When all votes were recorded: An empirical assessment of selection bias in scaled preferences and party cohesion scores in the European Parliament

Philipp Broniecki and Bjørn Høyland

Department of Political Science, University of Oslo

EUSA 4-6 May, 2023

Abstract

Rollcall votes are commonly used to scale preferences of individual legislators, to compute legislator loyalty and vote-level party unity scores. One major difficulty is that rollcall votes may be requested for strategic reasons and strategic selection of rollcalls may bias downstream quantities of interest such as loyalty, unity and ideal points. From May 2020 to March 2022 – during the ninth European Parliament – most voting was recorded to allow Members of the European Parliament (MEPs) to verify their vote choices during the Parliaments remote voting regime that was applied amidst the Covid pandemic. During this time, more votes were taken by rollcall than during the previous legislative term — 10,702 compared to 10,253. We compare scaled preference estimates, legislator loyalty and party unity from the remote voting period when 88,7% percent of votes were recorded to the same estimates using data from the remainder of the ninth European Parliament (currently 3,497; 45% rollcalls). This setup allows us to empirically test how severely scaled preference estimates, legislator loyalty and party unity scores are affected by a selection bias in rollcall voting in the European Parliament. We contribute to a large literature on the representativeness of rollcall votes.

Introduction

Rollcall votes (RCVs) are widely used in legislative studies to compute party cohesion, to estimate representative's ideal points and to identify the main dimensions of political conflict. However, in many legislatures RCVs are not taken on all votes; instead they are required on some votes and in addition, they can be requested by political parties or a group of representatives. Theoretical contributions caution of selection bias induced by RCVs (Carrubba et al., 2006; Ainsley et al., 2020). However, recent empirical contributions have not found systematic and substantial biases (Yordanova and Mühlböck, 2015; Kaniok and Mocek, 2017; Hix, Noury and Roland, 2018). We contribute to the literature on rollcall votes with an empirical analysis that demonstrates substantial bias in estimates of party cohesion and ideal points in the ninth European Parliament (2019–2024).

During the period May 2020 to March 2022, in the midst of the Covid pandemic, the share of rollcall votes in the European Parliament (EP) rose from 44.5 to 88.7 percent. During this period, voting took place remotely and RCVs ensured that Members of the European Parliament (MEPs) could verify their vote choices. We find substantial and heterogenous differences between party group estimates of cohesion for the 'Remote voting period' and the remainder of the term. For ideal point estimates, we again find substantial differences between estimates from the Remote voting and regular periods. However, in contrast to the cohesion estimates, these differences are systematic. The second dimension, taken here to be the EU integration dimension, is compressed. Legislators across the political spectrum moved towards the centre. Overall, our findings suggest that estimates of party cohesion can be severely biased if RCVs are not a representative sample, as is the case, for example, in the EP. With respect to legislator ideal points we also find bias. Differences on the "EU anti-pro" dimension are overstated.

To identify bias in cohesion scores, we estimate mixed effects models that let us identify uniform remote voting effects as well as party group specific effects of remote voting. We first estimate legislator loyalty and participation. Loyalty is measured at the legislator-level and it is the share of votes MEPs voted with their groups. Second, we estimate unity which is measured at the party-group and vote-level. Unity measures party group cohesion for each group and each vote. To identify bias in ideal point estimates, we split our sample into a remote voting part and a regular voting part. For both samples, we estimate Bayesian IRT models to generate ideal points. We compare face validity and the differences in the estimates descriptively first and second we estimate mixed effects model that let us identify uniform effects and party group specific effects of remote voting.

In the following, we discus the extant literature on rollcall vote induced bias. Next, we summarise the Parliament's Rules of Procedure that govern voting in the EP. We then proceed to describe the stark differences between RCVs during the regular and remote voting periods. Our analysis proceeds in three parts. First, we analyse MEP loyalty and participation. Second, we analyse party group unity. Third, we analysis MEP ideal point estimates. We conclude by summarising our findings and their implications for empirical work based on rollcalls.

RCV selection bias: Theory and evidence

Theoretical contributions on rollcall voting suggest selection bias in measures that are generated from RCVs. Carrubba, Gabel and Hug (2008) argue that party leaders request rollcalls to discipline their rank-and-file. They demonstrate in a formal model that party cohesion can be biased to varying degrees and in ways that depend on the legislative setting and factors such as, for example, party size and heterogeneity.

Ainsley et al. (2020) show the effects of position taking as a motivation for rollcall requests. Based on a formal model they demonstrate that cohesion scores and ideal points will be biased in unpredictable ways even in a benign setting were agenda-setting power is low, a single legislator can request rollcalls and only the rollcall vote request is strategic but not the subsequent vote decision.

Empirical assessments of rollcall votes have found evidence for bias. Carrubba et al. (2006) show that RCVs are not a representative sample of all votes with respect to the method of vote, the group that requests the RCV, the responsible committee, and the type of motion. Based on a survey among MEPs, Trumm (2015) finds that MEPs defect less in rollcalls. Høyland (2010) and Hug (2016) find evidence of party pressure in their analysis of RCVs. Finke (2015) and Thierse (2016) also find evidence for bias in RCVs but argue that the patterns are in line with position taking.

Recently, several empirical studies have concluded the severity of selection bias for downstream quantities of interest such as party cohesion is either insignificant or significant but not substantial. Kaniok and Mocek (2017) replicate Carrubba et al. (2006) and find that RCVs constituted a representative sample of all votes in that year. Yordanova and Mühlböck (2015) as well as Hix, Noury and Roland (2018) use data from the EP to address the effect of party discipling on cohesion scores. Based on the discipling logic, one would expect that legislators are more loyal when their parties request RCVs. Therefore, RCV-based cohesion scores should be upwards biased. Yordanova and Mühlböck (2015) use a 2009 rule change that made RCVs on all final legislative votes required and compare party group cohesion of the big center right and center left groups in a matched sample of votes before and after the rule change. They show that cohesion was larger after the reform when RCVs became required and conclude that RCV-based cohesion scores may provide an over-estimate of voting loyalty rather than an underestimate. Hix, Noury and Roland (2018) investigate the same rule change and with a different estimation strategy do not detect a difference in party cohesion before and after the rule change.

Overall, rollcall votes are widely used in legislative politics to estimate various quantities of interest (Ainsley et al., 2020). At a theoretical level, there is agreement that rollcall vote samples are biased. At the empirical level, early contributions seemed to confirm this expectation. However, recently, and in specifically for the the European Parliament, scholars have concluded that the bias in quantities of interest such as party cohesion could be inconsequential. We contribute to this debate with new data from the ninth term of the EP. During a remote voting period amidst the Covid pandemic around 90 percent of all votes were RCVs. We demonstrate substantial and heterogenous differences between cohesion scores from the remote and regular voting periods. We also demonstrate large biases in ideal point estimates based on our data.

RCV rules during the ninth EP

In the EP, voting takes place either by show of hands or electronically via voting boxes placed on the desks in plenary. MEPs authenticate themselves via their ID cards. The term electronic vote, in the EP, refers to votes where aggregate outcomes are reported. Secret votes only report vote outcomes and rollcall votes record the vote choice of each individual.

Rollcalls are required for final votes on legislative reports (currently roughly a tenth of a percent of all votes). For all other votes, RCVs must be requested by a political group or one twentieth of the EPs component members (currently 35 MEPs). Amidst the Covid pandemic, the EP instated a remote voting procedure by email. “To ensure that MEPs could verify that their votes were correctly registered, it was agreed that all votes, unless otherwise specified in the rules of procedure, should be taken by roll call.

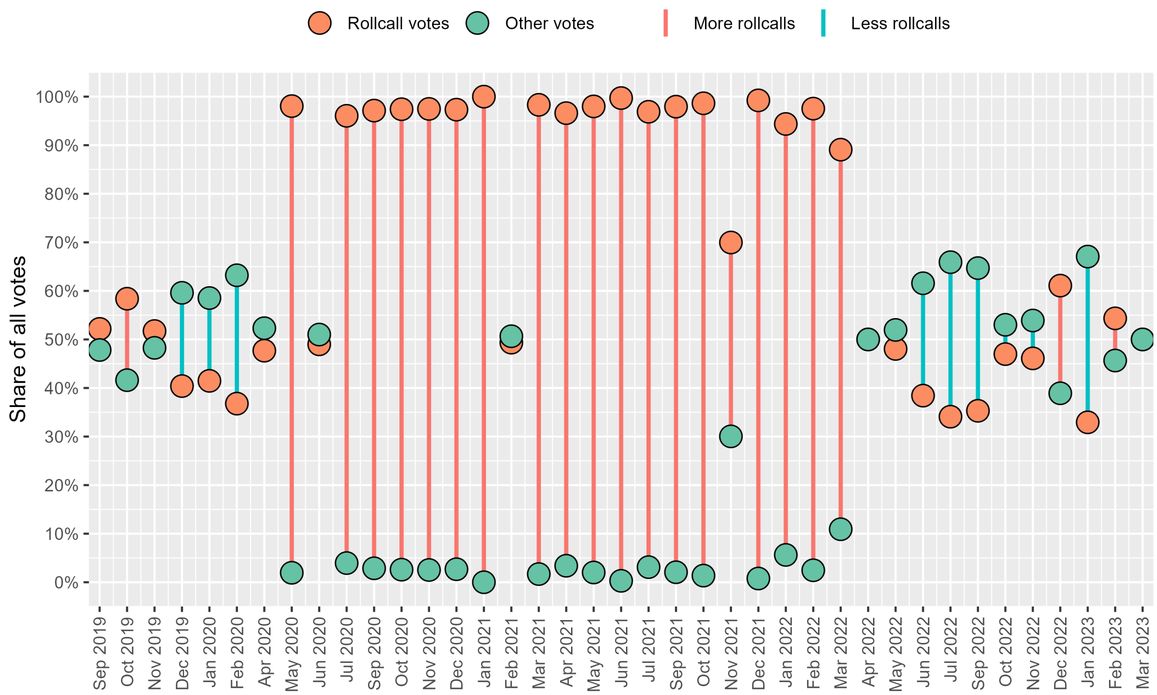
These changes were formalized in changes to the Rules of Procedure in September 2021 (Rule 237c), where it specified that the remote voting procedure should be technical neutral, but must enable MEPs to verify that their votes are counted as cast" (Broniecki and Høyland, 2023). For a more detailed overview of the rules regarding RCVs and the changes thereof, refer to Broniecki and Høyland (2023).

European Parliament voting data

We web-scraped our data from the European Parliament (europarl.europa.eu). Specifically, we collected the plenary minutes, the rollcall votes results tables and the results of all votes tables from the plenary sessions. Our data differs from conventional data on rollcall votes in the European Parliament, such as VoteWatch data, because in addition to rollcall votes, it also includes other votes that were not taken by rollcall. Our data allows us to compare the proportion of rollcall votes to all votes taken in plenary.

We validated our data by collecting the rollcall votes results data which includes rollcalls only as well as the all vote results table. As of 16 April 2023, the rollcall votes data contains 15,328 rollcalls during the ninth term of the European Parliament. Of these, we were able to match 14,412 in the results of all votes table (94%). The discrepancy is caused by agenda votes and some special votes that are not recorded in the overall votes table. For example, the rollcall taken on 13 December 2022 to terminate EP Vice-President Eva Kaili's term over alleged corruption is not recorded in the all votes table. Despite these differences between the data sources, we were able to match 94% of the rollcalls including all legislative and all final votes which overall gives us confidence in the accuracy of our data. We dropped the votes that were not recorded in both tables from our analysis. Furthermore, we dropped votes with missing references where we could not determine whether the vote was legislative,

Figure 1: Share of Rollcall Votes and all Other Votes 2019–2023



non-legislative or budgetary. Finally, we dropped all lapsed votes. After subsetting, our data contains contains 20,231 votes and 14,412 rollcalls (71%).

Description of voting data

The Covid pandemic changed voting patterns in the EP as Figure 1 illustrates. The share of rollcall votes abruptly increases in May 2020 and remains high until March 2022 after which it returns to pre-crisis levels.

We split our sample into a remote voting period which starts in May 2020 and ends in March 2022 and a regular period which includes the time before and after the remote voting period. Throughout EP 9 the share of rollcalls is 71.2%. This would be a large increase from EP 8 where the share of rollcalls was 32.5%[1] However, this masks the differences between the periods. During the Covid period, the share of rollcalls was 88.7% compared to 44.5% during the regular period. The difference is most extreme for legislative votes where the covid period share is 94.8% compared to 41% for the remainder of the term.

Table 1: Share of Rollcall Votes during Normal Operations and Remote Voting

Vote type	Regular Period	Remote Voting Period	Both Periods
Non-legislative	45.6 %	87.1%	71.7%
Legislative	41.0%	94.8%	70.1%
Budgetary	58.0%	78.6%	70.3%
All	44.5%	88.7%	71.2%

Overall, it is clear that the remote voting period is very different from the remainder of the term. During most remote voting months, the share of rollcalls is more than 90%. There are a few noticeable exceptions. In June 2020 and in February 2021 the share of rollcalls is slightly below 50%. There were different periods of lockdown in MEPs' home countries. We are not sure, at this point, what the explanation for these exceptions is and whether MEPs were present in person or not during that time. The overall number of votes in June 2020 and February 2021 was high but not exceptional (1141 and 904 votes respectively). In November 2021, the share of rollcalls dipped to 70% (253 votes). We do not have an explanation for these remote voting period exceptions, but we exclude them from our analysis in a robustness check.

[1] We collected EP 8 data similar to EP 9 data. The overall number of votes in EP 8 is 30,434. We dropped votes without references as well as lapsed votes.

Analysis

In the following, we assess whether selection effects induced by the votes that are decided by rollcall affect estimates of legislator loyalty, vote-level party unity, as well as ideal point estimates in EP 9.

Party group loyalty and participation

We assess the loyalty of each MEP including and excluding abstentions as well as the participation levels of MEPs. We construct three dependent variables. The first dependent variable *loyalty* is the average number of votes that an MEP voted with the party group. Abstentions and absences are dropped from the data. The second dependent variable *loyalty with abstentions* is similar but includes abstentions and drops absences from the data. In general, the degree of loyalty is much lower when we include abstentions. The third dependent variable *participation* is the average across all votes where an MEP voted yes, no or abstained. We normalise all three variables to mean zero and standard deviation one. Our unit of observation is a combination of MEP, party group and period.

To determine the difference between the remote period and the regular period in EP 9, we estimate linear mixed effects models where the fixed part is the predictor variable: whether voting took place during the remote period or not. We include random effects for MEPs and EPGs.

Table 2 summarises our results. The main explanatory variable *Remote voting period* which is an indicator of whether voting took place during the remote period or not is significant in the loyalty models but not in the participation model. With respect to the relevance of statistical significance for our analysis, we note that we include nearly all rollcall votes taken in EP 9 except agenda votes and some special

Table 2: Mixed Effects Models on Party Group Loyalty & Participation

	Loyalty	Loyalty with abstentions	Participation
Remote voting period	0.15^{***}	0.10^{***}	0.02
	(0.03)	(0.02)	(0.02)
AIC	4113.03	2654.70	2826.34
BIC	4140.40	2682.07	2853.71
Log Likelihood	−2051.52	−1322.35	−1408.17
Num. obs.	1761	1761	1761
Num. groups: mepid	831	831	831
Num. groups: epg	8	8	8
Var: mepid (Intercept)	0.27	0.16	0.30
Var: epg (Intercept)	0.51	0.81	0.60
Var: Residual	0.38	0.15	0.12

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$. Linear mixed effects models. Fixed effects for the remote voting period. Random effects for MEPs and European Party groups.

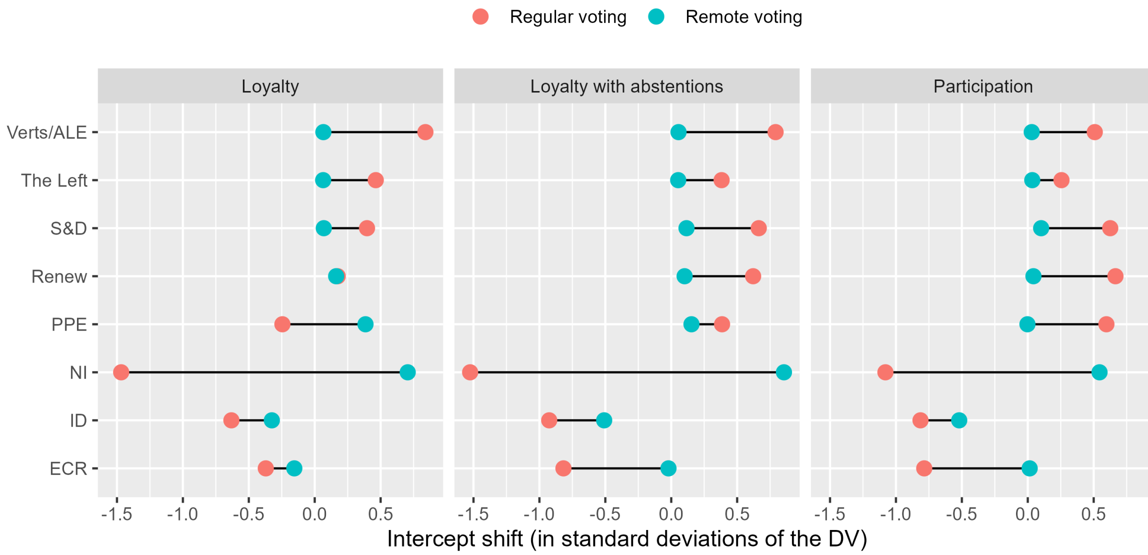
votes. Hence, we may ignore standard errors and interpret the coefficients as the actual differences between the remote period and the regular period in EP 9.

Loyalty, excluding abstentions, was on average across all party groups and MEPs, 0.15 standard deviations higher during the remote period than during the regular period. When we include abstentions, party group loyalty was 0.10 standard deviations higher on average during the remote period. With respect to participation, there is no substantial difference between the remote period and the regular period.

Our results demonstrate general differences between the remote period and the regular period for estimates of party group loyalty. Overall, loyalty is higher when the proportion of rollcalls is higher during the remote period.

Our overall results mask substantial variation at the party group level as Table 2 also illustrates. Most variation is at the party group level. To investigate party group level differences, we re-estimate the model such that it includes party group level random effects that vary by period. Figure 2 illustrates the differences between the remote period and the regular operations periods for each party group.

Figure 2: Party Group Level Differences in Loyalty & Participation



MEPs belonging to the Greens (Verts/ALE), the Left Group and the Socialists & Democrats were less loyal during Covid period. The differences are large at between 0.9 and 0.5 standard deviations in party group loyalty. For the liberal renew group, there is seems to be no difference between the remote period and the regular period. For the PPE, Identity and Democracy and the European Conservatives and Reformists, loyalty was higher during the remote period. The difference for the centre-right PPE is 0.6 standard deviations.

The results between loyalty and loyalty including abstentions also differ in some respects. Loyalty is now lower during the remote period for all groups except the non-attached, Identity and Democracy and ECR. Similarly, looking at participation we observe that participation was lower during the remote period for all groups except the non-attached members, the ID and ECR groups.

Overall, our analysis shows that there are substantial and heterogenous differences in party group loyalty during the remote period and during the regular period. During remote voting, the share of rollcalls was over 90 percent most of the time. If the Covid period is a good test for a hypothetical scenario in which most of the votes in the EP are recorded then one would conclude, that currently selection effects bias our sample of rollcalls in substantial but also heterogenous ways. One could not conclude, for example, that party loyalty scores based on rollcalls are generally an underestimate or an overestimate of real unobserved party loyalty.

Vote-level cohesion

In the following analysis, we asses vote-level cohesion. Our unit of analysis is a vote and party group combination. For each combination of vote and party group, we calculate a cohesion score based on Hix, Noury and Roland (2007, p. 91):

$$\frac{max(Y_i, N_i, A_i) - \frac{1}{2}[(Y_i + N_i + A_i) - max(Y_i, N_i, A_i)]}{Y_i + N_i + A_i}$$

where Y_i denotes the number Yes votes, N_i the number of No votes and A_i the number of abstentions. Larger scores correspond to higher party group cohesion. We normalise the cohesion score to mean 0 and standard deviation 1. The normalised cohesion score is the dependent variable. Our main predictor variable *Remote voting period* indicates whether the vote took place in the remote voting period or not. In addition, we control for an indicator of whether the vote was a final vote or not and the type of the vote. Our baseline vote type is non-legislative and our model includes binary variables for whether the vote was legislative or budgetary.

Table 3 summarises our findings. Vote-level cohesion was 0.06 standard deviations higher during the remote voting period. In addition, cohesion was 0.03 standard deviations higher during legislative votes compared to non-legislative votes. Table 3 also reveals a substantial amount of variation on the party group level. To investigate

Table 3: Mixed Effects Models on Vote-Level Cohesion

	Vote-level Cohesion
Remote voting period	0.06 (0.01)***
Final vote	0.01 (0.01)
Legislative vote	0.03 (0.01)***
Budgetary vote	−0.01 (0.02)
AIC	287799.30
BIC	287866.88
Log Likelihood	−143892.65
Num. obs.	115173
Num. groups: epg	8
Var: epg (Intercept)	0.33
Var: Residual	0.71

***$p < 0.001$; **$p < 0.01$; *$p < 0.05$. Linear mixed effects model. Fixed effects for the remote voting period, final votes and vote type. Vote type categories: Legislative vote, budgetary vote and the baseline category non-legislative vote. Random effects for European Party groups. Dependent variable normalised to mean 0 and standard deviation 1.

party group level differences we re-estimate the model such that it includes party group level random effects that vary by period on samples that are split by the vote type: Non-legislative, legislative and budgetary.

Figure 3: Party Group Level Differences in Cohesion by Vote Type

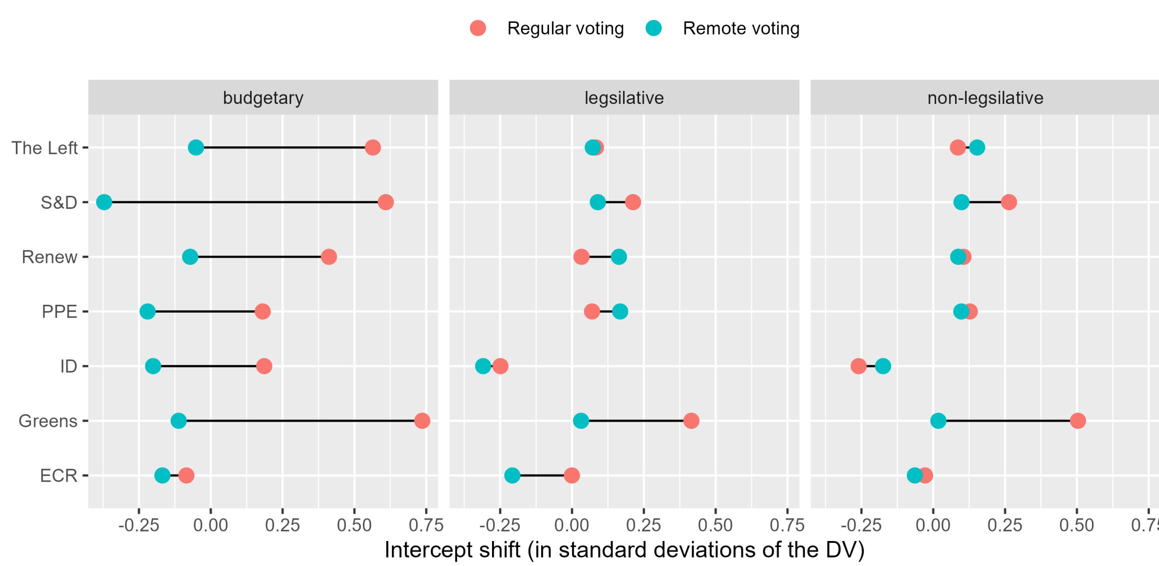


Figure 3 illustrates the differences between the remote period and the regular operations periods for each party group. The panels show budgetary votes, legislative and non-legislative votes. For budgetary votes, cohesion is consistently lower during the remote period. However, the degrees of change vary substantially. The changes are largest for the Greens and the Socialists and Democrats at around one standard deviation. For Renew, the PPE and the ID group, the changes are roughly half a standard deviation and for the ECR group, the change is just negligible. The changes for legislative votes and non-legislative votes are smaller. The Greens, however, stand out as being around half a standard deviation less cohesive on non-legislative votes and 0.44 standard deviations less cohesive on legislative votes. We excluded non-attached members from Figure 3. Similar to loyalty, the changes are very large for the non-attached MEP. Voting cohesion dropped between 1.5 and 2 standard deviations, suggesting that there was no coordination during the remote period between those representatives.

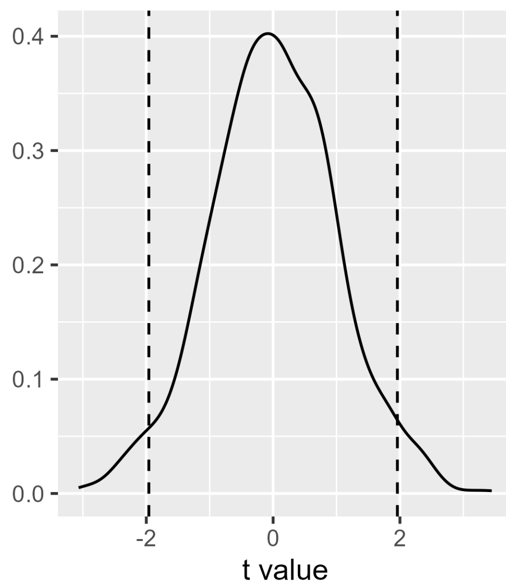
All in all, our analysis suggests large heterogenous differences between the remote voting period and the rest of the term for party cohesion estimates and loyalty scores. During the regular period, the share of rollcall votes was 44.5% and during the remote period, it was 88.7%. We argue that the remote period is a test case for a scenario where most legislation is taken by rollcall in the European Parliament. The differences in our estimates between the periods suggests that rollcall-based estimates of party cohesion or loyalty are severely biased by the selection of rollcall votes.

We cannot be certain that votes during the remote period where not very different from the regular period, i.e. that the remote period is a fair test case for measuring selection bias induced by rollcalls. However, if for example all legislation during the remote period had been urgent and uncontroversial, we would expect to see increased cohesion across the board for most party groups. The fact that the period-to-period changes are heterogenous decreases our confidence in rollcall-based measures of cohesion when the share of rollcalls is slow as is typically the case in the EP.

Placebo test for the remote voting period

In the following, we illustrate that the large changes in party cohesion are unlikely to be chance findings. We test whether we would observe similar differences period-to-period were we to split our data into a placebo period and a control period where the share of placebo cases reflects the share of votes in the actual remote voting period (75%).

Figure 4: Placebo Effect of the Remote Voting Period



Note: Distribution of t values of the remote voting period fixed effect from 1000 simulations. We randomly split the data into 75% placebo cases and 25% control cases reflecting the share of actual remote voting period votes. We estimate the same regression on vote-level cohesion as shown in Table 3.

Figure 4 shows the distribution of t-values of the fixed effect of the remote period obtained from 1000 simulations. We estimate the same model as shown in Figure 3 where we regress vote-level cohesion on the remote period, a final vote indicator, the vote type and party group random effects. The distribution is consistent with a null effect. It is centred on zero, 3% of the effects are significant and negative and 3% of the effects are significant and positive. To compare, the t-value we obtained in our model is 11.1.[2] Overall, the placebo test increases our confidence that the observed change in party cohesion between the periods is not a chance finding.

Ideal point estimates

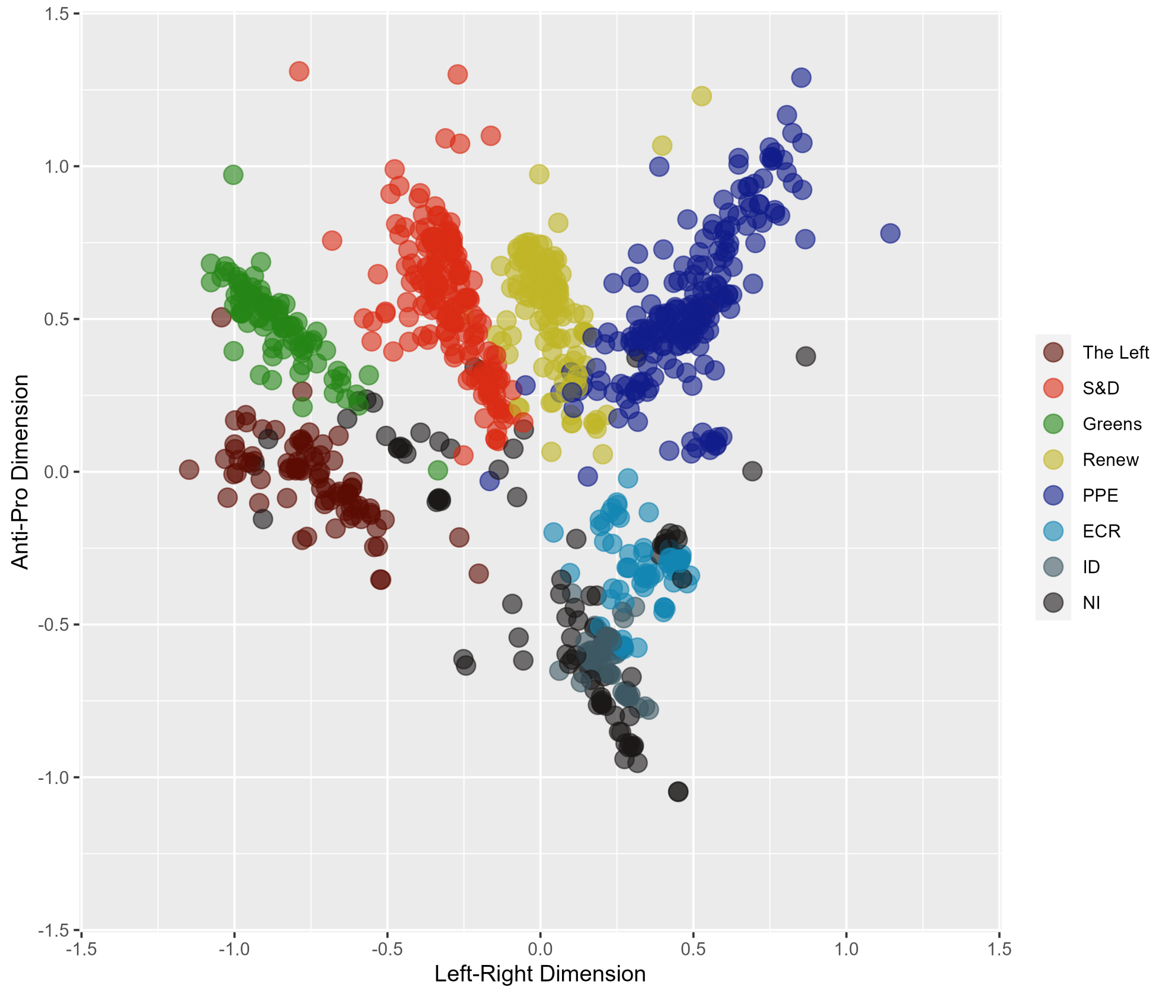
In the following, we estimate legislator ideal points using the Bayesian IRT approach (Martin and Quinn, 2002; Clinton, Jackman and Rivers, 2004) and the pscl package (Jackman, 2017). We estimate ideal points in two dimensions. To identify the model, we set prior means for all legislators. On the first dimension, priors for MEPs from the Left are set to -1, MEPs from ECR to 1 and all others are set to 0. On the second dimension, we set members of the ID groups, prior means to -1, MEPs from Renew to 1 and all others to 0. We label the first dimension *left-right* and the second dimension *anti-pro*.

We estimate our model for the two split samples. The first includes only remote voting period votes and the second includes only votes from the regular period. We set the maximum number of iterations to 10,000, the burnin period to 5000 and thinning to 100. Figure 5 illustrates the results for the regular period. Face validity seems to be good for this model. MEPs cluster together in their party groups and the relative positions of the party groups in the political space correspond to our intuition. One exception is that the Left Group is not distinguishably more left on the left-right dimension than the Greens. An interesting observation is that most within party group variation for all groups is on the European anti-pro dimension rather than on the left-right dimension.

Figure 6 illustrates the results for the remote period. The party group clusters

[2] In Table 3, the t value appears to be 6 which is due to rounding error.

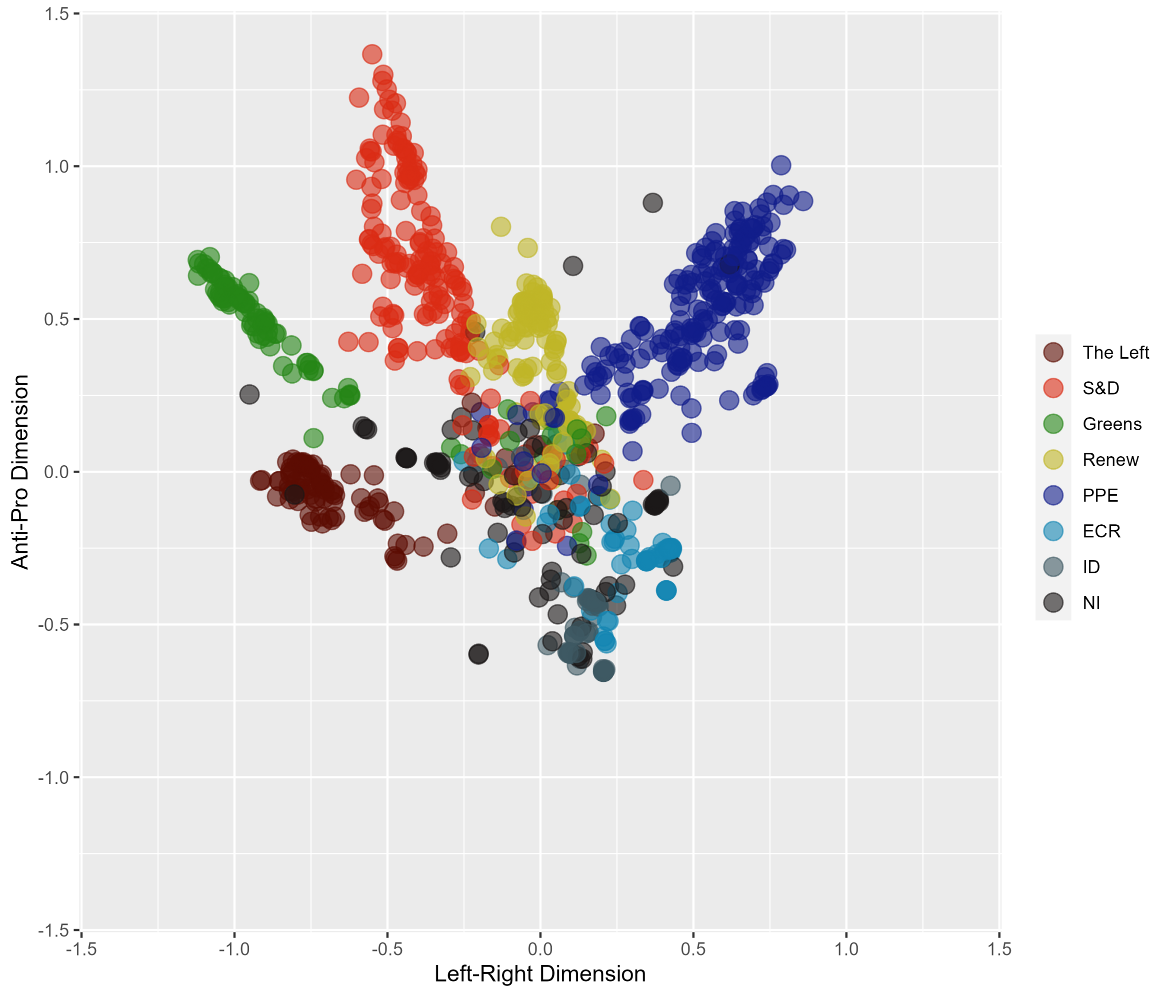
Figure 5: Ideal Point Estimates Regular Period



Note: Mean ideal point estimates from a Bayesian IRT model with 10,000 iterations, 5000 burnin and thinning 100.

are less distinguishable in the centre of the political space. Eurosceptic MEPs from the ID and ECR groups as well as the non-attached members appear less Eurosceptic. One explanation for this difference could be that voting during the remote period was from regular voting. A second explanation would that the sample of rollcalls that we usually have exaggerates differences on the Eurosceptic dimension because of rollcall vote selection bias. A third explanation would be the dimension that we labelled anti-pro was actually about something else during the remote period. Face validity on the

Figure 6: Ideal Point Estimates Remote Voting Period

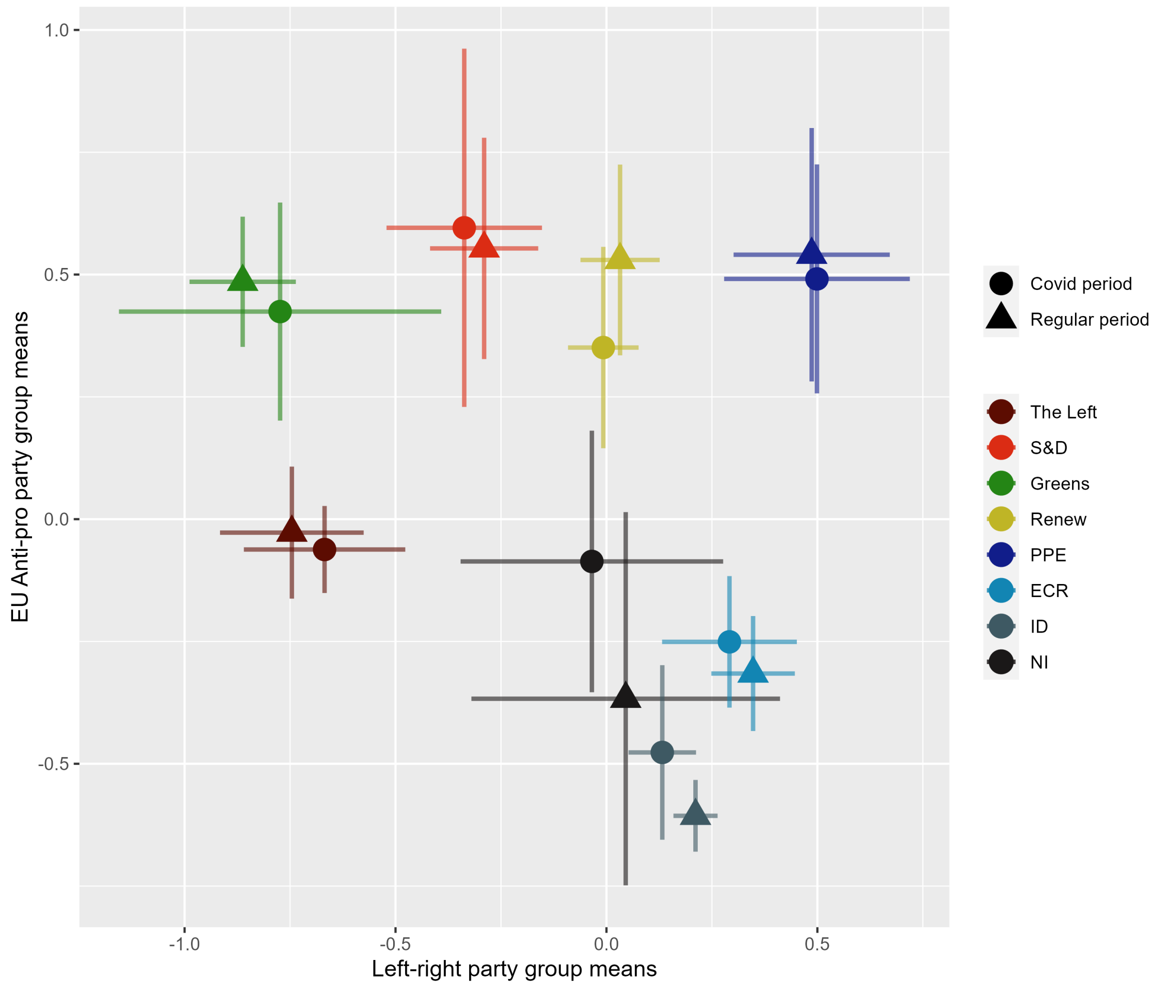


Note: Mean ideal point estimates from a Bayesian IRT model with 10,000 iterations, 5000 burnin and thinning 100.

left-right dimension is still high with the exception that many Greens are to the left of MEPs from the Left group.

The ideal point estimates are substantially different but the differences appear perhaps bigger at first sight than they are. Figure 7 illustrates the changes of the party group means and the sizes of the group standard deviations on the left-right and anti-pro dimensions. The standard deviations are slightly higher during the remote period but on the left-right dimension the shift is samller than on the anti-pro dimension. The

Figure 7: Changes of Party Group Means and Standard Deviations



Note: Party group means and standard deviations for the regular and remote voting periods. Horizontal bars are party group and period standard deviations on the left-right dimension. Vertical bars are the period and group standard deviations on the anti-pro dimension.

biggest shift occurred for the non-attached members who moved to the centre. The ECR, ID, Renew, The Left, and the Greens all moved towards the centre. For each of these groups, the shift is noticeable but not that large. The largest shift is about half a standard deviation for the ID group. The estimates for the European People's party and for the Left Group are quite stable.

To test for a systematic effect of the remote voting period, we estimate mixed effects models, on the squared left-right and squared anti-pro dimensions. We include a fixed

Table 4: Mixed Effects Models on Folded Left-Right and Anti-Pro Dimensions

	Left-right squared	Anti-pro squared
Remote voting period	-0.02	-0.23^{***}
	(0.03)	(0.04)
AIC	3871.58	5161.27
BIC	3899.34	5189.04
Log Likelihood	-1930.79	-2575.64
Num. obs.	1906	1906
Num. groups: mepid	831	831
Num. groups: party	8	8
Var: mepid (Intercept)	0.17	0.11
Var: party (Intercept)	0.99	0.73
Var: Residual	0.30	0.75

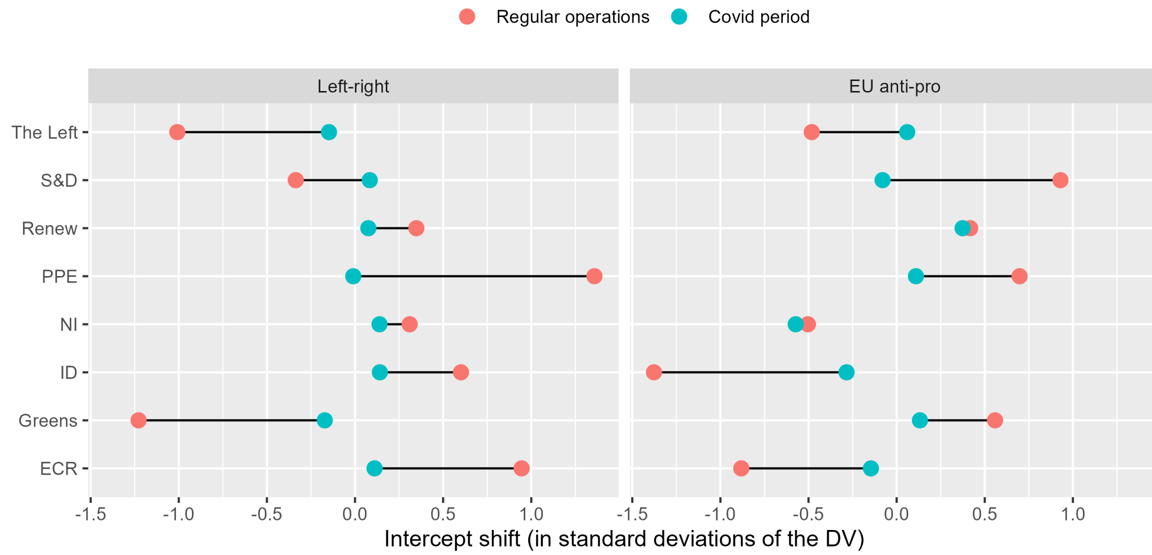
$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$. Linear mixed effects models. Fixed effects for the remote voting period. Random effects for European party groups. Dependent variables squared. A negative coefficient implies a move to the centre of the political space.

effect for the remote period and random effects for the European party groups and the MEPs. Squaring the dimensions will fold them such that a positive coefficient implies a move towards the extremes of the political space and negative coefficients imply a move towards the centre of the political space.

Table 4 illustrates the systematic effect of the remote voting period. On average the political groups moved towards the centre on the anti-pro dimension. To better compare shifts in ideal point estimates to the shifts in cohesion scores, we estimate a linear random effects model where we regress each of the dimensions on an MEP random effect as well as a party group random effect that varies by period. We normalize ideal points on both dimensions to mean 0 and standard deviation 1.

Figure 8 illustrates the changes. The changes in the party group random effect are measured in standard deviation shifts of the Left-right and EU anti-pro dimensions. The shifts are equally large if not larger than the shifts we observed for the cohesion scores. In contrast to the findings for the cohesion scores, the effects are systematic.

Figure 8: Changes of Party Group Means and Standard Deviations



Note: Party group means and standard deviations for the regular and remote voting periods. Horizontal bars are party group and period standard deviations on the left-right dimension. Vertical bars are the period and group standard deviations on the anti-pro dimension.

Overall, the party groups move towards the centre on both dimensions. This effect could imply that if all votes are recorded, the signaling and discipling effects are smaller. All in all, from these differences we conclude that estimates of legislator ideal points are biased if the share of rollcall votes is low.

Conclusion

Rollcall votes are widely used in legislative studies to estimate different quantities of interest such as party cohesion and legislator ideal points. Recent studies suggest that the downstream effects of selection bias in rollcalls on these quantities of interest are low. In this paper, we demonstrate that this is not the case for the ninth European Parliament (EP).

We collected data on all votes taken in the EP during 2019 to 2023. During the legislative term amidst the Covid pandemic, the EP instated a remote voting procedure that last roughly from May 2020 to March 2022. The share of rollcalls during this period rose from 44.5% to 88.7%. We take the remote voting period as a case where nearly all votes are rollcalls and hence during that period, rollcalls are a representative unbiased sample of all votes.

We compared legislator-level loyalty, party-vote-level cohesion and legislator ideal points between these periods. We estimated mixed effects models that let us identify uniform as well as differential party-group effects of the remote voting period on these quantities of interest. Loyalty and party unity/cohesion are very different during the remote and regular voting periods. We conclude that cohesion scores based on rollcalls can be severely biased. What is more, the biases that we found were different for different party groups. We also found substantial differences between the ideal point estimates from the remote voting and the regular periods. However, in contrast to estimates of loyalty or cohesion, the ideal point estimates were systematically different. The EU anti-pro dimension was compressed. One might argue that RCVs overstate differences between the legislators on the anti-pro dimension.

In general, party loyalty and unity/cohesion are more dangerously biased because the scores were affected in unpredictable ways across the party groups. The differences in our ideal point estimates were also substantial but the bias was systematic. In addition, we used a Bayesian IRT framework to estimate our ideal points. In principle, one could mitigate selection bias in rollcalls to some extent by including prior knowledge such as expert party assessments. Overall, our findings suggest that worries about selection bias in empirical analyses of rollcalls should be taken seriously.

References

Ainsley, Caitlin, Clifford J. Carrubba, Brian F. Crisp, Betul Damirkaya, Matthew J. Gabel and Dino Hadzic. 2020. "Roll-Call Vote Selection: Implications for the Study of Legislative Politics." *American Political Science Review* 114(3):691 – 706.

Broniecki, Philipp and Bjørn Høyland. 2023. "Requested or Required: The Changing Composition of Roll Call Votes in the European Parliament." Presented at the 18th biannual EUSA Conference, Pittsburgh, USA.

Carrubba, Clifford J., Matthew Gabel, Lacey Murrah, Ryan Clough, Elizabeth Montgomery and Rebecca Schambach. 2006. "Off the Record:Unrecorded Legislative Votes, Selection Bias and Roll-Call Vote Analysis." *British Journal of Political Science* 36(4):691 – 704.

Carrubba, Clifford, Matthew Gabel and Simon Hug. 2008. "Legislative Voting Behavior, Seen and Unseen: A Theory of Roll-Call Vote Selection." *Legislative Studies Quarterly* XXXIII(4):543 – 572.

Clinton, Joshua, Simon Jackman and Douglas Rivers. 2004. "The Statistical Analysis of Roll Call Data." *American Politcal Science Review* 98(2):355 – 370.

Finke, Daniel. 2015. "Why do European political groups call the roll?" *Party Politics* 21(5):750 –762.

Hix, Simon, Abdul G Noury and Gérard Roland. 2007. *Democratic politics in the European Parliament*. Cambridge University Press.

Hix, Simon, Abdul Noury and Gerard Roland. 2018. "Is there a selection bias in roll call votes? Evidence from the European Parliament." *Public Choice* 176(1/2):211 – 228.

Hug, Simon. 2016. "Party pressure in the European Parliament." *European Union Politics* 17(2):201 – 218.

Høyland, Bjørn. 2010. "Procedural and party effects in European Parliament roll-call votes." *European Union Politics* 11(4):597 – 613.

Jackman, Simon. 2017. *pscl: Classes and methods for R developed in the Political Science Computational Laboratory*. Sydney, New South Wales, Australia: United States Studies Centre, University of Sydney. R package version 1.5.2.
URL: *https://github.com/atahk/pscl/*

Kaniok, Petr and Ondrej Mocek. 2017. "Roll Call Votes in the European Parliament: a good sample or a poisoned dead end?" *Parliaments, Estates & Representation* 37(1):75 – 88.

Martin, Andrew D. and Kevin M. Quinn. 2002. "Dynamic ideal point estimation via Markov chain Monte Carlo for the US Supreme Court, 1953–1999." *Political Analysis* 10(2):134–153.

Thierse, Stefan. 2016. "Going on record: Revisiting the logic of roll-call vote requests in the European Parliament." *European Union Politics* 17(2):219 – 241.

Trumm, Siim. 2015. "Voting Procedures and Parliamentary Representation in the European Parliament." *Journal of Common Market Studies* 53(5):1126 – 1142.

Yordanova, Nikoleta and Monika Mühlböck. 2015. "Tracing the selection bias in roll call votes: party group cohesion in the European Parliament." *European Political Science Review* 7(3):373 – 399.