

Appendix of: Why and when do (Czech) 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

Contents

| | |
|--|----------|
| 1 Model 1: Effect of presence of dissenting opinion on the length of majority argumentation | 1 |
| 1.1 Caselaw partitioning classification | 1 |
| 1.2 Statistical model | 1 |
| 2 Model 2: Effect of workload on the dissenting behavior | 4 |
| 2.1 Model diagnosis | 4 |

Contents

This is an appendix, in which we diagnose and compare the models from our main article and explain our choices in higher detail.

1 Model 1: Effect of presence of dissenting opinion on the length of majority argumentation

1.1 *Caselaw partitioning classification*

1.2 *Statistical model*

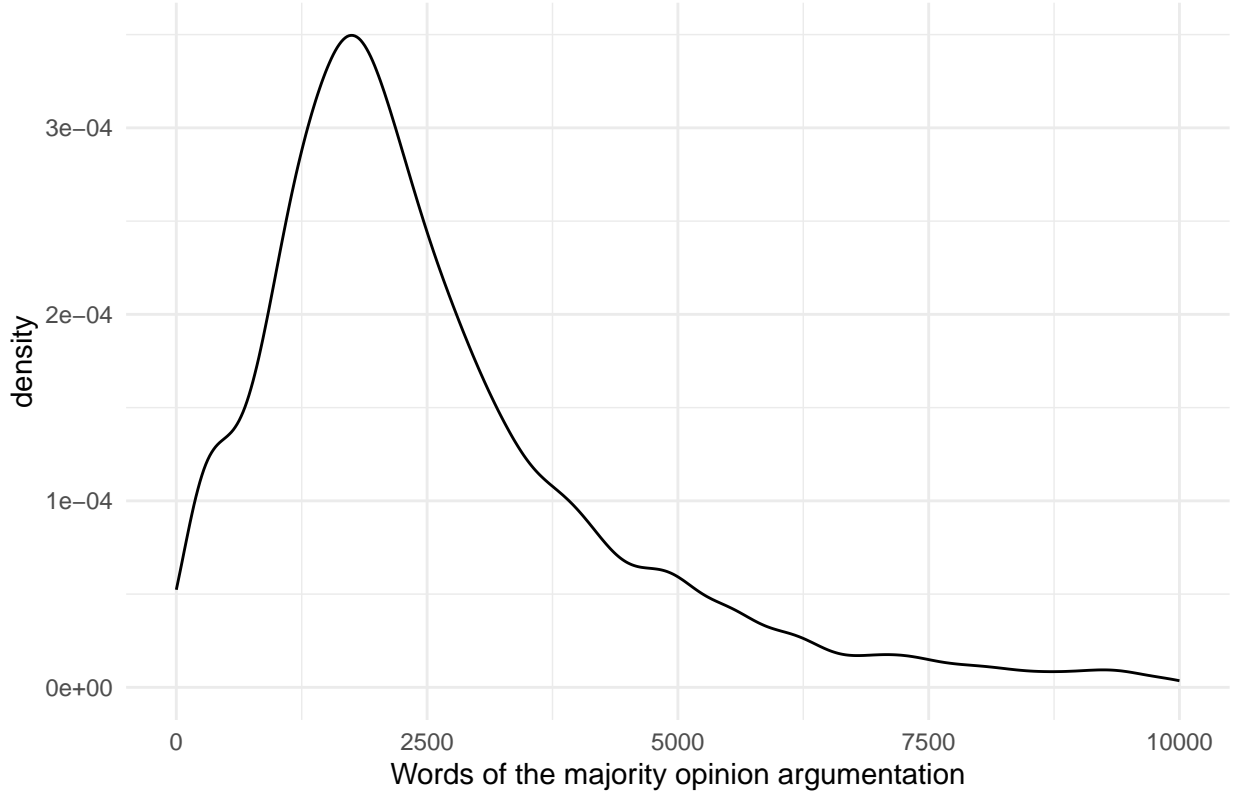
1.2.1 *Model specification*

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.

Distribution of number of words of argumentation of the majority



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. All our priors follow a normal distribution, the intercept being centered around the population mean. The remaining priors were kept uninformative, because we simply have no previous knowledge on the CCC.

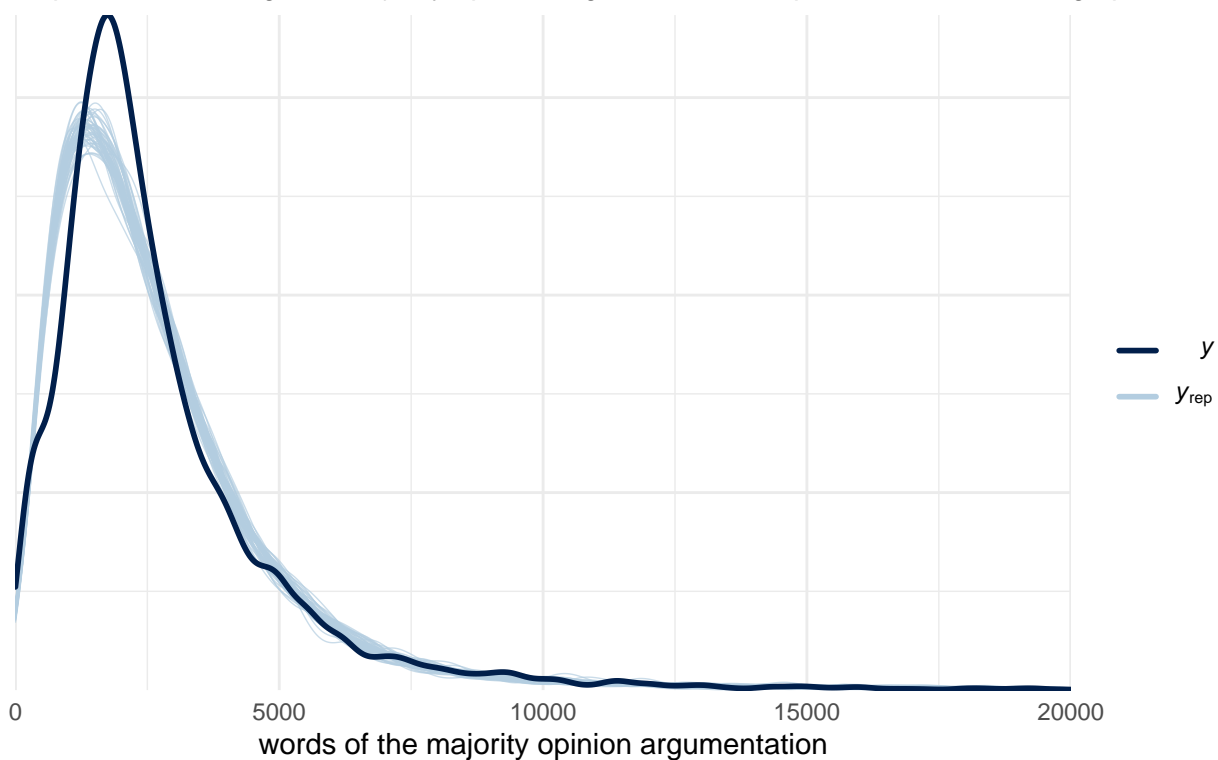
1.2.2 Model diagnosis

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. 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



2 Model 2: Effect of workload on the dissenting behavior

2.1 Model diagnosis

We tried two models, a completely pooled and hierarchical model clustered around the judges. 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 the individual judges) 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.

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.

Model 3: Collegiality costs of dissenting at the CCC ## 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

