Revision notes for "Intergroup Bias in Parliamentary Rule Enforcement"

Frederik Hjorth

March XXX, 2016

Revision memo

Dear Editors of Political Research Quarterly,

Thank you for giving me the opportunity to revise my manuscript "Intergroup Bias in Parliamentary Rule Enforcement" (Manuscript ID PRQ-2015-0341). I greatly appreciate the time and effort taken by yourself and the four reviewers to review the manuscript. The comments and suggestions by yourself and the reviewers have been extremely helpful and I have made my best effort to accommodate the suggestions in the revised version of the manuscript. I believe the changes have made for a stronger, clearer, and more persuasive paper.

Below I address the specific comments made by each of the reviewers in turn. These sections are labeled 1.1 for R1's first comment, 1.2 for R1's second comment, and so on.

In response to several of the reviewers' comments, I have added new figures and tables to the manuscript, some in the main text, some in the appendix. For ease of reading, I reproduce these tables and figures below, along with other material only shown in this note.

Reviewer 1

1.1: Data to justify 'as if' randomness

R1 requests more empirical data in support of the claim that assignment of chairpersons to debates is "as if" random. Specifically, R1 asks for a graph showing the percentage of seats controlled by the chairperson on the x-axis and the percentage of times chairing a debate on the y-axis, noting that the points in such a graph should fall along a 45 degree line.

I thank R1 for this suggestion. However, I do not believe the predicted pattern would constitute evidence in support of "as if" random assignment. The reason for this is that members of the leadership (with some exceptions for the head chairperson) are treated equally, with no

consideration given to their party's share of seats in parliament. Hence, a chairperson from the 5th largest party has the same obligations as the chairperson from the largest party (again, with some exceptions for the head chairperson). Seat share should thus not be a predictor of chairperson activity. I want to stress that the manuscript was not sufficiently clear on this point, and I have added a paragraph in section 3.1 making this point more explicitly.

As additional evidence for the "as if" random assignment of chairpersons to debates, I have added section A.1, "Predicting chairman activity", to the appendix. Here I 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. The results are also shown below.

Table 1: Model predicting number of remarks enforced by chairmen

	Number of remarks enforced		
	(1)	(2)	(3)
Head chairman	260.01^{**}	235.37^{*}	736.62^{***}
	(121.61)	(126.67)	(249.60)
No. of debates in leadership	169.95^{***}	167.01^{***}	
	(19.52)	(20.06)	
No. of party seats		2.31	6.68
		(2.88)	(6.34)
Constant	-153.72^{**}	-204.11^{**}	14.28
	(54.74)	(83.63)	(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)

$^{*}p < .1$; $^{**}p < .05$; $^{***}p < .01$

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.

1.2: Abrupt conclusion

R1 notes that the conclusion is abrupt and should include more discussion of the implications of the finding.

I have extended the conclusion section (section 6) with more discussion of the findings's implications. For the specific wording of the discussion, I refer to the manuscript. In short, I argue that the findings imply that intergroup bias can occur even in institutional and cultural settings that should work strongly against bias (the same reasons for which I argue the setting is a 'least likely' case). At the same time, the observed effect is indeed fairly small, suggesting that the aforementioned factors do go some way towards mitigating bias. Lastly, and perhaps most importantly, they illustrate the importance of balancing institutions. As noted in section 5.2, the effects on speaking time for each chairperson roughly wash out within each debate due to the rotation of chairpersons. This does not nullify the finding per se, but illustrates how the principle of rotating chairperson roles can result in (approximate) equity in the aggregate even in the presence of bias at the individual level (see also the discussion in section 3.3 below). I think this added discussion has strenghtened the manuscript, and I thank R1 for the suggestion.

1.3: Minor points

Lastly, R1 mentions some minor points. I have fixed the word in footnote 1 to 'modal', as was indeed intended. R1 also notes that section 5.2 seems unnecessary, which I agree with in principle. However, section 5.2, which discusses the magnitude of the effect, speaks to a criticism sometimes raised against the finding: namely, that the observed effect is fairly small. This is indeed the case, and in section 5.2, I make an effort to be explicit about this point so as not to appear to overstate the magnitude of the effect. Hence, I have opted to keep section 5.2, though I agree that it could be omitted if need be.

Reviewer 2

2.1: Data to justify 'as if' randomness

R2 also suggests additional evidence justifying the assumption of "as if" randomness. Specifically, R2 expresses concern that chairmen might self-select into particular issues of high importance to the party. If party members also take more time discussing these issues, this self-selection would produce a spurious finding of own-group bias. As a way of countering this concern, R2 suggests showing that the issues debated are similar in copartisan and non-copartisan matches.

This is indeed a relevant concern, and I thank R2 for pointing it out. Because of the structure of the data, which only contains information about the duration of each remark and not its content, I cannot compare the issues covered by different speakers. However, two aspects of how debates are organized indicate that issue-based self-selection is not likely. First of all, opening or closing debates are not about any one issue, but cover a multitude of issues, mostly based on issue disagreements between parties. MP's may debate one issue for a few remarks, and then shift to another issue. Second of all, the schedule of which chairmen oversee the debate is set in advance of the debate. Hence, chairmen are not able to opt in to the role when

a particular issue comes up in the debate. The combination of these two aspects suggests self-selection based on issues is unlikely. I have added a paragraph in section 3.1 of the manuscript making this point explicitly.

For additional evidence in favor of the "as if" randomness assumption, see section 1.1 in this note.

2.2: Debate-specific fixed effects

As additional evidence in favor of the "as if" randomness assumption, R2 suggests employing debate-specific fixed effects. As mentioned in the previous section, each debate covers several issues, so debate fixed effects should not be interpreted as holding issue content constant. Nevertheless, debate-specific fixed effects is a valuable robustness check of the results. Table 7 in section A.2 in the appendix shows the main results with debate-specific fixed effects added in every specification. The results are robust to including debate-specific fixed effects. The table is also shown below. In the added section 5.4 "Additional tests", I mention this robustness test.

Table 2: Results with debate-specific fixed effects

	Speaking time (seconds)				
	(1)	(2)	(3)	(4)	(5)
Copartisan	2.74***	2.63***	2.57***	1.42**	1.65***
	(0.88)	(0.90)	(0.93)	(0.55)	(0.53)
Time of day		−0.14	−0.13	−0.08	−0.08
		(0.13)	(0.13)	(0.13)	(0.14)
Gender (female)			0.53	0.04	0.11
			(0.51)	(0.51)	(0.58)
Intercept	58.53***	60.92***	60.58***	46.38***	47.65***
	(1.83)	(3.16)	(3.16)	(3.41)	(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

2.3: More information about speaking assignment rules

R2 asks for more information about the rules about how the chairman decides who gets to speak in what order. As R2 notes, chairman may be able to allocate more time to his/her own

party by just calling on more copartisans to speak, and the original version of the manuscript does not discuss this possibility.

I have added a discussion of this issue at the end of section 5.3. In short, chairmens are able to call on copartisans more than non-copartisans, but to a limited extent, since MP's sign up to speak using an electronic queue system. Chairmen can change the order of the queue, but do so only to a limited extent, and only (formally) with the purpose of giving party spokespersons preference. It is important to note that even if chairmen do give copartisans preference, this would not confound my estimate, since I estimate bias in speaking time *conditional* on speaking. I thank R2 for raising this issue, which I had not originally thought of.

2.4: Discuss the odd shape of Figure 2

Lastly, R2 points to the somewhat puzzling bimodal shape of the distribution of non-copartisan response times shown in Figure 2. The shape already receives some discussion in the text, but as R2 rightly points out, it is not quite the shape one would expect given just some average bias by chairmen.

To explore this issue further, I have changed the figure to present the distribution for non-copartisans separately for members of the chairman's own parliamentary bloc and for members of the other bloc. Figure 1 shows the revised figure.

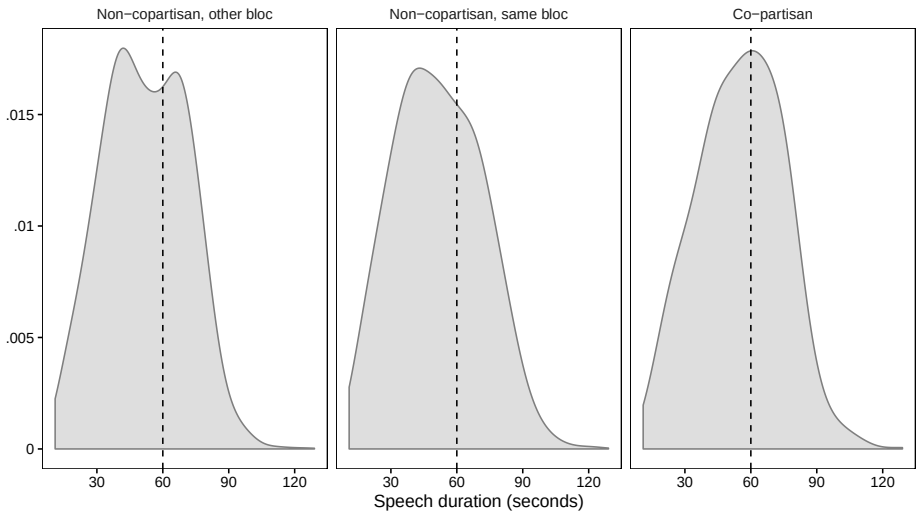

Figure 1: Distribution of speaking times for non-copartisans outside of the chairman's political bloc (left panel), non-copartisans from the chairman's own political bloc (middle panel) and copartisans (right panel).

Figure 1 makes clear that while the distribution of speaking times for same-bloc non-copartisans has the predicted shape, the shape for other-bloc non-copartisans is clearly

slightly bimodal. In terms of explaining this slight bimodality, I lean towards the explanation proposed by R4 (see section 4.1 in this note for more discussion of this). In short, some non-copartisans may anticipate differential rule enforcement and thus constrain their remark, though I acknowledge that this leaves unanswered what accounts for this heterogeneity. This does in turn suggest that some of the observed difference reflects self-censorship. I have added a discussion of this to the beginning of section 5 of the paper.

Reviewer 3

3.1: Analysis of distinction between rational choice and social identity processes

R3 argues that while there is some discussion of the distinction between rational choice and social identity processes, there is no analysis shedding light on this distinction. I do not agree with this assessment. Section 5.3 of the manuscript, testing the role of political moderators, has the purpose of exploring this distinction. As I argue, the fact that the bias is diminshed when looking only within the chairman's own bloc is suggestive of a rational mechanism, though the test is not conclusive. I appreciate the comment, and I have rewritten section 5.3 to make the purpose of the section clearer.

3.2: Partisanship as a 'soft' test

R3 notes that relative to a minimal group experiment, demonstrating an effect of partisanship is a relatively 'soft' test of social identity. As R3 writes, "parties for MPs are not minimal groups".

This point is well taken, and the test here is indeed softer than the test in Tajfel et al.'s original minimal group experiments. However, the contribution of this study is to test for intergroup bias in an ecologically valid setting. In contrast to experimentally assigned minimal groups, this study shows discrimination in naturally occurring groups. It is in the context of naturally occurring groups that I argue the setting constitutes a 'hard' test, insofar as it is characterized by transparency, limited political polarization and low levels of corruption. Hence, the study should be seen as a complement to minimal group experiments.

3.3: Pattern of queues of brief remarks

R3 argues that "the fact that certain talks - most probably on not bi-partisan but polarizing issues - leads to a queue of brief remarks point to another impact not accounted for". I am unsure of which aspect of the debates R3 refers to here, but on the question of chairman self-selection into issues, see the discussion in section 2.1 above.

3.4: Cancelling out of co-partisanship effect in the aggregate

R3 notes that the (small) impact of copartisanship cancels out in the aggregate because the chairmanship role rotates between parties. I thank R3 for highlighting this point, and I have added a mention of this to section 5.2 of the manuscript. I have also expanded the discussion of this aspect in the conclusion, cf. section 1.2 above.

3.5: Generalizability

R3 takes issue with the claim of generalizability of the result. I appreciate this point, and I fully agree that the result is not generalizable in the sense of representing in any way a random sample of theoretically relevant settings. Instead, the result allows for 'least likely' case inference because intergroup bias is *ex ante* relatively unlikely in the setting studied here. In other words, the argument concerns theoretical rather than statistical inference.

I have rewritten the conclusion to make this distinction clearer. Specifically, the revised version of the conclusion does not use the word 'generalizability'. I thank R3 for pointing out this potential confusion.

3.6: Intentionality of social identity theory

R3 notes that social identity theory does not require an "honest" wish to behave in line with rules, which my discussion on p. 4 indicates. I thank R3 for pointing out this potentially misleading phrase. The purpose of the sentence in question was to highlight that social identity theory predicts intergroup bias even if *chairmen* may have an honest wish to treat speakers neutrally. It was not meant to imply that this is a specific theoretical requirement of social identity theory. I have rewritten the sentence in question to make this distinction clearer.

3.7: Part 5 vs. Social Identity Theory

R3 argues that part 5 of the manuscript is not compatible with the social identity perspective. Part 5 of the manuscript is the entire results section, and it is unclear to which part R3 is specifically referring. However, section 5.3 in the manuscript does explore the distinction between rational choice and social identity processes (see also section 3.1 in this note). Though the test is not conclusive, it is indeed the case that the evidence is somewhat suggestive of a rational (as opposed to a social identity) mechanism. In my view the manuscript as it is is reasonably clear on this point.

3.8: Intergroup bias vs. bias

R3 argues that though the manuscript claims to show intergroup bias, "it doesn't, it shows only bias". I am unsure of precisely the distinction referred to here, but I read this comment as

interpreting "intergroup bias" to refer specifically to a social identity process. That is not the intended meaning. "Intergroup bias" does not specifically imply a social identity process, but is perfectly compatible with a rational process. In other words, a chairman who fully rationally discriminates against non-copartisan speakers (say, with the purpose of political gain) can also be said to show intergroup bias. It is my understanding that this represents a conventional use of the phrase "intergroup bias".

3.9: Reporting model significance

R3 requests that Table 2 and others report model significance. I omitted model significance statistics in the original version only for pragmatic, presentational reasons, since including them made the results tables too wide to fit on the page. In the revised version, I have now squeezed the tables together a bit, and all the regression tables now report model significance (a χ^2 statistic). As shown, all model significance tests are strongly significant, i.e. the null of no improvement in fit compared to an intercept-only model can be rejected in all models.

3.10: Significance of 'prime minister' variable

R3 asks about the meaning of the significance of the "Prime Minister" (PM) variable. The question is very reasonable, and I should have elaborated on the variable in the original version of the manuscript. In short, the PM variable captures remarks given when the MP speaking is also the sitting PM. The variable was included in the original manuscript in order to show that the effect was not driven by PMs, who may by virtue of their office be able to self-select into speaking slots (as opposed to other MP's, cf. section 2.3 in this note). As is clear from the original results, the PM variable is not a plausible confounder, as the copartisan variable remains significant throughout, and the PM variable is itself not consistently significant.

In the revised version of the manuscript, I have chosen not to include the PM variable. I have done so because upon further consideration I believe it is not theoretically relevant to distinguish between PM's and MP's in the context of this study: outside of their formal PM speeches (which are already excluded from the data), PM's have no special role or privileges in debates, and so they partake in debates only in their capacity of MP. For this reason, I have decided not to distinguish between PM's and other MP's in the revised version of the manuscript.

In any case, PM's do not drive the results of the paper. (They are not particularly active during debates, and account for only about 6 percent of all remarks). In order to show this, Table 8 of the appendix presents the results when PM remarks are excluded. As shown, the results are substantively unchanged. The table is also shown below.

I thank R3 for directing my attention to this variable, which was insufficiently explained in the original version of the manuscript.

Table 3: Results excluding prime ministers

	Speaking time (seconds)				
	(1)	(2)	(3)	(4)	(5)
Copartisan	3.42***	3.14***	3.13***	2.11**	2.67***
	(1.00)	(1.16)	(1.12)	(1.00)	(0.72)
Time of day		−0.14	−0.12	−0.09	−0.06
		(0.14)	(0.15)	(0.14)	(0.14)
Gender (female)			1.39**	0.66	0.72
			(0.58)	(0.61)	(0.60)
Debate type (Opening)			−0.72	−0.38	0.34
			(1.08)	(0.93)	(0.92)
Intercept	51.45***	53.74***	53.21***	34.28***	36.66***
	(0.93)	(2.70)	(2.83)	(2.37)	(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

Reviewer 4

4.1: Interpretation

R4 discusses my interpretation of the observed effect, noting (entirely correctly) that I do not provide direct evidence that chairmen gavel non-copartisans faster than co-partisans. As R4 notes, excluding this possibility leaves two possible mechanisms:

1. MP's self-select strategically into the debate schedule
2. Copartisan chairmen induce an unconscious bias on part of the speaker by means of subtle signals such as body language

R4 argues that the second interpretation "is [not] equivalent to debates being enforced unequally – the authors' interpretation".

I thank R4 for these considerations. However, I disagree with R4's argument in a few specific respects. First of all, while it is correct that I do not provide direct evidence of differential formal enforcement by chairmen, I do not believe this is sufficient grounds for concluding that it does not take place. Though gaveling is not recorded in the transcripts, it does happen (as can be witnessed in video recordings of debates), and so should be considered an at least possible mechanism by which the observed bias in favor of copartisans arises.

Of the two alternative mechanisms proposed by R4, the first is that MP's self-select into speaking based on the partisanship of the presiding chairman. This is indeed an important

concern, and the possibility of self-selection cannot be dismissed out of hand. In the original manuscript, the issue is partly covered by the concluding discussion of whether MP's anticipate biased rule enforcement. In the revised manuscript, I have expanded this discussion to explicitly mention the possibility of selection bias.

The other mechanism proposed by R4 is that subtle signals from the chairman such as body language, unconsciously perceived by the speaker, give rise to the observed difference, possibly in a differential way which gives rise to the slightly bimodal distribution of speaking times in the left panel of Figure 1 above. As R4 notes, "body language on the part of the presiding officer might cause some MPs to slow down and other to speed up".

I agree entirely with R4 that this is a plausible mechanism. I only disagree with R4 in R4's assessment that this does not amount to unequal enforcement of the rules. Though qualitatively different from consciously biased enforcement (i.e., the first mechanism discussed by R4), it still at the end of the day amounts to unequal enforcement insofar as it affects speakers differently based on their copartisanship with the chairman. In other words, "unequal enforcement" (or "bias") is a behavioral characterization, not a motivational one, and as such can still meaningfully be said to exist even if not necessarily consciously perceived by either chairman or speaker.

I want to stress that I found R4's discussion of this illuminating, and in the revised version of the manuscript I have extended the discussion in section 6 of how the finding should be interpreted. Specifically, the rewritten section 6 includes a discussion of the interpretation of the effect with a particular focus on the distinction between consciously vs. unconsciously unequal enforcement suggested by R4.

4.2: Unexploited information about effect heterogeneity

R4 notes that there is unexploited information about effect heterogeneity across chairmen and parties in the data.

I thank R4 for this suggestion, which I had not originally thought of. To test for effect heterogeneity, I have estimated random effects models with the same specification as model 5 in the main results, but allowing for the coefficient on the copartisan variable to vary by either chairman party or individual chairmen. The results are summarized in Figure 2 and Figure 3. In both figures, the parties are ordered from left (top) to right (bottom) based on voters' estimates of party positions.

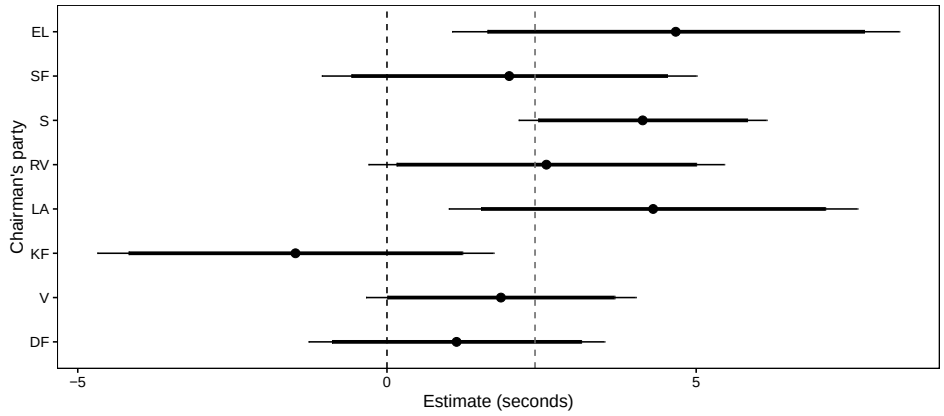

Figure 2: Estimates from a random effects model with varying slopes by chairman party. Effects are ordered by party position from left (top) to right (bottom).

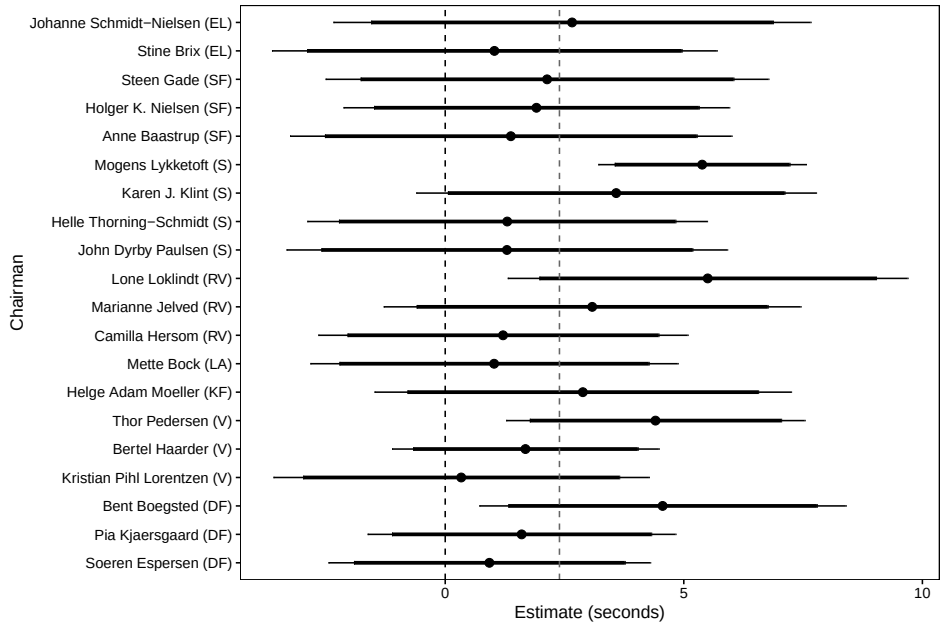

Figure 3: Estimates from a random effects model with varying slopes by chairman. Effects are ordered by party position from left (top) to right (bottom) and by descending coefficient value within parties.

Two lessons stand out in Figure 2 and Figure 3. First of all, looking at Figure 2, there appears to be no clear pattern with respect to effect heterogeneity across parties. If anything there may be a slight trend towards higher coefficients for more leftist parties, but the variance of the estimates is too high to conclude that with any certainty.

Second of all, though this trend is largely corroborated by Figure 3, there is one important

exception: one chairman, Mogens Lykketoft of the Social Democratic party (sixth from the top) is estimated to be significantly more biased in favor of copartisans than the average amount. This is significant because this study was originally inspired by allegations made by opposition parties in the summer of 2013 that the presiding head chairman was biased against the opposition (e.g., this article (in Danish)). The presiding head chairman was precisely Mogens Lykketoft. The theoretically important point of the paper is in my view the average bias, not the personal angle, but the particularly large estimate for Lykketoft lends some additional face validity to the results.

For good measure, I have also tried to estimate model 5 in the main results on separate subsets of the data excluding each individual chairmen, to see if the results are driven by any single chairman. The estimates are shown in Figure 4

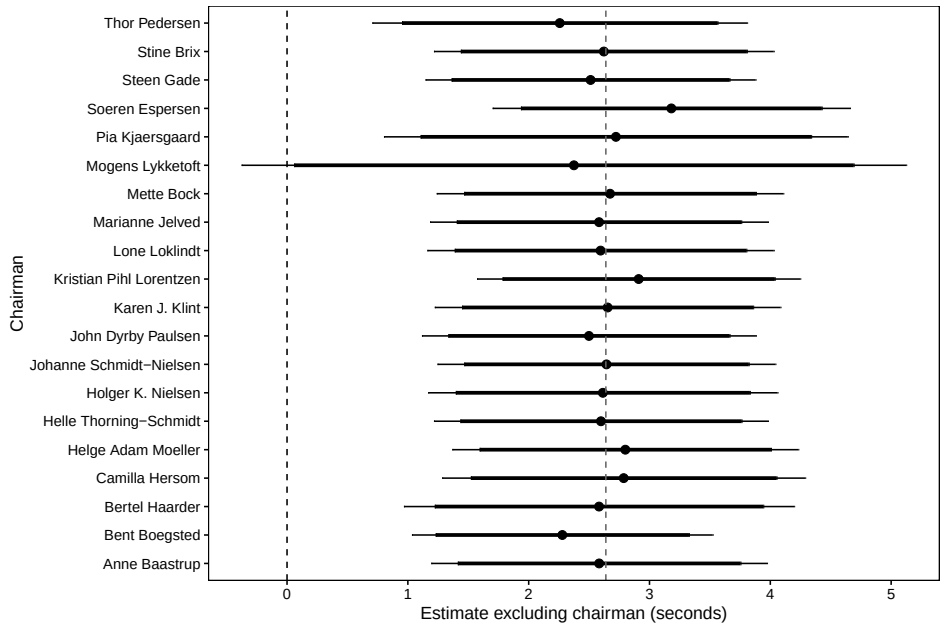

Figure 4: Estimates from model 5 in the main results on separate subsets of the data excluding each individual chairman. Effects are ordered by party position from left (top) to right (bottom) and by descending coefficient value within parties.

As shown in Figure 4, the results are robust to excluding each chairmen, demonstrating that the finding is not merely attributable to any single individual. The estimate for Mogens Lykketoft is less precise than the others because he alone accounts for nearly half of the observations, but the point estimate is in line with the others.

I have included Figure 2 and Figure 3 in the paper's appendix, and I describe the figures in the new section "Additional tests".