European Political Science Review (2015), 7:3, 373–399 © *European Consortium for Political Research*
doi:10.1017/S175577391400023X
First published online 28 November 2014

Tracing the selection bias in roll call votes: party group cohesion in the European Parliament

NIKOLETA YORDANOVA[1*a] AND MONIKA MÜHLBÖCK[2a]

[1]*School of Social Sciences, University of Mannheim, Mannheim, Germany*
[2]*Department of Economic Sociology, University of Vienna, Vienna, Austria*

Legislative politics scholars rely heavily on roll call vote (RCV) data. However, it has been claimed that strategic motives behind RCV requests lead to overestimating party group cohesion and, thus, biased findings on legislative behaviour. To explore this claim, we distinguish between two types of bias, a 'behavioural bias' and a 'selection bias'. A recent rule change in the European Parliament, making RCVs mandatory on all final legislative votes, presents the unique opportunity to evaluate the latter. We compare party group cohesion in requested and mandatory RCVs by examining final legislative votes before and after the rule adoption using amendment RCVs (which still need to be requested) as a benchmark. The analysis shows that group cohesion is higher whenever RCVs are not just requested on *some* but mandatory on *all* votes. Hence, there is indeed a 'selection bias' in RCV data. Yet, somewhat contrary to former claims, relying on requested RCVs leads to underestimation of the cohesion party groups would have had were all votes automatically roll called. We argue that this is mainly because requests occur on more contentious votes.

Keywords: roll call votes; party groups; European Parliament

Introduction

Our knowledge about legislative behaviour, coalition formation, the dimensionality of conflict in legislative assemblies, as well as party group cohesion, is to a large extent based on roll call vote (RCV) data. In RCVs, the voting behaviour of each representative is recorded in the minutes. However, in many legislatures only a fraction of the votes is taken by roll call (Carey, 2009; Hug, 2009; Stecker, 2010). The sample does not constitute a random selection of votes as RCVs are usually requested by party groups. This selection mechanism has drawn extensive scholarly attention. Two main theoretical explanations of RCV requests have been proposed in recent years: disciplining party members (e.g. Carrubba *et al.*, 2008) and signalling a certain position (Thiem, 2006; Finke and Thiem, 2010; Ainsley and

[a] The order of the authors' names reflects the principle of rotation. Both authors have contributed equally to all work.

* E-mails: yordanova@uni-mannheim.de, mo.muehlboeck@univie.ac.at

Maxwell, 2012). Scholars have claimed that both can lead to a bias in RCV data and, therefore, biased findings such as overestimation of party group cohesion (Carrubba *et al.*, 2006; Thiem, 2006).

However, empirical evidence to evaluate this claim is scarce because without information on the votes which are not recorded, we cannot know the actual overall voting cohesion, i.e. how legislators vote in both RCVs and non-RCVs.[1] We thus lack a direct comparison and empirical evaluation of the implications of the different theories on RCV requests. To fill in this gap, we first distinguish conceptually between the different types of bias: (1) a 'behavioural bias' caused by a change in vote choice if the vote is roll called, and (2) a 'selection bias' caused by selecting an unrepresentative sample of votes for roll call. Focusing on the study of the alleged selection bias in RCVs, we then address the question: 'What level of party group cohesion would we observe if all votes were roll called rather than only those selected, potentially strategically, by party groups?'.

Recently available data enables us to do so. In particular, following a strategy suggested by Høyland (2010), we take advantage of a change in the Rules of Procedure of the European Parliament (EP) in June 2009, which made a roll call mandatory on all final legislative votes. While before the rule change RCVs had to be requested by a group of Members of the European Parliament (MEPs), after the rule change voting behaviour became observable in the full population of final legislative votes. We can thus explore the difference between the average party group cohesion in a sample of requested RCVs and the average party group cohesion measured if voting behaviour is observable for the full population of votes, using as a benchmark amendment legislative votes,[2] on which roll calls still need to be requested.

We examine all legislative votes in the first 18 months of the 6th (EP6) and the 7th (EP7) European Parliament, at the start of which the new rule entered into force. Using fractional logistic regression models (Papke and Wooldridge, 1996) in combination with exact matching (Iacus *et al.*, 2009), thereby controlling for a number of factors, we demonstrate that the observable group cohesion on legislative votes has increased in EP7, and significantly more so on final than on amendment votes. The explanation for this finding we offer is that the now mandatory final legislative RCVs include a great number of non-contentious votes on which no roll call would have been requested under the old rules.

Based on our results, we offer the first direct empirical evidence that there is indeed a selection bias in RCV data. However, somewhat contrary to previous claims in the literature, we find that when relying on a sample of requested final legislative RCVs in the EP we actually underestimate the cohesion we would observe if all such votes were roll called.

[1] For an exception, see Hug (2006, 2009) on the Swiss parliament.

[2] For convenience sake, we refer to all non-final legislative votes as amendment votes throughout the paper.

https://doi.org/10.1017/S175577391400023X Published online by Cambridge University Press

The paper proceeds with a short overview of the literature on RCVs. We then examine the implications of different theories with respect to the two types of bias. Thereafter, we deduce hypotheses concerning the selection bias and explain how we can test them with our data. Subsequently, we test our hypotheses and conclude with a discussion of the findings and their implications for the study of legislative behaviour.

Party group cohesion in RCVs: the case of the EP

In many parliaments, a group of representatives is entitled to request a vote to be held by roll call, whereby the individual vote choice of each legislator is recorded (for an overview of rules in different parliaments, see Hug, 2006). Researchers have tackled the question of why RCVs are requested, looking for differences between legislatures (Carey, 2009; Stecker, 2010), time periods (Lynch and Madonna, 2013), party groups (Saalfeld, 1995), and vote-specific attributes (Carrubba *et al.*, 2006; Finke and Thiem, 2010). They agree that party groups are unlikely to request roll calls randomly but rather do so for strategic reasons. Consequently, the sample of RCVs is unlikely to be representative of all votes and findings about legislative behaviour based on RCV data may be biased. Specifically, the observed party group cohesion in these recorded votes may be affected by the motives behind requesting them.

While the potential bias applies to all parliamentary assemblies where RCVs are requested by individual party groups, the EP is an often cited case. This may be attributed to two factors. First, the number of RCVs is comparably high – according to Hix *et al.* (2007), about one-third of all EP votes are taken by roll call. This provides enough cases for most statistical analyses. Yet, the number is small enough for such analyses to be seriously affected by a potential selection bias. Second, the observed party group cohesion in EP RCVs is puzzlingly high (Attina, 1990; Raunio, 1997; Kreppel and Tsebelis, 1999; Hix *et al.*, 2007; Hix and Noury, 2009), given the groups' internal ideological heterogeneity, national fractionalization, and weak disciplining powers (Thiem, 2007, but see Yoshinaka *et al.*, 2010 and Yordanova, 2011).[3] This has raised suspicion that the findings concerning cohesion might be skewed upwards. If this were indeed the case, the bias in RCVs might seriously distort our perception of party group behaviour.

Carrubba *et al.* (2006) were the first to address this problem by analysing all the parliamentary votes taken in the EP between July 1999 and June 2000. They found that roll calls tend to be requested disproportionately more often on particular issues and by some party groups than others. Furthermore, while around 50% of all votes on non-legislative resolutions became RCVs, only 7% of all votes on

[3] Group cohesion in the EP has also been explained with persuasion and coordination within the groups (Ringe, 2010: 213) as well as the functional needs of MEPs to secure collective policy (Hix *et al.*, 2007: 88–91).

https://doi.org/10.1017/S175577391400023X Published online by Cambridge University Press

legislative proposals were held by roll call in their sample. This corroborates the proposition that RCVs do not constitute a random sample of all votes.

Based on their observations, Carrubba *et al.* (2006) argued that RCVs are subject to a bias that may affect findings. Høyland (2010) supported this latter conclusion by showing that results based on RCVs vary between legislative and non-legislative votes. Suggesting that the European party groups (EPGs) use roll calls strategically to enforce cohesion among their MEPs, Carrubba *et al.* (2006) infer that RCV analyses overestimate cohesion within party groups.

Arguing instead that party group leaders have nearly no means to sanction group members for divergent behaviour, Thiem (2006, 2007) holds that disciplining cannot be the main reason for RCV requests. She claims that the high level of group cohesion observed in RCVs is rather due to the fact that party leaders only request a roll call on votes on which they expect their own MEPs to display a cohesive position so as to send a signal to extra-parliamentary audiences.

In a subsequent study, Carrubba *et al.* (2008) developed a game theoretical model of RCV selection, predicting that the observed cohesion in requested RCVs depends on the size and the heterogeneity of party groups as well as on the status quo. They later provided initial empirical tests to support their hypotheses (Carrubba *et al.*, 2009; Hug, 2012a). Unfortunately, it is difficult to deduce a global expectation about the direction of any bias in RCVs based on their model as it relates to very specific circumstances.

A recent paper by Hug (2012b) suggests that there are differences in voting behaviour between EP6 and EP7, which can be attributed to enforced party control resulting from the rule change concerning mandatory RCVs. Yet, this analysis does not offer specific conclusions concerning the nature of a potential bias in RCV data.

Finally, in a paper concurrent to the present study, Hix *et al.* (2014) argue against the existence of a bias in requested RCVs. Their study employs a research design and data similar to ours, whereby final votes and amendment votes are compared in EP6 and EP7. However, there are substantial differences in the analysis. While they rely on OLS regression, we use a fractional logit model, which – as we explain in more detail later on – is essential to account for the characteristics of the dependent variable and to obtain valid predictions (see also Stecker, 2011). We further perform our analysis on a matched data of EP6 and EP7 votes to balance the two samples and minimalize the risks of biased regression results and high model dependence. Finally, alongside the policy area and the RCV requesting party, we include further controls at vote, proposal, and party group level that may be affecting the relationship between party group cohesion on the one hand, and the parliamentary term and type of vote (final or amendment) on the other hand.

To summarize, most of the literature suggests that relying on RCV data in the EP leads to overestimating party group cohesion, with the exception of Hix *et al.* (2014) who argue that there is no such bias. To evaluate whether RCV data is indeed biased, however, we first need to define the type of bias concerned and to be clear about the question 'biased in relation to what?'.

A bias concerning which cohesion? Disentangling the different theories on RCVs

In this section, we examine in detail the claim that party group cohesion is overestimated in RCVs by studying the assumptions and implications of different theories in the literature. For all theories, we distinguish between two types of bias they might imply: (1) bias due to behavioural change imposed by the fact that the vote is taken by a roll call (e.g. if legislators vote with their party group in a roll call, while they would have voted differently had the vote not been recorded) and (2) bias due to the selection of particular types of votes for roll calls (e.g. those on which the requesting party wants to signal to external constituents its cohesive position).

A 'behavioural bias' exists if the voting behaviour of a legislator in a RCV differs from how the legislator would have voted if the vote had not been roll called. Thereby, the fact that a vote is recorded affects the cohesion of party groups in the vote. To assess such a behavioural bias, we would thus need to ask: 'What would have been the party group cohesion if the vote had not been taken by roll call?'. We will call this the '*a priori* cohesion'. The *a priori* cohesion would be over- or underestimated by looking at the observed cohesion in RCVs if legislators are, respectively, more or less likely to vote with their party groups if a vote is roll called than if it is not roll called. Alternatively, there is no behavioural bias.

A 'selection bias' exists if the votes that get selected as RCVs differ significantly from votes that do not get selected as RCVs in terms of voting behaviour. As a result, the sample of RCVs is not representative of the total population of votes – in this case not because legislators change their voting behaviour but because votes with special characteristics are selected. To assess such a selection bias, we need to ask: 'What would have been the average party group cohesion if all votes had been taken by roll call?' We call this the 'all-RCV cohesion', which refers to the average cohesion we would observe if roll calls took place on all votes, instead of being requested only on a sample. The all-RCV cohesion is observable if roll calls are mandatory as a rule. It would be over- or underestimated by looking at the observed cohesion in the sample of requested RCVs if party group cohesion is in general higher or lower than average in the votes on which party groups request roll call. If neither of these applies, there would be no selection bias.

It is important to know about the existence and size of any behavioural bias and/or any selection bias in RCVs as this allows us to infer the 'actual voting cohesion' (i.e. the average party cohesion in all votes, whether roll called or not) from the cohesion we observe in the sample of requested RCVs.[4] This constitutes a prerequisite to assess the theories concerning RCVs as they implicitly include various expectations concerning over- or even underestimation of cohesion. So far, however, theoretical

[4] More formally, we can define the actual voting cohesion as follows: actual voting cohesion = observed cohesion $\frac{\text{No. of RCVs}}{\text{No. of votes}}$ + (observed cohesion – behavioural bias – selection bias) $\frac{\text{No. of nonRCVs}}{\text{No. of votes}}$, where the behavioural bias and the selection bias can be positive, negative, or null. Thus, to estimate the actual voting cohesion from the cohesion observed in RCVs, we need to know the size and the direction of both biases.

https://doi.org/10.1017/S175577391400023X Published online by Cambridge University Press

accounts of RVCs have not been explicit about the kind of bias they imply. Hence, rather than simply providing a summary of the previous literature, below we try to disentangle the different theories by focusing on their key assumptions and implications. Yet, it should be noted that we do not regard the theories as mutually exclusive. Combinations of the effects they predict are possible (and even likely) and will be discussed later on.

Disciplining

We start with examining the implications of the disciplining theory (Saalfeld, 1995; Carrubba *et al.*, 2006). The underlying assumption behind this theory is that party groups call RCVs to discipline their members and ensure voting cohesion. Without cohesiveness, party groups cannot build the needed parliamentary majorities to achieve their policy goals (Carey, 2007). Recorded votes allow group leaders to monitor the voting behaviour of group members. They can then reward loyal members and punish disloyal ones with a set of 'carrots and sticks' (Cox and McCubbins, 2007). To avoid punishment, members should refrain from voting against their party group's line in RCVs.

Hence, the disciplining theory is based on the premise that legislators are likely to change their voting behaviour if a vote is held by a roll call, in which case they would vote with group leaders. This suggests that the observed cohesion is a function of the *a priori* cohesion before disciplining and a positive disciplining effect of the vote being taken by a roll call. If disciplining takes place, the observed cohesion is higher than the *a priori* cohesion. In other words, analyses based on RCVs would overestimate the cohesion party groups would have had if these votes had not been recorded.

Following the theory, legislators may get disciplined in both mandatory and requested RCVs simply because their vote choices are visible to their party group leaders and we have no reason to expect that group cohesion will, on average, be lower or higher in mandatory than in requested RCVs. Thus, we should observe no change in group cohesion if all votes become RCVs compared with a situation where only some votes are requested for a roll call. As a result, while the disciplining theory implies overestimation of the *a priori* cohesion, it does not imply any bias concerning the all-RCV cohesion.

However, one may argue that not all parliamentarians are equally disciplined in a roll call. If they knew that it was their own party group, which requested the RCV, they can perceive that the vote is important for the group leaders, that they are closely watched and likely to be punished if they go against the group line. Thus, we can expect an enhanced disciplining effect for the members of the requesting group(s) because they know that their group leaders are trying to discipline them. In other words, this 'refined disciplining theory', as we term it, additionally stipulates that if a vote is held by a roll call, the members of a requesting party group are more likely to vote with their group than the members of the other party groups. This suggests

there is not only a disciplining effect of the roll call affecting all groups, but also an additional disciplining effect for the group that requests the roll call.

The refined disciplining theory, thus, implies that group cohesion is on average higher in requested than in mandatory RCVs. This is because of the additional positive disciplining effect for the members of the requesting group, which cannot be induced in mandatory RCVs. If RCVs were mandatory on all votes, we would thus observe a lower average level of cohesion than we do in the sample of requested RCVs. Consequently, according to the refined disciplining theory, cohesion in requested RCVs overestimates the *a priori* cohesion as well as the all-RCV cohesion.

Compliance with external principals

It is also conceivable that the visibility of the vote choice affects the behaviour of legislators in the opposite direction than implied by the disciplining theory, leading to the observed cohesion being lower than the *a priori* cohesion. We call this the 'external principals effect'. The underlying logic is that when the voting behaviour of parliamentarians is visible, party groups are not their only principals who can observe it. RCVs also enable the general public to monitor their representatives. Hence, legislators may choose to comply with the preferences of their constituents rather than the positions of their party groups if the two clash (Carey, 2009, 88f).

The external principals conjecture implies that the level of observed cohesion in RCVs is lower than the *a priori* cohesion (had there been no roll call). This is because representatives are expected to vote more often against their party groups in recorded votes due to the negative external principals effect. Hence, relying on RCV data will lead to underestimating the *a priori* cohesion. It should not lead to a bias in the all-RCV cohesion, though, because the external principals effect affects requested and mandatory RCVs in the same way.

One might argue, however, that in requested RCVs only members of the non-requesting party groups will be tempted to comply with external principals, while the observed cohesion of the requesting group as compared with its *a priori* cohesion does not decrease. Else, the latter group has little incentive to request the RCV. This 'refined external principals effect' would still, on average, lead to underestimating the *a priori* cohesion of party groups. However, it leads to a different expectation about how cohesion in a sample of requested RCVs compares with cohesion in the population of mandatory RCVs (i.e. cohesion when all votes are roll called by rule). Specifically, the latter is expected to be lower because there are no requesting party groups that would be left unaffected by the external principals effect. Hence, according to the refined external principals conjecture, the observed cohesion in requested RCVs overestimates the all-RCV cohesion.

Signalling a position

The signalling theory (cf. Thiem, 2006) provides the main alternative explanation of why party groups request RCVs besides disciplining. In this perspective, the

emphasis lies on the rewards legislators can reap from position-taking. While they cannot always be in the legislative majority, they can use voting as a tool to signal their position to external actors, whom they may depend on for their re-election (Mayhew, 1974). Therefore, it may be rational for a party group to request a RCV to signal a strong and cohesive position on an issue even if it loses the vote (Ainsley and Maxwell, 2012). Additionally, signalling can also be used to embarrass another group by exposing its unpopular position to the public.

Contrary to the disciplining theory and the external principals conjecture, the signalling theory does not assume that parliamentarians change their voting behaviour if their vote choice is recorded (Thiem, 2006). Therefore, it does not imply a difference between the *a priori* cohesion of a group on a given item and its observed cohesion if there were a RCV on this item.

Yet, signalling still introduces a bias in the analysis of group cohesion. Since a party group is only able to send a clear signal to its voters if it is not internally divided, it will call a RCV if it expects its own MEPs to vote cohesively. Likewise, a group can only be embarrassed for its position, if it has a clear position (i.e. if it is cohesive).[5]

As a consequence, because for each RCV at least one party group must have had a highly cohesive *a priori* position so as to request the vote (or to be embarrassed), on average, group cohesion will be higher in requested RCVs than in non-RCVs. Following this logic, party group cohesion is expected to be on average higher in the sample of requested RCVs than in the population of mandatory RCVs. Thus, according to the signalling theory, looking only at requested RCVs leads to overestimating the all-RCV group cohesion.

Exposing incohesive voting

Groups might also choose to request a roll call to expose the incohesive voting of other party groups. Internal party divisions send a negative signal to the voters that the party is ineffective, which affects the electoral fortunes of its members and can lead to a lower vote share for the party in the next elections (Cox and McCubbins, 2007). This may be a rationale for party groups to request roll calls on votes that are divisive for other groups in order to signal to voters the inability of these groups to put their act together.

Exposing incohesiveness is similar to the signalling theory in assuming that the *a priori* cohesion of party groups is simply revealed and no behavioural change occurs. In contrast to the signalling theory, though, exposing incohesion implies that cohesion in the sample of requested RCVs is lower than cohesion in the population of mandatory RCVs. This is because those votes are requested to be roll called on which at least one party group – the one to be embarrassed – has

[5] The possibility that a party group may request a RCV to expose another group's incohesiveness is considered in the next section.

a low cohesion. Hence, relying on the sample of requested RCVs results in an underestimation of the all-RCV cohesion.

RCV requests on contentious issues

There is yet another option that might lead to a bias concerning cohesion in requested RCVs. RCV requests may be more likely on contentious issues. This option has so far been considered in connection with disciplining effects (Carrubba *et al.*, 2008), as disciplining of party members is only necessary if there has been some internal conflict to begin with (cf. Hug, 2009). However, RCV requests on more contentious issues might also be a by-product of the other theories outlined above. For instance, signalling a position does only make sense if a party group holds a distinct position to other groups, implying that there is a conflict between groups. This in turn might very well result in conflict within groups. Especially in the EP where ideological distances within EPGs are often larger than between EPGs, it is likely that conflict lines do not run exactly between groups but rather through groups. Furthermore, for the external principals effect to lead representatives to vote against their party, there needs to be a clash between the external principals and the party. And finally, exposing incohesive voting behaviour of other party groups implies that there has been some dispute within these groups. In short, requested RCVs are necessarily associated with some kind of *a priori* conflict. For issues that are not disputed at all (i.e. technical matters where everyone is in favour), there is not much reason why any party group should bother to record the outcome. Hence, in this final conjecture, we hold that RCVs are requested on contentious issues.

The conjecture itself contains no implications concerning a behavioural bias, that is, whether legislators change their voting behaviour due to the fact that a vote is roll called. This depends on whether the selection of contentious votes occurs in combination with disciplining or the external principals effect. However, it does imply a selection bias. Specifically, the all-RCV cohesion will be underestimated in requested RCVs because the latter take place on more contentious issues where the cohesion of all party groups tends to be, on average, lower. In other words, since votes in which there is high disagreement are preselected for a roll call, we can expect the observed cohesion on these votes to be lower than in the votes that do not become RCVs (in which nearly all legislators are either in favour or against and, therefore, no group is interested in their voting behaviour). This conjecture therefore predicts lower cohesion in the sample of requested than in the population of mandatory RCVs, as the latter includes also votes that are not contentious and would not have been requested for roll call.

Summary of the different theories and their implications

To sum up, the theories on RCV requests have quite different implications concerning the potential bias in party cohesion in requested RCVs, especially if we

https://doi.org/10.1017/S175577391400023X Published online by Cambridge University Press

Table 1. Theories on RCV requests and the bias in party cohesion they imply

	Type of bias		Bias concerning	
Theory	Behavioural	Selection	*A priori* cohesion	All-RCV cohesion
Disciplining	✓	–	Overestimation	None
Refined disciplining	✓	✓	Overestimation	Overestimation
External principals	✓	–	Underestimation	None
Refined external principals	✓	✓	Underestimation	Overestimation
Signalling	–	✓	None	Overestimation
Exposing incohesiveness	–	✓	None	Underestimation
Contentious issues	–	✓	None	Underestimation

RCV = roll call vote.

distinguish between the types of cohesion affected. Table 1 presents an overview of the theories and the bias each of them implies concerning the *a priori* and the all-RCV cohesion.

Based on this conceptualization, in the remainder of the paper we focus on evaluating the potential selection bias in requested RCVs and the question whether the all-RCV cohesion is overestimated, as for example claimed by Thiem (2006, 2007). Of course, knowing about any behavioural bias is also necessary in order to be able to conclude whether the cohesion observed in requested RCVs constitutes an unbiased estimate of the actual voting cohesion. Unfortunately, while the change in the EP's Rules of Procedure enables us to test whether there is a selection bias, it does not enable us to examine the potential behavioural bias. Nevertheless, establishing whether there is a bias concerning the all-RCV cohesion in requested RCVs constitutes an important first step towards addressing the larger question of whether and, if so, how inferences regarding the actual voting cohesion are biased. Moreover, the question whether the all-RCV cohesion is over- or underestimated is interesting in itself. Due to the increased availability of electronic voting devices in parliaments and efforts to increase transparency, we can expect other parliaments to follow the example of the EP and make RCVs mandatory on more votes (Carey, 2009).

Hypotheses concerning the potential selection bias

In the last section, we explained how the different theories lead us to different expectations concerning a potential bias in the observed cohesion *vis-à-vis* the all-RCV cohesion. If only disciplining in its basic form took place there would be no bias, that is, all-RCV cohesion would be neither over- nor underestimated by the observed cohesion. The same holds for the basic form of the external principals effect. However, if only refined disciplining or only the refined external principals effect took place, the all-RCV cohesion would be overestimated. Contrary, if only the contentious votes conjecture or only exposing incohesiveness were taking place,

https://doi.org/10.1017/S175577391400023X Published online by Cambridge University Press

the all-RCV cohesion would be underestimated. Yet, it is quite unlikely that 'only' one of the theories applies. For example, as mentioned above, a contentious issue might be selected for a RCV due to disciplining reasons. Or, some RCVs might be called by a party group to signal its own cohesive position, while others might be called to expose the incohesive position of another group. Hence, the effects on cohesion predicted by the different theories may occur simultaneously or even cancel each other out (in the aggregate). Therefore, we have to take into account different combinations of these effects to assess whether there is a general selection bias in requested RCVs, and if so, whether this means that the all-RCV cohesion is over- or underestimated.

First, it is possible that the all-RCV cohesion of party groups is neither over- nor underestimated by their average observed cohesion in requested RCVs, that is, overall there may be no selection bias (H1). Considering all potential combinations of the different theories, this would happen if one of the following options applies: (a) none of the effects of the theories outlined in the last section is significant; (b) the theories implying only a behavioural bias but no selection bias hold (i.e. only disciplining takes place, only the external principals effect takes place, or only disciplining and the external principals effect take place); or (c) the possible effects on the all-RCV cohesion predicted by the remaining theories, all of which imply some kind of a selection bias, cancel each other out.

Second, the all-RCV cohesion may in general be overestimated (H2). This would happen under the condition that the combined effects of the theories predicting overestimation are stronger than the combined effects of the theories predicting underestimation, that is, if the refined disciplining, signalling, and refined external principals effects together outweigh any effects associated with a potential selection of contentious issues and RCV requests aimed at exposing incohesion.

Finally, the all-RCV cohesion may in general be underestimated (H3). This would happen under the condition that the combined effects of the theories predicting underestimation are stronger than the combined effects of the theories predicting overestimation, that is, if the contentious issues and exposing incohesion effects together outweigh any effects associated with refined disciplining, signalling, and the refined external principals theory.

In this paper, owing to a recent change in the EP Rules of Procedure concerning RCVs, we are able to test whether on average the all-RCV cohesion in the EP is over- or underestimated by the observed cohesion in requested RCVs or whether there is no such bias. Traditionally, roll calls have been taken if requested in advance by at least one party group or by a certain number of individual MEPs [40 in EP7, Rule 167.1(1); and 37 in EP6, ex Rule 160(1)]. However, since a revision of the EP's rules in 2009, all final legislative votes are taken by roll call (European Parliament, 2009, Rule 166).[6] The rule change enables us to observe voting behaviour of MEPs in the

[6] This rule change was brought about as a result of a tactical manoeuvre of the Independence/Democracy (ID) group. In January 2008, the ID group started to request roll calls, first, on all plenary votes and then, from

https://doi.org/10.1017/S175577391400023X Published online by Cambridge University Press

population of final legislative votes in EP7, while in EP6 we could only observe voting behaviour on the sample of requested RCVs. This presents a unique opportunity to compare the all-RCV cohesion (in the now mandatory final legislative votes) with the cohesion observed in requested RCVs and to estimate the direction and the size of the selection bias in requested RCVs.

However, simply comparing cohesion on final votes before and after the rule change would require the strong and rather unrealistic assumption that exogenous events such as the EP elections in 2009 (affecting party group composition) or the socialization of MEPs from the 12 new member states in the parliamentary work during the course of the EP6 (leading to potentially more harmonious party groups)[7] had no impact on group cohesion. Therefore, we control for any such systematic inter-parliamentary differences by considering also cohesion on amendment votes and making a relative rather than an absolute comparison of the change in group cohesion between the two legislative terms.[8] Thereby, we make the assumption that both final and amendment votes were equally affected by general inter-parliamentary developments so that any remaining difference in the extent to which cohesion in the two types of vote changed between EP6 and EP7 can be attributed to the rule revision. This implies that without the rule revision the change in the selection bias would have been the same for final and amendment votes so that we would have observed the same increase or decrease in cohesion for final and amendment votes. Furthermore, we assume that the rule change concerning RCVs

February onwards, on nearly all final votes. While the ID group argued that this was to increase legislative transparency, the other groups suspected that it was to disturb the legislative process because voting by roll call takes longer than voting by show of hands, thereby prolonging plenary sessions. Eventually, a compromise was found and the EP decided to change its Rules of Procedure so that all final legislative votes should be held by roll call from 2009 onwards, thereby increasing transparency without undermining efficiency too much (information based on interviews with two PSE MEPs in November 2009).

[7] Indeed, one may argue that EP6 is an outlier in terms of party disharmony due to the numerous non-socialized MEPs from the new member states. Yet, group cohesion was still higher in EP6 than in EP5, following the general trend of ever growing group cohesion between parliamentary terms (Hix and Noury, 2009).

[8] Still, there may be differences in the characteristics of the votes in the EP6 and EP7 samples, for instance due to the entry into force of the Lisbon Treaty. To address this, matching techniques are used in the analysis section. Note, however, that even with the matching we are not able to account for every potentially relevant difference between EP6 and EP7. As a partial remedy, in Appendix E we conduct an additional robustness test that partially recreates the conditions in our main analysis. Specifically, we compare group cohesion within EP6 before and after January 2008 when the ID group requested RCVs on almost all final legislative votes, nearly allowing us to observe the all RCV cohesion for such votes (see Footnote 5 above). This comparison is also not perfect because party group composition has changed between the two periods due to the 2008 enlargement round, the first years of a legislative period might differ from the last year concerning the legislative agenda and we do not observe voting behaviour on all final legislative votes but only on around 90%. Nevertheless, the data provides us with additional evidence concerning a potential bias in RCVs, which corroborates our findings from the main analysis below. Alternatively, we could have estimated a model predicting which final votes in EP7 would have been requested for roll call under the old rules and compare the sample of hypothetically requested final RCVs to the actual population of final RCVs. While we started out with this option, we found that the predictions of which votes would have been roll called were subject to a very high level of uncertainty.

on final votes did not affect the logic behind RCV requests on amendment votes. Note, however, that we do not assume cohesion or the selection bias to be the same on amendment and final votes within a parliamentary term. This implies that our findings will relate to final legislative votes in the EP and not necessarily to amendment votes. Generalization beyond final legislative votes or beyond the EP is only possible if a similar logic behind RCV requests applies to other types of votes and in other legislative assemblies. However, this yet needs to be shown.

Following this argumentation, if there was no selection bias on final votes we should observe no change in group cohesion on final legislative votes relative to cohesion on amendment legislative votes between EP6 (when RCVs were requested by party groups) and EP7 (when the EP rule change made roll calls mandatory on final votes). While the absolute level of group cohesion might have changed between the two terms as discussed above, this should have affected amendment and final votes to the same extent. Conversely, if the all-RCV cohesion was overestimated by the observed cohesion in the sample of requested RCVs, now that all final legislative votes are recorded, the observed group cohesion on final votes should decrease in relation to cohesion in amendment votes. Finally, if the all-RCV cohesion was underestimated, group cohesion on final votes should increase relative to cohesion on amendment votes.

We can now formulate our hypotheses as follows:

H1 (no bias in the all-RCV cohesion): On average, after the rule change the observed party group cohesion on final legislative votes is expected to have *neither increased*, *nor decreased* relative to cohesion on amendment legislative votes.

H2 (overestimated all-RCV cohesion): On average, after the rule change the observed party group cohesion on final legislative votes is expected to have *decreased* relative to cohesion on amendment legislative votes.

H3 (underestimated all-RCV cohesion): On average, after the rule change the observed party group cohesion on final legislative votes is expected to have *increased* relative to cohesion on amendment legislative votes.

Analysis

To test our hypotheses, we have combined data generously provided by Hix *et al.* (2007) with additional data we collected from the EP's Legislative Observatory and the official plenary minutes. The data covers all the legislative RCVs that took place in the first 1.5 years of EP6 and EP7 (until December 2010 inclusive) – 701 votes in total.[9] An observation in our data set corresponds to a party group in a given

[9] Data on the first 1.5 years is used to increase the comparability of the samples in the two legislative terms.

https://doi.org/10.1017/S175577391400023X Published online by Cambridge University Press

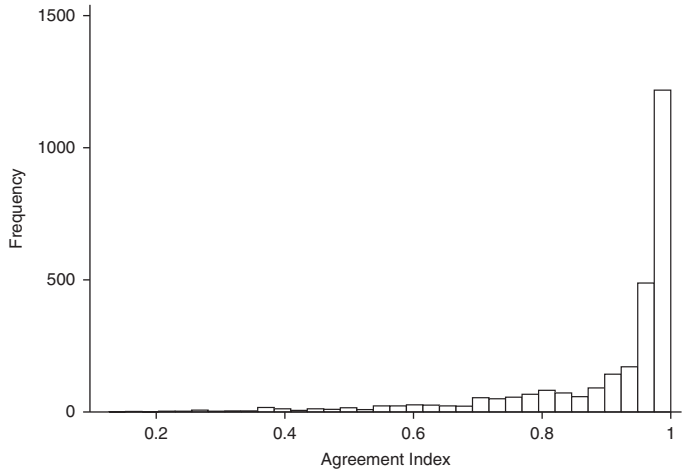

Figure 1 Distribution of the dependent variable.

legislative RCV. We only examine the cohesion of the four biggest party groups that have existed during both legislative terms, namely the European People's Party (PPE; which was called PPE-DE before 2009), the Socialist Group (PSE; now called S&D), the Liberal Group (ALDE), and the European Free Alliance Greens (Verts/ALE). Compared with others, these four EP groups remained relatively consistent between EP6 and EP7. Yet, we still need to account for changes in group composition due to the accession of Bulgaria and Romania in 2007 and the 2009 elections when, for instance, the British conservatives left the PPE to form their own group. Furthermore, some individual MEPs changed group affiliation within the legislative periods. These changes are traced using additional dynamic data on group membership (Høyland *et al.*, 2009).

To measure our dependent variable, namely party group cohesion, we use the Agreement Index as defined by Hix *et al.* (2007: 91).[10] The advantage of this measure is that it takes into account abstentions. It assumes values between 0 and 1, where 0 signifies complete discord in the group and 1 signifies that all group members voted the same way (see Figure 1 for the distribution of the Agreement Index in our data set).

$$\mathrm{AI}_{ij} = \frac{\max(\mathrm{Yes}_{ij}, \mathrm{No}_{ij}, \mathrm{Abst}_{ij}) - \frac{1}{2}\left[(\mathrm{Yes}_{ij} + \mathrm{No}_{ij} + \mathrm{Abst}_{ij}) - \max(\mathrm{Yes}_{ij}, \mathrm{No}_{ij}, \mathrm{Abst}_{ij})\right]}{\mathrm{Yes}_{ij} + \mathrm{No}_{ij} + \mathrm{Abst}_{ij}}$$

[10] For robustness check, we also performed our analysis with an alternative measure of group cohesion ($\mathrm{Cohesion}_{ij} = \frac{\max(\mathrm{Yes}_{ij}, \mathrm{No}_{ij})}{\mathrm{Yes}_{ij} + \mathrm{No}_{ij}}$, for each group i on each item j, see also Kreppel and Tsebelis, 1999), which led to almost identical results (not reported).

https://doi.org/10.1017/S175577391400023X Published online by Cambridge University Press

Table 2. Average cohesion on amendment and final legislative votes in the first 1.5 years of the 6th and the 7th EP

	Amendment			Final		
EPG	EP6	EP7	Difference	EP6	EP7	Difference
PPE-DE	0.842	0.873	0.031	0.885	0.978	0.093
PSE/S&D	0.874	0.918	0.043	0.948	0.982	0.033
ALDE	0.855	0.850	−0.005	0.921	0.975	0.054
Verts/ALE	0.913	0.964	0.051	0.906	0.969	0.063
All 4 EPGs	0.871	0.901	0.030	0.915	0.976	0.061
Number of votes	342	163		63	133	

EP = European Parliament; EPG = European party groups.

One note of caution is necessary. As party group cohesion is measured as a fraction bound between 0 and 1, it may not be possible to observe the same increase in cohesion on both amendments and final votes between the parliamentary terms, especially when assuming values close to 1. For example, if a group's average cohesion on amendments rises from 0.94 in EP6 to 0.99 in EP7, and the cohesion on final votes (which is generally higher than on amendments) was already 0.97 in EP6, it is virtually impossible to observe the same (0.05) or higher increase in cohesion on final votes as on amendments. In this case, our approach to use the difference between cohesion on final and amendment votes in the descriptive statistics we present below makes it easier to find evidence for Hypothesis 2 than for Hypothesis 1, and most difficult to find evidence for Hypothesis 3. Yet, we actually find support for Hypothesis 3.

Specifically, Table 2 shows the average observed cohesion of the four biggest party groups by type of vote (final or amendment) and legislative term (EP6 or EP7) for the first 1.5 years of each term. As can be seen in the last row, the average Agreement Index across the four party groups on both amendments and final legislative votes has increased between EP6 and EP7. The increase, however, is higher for final votes than for amendment votes, a finding which is even more pronounced if we hold the proposal constant (see Appendix A). This provides evidence in support of Hypothesis 3, stipulating that the all-RCV cohesion is underestimated, even though, as noted above, the nature of the dependent variable renders it more difficult to accept this hypothesis. With the new rule of mandatory RCV on all final legislative votes, cohesion has increased on these votes, and this is not due to a general shift in group cohesion between EP6 and EP7 as we do not observe the same increase in cohesion on amendment legislative votes.

We now turn to examining the statistical significance of this effect in a regression model where we also control for a number of potentially confounding factors. We rely on fractional logistic regression, which was designed for dependent variables that can only assume values in the interval [0,1] (see Papke and Wooldridge, 1996,

https://doi.org/10.1017/S175577391400023X Published online by Cambridge University Press

for a detailed description and specification of the functional form). Relying on no strong distributional assumptions, this method is more suitable for modelling our dependent variable (Agreement Index) – bound between 0 and 1, and further clustering at 1 – than simple OLS regression whose assumption of homoscedasticity would be violated. Also, beta-regression is inappropriate as while it can account for skewness of the dependent variable it assumes the dependent variable to follow a beta-distribution, meaning that it could only take values in the interval (0,1), excluding the values 0 and 1 (Papke and Wooldridge, 1996). We report robust standard errors, which are required to estimate the fractional logistic regression. Clustering the errors by proposal and vote, instead, produces the same results (not reported).

In line with our hypotheses, we aim at testing whether there is a difference in group cohesion on final relative to amendment legislative votes between EP6 (when RCVs were called by EPGs) and EP7 (when roll calls on final legislative votes became mandatory). Thus, we are mainly interested in the effect of the interaction between the variable EP7 [indicating whether the vote took place in EP7 (1) or in EP6 (0)] and the variable Final vote. An insignificant coefficient of the interaction term would lend support to Hypothesis 1, a significantly negative coefficient would support Hypothesis 2, while a significantly positive coefficient would support Hypothesis 3.

One may argue that votes in EP6 and EP7 are not comparable. For instance, the powers of the EP increased further with the entry into force of the Lisbon Treaty in 2009 – more policy areas fall under the co-decision (now ordinary) legislative procedure. Furthermore, due to the rule change there are, of course, many more final legislative RCVs in EP7 than there were in EP6, causing an imbalance in the data. These differences between the two legislative terms may introduce bias in the regression results and high model dependence (Ho *et al.*, 2006; Iacus *et al.*, 2011). To balance the samples of RCVs from EP6 and EP7, we opted for exact matching based on three variables accounting for the vote characteristics – Final vote, Procedure, and Committee (i.e. policy area, which we use as a crude proxy for characteristics such as status quo and proposal location) – with EP7 as the treatment and EP6 as the control group of cases. We used the CEM package in R developed by Iacus *et al.* (2009). As part of this procedure, first, all the cases were divided into strata, with cases within a stratum having identical values for all the three variables. Second, based on the relative number of observations from the two groups, weights were calculated such that observations from EP6 in a given stratum were weighted up if they were fewer than the number of observations from EP7 and down if they were more (in case where both groups had equal number of observations, all received a weight of 1). Where only one group (EP6 or EP7) was represented in a stratum, all observations in this stratum received zero weights, leading to the exclusion from the analysis of 81 amendment and five final votes from EP6 as well as 40 amendments and 45 final votes from EP7 (24% of our total number of votes). We use the weights in the subsequent analysis so that final (as well as amendment) votes or those falling under a given procedure or policy

area are not over- or underrepresented in either EP6 or EP7. Hence, after weighting the observations, the sample is perfectly balanced with respect to the three variables. Replicating the descriptive statistics from Table 2 on the matched data in Appendix B confirms that there was a higher increase in cohesion on final than on amendment votes between EP6 and EP7.

Additionally, in the analysis we control for several variables, which might have an impact on the effects of our main independent variables (see Appendix C for the descriptive statistics). Based on some of the theoretical approaches (refined disciplining, signalling, and refined external principals), one could expect that if a party group were the 'RCV requester', it would be more cohesive. 'RCV requests on proposal', measuring the number of roll calls the group requested on amendments to the respective proposal, is included as an indicator of the salience the group attaches to the proposal. An indicator of the general salience of the vote is the 'Participation' level, which gives the proportion of MEPs that took part in the vote out of the total number of MEPs serving during the respective parliamentary term. 'EPG had rapporteur' shows whether the rapporteur belonged to the party group. The fact that group composition has changed within and between parliamentary terms is, on the one hand, accounted for by controlling for group size ('EPG size') and internal fragmentation ('EPG fractionalization index').[11] On the other hand, we estimated models including dummy variables for the different groups (with ALDE serving as the reference category). Half-year controls are included, too, because group cohesion might change during the electoral cycle (cf. Lindstadt *et al.*, 2011). Finally, we accounted for differences between the legislative procedures of the proposals and their policy areas (using the responsible standing committees as a proxy).

Table 3 displays the results of the analysis on the matched data for different model specifications with Agreement Index as the dependent variable. The main results prove to be robust across model specifications (and also do not change when using the original non-matched data, see Appendix D, or the period from February 2008 to end of EP6 for comparison, see Appendix E). The primary finding is that the interaction between 'EP7' and 'Final vote' is significantly positive, showing that the increase in the level of group cohesion between EP6 and EP7 has been significantly higher for final than for amendment legislative votes. Specifically, our models predict a significantly higher increase of our dependent variable 'Agreement Index' for final votes (0.066 according to Model 5) than for amendment votes (0.040 according to Model 5) between EP6 and EP7. In addition, we see that the coefficient of 'Final vote' is not significant in Models 3–5, indicating that in EP6, there is no significant difference between the level of cohesion on amendment votes and the level of cohesion on final votes once we control for several other factors, while the

[11] $\textit{Fractionalization}(\mathrm{EPG}_i) = 1 - \left(\sum_k \frac{\mathrm{size}(\mathrm{NP}_k)}{\mathrm{size}(\mathrm{EPG}_i)} \right)^2$ for national parties k of group i (see also Hix *et al.*, 2007: 97). The fractionalization index assumes the value 0 if the party group consists of only one national party, while it approaches 1 if the group consists of many small national parties.

https://doi.org/10.1017/S175577391400023X Published online by Cambridge University Press

Table 3. Fractional logistic regression of party group Agreement Index in EP6 and EP7 using the matched data

	Model 1	Model 2	Model 3	Model 4	Model 5
EP7	0.581 (0.103)**	0.582 (0.103)**	0.625 (0.099)**	0.612 (0.097)**	0.611 (0.100)**
Final vote	0.512 (0.219)*	0.501 (0.221)*	0.443 (0.230)	0.148 (0.267)	−0.014 (0.243)
EP7 × Final vote	0.824 (0.270)**	0.862 (0.279)**	0.805 (0.284)**	0.692 (0.286)*	0.707 (0.267)**
RCV requester		0.137 (0.141)	−0.061 (0.155)	−0.074 (0.146)	−0.031 (0.134)
RCV requests on proposal			−0.001 (0.006)	−0.001 (0.005)	−0.002 (0.006)
Participation			−3.225 (0.846)**	−2.820 (0.905)**	−2.701 (0.949)**
EPG had rapporteur			−0.078 (0.228)	−0.068 (0.215)	−0.037 (0.156)
EPG fractionalization index			−16.228 (2.633)**		
EPG size			0.002 (0.001)*		
PPE-DE				0.375 (0.198)	0.366 (0.151)*
PSE				0.591 (0.176)**	0.597 (0.143)**
Verts/ALE				1.042 (0.161)**	1.053 (0.163)**
1st half year				0.655 (0.277)*	0.672 (0.294)*
2nd half year				−0.092 (0.103)	−0.088 (0.105)
Procedure dummies				Yes	Yes
Committee dummies					Yes
Constant	1.813 (0.064)**	1.785 (0.073)**	19.110** (2.390)	2.741 (0.991)**	2.767 (1.025)**
Observations	2120	2120	2120	2120	2120
Log-likelihood	−510.6	−510.4	−495.5	−488.3	−472.9

EP = European Parliament; EPG = European party groups; RCV = roll call vote.
Robust standard errors in parentheses; $^{**}P < 0.01$, $^{*}P < 0.05$.

additive effect of 'Final vote' and the interaction term is significantly positive in all models (e.g. in Model 5 the combined term is 0.707 – 0.014 = 0.693; $P = 0.000$), meaning that in EP7 voting cohesion on final votes is significantly higher than voting cohesion on amendment votes (Brambor *et al.*, 2005). Thus, the analysis supports Hypothesis 3 and contradicts the remaining hypotheses.

As for the control variables, we find no significant effect on party group cohesion of the number of RCV requests on a proposal by a party group but a significant negative effect of the participation level in a vote. The variable 'RCV requester' and the variable indicating whether the group had the rapporteur or not do not have significant effects, in contrast to the fractionalization index variable. The more fractionalized a group is, the less cohesive do its MEPs vote.

To sum up, our empirical analysis shows that cohesion on RCVs has increased between EP6 and EP7, and significantly more so on final votes than on amendment votes. The results, hence, support Hypothesis 3, suggesting that the all-RCV cohesion, that is, the cohesion we would observe if all votes became RCVs, is underestimated in requested RCVs. As we can reject Hypothesis 1 (predicting on average no bias concerning the all-RCV cohesion), we can further infer that there was a selection bias in roll calls on final legislative votes before the rule change. This also implies that (a) at least one of the theories predicting such a bias is of empirical relevance, (b) whether disciplining and the external principals effects occur or not, they cannot be the only effects taking place, and (c) the possible effects on all-RCV cohesion do not cancel each other out to produce unbiased results on average. In fact, the combined effect of the theories predicting underestimation is stronger than the combined effect of the theories predicting overestimation of the all-RCV cohesion. More specifically, we can conclude that the contentious issues and exposing incohesion effects together outweigh any effects associated with refined disciplining, signalling, and the refined external principals theory. Thus, more generally our findings lend support to the claim that roll calls are mainly requested on votes that divide party groups internally, be it, for example, to embarrass other groups or to discipline party members.

Conclusion

In this paper, we examine the claim that relying on requested RCVs may lead to overestimation of group cohesion (Carrubba *et al.*, 2006; Thiem, 2006). We argue that to evaluate this claim, we should first distinguish between two types of possible bias: a behavioural bias, implying that the observed cohesion of party groups in a RCV differs from their *a priori* cohesion (i.e. the cohesion they would have displayed had the vote not been roll called) and a selection bias, meaning that the observed cohesion in the sample of requested RCVs differs from the all-RCV cohesion (i.e. the cohesion we would observe if roll calls were mandatory on all votes). We focus on the selection bias, thereby studying whether the all-RCV cohesion is over- or rather underestimated when looking at the observed cohesion in requested RCVs.

To examine the potential selection bias in RCVs, we took advantage of the unique opportunity that a change in the parliamentary rules in 2009 offered by making roll call voting mandatory on all final legislative votes. The empirical analysis shows that after the rule change party group cohesion in the roll calls on final legislative votes (now taking place on the population of such votes) has increased in relation to cohesion in the roll calls on amendment legislative votes (still taking place only on a sample of such votes, whenever requested). This implies that requested RCVs on final legislative votes were indeed subject to a selection bias and suggests that relying on the cohesion in requested RCVs generally leads to underestimation of the all-RCV cohesion.

We cannot be sure that our results (and thus our conclusions) hold for other types of vote than final legislative votes. For example, it is possible that RCVs on amendment votes are called for different reasons than RCVs on final votes. As a consequence, RCVs on amendment votes might display a different sort of bias. Yet, if a similar logic were to apply to RCV requests on amendments, we expect to find an even higher increase in the observed group cohesion on amendment RCVs should they also become mandatory because cohesion on amendments is generally lower than on final votes (allowing more room for improvement). This would indicate an even more pronounced selection bias for non-final RCVs.

Our finding of a selection bias in requested RCVs allows us to draw first indirect conclusions concerning the actual voting cohesion, that is, the cohesion in all votes, whether roll called or not. Whether the actual voting cohesion is in fact over- or underestimated depends on the size and the direction of both the behavioural and the selection bias. Based on our results, we can say that due to a selection bias the actual voting cohesion of party groups is on average underestimated in requested RCVs, unless there is a behavioural bias leading to such a strong overestimation that it outweighs the selection effect, which points towards underestimation (i.e. a behavioural bias with a different sign and at least the same size as the selection bias).

Furthermore, we can also draw some conclusions concerning the different theories on RCV requests. While it is possible and even highly likely that various effects on cohesion in RCVs, caused by disciplining, signalling, or external principals, simultaneously take place, the only theories that account for our findings are the ones predicting the selection for roll call of issues that are contentious within party groups. We show that the effect of the selection of contentious issues (in general, as well as more specifically to expose incohesion of other groups) outweighs any potential further effects by the other theories concerning the selection bias. Thus, the potential positive effects on the observed cohesion in requested RCVs (as compared with cohesion if all votes were taken by roll call), as highlighted in the context of the refined disciplining theory, the refined external principal theory or the signalling theory, are weaker than the negative effect of selecting more contentious issues for roll calls.

Finally, the finding that the all-RCV cohesion is underestimated in final votes is interesting in itself. It shows that the high party group cohesion we observe in the EP is not due to a selection bias in RCVs. In fact, according to our analysis, the mean observed group cohesion in EP6 would have been even higher had all votes been

taken by RCV back then. This may be relevant for other parliaments, too. As electronic voting devices are becoming more common in legislatures, taking RCVs is greatly facilitated. Additionally, the pressure for more transparency might lead parliaments to follow the route of the EP and to make RCVs mandatory on final votes. Hence, if our results for the EP translate to other parliaments we might witness an increase in the observed voting cohesion over time. Thus, by presenting some expectations about the future development of legislative behaviour, we hope to not only have added to the scientific discussion on a potential bias in RCV data, but also to the general understanding of legislative institutions.

Acknowledgements

The authors are grateful to Simon Hix for providing them with the data on the RCVs in the first 1.5 years of the 7th EP before its official release and to Bartlomiej Grychtol for helping them with further automated data collection. Furthermore, they would like to thank the three anonymous referees, Gerald Schneider, Simon Hug, Bjørn Høyland, and the Salzburg 'Mittwochsrunde' for helpful comments on previous versions of the paper. An earlier version of this paper was presented at the general conference of the European Political Science Association, Berlin, 2012.

References

Ainsley, C. and L. Maxwell (2012), 'What can we learn from roll call votes? Strategic signalling incentives and the decision to call roll call votes', Working paper, Emory University.

Attina, F. (1990), 'The voting behaviour of the European Parliament members and the problem of the Europarties', *European Journal of Political Research* **18**(4): 557–579.

Brambor, T., W.R. Clark and M. Golder (2005), 'Understanding interaction models: improving empirical analyses', *Political Analysis* **14**(1): 63–82.

Carey, J.M. (2007), 'Competing principals, political institutions, and party unity in legislative voting', *American Journal of Political Science* **51**(1): 92–107.

Carey, J. (2009), *Legislative Voting and Accountability*, Cambridge: Cambridge University Press.

Carrubba, C., M. Gabel and S. Hug (2008), 'Legislative voting behavior, seen and unseen: a theory of roll-call vote selection', *Legislative Studies Quarterly* **33**(4): 543–572.

—— (2009), 'Voting at the surface: roll call votes in the European Parliament'. In 2nd Conference on the Political Economy of International Organizations, Geneva, Switzerland, January 29–31 .

Carrubba, C.J., M. Gabel, L. Murrah, R. Clough, E. Montgomery and S. Rebecca (2006), 'Off the record: unrecorded legislative votes, selection bias and roll-call vote analysis', *British Journal of Political Science* **36**: 691–704.

Cox, G.W. and M.D. McCubbins (2007), *Legislative Leviathan: Party Government in the House*, 2nd edn., New York: Cambridge University Press.

European Parliament (2009), Rules of Procedure of the European Parliament. Retrieved 12 November 2014 from http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//NONSGML+RULES-EP+20100705+0+DOC+PDF+V0//EN&language=EN.

Finke, D. and J. Thiem (2010), 'Two principals, two motives: the amendment process and roll call votes in the European Parliament'. In Proceedings of the Fifth Pan-European Conference on EU Politics, June 23–26, Porto.

Hix, S. and A. Noury (2009), 'After enlargement: voting patterns in the sixth European Parliament', *Legislative Studies Quarterly* **34**(2): 159–174.

——, A.G. Noury and G. Roland (2007), *Democratic Politics in the European Parliament*, Cambridge: Cambridge University Press.

——, A. Noury and G. Roland (2014), Is there a strategic selection bias in roll call votes in the European Parliament?', Manuscript. Retrieved 12 November 2014 from http://personal.lse.ac.uk/hix/Working_Papers/Hix-Noury-Roland_Selection_EP_22July2014.pdf.

Ho, D., K. Imai, G. King and E. Stuart (2006), 'Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference', *Political Analysis* **15**(3): 199–236.

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

——, I. Sircar and S. Hix (2009), 'Forum section: an automated database of the European Parliament', *European Union Politics* **10**(1): 143–152.

Hug, S. (2006), 'Selection effects in roll call votes', *CIS Working Paper No. 15, 2006*.

—— (2009), 'Selection effects in roll call votes', *British Journal of Political Science* **40**(1): 225–235.

—— (2012a), 'Roll call votes in the European Parliament'. In Proceedings of the Sixth Pan-European Conference on EU Politics, September 13–15, Tampere.

—— (2012b), 'The European Parliament after Lisbon (and before)'. In International Conference 'Beyond Lisbon Treaty', Institute of European and American Studies (IEAS), Academia Sinica, September 7–8, Taipei, Taiwan.

Iacus, S.M., G. King and G. Porro (2009), 'Cem: software for coarsened exact matching', *Journal of Statistical Software* **30**(9): 1–27.

—— (2011), 'Causal inference without balance checking: coarsened exact matching', *Political Analysis* **20**(617): 1–24.

Kreppel, A. and G. Tsebelis (1999), 'Coalition formation in the European Parliament', *Comparative Political Studies* **32**: 933–966.

Lindstadt, R., J.B. Slapin and R.J. Vander Wielen (2011), 'Balancing competing demands: position taking and election proximity in the European Parliament', *Legislative Studies Quarterly* **36**(1): 37–70.

Lynch, M.S. and A.J. Madonna (2013), 'Viva voce: implications from the disappearing voice vote', *Social Science Quarterly* **94**(2): 530–550.

Mayhew, D.R. (1974), *Congress: The Electoral Connection*, New Haven and London: Yale University Press.

Papke, L.E. and J.M. Wooldridge (1996), 'Econometric methods for fractional response variables with an application to 401(k) Plan participation rates', *Journal of Applied Econometrics* **11**(6): 619–632.

Raunio, T. (1997), *The European Perspective: Transnational Party Groups in the 1989–1994 European Parliament*, Aldershot: Ashgate.

Ringe, N. (2010), *Who Decides, and How? Preferences, Uncertainty, and Policy Choice in the European Parliament*, Oxford: Oxford University Press.

Saalfeld, T. (1995), 'On dogs and whips: recorded votes', in H. Doring (ed.), *Parliaments and Majority Rule in Western Europe*, New York: St. Martin's Press, pp. 528–565.

Stecker, C. (2010), 'Causes of roll-call votes supply: evidence from the German loander', *The Journal of Legislative Studies* **16**(4): 438–459.

—— (2011), 'Bedingungsfaktoren der fraktionsgeschlossenheit. eine vergle-ichende analyse der deutschen loanderparlamente', *PVS* **52**(3): 424–447.

Thiem, J. (2006), 'Explaining roll call vote request in the European Parliament', *Mannheimer Zentrum für Europäische Sozialforschung Working Paper No. 90*.

—— (2007), 'A supranational elite (in the making)? Agency relations in the European Parliament and the role of national parties'. In ECPR Joint Sessions, May 7–12, Helsinki.

Yordanova, N. (2011), 'The effect of inter-institutional rules on the division of power in the European Parliament: allocation of consultation versus codecision reports', *West European Politics* **34**(1): 97–121.

Yoshinaka, A., G. McElroy and S. Bowler (2010), 'The appointment of rapporteurs in the European Parliament', *Legislative Studies Quarterly* **35**(4): 457–486.

Appendix A

Evaluating whether the increase in cohesion between EP6 and EP7 is indeed higher on final than on amendment votes when holding the proposal constant requires a different set-up of the data. The figures in Table 2 currently present the average group cohesion on all final and amendment legislative votes, ignoring whether they occurred on the same proposals or not. This may render the comparison flawed if cohesion is higher on final votes whenever no RCVs were taken on amendments belonging to the same proposal. This may happen, for instance, because such a proposal is rather technical and uncontroversial or one on which the Parliament simply made no amendments. To address this problem, in the table below we compare, instead, the group cohesion on each amendment to group cohesion on the corresponding final vote of the respective proposal. Thus, we can examine the difference in group cohesion on final and amendment votes in EP6 and EP7 while controlling for the particularities of given legislative acts.

First, the table confirms that group cohesion has increased on both amendment and final votes between EP6 and EP7. Second, in fact cohesion on final votes appears to be, on average, lower than on the corresponding amendment votes in EP6, contrary to the conventional wisdom based on considering all amendment and final votes. The opposite holds true in EP7. Finally, the gap between group cohesion on final and amendment votes has widened in EP7 (see last row of the table). Thus, the average cohesion on final legislative votes has increased with the EP's rule change, and more so than we would expect due to the general shift in the cohesion level between EP6 and EP7 as shown by the lower cohesion increase on amendments between the two legislative terms.

Average cohesion on amendment and their corresponding final legislative votes in the first 1.5 years of the 6th and the 7th EP of the four biggest party groups

	EP6	EP7	Diff. EP7 – EP6
Amendment	0.853	0.914	0.061
Final	0.773	0.950	0.177
Diff. Fin. – Am.	– 0.080	0.036	0.116
Number of votes	202	77	

EP = European Parliament.

https://doi.org/10.1017/S175577391400023X Published online by Cambridge University Press

Appendix B

Weighted average cohesion (*after matching*) on amendment and final legislative votes in the first 1.5 years of EP6 and EP7

	Amendment			Final		
EPG	EP6	EP7	Difference	EP6	EP7	Difference
PPE-DE	0.848	0.919	0.072	0.891	0.982	0.091
PSE/S&D	0.866	0.915	0.049	0.947	0.979	0.032
ALDE	0.802	0.856	0.054	0.879	0.975	0.097
Verts/ALE	0.923	0.975	0.052	0.927	0.970	0.043
All 4 EPGs	0.860	0.916	0.057	0.911	0.977	0.066
No. of votes with non-0 weights	261	123		58	88	
Sum of vote weights	186	123		133	88	

EP = European Parliament; EPG = European party groups.

https://doi.org/10.1017/S175577391400023X Published online by Cambridge University Press

Appendix C

Descriptive statistics of the variables in the analysis

	Min	Max	Mean	*N*
EPG Agreement Index	0.127	1	0.902	2804
EP7	0	1	0.422	2804
Final vote	0	1	0.280	2804
RCV requester	0	1	0.161	2804
RCV requests on proposal	0	41	3.765	2804
Participation	0.104	0.93	0.852	2804
EPG had rapporteur	0	1	0.204	2804
EPG fractionalization index	0.859	0.94	0.912	2804
EPG size	42	268	148.556	2804
PPE-DE	0	1	0.25	2804
PSE	0	1	0.25	2804
Verts/ALE	0	1	0.25	2804
ALDE	0	1	0.25	2804
1st half year	0	1	0.118	2804
2nd half year	0	1	0.277	2804
Consent procedure	0	1	0.001	2804
Assent procedure	0	1	0.056	2804
Consultation procedure	0	1	0.184	2804
Codecision procedure	0	1	0.756	2804
Cooperation procedure	0	1	0.003	2804
Agriculture	0	1	0.039	2804
Budgets	0	1	0.014	2804
Civil liberties, justice and home affairs	0	1	0.087	2804
Constitutional affairs	0	1	0.007	2804
Culture and education	0	1	0.011	2804
Development	0	1	0.001	2804
Economic and monetary affairs	0	1	0.061	2804
Employment and social affairs	0	1	0.067	2804
Environment, public health and food safety	0	1	0.342	2804
Fisheries	0	1	0.054	2804
Foreign affairs	0	1	0.009	2804
Industry, research and energy	0	1	0.021	2804
Internal market and consumer protection	0	1	0.003	2804
International trade	0	1	0.017	2804
Legal affairs	0	1	0.037	2804
Regional development	0	1	0.073	2804
Transport and tourism	0	1	0.097	2804
Women's rights and gender equality	0	1	0.007	2804

EP = European Parliament; EPG = European party groups; RCV = roll call vote.

https://doi.org/10.1017/S175577391400023X Published online by Cambridge University Press

Appendix D

Fractional logistic regression of party group Agreement Index in EP6 and EP7 using the original data

	Model 1	Model 2	Model 3	Model 4	Model 5
EP7	0.298 (0.078)**	0.297 (0.078)**	0.339 (0.077)**	0.379 (0.079)**	0.504 (0.091)**
Final vote	0.468 (0.126)**	0.471 (0.126)**	0.385 (0.132)**	0.290 (0.144)*	0.355 (0.142)*
EP7 × Final vote	1.036 (0.190)**	1.021 (0.190)**	0.970 (0.194)**	0.937 (0.193)**	0.659 (0.197)**
RCV requester		−0.061 (0.084)	−0.239 (0.089)**	−0.224 (0.092)*	−0.232 (0.091)*
RCV requests on proposal			−0.000 (0.003)	0.001 (0.004)	0.001 (0.004)
Participation			−3.834 (0.637)**	−3.793 (0.674)**	−3.529 (0.678)**
EPG had rapporteur			0.215 (0.080)**	0.222 (0.082)**	0.158 (0.078)*
EPG fractionalization index			−10.853 (1.480)**		
EPG size			−0.000 (0.000)		
PPE-DE				−0.068 (0.090)	−0.048 (0.087)
PSE				0.302 (0.085)**	0.311 (0.085)**
Verts/ALE				0.751 (0.105)**	0.742 (0.104)**
1st half year				0.276 (0.150)	0.347 (0.164)*
2nd half year				−0.157 (0.070)*	−0.124 (0.080)
Procedure dummies				Yes	Yes
Committee dummies					Yes
Constant	1.912 (0.037)**	1.924 (0.040)**	15.159 (1.393)**	5.107 (0.618)**	3.876 (0.659)**
Observations	2804	2804	2804	2804	2804
Log-likelihood	−685.7	−685.6	−671.3	−668.8	−657.9

EP = European Parliament; EPG = European party groups; RCV = roll call vote.
Robust standard errors in parentheses; $^{**}P < 0.01$, $^{*}P < 0.05$.

Appendix E

Fractional logistic regression of party group Agreement Index in EP6 pre-2008 and post-January 2008 using the original data[a]

	Model 1	Model 2	Model 3	Model 4	Model 5
Post-2008	0.150 (0.034)**	0.150 (0.034)**	0.204 (0.034)**	0.169 (0.034)**	0.275 (0.036)**
Final vote	0.657 (0.069)**	0.657 (0.069)**	0.641 (0.072)**	0.713 (0.074)**	0.765 (0.075)**
Post-2008 × Final vote	0.294 (0.092)**	0.295 (0.093)**	0.296 (0.095)**	0.245 (0.096)*	0.122 (0.099)
RCV requester		0.005 (0.042)	−0.145 (0.045)**	−0.138 (0.046)**	−0.166 (0.045)**
RCV requests on proposal			0.003 (0.002)	0.001 (0.002)	0.002 (0.002)
Participation			−0.001 (0.000)**	−0.001 (0.000)**	−0.001 (0.000)**
EPG had rapporteur			0.042 (0.039)	0.049 (0.039)	0.049 (0.039)
EPG fractionalization index			−6.985 (0.732)**		
EPG size			0.001 (0.000)**		
PPE-DE				0.073 (0.044)	0.076 (0.043)
PSE				0.237 (0.045)**	0.240 (0.045)**
Verts/ALE				0.491 (0.052)**	0.497 (0.052)**
Procedure dummies				Yes	Yes
Committee dummies					Yes
Constant	1.802 (0.021)**	1.801 (0.023)**	8.688 (0.697)**	2.491 (0.222)**	1.788 (0.245)**
Observations	8300	8300	8300	8300	8300
Log-likelihood	−2274	−2274	−2260	−2257	−2234

EPG = European party groups; RCV = roll call vote.
Robust standard errors in parentheses; $^{**}P<0.01$, $^{*}P<0.05$.
[a]For the period February 2008 until the end of EP6, about 90% of all final legislative votes were roll called due to requests of the ID group. Hence, we come close to the all RCV cohesion in final votes for the post-January 2008 sample, similarly to the situation after the 2009 EP rule change. We thus use the pre-2008/post-January 2008 comparison as a robustness check of our findings.