Lucky You: Your Case is Heard by a Seasoned Panel

Panel Effects in the German Constitutional Court

Christoph Engel

MAX PLANCK SOCIETY

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

Christoph Engel

February 2021

This version: June 2022

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

Abstract

Panel effects have been widely studied in randomly composed panels. However for many courts, panel composition stays constant. Then judges become familiar with each other. They know what to expect from each other. Mutual trust may develop. A local culture may emerge. If rejection is the default, familiarity is likely to help plaintiffs, as familiar panels can be more effective, and more self-confident. In the German Constitutional Court, the effect of familiarity on three success measures can be causally identified: success on the merits, with the request for a preliminary ruling, or with a procedural request. Justices experience multiple, exogenous recompositions of their chamber. In the logic of regression discontinuity, the effect of familiarity on the alternative measures for success can be identified if these recompositions lead to a clear decrease in familiarity with the other members of the chamber.

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

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

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.

Lawyers and political scientists have been keenly interested in the behavioral effects of shifting judicial decision making to a bench of multiple judges. Yet almost all of the evidence comes from court panels that are composed ad hoc, often even at random. This obviously helps with identification. Panel characteristics are credibly exogenous. Yet in judicial practice, court panels are often not formed ad hoc. In the US, this notably holds for the Supreme Court. Justices are appointed for life. The European Court of First Instance[1] as well as the European Court of Justice decide in chambers with fixed composition[2], as do, for instance, the German[3] and the French courts[4]. Fixed panels have pragmatic advantages. Benches may specialize on certain areas of law[5]. Benches may be balanced along lines that might bias outcomes, like ideological position or gender[6]. But fixed benches do also have a mechanical effect. The decision-making body stays together for an extended period of time. This creates familiarity. A rich literature in management, organization, and social psychology has shown that familiarity matters. But how does it matter for judicial decision-making?

If familiarity has an impact on outcomes, this might also hold for jurisdictions with randomly composed panels. For these panels are composed from a larger, but not an open set of judges, say all the judges from one of the Circuits of the U.S. Court of Appeals. The smaller the Circuit, and the longer two judges have been on the Court, the more it is likely that they have been sitting on the same panel before. Yet if panels are composed ad hoc, isolating the causal effect of familiarity is hard. The German Constitutional Court provides the opportunity to do this. The bulk of cases heard by the 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. Chambers stay together for a protracted period of time, typically multiple years. Even if chambers are recomposed, individual justices keep their cases. Chamber composition varies, between justices, and over time. This creates variance in

---

[1] https://curia.europa.eu/jcms/jcms/Jo2_7038/en.

[2] https://curia.europa.eu/jcms/jcms/Jo2_7029/en.

[3] See, for illustration, the distribution of judges to Senates at the highest German court in matters of civil and criminal law, the *Bundesgerichtshof*, https://www.bundesgerichtshof.de/SharedDocs/Downloads/DE/DasGericht/GeschaeftsvertPDF/2021/geschaeftsverteilung2021.html;jsessionid=DB53B849220B1D01E695E7B20D5E33EB.1_cid286?nn=10742208.

[4] See, for illustration, the distribution of judges in the *Court de Cassation,* https://www.courdecassation.fr/institution_1/composition_56/.

[5] The German *Bundesgerichtshof* for instance has a Senate specializing on corporate law (II. Senate), and another on medical malpractice (VI. Senate), see above footnote 3.

[6] See below section 3 for the German *Bundesverfassungsgericht*.

terms of familiarity. This variance makes it possible to empirically study the effect of familiarity.

Yet is the degree of familiarity causal for outcomes? Panels do not only differ by joint experience, but also by the gender composition, the tenure and age of the justices, and by the political party that has selected them. More importantly even, the measure of familiarity used in this paper is correlated with some of these demographic variables. It is therefore important to isolate the effect of familiarity. This is made possible by the fact that chamber composition regularly changes. These changes are beyond the control of the individual justice, and only partly predictable. This provides scope 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.

Social psychology and organization science have demonstrated a cognitive and a motivational effect of familiarity in multi-person decision-making bodies. Experienced bodies can handle more work. They can afford to be more forthcoming with procedural requests of the complainant, although this normally means more work. Experience in deciding together results in members better understanding and predicting each other. This increases confidence and provides scope for the development of mutual trust. On this channel, familiarity empowers the chamber to take more daring decisions. In the German Constitutional Court, overall less than 3% of constitutional complains are successful on the merits.[7] Rejection is the default. Holding for the complainant, either on the merits or with a request for a preliminary injunction, is a salient event that exposes the chamber to criticism: by other members of the court, the wider legal community, policy makers, or the media. This explains why, in the German Constitutional Court, complainants benefit from higher familiarity. Multiple indicators of success decrease discernibly and significantly after the recomposition of a chamber leads to an exogenous reduction in team familiarity.

The remainder of the paper is organised as follows: the next section discusses panel effects and derives hypotheses from the literature. Section 3 introduces the institutional framework of the German Constitutional Court. Section 4 explains the character and scope of the data. Section 5 defines the identification strategy. Section 6 reports results. Section 7 concludes with discussion.

2. Literature

Panel effects across jurisdictions. Up till now, the debate regarding panel effects in courts has been US centric. Exceptions include 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 come with different legal cultures. Arguably the courts in, for instance, Germany are much less politicised. They also

---

[7] For details see below Sections 3 and 4.

come with very different institutional arrangements. This paper investigates an arrangement that potentially has a strong behavioral effect: keeping panel composition constant over time.

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, Wendel 2020). In an earlier attempt, I have found that familiarity 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.

Ad hoc panels. There is a rich empirical literature on panel effects. It has focused on ad hoc panels. If composition is at random, as most importantly in the US Federal Courts of Appeal, from an empirical perspective this is fortunate. Panel composition is exogenous. Exploiting the exogeneity, the literature has causally identified multiple panel effects, like moderating ideological (Miles and Sunstein 2006, Kastellec 2011), gender (Boyd, Epstein et al. 2010) or racial bias (Kastellec 2020). Shifting jurisdiction from single judges to panels matters.

Some explanations are cognitive. A "counterjudge" may alert the majority of the panel to a concern that they feel unable to overlook (Sommers 2006, Spitzer and Talley 2013), induces them to better guard against the risk of implicit bias (Sommers 2006), or provides them with credible expertise (Boyd, Epstein et al. 2010). Alternative explanations could be called cultural. The fact that judges decide as a group may activate a norm of collegiality (Edwards 1998, Edwards 2003). These effects could also be at work if panel composition is held constant for a longer period.

If panels are randomly composed on a case-by-case basis, the degree of familiarity among the judges on the bench is held low. But as long as these judges are from the same district, they at least in the longer run have a chance to repeatedly decide together with another member of the court. The impact on familiarity is considerably smaller than in a fixed panel. The frequency of interaction is much lower. It is quite unlikely that the complete panel has had the same composition before. Still there is scope for the evolution of a stripped-down version of familiarity. Even this small degree of familiarity is absent if a judge from a district court, or from another circuit of the appellate courts, is ad hoc added to the appellate panel, as is possible in the U.S. federal court appellate system (Wasby 2018, Levy 2019), and if panels of three in the U.S. district courts are composed for the isolated purpose of hearing a claim of voting rights violation (Solimine 1996, Cox and Miles 2008, Mak, Sidman et al. 2021). This practice has been investigated for its consequences on the consistency of a circuit's jurisprudence (Wasby 1980), for a greater preparedness of the ad hoc members to accommodate the ideological preferences of the permanent judges on their bench (Collins and Martinek 2011), and has been criticized for its potential to hamper collegiality (Saphire and Solimine 1994).

Long-term panels. Arguably, in long-term panels, there are multiple additional behavioral channels for familiarity to matter. Very recently, the literature has begun to be interested in the behavioral effects of keeping the composition of a court constant over time. Hinkle, Nelson et al. (2022) show that familiarity between an appellate court and judges of a lower court tampers ideological divides. Swalve (2022) shows that familiarity increases procedural effort, in terms of holding an oral hearing, and writing a longer opinion.

Familiarity has for long been an issue in management, organization, and social psychology. I rely on this literature to derive hypotheses. Familiarity has multiple facets (Hanft 2002, Espinosa, Slaughter et al. 2007). For the purposes of this paper, team familiarity is critical, i.e. an individual's prior shared work experience (Reagans, Argote et al. 2005, Huckman, Staats et al. 2009, Staats 2012). Judicial panels are "concoted" groups, i.e. groups exogenously composed (Harrison, Mohammed et al. 2003: 636). Familiarity has been shown to increase performance. Familiar teams have been more productive (Goodman and Leyden 1991, Shah and Jehn 1993, Jehn and Shah 1997, Littlepage, Robison et al. 1997), more accurate (Moore and Geuss 2020), less prone to accident (Goodman and Garber 1988, Kurmann, Keller et al. 2014), and more creative (Sosa 2011). Conversely performance has been observed to go down if membership changes (Arrow and McGrath 1993, Hollenbeck, Ilgen et al. 1995).

Yet the beneficial effect of familiarity is not ubiquitous. Occasionally, familiarity has even been reported to have a negative effect on performance, for instance in a collaborative learning task (Janssen, Erkens et al. 2009) or in an experimental survival task (Kim 1997). In other studies, familiarity did not make a significant difference. This for instance held for teams of engineers reviewing product design (Wetmore III, Summers et al. 2010), or for software development if the task was complex (Espinosa, Slaughter et al. 2007).

The quality of judicial decisions is not as clearly defined as in most of the tasks for which the effect of familiarity has been investigated. It is therefore not obvious that an analogy to the findings about the performance of familiar teams can be drawn. To predict the effect of familiarity on judicial decision-making, one has to dig deeper. One must understand the channels on which familiarity may have an effect on the decisions group make.

At the highest level, cognitive, motivational and institutional effects of familiarity can be distinguished. On the cognitive side, joint decision-making experience is informative. Panel members acquire "meta-knowledge" (Gruenfeld, Mannix et al. 1996). They gain a better understanding of other team members' competencies (Littlepage, Robison et al. 1997), reputation, work performance, dependability and attention to detail (Maynard, Mathieu et al. 2019: 15). They know who knows what (Huckman and Staats 2011, Maynard, Mathieu et al. 2019: 8). They also better know other team members' attitudes (Sanbonmatsu, Uchino et al. 2011, Sanbonmatsu, Uchino et al. 2012), preferences, habits, values (Rockett and Okhuysen 2002), beliefs, likes and dislikes, education and employment history, family situation (Maynard, Mathieu et al. 2019: 15) and other personal characteristics (Mohammed and Dumville 2001: , 6). This knowledge reduces perceived uncertainty in their dealings with each other (Espinosa, Slaughter et al. 2007: 616). The actions of other team members become more predictable (Okhuysen 2001: 796). Team members are better able to interpret each others' actions (Yoon and Rolland 2012: 1135).

On the motivational side, better information translates into higher "psychological safety" (Edmondson 1999, Tucker 2007, Siemsen, Roth et al. 2009, Newman, Donohue et al. 2017). Anxiety about social acceptance is reduced (Gruenfeld, Mannix et al. 1996, Hinds, Carley et al. 2000). The members of familiar groups are less likely to betray each other (Mason and Clauset 2013) and to take offense (Maynard, Mathieu et al. 2019: 8), or to be upset (Sanbonmatsu, Uchino et al. 2012). They are less likely to be in conflict with each other (Shah

and Jehn 1993, Jehn and Mannix 2001, Sanbonmatsu, Uchino et al. 2012), and more likely to separate task conflict from interpersonal conflict (Arrow and McGrath 1993, Jehn and Shah 1997, Killumets, D'Innocenzo et al. 2015: 235). Possibly, trust develops (Jones and George 1998, Harrison, Mohammed et al. 2003: 640, Jarvenpaa, Shaw et al. 2004, Espinosa, Slaughter et al. 2007, Huckman, Staats et al. 2009, Yoon and Rolland 2012), or even affect (Rockett and Okhuysen 2002). Familiar group members are more likely to be prosocial with each other (Mason and Clauset 2013). They are "experiencing relatedness, feeling connected and supported by others" (Yoon and Rolland 2012: 1135). The group becomes more cohesive (Shah and Jehn 1993, Harrison, Price et al. 1998). In response, team members experience increased self-efficacy and competence (Yoon and Rolland 2012).

The longevity of a panel provides scope for the development of informal institutions. These institutions have been characterized as the group's "transactive memory system" (Liang, Moreland et al. 1995, Faraj and Sproull 2000, Lewis 2003, Smith-Jentsch, Kraiger et al. 2009). A metaphorical analogy to musicians getting into sync with each other has been drawn, calling the effect "social entrainment" (McGrath and Kelly 1986, Harrison, Mohammed et al. 2003: 642). The group increasingly develops the "understanding of an entity" (Yoon and Rolland 2012: 1135). It builds social capital (Chillemi and Gui 1997, Staats 2012: 620). The group may create a collective learning system (Lewis, Lange et al. 2005). There is room for the establishment of group norms (Okhuysen 2001: 796, Adams, Roch et al. 2005), routines (Killumets, D'Innocenzo et al. 2015: 235), heuristics (Reagans, Argote et al. 2005: 872), a "style of play" (Mason and Clauset 2013: 376) and other task specific processes (Marlow, Lacerenza et al. 2018: 147). Familiar groups increasingly coordinate tacitly (Huckman and Staats 2011: 311). They are more committed to group objectives (Jehn and Shah 1997).

Familiar groups communicate more (Gruenfeld, Mannix et al. 1996), and more effectively (Jarvenpaa, Shaw et al. 2004, Narayanan, Balasubramanian et al. 2011, Marlow, Lacerenza et al. 2018). This makes them more open to learning from each other (Gruenfeld, Mannix et al. 1996, Dalal, Nolan et al. 2017). They are more likely to share information that initially only one team member holds (Marlow, Lacerenza et al. 2018), and to rely on this information (Staats 2012: 620). They feel more comfortable with expressing disagreement (Gruenfeld, Mannix et al. 1996), and with being criticized (Shah and Jehn 1993). This leads to an "ideal conflict profile" (Jehn and Mannix 2001: 238). Familiar groups are more willing to experiment, to take risks, and to innovate (Edmondson 1999, Staats 2012: 620). Familiar groups establish a better division of labor (Liang, Moreland et al. 1995, Reagans, Argote et al. 2005). They make sure the right member is assigned to the right task (Faraj and Sproull 2000, Huckman and Staats 2011). They exploit within group diversity more productively (Harrison, Price et al. 1998, Harrison, Mohammed et al. 2003). They react faster and more appropriately to changes in the task (Huckman and Staats 2011), and are better able to handle complex tasks (Espinosa, Slaughter et al. 2007).

Implications for judicial decision-making. Familiar teams can hence handle more, and more challenging, work. I expect this to also hold for judicial teams, i.e. court panels. It has been shown with experimental methods that judges aim at doing a good job, rather than maximizing income or leisure (Engel and Zhurakhovska 2017). And it has been shown with observational data that courts do a better job when given more time (by an increase in judicial personnel): they spend more time per case, hear more witnesses, are less likely to revert to

summary procedure, and write longer opinions (Engel and Weinshall Margel 2020). Familiarity enables a better division of labor, and the establishment of panel specific decision-making routines. Panel members have a better sense for potentially contentious issues, and can steer clear of them. On all these channels, familiarity enables the panel to handle more cases in the same time. If the case load is exogenous, this gives them more time per case. An experienced panel can afford a more laborious procedure.

Now constitutional jurisprudence is asymmetric by design. This in particular holds for the object of study of the present paper: constitutional complaints brought by a citizen against government. Since the foundation of the German Constitutional Court in 1951 and the end of 2019[8], only 2.3% of all constitutional complaints have been successful[9]. Constitutional scrutiny is meant to be exceptional. Under the rule of law, government, the legislature, and the lower branches of the judiciary, can ordinarily be expected to obey the constitution. Consequently, dismissal is the implicit default. It takes a lot for the court to be convinced that, indeed, it must intervene. Moreover, the court faces a huge caseload[10]. To shoulder the caseload, the Court has even been given power to decide without giving reasons[11]. Efficacy is an important concern in court practice. For all these reasons, granting a procedural request is already a favor to the complainant which the panel must be able to afford. The more panel members are familiar with each other, the more they are likely to sense the capacity to do so. Granting a procedural request is a proxy for the panel devoting more effort to the case.

This holds *a fortiori* for granting a preliminary injunction. Such an injunction does not preempt the final ruling. The court may still reject the complaint when deciding on the merits. But if it grants the request, the panel has to see the case twice: at this intermediate stage, and when eventually deciding it. The court must have the capacity to invest the extra effort. Moreover, a preliminary injunction at least temporarily prevents government from acting as desired. Legally, the court has power to issue the injunction. But the court starts a conflict with the affected branch of government. Ultimately, the authority of the court hinges on government. Were government too much annoyed, it could approach the legislator with the intention to curtail the court's powers. Recent experiences from Hungary and Poland demonstrate that this is not a merely theoretical concern. Consequently, if it grants the injunction, the court engages political capital. This is how the motivational effect of familiarity could matter. 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). The greater familiarity, the more a member who believes the injunction is warranted may feel confident arguing for it. Mutual trust may help the court becoming more daring. Justices may also more trust their ability to couch the intervention in terms that make it palpable for more cautious members of their chamber, or for the defendant.

These considerations hold *a fortiori* for a decision in favour of the complainant on the merits. As success on the merits is the rare exception, it is much easier for the court to dismiss, rather than to grant the request for an intervention. The court needs both the resources to do so,

---

[8] i.e. the end of observation in the present data.
[9] https://www.bundesverfassungsgericht.de/DE/Verfahren/Jahresstatistiken/2019/gb2019/A-I-1.pdf?__blob=publicationFile&v=2.
[10] In the end of 2019, 3472 cases were pending, to be decided by 16 justices.
[11] § 93c I 1 Bundesverfassungsgerichtsgesetz.

and the political will. There is a norm of deciding narrowly, as each individual decision engages the authority of the court at large. Arguably, it takes courage for the panel to become proactive. If it does, the ruling is salient, and considerably more likely to expose the chamber to criticism from other members of the court, from the legal community, from policy makers, or from the media. This expectation resonates with a finding from Israeli courts. If an intervention eases the caseload, the courts exert extra effort. This extra effort is predominantly to the advantage of plaintiffs (Engel and Weinshall Margel 2020).

Both on the productivity and on the trust channel, it can therefore be hypothesized:

Hypothesis: The more pronounced the joint experience of the justices on the panel, the more the panel is likely
a) to grant the constitutional complainant a procedural request,
b) to grant a preliminary injunction if favour of the complainant,
c) to decide in favour of the complainant on the merits.

3. 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. 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, or 96.54%, have been constitutional complaints. On Dec 31, 2019, 3,300 constitutional complaints have been pending.[12] The court does not have the power of *certiorari*; it officially has to decide each case. Over the years, the court has however developed a whole panoply of techniques for implicit docket control (Engel 2020).

The court is composed of two Senates with eight justices each. If the case originates in a constitutional complaint, 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.[13] Decision by chamber however presupposes unanimity.[14] 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.

---

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

[13] § 93c I 1 Bundesverfassungsgerichtsgesetz; before this rule has been introduced, if the chamber wanted to accept the complaint, it had to refer the case to the senate.

[14] § 81a, 1 Bundesverfassungsgerichtsgesetz; § 93d III 1 Bundesverfassungsgerichtsgesetz.

Justices have a fixed tenure of 12 years, which is non-renewable. They can resign earlier. But this is a rare event.[15] 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.[16] This political compromise is stabilized by the legal rule that new justices need a majority of 2/3 to be appointed. Thus far, the far right AfD has been denied a position.[17]

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.[18] Officially, each Senate decides every year about the composition of each of three chambers for the coming year.[19] In practice, chamber composition tends to be constant for a longer period. After three years, the statute wants composition to change, at the latest.[20] As Senates have 8 members, not 9, one Justice must sit on two chambers. 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 more than 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.

---

[15] One of these exceptions 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.

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

[17] It has only been represented for the first time in Parliament in 2017 though.

[18] 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 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.

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

[20] § 15a I 2 Bundesverfassungsgerichtsgesetz.
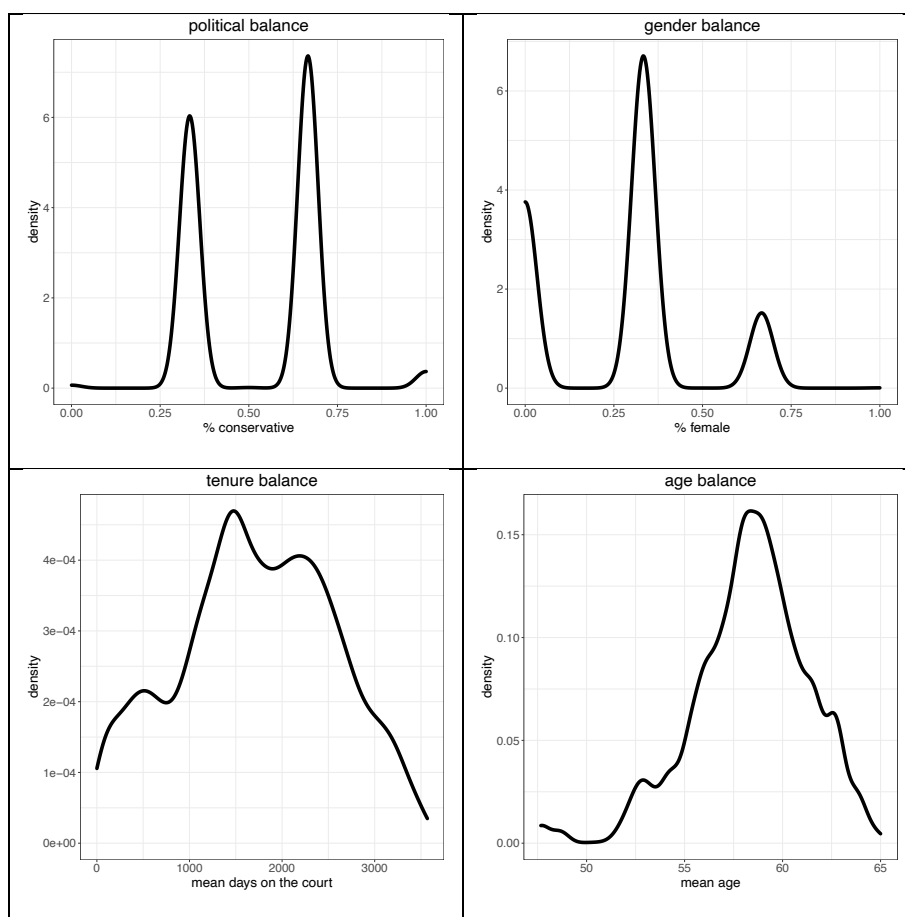
Figure 1
Chamber Composition
all 6359 chamber decisions posted on the court's website until July 2020
density plots
conservative: fraction of justices on the chamber that have been picked by CDU or FDP
female: fraction of female justices on the chamber

4. Data

Sample. Since its inception, and until Dec 31, 2019, the German Constitutional Court has heard 243,494 cases.[21] 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.[22] This paper uses a dataset that originates from scraping all decisions that have been posted online, until July 2020. The paper only uses decisions made by a chamber (not a senate or the plenary). It only uses decisions about constitutional complaints. Online, there are 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

---

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

[22]

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

available online. As I was concerned about a potential selection bias, I have had an interview with one of the current justices, Justice Baer. She has given me permission to write that only a very small fraction of the cases not posted online, below 1%, come with reasons. All other cases are not posted because there would be nothing to post: the court has exploited the power to either reject or even accept a constitutional complaint without any written reasons. The dataset consists of 6,466 decisions posted online have originated in a constitutional complaint and have been decided by a chamber. These are 85% of all cases that the court has posted online.

Dependent variables. 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. Almost half of the constitutional complaints heard by a chamber are successful. 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. If non-acceptance is taken into account, only 2.35% of the posted constitutional complaints are successful. As acceptance without reasons is much rarer than rejection without reasons, 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.

	failure	success
merit	114	94
merit + non-acceptance	3901	94
preliminary injunction	644	197
procedural decision	367	1766
combined	825	1882

Table 1
Success in the Constitutional Court
constitutional complaints decided by a chamber

Independent variable. The main explanatory variable investigated in this paper is the degree of familiarity among the justices currently on the bench. For each of the justices and each decision, the familiarity score is calculated as follows:

1. remove decisions before 1.1.1998
2. remove decisions in which Justice X has not participated

3. separately for each decision, generate a dummy variable that is 1 if Justice X and any Justice Y with whom she has ever jointly been on the bench have jointly decided the case at hand
4. generate the running count of each of the variables generated in step 3
5. generate a running count of all decisions in which Justice X has participated
6. for each decision, divide step 4 by step 5, to get the fraction of cases both justices have jointly decided up till this point
7. per case sum up all variables defined in step 6, and divide them by the number of justices on the bench, -1

As familiarity may not only result from jointly sitting on a chamber, for these calculations I also consider decisions taken by the senate of 8 justices or the rare cases in which the plenary of 16 justices has decided.

Let me illustrate the approach in an example. The upper panel of Figure 2 shows with which other justices Justice Baer has jointly decided as a member of a chamber. One sees a clear pattern. She has started in a chamber with Justices Schluckebier and Ferdinand Kirchhof. While the latter has remained a member of her chamber, early in 2013 Justice Schluckebier has been replaced by Justice Masing. In 2017 Justice Ferdinand Kirchhof has left the court. At this point, the chamber has been completely recomposed, and Justice Baer has for about two years been jointly deciding with Justices Britz and Eichberger. For a few months, Justice Radtke, who newly joined the court in replacement of Justice Eichberger, has been in Justice Baer's chamber. Finally since 2019, Justice Baer is deciding together with the newly appointed Vice President of the Court, Justice Harbarth, and Justice Ott.

The lower panel of Figure 2 shows how these regular recompositions of the chambers translate into familiarity scores. The recompositions have little effect on the familiarity score when deciding as a member of the full Senate. This is expected, as in the Senate there are also those justices with whom a justice has been in a chamber sometime in the past, and those with whom she has only decided in the Senate. By contrast, for chamber decisions, regular recompositions have a strong effect. As long as composition is unchanged, the score increases, and it drops pronouncedly in the moment of recomposition.

Consider the numbers at the first recomposition, on Dec. 30, 2012. Up till this day, Justice Baer had participated in 141 decisions. In 137 of them she had jointly decided with Justice Schluckebier, i.e. in 97.16% of all decisions. In 130 cases, or 92.20% of all decisions, she had jointly decided with Justice Ferdinand Kirchhof. Adding these two percentages up, and dividing the sum by the number of justices presently on the bench – 1, results in the score of .947.
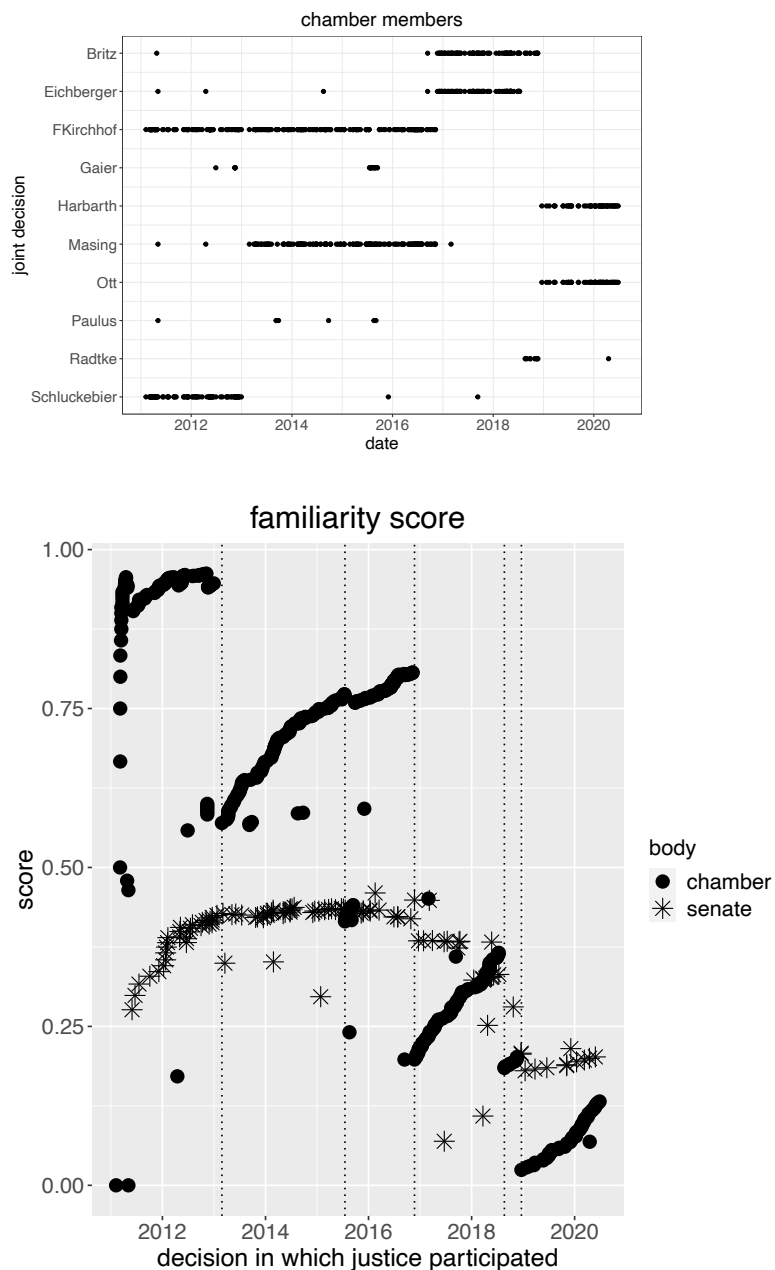
Figure 2
Familiarity Scores for Justice Baer
upper panel: dot for each chamber decision in which Justice Baer has jointly decided with the respective other justice
lower panel: dot (for chamber decisions) or star (for senate decisions) for every decision in which Justice Baer has participated
score: total of all bilateral scores for all justices who were on the bench, other than Justice Baer, divided by number of justices on the
bench – 1 (for detail see text of this section)

In Appendix A1, for all justices the development of their familiarity score over time is reported. As the analysis of this paper is confined to chamber decisions, these graphs only present the score when deciding as a member of a chamber. Justices occasionally decide in differently composed chambers, as they have to replace one of their colleagues. As each Senate consists of 8 justices, but 3 chambers of 3 justices, one justice must sit in two chambers. This is often, but not always, the presiding justice of the respective senate. For both reasons, it is hard to define breakpoints by an abstract criterion. To make sure the breakpoints are meaningful, I have relied on the power of human vision to recognize patterns,

and have manually defined breakpoints. I have required a discernible reduction in familiarity which is lasting. Breakpoints are represented in Appendix A1 by dotted vertical lines. This procedure is designed to err on the cautious side. If in doubt, I have refrained from defining a breakpoint. This explains why, not so rarely, chamber recomposition does not lead to a clear breakpoint for all affected justices.

Four justices (Grasshof, Kruis, Seibert, Seidl) have left the court so early that no break in their familiarity score can be found. 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 estimation, the data is normalised. The analysis uses chamber decisions from a window starting t decisions before the break up till t 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.

5. Identification Strategy

Potential Endogeneity. This paper is interested in the effect of familiarity among the deciding justices on a series of measures for the success of a constitutional complaint. Now, the mean degree of familiarity (for all justices participating in the respective decision, 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 = .36***), and it is negatively correlated with the fraction of justices selected by one of the conservative parties CDU or FDP (r = -.14***). 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. This is why simply regressing the respective success measure on the degree of familiarity would be problematic.

Regression discontinuity. Yet the very fact that chambers are regularly recomposed provides an opportunity for identifying a causal effect of familiarity. This is because, for the individual justice, the resulting 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 most an influence on chamber composition. Even this potential 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. This provides the opportunity to identify a LATE, a local average treatment effect. In the spirit of regression discontinuity, it is possible to causally identify the effect of familiarity near the

respective breakpoint. One can treat recomposition as an exogenous shock. If one finds an effect on the probability of success in the neighbourhood of this breakpoint, it must have been caused by the break, and the reduction in familiarity to which it has led.

Potential Concern with Measurement. Another potential concern is less important, but still worth addressing, to be on the safe side. The paper follows the literature on team familiarity and measures the degree of team familiarity by the previous experience in joint decision making (see e.g. Huckman, Staats et al. 2009: 90, Huckman and Staats 2011: 317, Staats 2012: 623, Marlow, Lacerenza et al. 2018: 151). Now mechanically familiarity measured this way is high soon after a justice joins the court[23]. As tenure is 12 years, and most justices are affected by multiple breakpoints, this mechanical effect likely largely washes out anyhow. Yet regression discontinuity also takes care of this potential confound. If the confound exists, it results from the individual justice's tenure. Near the respective breakpoint, her tenure is held constant.

Bandwidth. The analysis 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 this 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.[24] The main specification uses a window of 10 decisions before and 10 after the respective breakpoint. In Appendix A2, results for wider windows of 20 or 30 decisions before and after the respective breakpoint are reported.

Duplicate 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.[25] 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. Moreover, the regressions use 10 decisions before and 10 decisions after the breakpoint. These decisions are very unlikely to be the same for the justices affected by the break.[26] Yet to be on the safe side, results from using data all 135 breakpoints are presented side-by-side with results from a narrower dataset that removes duplicates, so that only data around 87 unique breakpoints is used.

---

[23] As illustrated by Figure 5.

[24] 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*.

[25] Recall that the cautious approach only uses breakpoints that are clearly discernible from the graphs in Appendix A1. Not so rarely, this is the case for one justice who is affected by the recomposition, but not for another.

[26] This could only happen if two Justices remain together, and get a new colleague. For the most part, the Senates have used the opportunity to a more wide-reaching recomposition of their chambers.

6. Results

Main result. Figure 3 is the main result of the paper. As the figure shows, 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. 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.[27]
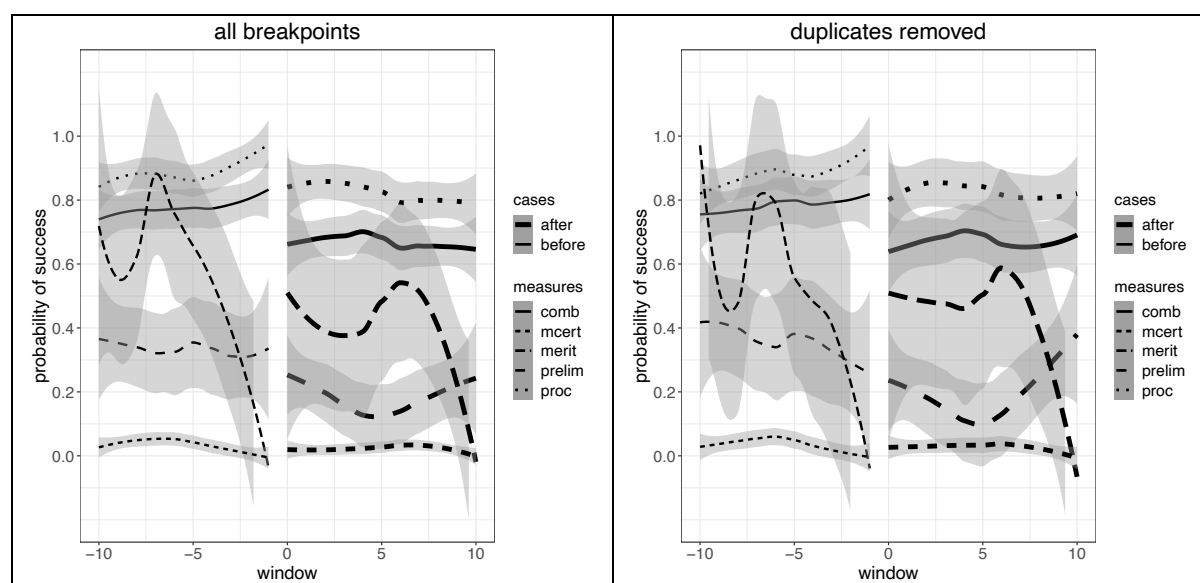
Figure 3
Effect of Familiarity on Success Near Breakpoints
decisions normalized about the breakpoints (see Appendix A1 and Section 4 for the definition of breakpoints)
window: 10 decisions before and after the breakpoint
success variables: merit: on the merits; mcert: merit, including the decision not summarily reject the complaint as inadmissible or obviously unfounded; prelim: granting a preliminary ruling; proc: granting a procedural request; comb: succeed either on the merits, or with a request for a preliminary injunction, or with a procedural request
all five success variables are dummies. Hence on the y-axis, fractions are reported
results are reported with 95% confidence intervals, using `R ggplot2 geom_smooth,` which employs a cubic spline for smoothing
left panel: any breakpoint reported in Appendix A1 is used
right panel: per breakpoint, only the familiarity score for one justice is used

Table 2 provides statistical tests. Models 1 and 3 regress each of the five success variables defined in Section 4 on a dummy that is 1 if the decision has been taken after the breakpoint, and 0 otherwise. Arguably, this is the most reliable explanatory variable, as it is not affected by the orthogonal distribution of other impacts on court decision making, like the domain of life, the quality of legal advice of which the complainant has benefitted, or the political

---

[27] 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.

salience of the case. Decisions do not get more weight if such an orthogonal feature of the case happened to occur near or a little further away from the breakpoint. Yet as a complement, models 2 and 4 also offer results that take the distance from the breakpoint into account.

The psychological literature suggests that familiarity makes teams more effective. In the context of the Constitutional Court, this could translate into scope for a more involved and laborious procedure. In line with this explanation, in all four specifications after chamber recomposition the probability that the court grants the complainant a procedural request is substantially and significantly lower.[28] Lower effectiveness is also a plausible reason why all four specifications show that the probability of granting a request for a preliminary injunction is significantly lower after the drop in familiarity.

For explaining an effect of familiarity on the decision on the merits, one needs a motivational channel. Arguably greater familiarity enables a chamber to become more courageous, as justices better trust each other, and as they have more confidence in their individual expectations about the reactions of their colleagues to proposing a bolder intervention. Models 1-4 support this line of reasoning. In all specifications, the effect of a reduction in familiarity on success on the merits is significantly negative. Yet the result for the dependent variable that includes certiorari is less clearly established. At conventional levels, the effect of familiarity is only significant when using the continuous explanatory variable, and including duplicate breakpoints. There are two more weakly significant findings pointing into the same direction, though.

As explained in section 4, decisions on the merits are rare in the first place. Consequently, for estimating an effect near the respective breakpoint, there is only relatively little data. This is an additional reason for treating results on the merits with caution. By contrast when combining success on the merits, on a preliminary ruling, or on a procedural request, in all four models there is a strong and highly significant effect of familiarity.

	all breakpoints		duplicate breakpoints removed				
merit	model 1	model 2	model 3	model 4	model 5	model 6	model 7
after	-.2064*		-.2354*		-.0461	-.0529	.3536
	(.0927)		(.1149)		(.0582)	(.1110)	(.2380)
window		-.0210**		-.0275**			-.0517**
		(.0068)		(.0084)			(.0182)
paragraphs						.0089**	
						(.0028)	
tenure						.00016[+]	
						(.00009)	
left						-.8549*	
						(.3565)	
cons	.5916***	.4858***	.5948***	.4657***	.0253	.4100	.2769[+]
	(.0806)	(.0637)	(.0969)	(.0764)	(.0431)	(.2184)	(.1481)
N	104	104	73	73	73	73	73

[28] Note, however, that the effect is only weakly significant in model 4, i.e. with the reduced dataset and the continuous explanatory variable.

merit + certiorari							
after	-.0125+ (.0074)		-.0081 (.0095)		-.0111 (.0089)	-.0030 (.0097)	.0249 (.0195)
window		-.0014* (.0006)		-.0013+ (.00078)			-.0031+ (.0016)
paragraphs						.0015*** (.0003)	
tenure						.00001 (.000007)	
left						-.0462 (.0282)	
cons			.0333*** (.0083)	.0291*** (.0066)	.0058 (.0065)	-.0076 (.0193)	.0160 (.0123)
N	1877	1877	1213	1213	1213	1213	1213
preliminary							
after	-.1371** (.0425)		-.1496** (.0529)		-.0881* (.0418)	-.1220* (.0531)	-.1491 (.1018)
window		-.0112** (.0036)		-.0109* (.0045)			-.00004 (.0086)
paragraphs						-.0042** (.0015)	
tenure						-.00006+ (.000037)	
left						.4385*** (.1331)	
cons	.3508*** (.0371)	.2734*** (.0281)	.3748*** (.0469)	.2900*** (.0361)	.0499 (.0314)	.3296** (.1064)	.3747*** (.0681)
N	409	409	270	270	270	270	270
procedural							
after	-.0663** (.0212)		-.0691* (.0262)		-.0736** (.0238)	-.0429+ (.0246)	-.1091* .0523)
window		-.0046** (.0018)		-.0040+ (.0022)			.0044 (.0044)
paragraphs						.0087*** (.0008)	
tenure						.00002 (.00002)	
left						.1194 (.0752)	
cons	.8844*** (.0215)	.8495*** (.0187)			.0365* (.0168)	.5581*** (.0554)	.9053*** (.0343)
N	1040	1040	689	689	689	689	689
merit, preliminary or procedural							
after	-.1093*** (.0244)		-.1148*** (.0296)		-.1103*** (.0276)	-.0929*** (.0277)	-.1494* (.0587)
window		-.0075*** (.0020)		-.0074** (.0025)			.0033 (.0049)
paragraphs						.0108*** (.0009)	

tenure						.000009 (.00002)	
left						.2104* (.0817)	
17	.7767*** (.0224)	.7188*** (.0186)	.7816*** (.0270)	.7213*** (.0226)	.0566** (.0198)	.4116*** (.0606)	.8000*** (.0383)
N	1315	1315	869	869	869	869	869

Table 2
Effect of Familiarity on Success Near Breakpoints
10 decisions before and after the breakpoint of chamber in which Justice participated who experienced breakpoint
Models 1-2: data around all 135 breakpoints
Models 3-5: 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
tenure: mean number participating justices have been on the court
left: fraction of justices nominated by SPD or Greens
standard errors in parenthesis
*** p < .001, ** p < .01, * p < .05, + p < .1

The evidence thus supports the main

Result:
The more justices are familiar with each other, the more it is likely that a constitutional complaint
a) is successful on the merits,
b) leads to a preliminary injunction in favour of the complainant,
c) leads to a procedural decision in favour of the complainant.

Robustness. Breakpoints are beyond the control of the respective justice. Near the breakpoint, the effect of the break is therefore causally identified. At each breakpoint, the degree of familiarity drops. Yet mechanically, panel composition changes as well. The newly composed chamber has a different mean tenure, usually also a different mean age, and often the gender composition, and the political parties that have selected the Justices change as well. Moreover what looks like a change in familiarity might actually be the effect of earlier interaction in the framework of the wider Senate from which the new chamber members are selected. One of the new chamber members might just be a difficult person, to only mention a few options. Yet there is a straightforward way to disentangle such effects from familiarity. Model 5 adds a fixed effect for each breakpoint.[29] This removes all changes in panel composition that are held constant across the observations before and after the breakpoint, and isolates the pure effect of the change, i.e. the difference in familiarity. With breakpoint fixed effects, the effect of a drop in familiarity on success on the merits turns insignificant.

---

[29] Technically it demeanes all dependent variables at the respective breakpoint, i.e. it replace success in each of the 21 cases by the difference between success in the case at hand – mean success in all 21 cases. By this procedure, mechanically, all changes in chamber demographics drop out, as they are held constant.

This is a further reason to treat this result with caution. However the effects on success with a request for a preliminary injunction, on procedural matters and on the combined success measure remain significant and sizeable.

Some of the potential additional effects at the breakpoints are observable. Model 6 adds them as controls. 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 for the complainant 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. 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 of the chambers, it should disappear when controlling for this proxy.

One consistent explanation for the effect of familiarity on success in procedural matters, with a preliminary injunction, or on the combined success measure, is lower effectiveness of the newly composed team. An alternative cause of effectiveness might be the time the members of the chamber have spent on the court. Arguably, the longer they have been on the court, the more they are experienced and effective. Note that the relationship between recomposition and the degree of experience on the court is not mechanical. Mean experience may even increase, if a very experienced justice joins a chamber of more junior justices. As the moment is observed when each justice has joined the court, the mean tenure of the three justices on the chamber is available, and added to model 6 as a control for the potential effect of professional expertise.

Finally, ideology might matter. It is conceivable that more liberal justices are more inclined to decide in favour of the complainant. This consideration is more intuitive for success on the merits, but liberal justices might also want to help the complainant on procedural grounds, or with a preliminary ruling. As chambers consist of three justices, it is quite likely that the majority shifts from liberal to conservative, or vice versa.[30] As chamber composition is observed, model 6 further controls for the fraction of justices selected by either the social democrats or the greens ("left").

These controls do have explanatory power. All five success measures substantially and significantly increase the longer the ruling. If the court takes the case more seriously, complainants have a greater chance to win. If more liberal justices are on the bench, a

---

[30] Recall from Figure 1 that the court tries to avoid all-liberal or all-conservative chambers.

preliminary injunction is more likely, as is success on the combined measure. The mean degree of experience on the court only has weakly significant, small effects. Of course these are only controls, not causal effects. Most importantly, with these controls, the effect of familiarity on success with preliminary injunctions, on procedural matters,[31] and on the combined success measure remain significant.

Technically, the previous analysis treats the breakpoints as truly exogenous. Regression discontinuity is more cautious. It allows for treatment (the drop in familiarity, in the present context) to be partly determined by covariates. Then the continuity assumption is critical for identification. Technically, continuity is established with the help of variables that control for the effect of these covariates (for background see Lee and Lemieux 2010, Lee and Lemieux 2014). Arguably, the breakpoint fixed effects, introduced in model 5, serve this function. Alternatively, and even more in line with a classic regression discontinuity design, such continuous influences can be captured by continuous time, normalized about the respective breakpoint. These estimations are reported in model 7. Even when using this cautious approach, there is a significant effect of treatment ("after") on procedural success, and on the combined success measure.

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 to matter 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 3 with Figure 6, 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 long-term trends. Success on the merits is very volatile, most likely due to the very small number of observations.

Comparing Table 2 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 only weakly significant, while it was significant at conventional levels 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. 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 success.

Global association between familiarity and success. The fact that chamber recompositions are beyond the control of chamber members makes it possible to identify the causal effect of (a drop in) familiarity on various dimensions of success with a constitutional complaint. Yet this identification strategy is confined to a local effect, in the proximity of the respective breakpoint. A global effect of familiarity on success cannot be identified. It still is interesting

---

[31] $p = .08136$.

to report the association. Figure 4 does so for the maximum familiarity score among the three members of the chamber. Interestingly, the association is most pronounced for success on the merits, and is also sizeable for success with procedural requests, and for the combined measure. The relationship is less clear if one adds the decision to hear the case to the measure for success on the merits, and for preliminary rulings. These visual impressions are supported by regression analysis, Table 3 The fact that the highest degree of familiarity is so pronouncedly correlated with success suggests that this justice tends to have a formative influence on the team spirit. To the extent that originally views within the chamber diverge, the justice who has been most closely related to his two colleagues in the past can be most instrumental in forging a compromise.
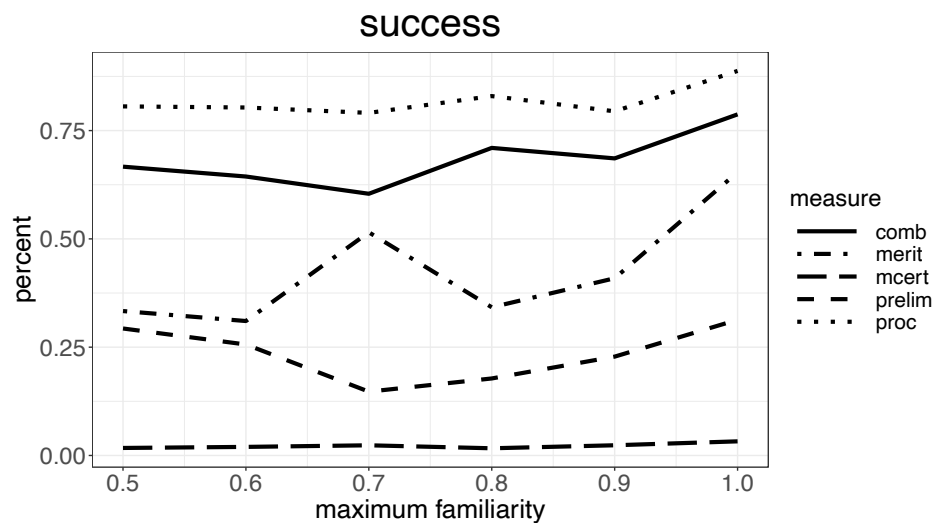

Figure 4
Global Association between Maximum Familiarity and Success
success variables: merit: on the merits; mcert: merit, including the decision not summarily reject the complaint as inadmissible or obviously unfounded; prelim: granting a preliminary ruling; proc: granting a procedural request; comb: succeed either on the merits, or with a request for a preliminary injunction, or with a procedural request
all five success variables are dummies. Hence on the y-axis, fractions are reported

	merit	mcert	prelim	proc	comb
maximum familiarity	.641***	.030+	.018	.138**	.300***
	(.213)	(.015)	(.087)	(.051)	(.054)
cons	-.035	.001	.221***	.722***	.468***
	(.165)	(.012)	(.064)	(.040)	(.042)
N	208	3995	841	2133	2707

Table 3
Global Association between Maximum Familiarity and Success
Linear Probability Models, separately for each success variable
success variables: merit: on the merits; mcert: merit, including the decision not summarily reject the complaint as inadmissible or obviously unfounded; prelim: granting a preliminary ruling; proc: granting a procedural request; comb: succeed either on the merits, or with a request for a preliminary injunction, or with a procedural request
maximum familiarity: highest familiarity score for one of the three justices on the chamber
standard errors in parenthesis
*** $p < .001$, ** $p < .01$, * $p < .05$, + $p < .1$

7. 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.

Back of the envelope calculations based on Table 3 suggest that chamber composition matters, though. These regressions work with the familiarity score introduced in section 4. Among the three justices on the chamber, the regressions use the highest score. It is the score for the justice who has most frequently decided jointly with one or both of the remaining chamber members. In the data, this maximum is at least 15.44 (the justice with the highest familiarity has on average been together with his two colleagues in 15.44% of all cases she has heard), and at most 99.05. If the maximum is .2, the complainant stands a 9.4% chance to succeed on the merits, provided the chamber has accepted the case for a decision on the merits in the first place. If the maximum familiarity score on the bench is .8, the odds of winning on the merits jump to 47.83%. This is an increase in the odds of winning by 38.48 pp! If one takes non-acceptance into account, the odds of winning are much lower. Still complainants have an almost 4 times higher chance to get a favorable decision on the merits. The probability is 0.68% when the highest degree of familiarity is .2, and increases to 2.46% if the highest degree of familiarity is .8. Chances are high in the first place that the chamber grants a procedural request.[32] It is 74.99% if the maximum degree of familiarity is .2, but it increases by 8.3 pp to 83.27% if the maximum degree of familiarity is .8. Finally if one counts a favorable decision on the merits, on a preliminary ruling or on a procedural request as a success, the probability increases from 52.77% if the maximum degree of familiarity is .2 to 70.78% if the maximum degree of familiarity is .8. It thus increases by 18.01 pp.

These numbers alert to a policy concern. But the evidence from which these numbers are calculated is correlational. As such, these results do not prove that familiarity is indeed the source of the pro-complainant bias, rather than any other (observed or unobserved) feature of the decision-making body, the case, the parties or the general political climate with which familiarity is correlated. This is why it is important that regular recompositions of the chambers frequently lead to a pronounced reduction of team familiarity. These recompositions are essentially beyond the control of individual justices. Consequently near clearly discernible breakpoints a causal effect of familiarity can be identified. This analysis of local effects supports the observed global effects. The drop in familiarity leads to a sizeable and significant reduction in the probability that a constitutional complaint is successful. Results are most robust for procedural requests, and for the combined measure. But in multiple specifications, there is also a significant (local) effect on the probability of granting a preliminary injunction, and of success on the merits.

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

---

[32] Provided the chamber at all gives written reasons, and then posts the decision online, see above section 3.

fraction of all decisions. Ultimately, only complete data could prove that the results reported in this paper do not result from selection. As almost all unpublished decisions also come without written reasons, even seeing the decisions would not help. One would need the advisory opinions prepared by the clerks – which the court keeps confidential. 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 hypothesis section, 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 (in the chambers) 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.

This paper documents the pronounced effect of familiarity on the decisions taken by judicial panels. The hypothesis section explains the cognitive, motivational and institutional channels on which the effect might obtain. But it is beyond the scope of the present paper to isolate one of these channels, or to measure their relative importance. This must be left for future work. Very likely, observational data will not be best suited for the purpose. It might be preferable to supplement the present analysis (that has documented the existence and the size of familiarity effects) with experiments focusing on individual channels.

In a way, the German Constitutional Court already addresses the policy concern. Chambers are regularly recomposed. Justices are not allowed to stay together on a chamber for a large part of their 12 year term. In practice, chamber composition changes at least every third year. Precisely since familiarity makes a judicial bench more effective, the court must strike a balance. It is not desirable that complainants stand a lower chance after, beyond their control, the competent chamber has been recomposed. But it is also not desirable that the court remains below its potential since justices have too little joint experience. Hence the main contribution of this paper to judicial policy making in Germany is largely confined to making the concern explicit, so that it can be reflected in the decision of the court about when and how to reassign justices to decision-making panels. This is different in jurisdictions that do not have fixed panels in the first place. What may appear like a cautious practice in the interesting of bolstering judicial impartiality turns out to be an intervention to the detriment of complainants, and plaintiffs more broadly. If panels are routinely composed ad hoc and at

random, this practice hurts those whom the judicial system is meant to help: those who address the court in hopes of receiving legal redress.

References

ADAMS, SUSAN J, SYLVIA G ROCH AND ROYA AYMAN (2005). "Communication Medium and Member Familiarity. The Effects on Decision Time, Accuracy, and Satisfaction." Small Group Research **36**(3): 321-353.

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

ARROW, HOLLY AND JOSEPH E MCGRATH (1993). "Membership Matters. How Member Change and Continuity Affect Small Group Structure, Process, and Performance." Small Group Research **24**(3): 334-361.

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.

CHILLEMI, OTTORINO AND BENEDETTO GUI (1997). "Team Human Capital and Worker Mobility." Journal of Labor Economics **15**(4): 567-585.

COLLINS, JR, PAUL M AND WENDY L MARTINEK (2011). "The Small Group Context. Designated District Court Judges in the US Courts of Appeals." Journal of Empirical Legal Studies **8**(1): 177-205.

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

DALAL, DEV K, KEVIN P NOLAN AND LAUREN E GANNON (2017). "Are Pre-assembly Shared Work Experiences Useful for Temporary-team Assembly Decisions? A Study of Olympic Ice Hockey Team Composition." Journal of Business and Psychology **32**(5): 561-574.

EDMONDSON, AMY (1999). "Psychological Safety and Learning Behavior in Work Teams." Administrative Science Quarterly **44**(2): 350-383.

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 KEREN WEINSHALL MARGEL (2020). "Manna from Heaven for Judges. Judges' Reaction to a Quasi-Random Reduction in Caseload." Journal of Empirical Legal Studies **17**: 722-751.

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.

ESPINOSA, J ALBERTO, SANDRA A SLAUGHTER, ROBERT E KRAUT AND JAMES D HERBSLEB (2007). "Familiarity, Complexity, and Team Performance in Geographically Distributed Software Development." Organization Science **18**(4): 613-630.

FARAJ, SAMER AND LEE SPROULL (2000). "Coordinating Expertise in Software Development Teams." Management Science **46**(12): 1554-1568.

GOODMAN, PAUL S AND STEVEN GARBER (1988). "Absenteeism and Accidents in a Dangerous Environment. Empirical Analysis of Underground Coal Mines." Journal of Applied Psychology **73**(1): 81-86.

GOODMAN, PAUL S AND DENNIS P LEYDEN (1991). "Familiarity and Group Productivity." Journal of Applied Psychology **76**(4): 578-586.

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.

GRUENFELD, DEBORAH H, ELIZABETH A MANNIX, KATHERINE Y WILLIAMS AND MARGARET A NEALE (1996). "Group Composition and Decision Making. How Member Familiarity and Information Distribution Affect Process and Performance." Organizational Behavior and Human Decision Processes **67**(1): 1-15.

HANFT, TRACEY ROCKETT (2002). Familiarity in Organizations, The University of Texas at Dallas.

HARRISON, DAVID A, SUSAN MOHAMMED, JOSEPH E MCGRATH, ANNA T FLOREY AND SCOTT W VANDERSTOEP (2003). "Time Matters in Team Performance. Effects of Member Familiarity, Entrainment, and Task Discontinuity on Speed and Quality." Personnel Psychology **56**(3): 633-669.

HARRISON, DAVID A, KENNETH H PRICE AND MYRTLE P BELL (1998). "Beyond Relational Demography. Time and the Effects of Surface-and Deep-level Diversity on Work Group Cohesion." Academy of Management Journal **41**(1): 96-107.

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.

HINDS, PAMELA J, KATHLEEN M CARLEY, DAVID KRACKHARDT AND DOUG WHOLEY (2000). "Choosing Work Group Members. Balancing Similarity, Competence, and Familiarity." Organizational Behavior and Human Decision Processes **81**(2): 226-251.

HINKLE, RACHAEL K, MICHAEL J NELSON AND MORGAN LW HAZELTON (2022). "How Interpersonal Contact Affects Appellate Review." Journal of Politics *******: ***.

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.

HOLLENBECK, JOHN R, DANIEL R ILGEN, DOUGLAS J SEGO, JENNIFER HEDLUND, DEBRA A MAJOR AND JEAN PHILLIPS (1995). "Multilevel Theory of Team Decision Making. Decision Performance in Teams Incorporating Distributed Expertise." Journal of Applied Psychology **80**(2): 292-316.

HUCKMAN, ROBERT S AND BRADLEY R STAATS (2011). "Fluid Tasks and Fluid Teams. The Impact of Diversity in Experience and Team Familiarity on Team Performance." Manufacturing & Service Operations Management **13**(3): 310-328.

HUCKMAN, ROBERT S, BRADLEY R STAATS AND DAVID M UPTON (2009). "Team Familiarity, Role Experience, and Performance. Evidence from Indian Software Services." Management Science **55**(1): 85-100.

JANSSEN, JEROEN, GIJSBERT ERKENS, PAUL A KIRSCHNER AND GELLOF KANSELAAR (2009). "Influence of Group Member Familiarity on Online Collaborative Learning." Computers in Human Behavior **25**(1): 161-170.

JARVENPAA, SIRKKA L, THOMAS R SHAW AND D SANDY STAPLES (2004). "Toward Contextualized Theories of Trust. The Role of Trust in Global Virtual Teams." Information Systems Research **15**(3): 250-267.

JEHN, KAREN A AND ELIZABETH A MANNIX (2001). "The Dynamic Nature of Conflict. A Longitudinal Study of Intragroup Conflict and Group Performance." Academy of Management Journal **44**(2): 238-251.

JEHN, KAREN A AND PRITI PRADHAN SHAH (1997). "Interpersonal Relationships and Task Performance. An Examination of Mediation Processes in Friendship and Acquaintance Groups." Journal of Personality and Social Psychology **72**(4): 775-790.

JONES, GARETH R AND JENNIFER M GEORGE (1998). "The Experience and Evolution of Trust. Implications for Cooperation and Teamwork." Academy of Management Review **23**(3): 531-546.

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

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.

KILLUMETS, ELAR, LAUREN D'INNOCENZO, M TRAVIS MAYNARD AND JOHN E MATHIEU (2015). "A Multilevel Examination of the Impact of Team Interpersonal Processes." Small Group Research **46**(2): 227-259.

KIM, PETER H (1997). "When What You Know Can Hurt You. A Study of Experiential Effects on Group Discussion and Performance." Organizational Behavior and Human Decision Processes **69**(2): 165-177.

KURMANN, ANITA, S KELLER, F TSCHAN-SEMMER, J SEELANDT, NK SEMMER, DANIEL CANDINAS AND GUIDO BELDI (2014). "Impact of Team Familiarity in the Operating Room on Surgical Complications." World Journal of Surgery **38**(12): 3047-3052.

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.

LEVY, MARIN K (2019). "Visiting Judges." California Law Review **107**: 67-140.

LEWIS, KYLE (2003). "Measuring Transactive Memory Systems in the Field. Scale Development and Validation." Journal of Applied Psychology **88**(4): 587-604.

LEWIS, KYLE, DONALD LANGE AND LYNETTE GILLIS (2005). "Transactive Memory Systems, Learning, and Learning Transfer." Organization Science **16**(6): 581-598.

LIANG, DIANE WEI, RICHARD MORELAND AND LINDA ARGOTE (1995). "Group Versus Individual Training and Group Performance. The Mediating Role of Transactive Memory." Personality and Social Psychology Bulletin **21**(4): 384-393.

LITTLEPAGE, GLENN, WILLIAM ROBISON AND KELLY REDDINGTON (1997). "Effects of Task Experience and Group Experience on Group Performance, Member Ability, and Recognition of Expertise." Organizational Behavior and Human Decision Processes **69**(2): 133-147.

MAK, MAXWELL, ANDREW H SIDMAN, VINCENT PALMERI, NICO DENISE AND RUBEN HUERTERO (2021). "Judges' Race and the Voting Rights Act. Perceived Expertise in Three-Judge District Court Panels." Justice System Journal **42**: 375-393.

MARLOW, SHANNON L, CHRISTINA N LACERENZA, JENSINE PAOLETTI, C SHAWN BURKE AND EDUARDO SALAS (2018). "Does Team Communication Represent a One-size-fits-all Approach?:A Meta-analysis of Team Communication and Performance." Organizational Behavior and Human Decision Processes **144**: 145-170.

MASON, WINTER AND AARON CLAUSET (2013). Friends ftw! Friendship and Competition in Halo: Reach. Proceedings of the 2013 Conference on Computer Supported Cooperative Work.

MAYNARD, M TRAVIS, JOHN E MATHIEU, LUCY L GILSON, DIANA R. SANCHEZ AND MATTHEW D DEAN (2019). "Do I Really Know You and Does it Matter? Unpacking the Relationship Between Familiarity and Information Elaboration in Global Virtual Teams." Group & Organization Management **44**(1): 3-37.

MCGRATH, JOSEPH E AND JANICE R KELLY (1986). Time and Human Interaction. Toward a Social Psychology of Time, Guilford Press.

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.

MOHAMMED, SUSAN AND BRAD C DUMVILLE (2001). "Team Mental Models in a Team Knowledge Framework. Expanding Theory and Measurement Across Disciplinary Boundaries." Journal of Organizational Behavior **22**(2): 89-106.

MOORE, SHANNON M AND MICHAEL N GEUSS (2020). "Familiarity with Teammate's Attitudes Improves Team Performance in Virtual Reality." PloS one **15**(10): e0241011.

NARAYANAN, SRIRAM, SRIDHAR BALASUBRAMANIAN AND JAYASHANKAR M SWAMINATHAN (2011). "Managing Outsourced Software Projects. An Analysis of Project Performance and Customer Satisfaction." Production and Operations Management **20**(4): 508-521.

NEWMAN, ALEXANDER, ROSS DONOHUE AND NATHAN EVA (2017). "Psychological Safety. A Systematic Review of the Literature." Human Resource Management Review **27**(3): 521-535.

OKHUYSEN, GERARDO ANDRÉS (2001). "Structuring Change. Familiarity and Formal Interventions in Problem-solving Groups." Academy of Management Journal **44**(4): 794-808.

REAGANS, RAY, LINDA ARGOTE AND DARIA BROOKS (2005). "Individual Experience and Experience Working Together. Predicting Learning Rates from Knowing Who Knows What and Knowing How to Work Together." Management Science **51**(6): 869-881.

ROCKETT, TRACEY L AND GERARDO A OKHUYSEN (2002). "Familiarity in Groups. Exploring the Relationship between Inter-member Familiarity and Group Behavior." Toward phenomenology of groups and group membership **4**: 173-201.

SANBONMATSU, DAVID M, BERT N UCHINO AND WENDY BIRMINGHAM (2011). "On the Importance of knowing your Partner's Views. Attitude Familiarity is Associated with Better Interpersonal Functioning and Lower Ambulatory Blood Pressure in Daily Life." Annals of Behavioral Medicine **41**(1): 131-137.

SANBONMATSU, DAVID M, BERT N UCHINO, KEVIN K WONG AND JOON YONG SEO (2012). "Getting Along Better. The Role of Attitude Familiarity in Relationship Functioning." Social Cognition **30**(3): 350-361.

SAPHIRE, RICHARD B AND MICHAEL E SOLIMINE (1994). "Diluting Justice on Appeal. An Examination of the Use of District Court Judges Sitting by Designation on the United States Courts of Appeals." University of Michigan Journal of Law Reform **28**: 351-408.

SHAH, PRI PRADHAN AND KAREN A JEHN (1993). "Do Friends Perform Better than Acquaintances? The Interaction of Friendship, Conflict, and Task." Group Decision and Negotiation **2**(2): 149-165.

SIEMSEN, ENNO, ALEDA V ROTH, SRIDHAR BALASUBRAMANIAN AND GOPESH ANAND (2009). "The Influence of Psychological Safety and Confidence in Knowledge on Employee Knowledge Sharing." Manufacturing & Service Operations Management **11**(3): 429-447.

SMITH-JENTSCH, KIMBERLY A, KURT KRAIGER, JANIS A CANNON-BOWERS AND EDUARDO SALAS (2009). "Do Familiar Teammates Request and Accept More Backup? Transactive Memory in Air Traffic Control." Human Factors **51**(2): 181-192.

SOLIMINE, MICHAEL E (1996). "The Three-Judge District Court in Voting Rights Litigation." University of Michigan Journal of Law Reform **30**: 79-146.

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.

SOSA, MANUEL E (2011). "Where do Creative Interactions Come from? The Role of Tie Content and Social Networks." Organization Science **22**(1): 1-21.

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.

STAATS, BRADLEY R (2012). "Unpacking Team Familiarity. The Effects of Geographic Location and Hierarchical Role." Production and Operations Management **21**(3): 619-635.

SWALVE, TILKO (2022). "Does Group Familiarity Improve Deliberations in Judicial Teams? Evidence from the German Federal Court of Justice." Journal of Empirical Legal Studies **19**(1): 223-249.

TUCKER, ANITA L (2007). "An Empirical Study of System Improvement by Frontline Employees in Hospital Units." Manufacturing & Service Operations Management **9**(4): 492-505.

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

WASBY, STEPHEN L (1980). "Extra Judges in a Federal Appellate Court. The Ninth Circuit." Law and Society Review **15**: 369-388.

WASBY, STEPHEN L (2018). Borrowed Judges. Visitors in the U.S. Courts of Appeals.

WENDEL, LUISA (2020). "Welche Grundrechte führen zum Erfolg? Eine quantitative, korpusgestützte Untersuchung anhand von Entscheidungen des Bundesverfassungsgerichts." JuristenZeitung **75**(13): 668-679.

WETMORE III, WILLIAM R, JOSHUA DAVID SUMMERS AND JOEL S GREENSTEIN (2010). "Experimental Study of Influence of Group Familiarity and Information Sharing on Design Review Effectiveness." Journal of Engineering Design **21**(1): 111-126.

YOON, CHEOLHO AND ERIK ROLLAND (2012). "Knowledge-sharing in Virtual Communities. Familiarity, Anonymity and Self-determination Theory." Behaviour & Information Technology **31**(11): 1133-1143.
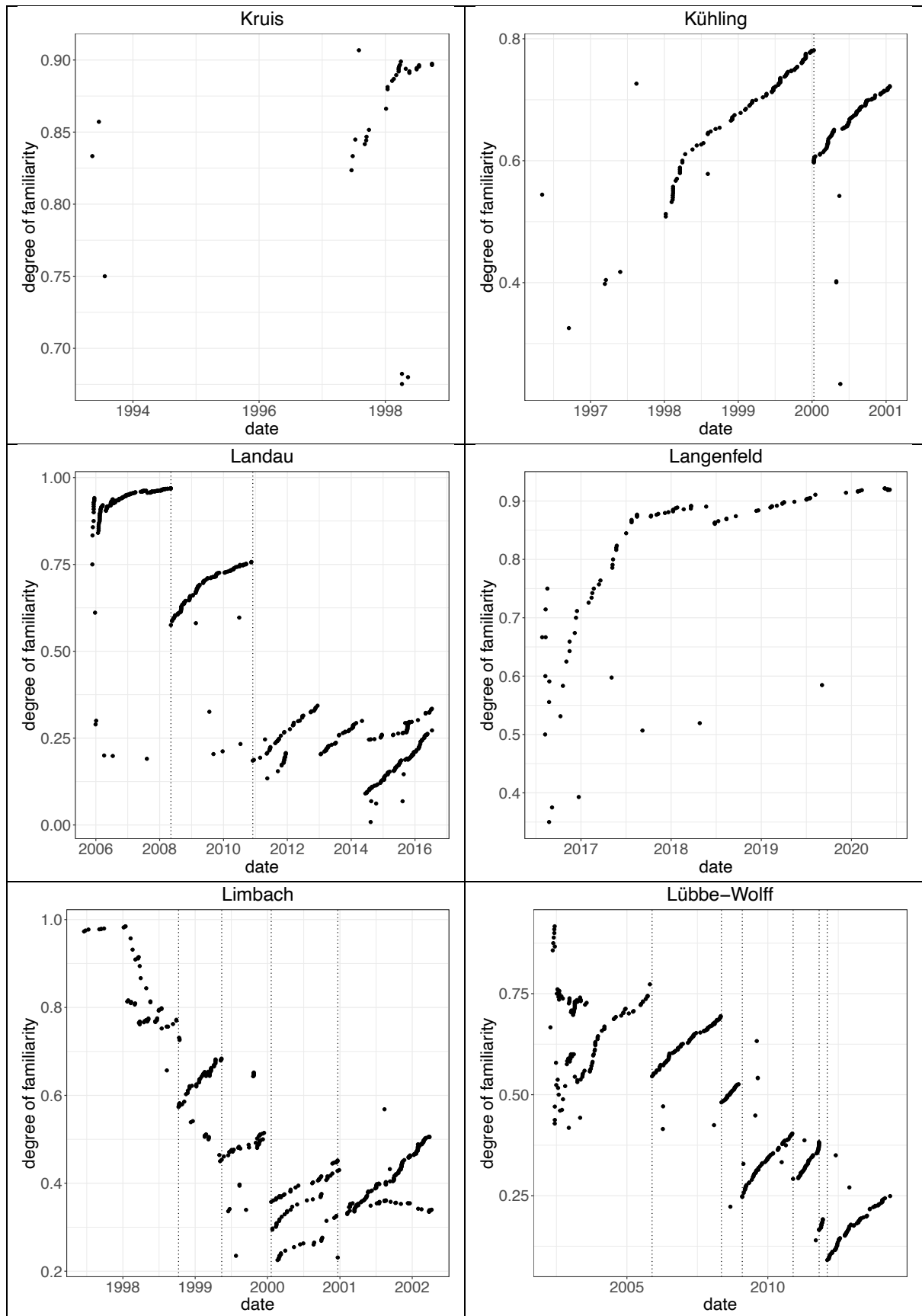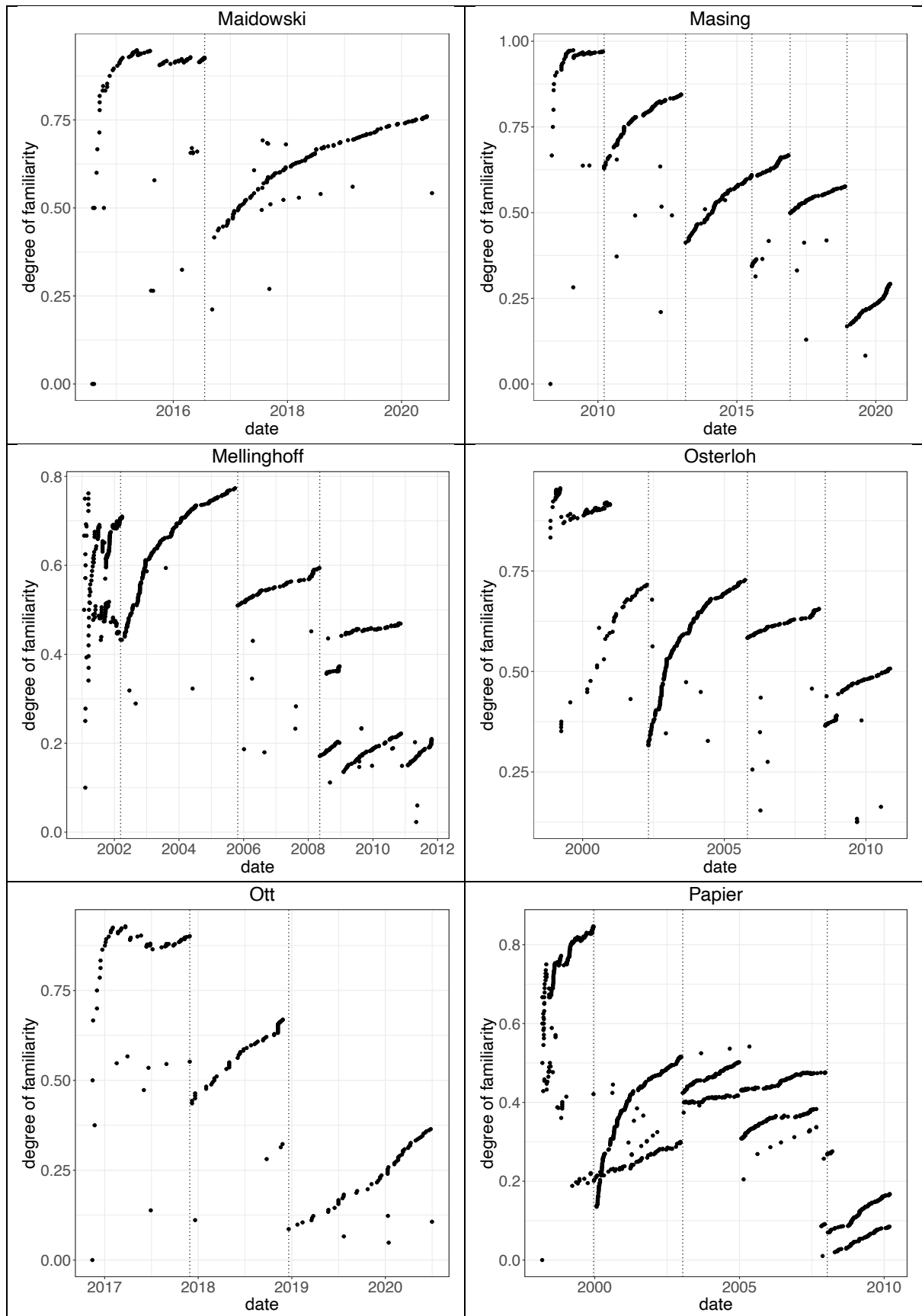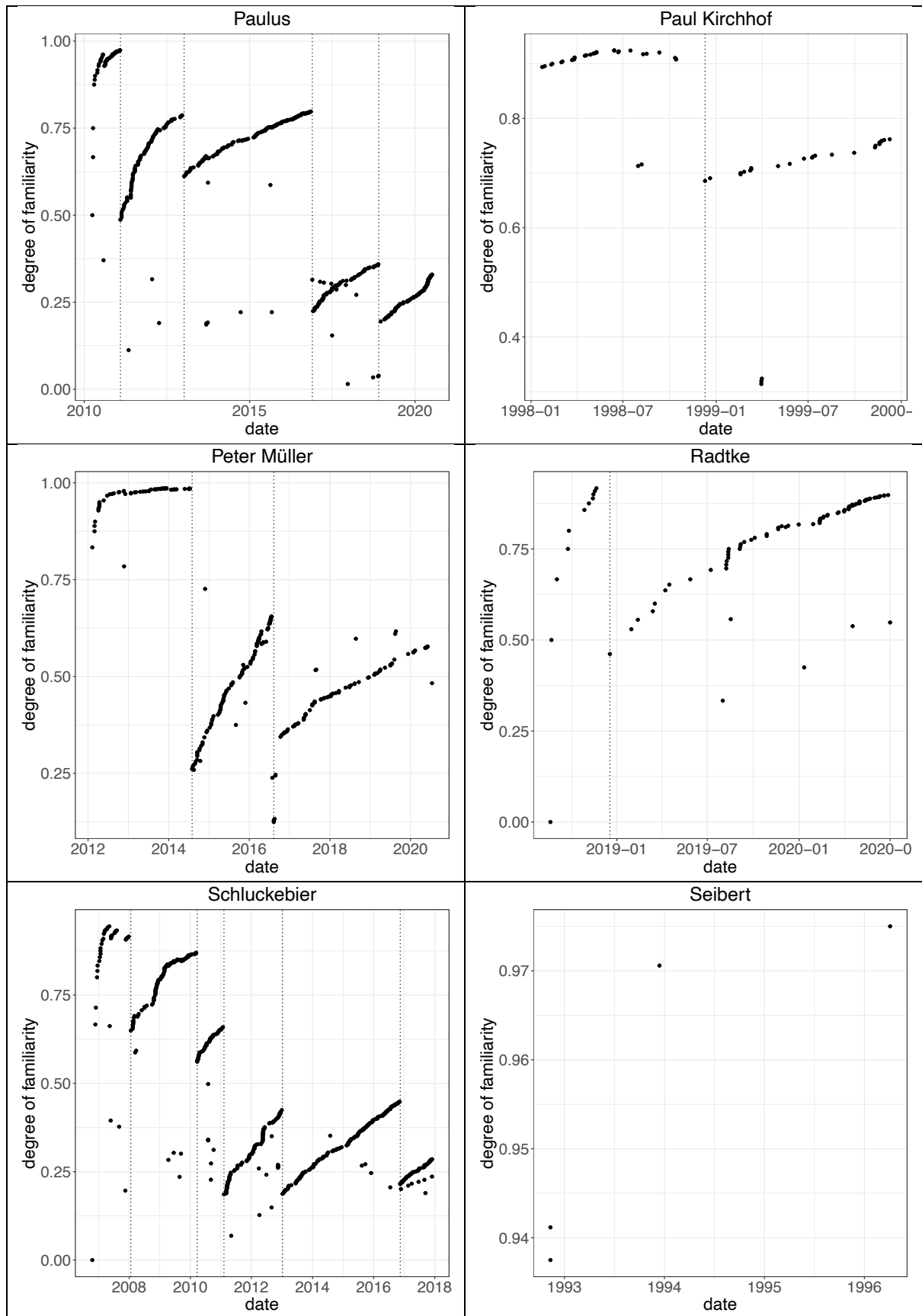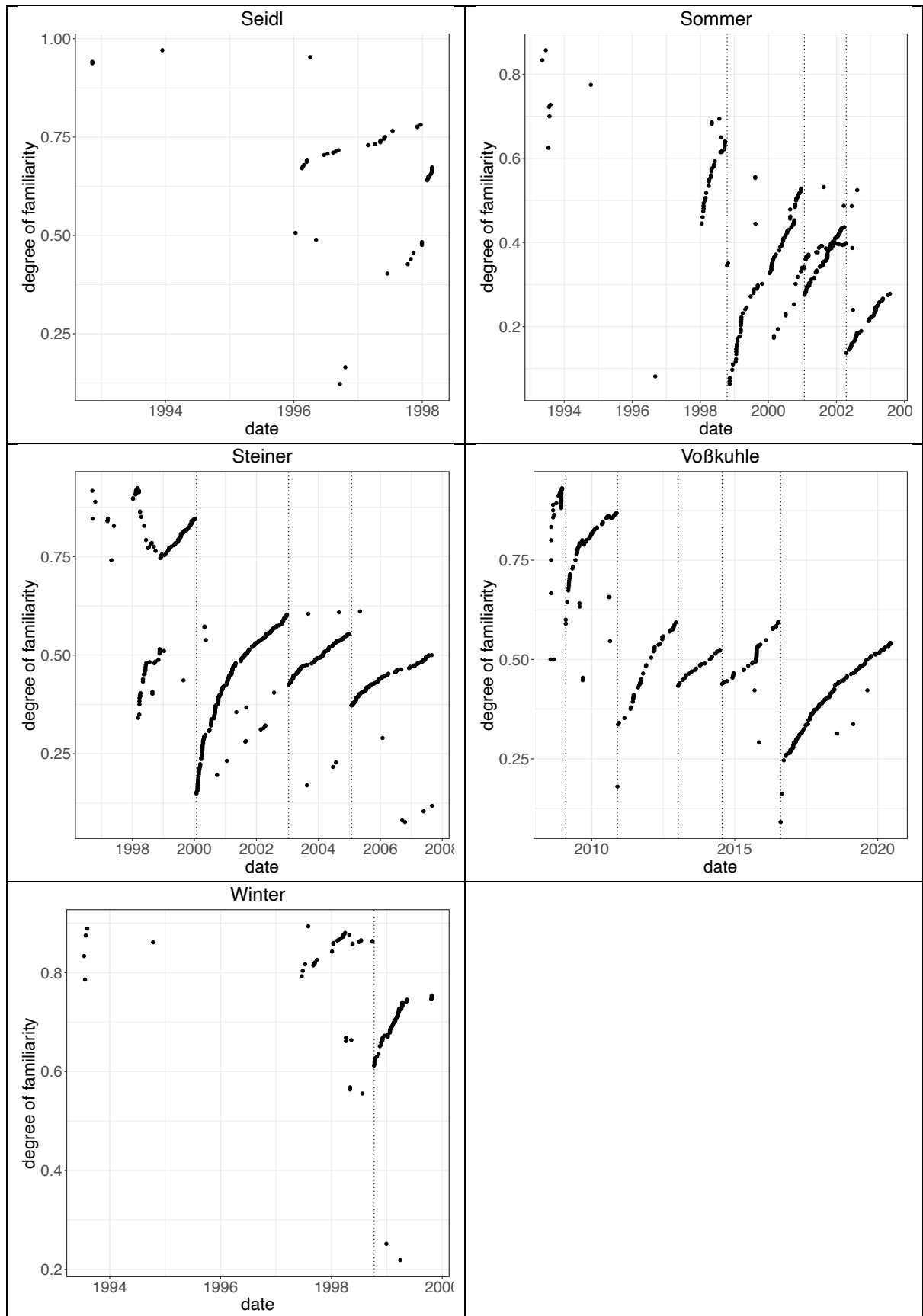
**Appendix
A1: Breakpoints**
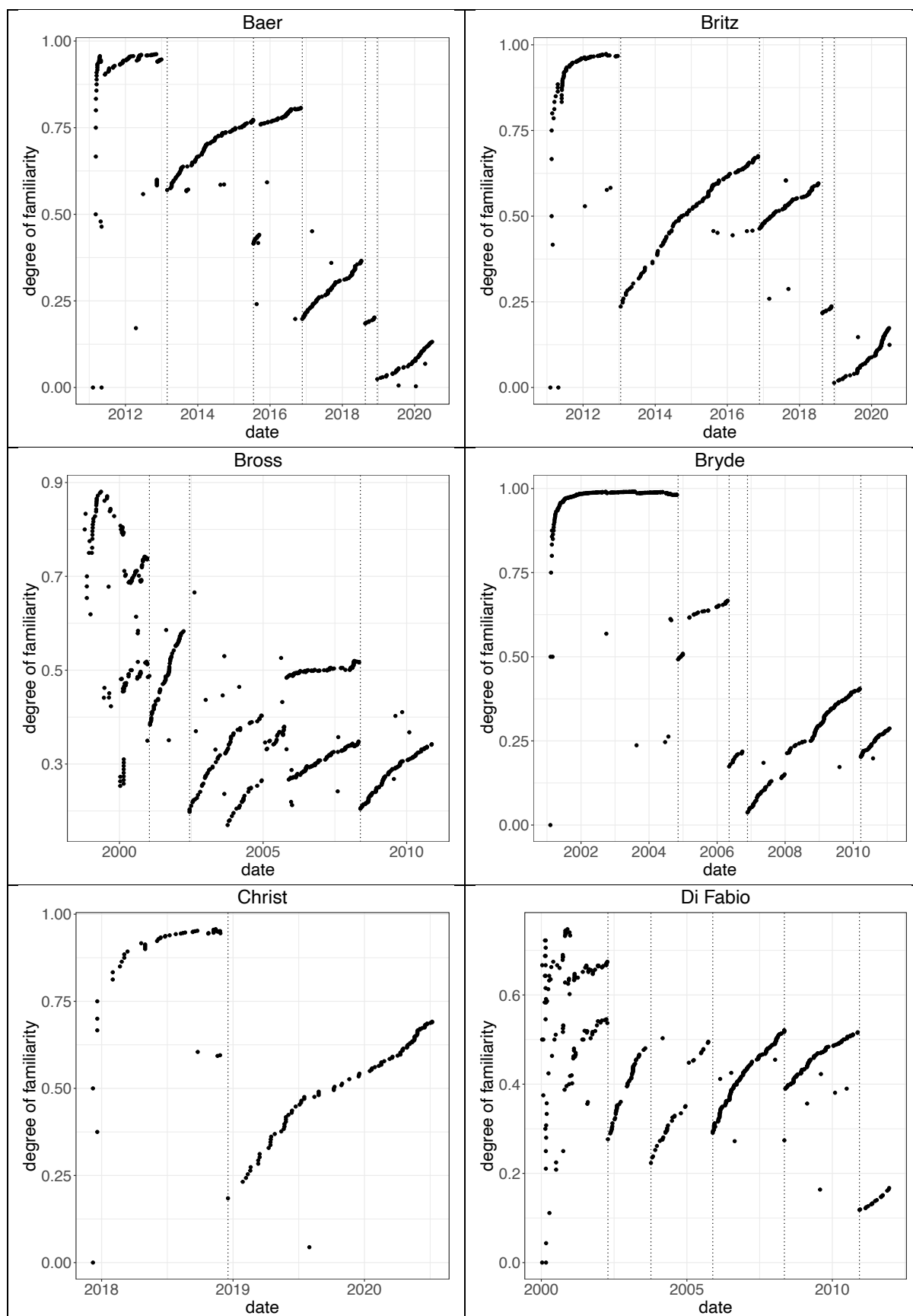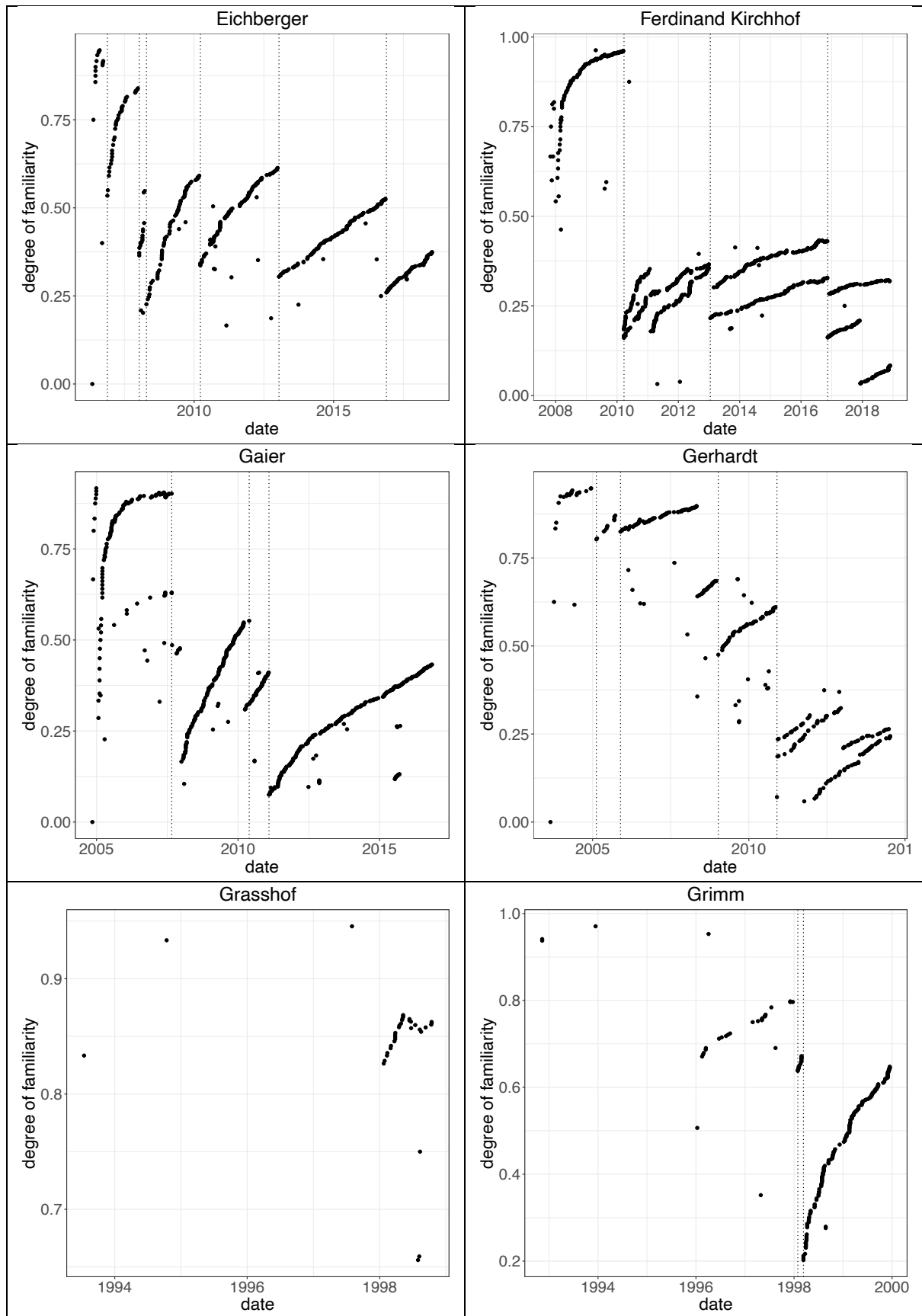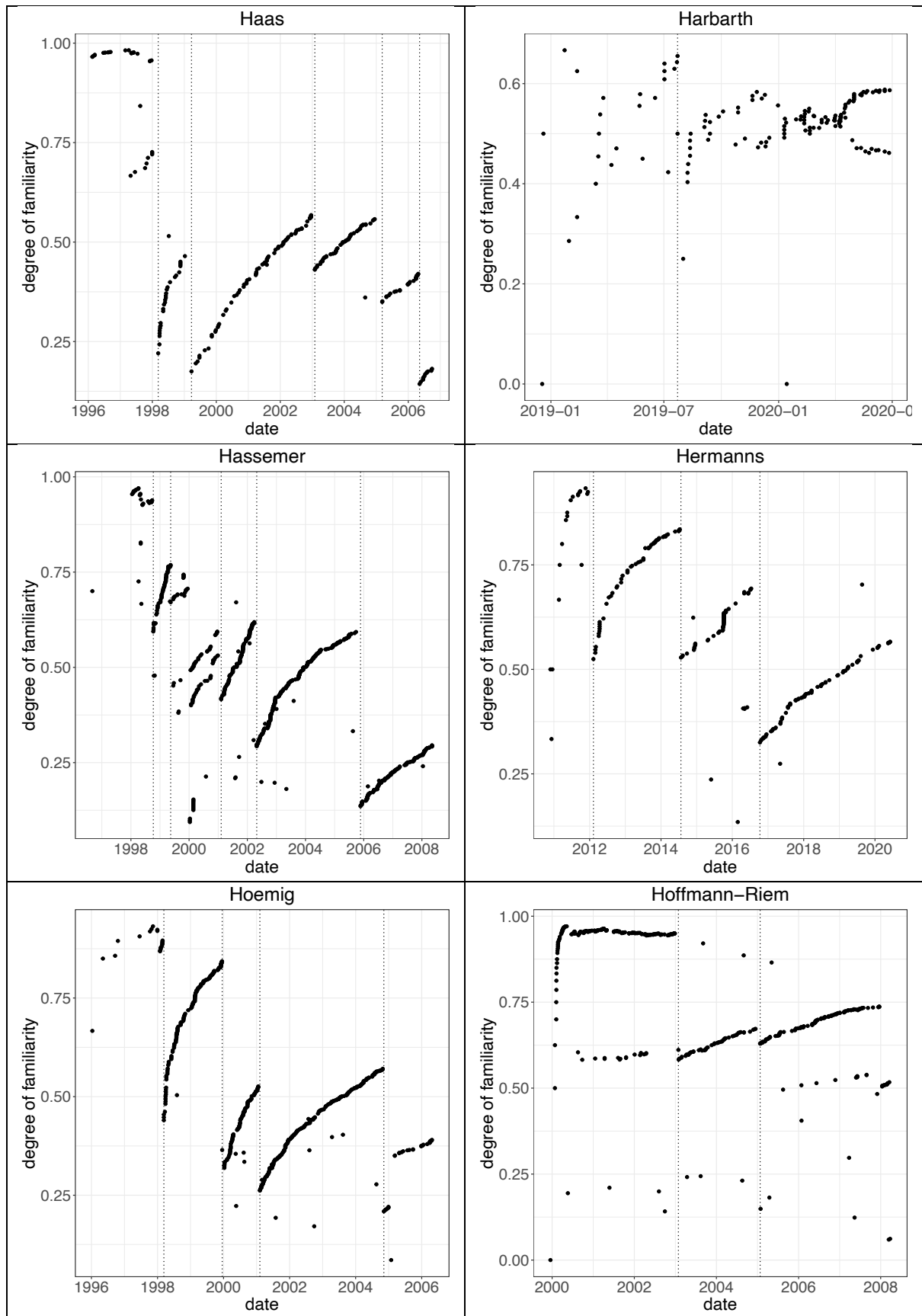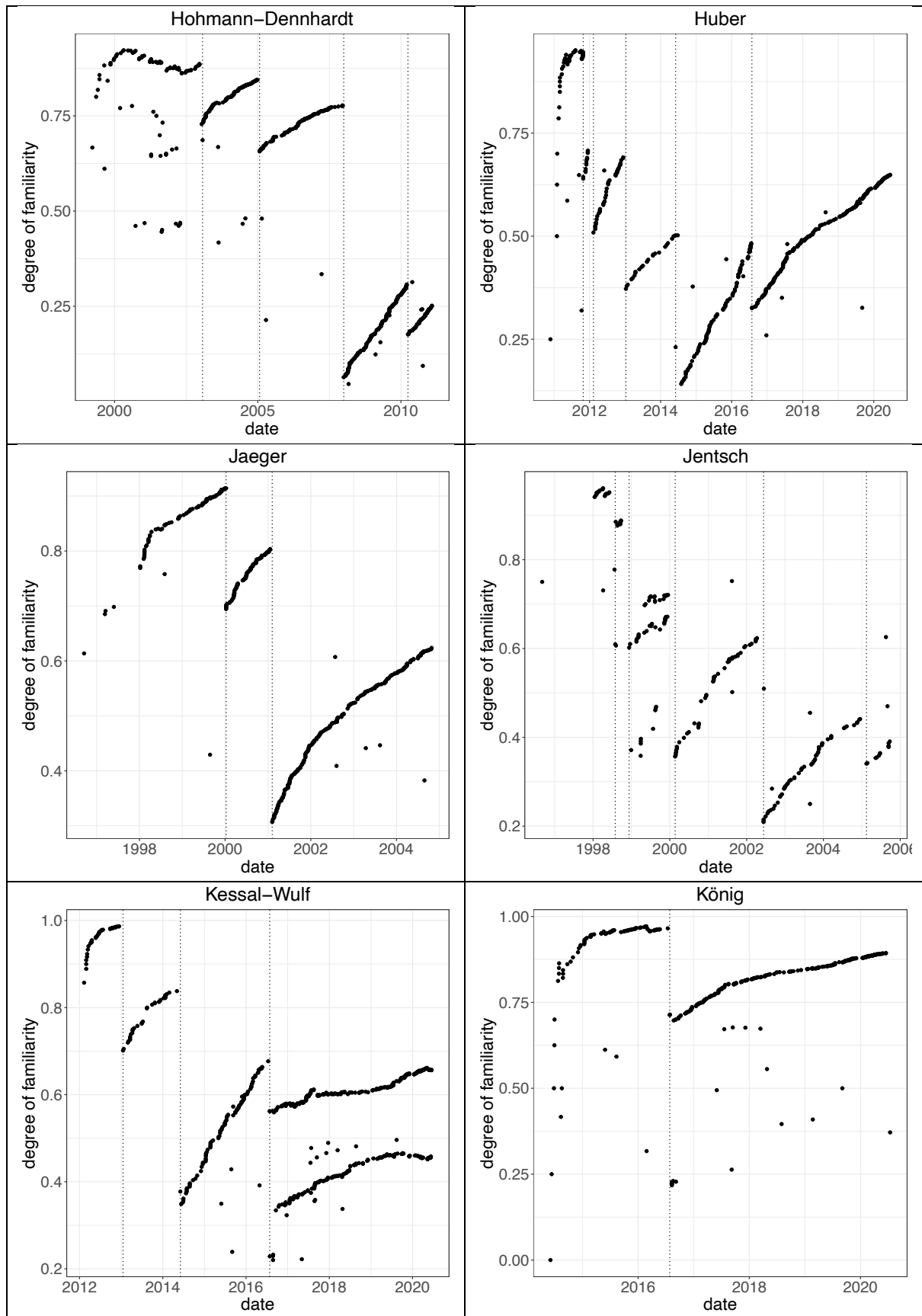
Figure 5
Breakpoints per Justice
blue dots: senate, red dots: chamber

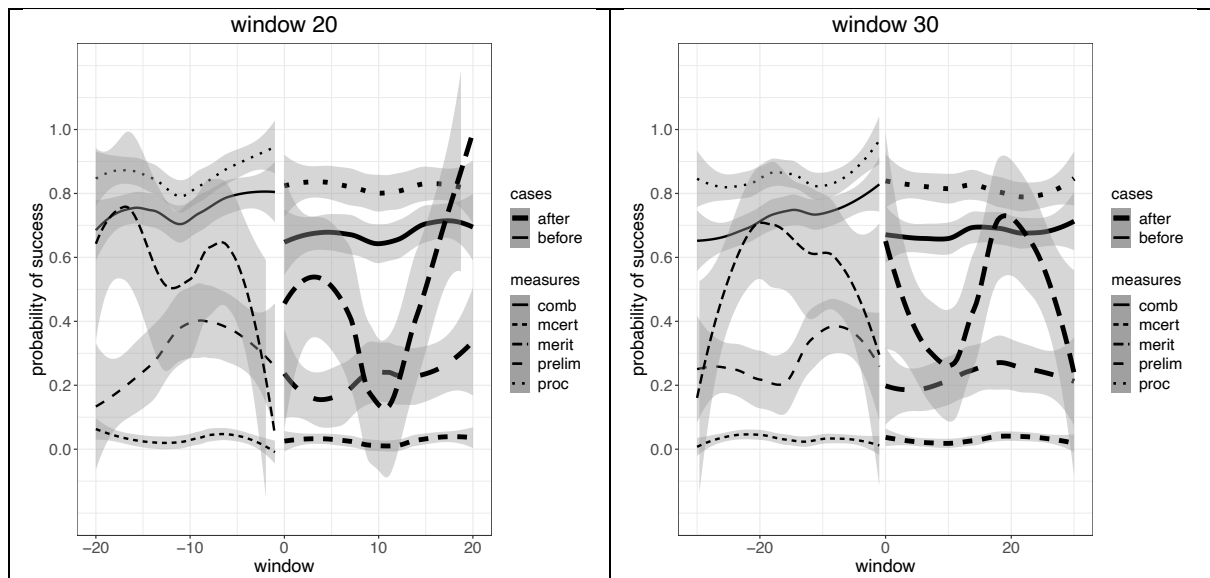dotted lines: breakpoints used for the analysis

A2: Wider Windows

Figure 6
Regression Discontinuity with Wider Windows
decisions normalized about the breakpoints (see Appendix A1 and Section 4 for the definition of breakpoints)
window: 20/30 decisions before and after the breakpoint
success variables: merit: on the merits; mcert: merit, including the decision not summarily reject the complaint as inadmissible or
obviously unfounded; prelim: granting a preliminary ruling; proc: granting a procedural request; comb: succeed either on the merits, or
with a request for a preliminary injunction, or with a procedural request
all five success variables are dummies. Hence on the y-axis, fractions are reported
duplicate breakpoints removed
results are reported with 95% confidence intervals, using R `ggplot2 geom_smooth,` which employs a cubic spline for smoothing

	window 20	window 30
merit		
after	-.1636[+] (.0863)	-.1135[+] (.0669)
cons	.5893**** (.0711)	.5715*** (.0614)
N	134	210
merit + certiorari		
after	-.0042 (.0070)	-.0030 (.0058)
cons	.0342*** (.0075)	.0345*** (.0062)
N	2346	3503
preliminary		
after	-.0720[+] (.0383)	-.0437 (.0317)
cons	.3090*** (.0364)	
N	501	747
procedural		
after	-.0467* (.0193)	-.0297[+] (.0167)
cons	.8623*** (.0201)	
N	1349	1924
merit, preliminary or procedural		
after	-.0694** (.0215)	-.0389* (.0183)
cons		
N	1699	2434

Table 4

Regression Discontinuity with Wider Windows

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

*** $p < .001$, ** $p < .01$, * $p < .05$, [+] $p < .1$