

Scaling Lower Court Decisions*

CHRISTIAN ARNOLD[†]
Cardiff University

BENJAMIN G. ENGST[‡]
University of Mannheim

THOMAS GSCHWEND[§]
University of Mannheim

June 18, 2020

ABSTRACT

Not all judges necessarily take a decision in exactly the same way. Some legal decisions are quite ‘hard’, others are rather ‘soft’. To tell them apart, scholars and practitioners have an interest in measuring and comparing legal decisions within a well-defined case-space. So far, this required a close reading of each decision and hence a considerable human effort. The only area amenable to automation were courts with judges who not only vote on verdicts, but who also make these voting records available—such as for example the U.S. Supreme Court. This paper introduces a new measurement approach that does not rely to voting data, but uses the citations in written legal decisions instead. Modelling the frequency of how often a decision is citing a legal source, we estimate the latent positions of both, decisions and sources in a joint case-space. We showcase our model in the context of *forum shopping* and *forum selling* in a particularly unlikely case, Germany’s *Landgerichte* (lower courts). A method for easily and automatically measuring decisions in a joint case-space is not only a mere academic exercise, but is instrumental to anyone who studies legal decision making at large scale.

*NOTE: Previous versions of this paper have been presented at the *Second Conference on Empirical Legal Studies in Europe*, Leuven, 2018 and at the *Conference on Data Science and Law*, Zurich, 2019. *JURIS* generously supported this research by allowing access to their data. We thank Pablo Barberá, Jens Frankenreiter, J. Andrew Harris, Kevin Quinn, Nils Schaks and Philip Schroeder for helpful comments. We also thank Felicia Riethmüller for her insightful and instrumental research assistance. Arnold is grateful for support through Cardiff’s “Darlithwyr Disglair Development Scheme”. Gschwend and Engst acknowledge financial support for SFB 884 (project C4) at the University of Mannheim from the *German Research Foundation* (DFG). All data and computer code necessary to replicate the results in this analysis will be made publicly available on our webpages on publication.

[†]Lecturer, School of Law and Politics, Cardiff University, Museum Avenue Cardiff, CF10 3AX. (arnoldc6@cardiff.ac.uk).

[‡]PostDoc, Collaborative Research Center 884, University of Mannheim, B6, 30-32, D-68131, Mannheim, Germany (engst@uni-mannheim.de).

[§]Professor, Department of Political Science, University of Mannheim, A5, 6, D-68131, Mannheim, Germany (gschwend@uni-mannheim.de).

1 The Promise of Scaling Lower Court Decisions

How similar or how different is the legal reasoning in comparable legal decisions? The answer to this question is a key concern to law practitioners and law scholars alike. Some might want to quickly identify particularly controversial decisions. Others seek to get a more comprehensive view on a large set of decisions—for example to understand patterns across different courts or the development of a legal doctrine over time. Existing approaches for comparing decisions at large scale tackle the issue mostly on the basis of voting behavior of the respective judges in a court (e.g. [Clark and Lauderdale, 2010](#)). It turns out, however, that in cross-national comparison judges do not always take a vote—and if they do, courts only rarely publish the individual votes ([Kelemen, 2013](#); [Raffaelli, 2012](#)). Thus, only courts in some countries qualify for an analysis on the basis of this approach (e.g. [Dalla Pellegrina, Garoupa and Gili, 2020](#)).

We introduce a scaling model that estimates the location of legal decisions in a common case-space. Instead of published votes, we use information about the legal sources cited in decisions. Citations are commonly available at all levels in every legal system because judges need these references to justify their reasoning. Given the ubiquity of citation data, our approach lends itself to a broad range of new empirical settings.

Our model estimates that a decision i is citing a particular legal source j . We assume that the closer the decision i to a legal source j in a common case-space—and hence the more amenable a source to the legal reasoning in the decision—the more likely are judges to cite this legal source and the more often they do so. When implementing the model, we first pre-select decisions that actually can be compared on legal grounds. While curating a concise set of legal documents can of course be done manually, we show how to help scale the human effort with easy-to-deploy algorithms from information retrieval. With our model and the algorithm at hand, we then rise to a formidable challenge. Accessing the legal data base *Juris*,¹ we investigate systematic tendencies in

¹Perhaps surprisingly, court rulings in Germany are not freely available. *Juris* GmbH is a publishing company that provides a database of legal documents and information on which we rely on in this paper. The *JURIS* data is very well suited for our endeavor, because it comprises a complete and already digitized corpus of written rulings of all available German court decisions. The annotation of the corpus with meta data makes this data even more valuable. In short, the extraordinary data quality allows us to develop state-of-the-art measurement models to tap into this so far

judicial decision making in an unlikely case. Germany has a civil law system that is renowned for a particularly impartial and objective way of creating legal decisions (Langbein, 1985). However, a recent debate on ‘forum-selling’ and ‘forum-shopping’ at Germany’s *Landgerichte* (lower courts) suggests that there are areas where courts are systematically biased (Klerman and Reilly, 2016; Bechtold, Frankenreiter and Klerman, 2019). Using our model to study cases from press law and antitrust, we indeed corroborate these findings at Germany’s lower courts in Köln and Hamburg for press law, and Dortmund and Mannheim for antitrust.

We contribute to the literature in three ways. First, we introduce a new model that allows scholars and legal practitioners to systematically compare decisions in the same case-space. Citation data is particularly useful because this information is widely available for courts in many countries and at all levels. In the context of estimating this model, we also propose a new solution for model identification and solve rotational invariance in the light of *ex-ante* ignorance about the location of decisions or sources. Second, we showcase how to scale measurement efforts when analyzing decisions from a large legal corpus of lower court decisions. Tools from information retrieval allow to easily identify decisions that lie in the same case-space and that actually can be compared to one another. Third, the substantive application of our new scaling model informs recent debates on ‘forum-selling’ and ‘forum-shopping’ and leads the way to investigating the issues quantitatively and at large scale.

Capable of easily mapping so far uncharted areas of law, our automated approach opens the door to a whole range of new applications. Scholars and practitioners—independent of their background in law, business, economics, social sciences or politics—may all want to understand whether systematic tendencies in judicial decision making exist at large scale, where they are coming from, and how they effect other legal and political rule making. We are hopeful that our model proves to be a fine addition to their analytical toolbox.

The paper proceeds as follows. We begin with surveying existing approaches capable of mapping legal decisions in a case-space. The subsequent section explains why citations leave a

unused data source.

trace about a judge’s tendency in rule making. We then translate this reasoning into a statistical model, and finally study ‘forum-selling’ in Germany’s lower courts in press law and antitrust. A final section concludes.

2 Measuring the Content of Legal Decisions

Comparing legal decisions to one another implicitly accepts the notion that the decisions reside on a continuum: very strict decisions on one end, very lax decisions on the other, and all others somewhere in between. This spatial notion has been formalized into an analytical framework that was originally intended to study the influence of politics in high courts such as the U.S. Supreme Court (for an overview e.g. [Clark and Lauderdale, 2010](#); [Lax, 2011](#)). But while the resulting case-space model ([Kornhauser, 1992a,b](#); [Landa and Lax, 2008](#); [Lax, 2011, 2012](#)) offers exciting analytical possibilities, one of its most significant limitations is an empirical one. So far there are no tools to easily and reliably map a large number of real-world cases into such a case-space.

Current empirical strategies offer room for improvement. Distinguishing cases on the basis of a court’s verdict (e.g. guilty-innocent, affirm-reverse) merely allows to tell on which side of a judge’s threshold the case falls—but not where on the spectrum this legal decision is actually located. Broad ideological categories (e.g., liberal-conservative) are also of limited use, since they fail to honor the respective legal context. Nevertheless, some equate the latent political—not the legal—position of a politically nominated (median) judge with a judge’s resulting legal decisions (e.g. [Brouard and Hönnige, 2017](#); [Carrubba et al., 2012](#); [Hönnige, 2007](#); [Sternberg et al., 2015](#)). The judge’s displayed political position might not even be her own, but can be “inherited” from other actors who nominated the (median) judge such as parties, senators or presidents (e.g. [Epstein et al., 2007](#); [Hönnige, 2009](#)). An alternative approach is to closely analyze the decisions’ content. Law scholars embraced content analysis methods a while ago ([Hall and Wright, 2008](#)). However, when hand-coding potentially a large number of written decisions, reliability can indeed be an issue. In addition, since the task is labour intensive, it is hard to scale—particularly so when time

and resources are scarce.

Computers are the analyst's best friend and there are various ways in which humans can make the most of machines to support them locating decisions in a case-space at large scale. Of course, the words of a decision carry important parts of its meaning. Existing approaches pay close attention to studying the language of legal decisions and have been exploring supervised and unsupervised text-scaling methods (Evans et al., 2007; McGuire and Vanberg, 2005; Dyevre, 2019). In the light of the specific requirements of legal terminology, further research in this direction promises important progress. Another option is to leverage the legal sources judges cite in their decisions (Whalen, 2016). The resulting citation networks allow for example to uncover the relevance of a legal decision (Fowler et al., 2007; Winkels, de Ruyter and Kroese, 2011). Citation patterns have also been used to show that judges chose legal sources neither randomly nor independently of their preferences. Frankenreiter (2017) exploits the institutional setting at the ECJ where—unlike in other courts—two opinions are drafted, one by the *Advocate General* and another one by the *Judge Rapporteur*. He finds that judges tend to cite decisions of judges appointed by Member State governments that have similar preferences regarding European integration. Finally, we are not the first to use citation patterns for explicitly mapping decisions into a case space. Clark and Lauderdale (2010) analyze search and seizure cases and freedom of religion opinions before the US Supreme Court between 1953 and 2006. But while Clark and Lauderdale have to consider the judges' voting behavior to estimate valid positions in the case-space, we show how to make do without such data. Given that in cross-national comparison highest courts rarely, or never, publish judges' votes (Kelemen, 2013; Raffaelli, 2012), we believe that our approach offers a welcome contribution to the existing toolkit for locating decisions in a case-space.

In short, current approaches that locate decisions in the case space too often work with crude proxies. There are unexplored potentials in going beyond the human effort of reading and coding each single decision. While first work exists that makes use of the information in citation patterns, there is yet no model that allows to map a legal decision in the case space in a convenient way.

3 Locating Decisions Using Citation Data

The sources a judge is citing in a decision are not only a technical requirement; citations may also reveal legal preferences. To ultimately understand how to tell a ‘hard’ from ‘soft’ decision and how to exactly quantify this statement we first take a closer look at the case-space framework. In a second step, we will then understand why different legal preferences express themselves in the citations of a legal decision and the consequences arising from it.

3.1 The Case Space

The case-space model comes with a number of core concepts (Cameron and Kornhauser, 2017a,b). A legal case can be defined as a “concrete, fact-ridden dispute between two (or perhaps more) parties (Cameron and Kornhauser, 2017a, 2)”. The court has to resolve the dispute—and it does so by applying the law as a rule to the facts of the case. All possible cases on the same subject matter reside on a one-dimensional case-space. For example, if the police catches car drivers at a certain velocity, it is possible to map all cases into this case-space on the basis of their speed. Now, for the judge to be able to take a legal decision, she requires a rule that tells her when the defendant has to be convicted. In the example, there needs to be a law that defines the maximum velocity. A driver below that threshold will not be convicted, a case with a driver above that speed will be. Finally, judges may hold own preferences next to the law. An environmentalist judge and a judge who is a sports-car *aficionado* might differ in what they personally consider an appropriate speed limit—irrespective of the law itself.

Turning facts into a disposition on the basis of a rule seems straightforward in the simplified speeding example. And if adjudicating were a simple matter of applying a rule to clear facts, an algorithm would be qualified enough to sit on the bench. The legal reality, however, is far more complex. Judges have to translate evidence into legal facts. They then need to select the appropriate rule and apply it to these legal facts. Typically, a judge has to evaluate legal facts with a rule on more than one single issue to arrive at a verdict, which is why judges ultimately need to weigh

and aggregate all issues when they finally speak the law. And given that this process requires interpretation and evaluation, the door is open to human bias—be it conscious or unconscious. Two judges facing the very same case therefore may come to different conclusions and this is increasingly likely the more complex the case.

Human bias leaves a trace in a decision. When judges evaluate facts on the basis of a rule—be it laws, existing precedent or legal scholarship—they have to provide reason for their judgement and need to cite the respective source. The set of citations in a decision are therefore indicative for a judges' overall legal doctrine (Choi and Gulati, 2008). Two judges who apply the same set of rules to assess the same legal facts will come to the same conclusion citing similar legal sources. If the judges decide a case differently, they have to apply different rules to the legal facts and hence will cite different legal sources. Therefore, the legal sources in a set of decisions of the same case-space hold the necessary information to locate the positions of the decisions in that case-space.

3.2 How Do Judges Cite?

If judges' legal preferences leave a trace in the citations, it should be possible to infer these preferences using these citations. Decisions that are similar—i.e. decisions that are in close proximity in a case-space—also use a similar argument structure. They cite similar legal sources and will do so with a similar frequency. In result, similar decisions display a typical combination of legal sources that substantiate a particular legal view. The legal sources alone already contain important information about the kind of legal argument that the judge advanced in a decision (Choi and Gulati, 2008; Clark and Lauderdale, 2012).

How do judges pick the legal sources they cite in a decision? Three different behavioral mechanisms drive what and how often a source gets cited in a legal decision. First, the substance matter dictates an authoritative core set of legal sources any judge would agree to refer to. Those core legal sources help defining the legal substance of what is at stake. In result, legal decisions in a certain domain display a similar set and a similar frequency of these core legal sources. Second,

there are idiosyncratic legal sources that appear in only one written decision. Judges may cite idiosyncratic legal sources to account for the particular characteristics of a case. Finally, a judge's citation choice reflects tendencies in her rulings, i.e. a particular hard or a particular soft stance on a case. A judge is more likely to refer to those legal sources that are in line with her reasoning when drafting a decision (Choi and Gulati, 2008). On the one hand, sources that are in line with the spirit of her own decision helps her making a legally sound argument, which is why she is much more likely to quote these benign legal sources. On the other hand, a judge tends to refrain from citing legal sources that depart from her own opinion. They generate cognitive dissonance and, eventually, cause extra effort, i.e. more text to write, when justifying her argument. She refers to dissenting legal sources—if at all—only in passing. In effect, a judge eager to present a cohesive legal argument more often cites legal sources that are in line with her argument than legal sources that are opposing or questioning her argument. Legal decisions on one side of a spectrum will resort to one particular set of legal sources. In contrast, legal decisions on the other side of a spectrum will cite those sources that help the respective judge make the argument in the other way. In short, similar legal decisions cite similar legal sources.

Among these three kinds of citations, only the latter one offers information regarding the position of a decision in the case-space. Core legal sources define the substance of a decision. When estimating the location of a decision, these core sources are less informative for the eventual location of the decision in a case-space. Next, idiosyncratic legal sources neither reveal information about the location of a decision nor the substantive identification of a particular case-space, simply because they cannot be compared to one another within the context of a citation network. Instead, information that helps identify the location of legal decisions stems from all other legal sources—that is from legal sources that are cited with different frequency by more than one decision. Decisions with similar legal reasoning refer to the same legal source in a similar way and therefore cause informative legal sources to cluster. Or, put differently, legal sources that can be used to substantiate one line of reasoning appear more often together in decisions that take this stance. In contrast, a judge who departs from a standard argumentation in her line of reason-

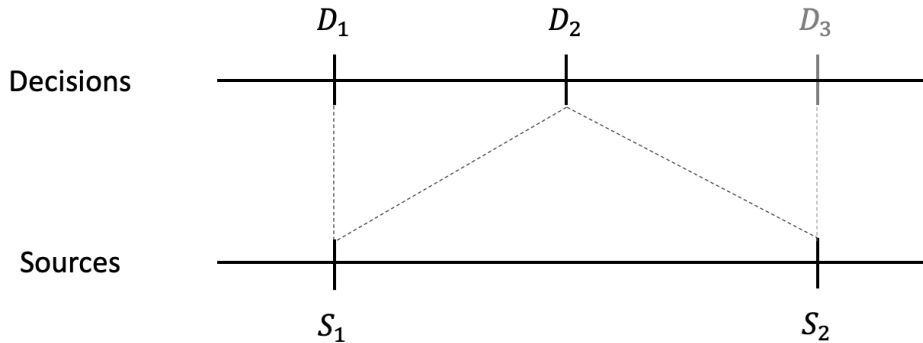


Figure 1: Decisions and Cited Sources in the Case Space. The Position of Decision D_2 Citing Source S_2 Is Pulled Away. Adding Decision D_3 Shifts the Size of the Decision Case-Space and the Relative Position of D_2 .

ing is most likely to bolster her argument with other legal sources—anything that somehow can substantiate the claims made in the decision that pushes this decision away from the well-known authoritative argument in land-mark decisions.

To illustrate, we introduce a simple visualisation in Figure 1. Suppose one decision D_1 cites legal sources S_1 several times and another decision D_2 cites legal sources S_1 and S_2 . We would conclude that D_2 should lie somewhere closer to the position of S_2 than D_1 . In our model, we will account for these co-occurrences to identify decision locations in a given case-space.²

3.3 Implications and Features

There are a number of implications when estimating decision locations using citation networks. First, all locations are relative to one another—which means that adding further decisions to or subtracting decisions from the existing set may change existing locations. Again, imagine the same legal decision D_1 citing the legal source S_1 and decision D_2 citing legal sources S_1 and S_2 . Suppose we add a third decision D_3 citing only legal source S_2 . We would conclude that the third decision D_3 mirrors the position of source S_2 and is hence further extreme than the legal decision D_2 . What effectively happened is that D_1 and D_2 do not represent the minimum and maximum

²This is comparable to scaling models in the context of roll-call votes in legislative politics (e.g. Clinton, Jackman and Rivers, 2004). There, ideologically similar legislators vote ‘yes’ (or ‘no’) on much the same roll-call votes. Those who only sometimes vote ‘yes’ or ‘no’ are identified to lie somewhere in between those two groups of legislators.

of the underlying case-space any longer. Instead, the extreme locations of the new space are now defined by D_1 and D_3 . If we define the decisions' case-space to have a standardized metric—for example z-scores—then the numerical value from D_2 would change to a more central position once the decision D_3 is included into the set of decisions that are being analyzed.

Are all citations necessarily indicative of a legal preference for one side of a legal argument or another? Not every legal source is cited in a way to support an argument. Judges also develop counter-arguments and cite to dispute a legal reasoning. Similar to the analysis of the use of precedent in decisions of the ECJ (Jacob, 2014), these negative citations are nevertheless used to bolster the credibility of the decision, demonstrating the argument's coherence or fending-off potential revisions from higher courts. Judges may distinguish their legal argument from existing ones in order to show why a certain legal belief is not binding or valid by entertaining a potential counterargument or introducing a more nuanced view. Positive and negative citations will not be mentioned with the same frequency, however. Given limited time and resources, judges are more likely to cite a legal source in line with her legal reasoning than a legal source that runs against it. If they do quote a legal source that mentions a counter-argument, judges will refer to this source only on a few occasions. In result, judges are not only more likely to cite legal sources supporting their argument at all, judges will also refer to them more frequently.³

In contrast to previous research (e.g. Clark and Lauderdale, 2010) we consider all citations alike and do not select particular legal sources, such as substantial but not procedural legal sources. It is not necessary to *ex-ante* distinguish between positive and negative citations and between different types of legal sources to estimate the location of lower court decisions in an appropriate case-space. Positive and negative citations are already accounted for with different expected frequencies in the data generating process. Negative citation should occur less often than positive citations. Furthermore, irrespective of the type of the legal sources, the frequency with which lower court judges use certain legal sources, substantive or procedural, will always be informative if it reflects a choice of the judge.⁴

³Our work is thus similar to Clark and Lauderdale (2012) who also analyze citation counts.

⁴We also do not distinguish between legal sources for pragmatic reasons. We explicitly want to show that it is

4 A Case-Space Estimator for Decisions and Cited Legal Sources

Now that we developed an understanding for how judges chose their legal sources, we translate these insights into an appropriate measurement model. The data we observe consists of a $n \times m$ matrix \mathbf{Y} of citation counts—also the *decision-source matrix*—where element y_{ij} represents the frequency how often a judge writing decision $i \in \{1, \dots, n\}$ is citing legal source $j \in \{1, \dots, m\}$. We use a Poisson distribution as a typical probability model for count data to capture how often decisions are citing particular legal sources. The systematic component of our model rests on two assumptions. (1) Each decision has a fixed location along a unidimensional case-space. (2) A written decision will cite a legal source more often the closer their locations in this case-space. We express the probability of observing any particular distribution of legal source citations as

$$y_{ij} = \text{Poisson}(\lambda_{ij}) \tag{1}$$

$$\lambda_{ij} = \exp(\alpha_j + \beta_i - \gamma \|\theta_i - \phi_j\|^2). \tag{2}$$

A quadratic loss function operationalizes both assumptions. The distance between a decision i and a legal source j is expressed as $\|\theta_i - \phi_j\|^2$, where $\theta_i \in \mathbb{R}$ is the location of decision i and $\phi_j \in \mathbb{R}$ is the location of legal source j . The parameter γ captures the overall sensitivity of this difference in positions and ultimately measures its impact on the number of citations in each decision.

Following [Barberá \(2015\)](#), our model provides a richer parameterization of the systematic component than [Clark and Lauderdale \(2010\)](#): the parameters α_j and β_i explicitly capture the idiosyncrasies of the legal decisions and the cited legal sources. The parameter α_j expresses the authority of a legal source j . Some legal sources are by default cited more often than others, simply because they are more relevant. Similarly, β_i captures decision-specific differences. Some decisions cite on average more legal sources than others, e.g. because they are longer.

possible to retrieve the positions of legal decisions with minimal effort.

Identification is a core concern when estimating latent variable models. These models typically have more parameters than observations which results in infinitely many ‘correct’ solutions. Shifting, rotating or scaling one parameter can be easily offset by inversely shifting, rotating or scaling another related parameter. A unique solution needs to address all of these concerns. First, the model has to be anchored in the parameter space so that it does not float around. Second, identification also requires a well defined scale and the model must not be allowed to arbitrarily stretch. But even if local identification is guaranteed, it would still be possible to, third, symmetrically rotate the model in the parameter space by inverting all parameters.

Bayesian MCMC estimators address identification challenges by specifying prior distributions (Gelman and Hill, 2007; Jackman, 2009). Following Barberá (2015), we identify our model locally with a standard normal distribution for α_j and θ_i , thus solving the local identification problem.

$$\alpha_j \sim N(0, 1) \tag{3}$$

$$\beta_i \sim N(\mu_\beta, \sigma_\beta) \tag{4}$$

$$\phi_j \sim N(\mu_\phi, \sigma_\phi) \tag{5}$$

$$\theta_i \sim N(0, 1) \tag{6}$$

Global identification is a bit more challenging. Political scientists who estimate locations of political actors and legislative proposals in a common space face a similar problem when they repurpose item response theory (e.g., Clinton, Jackman and Rivers, 2004; Poole and Rosenthal, 2007). They address rotational invariance in at least two ways. Authors either globally identify their model. Highly informative priors on a well known political actor clearly determine who belongs to “the right” or “the left” (Clinton, Jackman and Rivers, 2004; Martin and Quinn, 2002). In a similar vein, in the case of multi-dimensional scaling where prior intuitions about political actors might not be as well defined, Jackman (2001) proposes to set priors on well understood legislative

proposals instead. The other option is to not identify a model globally at all. Since the model can flip, the respective posterior distribution can have two modes. While in theory, the sampler could visit both sides, in practice, however, this concern typically turns out not to be an issue as long as there is a reasonably large number of political actors whose locations can be reasonably well distinguished (Jackman, 2001, 2004). It is then sufficient to initialize the sampler around an educated guess. Even with symmetric—and thus for rotational invariance uninformative—priors the algorithm quickly converges to the correct posterior (Jackman, 2009).

Regarding the first strategy, fixing certain decisions or legal sources is not possible in our context. *Ex-ante*, an analyst is neither certain about a clear location of certain decisions nor about the location of the cited legal sources. Avoiding any unjustified bias from priors, the model should identify the parameters on the basis of the citation data only. We therefore begin by generating starting values empirically from a pre-run of the model where we draw 8 000 times from the posterior distribution. Because each chain explores only one side of the bi-modal posterior, the chains have to be collapsed to one side. Whenever the median of a chain has a different sign than the median from the first chain, we multiply all draws for this parameter with 1 to map the posterior on the same side. Averaging over all posterior draws from this first run, we obtain a good proxy for a starting value for the second run. In this second run, we then draw overall 80 000 times from the posterior across multiple chains to effectively explore its central tendency and variation.

5 Forum-Shopping and Forum-Selling of Lower Courts in Germany

With the model in hand, we now study decisions by German *Landgerichte* (district courts) and *Oberlandesgerichte* (regional courts of appeal) to see whether we can uncover systematic evidence for forum shopping and forum selling. While it has been documented that plaintiffs engage in ‘forum shopping’—and judges in ‘forum selling’—if respective institutional incentives are in place (e.g. Bechtold, Frankenreiter and Klerman, 2019), Germany is a particularly unlikely case for this

kind of behavior: Its civil law system is known to generate decisions in a bureaucratic way, and most importantly, without recording the individual votes of the participating judges who can be seen as apolitical career civil servants. As a matter of fact, Germany is often cited as the exemplary case of an impartial and objective civil law system (Langbein, 1985). Extending research beyond the well-known U.S. context, Bechtold, Frankenreiter and Klerman (2019) interview attorneys, judges and court officials and document mechanisms through which forum selling in fact also occurs in Germany. In areas like press law and antitrust where German citizens can choose the court they want to file their case with, litigants indeed strategically “shop” for the forums that suit their purpose. In return, courts who wish to establish themselves as a go-to-place in a certain legal area rely on different means to be particularly appealing. Earning a reputation for a tendency in rule making is a successful strategy when “selling” the own forum—an observation also echoed by recent press reports (Dahlkamp and Schmid, 2014; van Lijnden, 2016). Following the lead of Bechtold, Frankenreiter and Klerman (2019), we therefore decide to study case-spaces in press law and antitrust. Our quantitative analysis will test expectations generated by their case studies on a large sample of decisions statistically.

5.1 Collecting Data for a Case-Space

How to identify an appropriate set of legal decisions that can be compared to one another? And how to turn citation patterns into a document-source matrix that can be fed to the measurement model? An analyst first has to find, read, and curate cases that relate to the exact same legal topic. She then determines all legal sources the judge was citing. Finally, she counts all unique sources per decision and turns them into a table: Each row corresponds to a decision and each column carries the count for the respective citations of a legal source.

Instead of completely relying on manual labour for this task, we show how to use technology to save oneself this human effort and, at the same time, keep this process transparent and replicable. We let the machines do the work for collecting a well defined set of decisions (step 1), for identifying the citations (step 2), and finally for generating the decision-source matrix (step 3).

Table 1: Overview Over Data Collection.

	Basic Search	Enhanced Sample
Custom Indexed Data Base	d_{1A} : Press Law: Infringement d_{2A} : Press Law: Compensation	d_{1B} : Press Law: Infringement d_{2B} : Press Law: Compensation
Juris Homepage	d_3 : Antitrust	

Table 1 documents the different approaches we take.

Our data is from the legal database *Juris*⁵ which comprises a comprehensive digital collection of all available German court decisions. Each of *Juris*' records contains not only the complete text, but also further meta data such as titles, dates or information about the respective court. The data provides information on the further fate of a legal proceeding within the judicial hierarchy, all citations a legal decision makes (backward citation), as well as other legal decisions that refer to a certain decision (forward citation). The annotation of the corpus of decisions with meta data makes the further processing particularly easy.

Granted access to the backend of the *Juris* database, we had access to the data via a data base (*MongoDB*) that was indexed with a *Lucene* based search engine (*ElasticSearch*). This infrastructure allowed us to make full use of search engine functionalities and query this corpus as we saw fit. While collected the data for cases on press law using this infrastructure, we also used a second approach for the sake of reproducibility. For the antitrust case, we accessed the data through the frontend, making use of the *Juris*' own search functionality on their homepage.

We first identify a set of decisions that all belong to the same case-space. For the application to press law, we define a number of key words that describe the two areas we are interested in: decisions regarding claims for compensation and decisions regarding claims for injunction.⁶ Acknowledging human limitations in devising dictionaries for direct document selection (e.g. [Beauchamp, 2017](#); [King, Pan and Roberts, 2013](#); [Puglisi and Snyder Jr, 2011](#)), we strive to mitigate this bias.

⁵The data base is comparable to services like *Westlaw* for the British or the U.S. context.

⁶Please see appendix for the dictionaries.

The key terms are a seed query for the search engine and allow to retrieve a large collection of 100 documents that are the best ‘fit’ to the search terms.⁷ A trained human coder then uses reasoning to identify those cases that truly belong to the same case-space. In result, we select nine cases on privacy infringements through the publication of photos in print media and a respective claim for compensation (dataset d_{1A}). In addition, we collect six decisions on privacy infringement through criminal act allegation in the media and the corresponding claim for injunction (dataset d_{2A}).

For antitrust law we proceed in a similar way—this time, however, we access the database through the *Juris* online portal. We search with key words on antitrust⁸ for decisions from German regional courts (*Landgerichte*). Again, a trained coder reviewed all resulting 64 decisions and clustered them into reasonable topics. We identified 19 decisions in the same case-space regarding damage claims following from a cartel’s action (dataset d_3).

The keyterm queries to identify legal decisions from the same case-space result only in a modest amount of suitable decisions. For the two case-spaces in press law we therefore enlarge our sample with the help of the search engine. This time we take the titles of the decisions from the two already identified sets d_{1A} and d_{2A} as query terms.⁹ The top 25 most similar results are assumed to belong to the same case-space—and we use those decisions as the enlarged set for claims for compensation (dataset d_{1B}) and enlarged set of decisions on privacy infringement through criminal act allegation in the media and the corresponding claim for injunction (dataset d_{2B}).

Once the decisions in each set are defined, identifying the citations is straightforward: *Juris* already enriched all decisions with meta data. When accessing the original *Juris* data for press law with the own search engine, all sources of a decision are readily available as a list. In turn, when querying the data via *Juris*’ homepage, downloading each decision’s HTML page allows to extract citations on the basis of embedded hyperlinks.¹⁰

⁷Please see the appendix for a more in-depth explanation how to calculate this similarity.

⁸Please see appendix for the dictionary.

⁹An example for the collection on compensation cases is decision ‘LG Köln Aktz: 28 O 567/14’ with the following (German) title: “*Unterlassungsanspruch hinsichtlich der Bildnisveröffentlichungen wegen Verletzung des allgemeinen Persönlichkeitsrechts*”. For the collection of infringement cases, an example is decision ‘LG Heidelberg Aktz: 2 O 162/13’ entitled “*Störerhaftung des Betreibers einer Internet-Suchmaschine: Anzeige von Links durch die Suchmaschine zu Internetseiten Dritter mit persönlichkeitsrechtsverletzenden Inhalten*”.

¹⁰Some challenges, however, remain. Some decisions use the ‘Ibid.’ citation style (‘a.a.O’ in the German context)

As a third and final step, we count all sources in all legal decision. We define a legal source by its section (*Paragraph*) and paragraph (*Absatz*). The same law and section, but with a different paragraph counts as another legal source. Legal sources can be references to German Civil Code (e.g., a particular section of the BGB), to criminal law (e.g., a section in the StGB), to code of civil procedure (e.g., a section in the ZPO) or to a section in the German constitution (the *Grundgesetz*, GG); but also previous decisions of both lower courts (e.g. a decision written at the OLG Hamburg) and higher courts such as the Federal Court of Justice (BGH) or the German Federal Constitutional Court (BVerfG) as a specialized court of higher order. We also consider mentions of academic articles. Using the meta data, we then construct the respective decision-source matrix Y_{ij} that identifies all cited legal sources (j) for each pre-selected decision (i). When counting the sources, we leave out all citations to idiosyncratic legal sources, meaning any legal source which is cited by only one decision.¹¹

5.2 Application to Press Law

Let us now analyze the legal decisions in press law. [Jürgens \(2014\)](#) gathers data on press law caseload at Germany's regional courts. He finds that between 2010 and 2012, Berlin (28.67%), Hamburg (22.32%), and Cologne (*Köln*) (11.94%) adopt the bulk of all legal decisions with the rest of the caseload shared among all other courts. Authors disagree on how to interpret these data. Some argue, that there are systematic tendencies in rule-making favoring litigants who defend themselves against the press at these three places ([Dahlkamp and Schmid, 2014](#); [Höch, 2018](#); [Jürgens, 2014, 2016](#); [Kompa, 2012](#); [Sajuntz, 2014](#); [van Lijnden, 2016](#)). The high case load is a result of litigants who—aware of these tendencies—select their courts strategically. Adding to the data from case loads, authors substantiate their claims with anecdotal evidence from either the court in Cologne ([Dahlkamp and Schmid, 2014](#); [Jürgens, 2014, 2016](#)) or Hamburg ([Höch, 2018](#);

and the *Juris* data does not fully annotate this information (see also [Coupette and Fleckner, 2018](#)). The count of existing links between a decision and a legal source is potentially lower if a judge uses this citation style. If anything, this bias makes our final estimates more conservative.

¹¹These idiosyncratic legal sources do not offer any information about locations in the context of our model and leaving them out speeds-up estimation.

Jürgens, 2014, 2016; Kompa, 2012). Others contend that this interpretation is wrong. The high number of cases in Berlin, Hamburg or Cologne is rather the result of a concentration of media companies—and that includes an eco-system of specialized press lawyers who can serve them (Dölling, 2015; Höcker and Brost, 2015). Anecdotal evidence suggests that this is the case in particular for Berlin (Dölling, 2015). Answering the call for an analysis that “meets the standards of an empirical science” (Dölling, 2015, 130, *own translation*), we investigate these claims with our measurement model. Are the courts in Hamburg and Cologne really more likely to adopt decisions that are friendly to litigants and hostile towards the press?

5.2.1 Estimates Of Decision Locations

We first investigate the two sets of decisions d_{A1} and d_{B1} we collected with a key search query. The resulting position estimates are in line with *ex-ante* expectations based on expert knowledge and media reports. Figure 2 depicts the estimated decision locations. The figure displays all decision locations in our samples from Cologne or Hamburg in red and locations of any other court decisions in blue. A point represents the respective median of the posterior draws. Uncertainty bars around the estimate depict the central 90% credible interval. The box plots on top summarize the distribution of estimated locations from Cologne and Hamburg *vis-à-vis* other courts. For the decisions related to compensation in the panel on the left, we observe two outliers—both from Cologne. However, even though Hamburg and Berlin are also known as friendly towards compensation, the decisions do not show systematic differences. In the decisions related to injunction on the right panel, the decisions from Cologne and from Hamburg cluster, i.e. they are mapped onto a similar location in the case-space. Heidelberg, the only other court in the sample, is distinctly situated on the right. To sum-up, we find empirical evidence that lends itself to the experts’ and journalists’ expectations. In the context of press law, lower courts throughout Germany show systematic differences in their judgments on comparable cases when it comes to compensation or injunction in privacy infringement cases.

In the light of a limited empirical sample, we use the titles from the decisions in d_{1A} and d_{2A}

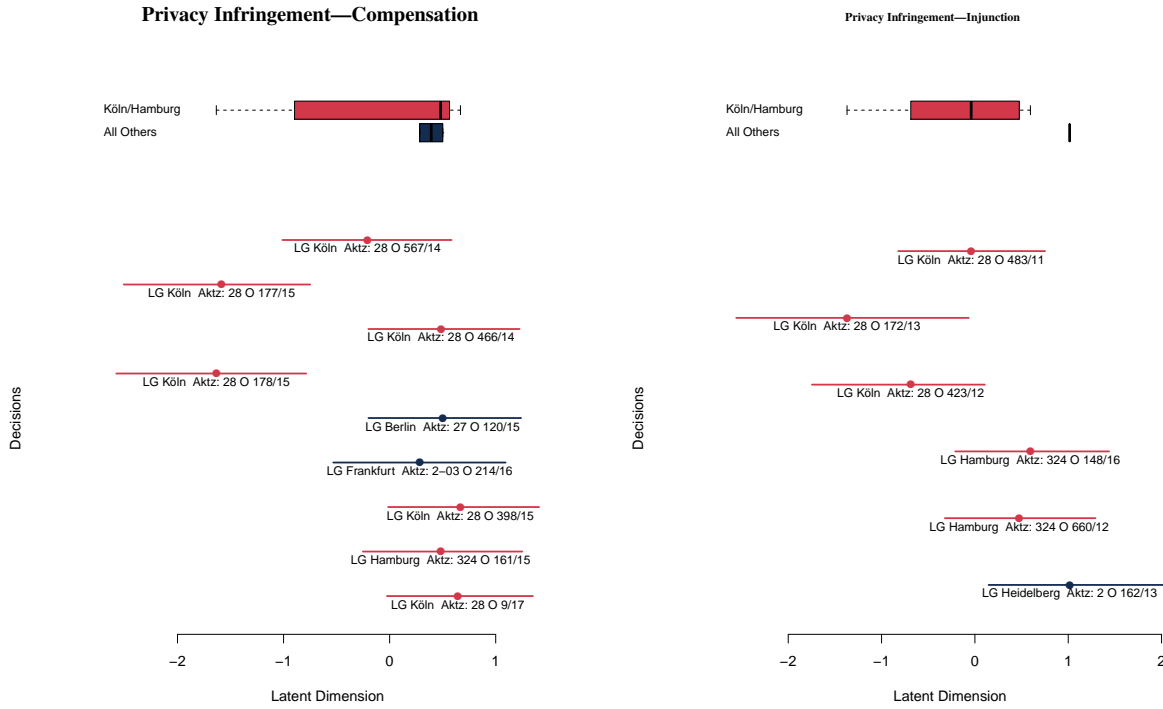


Figure 2: Estimated Locations of Written Decisions ($\hat{\theta}_i$). Pre-selected Set of Decisions. Points Indicate the Median of the Posterior Draws. The Bars Represent the Central 90% Credible Interval.

to expand the respective sets of decisions while keeping the concise and well defined case-spaces. Based on these 25 most (cosine) similar decisions each, we generate the decision-source matrices for d_{1B} and d_{2B} , again omitting idiosyncratic references and estimating the decision locations with the same Bayesian setup. Figure 3 displays the median estimates with a circle and the bars indicate the core 90% credible interval. In the first set of legal decisions—the collection on compensation—the courts in Cologne and Hamburg cluster on one end of the spectrum; all other courts can be found on the other end. Decisions on injunctions show more mixed positions. However, as the boxplots on top of the figure show, there is still a trend for the courts of Hamburg and Cologne to be on one end of the spectrum, and all others to be on the other end. Given the size of the credible intervals, though, the locations of many decisions cannot be distinguished on substantive grounds.

To sum up, techniques known from information retrieval help us to automatically enlarge our initial sample of selected seed decisions with similar decisions that can be validly mapped into the same case-space. The estimated locations are again in line with anecdotal evidence from media

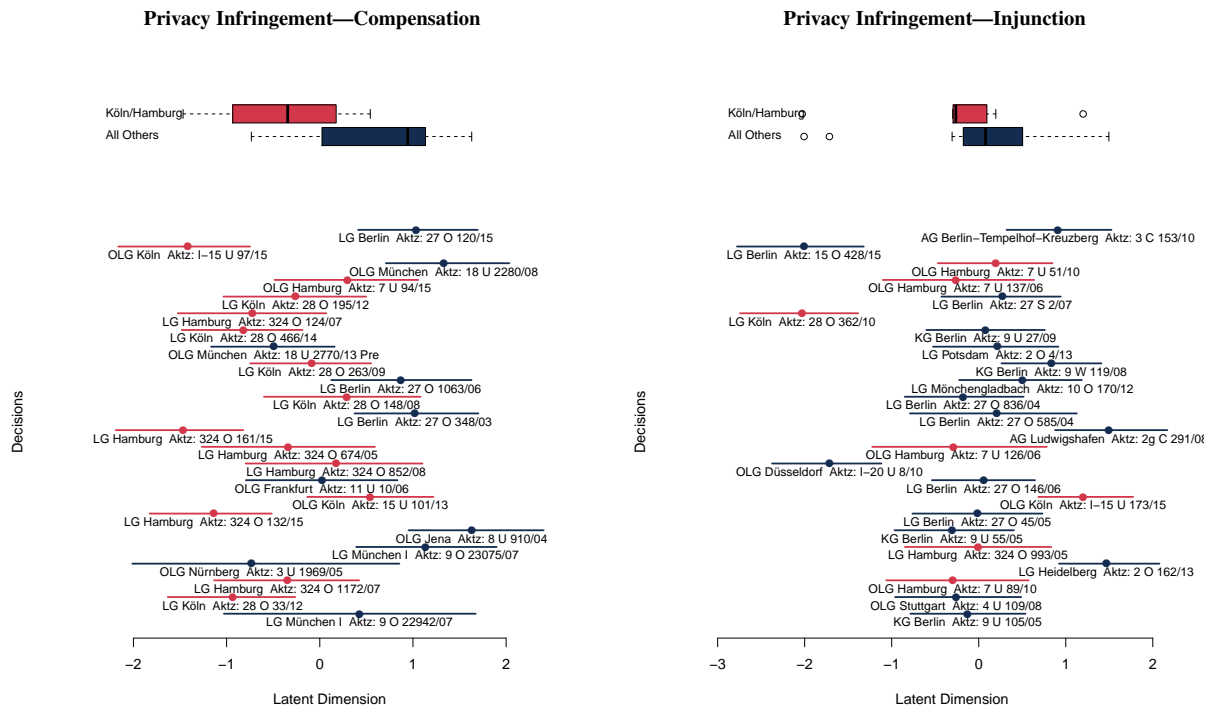


Figure 3: Estimated Locations of Written Decisions ($\hat{\theta}_i$). Set of Decisions Using a ‘More-Like-This’ Query. Points Indicate the Median of the Posterior Draws. The Bars Represent the Central 90% Credible Interval.

reports and from experts.

5.2.2 Case Study to Assess the Validity

We assess the validity of our estimates and study whether the results of our model hold up to legal interpretation. We manually compare the substantial content of three decisions from the extended set of cases on privacy infringement and the claim for compensation (d_{1B}): two decisions that are scaled at similar positions (LG Hamburg, 324 O 161/15 and LG Cologne, 28 O 466/14) and one decision that is scaled at the opposite end (LG Munich, 9 O 23075/07). Figure 4 shows their differences on the latent dimension.¹² The figure displays the respective medians and the 95% credible interval. Our model predicts a high probability that the position of LG Munich, 9 O 23075/07 is different from LG Hamburg, 324 O 161/15 and LG Cologne, 28 O 466/14. In return, the probability that LG Hamburg, 324 O 161/15 and LG Cologne, 28 O 466/14 are different from one another is relatively low. If our approach is valid, then similar legal decisions will derive their legal argument using similar legal sources and the legal decision that is distinct should rely on different legal sources.

The median estimate of the Hamburg decision (LG Hamburg, 324 O 161/15) is to the left of the case-space in Figure 3. The litigant in the Hamburg decision requests a compensation for the repeated publication of pictures of her taken while visiting her hospitalized husband—a famous Formula One driver. The litigant used various legal means to stop the defendant from publishing pictures before referring to the *Landgericht* (324 O 161/15, Mn 6).¹³ The defendant requests to dismiss the lawsuit arguing, for example, that the contemporary interest in the hospitalization was not limited towards the Formula One driver but would include how the spouse addresses the stroke of fate (324 O 161/15, Mn 27-29). The district court concluded that publishing the pictures violated the litigant in her general personality rights derived from the German Civil Code (BGB) in connection with the German Constitution (§ 823 I BGB in connection with Art. 2 I

¹²To calculate these first differences, we subtract the respective position estimates for all posterior draws from one another.

¹³We refer to sections of interest in court decisions using margin numbers (Mn) as shown in the JURIS database.

Privacy Infringement—Compensation First Differences of Three Decisions

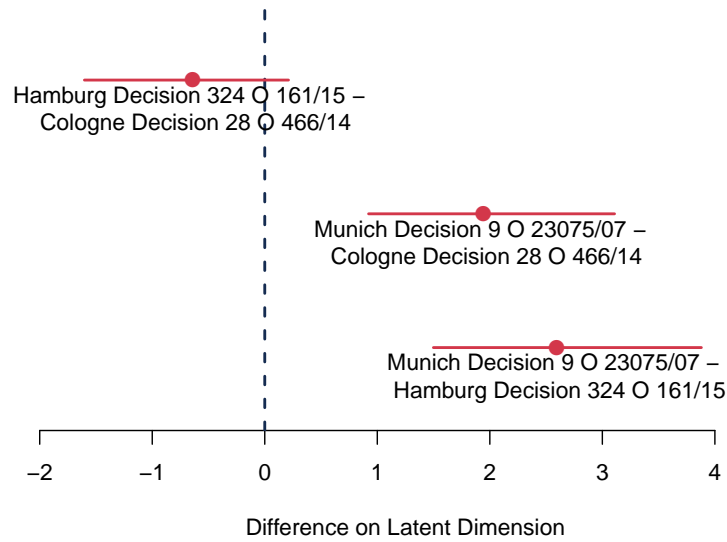


Figure 4: Estimated Differences in the Location ($\hat{\theta}_i$) for Hamburg Decision 324 O 161/15, Cologne Decision 28 O 466/14 and Munich Decision 9 O 23075/07. Points Indicate the Median of the Difference of the Respective Posterior Draws. The Bars Represent the Central 95% Credible Interval of the Difference of the Respective Posterior Draws.

and Art. 1 I GG). According to the court, the publication of a picture does not *per se* violate a person's general personality rights (LG Hamburg, 324 O 161/15, 35). Instead, publishing can be justified when it is documenting contemporary events in line with the German law regulating art and copyright questions (§ 23 I *Kunsturhebergesetz*, [KUG]). Subsequently, the court had to balance the protection of the private sphere of the individual according to the German Constitution (Art. 1 I GG and Art. 2 I GG) and the European Convention on Human Rights (Art. 8 I ECHR) against the freedom of the press to document contemporary events of importance in line with the German Constitution (Art. 5 I GG) and the ECHR (Art. 10 I ECHR).

To justify the litigant's claim, the district court of Hamburg referred to case-law by the German Federal Court (BGH, VI ZR 51/06; VI ZR 272/06). The BGH had argued in the past that the repeated and tenacious publication of pictures can become a very serious infringement of a litigant's privacy and respective violations require financial compensation (BGH, VI ZR 223/94). In particular, the infringement of a person's privacy is wider when media outlets publish pictures compared to textual articles (BGB, VI ZR 230/08). Moreover, the district court of Hamburg highlighted that the litigant had used legal measures against the defendant to hinder the publication of pictures. Hence, the defendant was well aware of the litigant's disagreement to publish pictures showing her in a personal, very exceptional situation. The infringement of the litigant's privacy was not justified and requires the defendant to pay financial compensation. In short, the court of Hamburg heavily relied on case-law published by the BGB.

If our scaling approach is valid, then the citation pattern of the district court of Cologne in LG Köln, 28 O 466/14 should be similar. The median position of the Cologne decision is to left of the case-space in Figure 3, and similar to the median position of the Hamburg decision. The estimated location of both decisions are not systematically different from one another in the common case-space.

The litigant in the Cologne decision is an actress who requests an act of omission and a financial compensation for the online and offline publication of an article together with a picture speculating about a second pregnancy of her. The litigant argues that the picture was taken in

a private moment and violates her general personality rights. The litigant is also anxious to not share information about her personal life. Instead she succeeded with similar legal claims to not publish pictures during her first pregnancy (28 O 466/14, Mn 7). The defendant requests to dismiss the lawsuit arguing in favor of the contemporary importance of the picture which is in the public interest (28 O 466/14, Mn 15-16). These general case characteristics of the Cologne decision are similar to the Hamburg decision. In both decisions the litigants claimed their pictures were taken in private moments violating their personality rights. The defendants rejected claims arguing in favor of the contemporary importance of the pictures.

A closer reading of the Cologne decision shows that the district court refers to the same legal norms and similar case-law compared with the Hamburg decision to settle the dispute. Subsequently, the district court of Cologne argues in favor of the litigant's general personality rights derived from the BGB, the GG and the KUG (§§ 1004 and 823 II GCC in connection with Art. 2 I GC and Art. 1 GC as well as §§ 22, 23 KUG). Moreover, the court weighs—similar to the district court of Hamburg—the protection of the private sphere (Art. 1 I GG and Art. 2 I GG together with Art. 8 I ECHR) against the freedom of the press (Art. 5 I GG together with Art. 10 I ECHR) on the basis of the European Convention on Human Rights and the German Constitution (28 O 466/14, Mn 20-23).

To argue the case, the court in the Cologne decision relies on case-law published by a number of courts and especially the BGH. While not all decisions referred to by the district court of Cologne were of relevance to the court in Hamburg, both courts heavily relied on case-law published by the BGH. In particular, there is overlap in two BGH rulings repeatedly cited by both courts: VI ZR 223/94 and VI ZR 51/06. These decisions were used to justify the litigant's claim against the defendant.

In sum, the Hamburg decision and the Cologne decision are scaled at similar ends of the case-space in Figure 3. The general case characteristics and the legal outcomes are similar in both cases. In addition, the legal norms and the case-law used to argue the cases widely overlap.

The median position of the *Munich decision* (LG Munich, 9 O 23075/07) is scaled at the

opposite end of the case-space in Figure 3 when compared to the median positions of the Hamburg and Cologne decision. If our scaling approach is accurate, we should find that the legal arguments developed in the Munich decision are based on different legal sources than the arguments in the Hamburg and Cologne decision. Moreover, while the cases should address similar scenarios, we might find variation in some case characteristics.

The litigant in the Munich decision—a famous actress—requests financial compensation for the publication of secretly taken pictures showing her going for a walk with her newborn. This scenario is comparable to the scenarios in the Hamburg and Cologne decision. However, different to the latter two decisions the defendant in the Munich decision had already given a declaration to refrain from further publications and had payed previous legal fees of the litigant. Nevertheless, the litigant still requested compensation for the pictures already published. The actress argued that the pictures violated her general personality rights, especially as she was in a private moment with her newborn (9 O 23075/07, Mn 3). The defendant requested that the district court rejects the claim. The defendant had already signed a declaration to cease and desist and argued that the litigant is a public figure which is why the pictures were of contemporary interest. The newborn was not recognizable in the pictures (9 O 23075/07, Mn 10-11). The litigant's and the defendant's requests are comparable to the scenarios described in the Hamburg and Cologne decision. Nevertheless, the district court dismissed the litigant's request in the Munich decision and saw no right to receive compensation (9 O 23075/07, Mn 13-14).

Interestingly, the court argued that the publication of the pictures violated the litigant's rights derived from the German law regulating art and copyright questions (§§ 22 and 23 KUG; 9 O 23075/07, Mn 15). The courts in Hamburg and Cologne presented similar thoughts. Thus, all three decisions seem to be comparable in an appropriate case-space. However, the judges in the Munich decision do not derive a financial compensation from the violation. Instead, the court's line of reasoning is different to the ones presented by the district courts in Hamburg and Cologne. The court in Munich referred to different legal norms than the other two courts.

The financial compensation for a violation of someone's personality rights after publishing

a picture is commonly based on Art. 1 and Art. 2 I GG in conjunction with § 823 I BGB—the regulations referred to in the Hamburg and Cologne decision. Nevertheless, in the Munich decision the judges rather cite decisions by the BGH which the other district courts do not quote. Accordingly, the BGH had outlined that financial compensation requires a “very serious infringement” (9 O 23075/07, Mn 17; own translation) of someone’s personality rights. However, the unjustified publication of a picture lowers the legal barriers to receive financial compensation (VI ZR 56/94; VI ZR 255/03). Nevertheless, the district court of Munich concludes that the publication of the actresses’ picture does not qualify as serious infringement of her privacy. Instead, the scenario is different to the ones in decisions by the BGH. The legal norm of relevance to the court in Munich is the newly edited § 253 II BGB. This norm allows to grant financial compensation for various physical and psychological violations but—according to the court—is not intended to justify universal compensation. Case-law of courts which had to assess very serious infringements supports this view. Subsequently, in the case at hand the district court of Munich does not see any justification for financial compensation (9 O 23075/07, Mn 16-22).

The district court in Munich faced a scenario similar to the courts in Hamburg and Cologne. Nevertheless, while the case characteristics are comparable the court in the Munich decision argues the case differently compared to the district courts in Hamburg and Cologne. Eventually the court in Cologne derives different legal consequences rejecting the litigant’s claim. Our model finds that the median position of the Munich decision is located at the opposite end to the median positions of the Hamburg and Cologne decision in Figure 3. Subsequently, the comparisons of the three cases supports the validity of our approach to estimate similar locations of the Hamburg and the Cologne decisions, which are very different from the Munich decision.

Figure 5 summarizes similarities and differences between the three decisions assessed throughout the case study. The figure is read from top to bottom. Following the solid line (Hamburg decision), dashed line (Cologne decision) and dotted line (Munich decision) leads to the legal norms and the case-law which are used to develop the judicial arguments over the course of the different decisions. The figure illustrates that our scaling approach allows to differentiate between different

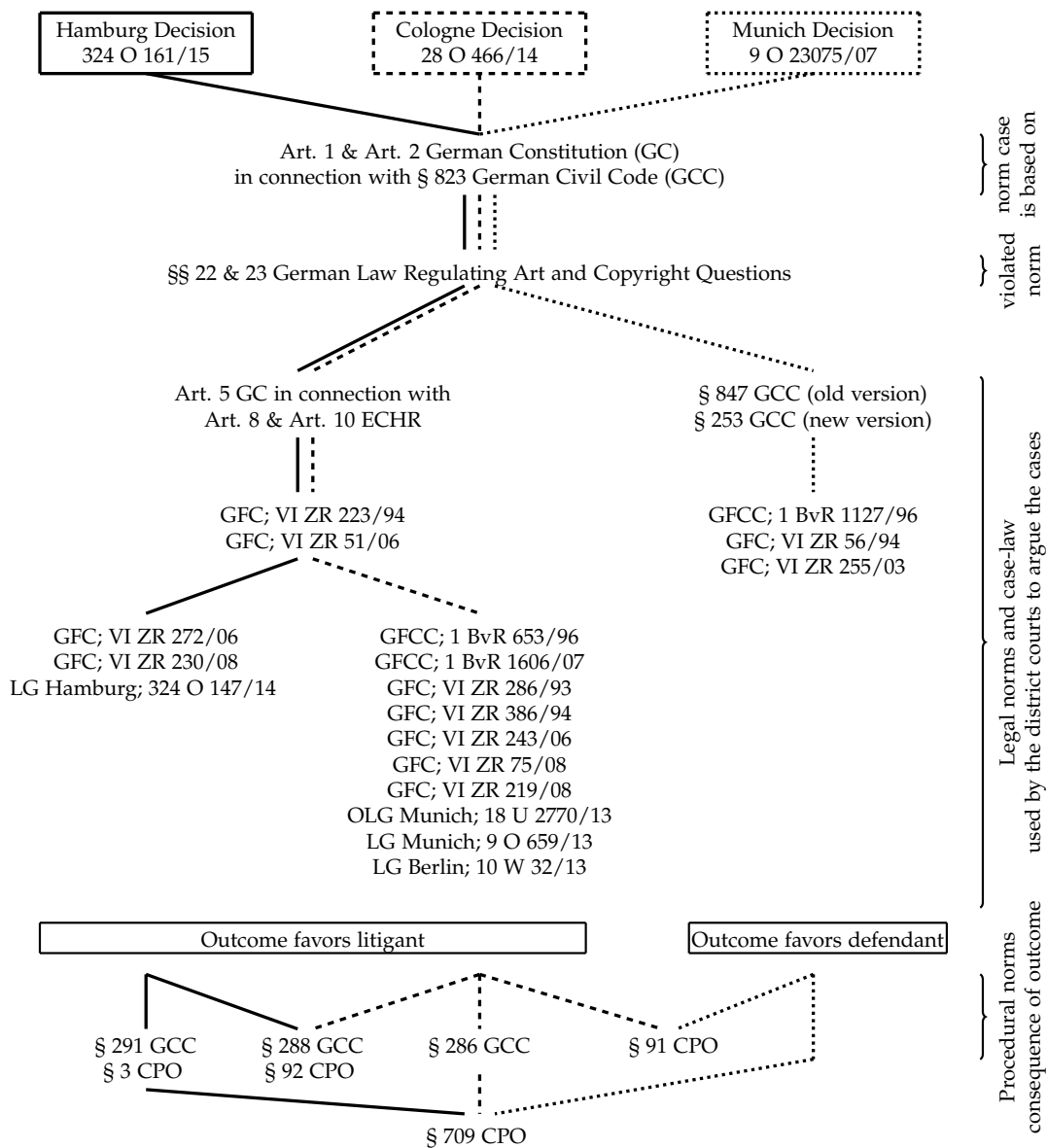


Figure 5: Summary of Legal Argumentation in Three Decisions from the Case-Space on Privacy Infringement and Claims for Compensation in Figure 3

nuances in the main body of the decisions, where courts argue the cases. In the initial parts of the decisions there is overlap between the legal norms on which a case is based on. This is not surprising as these norms are presented by the litigant to initiate the cases based on similar scenarios. However, the Hamburg and Cologne decision argue based on similar legal norms and case-law while the Munich decision has nothing in common with the other two decisions. This speaks to the validity of our scaling approach according to which the median scores estimated for the Hamburg and Cologne decision are located at similar ends of our case-space while the Munich decision is placed at the opposite end (Figure 3). Moreover, the referral to different norms and case-law lead to different outcomes favoring either the litigant or the defendant.

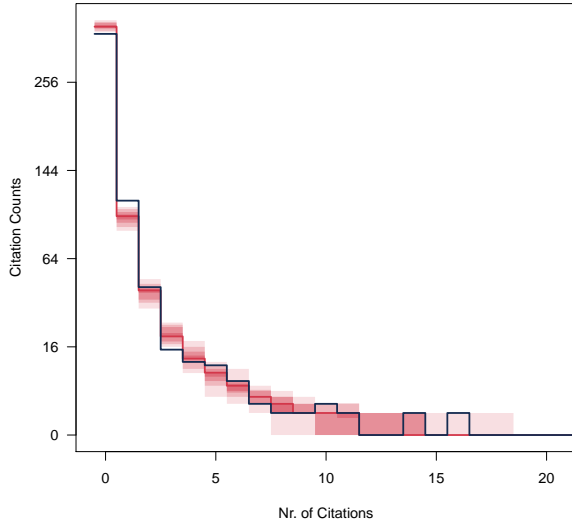
5.2.3 Assessment of Model Fit using Posterior Predictive Checks

Finally, we also want to assess whether the model actually fits the data generating process that we posit using posterior predictive checks. We therefore predict the outcome variable on the basis of the last 50 draws from the posterior sampling chain and plot it against the original data. Figure 6 displays the results for all models. Each figure represents the predicted citation counts \hat{y}_{ij} for all decision-source pairs as a histogram. The figures chart the respective counts horizontally and show the square of their occurrence vertically.¹⁴ Credible intervals are at the core 80%, 60%, 40% and 20% of the distribution. The solid lighter line depicts the median prediction across all models. The darker line plots the observed data.

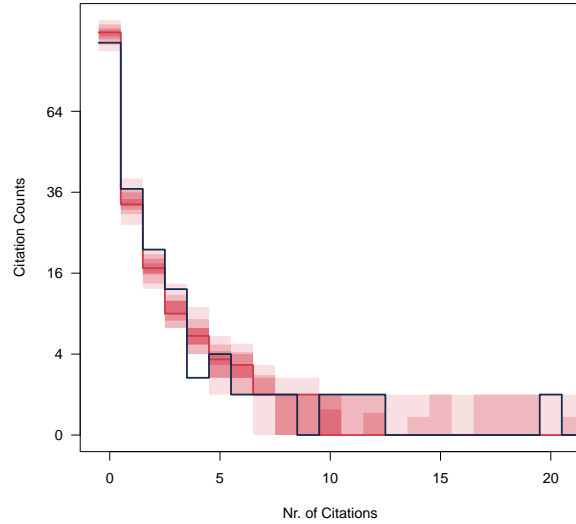
Our predictions mostly match the observed data and we conclude that the models do a good job in predicting the observed counts. For the set of cases that claim compensation for privacy infringements, both models do very well as long as the respective decision-source pair is cited frequently. We are thus confident that our systematic and our stochastic component correctly model the true underlying data generating count process.

¹⁴In line with [Clark and Lauderdale \(2012\)](#), we use the counts' squared values for better visibility.

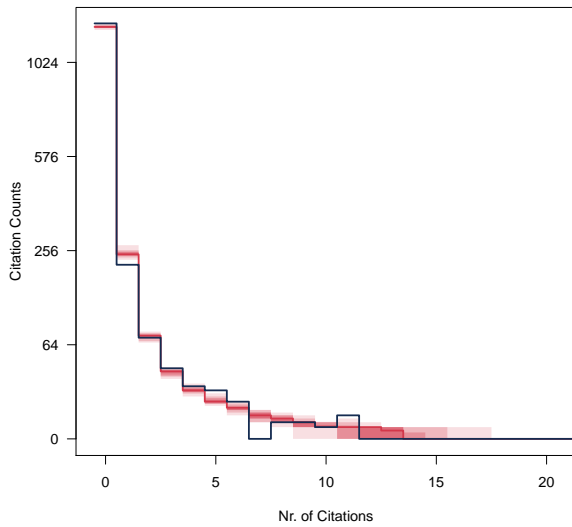
**Privacy Infringement—Compensation
(Curated Sample)**



**Privacy Infringement—Injunction
(Curated Sample)**



**Privacy Infringement—Compensation
(Expanded Sample)**



**Privacy Infringement—Injunction
(Expanded Sample)**

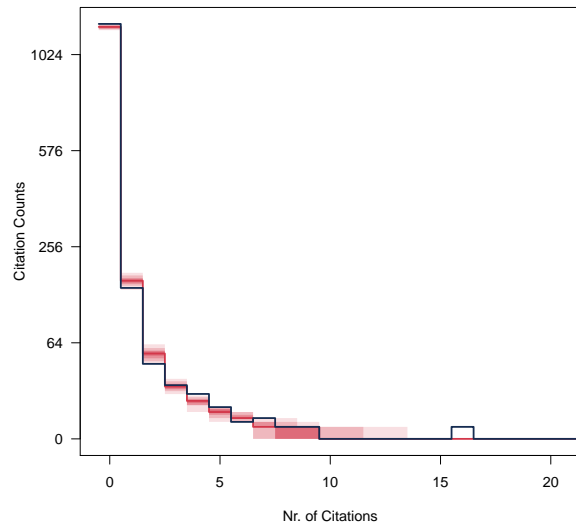


Figure 6: Posterior Predictive Checks. The Figures Visualize the Citation Counts of All Possible Decision Source Pairs in Form of a Histogram. Counts Are Displayed Horizontally and Their Respective Frequencies Is Shown Vertically. Predictions are in Red and stem from the Last 50 Draws of the Sampling Chain with Credible Intervals at 80%, 60%, 40% and 20%. Median Predicted Counts Are Represented with a Red Line, the Observed Data with a Blue Line.

5.3 Application to Antitrust Law

Antitrust is yet another area in German law that offers incentives for judges to behave strategically. Plaintiffs can file their case either at the seat of a cartel member or at the place where mischief occurred. Often, cartels operate nationally—which means that the plaintiff can choose from the overall 24 courts when filing the complaint. Judges themselves are interested in attracting cases for their highly visible and exciting nature.

[Bechtold, Frankenreiter and Klerman \(2019\)](#) indeed find some evidence in favor of forum shopping for the regional court in Mannheim. However, they are more cautious in their conclusions with regards to other cases. According to their interviews, the three regional courts in Mannheim, Cologne and Dortmund are particularly plaintiff friendly. In contrast, there are other courts such as Kiel, Leipzig, Düsseldorf or Munich that are less so. We are keen to examine this area with our model and hope to contribute to this ongoing debate by examining a larger set of cases quantitatively.

Figure 7 plots the result for our twenty cases. The courts in Mannheim and Dortmund tend to adopt decisions that are on one side of the spectrum. Decisions from Cologne, however, seems to have a more central position in our sample. Other courts either occupy the centre or display estimated locations of decisions on the centre-right. The box plots nicely summarize the core message of our findings: we do find systematic differences for the case-space in antitrust. A formal t-test indeed corroborates that the difference in the means of the estimated decisions from court decisions of the Dortmund/Mannheim group is different from the estimated locations of decisions of all others—even in this limited sample.

6 Conclusion

Legal scholars and political scientists are not only interested in understanding the decisions judges make, but also seek to analyze the policy implications of written decisions. Previous strategies rely on voting data of the involved judges ([Clark and Lauderdale, 2010](#)). But in cases where no such

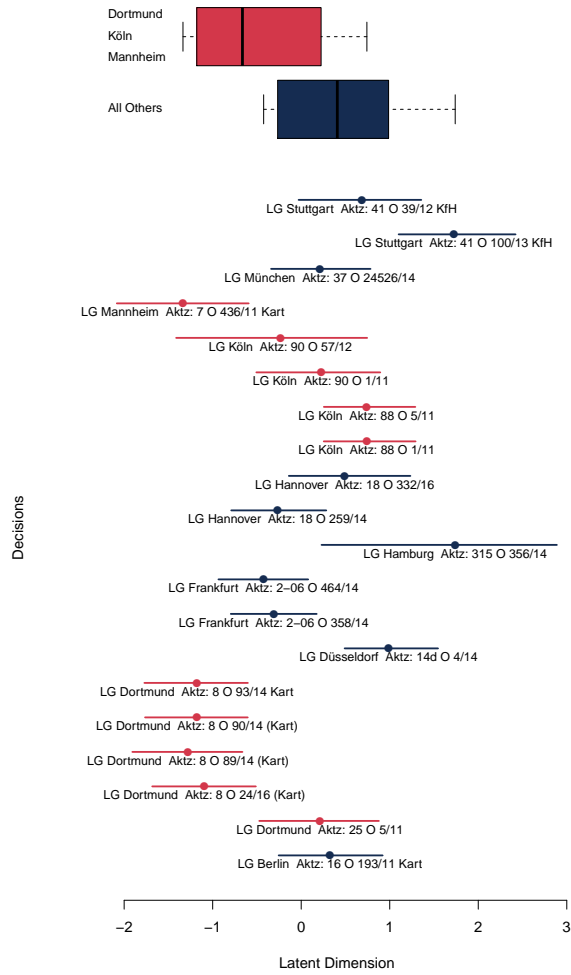


Figure 7: Estimated Locations ($\hat{\theta}_i$) of Written Decisions on Antitrust. Points Indicate the Median of the Posterior Draws. The Bars Represent the Central 90% Credible Interval.

records exists, empirical strategies to scale decisions in a common space are still missing. We show how to estimate and consequently compare positions of various actors in a common case-space of court decisions using citations networks. Our model rests on the notion that ‘similar’ decisions, i.e. decisions with similar locations in a common space, should cite similar legal sources using similar citation patterns.

To showcase the abilities of our model, we study a particularly challenging case: the German legal system. Qualitative evidence indicates that German lower courts (*Landesgerichte*) have a systematic bias in some legal areas, engaging in *forum selling* in press law and antitrust (Klerman and Reilly, 2016; Bechtold, Frankenreiter and Klerman, 2019). We corroborate these findings. Indeed, there is evidence that some lower courts—in press law Köln and Hamburg, and for antitrust Dortmund and Mannheim—do take systematically biased decisions. Our detailed legal analysis of three decisions from claims for compensation regarding privacy infringements corroborates our measurements. We can replicate our findings for different sets of decisions. In addition, our paper also shows how to find appropriate decisions in a large legal corpus that all belong to the same case-space, thus leading the way to easily scaling our method for analysing larger sets of legal decisions.

Of course, there are limits to what our model can do. Just like any other item-response theory approach—for example when locating parliamentarians on a left-right spectrum—all units that are being scaled need to reside in the same latent space. For our legal application, this means that all decisions have to concern the same legal matter. This includes regulatory contexts: Any major legislative change that affects the legal sources available to the judge is likely to introduce bias.

We hope that our model is a welcome addition to the toolbox for an interdisciplinary group of scholars who wish to empirically study judges and their decisions. Legal scholars might be interested in unveiling the development of doctrine across time and space. Political scientists might want to understand the nexus between institutions and power. Others, like economists or sociologists, might rather be concerned with the effects of legal decisions on business or societies

at large. Empirical preference measures of legal decisions are the foundation for the development and testing of powerful analytical models and may find their application in any of these fields.

Lastly, our insights has the potential to be useful beyond academia. Locating written lower court decisions in a common space has also important practical implications for law firms. Suppose a lawyer who is in favor of an extreme position in an ongoing case is trying to build an argument. Using our scaling model, she can obtain a quick overview over the tendency in each ruling she selects. This does not only help her identify the locations of those who are in favor of her own opinion. She could also study the arguments of the opposite side to be able to anticipate and preempt them. Thus, locating actual decisions in a case-space is very helpful for all who want to quickly spot similar or opposing decisions. Using our model, they can retrieve an easy roadmap that can guide them in their professional effort.

References

- Barberá, Pablo. 2015. “Birds of the Same Feather Tweet Together: Bayesian Ideal Point Estimation Using Twitter Data.” *Political Analysis* 23(1):76–91.
- Beauchamp, Nicholas. 2017. “Predicting and Interpolating State-Level Polls Using Twitter Textual Data.” *American Journal of Political Science* 61(2):490–503.
- Bechtold, Stefan, Jens Frankenreiter and Daniel Klerman. 2019. “Forum Selling Abroad.” *Southern California Law Review* .
- Brouard, Sylvain and Christoph Hönnige. 2017. “Constitutional courts as veto players: Lessons from the United States, France and Germany.” *European Journal of Political Research* 56(3):529–552.
- Cameron, Charles M. and Lewis A. Kornhauser. 2017a. “Chapter 2: What Courts Do . . . And How to Model It.”.
- Cameron, Charles M and Lewis A Kornhauser. 2017b. Theorizing the U.S. Supreme Court. In *Oxford Research Encyclopedia of Politics*.
- Carrubba, Cliff, Barry Friedman, Andrew D. Martin and Georg Vanberg. 2012. “Who Controls the Content of Supreme Court Opinions?” *American Journal of Political Science* 56(2):400–412.
- Choi, Stephen J and G Mitu Gulati. 2008. “Bias in Judicial Citations : A Window into the Behavior of Judges ?” *Journal of Legal Studies* 37(1):87–129.
- Clark, Tom S. and Benjamin E. Lauderdale. 2010. “Locating Supreme Court Opinions in Doctrine Space.” *American Journal of Political Science* 54(4):871–890.
- Clark, Tom S. and Benjamin E. Lauderdale. 2012. “The Genealogy of Law.” *Political Analysis* 20(3):329–350.
- Clinton, Joshua D., Simon Jackman and Douglas Rivers. 2004. “The Statistical Analysis of Roll Call Data.” *The American Political Science Review* 98(2):355–370.
- Coupette, Corinna and Andreas M. Fleckner. 2018. “Quantitative Rechtswissenschaft Sammlung, Analyse und Kommunikation juristischer Daten.” *JuristenZeitung* 73(8):379–389.
- Dahlkamp, Jürgen and Barbara Schmid. 2014. “Köln nimmt das alles.” *Der Spiegel* (42):150–153.
- Dalla Pellegrina, Lucia, Nuno Garoupa and Marian Gili. 2020. “Estimating Judicial Ideal Points in Bi-Dimensional Courts: Evidence from Catalonia.” *Journal of Empirical Legal Studies* n/a(n/a).
- Dölling, Birger. 2015. “Der fliegende Gerichtsstand im Presserecht Spielball der Interessen ?” *Neue Juristische Wochenschrift* 43:124–130.
- Dyevre, Arthur. 2019. “The Promise and Pitfalls of Automated Text-Scaling Techniques for the Analysis of Judicial Opinions.” *Working Paper* .

- Epstein, Lee, Andrew D. Martin, Jeffrey A. Segal and Chad Westerland. 2007. "The Judicial Common Space." *Journal of Law, Economics, & Organization* 23(2):303–325.
- Evans, Michael, Wayne McIntosh, Jimmy Lin and Cynthia Cates. 2007. "Recounting the courts? Applying automated content analysis to enhance empirical legal research." *Journal of Empirical Legal Studies* 4(4):1007–1039.
- Fowler, James H., Timothy R. Johnson, James F. Spriggs, Sangick Jeon and Paul J. Wahlbeck. 2007. "Network analysis and the law: Measuring the legal importance of precedents at the U.S. Supreme Court." *Political Analysis* 15(3):324–346.
- Frankenreiter, Jens. 2017. "The Politics of Citations at the ECJ-Policy Preferences of E.U. Member State Governments and the Citation Behavior of Judges at the European Court of Justice." *Journal of Empirical Legal Studies* 14(4):813–857.
- Gelman, Andrew and Jennifer Hill. 2007. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge: Cambridge University Press Cambridge.
- Hall, Mark A and Ronald F Wright. 2008. "Systematic Content Analysis of Judicial Opinions." *California Law Review* 96(1):63–122.
- Höch, Dominik. 2018. "Alles neu, alles anders? Das BVerfG nimmt sich der Gerichtspraxis im Presserecht an."
- Höcker, Ralf and Lucas Brost. 2015. "Kompetenz zahlt sich aus - Zur Notwendigkeit des fliegenden Gerichtsstandes im Nischenrechtsgebiet Presserecht." *IP - Rechtsberater* (6):138–141.
- Hönnige, Christoph. 2007. *Verfassungsgericht, Regierung und Opposition. Die vergleichende Analyse eines Spannungsdreiecks*. Wiesbaden: VS Verlag für Sozialwissenschaft.
- Hönnige, Christoph. 2009. "The Electoral Connection: How the Pivotal Judge Affects Oppositional Success at European Constitutional Courts." *West European Politics* 32(5):963–984.
- Jackman, Simon. 2001. "Multidimensional Analysis of Roll Call Data via Bayesian Simulation: Identification, Estimation, Inference, and Model Checking." *Political Analysis* 9(3):227–241.
- Jackman, Simon. 2004. "Bayesian Analysis for Political Research." *Annual Review of Political Science* 7:483–505.
- Jackman, Simon. 2009. *Bayesian Analysis for the Social Sciences*. New York: Wiley.
- Jacob, Marc A. 2014. *Precedents and Case-Based Reasoning in the European Court of Justice: Unfinished Business*. Cambridge: Cambridge University Press.
- Jürgens, Uwe. 2014. "Abgestürzte Gerichtsstände Der fliegende Gerichtsstand im Presserecht." *Neue Juristische Wochenschrift* 42:3061–67.
- Jürgens, Uwe. 2016. "Turbulenzen im Presseprozessrecht. Der rechtswidrige Standardverzicht auf mündliche Verhandlungen im einstweiligen Rechtsschutz." *Kommunikation und Recht* (1):7–12.

- Kelemen, Katalin. 2013. "Dissenting Opinions in Constitutional Courts." *German Law Journal* 14(8):1345–1371.
- King, Gary, Jennifer Pan and Margaret E Roberts. 2013. "How censorship in China Allows Government Criticism But Silences Collective Expression." *American Political Science Review* 107(2):326–343.
- Klerman, Daniel and Greg Reilly. 2016. "Forum Selling." *Southern California Law Review* 89(2):241–316.
- Kompa, Markus. 2012. "Experiment zum fliegenden Gerichtsstand."
- Kornhauser, Lewis A. 1992a. "Modeling Collegial Courts I: Path-Dependency." *International Review of Law and Economics* 12(2):169–185.
- Kornhauser, Lewis A. 1992b. "Modeling Collegial Courts II: Legal Doctrine." *Journal of Law, Economics, & Organization* 8(3):441–470.
- Landa, Dimitri and Jeffrey R Lax. 2008. "Disagreements on Collegial Courts: A Case-Space Approach." *Journal of Constitutional Law* 102:305–329.
- Langbein, John H. 1985. "The German Advantage in Civil Procedure." *The University of Chicago Law Review* 52(4):823.
- Lax, Jeffrey R. 2011. "The New Judicial Politics of Legal Doctrine." *Annual Review of Political Science* 14(1):131–157.
- Lax, Jeffrey R. 2012. "Political constraints on legal doctrine: How hierarchy shapes the law." *Journal of Politics* 74(3):765–781.
- Manning, Christopher D., Prabhakar Raghavan and Hinrich Schütze. 2009. Scoring, term weighting, and the vector space model. In *Introduction to Information Retrieval*. Cambridge: Cambridge University Press chapter 6, pp. 118–132.
- Martin, Andrew D. and Kevin M. Quinn. 2002. "Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the U.S. Supreme Court, 1953-1999." *Political Analysis* 10(2):134–153.
- McGuire, Kevin T. and Georg Vanberg. 2005. "Mapping the Policies of the U.S. Supreme Court: Data, Opinions, and Constitutional Law." *Working Paper* .
- Poole, Keith T. and Howard Rosenthal. 2007. *Ideology and Congress*. Transaction Publishers.
- Puglisi, Riccardo and James M Snyder Jr. 2011. "Newspaper Coverage of Political Scandals." *The Journal of Politics* 73(3):931–950.
- Raffaelli, Rosa. 2012. "Dissenting opinions in the Supreme Courts of the Member States." *European Parliament: Directorate General for Internal Policies* (PE 462.470).
- Sajuntz, Sascha. 2014. "Die Entwicklung des Presse- und Äußerungsrechts in den Jahren 2012/2013." *Neue Juristische Wochenschrift* pp. 25–30.

- Sternberg, Sebastian, Thomas Gschwend, Caroline E. Wittig and Benjamin G. Engst. 2015. "Zum Einfluss der öffentlichen Meinung auf Entscheidungen des Bundesverfassungsgerichts: Eine Analyse von abstrakten Normenkontrollen sowie Bund-Länder-Streitigkeiten 1974 - 2010." *Politische Vierteljahresschrift* 56(4):570–598.
- van Lijnden, Constantin. 2016. "Fliegende Richter." *Die Zeit* (28):1–5.
- Whalen, Ryan. 2016. "Legal Networks: The Promises and Challenges of Legal Network Analysis." *Michigan State Law Review* (539):539–566.
- Winkels, Radboud, Jelle de Ruyter and Henryk Kroese. 2011. "Determining Authority of Dutch Case Law." *Legal Knowledge and Information Systems* 235:103–112.

Scaling Lower Court Decisions

Appendix

A Proof of Concept for the Model

We devise toy examples to show that our model indeed correctly picks up positions in a way we expect. Let us begin with a baseline experiment. The artificial data we constructed in Table A.1 specifies whether a legal decision cites a certain source or not. There are five decisions that refer to five different legal sources. If the decision refers to a legal source, it carries a 1, otherwise it has a 0. The way the data is set up, it is straightforward to see that all decisions should be distributed symmetrically and at equal distances in the case space.

Table A.1: Toy Data 1: Does a Decision Cite a Source?

	Source 1	Source 2	Source 3	Source 4	Source 5
Decision 1	1	1	0	0	0
Decision 2	0	1	1	0	0
Decision 3	0	1	1	1	0
Decision 4	0	0	1	1	0
Decision 5	0	0	0	1	1

We then measure the ideal points using our model—the only change is to select the appropriate link function for this kind of data, i.e. a logit link function. The model in the main paper is not completely identified and it suffers from rotational invariance. To account for this issue, we estimate the model using multiple chains and then calculate the median of all chains for each parameter. If the sign of a chain differs from the first chain, we multiply all values in that chain with -1 . Overall, we take 80 000 draws from the posterior.

The experiment is encouraging: Figure A.1 shows that as expected all estimates are distributed with equal distance in the case space. Credible intervals are also well behaved.

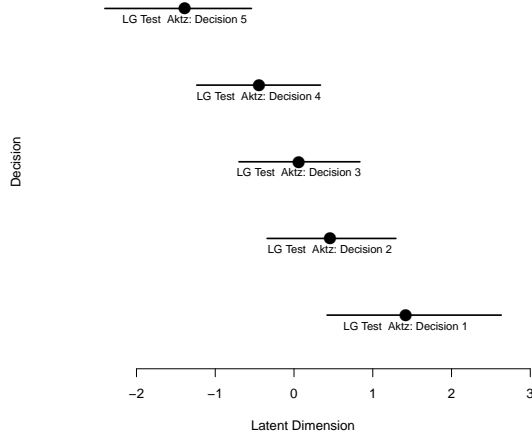


Figure A.1: Estimating the Position of Decisions on the Basis of the Data in Table A.1

Recording only whether a court cites a legal source or not is quite a strong assumption. It seems much more realistic to also think about how often a court is citing a sources—which is what we actually do in the paper. While a decision considers a dissenting legal source, it might cite it only once or twice. But a legal source that is relevant will be cited much more often. We therefore also estimate ideal points on the basis of the more realistic data structure in A.2. All decisions

Table A.2: Toy Data 2: How Often Does a Decision Cite a Source?

	Source 1	Source 2	Source 3	Source 4	Source 5
Decision 1	10	5	1	1	1
Decision 2	1	10	5	1	1
Decision 3	1	4	7	4	1
Decision 4	1	1	5	10	1
Decision 5	1	1	1	5	10

cite all sources—however they do so with a different frequency. The resulting citation graph with weighed edges can not be estimated with a logit link function. Count models such as the poisson link function allows to appropriately take account of the data generating process.

We use the exact same setup to estimate the bayesian Model—including also the chain flipping to solve rotational invariance—and sample overall 80’000 draws from the posterior. Results

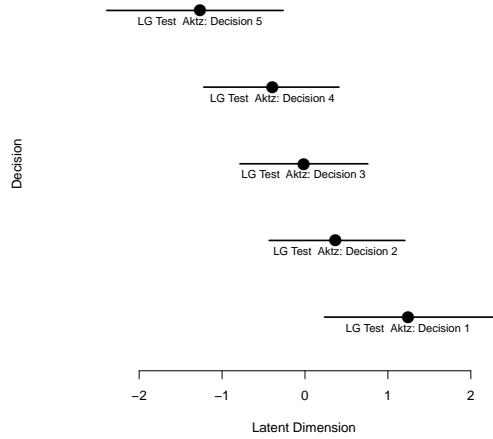


Figure A.2: Estimating the Position of Decisions on the Basis of the Data in Table A.2.

in figure A.2 show that the model retrieves ideal points that reflects the data. Credible intervals indicate that the model is capable to handle the count input data from table A.2 well.

B Enhancing the Query

The first step in data collection is to identify a suitable set of decisions that belong to the same legal case-space.

B.1 Dictionaries for Querying the Data Base

For press law, we are accessing Juris data directly with a *Lucene* based search engine (*Elastic-Search*). We define the following two dictionaries for the initial seed query. Table B.1 relates to claims for compensation. In a similar vein, Table B.2 concerns claims that demand injunction.

B.2 Dictionaries for Querying Juris on their Homepage

To analyze cartel cases, we query the Juris homepage directly. The two terms that make up the dictionary are *Kartellrecht* and *Schadensersatz*. The dictionary is much shorter than for press law,

Table B.1: Dictionary to Query the Data Base for Decisions on Privacy Infringement Claiming Compensation.

Presserecht, Presse, Pressefreiheit, Presseerzeugnis, Äußerung, Interview, Darstellung, Medien, Meinungsfreiheit, Meinung, Meinungsäußerung, Persönlichkeitsrechtsverletzung, Persönlichkeitsrecht, Schadensersatz, Schadensersatzanspruch, Schadensersatzberechnung, Schadensersatzklage, Schadensersatz, Schadensfeststellung, Schadenshöhe

Table B.2: Dictionary to Query the Data Base for Decisions on Privacy Infringement Demanding Injunction.

Presserecht, Presse, Pressefreiheit, Presseerzeugnis, Äußerung, Interview, Darstellung, Medien, Meinungsfreiheit, Meinung, Meinungsäußerung, Persönlichkeitsrechtsverletzung, Persönlichkeitsrecht, Unterlassung, Unterlassungsanspruch, Unterlassungsklage, Unterlassungsverfügung, Unterlassungserklärung, Unterlassungsangebot, Unterlassungspflicht, Unterlassungsantrag, Unterlassungsverpflichtung

because we have to abide by the more limited query functionality of the *Juris* frontend. We also restrict the search to lie between 01.01.2012-01.01.2018, with the author being a regional court (*Landgericht*).

B.3 Calculating the Similarity between Query Terms and Documents

For our four analyses on press law, we query the data base with the search engine. We first use a seed list with terms we are interested in (Table B.1 and Table B.2). But to expand our sample, we also query the data base with the titles of the decisions that are in these two sets. How does the search engine evaluate the similarity between the input we provide and the corpus in the data base?

This is what is happening from a technical point of view.¹ When looking for similar legal documents in the data base, the algorithm considers the complete text of all decisions. In a first step, the data has to be converted into a format that allows computers to calculate the similarity between the query document and all other documents in the data base. All documents in the data base can be represented with one large term-document matrix that contains the whole vocabulary of the corpus on the first dimension. The second dimension holds the word counts of each document

¹For a more in-depth treatment, see [Manning, Raghavan and Schütze \(2009\)](#).

in the corpus.² Each document can therefore be represented with a vector \vec{v}_i that contains all counts for all words. This vector \vec{v}_i is as long as the size of the vocabulary in the corpus. It turns out that the cosine between two vectors \vec{v}_q and \vec{v}_d is a very good measure to calculate the similarity between a query document q and any other document d in a corpus. This cosine similarity is calculated as

$$\text{cosine similarity}(q, d) = \frac{\vec{v}_q \cdot \vec{v}_d}{|\vec{v}(q)| \cdot |\vec{v}(d)|}$$

with the numerator being the scalar product between the two vectors \vec{v}_q and \vec{v}_d . The denominator standardizes both vectors with the product of their Euclidian lengths.³ Implementing this measure, we concatenate the titles from the query set to one single query document d and find the most similar legal decisions available in the data base using the cosine similarity between the tf-idf weighted vectors \vec{v}_q for the query set and \vec{v}_d for all others.

C Decision Source Matrix

C.1 Press Law: Hand Selected

Figure C.1 provides an overview over the resulting decision-source matrices Y_{ij} . Each row represents a legal decision, each column a legal source. The darker the color, the more often does a decision cite a particular legal source. For the cases on compensation, the left figure displays a number of long vertical lines that indicate a high degree of overlap: These decisions cite similar legal sources—but some more and others less often. In addition, there are a number of legal sources that get cited by only a few legal decisions. The decision-source matrix for the second application looks quite similar, with some legal sources being standard sources and others that are picked up by

²In practice, these counts are being weighted calculating the *term frequency-inverse document frequency (tf-idf)*. This score corrects the pure counts of words in a document with the goal of generating a score that reflects the importance of a word in a document. The score takes into account how often a word occurs relative to all other words and it also corrects for the different lengths of documents.

³The *Lucene* scoring built into the *ElasticSearch* search engine further refines this measure and allows for more fine grained specifications of search queries. However, these scores are not relevant here. For further information please refer to https://lucene.apache.org/core/4_9_0/core/org/apache/lucene/search/similarities/TFIDFSimilarity.html (last accessed April 2020).

a subset of the legal decisions, only. In conclusion, the respective case-spaces seem to be narrowly defined. The decisions draw on a similar set of legal sources, allowing to meaningfully locate and compare the decisions in the respective uni-dimensional space.

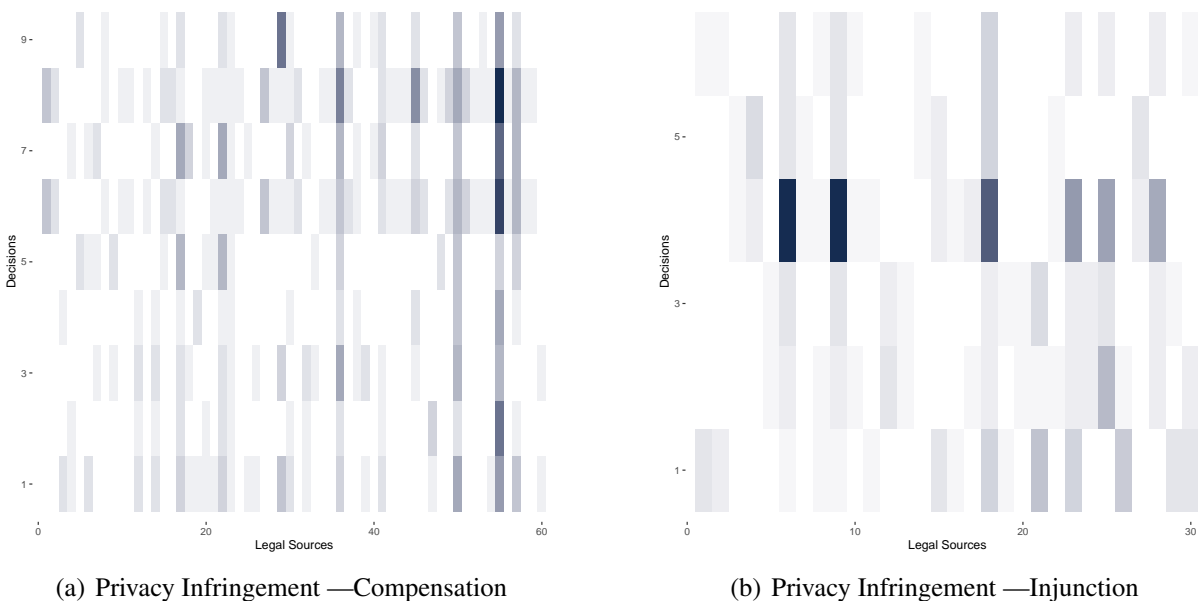


Figure C.1: Exact Query. Coding of the Decision-Source Matrix on the Basis of Metadata. The Darker the Shading, the More Often a Legal Decision Cites a Legal Source.

C.2 Press Law: Hand Selected and Query Expansion

For the decisions on compensation, the matrix is quite well-behaved and shows a substantial degree of overlap between those decisions and the cited legal sources. We can clearly identify five legal sources that are being cited by a large number of decisions. In addition, there are legal sources that are being cited only by some decisions, which in the end provide most of the analytic leverage. The decision-source matrix of the second set of written decisions on injunction shows less overlap. Only two legal sources are apparently widely cited.

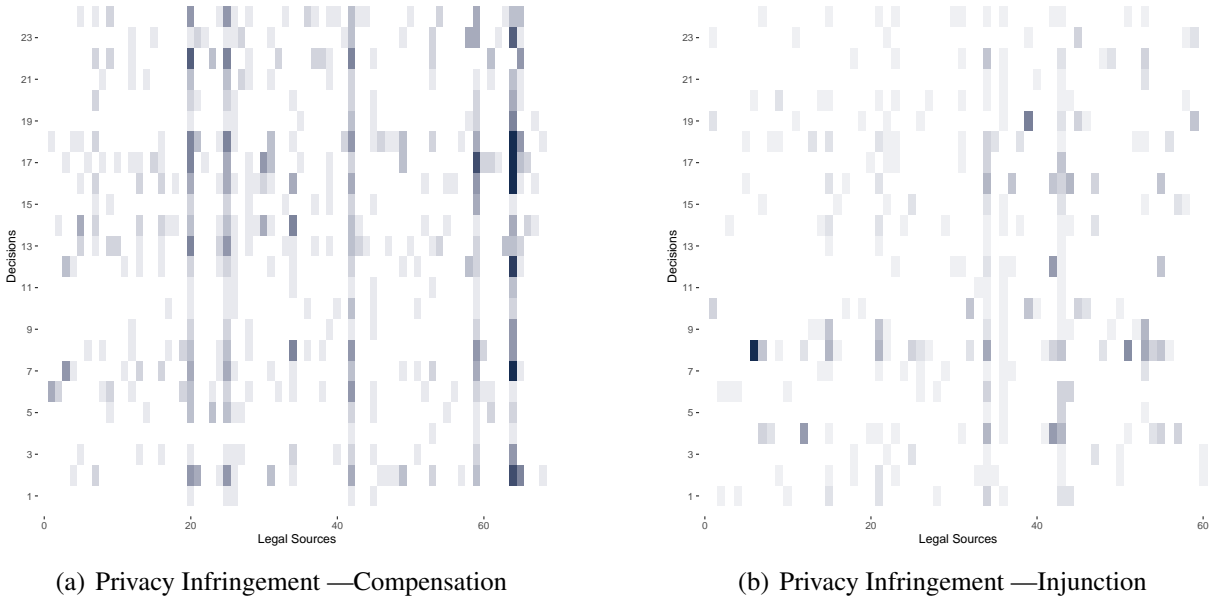


Figure C.2: Set of Decisions Using a ‘More-Like-This’ Query. Coding of the Decision-Source Matrix on the Basis of Metadata. The Darker the Shading, the More Often a Legal Decision Cites a Legal Source.

C.3 Anti Trust: Hand Coded

We also chart the decision-source matrix for the set of decisions in antitrust in Figure C.3. This set has considerably more legal sources than the other sets in press law. Due to the annotation process of the legal sources, they are ordered chronologically—hence the sparse triangle in the lower right corner. Clearly, there seems to be a core doctrine that courts typically refer to. It is easy to identify it on the left of the figure in darker shading.

D Estimating the Location of Sources

In addition to estimates for the location of a written decision, our model also estimates the location of the cited legal sources within the same case-space. As an illustration, Figure D.1 shows the estimated locations ($\hat{\phi}_i$) of every cited legal source for the two sets of decision in press law without query extension. The other sets of decisions contain even more legal sources which would make them difficult to display at the detailed level.

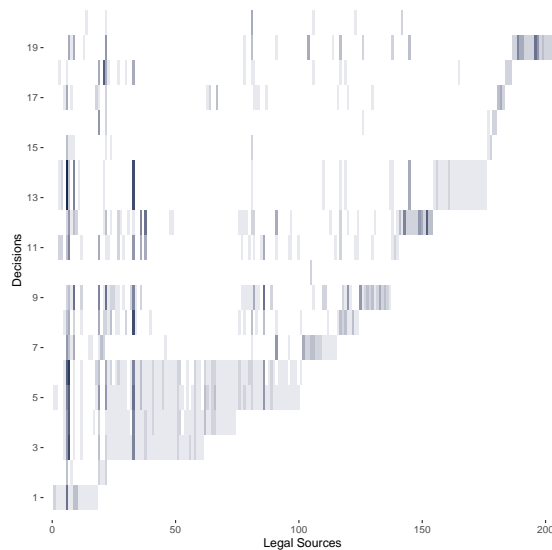


Figure C.3: Set of Decisions Accessed via *Juris* Homepage. Coding of the Decision-Source Matrix on the Basis of Metadata. The Darker the Shading, the More Often a Legal Decision Cites a Legal Source.

Thus, we do not only scale decisions but also legal sources in the same case-space. This can facilitate more fine-grained substantive interpretations of the legal argumentation that is developed within those decisions because the type and frequency of the cited legal sources provide additional information that has not been leveraged before. For our application of compensation cases we see that we could identify a greater variety of legal sources that cluster around -1 and $.5$ respectively than in our application of the injunction cases which are more spread out and, hence seemingly more informative about the location of the citing decisions.

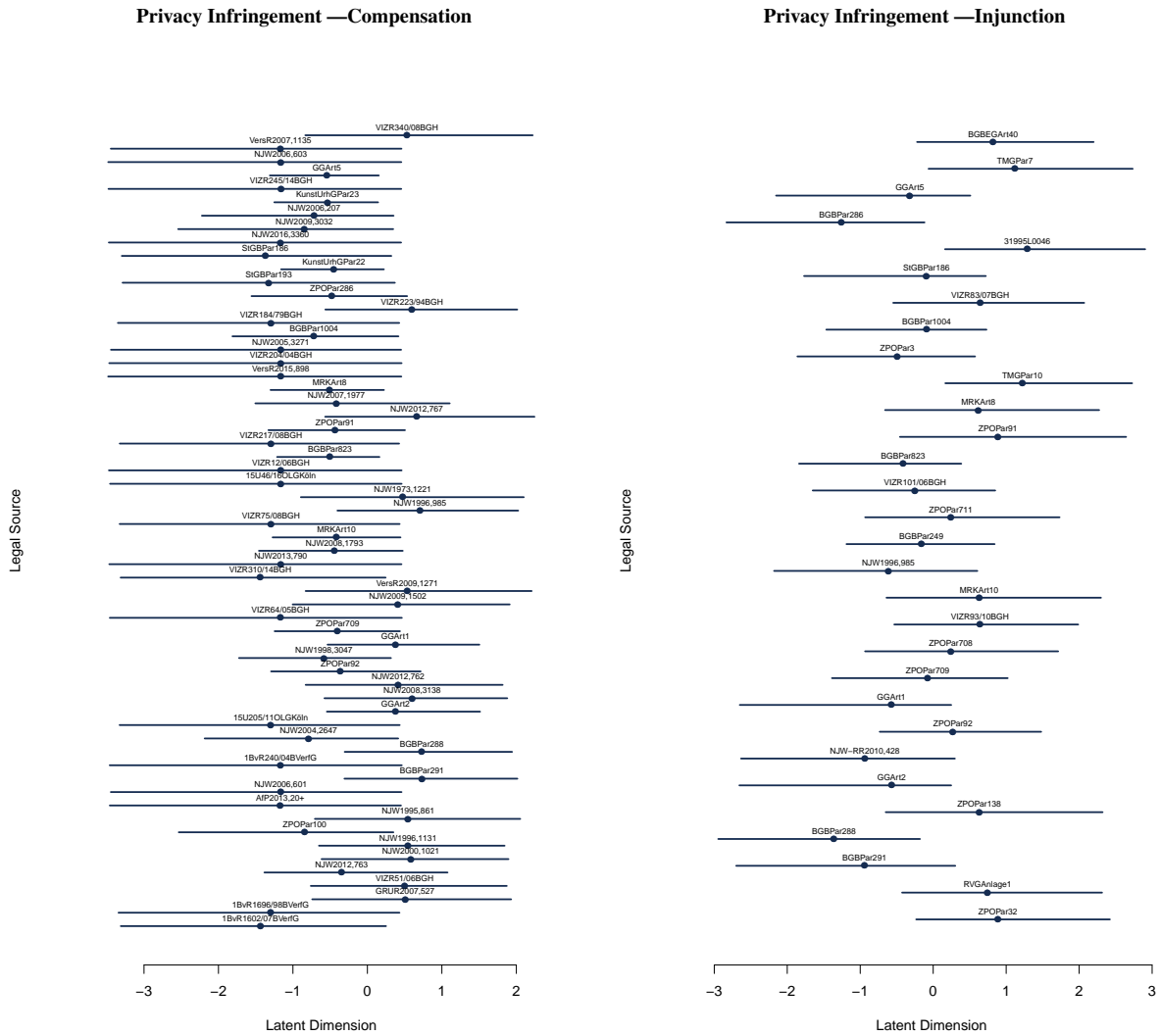


Figure D.1: Estimated Locations of Cited Legal Sources ($\hat{\phi}_j$). Sets of Decisions Without Query Extension (d_{1A} , d_{1B}). Points Indicate the Median of the Posterior Draws. The Bars Represent the Central 90% Credible Interval.