

Intergroup Bias in Parliamentary Rule Enforcement

Frederik Hjorth*

*Corresponding author. Department of Political Science, University of Copenhagen, Øster Farimagsgade 5, DK-1353 Copenhagen K, (+45) 26 27 24 41, fh@ifs.ku.dk

Abstract

Political actors are often required to enforce rules without giving in-groups special treatment. But are such institutional roles likely to be successful? Here, we exploit a special case of exogenously assigned intergroup relations: debates in the Danish parliament, in which parliament chairmen drawn from parliamentary parties enforce speaking time. Analyzing 5,756 speeches scraped from online transcripts, we provide evidence that speech lengths are biased in favor of the presiding chairman's party. On average, speakers of the same party as the presiding chairman give 5 percent longer speeches and are 5 percent more likely to exceed the speaking time limit. The paper contributes to the extant literature by demonstrating political intergroup bias in a natural setting, suggesting that group loyalties can supersede institutional obligations even in a 'least likely' context of clear rules, complete observability, and a tradition of parliamentary cooperation.

1 Introduction

In a variety of contexts, institutions accord individuals formal roles which require them to enforce rules governing the competing interest of societal groups. In some cases, by either design or necessity, these individuals are drawn from one of the groups subject to these rules. Hence, in order to enforce the rules in an unbiased manner, assuming the role requires the individual to disregard his or her group loyalties.

But are such institutional roles and norms likely to be successful? In this study, we exploit a natural experiment in rule enforcement where the enforcer belongs to a group subject to enforcement. The empirical setting is parliamentary debates in Denmark, in which speaking times, governed by a simple, uniform rule, are enforced by chairmen drawn from major political parties.¹ Since combinations of speakers and chairmen are effectively random, the setting allows for identifying whether chairmen disregard partisan loyalties in enforcing speaking time rules.

The evidence suggests they do not. Across a variety of specifications, chairpersons accord significantly more speaking time to speakers of their own party (*copartisans*) than speakers from other parties (*non-copartisans*). On average, copartisans are accorded around 3 seconds more speaking time per speech, of which there can be up to 748 during a single debate. The effect is small, but non-negligible, corresponding to around 5 percent more speaking time allocated to co-partisan speakers per speech. The difference is concentrated around the formal limit of 60 seconds, such that copartisan speakers are on average 5 percent more likely to exceed the limit. The results demonstrate that partisanship can subtly bias the enforcement of rules, even in a political environment characterized by clear rules of universality, complete observability, low societal-level corruption, and a strong tradition of parliamentary cooperation.

The study adds to a broader social science literature on the influence of group identities on social and political behavior. In particular, we provide rare evidence of elected officials exhibiting intergroup bias in a highly controlled, natural setting. We argue that the political context makes for an empirical setting in many ways ‘least likely’ with respect to the observed

results. We discuss the question of the operative psychological mechanism towards the end of the paper. First, we describe the main strands of scholarship with which we engage.

2 Existing literature

Substantively, the study ties into two analytically distinct strands of research. First of all, research from what one might call the *social identity* perspective promotes the idea that, when perceiving the world in terms of ingroups and outgroups, humans are prone to discriminating against outgroup members and treating ingroup members favorably (Hewstone et al., 2002). In lab experiments conducted under the “minimal group paradigm”, Tajfel and Turner (1979, 1986) show that this results holds even when the groups in question are minimally cohesive, put together just for the purpose of the experiment. By finding group-based discrimination in a setting where groups were clearly entirely arbitrary and cultivated with a near-minimum of socialization, the studies demonstrated the pervasive influence of social identity on behavior, a widely replicated tendency (Brewer, 1979; Abrams and Hogg, 1990; Brown, 2000; Chen and Li, 2009).

Within the social identity perspective, a rich literature closely related to this study looks at racial bias in rule enforcement. Some studies utilize data on referee judgments in sports. For example, Price and Wolfers (2010) show that NBA referees award more fouls against opposite-race players. Other studies provide evidence of racial bias in policing (Donohue and Levitt, 2001; Antonovics and Knight, 2009), jury sentencing (Anwar et al., 2012), or political constituency service (Butler and Broockman, 2011). Though much more empirically oriented than the social psychological literature described above, the racial bias literature is conceptually within the social identity perspective insofar as it tends to assume that the racial bias operates implicitly.

Second of all, the study ties into a strand of research one might call *partisan governance*.

This literature demonstrates various ways in which political actors privilege fellow partisans in the execution of policy. For example, some studies have shown that the distribution of state-level funds sometimes reflects partisan political concerns rather than social welfare or economic efficiency criteria (Dahlberg and Johansson, 2002; Stratmann and Baur, 2002; Larcinese et al., 2008), a phenomenon colloquially known as ‘pork-barrel politics’. Similarly, in systems with politically appointed law enforcement, some studies suggest that prosecution patterns reflect partisan concerns (Gordon, 2009).

The distinction between these two strands is consequential, since it has bearing on how to theoretically interpret the observed bias. From the partisan governance perspective, intergroup bias in parliamentary debates is simply strategically deployed preferential treatment: chairmen have a political interest in tipping the scales in favor of their fellow partisans, so they do. In contrast, from the social identity perspective, partisan bias is the implicit operation of a powerful social identity, and while chairmen may give preferential treatment to fellow partisans, it is against an honest wish to behave neutrally.

As suggested, the empirical setting of this study does not allow us to conclusively distinguish between these two accounts, though we dig into the question towards the end of the paper. However, in other respects, we provide novel evidence. Most importantly, the discretionary enforcement of speaking time rules during debates in the Danish parliament provides us with a setting structurally similar to the original experiments of the minimal group paradigm, only with elected officials acting in a natural setting. By using unobtrusively obtained data, we eschew concerns about experimenter effects or lack of external validity.

Furthermore, by identifying partisan intergroup bias in a context of little ideological or affective polarization, we strengthen the case that neither are necessary conditions for rule enforcement to be subject to intergroup bias. As we will argue, these contextual features all work against an expectation of intergroup bias, making the setting a case for ‘least likely’ case inference (Levy, 2008): if intergroup bias is detectable here, it is likely to be present in most

other similar settings.

We also make two methodological contributions. First of all, we add to a small, but growing literature utilizing automatically recorded data to study social processes (King, 2011). Specifically, we gather data on speaking times in parliamentary debates by scraping out time stamps hidden in online transcripts of parliamentary proceedings. The increasing embeddedness of social life in digital processes provides social science with novel opportunities to test theories using highly precise, unobtrusively obtainable data.

Second of all, we add to a relatively sparse research literature using floor debates to study political behavior. The consensual view in the comparative politics literature is that floor debates in and of themselves serve no purpose of interest (Gallagher et al., 2005). Challenging this consensus, Martin and Vanberg (2008) use variation in the length of floor debates to show that coalition governments devote more time to debating bills on issues on which government is divided, particularly as election day approaches. The present study extends the idea of using data from floor debates, demonstrating that floor debates can reveal patterns of not just political behavior but social behavior broadly construed. Because they pit actors with relatively well-defined identities and incentives against each other under a clear and commonly known set of rules, floor debates are an attractive empirical setting for studying social behavior under institutional constraint.

The paper is structured as follows. In section 3, we present the empirical setting. Section 4 presents the data obtained, and the statistical model we use to identify intergroup bias. Section 5 presents the results of the analysis. Finally, section 6 draws up the main conclusion and discusses perspectives for future research. In this last section we also discuss two important questions this study is not able to settle: the issue of intentionality and the role of self-censorship.

3 Empirical setting

3.1 Parliamentary debates in the *Folketing*

The empirical setting of this study is the Danish parliament, the *Folketing*. The Folketing's legislative calendar is bracketed by two major debates: the *opening debate*, which takes place each year in early October, and the *closing debate*, which takes place the following year around the end of May. Major debates each begin with a speech by the Prime Minister, followed by speeches by party spokespersons in descending order of party size. However, the majority of the typical debate, which lasts between 12 and 16 hours, is spent on brief back-and-forth arguments between MP's and shifting party spokespersons. Hence, for the entire debate, the right to speak alternates quickly and continuously between parties.

Debates are governed by restrictive rules about speech time. The rules are as follows: each party spokesperson is allowed a speech up to 10 minutes long. After each speech, MP's make a number of so-called 'brief remarks' (*korte bemærkninger*) of up to 60 seconds, to which the spokesperson who gave the speech typically responds with his or her own brief remark. Debate continues around half an hour, after which the next party spokesperson gives his or her speech. Debates typically feature 600-800 brief remarks.

Because of the constant alternation of the right to speak, MP's do not schedule brief remarks in advance, but sign up to give remarks 'on the fly' using an electronic queue system.² Furthermore, the content of the debate at any given time is not known in advance, but is largely dictated by the content of the spokespersons' speeches. The opportunity for an individual MP to strategically select when to give remarks based on either chairman partisanship or debate content is thus minimal.

The identification strategy of this study rests on the fact that debate rules are enforced in as-if-random order by 'chairs', i.e. members of the Folketing presidency, which consists of senior members from each of the five biggest parties. During a debate, each chairman oversees

a considerable amount of debate involving MP's from his or her own party (in the following called *copartisans*) as well as MP's from other parties (*non-copartisans*).³

The 'head chairman', i.e. the head of Parliament, starts out overseeing each debate, but the order of chairmen is otherwise arbitrary, determined based on scheduling convenience. The number of remarks overseen by each chairman is thus not a function of his or her party's size in parliament, but only on membership of the presidency and availability on the day of the debate. The specific schedule of presiding chairmen is set in advance of each debate, meaning that chairmen cannot self-select into shifts based on events during the debate.

3.2 Denmark as 'least likely' setting

A number of characteristics of opening and closing debates in the Folketing make them a plausibly 'least likely' setting for intergroup bias in speaking time rule enforcement. Some of these characteristics relate to the specific setting of the debates. As described, the rules governing speech time are simple, leaving chairmen little room to excuse uneven rule enforcement. Second, enforcement is monitored, as debates are not only viewed by all members of parliament, but also attending journalists and viewers of the televised broadcast of the debate.

Other characteristics relate to the broader political context. By most measures, Denmark ranks among the world's least corrupt countries, with a public sector characterized by Weberian norms of impartial governance (Treisman, 2000). These legal-rational norms should work against chairmen letting partisan affiliation influence rule enforcement. Lastly, due to its cooperative parliamentary tradition, Denmark is typically classified as a 'consensus' type democracy in Lijphart's classic typology (Lijphart, 1999; Esaiasson and Heidar, 2000). Though it lacks the oversized governing coalitions common in ideal typical consensus democracies, parliamentary politics in Denmark has a strong tradition of cross-party cooperation Green-Pedersen and Thomsen (2005). This consensual tradition similarly works against an expectation of partisan bias.

These formal and contextual characteristics should not be taken to imply that rule enforcement is likely to be perfectly neutral. As argued, the extant literature provides ample reasons to expect some measure of intergroup bias. But the combination of these characteristics arguably makes for a ‘least likely’ setting in terms of the magnitude of biased parliamentary rule enforcement. The inferential implication is that evidence of intergroup bias found here is informative of the presence of intergroup bias in other contexts (Levy, 2008). Our estimates thus provide a plausible lower bound on what is likely to be obtained elsewhere.

3.3 Hypothesis

The hypothesis tested here is whether chairpersons, enforcing common rules about speech time, bias their enforcement to the benefit of copartisans. Specifically, the hypothesis tested is as follows:

Hypothesis: Members’ speeches under copartisan chairmanship are on average longer than members’ speeches under non-copartisan chairmanship.

The assumption implicit in the hypothesis is that any observable difference between speech lengths under copartisanship vis-à-vis non-copartisanship can be causally attributed to copartisanship and not some omitted factor. This assumption is credible as long as the partisanship of the chair is *as-if randomly* assigned to individual speakers (Dunning, 2007). As mentioned, the order of the chairmen’s shifts is itself allocated as a matter of convenience in terms of chairmen’s other private and professional obligations. The order of chairmen’s shifts is thus in itself plausibly random with respect to the debate schedule. This assumption is supported by the seemingly haphazard order of shifts observable in the data.

By focusing on the lengths of speeches given by copartisan and non-copartisan members, we bracket the issue of whether chairmen exert bias by calling on copartisan members more. Though chairmen are allowed to change the order of speakers defined by the electronic queue

system, they formally do so only to give preference to spokespersons, who typically dominate debates. More importantly, even chairmen do use this discretionary power to call on copartisans more often, it will not confound the bias estimate in this study, since we estimate bias in speaking time conditional on being called upon to speak.

An important challenge to identification remains, namely the possibility that speakers self-censor. Since speakers know the partisan affiliation of the chair, it is possible that they deliberately shorten their speeches under non-copartisan enforcement, anticipating that the non-copartisan chairman is biased against them.

The extent of self-censorship is difficult to assess even based on qualitative scrutiny of archived recordings of floor debates. Recordings of debates indicate that chairpersons rarely interrupt MP's to stop speeches, which in turn suggests extensive self-censorship. However, chairpersons will often stand up from their chair as a sign to the speaker that their time is up, a signal invisible in televised recordings. Neither are actual speech lengths conclusive. Though speech lengths shorter than 60 seconds would seem to suggest that speakers self-censor, chairpersons are allowed to interrupt speakers before their time is up in the interest of allowing for more remarks. Hence, though each of these factors suggest at least some self-censorship, its precise extent is difficult to assess from the data at hand. Section 6 discusses the implications of this for whether the observed pattern can be characterized as intergroup bias.

4 Data and measurement

4.1 Obtaining speech data

The data used in this study is obtained by extracting detailed information about individual speeches from online transcripts of opening and closing parliamentary debates since the opening debate in 2010. The unique feature of this data is that, beginning in 2010, the lengths of speeches are recorded in time stamps down to the number of seconds, yielding highly precise measures

of how much speaking time each debate participant is given. Though invisible in the rendered transcripts, the time stamps are available as meta tags embedded in the underlying html code. We accessed transcripts at the parliament's official website (<http://www.ft.dk>) and scraped the html code for time stamps and other information using the package `rvest` for the statistical software R (R Core Team, 2013). The code for scraping out time stamps is provided in the replication materials.⁴

4.2 Summary statistics

The full data set contains information about speaker and length for 12,048 speeches. Around half of these are merely transcriptions of chairpersons calling the next speaker in line, yielding 6,012 actual speeches by members of parliament. 14 of these are official speeches by the PM, and 144 are official speeches by party spokespersons. Brief remarks, the focus of this paper, constitute the remaining 5,756 speeches equivalent to around 96 percent of all speeches in the data. Table 1 summarizes the available data.

[Table 1 about here.]

4.3 Modeling strategy

The key variables in the data are a measure of the length of each of the 5,756 speeches in seconds and a dummy indicator of whether the member delivering the speech is a copartisan of the presiding chairman. Hence, the hypothesis presented above can be stated in its simplest regression model form as follows:

$$SpeechLength_i = f(\alpha + \beta \times Copartisan_i + \mathbf{X}_i \times \gamma) \quad (1)$$

Where $SpeechLength_i$ is an either continuous or dichotomous measure of the length of speech i , $Copartisan_i$ is a dummy indicator of whether the member delivering speech i is

a copartisan of the presiding chairperson, and \mathbf{X}_i is a vector of additional control variables including various fixed effects terms with coefficients γ . Depending on the measurement of $SpeechLength_i$, $f(\cdot)$ is either linear or an inverse logit. The hypothesis tested is whether the coefficient for $Copartisan_i$, i.e. β , is positive.

5 Results

We estimate equation 1 with both linear regression for the continuous speech length measure and logistic regression for the dichotomous measure. Before proceeding to the regression results, Figure 1 plots the raw distribution of speech lengths for copartisans and non-copartisans. For non-copartisans, distributions are presented separately based on bloc affiliation relative to the chairman. We revisit the role of bloc affiliation in section 5.3 below.

[Figure 1 about here.]

As seen in Figure 1, the frequency of non-copartisan speeches drops markedly as speech lengths approach the 60-second limit. By contrast, copartisan speeches are almost perfectly symmetrically distributed around the limit. The figure lends *prima facie* support to the notion that speech lengths depend on the partisan relation between chair and speaker.

A noticeable detail in Figure 1 is that for non-copartisans not in the chairman’s political bloc (i.e., left panel), the distribution of speaking times is slightly bimodal. The reason for this slight bimodality is not clear. It is indicative of some heterogeneity in the observed effect, such that the remarks of some non-copartisans are censored to well below the cutoff, whereas others are unaffected. We do not delve further into what underpins this heterogeneity in this paper.

Table 2 presents estimations of various specifications of equation 1. Model 1 presents a bivariate regression of speech length on copartisanship; models 2-5 add various controls. Across all model specifications, the estimated effect of copartisanship is robustly significant and in the predicted direction.

[Table 2 about here.]

The results thus support the hypothesis that copartisanship is associated with longer speeches. The estimated effect of copartisanship ranges from 3.3 additional seconds (without any controls) to 2 seconds (with speaker party fixed effects), with a median of 3.1 seconds. In absolute terms, the estimated effects thus seem fairly small, but the effect is equivalent to an appreciable ‘copartisanship premium’ of around 5 percent of the expected speech length.

The models presented in table 2 also control for two other possible confounders. For one, one may suspect that the observed copartisanship effect reflects the fact that members of major parties, which hold the chairman positions, speak earlier in the day, and have more energy to use the full extent of their speaking time. Models 2-5, which include a speech order control, suggest that there is in fact a slight ‘fatigue effect’: speeches later in the day are very slightly, though insignificantly, shorter. The latest speeches in the data are on average around two seconds shorter than the earliest. But controlling for order does not affect the significance of the estimated effect of copartisanship.

Second, one could suspect that the effect is driven by minor parties, which do not hold chairman positions, being awarded less speaking time as a matter of sheer proportionality. Such a difference could produce a spurious copartisanship effect. Model 5 accounts for this by estimating the model for only parties which hold chairman positions. As shown, the effect is undiminished.

5.1 Robustness check: binary speaking time measure

If the difference in speaking times observed for copartisans vs. non-copartisans reflects biased enforcement of speaking time rules, we should expect the difference to be concentrated around the 60-second threshold. In contrast, if the observed difference instead reflected fewer copartisan speeches lasting five seconds and more lasting 10 seconds, it would not make sense to attribute

the difference to biased rule enforcement. Chairmen can only meaningfully affect speaking times in the neighborhood of the 60-second threshold.

The distributions plotted in Figure 1 strongly suggest that the difference is indeed concentrated around the 60-second threshold. As a way of formally testing the idea, we dichotomize the dependent variable to indicate whether the speaker has exceeded the 60-second limit or not. Table 3 presents the results from a series of logistic regression models, specifications identical with those in Table 2, estimated with this new dependent variable.

[Table 3 about here.]

As the table shows, the results using the dichotomized dependent variable are substantively similar. The copartisanship coefficient remains significant across all specifications.

5.2 Making sense of the magnitude of bias

As a way of providing more intuitive quantities of interest, figure 2 plots average marginal effects for all the copartisanship estimates in tables 2 and 3. As shown in the logit model estimates in the left hand panel, a speaker of the same party as the chairman has an on average five percent higher chance of exceeding the speaking time limit, an estimate that is fairly stable across specification. In the linear model estimates in the right hand panel, the median effect estimate reflects an around 3 seconds longer speaking time for copartisans.

[Figure 2 about here.]

There are several ways to make sense of the magnitude of this effect. One is in terms of standardized measures of effect size such as Cohen's d , which assesses effects relative to the standard deviation of the dependent variable. Since the standard deviation of speaking time is 28.6, Cohen's d for the estimated effect is around 0.1, typically classified as a 'small' effect (Cohen, 1992). This corresponds well to the intuitive sense that 3 seconds out of a speech of

around 60 seconds is indeed a relatively small fraction. At the same time, this characterization somewhat understates the magnitude of the effect, since it ignores the repeated nature of the speeches. Given that each chairman oversees 128 speeches per debate on average, the cumulative speaking time bias per chairman per debate comes to around six minutes. In other words, though the speaking time bias is small at the level of the individual speech, it adds up when viewed from the perspective of each chairman. At the same time, the continual rotation of chairmen means that in the aggregate, these biases are likely to roughly cancel out. Overall, the magnitude of bias would seem to be best characterized as small, but non-negligible.

5.3 Political moderators: Distance or party group affiliation?

Having established the presence of biased rule enforcement, we return to the question outlined in the introduction: what is the most plausible psychological mechanism? As a way of getting at this question, we test if chairmen appear responsive to variation in the political affiliation of speakers which should be relevant to a politically strategic actor. If the partisan governance account is correct, chairmen would plausibly be most strongly biased against politically distant speakers. Conversely, if the social identity account is correct, chairmen would presumably treat non-copartisans roughly equally, seeing as they are all out-group members, irrespective of political distance. The extent to which the observed effect is moderated by political factors is thus at least suggestive of the mechanism underpinning the observed bias.

The standard way to test if political distance moderates the observed relationship would be to set up an interaction between copartisanship and political distance between chairman and speaker. However, this is statistically infeasible, since copartisanship and political distance are highly conditionally dependent: for all copartisans, political distance is by definition zero, and for all positive values of political distance, copartisanship is by definition zero. Hence, a regression model with copartisanship interacted with political distance is inestimable.

We therefore take a different approach, testing the moderating effect of political distance

by reestimating the full model (model 4 in table 2) for various subsets of the data where the political distance between chairman and speaker is artificially restricted. First, we calculate an approximate distance measure based on averaged voter estimates of party positions from the most recent national election survey (Stubager et al., 2013).⁵ For each speaker, we calculate the distance measure as the absolute value of the distance between the parties of speaker and chairman. We then reestimate the full model using only data from chairman-speaker pairs where the distance is less than or equal to the median of all observed distances. Second, to test the robustness of the distance measure, we apply a similar restriction using a rank-transformed (ordinal) version of the distance measure. Finally, we also estimate a model restricting the data to cases where the chairman-speaker pair belong to the same political bloc. Table 4 shows the results of each of these restrictions alongside the full model. We plot the estimates in figure 3.

[Table 4 about here.]

[Figure 3 about here.]

As shown, all of the estimated effects of copartisanship on speaking time are clearly smaller than in the full model, and none of them are statistically significant. The difference should not be overstated: all of the estimates still have the predicted sign, and there is considerable overlap between the confidence intervals. Still, the results do strongly suggest that the effect of copartisanship is weaker when copartisans are compared to politically proximate non-copartisans. As argued above, this pattern is more consistent with the partisan governance account, according to which chairmen should be attuned to how politically distal the speaker is.

Empirically speaking, the conclusion from these moderation tests is reasonably clear: though chairmen exhibit bias against non-copartisans on average, this bias is much less pronounced when comparing copartisans only with non-copartisans who are politically proximate and/or belong to the same political bloc. Strictly speaking, the bias thus appears to be not one of partisanship per se, but of political proximity and/or bloc affiliation.

However, this empirical result is not necessarily highly informative of the merits of the competing theoretical perspectives. The fact that chairmen are sensitive to the political distance and bloc affiliation of the speaker would seem to speak in favor of the partisan governance perspective, since making this distinction furthers a broader political agenda. On the other hand, a social identity perspective might argue that, in parliamentary politics in multiparty systems, political blocs and groups of ideologically similar parties are themselves meaningful social identities. Thus, though the results do lend some support to the partisan governance account, it seems safest to conclude that the tests cannot fully adjudicate between the two theorized mechanisms.

5.4 Additional tests

In the appendix, we present some additional results and robustness checks which serve to shore up potential concerns with the empirical analysis and shed light on heterogeneities unexplored in the main analysis. The appendix is divided into three sections. Section A.1 presents robustness checks of the results from Table 2. We show that the results are robust to restricting the data only to members of leadership parties (Table 5), excluding prime ministers (Table 6), and adding debate-specific fixed effects (Table 7). Section A.2 shows results from a model predicting how many remarks chairmen oversee (Table 8). The main determinant of the number of remarks enforced by a chairmen is the chairman's tenure, i.e. for how many debates he or she was in the leadership. In contrast, party seat share is uncorrelated with activity. Lastly, Section A.3 presents analyses of effect heterogeneity, showing the effect of copartisanship estimated separately for each party (Figure 4) and chairman (Figure 5) in the data. The analysis suggest limited overall effect heterogeneity and no discernible political pattern.

6 Conclusion

The key finding from this study is that speaking time rules in Danish parliamentary debates are not enforced equally. Specifically, speeches are around 3 seconds longer on average when the person speaking shares party affiliation with the chairman enforcing the rules. This corresponds to a roughly five percent greater chance for copartisans of exceeding the 60-second speaking time limit. Since the official rules call for neutral enforcement of the limit, this difference can be attributed to intergroup bias on the part of the rule enforcer, i.e. the chairman. The intergroup bias explanation is consistent with a wealth of laboratory experimental evidence demonstrating the impact of group identity on distributional preferences.

We contribute to this literature by showing intergroup bias among elected political officials in a natural setting. The generalizability of the finding is further bolstered by the fact that the empirical setting is characterized by clear rules of universality, complete observability, low societal-level corruption, and a strong tradition of parliamentary cooperation – thus in several respects a ‘least likely’ case for intergroup bias in rule enforcement.

The magnitude of the observed effect is small, but non-negligible, corresponding to around 5 percent of the average speaking time or around 5 percent additional chance of exceeding the formal threshold of 60 seconds. Equally importantly, the rotation of chairmen throughout the typical debate means that accumulated advantages in speaking time for any given party will tend to cancel out in the long run. This implies that the total impact of the observed bias in parliamentary life is minimal.

However, from a theoretical perspective it is important to note that the principle of rotating chairman responsibilities is not an accident, but a particular institutional design. Though the principle may be motivated by mere convenience, it has real distributional effects. Even in the presence of (limited) biased rule enforcement, the rotation principle ensures that biases roughly cancel out over the course of a debate. The study thus implies that while institutional features such as simple rules and observability do not completely eradicate biased rule enforcement, a

principle of rotation among rule enforcers can ensure that the impact of enforcement remains approximately balanced in the aggregate.

Two concerns about the nature of the observed effect linger. First of all, the data is uninformative as to whether the observed difference reflects actual biased enforcement on behalf of chairmen as opposed to speakers voluntarily cutting their speeches short in anticipation of biased enforcement. In the latter scenario, copartisans could theoretically gain an advantage solely by anticipating more lenient rule enforcement, even if rules are in fact enforced equally. The effect could even arise from a hybrid scenario in which non-verbal signals from the chairman such as body language cause non-copartisan MP's to censor their remarks. The possibility of self-censorship remains an important caveat.

Second, and theoretically perhaps most crucially, the design cannot identify the motivational basis of the observed effect. In the introduction, we presented two competing perspectives on biased rule enforcement. The first, which we have called the social identity perspective, construes bias as operating largely implicitly, based on the enforcers identification with a socially salient ingroup. The second, which we have called the partisan governance perspective, construes bias as motivated by a strategic wish to favor members of one's own coalition.

The result that the observed effect of copartisanship is sharply diminished when restricting the sample to politically proximate speakers suggests that political distance is an important moderator. All else equal, this speaks in favor of the partisan governance perspective. At the same time, political parties as well as political blocs simultaneously structure political interests and group socialization. For this reason, the observational design of this study cannot fully disentangle these two mechanisms.

To be sure, this theoretical distinction is not necessarily a perfectly exclusive one: biased behavior may reflect a murky combination of strategic and implicit motives. Still, the distinction is too important to gloss over. A promising avenue for future research is thus to find ways to distinguish more clearly which of the two mechanisms dominate. In a broader sense, future

research would do well to examine other ways in which psychological phenomena, replicated in laboratory studies many times over, manifest themselves in political behavior observable in natural settings.

Notes

¹In the interest of fluency, the paper refers to both chairmen and -women using male pronouns, reflecting the modal speaker gender.

²The source of this, and other details below about the workings of debates, is personal correspondence with *Folketinget*'s legal office.

³The distinction borrows from Habyarimana et al. (2007), who distinguish between *coethnics* and *non-coethnics* in experimental studies of group identity and rule enforcement.

⁴Available upon publication at <http://github.com/fghjorth/parlbias>.

⁵Respondents in the 2011 Danish National Election Study were asked to place parties on a 0 to 10 scale. From left to right, the party position estimates are: Unity List (1.49), Socialist People's Party (2.79), Social Democrats (3.84), Social Liberals (5.01), Liberal Alliance (7.14), Conservative People's Party (7.23), Liberals (7.35), Danish People's Party (7.74)

References

- Abrams, D., Hogg, M.A., 1990. *Social identity theory: Constructive and critical advances*. Springer-Verlag, New York.
- Antonovics, K., Knight, B.G., 2009. A New Look at Racial Profiling: Evidence from the Boston Police Department. *Review of Economics and Statistics* 91, 163–177.
- Anwar, S., Bayer, P., Hjalmarsson, R., 2012. The Impact of Jury Race in Criminal Trials. *The Quarterly Journal of Economics* 127, 1017–1055.
- Brewer, M.B., 1979. In-group bias in the minimal intergroup situation: A cognitive-motivational analysis. *Psychological Bulletin* 86, 307–324.
- Brown, R., 2000. Social identity theory: past achievements, current problems and future challenges. *European Journal of Social Psychology* 30, 745–778.
- Butler, D.M., Broockman, D.E., 2011. Do Politicians Racially Discriminate Against Constituents? A Field Experiment on State Legislators. *American Journal of Political Science* 55, 463–477.
- Chen, Y., Li, S.X., 2009. Group Identity and Social Preferences. *The American Economic Review* 99, 27.
- Cohen, J., 1992. A power primer. *Psychological bulletin* 112, 155–9.
- Dahlberg, M., Johansson, E., 2002. On the Vote-Purchasing Behavior of Incumbent Governments. *American Political Science Review* 96, 27–40.
- Donohue, J.J., Levitt, S.D., 2001. The Impact of Race on Policing and Arrests*. *The Journal of Law and Economics* 44, 367–394.

- Dunning, T., 2007. Improving Causal Inference: Strengths and Limitations of Natural Experiments. *Political Research Quarterly* 61, 1065912907306470–293.
- Esaiasson, P., Heidar, K., 2000. *Beyond Westminster and Congress: The Nordic Experience*. Ohio State University Press, Columbus, OH.
- Gallagher, M., Mair, P., Laver, M., 2005. *Representative Government in Modern Europe*. 4 ed., McGraw-Hill Humanities/Social Sciences/Languages.
- Gordon, S.C., 2009. Assessing Partisan Bias in Federal Public Corruption Prosecutions. *American Political Science Review* 103, 534.
- Green-Pedersen, C., Thomsen, L.H., 2005. Bloc Politics vs. Broad Cooperation? The functioning of Danish minority parliamentarism. *The Journal of Legislative Studies* 11, 153–169.
- Habyarimana, J., Humphreys, M., Posner, D.N., Weinstein, J.M., 2007. Why Does Ethnic Diversity Undermine Public Goods Provision? *American Political Science Review* 101.
- Hewstone, M., Rubin, M., Willis, H., 2002. Intergroup bias. *Annual review of psychology* 53, 575–604.
- King, G., 2011. Ensuring the data-rich future of the social sciences. *Science* (New York, N.Y.) 331, 719–21.
- Larcinese, V., Rizzo, L., Testa, C., 2008. Allocating the U.S. Federal Budget to the States: The Impact of the President. *The Journal of Politics* 68.
- Levy, J., 2008. Case Studies: Types, Designs, and Logics of Inference. *Conflict Management and Peace Science* 25, 1–18.
- Lijphart, A., 1999. *Patterns of Democracy*. Yale University Press , 143–170.

- Martin, L.W., Vanberg, G., 2008. Coalition Government and Political Communication. *Political Research Quarterly* 61, 502–516.
- Price, J., Wolfers, J., 2010. Racial Discrimination Among NBA Referees. *The Quarterly Journal of Economics* 125, 1859–1887.
- R Core Team, R., 2013. R: A Language and Environment for Statistical Computing. URL: <http://www.r-project.org>, doi:10.1007/978-3-540-74686-7.
- Stratmann, T., Baur, M., 2002. Plurality Rule, Proportional Representation, and the German Bundestag: How Incentives to Pork-Barrel Differ Across Electoral Systems. *American Journal of Political Science* 46, 506–514.
- Stubager, R., Hansen, K.M., Andersen, J.G., 2013. Krisevalg. Økonomien og folketingsvalget 2011. Jurist- og Økonomforbundets Forlag, København.
- Tajfel, H., Turner, J.C., 1979. An integrative theory of intergroup conflict, in: Austin, W.G., Worchel, S. (Eds.), *Differentiation between social groups: Studies in the social psychology of intergroup relations*. Brooks/Cole, pp. 33–47.
- Tajfel, H., Turner, J.C., 1986. *The social identity theory of intergroup behavior*, Nelson-Hall Publishers, pp. 7–24.
- Treisman, D., 2000. The causes of corruption: a cross-national study. *Journal of Public Economics* 76, 399–457.

A Appendix

A.1 Robustness checks of main results

Here we present additional robustness checks of the main results. Table 5 presents results restricting the sample to remarks given by members of parties in the parliamentary leadership (i.e., the five biggest parties by number of seats).

[Table 5 about here.]

Table 6 presents the main results excluding brief remarks given by the sitting PM. Note that all analyses cover only brief remarks, so formal speeches given by the PM are excluded to begin with. Here we excluded brief remarks given by the PM in their capacity of also being an MP. These remarks are relatively rare, accounting for only 377 remarks or about 6 percent. A potential concern could be that the results are confounded by appearances by the PM, whom the chair might show extra deference to. As shown in Table 6, the results are robust to excluding PM remarks.

[Table 6 about here.]

Table 7 presents the main results adding fixed effects for individual debates. Note that since each debate covers several issues, debate fixed effects should not be interpreted as holding issue content constant.

[Table 7 about here.]

A.2 Predicting chairman activity

Here we present results from regression models predicting how many debates chairpersons control. The key result from these models is that the main predictor of number of debates controlled is how long a chairperson has been a member of the leadership. There is an additional small premium for being the head chairperson, which likely reflects the fact that by convention, the president always takes the first shift controlling a debate. In contrast, number of seats has no significant association.

[Table 8 about here.]

Table 1 shows that the main determinant of the number of remarks enforced by a chairmen is chairman's tenure, i.e. for how many debates he or she was in the leadership. In contrast, party seat share is uncorrelated with activity. The results complement the qualitative evidence from correspondence with the parliament leadership that chairperson activity is largely driven by 'supply-side' factors, i.e. chairpersons' availability on the day of the debate, and not by party specific factors such as their parliamentary power.

A.3 Effect heterogeneity across parties and chairmen

Here we assess the heterogeneity of the observed effect across parties and individual chairmen. In both cases, estimates are drawn from a random effects model corresponding to model 3 in Table 2 where we have allowed the slope of the copartisanship variable to vary by either parties or individual chairmen.

Figure 4 presents effect heterogeneity across chairmen's parties.

[Figure 4 about here.]

Figure 5 presents effect heterogeneity across individual chairmen.

[Figure 5 about here.]

Tables

Table 1: Types of speeches in opening and closing debates in the Folketing

Type	Number	Share (numeric)	Share (time-weighted)
Brief remark	5756	95.74	82.4
PM speech	14	0.23	2.3
Spokesperson speech	144	2.40	15.0
Total	6012	100.00	100.0

Table 2: OLS models of speaking time

	Speaking time (seconds)				
	(1)	(2)	(3)	(4)	(5)
Copartisan	3.29*** (0.93)	3.27*** (0.98)	3.14*** (1.02)	2.03** (1.03)	2.64*** (0.71)
Time of day		-0.03 (0.17)	-0.02 (0.17)	0.03 (0.17)	0.04 (0.18)
Gender (female)			1.12 (0.75)	0.48 (0.76)	0.64 (0.75)
Debate type (Opening)			-0.92 (1.13)	-0.63 (1.03)	0.08 (1.06)
Intercept	51.81*** (0.92)	52.28*** (2.96)	52.20*** (3.14)	32.63*** (3.06)	35.93*** (4.91)
N	5,756	5,756	5,756	5,756	5,756
Speaker party FE				✓	✓
Chair party FE					✓
R ²	0.004	0.004	0.01	0.04	0.05
Adjusted R ²	0.004	0.004	0.01	0.04	0.05
chi ²	25.14*** (df = 1)	25.33*** (df = 2)	33.27*** (df = 4)	224.15*** (df = 13)	314.18*** (df = 20)

*p < .1; **p < .05; ***p < .01

Table 3: Logit models of exceeding standard speaking time

	Dummy: speaking time exceeds 60 seconds				
	(1)	(2)	(3)	(4)	(5)
Copartisan	0.24*** (0.07)	0.24*** (0.07)	0.23*** (0.07)	0.15* (0.08)	0.20** (0.08)
Time of day		0.002 (0.01)	0.004 (0.01)	0.01 (0.01)	0.01* (0.01)
Gender (female)			0.16*** (0.06)	0.08 (0.06)	0.10 (0.06)
Debate type (Opening)	-0.60*** (0.03)	-0.65*** (0.12)	-0.69*** (0.13)	-3.17*** (0.74)	-2.89*** (0.74)
N	5,756	5,756	5,756	5,756	5,756
Speaker party FE				✓	✓
Chair party FE					✓
Log Likelihood	-3,767.39	-3,767.33	-3,761.94	-3,679.05	-3,647.81
AIC	7,538.78	7,540.66	7,533.87	7,386.10	7,337.63

*p < .1; **p < .05; ***p < .01

Table 4: Tests of political moderators

	Speaking time (seconds)			
	Full	Distance \leq median (int.)	Distance \leq median (ord.)	Same bloc
	(1)	(2)	(3)	(4)
Copartisan	2.64*** (0.71)	1.09 (1.25)	1.13 (1.41)	0.61 (1.01)
Time of day	0.04 (0.18)	0.05 (0.22)	0.03 (0.19)	0.01 (0.09)
Gender (female)	0.64 (0.75)	-0.44 (1.17)	-0.38 (1.00)	-0.27 (0.80)
Debate type (Opening)	35.93*** (4.91)	59.71*** (5.09)	60.40*** (4.49)	37.36*** (3.73)
N	5,756	2,853	2,987	2,950
Speaker party FE	✓	✓	✓	✓
Chair party FE	✓	✓	✓	✓
R ²	0.05	0.05	0.07	0.08
Adjusted R ²	0.05	0.04	0.06	0.07
chi ²	314.18*** (df = 20)	149.17*** (df = 18)	204.15*** (df = 18)	240.26*** (df = 20)

*p < .1; **p < .05; ***p < .01

Table 5: Results for only members of leadership parties

	Speaking time (seconds)				
	(1)	(2)	(3)	(4)	(5)
Copartisan	3.09*** (0.98)	3.07*** (1.01)	2.95*** (1.04)	2.03** (1.03)	2.64*** (0.70)
Time of day		−0.03 (0.17)	−0.02 (0.17)	0.03 (0.17)	0.05 (0.18)
Gender (female)			1.02 (0.71)	0.54 (0.76)	0.73 (0.75)
Debate type (Opening)			−0.73 (1.07)	−0.69 (1.03)	0.05 (1.07)
Intercept	52.02*** (0.93)	52.51*** (2.90)	52.35*** (3.13)	58.26*** (3.54)	57.10*** (4.92)
N	5,528	5,528	5,528	5,528	5,528
Speaker party FE			✓		✓
Chair party FE					✓
R ²	0.004	0.004	0.01	0.03	0.05
Adjusted R ²	0.004	0.004	0.004	0.03	0.04
chi ²	22.08*** (df = 1)	22.29*** (df = 2)	28.01*** (df = 4)	177.16*** (df = 11)	267.66*** (df = 18)

*p < .1; **p < .05; ***p < .01

Table 6: Results excluding prime ministers

	Speaking time (seconds)				
	(1)	(2)	(3)	(4)	(5)
Copartisan	3.42*** (1.00)	3.14*** (1.16)	3.13*** (1.12)	2.11** (1.00)	2.67*** (0.72)
Time of day		-0.14 (0.14)	-0.12 (0.15)	-0.09 (0.14)	-0.06 (0.14)
Gender (female)			1.39** (0.58)	0.66 (0.61)	0.72 (0.60)
Debate type (Opening)			-0.72 (1.08)	-0.38 (0.93)	0.34 (0.92)
Intercept	51.45*** (0.93)	53.74*** (2.70)	53.21*** (2.83)	34.28*** (2.37)	36.66*** (4.11)
N	5,379	5,379	5,379	5,379	5,379
Speaker party FE				✓	✓
Chair party FE					✓
R ²	0.004	0.005	0.01	0.04	0.06
Adjusted R ²	0.004	0.005	0.01	0.04	0.05
chi ²	22.50*** (df = 1)	26.62*** (df = 2)	35.22*** (df = 4)	222.21*** (df = 13)	308.80*** (df = 20)

*p < .1; **p < .05; ***p < .01

Table 7: Results with debate-specific fixed effects

	Speaking time (seconds)				
	(1)	(2)	(3)	(4)	(5)
Copartisan	2.74*** (0.88)	2.63*** (0.90)	2.57*** (0.93)	1.42** (0.55)	1.65*** (0.53)
Time of day		-0.14 (0.13)	-0.13 (0.13)	-0.08 (0.13)	-0.08 (0.14)
Gender (female)			0.53 (0.51)	0.04 (0.51)	0.11 (0.58)
Intercept	58.53*** (1.83)	60.92*** (3.16)	60.58*** (3.16)	46.38*** (3.41)	47.65*** (3.89)
N	5,756	5,756	5,756	5,756	5,756
Speaker party FE			✓		✓
Chair party FE					✓
Debate FE	✓	✓	✓	✓	✓
R ²	0.09	0.09	0.09	0.12	0.12
Adjusted R ²	0.09	0.09	0.09	0.12	0.12
chi ²	552.21*** (df = 10)	557.29*** (df = 11)	558.43*** (df = 12)	725.40*** (df = 21)	760.44*** (df = 28)

*p < .1; **p < .05; ***p < .01

Table 8: Model predicting number of remarks enforced by chairmen

	Number of remarks enforced		
	(1)	(2)	(3)
Head chairman	260.01** (121.61)	235.37* (126.67)	736.62*** (249.60)
No. of debates in leadership	169.95*** (19.52)	167.01*** (20.06)	
No. of party seats		2.31 (2.88)	6.68 (6.34)
Constant	-153.72** (54.74)	-204.11** (83.63)	14.28 (177.87)
N	20	20	20
R ²	0.89	0.90	0.46
Adjusted R ²	0.88	0.88	0.39
Residual Std. Error	162.87 (df = 17)	164.60 (df = 16)	368.67 (df = 17)
F Statistic	71.52*** (df = 2; 17)	46.90*** (df = 3; 16)	7.12*** (df = 2; 17)

*p < .1; **p < .05; ***p < .01

Figures

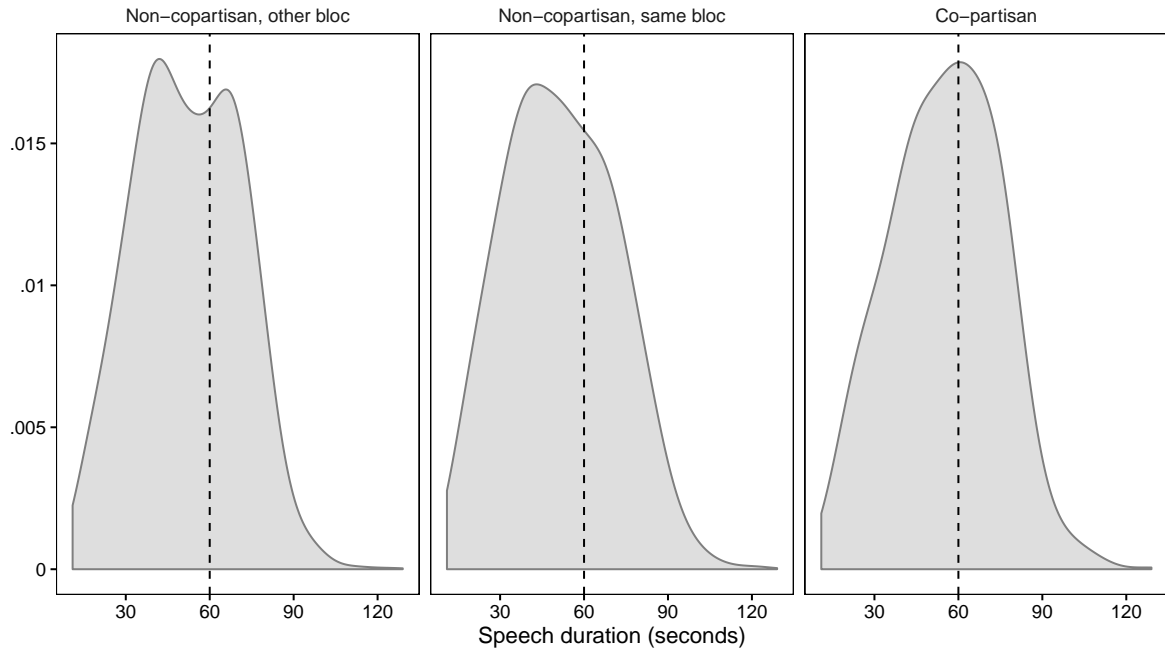


Figure 1: Distribution of speaking times for non-copartisans outside of the chairman's political bloc (left panel), non-copartisans from the chairman's own political bloc (middle panel) and copartisans (right panel). The dashed vertical line indicates the 60-second limit for short speeches. Non-copartisans drop off before reaching the limit. No such pattern is apparent for copartisans.

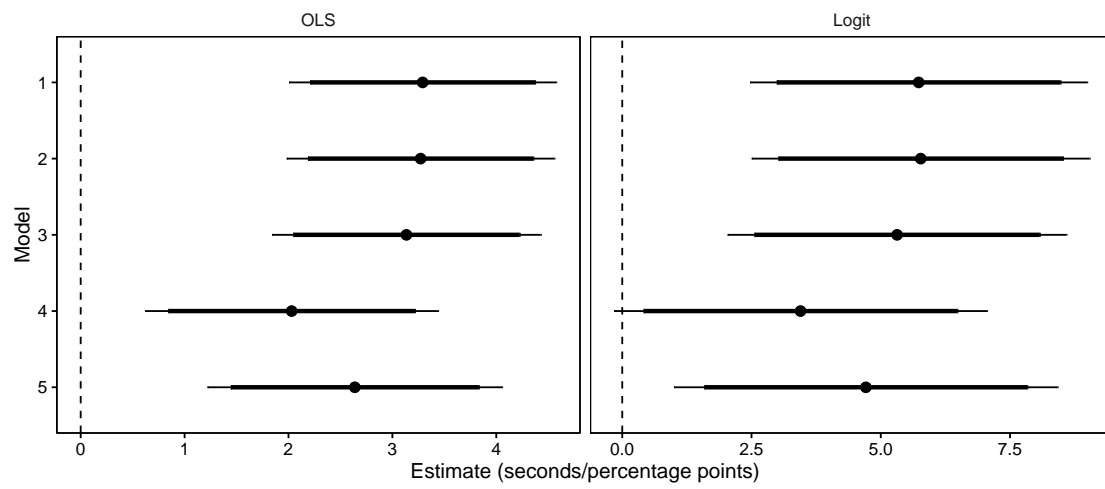


Figure 2: Estimated effect of copartisanship on speaking time (left panel) and probability of exceeding 60 seconds (right panel) for each of the five models estimated. Error bars represent 90 percent (thick lines) and 95 percent (thin lines) confidence intervals. Logit model estimates represent marginal effects expressed in percentage points.

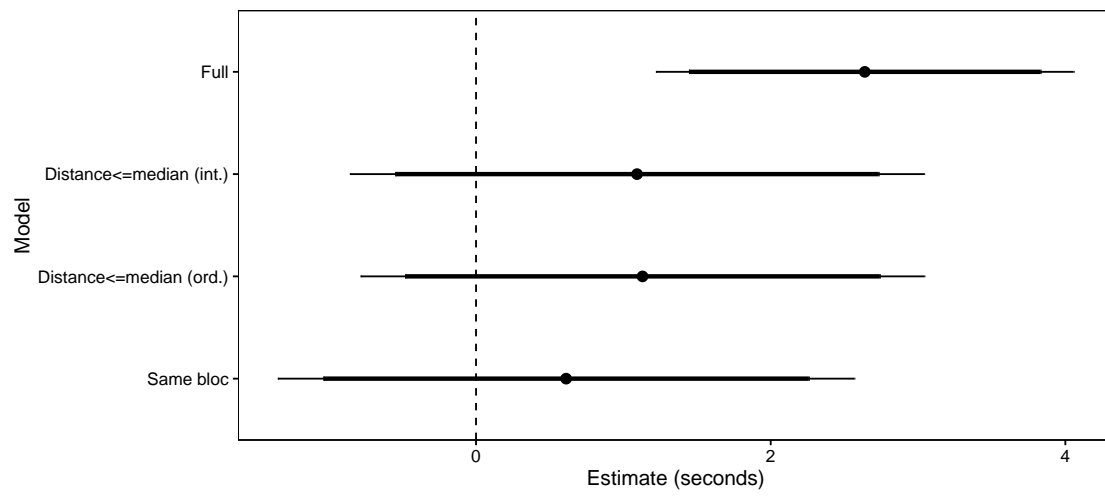


Figure 3: Estimated effect of copartisanship on speaking time under various restrictions on political distance between chairman and speaker. Error bars represent 90 percent (thick lines) and 95 percent (thin lines) confidence intervals.

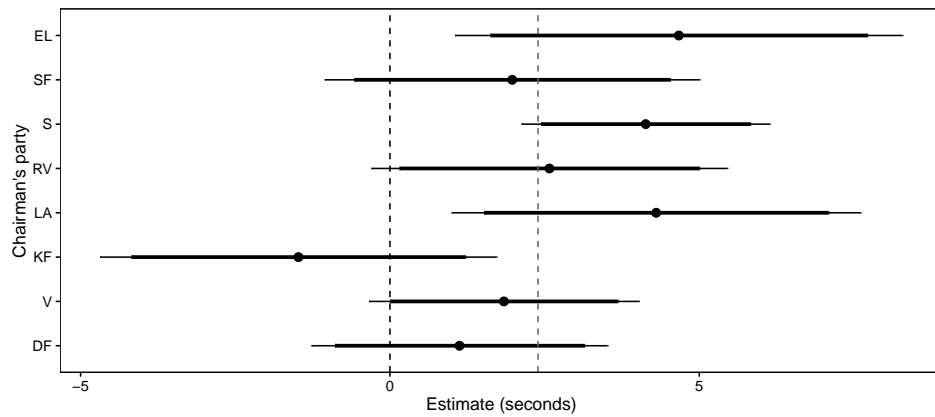


Figure 4: Effects of copartisanship estimated separately for each chairman's party. Estimates are drawn from a random effects model corresponding to model 3 in Table 2 with varying slopes for chairman party. The grey dashed vertical line shows the average copartisanship effect (i.e., the fixed effect). Parties are ordered top to bottom according to party ordinal left/right ranking. Error bars represent 90 percent (thick lines) and 95 percent (thin lines) confidence intervals.

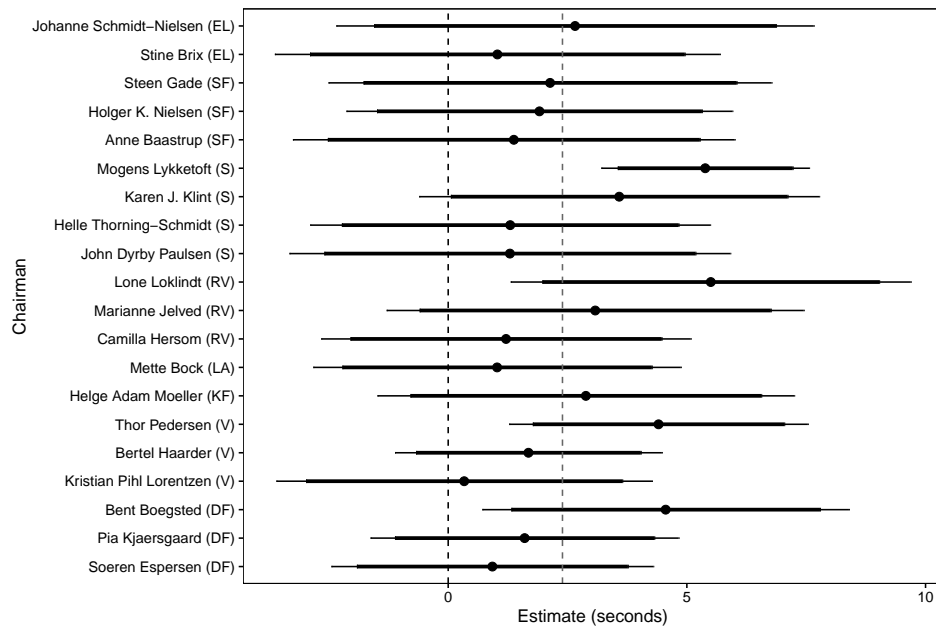


Figure 5: Effects of copartisanship estimated separately for each chairman in the data. Estimates are drawn from a random effects model corresponding to model 3 in Table 2 with varying slopes for individual chairmen. The grey dashed vertical line shows the average copartisanship effect (i.e., the fixed effect). Chairmen are ordered top to bottom according to party ordinal left/right ranking and within parties according to estimated effect. Error bars represent 90 percent (thick lines) and 95 percent (thin lines) confidence intervals.