cases. The final column reports a measure for the variance of bench composition. For each case, I calculate the fraction of cases the deciding justices have heard in this composition throughout the year. The theoretical maximum of this measure is 1. That would imply that the same justices have been together for every case they have heard in the year. For the purposes of Table 3 only, for each justice I average this number over all cases she has heard. The measure shows that most of the time most justices have decided jointly with colleagues with whom they have taken many other decisions throughout the year. Yet not so rarely they also had to come to terms with colleagues with whom they interacted less frequently. And there was variance between justices in this respect. Some of them faced an unusual bench composition considerably more often than others.

	tenure	age	female	professor	party	cases	joint
Baer	-2	47	1	1	Green	45	.827
Britz	-2	43	1	1	SPD	53	.906
Bryde	119	68	0	1	Green	9	.566
di Fabio	132	57	0	1	CDU	41	.746
Eichberger	68	58	0	0	CDU	39	.770
Gaier	73	57	0	0	SPD	63	.868
Gerhardt	89	63	0	0	SPD	56	.782
Hermanns	0	52	1	0	SPD	41	.743
Hohmann-Dennhardt	143	61	1	0	SPD	11	.655
Huber	0	52	0	1	CDU	87	.849
Kirchhof	38	61	0	1	CDU	80	.822
Landau	62	63	0	0	CDU	54	.759
Lübbe-Wolff	104	58	1	1	SPD	87	.849
Masing	32	52	0	1	SPD	38	.774
Mellinghoff	119	57	0	0	CDU	78	.855
Paulus	9	43	0	1	FDP	63	.868
Schluckebier	50	62	0	0	CDU	51	.797
Voßkuhle	31	48	0	1	SPD	44	.749

Table 3
Bench Characteristics

As I have pointed out, I only observe benches (and their composition), not individual justices. Therefore for the following analysis, for each case I average each characteristic over the justices that have been sitting on the bench. As Table 4 shows, some of these characteristics are highly correlated. If the fraction of justices nominated by the Social Democrats or the Greens was high, on average the justices on the bench were less long on the court, have decided less cases during the year, have been less likely to be professors or female, and were younger. Conversely, the judges who have been most frequently together on the same bench have dealt with more cases, were more likely to be professors, and were younger. If the judges on a bench had higher tenure, they dealt with more cases, were older and less likely to be professors. Those who dealt with more cases were more likely to be professors and female. Professors were more likely to be female, and female justices were elder.

	left	joint	tenure	workload	professor	female	age
left	1	-.033	-.190**	-.487***	-.476***	-.461***	-.554***
joint		1	-.115$^+$	.386***	.285***	-.072	-.309***
tenure			1	.369***	-.448***	-.009	.547***
workload				1	.440***	.227***	-.024
professor					1	.135*	-.316***
female						1	.375***
age							1

Table 4
Correlation of Bench Characteristics
left: fraction of justices on the decision body that have been nominated by the SPD or the Greens
joint: mean of fraction of decisions taken jointly by justices on the decision body
tenure: mean number of months justices on the decision body have been on the court
workload: mean of sum of cases decided by justices on the decision body
professor: fraction of justices on the decision body who are professors
female: fraction of justices on the decision body who are female
age: mean age of justices on the decision body
*** p < .001,** p < .01, * p < .05, $^+$ p < .1

If explanatory variables are highly correlated, multiple regression can be misleading. The fact that one explanatory variable turns out insignificant when controlling for the other, and vice versa, may only imply that too much of the variance in the dependent variable is already explained by the respective other independent variable. I therefore first check whether each bench characteristic in isolation significantly explains how seriously the court takes a case, but thereafter also present multiple regressions that control for all other bench characteristics.

Bench characteristics are not the only possible determinant of the court taking a case seriously. Arguably, case characteristics matter as well. Since this paper is interested in the effect of bench characteristics, I only check, however, whether the effect of the respective bench characteristic changes if I control for observable case characteristics, namely: whether the constitutional complaint is directed against a statute or the decision by a (subject matter) supreme court, rather than the decision of a lower court or an administrative agency;[15] whether the complainant has been represented by a lawyer, or even by one of the highly visible big law firms or a professor;[16] how many briefs by third parties the court has received (as an indicator for the interest of outsiders in the outcome of the case); which subject matter was at stake.[17]

In these estimations, the dependent variable is constructed. It is the predicted value from exploratory factor analysis: how seriously did the court take the case in question. Bench characteristics are also specific for individual cases. The analysis therefore twice looks at characteristics of each case: first for constructing the dependent variable, and then for explaining it with bench characteristics. One may be concerned that this procedure puts too much stress on the reliability of the constructed seriousness score. As a safeguard, I therefore also estimate structural equation models that simultaneously generate and explain the dependent variable. To the extent that I exclusively look at decisions taken by chambers, this alternative procedure has an additional advantage. As explained in the previous section, numerous indicators for the court taking the case seriously are by design only available if the decision has been taken by a senate. In the structural model that exclusively looks at decisions taken by chambers, I do not use these indicators for generating the dependent variable.

---

15   Both explanatory variables are dummies.
16   Both explanatory variables are dummies.
17   Dummy variables for cases that have a business context, concern social security or welfare, family relations, disputes between private parties, concern democratic institutions, disputes between citizens and the state, or tax law. The residual category is cases that fall in none of these case categories.

	all complaints one bench characteristics		chamber only one bench characteristic		all complaints multiple regression		chamber only multiple regression		all complaints SEM		chamber only SEM	
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12
left	.454 (.386)	-.016 (.258)	.118 (.120)	.014 (.120)	-.306 (.693)	.014 (.511)	-.146 (.286)	-.066 (.268)	.018 (.046)	.001 (.034)	.125 (.184)	.063 (.161)
joint	-5.414*** (.423)	-2.712*** (.364)	-.205 (.271)	-.155 (.259)	-6.046*** (.493)	-3.374*** (.413)	.197 (.335)	-.125 (.319)	.452* (.208)	.278* (.123)	-.174 (.218)	.036 (.191)
tenure	.005 (.003)	.002 (.002)	-.001 (.001)	.001 (.001)	.005 (.005)	.000 (.004)	.001 (.002)	.001 (.002)	-.001 (.000)	.000 (.001)	-.001 (.001)	-.000 (.001)
workload	-.015** (.005)	-.001 (.004)	-.004** (.0015)	-.002 (.002)	.005 (.008)	.012$^+$ (.006)	-.006$^+$ (.003)	-.001 (.003)	-.001 (.001)	-.001 (.001)	.004$^+$ (.002)	.000 (.002)
professor	-1.312*** (.488)	-.122 (.319)	-.173 (.157)	-.123 (.149)	-.417 (.777)	-.260 (.578)	.064 (.003)	-.033 (.306)	.036 (.053)	.025 (.039)	.028 (.209)	.107 (.185)
female	-.251 (.260)	.008 (.179)	-.118 (.080)	-.149$^+$ (.083)	-.287 (.250)	-.244 (.195)	-.077 (.103)	-.195$^+$ (.103)	.019 (.019)	.016 (.015)	.057 (.067)	.134 (.066)
age	.025 (.016)	.008 (.011)	.001 (.005)	.004 (.005)	-.050 (.030)	-.016 (.023)	-.001 (.013)	.005 (.012)	.004 (.003)	.001 (.002)	-.001 (.008)	-.005 (.007)
case controls	no	yes	no	yes	no	yes	no	yes	no	yes	no	yes
cons	suppressed	suppressed	suppressed	suppressed	7.931** (2.283)	2.942$^+$ (1.762)	.046 (1.007)	-.404 (.971)	no	no	no	no
N	231	231	212	212	231	231	212	212	231	231	212	212

Table 5
Bench Characteristics as Explanations for Cases Being Taken More Seriously
Models 1-8: OLS, dv: predicted value from factor analysis (Table 2)
Models 9-12: structural equations models that simultaneously generate seriousness index and explain it with the list of bench characteristics (and case characteristics, if applicable)
variable names for bench characteristics as in Table 4
Models 2, 4, 6, 8, 10, 12: regression additionally controls for case characteristics (see text for detail)
Models 1-2, 5-6, 9-10 include constitutional complaints decided by a senate; Models 3-4, 7-8, 11-12 only cover decisions by chambers
Models 1-4: coefficients from regressions with a constant (not reported)
*** p < .001, ** p < .01, * p < .05, $^+$ p < .1

Table 5 shows that, irrespective of in the specification of the regression, there is very little sign of any bias in the court's decision to take a case more seriously. In the international discourse the one characteristic of the bench that has attracted most interest and effort is political orientation. Across all specifications, I never find a significant effect. This also holds for tenure and age. If the justices have been longer on the court, or if they are elder, this does not matter for how seriously the court takes a case. In two specifications that control for case characteristics (Models 4 and 8), I find a weakly significant negative gender effect. But the effect disappears if I simultaneously generate and explain the dependent variable (Models 10 and 12). If I use no other explanatory variable and all constitutional complaints, I find a significant negative effect of professors being on the bench (Model 1). But the effect already disappears if I control for case characteristics (Model 2). In two specifications I find a significant negative effect of workload: if the justices on the specific bench see more cases throughout the year, they take the individual case less seriously (Models 1 and 3). But the effect turns positive in the structural equation models for chambers only (Model 11) and, most importantly, it is always insignificant if I add case controls.

The most surprising finding is for the measure of joint experience. If I also consider constitutional complaints on which a senate has decided, I find a highly significant, negative effect, both if I only use this one bench characteristic for explanation (Models 1 and 2) and in multiple regression (Models 5 and 6). This suggests that, the more the bench is homogeneous in terms of having joint decision experience, the less effort they allocate to the individual case. However if I simultaneously generate the dependent variable and explain it in a structural equation model, the effect reverses (Models 9 and 10). And, most importantly, it disappears whenever I exclusively consider the cases decided by chambers. The seeming effect of joint experience is an artifact. It just captures the fact that, by design, justices have much less joint experience if they decide as a senate. This follows from the fact that the great majority of decisions are taken by chambers, and that chamber composition is fairly stable over time.

VI. Conclusions

While there is an extended quantitative literature on the US Supreme Court, and growing interest in the quantitative analysis of the jurisprudence of Supreme Courts in countries like Israel, Taiwan, or Australia, the jurisprudence of the German Constitutional Court thus far has hardly been touched by quantitative scrutiny. This paper is an attempt at helping close this gap in the literature.

The German case is interesting for multiple reasons. The court decides more than 85% of all cases in chambers of three, rather than in senates of eight justices. Consequently in a single year there is already considerable variation with respect to a multitude of bench characteristics. These characteristics include by which political party the justice has been selected, gender, age, and whether the justice is a law professor. Moreover, several measures for the com-

position of the respective bench may be calculated: how long have the justices on average been on the court, how many cases have they on average heard throughout the year, and how frequently have they together been on the bench? The German case is also interesting since the court legally must hear all cases that are referred to it.

While the German Constitutional Court may not refuse to decide a case, it has found multiple ways to take a case more or less seriously. For the scientific analyst, this provides an opportunity to observe judicial discretion, and to generate a continuous measure of how seriously the court has taken the individual case. This index is generated using factor analysis. The following indicators contribute to the generation of a measure for taking a case more or less seriously: whether the court has referred the case to a senate, whether the court has formulated the implications of the case for the evolution of constitutional law in the form of black letter rules, whether the court has solicited amicus briefs, whether it has held an oral hearing, whether the court has issued a press release, how many paragraphs the decision is long, and whether the court has applied the proportionality test. Generating predicted values from this statistical model, one sees that there is considerable variance. While the court devotes little effort to most cases, there are quite a few cases it takes more seriously. This not only holds for cases referred to the senate, but encompasses a considerable number of chamber decisions.

Is this good or bad news? The resources of the court are limited. The most severe limit is the number of justices. The statute defines that the court is composed of two senates with eight justices each. If the court devotes most energy to the cases that deserve most attention, this is normatively desirable. But the court's decision to take a case seriously might also be determined by factors that are normatively less desirable. This decision of the court might be biased. The literature has been most concerned about bias originating in bench characteristics. For the German constitutional court, multiple bench characteristics are observed. Yet I find virtually no sign of any of these characteristics influencing the courts decision how seriously to take a case.

In terms of statistical techniques, my analysis is limited. I can only use data from cases the court itself has decided to make public on its website. These cases are only a small fraction of the total of cases decided in my year of observation. I cannot exclude that the decision how much effort to devote to the remaining cases has been biased. Moreover I only observe bench characteristics. They are not randomly assigned. I can therefore not exclude that other characteristics of the bench that I cannot observe do actually have an effect. Finally frequentist statistics are problematic in the first place. Arguably the decision how much energy to devote on one case is not independent from the decision to devote energy on other cases. Arguably this also holds across chambers from the same senate (the justices must decide how much energy to reserve for choices taken by the plenum) and also for decisions taken by the respective other senate (the institutional culture of the court is strong, and the court quite openly cares about its reputation with government and the public at large). Essentially I therefore have a single observation. I acknowledge all these limitations. For the purposes of my analysis they are insurmountable. But I do not therefore consider the exercise pointless. Readers should be aware

that the evidence is only suggestive. I could at most have found a smoking gun. But I find it remarkable that there is virtually no sign of smoke. My evidence suggests that, when the German constitutional court decides how much energy to devote to a single case, it is not biased by bench characteristics.

References

ASHENFELTER, ORLEY, THEODORE EISENBERG AND STEWART J. SCHWAB (1995). "Politics and the Judiciary. The Influence of Judicial Background on Case Outcomes." *Journal of Legal Studies* **24**: 257-281.

BOYD, CHRISTINA L., LEE EPSTEIN AND ANDREW D. MARTIN (2010). "Untangling the Causal Effects of Sex on Judging." *American Journal of Political Science* **54**(2): 389-411.

BRASLOW, LAURA AND ROSS E. CHEIT (2011). "Judicial Discretion and (Un) equal Access: A Systematic Study of Motions to Reduce Criminal Sentences in Rhode Island Superior Court (1998–2003)." *Journal of Empirical Legal Studies* **8**(1): 24-47.

CAMERON, CHARLES M. AND JEE-KWANG PARK (2009). "How Will They Vote? Predicting the Future Behavior of Supreme Court Nominees, 1937–2006." *Journal of Empirical Legal Studies* **6**: 485-511.

CARROLL, ROYCE AND LYDIA TIEDE (2011). "Judicial Behavior on the Chilean Constitutional Tribunal." *Journal of Empirical Legal Studies* **8**: 856-877.

CHOI, STEPHEN J., G. MITU GULATI, MIRYA HOLMAN AND ERIC A. POSNER (2011). "Judging Women." *Journal of Empirical Legal Studies* **8**: 504-532.

CLERMONT, KEVIN M. AND THEODORE EISENBERG (2007). "Xenophilia or Xenophobia in US Courts? Before and After 9/11." *Journal of Empirical Legal Studies* **4**(2): 441-464.

EISENBERG, THEODORE (1974). "Congressional Authority to Restrict Lower Federal Court Jurisdiction." *Yale Law Journal* **83**: 498-533.

EISENBERG, THEODORE AND GEOFFREY P. MILLER (2009). "Reversal, Dissent, and Variability in State Supreme Courts. The Centrality of Jurisdictional Source." *Boston University Law Review* **89**: 1451-1504.

EPSTEIN, LEE, JACK KNIGHT AND ANDREW D. MARTIN (2003). "The Norm of Judicial Experience and its Consequences for Career Diversity on the U.S. Supreme Court." *California Law Review* **91**: 903-966.

FREEBORN, BETH A. AND MONICA E. HARTMANN (2010). "Judicial Discretion and Sentencing Behavior. Did the Feeney Amendment Rein in District Judges?" *Journal of Empirical Legal Studies* **7**(2): 355-378.

HALL, MELINDA GANN (1992). "Electoral Politics and Strategic Voting in State Supreme Courts." *Journal of Politics* **54**: 427-446.

HARVEY, ANNA AND BARRY FRIEDMAN (2009). "Ducking Trouble. Congressionally Induced Selection Bias in the Supreme Court's Agenda." *Journal of Politics* **71**(2): 574-592.

HUBER, GREGORY A. AND SANFORD C. GORDON (2004). "Accountability and Coercion: Is Justice Blind When It Runs for Office?" *American Journal of Political Science* **48**: 247-263.

KEELE, DENISE M., ROBERT W. MALMSHEIMER, DONALD W. FLOYD AND LIANJUN ZHANG (2009). "An Analysis of Ideological Effects in Published Versus Unpublished Judicial Opinions." *Journal of Empirical Legal Studies* **6**(1): 213-239.

KING, GARY, ORI ROSEN AND MARTIN A. TANNER (2004). *Ecological Inference. New Methodological Strategies*, Cambridge Univ Pr.

SCHUBERT, GLENDON (1965). *The Judicial Mind. The Attitudes and Ideologies of Supreme Court Justices 1946–1963*. Evanston, Northwestern University Press.

SEGAL, JEFFREY A. AND HAROLD J. SPAETH (1993). *The Supreme Court and the Attitudinal Model*. New York, Cambridge University Press.

SEGAL, JEFFREY A. AND HAROLD J. SPAETH (2002). *The Supreme Court and the Attitudinal Model Revisited*, Cambridge Univ Pr.

SHEPHERD, JOANNA M. (2009). "The Influence of Retention Politics on Judges' Voting." *Journal of Legal Studies* **38**: 169-206.

SHIKANO, SUSUMU AND ANNA-KATHARINA KOCH (2010). Estimation of Political Actors' Positions Using Unfolding Item Response Models. Ideological Location of Judges of Federal Constitutional Court of Germany.

SOLBERG, RORIE SPILL AND STEFANIE A. LINDQUIST (2006). "Activism, Ideology, and Federalism. Judicial Behavior in Constitutional Challenges Before the Rehnquist Court, 1986–2000." *Journal of Empirical Legal Studies* **3**: 237-261.

SUNSTEIN, CASS R., DAVID SCHKADE AND LISA MICHELLE ELLMAN (2004). "Ideological Voting on Federal Courts of Appeals. A Preliminary Investigation." *Virginia Law Review* **90**: 301-354.

THOMPSON, DAVID C. AND MELANIE F. WACHTELL (2008). "An Empirical Analysis of Supreme Court Certiorari Petition Procedures. The Call for Response and the Call for the Views of the Solicitor General." *George Mason Law Review* **16**: 237-302.

VANBERG, GEORG (2004). *The Politics of Constitutional Review in Germany*, Cambridge University Press.