

Why and when do judges dissent: an empirical analysis of the Czech Constitutional Court

ŠTĚPÁN PAULÍK, Humboldt Universität zu Berlin, stepan.paulik.1@hu-berlin.de
GOR VARTAZARYAN, Charles University, gorike2000@gmail.com

The decision of a judge to dissent or not to dissent opens up avenue for strategical considerations. Building on the economic-strategic account of judicial behavior developed by Lee Epstein, Richard A. Posner and William M. Landes, we develop and test multiple hypotheses on the Czech Constitutional Court. To test the hypotheses, we utilize Bayesian regression analyses. We find that the workload of a judge does not affect their dissenting behavior as previous research in the US context suggests. We also find that a dissent imputes substantial costs on the majority that produces longer arguments to address a dissent. The effect is stronger the more disagreement there is on the bench. Lastly, we cast doubt upon the theory that dissents bring about significant collegiality costs to the dissenter

Keywords: empirical legal research, courts, dissents, judicial behavior, political science, Bayesian statistics, regression analysis

Empirical legal research has been slowly but surely finding it's outside the predominant US context. Historically though most of the empirical studies have been conducted in the US, especially the Supreme Court, context (such as Boyd, Epstein, and Martin 2010; Carrubba et al. 2012; Epstein, Landes, and Posner 2011). We now know that the judicial decisions are what judges had for breakfast or that judges are always pursuing political goals or enhancing their career. What we do not know is the extent to which these conclusions carry over to other legal systems and context.

In our article, we set out to replicate one of the empirical legal research studies conducted in the US context elsewhere. More specifically, the main goal of our article is to replicate the method and the findings of Epstein, Posner and Landes (Epstein, Landes, and Posner 2011) on the Czech Constitutional Court ("CCC") and to test whether the same conclusions hold in a different context

The authors study under which circumstances do US judges generally dissent. They build a formal economic model based on the strategic account of judicial behavior. The study's empirical part is firmly rooted in their theoretical part: they draw and test multiple hypotheses, which flow from their theory. In particular, they test the the dependence of dissent rate on workload, the dependence of dissent rate on size of courts, the dependence of dissent rate on the ideological distance, and the dependence of length of majority argumentation on the presence of a dissenting opinion. We will test similar hypotheses on the CCC based on the same theories.

We adapt the theory to a different context. Based on the adaption, we draw hypotheses that

resemble the original ones. We test whether the length of majority argumentation depends on the presence of one or more dissents due to, whether the workload of a judge affects their dissenting behavior, whether relationships from plenary sessions carry over to 3-member panel proceedings, and, lastly, whether the dissenting behavior of judges changes at the start and end of their terms.

To test the hypotheses, we utilize Bayesian regression analyses. We find that the workload of a judge does not affect their dissenting behavior as previous research in the US context suggests. We also find that a dissent imputes substantial costs on the majority that produces longer arguments to address a dissent. The effect is stronger the more disagreement there is on the bench. Lastly, we cast doubt upon the theory that dissents bring about significant collegiality costs for the dissenter.

Our article proceeds as follows. We start out with a theory.¹ We explain the main differences between the expectations based on the theory in the CCC context in comparison to the SCOTUS context and based on that we draw the hypotheses for the empirical part. We then explain the choice of our broad methodological framework: the Bayesian statistics. We proceed to test the hypotheses in empirical part divided into sections one per each research question. Lastly, we discuss the results.

Theory

In general, there are multiple accounts of behavior of judges'. The first that had dominated until ~the end of 20th century posited that judges are policy oriented. A lot of research has been conducted on whether, how and to what extent do judges indeed seek to advance the policies they desire (Berdejó and Chen 2017; Clark and Lauderdale 2010; Dworkin 1980; Kastellec 2016; Moyer and Tankersley 2012).

However, as of recently, the perspective on judges has shifted. Judges are now allegedly strategic actors. Instead of pushing their political agenda, they now push their self-interest in terms of career progression, higher income, or lesser workload (Carrubba et al. 2012; Epstein and Knight 2000).

In their article Epstein et. al test multiple hypothesis about the circumstances under which do the judges of the US courts dissent or under which they decide to not dissent (dissent aversion). Their analysis is based on the strategic-economic framework of self-interested strategically

¹We need to develop it more in depth, Gor is currently working on it

motivated judges. They presume that judges “leisure preferences, or, equivalently, effort aversion, which they trade off against their desire to have a good reputation and to express their legal and policy beliefs and preferences (and by doing so perhaps influence law and policy) by their vote, and by the judicial opinion explaining their vote, in the cases they hear.”

The authors find the strategic aspect of dissenting in how a judge squares their decision to either dissent or to avert a dissent based on the costs and benefits of a dissent. The authors claim that “[s]ince writing a dissenting opinion requires effort, which is a cost, a judge will not dissent unless he anticipates a benefit from dissenting that offsets his cost.” The majority also accrues costs from dissenting. In the words of Epstein et al.: “[d]issenting imposes an effort cost on the majority as well and sometimes a reputation cost too, if the dissenting opinion criticizes the majority forcefully. To minimize the dissenter’s criticisms and retain the vote of the other judge in the majority (in a panel of three judges, the normal number of judges who decide a case in the federal courts of appeals), the author of the majority opinion often will revise his opinion to meet, whether explicitly or implicitly, the points made by the dissent.”

The benefits are the potential to undermine the majority opinion when the dissent is influential and the enhanced reputation that the judge enjoys. The dissenting opinion may be cited in the future by other judges.

The theories they presume and hypotheses they test rest on this framework: in the policy-oriented framework, it would not make sense to expect judges less as their workload increases. They would still seek a way to advance their political agenda and research has shown that dissenting opinions usually correspond to exactly just that (Clark and Lauderdale 2010). However, in the strategic account, the higher the workload of a judge, the more pressing the effort costs of a dissent. Similarly, if a dissenting opinion imputes costs on the majority, we can theoretically expect it to respond to the dissent with a more thorough or detailed argumentation in the majority opinion. It is safe to say that the research questions are grounded in theory.

Our hypotheses and research design differ. Our article focuses only a single court, the CCC, rather than focusing on multiple courts. It also does not replicate the political distance research question. It follows that our research questions are slightly different mainly for two reasons.

Firstly, a major obstacle in conducting and carrying the US research elsewhere is data availability. We narrow our object of analysis to the CCC because there is the largest variation in the

dissenting behavior (unlike on the Supreme Administrative Court). While there is not yet a full fledged dataset on the CCC (like the SCOTUS or CJEU Brekke et al. 2023), we have managed to build a complete yet unreleased dataset on the CCC, which includes the metadata about the cases as well as the text corpus. However, Epstein, Posner and Landes include in their analysis the ideological distance between judges. The ideological distance serves as one of the explanatory variables for dissent aversion. The measures of ideological position of judges mainly rely on information about their voting behavior. Regrettably such information is in continental legal systems typically not made public: the votes in cases are kept secret. Therefore, it is near impossible to construct a measure of the political position of judges without knowing how they voted in each case.

Secondly, we utilize variation of different institutional settings, which allows us to answer very similar questions. While Epstein et al. explore the variation between SCOTUS and Federal Courts, we're able to explore variation within the CCC. The CCC is structured so that it can either decide cases in 3-member panels or a plenary session. Moreover, we are able to use the variation of a judge rapporteur.

Thus, on the one hand, we narrow our inquiry in comparison to Epstein et al. and omit the research questions that require the knowledge of the ideological position of the judges. On the other hand, we are able to utilize the variation at the CCC that's not present at the SCOTUS: namely the chamber system and the limited terms of CCC judges. As a result of that we are able to add two additional research questions.

Literature on judicial coalitions at the CCC has revealed that the third period of CCC is the most polarized, and there are clearly two big coalitions of judges that clash against each other. However this literature focuses only on plenary decisions and is of rather superficial descriptive nature. We hypothesize that the relationship from the plenary sessions carry over to the 3-member panel hearings. Our hypothesis is that senates composed of judges from both coalitions will be more likely to show disagreement in the form of dissenting opinion. If this shows to be true, it would provide further evidence to the two coalition theory of the CCC (Chmel 2021; Vartazaryan 2022).

Epstein et al. base their theory on an economic model of dissenting, in which a decision to dissent imposes collegiality costs on the dissenter. They then test whether the variation of panel

sizes impacts the likelihood of dissent, as bigger chambers impose lower collegiality costs. Similarly, one of the profits of a dissent may be increased reputation costs (as explained below). We test whether the dissenting behavior of CCC judges changes at the end of their term. The CCC judges enjoy a 15 years long limited term, which they may repeat if they go through a re-election before the Senate. Our hypothesis is that as the collegiality costs decrease before the end of the judges' terms, their likelihood to dissent may go up. We will also analyse whether judges changed their dissenting behavior depending on whether they decided to face a re-election or not.

We therefore test the following hypotheses:

RQ1: The presence of dissent positively affects the length of the majority argumentation.

RQ2: The higher the workload of a judge, the lower their dissent rate as a result of higher effort costs.

RQ3: The judges are more likely to dissent at the end of their terms as the collegiality costs decrease.

RQ4: Judicial coalitions formed in the plenary proceedings affect the likelihood to dissent in 3-member panels²

Method

Our goal, as we have noted, is to replicate the Lee Epstein et al. study in the context of the CCC. The original study employs quantitative methods, namely a linear regression (including a log transformation in one case) on observational data.

We will deviate in some aspects from the original study. Firstly, we utilize the *Bayesian* rather than *frequentist* framework of statistics. Secondly, we adapt the statistical models to a different regulatory and institutional setting of the CCC. Thirdly, we change the statistical models so that only relevant explanatory and confounding variables remain. We now discuss the reasoning behind opting for Bayesian statistics as the framework for our quantitative analysis. We address the adaptation of the models and the selection of variables for each research question separately.

²We will add this one later

Bayesian framework

Without delving too much into the Bayesian versus frequentist statistics, we opt for the Bayesian framework for it, we believe, reflects better our understanding of probability and scientific inquiry. There are two major differences in understanding of concepts between the two approaches towards statistics: that of role of prior knowledge and that of probability.

In the frequentist framework, prior knowledge does not play too much of a role and the inference is shaped solely by the observed data, whereas in the Bayesian framework prior knowledge is updated with new data to form new posterior conclusions. In other words, the Bayesian statistician concerns themselves not with the uncertainty of the data but also with how it fits into his prior knowledge.

That is reflected in different understandings of probability. The frequentist understanding of probability refers to the long-run relative frequency of a repeatable event. In other words, the main concern of frequentist statistics is what would the frequency of any event be if we could repeat it as many times as possible. The Bayesian probability measures the relative plausibility of an event (Johnson, Ott, and Dogucu 2022).

Science in general is based on the frequentist framework. The typical quantitative studies are driven by finding a low enough p-value, i.e., the measure of probability of having observed data as or more extreme than the observed data if in fact the original null hypothesis is incorrect. In simple terms, the search for statistical significance is a search for data so unlikely to have occurred due to chance, even if we could gather them again and again.

The Bayesian framework rather than measuring the uncertainty about observed data measures the uncertainty of the parameters of interests, given the observed data and our prior knowledge. In simple terms, the Bayesian statistician puts into doubt their conclusions about parameters of a certain model, given the observed data and their prior knowledge. Mathematically, the uncertainty is reflected in the fact that the posterior parameters are drawn from a posterior distribution of the model and are just an approximation of thereof in the form of probability density function rather than a single value. The posterior distribution of a parameter comes from simulating in our case 40000 (4 chains*10000 simulations) possible posterior models via the Monte Carlo Markov Chain simulations.

The Lee Epstein et al. original study employs the frequentist framework and hence they report statistically significant relationships between the variables of interest. Our approach will be Bayesian. We will firstly see whether the employed models actually make sense for the data by running posterior predictive checks and afterwards we will draw the parameters from the posterior distribution to see whether their distribution indeed looks similar to the ones reached by Epstein et al.

Effect of presence of dissenting opinion on the length of majority argumentation

According to Epstein, Landes and Posner, “[a] dissent imposes an effort cost on the majority because the author of the majority opinion is likely to revise his opinion to address the objections raised by the dissent. This suggests that the majority opinions will be longer when there is a dissent.”

To test this hypothesis, Epstein, Landes and Posner collected roughly 446 SCOTUS and 1025 US court of appeals decisions. They then create regression models to see whether the presence of at least one dissenting opinion affected the length of the majority opinion. The regression model of Epstein et al. naturally controls for multiple covariates, so that bias is eliminated and a causal interpretation of the result is made possible. The outcome variable of their model was the number of words of the majority opinion. The explanatory and control variables in their model are:

- (1) whether the decision included oral hearing,
- (2) and (3) a dummy variable for presence of 1 or 2 and more dissents (the explanatory variable),
- (3) whether the majority mentioned the dissent,
- (4) a presence of concurring opinion,
- (5) a dummy for subject matter,
- (6) a dummy for the term of the court,
- (7) importance of the case, proxied by the number of references to the case in caselaw of SCOTUS and courts of appeal

Based on application of their model to the SCOTUS data, the trio of authors found a statistically significant and positive relationship between presence of at least two dissenting opinions and the length of the majority decision. Unsurprisingly, the authors also found a statistically significant positive relationship between the outcome variable and the importance of the case.

We test a very similar hypothesis with a slight tweak. We assume that what becomes longer as a result of a presence of a dissent is not the majority decision as such but its argumentation part, i.e., we are not including, for example, the heading or the procedural history, which both make up the majority decision. A following hypotheses can be distilled:

RQ₁: The presence of dissent positively affects the length of the majority argumentation.

Adapting the Epstein et al. regression model to the CCC context

There are multiple obstacles we had to overcome to conduct a similar model on the CCC. Firstly, while we are aware that the importance or salience of a case is probably the key confounding variable, measuring as a post-treatment amount of references of that decision might introduce further bias. Secondly, we think inclusion of some of the covariates is unnecessary and may even constitute a bad control, insofar we do not think there is any potential avenue for confounding. Here we present a brief overview of issues we have with the model and how we solve them.

Conceptualizing importance of a case as a control variable

In general, when utilizing a regression design with observational data, as is our case, the researcher must satisfy certain conditions to be able to interpret regression results as a causal relationship. The first is the stable unit treatment value assumption (SUTVA) and the other is usually referred to as *conditional independence assumption*. This condition requires that the assignment of treatment T to unit i is independent of any covariates X of the unit i that also influence their outcome Y . The CIA may be formalized as follows:

$$Y_{1i}, Y_{0i} \perp T_i | X_i$$

The notation follows the potential outcomes framework. $\{Y_{1i}, Y_{0i}\}$ refer to the outcome of a unit i with or without treatment, in our case the presence of at least one dissenting opinion. There are, in general, two types of causes of bias: a confounding variable, which breaks the CIA, and reverse causality. A confounding variable is such that

- (1) has an effect on treatment status,
- (2) has an effect on the outcome over and above its effect on the treatment status.

Not controlling for confounding variables causes an omitted variable bias that in turn precludes causal interpretation of the regression.

While at first glance it may thus seem that researchers should throw in as many covariates as possible, that is in reality not the case. There are examples of bad or unnecessary controls that are themselves an outcome of the treatment (for a detailed discussion see Angrist and Pischke 2009, 2014; Montgomery, Nyhan, and Torres 2018). Among them are post-treatment variables, which imply that all control variables must occur before the treatment takes place.

We believe that importance of a decision is a potential confounding variable as it clearly may impact both the length of a judgment as well as the likelihood of a dissent. However, proxying it by the number of citations in ensuing caselaw may present a bad post-treatment control variable because it occurs and is measured after the decision to dissent or not to dissent has been made. We believe that in our context, the *formation of the CCC* and the *type of the decision* are better pre-treatment proxies for the importance of a decision.

Institutionally, the CCC can decide cases either in 3-member panels or in the whole plenary session. Put simply, we assume that more important cases are decided in plenary rather than 3-member panel formation. The plenary is more likely to rule on merits and its decisions are, therefore, on average longer. Moreover, the dissent rate in the plenary decisions is also higher. Thus, the formation of the CCC has a confounding potential as it reflects the importance of the case being decided.

We confirm our intuition by comparing our metric of importance of a case to Epstein’s metric in the CCC context. While plenary decisions make up only 1.5 % of all CCC decisions, they make up 15 % references in the CCC caselaw. Unlike the Epstein metric, our metric is in any case determined before the decision to dissent or not to dissent is made. Thus, in our model, we’re including a dummy variable for the formation of a court, i.e., whether the decision was made in a 3 member panel or a full court plenary. We find similar numbers regarding the distribution of decisions on the merits and on the procedure.

formation	Count_total	Percent_total	Count_cited	Percent_cited
Panel	88428	98.4 %	166583	84.8 %
Plenary	1417	1.6 %	29823	15.2 %

Similarly, the type of decision affects both the length of argumentation as well as the likelihood of dissent. The CCC can decide either procedural (“usnesení”) or on merits (“nález”). The latter type of decisions are on average longer and contain disproportionately more dissents than procedural decisions. We assume that the type of decision to some extent also reflects the importance of a case.

Unnecessary or untransferrable control variables

Secondly, we believe not all variables in the Epstein et al. model have potential for confounding or are transferable to the CCC context. The majority mentioning the dissent does not impact the judge’s decision to or not to dissent. What’s more, a mention of a dissent by the majority can only follow after a decision to dissent was made. Thus, we omit it from our model. Oral hearings are few and far between in the CCC context. It therefore makes no sense to control for the presence of oral hearing in the context of the CCC.

Data collection and method

The data used for this analysis includes the complete dataset of 87845 CCC decisions, with short procedural decisions being filtered out from the dataset. The Czech Apex Court dataset was built by Štěpán Paulík and includes complete database of decisions of the CCC, SAC and the Czech Supreme Court, including comprehensive metadata, text corpus, as well as additional information mined from the texts or publicly available sources, such as references, compositions, or background information of the judges. The analysis was limited up to December 2022.

While the metadata of decisions contain information about a presence of a dissenting opinion, regrettably, the CCC decisions are not neatly split up like their SCOTUS counterparts. The text of the decision contains both the majority opinion as well as the dissent. That is why we relied on machine learning to extract the information about presence, position and length of dissenting opinion as well as the majority argumentation in the CCC decisions.

Utilizing machine learning unlocked one more avenue to further improve the Epstein et al. model. Instead of conceptualizing the length of the majority opinion as the whole opinion, we narrowed our inquiry only to the majority argumentation rather than including, for example, the heading or the facts of the case in the length of a majority opinion variable. Therefore, we believe our model better

reflects the relationship between the presence of a dissent and the majority legal argumentation.

To extract the length of dissents and length of majority argumentation, multiple supervised classification algorithms were trained following similar structure-mining attempts within the Czech context (Eliášek, Kól, and Švaňa 2020; Harašta et al. 2019; and elsewhere Lüders and Stohlman 2023). A sample of 200 decisions was manually annotated on a paragraph level. The paragraphs were then represented either as dense doc2vec vectors based on word2vec model of the whole CCC text corpus or as sparse document-term-matrix with the td-idf values for each word. Positional encoding of a paragraph was added to both representations of text. Because there was a large unbalance between the classes, oversampling algorithm SMOTE was applied to balance out the dataset, which is a standard practice when working with smaller datasets (Fjelstul 2021).

Our classification followed in two stages. In the first step, the dissenting opinions were classified from the rest of the decision so that the positional encoding remained consistent across all cases. Otherwise, in the decision without a dissenting opinion, usually a conclusion or court argumentation was at the end of the decision, whereas in the decisions with a dissent, the dissenting opinion was at the end of the decision. That created confusion with the positional encoding. Our first stage classification allowed us to separate the dissenting opinion from the rest of the decision and then recalculate the positional encoding for the remaining paragraphs. In the second step, the remaining text as well as the decisions that did not contain a dissenting opinion were classified into an inner structure consisting of:

- (1) heading
- (2) verdict
- (3) procedure history
- (4) complainant arguments
- (5) court arguments
- (6) conclusion
- (7) information on further legal remedies
- (8) signature

A couple of various models were trained and compared and, in line with the findings of Eliasek and Lüders articles, the choice was narrowed down to two algorithms - Support Vector Machines

(Gandhi 2018) with the sparse td-idf representation and Gradient Boosted Decision Trees with the dense doc2vec embeddings (Maklin 2019). More complex algorithms did not provide any improvement in accuracy at the cost of higher computing costs. The benchmark for both classifiers was the zero rule, i.e., the proportion of the majority class as a zero rule classifier would presume that all occurrences are of that class and would be right the proportion amount of the time.

In the end, the XGBoost algorithm combined with the doc2vec embeddings boasted the highest benchmark values (precision, accuracy, FScore). The precision of the first stage classifier was ~85 %, slightly above the 75 % zero rule benchmark. The precision of the second stage classifier was ~82 %, well above the 37 % zero rule benchmark. Afterwards, both classification models were trained on all annotated data and used to predict classes of the whole dataset.

Result

The model

Building on the theory and the Epstein model, in our model, we included the number of words of the court argument part of a decision as the outcome variable. Regarding the explanatory variables, we opted for

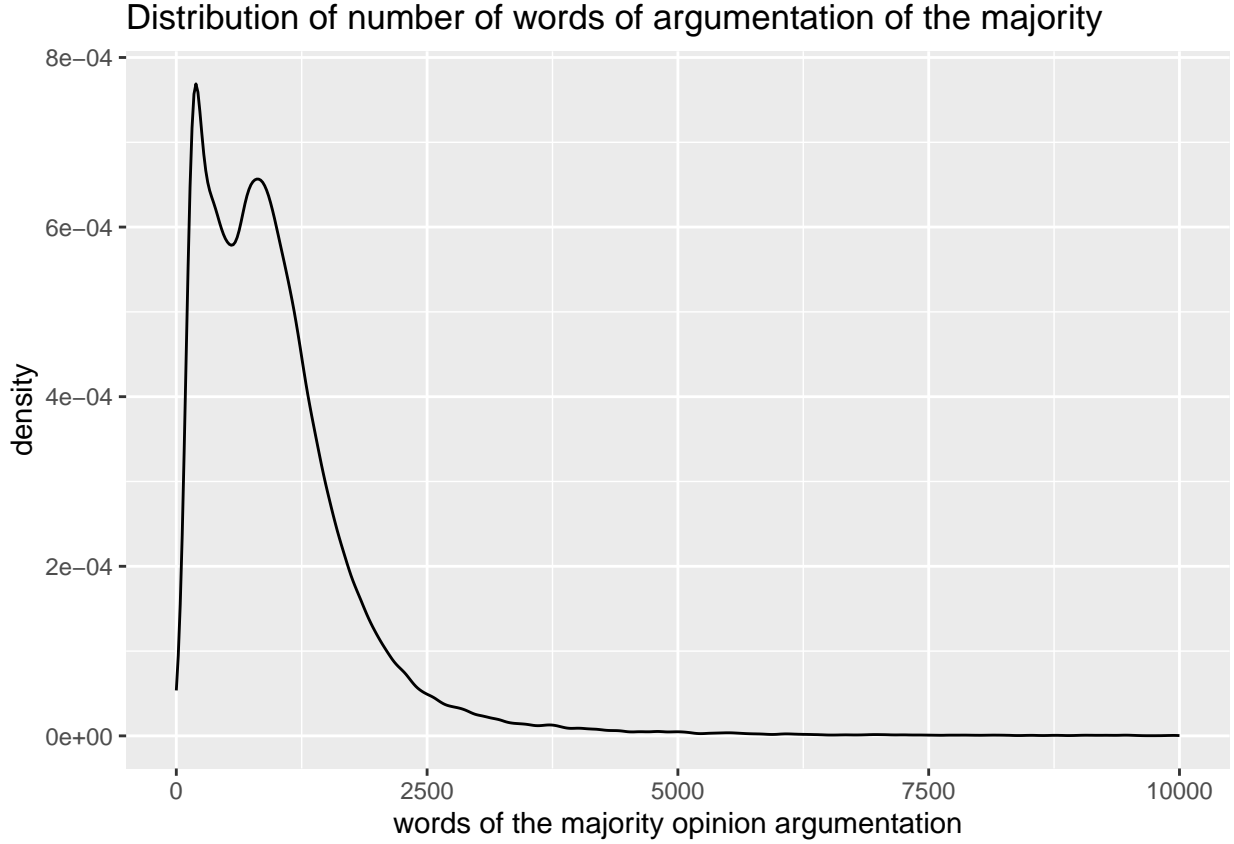
- (1) a dummy variable signifying a presence of one dissent as an explanatory variable,
- (2) a dummy variable signifying a presence of two or more dissents as an explanatory variable,
- (3) the formation of the CCC as a control variable,
- (4) the type of the decision as a control variable,
- (5) year of the decision as a control variable.

We conducted our analysis employing Bayesian regression implemented by the software Stan in R. We opted for a completely pooled model as the data did not contain any inherent structure (there were no clusters). At first glance, we assumed that

$$Y|\lambda \sim Pois(\lambda)$$

because our outcome variable of interest is a discrete count and the density plot of the length

of court argumentation suggests so.



However, as the posterior checking revealed, the Y was actually overdispersed and the Poisson regression was not able to capture the overdispersion. Therefore, we instead opted for the Negative Binomial model, which allows for relaxing the assumption of equality of variance of Y to its expected value. Thus, the explanatory variable, the number of words of argumentation of the CCC Y

$$Y_{words}|\mu, r \sim NegBin(\mu, r)$$

As for the priors, we based the priors on the Epstein results as well as a cursory exploratory peak into the data.

Diagnosis

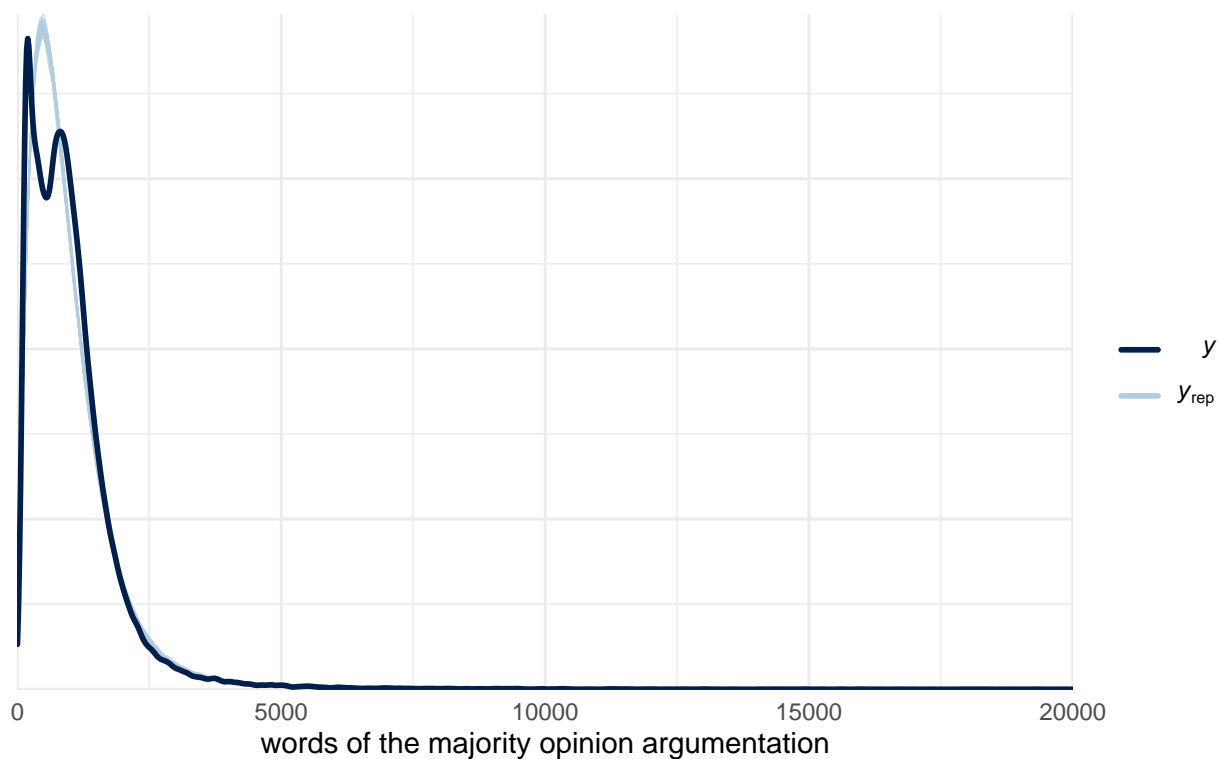
We ran the model via Stan with 4 Monte Carlo Markov Chains (MCMC) of 20000 iterations each, the first 10000 warm up iterations being discarded. The trace plot shows that the chains were stable and probed plausible parameter values, the density plots of the MCMC show that all 4 chains exhibited similar behavior, and the autocorrelation between the iterations always dropped

quickly and that the chains were moving around the potential parameter values quickly.

The posterior diagnosis confirms that although the simulations are not perfect, they do reasonably capture the features of the observed number of words of court arguments. In other words, we selected the correct model and the priors are not too off either. Thus, our Negative Binomial regression assumptions are reasonable.

Posterior predictive check of the Negative Binomial model

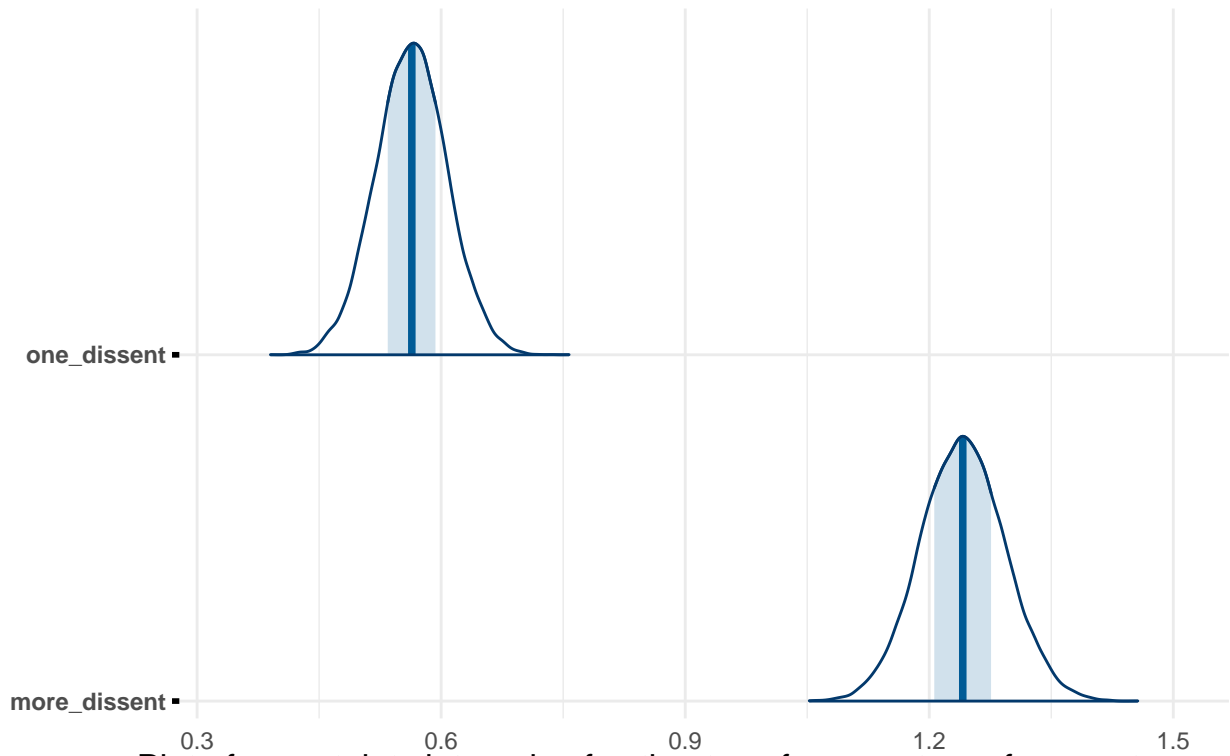
Dependence of length of majority opinion argumentation on presence of dissenting opinion



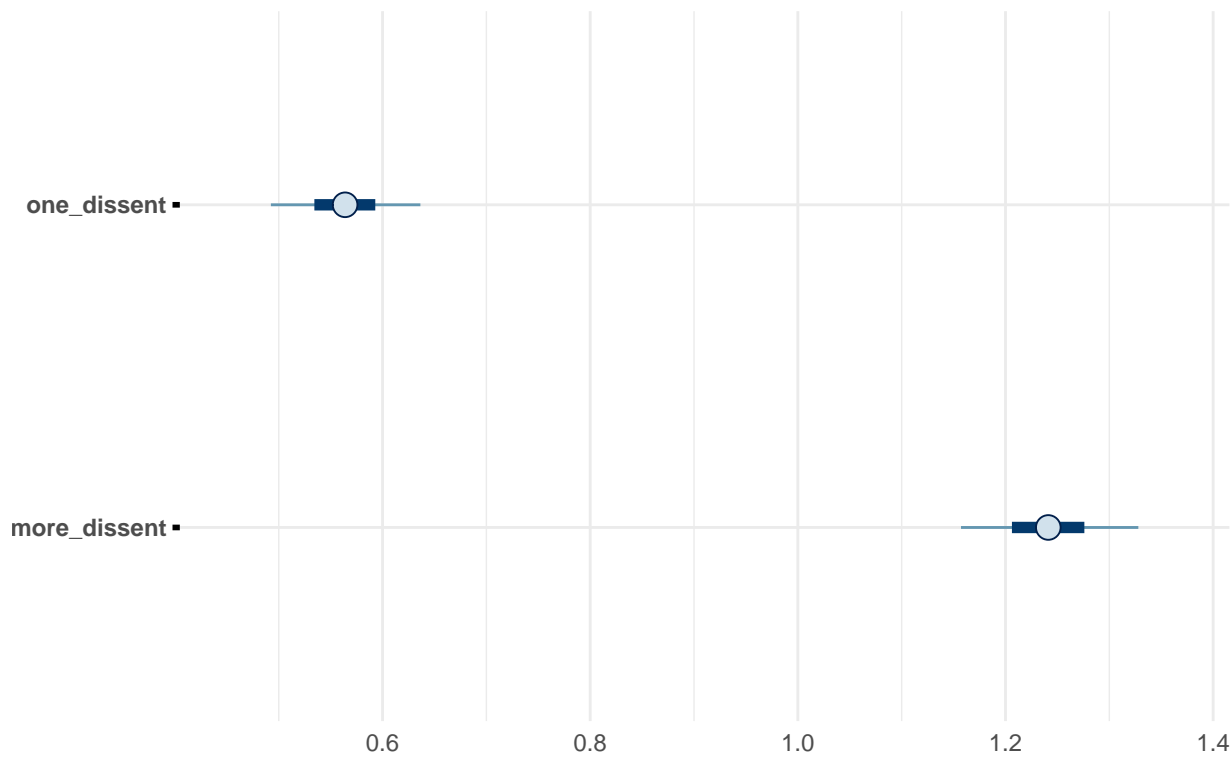
Posterior interpretation

Parameters of all variables of interest are significantly different from 0 as revealed by the density and whisker plots with 80 % and 95 % posterior credible intervals.

Density plot of estimates of parameters of one or two and more disser
The inner area is for 50 % posterior credible interval, the outer for 95 %



Plot of uncertainty intervals of estimates of parameters of one or two α
The inner whisker is for 50 % posterior credible interval, the outer for 95 %



Even after controlling for all potential observable confounding variables, the regression table

looks as follows

term	estimate	std.error	conf.low	conf.high
(Intercept)	7.82	0.01	7.81	7.84
one_dissent	0.56	0.04	0.51	0.62
more_dissent	1.24	0.05	1.18	1.31
formationPlenum	0.27	0.02	0.25	0.30
type_decisionUsnesení	-1.03	0.01	-1.04	-1.01

In other words, the presence of one dissent implies an average $e^{0.561}$ increase of words in the court arguments part of judgment. Put in terms of percentage, a presence of one dissent increases the length of the argumentation by 75 %. The presence of two or more dissents implies an average $e^{1.239}$ increase of words in the court arguments part of judgment. That is a staggering ~234 % increase in the length as a result of presence of two or more dissenting opinions. To this end, the result of our study is in line with that of Epstein et al.: a presence of dissenting opinion increases the length of the majority opinion argumentation.

We believe there are two potential ways to explain this behavior. Either the majority simply takes the dissenting opinion seriously and addresses the arguments raised in them or the presence of a dissenting opinion reflects a deeper disagreement between judges that would have taken place during the deliberation. Based on our knowledge of the inner organisation of the court, the deeper disagreement explanation would fit the plenary proceedings more accurately as a more thorough debate usually takes place than in the 3-member panel proceedings. Such a substantive explanation for our findings is supported by the fact that decisions originating in the plenary proceedings are disproportionately over-represented among decisions that contain at least one dissenting opinion. This explanation is further supported by the larger effect of having 2 or more dissents over just the 1 dissent. More dissenting judges simply imply higher degree of disagreement on the bench.

While unnecessary, we briefly go over the remaining control variable parameters. Unsurprisingly, the decisions on the procedure contain much shorter court argumentation on average. The same applies to the formation as the decisions in 3-member panels also imply much shorter decisions.

Effect of workload on the dissenting behavior

Adapting the Epstein et al. regression model to the CCC context

According to Epstein et al. “[t]he economic theory of judicial behavior predicts that a decline in the judicial workload would lower the opportunity cost of dissenting and increase the frequency of dissents, and also that the greater the ideological heterogeneity among judges the more likely they are to disagree and so the higher the dissent rate will be.” The authors then test both of these hypotheses - whether the workload and the political distance between judges affects dissent rate.

We cannot regrettably measure the ideological distance among CCC judges.³ We omit that from our study and stick only to the effect of changes in the workload of a judge. Our second research question may be formulated as

RQ₂: The higher the workload of a judge, the lower their dissent rate as a result of higher effort costs.

The authors find a positive relationship between the log of dissent rate, i.e., number of dissents divided by the number of cases and log of caseload, i.e., the log of total number of cases decided after oral arguments. The authors find that on the SCOTUS, “a 10 percent decrease in the caseload increases statistically significantly the dissent rate by about 3.3 percent” at a p-value of less than 0.05 Epstein, Landes, and Posner (2011).

The authors again control for multiple variables, which we believe to be unnecessary. They control for ideological differences between judges. It is hard to see how the ideological difference between judges could affect caseload, although it undoubtedly carries an explanatory value for dissent rate variance, as the authors indeed conclude.

Model

Our model is built slightly differently. We utilize the variance in the caseload among judge rapporteurs to measure workload. Hence, we operationalized caseload in two ways:

1. the number of cases submitted and assigned to each judge rapporteur per year

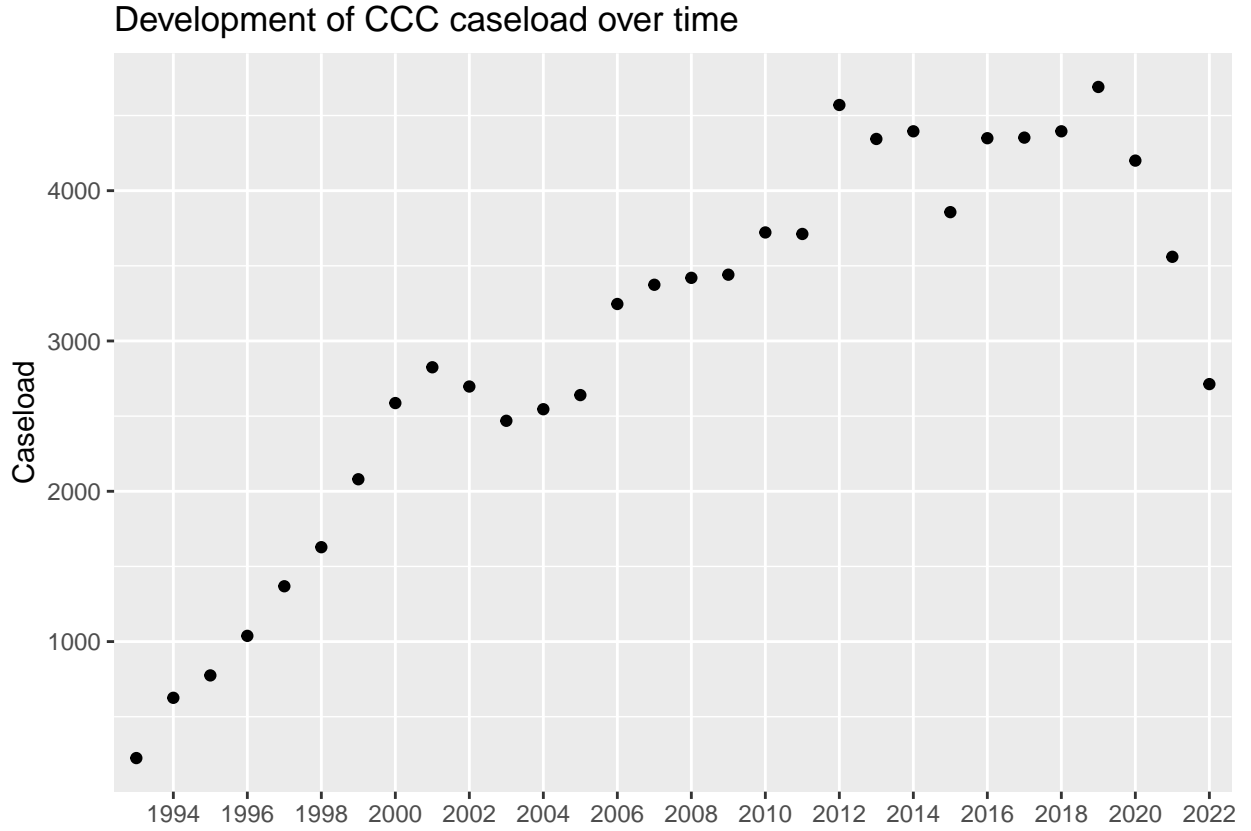
³At least not yet. One of the authors is currently working on a study that estimates the positions of opinions and, by extension, of judges. The extension to judges is built on the assumption that dissenting opinions correspond the most to the judges’ ideological or political positions.

2. as the yearly rate of change of thereof.

In other words, the caseload here is that of an individual judge of decisions that they have to decide and write, or the rate of change of thereof. That in our eyes reflects the idea of workload of each judge on the court. The more cases the judges have to author, the busier they are.

Similarly, instead of measuring dissent rate on the Constitutional court in general, we measure the number of dissents written by each judge either in absolute terms or as a rate relative to their caseload (similarly to Epstein et al.) Ideally, we would measure both variables as rate of change. However, there are many observations of the number of dissents with the value of 0. In these cases, the rate of change would either be infinite or not a number (as a 0 would appear either in the denominator or numerator of the rate of change formula). Getting rid of the zeroes would imply a rather complex transformation (for a more detailed overview of the possible transformations see Hyndman 2010). We include time as a control variable because we presume that the workload of judges increases over time and so may the number of dissents.

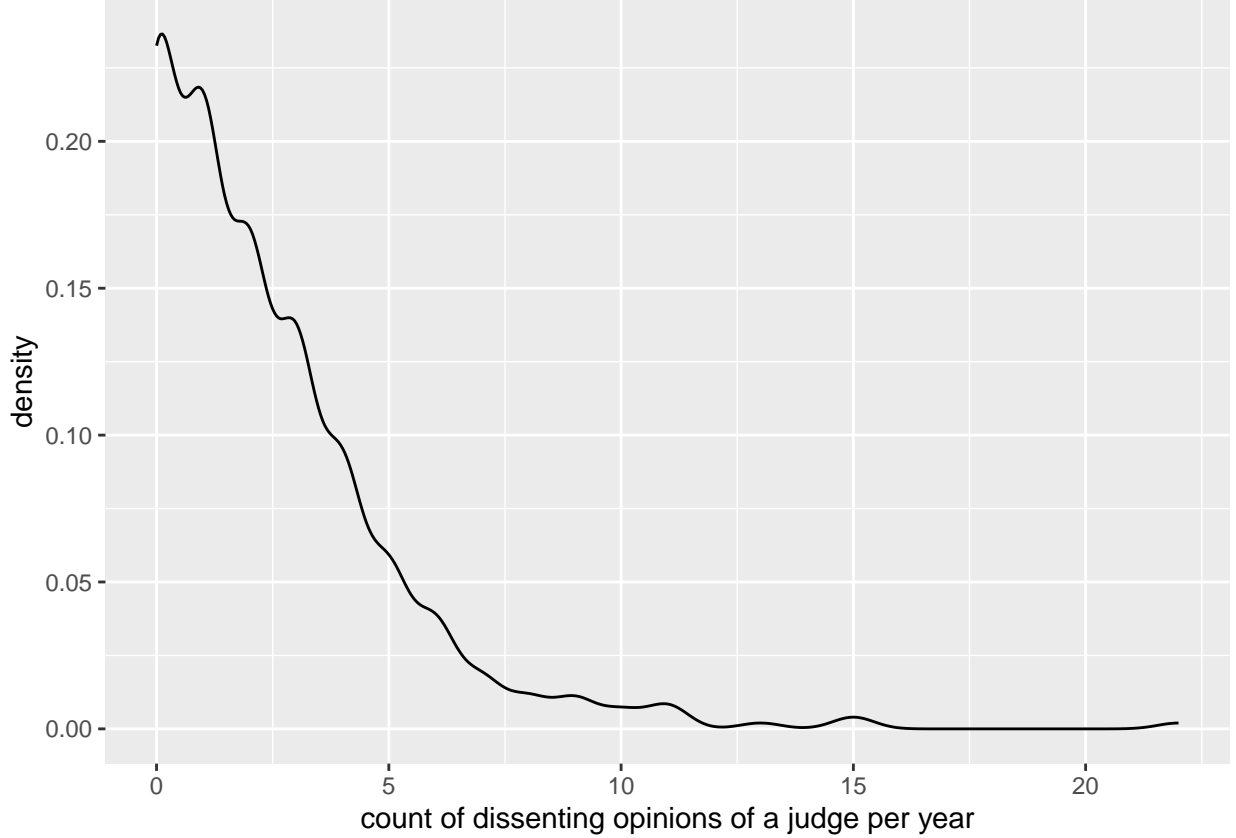
To address potential sources of bias in our regression analysis, we consider the caseload of a judge to be assigned as good as random. The cases once submitted to the CCC get assigned to individual judges based on the alphabetic order of their surnames. There is no intentional case selection in play. Therefore the assignment of the treatment, the workload of a judge, is independent of other covariates and so is the outcome of interest. The same applies to our last research question.



We opt for the Bayesian Poisson regression model. A quick precursory look at our data as well as the fact that the dependent variable is a *count* reveals that our Y , the count of dissenting opinions of a judge per year:

$$Y|\lambda \sim Pois(\lambda)$$

Plotting the density of the number of dissents confirms our intuition. The number of dissenting opinions clearly follows a Poisson distribution that is not over-dispersed to justify Negative Binomial model as in the previous section.



Result

Model diagnosis

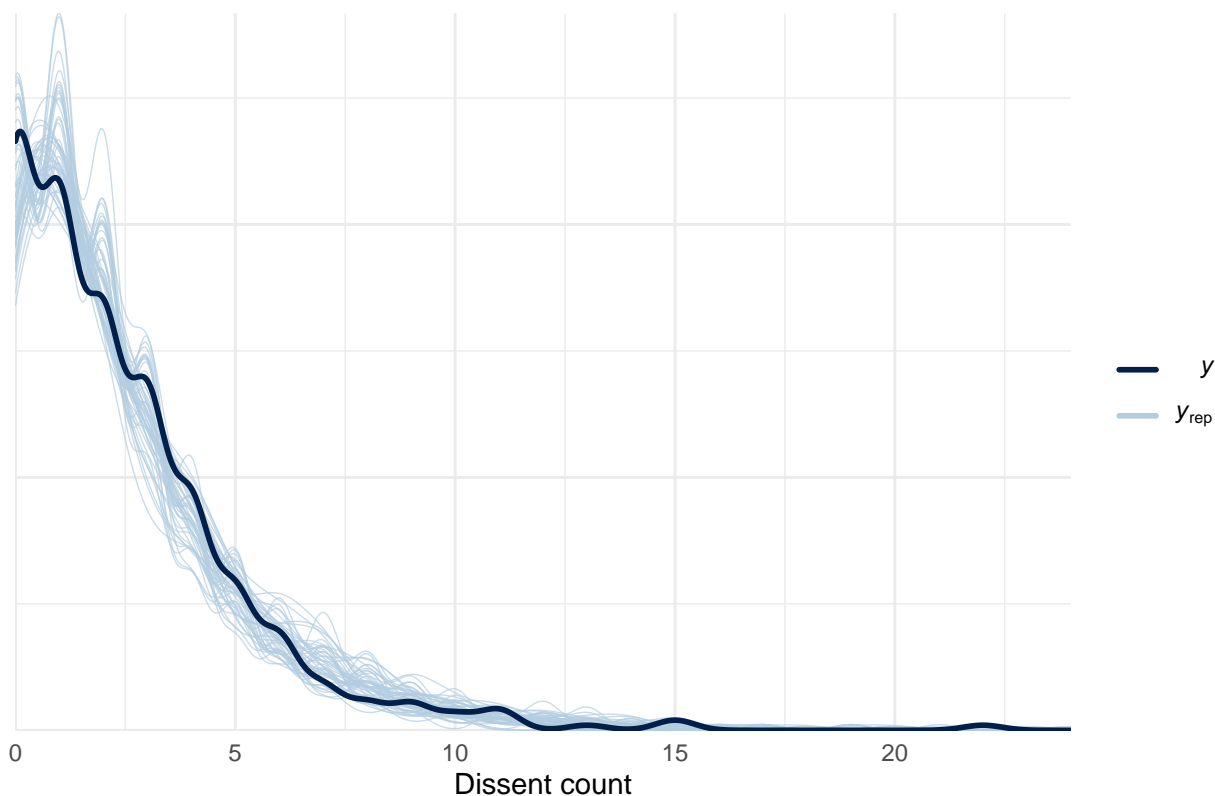
We tried two models. Firstly, we tried a completely pooled Poisson regression. Secondly, we also tried a hierarchical Poisson regression. The main difference between the two models is that the former model completely ignores individual intercept. The latter allows for differentiating intercepts between the groups (in our case judge rapporteurs) and the global intercept. The global parameter of interest is then informed both by the global trends as well as the individual intercepts. That can usually lead to higher accuracy in case of structured or time series data at the cost of higher computational expenses.

We ran both the models via Stan with 4 Monte Carlo Markov Chains (MCMC) of 20000 iterations each, the first 10000 warm up iterations being discarded. We did a diagnosis of all the models. In all cases, the trace plots show that the chains were stable and probed plausible parameter values, the density plots of the MCMC show that all 4 chains exhibited similar behavior, and

the autocorrelation between the iterations always dropped quickly and that the chains were moving around the potential parameter values quickly.

The posterior diagnosis reveals that although the simulations are not perfect, they do reasonably capture the features of the observed dissents count data. In other words, we selected the correct model and the priors are not too off either. Thus, our Poisson regression assumptions are reasonable.

Posterior predictive check of the absolute terms hierarchical model



We now compare the pooled against the hierarchical models. The former model got the posterior prediction 1.2 of the number of dissents per judge wrong (0.84 standard deviations off), whereas the former model got the posterior prediction wrong only by 0.94 of the number of dissents per judge (0.7 standard deviations off), with 99 % of the predictions falling within the 95 % posterior credible interval. 6-fold cross-validated check reveals that neither of the models overfitted. Thus, while the hierarchical model is slightly more computationally expensive, it yields better results.

Interpreting the posterior

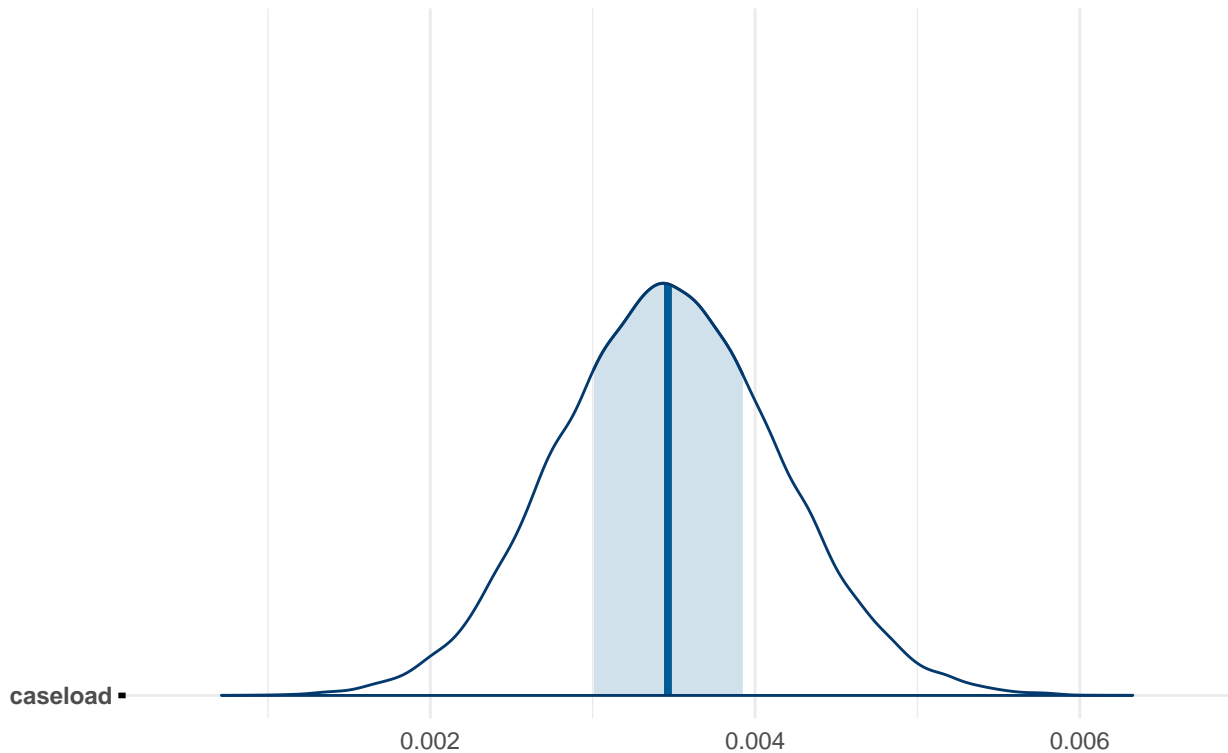
Interpreting the posterior is quite interesting. In neither case was the parameter of caseload in any way negative, as was found by Lee Epstein.

term	estimate	std.error	conf.low	conf.high
(Intercept)	-3.53	1.42	-5.65	-2
caseload	0.00	0.00	0.00	0

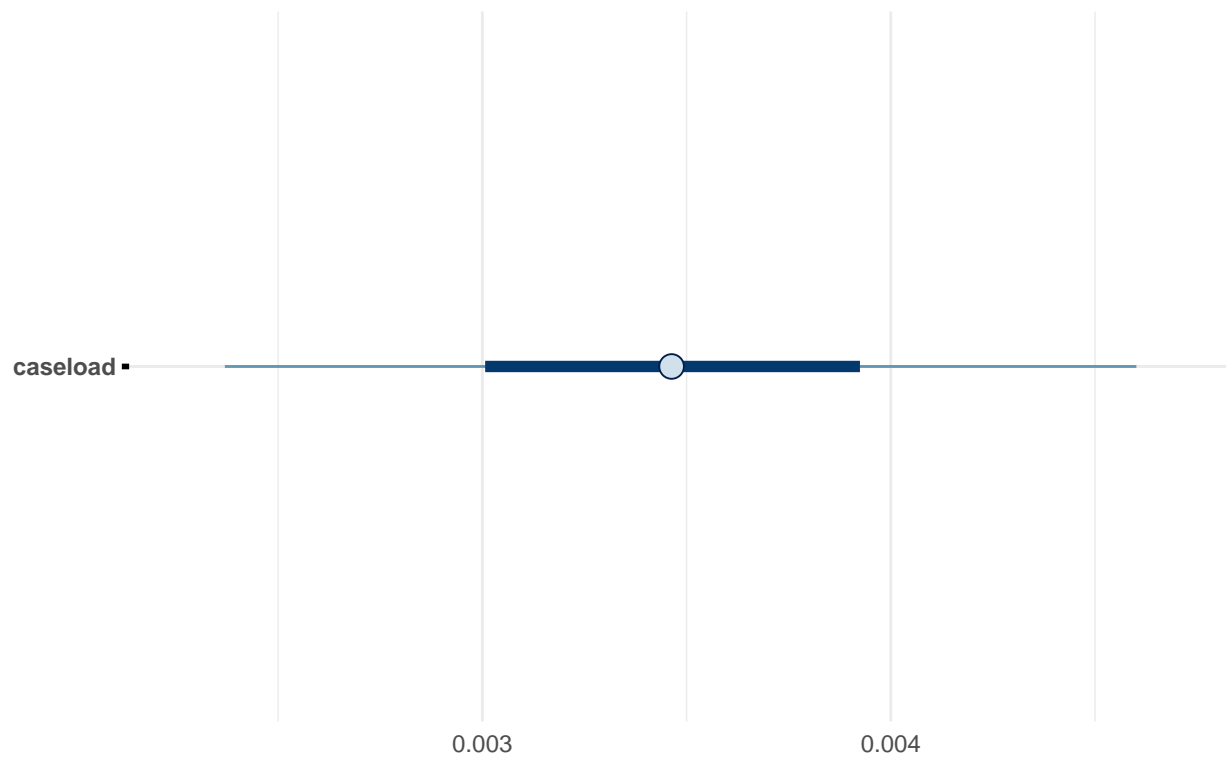
Even with the 80 % posterior credible interval, the caseload parameter is very slightly positive and all 4 chains converge on it. In other words, there is not a negative relationship between dissent rate and caseload on the CCC as on SCOTUS. Although the relationship statistically significantly differs from 0, it is hard to interpret the substantive significance - with each new case a judge gets assigned, their amount of dissents increases by 0.35 % ($e^{0.0035}$).

Density plot of estimates of parameters of workload of a judge

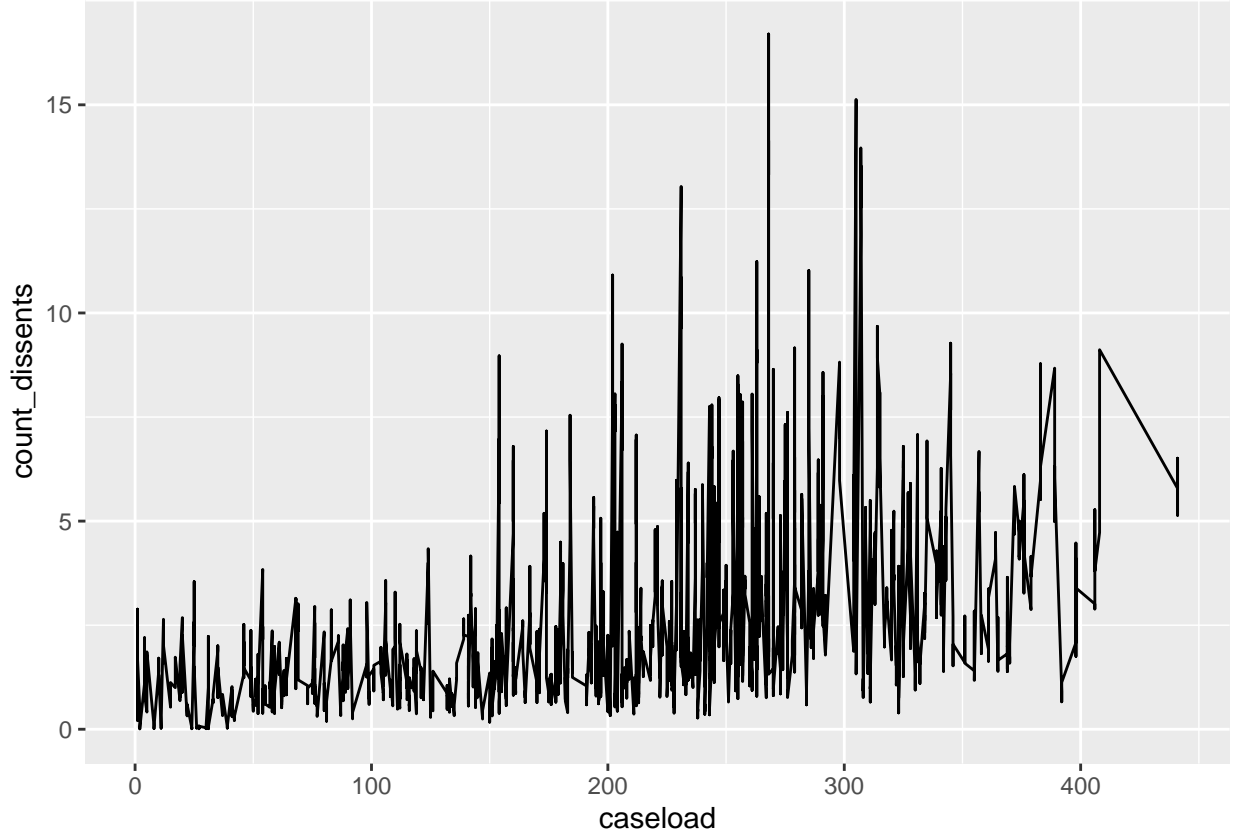
The inner area is for 50 % posterior credible interval, the outer for 95 %



Plot of uncertainty intervals of estimates of parameters of workload of a j
The inner whisker is for 50 % posterior credible interval, the outer for 95 %



What is more, when we draw 4 plausible posterior models, the models seem to be rather all over the place. It is then hardly possible to draw any definite conclusions regarding the slope of the explanatory variable.



The result from our analysis is in line with our intuition. The CCC judges may not act strategically or economically as legal scholars seem to posit based on the US studies. The parameters are in either - the rate as well as the absolute number - models positive in contrast to the result of Epstein et al.

Collegiality costs of dissenting at the CCC

Theory

Epstein et al. address the issue of collegiality costs arising for a dissenting judge: “The effort involved in these revisions, and the resentment at criticism by the dissenting judge, may impose a collegiality cost on the dissenting judge by making it more difficult for him to persuade judges to join his majority opinions in future cases.” Based on this theory, they predict and indeed empirically confirm that “dissents will be less frequent in circuits that have fewer judges because any two of its judges will sit together more frequently and thus have a greater incentive to invest in collegiality.”

While it is hard for us to see how a variation between the number of members in the plenary

session and 3-member panels could be isolated from a plethora of potential confounding variables, we are able to make use of the limited term of CCC judges. We test, based on the Epstein et al. theory, whether judges that are at the start of their term, and thus are aware that they will “sit together more frequently” invest in collegiality by averting dissents and whether when their term draws to an end, they give way to their disagreement. This presumes that the outlook of sharing the 10 year term with your colleagues at the beginning of judges’ terms increases the collegiality costs of dissenting, whereas at the end of their terms, the collegiality costs decrease with the end of the shared term looming on the horizon.

In the following section, we test the hypothesis:

RQ3: The judges are more likely to dissent at the end of their terms as the collegiality costs decrease.

Model

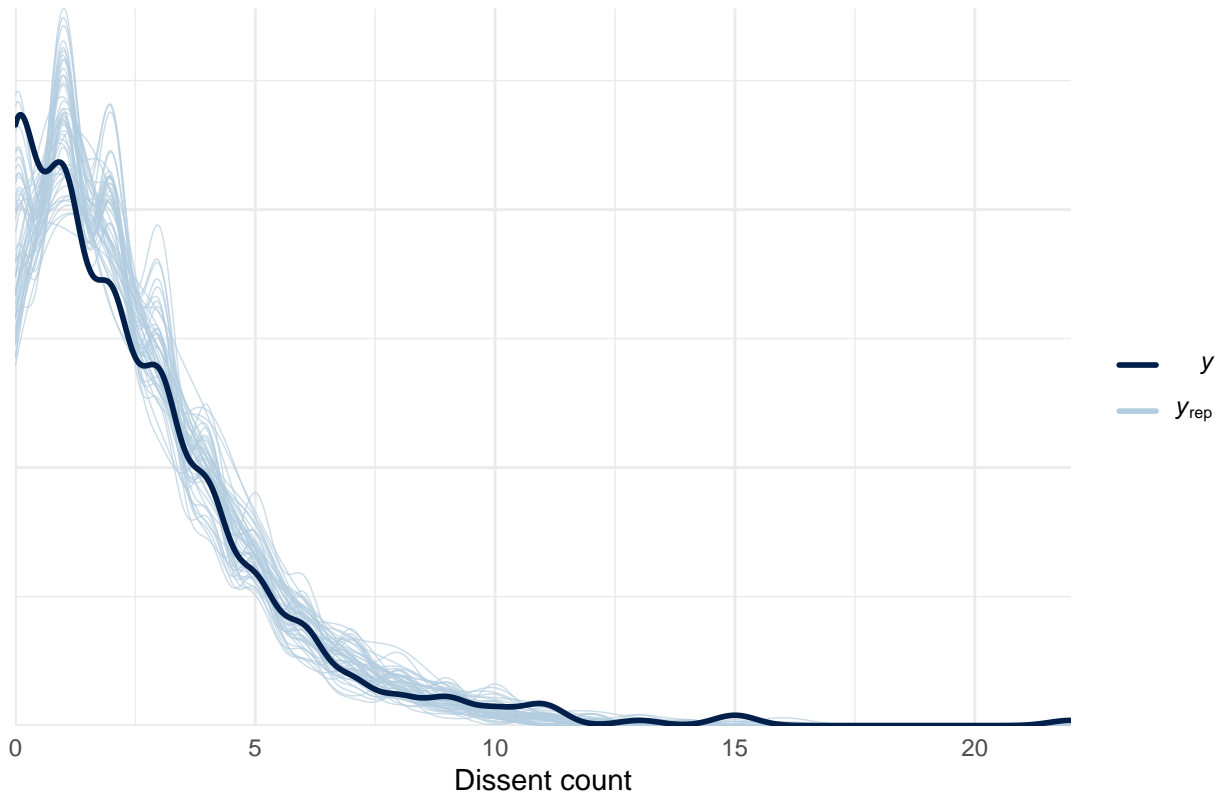
We build on our previous model. We now know that the number of dissents of a judge rapporteur follows a Poisson distribution. We use a hierarchical model pooled on the judges. We have no knowledge whatsoever about the effect of start or end of term on the number of dissents, thus, we use only weakly informative priors. We have addressed the potential sources of bias with the workload as an explanatory variable above.

Result

Model diagnosis

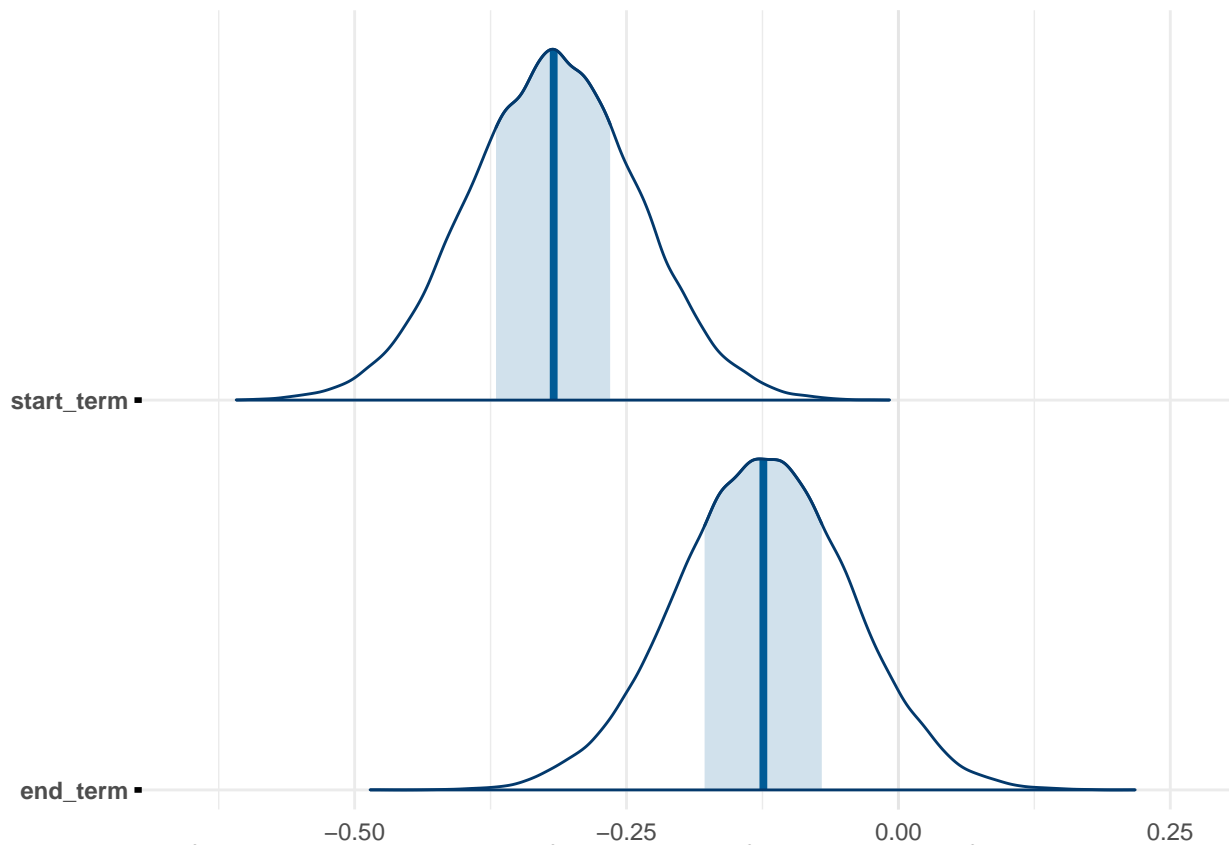
We ran the model via Stan with 4 Monte Carlo Markov Chains (MCMC) of 20000 iterations each, the first 10000 warm up iterations being discarded. We did diagnosis of all the models. The trace plots reveal that the chains were stable and probed plausible parameter values, the density plots of the MCMC show that all 4 chains exhibited similar behavior, and the autocorrelation between the iterations always dropped quickly and that the chains were moving around the potential parameter values quickly. The posterior predictive check again reveals that our posterior model reasonably captures the underlying data.

Posterior predictive check of the start and end of terms model

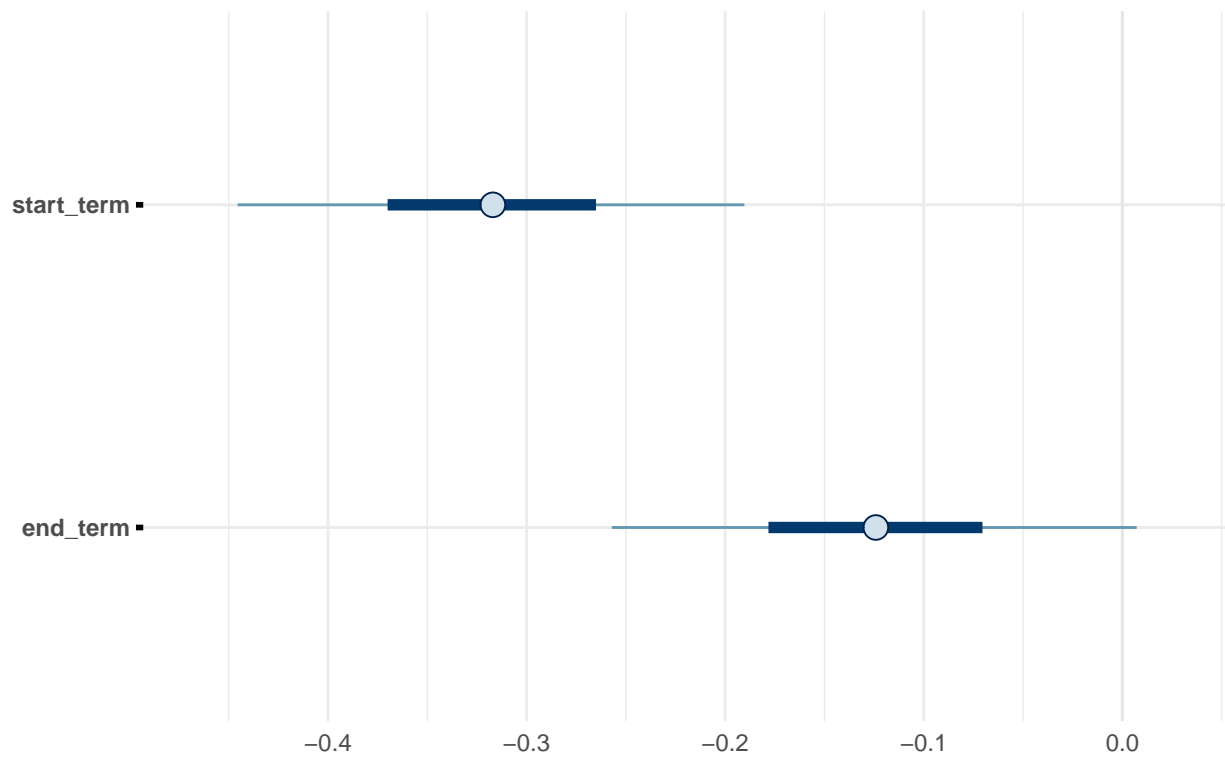


Interpreting the posterior

The parameters of both variables of interest turned out to be significantly negative.



Plot of uncertainty intervals of estimates of parameters of start and end
The inner whisker is for 50 % posterior credible interval, the outer for 95 %



On the one hand, consistently with the Epstein theory of collegiality costs of dissents, the

likelihood of dissenting is lower by 27.2 % ($e^{-0.317}$) during the first two years of judges' terms. On the other hand, contrary to our theoretical expectations, the likelihood of dissent drops by 11.7 % ($e^{-0.124}$).

term	estimate	std.error	conf.low	conf.high
(Intercept)	0.7890757	0.1310522	0.6187469	0.9563755
end_term	-0.1241316	0.0798892	-0.2275552	-0.0223108
start_term	-0.3169851	0.0777664	-0.4172039	-0.2186298

It is again difficult to square the theory of Epstein et al. with the results of our analysis. It is either that the theory of collegiality costs of dissents is not entirely accurate - perhaps it does not bring about any collegiality costs. Or it may not apply to the CCC because the CCC judges do not give weight to the collegiality costs. In either way, our results were not wholly consistent with the theory of collegiality costs of dissenting.

Discussion

We successfully transplanted a research design from the US context to the European context. We had to adjust it to the extent that data availability precluded us for posing certain research questions or applying certain methods. Our results are inconclusive as to the potential to transfer conclusions from empirical legal research conducted in the US context elsewhere.

On the one hand, we reached a similar conclusion regarding the costs a dissenting opinion imputes on the majority. In the CCC context the majority takes dissents seriously and appears to address them in their opinions. This conclusion is further supported by the fact that the deeper the disagreement seems to run, the more seriously it is taken by the majority.

On the other hand, the CCC judges do not seem to act as strategically as their US counterparts, at least in the context of dissenting opinions. Increase in their workload (operationalized both as the absolute number of cases assigned to a judge as well as the rate of change of thereof) does not appear to decrease the likelihood of that judge dissenting, as Epstein et al. concluded.

Our last remark addresses our research design in general: the regression analysis. We are aware that given the potential outcome frameworks, it is difficult to sustain all the assumptions of

regression research design. Experimental or quasi-experimental research design such as difference-in-differences or discontinuity designs should be the golden standard of social science (Bueno de Mesquita and Fowler 2021). The relative unmalleability of law in general, but rather conservative institutions such as courts in particular, leaves little space for experimental design and the general applicability of law within a legal system leaves little space for quasi-experimental design. That is not to say that it is impossible. Although we tried our best to think of and to address all potential sources of bias and whether the assumptions of the models of choice were met, we are aware of limitations of regression-based research design and therefore our conclusions should be taken with a grain of salt.

- Angrist, Joshua D., and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press. <https://books.google.com?id=YSAzEAAAQB AJ>.
- . 2014. *Mastering 'Metrics: The Path from Cause to Effect*. Princeton University Press.
- Berdejó, Carlos, and Daniel L. Chen. 2017. “Electoral Cycles Among US Courts of Appeals Judges.” *The Journal of Law and Economics* 60 (3): 479–96. <https://doi.org/10.1086/696237>.
- 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. <https://www.jstor.org/stable/25652213>.
- Brekke, Stein Arne, Joshua C. Fjellstul, Silje Synnøve Lyder Hermansen, and Daniel Naurin. 2023. “The CJEU Database Platform: Decisions and Decision-Makers.” *Journal of Law and Courts*, January, 1–22. <https://doi.org/10.1017/jlc.2022.3>.
- Bueno de Mesquita, Ethan, and Anthony Fowler, eds. 2021. *Thinking Clearly with Data: A Guide to Quantitative Reasoning and Analysis*. 1st. edition. Princeton: Princeton University Press.
- 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. <https://doi.org/10.1111/j.1540-5907.2011.00557.x>.
- Chmel, Jan. 2021. *Co Ovlivňuje Ústavní Soud a Jeho Soudce? /*. Vydání první. Teoretik (Leges). Leges,.
- Clark, Tom S., and Benjamin Lauderdale. 2010. “Locating Supreme Court Opinions in Doctrine Space.” *American Journal of Political Science* 54 (4): 871–90. <https://doi.org/10.1111/j.1540->

5907.2010.00470.x.

- Dworkin, Ronald M. 1980. *Political Judges and the Rule of Law*. London: British Academy.
- Eliášek, Martin, Jakub Kól, and Miloš Švaňa. 2020. “Automatická Klasifikace Významových Celků v Judikatuře.” *Revue Pro Právo a Technologie* 11 (21): 3–20. <https://doi.org/10.5817/RPT2020-1-1>.
- Epstein, Lee, and Jack Knight. 2000. “Toward a Strategic Revolution in Judicial Politics: A Look Back, A Look Ahead.” *Political Research Quarterly* 53 (3): 625–61. <https://doi.org/10.1177/106591290005300309>.
- Epstein, Lee, William M. Landes, and Richard A. Posner. 2011. “Why (And When) Judges Dissent: A Theoretical And Empirical Analysis.” *Journal of Legal Analysis* 3 (1): 101–37. <https://doi.org/10.1093/jla/3.1.101>.
- Fjelstul, Joshua. 2021. “How the Chamber System at the CJEU Undermines the Consistency of the Court’s Application of EU Law.” *Journal of Law and Courts*, November, 717422. <https://doi.org/10.1086/717422>.
- Gandhi, Rohith. 2018. “Support Vector Machine — Introduction to Machine Learning Algorithms.” Towards Data Science. July 5, 2018. <https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca47>.
- Harašta, Jakub, Jaromír Šavelka, František Kasl, and Jakub Míšek. 2019. “Automatic Segmentation of Czech Court Decisions into Multi-Paragraph Parts.” *Jusletter IT* 4 (M). <https://is.muni.cz/publication/1534440/cs/Automatic-Segmentation-of-Czech-Court-Decisions-into-Multi-Paragraph-Parts/Harasta-Savelka-Kasl-Misek>.
- Hyndman, Rob. 2010. “Transforming Data with Zeros.” Rob Hyndman. August 13, 2010. <https://robjhyndman.com/hyndsight/transformations/>.
- Johnson, Alicia A., Miles Q. Ott, and Mine Dogucu. 2022. *Bayes Rules!: An Introduction to Applied Bayesian Modeling*. CRC Press. <https://books.google.com?id=pISQzgEACAAJ>.
- Kastellec, Jonathan P. 2016. “Empirically Evaluating the Countermajoritarian Difficulty: Public Opinion, State Policy, and Judicial Review Before *Roe v. Wade*.” *Journal of Law and Courts* 4 (1): 1–42. <https://doi.org/10.1086/683466>.
- Lüders, Kilian, and Bent Stohlman. 2023. “Proportionality as an Argument: Identification of a Judicial Decision Technique.” *DRAFT for 5th ANNUAL COMPTTEXT Conference*.

- Maklin, Cory. 2019. “Gradient Boosting Decision Tree Algorithm Explained.” Towards Data Science. July 21, 2019. <https://towardsdatascience.com/machine-learning-part-18-boosting-algorithms-gradient-boosting-in-python-ef5ae6965be4>.
- Montgomery, Jacob M., Brendan Nyhan, and Michelle Torres. 2018. “How Conditioning on Post-treatment Variables Can Ruin Your Experiment and What to Do about It.” *American Journal of Political Science* 62 (3): 760–75. <https://www.jstor.org/stable/26598780>.
- Moyer, Laura P., and Holley Tankersley. 2012. “Judicial Innovation and Sexual Harassment Doctrine in the U.S. Courts of Appeals.” *Political Research Quarterly* 65 (4): 784–98. <https://doi.org/10.1177/1065912911411097>.
- Vartazaryan, Gor. 2022. “Sít’ová Analýza Disentujících Ústavních Soudců.” *Právník*, no. 12.