Discussion Papers of the
Max Planck Institute for
Research on Collective Goods
2021/5

Lucky You: Your Case is Heard by a Seasoned Panel

Panel Effects in the German Constitutional Court

Christoph Engel

MAX PLANCK
SOCIETY

Discussion Papers of the
Max Planck Institute
for Research on Collective Goods 2021/5

Lucky You: Your Case is Heard by a Seasoned Panel
Panel Effects in the German Constitutional Court

Christoph Engel

February 2021

Max Planck Institute for Research on Collective Goods, Kurt-Schumacher-Str. 10, D-53113 Bonn
https://www.coll.mpg.de

Lucky You: Your Case is Heard by a Seasoned Panel
Panel Effects in the German Constitutional Court*

Christoph Engel

Abstract

The German Constitutional Court is radically different from the (mostly US) courts in which panel effects have been studied so widely. On the one hand, to a large extent, ideological and gender bias are neutralized by design. On the other hand, panels are not randomly composed. This makes it possible to study a panel effect that has escaped attention. It results from the degree of familiarity among the judges on the bench. The longer their joint experience on the bench, the higher the chances that a complaint is successful. Regression discontinuity allows for a causal interpretation, at least near regular, exogenous recompositions of the bench.

Keywords: panel effect, German Constitutional Court, familiarity, regression discontinuity

JEL: C12, D71, D73, D91, H11, K41

* Helpful comments by Stefanie Egidy and Sebastian Schneider on an earlier version are gratefully acknowledged. Elliott Ash, Christoph Gössmann and Philip Schmidt have been instrumental with scraping the data.

1. Introduction

Six eyes see more than two. Most jurisdictions rely on this maxim and entrust important legal decisions to panels of multiple judges. The magic number seems to be three. It strikes a balance between efficiency and diversity. Judges of different background, gender, race, experience and possibly ideological orientation may be represented. It is easy to define a majority, and hence to avoid an impasse. At the same time, compared with yet larger panels, the judicial system saves resources and may handle more cases with the existing judicial personnel.

A sizeable empirical literature demonstrates that shifting jurisdiction from single judges to panels matters. As the attitudinal model (Segal and Spaeth 1993, Segal and Spaeth 2002) features so prominently in the scholarly discourse in the US, political orientation has attracted most interest (Hettinger, Lindquist et al. 2004). The straightforward effect has indeed been found: if the majority of a panel has been appointed by a Republican president, on average decisions are more in line with conservative attitudes than if the majority has been appointed by a Democratic president (Songer 1982). Same for a majority of female judges when the case is about sex discrimination (Kim 2008, Haire, Moyer et al. 2013).

More interestingly, though, the decisions by mixed panels (on their prevalence see Kastellec 2011b) are more moderate than the decisions of panels all appointed by a president of the same party (Miles and Sunstein 2006, Kastellec 2011a), as are the opinions (Hinkle 2017). Likewise decisions are more sensitive to sex discrimination if one member of the bench is female (Farhang and Wawro 2004, Peresie 2004, Boyd, Epstein et al. 2010), and more sensitive to racial discrimination if one judge of colour is on the bench (Sommers 2006, Cox and Miles 2008, Kastellec 2013, Kastellec 2020). From the perspective of the attitudinal model, these findings create a puzzle: how can a minority judge have an effect? Most of the literature is devoted to finding explanations.

Arguing in terms of individual ideology maximisation, authors have carved out constraints. Outvoting a counterjudge can be too costly (Fischman 2011, Fischman 2013), as the majority has to put more work into drafting the decision (Epstein, Landes et al. 2011); the panel or the court at large may lose reputation and thereby legitimacy (Atkins and Green 1976); there is social pressure in the judiciary (Atkins 1973); there is a norm of consensus (Farhang and Wawro 2004); group dynamics amplify conformity (Sunstein, Schkade et al. 2007). Other authors have redefined the game. The parties may appeal the case, and the fact that the decision had not been taken unanimously may increase the prospect of the appeal, or *en banc* review (Beim, Hirsch et al. 2016), being successful. The decision is taken in the shadow of a judicial hierarchy. The minority judge may, in particular, "blow the whistle" and signal to the superior court that it should have a closer look (Cross and Tiller 1997, Kastellec 2007, Kastellec 2011a). To the extent that members of the panel have reason to anticipate future interaction, another redefinition of the game introduces a shadow of the future, and in particular scope for logrolling (Farhang and Wawro 2004, Hessick and McLaughlin 2013).

Alternative explanations are cognitive. The counterjudge alerts the majority of the panel to a concern that they feel unable to overlook (Sommers 2006, Spitzer and Talley 2013), for instance as doctrine leaves little room for manoeuvre (Tiller 2015), induces them to better guard

against the risk of implicit bias (Sommers 2006), or provides them with credible expertise. The latter argument is, for instance, brought forward to explain why the mere presence of a female judge (Boyd, Epstein et al. 2010), or a judge of colour (Sommers 2006), matters.

A third line of explanations could be called cultural. The fact that judges decide as a group may activate a culture of collegiality (Edwards 1998, Edwards 2003). Rather than insisting on ideological positions, judges may in good faith aim at doing the parties justice, finding the truth, and implementing the normative decisions of the legislator (experimental evidence is provided by Engel and Zhurakhovska 2017). In that perspective, a judge with different normative convictions, or different demographic characteristics, may bring an argument to the table that her colleagues find persuasive.

Recent, rigorous attempts have found multiple explanations to hold simultaneously, and have concluded: “Finally, this work highlights the importance of continuing to investigate the mechanisms behind panel effects, which appear to be complex, nuanced, and multifaceted” (Hinkle, Nelson et al. 2020: 279). The present paper responds to that request.

Up till now, the debate over panel effects in courts has been US centric. Exceptions concern Canada (Hausegger and Haynie 2003, Alarie, Green et al. 2015), South Africa (Hausegger and Haynie 2003), Israel (Grossman, Gazal-Ayal et al. 2016), and arbitration panels set up under the umbrella of the International covenant for the settlement of investment disputes (Kapeliuk 2012). These comparative investigations have largely followed the scholarly US tradition. Yet different legal orders do not only come with different legal cultures. Arguably the courts in, for instance, Germany are much less politicised. They also come with very different institutional arrangements.

Germany has a separate court for constitutional scrutiny. The bulk of cases heard by the Federal Constitutional Court are constitutional complaints brought by individuals. At least officially the court does not have the power of certiorari, and indeed hears thousands of cases per year. To manage the caseload, two senates of eight justices for the most part split into panels ("chambers") of three. The court has always made an effort to construct these panels as diversely as possible, in multiple dimensions. Most cases are not randomly assigned, but each justice has a docket, defined by some subject matter. Chambers stay together for a protracted period of time, typically multiple years. Even if chambers are recomposed, individual justices keep their cases.

These institutional features make it possible to empirically study arguable determinants of collegiality. Each justice goes through periodic changes in joint experience on the bench. Once the chamber has been recomposed, it must find its rhythm. The justices know that they will have to collaborate closely for the years to come. They increasingly observe the working style, the legal competence, specific strengths and weaknesses of their new colleagues. They learn which of their arguments other justice on this bench see as compelling, and how they react to their persuasive strategies. The consistent composition of the chamber facilitates the formation of mutual trust. The chamber may forge a local culture, and may increasingly become self-confident.

Observing only the outcomes, i.e. the decisions taken by the chamber, does not make it possible to discriminate between these channels, or to isolate yet other channels through which joint experience matters. Yet as a first step it is interesting, and also normatively important, that joint experience does indeed matter for outcomes. There are multiple ways (introduced below) of defining whether a constitutional complaint has been successful. Most of these dependent variables do indeed react to the degree of familiarity resulting from joint experience on the bench. The complainant benefits. She is more likely to be successful.

Yet is the degree of familiarity causal for outcomes? As explained, chamber composition regularly changes. These changes are beyond the control of the individual justice, and only partly predictable. This provides scope for regression discontinuity, and for establishing a causal effect near the point in time when, for the individual justice, familiarity suddenly drops. It turns out that there is indeed a substantial and significant local effect, on multiple indicators of success.

The small empirical literature on the German Constitutional Court has not looked at panel effects (Vanberg 2004, Engst, Gschwend et al. 2017, Engst, Gschwend et al. 2020, Lang 2020). In an earlier attempt, I have found that panel homogeneity is associated with a composite index for the court taking a case more seriously (Engel 2020). That paper uses a much smaller dataset, analysis is not causal, and I investigate a different dependent variable.

The remainder of the paper is organised as follows: the next section introduces the institutional framework of the German Constitutional Court. Section 3 explains the character and scope of the data. Section 4 reports results. Section 5 concludes with discussion.

2. The German Constitutional Court

The German Constitutional Court is the highest court of the land. There are specialised supreme courts by subject matter, like the Federal Administrative Court or the Federal Tax Court, including the "Federal Court", which is actually a specialised court for private and criminal law. Any case decided by the final court having jurisdiction in the subject matter can be brought before the Constitutional Court. Citizens can in principle also directly complain to the Constitutional Court about a statute or other legislative act. The Constitutional Court has the monopoly over declaring legislation unconstitutional. If a lower court comes to the conviction that a provision of a statute is unconstitutional, and if this question is critical for deciding the case in front of that court, it may refer the question to the Constitutional Court. The Constitutional Court has also jurisdiction if a citizen or a political party questions the legality of an election. Municipalities may bring alleged violations of the constitutional protection of their autonomy to the Constitutional Court. The Constitutional Court is competent to decide whether an alleged rule of customary international law exists and is part of the law of the land. Each of the Länder may bring a case against other Länders or the Federation. One of the Länder, or a third of the members of Parliament, may challenge an act of legislation. The Constitutional Court decides disputes between federal institutions, like the second chamber that represents the Länder, Parliament, or government.

These are only the practically most important types of cases the court has jurisdiction to see.[1] Since the first day of deciding a case in Sep 7, 1951 and Dec 31, 2019 the court has had to deal with 243,494 cases, of which 235,057 have been constitutional complaints. On Dec 31, 2019 3,472 cases (3,300 constitutional complaints) have been pending.[2] Some types of cases (like the impeachment of the President of Republic) have never been brought, and others (like limiting an individual's exercise of fundamental rights: 3 cases ever since) have been exceedingly rare.[3] The court does not have the power of *certiorari*; it officially has to decide each case. Over the years, the court has developed a whole panoply of techniques for implicit docket control (Engel 2020). The most important techniques have, however, been gradually been given to the court by the legislator.

The court is composed of two senates with eight justices each. If the case, however, originates in a constitutional complaint (5,158 cases have been brought in 2019),[4] or has been referred to it by a lower court (20 cases have been referred in 2019),[5] the decision may be taken by a chamber of only three justices. While chambers originally only had jurisdiction to reject complaints, they now also have jurisdiction to accept them.[6] Decision by chamber however presupposes unanimity.[7] Otherwise the case is propelled to the senate sitting *en banc*. The senate also decides if the chamber is of the opinion that the subject matter is so important that the full senate should decide. In 2019, no more than 8 constitutional complaints have actually been decided by the respective senate.

Justices have a fixed tenure of 12 years, which is non-renewable. They can resign earlier. But this is a rare event.[8] Justices are appointed, half of them by Parliament, and half of them by the second chamber of parliament, i.e. the representatives of the Länder. In practice slots are assigned to political parties, with half of the slots going to the (more conservative) CDU or FDP, and half of them going to the (more liberal) SPD or the Greens.[9] Thus far, the far right AfD has been denied a position.[10]

In court practice, each justice has a fixed docket. For the most part, the docket is defined by subject matter. Essentially, the justice inherits the docket from the justice whom she replaces, and the docket changes at most very little during her time on the court.[11] Officially, each senate

1 For a more exhaustive list see § 13 Bundesverfassungsgerichtsgesetz.

2 https://www.bundesverfassungsgericht.de/DE/Verfahren/Jahresstatistiken/2019/gb2019/A-I-1.pdf?__blob=publicationFile&v=2.

3 https://www.bundesverfassungsgericht.de/DE/Verfahren/Jahresstatistiken/2019/gb2019/A-I-4.pdf?__blob=publicationFile&v=3.

4 §§ 93a – 93d Bundesverfassungsgerichtsgesetz.

5 § 81a Bundesverfassungsgerichtsgesetz.

6 § 93c I 1 Bundesverfassungsgerichtsgesetz.

7 § 81a, 1 Bundesverfassungsgerichtsgesetz; § 93d III 1 Bundesverfassungsgerichtsgesetz.

8 One of these exception is Justice Mellinghoff, who has resigned little more than a year before the end of his 12 year term, to become President of the Federal Tax Court.

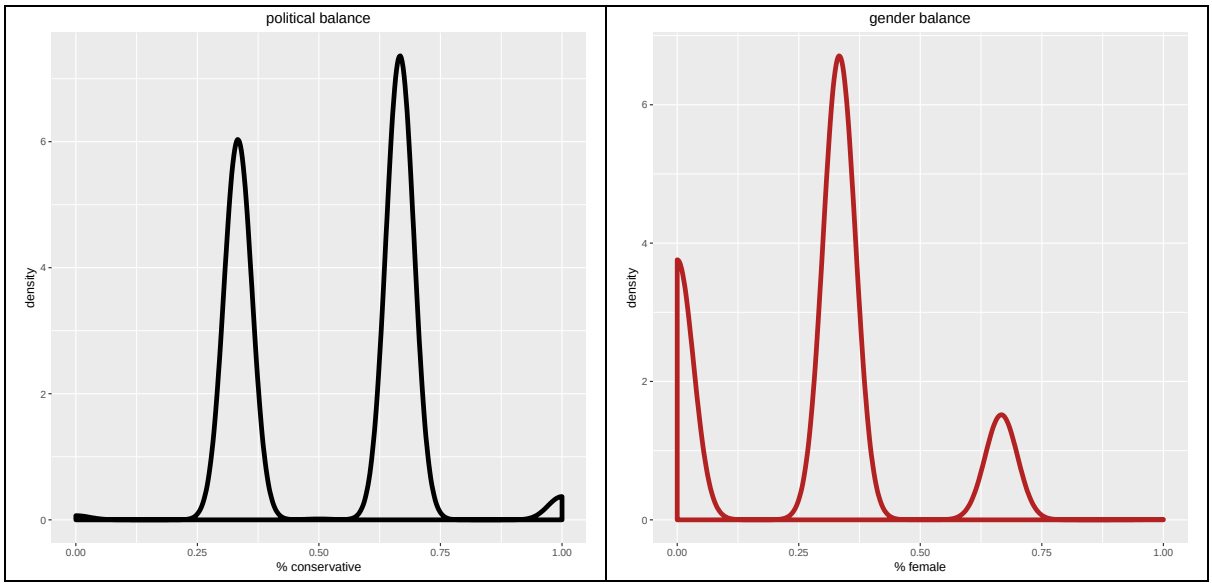
9 For a complete all-time list, see https://de.wikipedia.org/wiki/Liste_der_Richter_des_Bundesverfassungsgerichts.

10 It has only been represented for the first time in Parliament in 2017 though.

11 To illustrate, when Justice Baer joined the court in 2011, she replaced Justice Bryde. Her initial docket was identical with Justice Bryde's docket, except for one (of 10) issue areas. While Justice Bryde was also responsible for subsidies to students, this matter was shifted to Justice Schluckebier. Justice Baer's docket for 2020 is the same as the one she had when joining the court, except for social security, which is now part of Justice

decides every year about the composition of each of three chambers for the coming year.[12] In practice, chamber composition tends to be constant for a longer period. After three years, the statute wants composition to change, at the latest.[13] As senates have 8 members, not 9, one justice must sit on two chambers. In the past, this has often been the presiding justice.

Figure 1 shows that the court makes an effort to balance chamber composition. Of 6,359 chamber decisions posted on the court's website, no more than 169 have been taken by an all-Conservative panel, and no more than 30 by an all-liberal panel. All other chambers were ideologically mixed. All-male chambers are more frequent (31.36% of all decisions). But today exactly half of the justices are female, so that the gender balance in the chambers is bound to improve. The court also makes an effort to balance tenure. Typically, one justice who has been on the bench for a long time decides together with one or two justices who have recently joined the court. Age tends to be balanced as well.



Britz' docket, most likely as there are too many cases on welfare for asylum seekers, which remains part of Justice Baer's docket. The yearly decisions of the court about the justices' dockets are available at https://www.bundesverfassungsgericht.de/DE/Verfahren/Geschaeftsverteilung/archiv_geschaeftsverteilung.html.

12 The decisions are available here https://www.bundesverfassungsgericht.de/DE/Verfahren/Geschaeftsverteilung/archiv_geschaeftsverteilung.html.

13 § 15a I 2 Bundesverfassungsgerichtsgesetz.

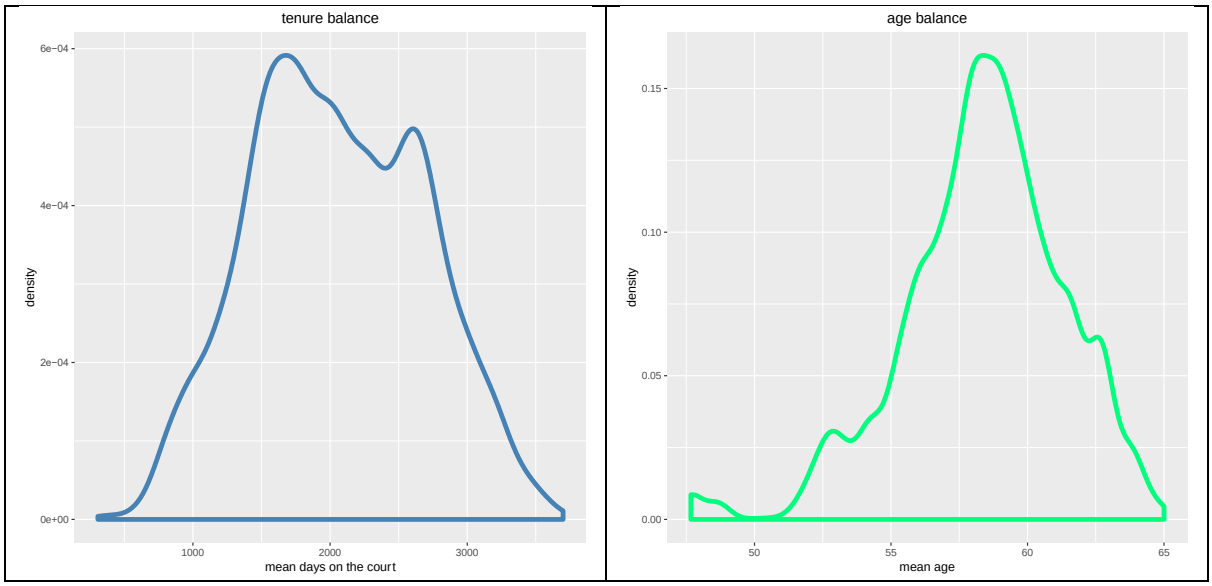


Figure 1
Chamber Composition
all 6359 chamber decisions posted on the court's website until July 2020
density plots
conservative: Justice has been picked by CDU or FDP

3. Data

Since its inception, and until Dec 31, 2019, the German Constitutional Court has heard 243,494 cases.[14] The court has routinely published decisions. Yet printed reports have always been selective, and in particular only cover a small portion of chamber decisions. Since 1998 the court regularly posts decisions on its website.[15] This paper uses a dataset that originates in scraping all decisions that have been posted online, until July 2020. It includes a small number of (landmark) decisions from earlier years that the court has added to the online collection. Senate rulings are covered fairly comprehensively.[16] Online, there are also many more chamber decisions than printed. Still the coverage of chamber decisions remains substantially incomplete. For instance of the 4,754 chamber decisions taken in 2019, only 224, i.e. less than 5%, are available online. One has reason to believe that the court has legitimately considered the remaining decisions to be of little importance. But one cannot check back.

Online, 7,614 decisions have been posted. 1,247 of them have been decided by one of the two senates, sitting *en banc*. There are 8 plenary decisions in the dataset.[17] The remaining 6,359

14 https://www.bundesverfassungsgericht.de/DE/Verfahren/Jahresstatistiken/2019/gb2019/A-I-1.pdf?__blob=publicationFile&v=2.

15 https://www.bundesverfassungsgericht.de/SiteGlobals/Forms/Suche/Entscheidungensuche_Formular.html?language_=de.

16 In 2019, the two senates have taken 54 decisions (plus another 15 preliminary and subsidiary decisions). In 2018 the senates have taken 33 decisions (plus 41 preliminary and subsidiary decisions), https://www.bundesverfassungsgericht.de/DE/Verfahren/Jahresstatistiken/2019/gb2019/A-III.pdf?__blob=publicationFile&v=2. Online, 2019 51 decisions are posted, and 53 decisions in 2018.

17 The plenary decides if one senate plans to overrule the other senate.

decisions have been taken by a chamber. 6,466 decisions have originated in a constitutional complaint. There are 384 requests for preliminary injunctions.[18] 322 cases have been referred to the constitutional court by lower courts, arguing that a statute violates the constitution. There are 200 election complaints. The total number of disputes between federal authorities (100) and between the Federation and one of the Länder (19) is small. But some of these cases have been very important.

The first line in Table 1 might suggest that those who bring a case before the constitutional court stand a fair chance to win on the merits. More than a quarter of all cases is successful, and almost half of the constitutional complaints heard by a chamber. Yet this impression is misleading. A very large majority of cases is summarily rejected, as some procedural requirement has not been met, or as the case does obviously not have merit. Little more than 5% of the constitutional complaints submitted to one of the chambers pass this hurdle. And if non-acceptance is taken into account, only 2.35% of the posted constitutional complaints are successful. Very likely almost all of the non-posted complaints were unsuccessful as well. Effectively the constitutional court only intervenes very rarely.

Requests for preliminary injunctions are rare in the first place. But if the court hears such a request, and thereby grants the urgency of the case, chances for success are brighter. Almost a quarter of the requests for preliminary injunctions filed with a constitutional complaint that the court has not declared inadmissible are granted. In procedural matters, success is much more likely. Almost 83% of all procedural decisions regarding the way how the court deals with a constitutional complaint end in favour of the complainant. This is very likely evidence of a selection effect. The court has posted the decision because it wanted to clarify some procedural matter. As the fraction of success in procedural matters is high, so is success in either the merits, the request for a preliminary injunction, or procedural matters, i.e. in the combined category.

	all cases		chamber		chamber constitutional complaint	
	failure	success	failure	success	failure	success
merit	902	250	398	94	114	94
merit + non-acceptance	4881	250	4343	94	3902	94
preliminary injunction	982	308	899	249	644	197
procedural decision	473	2125	391	1766	367	1766
combined	1577	2349	1100	1937	825	1882

Table 1
Success in the Constitutional Court

The main explanatory variable investigated in this paper is the degree of familiarity among the justices currently on the bench. Familiarity is calculated the following way: for each justice a dummy is 1 if the justice has been on the bench for the decision in question. A second set of dummies, for each decision in the dataset, matches each justice with each other. For each

18 They may, but need not, have originated in a constitutional complaint.

justice and each point in time, these bilateral dummies are summed up. This sum is divided by the total number of cases the justice has decided up till this point in time, with whichever other justices on the bench. Using this procedure, there is a score for each pair of participating justices. Per justice and point in time, these scores are summed up, and divided by the number of justices participating in the present decision. Figure 2 reports the distribution of the mean, minimum and maximum degree of these familiarity scores.

As Figure 2 shows, the shape of these three distributions is comparable whether or not senate decisions are included, and whether or not all types of cases are covered, or only constitutional complaints. Yet if senate decisions are included, the distributions shift to the left, as now the full senate is covered, not only the chamber of three.

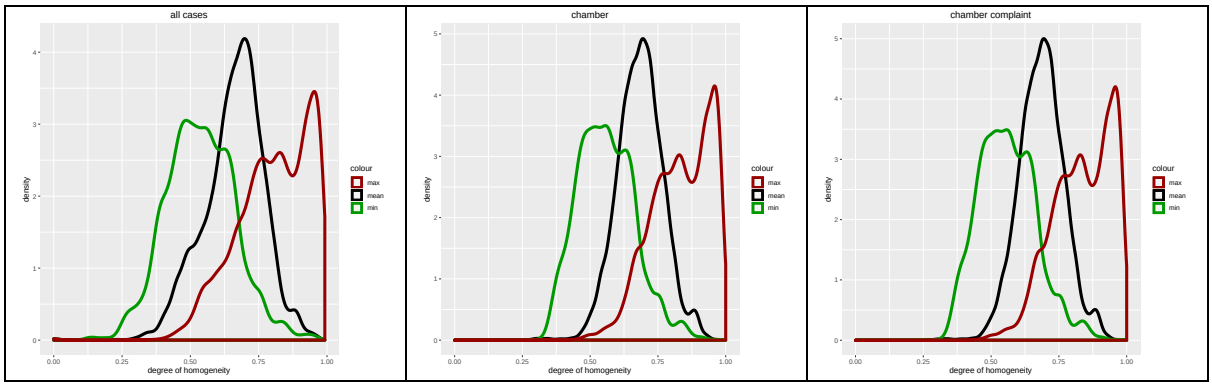


Figure 2
Degree of Minimum, Mean, and Maximum Familiarity

4. Results

a) Association between the Degree of Familiarity and Success of a Constitutional Complaint

As Figure 3 shows, the degree of familiarity with colleagues on the bench matters for the probability that a constitutional complaint is successful. Overall, a higher degree of familiarity is associated with a higher probability of success. The effect is clearest if one looks at the maximum degree of familiarity. It originates in the justice who has been most consistently deciding together with one of the other two justices in the past. Yet for most indicators of success, there is also a discernible effect of the mean degree of familiarity, and of its minimum.

Descriptively, the effect is however not monotonic. Peaks at a low degree of familiarity are not particularly relevant as they result from a fairly small number of cases (see Figure 2). Very high degrees of familiarity are more frequent (see again Figure 2). Hence descriptives suggest the possibility of a non-linear reaction to an increase in the degree of familiarity. This is what the regressions of Table 2 test, by adding polynomial terms.

Recall that the court only very rarely actually decides on the merits of a constitutional complaint. But if it does, complainants stand a fair chance to win (Table 1). The regressions in the first panel of Table 2 show that the maximum degree of familiarity is by far the strongest predictor of success on the merits. It is highly significant when explaining success with only the minimum and the maximum degree of familiarity on the bench. Only the linear term turns out significant. If one adds the mean degree of familiarity to the specification, the significance level of the maximum degree of familiarity drops to the 10% level. This is expected as the maximum contributes to the calculation of mean, so that both regressors partly capture the same effect. In the regressions explaining success on the merits with the minimum and the maximum, and polynomials thereof, there is also a significant negative effect of the square of the minimum. This shows that not only the maximum matters, so does the spread. If a new justice has been added to the chamber of two justices with a long history of joint decision-making, the two justices staying on the bench have less of an impact on the outcome.

Results essentially stay the same if one also considers the complaint as a failure if it has been summarily rejected (second panel of Table 2). This is important as a robustness check as, then, the number of observations jumps from little more than 200 to almost 4000.

On success with a request for a preliminary injunction, the mean degree of familiarity has a straightforward positive effect. It shows up in the third panel of Table 2 when only considering this explanatory variable, but also when controlling for the minimum and the maximum degree of familiarity. The square and cubic terms also have a significant positive effect, indicating that the success rate not only increases linearly in the mean degree of familiarity, but exponentially. When only using the minimum and the maximum degree of familiarity for explanation, there is also a significant and sizeable effect of the minimum degree and its square, again indicating a non-linear relationship. By contrast the maximum degree of familiarity only positively explains success in a preliminary procedure in the square and the cubic terms. Consequently for preliminary injunctions it is more important that the lowest degree of familiarity is not too low. A preliminary injunction stands a lower chance to be successful after a new justice has joined a chamber. The maximum degree of familiarity, i.e. the justice with the strongest familiarity with her peers, only matters if this degree is very high. This suggests that the most important feature of the bench is the spread in familiarity between the justice with the lowest and the justice with the highest degree of familiarity. When adding the mean degree of familiarity, this measure turns out to be most important. With the exception of the cubic term of the maximum degree of familiarity, all effects of the minimum and the maximum degree turn out at most weakly significant.

Ultimately, complainants of course care about the merits of the case. Procedure is only a way of getting their cases heard, and of improving the probability of success on the merits. Still it is interesting to consider the association between the familiarity of the justices on the bench and success in procedural matters. As the penultimate panel of Table 2 shows, on this success variable the mean degree of familiarity has a straightforward positive effect. It shows up in all specifications. It is not qualified by either the square or the cubic terms. When exclusively considering the minimum and the maximum degree of familiarity, there is also a consistent and strong effect of the maximum, as well as of the square of the maximum. The probability of

success in procedural matters increases exponentially in the maximum degree of familiarity of the bench. There is also a much smaller, and only weakly significant, effect of the minimum degree of familiarity. When adding all three measures to the specification, except for the square of the maximum degree of familiarity, all effects of the minimum and the maximum vanish. For this dependent variable, mean familiarity tells the story.

As success in the constitutional court is multifaceted, the last panel of Table 2 considers whether a complainant has been either successful on the merits, with a request for a preliminary injunction, or at least in procedural matters. This composite dependent variable is again straightforwardly explained by the mean degree of familiarity on the bench. If one exclusively considers the minimum and the maximum degree of familiarity, both significantly and positively explain success, as does the square of the maximum degree of familiarity. Hence the probability of success is again exponentially growing in the maximum degree of familiarity. When adding the mean degree of familiarity, the only remaining effect that is significant at conventional levels is the square of the maximum degree of familiarity. Essentially success in this composite manner is explained by the mean degree of familiarity.

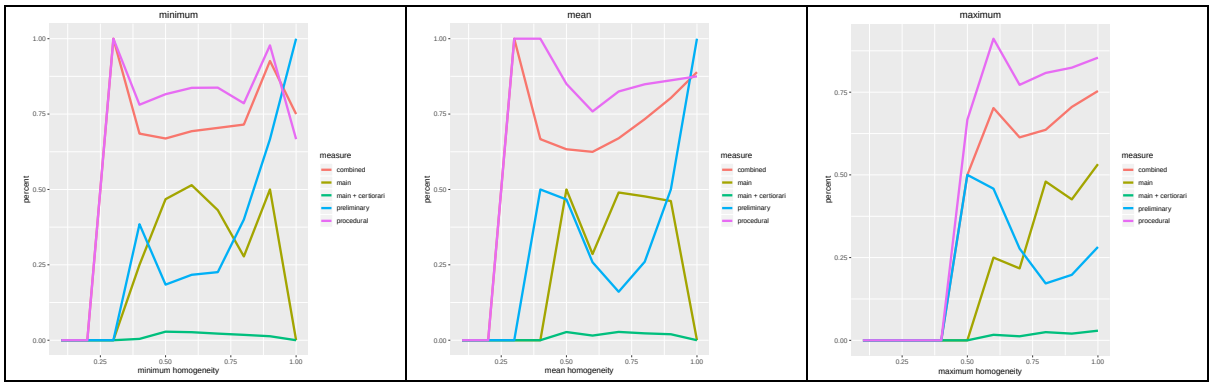


Figure 3
Association between Familiarity of Justices on the Bench and Success in Court Constitutional Complaints Heard by a Chamber

merit									
	model 1	model 2	model 3	model 4	model 5	model 6	model 7	model 8	model 9
mean	.333 (.431)	.386 (.496)	.386 (.496)				-.692 (1.450)	-.511 (1.720)	-.549 (1.860)
$mean^2$		-1.020* (.496)	-1.020* (.496)					-.363 (1.060)	-.569 (1.240)
$mean^3$			-.224 (.496)						-1.040 (1.160)
min				-.543 (.350)	-.663 (.495)	-.729 (.495)	-.158 (.880)	-.191 (1.380)	-.222 (1.520)
min^2					-1.270* (.538)	-1.270* (.536)		-1.020 (.887)	-.468 (1.250)
min^3						.118 (.498)			.834 (.948)
max				.976** (.320)	1.870*** (.513)	1.890*** (.516)	1.290^{+} (.732)	2.000^{+} (1.150)	1.991^{+} (1.170)
max^2					.838 (.521)	.816 (.522)		.962 (.646)	.563 (.713)
max^3						.806 (.493)			1.001^{+} (.555)
cons	.226 (.295)	.452*** (.034)	.452*** (.034)	-.060 (.313)	.452*** (.034)	.452*** (.033)	-.066 (.314)	.452*** (.034)	.452*** (.034)
adj. R^2	-.002	.014	.010	.040	.060	.065	.037	.052	.054
N	208	208	208	208	208	208	208	208	208

merit + certiorari									
	model 1	model 2	model 3	model 4	model 5	model 6	model 7	model 8	model 9
mean	-.006 (.029)	-.029 (.152)	-.029 (.152)				-.149 (.093)	-.633 (.515)	-.554 (.522)
$mean^2$		-.235 (.152)	-.235 (.152)					-.041 (.280)	.024 (.321)
$mean^3$			-.035 (.152)						.095 (.240)
min				-.030 (.023)	-.180 (.156)	-.184 (.434)	.053 (.056)	.275 (.423)	.194 (.234)
min^2					-.317* (.155)	-.311* (.157)		-.259 (.237)	-.339 (.334)
min^3						.151 (.154)			.078 (.196)
max				.047* (.023)	.369* (.155)	.366* (.155)	.112* (.046)	.702* (.314)	.693* (.325)
max^2					.116 (.156)	.091 (.158)		.095 (.192)	.076 (.196)
max^3						.024 (.154)			.003 (.169)
cons	.027 (.020)	.024*** (.002)	.024*** (.002)	.001 (.022)	.024*** (.002)	.024*** (.002)	.003 (.022)	.024*** (.002)	.024*** (.002)
adj.R^2	-.0002	.0001	-.0001	.001	.001	.001	.001	.001	.001
N	3,996	3,996	3,996	3,996	3,996	3,996	3,996	3,996	3,996

preliminary									
	model 1	model 2	model 3	model 4	model 5	model 6	model 7	model 8	model 9
mean	.349* (.173)	.856* (.413)	.856* (.411)				1.020^{+} (.573)	3.720** (1.420)	3.120* (1.420)
mean^2		2.730*** (.413)	2.730*** (.411)					1.930** (.709)	3.170*** (.808)
mean^3			1.130* (.411)						1.600* (.657)
min				.341 (.150)	1.590*** (.435)	1.690*** (.436)	-.199 (.338)	-1.330 (1.050)	-1.330 (1.050)
min^2					1.420*** (.428)	1.530*** (.429)		.354 (.576)	-.962 (.774)
min^3						-.264 (.420)			-.945^{+} (.512)
max				-.030 (.130)	-.415 (.424)	-.430 (.423)	-.511^{+} (.300)	-2.350* (.968)	-1.370 (1.010)
max^2					2.090*** (.440)	2.130*** (.443)		1.220* (.548)	.932^{+} (.554)
max^3						1.140** (.419)			.966* (.439)
cons	.0003 (.117)	.234*** (.014)	.234*** (.014)	.069 (.124)	.234*** (.014)	.234*** (.014)	.075 (.124)	.234*** (.014)	.234*** (.014)
adj.R^2	.004	.052	.059	.004	.050	.056	.006	.060	.073
N	841	841	841	841	841	841	841	841	841

procedural									
	model 1	model 2	model 3	model 4	model 5	model 6	model 7	model 8	model 9
mean	.340*** (.100)	1.290*** (.377)	1.290*** (.377)				.712* (.310)	3.080* (1.230)	2.970* (1.250)
mean^2		.326 (.377)	.326 (.377)					.195 (.690)	-.298 (.836)
mean^3			-.435 (.377)						-.860 (.635)
min				.088 (.079)	.668^{+} (.387)	.684^{+} (.388)	-.280 (.178)	-1.410 (.979)	-1.110 (1.020)
min^2					-.120 (.390)	-.085 (.392)		-.399 (.580)	.312 (.799)
min^3						.151 (.380)			.694 (.511)
max				.206** (.077)	1.020** (.387)	1.010** (.387)	-.126 (.163)	-.765 (.812)	-1.030 (.833)
max^2					1.220** (.390)	1.200** (.393)		1.210* (.478)	1.130* (.483)
max^3						.344 (.380)			.527 (.410)
cons	.595*** (.069)	.828*** (.008)	.828*** (.008)	.606*** (.074)	.828*** (.008)	.828*** (.008)	.602*** (.074)	.828*** (.008)	.828*** (.008)
adj.R^2	.005	.005	.005	.003	.007	.007	.005	.009	.009
N	2,133	2,133	2,133	2,133	2,133	2,133	2,133	2,133	2,133

merit, preliminary or procedural									
	model 1	model 2	model 3	model 4	model 5	model 6	model 7	model 8	model 9
mean	.664*** (.106)	2.860*** (.457)	2.860*** (.457)				.970** (.339)	4.600** (1.520)	4.760** (1.530)
$mean^2$		$.814^{+}$ (.457)	$.814^{+}$ (.457)					.476 (.845)	.958 (1.000)
$mean^3$			.083 (.457)						.401 (.767)
min				.182* (.086)	1.330** (.472)	1.370** (.473)	$-.325^{+}$ (.196)	-1.810 (1.190)	-2.070^{+} (1.240)
min^2					.433 (.474)	.491 (.476)		-.037 (.697)	-.446 (.955)
min^3						.124 (.462)			.022 (.603)
max				.440*** (.082)	2.360*** (.469)	2.350*** (.469)	-.010 (.177)	-.246 (1.000)	-.122 (1.030)
max^2					1.590*** (.477)	1.570*** (.480)		1.500* (.595)	1.380* (.602)
max^3						.560 (.462)			.494 (.504)
cons	.242*** (.073)	.695*** (.009)	.695*** (.009)	.227** (.078)	.695*** (.009)	.695*** (.009)	.222** (.078)	.695*** (.009)	.695*** (.009)
adj.R^2	.014	.015	.014	.013	.017	.017	.016	.020	.019
N	2,707	2,707	2,707	2,707	2,707	2,707	2,707	2,707	2,707

Table 2
Association between Familiarity of Justices on the Bench and Success in Court Constitutional Complaints Heard by a Chamber
linear probability models
mean: mean degree of homogeneity on the bench
min: minimum degree of homogeneity
max: maximum degree of homogeneity
standard errors in parenthesis
*** p < .001, ** p < .01, * p < .05, $^{+}$ p < .1

The story is therefore nuanced. It can be summarised as follows:

Result 1:

a) The success of a constitutional complaint on the merits, in a request for a preliminary injunction, and in procedural matters, is associated with the degree of familiarity among the justices in the chamber.
b) Success on the merits and in procedural matters is associated with the maximum degree of familiarity.
c) Success in a preliminary injunction is associated with the minimum degree of familiarity.

b) Causal Effect Near Changes in Chamber Composition

Potential Endogeneity This result points to a normative concern: one would not want the success of a constitutional complaint to be predicted by a variable beyond the control of the complainant. Yet as always, correlation does not prove causation. Actually, the mean degree of

homogeneity (in constitutional complaints heard by a chamber) is significantly and positively correlated with the fraction of female justices on the bench (r = .13***) and with their mean age (r = .37***), and it is negatively correlated with the fraction of justices selected by one of the conservative parties CDU or FDP (r = -.13***). One of these demographic variables might drive the effect. Correlation with unobserved variables can of course not be ruled out either. As reporting is selective, there may be measurement error. At least in the long run, there could even be reverse causality. In the periodic recomposition of chambers, the respective senate might take the expected effect on the success of constitutional complaints into account.

Regression Discontinuity Yet the data provides an opportunity to test whether the effect of homogeneity on the success of a constitutional complaint is indeed causal, at least for a more narrow set of observations. The left panel of Figure 4 shows why and how the familiarity score of each justice varies over time, using Justice Baer for illustration. When she joined the court in 2011, for about two years she was in a chamber with Justices Ferdinand Kirchhof and Schluckebier. After Justice Schluckebier left the court, for four years she was in a chamber still with Justice Ferdinand Kirchhof, but also with Justice Masing. Thereafter, she was together with two other Justices, Britz and Eichberger. For a few months, she was together with Justices Britz and Radtke, and finally is in a chamber with Justices Harbarth and Ott. The right panel of Figure 4 demonstrates in which ways these periodic recompositions of the chamber affect her familiarity score. For chamber decisions, recomposition typically leads to a sudden drop in familiarity.

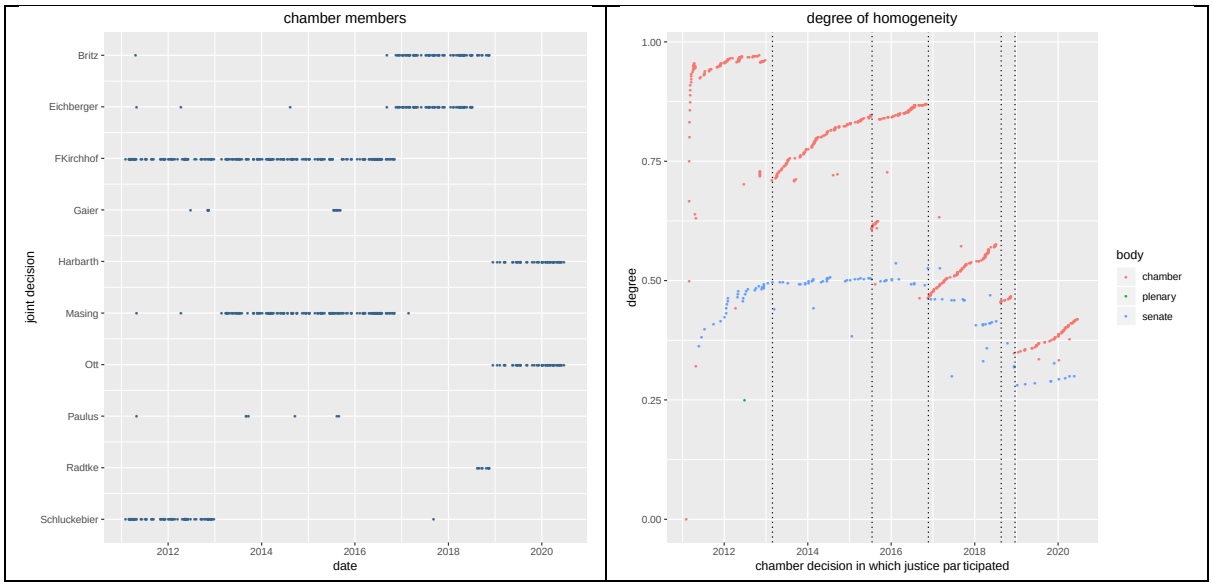


Figure 4
Periodic Breaks in Homogeneity: Justice Baer

For the individual justice, this drop in familiarity can be regarded as (nearly) exogenous. It is true that chamber composition is decided by the senate, and that the individual justice takes part in the decision-making process. Yet the justice has at best an influence on chamber composition. As explained above, even this influence is very limited. The primordial concern is ideological balance. To the extent feasible the senates also aim at balancing tenure and gender.

The main driving force for recomposition is the fixed 12 year term, which is beyond any justice's control, as is the identity of their newly appointed colleague. Most importantly, no justice can prevent recomposition from happening, and she has no influence on the point in time when recomposition takes place. She has to come to terms with the fact that, for a considerable amount of time, she will have to closely collaborate with new colleagues.

The fact that recomposition is beyond the control of individual justices makes it possible to implement a regression discontinuity design (for background see Lee and Lemieux 2010, Lee and Lemieux 2014). One can treat recomposition as an exogenous shock which allows to causally estimate a local treatment effect. If one finds an effect on the probability of success in the neighbourhood of this breakpoint, it must have been caused by the break.

The right panel of Figure 4 shows that, in the familiarity scores of Justice Baer, there are five clearly discernible breakpoints (marked with dotted vertical lines). The same procedure can be used for all justices. Results are displayed Appendix A1. As this is where the discontinuity in familiarity originates, and as the analysis in this paper is confined to chamber decisions (on constitutional complaints), breaks are defined with respect to chambers, not with respect to senate or plenary decisions. Four justices (Grasshof, Kruis, Seibert, Seidl) have left the court so early that no break in their familiarity score can be discerned. Justice Langenfeld has joined the court so recently that no break in familiarity has occurred. For the remaining 42 justices, at least one break point can be found in the data. For most of them, during the period of observation more than one break can be observed.

There is a total of 135 breakpoints in the dataset. Now these breakpoints did not occur for all justices at the same point in time. This is why, for regression discontinuity, the data is normalised. The analysis uses chamber decisions from a window starting 10 decisions before the break up till 10 decisions thereafter. In implementing this normalisation, only decisions of the chambers are used of which the justice who experienced the break has been a member, before and after the break.

Regression discontinuity exploits the fact that, in the neighbourhood of a breakpoint, the exogenous change in the composition of the chamber is the dominant influence. If there is a significant difference between observations before and after the breakpoint, one may confidently infer that the difference has been caused by the break. Now the composition of the chamber is not the only factor that varies over time. So do the parties, the areas of law, or the political salience of the case, to only list a few of those factors. This is why the window should not be too narrow. One would not see any effect, not because the effect has not been present, but because this variance among cases makes the data too noisy. On the other hand, the window should not be too wide, as one then would have to be concerned that an observed change in outcomes is caused by other systematic changes over time that have occurred during the relevant period of time.[19]

19 Note, however, that the regressions work with multiple windows, at different points in calendar time. This feature of the dataset makes it less likely in the first place that alternative, longer-term influences are systematic *with respect to the breakpoints*.

43 of the 135 breakpoints only affect a single justice, 40 affect two of them simultaneously, and 4 even affect three justices at a time. One may therefore worry that a dataset covering windows around all 135 breakpoints is misleading as 40 windows feature twice in the dataset, and 4 even three times. One may, however, object that a break that affects more than one justice simultaneously should also carry more weight in estimating the local effect. In the following, results from using data all 135 breakpoints are therefore presented side-by-side with results from a narrower dataset that removes duplicates, so that only data around 87 unique breakpoints is used. Figure 5 shows that descriptives look very similar whether one uses the wider or the narrower dataset. As the latter uses about a third less data, some of the confidence intervals are a bit wider though.

Local Effect Descriptively, the drop in the success probability after the break is most pronounced for procedural decisions, and for the success measure that combines the decision on the merits, preliminary injunctions, and procedural decisions.[20] Table 3 reports statistical tests. As coefficients then have the intuitive interpretation as the marginal effect of the break on the probability of success, linear probability models are estimated.

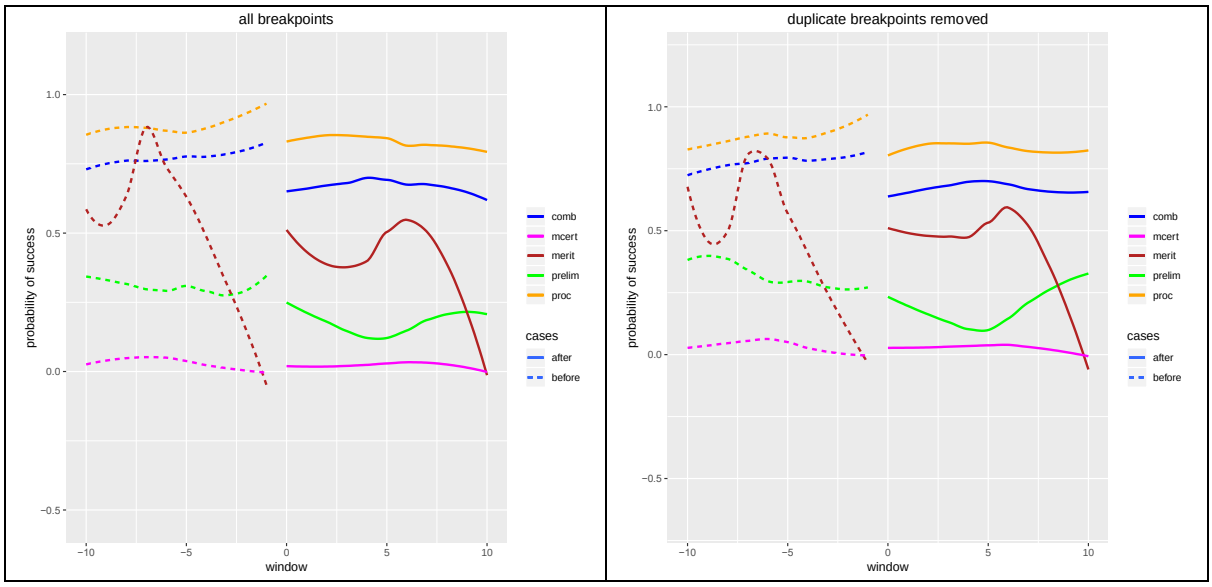


Figure 5
Regression Discontinuity

If one keeps duplicates in the dataset, one not only weighs effects by the number of justices affected by the breakpoint. One must also take into account that the duplicate observations are not independent. This correction is performed with the help of a justice random effect.

20 The somewhat unruly shape of the success on the merits results from the fact that the court only very rarely decides on the merits. Hence these lines are the mean of very few observations, which is why the particularities of the individual cases are not averaged out.

Occasionally, the Hausman test turns out significant. In these specifications, the justice random effect is replaced by a justice fixed effect.[21] Dependence also results from the fact that, on either side of the breakpoint, there are multiple decisions taken by the same chamber. This is why the models removing duplicates are also estimated with justice random effects (or justice fixed effects if the Hausman test turns out significant).

There are two options for estimating the local effect of chamber recomposition. One may estimate the effect of a dummy that is 1 for decisions taken by the newly composed chamber. Alternatively one may treat the distance from the breakpoint as a continuous variable. This specification allows the effect to unfold over time. As Table 3 shows, which specification one chooses matters for the effect of chamber recomposition on success on the merits. There is no effect (neither with the wider nor with the narrower dataset) if one explains decisions with the dummy variable. There is, however, a sizeable and significant negative effect on success on the merits if one uses the continuous explanatory variable. The same holds for the effect of recomposition on a variable that also codes it as failure if the constitutional court has treated the case as inadmissible or obviously unfounded.

For the remaining three dependent variables, the choice of the explanatory variable is not critical. One finds a significant negative effect of recomposition on requests for a preliminary injunction, procedural decisions, and on a combined measure for success on the merits, in a preliminary injunction, or in procedural matters. With one exception, for these three dependent variables it does also not matter whether one removes duplicate observations from the dataset or not. The only exception is the effect of the continuous variable on success in procedural matters. If one keeps duplicates, one finds a significant effect, while the effect turns insignificant in the smaller dataset.

merit	all breakpoints		duplicates removed		
	model 1	model 2	model 3	model 4	model 5
after	-.1779^{+} (.0951)		-.1868 (.1170)		-.1126 (.1104)
window		-.0190** (.0069)		-.0234** (.0086)	
paragraphs					.0098*** (.0027)
cons	.5603*** (.0806)	.4671*** (.0622)	.5548*** (.0954)	.4489*** (.0733)	.2757* (.1184)
N	103	103	73	73	73
merit + certiorari					
after	-.0099 (.0071)		-.0063 (.0094)		-.0034 (.0093)
window		-.0013* (.00057)		-.0013^{+} (.00076)	
paragraphs					.0014***

21 As fixed effects estimation works with demeaning, these regression have no constant – which is also how they can be recognized in Table 3.

					(.0003)
cons			.0318*** (.0083)	.0285*** (.0066)	.0038 (.0097)
N	1921	1921	1220	1220	1220
preliminary					
after	-.1185** (.0416)		-.1287* (.0522)		-.1015^{+} (.0545)
window		-.0093** (.0035)		-.0087* (.0044)	
paragraphs					-.0044** (.0015)
cons	.3224*** (.0372)	.2546*** (.0278)	.3413*** (.0469)	.2681*** (.0356)	.4072*** (.0523)
N	414	414	270	270	270
procedural					
after	-.0603** (.0208)		-.0630* (.0258)		-.0478* (.0237)
window		-.0040* (.0017)		-.0035 (.0022)	
paragraphs					.0089*** (.0008)
cons	.8829*** (.0215)	.8512*** (.0188)			
N	1056	1056	690	690	690
merit, preliminary or procedural					
after	-.1061*** (.0242)		-.1115*** (.0298)		-.0938*** (.0273)
window		-.0070*** (.0020)		-.0072** (.0025)	
paragraphs					.0112*** (.0009)
cons	.7705*** (.0232)	.7142*** (.0195)	.7760*** (.0267)	.7174*** (.0220)	.5134*** (.0328)
N	1336	1336	870	870	870

Table 3
Regression Discontinuity

10 decisions before and after the breakpoint of chamber in which Justice participated who experienced break-point
Models 1-2: data around all 135 breakpoints
Models 3-4: if two or three Justices had the same breakpoint, only one of them is kept (87 breakpoints)
Linear Probability Models
Justice random effects
If Hausman test turns out significant, corresponding model with Justice fixed effects reported
(in this case, constant drops out by demeaning)
after: dummy variable that is 1 for decisions taken at or after breakpoint
window: continuous variable ranging from -10 .. 10
paragraphs: the number of paragraphs posted on the court's website
standard errors in parenthesis
*** p < .001, ** p < .01, * p < .05, $^{+}$ p < .1

With these qualifications, the following causal effects can be established:

Result 2:
The recomposition of a chamber of the German Constitutional Court reduces the probability that a constitutional complaint

a) is successful on the merits,
b) is not rejected, either on the merits, or for being obviously unfounded, or for being inadmissible,
c) leads to a preliminary injunction in favour of the complainant
d) leads to a procedural decision in favour of the complainant,
e) is either successful on the merits, leads to a preliminary injunction, or to a procedural decision in favour of the complainant.

Robustness Recomposition is beyond the control of individual justices. But they can see it coming. Hence the old chamber may try to complete proceedings before recomposition. The new chamber might first decide cases that are easier to settle. Hence the local effect might result from strategic reactions of chambers to the exogenous interference with chamber composition. Note, however, that this explanation would not invalidate the normative concern. It would just change the channel. The privilege would result from the fact that the previously composed chamber seizes the opportunity to decide in favour of the complainant before it can no longer do so. Moreover recall that, descriptively, there are also pronounced global effects (see Figure 3 and Table 2). Even if strategic choices near the breakpoint had a local effect, this could not explain the global effects.

Actually, in the data there is a proxy. Recall that, when deciding about constitutional complaints, the court has power to decide without giving reasons. Also only a small fraction of the decisions is posted on the website. Hence it is meaningful that the court justifies its decision, and makes this justification public. Arguably, the more elaborate these reasons, the more the court was struggling with a normative issue. Now recall that a very large majority of constitutional complaints fails. Hence the less the outcome is obvious, the higher the chances for (at least partial) success. The number of paragraphs spent on the respective decision is available in the dataset. If the local effect is chiefly driven by strategic moves, the local effect should disappear when controlling for this proxy.

Model 5 of Table 3 shows that, indeed, for all dependent variables the number of paragraphs has a significant effect on success. The effect is positive, except for the success of a request for a preliminary injunction. Yet whenever the break had a significant effect without controlling for this proxy (Model 3) it also has a significant effect with the control variable included. The only qualification concerns preliminary injunctions. While the negative effect is significant at the 5% level without the control, the p-value goes up to .0637 when controlling for the number of paragraphs.[22] Hence one cannot rule out strategic moves of the court around the break-

22 In the corresponding random effects model, the coefficient is -.1057, p = .0434.

points. But even when taking this possibility into account, there is a significant residual negative effect. Familiarity matters as such, not only through the strategic reactions of chambers to recomposition.

As explained above, when choosing the width of the window, one must strike a balance between local noise (that is removed by using a wider window) and unobserved intervening variables (that are the more likely the wider the window). As explained, a window covering 10 decisions (in which this justice has participated) before and after the break seems to be well balanced. Yet as a further robustness test, Appendix A2 collects information about two wider definitions of the window: 20, or 30 decisions before and after the break.

Comparing Figure 5 with Figure 7, one sees a few descriptive effects: In the procedural and combined success rates, there is a negative dip about 10 decisions before the break, and in preliminary injunctions there is a downward dip. But the 30 decision window shows that these are fluctuations, not longterm trends. Success on the merits is very volatile, most likely due to the very small number of observations.

Comparing Table 3 with Table 4, it becomes apparent that results remain similar, also when using a wider window. With window 20, the effect of the break on the success rate with preliminary injunctions is insignificant, while it was significant with window 10. However, for success in procedural matters and the combined success measure the break has a significantly negative effect with either width of the window. There is now even a weakly significant effect ($p = .0728$) of the break on success on the merits, that was absent with window 10. With window 30, the effect of the break on success in procedural matters is insignificant, but there is a weakly significant effect ($p = .0816$) of success with preliminary injunctions. Hence, unsurprisingly, the width of the window matters. Not all effects replicate. But a window of width 10 does not seem to be the exception that proves the rule of no effect. Checking alternative windows increases the confidence in the causal effect of familiarity on the outcome of the case.

5. Discussion

Not every complainant has a valid case. Not every request for a preliminary injunction is well founded. Not every procedural plea has substance. Per se, the fact that a complaint was not successful on the merits, has not led to a preliminary injunction, or that the court has rejected a procedural request, is no reason for concern. But constitutional complaints should fail because they are unfounded, not because the complainant had bad luck. Which chamber hears her case is beyond the control of the complainant. If the court has not summarily rejected the case, the probability to win on the merits is about 18% higher if the case is decided before chamber recomposition. A preliminary injunction is about 12% more likely to be granted. A procedural request is about 6% more likely to be successful. There is no normative justification for the odds of success to depend on a measure that, at its face, should only help the court shoulder a higher workload while maintaining impartiality.

These remarks have a statistical correlate. The observation that a complainant has failed to succeed in an individual case is almost meaningless. To make it meaningful, one would need an objective measure for the normatively desirable probability of success, which is not available. Yet over the totality of 5848 cases in the complete dataset, over the totality of 10 cases each before and after 135 breakpoints, and over the totality of 10 cases each before and after 87 unique breakpoints, relying on the central limit theorem, variance in the objective probability of success should balance out. What the analysis of the complete data documents, and what the analysis near the breakpoints identifies, is the marginal effect of (a drop in) familiarity on outcomes.

The main limitation of the present analysis results from the reporting practice of the court. While online many more cases are reported than in print, the reported cases are still only a fraction of all decisions. Ultimately, only complete data could prove that the results reported in this paper do not result from selection. But selection would require that before recomposition systematically more successful cases are reported, and after recomposition systematically more failed cases. There is no plausible motive for such a bias in reporting practice. Moreover the local effect near recomposition remains significant when controlling for the number of paragraphs. If the effect resulted from selection, it should translate into the degree of elaboration in the individual decision.

With the present data from the German Constitutional Court one cannot isolate one of the channels, discussed in the introduction, on which a single member of a panel of three judges may influence outcomes. Yet a number of the mechanisms that can consistently explain panel effects as observed in the US Court of Appeal cannot explain the effects documented in this paper. This is not to say that these effects can be ruled out. The fact that the German Constitutional Court appears so unbiased ideologically might precisely result from the practice of almost always having a "counterjudge" on the bench. Gender balance within chambers might be important for the court being perceived as sensitive towards discrimination. Unanimity rule makes dissent very costly. The senate does not want to deal with multiple cases just because one justice cannot come to terms with the other members of her chamber. There is certainly a norm of consensus and conformity pressure. The court has always been concerned about its perceived legitimacy, which is considered the most important source of its considerable political power. Yet none of these mechanisms can explain why the odds of success are substantially lower after the recomposition of a chamber.

A plausible candidate for an explanation is, by contrast, collegiality. It has been argued that, everything else held constant, collegiality should be more pronounced the more often judges have interacted in the past (Hinkle, Nelson et al. 2020: 282), and the longer the prospect for their future interaction. Familiarity helps building trust, and makes individuals more open to being persuaded by a different reading of the facts, or to the normative question posed by the case (Edwards 1998, Edwards 2003). Greater familiarity may also heighten the sense of accountability (Sommers 2006). It may help individuals to build trust. They might also become more self-confident. Isolating these channels, and discriminating between them, is a promising topic for future research.

Panel effects have spawned a large literature. This paper does not only add the German case. The radically different organisation of panels in the German Constitutional Court casts light on a qualitatively different type of panel effects. In the court, for the most part cases are not randomly assigned to panels. Panels (chambers) keep their composition for an extended period of time, typically multiple years. This is why, mechanically, the degree of joint experience grows over time, and faces a sudden drop whenever chambers are recomposed. These recompositions are largely beyond the control of individual justices. In the temporal proximity of recomposition, one can therefore causally identify the effect of familiarity, defined as the degree of joint experience on the bench. The data shows that the effect of familiarity is substantial, and to the benefit of constitutional claimants. This effect cannot be explained by many of the mechanisms hypothesised in the literature interested in understanding the impact of minority judges. The German evidence provides additional support for the cognitive and motivational effect of collegiality.

References

ALARIE, BENJAMIN, ANDREW GREEN AND EDWARD M IACOBUCCI (2015). "Panel Selection on High Courts." University of Toronto Law Journal **65**(4): 335-381.

ATKINS, BURTON M (1973). "Judicial Behavior and Tendencies Towards Conformity in a Three Member Small Group. A Case Study of Dissent Behavior on the US Court of Appeals." Social Science Quarterly: 41-53.

ATKINS, BURTON M AND JUSTIN J GREEN (1976). "Consensus on the United States Courts of Appeals: Illusion or Reality?" American Journal of Political Science: 735-748.

BEIM, DEBORAH, ALEXANDER V HIRSCH AND JONATHAN P KASTELLEC (2016). "Signaling and Counter-Signaling in the Judicial Hierarchy: An Empirical Analysis of En Banc Review." American Journal of Political Science **60**(2): 490-508.

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.

COX, ADAM AND THOMAS J MILES (2008). "Judging the Voting Rights Act." Columbia Law Review **108**: 1-54.

CROSS, FRANK B AND EMERSON H TILLER (1997). "Judicial Partisanship and Obedience to Legal Doctrine. Whistleblowing on the Federal Courts of Appeals." Yale Law Journal **107**: 2155-2176.

EDWARDS, HARRY T (1998). "Collegiality and Decision Making on the DC Circuit." Virginia Law Review **84**: 1335-1370.

EDWARDS, HARRY T (2003). "The Effects of Collegiality on Judicial Decision Making." University of Pennsylvania Law Review **151**(5): 1639-1690.

ENGEL, CHRISTOPH (2020). Does Efficiency Trump Legality? The Case of the German Constitutional Court. Selection and Decision in Judicial Process Around the World. Empirical Inquiries. Y.-c. Chang. Cambridge, Cambridge University Press**:** 261-286.

ENGEL, CHRISTOPH AND LILIA ZHURAKHOVSKA (2017). "You Are In Charge. Experimentally Testing the Motivating Power of Holding a Judicial Office." Journal of Legal Studies **46**: 1-50.

ENGST, BENJAMIN G, THOMAS GSCHWEND, NILS SCHAKS, SEBASTIAN STERNBERG AND CAROLINE WITTIG (2017). "Zum Einfluss der Parteinähe auf das Abstimmungsverhalten der Bundesverfassungsrichter–eine quantitative Untersuchung." JuristenZeitung **72**(17): 816-826.

ENGST, BENJAMIN G, THOMAS GSCHWEND AND SEBASTIAN STERNBERG (2020). "Die Besetzung des Bundesverfassungsgerichts. Ein Spiegelbild gesellschaftlicher Präferenzen?" Politische Vierteljahresschrift **61**(1): 39-60.

EPSTEIN, LEE, WILLIAM M LANDES AND RICHARD A POSNER (2011). "Why (and When) Judges Dissent. A Theoretical and Empirical Analysis." Journal of Legal Analysis **3**(1): 101-137.

FARHANG, SEAN AND GREGORY WAWRO (2004). "Institutional Dynamics on the US Court of Appeals. Minority Representation under Panel Decision Making." Journal of Law, Economics, and Organization **20**(2): 299-330.

FISCHMAN, JOSHUA B (2011). "Estimating Preferences of Circuit Judges. A Model of Consensus Voting." Journal of Law and Economics **54**(4): 781-809.

FISCHMAN, JOSHUA B (2013). "Interpreting Circuit Court Voting Patterns. A Social Interactions Cramework." Journal of Law, Economics, and Organization **31**(4): 808-842.

GROSSMAN, GUY, OREN GAZAL-AYAL, SAMUEL D PIMENTEL AND JEREMY M WEINSTEIN (2016). "Descriptive Representation and Judicial Outcomes in Multiethnic Societies." American Journal of Political Science **60**(1): 44-69.

HAIRE, SUSAN B, LAURA P MOYER AND SHAWN TREIER (2013). "Diversity, Deliberation, and Judicial Opinion Writing." Journal of Law and Courts **1**(2): 303-330.

HAUSEGGER, LORI AND STACIA HAYNIE (2003). "Judicial Decisionmaking and the Use of Panels in the Canadian Supreme Court and the South African Appellate Division." Law & Society Review **37**(3): 635-657.

HESSICK, F ANDREW AND JATHAN P MCLAUGHLIN (2013). "Judicial Logrolling." Florida Law Review **65**: 443-492.

HETTINGER, VIRGINIA A, STEFANIE A LINDQUIST AND WENDY L MARTINEK (2004). "Comparing Attitudinal and Strategic Accounts of Dissenting Behavior on the US Courts of Appeals." American Journal of Political Science **48**(1): 123-137.

HINKLE, RACHAEL K (2017). "Panel Effects and Opinion Crafting in the US Courts of Appeals." Journal of Law and Courts **5**(2): 313-336.

HINKLE, RACHAELK, MICHAEL J NELSON AND MORGAN LW HAZELTON (2020). "Deferring, Deliberating, or Dodging Review. Explaining Counterjudge Success in the US Courts of Appeals." Journal of Law and Courts **8**(2): 277-300.

KAPELIUK, DAPHNA (2012). "Collegial Games. Analyzing the Effect of Panel Composition on Outcome in Investment Arbitration." Review of Litigation **31**: 267-312.

KASTELLEC, JONATHAN P (2007). "Panel Composition and Judicial Compliance on the US Courts of Appeals." Journal of Law, Economics, & Organization **23**(2): 421-441.

KASTELLEC, JONATHAN P (2011a). "Hierarchical and Collegial Politics on the US Courts of Appeals." Journal of Politics **73**(2): 345-361.

KASTELLEC, JONATHAN P (2011b). "Panel Composition and Voting on the US Courts of Appeals Over Time." Political Research Quarterly **64**(2): 377-391.

KASTELLEC, JONATHAN P (2013). "Racial Diversity and Judicial Influence on Appellate Courts." American Journal of Political Science **57**(1): 167-183.

KASTELLEC, JONATHAN P (2020). "Race, Context, and Judging on the Courts of Appeals: Race-Based Panel Effects in Death Penalty Cases." Justice System Journal: 1-22.

KIM, PAULINE T (2008). "Deliberation and Strategy on the United States Courts of Appeals. An Empirical Exploration of Panel Effects." University of Pennsylvania Law Review **157**: 1319-1382.

LANG, ANDREJ (2020). "Der Verhältnismäßigkeitsgrundsatz in der Rechtsprechung des Bundesverfassungsgerichts. Eine rechtsempirische Untersuchung mit rechtsvergleichenden Perspektiven." Archiv des öffentlichen Rechts **145**: 75-132.

LEE, DAVID S AND THOMAS LEMIEUX (2010). "Regression Discontinuity Designs in Economics." Journal of Economic Literature **48**(2): 281-355.

LEE, DAVID S AND THOMAS LEMIEUX (2014). Regression Discontinuity Designs in Social Sciences. Regression Analysis and Causal Inference. H. Best and C. Wolf**:** 301-326.

MILES, THOMAS J AND CASS R SUNSTEIN (2006). "Do Judges Make Regulatory Policy? An Empirical Investigation of Chevron." University of Chicago Law Review: 823-881.

PERESIE, JENNIFER L (2004). "Female Judges Matter. Gender and Collegial Decisionmaking in the Federal Appellate Courts." Yale Law Journal **114**: 1759-1790.

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

SEGAL, JEFFREY ALLAN AND HAROLD J. SPAETH (2002). The Supreme Court and the Attitudinal Model Revisited. Cambridge, UK ; New York, Cambridge University Press.

SOMMERS, SAMUEL R (2006). "On Racial Diversity and Group Decision Making. identifying Multiple Effects of Racial Composition on Jury Deliberations." Journal of Personality and Social Psychology **90**(4): 597-612.

SONGER, DONALD R (1982). "Consensual and Nonconsensual Decisions in Unanimous Opinions of the United States Courts of Appeals." American Journal of Political Science **26**: 225-239.

SPITZER, MATTHEW AND ERIC TALLEY (2013). "Left, Right, and Center. Strategic Information Acquisition and Diversity in Judicial Panels." Journal of Law, Economics, & Organization **29**(3): 638-680.

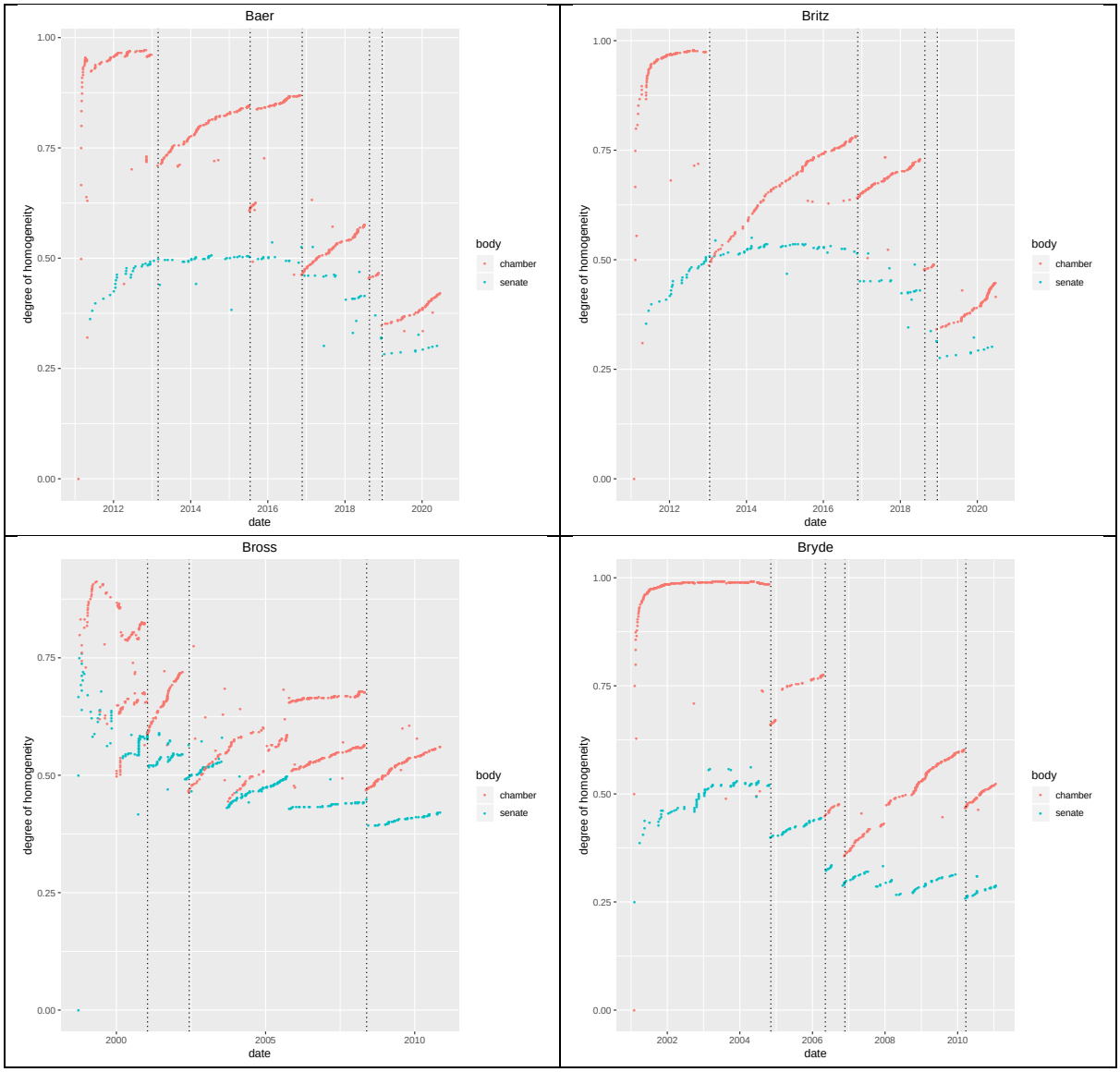
SUNSTEIN, CASS R, DAVID SCHKADE, LISA M ELLMAN AND ANDRES SAWICKI (2007). Are Judges Political? An Empirical Analysis of the Federal Judiciary, Brookings Institution Press.

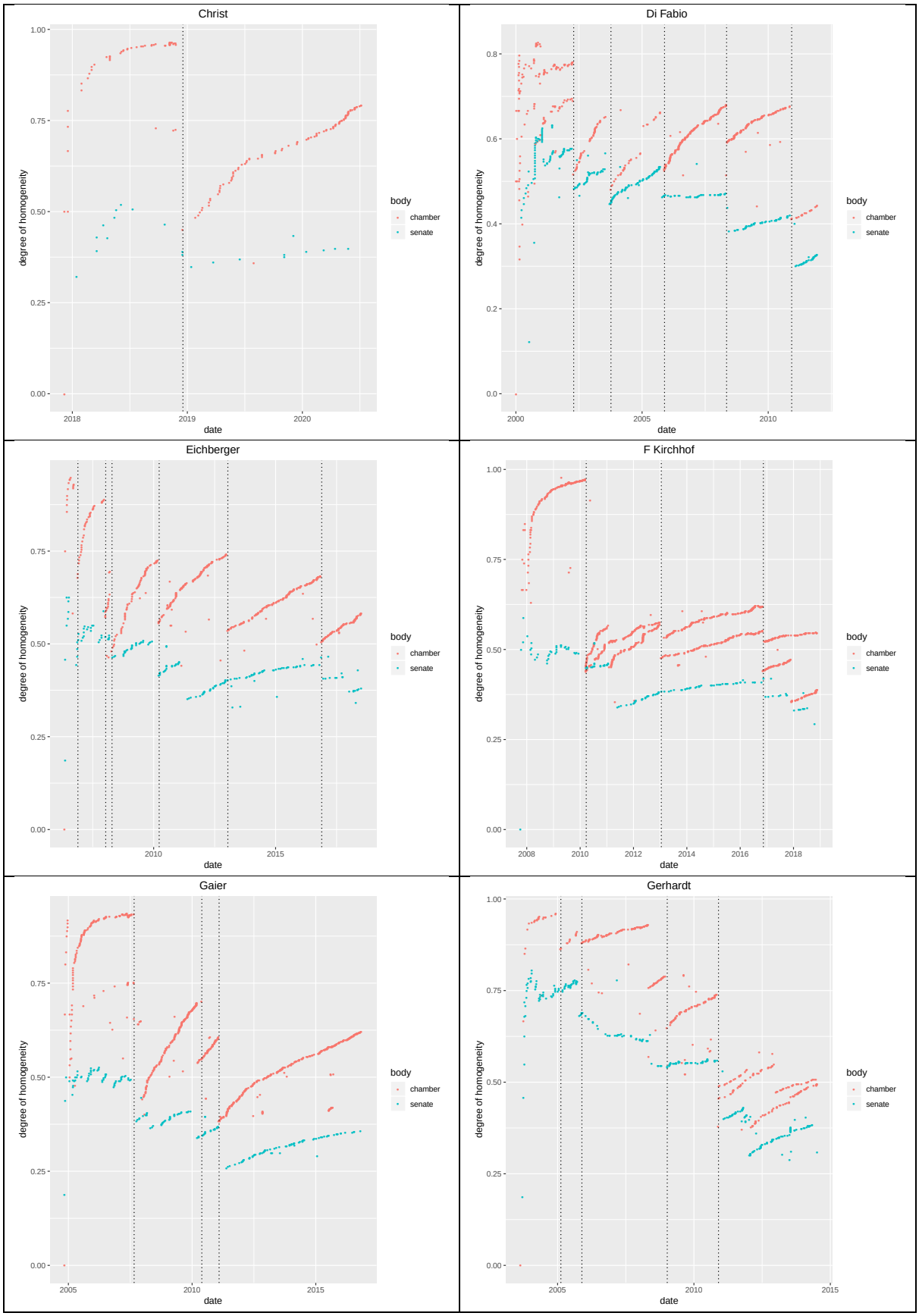
TILLER, EMERSON H (2015). "The Law and Positive Political Theory of Panel Effects." Journal of Legal Studies **44**(S1): S35-S58.

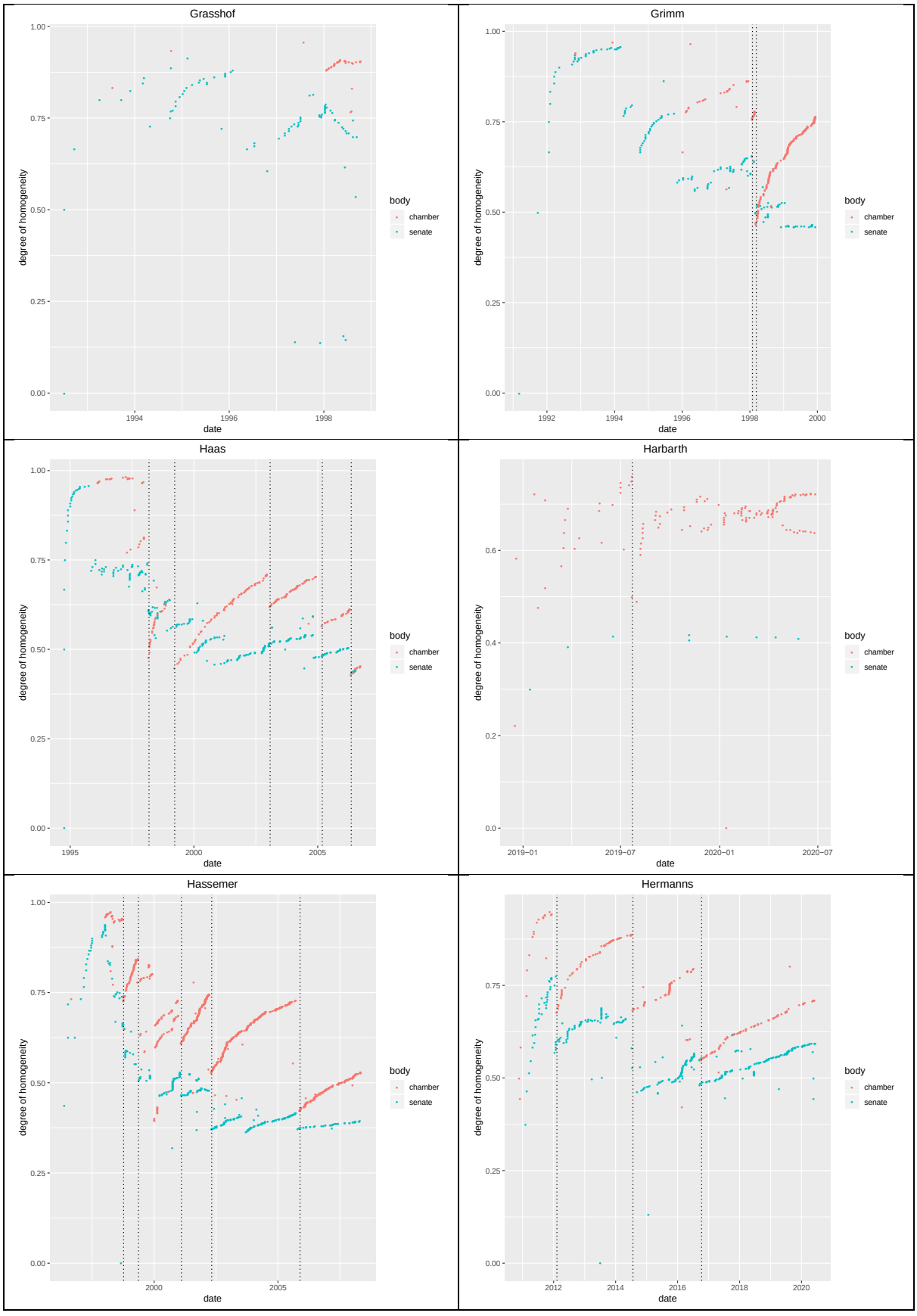
VANBERG, GEORG (2004). The Politics of Constitutional Review in Germany, Cambridge University Press.

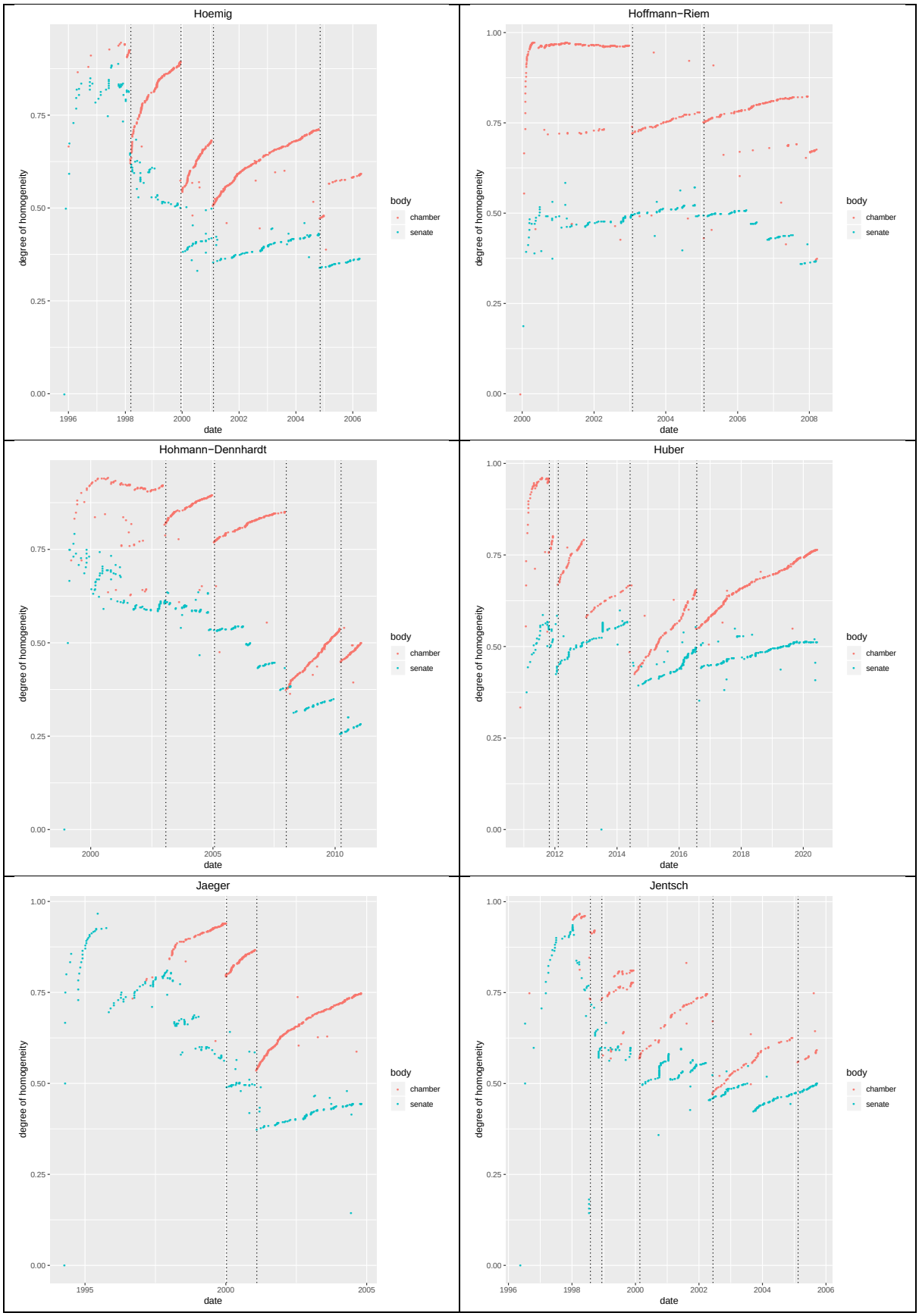
Appendix

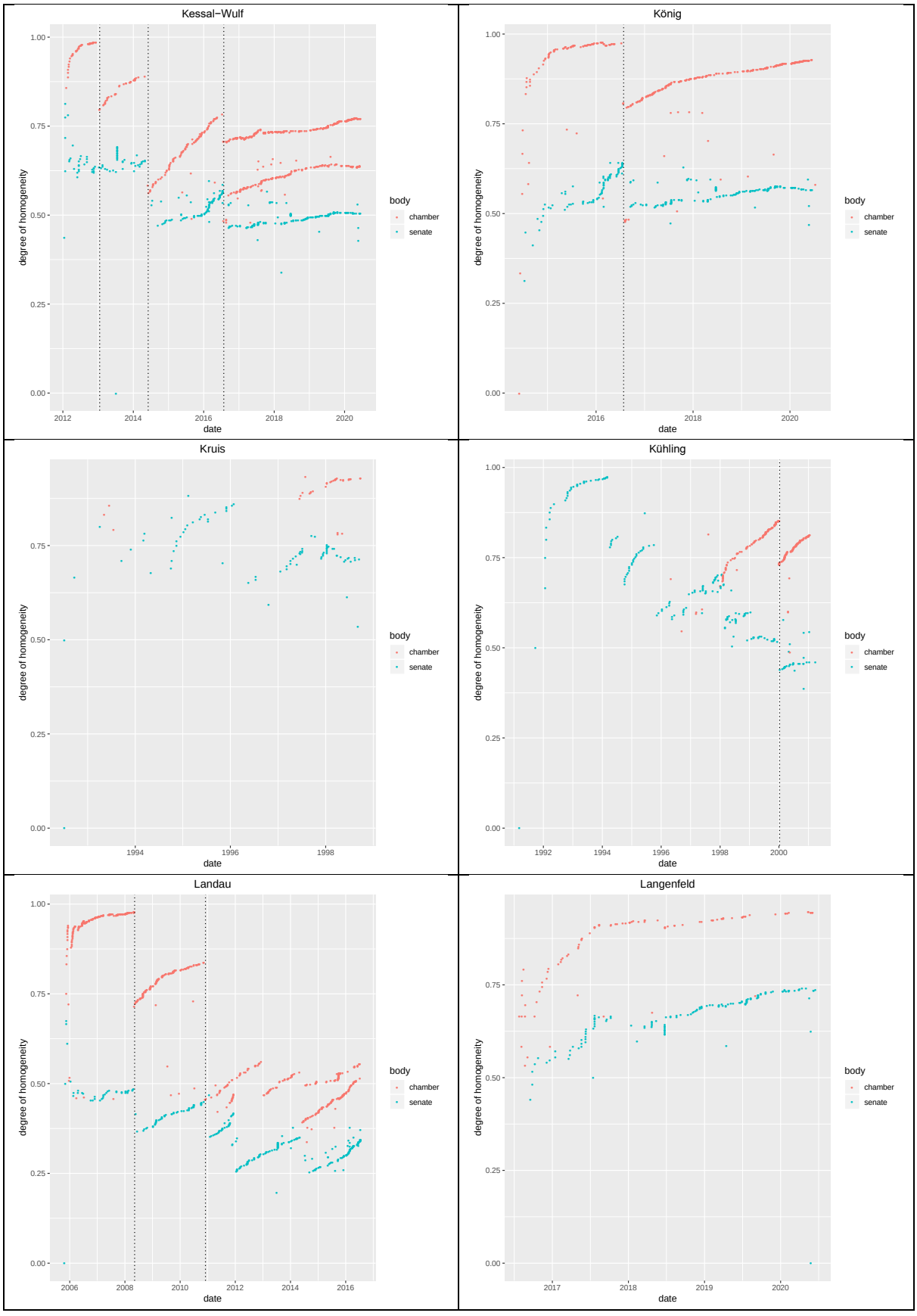
A1: Breakpoints

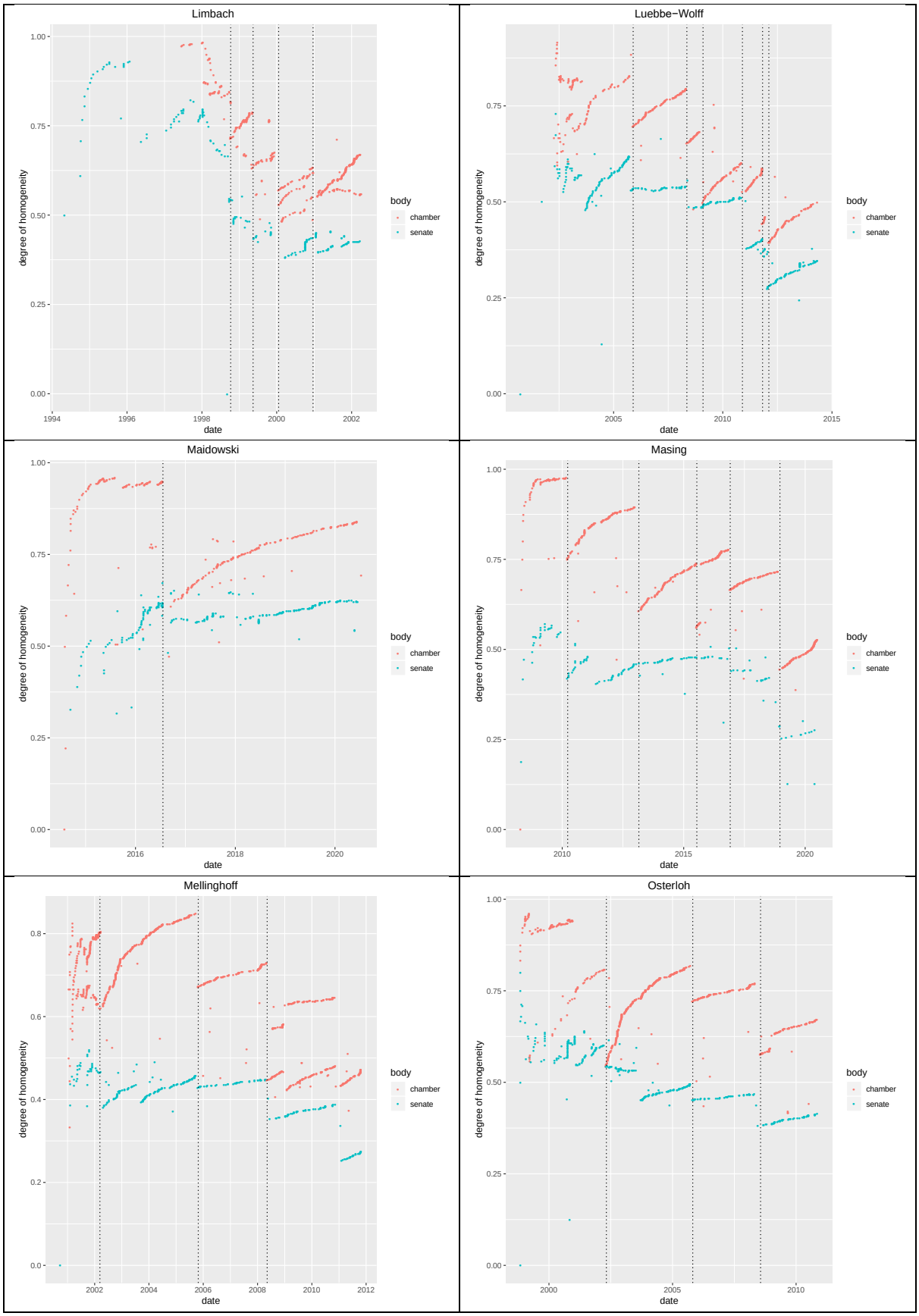


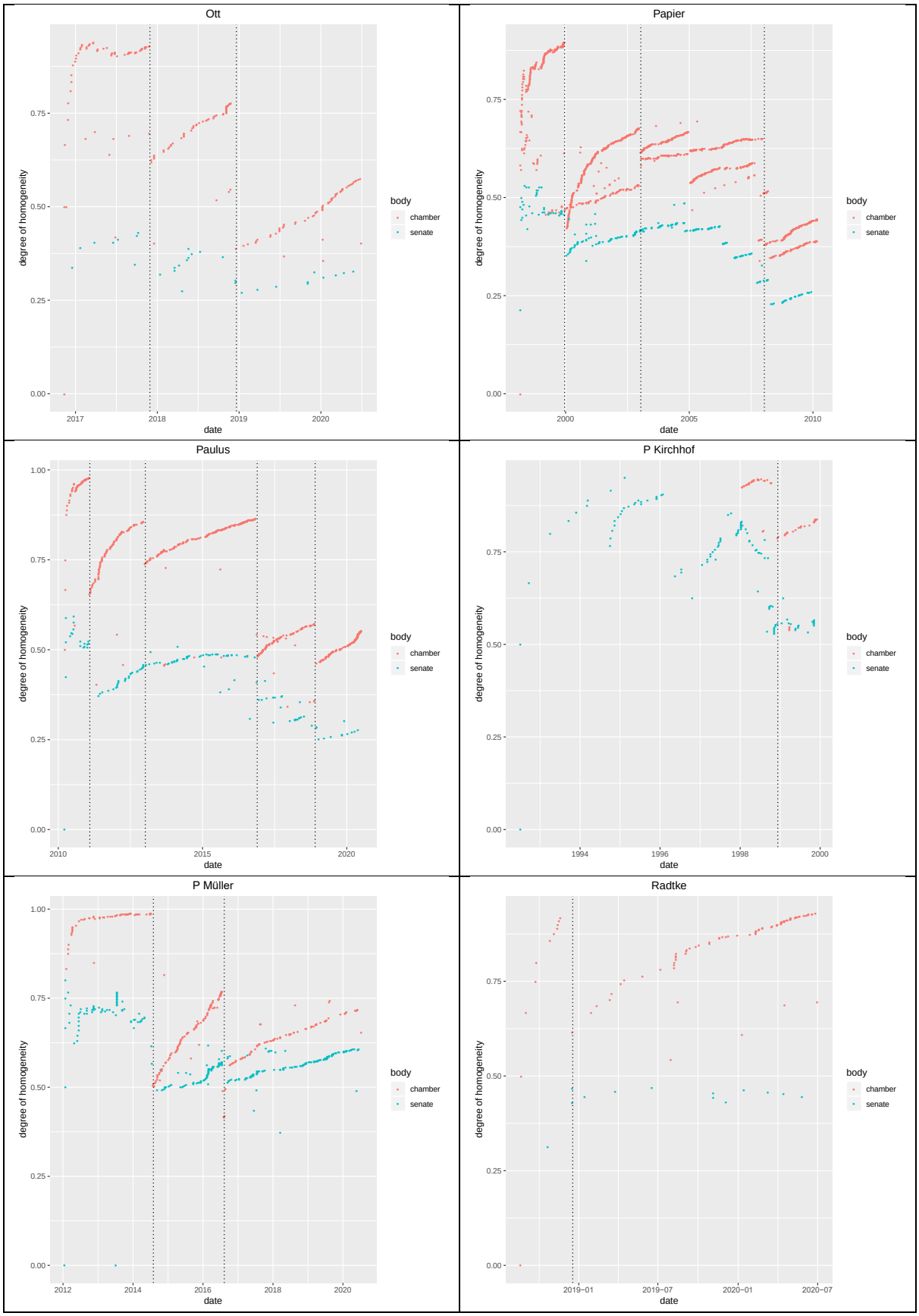

Christ
Di Fabio
Eichberger
F Kirchhof
Gaier
Gerhardt
degree of homogeneity
date
body
chamber
senate

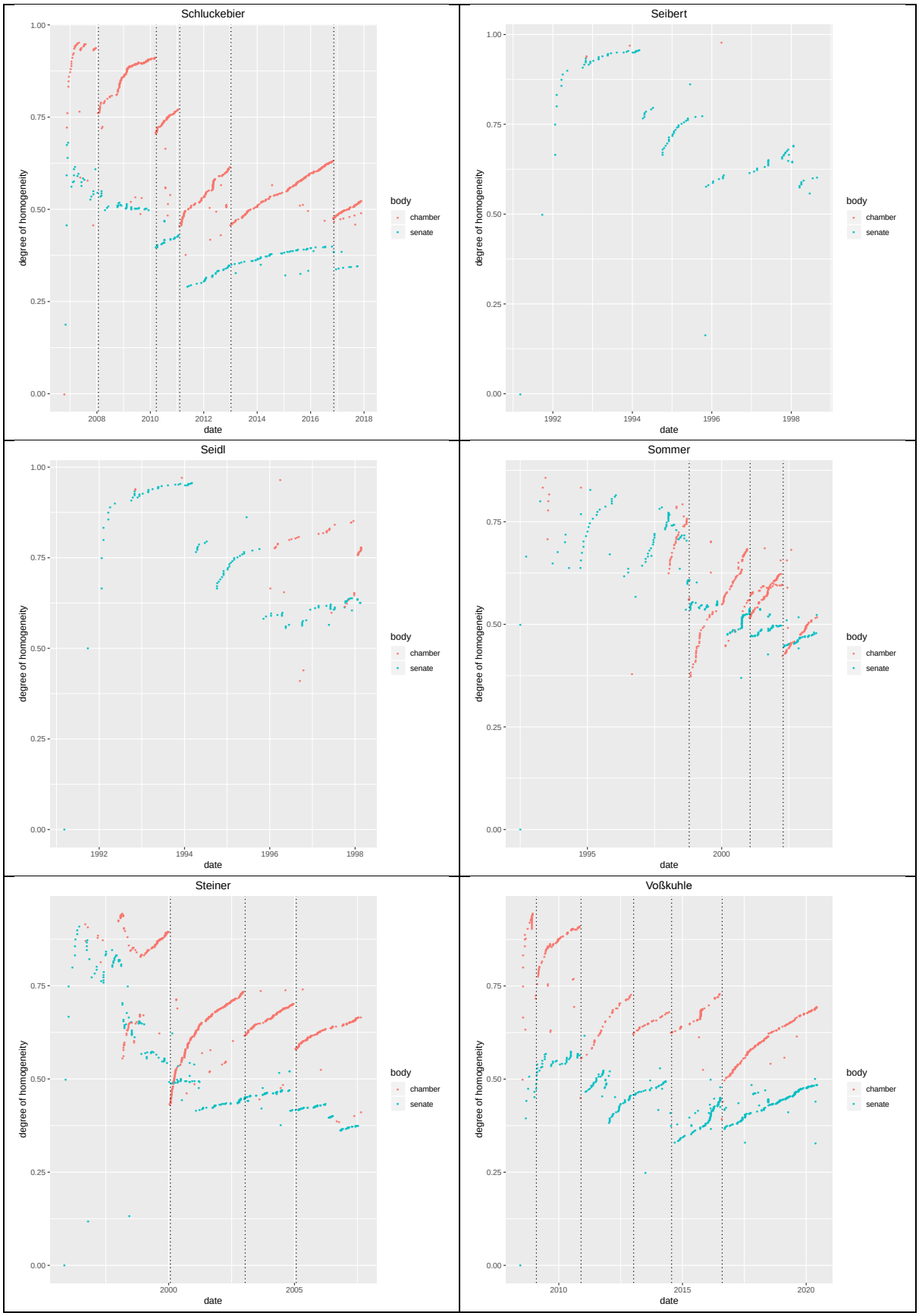

Grasshof
Grimm
Haas
Harbarth
Hassemer
Hermanns
degree of homogeneity
date
body
chamber
senate


Hoemig
Hoffmann−Riem
Hohmann−Dennhardt
Huber
Jaeger
Jentsch
degree of homogeneity
date
body
chamber
senate


Kessal−Wulf
König
Kruis
Kühling
Landau
Langenfeld
degree of homogeneity
date
body
chamber
senate


Limbach
Luebbe−Wolff
Maidowski
Masing
Mellinghoff
Osterloh
degree of homogeneity
date
body
chamber
senate


Ott
Papier
Paulus
P Kirchhof
P Müller
Radtke
degree of homogeneity
date
body
chamber
senate


Schluckebier
Seibert
Seidl
Sommer
Steiner
Voßkuhle
degree of homogeneity
date
body
chamber
senate

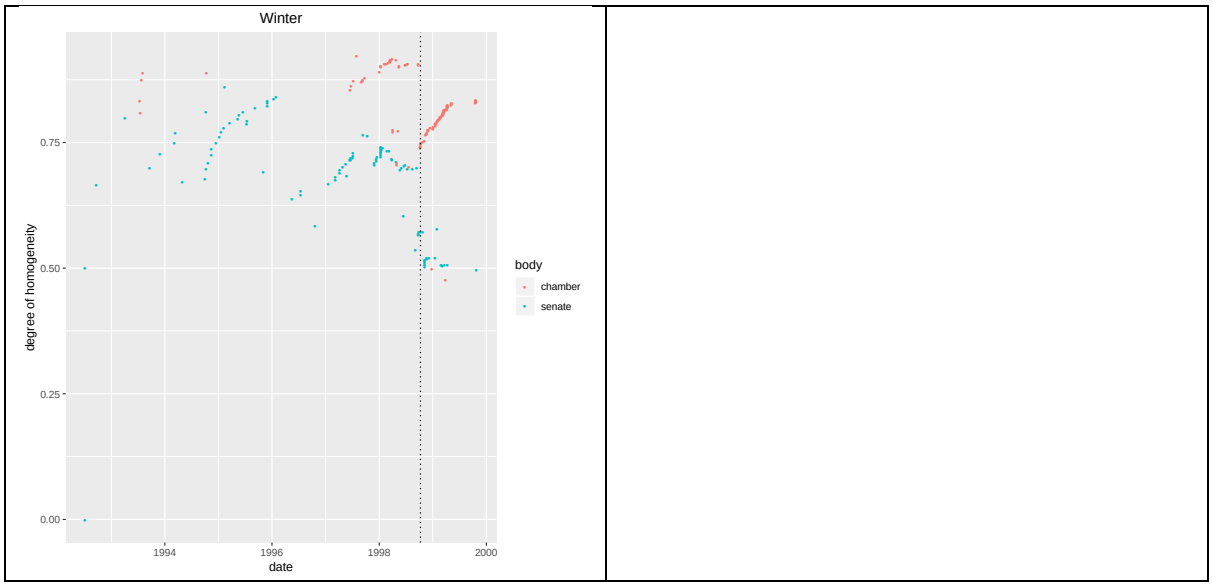


Figure 6
Breakpoints per Justice

A2: Wider Windows

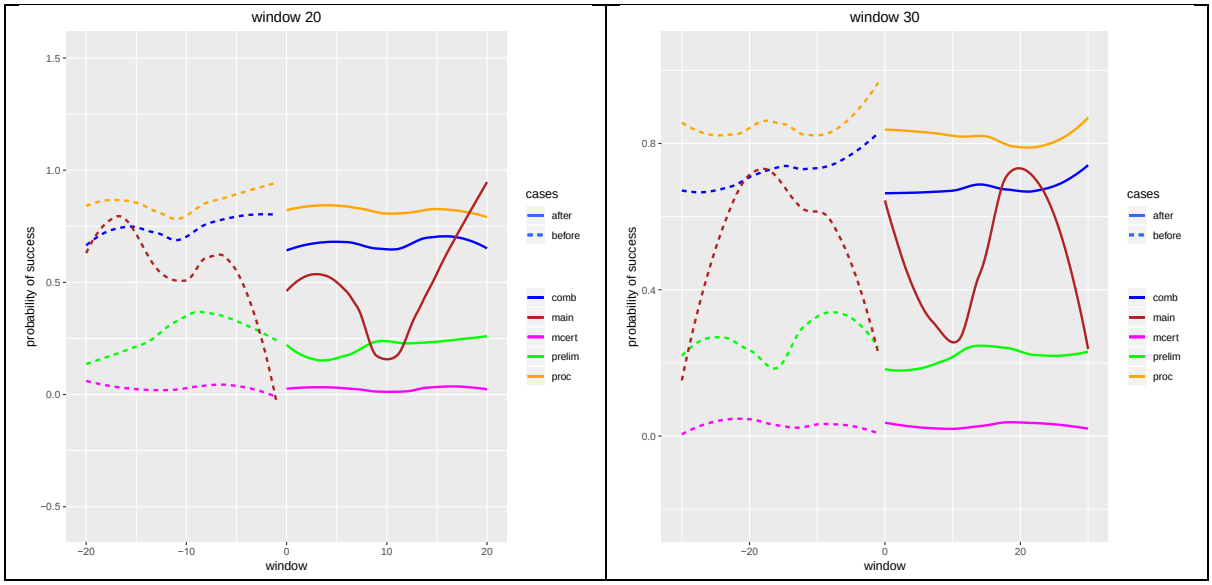


Figure 7
Regression Discontinuity with Wider Windows

	window 20	window 30
merit		
after	-.1569^{+} (.0867)	-.1030 (.0669)
cons	.5822**** (.0715)	.5702*** (.0619)
N	133	207
merit + certiorari		
after	-.0039 (.0068)	-.0032 (.0057)
cons	.0331*** (.0075)	.0339*** (.0062)
N	2391	3560
preliminary		
after	-.0560 (.0377)	-.0541^{+} (.0310)
cons	.2763*** (.0360)	.2822*** (.0302)
N	495	739
procedural		
after	-.0431* (.0195)	-.0271 (.0166)
cons	.8596*** (.0201)	
N	1337	1922
merit, preliminary or procedural		
after	-.0687** (.0218)	-.0382* (.0183)
cons		
N	1681	2424

Table 4
Regression Discontinuity

20 or 30 decisions before and after the breakpoint of chamber in which Justice participated who experienced breakpoint
if two or three Justices had the same breakpoint, only one of them is kept (87 breakpoints)
Linear Probability Models
Justice random effects
If Hausman test turns out significant, corresponding model with Justice fixed effects reported
(in this case, constant drops out by demeaning)
after: dummy variable that is 1 for decisions taken at or after breakpoint
standard errors in parenthesis
*** p < .001, ** p < .01, * p < .05, $^{+}$ p < .1