DISCUSSION PAPER SERIES

No. 9418

INFERRING HAWKS AND DOVES FROM VOTING RECORDS

Sylvester C W Eijffinger, Ronald J Mahieu
and Louis Raes

INTERNATIONAL MACROECONOMICS

Centre for Economic Policy Research

www.cepr.org

Available online at: www.cepr.org/pubs/dps/DP9418.asp

ISSN 0265-8003

INFERRING HAWKS AND DOVES FROM VOTING RECORDS

Sylvester C W Eijffinger, CentER and EBC, Tilburg University and CEPR
Ronald J Mahieu, CentER and Netspar, Tilburg University
Louis Raes, CentER and EBC, Tilburg University

Discussion Paper No. 9418
April 2013

CEPR Discussion Paper No. 9418

April 2013

ABSTRACT

Inferring hawks and doves from voting records*

In this paper we estimate spatial voting models for the analysis of the voting record of the monetary policy committee of the Bank of England. We use a flexible Bayesian approach for estimating such models. A simple modification to the standard spatial model as well as a variety of model checks are proposed to deal with the specifics of the data available. We provide evidence that extreme policy preferences are to be found among the external members. We also consider the variation in policy preferences according to career backgrounds. The median voter preference is similar for different backgrounds, except for those with a background in the industry where the median voter is more hawkish. The heterogeneity in policy preferences is the largest among academics and those with a background in the industry. The range of policy preferences is much smaller among other groups, in particular among monetary policy committee members with central bank experience who exhibit the lowest heterogeneity in policy preferences.

JEL Classification: C11, E58 and E59
Keywords: central banking, ideal points and voting record

Sylvester C W Eijffinger
CentER and EBC
Tilburg University
PO Box 90153
5000 LE Tilburg
THE NETHERLANDS

Email: s.c.w.eijffinger@uvt.nl

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=120432

Ronald J Mahieu
CentER and EBC
Tilburg University
PO Box 90153
5000 LE Tilburg
THE NETHERLANDS

Email: r.j.mahieu@uvt.nl

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=121958

Louis Raes
CentER and EBC
Tilburg University
PO Box 90153
5000 LE Tilburg
THE NETHERLANDS

Email:
l.b.d.raes@tilburguniversity.edu

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=172719

*We would like to thank Bjørn Høyland, Wieland Müller, Christopher Spencer, Steven Verheyen, Nick Vivyan for discussion and advice, as well as seminar discussants at the Tilburg Economic Workshop. Comments are welcome.

Submitted 29 March 2013

1 Introduction

The past decade(s) there have been remarkable changes in the decision making process in central banks around the globe. Notably, many central banks have given the authority for the policy decisions to independent committees whereas most were previously run by individual governors.[1]

The institutional arrangements differ considerably across central banks. These institutional details concern the composition of such a committee, the way decision are reached, the transparency of procedures, etc. In this paper we focus on the composition of a monetary policy committee. We seek to understand the heterogeneity on policy preferences among the individual members. To do this we estimate a spatial voting model for the monetary policy committee of the Bank of England. Our spatial framework is characterized by a single dimension. This dimension captures an ideological continuum from being extremely dovish to being extremely hawkish. The estimation of such a spatial model yields ideal points or the revealed policy preferences of the monetary policy committee members, which can be depicted as points on the dove-hawk dimension. Armed with these individual preferences we are able to tackle a variety of questions posed in the literature on decision making in monetary policy committees.

Existing research on this topic uses mainly one of two approaches. At the one hand researchers have tried to estimate aggregate and individual interest rate rules. Besley, Meads, and Surico (2008) estimate reaction functions for the individual committee members and assess the extent to which these capture the heterogeneity in voting patterns. The authors group the committee members according to career background (e.g. academia vs non-academia) and according to their appointment within the committee (external or internal member). The parameters of the individuals are then compared across groups. The authors find that while there is substantial heterogeneity in voting patterns, the individual reactions functions are fairly homogenous with no significant differences between members according to the background characteristics considered. Other examples of this approach studying the Bank of England are Besley, Meads, and Surico (2008), Riboni and Ruge-Murcia (2008) and Harris and Spencer (2009).

The other predominant approach builds upon a regression framework where the dependent variable captures the votes casted by members. This dependent variable is then regressed on relevant meeting characteristics (variables capturing economic conditions) and voter characteristics (backgrounds of the individual voters). As an example, Harris, Levine, and Spencer (2011) examine the frequency and type of dissenting votes in the monetary policy committee at the Bank of England. While they find strong heterogeneity in voting patterns, they find only a weak role for career experience in determining the decision to dissent. These findings stand in contrast to the large literature studying votes at the FOMC suggesting that career backgrounds do matter as well as political influence through appointment, see the discussion in Harris, Levine, and Spencer (2011) and Chappell, Havrilesky, and McGregor (1993) for

[1] Blinder (2004) refers to this as the *quiet revolution*. Pollard (2004) surveyed 88 central banks and found that already 79 of the surveyed central banks take decisions by committee.

early evidence on the appointment channel.

We use an alternative way to investigate voting behavior of monetary policy committee members. The approach builds on methodological advancements in other disciplines where researchers have investigated voting behavior of legislators and judges. The methodology yields a convenient to use and intuitive measure of policy preferences. Moreover, the (Bayesian) method yields a joint posterior distribution of all the parameters. This makes our approach much more flexible then the approaches discussed above. We demonstrate this flexibility throughout the paper by investigating and making inference about derived quantities in a way that would be nearly infeasible with the approaches discussed above.

This paper has two goals. First, we introduce the methodology for estimating ideal points of voting members in a spatial voting framework. We argue that such an approach is best suited for investigating voting records. While economists have greatly contributed to the development of the theoretical underpinnings of the spatial voting model since the pioneering work by Black (1948), empirical implementations of the spatial voting model remain scant within economics. An early contribution is the paper by Heckman and Snyder (1997) which is an influential methodological contribution with an application to legislative data but by now outdated for the purpose of estimating ideal points, see also Clinton, Jackman, and Rivers (2004). Henry and Mourifié (2013) test the spatial voting model in the context of US national elections. The few empirical papers we are aware of, consider nearly always applications in politics and the analysis of judicial votes.[2]

In this paper we focus on voting data of monetary policy committees. There is one other paper we are aware of that empirically investigates voting at a monetary policy committee in a spatial voting framework, see Hix, Hoyland, and Vivyan (2010). This research by political scientists was published in a political science journal and consequently went largely unnoticed by the economic community. Their research can be seen as a direct predecessor to this paper. Their analysis looked at the voting process through the lens of political economy. We add to their work by connecting the ideal point approach with the existing research in (monetary) economics, adapt the methodology and widen the scope of the approach. After presenting a concise introduction to this approach, we show how this methodology can (and should) be adapted to the data available in the study of decisions by monetary policy committees. We also present a variety of tools for model checking which we feel are important when employing this type of models.

Second, we use this approach to analyze the voting records of the Bank of England. The Bank of England

[2]Recently, the empirical analysis of voting data regained attention. Authors have proposed different equilibrium models of decision making. The idea is that some of these traditional areas of application (such as voting in supreme courts) have peculiar features (e.g. strategic voting) which require a modified methodology. These papers typically do not build upon a spatial voting model. As an example, Iaryczower and Shum (2012) examine voting behavior in the US supreme court and build an equilibrium model of decision making to quantify the value of information. Iaryczower, Lewis, and Shum (2013) also look at the US supreme courts and investigate the trade-off between politicians and bureaucrats. While the *standard* spatial voting model assumes sincere voting, this assumption can be relaxed. In the application we consider in this paper there is no particular reason why strategic voting would be an issue. This also discussed in Blinder (2007). Explicitly allowing for strategic voting in our framework could be done along the lines of Clinton and Meirowitz (2004).

has been the subject of an extensive literature (see earlier references). Here we review some of the claims in this literature. Specifically we investigate the differences in revealed policy preferences of individual members. We compare the preferences of members according to their backgrounds and according to appointment. The difference in voting behavior of career central bankers and outsiders has been the topic of great interest and we investigate this is in detail. Our results confirm some findings in the literature while providing an understanding why these results emerge. As an example, we confirm the results reported by Besley, Meads, and Surico (2008) and Harris, Levine, and Spencer (2011) that it is not the case that the internals are more dovish than hawkish. We also add to the literature. We find more heterogeneity in policy preferences among external members. Extreme policy preferences (on either side) are much more likely to be held by external members than by internal members. To explore this we estimate the probability of holding the most dovish and the most hawkish policy preference. Since policy preferences are estimated with uncertainty, so is any rank ordering. However five out of the six members who have a non-negligible probability of having the most dovish viewpoint are externals (we estimate that there is more than 99.5% chance that any of these holds the most dovish viewpoint). Similarly, from all the monetary policy members who have a non-negligible probability of having the most hawkish viewpoint, we find that only one of them is an internal member. That is sir Andrew Large for whom we estimate a probability of about 10% to be the most hawkish. We also investigate whether voting members with different career backgrounds tend to hold different policy preferences. To evaluate this, we divide all monetary policy members in different categories according to their career backgrounds. We compare the median voter policy preference for each category of the monetary policy committee members. We find that the median voter in each of this categories is very similar (and similar to the overall median voter) except for those with an career experience in the industry (excluding the financial industry). We then compare the heterogeneity in policy preferences in different groups. We find that in particular monetary policy committee members with a background in academia and the industry exhibit a large heterogeneity in policy preferences. In contrast, monetary policy committee members with a central bank background exhibit the lowest heterogeneity in preferences.

In this paper we rely mostly on the voting record with respect to the policy rate. Since March 2009, the Monetary Policy Committee also votes on asset purchases. Adding these voting records allows us to refine our estimates of the ideal points. Substantively these additional votes sharpen the conclusions based only on votes on the policy rate but do not alter them qualitatively.

The structure of the paper is the following. We start in Section 2 with explaining what a spatial voting model is and how such a model can be estimated. In this section we also discuss issues related to identification. Then in Section 3, we introduce the data. We explain how the dataset was constructed and comment on the raw data. We argue that a modification of the standard approach is needed because the dataset is small which may exacerbate the influence of outliers. In Section 4 we provide the estimation results. We present some model checks which make clear where the model fits well and where the model performs worse. Here we also compare the model with some alternative specifications. In Section 5 we use the estimated ideal points to evaluate some claims in the literature. We compare internal and external monetary policy committee members. In Section 6, we explore whether policy preferences are driven by career backgrounds. In Section 7 we extend our data with the voting record on asset purchases. In Section 8 we conclude.

2 Voting records and ideal point estimation

The approach presented in this paper starts from voting records from central bank committee deliberations. Our goal is to estimate the policy preferences of each member. To do so, we borrow from statistical methods developed for analyzing political roll calls and decisions at judicial courts. These methods are rooted in the psychometric literature and in educations research but can be motivated by a spatial voting model. The spatial voting model itself has its roots in political economy, see Black (1948), but was further developed in political science, see Enelow and Hinich (1984).

2.1 Ideal points and a spatial voting model

The data we analyze consist of voting records of monetary policy deliberations. For a given central bank, we observe the votes casted on a policy rate. The data consist of monetary policy committee members to whom we refer as voters $n = 1, \ldots, N$ voting on policy choices $t = 1, \ldots, T$. Each policy choice t presents the present voting members with a choice between a *dovish* position ψ_t and a *hawkish* position ζ_t, locations in a one-dimensional Euclidean policy space $\mathbb{R}$. A voter n choosing the hawkish position ζ_t on policy choice t is denoted as $y_{nt} = 1$. If voter n chooses the dovish position ψ_t, we code this as $y_{nt} = 0$.

It is important to realize that both choices ζ_t and ψ_t are functions of a policy rate and variables capturing the contemporaneous economic conditions summarized in I_t the information set at policy choice t. However both choices differ only in the policy rate with ζ_t being the more restrictive choice i.e. the higher policy rate of the two. Assume that voters have quadratic utility functions over the policy space such that $U_n(\zeta_t) = -\|x_n - \zeta_t\|^2 + \eta_{nt}$ and $U_n(\psi_t) = -\|x_n - \psi_t\|^2 + \nu_{nt}$, where $x_n \in \mathbb{R}$ is the *ideal point* or the underlying monetary policy preference of voter n, η_{nt} and ν_{nt} are the stochastic elements of utility and $\|.\|$ denotes the Euclidean norm.[3]

Utility maximization implies that $y_{nt} = 1$ if $U_n(\zeta_t) > U_n(\psi_t)$ and $y_{nt} = 0$ otherwise.

[3]This framework is readily extended to a multidimensional policy space $\mathbb{R}^d$ with d-dimensional ideal points and positions. All methods presented here are valid in the multidimensional case. However the intuition quickly becomes more difficult and identification harder as we discuss in subsection 2.2.

To derive an item response specification, we need to assign a distribution to the errors. Assuming a type-1 extreme value distribution leads to a logit model with unobserved regressors x_n corresponding to the ideal points of the voters:[4]

$$\begin{aligned} P(y_{nt} = 1) &= P\left(U_n(\zeta_t) > U_n(\psi_t)\right) \\ &= P\left(\nu_{nt} - \eta_{nt} < \|x_n - \psi_t\|^2 - \|x_n - \zeta_t\|^2\right) \\ &= P\left(\nu_{nt} - \eta_{nt} < 2(\zeta_t - \psi_t)x_n + \psi_t^2 - \zeta_t^2\right) \\ &= \text{logit}^{-1}(\beta_t x_n - \alpha_t) \end{aligned} \tag{1}$$

The last line follows by substituting $2(\zeta_t - \psi_t)$ with β_t and substituting $\zeta_t^2 - \psi_t^2$ with α_t. To understand these coefficients, start by considering the situation where β_t equals 1. Then the model reduces to:

$$P(y_{nt} = 1) = \text{logit}^{-1}(x_n - \alpha_t). \tag{2}$$

Figure 1 provides an illustration of this model as it might be estimated for two voters and two different meetings.

Voter 1 has an ideal point x_1 slightly smaller than zero, whereas voter 2 has an ideal point x_2 larger than 2. The dove-hawk dimension runs from dovish to hawkish and so x_2 would be a clear hawk here. Consider voter 2 voting in the meeting with vote-difficulty parameter α_2. We have that $x_2 - \alpha_2 > 0$, so we find that $\text{logit}^{-1}(x_2 - \alpha_2) > 0.5$. This means that voter two when voting in meeting 2 has an estimated probability higher than 0.5 to choose for the hawkish policy option.
Consider meeting 1 for which we have estimated a vote-difficulty parameter α_1 which is to the far left. For both voters we find that $x_n - \alpha_1 > 0$ hence for both voters we would expect a better-than-even chance of voting for the hawkish policy choice. These examples show that the vote-difficulty parameter captures meeting characteristics and determines how likely it is a priori that voters vote for the dovish or the hawkish policy choice.
Now consider voter 1 voting in meeting 2. The ideal point of voter 1 and meeting 2 are estimated to be the same, so we find that $\text{logit}^{-1}(x_1 - \alpha_2) = \text{logit}^{-1}(0) = 0.5$. There is an equal probability that the voters chooses the hawkish or the dovish choice.
Finally consider now the effect of β_t or the *discrimination parameter*. This parameter captures the extent to which preferences in in the dove-hawk dimension determine the choice between two competing policy rate proposals. Say we find that for a certain meeting t, β_t equals zero. Then $\beta_t x_n$ equals zero so the preferences in the underlying dove-hawk dimension do not determine the choice between competing policy rate proposals. Analogously, a negative β_t imply that doves (hawks) in the underlying have a

[4]The logit specification seems to be the more popular approach but we could just as well have assumed a joint normal distribution for the errors which results in a probit specification with unobserved regressors x_n. An example of the latter approach is Clinton, Jackman, and Rivers (2004). Substantially this does not matter.

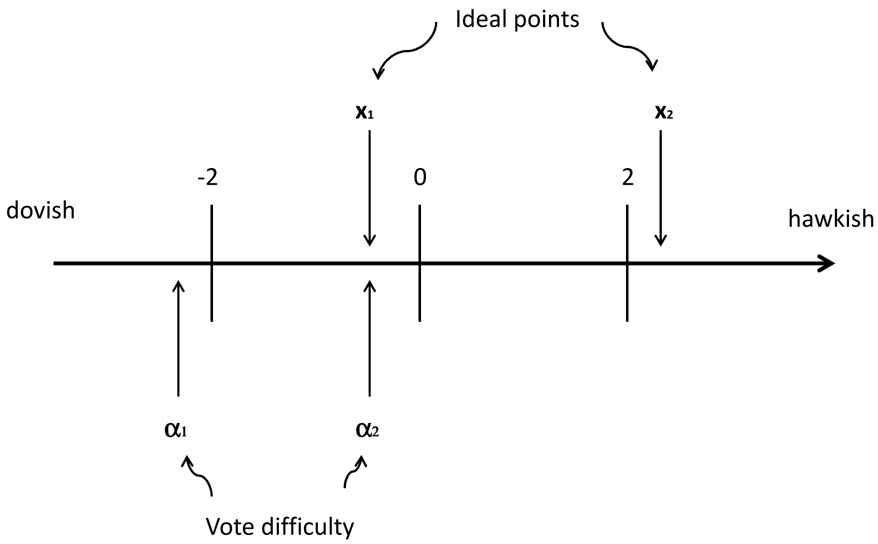

Figure 1: This figure illustrates model 2. On the latent dove-hawk dimension two ideal points x_1, x_2 (voters) and two vote-difficulty parameters α_1, α_2 (meetings) are shown. If the ideal point of voter n is larger than the vote-difficulty parameters α_t, then it is more likely that voter n votes for the hawkish policy choice. In this example, voter 1 is as likely to vote hawkish as to vote dovish on the policy choice represented by α_2.

higher probability of choosing the hawkish (dovish) policy choice. We return to the intuition behind this parameter in 4.1 where we discuss estimates of this parameter in the context of our empirical application.

2.2 Identification

Model 1 is not identified in two different ways. The first way can be seen in Figure 1, the probabilities depend on the relative position of ideal points and vote-difficulty parameters. We could add a constant to $\beta_t x_n$ and to α_t and the predictions would not change. This is referred to as additive aliasing. Analogously we could multiply $\beta_t x_n$ and α_t by a constant. This is referred to as multiplicative aliasing. In a unidimensional spatial context, identification is easier resolved than in a multidimensional model, see Rivers (2003) for a detailed description of the issues involved in general spatial models. In a unidimensional model in principle two linearly independent a priori restrictions are sufficient. For example we could simply fix two ideal points at arbitrary position, e.g. one voter at -1 and another voter at +1. Fixing two voters in this way forces the model to estimate the ideal points of the other votes relative to these two voters. The results may be hard to interpret depending on the choice of ideal points.[5] Another, more often used approach is to constrain the ideal points to have mean zero and a standard deviation of one (when using normal priors on the ideal points). This facilitates interpretation this ensures only local identification, see Clinton, Jackman, and Rivers (2004). This means that the left-right direction still can be reversed. To achieve (global) identification one needs to fix the direction. In this paper we use this approach of local identification since it is well established and reasonable. To achieve global identification we explore two different approaches. We present these in the next section after having discussed the data.

3 The data

In this paper we study the voting records of the monetary policy committee of the Bank of England. This committee is classified as individualistic in the classification scheme of Blinder (2007). Such an individualistic committee is characterized by members who express their opinions and act on them. The important advantage for our purposes is that *"the vote of an individualistic committee conveys genuine information"* (Blinder (2007)).

This facilitates our analysis as we can safely assume that the votes are in fact a reflection of the preferences of the voting members. Given the individualistic nature of the monetary policy committee, the voting records are characterized by a fairly high degree of dissent. In over 60% of the 190 Monetary Policy Committee meetings held between June 1997 and February 2013, a decision was taken by non-unanimous votes.

In this paper we drop the unanimous votes as these are uninformative for our purposes. The remaining votes were coded as decisions over two alternatives. Table 1 clarifies the coding with two examples. Example 1 is the situation where there were only two policy choices voted for in a given meeting. Nickell voted in that particular meeting for a lowering of the interest rate with 25 basis points, whereas the other voters preferred to keep the interest rate unchanged. In this case Nickell chose for the dovish option so his vote is coded as 0, whereas the others chose the hawkish choice and therefor their vote is coded as 1. If a meeting involved a choice with more than two interest rates, we coded these as a series of choices over pairwise alternatives. Consider example 2 in Table 1. At this meeting, Walton voted in favor of raising the policy rates by 25 basis points, Nickell voted in favor of lowering the policy rate by 25 basis points and the other voters preferred to keep the rate unchanged. We coded these once as a choice between raising the policy rate and maintaining *or* lowering the rate at the other hand. We coded these votes a second time but now as the choice between raising the policy rate *or* maintaining it and lowering the rate.

[5]For example if we would fix two ideal points to be -1 and +1 of voters who have a very dovish voting record. Then the other ideal points would be stretched out below -1 or above +1 and it would be hard to figure out when ideal points become centrist or hawkish.

Example 1: April 6 2006			Example 2: May 4 2006		
Name	Vote Casted	Coded once as	Vote Casted	Coded once as	Coded a second time as
King	+0	1	+0	0	1
Lomax	+0	1	+0	0	1
Tucker	+0	1	+0	0	1
Bean	+0	1	+0	0	1
Barker	+0	1	+0	0	1
Nickell	-0.25	0	-0.25	0	0
Walton	+0	1	+0.25	1	1
Gieve	+0	1	+0	0	1

Table 1: This table explains how the data was code. Example 1 shows the situation where there were only two alternatives favored. In example 2, votes were split among three policy choices.

Our raw dataset contained the votes by 32 Monetary Policy Committee members casted at 190 meetings. We recoded the recorded votes then in the way described above. We dropped the unanimous meetings. Then we dropped voting records of the Monetary Policy members who voted two times or less. This leaves us with 117 meetings and 29 Monetary Policy Committee members, henceforward referred to as voters.[6] However, not all voters vote at each meeting since the Monetary Policy Committee contains at nine voters, we can at most observe nine votes at each meeting. This implies that the 29-by-117 matrix of votes is in fact mostly empty and only 1038 entries are filled.

At this point it is useful to get a feel for the raw data. In Figure 2 on the horizontal axis we have to the total votes casted and on the vertical axis the number of votes we coded as dovish for a given voter. The straight line indicates the combinations where exactly half of the votes is coded as being hawkish and half the votes is coded as being dovish. The graph shows that there is a wide variation in the number of votes we observe for the different voters. We have 117 votes of Mervyn King in our data set while we have only 10 votes of David Walton. In the graph we labeled five voters which we use as a reference throughout this paper. As mentioned, King has the largest number of votes and has casted more votes classified as hawkish. Sentance is another example of someone who has voted predominantly hawkish, Blanchflower on the other hand voted exclusively dovish. Nickell and Buiter seem to have a centrist voting record. Both voted about equally dovish as hawkish and for both we observe a reasonable amount of votes.

Another way to look at the raw data is to consider the votes casted in minority. Two graphs are presented in Figure 3. In the left graph we present the dovish and hawkish votes (zeros and ones in our dataset) for all members *when they voted with the minority.* In the left graph we have on the horizontal axis for each voter the number of hawkish votes. On the vertical axis we have the number of dovish votes casted when voting with minority. The closer voters are to the left bottom corner, the less they voted with the minority (in absolute numbers). What we notice is that voters are either close to the

[6]Henceforward we use the term meeting for one recorded voting session in our dataset. This means that a *real* monetary policy meeting could occupy two meetings in our dataset when three policy choices were considered. The uncoded data did not contain records of votes where four or more policy choices were considered.

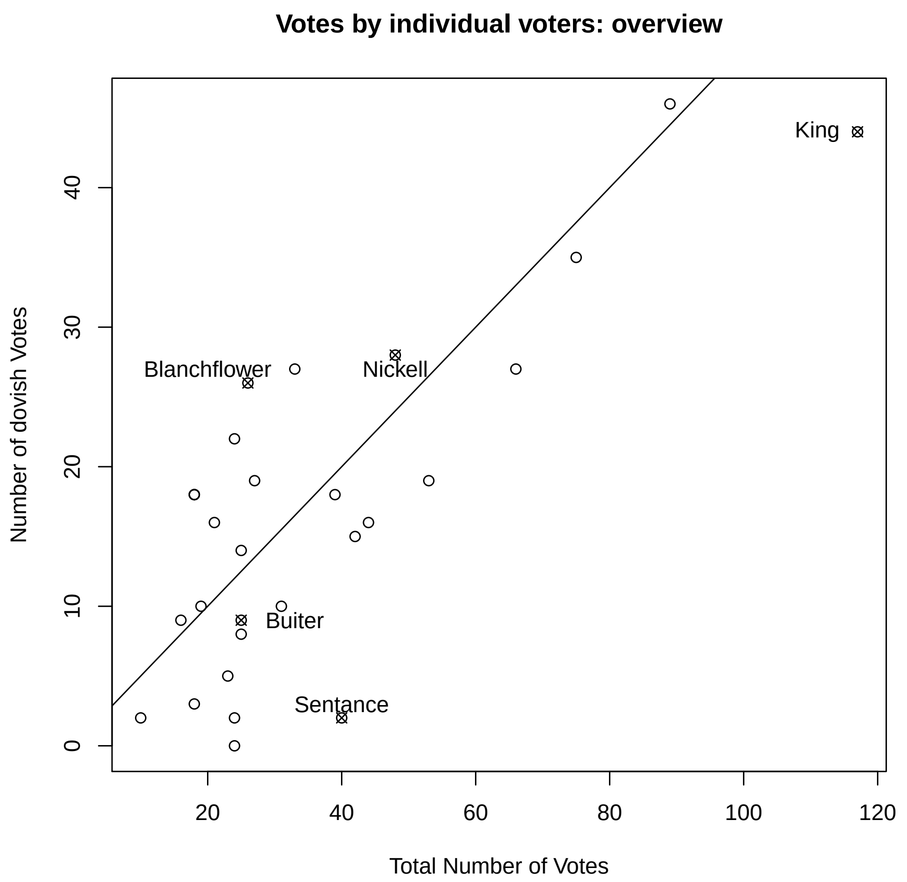

Figure 2: Here we present for each voter the total number of votes (horizontal axis) versus the number of votes coded as the dovish choice. The straight line indicates the combinations where exactly 50% of the total votes is coded as being dovish and 50% as being hawkish.

vertical axis or close to the horizontal axis. This means that if voters dissent with the majority, they tend to do this mainly on the dovish side or mainly on the hawkish side. This makes sense if votes align well with an underlying dove-hawk dimension. A hard-nosed inflation hawk is unlikely to dissent against the majority by voting for a more dovish policy rate. For example we can see that David Blanchflower dissented only by voting for dovish proposals and similarly Andrew Sentance only dissented by voting for hawkish policy rates. We also see some voters who have dissented in either direction. Three of them stand out and we have labeled the dots representing them in the graph: Nickell, Buiter and King.

The right graph tells the same story as the left graph. We plotted the same but now as fractions of the total votes we have for each voter in our dataset. On the horizontal axis we have for each voter the percentage of total votes which were casted in minority and for a hawkish proposal. On the vertical axis we have for each voter the percentage of total votes which were casted for in minority and for the dovish proposal. In this graph Willem Buiter stands even more out than in the left graph. We see that he votes with the minority in over 60% of the meetings we have of him in our dataset. The only voter who does better in this regard is Blanchflower. Nearly 30% of Buiter's votes were in favor of the dovish

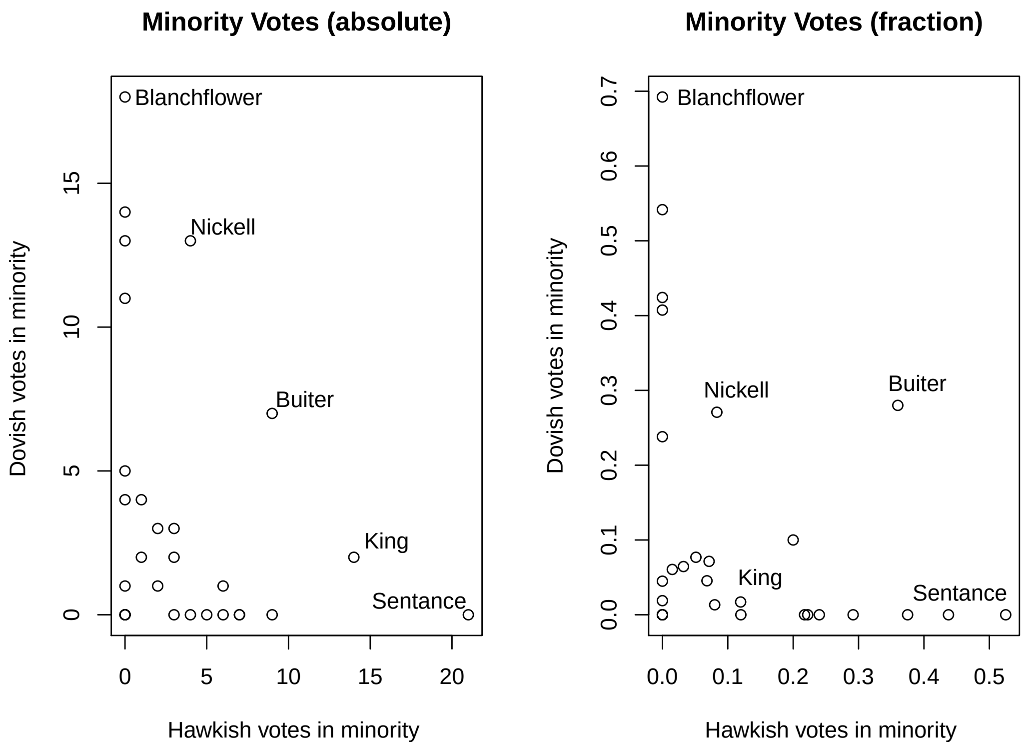

Figure 3: Here we present two graphs where the dovish votes are plotted against the hawkish votes for each voter. We only consider minority votes in both graphs. In the right graph we have scaled each point by the total number of votes casted by each individual voter.

majority while he also voted nearly 40% of the time for the hawkish minority proposal. Taken together these graphs already suggests some candidates who are likely to be dovish or hawkish, like Blanchflower and Sentance. Additionally we see that there may be a good chance that the votes casted by Buiter do not align well with an underlying dove-hawk dimension.

3.1 Outliers and few observations: a robust modification

Before proceeding to the empirical analysis we would like to motivate and present a modification of the standard spatial model. The model we presented in 2.1 has quickly become the standard approach for estimating ideal points of legislators.[7]

However, the data we have in the context of Monetary Policy Committees is more limited than the roll call data available to researchers investigating votes in the Senate or Congress for which these methods were developed. Consider for example the seminal article by Clinton, Jackman, and Rivers

[7]The canonical method for inferring ideal points was based on an unfolding procedure and known as NOMINATE. In recent years the Bayesian approach of estimating ideal points has become the preferred method. A discussion and comparison of both approaches can be found in Clinton and Jackman (2009). Without going into detail on the comparison we mention two important reasons for choosing the Bayesian approach. First of all, the NOMINATE approach we drop additional voters and meetings because they are inappropriate for the algorithm (too many votes or too lopsided meetings). Second the Bayesian approach facilitates inference over derived quantities. We make use of this in further sections of this paper. A classic introduction to the statistics of standard spatial model as presented here is Clinton, Jackman, and Rivers (2004).

(2004) where the authors fit the standard spatial voting model to the roll calls from the 106th U.S. House of representatives. This gives the authors 444.326 individual voting decisions. Compare this with our sample of only 1038 individual voting decisions and the huge variation in number of observed votes. Under a correctly specified model this should not matter as less votes would just increase the uncertainty.

But there is another problem. Logit and probit models are not robust to outliers. This was already shown by Pregibon (1982) and more recently by Liu (2004). In this context the term outlier refers to an observation of an outcome that is highly unexpected given the linear predictor. Bafumi, Gelman, Park, and Kaplan (2005) provide the following example. Say we have estimated a logit model $P(y_i = 1) = \text{logit}^{-1}(X_i\beta)$ and we have for a particular observation i, $X_i\hat{\beta} = 10$. Then $\text{logit}^{-1}(10) = 0.99995$ so the observation $y_i = 0$ would be an outlier. The many missing entries in our votes matrix and the fact that we observe only a limited number of votes for some voters potentially aggravate the problem of outliers. Additionally we saw earlier that some voters have a voting profile which may not align well with an underlying dove-hawk dimension. Only a few outliers could already bias our parameter estimates.

A modification to the standard voting model to become more robust against outliers (in the sense explained earlier) is proposed in Bafumi, Gelman, Park, and Kaplan (2005). To understand this, consider model 1 we derived earlier. Bafumi, Gelman, Park, and Kaplan (2005) propose to add a level of error ϵ_0 and ϵ_1 as follows:

$$P(y_{nt} = 1) = \epsilon_0 + (1 - \epsilon_0 - \epsilon_1)\text{logit}^{-1}(\beta_t x_n - \alpha_t). \tag{3}$$

Now, every person has an immediate probability of success ϵ_0 and of failure ϵ_1. The initial item-response model applies then to the remaining outcomes. This simple modification makes the standard spatial voting model more robust and is straightforward to implement. When we do model checking (see 4.2) we are going to explicitly compare the performance of this modified model presented here with the standard spatial voting model we presented earlier.

Our approach is Bayesian so we need to specify priors. In the literature on ideal points the *local identification* approach as we outlined earlier is considered to be the least restrictive. We follow this approach and choose therefore standard normal priors for the ideal points. However we then still need to fix the dove-hawk direction, see our discussion in 2.2. We propose two different ways of doing this. In this way, we can compare the resulting ideal points and ensure that our empirical analysis is not sensitive to the assumptions we make in order to achieve global identification. The two different sets of priors are summarized in table 2.

Our preferred choice of priors to which we refer to as the *baseline* prior choice ensures that the discrimination parameters β_t cannot be negative. Remember that we coded votes as 0 or 1 where in each given meeting a vote coded as zero was the most dovish (the least restrictive) of two proposed policy rates. Restricting β_t to nonnegative values makes then sense. It implies that we explicitly model

Parameter	Baseline Prior
ϵ_0	$\sim$ Uniform$(0, 0.1)$
ϵ_1	$\sim$ Uniform$(0, 0.1)$
α_t	$\sim N(0, 4)$
x_n	$\sim N(0, 1)$
β_t	$\sim N(1, 4)$ truncated at 0

Parameter	Alternative Prior
ϵ_0	$\sim$ Uniform$(0, 0.1)$
ϵ_1	$\sim$ Uniform$(0, 0.1)$
α_t	$\sim N(0, 4)$
x_n	$\sim N(0, 1)$
with $x_{blanchflower}, x_{wadhwani}$ truncated above at 0	
with $x_{large}, x_{sentance}$ truncated below at 0	
β_t	$\sim N(1, 4)$

Table 2: This table provides an overview of the two sets of prior choices we work with throughout the paper. Our preferred choice is labeled as baseline prior. The alternative prior serves as a check.

the directionality of each vote which is clear in this application. We restrict β_t to be nonnegative by choosing for each β_t a diffuse normal prior with positive mean which is truncated at zero.[8] When we present different model checks, we show a straightforward way to check the this assumption, see 4.1. For the vote-difficulty parameters α_t we also choose diffuse priors. The prior choice for ϵ_0 and ϵ_1 follow the recommendations of Bafumi, Gelman, Park, and Kaplan (2005). This prior choice restricts the values of these parameters to lie in the interval $[0, 1]$. This is not restrictive because if we would find values which are close to 0.1 suggesting an immediate chance of success or failure of 10% then a logit-type model should not even be used as an approximation, see Bafumi, Gelman, Park, and Kaplan (2005).

The *alternative* prior choice relaxes the assumption on the discrimination parameters. We do not truncate the normal distribution and so the discrimination parameters could take on negative values. To achieve global identification we restrict the support of the priors of certain ideal points of voters which are *obvious* candidates for being Hawk or Dove. Specifically we restrict the ideal points of Wadhwani and Blanchflower to be negative and the ideal points of Large and Sentance to be positive. This seems reasonable since we coded 22 of the 24 votes by Wadhwani in our dataset as 0 and all 26 votes by Blanchflower. At the other hand we coded 38 of the 40 votes by Sentance and 22 of the 24 votes by Large as 1. Just as with the baseline prior, we can check the reasonableness of this prior.

Simulation from the posterior is done by means of slice sampling, see Neal (2003), as implemented in the MCMCpack package, see Martin, Quinn, and Park (2011) for details. The MCMC algorithm ran for 1 million iterations. We discarded the first 200.000 draws and thinned the remaining iterations by a factor 80 to keep 10.000 draws. We estimated four models in total, both the standard spatial voting

[8]Cromwell's rule states that if a particular region of the parameter space has zero prior probability then it also has zero posterior probability, see Jackman (2009). So by restricting the support of the prior to the positive real line we have in fact restricted the β_t parameters to be nonnegative.

model and the robust model under the baseline prior and the alternative prior. Standard convergence tests suggested convergence and good mixing. In the appendix of to this paper we report these diagnostic tests in detail. We provide trace plots for three ideal points, we consider the stickiness of the Markov chain by looking at the autocorrelation of successive draws of some parameter, we consider the appropriate length of the Markov chain with the diagnostic proposed by Raftery and Lewis (1992) and we investigate whether the chain is stationary by calculating the diagnostic proposed in Geweke (1992). In the same appendix we also present results of a thorough sensitivity analysis where we look at the sensitivity of our findings to alternative prior choices. We do this by starting with model 3 and the baseline prior choice and then specifying different priors for the parameters. We re-estimate the model with four alternative specifications. We conclude from these tests that our results are robust.

4 Ideal points at the Bank of England

We start by estimating the ideal points of the Monetary Policy Committee members with model 3 and the baseline constraints. Figure 4 presents the ideal points of the 29 voters along with the uncertainty in the estimates.

Inspection of the figure shows that we could roughly classify the monetary policy members as follows. The doves are Blanchflower, Wadhwani, Julius, Allsopp and Nickell. The group of Hawks consists of Sentance, Besley, Large, Budd, Weale, Dale and King. The other voters are then centrist. In this group, we find that for Fisher, Posen, Miles and Bell the 95% credibility overlaps barely with zero and we would be inclined to label these as Dovish as well.

Before we proceed with exploring various ideas on the voting behavior of the monetary policy committee members it is important to verify that the estimation results pass some checks. These checks give an insight in how well the model fits the data and what the impact of certain assumptions are. We devote a few pages to model checking because we feel that this is an under appreciated yet critical aspect of (Bayesian) data analysis. One of the goals of this paper is to give an exposition on the methodology of ideal-point estimation to economists who are less familiar with this approach. Model checking is in our opinion a crucial element. The idea of model checking is not unique to Bayesian data analysis and has been used by researchers working with complex stochastic models in a variety of fields. A useful reference discussing the philosophical aspects and containing plenty of references is Gelman and Shalizi (2013). Practical advice and specific procedures on which we draw in this paper are described in Gelman, Goegebeur, Tuerlinckx, and Van Mechelen (2000), Bafumi, Gelman, Park, and Kaplan (2005) and Gelman, Carlin, Stern, and Rubin (2003).

In the following subsection we undertake three checks. We start by gauging the impact of the identifying constraints, respectively the constraint on the discrimination parameters and on the ideal points. We find

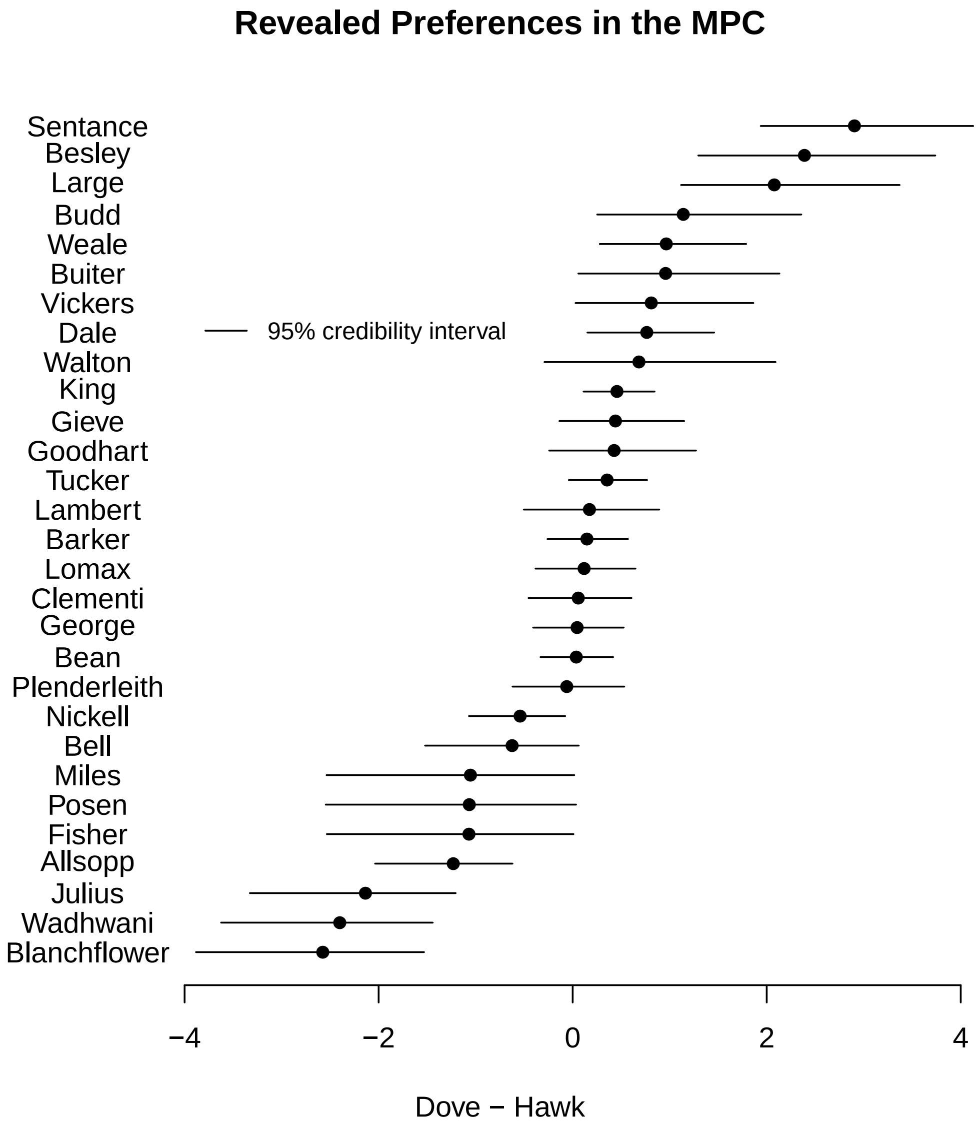

Figure 4: This figure is a graphical representation of the estimated ideal points of the monetary policy committee members. A point indicates the estimate of the ideal point, the thin line represents the 95% credibility interval.

that both constraints are reasonable and not outright rejected by the data. Then we compare the model specification we prefer, that is the simple modification as given in equation 3, with the standard spatial model. We show that the robust modification gives results which are less sensitive to the prior choice. This check shows that our results do not depend on the identifying assumption we make. Additionally it provides evidence in favor of our specification choice over the standard spatial model. Both these checks give an insight in respectively the impact of our prior and our model choice. Then we investigate prediction errors. These provide a more rigorous idea of the model fit.

4.1 The constraints on the priors

In this subsection we inspect the constraints on the priors we used to fix the dove-hawk direction. We start by inspecting the constraint on the discrimination parameter. Our baseline prior constrains β_t to be positive. This constraint on the discrimination parameter is from a theoretical point of view justified. In our framework, when given the choice between a lower and a higher policy rate, someone who is more hawkish should be more inclined to choose the higher policy rate then someone who is less hawkish. Here we check whether this is confirmed by the data. There are two reasons why we could find a negative discrimination parameter when we would *not* explicitly constrain the discrimination parameter as we do in our baseline prior choice. First of all we could have miscoded the data. As such investigating the negative discrimination parameters under the alternative prior choice offers an additional check of the data. Secondly, negative discrimination parameters may result because of *switching coalitions*. When hawkish voters vote in the dovish direction or dovish voters vote in the hawkish direction this situation can arise. Figures 2 and 3 revealed that voters do in fact vote in both directions. Moreover some tend to vote for the minority in either direction, something which may be suggestive of switching coalitions. In figure 5 we plot a random draw from the posterior distribution of the discrimination parameters. In the left graph we plot a random draw from our model with the baseline prior and in the right graph from the model under the alternative prior choice. The left graph shows that a few discrimination parameters are close to zero under the baseline prior choice. The right plots reveals which discrimination parameters are suspect in particular. We see that when we put a constraint on the discrimination parameters, some would take on negative values. The graph suggests that these negative discrimination parameters are clustered in two periods. Consider as an example the first period which coincides with the tenure of Willem Buiter. After having developed a more hawkish voting profile, he voted in a number of meetings for the dovish choice whereas the majority of the other members including, those who tend to be more dovish, voted for the hawkish option. This explains why the discrimination parameter is negative in the unconstrained specification. Our prior choice does not allow for negative discrimination parameter but now these votes receive a very low weight. Our preferred model with the baseline constraints effectively puts a value very close to zero on these particular meetings. The interpretation is that under our model the votes in these meetings do not reflect the latent policy preferences of the voters. Overall, this graph tells us that the model fits fairly well. The impact of the choice of identifying constraint on the estimated ideal points is mild. The major impact of the identifying constraint is the ideal point of Buiter which is substantially more hawkish under the alternative prior choice as shown in figure 7.

The alternative prior choice also incorporates some prior knowledge into our framework. Based on the outspoken voting records of Blanchflower, Wadhwani (both very dovish) and Sentance, Large (both very hawkish), we restricted the priors on their ideal points to take on only non-positive values for the former two. The latter two ideal points were restricted to take on only non-negative values. In principle only two

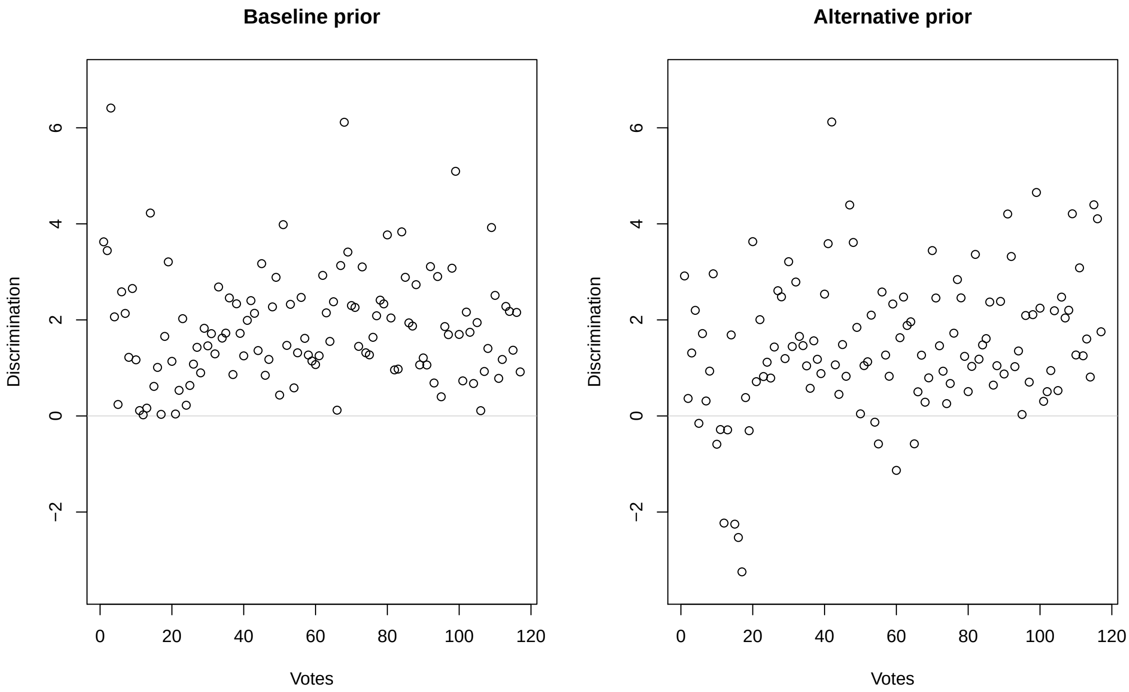

Figure 5: These graphs present a random draw from the posterior distribution of the discrimination parameters β_t, plotted across meetings. The left graph shows a random draw when under the baseline prior choice and the right graph under the alternative prior choice.

constraints instead of four should be sufficient. Constraining four ideal points instead of two improved our simulation. Here we investigate whether constraining these ideal points was in fact warranted. To do this we can inspect the posteriors of the constrained ideal points. Figure 6 presents a histogram for each constrained ideal point. In the top row we see that the coefficient for Sentance is clearly separated from zero while the coefficient for Large is also reasonably well separated. In the bottom row we see that the coefficients for both Wadhwani and Blanchflower are well separated. This suggests that these alternative prior choices were another reasonable way to break the reflection invariance.

4.2 Comparing the robust and standard spatial model

To check whether our results are not dependent on our prior choice and identification scheme we compare the estimated ideal points under both identifying prior choices. We do this once for the robust modification and once for the standard model. In the left graph of Figure 7 both prior choices are compared for the robust model. When we obtain the exact same estimates for the ideal points, all dots should lie on the diagonal. We find that the estimated ideal points are close to the diagonal except for the ideal point of Willem Buiter. His ideal point depends on the prior choice to some extent although it remains positive.

We do the same check and compare the ideal points found under both prior choices when we estimate

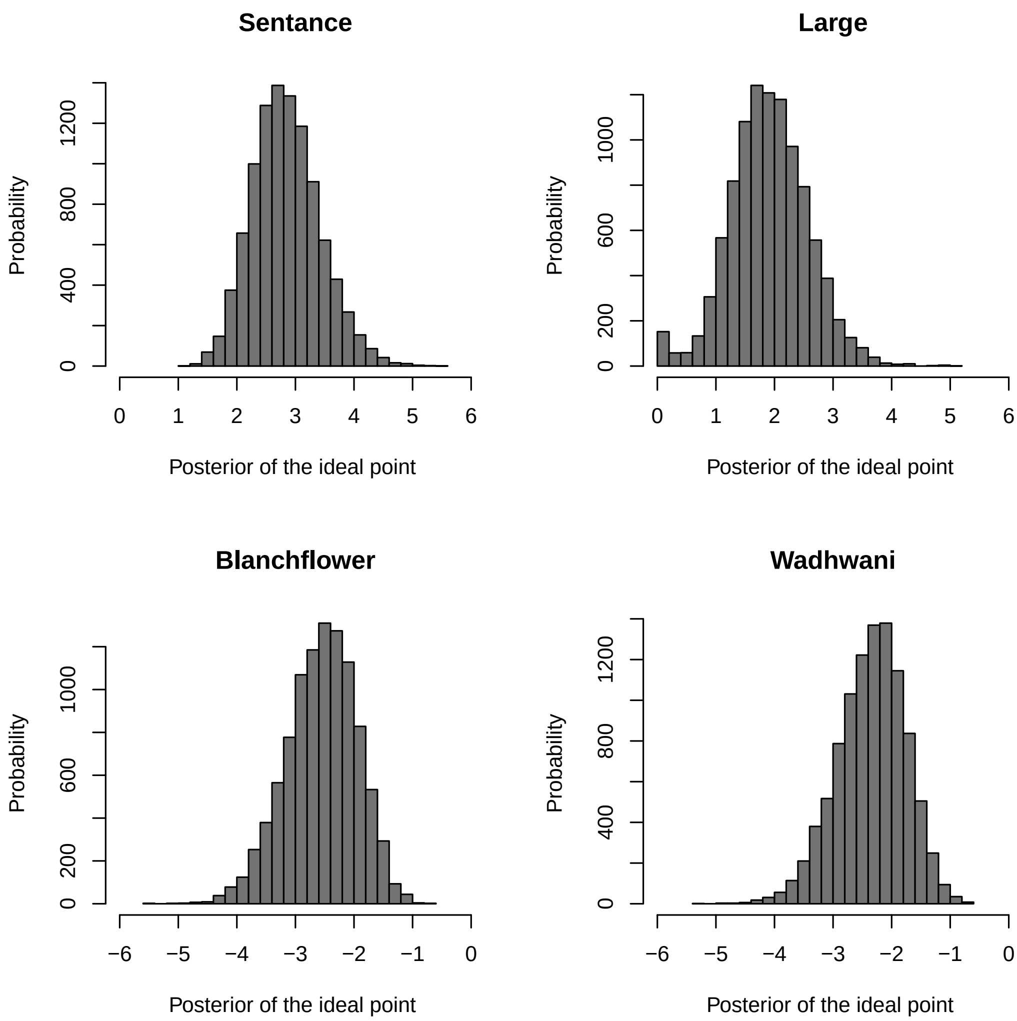

Figure 6: These four histograms show to what extent the posterior distribution of each constrained ideal point is in line with the priors we specified for these ideal points. The graphs suggest that the posteriors are well separated from zero, providing support for the priors.

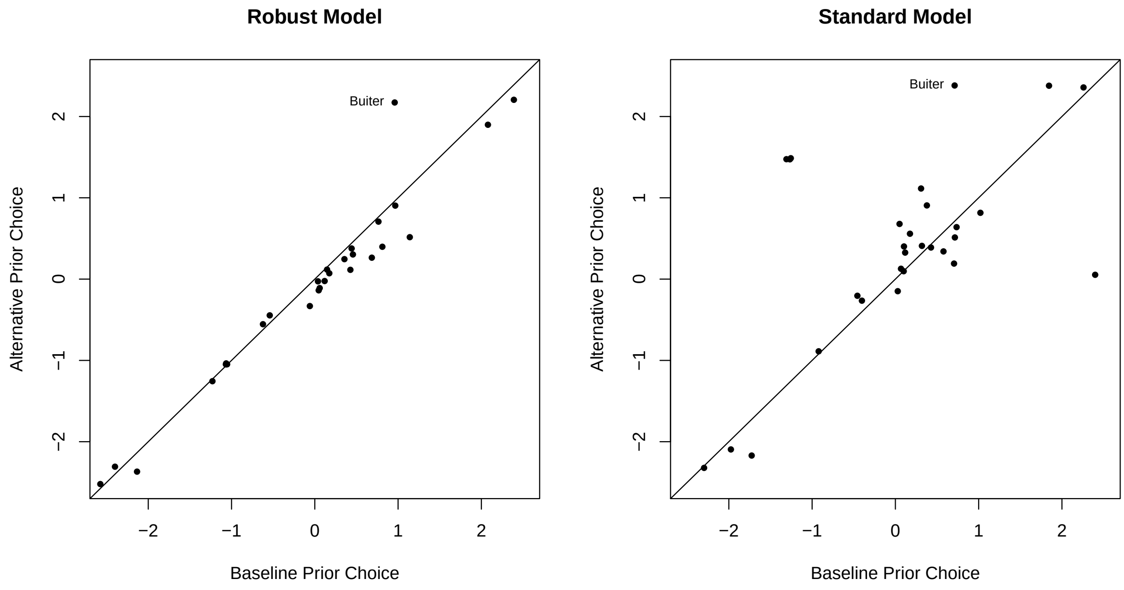

Figure 7: In these graphs we compare the ideal points found under the baseline prior choice (horizontal axis) and the alternative prior choice (vertical axis). When the ideal points are exactly the same under both prior choices, the dots should all lie on the 45 degree line. The further away from this line, the more the estimate is sensitive to the choice of prior. In the left graph this comparison is presented for our preferred specification. All dots are close to the diagonal line except for the ideal point of Buiter. The right graph displays the comparison for the standard spatial voting model. We see that the estimates are far less stable.

the standard spatial model. These results can be found in the right graph of Figure 7. We see that in the standard spatial voting model (without the small modification) the estimates are far less stable. Several dots lie relatively far away from the diagonal, indicating that the ideal points change according to the identifying assumption. Some even switch from clearly dovish to clearly hawkish. This supports our choice for Model 3.

4.3 Prediction errors

Here we check the performance of our model by looking at prediction errors. Define a prediction error pe_i as:

$$
\begin{aligned}
pe_i &= 1 \quad \text{if} \quad \mathbb{E}(y_i) > 0.5 \text{ and } y_i = 0, \text{ or } \mathbb{E}(y_i) < 0.5 \text{ and } y_i = 0 \\
&= 0 \quad \text{otherwise.}
\end{aligned} \tag{4}
$$

Using the prediction errors we can quickly check the *error rate* or the proportion of times the prediction is wrong. The error rate for our preferred model is slightly over 8%. This is substantially lower than the error rate of the null model, that is a model where we give each outcome a probability equal to the proportion of 1's in our dataset. The error rate of the null model is over 46% whih suggests that the model we propose predicts fairly well.

We can also use the prediction errors to consider the *excess error rate* i.e. the proportion of errors beyond what is expected. If the model is true, the probability of error is:

$$\mathbb{E}(e_i) = \min(\text{logit}^{-1}(\beta_t x_n - \alpha_t), 1 - \text{logit}^{-1}(\beta_t x_n - \alpha_t)). \tag{5}$$

The excess error rate ee_i is then: $ee_i = pe_i - \mathbb{E}(e_i)$.

To be able to interpret these prediction errors, we average these over voters. We plot the excess error rate for each voter. The results hereof can be found in the graphs on in the left column of Figure 8. Each row presents the results based on another draw from the posterior. These errors are referred to as *realized error rates*, see Gelman (2004). In the right column we plot the excess errors which we could expect if the model were true. This provides a *reference* to compare the graphs in the left column against. We do this by generating voting data under the assumption that our model is true. The replicated datasets are generated by drawing from a binomial distribution $B(n, p)$ with $n = 1$ and p equal to the probability of voting hawkish according to our model. The probability p here is calculated from the same draws we used to construct the realized error rates. We then calculate the excess error rates as we did in the left column.

In Figure 8 each graph in the left column presents the excess error rate while the graph in the right column provides a reference -what could we expect if the model exactly captures the data generating process. By doing this for different parameter draws we capture the uncertainty in the posterior. Comparing the different plots in the right column, it seems that there are no particular voters for which we expect persistently high excess error rate. The left column provides the fit of our model for the different voters. We see that the fit for the ideal point of Buiter is the worst. Also the ideal point of Walton seems to be slightly worse than what would expect on the basis of our model. All things considered, this model check suggests that our model fits well. The realized excess error rates are low and of the same order of magnitude as the reference errors. Even for Buiter we only find realized error rates around 0.3 which is still reasonable given the reference distribution.

5 Internals and Externals

The model checks in the previous section suggest that the model well. In this section we use our results to investigate whether groups of members systematically differ in their policy preferences. First we consider the differences between internals and externals, that is internal members who have a full-time executive position at the Bank and external members who have no executive responsibilities. Besley, Meads, and Surico (2008) make the same distinction. Harris, Levine, and Spencer (2011) split the monetary policy committee members in three groups distinguishing between external members, internal

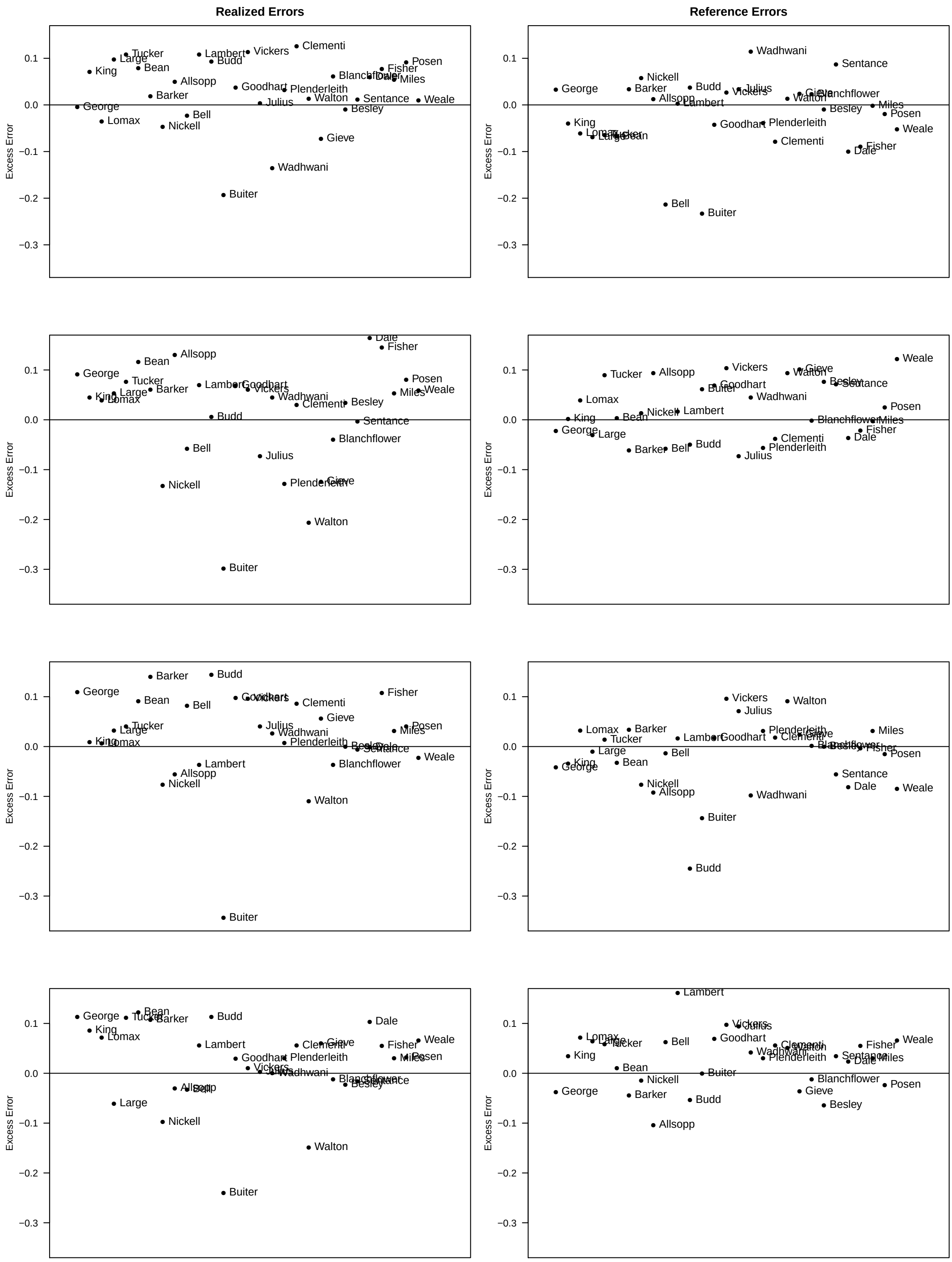

Figure 8: Here we present the excess error rate in real and replicated votes by the monetary policy committee members. Each row corresponds to a random draw of the posterior. In each plot the average excess error rate is plotted for each voter. The plots on the right show what we could expect if the model were true, whereas the plots on the left show the realized excess error.

Name	Internal or External	Political or Non-political	Finance	Industry	Government	Academia	Bank	NGO
George, Eddie	Internal	Political	N	N	N	N	Y	Y
King, Mervyn	Internal	Both	N	N	N	Y	Y	N
Clementi, David	Internal	Political	Y	N	N	N	N	N
Large, Andrew	Internal	Political	Y	Y	Y	N	N	N
Lomax, Rachel	Internal	Political	N	N	Y	N	N	Y
Gieve, John	Internal	Political	N	N	Y	N	N	N
Plenderleith, Ian	Internal	Non-political	N	N	N	N	Y	Y
Dale, Spencer	Internal	Non-political	N	N	N	N	Y	N
Fisher, Paul	Internal	Non-political	N	N	N	N	Y	N
Vickers, John	Internal	Non-political	N	N	N	Y	N	N
Bean, Charles	Internal	Non-political	N	N	Y	Y	N	N
Tucker, Paul	Internal	Non-political	Y	N	N	N	Y	N
Buiter, Willem	External	Political	N	N	N	Y	N	N
Goodhart, Charles	External	Political	N	N	Y	Y	Y	N
Julius, DeAnne	External	Political	N	Y	N	Y	N	Y
Budd, Sir Alan	External	Political	Y	N	Y	Y	N	N
Wadhwani, Sushil	External	Political	Y	N	N	Y	N	N
Nickell, Stephen	External	Political	N	N	N	Y	N	N
Allsopp, Christopher	External	Political	N	N	N	Y	Y	N
Barker, Kate	External	Political	N	Y	N	N	N	N
Bell, Marian	External	Political	Y	N	Y	N	N	N
Lambert, Richard	External	Political	Y	N	N	N	N	N
Walton, David	External	Political	Y	N	N	N	N	N
Blanchflower, David	External	Political	N	N	N	Y	N	N
Besley, Tim	External	Political	N	N	N	Y	N	N
Sentance, Andrew	External	Political	N	Y	N	Y	N	N
Weale, Martin	External	Political	N	N	N	Y	N	N
Miles, David	External	Political	Y	N	N	Y	N	N
Posen, Adam	External	Political	N	N	N	Y	Y	N

Table 3: The information provided in this table draws upon tables provided by Harris, Levine, and Spencer (2011) and Hansen, McMahon, and Rivera (2012) which we updated for members who joined the monetary policy committee after May 2007. Y (N) stands for Yes (No) and means that the monetary policy member does (not) have career experience in that sector. Some classifications required judgement calls. In particular we classified Weale and Posen to have academic experience. Weale spent time at the NIESR which we consider to be an academic institution. Similarly we labeled Posen's experience at the Peterson Institute, combined with his publication record as academic experience.

members who are politically appointed and internal members who are not politically appointed. In table 3 we provide an overview of all the monetary policy committee members we consider in our analysis as well as some information on career backgrounds. The info in this table comes from Harris, Levine, and Spencer (2011) which we updated for the monetary policy committee members who joined after 2007. The classification requires some judgement calls. We tried to only take major appointments into account and so we disregarded consulting roles or special advisory positions.

So how can we investigate whether there are meaningful differences or group patterns? At first sight it seems that externals and internals do not seem to be easily classified as either dovish or hawkish. Outspoken Doves such as Blanchflower and Wadhwani as well as clear hawks such as Sentance and Besley both belong to the external group. What does seem to be the case is that the internals tend to take the centrist position. Of the politically appointed internally appointed members only Large and King are hawks. Also the other internally appointed members tend to take the centrist position and only Dale has a hawkish ideal point. Remarkably all the doves belong to the external group. These conclusions all stem from looking at figure 4. To verify that it are indeed the external members who have the most outspoken ideal points, we want to infer the *rank* of the estimated voting scores.

Our Bayesian simulation results facilitate this analysis. As explained in Jackman (2009), given the joint posterior density over the ideal points $\mathbf{x} = (x_1, \dots, x_{29})$ we can induce a posterior density over any quantity of interest that is a function of the $\mathbf{x}$.

5.1 The most dovish and hawkish voters

To obtain a posterior density over the order statistics for each voter we use the following procedure.[9] For each MCMC draw $k = 1, \dots, 10.000$, we order the ideal points $x_n = x_1, \dots, x_{29}$ and assign a rank r to the sampled ideal points. Denote the ranks r at each iteration of the MCMC algorithm as $\mathbf{r^{(k)}} = (r_1^{(k)}, \dots, r_{29}^{(k)})'$. Each element of $\mathbf{r^{(k)}}$ is an integer $r_n^{(k)} \in \{1, \dots, 29\}$. The probability that voter n occupies rank r is thus $\sum_{k=1}^{10000} \mathbb{1}_{(\in\{r\})} r_n^{(k)}$ By computing these ranks over the iterations of our sample, we compute a posterior mass function over the possible ranks.

Using this procedure we also compute the probability of being the *most* dovish and the *most* hawkish. This is done by estimating the posterior probabilities that a monetary policy committee member occupies rank 1 and rank 29 when ranking the members from dovish to hawkish. The results can be found in Figures 9 and 10.

In Figure 9 we present the marginal posterior probabilities as histograms of the six voters with the highest probability of being *the most dovish*. The probability that one of these six monetary policy committee members is the most is dovish is more than 99.5%. Five out of six of these monetary policy members are externals.

In Figure 10 we present the marginal posterior probabilities as histograms of the six voters with the highest probability of being *the most hawkish*. The probability that one of these six monetary policy committee members is the most hawkish is more than 99.5%. Five out of six of these monetary policy members are externals.

Taken together our analysis suggests that the most dovish and most hawkish members at the monetary policy committee of the Bank of England so far were externals. Internal members tend to occupy more the middle ground. To our knowledge this observation is new in the literature and aligns well with previous findings. Besley, Meads, and Surico (2008) and Harris, Levine, and Spencer (2011) could not classify external or internal members as either more dovish or more hawkish. Our results confirm their conclusion. Recent work by Hansen, McMahon, and Rivera (2012) suggests that internals make more precise assessments of the economy and that they tend to be more hawkish. Our analysis does not shed light on the precision of their assessments but the claim that internals are more hawkish cannot be confirmed by our analysis. Our procedure takes preferences in a relative manner into account. That is,

[9] This procedure can be found with additional detail in Jackman (2009), chapter 9.

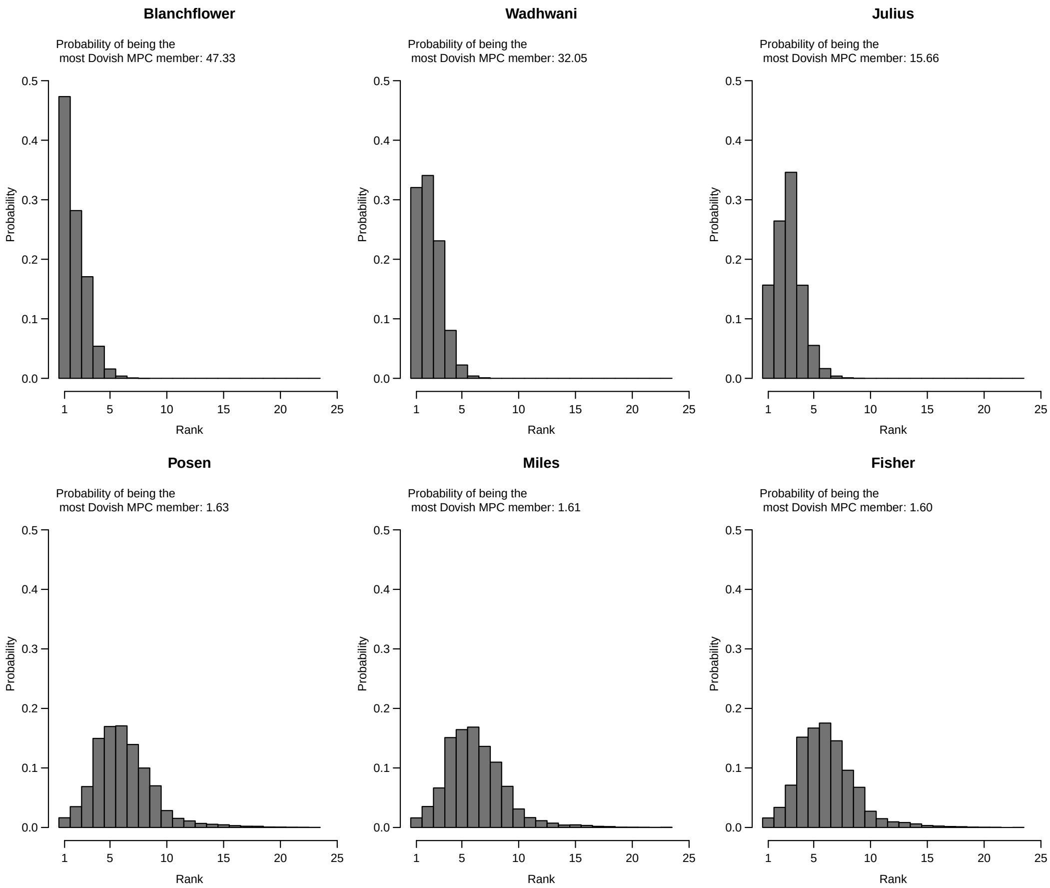

Figure 9: These graphs represent the posterior mass functions over order statistics of ideal points. Additionally the probability of being the most dovish, which is the proportion of times we see the ideal point occupying the lowest rank over our 10.000 samples, is mentioned. These are the six voters with the highest probability of being the most dovish.

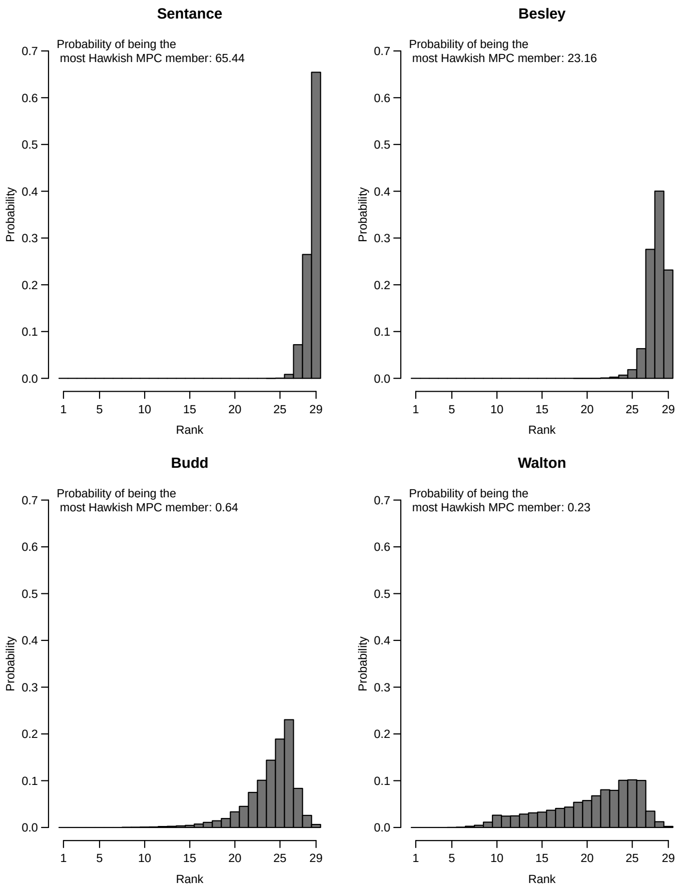

Figure 10: These graphs represent the posterior mass functions over order statistics of ideal points. Additionally the probability of being the most hawkish, which is the proportion of times we see the ideal point occupying the highest rank over our 10.000 samples, is mentioned. These are the six voters with the highest probability of being the most hawkish.

the choice between different policy rates is always conditional. The estimated ideal points for our clear doves and hawks are well separated from zero and so we are confident that the conclusions hold. Internals do not vote clearly more hawkish or dovish then externals. But externals do tend to have members among their ranks with more outspoken policy preferences. This finding resonates with the views expressed in Gerlach-Kristen (2003) and Harris and Spencer (2009) that differences between internals and externals could arise because of an organizational consensus among internals. Related to this, there may be career concerns among internals which are less relevant for external members. Gerlach-Kristen (2003) suggests that externals may even be incentivized to gain media attention. Our analysis does not focus on the act of dissenting per se but rather on the revealed policy preferences. Our results do suggest that internals indeed tend to have policy preferences which are less heterogenous.

6 Career backgrounds

Related to the differences between internals and externals, we are also interested in background effects. The intuition is that career background may persistently influence the policy preferences of monetary policy committee members. This notion comes from the literature investigating voting at the FOMC where such effects have been suggested. The aforementioned study by Harris, Levine, and Spencer (2011) only finds *weak* (often counterintuitive) influences of career backgrounds when analyzing the records of dissents. Besley, Meads, and Surico (2008) consider less career background characteristics when comparing coefficients of reaction functions. They do not find a meaningful pattern.

We investigate the differences in policy preferences by comparing the entire group of voters and the voters with experience in (1) the finance industry (including banks), (2) industry in general (excluding the financial industry), (3) government (civil service or working for any government), (4) academia (only an appointment post doctoral education counts; most voters have obtained a Ph.D.), (5) at the Bank of England, (6) at an NGO. An overview is provided in Table 3. These groups overlap so some voters belong to multiple groups.

We compare these groups by comparing the median voters within each group. For each iteration k we rank the voters within the different groups and select the median voter of group $l = 1, \ldots, 7$ (the six subgroups listed above and the entire group of voters). Let $x_{l,med}^{(k)}$ denote the ideal point of the median voter in iteration k. We then have for each group l a sample of 10000 simulation draws of the ideal point of the ideal voter. Similar to our earlier inferences we can construct an estimate of the median voter ideal point and corresponding uncertainty. In the left graph of Figure 11 we present the ideal points of the seven median voter ideal points. The median voters of the six subgroups we listed above and the median voter of the entire group (all voters in our dataset). The figure reveals that the median voter from the group with industry experience is more hawkish than the median voters out of the other groups, including the overall median voter. Monetary Policy Committee members with NGO experience tend to be a bit more dovish but they are only credibly distinguishable from the group with industry experience. Harris, Levine, and Spencer (2011) found that industry experience and work experience at the Bank of England promote tightness dissents -both findings were deemed to be counterintuitive. Our findings suggest while we find a more hawkish policy preferences among those with industry experience, those with work experience at the Bank of England do not hold more hawkish policy preferences. Experience in government, academia or in finance does not seem to systematically shift the policy preferences. These results are in line with Harris, Levine, and Spencer (2011) and Besley, Meads, and Surico (2008) who could not find systematic differences in the estimated parameters of reaction functions when comparing voting members with academic experience and without.

We also consider the heterogeneity in policy preferences in the different groups. To do this we estimate the dispersion of ideal points in the different groups with a procedure similar to our estimations of the median voters in each group. For each draw of the simulation, we calculate the standard deviation of ideal points, std_l, for each group separately. The results can be found in the right graph of Figure 11.

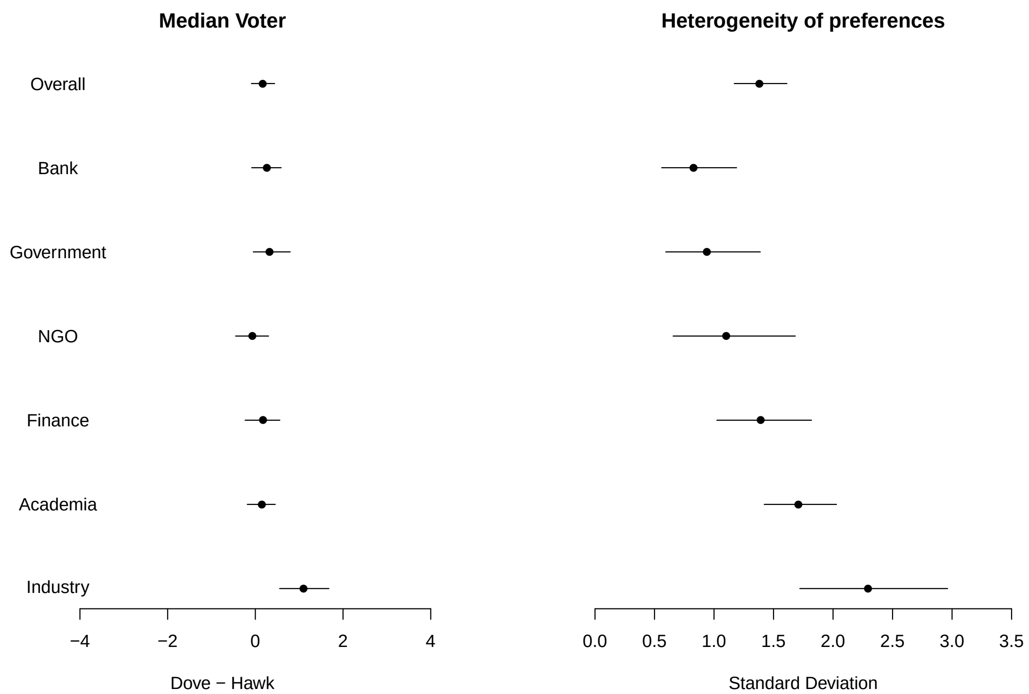

Figure 11: In this figure we present the ideal point of the median voter of different groups of voters. These groups of voters are constructed according to the career backgrounds displayed in Table 3. Voters may belong to multiple groups.

We find that there is a larger heterogeneity among the monetary policy committee members with a background in the industry and academia than those with career experience in a central bank or at the government. To estimate the probability that the heterogeneity in group A, std_A, is larger than in group B, std_B, we can generate a binary variable $D^{(k)}_{A>B}$ which takes the value of 1 when $std_A > std_B$ and zero otherwise. We can then compute the probability that the heterogeneity in group A is then larger than in group B: $P(std_A > std_B) = \sum_{k=1}^{10000} D^{(k)}_{A>B}$. The results of this calculation indicate that the heterogeneity among voters with an academic background is larger than the heterogeneity among (i) voters with a background at a central bank (> 99%), (ii) voters with a career background at the government (> 99%), (iii) voters with an NGO background (> 98%), (iv) voters with a background in the financial industry (> 93%). The heterogeneity among the voters with a background in the industry is the largest but it should be noted here that this group is very small compared to the other groups. These findings align well with the findings on internal and external members. Voters with an academic background are predominantly found among the external members. Earlier we showed that this group tends to have more extreme policy preferences. Inspection of Figures 9 and 10 shows that these are quite often voters with an academic background. One explanation could be that academics may have developed an own, idiosyncratic view on what monetary policy should do and are subsequently more pronounced in their opinions and preferences. Voters with other career backgrounds, be it in government, at a central bank or in the financial industry, may share a sort of consensus view and hence have less heterogenous policy preferences.

7 Data on Asset Purchases

Since March 2009 the Monetary Policy Committee of the Bank of England also votes on asset purchases financed with central bank reserves. We can integrate these voting data in our framework by coding the votes in a similar way as we did for the policy rate votes. First we drop unanimous decisions. Then when confronted with two alternatives of asset purchases, we code the lowest amount as the hawkish alternative (i.e. 1) and the highest amount of asset to be purchased as the dovish alternative (i.e. 0). Meetings with three alternatives are then coded as votes over consecutive pairs of alternatives as we have done for the policy rate.

These new vote data provide additional information to identify the ideal points of the voters. However some remarks are in order. First, we only have voting data on asset purchases for a limited period, hence only for a subset of all voters in the dataset. Second, we assume that a vote on the asset purchase program and the policy rate can be used more or less on equal footing in our spatial voting model. While we our model is flexible and gives different weights to votes in different meetings, this assumption underlies the data construction. In our opinion this is not a stretch. In figure 12 the estimated ideal points are shown.

If we compare these with the estimates we obtained earlier the results do not change to much. The most noticeable changes are related to Miles, Fisher and Posen. Previously we had a similar voting

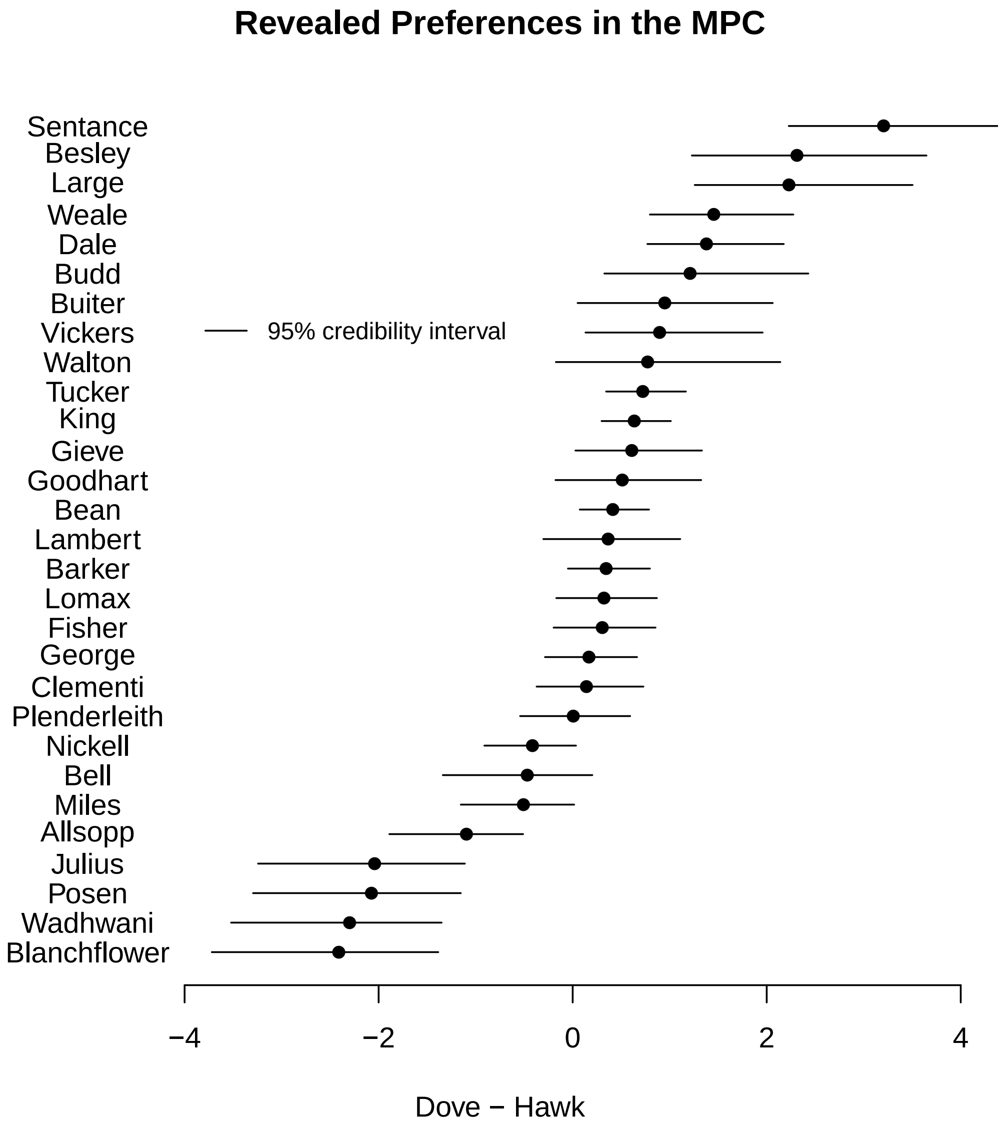

Figure 12: This figure is a graphical representation of the estimated ideal points of the monetary policy committee members. A point indicates the estimate of the ideal point, the thin line represents the 95% credibility interval.

record at our disposal and so we could not distinguish these. Moroever their voting record was fairly uninformative and so we found large uncertainty about their ideal point. The additional votes allow us to discriminate more clearly and we find now that Posen becomes one of the most outspoken Doves. Miles and Fisher are now more centrist with Fisher leaning more towards the hawkish side than Miles. This makes sense when we look at the vote data on asset purchases. Posen voted about 85% of the times for the dovish option in asset purchase decisions (more asset purchases) whereas Fisher did so only about 15% of the time. Miles voted half the time for the dovish choice and half the time for the hawkish choice. At the same time, the uncertainty on the ideal points of Miles and Fisher became much smaller whereas the uncertainty on the ideal point of Posen did not change that much. The reason for this is that in the voting data on asset purchases, Posen had the most dovish profile and often voted alone for the most dovish option. This allows us to place Posen of the left of the other on which we have asset purchase voting data but the extent to which we can do this remains uncertain.[10] It is true that the order of the other voters also changes a bit as we rank the estimated ideal points according to the median of the posterior. These medians are very close for a large group of voters in the middle and therefore sensitive to small changes.

With respect to the substantive results we presented in sections 5 and 6, these do not change. In fact our estimates support all our claims even *more*. As an example, previously we obtained a small probability that an internal would be the most dovish of all committee members. With the additional data we place Fisher more to the hawkish side. Now we only find external members in the group of outspoken doves. This is in support of our finding that pronounced preferences are found among the externals whereas internal members tend to take the middle ground. To be conservative, we preferred to show the results based on only the policy rate vote data in detail.

8 Conclusion

The spatial voting model provides an appealing way of inferring policy preferences from voting records. This approach has widespread acceptance in research outside of economics but remains rather unknown within economics. This paper introduces a Bayesian incarnation of the spatial voting model to economists. We hope to have convinced that this approach is very flexible and powerful. Flexible in the sense that we can modify the model in a variety of ways. In the paper we proposed a small modification to make the model more robust but other modifications are conceivable and may provide avenues for further research.

Another modification to deal with unpredictable voters was suggested by Lauderdale (2010). It is also possible to relax some of the assumptions of the spatial voting model. The model can be made dynamic as in Martin and Quinn (2002) although this poses some demands on the data which may be hard to satisfy. Another extension consists of considering more than one dimension. This could help in explaining the behavior of voters for which the model fit less well or those which seem to be characterized by more than just an ideal point on an underlying dove-hawk dimension.

The approach is also powerful. It delivers the joint probability of all parameters and hence we can quickly devise tests and explores ideas. As an example, in this paper we investigated whether the heterogeneity varies in groups of voters with different career backgrounds. We could estimate the heterogeneity (mea-

[10] Wide uncertainty intervals in this type of analysis is typically the result of a) little data and b) uninformative data. Extreme vote sessions (also called lopsided votes) where nearly all voters vote for one choice and one voter votes for another choice are not very informative. Such a vote session allows to place the one dissenting voter to the left or right of the others but we do not learn much more. See also the discussion in Jackman (2009) p.461.

sured by the standard deviation of ideal points) while accounting for the uncertainty in the estimates of ideal points. We obtained uncertainty in this measure of heterogeneity and could quickly verify whether the heterogeneity in one group is larger than in another group. The underlying idea is easily amenable to explore other hypotheses. Substantively we confirmed a few results in the literature and presented some findings which are to our knowledge new. We confirmed that it is not the case that internals are more hawkish than externals (or vice versa). But we noticed that extreme policy preferences are prevalent among externals. Related to this, we showed that some groups of voters have more heterogenous policy preferences than other groups. It is remarkable that academics have more differing policy preferences than other groups (except for those with an industry background), a finding which is new to our knowledge. These findings are important in the debate on the relevance (or advantage) of having externals in a central bank committee. The Bank of England is known to be an individualistic monetary policy committee. Our results suggest that within our sample, the academics have the highest degree of individualism in central bank committees. In so far this is desirable in the constitutional design of central bank committees this should be taken into account.

Not for publication

A Checks of convergence for the results in the paper

In this section we provide some basic convergence checks for the main model in our paper of which the results were presented in figure 4.

We perform three checks. First we provide three trace plots. These provide a visual inspection of the Markov chains. Then we use the diagnostic check developed in Raftery and Lewis (1992) to learn about the required run length. This diagnostic informs us that we should use many iterations. Finally we apply a diagnostic developed in Geweke (1992), to see whether the Markov Chain has converged. The latter two checks were carried out with the CODA package developed by Plummer, Best, Cowles, and Vines (2006).

Trace plots

In Figure 13 we show the trace plots of the MCMC iterations for the ideal points of Buiter (a 'special' case in the paper), Blanchflower (a clear dove) and Sentance (a clear hawk). These trace plots allow for visual inspection of the Markov chain and suggest reasonable good mixing of the Markov chain.

Raftery and Lewis's diagnostic

We started by applying the procedure described in Raftery and Lewis (1992). This procedure is a run length diagnostic. The diagnostic used here calculates the number of iterations required to estimate the 2.5%-quantile within an accuracy of +/- 0.005 with a probability of 95%. The Raftery diagnostic reveals that our Markov Chain is rather sticky and requires a fairly long run with a minimum of about 3800 iterations but ideally over 300000 iterations. We choose to undertake much longer runs in this study to remain on the safe side and so we used a Markov chain of length 1000000 of which discarded the first 200000 iterations. Afterwards we thinned the remaining iterations by a factor 80 to keep 10000 draws (see also the trace plots).

Geweke's diagnostic

This diagnostic proposed in Geweke (1992) compares the means of two parts of a Markov Chain. Here these are the first 1000 draws and the last 5000 draws. Geweke's diagnostic is a Z-score (the difference between the two sample means divided by the estimated standard deviation) and has an asymptotically standard normal distribution. The resulting Z-statistics can be found in table 4. None of the Z-statistics are larger than 1.96 in absolute value which indicates the according to this diagnostic our MCMC is stationary.

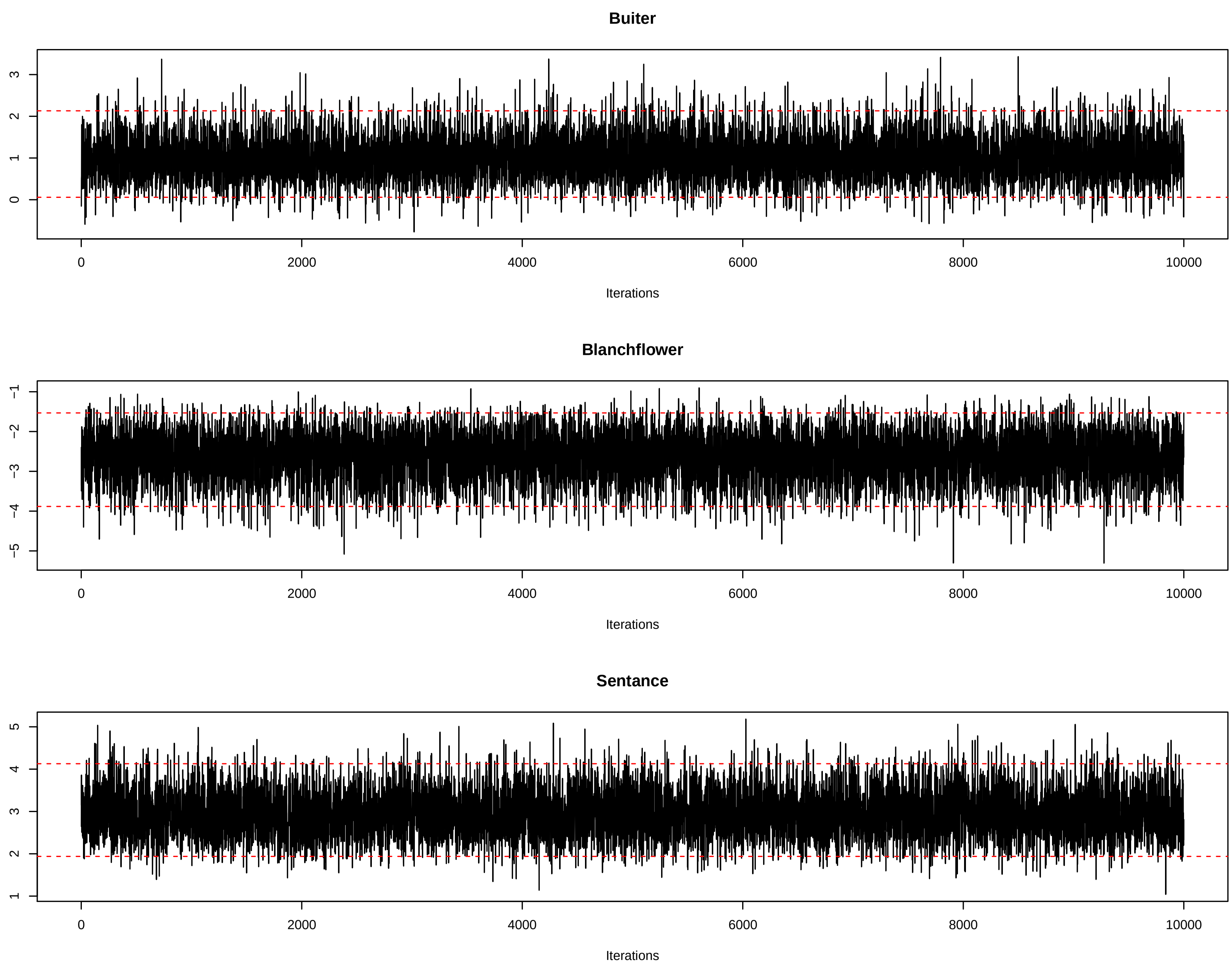

Figure 13: This figure shows three trace plots, the MCMC iterations for the ideal points of Buiter, Blanchflower and Sentance. The dashed red line indicates the 2.5 and 97.5 percentiles of the MCMC samples and provide an estimate of the 95% posterior credible interval.

Name	Z-stat	Name	Z-Stat
George	0.55	King	1.1
Lomax	1.09	Large	1.4
Tucker	-1.38	Bean	-0.43
Barker	-0.43	Nickell	-0.95
Allsopp	-0.67	Bell	-0.99
Lambert	-0.02	Budd	0.84
Buiter	-0.76	Goodhart	-0.42
Vickers	-0.39	Julius	0.65
Wadhwani	-0.86	Plenderleith	0.44
Clementi	0.41	Walton	-1.83
Gieve	-0.2	Blanchflower	1.07
Besley	-1.11	Sentance	-0.8
Dale	-0.51	Fisher	-0.46
Miles	0.79	Posen	0.48
Weale	0.19		

Table 4: This table presents the resulting Z-statistics from applying Geweke's convergence diagnostic to the MCMC samples for the ideal points.

B Sensitivity to prior choices

In the paper we have discussed in some detail the sensitivity of the results to the *identifying* priors. Here we present an additional sensitivity analysis. We only analyze the sensitivity for the baseline model (see paper). The priors we consider are the vote parameters α_t and β_t and the priors on the ϵ parameters. We also show that, except for the truncation, the β_t prior is also robust to other choices. Replication materials for the paper and the results mentioned in the appendix will be put in an online depository upon publication of this paper.

To demonstrate the sensitivity to the priors we show the results from three alternative alternative prior choices.

Table 5 provides an overview of the three sensitivity checks we undertake along with our baseline prior choice as we mentioned in the paper. We show that in fact this does not alter our the findings reported in the paper, confirming that the priors we specified in the paper are not overly restrictive.

Parameter	Sensitivity Check
ϵ_0	$\sim$ Uniform$(0, 0.1)$
ϵ_1	$\sim$ Uniform$(0, 0.1)$
α_t	$\sim N(0, 2.5)$
x_n	$\sim N(0, 1)$
β_t	$\sim N(1, 2.5)$ truncated at 0

Parameter	Baseline specification (see paper)
ϵ_0	$\sim$ Uniform$(0, 0.1)$
ϵ_1	$\sim$ Uniform$(0, 0.1)$
α_t	$\sim N(0, 4)$
x_n	$\sim N(0, 1)$
β_t	$\sim N(1, 2.5)$ truncated at 0

Parameter	Sensitivity Check
ϵ_0	$\sim$ Uniform$(0, 0.1)$
ϵ_1	$\sim$ Uniform$(0, 0.1)$
α_t	$\sim N(0, 10)$
x_n	$\sim N(0, 1)$
β_t	$\sim N(1, 10)$ truncated at 0

Parameter	Sensitivity Check
ϵ_0	$\sim$ Uniform$(0, 0.2)$
ϵ_1	$\sim$ Uniform$(0, 0.2)$
α_t	$\sim N(0, 10)$
x_n	$\sim N(0, 1)$
β_t	$\sim N(1, 10)$ truncated at 0

Table 5: This table provides an overview of the three sensitivity checks we undertake along with the baseline specification of the paper.

Graphical results of different priors

In Figure 14 we show how different priors affect the ϵ parameters. From left to right we increasingly widen the priors on the vote parameters. The histogram presents the posterior marginal density whereas the red line represents the prior on that specific parameter. The second column shows the results used in the paper. What we can see is that as we make the prior on the vote parameters more diffuse, the density of the ϵ parameters become wider. However, when looking at the scale, we see that this effect is very small.

In Figures 15 and 16 we show the effect on the vote parameters. As we specify more diffuse priors, the posteriors also become more diffuse but the effect is not too large once again. At the same time we see that the effect of changing the prior on the ϵ parameters is negligible.

In Figure 17 we see the impact on the ideal points of Blanchflower and Sentance. We show these two voters because they are on the opposite ends of the dove-hawk dimension. Moreover, they have a wide marginal posterior density to start with. What we can see is that these ideal points are fairly stable. The change in the median in both cases is very small. This emphasizes that the sensitivity of the ideal point to the priors on the vote parameters and the ϵ parameters is very mild.

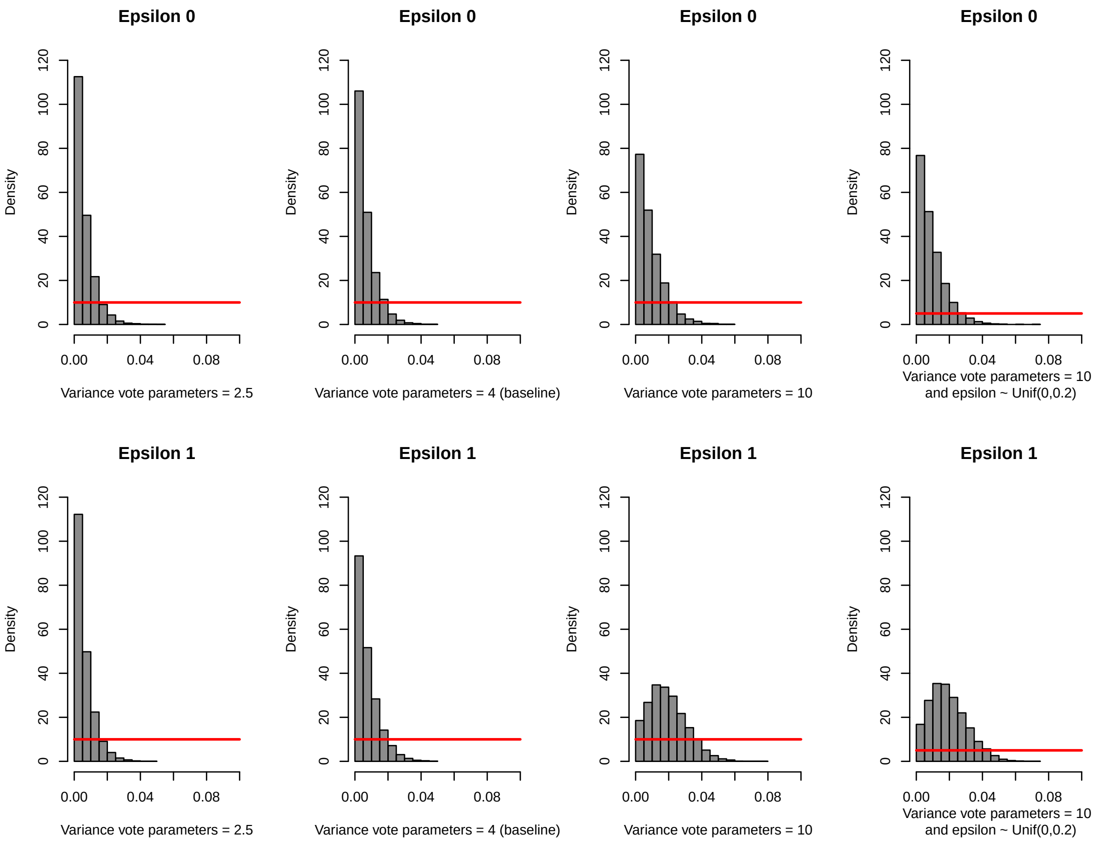

Figure 14: Here we present marginal posterior densities (histogram) of the ϵ parameters under different prior specifications (the red line). From left to right we make the priors more diffuse. The left most graph uses a prior on the vote parameters with a smaller variance i.e. less diffuse as we used in the paper. The second graph presents the results in the paper (baseline specification), the third graph comes from a simulation with a more diffuse prior on the vote parameters (variance of 10) and the right most graph also increases the prior on the ϵ parameters. The top row presents the results for ϵ_0, the bottom row for ϵ_0.

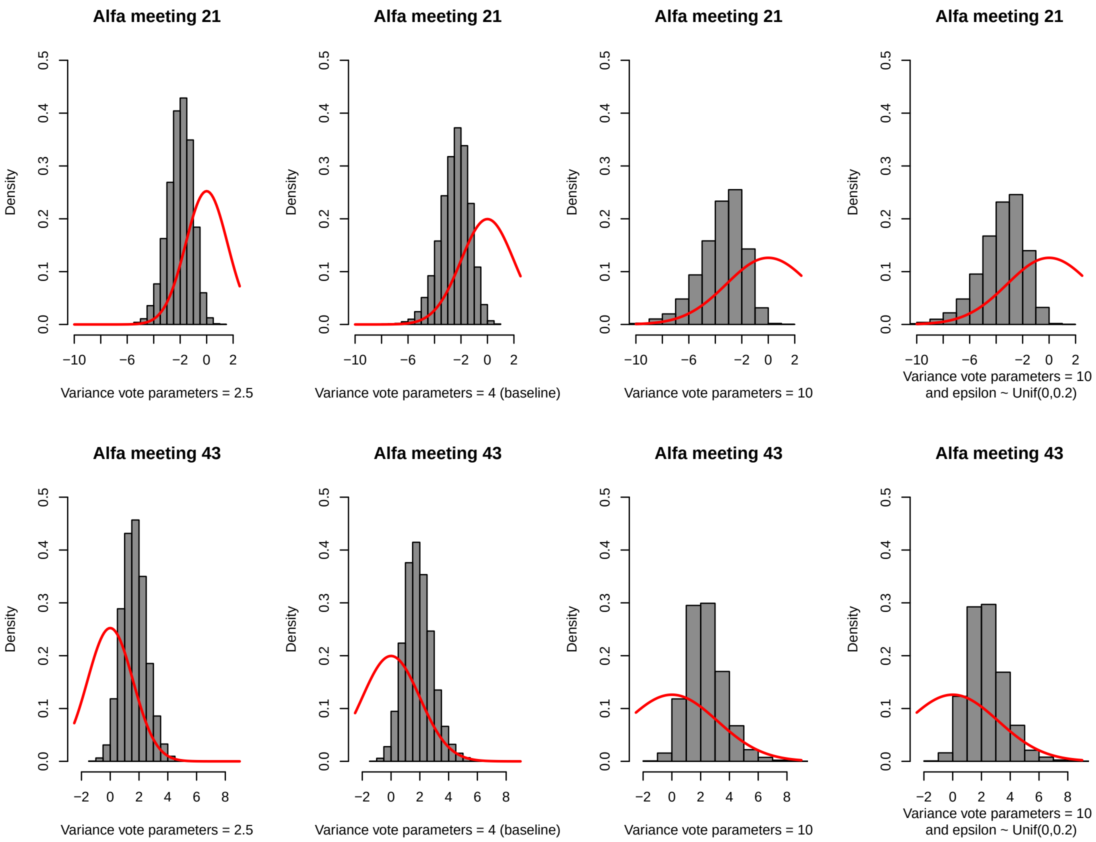

Figure 15: Here we present marginal posterior densities (histogram) of two α parameters under different prior specifications (the red line). From left to right we make the priors more diffuse. The left most graph uses a prior on the vote parameters with a smaller variance i.e. less diffuse as we used in the paper. The second graph presents the results in the paper (baseline specification), the third graph comes from a simulation with a more diffuse prior on the vote parameters (variance of 10) and the right most graph also increases the prior on the α parameters. The top row presents the results for α_{21}, the bottom row for α_{43}.

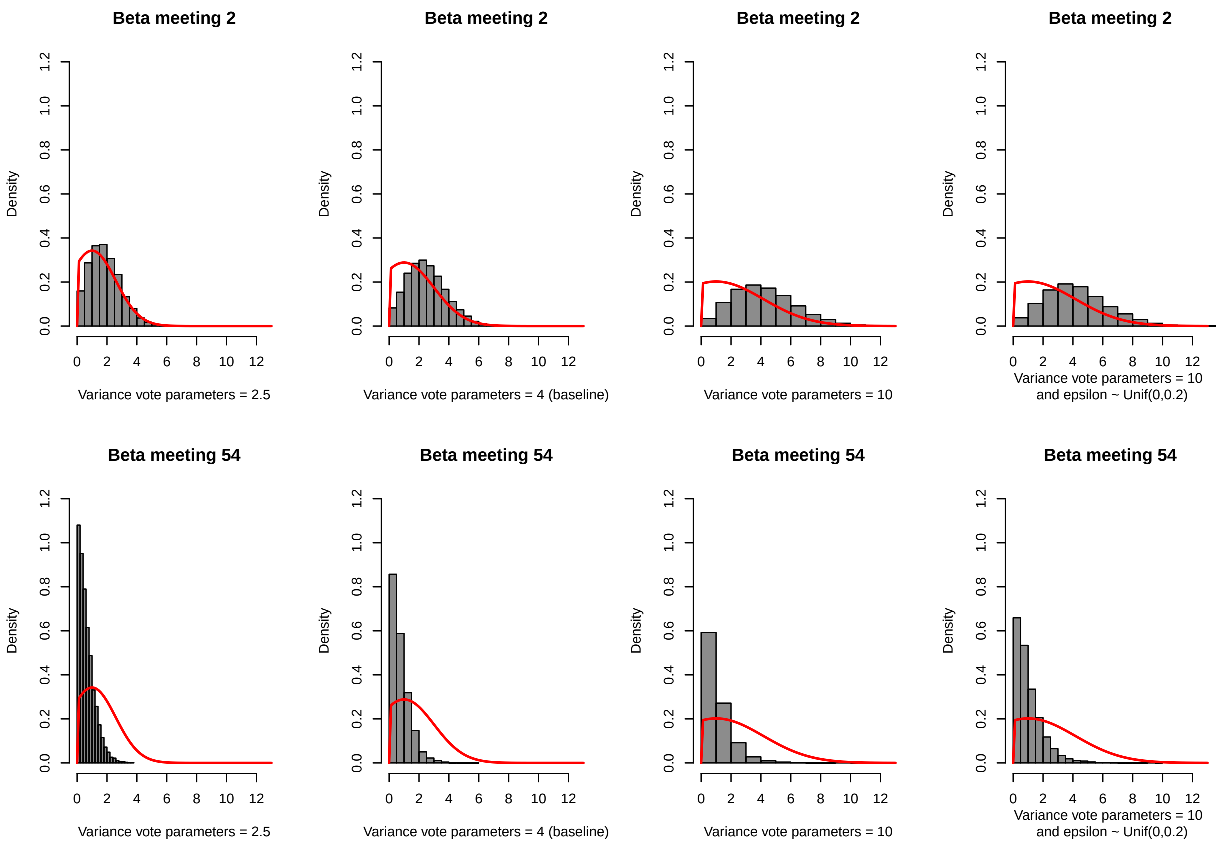

Figure 16: Here we present marginal posterior densities (histogram) of two β parameters under different prior specifications (the red line). From left to right we make the priors more diffuse. The left most graph uses a prior on the vote parameters with a smaller variance i.e. less diffuse as we used in the paper. The second graph presents the results in the paper (baseline specification), the third graph comes from a simulation with a more diffuse prior on the vote parameters (variance of 10) and the right most graph also increases the prior on the β parameters. The top row presents the results for β_2, the bottom row for β_{54}.

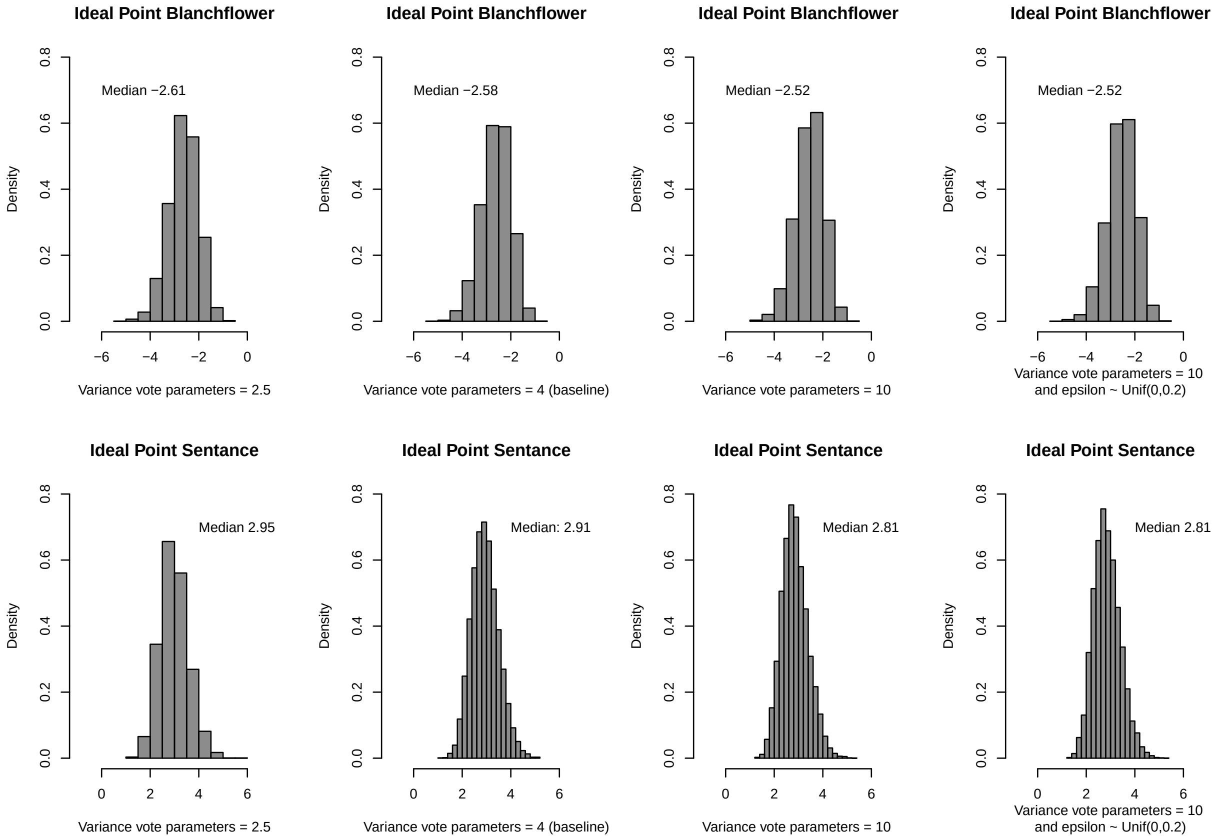

Figure 17: Here we present marginal posterior densities (histogram) of two ideal points (Blanchflower and Sentance under different prior specifications (the red line). From left to right we make the priors more diffuse. The left most graph uses a prior on the vote parameters with a smaller variance i.e. less diffuse as we used in the paper. The second graph presents the results in the paper (baseline specification), the third graph comes from a simulation with a more diffuse prior on the vote parameters (variance of 10) and the right most graph also increases the prior on the ideal points.The top row presents the results for Blanchflower, the bottom row for Sentance.

References

Bafumi, J., A. Gelman, D. K. Park, and N. Kaplan (2005): "Practical Issues in Implementing and Understanding Bayesian Ideal Point Estimation," *Political Analysis*, 13, 171–187.

Besley, T., N. Meads, and P. Surico (2008): "Insiders versus Outsiders in Monetary Policymaking," *American Economic Review*, 98(2), 218–23.

Black, D. (1948): "On the Rationale of Group Decision-making," *Journal of Political Economy*, 56(1).

Blinder, A. (2004): *The Quiet Revolution: Central Banking Goes Modern*. Yale University Press, New Haven, CT.

Blinder, A. (2007): "Monetary policy by committee: Why and how?," *European Journal of Political Economy*, 23(1), 106–123.

Chappell, Henry W, J., T. M. Havrilesky, and R. R. McGregor (1993): "Partisan Monetary Policies: Presidential Influence through the Power of Appointment," *The Quarterly Journal of Economics*, 108(1), 185–218.

Clinton, J., S. Jackman, and D. Rivers (2004): "The Statistical Analysis of Roll Call Data," *American Political Science Review*, 98(02), 355–370.

Clinton, J. D., and S. Jackman (2009): "To Simulate or NOMINATE?," *Legislative Studies Quarterly*, 34(4), 593–621.

Clinton, J. D., and A. Meirowitz (2004): "Testing explanations of strategic voting in legislatures: A reexamination of the compromise of 1790," *American Journal of Political Science*, 48(4), 675–689.

Enelow, J. M., and M. J. Hinich (1984): *The spatial theory of voting : an introduction*. Cambridge University Press, Cambridge [Cambridgeshire] ; New York :.

Gelman, A. (2004): "Exploratory data analysis for Complex Models," *Journal of Computational and Graphical Statistics*, (13), 755–779.

Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin (2003): *Bayesian Data Analysis, Second Edition (Chapman & Hall/CRC Texts in Statistical Science)*. Chapman and Hall/CRC, 2 edn.

Gelman, A., Y. Goegebeur, F. Tuerlinckx, and I. Van Mechelen (2000): "Diagnostic checks for discrete data regression models using posterior predictive simulations," *Journal of the Royal Statistical Society Series C*, 49(2), 247–268.

Gelman, A., and C. Shalizi (2013): "Philosophy and the practice of Bayesian statistics," *British Journal of Mathematical and Statistical Psychology*, 66, 8–38.

Gerlach-Kristen, P. (2003): "Insiders and Outsiders at the Bank of England," *Central Banking*, 14(1), 96–102.

Geweke, J. (1992): "Evaluating the accuracy of sampling-based approaches to calculating posterior moments.," *Bayesian Statistics 4, edited by Bernardo, Berger, Dawid and Smith.*

Hansen, S., M. McMahon, and C. V. Rivera (2012): "How Experts Decide: Identifying Preferences versus Signals from Policy Decisions," working paper.

Harris, M., P. Levine, and C. Spencer (2011): "A decade of dissent: explaining the dissent voting behavior of Bank of England MPC members," *Public Choice*, 146(3), 413–442.

Harris, M. N., and C. Spencer (2009): "The Policy Choices and Reaction Functions of Bank of England MPC Members," *Southern Economic Journal*, 76(2), 482–499.

Heckman, J. J., and J. M. Snyder (1997): "Linear Probability Models of the Demand for Attributes with an Empirical Application to Estimating the Preferences of Legislators," *RAND Journal of Economics*, 28(0), s142–s189.

Henry, M., and I. Mourifié (2013): "Euclidean Revealed Preferences: Testing the Spatial Voting Model," *Journal of Applied Econometrics*, 28(3).

Hix, S., B. Hoyland, and N. Vivyan (2010): "From doves to hawks: A spatial analysis of voting in the Monetary Policy Committee of the Bank of England," *European Journal of Political Research.*

Iaryczower, M., G. Lewis, and M. Shum (2013): "To elect or to appoint? Bias, information, and responsiveness of bureaucrats and politicians," *Journal of Public Economics*, 97(C), 230–244.

Iaryczower, M., and M. Shum (2012): "The Value of Information in the Court: Get It Right, Keep It Tight," *American Economic Review*, 102(1), 202–37.

Jackman, S. (2009): *Bayesian Analysis for the Social Sciences.* Wiley Series in Probability and Statistics, John Wiley & Sons.

Lauderdale, B. (2010): "Unpredictable voters in ideal point estimation.," *Political Analysis*, 2(18), 151–171.

Liu, C. (2004): "Robit Regression: A simple Robust Alternative to Logistic and Probit regression," *Applied Bayesian Modeling and Causal Inference from an Incomplete-Data Perspective, ed. A. Gelman and X.L. Meng*, Ch 21.

Martin, A. D., and K. M. Quinn (2002): "Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the U.S. Supreme Court, 19531999," *Political Analysis*, 10(2), 134–153.

Martin, A. D., K. M. Quinn, and J. H. Park (2011): "MCMCpack: Markov Chain Monte Carlo in R," *Journal of Statistical Software*, 42(9), 22.

Neal, R. (2003): "Slice Sampling," *Annals of Statistics*, 31, 705–767.

Plummer, M., N. Best, K. Cowles, and K. Vines (2006): "CODA: Convergence Diagnosis and Output Analysis for MCMC," *R News*, 6(1), 7–11.

Pollard, P. (2004): "Monetary policy-making around the world," *Presentation, Federal Reserve Bank of St.Louis.*

Pregibon, D. (1982): "Resistant fits for some commonly used logistic models with medical applications," *Biometrics*, 38, 485–498.

Raftery, A., and M. Lewis, S (1992): "One long run with diagnostics: Implementation strategies for Markov chain Monte carlo.," *Statistical Science*, 7, 493–497.

Riboni, A., and F. J. Ruge-Murcia (2008): "Preference Heterogeneity in Monetary Policy Committees," *International Journal of Central Banking*, 4(1), 213–233.

Rivers, D. (2003): "Identification of Multidimensional Spatial Voting Models," *typescript, Stanford University.*