How Experts Decide :
Identifying Preferences versus Signals from Policy Decisions

Stephen Hansen and Michael McMahon

No 963

WARWICK ECONOMIC RESEARCH PAPERS

DEPARTMENT OF ECONOMICS

THE UNIVERSITY OF
WARWICK

How Experts Decide: Identifying Preferences versus Signals from Policy Decisions*

Stephen Hansen
Universitat Pompeu Fabra
Barcelona GSE
stephen.hansen@upf.edu

Michael McMahon
University of Warwick
CEP, LSE
m.mcmahon@warwick.ac.uk

May 11, 2011

Abstract

A large theoretical literature assumes that experts differ in terms of preferences and the distribution of their private signals, but the empirical literature to date has not separately identified them. This paper proposes a novel way of doing so by relating the probability a member chooses a particular policy decision to the prior belief that it is correct. We then apply this methodology to study differences between internal and external members on the Bank of England's Monetary Policy Committee. Using a variety of proxies for the prior, we provide evidence that they differ significantly on both dimensions.

Keywords: Bayesian decision making, Committees, Monetary policy

JEL Codes: D81, D82, E52

*This paper extends part of the research previously circulated under the title "What Do Outside Experts Bring To A Committee? Evidence From The Bank of England". We would like to thank Tim Besley, Pablo Casas, Fabio Canova, Francesco Caselli, Antonio Ciccone, Greg Crawford, Thomas Cunningham, Nathan Foley-Fisher, Francesco Giavazzi, Charles Goodhart, Clare Leaver, Gilat Levy, Massimo Motta, Andrew Oswald, Morten Ravn, Karl Schlag, Kevin Sheedy, Jón Steinsson, and Thijs van Rens for insightful comments. We have also benefited from the suggestions of seminar participants at the Australia National University, the Bank of England, London School of Economics, Monash University, Queen Mary University, University College Dublin, Universitat Pompeu Fabra, University of Strathclyde, and University of Warwick. We are also very grateful to Grainne Gilmore and her colleagues at the Times of London Newspaper for making available the data on the Times MPC, and to 26 City economists for taking the time to complete our survey.

1 Introduction

Consider an expert who must take one of two decisions whose consequences are uncertain. Two individual characteristics can potentially guide her decision. First, there are her preferences over the outcomes associated with each decision. Second, in an idea going back to Condorcet (1785), she might receive private signals that allow her to update her beliefs about the probabilities of the different outcomes being realized. Many papers in the theory literature, especially in the area of committees,[1] explicitly model both dimensions and allow each to influence decision making.

One example of a situation in which these considerations are relevant is monetary policy, where the exact effect of implementing a given interest rate is, by its nature, uncertain and decision making power is increasingly being transferred to independent committees of experts.[2] Different monetary policy experts presumably differ both in their utility from different inflation outcomes—frequently described in terms of "hawks" and "doves"—as well as in their assessment of the likelihood that inflationary shocks have hit the economy. While both the preference and signal dimensions are potentially important in general for explaining expert voting behavior, we are not aware of empirical work that separately identifies heterogeneity in each dimension separately from data on individual policy choices. This paper attempts to do so in the context of monetary policy making at the Bank of England, although, as we explain below, our results could potentially be extended to others.

One issue that separately controlling for these dimensions can help resolve is how the policy choices of Bank executives differ from those of outside experts whose only responsibility is to serve on the committee. Some central banks, like the Federal Reserve Bank in the US and the Riksbank in Sweden, use committees composed solely of the former, while the Bank of England's Monetary Policy Committee (MPC) is made up of five Bank executives (so-called *internal* members) and four outside experts (so-called *external* members).[3] Whether internal or external members differ in terms of preferences or information clearly has different implications for their inclusion. Several papers in the literature (Gerlach-Kristen 2003, Bhattacharjee and Holly 2005, Spencer 2006, Besley, Meads, and Surico 2008, Harris and Spencer 2008) have found internal-external differences within the MPC, but they generally attribute these differences to externals being more dovish than internals. Controlling for signal distributions allows us to establish whether such a difference really exists and to measure its magnitude, as well as to ex-

[1]See Gerling, Gruner, Kiel, and Schulte (2005) for a survey.

[2]90 percent of 88 central banks surveyed by Pollard (2004) use committees to decide interest rates.

[3]Some outside experts are appointed directly to executive positions and therefore become internal members.

plore informational differences.

After providing some detailed background information on the MPC (section 2), we explicitly model a monetary policy expert's interest rate choice as a Bayesian decision problem (section 3). In the model, an expert must choose one of two interest rates—one higher than another—and there is a prior probability that an inflationary shock which requires the higher interest rate has hit the economy. Members can differ both in terms how much evidence they need that this shock has hit before they choose the high rate, which one can interpret as capturing preference differences,[4] as well as the precision of a private signal they receive that is correlated with the magnitude of the shock. Whether or not such a signal literally represents more information is a matter we do not seek to resolve; the point is to model any systematic differences in how members arrive at their posterior beliefs given the prior.

We then show that differences in preferences and signal distributions are separately identifiable using information on the prior. The logic is as follows: if a member requires more evidence for the high shock to vote high (i.e. he is more dovish than another), then he will tend to vote for lower rates no matter what the value of the prior. In contrast, if a member has a more precise private signals than another, then he is less influenced by the prior and will tend to vote for the high rate more often when the prior favors the low state and for the low rate more often when the prior favors the high state; informally speaking, he is less likely to "follow the crowd".

In order to take our model to the data, we need to construct a convincing proxy of the prior in each period (section 4). We combine three sources of data: a monthly Reuters survey of financial economists working in London; a monthly meeting of an informal MPC assembled from monetary policy experts by the Times of London newspaper; and options prices. We then use common factor analysis to extract the principal component driving the variation in all three sources and use it as our measure of the public prior.

Our empirical analysis begins by showing that the frequency with which internal members vote high relative to externals varies significantly for different values of the prior. To uncover why this is the case, we structurally estimate the parameters in internal and external members' decision rules generated by our model. We find that (1) significant differences exist in terms of preference conflict and signal distributions and that (2) allowing for variation in signal distributions significantly better fits the voting data. This means that the hawk-dove distinction between internals and externals discussed in the

[4]In fact we assume, as does the Bank of England, that all members have the same objective function: to hit the inflation target. We present an explicit model of the macroeconomy in order to show how differences in the evidence threshold nevertheless arise from different (fixed) beliefs about the structural parameters of the economy. We will term this a member's "philosophy" rather than preferences, but for informal purposes one can think of them as equivalent.

literature is robust to controlling for signal heterogeneity, although we argue that we more accurately measure the magnitude of the difference. The more novel result is that fundamental differences exist in how members form their posterior beliefs. Appropriately modelling and controlling for differences in the distributions of private signals is important to more accurately fit the data than is possible by considering preference differences alone.

These results open up two potentially fruitful areas for research. First, our results show that the behavior of MPC members is richer than previously understood. While much of the discussion surrounding monetary policy makers appeals to the hawk-dove dichotomy, our paper shows there is a wholly different dimension that is hardly ever discussed and is nevertheless important for explaining policy making. Looking deeper into exactly why external members form (or behave as if they form) posterior beliefs that are less influenced by the public's beliefs than internals is the next logical step in terms of the monetary literature. In terms of the broader economics literature, our approach—with only slight modifications—could be used to estimate preferences and signal distributions in any context where one has individual decision data along with a credible measure of public beliefs about what the right decision is. Examples might include stock analysts' decisions to downgrade investment recommendations following the arrival of a common negative asset shock and competition commissions' decisions about whether to punish anti-competitive behavior following an investigation. Although these broader issues are beyond the scope of this paper, we hope it can stimulate thinking about them.

2 The Monetary Policy Committee

Until 1997 the Chancellor of the Exchequer (the government minister in charge of the UK Treasury) had sole responsibility for setting interest rates in the UK. Immediately after the UK general election in 1997, the new Labour government established an independent committee of experts, the MPC, for setting interest rates. Its remit, as defined in the Bank of England Act (1998) is to "maintain price stability, and subject to that, to support the economic policy of Her Majesty's government, including its objectives for growth and employment." In practice, the committee seeks to achieve a target inflation rate of 2%, based on the Consumer Price Index.[5] If inflation is greater than 3% or less than 1%, the Governor of the Bank of England must write an open letter to the Chancellor explaining why. The inflation target is symmetric; missing the target in either direction is treated with equal concern.

The MPC first convened on 6 June 1997, and has met every month since. Throughout

[5]This target changed from the RPIX to the CPI measure of inflation in January 2004, with a reduction in the inflation target from 2.5% to 2%.

the paper we analyze the MPC voting records between June 1997 and March 2009 which is when the main focus of the decision (temporarily) shifted to asset purchase decisions related to quantitative easing.[6] The voting records indicate both the proposed interest rate decision (such as +50 basis points), as well as the alternative preference for those who do back the proposal (such as +25bps).[7]

The MPC has nine members; five of these are internal members serving as executives of the Bank of England: the Governor, two Deputy Governors, the Chief Economist, and the Executive Director for Markets. The Chancellor also appoints four external members (subject to approval from the Treasury Select Committee) from outside the Bank. There are no restrictions on who can serve as an external member; external members have come from many different backgrounds.[8] Bar the governors, all members serve three year terms; the governors serve five year terms. When members' terms end, they can either be replaced or re-appointed. Table 1 lists the members that served on the MPC during our sample and whether they are internal or external. Our sample contains a total of 13 internal and 14 external members.

Table 1: MPC Members

Internal Member	Tenure	External Member	Tenure
Howard Davies	06/97 - 07/97	Willem Buiter	06/97 - 05/00
Edward George	06/97 - 06/03	Charles Goodhart	06/97 - 05/00
Mervyn King	06/97 - 03/09	DeAnne Julius	11/97 - 05/01
Ian Plenderleith	06/97 - 05/02	Alan Budd	12/97 - 05/99
David Clementi	11/97 - 08/02	Sushil Wadhwani	06/99 - 05/02
John Vickers	06/98 - 11/00	Christopher Allsopp	06/00 - 05/03
Charles Bean	10/00 - 03/09	Stephen Nickell	06/00 - 05/06
Paul Tucker	06/02 - 03/09	Kate Barker	06/01 - 03/09
Andrew Large	11/02 - 01/06	Marian Bell	06/02 - 06/05
Rachel Lomax	07/03 - 06/08	Richard Lambert	06/03 - 03/06
John Gieve	01/06 - 03/09	David Walton	07/05 - 06/06
Spencer Dale	07/08 - 03/09	Tim Besley	06/06 - 03/09
Paul Fisher	03/09 - 03/09	David Blanchflower	06/06 - 03/09
		Andrew Sentance	10/06 - 03/09

[6]These data are available from the Bank of England (2010a). We use each regular MPC meetings in this period but we drop from the dataset the (unanimous) emergency meeting held after 9/11.

[7]Before June 1998 there is information about whether members preferred higher or lower interest rates compared with the decision, but not about their actual preferred rate. In these cases, we treat a member's vote as either 25 basis points higher or lower than the decision, in the direction of disagreement. Given how we use the voting data, discussed below, this assumption has no implications for our analysis.

[8]Of course there are restrictions, to avoid conflicts of interest, on what the members can do while serving on the committee.

Each member is independent in the sense that they do not represent any interest group or faction. The Bank encourages members to simply determine the rate of interest that they feel is most likely to achieve the inflation target,[9] and majority vote determines the outcome. As such, the observed votes should reflect members' genuine policy preferences. Consistent with its one-person one-vote philosophy, the MPC displays substantial dissent. 64% of the 142 meetings in the sample have at least one deviation from the committee majority. Figure 1 shows the level of interest rates that the MPC has implemented, the votes of each member around this, and highlights the periods of interest rate loosening.[10] Figure 2 shows how many votes were cast in each meeting in opposition to the final decision. Within the set of non-unanimous meetings, 5-4 and 6-3 decisions are not uncommon.

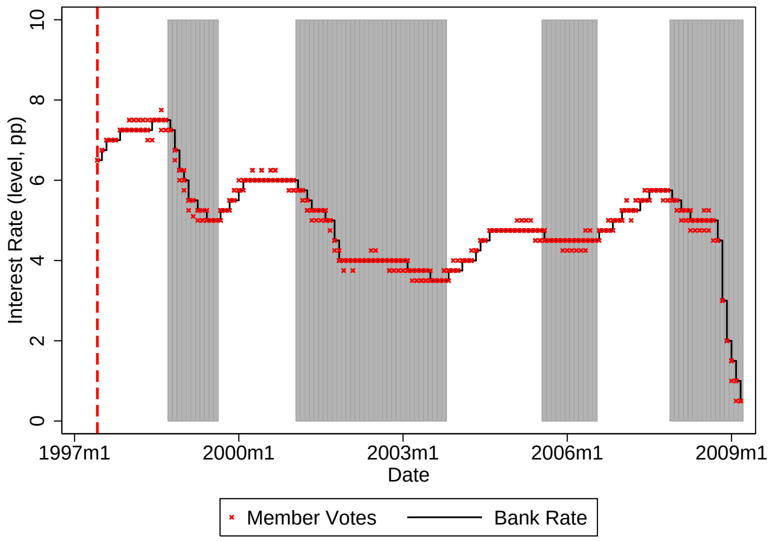


Figure 1: Votes and Decisions

The MPC meets on the first Wednesday and Thursday of each month. In the month between meetings, members receive numerous briefings from Bank staff and regular updates of economic indicators. On the Friday before MPC meetings, members gather for a meeting in which they are given the latest analysis of economic and business trends. Then

[9] According to the Bank of England (2010b)

> Each member of the MPC has expertise in the field of economics and monetary policy. Members are not chosen to represent individual groups or areas. They are independent. Each member of the Committee has a vote to set interest rates at the level they believe is consistent with meeting the inflation target. The MPC's decision is made on the basis of one-person, one vote. It is not based on a consensus of opinion. It reflects the votes of each individual member of the Committee.

[10] The loosening cycle is defined as the period from the first cut in interest rates until the next increase in interest rate.

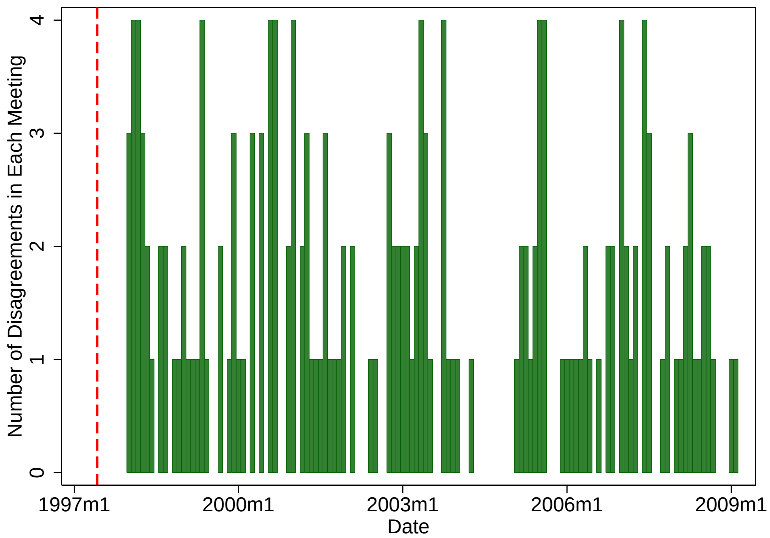


Figure 2: Deviations

on Wednesday members discuss their views on several issues. The discussion continues on Thursday morning, when each member is given some time to summarize his or her views to the rest of the MPC and to suggest what vote they favor (although as Lambert (2006) notes they can, if they wish, wait to hear the others views before committing to a vote). This process begins with the Deputy Governor for monetary policy and concludes with the Governor, but the order for the others is not fixed. To formally conclude the meeting, the Governor proposes an interest rate decision that he believes will command a majority. Each member then chooses whether to agree with the Governor's proposal, or dissent and state an alternative interest rate. The MPC decision is announced at noon, and two weeks after each meeting, attributed members' votes are published as part of otherwise unattributed minutes.

3 Model

We now lay out a parsimonious model that captures the basic elements of MPC members' decision making that, by highlighting the drivers of their behavior, will serve as the basis of our empirical analysis. We assume that the MPC, in each period t, must decide $v_t \in \{0, 1\}$, where 0 represents the lower of two possible rate changes and 1 the higher.[11] The restriction that a vote must take one of two values is not as restrictive as it might first appear since there are three unique votes in only 7 of the 142 meetings in our sample

[11] We do not attempt to model the two rates over which voting occurs; Riboni and Ruge-Murcia (2010) has recently studied this issue.

and in no meeting are there four or more unique votes.

We model inflation in period t as $\pi_t \sim F[\,\pi_t \mid v_t\,]$ where $\mathbb{E}[\,\pi_t\,] = \theta + \omega_t - v_t$ and $V(\pi_t) < \infty$.[12] The objective function of MPC members is to pick the interest rate that they view as consistent with hitting an inflation target π^*; specifically, in period t member i chooses a vote to solve

$$\max_{v_{it} \in \{0,1\}} \mathbb{E}\big[-(\pi_t - \pi^*)^2\,\big] \tag{1}$$

where member i treats inflation as coming from the distribution $\pi_t \sim F[\,\pi_t \mid v_{it}\,]$.

The first driver of inflation is $\omega_t \in \{0, 1\}$. This represents economic conditions relevant to inflation at time t such as the magnitude of a demand shock or the output gap of the economy. We call $\omega_t = 1$ the "high inflation state" and $\omega_t = 0$ the "low inflation state". These states capture that in some times the economy faces more inflationary pressures than others. Member i has belief $\widehat{w}_{it}$ on the expected value of ω_t.

The second driver of inflation is $\theta \in \big[\underline{\theta}, \overline{\theta}\big]$. This represents structural sources of inflation such as the natural rate of interest or monetary policy transmission mechanism. Member i has belief $\widehat{\theta}_i$ on θ. We do not allow this belief to vary with time; one can imagine that information about θ does not arrive every period and so members' beliefs on it evolve slowly, if at all. It therefore represents a fixed source of disagreement among members and can capture the hawk-dove distinction frequently made in the monetary literature; ceteris paribus, a member with a higher $\widehat{\theta}_i$ believes future inflation to be more likely than one for whom it is lower. We wish to remain agnostic about the specific dimension of disagreement that $\widehat{\theta}_i$ represents and therefore refer to it as a members' *philosophy*.

Given the objective function, the goal is to choose the vote that will make expected inflation closest to the target π^*. This implies choosing $v_{it} = 1$ if and only if [13]

$$\widehat{\omega}_{it} \geq 1 - \Big(\widehat{\theta}_i - \pi^* + 0.5\Big) \equiv 1 - \theta_i. \tag{2}$$

In words, member i votes for the high rate only when his belief that the economy is in the high inflation state is sufficiently large. Rather than consider the quantity $\widehat{\theta}_i$, we consider its linear transformation θ_i. Notice that heterogeneity in θ_i would also arise

[12] We think of π_t as the distribution of inflation outcomes at the two year horizon since that is the horizon at which committee members are told to direct monetary policy.

[13]
$$-\mathbb{E}\Big[(\pi_t - \pi^*)^2\Big] = -\mathbb{E}\big[\,\pi_t^2\,\big] + 2\pi^*\mathbb{E}[\,\pi_t\,] - (\pi^*)^2 = -V[\pi_t] - (\mathbb{E}[\,\pi_t\,] - \pi^*)^2.$$

So $v_{it} = 1$ is chosen if and only if

$$-\Big(\widehat{\theta}_i - 1 + \widehat{\omega}_{it} - \pi^*\Big)^2 > -\Big(\widehat{\theta}_i + \widehat{\omega}_{it} - \pi^*\Big)^2$$

if inflation targets are person-specific, which would be the case if members literally had different preferences. Since every member used in our sample sometimes votes for high and sometimes for low rates, we impose the restriction $\theta_i \in (0, 1)$. Every member votes for the high rate when they are certain that the high state has hit and for the low rate when they are certain the low state has hit; however, they differ in how high the belief $\widehat{\omega}_{it}$ must be in order to vote high. This is how philosophical differences manifest themselves in our model.

The model is a reduced form version of at least two commonly-used models. First, one can show that our decision rule is equivalent to one that arises in a standard New Keynesian model of the macroeconomy (Clarida, Galí, and Gertler 1999, Galí 2008). Appendix A.3 presents this model and shows its equivalence to the above in terms of policy choices. Second, one can view θ_i as a preference parameter over type I and type II errors along the lines of Feddersen and Pesendorfer (1998) and again derive a behavioral equivalence to the above model.

While no information arrives about θ that allows members to update their beliefs $\widehat{\theta}_i$, we allow them to form their beliefs on the inflationary state according to Bayes' Rule in the standard way for decision problems under uncertainty. They all begin with a prior belief $q_t = \Pr[\,\omega_t = 1\,]$ that the economy is in the high inflation state that can be interpreted as conventional wisdom about inflationary pressures. In addition members receive private signals $s_{it} \sim N(\omega_t, \sigma_i^2)$ that allow them to update their prior to arrive at belief $\widehat{\omega}(s_{it})$. Heterogeneity in σ_i represents heterogeneity in the distribution of private signals. In principle, one could allow σ_i to evolve with time; we will discuss the restriction of time-invariance in section 6 below.

As a behavioral matter, a member with a lower σ_i will put more weight on his signal and less weight on the prior q_t. To see this in a simple way, note that a member's posterior belief has the following relationship with his signal:

$$\ln\left(\frac{\widehat{\omega}(s_{it})}{1-\widehat{\omega}(s_{it})}\right) = \ln\left(\frac{q_t}{1-q_t}\right) + \ln\left(\frac{f_1}{f_0}\right)$$

where $f_1 \sim (1, \sigma_i^2)$ is the distribution of s_{it} conditional on $\omega_t = 1$ and $f_0 \sim (0, \sigma_i^2)$ is the distribution of s_{it} conditional on $\omega_t = 0$. Further simplifying, using the assumption that s_{it} is normally-distributed, yields

$$\ln\left(\frac{\widehat{\omega}(s_{it})}{1-\widehat{\omega}(s_{it})}\right) = \ln\left(\frac{q_t}{1-q_t}\right) + \frac{2s_{it}-1}{2\sigma_i^2}. \tag{3}$$

In other words, if a member receives uninformative signals ($\sigma_i = \infty$), his belief will be determined only by the prior ($\widehat{\omega}(s_{it}) = q_t$ for all s_{it}). As the signal becomes more informative (σ_i declines), he puts more weight on the signal relative to the public prior.

The committee voting literature take seriously the idea that a member with a lower σ_i has more private information. While this may be the case on the MPC, other interpretations are possible. For example, a member with a lower σ_i may put more weight on their own forecast relative to standard forecasts.[14] We do not wish to take a stand on the underlying source of this information difference, but rather to simply show that it has contributed to voting differences.

We assume that member i votes high whenever $\widehat{\omega}_{it}(s_{it}) \geq 1-\theta_i$; that is, he maximizes his expected utility conditional on s_{it} and q_t. In other words, we model members' votes as individual decision problems rather than as equilibrium outcomes of a strategic voting game. The primary strategic concern in voting games is that members must use the information conveyed in the configurations of other members' votes in which they are pivotal to the decision to update their beliefs. These considerations are most relevant under super-majority rules; since the MPC operates under majority rule we do not view the additional complication of equilibrium play as having enough empirical relevance to justify its inclusion.[15]

We make no attempt to model the effect of pre-vote communication on beliefs $\widehat{\omega}_{it}$. The theory literature in this area is itself quite nascent, so the paper leaves the issue aside. One might imagine that the ability to reveal private information during discussion should reduce heterogeneity in posterior beliefs, not increase them. If this is the case, not accounting for communication should bias us against finding differences in private information, not towards finding such differences.

Our first result pins down member i's period t decision rule as a function of his signal.

Proposition 1

$$v(s_{it}) = \begin{cases} 1 & \text{if } s_{it} \geq s^*_{it} \\ 0 & \text{if } s_{it} < s^*_{it} \end{cases}$$

where

$$s^*_{it} = \frac{1}{2} - \sigma_i^2 \ln\left(\frac{\theta_i}{1-\theta_i}\frac{q_t}{1-q_t}\right). \tag{4}$$

Because the normal distribution satisfies the monotone likelihood ratio property, higher realizations of s_{it} provide more evidence of inflation, so members adopt a cutoff rule (with

[14]This interpretation is preferred by Blinder (2007), who points out that members do not differ in the amount of information they have, but in their views about the implications of the same data for inflationary pressures.

[15]A recent experiment reported in Goeree and Yariv (2010) shows that under majority rule members vote with the signal in over 90% of cases, a percentage that drops markedly as decision rules become biased against one of the voting outcomes.

s^*_{it} as the cutoff).

The principle value of the model is that it allows one to separately identify the effect of differences in philosophy and differences in private information on observed voting behavior, which we now discuss. Suppose that θ_i increases so that member i requires less evidence of the inflationary shock to vote for the high rate. The associated change in s^*_{it} is

$$\frac{\partial s^*_{it}}{\partial \theta_i} = -\frac{\sigma_i^2}{\theta_i(1-\theta_i)} < 0 \tag{5}$$

which immediately implies that the probability a member votes more often is decreasing for all values of the prior.

Now suppose that σ_i^2 increases. The associated change in s^*_{it} is

$$\frac{\partial s^*_{it}}{\partial \sigma_i^2} = -\ln\left(\frac{\theta_i}{1-\theta_i}\frac{q_t}{1-q_t}\right) \lesseqqgtr 0 \iff q_t \gtreqless 1-\theta_i. \tag{6}$$

When q_t favors the high state, an increase in σ_i^2 reduces the threshold that the signal must reach for the member to vote high, increasing the range of values of the signal for which the member votes high. Exactly the opposite logic works when q_t is low. Intuitively speaking, when σ_i^2 increases, the member becomes more likely to "follow the crowd", and moves the threshold towards accepting whichever rate the prior does.

Although the derivative in (6) does not immediately imply anything about changes in the probability of voting high since the distribution of signals also depends on σ_i, it still provides the intuition for the following result.

Proposition 2 *When θ_i increases, the probability that member i chooses $v_{it} = 1$ increases for all q_t; when σ_i^2 increases, the probability that member i chooses $v_{it} = 1$ increases if and only if $q_t \geq q^* \in (0,1)$.*

In short, if two members have the same distribution of private signals but differ in terms of philosophy, then the more hawkish member votes for systematically higher rates for all values of the prior. In contrast, if two members have the same philosophy but differ in the signal distribution, the member with a lower σ will vote for systematically higher (lower) rates than the member with less expertise when prior beliefs are low (high).

Figure 3 gives a graphical interpretation of proposition 2. For all combinations of θ_i and σ_i, increases in the prior q_t increase the probability of voting high. An increase in θ_i induces a shift in the predicted probability, while an increase in σ_i induces a rotation. The remainder of the paper will exploit the intuition behind proposition 2 to examine differences in private information between internal and externals.

(a) Change in Philosophy (b) Change in Signal Distribution

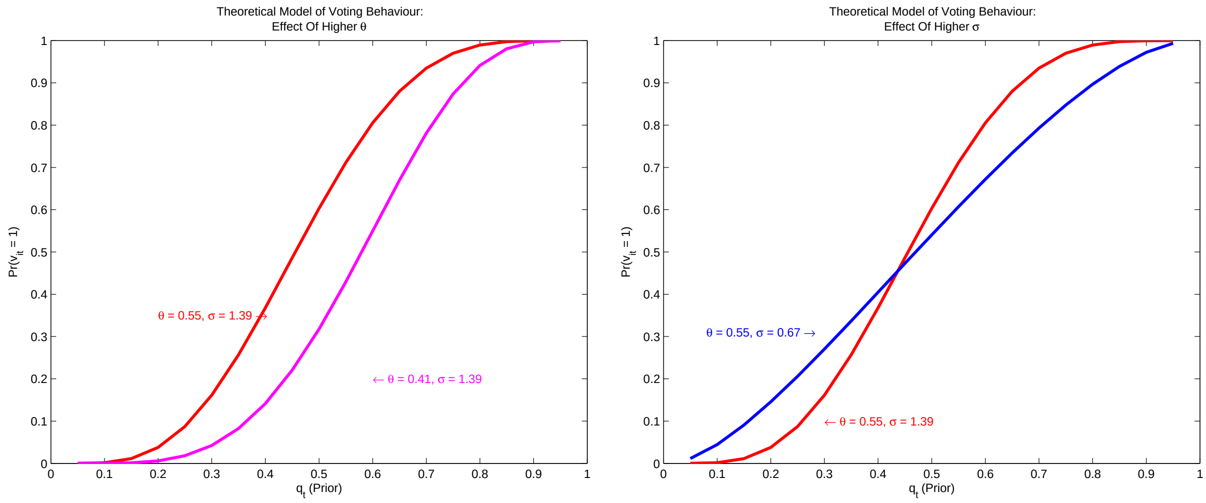


Figure 3: Distinguishing Information and Preferences

4 Constructing Public Beliefs

The main empirical challenge we confront is that we do not observe q_t, but there are several observable sources of data on public beliefs that provide reasonable proxies for q_t. The three different measures that we use are:

1. A survey of market economists conducted by Reuters (Reuters q $\equiv \widehat{q}_t^R$).
2. The voting record of the Times MPC (Times q $\equiv \widehat{q}_t^T$).[16]
3. Implied probabilities from short sterling option prices (Market q $\equiv \widehat{q}_t^M$).

This section describes the construction of each, their strengths and weakness, and how they can be combined to derive a single common factor that one can view as the prior.

Of course, in order to construct an empirical measure of q_t, one needs to know the set of rates over which voting occurs in period t. More generally, in order to test our model, we need an empirical counterpart of v_{it}, which we denote by $\widehat{v}_{it}$. As discussed above, we model the choice facing MPC members as being over two alternative interest rates (one higher than the other) while members have, since the beginning of the MPC, voted for nine unique interest rate changes.[17] In periods with two unique votes by the MPC members (64% of the meetings), we set $\widehat{v}_{it} = 1$ if and only if member i voted for

[16] We also got access to a similar independent committee organized by the Institute of Economic Affairs (IEA). While this has the advantage of being in existence since before to the start of the MPC, it only meets every 3 months, it has a rotating attendance of members, and on many occasion only made an assessment of the current stance of policy rather than formally voting on an ideal policy choice.

[17] These are -150bps, -100bps, -75bps, -50bps, -40bps, -25bps, no change, +25bps and +50bps.

the higher of the two rates under consideration in period t. This approach means that a vote of +25bps can be mapped to $\widehat{v}_{it} = 1$ (if the alternative was no change) or $\widehat{v}_{it} = 0$ (if the alternative was +50bps). The problem comes in periods with unanimous votes since it is not clear whether the alternative under consideration is higher or lower. To resolve this difficulty, we make use of the Reuters data, as explained below.

4.1 Reuters survey data

In the days leading up to the MPC meeting, Reuters surveys around 30-50 market economists from financial institutions in London and asks them to predict the outcome of MPC voting by writing a probability distribution over possible interest rate choices.[18] Because of the fairly large cross-sectional sample size and the prominence of the participating institutions, the average beliefs in the survey data can be taken as a good measure of conventional wisdom.

In periods with two unique votes by the MPC members we set $\widehat{q}_t^R$ equal to the average probability placed on the higher observed rate over the total average probability placed on both observed rates. In periods with one observed vote, we use the survey to identify the two rates on which the market places the highest average probability and then set $\widehat{v}_{it} = 1$ if an only if i votes for the higher of these two rates.[19] We then set $\widehat{q}_t^R$ equal to the average probability placed on the higher rate over the total average probability placed on both rates. The Reuters data is the only one of our measures to provide a full probability mass function over voting alternatives, so it is the natural choice for selecting the two most likely outcomes in periods with unanimous voting.

Table 2 illustrates a hypothetical example of the survey and how we use it. The two outcomes with the highest average probability are a rise of 50 basis points and a rise of 25 basis points. So, if we observed a unanimous vote of +25, our proxy measure for the prior would be $\widehat{q}_t^R = \frac{23.75}{23.75+71.25} = 0.25$ and we would set $\widehat{v}_{it} = 0$ for all members. In fact the survey data has somewhat different formats for different sub-samples within our data and in some cases respondents were not able to write their beliefs over a full probability distribution. In these cases we follow a slightly different methodology, but the construction is similar. Appendix B.1 contains the full details of the construction in

[18] As Reuters did not have the survey results stored in their database, they were unable (or unwilling) to provide the data for us. Instead, we have been able to collate copies of the survey results for most periods in the sample; the two exceptions are February 2000 and March 2000. In addition, we are unable to use the data for periods April 2000, August 2008, and November 2008 (details of why are in the appendix). This leaves 137 out of the 142 months in our sample in which we can construct $\widehat{q}_t^R$ from this survey.

[19] We confirm that the unanimous decision reached by the MPC is one of the interest rates on which the market puts highest probability, which is itself an important test of the quality of the Reuters survey.

all periods, including how we treated anomalies in the data.

Table 2: Example of Survey Data

	+50bps	+25bps	0	-25bps	-50bps
UBS	15%	80%	5%		
Goldman Sachs	20%	75%	5%		
JP Morgan	45%	45%	10%		
AIB	15%	85%			
Average	23.75%	71.25%	5%		

Of course, $\widehat{q}_t^R$ is not a perfect measure of the unobservable q_t. For example, the data on which respondents from their beliefs is a subset of that available to the MPC; the committee is regularly given advance access to data that will only be released subsequently to the wider public, and the information about the MPC's own quarterly forecast is not known by the market respondents until the middle of the month in which it is published (about 2 weeks after the decision). Perhaps the most serious potential problem is that $\widehat{q}_t^R$ predicts the outcome of MPC voting, not the probability of the realization of the underlying state variable. We take this problem very seriously and will return to it below (see in particular sections 4.4 and 6.2).

4.2 Times MPC data

The second data source on which we draw is voting data from the Times MPC. This MPC is a committee of experts—many of whom are former or future MPC members—that votes monthly on interest rates as part of a feature for the Times of London newspaper. Importantly, they are giving opinions on what they think interest rates should be, not on what they think the committee will choose. We construct $\widehat{q}_t^T$ as the fraction of Times MPC votes for the higher of the two rates under consideration (identified either directly or indirectly via the Reuters data). Unfortunately, the Times MPC votes are a discrete measure of views and may not capture precise movements in probabilities well. Additionally, the Times MPC only began meeting November 2002, so using it means we lose 65 out of 142 meetings (561 out of 1,246 observations).

4.3 Option price data

The third data source on which we draw comes from the cross-section of prices for short sterling futures options the first day (Wednesday) of the MPC meeting.[20] Short sterling

[20]Full details about these data, as well as the data itself, are provided by the Bank of England (see Bank of England (2011a), Bank of England (2011b) and Lynch and Panigirtzoglou (2008)); here we

futures contracts are standardized futures contracts which settle on the 3 month London Interbank Offered Rate (LIBOR) on the contract delivery date. Short sterling futures options are a European option (it only settles on the delivery date and not before) on the short sterling futures contract but since the futures contract settles on 3 month LIBOR, the option is effectively an option on 3 month LIBOR. The Bank of England computes the expected value of 3 month LIBOR consistent with a risk neutral trader being willing to hold the option at each observed price; this yields a distribution over risk-neutral traders' beliefs on 3 month LIBOR. The Bank then publishes the $0.05, 0.15, \ldots, 0.95$ percentiles of this CDF. Figure 4 displays the spot 3 month LIBOR rate as well as various percentiles of this CDF. We subtract the actual value of LIBOR on the Wednesday of the MPC meeting (before the decision is made on the Thursday) to express the CDF in terms of traders' beliefs on changes in 3 month LIBOR. We denote this transformed CDF as F, and an individual belief about a change x.

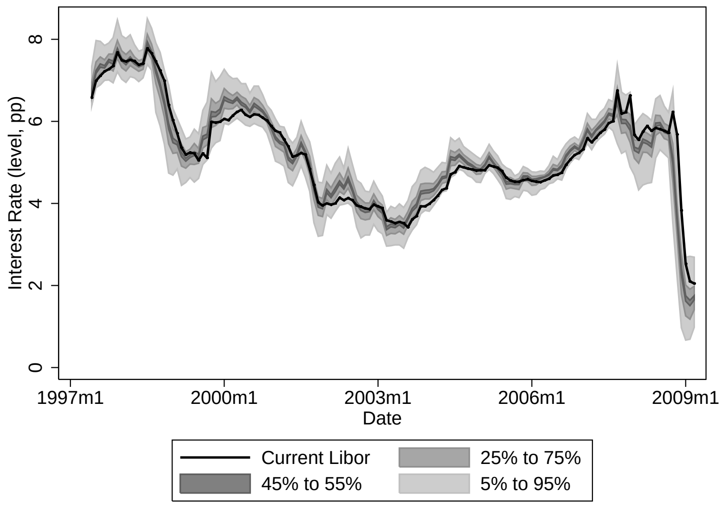


Figure 4: Short-Sterling Implied PDF and 3 Month LIBOR

Since base rate changes are made in discrete 25 point movements while traders' beliefs are continuous, we consider beliefs that lie within 12.5 basis points on either side of the corresponding change to be beliefs associated with that change being more likely. So, for example, our ideal proxy measure for q_t in a period in which no change and a 25 point rise are on the agenda would be

$$\frac{\Pr\left[\,12.5 \leq x \leq 37.5\,\right]}{\Pr\left[\,-12.5 \leq x \leq 37.5\,\right]} = \frac{F(37.5) - F(12.5)}{F(37.5) - F(-12.5)}.$$

simply outline the assumptions underlying these data, why they provide useful information for our needs and the weaknesses of the data for our use.

Unfortunately we only observe certain percentiles of F. To correct for this, we linearly interpolate between the two percentiles in which a rate change falls. So, for example, suppose that we observe that $F(10) = 0.1$ and $F(20) = 0.3$; then we would construct $F(12.5) = 0.15$. Using this method we are able to build $\widehat{q}_t^M$ for all but four time periods, meaning that we observe the markets data for 138 meetings out of 142 possible meetings in our sample.[21]

$\widehat{q}_t^M$ undoubtedly has several problems, and is perhaps the weakest proxy we construct. First, the sterling options that go into the constructed probabilities are based on LIBOR rather than the interest rate that MPC members choose (Bank of England base rate). Second, it is based on beliefs about the 3-month interest rate and as such reflects expected changes over the next three meetings not simply the one immediately following. Third, the beliefs are associated with risk-neutral traders; to the extent that actual traders are risk averse the beliefs backed out from option price data will be biased downward (yields will be biased upward) by the presence of a risk-premium in the observed market data. Finally, as with the Reuters data, $\widehat{q}_t^M$ captures predictions about LIBOR, not underlying inflation states. Despite these problems it is a useful measure because it aggregates the opinions of a large number of agents (all traders in the sterling options market) and, unlike with the Reuters and Times MPC data, these opinions are backed by real money and so potentially less subjective and manipulable.

4.4 Common Factors

So far this section has presented three alternative proxies for q_t. While each has specific weaknesses, we claim that beliefs about the underlying inflationary state should drive at least some of the variation in all three. To examine how much of the variation it drives, whether there are any other sources of variation, and to aggregate the three measures into a common measure, we adopt a factor analysis approach. Specifically, we adopt the model specification

$$\begin{aligned}
\Phi^{-1}\left(\widehat{q}^R\right) &= \alpha_R + \eta_{R,1} f_1 + \eta_{R,2} f_2 + \eta_{R,3} f_3 + \epsilon_R \\
\Phi^{-1}\left(\widehat{q}^M\right) &= \alpha_M + \eta_{M,1} f_1 + \eta_{M,2} f_2 + \eta_{M,3} f_3 + \epsilon_M \\
\Phi^{-1}\left(\widehat{q}^T\right) &= \alpha_T + \eta_{T,1} f_1 + \eta_{T,2} f_2 + \eta_{T,3} f_3 + \epsilon_T
\end{aligned}$$

where Φ^{-1} is the inverse of the standard normal cumulative density function; f_1 is the principal common factor; f_2 and f_3 are other common factors; and ϵ are normally dis-

[21]The missing periods are February and March 2000, August 2008, and November 2008. These data are missing due to thin or illiquid short-sterling options markets which mean that no options pdf data is available on the days in the run up to MPC meeting.

tributed error terms. It is necessary to transform the raw measures because we need the left hand side variables to have the same range as the right hand side ($-\infty$ to ∞).[22] We believe that f_1 is likely to capture the underlying public prior, or, in the language of our model, q_t. We will return to this assumption below; for now we stay agnostic and simply perform the factor analysis as a statistical procedure.

Factor analysis allows us to estimate f_1 as well as to determine whether f_2 and f_3 also have explanatory power and, if so, in which of the variables. Applying the Kaiser criterion (Kaiser 1960), which suggests keeping only those factors with eigenvalues greater than 1, we find a single factor that captures co-movement in the variables; the associated eigenvalue is 1.5. The eigenvalue associated with the second common factor is essentially 0, meaning it contributes nothing to the common movement. This result has important implications for understanding the effect on $\widehat{q}^M$ and $\widehat{q}^R$ of traders and Reuters respondents predicting the decision and not the inflation shock. First, since f_1 is recovered from $\widehat{q}^T$ as well as $\widehat{q}^M$ and $\widehat{q}^R$, it is not capturing any force associated with predicting decisions. So the fact that f_2 explains no additional variation in $\widehat{q}^M$ and $\widehat{q}^R$ means that the way in which individuals map q_t into a decision prediction does not vary systematically over our sample period. At the same time, even if this mapping is non-linear and complex, our factor analysis approach should strip out the underlying variable driving the volatility in the observed $\widehat{q}^M$ and $\widehat{q}^R$, which we argue is precisely q_t. We return to these issues in section .

Meanwhile, the percentage of the variance that the common factor explains in $\widehat{q}^M$, $\widehat{q}^T$, and $\widehat{q}^R$ is, respectively, 26, 50, and 74. This is consistent with the common factor reflecting q_t. As explained above, the market's beliefs about three-month LIBOR are only indirectly related to their beliefs about one-month Bank Rate; the Times data is a discrete measure that may not move much with q_t in some cases; but the Reuters respondents directly state continuous beliefs about the next decision.

We use the estimated common factor $\widehat{f_1}$ to construct a measure of q_t by using the reverse transformation $\widehat{q}^{RMT} = \Phi(\widehat{f_1})$. Since $\widehat{q}^{RMT}$ reflects the common component in the three measures, we view it as a better measure of public beliefs than any raw data series on its own. For this reason, it will serve as one of our core measures in our estimations below.

As $\widehat{q}^T$ is only available from November 2002, we repeat the factor analysis on just $\widehat{q}^R$ and $\widehat{q}^M$ from 1997 in order to construct a longer series of beliefs. We again find a single common factor according to the Kaiser criterion, which we use to construct a measure of public beliefs $\widehat{q}^{RM}$ by applying the Φ reverse transformation as above. Over the period

[22] Since $\Phi^{-1}(0) = -\infty$ and $\Phi^{-1}(1) = \infty$ we replace 0 values in the raw measures with 0.001 and 1 values with 0.999.

in which we can construct both $\widehat{q}^{RMT}$ and $\widehat{q}^{RM}$ they have a correlation of 0.94 (which is statistically significantly different from zero at 99% level).[23]

Table 3 below provides descriptive statistics for all five beliefs measures this section has presented; figure 5 shows these as time series (as one can see from 5b $\widehat{q}^{RMT}$ and $\widehat{q}^{RM}$ are very similar); and figure 6 plots the histograms for each belief measure derived from the common factor analysis.

Table 3: Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Reuters q ($\widehat{q}^{R}$)	0.5	0.3	0	1	137
Times q ($\widehat{q}^{T}$)	0.4	0.4	0	1	77
Market q ($\widehat{q}^{M}$)	0.6	0.1	0.2	1	138
Extracted Belief RMT ($\widehat{q}^{RMT}$)	0.5	0.3	0.1	1	75
Extracted Belief RM ($\widehat{q}^{RM}$)	0.5	0.2	0.1	1	137

(a) Proxies for the Prior q_t (b) Extracted Beliefs

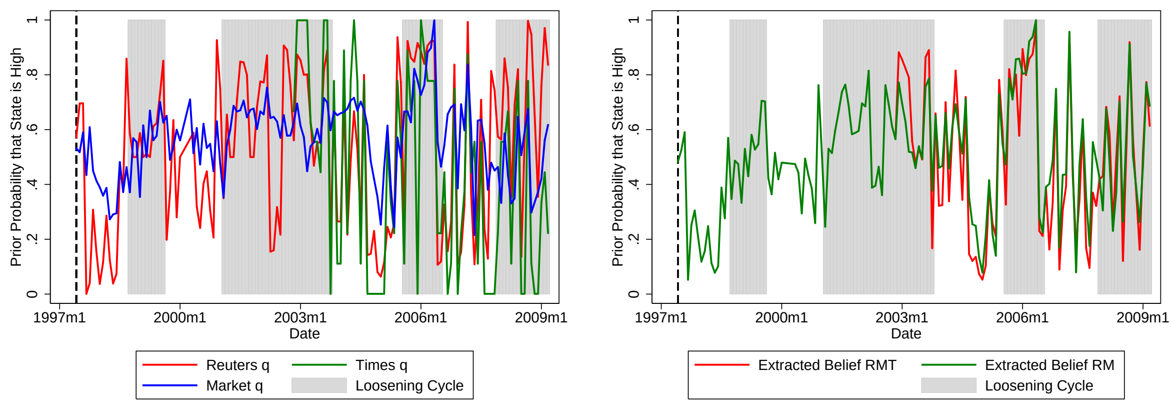


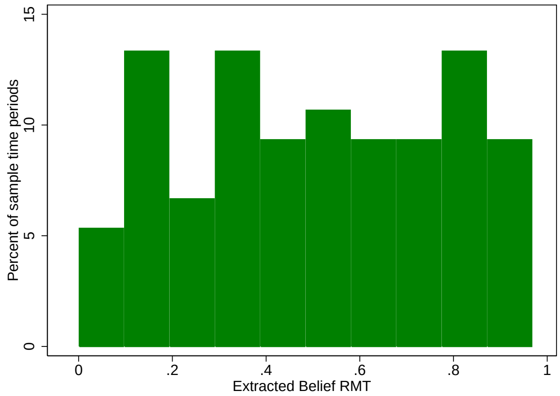
Figure 5: Key Empirical Variables

4.4.1 Is the transformed principal common factor a belief?

We have so far extracted a single common factor that drives most of the variation in the three raw measures of public beliefs. This fact alone says nothing about whether this

[23]One concern may be that none of the measures captures the information that arises in discussions during the MPC meeting prior to the point at which they vote. This information becomes public to the MPC members and should be included in the theoretically relevant measure of the common prior. To attempt to address this, we have also created a common factor using the three variables above as well as the actual votes of the MPC members in the meeting. This exercise shows again that there is only a single common factor among the 4 variables. Moreover, this common factor is highly correlated with "Common factor 1" which is constructed without using the information from the discussion (correlation is 0.87 and statistically significantly different from zero at 99% level). We do not present, or use, this measure here but it is available from the us on request.

(a) $\widehat{q}^{RMT}$



(b) $\widehat{q}^{RM}$

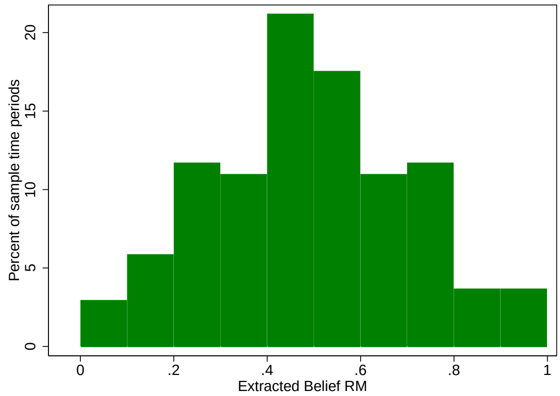


Figure 6: Histogram of the Extracted Beliefs

common factor, or more precisely its transformation into $\widehat{q}^{RMT}$ (and $\widehat{q}^{RM}$ in the longer series), is actually picking up public beliefs or instead some other common component. While we cannot directly observe whether $\widehat{q}^{RMT}$ and $\widehat{q}^{RM}$ are measuring q_t we can at least say whether they are correlated with individual votes in a way that our model predicts q_t should be.

A preliminary test of the consistency of $\widehat{q}^{RMT}$ and $\widehat{q}^{RM}$ is whether they predict whether or not an individual votes for the high rate. According to our model, a higher q_t should mean a greater probability of voting high, regardless of the values of θ and σ. To check this, we estimate the model

$$\widehat{v}_{it} = \alpha + \psi_1 \cdot \widehat{q}_t^x \tag{7}$$

for $x \in \{RMT, RM\}$ via panel logit both with and without individual fixed effects. As table 4 shows, both $\widehat{q}^{RMT}$ and $\widehat{q}^{RM}$ are statistically significant predictors of whether or not an individual votes high.

A more subtle property of q_t in the model relates to individual disagreement. As q_t tends to 0 (1) in the model, it is easy to see that the probability that all members vote for the low (high) rate tends to one. In other words, when the public prior contains very strong evidence that one or the other inflationary shock has hit the economy, no member should dissent from the majority. On the other hand, for intermediate values of q_t, there is less certainty in the public prior so dissent is more likely than when q_t takes on extreme values. Empirically, if $\widehat{q}^{RMT}$ and $\widehat{q}^{RM}$ are indeed valid measures of q_t, we would expect them to have a concave relationship with whether or not a member deviates from the majority. To check this we estimate the model

$$D\left(\text{Deviates from MPC}_{it}\right) = \alpha + \psi_1 \cdot \widehat{q}_t^x + \psi_2 \cdot \left(\widehat{q}_t^x\right)^2 + \varepsilon_{it} \tag{8}$$

Table 4: Checks on the Behavior of the Extracted Beliefs: Estimates of Equation (7)

	(1) $\widehat{v}_{it}$	(2) $\widehat{v}_{it}$	(3) $\widehat{v}_{it}$	(4) $\widehat{v}_{it}$
$\widehat{q}^{RMT}$	6.35***	8.58***		
	(0.00)	(0.00)		
$\widehat{q}^{RM}$			6.74***	8.56***
			(0.00)	(0.00)
Constant	-3.04***	-4.14***	-3.27***	-4.85***
	(0.00)	(0.00)	(0.00)	(0.00)
Observations	667	659	1,201	1,198
Estimation	Panel Logit	Panel Logit	Panel Logit	Panel Logit
Member effects	None	FE	None	FE

p-value in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 5: Checks on the Behavior of the Extracted Beliefs: Estimates of Equation (8)

	(1) D(Deviate)	(2) D(Deviate)	(3) D(Deviate)	(4) D(Deviate)
$\widehat{q}^{RMT}$	6.25***	6.99***		
	(0.00)	(0.00)		
$(\widehat{q}^{RMT})^2$	-6.09***	-6.98***		
	(0.00)	(0.00)		
$\widehat{q}^{RM}$			3.16*	3.30*
			(0.07)	(0.08)
$(\widehat{q}^{RM})^2$			-3.40**	-3.29*
			(0.05)	(0.07)
Constant	-3.09***	0.57	-2.40***	-1.90***
	(0.00)	(0.64)	(0.00)	(0.01)
Observations	667	617	1,201	1,087
Estimation	Panel Logit	Panel Logit	Panel Logit	Panel Logit
Member effects	None	FE	None	FE

p-value in parentheses
*** p<0.01, ** p<0.05, * p<0.1

for $x \in \{RMT, RM\}$, where $D\,(\text{Deviates from MPC}_{it})$ is a dummy variable equal 1 if member i deviates from the majority in time t, again via panel logit. As table 5 shows, both the linear and quadratic terms in $\widehat{q}_t^x$ are highly significant and consistent with concavity.

5 Testing the Model

A key prediction of our model is that signal differences between external and internal members should manifest themselves as different tendencies to vote for high interest rates depending on the value of the public prior. This section first examines whether external and internal members have a different relative tendency to vote high depending on the value of public beliefs, and find this is case. To determine whether this arises from differences in σ or θ parameters we move beyond reduced form analysis, structurally estimate the parameters in internals' and externals' decision rules, and find strong evidence of signal differences.

5.1 Reduced-form evidence

Before examining how voting behavior depends on the prior, we first simply see how internal and externals differ without conditioning on public information. When we observe $\widehat{v}_{it} = 1$ we say that a member has chosen the "high" vote and when we observe $\widehat{v}_{it} = 0$ we say that a member has chosen the "low" vote. Figure 7 displays the percentage of internal and external votes that are high and low. It clearly shows that internal members vote for high rates more frequently. Differences of proportion tests, reported below the figure, show that the 12.7 gap in the percentage of high votes for internal members is statistically significant at 99.9% level. This appears to confirm the idea in the existing literature that there are strong philosophical differences between internal and external members, with external members sytematically favoring lower interest rates.

The first step in our empirical analysis is to examine voting differences conditional on public information to see if any strong patterns emerge from the data. To do so, we now split the sample periods into two using both $\widehat{q}^{RMT}$ and $\widehat{q}^{RM}$. "Low-q" periods are those in which the proxy for q_t is below 0.5 and "high-q" those in which it is above. Figure 8 replicates figure 7 with this stratification, clearly showing that conditioning on public information modifies the unconditional results. In particular, there is no statistical difference between the voting of internal and external members in low-q periods (differences of proportions tests are provided below the figures); it is only in high-q periods that internal members are more likely to vote for the high interest rate.

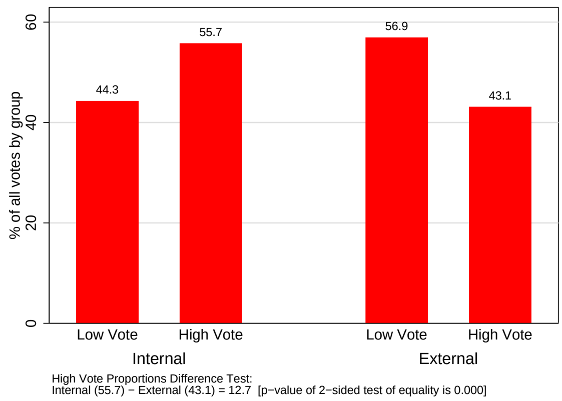


Figure 7: Histogram of $\widehat{v}_{it}$ by internal and external

Another way to look at this issue is to run a simple analysis of the probability that internal and external members vote for the higher interest rate for different values of $\widehat{q}^{RM}$. In table 6 we estimate:

$$\mathrm{D}(\widehat{v}_{it} = 1) = \alpha + \gamma_1 \cdot \mathrm{D}(\text{External} = 1) \tag{9}$$

where the sample is split into four quartiles based on the value of $\widehat{q}^{RM}$. Again, we find that the majority of the conflict arises in the upper two quartiles, with external members actually slightly more likely to vote for high rates than internal in the lowest quartile. The rest of the paper attempts to explain why the patterns in figure 8 and table 6 emerge. More precisely, if we attribute to all external members parameters θ^E and σ^E and all internal members parameters θ^I and σ^I, we are interested in what combination of these four parameters best explains the data.

(a) Using $\widehat{q}_t^{RMT}$

(b) Using $\widehat{q}_t^{RM}$

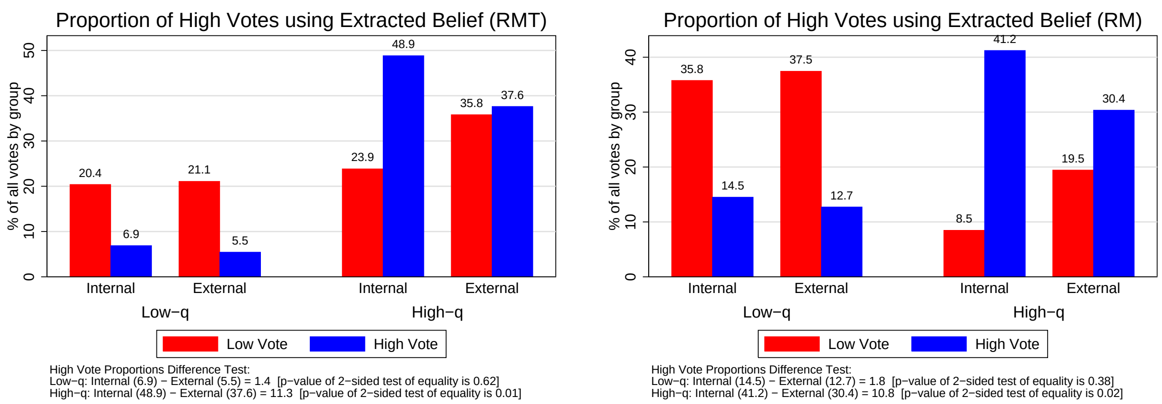


Figure 8: Histogram of $\widehat{v}_{it}$ by low and high q

Table 6: The likelihood of voting for the high rate by type across the quartiles of $\widehat{q}_t^{RM}$

	(1) 1st Quartile	(2) 2nd Quartile	(3) 3rd Quartile	(4) 4th Quartile
D(External)	0.22	-0.29*	-0.77***	-0.56***
	(0.199)	(0.057)	(0.000)	(0.000)
Constant	-1.11***	-0.18*	0.75***	1.07***
	(0.000)	(0.061)	(0.000)	(0.000)
Internal Prob	.13	.43	.77	.86
External Prob	.19	.32	.49	.69
Estimation	Probit	Probit	Probit	Probit
Range of $\widehat{q}_t^{RM}$	$\leq .342$	$> .342$ & $\leq .488$	$> .488$ & $\leq .703$	$> .703$

Robust p-values in parentheses
*** p<0.01, ** p<0.05, * p<0.1

It is first instructive to ask how philosophy differences alone might account for these findings. One might imagine that θ^E is small so that external members essentially always vote for low rates while internal members have less extreme preferences and vote for low rates when q_t is low and high rates when q_t is high. The awkwardness of explanations that rely solely on θ differences is that the voting behavior of internal and external members in terms of high votes is, on average, nearly *exactly* the same when q_t proxies are less than 0.5. While extreme dovishness among externals may generate an increased divergence between internals and externals in terms of the frequency of voting high as q_t increases, such a story would presumably show at least some conflict on average when q_t proxies are low.

Of course σ differences might also underlie the voting patterns. Suppose that in addition to an underlying philosophical difference ($\theta^E < \theta^I$) there were also an information difference $\sigma^E < \sigma^I$. Recalling proposition 2, such an informational difference would make, ceteris paribus, external members more likely to vote for high rates than internal members for low q_t and less likely to vote for high rates for high q_t. In other words, externals' lower σ would counteract the philosophical difference for low q_t and reinforce it for high q_t. This provides an alternative explanation to one that relies on philosophical differences alone.

Without moving away from reduced form analysis it is difficult to go beyond the above informal discussion of the effects of different parameter combinations. As such we now move towards structural estimation, which allows us to directly estimate which parameters best fit the data.

5.2 Structural evidence

Our model generates the likelihood function L_{it} over policy choices in time t as

$$L_{it} = \begin{cases} q_t \left(1 - \Phi\left(\frac{s^*_{it}-1}{\sigma_i}\right)\right) + (1-q_t)\left(1 - \Phi\left(\frac{s^*_{it}}{\sigma_i}\right)\right) & \text{if } v_{it} = 1 \\ q_t \Phi\left(\frac{s^*_{it}-1}{\sigma_i}\right) + (1-q_t)\Phi\left(\frac{s^*_{it}}{\sigma_i}\right) & \text{if } v_{it} = 0. \end{cases} \tag{10}$$

With our proxies $\widehat{v}_{it}$, $\widehat{q}^{RMT}$, and $\widehat{q}^{RM}$ we can simply directly estimate the model via maximum likelihood after resolving two estimation difficulties.[24]

First, for reasons we explain below, we cannot estimate θ and σ parameters for each member separately. Instead we model $\theta_i = \alpha_\theta + \beta_\theta D(\text{EXT})_i$ and $\sigma_i = \alpha_\sigma + \beta_\sigma D(\text{EXT})_i$ where $D(\text{EXT})_i$ is a dummy variable equal one if member i is external. In other words, we assume that voting parameters are constant within the subgroups of internal and external voters and compare across these groups. While one might worry about the effect of individual heterogeneity, we show in section 6 via a Monte Carlo exercise that assuming within-group homogeneity in the presence of within-group heterogeneity does not bias the estimates of the group mean.

The second difficulty in implementing the estimator is that the likelihood function is nearly flat for large and small values of q_t when σ is large. When σ increases, the probability of voting for the high rate as a function of q_t approaches a step function as figure 9 shows. This implies that our maximum likelihood estimator will have difficulty converging when q_t takes on extreme values. Therefore in our estimation procedures we drop periods in which $\widehat{q}^{RMT}$, or alternatively $\widehat{q}^{RM}$, lies below 0.15 or above 0.85; with $\widehat{q}^{RMT}$ we lose 24 out of 75 periods and under $\widehat{q}^{RM}$ we lose 20 out of 137 periods.[25]

Our first estimation uses $\widehat{q}^{RM}$ instead of $\widehat{q}^{RMT}$ since the former covers more time periods. The results of the estimation, using heteroskedastic-robust estimates of the variance covariance matrix, are shown in table 7. We report heterogeneity in θ and σ between external and internal members both in terms of gaps and ratios. The key result is that highly significant differences exist along both the θ and σ dimensions. External members' θ parameter is 0.41 as opposed to 0.55 for internals; the gap is highly significantly different from 0 and their ratio is highly significantly different from 1. This result is of independent interest since the hawk-dove distinction between internals and externals is robust to controlling for σ differences.

[24]Our estimation is done in Stata and the programs are available on request. In order to facilitate faster estimation, we actually estimate $\ln(\sigma_i)$ and $\ln\left(\frac{\theta_i}{1-\theta_i}\right)$ and then convert our estimates and standard errors back appropriately.

[25]The presence of many unanimous decision in the Times MPC pushes the common factor underlying $\widehat{q}^{RMT}$ into the extremes more often than $\widehat{q}^{RM}$, so even though they have the same mean the variance of the former is larger, meaning that we drop more time periods in percentage terms when we use it.

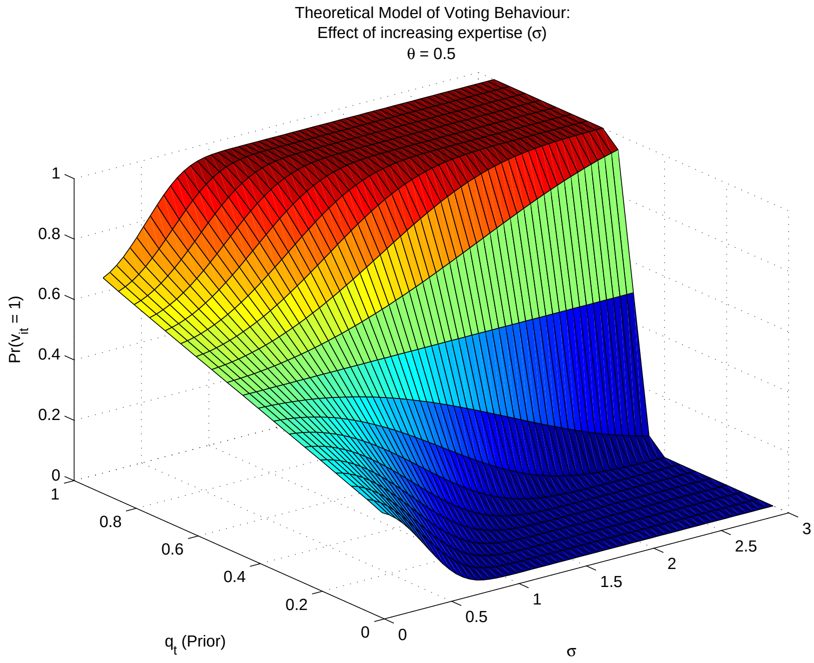


Figure 9: The Probability of Voting for the High Interest Rate

The result more central to our paper is on σ differences. The external σ parameter is almost half that of the internal σ with again a highly significant gap and ratio. This implies that externals and internals differ not just in terms of philosophy but also in terms of how influential the public prior is in their decision making; in short, external members are systematically more likely to contradict public beliefs than internal members. While our structural results alone do not identify the precise mechanism driving this different propensity to deviate from conventional wisdom, it does show that it is an important factor influencing observed voting behavior. One interpretation that readers might be tempted to make is that we have found more underlying variation in externals' θ parameters than in internals'. In fact, this issue is wholly orthogonal to the σ dimension, which instead measures differences in posterior belief formation, not heterogeneity in philosophies. In fact, section 6.1 shows that not controlling for additional variation on the θ dimension for external members does not bias the estimated σ difference.

Figure 10 is the estimated counterpart to figure 3; it is the predicted probability that internal and external members will vote for a high rate given the estimated values of the voting parameters. As can be seen, external members are on average less likely to vote for high rates, but this tendency is more pronounced as q_t increases. Comfortingly, this matches the reduced form results from above. In fact, the predicted probability differences from the structural estimates qualitatively track quite closely the actual reduced form probability differences presented in table 6 for alternative ranges of the prior. In the first

Table 7: Estimates of θ and σ using Extracted Belief RM ($\widehat{q}^{RM}$)

VARIABLES	(1) external=0	(2) external=1	(3) Difference	(4) Ratio
theta	0.55***	0.41***	0.14***	1.34***
	[0.53 - 0.58]	[0.34 - 0.48]	(0.000)	(0.001)
sigma	1.39***	0.67***	0.72***	2.07***
	[1.13 - 1.64]	[0.42 - 0.92]	(0.000)	(0.001)

Robust 95% confidence interval in brackets
Robust p-values in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Column (1)-(3), H0: Estimate = 0
Column (4) H0: Ratio = 1

quartile, external members are predicted to vote slightly more often for high rates than internals; in the second quartile this small tendency is reversed; in the third the predicted probability gap is quite large; and in the fourth it is again large but smaller than in the third. This is corresponds exactly to the pattern in table 6. Crucially, our structural estimates indicate that this emerges because external members tend to contradict the prior, not simply because they have dovish preferences.

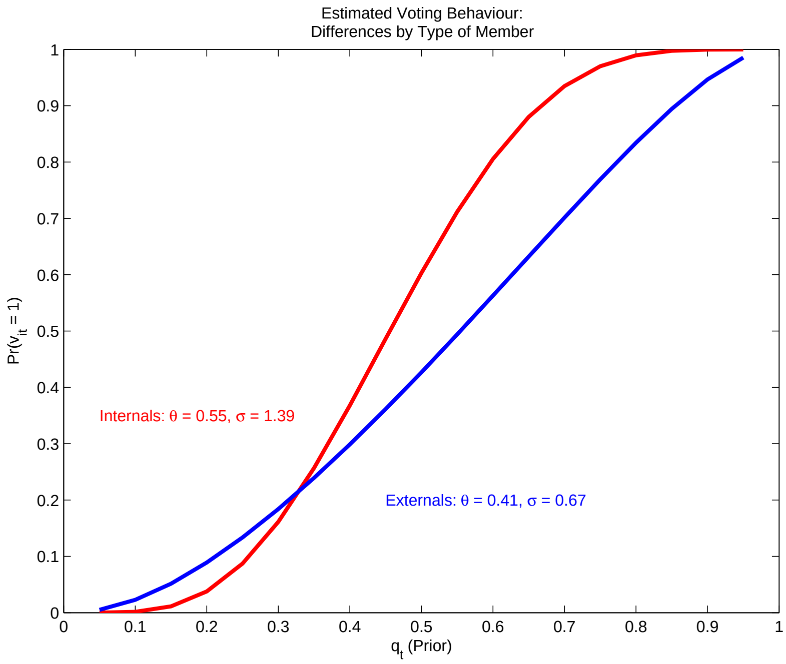


Figure 10: The Estimated Probability of Voting for the High Interest Rate

We can also run the same regressions on the restricted sample using $\widehat{q}^{RMT}$; the results are contained in table 8. The major difference between these results and those in table 7 is that the estimated magnitudes of the σ parameters are smaller. In fact, using $\widehat{q}^{RMT}$, all members are measured to be less influenced by public beliefs on average. This is because $\widehat{q}^{RMT}$ has a slightly different shape to $\widehat{q}^{RM}$ for reasons alluded to in footnote 25 above. The important message for our paper, however, is not to establish the exact magnitude of the parameters, but rather how they differ across different voters. In this sense, very little changes from the previous estimations: the ratios of θ and σ parameters is nearly the same as in table 7. One more point to note is that here all estimates are measured with more noise due to the smaller sample size, meaning the statistical significance of the results drops slightly. Now we find differences in both dimension are significant at the 10% level and the estimated ratios at the 15% level.

Table 8: Estimates of θ and σ using Extracted Belief RMT ($\widehat{q}^{RMT}$)

VARIABLES	(1) external=0	(2) external=1	(3) Difference	(4) Ratio
theta	0.61***	0.43***	0.18*	1.41
	[0.55 - 0.66]	[0.24 - 0.62]	(0.078)	(0.131)
sigma	0.90***	0.48**	0.42*	1.89
	[0.62 - 1.18]	[0.11 - 0.84]	(0.072)	(0.131)

Robust 95% confidence interval in brackets
Robust p-values in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Column (1)-(3), H0: Estimate = 0
Column (4) H0: Ratio = 1

The ultimate message of this paper is that the σ dimension is important for explaining the MPC voting record. Perhaps the most direct test of this hypothesis is to see whether allowing for σ variation improves the model's fit. The first two columns of table 9 show the results from two different MLE regressions. The first is simply a replication of the regression from above using $\widehat{q}^{RM}$. Now, a key question is: would the model explain less of the variation in the data if σ were constrained to be constant across internals and externals? Accordingly, the second regression in table 9 performs this constrained regression. Answering the previous question is equivalent to performing a likelihood ratio test; the hypothesis that the restriction of no σ heterogeneity is valid can be rejected with a p-value less than 0.001. This analysis can be replicated on the restricted sample with $\widehat{q}^{RMT}$ (the third and fourth columns of the table) and again the hypothesis no σ heterogeneity can be rejected, albeit at the smaller p-value of 0.034 due to the smaller sample size. These results highlight in a clear way that philosophy differences alone cannot account for the data as well as the combination of philosophical differences and σ differences.

Table 9: The effect of not allowing for differences in the σ

	(1) Unrestricted	(2) Just θ	(3) Unrestricted	(4) Just θ
θ^I	0.55***	0.57***	0.61***	0.64***
	(0.000)	(0.000)	(0.000)	(0.000)
θ^E	0.41***	0.44***	0.43***	0.46***
	(0.000)	(0.000)	(0.000)	(0.000)
σ^I	1.39***	1.01***	0.90***	0.71***
	(0.000)	(0.000)	(0.000)	(0.000)
σ^E	0.67***	1.01***	0.48***	0.71***
	(0.000)	(0.000)	(0.008)	(0.000)
Theta Ratio	1.34	1.28	1.41	1.38
Theta Ratio p-value[&]	[0.004]	[0.000]	[0.193]	[.017]
Sigma Ratio	2.07	1	1.89	1
Sigma Ratio p-value[&]	[0.008]	.	[0.243]	.
Log-Likelihood	-543.96	-554.23	-259.94	-262.18
LR test	.	20.55	.	4.47
Prob > chi2	.	0.000	.	.034
Proxy	$\widehat{q}^{RM}$	$\widehat{q}^{RM}$	$\widehat{q}^{RMT}$	$\widehat{q}^{RMT}$
Trimmed	$[0.15 - 0.85]$	$[0.15 - 0.85]$	$[0.15 - 0.85]$	$[0.15 - 0.85]$

p-value in parentheses
*** p<0.01, ** p<0.05, * p<0.1
[&] null hypothesis is difference from 1

5.3 Effect of mis-specifying σ

The final estimation exercise that we conduct is to examine how the measured gap in θ changes as one imposes different values of σ onto the model. The motivation for this is to examine whether explicitly estimating the σ parameter is important for correctly identifying differences in preferences. Even if one were interested in correctly computing the degree of preference conflict between two experts and not interested in their information differences, it might still be important to account for the latter in order to correctly estimate the former.

In order to examine this question, we estimate three different versions of our model with the proxy $\widehat{q}^{RM}$: one in which $\sigma = 0.4$ for both groups; one in which $\sigma = 1$ for both groups; and one in which $\sigma = 5$ for both groups. The pattern that emerges from the results, displayed in table 10, is that the estimated θ difference is decreasing in value of σ imposed on the model. The regression with $\sigma = 1$ (which corresponds to the estimated value of σ in table 9) estimates a θ gap of 0.13. When σ is set to 0.4 the estimated difference grows more than five times to 0.77, and when σ is set to 5 it is nearly cut in half to 0.067. Thus the estimated θ difference appears sensitive to mis-specification of σ, especially for small values. Hence empirical exercises that seek to measure preference differences in contexts where learning about the state of the world from private signals is relevant, the estimation of those signals' distribution is crucial.

Table 10: The effect of estimating θ using different assumptions for σ

(a) $\sigma = 0.4$, θ Estimates

	(1) D(External)=0	(2) D(External)=1	(3) Difference	(4) Ratio
θ	0.90***	0.13**	0.77***	6.99***
	[0.83 - 0.98]	[0.00088 - 0.26]	(0.000)	(0.000)

(b) $\sigma = 1$, θ Estimates

	(1) D(External)=0	(2) D(External)=1	(3) Difference	(4) Ratio
θ	0.57***	0.44***	0.13***	1.28***
	[0.54 - 0.60]	[0.41 - 0.48]	(0.000)	(0.000)

(c) $\sigma = 5$, θ Estimates

	(1) D(External)=0	(2) D(External)=1	(3) Difference	(4) Ratio
θ	0.54***	0.47***	0.067***	1.14***
	[0.52 - 0.56]	[0.45 - 0.50]	(0.000)	(0.000)

Robust 95% confidence interval in brackets
Robust p-values in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Column (1)-(3), H0: Estimate = 0
Column (3) H0: Ratio = 1

6 Robustness

To conclude our empirical analysis we return to two important issues mentioned in the text above. First we explore the consequences for our results if we allow for individual heterogeneity. Next we go deeper into the problems associated with two of our three raw measures of public beliefs reflecting predictions of MPC voting outcomes rather than predictions of the underlying inflationary conditions. Neither concern appears to have a substantial impact on our key findings.

6.1 Individual heterogeneity

In our analysis, we have not included member fixed effects for two main reasons. The first is that, as we make use of a probit-type estimation, inclusion of fixed-effects is potentially problematic (Baltagi 2005). Second, we are interested in the behavior of the average external and internal member on the MPC and not heterogeneity among individual members. While we accept that the θ and σ parameters may differ at the level of individual members, member fixed effects would soak up the variation in which we are interested. We would therefore need to calculate member-specific estimates and then examine the distribution of these member-specific parameters by internal and external grouping. The problem with this approach is that many external members have not served for long periods, and our likelihood estimator is unable to converge in many cases. The approach we pursue above means that we can draw on statistical power from the fact that we have over 500 external member votes and over 600 internal member votes.

Nevertheless, one may still be concerned that individual heterogeneity manifests itself in the form of biased estimates when we impose the restriction of a common θ and σ within the group of internals and externals. In order to address these worries, we carry out two robustness exercises. First, we confirm that our analysis is unchanged if we relax the assumption that each of our observations is conditionally uncorrelated. While the previous results used heteroskedastic-robust estimates of the variance-covariance matrix, these estimates do not allow for member-specific correlation of errors. In table 11, we use the $\widehat{q}^{RM}$ prior but instead cluster the errors by individual member; there is no change in the significance of the σ differences, and the significance of the θ ratio only falls to 11%.

Table 11: Robustness to Individual Heterogeneity: Estimates if cluster errors by member

	(1) Internal	(2) External	(3) Difference	(4) Ratio
θ	0.55***	0.41***	0.14*	1.34
	[0.52 - 0.59]	[0.27 - 0.56]	(0.060)	(0.101)
σ	1.39***	0.67***	0.72***	2.07***
	[1.23 - 1.55]	[0.31 - 1.03]	(0.000)	(0.009)

Errors Clustered by Member
95% confidence interval in brackets
p-values in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Column (1)-(3), H0: Estimate = 0
Column (4) H0: Ratio = 1

Second, we use Monte Carlo simulations to study the properties of our estimator when there is individual heterogeneity in both θ and σ parameters. The full details are contained in appendix C, and here we simply present the main findings. We first perform the same exercise as in section 5 and allow for across group differences but no within group differences; this is simply to verify that if the data exactly matched our modelling assumptions, our estimator would produce unbiased results, and indeed the simulations indicate this is the case.

Next we allow for each individual's θ and σ parameter to be drawn from a uniform distribution, and assume that internals' and externals' distributions differ in means but not in variances. We then impose within group homogeneity on the estimator and examine whether it accurately estimates the group average. Since our regression model is non-linear it is perhaps not clear ex ante that this will be true, but the Monte Carlo results nevertheless indicate that it is.

Our last exercise is designed to address a particular concern. Suppose that external members not only have a lower average θ but that the distribution from which their θ parameters are drawn has a higher variance than the one for internal members. If we fail to control for the additional variance in the external θ distribution, will it wrongly show up as a difference in σ parameters? More loosely speaking, is the only difference that our empirical results on σ capture the fact that external members have more heterogeneous philosophies? To answer this question, we perform another set of simulations in which we again assume that individual parameters are drawn from a uniform distribution, but that the external θ distribution is more variable than the internal. We again find that if we impose within group homogeneity on our estimator it accurately estimates group averages of θ and σ.

Note that these results also indicate that the imposition of time-independence of the θ and σ parameters is also innocuous. Any dynamics in these parameters would appear in exactly the same way that individual variation does in the Monte Carlo exercises. If σ and θ change through time, our estimator still appears to uncover the average difference between internals and externals.

6.2 The prior as a measure of voting outcomes

We now revisit the issue of the potential prediction bias in the Reuters survey. Consider an outside observer to the committee that holds the public prior q_t that the economy is in the high inflation state. In order to predict the outcome the committee will actually decide, a natural tendency would be to predict the vote of the average internal member, since there are more internals than externals. Figure 11 plots out the probability than an internal member votes high on the basis of the parameter values estimated in our baseline results (the red curve) against the 45 degree line. If the outside observer reported his prior belief directly, one could read his report from the 45 degree line. On the other hand, if he were reporting the predicted decision, one would read his report off of the red line. For example, suppose that $q_t = 0.3$. At this value of the prior, the internal member votes high with probability 0.17, meaning that if the observer were reporting predictions of outcomes rather than predictions of inflation shocks, there would be a downward bias of 0.12. On the other hand, if $q_t = 0.7$, the red line gives a value of 0.93, introducing an upward bias of 0.23. In other words, the prediction bias, if it exists, biases reported beliefs down when q_t is low and up when q_t is high.[26] This would be true if the outside observer predicted the outcome using the characteristics of the average member, or even the average external member.

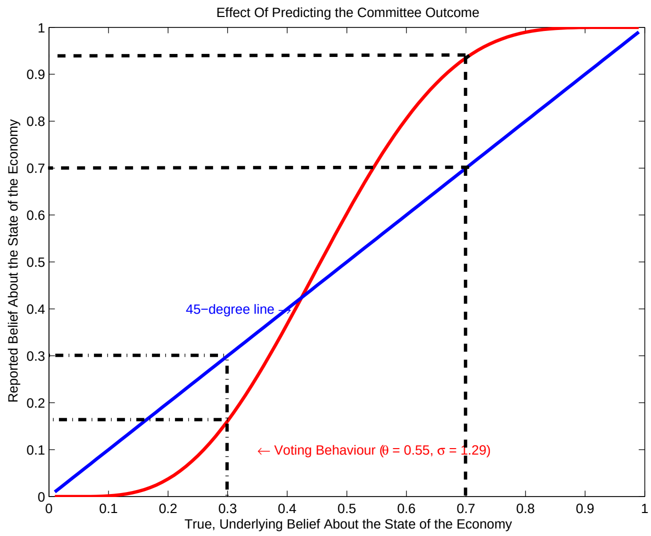


Figure 11: How q_t is related to $\widehat{q}_t$ when the respondent is predicting the outcome

The first important point that emerges from figure 11 is that the predicted probability that an internal member votes for the high rate is increasing in the prior probability that the high inflation state is realized—there is a one-to-one mapping between the two quantities. So ordering periods according to reported beliefs based on predicted outcomes produces the same ordering of periods that unbiased reports would. So the reduced form results above, even if based on biased belief data, will still be valid in a qualitative sense.

More importantly, we don't use the raw data series in our structural estimation but rather use our extracted beliefs based on common factor analysis. We have already argued that the single common factor that explains the covariation in our three raw measures should reflect the underlying common prior that goes into forming each of them. As such, any prediction bias in the Reuters should not arise in it.

One way of identifying whether this is the case is seeing whether $\widehat{q}^{RM}$ and $\widehat{q}^{RMT}$ correct $\widehat{q}^{R}$ in a way consistent with the cleaning of the anticipated prediction error. To examine this in a simple way, consider figure 12a, which is a scatter plot of $\widehat{q}^{RM}$ on the horizontal

[26] Of course, one may also think that the prediction curve shifts over the sample period; for example, if the committee composition changes then the optimal prediction may require a different imputed value of θ and σ to the average external member. Recall, however, that we found absolutely no evidence of such variability in the common factor analysis.

axis against the corresponding values of $\widehat{q}^R$ on the vertical. The vertical line corresponds to the point at which the red curve in figure 11 crosses the blue, or 0.45. The relationship between the two is broadly consistent with $\widehat{q}^{RM}$ being the average underlying belief on q_t and $\widehat{q}^R$ being the predicted committee decision made on the basis of it. Below 0.45, the majority of the observations of $\widehat{q}^R$ lie below the 45 degree line, whereas above 0.45 the majority lie above it, just as the model would predict. Figure 12b replicates this analysis on the basis of $\widehat{q}^{RMT}$ (now placing a vertical line at 0.37, which is where the prediction bias is estimated to change signs on the basis of the parameter estimates in table 8) and again finds broadly the same pattern.

(a) $\widehat{q}^{RM}$ vs. q^R (b) $\widehat{q}^{RMT}$ vs. q^R

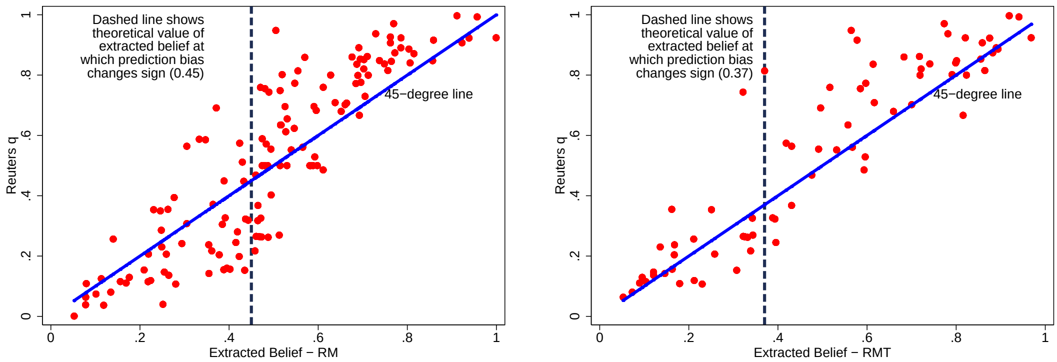


Figure 12: Scatter of raw Reuters (q^R) and Extracted Beliefs ($\widehat{q}^{RM}$ and $\widehat{q}^{RMT}$)

In order to more explicitly examine prediction bias, we also carried out our own (one-off) survey of market participants replicating the approach of Reuters by asking respondents to write down probability mass functions over different interest rate outcomes, but unlike Reuters we ask two questions. The first explicitly asks what the respondent thinks the MPC will do (which we call q^{will}) and the second asks what they believe should happen (q^{should}). We carried out the survey in the week of the April 2011 meeting and received 26 responses including from some of the most prominent market economists in the City of London; 14 of these are monthly respondents of the Reuters survey. Nearly all the probability was placed on no change and +25bps, which were the two votes observed in the subsequent meeting.

We find that the average q^{will} is 0.16 while the average q^{should} is 0.26, which is consistent with their being a prediction bias.[27] In order to see how our common factor analysis would

[27]Interestingly, among the subset of respondents that regularly appear in our Reuters survey, there is a negligible difference between the average responses to the two questions - the average q^{will} is 0.18 while the average q^{should} is 0.21. Thus, while there is no doubt that respondents are being asked to predict outcomes, how they go about doing so may not actually correspond to the rather complex process described above, but to simply reporting one's genuine belief.

treat this situation, we:

1. treat q^{will} as the $\widehat{q}^R$ (Reuter's response) for April 2011;
2. plug in the April 2011 values for $\widehat{q}^M$ and $\widehat{q}^T$ as described above;
3. perform our common factor analysis as in section 4 and generate $\widehat{q}^{RMT}$.

We find that $\widehat{q}^{RMT}$ for April 2011 is 0.25, which is much closer to the average q^{should} than the average q^{will}. Thus, in the presence of a discrepancy between the two responses, our factor analysis approach appears to deliver a result more or less in line with respondents' beliefs independent of prediction bias. Whilst only a one-off survey, this exercise reassures us that using multiple sources on the public prior helps to correct for any prediction bias which may be present in the Reuters survey.

7 Conclusion

This paper has both demonstrated that philosophy and signal differences are separately identifiable from policy choices and applied our model to MPC voting data to show that each contributes to fitting the data. Our findings complement the existing literature on MPC voting that emphasizes preference conflict between internal and external members by estimating that the latter have more dovish philosophies, but also extends it by estimating that externals' policy choices are on average less influenced by public beliefs than internals'. More generally, we have provided a clean way of estimating the parameters of Bayesian decision rules that are common in the theory literature, meaning that one could use our same approach in contexts other than monetary policy.

While we have taken care to highlight that internal and external members have different estimated signal distributions, we have not provided a precise interpretation of the result. In the spirit of discussion, we can offer a few ideas on what might drive it, but leave the task of discriminating among them to future research. In our model members begin with a public forecast of a shock and then combine it with their private forecast to form their posterior beliefs. The most literal interpretation of the result is that external members put more weight on their private forecast than internal members because their private forecasts provide relatively more insight on economic conditions. Note that this statement is not necessarily equivalent to saying that external members have more expertise than internals. It could be, for example, that internals' forecasts are incorporated into the market's belief more so than externals' if the Bank's public communications reflect more the former. Still, several other factors might drive the result.

First, our model implicitly assumes that members' signals are independent conditional on the prior but in practice this may not be the case. Our constructed prior is an aggregation of individual views that might be more correlated with internal members' than externals'. Informally speaking, the market and internals might "see the world" in the same way while externals exhibit more unique outlooks. Second, internal and external members might have utilities from contradicting the prior apart from those modeled. For example, internals members might fear contradicting the market's expectations more than external members, or external members might want to appear contrarian to increase their exposure or gain a reputation as experts.[28]

Whichever combination of these forces explains our results, we hope to have convinced the reader that an aspect of decision making that is often ignored in the empirical literature is in fact of first order importance. We view this as an important first step in the effort to tackle the difficult twin questions of why experts choose the policies that they do and how committees should be optimally composed.

[28]The possibility that career concerns lead members to react to the prior distribution to gain reputation has been emphasized by, among others, Ottaviani and Sorensen (2000), Levy (2004), Levy (2007), Visser and Swank (2007), and Meade and Stasavage (2008).

References

Baltagi, B. (2005): *Econometric Analysis of Panel Data*. John Wiley and Sons, New York, New York, third edn.

Bank of England (2010a): `http://www.bankofengland.co.uk/monetarypolicy/decisions.htm`, last accessed 15 September 2010.

——— (2010b): `http://www.bankofengland.co.uk/monetarypolicy/overview.htm`, last accessed 15 September 2010.

——— (2011a): "Notes On The Bank Of England Option-Implied Probability Density Functions," `http://www.bankofengland.co.uk/statistics/impliedpdfs/BofEOptionImpliedProbabilityDensityFunctions.pdf`, last accessed 12 March 2011.

——— (2011b): "Option-Implied Probability Density Functions for Short Sterling Interest Rates and the FTSE 100," `http://www.bankofengland.co.uk/statistics/impliedpdfs/index.htm`, last accessed 12 March 2011.

Bank of England Act (1998): `http://www.bankofengland.co.uk/about/legislation/1998act.pdf`, last accessed 15 September 2010.

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

Bhattacharjee, A., and S. Holly (2005): "Inflation Targeting, Committee Decision Making and Uncertainty: The Case of the Bank of England's MPC," Discussion paper, Department of Applied Economics, University of Cambridge.

Blinder, A. S. (2007): "Monetary policy by committee: Why and how?," *European Journal of Political Economy*, 23, 106123.

Clarida, R., J. Galí, and M. Gertler (1999): "The Science of Monetary Policy: A New Keynesian Perspective," *Journal of Economic Literature*, 37(4), 1661–1707.

Condorcet, M. d. (1785): *Essai sur l'application de l'analyse la probabilit des dcisions rendues la pluralit des voix (Essay on the Application of Analysis to the Probability of Majority Decisions)*. l'Imprimerie Royale, Paris.

Feddersen, T., and W. Pesendorfer (1998): "Convicting the Innocent: The Inferiority of Unanimous Jury Verdicts under Strategic Voting," *American Political Science Review*, 92(1).

Galí, J. (2008): *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework*. Princeton University Press, Princeton, USA and Oxford, UK.

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

Gerling, K., H. P. Gruner, A. Kiel, and E. Schulte (2005): "Information acquisition and decision making in committees: A survey," *European Journal of Political Economy*, 21(3), 563–597.

Goeree, J. K., and L. Yariv (2010): "An Experimental Study of Collective Deliberation," *Econometrica*, forthcoming.

Harris, M., and C. Spencer (2008): "Decade of dissent: explaining the dissent voting behavior of Bank of England MPC members," MPRA Paper 9100, University Library of Munich, Germany.

Kaiser, H. F. (1960): "The Application of Electronic Computers to Factor Analysis," *Educational and Psychological Measurement*, 20(1), 141–151.

Lambert, R. (2006): "Who sets our interest rates?," `http://www.bankofengland.co.uk/publications/other/monetary/inside_mpc.pdf`.

Levy, G. (2004): "Anti-herding and strategic consultation," *European Economic Review*, 48(3), 503–525.

——— (2007): "Decision Making in Committees: Transparency, Reputation, and Voting Rules," *American Economic Review*, 97(1), 150–168.

Lynch, D., and N. Panigirtzoglou (2008): "Summary statistics of option-implied probability density functions and their properties," Bank of England Working Paper 345.

Meade, E. E., and D. Stasavage (2008): "Publicity of Debate and the Incentive to Dissent: Evidence from the US Federal Reserve," *Economic Journal*, 118(528), 695–717.

Ottaviani, M., and P. Sorensen (2000): "Herd Behavior and Investment: Comment," *American Economic Review*, 90(3), 695–704.

Pollard, P. (2004): "Monetary Policy-Making around the World," Presentation, Federal Reserve Bank of St. Louis.

Riboni, A., and F. J. Ruge-Murcia (2010): "Monetary Policy by Committee: Consensus, Chairman Dominance, or Simple Majority?," *The Quarterly Journal of Economics*, 125(1), 363–416.

Spencer, C. (2006): "The Dissent Voting Behaviour of Bank of England MPC Members," Department of Economics Discussion Papers 0306, Department of Economics, University of Surrey.

Visser, B., and O. H. Swank (2007): "On Committees of Experts," *The Quarterly Journal of Economics*, 122(1), 337–372.

A Model Derivations and Micro-foundations

A.1 Proof of proposition 1

Proof. As is clear from equation (3) in the text, $\widehat{w}(s_{it})$ is strictly increasing in s_{it}. So the voter adopts a cutoff rule by choosing $v_{it} = 1$ if and only if $s_{it} \geq s_{it}^*$ where $\widehat{w}\left(s_{it}^*\right) = 1 - \theta_i$. This in turn is equivalent to $\ln\left(\frac{\widehat{w}\left(s_{it}^*\right)}{1-\widehat{w}\left(s_{it}^*\right)}\right) = \ln\left(\frac{1-\theta_i}{\theta_i}\right)$. Plugging in equation (3) immediately gives the result. ∎

A.2 Proof of proposition 2

Proof. The probability of member i voting for a high rate, given in (10), can be alternatively expressed as

$$q\left\{1 - \Phi\left[-\frac{1}{2\sigma} - \sigma\ln\left(\frac{q}{1-q}\frac{\theta}{1-\theta}\right)\right]\right\} + (1-q)\left\{1 - \Phi\left[\frac{1}{2\sigma} - \sigma\ln\left(\frac{q}{1-q}\frac{\theta}{1-\theta}\right)\right]\right\}$$

whose derivative with respect to σ is

$$\begin{aligned}
&-q\phi\left[-\frac{1}{2\sigma} - \sigma\ln\left(\frac{q}{1-q}\frac{\theta}{1-\theta}\right)\right]\left[\frac{1}{4\sigma^2} - \ln\left(\frac{q}{1-q}\frac{\theta}{1-\theta}\right)\right] \\
&-(1-q)\phi\left[\frac{1}{2\sigma} - \sigma\ln\left(\frac{q}{1-q}\frac{\theta}{1-\theta}\right)\right]\left[-\frac{1}{4\sigma^2} - \ln\left(\frac{q}{1-q}\frac{\theta}{1-\theta}\right)\right].
\end{aligned}$$

using the fact that by definition

$$\phi\left[\frac{s^*}{\sigma}\right] = \frac{q}{1-q}\frac{\theta}{1-\theta}\phi\left[\frac{s^*-1}{\sigma}\right]$$

allows this to be rewritten as

$$-q\phi\left[\frac{s^*-1}{\sigma}\right]\left[\frac{1}{4\sigma^2} - \ln\left(\frac{q}{1-q}\frac{\theta}{1-\theta}\right) - \frac{\theta}{1-\theta}\frac{1}{4\sigma^2} - \frac{\theta}{1-\theta}\ln\left(\frac{q}{1-q}\frac{\theta}{1-\theta}\right)\right]$$

which simplifies to

$$-q\phi\left[\frac{s^*-1}{\sigma}\right]\frac{1}{1-\theta}\left[\frac{1-2\theta}{4\sigma^2} - \ln\left(\frac{q}{1-q}\frac{\theta}{1-\theta}\right).\right]$$

so the derivative is positive if and only if

$$\ln\left(\frac{q}{1-q}\frac{\theta}{1-\theta}\right) > \frac{1-2\theta}{4\sigma^2}$$

which is clearly satisfied for all $q > q^* \in (0,1)$ since $\ln\left(\frac{q}{1-q}\right)$ is monotonically increasing from $-\infty$ to ∞ as q ranges from 0 to 1. ■

A.3 New Keynesian Model of Clarida, Galí, and Gertler (1999)

This section presents a basic version of Clarida, Galí, and Gertler's (1999) New-Keynesian model of monetary policy as presented in Galí (2008). We show that such a model leads to a similar interest rate rule as we derived in section 3. To begin with, we define the welfare-relevant output gap is $x_t \equiv y_t - y_t^e$, where y_t is the log of the stochastic component of output and y_t^e is (log) efficient level of output, and the natural level of output y_t^N to be the flexible price and wage level of output. Finally, define π_t as the deviation of period t inflation from its long-run level and i_t to be the nominal interest rate.

As is standard in the New Keynesian model, the economy is characterized by a dynamic IS curve and a New Keynesian Phillips Curve (NKPC):

$$x_t = \mathbb{E}_t[\, x_{t+1}\,] - \phi(i_t - \mathbb{E}_t[\, \pi_{t+1}\,] - r_t^e) \tag{A.1}$$

$$\pi_t = \beta \mathbb{E}_t[\, \pi_{t+1}\,] + \kappa x_t + u_t \tag{A.2}$$

Equation (A.1) is derived from a log-linearisation of the household Euler equation where $r_t^e = -log\beta + \frac{1}{\phi}\mathbb{E}_t\left[\, \Delta y_{t+1}^e\,\right]$ is the interest rate that supports the efficient level of output and ϕ is inversely related to the curvature of the consumption function. For simplicity we allow only for a shock to the NKPC, u_t; this shock is generally called a cost-push shock in the literature but it can be anything which drives a wedge between y_t^e and y_t^N such as changes in desired mark-ups, or anything that affects expected marginal costs.[29] This shock is assumed to follow an AR(1) stochastic process:

$$u_t = \rho_u u_{t-1} + \epsilon_t^s \tag{A.3}$$

The social loss to the representative household are approximated (using a second-order approximation) to be proportional to a function that is quadratic in inflation and output where λ is $\frac{\kappa}{\varepsilon}$ (and ε is the substitutability between goods in the consumption aggregator):[30]

$$\Gamma_t = \frac{\pi_t^2}{2} + \lambda \frac{y_t^2}{2} \tag{A.4}$$

[29]We can also include a shock to the IS curve, g_t, which is related to expectations of the relative changes in natural output and government spending. Such a shock enters the optimal interest rate linearly.

[30]We can also allow for an interest rate smoothing term.

We assume that the policymaker has discretion to choose the interest rate at time t, based on information available at the end of time $t-1$, so as to minimize the discounted future losses. As the central banker is free to reoptimise in every period, there is no way that they can credibly commit to any pre-specified path of future policy; this allows us to reduce the problem to a series of static problems in which the central bank takes private sector expectations as given. The discretionary problem to determine optimal monetary policy is:

$$\begin{aligned}
\underset{i_t}{\text{minimize}} \quad & \frac{\pi_t^2}{2} + \lambda \frac{y_t^2}{2} + \mathbb{E}_t \left[\underbrace{\sum_{\tau=1}^{\infty} \Gamma_{t+\tau}}_{\text{taken as given}} \right] \\
\text{subject to} \quad & \pi_t = \kappa x_t + \underbrace{\beta \mathbb{E}_t[\, \pi_{t+1} \,] + u_t}_{f_t \text{ taken as given}}
\end{aligned}$$

The first order necessary condition for an optimum is:

$$x_t = -\frac{\kappa}{\lambda}\pi_t \tag{A.5}$$

We can use equation (A.5), together with the other model equations, to derive the optimal interest rate:

$$i_t = r_t^e + \frac{1}{\phi}\left(\frac{1}{\kappa^2 + \lambda(1-\beta\rho_u)}\right)(\kappa + \lambda\phi)\, u_t + \frac{1}{\phi} g_t \tag{A.6}$$

Forcing the MPC member must choose between two interest rates $\bar{i}_t > \underline{i}_t$ (as we do in our main model and show is the case for most MPC meetings), the vote is given by:

$$v_{it} = \begin{cases} \bar{i}_t & \text{if } i_{jt} \geq \bar{i}_t - 12.5\text{bps} \\ \underline{i}_t & \text{otherwise} \end{cases} \tag{A.7}$$

$$\text{where } i_{jt} = r_{jt}^e + \frac{1}{\phi}\left(\frac{1}{\kappa^2 + \lambda(1-\beta\rho_u)}\right)(\kappa + \lambda\phi)\, u_t$$

Of course, this assumes that the central banker has perfect knowledge of the current period information whereas we are interested in the vote under imperfect information. As Clarida, Galí, and Gertler's (1999) Result 9 states: "With imperfect information, stemming either from data problems or lags in the effect of policy, the optimal policy rules are the certainty equivalent versions of the perfect information case. Policy rules must be expressed in terms of the forecasts of target variables as opposed to the ex post behavior." Therefore, under uncertainty about the state of the economy, where Ω_t is the information set at time t, the decision rule is as above except that we have:

$$r_{jt}^{e} + \frac{1}{\phi}\left(\frac{1}{\kappa^2 + \lambda(1-\beta\rho_u)}\right)(\kappa + \lambda\phi)\,\mathbb{E}[\,u_t \mid \Omega_t\,] \geq \bar{i}_t - 12.5\text{bps} \tag{A.8}$$

In other words, the member will vote for the higher interest rate ($\bar{i}_t$) if and only if:

$$\mathbb{E}[\,u_t \mid \Omega_t\,] \geq \phi\left(\frac{\kappa^2 + \lambda(1-\beta\rho_u)}{(\kappa + \lambda\phi)}\right)\left[\bar{i}_t - r_{jt}^{e} - 12.5\text{bps}\right] \tag{A.9}$$

This is analogous to equation (2) in our main model; the member votes high when their belief about the economic shock exceeds their member-specific cut-off.

B Data

B.1 Reuter's Survey Data

As described above, the Reuters Data that we have been able to gather is divided into three segments corresponding to data availability, and each period calls for a different construction methodology. Implicit in the construction is how we define $\widehat{D}(\text{High Vote})_{it}$.

June 1997 - June 1998: Mark I Construction

We have modal data, all of which takes a value of 0, and +25, except for October 1997 in which one bank reported +50 which we treat as +25.

- $\widehat{q}_t^R$ is computed as the total number of reports for +25 over the total number of reports.
- Anomalies
  - In the May 1998 MPC meeting there were 6 votes for no change, 1 vote for -25 (Julius) and 1 vote for +25 (Buiter). We treat the -25 vote as a low vote and proceed as above.
  - In the June 1998 MPC meeting there were 8 votes for +25 and 1 vote for -25 (Julius). Again, we treat -25 as a low vote.

July 1998 - December 2001: Mark II Construction

We have partial probability distribution data over rise, no change, and cut. In some periods only two of these three options are available.

- We compute the average probability placed on rise, cut, and no change.
- For all periods in which the observe votes that lie in the set -25,0,25 we treat the data as if the probability of rise is the probability of +25 and the probability of cut is the probability of -25.
- For periods in which the votes are +25 and 0 we construct $\widehat{q}_t^R$ as the total average probability put on +25 over the total average probability placed on +25 and 0.
- For periods in which the votes are 0 and -25 we construct $\widehat{q}_t^R$ as the total average probability put on 0 over the total average probability placed on 0 and -25.
- For periods in which all votes are for +25 we construct $\widehat{q}_t^R$ over the total average probability put on +25 over the total average probability placed on +25 and 0.
- For periods in which all votes are for -25 we construct $\widehat{q}_t^R$ over the total average probability put on 0 over the total average probability placed on 0 and -25.

- For periods in which all votes are for nochange we construct $\widehat{q}_t^R$ as the total average probability put on +25 over the total average probability placed on +25 and 0 if the total average probability placed on +25 is larger than the total average probability placed on -25, and construct $\widehat{q}_t^R$ as the total average probability put on 0 over the total average probability placed on 0 and -25 if the total average probability placed on -25 is larger than the total average probability placed on +25.

- Anomalies

  - In August 1998 there were 7 votes for no change, 1 vote for -25 (Julius) and 1 vote for 25 (Buiter). The total average probability placed on +25 is 0.39 and the total average probability placed on -25 is 0.001. So we treat +25 as high vote and compute as the total average probability place on +25.
  - In January 1999 there was one vote for no change (Plenderleith), 7 votes for -25, and one vote for -50 (Julius). We treat the vote for -25 and -50 as low votes and compute $\widehat{q}_t^R$ as the total average probability placed on nochange over the total average probability placed on nochange and cut.
  - In March 1999 8 people voted for nochange and 1 person (Buiter) voted for -40. We treat the -40 vote as a vote for -25 and proceed as above.
  - In April 2000 3 people voted for +25 and six for 0, but Reuters survey does not ask for probability of rise. We treat these data as missing in this period.
  - In January 2000 the votes were over +25 and +50 and we set $\widehat{q}_t^R = 0.5$.
  - In April 2001, May 2001, October 2001, and November 2001, the votes were over -25 and -50 and we set $\widehat{q}_t^R = 0.5$.

January 2002 - November 2008: Mark III Construction
We have full distribution data over +50, +25, 0, -25, -50.

- We compute the average probability placed on +50, +25, no change, -25 and -50.

- For periods in which there are two unique votes we take $\widehat{q}_t^R$ as the total average probability placed on the higher of the two votes over the total average probability placed on both votes.

- For periods in which all votes are for +50 we construct $\widehat{q}_t^R$ over the total average probability put on +50 over the total average probability placed on +50 and +25.

- For periods in which all votes are for -50 we construct $\widehat{q}_t^R$ over the total average probability put on -25 over the total average probability placed on -25 and -50.

- For periods in which all votes are for nochange we construct $\widehat{q}_t^R$ as the total average probability put on +25 over the total average probability placed on +25 and 0 if the total average probability placed on +25 is larger than the total average probability placed on -25, and construct $\widehat{q}_t^R$ as the total average probability put on 0 over the total average probability placed on 0 and -25 if the total average probability placed on -25 is larger than the total average probability placed on +25.

- We follow a similar procedure as the above for periods in which all votes are for +25 or -25.

- Anomalies

  - In May 2006 there were six votes for no change, one vote for +25 (Walton) and one vote for -25 (Nickell). The market put probability 0.08 on -25 and probability 0.03 on +25. So we take 0 to be high vote and compute $\widehat{q}_t^R$ as (meannochange + meanrise25)/(meancut25 + meannochange + meanrise25).
  - In April 2008 there were votes for 0, -25, and -50. We take 0 as a high vote and -25 and -50 as low votes, and compute the high vote as (meannochange)/(meancut25 + meannochange + meancut50).
  - In August 2008 there were votes over -25, 0, +25 but the market placed roughly equal probability on -25 and +25 so we set $\widehat{q}_t^R$ as missing.
  - In November 2008 there was unanimity on -150, which was not considered by respondents, and we set $\widehat{q}_t^R$ as missing.

December 2008 - March 2009: Mark IV Construction

We again have modal data.

- In December 2008 everyone vote to cut -100 and -100 was the lower bound on the modes. We set -100 as low vote and compute $\widehat{q}_t^R$ as all modes not equal to -100 over all modes.

- In January 2009 everyone votes to cut by -50 or -100 and these make up most modes. We set $\widehat{q}_t^R$ as the number of -50 modes over all modes equal to -50 or -100.

- In February 2009 everyone votes to cut -50 or -100. We take $\widehat{q}_t^R$ as the total number of modes not equal to -100 over the total number of modes.

- In March 2009 everyone votes to cut -50. We set -50 as the high vote and set $\widehat{q}_t^R$ as all modes -50 or greater over all modes.

C Monte Carlo Exercise

In this appendix we describe the Monte Carlo exercise that we carry out to explore the robustness of our estimation approach to individual heterogeneity in terms of θ and σ.

C.1 Design and Implementation

The first test of our estimator (case 1 below) is to determine its performance when our model exactly generates the data. In order to do this, we use our model to generate a dataset with 1350 observations with 11 internal members (serving 80 meetings each) and 12 external members (serving 50 meetings each), a division that roughly mimics our sample. We attribute to internal members $\theta_{\text{Int}} = 0.55$ and $\sigma_{\text{Int}} = 1.3$ and to external members $\theta_{\text{Ext}} = 0.4$ and $\sigma_{\text{Ext}} = 0.7$. To generate the jth vote of individual i, we proceed as follows:

1. we first draw q_{ij} from a $U\left[0.15, 0.85\right]$ distribution;
2. we then draw $\omega_{ij} \in \{0, 1\}$ from a Bernoulli distribution with parameter q_{ij};[31]
3. we draw s_{ij} from a $N(\omega_{ij}, \sigma_i^2)$ distribution
4. v_{ij} is determined by the voting rule derived in section 3.

We then run our estimator, with the generated values of q_{ij} and v_{ij} to estimate σ_{Int}, θ_{Int}, σ_{Ext}, and θ_{Ext}. We repeat this procedure 1,000 times storing the estimates as well as the true underlying values which started the procedure with.

In case 2 we generate the dataset exactly as above, except that we draw a θ and σ parameter for each individual in our sample. For internal (external) members, we draw θ from a $U[0.45, 0.65]$ ($U[0.3, 0.5]$) distribution and σ from a $U[1.0, 1.6]$ ($U[0.3, 1.0]$) distribution. We then run our estimator under the restriction that θ and σ are constant within the group of internals and externals.

In case 3 we proceed as in case 2 but introduce greater heterogeneity in the range of values that θ can take for external members. Specifically, we draw θ_{Ext} from a $U[0.2, 0.6]$ distribution. In this we interested in whether greater heterogeneity of preferences for the group of external members is picked up in our estimates of σ.

For each case we are interested in how far away the estimated values of σ_{Int}, θ_{Int}, σ_{Ext}, and θ_{Ext} are from their true values. In cases 2 and 3 we define $\sigma_{\text{Int}} = \frac{1}{11}\sum_1^{11}\sigma_i$ where i corresponds to the ith internal members. We apply similar definitions to the other three

[31]Notice that we draw a new state variable for each vote and thus do not construct a time series dimension; this should not effect the estimates in any systematic way and simplifies the programming.

Table C.1: Parameter Values for the Monte Carlo Exercises

	Case 1 Baseline		Case 2 Member Variation		Case 3 External heterogeneity	
	Internal	External	Internal	External	Internal	External
θ	0.55	0.4	0.45-0.65	0.3-0.5	0.45-0.65	0.2-0.6
σ	1.3	0.7	1.0-1.6	0.4-1.0	1.0-1.6	0.4-1.0
$\widehat{q}_t$	q_{ij}		q_{ij}		q_{ij}	

parameters of interest. For each of the 1,000 replications of each case, we calculate the following biases:

$$
\begin{aligned}
\text{Bias}[\widehat{\sigma}_{\text{Int}}] &= \widehat{\sigma}_{\text{Int}} - \sigma_{\text{Int}} \\
\text{Bias}[\widehat{\sigma}_{\text{Ext}}] &= \widehat{\sigma}_{\text{Ext}} - \sigma_{\text{Ext}} \\
\text{Bias}[\widehat{\sigma}_{\text{Int}} - \widehat{\sigma}_{\text{Ext}}] &= (\widehat{\sigma}_{\text{Int}} - \sigma_{\text{Int}}) - (\widehat{\sigma}_{\text{Ext}} - \sigma_{\text{Ext}}) \\
\text{Bias}\left[\frac{\widehat{\sigma}_{\text{Int}}}{\widehat{\sigma}_{\text{Ext}}}\right] &= \frac{\widehat{\sigma}_{\text{Int}}}{\widehat{\sigma}_{\text{Ext}}} - \frac{\sigma_{\text{Int}}}{\sigma_{\text{Ext}}} \\
\text{Bias}[\widehat{\theta}_{\text{Int}}] &= \widehat{\theta}_{\text{Int}} - \theta_{\text{Int}} \\
\text{Bias}[\widehat{\theta}_{\text{Ext}}] &= \widehat{\theta}_{\text{Ext}} - \theta_{\text{Ext}} \\
\text{Bias}[\widehat{\theta}_{\text{Int}} - \widehat{\theta}_{\text{Ext}}] &= (\widehat{\theta}_{\text{Int}} - \theta_{\text{Int}}) - (\widehat{\theta}_{\text{Ext}} - \theta_{\text{Ext}}) \\
\text{Bias}\left[\frac{\widehat{\theta}_{\text{Int}}}{\widehat{\theta}_{\text{Ext}}}\right] &= \frac{\widehat{\theta}_{\text{Int}}}{\widehat{\theta}_{\text{Ext}}} - \frac{\theta_{\text{Int}}}{\theta_{\text{Ext}}}
\end{aligned}
$$

C.2 Results

For each case, we report in Table C.2 the mean bias as well as select points in the distribution of the biases over the 1,000 replications (we use the 10th, 25th, 50th, 75th and 90th percentiles).

Our basic estimator accurately estimates all the various parameters of interest when there is no heterogeneity within types (case 1). Introducing individual heterogeneity (case 2), for which our estimator does not account, has a negligible effect on the point estimates of parameters or the differences and ratios between internal and external members. This is even the case when we introduce more preference heterogeneity among the external members (case 3); importantly for our analysis, such greater heterogeneity does lead us to wrongly identifiy informational differences.

Table C.2: Results of the Monte Carlo Exercises

Variable	**Mean**	**p10**	**p25**	**p50**	**p75**	**p90**
Case 1						
Bias$[\widehat{\sigma}_{\text{Int}}]$	-0.0	-0.2	-0.1	-0.0	0.0	0.1
Bias$[\widehat{\sigma}_{\text{Ext}}]$	-0.0	-0.2	-0.1	0.0	0.1	0.1
Bias$[\widehat{\sigma}_{\text{Int}} - \widehat{\sigma}_{\text{Ext}}]$	-0.0	-0.2	-0.1	-0.0	0.0	0.1
Bias$[\frac{\widehat{\sigma}_{\text{Int}}}{\widehat{\sigma}_{\text{Ext}}}]$	0.0	-0.4	-0.2	-0.1	0.1	0.4
Bias$[\widehat{\theta}_{\text{Int}}]$	0.0	-0.0	-0.0	0.0	0.0	0.0
Bias$[\widehat{\theta}_{\text{Ext}}]$	-0.0	-0.1	-0.0	-0.0	0.0	0.0
Bias$[\widehat{\theta}_{\text{Int}} - \widehat{\theta}_{\text{Ext}}]$	0.0	-0.0	-0.0	0.0	0.0	0.1
Bias$[\frac{\widehat{\theta}_{\text{Int}}}{\widehat{\theta}_{\text{Ext}}}]$	0.1	-0.1	-0.1	0.0	0.1	0.3
Case 2						
Bias$[\widehat{\sigma}_{\text{Int}}]$	-0.0	-0.2	-0.1	-0.0	0.0	0.1
Bias$[\widehat{\sigma}_{\text{Ext}}]$	-0.0	-0.2	-0.1	0.0	0.1	0.1
Bias$[\widehat{\sigma}_{\text{Int}} - \widehat{\sigma}_{\text{Ext}}]$	-0.0	-0.2	-0.1	-0.0	0.0	0.1
Bias$[\frac{\widehat{\sigma}_{\text{Int}}}{\widehat{\sigma}_{\text{Ext}}}]$	-0.0	-0.4	-0.2	-0.1	0.1	0.4
Bias$[\widehat{\theta}_{\text{Int}}]$	0.0	-0.0	-0.0	0.0	0.0	0.0
Bias$[\widehat{\theta}_{\text{Ext}}]$	-0.0	-0.1	-0.0	-0.0	0.0	0.0
Bias$[\widehat{\theta}_{\text{Int}} - \widehat{\theta}_{\text{Ext}}]$	0.0	-0.0	-0.0	0.0	0.0	0.1
Bias$[\frac{\widehat{\theta}_{\text{Int}}}{\widehat{\theta}_{\text{Ext}}}]$	0.1	-0.1	-0.1	0.0	0.1	0.3
Case 3						
Bias$[\widehat{\sigma}_{\text{Int}}]$	-0.1	-0.2	-0.1	-0.1	0.0	0.1
Bias$[\widehat{\sigma}_{\text{Ext}}]$	-0.0	-0.2	-0.1	-0.0	0.1	0.1
Bias$[\widehat{\sigma}_{\text{Int}} - \widehat{\sigma}_{\text{Ext}}]$	-0.0	-0.2	-0.1	-0.0	0.0	0.1
Bias$[\frac{\widehat{\sigma}_{\text{Int}}}{\widehat{\sigma}_{\text{Ext}}}]$	-0.0	-0.4	-0.2	-0.1	0.2	0.4
Bias$[\widehat{\theta}_{\text{Int}}]$	0.0	-0.0	-0.0	0.0	0.0	0.0
Bias$[\widehat{\theta}_{\text{Ext}}]$	-0.0	-0.1	-0.0	-0.0	0.0	0.0
Bias$[\widehat{\theta}_{\text{Int}} - \widehat{\theta}_{\text{Ext}}]$	0.0	-0.0	-0.0	0.0	0.0	0.1
Bias$[\frac{\widehat{\theta}_{\text{Int}}}{\widehat{\theta}_{\text{Ext}}}]$	0.1	-0.1	-0.1	0.0	0.1	0.3