BANK OF ENGLAND

Staff Working Paper No. 898

Uncertainty and voting on the Bank of England's Monetary Policy Committee

Alastair Firrell and Kate Reinold

December 2020

BANK OF ENGLAND

Staff Working Paper No. 898

Uncertainty and voting on the Bank of England's Monetary Policy Committee

Alastair Firrell[1] and Kate Reinold[2]

Abstract

Differences of opinion are a natural and vital part of monetary policy making by committee. With the appropriate stance for monetary policy both unobservable and uncertain, individual policymakers need to synthesise a wide range of information, including the views of other committee members. Using a novel measure of views that we construct from text analysis of the Bank of England Monetary Policy Committee's minutes and speeches, we show that both individual economic assessments and broader committee views are important in explaining individual voting. But in periods of high uncertainty both become more volatile and carry less weight in votes, consistent with the predictions of a simple voting model embedding a signal extraction problem. There is no increase in the dispersion of economic assessments in periods of uncertainty, nor in the mean dissent rate. Thus we show that interpreting the voting record as a reflection of policy uncertainty is unreliable, and highlight the value of individual committee members' communications — such as speeches — for conveying differences in view.

Key words: Central bank communication, committees, monetary policy, uncertainty.

JEL classification: D71, D81, E52, E58.

(1) Bank of England. Email: alastair.firrell@bankofengland.co.uk
(2) Bank of England. Email: kate.reinold@bankofengland.co.uk

The views expressed in this paper are those of the authors, and not necessarily those of the Bank of England or its committees. We are grateful to Saleem Bahaj, David Bradnum, Stephen Hansen, Sujit Kapadia, Lien Laureys, Riccardo Masolo, Michael McMahon, Roland Meeks, Paul Robinson, Misa Tanaka and Matthew Waldron for useful conversations, and Aidan Saggers for research assistance. We also thank participants at the 2019 EEA conference, the 2020 ESCoE conference on Economic Measurement, and an anonymous referee for comments that have materially improved the paper.

The Bank's working paper series can be found at www.bankofengland.co.uk/working-paper/staff-working-papers

Bank of England, Threadneedle Street, London, EC2R 8AH
Email enquiries@bankofengland.co.uk

ISSN 1749-9135 (on-line)

1 Introduction

'Disagreement among the [Monetary Policy] Committee is inevitable; it is also desirable because it represents the individual judgements of members, rather than an attempt to create a false consensus... Differences of view tell you more about the nature of the uncertainty confronting the MPC than the nature of the MPC itself.' King (2007)

Disagreement is seen in monetary policy committee deliberations across the world. The Bank of England's Monetary Policy Committee (MPC) averages almost one dissenting vote – a vote cast against the majority – per meeting. And differences in votes themselves are unlikely to capture the full range of opinions on a policy committee. But why do different policymakers favour different policy stances when they have the same remit? And might these mechanisms change during times of uncertainty?

A prominent hypothesis from the literature unpacking votes on the Bank of England's MPC is different assessments of the economy and so the appropriate policy rate.[1] Policymakers face a raft of incoming data, all measured with error. This creates a signal extraction problem in estimating the true state of the economy and from that the appropriate stance of policy, typically modelled using a Kalman filter or Bayesian updating. In such an environment, the views of other policymakers on a committee hold useful information about the state of the economy, and so play a role in individual voting. The relative weights assigned to different information about economic conditions would depend crucially on their signal-to-noise, and so could vary with the uncertainty of the environment. Moreover, depending on the source, periods of greater uncertainty could be – but is not always – associated with a change in the dispersion of views, and thus a change in dissenting votes.

But a key challenge in testing such predictions is that while we directly observe MPC member votes (the voting record is published after every meeting), individual assessments of economic conditions are unobserved.[2] The MPC publishes a 'best collective judgement' forecast for the economy that masks subtleties in individual MPC member views; these views are instead described qualitatively in MPC members' speeches and other communications.[3] Yet these differences could be quantitatively important for policy decisions.[4]

In order to fill this gap, we construct novel estimates of individual MPC-member economic sentiment from their speeches, as well as collective MPC sentiment from the minutes of the policy meetings. Starting with unstructured and multidimensional speeches, we distil a measure of the change in an individual policymaker's economic assessment. We identify combinations of economic terms and indicators of relative hawkishness/dovishness to create a *net change in sentiment index* for each speech made over the period 1997 to 2018. This gives us a numeric measure for each individual policymaker's view.[5] Applying the same technique to the more formal and structured MPC meeting minutes, we obtain a collective view.

We analyse the properties of these sentiment measures in 'normal' and uncertain times, and the role they play in MPC voting, using ordered probit regressions.[6] These measures corroborate

[1] E.g., Bhattacharjee and Holly (2005), Gerlach-Kristen (2006), Gerlach-Kristen (2008), Hansen et al. (2014).

[2] The FOMC publishes (unattributed) individual member forecasts. However, the FOMC is a more consensual committee than the Bank of England's MPC, with a relatively low dissent rate (Blinder (2004)).

[3] Weale (2015) and Stockton (2012) describe the use of MPC member speeches to express differences in view that are hidden by the 'best collective judgement' forecast.

[4] Hansen et al. (2014) model Bank of England MPC members' private information and find an important role for it in explaining policy votes. Indeed they find that private economic assessments play a larger role than differences in preferences.

[5] By interpreting the tone of speeches as a pure measure of an MPC members view of economic conditions, we abstract from the possibility that individual public communications are made more strategically, e.g. Vissing-Jorgensen (2019).

[6] We classify a quarter as 'uncertain' if either the option-implied volatility of the FTSE, or the disagreement of Consensus forecasters is higher than 0.75 standard deviations above its mean. We describe the selection of a

a number of results emerging from the simulation of a simple model of committee voting under noisy information (very close to Gerlach-Kristen (2008)). First, both individual and committee estimates of the state of the economy are important in explaining individual MPC member votes. This role for colleagues' views is consistent with the deliberative role of MPC discussions. Second, these estimates receive a smaller weight in uncertain times, consistent with the optimal downweighting of signals when volatility is higher. Individual estimates of the state of the economy are downweighted by somewhat more than the broader commitee estimate. Finally, we show that while estimates of the state of the economy become more volatile in uncertain times, the *dispersion* in views does not increase, and there is no significant difference in dissent rates. This pattern is predicted by our simulation analysis when the underlying economy is more volatile: optimal estimates of the state of the economy also become more volatile, but with no change to the relative information content of different signals, the dispersion does not increase.[7]

Differences in view and in votes are typically the focal point of media reporting of monetary policy decisions. Dissenting votes are used as a metric for whether the committee process is working: whether diverse perspectives are being combined to deliver better decisions. The quote that we began with from a former Bank of England Governor highlights the connection between uncertainty and disagreement. However, as our simulation exercise demonstrates, the relationship between uncertainty, estimates of the state of the economy and voting is nuanced. This means that drawing conclusions about committee dynamics from the vote alone is unlikely to be robust. Our text analysis shines a light on the breadth of views that underlie policy decisions and that may not always be captured in the vote.

Our paper relates to a number of other strands of literature. While our paper is novel in constructing individual Bank of England MPC members views from speech, the textual analysis of central bank communications as a means to uncover views on monetary policy is similar to Apel and Blix Grimaldi (2014), Malmendier et al. (2017) and Picault and Renault (2017) in their analyses of Sveriges Riksbank minutes, FOMC speeches and ECB press conferences respectively. We also apply the text measures to test predictions on committee behaviour through the lens of a small voting model in a novel way (rather than focussing on, for example, market reactions). A number of others have conducted econometric analysis of policymakers' voting records, such as Besley et al. (2008), Berk et al. (2010), Harris et al. (2011), positing the importance of internal or external status of committee members, career backgrounds, and preferences in explaining votes and dissents. And Bhattacharjee and Holly (2005), Gerlach-Kristen (2006), Gerlach-Kristen (2008), and Hansen et al. (2014) have all considered the implications of MPC members' individual economic assessments, with the latter attempting to model them explicitly. However, our direct measures of MPC member views allow us to control for a key missing variable in these studies. Finally, our work relates to Riboni and Ruge-Murcia (2010) and Riboni and Ruge-Murcia (2014) who develop models of committee voting and dissents, accounting for committee norms, such as a desire for consensus. We use their voting schema as the basis for our simulation analysis.

The remainder of the paper proceeds as follows. Section 2 presents a model of voting where the policymaker faces a signal extraction problem about the state of the economy, and notes its predictions for how estimates of the state and votes are formed in periods of high and low uncertainty. Section 3 discusses our data: the choice of uncertainty measure, the voting record of the MPC, methods for constructing measures of individual and collective views from the text of speeches and minutes, and the challenges addressed. Section 4 describes our econometric

measure of uncertainty in Section 3.

[7]Alternatively, the pattern of results could reflect an increase in the noise around estimates of the state accompanied by an increase in the weight placed on the views of others. The former, without the latter, would lead to more dispersion in views which we do not observe in practice.

strategy and results, and Section 5 concludes.

2 Committee voting and uncertainty

To frame our empirical analysis, we consider the predictions of a voting model that incorporates different committee member signals about the state of the economy. The key features of such a setup are that (i) the true state of the economy is not perfectly observed, and (ii) committee members each have different, noisy signals about the true state. From these it follows that (iii) the views of other committee members contain useful information about economic conditions.

These are reasonable assumptions. While policymakers on a monetary policy committee have access to the same data and staff analysis, they have different career backgrounds and limited bandwidth to analyse the wide array of incoming economic data. They may also have different models for how underlying variables like the output gap are determined, which naturally means more or less importance ascribed to different pieces of data. In such an environment, the views of colleagues on a committee hold useful information about appropriate monetary policy (Bhattacharjee and Holly (2005), Gerlach-Kristen (2008), Hansen et al. (2014)).

We take the setup of private information amongst committee members of Gerlach-Kristen (2008) as the basis for our exercises. The true state of the economy, S_t, is assumed to follow an autoregressive process, with persistence governed by ρ:

$$S_t = \rho S_{t-1} + u_t \tag{1}$$

and where $u_t \sim N\left(0, \sigma_u^2\right)$. Policymakers on the committee are unable to observe the true state of the economy. Policymaker i receives an unbiased but noisy signal of the state of the economy:

$$S_{i,t} = S_t + \nu_{i,t} \tag{2}$$

with $\nu_{i,t} \sim N\left(0, \sigma_\nu^2\right)$.[8] She can also observe the signals of other policymakers, but with additional noise, capturing the idea that her colleagues are unable to perfectly communicate their view of the world to her (nor can she to them). Policymaker i's view of policymaker j's signal, $S_{ij,t}$, is therefore subject to additional noise, $\omega_{ij,t}$:

$$S_{ij,t} = S_t + \nu_{j,t} + \omega_{ij,t} \tag{3}$$

where $\omega_{ij,t} \sim N\left(0, \alpha\sigma_\nu^2\right)$. We assume that $\alpha > 0$, implying that the noise around policymaker i's perception of policymaker j's signal is greater than that around her own. The error in communications and the noise around the signals are assumed to be uncorrelated.

In the face of multiple noisy signals of the state of the economy, the policymaker faces a signal extraction problem. Given the autocorrelation in the process, Gerlach-Kristen (2008) describes that the policymaker optimally employs a Kalman filter in determining her view of the state of the economy, $\hat{S}_{i,t}$. She weights the information from her own view of conditions with those of others on the committee according to their relative signal-to-noise.

Key parameters for the weight that different signals receive are σ_ν, α and σ_u. Greater noise in committee members' signals σ_ν reduces their informativeness about the true state of the economy, and therefore their optimal weight. Given that the increase in noise affects all signals, an increase in this parameter means that the policymaker places less weight on her own signal, but also the signals of others in forming $\hat{S}_{i,t}$. By contrast, a higher value of α implies that the views of others are observed with more noise (with no effect on the noise of her own signal) and therefore shifts the relative weight that the policymaker places on her own signals and others'.

[8] We assume that σ_ν is common to all committee members. Hansen et al. (2014) allow for different committee members to have different levels of signal precision, capturing the possibility of differing expertise.

A higher value for σ_u implies more volatility in the true state of the economy. A key difference from greater noise around signals is that this higher volatility would ideally be reflected in a higher volatility of policymaker i's estimate, $\hat{S}_{i,t}$.

We assume that the policymaker i selects her preferred interest rate $i^*_{i,t}$ based on her estimate of the state of the economy, $\hat{S}_{i,t}$, according to a multiplier β and with a preference for some persistence in rates (governed by θ).

$$
\begin{aligned}
i_{i,t} &= \theta i_{i,t-1} + (1-\rho_i)\,\theta \hat{S}_{i,t} \qquad (4)\\
i^*_{i,t} &= c_i + i_{i,t} \qquad (5)
\end{aligned}
$$

We allow for a policymaker-specific intercept, c_i, indicating that policymakers could vote differently because of differences in preferences as well as differences in economic assessments.[9] By assuming that the policymaker responds to a composite measure of the state of the economy, we abstract from the potential for multiple committee objectives that present trade-offs (for example as described in the remit of the Bank of England's MPC). These could warrant a different response to inflation and the output gap.

In order to capture two features of real-world committee policymaking – discrete votes and a desire for consensus – we follow the voting schema of Riboni and Ruge-Murcia (2010) and Riboni and Ruge-Murcia (2014). Policymakers come to their individual view on the appropriate interest rate according to equation 5. But votes are made in 25bps increments, and a supermajority is required for an interest rate decision to pass. Policymakers who disagreed with the passing policy rate do not dissent unless their difference in view is sufficiently large, leading to regions of inaction.[10]

While these arrangements do not necessarily reflect the *de jure* arrangements of a monetary policy committee, they can capture the *de facto* ones. By statute, the Bank of England's MPC operates under a majority voting system. But Riboni and Ruge-Murcia (2014) find that, conditional on their model, the voting record suggests some desire for consensus, such that a supermajority of one is required for a vote to pass. Similarly, Apel et al. (2010) cite a 'bargaining margin' that Riksbank MPC members operate under, driven by an effort to establish public confidence or avoid unease in financial markets.

2.1 Simulating votes under uncertainty

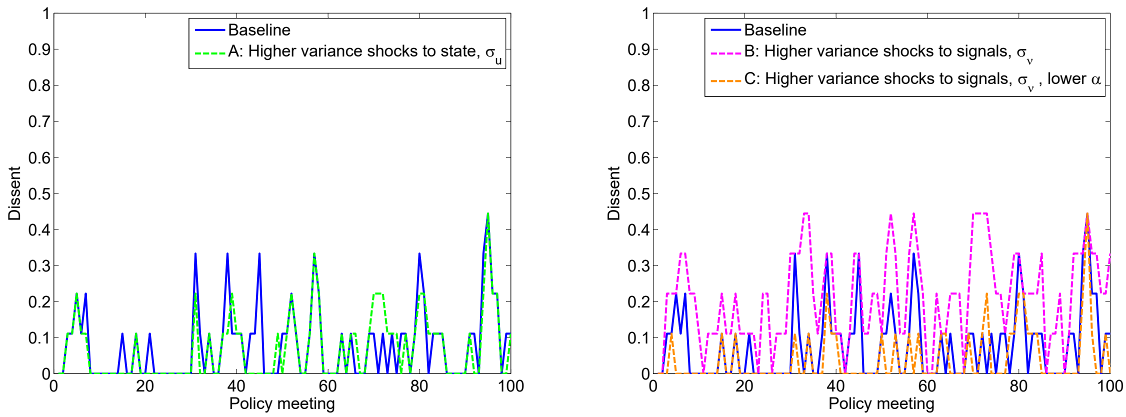
From a policymaker's perspective, a period of high uncertainty might be one in which underlying economic conditions are more volatile, but also one in which it is more difficult to extract a signal about the true state of the economy. To demonstrate the impact of changes in the volatility of the state of the economy or of noise around different signals, and to provide a benchmark against which to compare our empirical analysis, we simulate data from the model.

[9] Riboni and Ruge-Murcia (2014) show that differing degrees of asymmetry over inflation in policymakers' loss functions – i.e. hawkish or dovishness – enter as an intercept in the interest rate rule like this. While central banks typically have symmetric inflation targets, Hansen et al. (2014) argue that some asymmetry could reflect views of the appropriate trade-off between inflation and unemployment. And there is some empirical evidence to support an asymmetry in monetary policy reaction function. Surico (2003) explores the possibility using aggregate data. Besley et al. (2008) find differences in MPC votes consistent with differences in the intercept, compared to an alternative hypothesis of responsiveness to inflation relative to activity.

[10] Mechanically, the simulated MPC first votes on the direction of the interest rate change by majority. The magnitude of the change is then decided by gradually increasing the size of the change (in 25bps), until the size can no longer maintain a consensus. Policymakers then decide whether or not to dissent, depending on whether their difference in view exceeds some threshold (around 75bps). Similarly, Gerlach-Kristen (2008) assumes that a consensus of 60% is required, and that policymakers only dissent if preferred interest rates are sufficiently different. In this voting schema, it is possible to see $\geq 50\%$ dissent.

We simulate the model for 250 meetings, for a committee of $n = 9$ members, with a rotating membership. The model is calibrated to deliver simulated interest and dissent rates approximately in line with the MPC's voting record over our sample (1997 to 2018).[11]

Figure 1: Simulated dissent rates under alternative model calibrations



(a) Greater volatility of the true state (b) Greater noise to signals about the state

Notes: In the green line of panel (a), σ_u has been doubled relative to the baseline calibration. And in the magenta line of panel (b), σ_ν is doubled, while the orange line includes a doubling of σ_ν and a lower value of α (to 0.1).

First, we report the dissent rate implied by the same sequence of simulated shocks for different calibrations of the model (Figure 1)[12]. These calibrations are intended to capture shifts in periods of high uncertainty, driven either by an increase in the variance of the true state, or in the noise around particular signals. The blue lines in both panels report the dissent rate for the baseline calibration of the model.

The green line in the first panel reports simulated data from the model where the variance of shocks to the true state (σ_u) has been doubled (calibration A). This increases the variance of the true state, but not the dispersion between the estimates of different policymakers and so their votes (column 2, Table 1). There is therefore relatively little effect on the dissent rate.

By constrast, where higher uncertainty is modelled as an increase in the variance of the noise around signals (σ_ν), the dissent rate rises (magenta line in second panel, calibration B). The variance of policymakers' estimates of the state increase by a similar amount to calibration A (third column, Table 1), but they are also much more dispersed, leading to greater dispersion in votes (second column).

In the final calibration, the orange line in the second panel of Figure 1 shows the dissent rate when the variance of the shocks to the signals is increased by the same degree as calibration B, but α falls to 0.05 (calibration C). A lower α makes the signals of others a more reliable measure, increasing their signal-to-noise relative to calibration B. Despite the greater noise in the signals, the dissent rate falls as the policymaker places more weight on the views of others.

Finally, we look at regressions on data simulated from the model, which we will replicate with the actual MPC record in our empirical analysis. For the simulated dataset of 250 meetings, we assume that 30% occur in a period of higher uncertainty (in line with the uncertainty measure that we will discuss in Section 3), approximated by a change in calibration in each of the three ways described above. We adopt an ordered probit specification where the dependent variable is a binary indicator, $V_{i,t}$, of the change in policymaker i's vote, $\bar{i}_{i,t}$ (which once we've applied the voting schema, will be different from $i^*_{i,t}$). We code the dependent variable as -1, 0 and 1

[11] In particular, $\rho = 0.95$, $\sigma_u^2 = 0.000015$, $\sigma_\nu^2 = 0.00125$, $\alpha = 0.25$, $\beta = 2.5$, $\theta = 0.8$. The individual-specific intercepts are randomly drawn from a zero-mean uniform distribution with a width of 50bps. The new incumbent inherits their estimate of the state from their predecessor.

[12] Dissent, being votes against the majority, cannot exceed 4 out of 9 members ≈ 0.44.

Table 1: Ratio of standard deviation of signals and votes under higher uncertainty to baseline calibrations

	Stdev i_t	Stdev $i_{i,t} - i_t$	Stdev $\hat{S}_{i,t}$	Stdev $S_{i,t}$	Stdev $S_{ij,t}$
A: Higher var shocks to state, σ_u	1.37	1.00	1.22	1.05	1.04
B: Higher var shocks to signals, σ_ν	1.05	1.40	1.20	1.38	1.38
C: Higher σ_ν, lower α	1.05	0.89	1.16	1.38	1.30

representing looser, no change, tighter.

$$V_{i,t} = \begin{cases} -1, & \bar{i}_{i,t} < \bar{i}_{i,t-1}, \\ 0, & \bar{i}_{i,t} = \bar{i}_{i,t-1}, \\ 1, & \bar{i}_{i,t} > \bar{i}_{i,t-1}. \end{cases}$$

The ordered probit specification is frequently used in the literature for modelling committee voting given the use of discrete votes (e.g. 25bps) and the prevalence of 'no change' votes in the voting record. It allows for a non-linearity in how the explanatory variables affect the probability of different outturns.[13]

For the explanatory variables, we take the change in member i's estimate of the sentiment measure, $\Delta\hat{S}_{i,t}$, and construct a committee-wide view by averaging across the change in the estimate of each committee member, $\sum_{j=1}^{n}\frac{1}{n}\Delta\hat{S}_{j,t}$. A dummy variable U_t^H takes a value of one in the quarters in which we have changed the calibration to mimic higher uncertainty. We control for the policymaker's lagged vote, and the lagged passing vote. And we allow for interactions between sentiment and the uncertainty dummy variable, using the following specification:

$$V_{i,t} = \alpha_i + \rho_1 V_{i,t-1} + \rho_2 V_{t-1} + \left(\beta_1 + U_t^H\beta_2\right)\left(\Delta\hat{S}_{i,t} - \sum_{j=1}^{n}\frac{1}{n}\Delta\hat{S}_{j,t}\right) + \left(\beta_3 + U_t^H\beta_4\right)\sum_{j=1}^{n}\frac{1}{n}\Delta\hat{S}_{j,t} + U_t^H + \eta_{i,t} \tag{6}$$

Table 2 reports the results. In the first column, the variance of the shocks and signals is in line with the baseline calibration and the same for the full sample. The committee member's sentiment measure and the committee composite are both positive and significant in explaining votes, with the latter receiving a higher weight. In the remaining three columns, we allow for a period of higher uncertainty approximated by a change in the calibration in the three ways described above (A: higher variance of shocks to true state, σ_u; B: higher variance of shocks to signals, σ_ν; C: higher variance of shocks to signals and lower α).

For all of the calibrations the high uncertainty dummy is insignificant. The interaction between the dummy variable and the two sentiment measures is uniformly negative, indicating a lower weight being placed on own estimates of the state of the economy and those of others in periods of high uncertainty. The significance for the individual sentiment measure tends to be weaker (which is also true across other simulated data samples). For all specifications, taking into account the shifts in periods of higher uncertainty, the MPC sentiment measure tends to receive a higher weight than individual differences. While Table 1 showed that a lower α reduces the dispersion of signals, the final column in Table 2 shows that there is little impact on the relative weight on own views and those of others.

2.2 Predictions

Together, these simulation exercises suggest a number of predictions about voting in an environment of noisy signals about the state of the economy:

[13] It will also ease the analysis in the period that the MPC was voting over both interest rates and unconventional policy tools, given that the coding of the variable is naturally extended for multiple instruments.

Table 2: Ordered probit regression of policy vote on economic sentiment and uncertainty, from simulated data

	Baseline	A: Higher σ_u	B: Higher σ_ν	C: Higher σ_ν, lower α
Member i sentiment relative	6.10***	6.38***	7.95***	7.39***
to MPC ($\Delta\hat{S}_{i,t} - \sum_{j=1}^{n} \frac{1}{n}\Delta\hat{S}_{j,t}$)	(0.31)	(0.37)	(0.46)	(0.43)
MPC sentiment	7.05***	7.82***	9.70***	9.17***
($\sum_{j=1}^{n} \frac{1}{n}\Delta\hat{S}_{j,t}$)	(0.27)	(0.29)	(0.35)	(0.34)
High uncertainty (U^H)		-0.04	-0.08	0.05
		(0.13)	(0.13)	(0.13)
U^H * member i sentiment		-0.70	-1.89***	-1.10
rel. MPC		(0.61)	(0.70)	(0.81)
U^H * MPC sentiment		-1.54***	-3.20***	-2.82***
		(0.40)	(0.45)	(0.42)
Lagged vote, member i	0.48***	0.49***	0.49***	0.47***
($V_{i,t-1}$)	(0.05)	(0.05)	(0.05)	(0.05)
Lagged policy change	0.78***	0.78***	0.74***	0.76***
(within V_{t-1})	(0.05)	(0.05)	(0.05)	(0.05)
Lower cut-off	-0.61**	-0.62**	-0.59*	-0.62**
	(0.29)	(0.31)	(0.30)	(0.30)
Upper cut-off	0.15	0.10	0.07	0.09
	(0.29)	(0.31)	(0.30)	(0.30)
Member fixed effects	Yes	Yes	Yes	Yes
Sample	2,249	2.249	2,249	2,249

Notes: *, ** and *** indicate statistical significance at the 10%, 5% and 1% levels respectively. Individual and MPC sentiment have been standardised to be consistent with the measures from text. Standard deviations are clustered by MPC member and meeting. MPC members for whom there are fewer than 3 votes are dropped from the sample.

(i) Both individual and committee estimates of the state of the economy are positive and significant in explaining member votes. This is despite the fact that the committee member has already weighted the views of others when forming their estimate of the state of the economy.

(ii) A lower weight is placed on individual and committee-wide estimates of the state of the economy in periods of higher uncertainty in our voting regression, irrespective of the calibration of high uncertainty.

(iii) The implications for the dispersion of views and dissent *depends on the source of the uncertainty*. Greater volatility in the true state makes estimates of the state more volatile, but their dispersion does not increase, and nor does dissent. Noisier signals about the state of the economy do increase the dispersion in committee estimates of the state and see higher dissent rates. But a shift in the perception of relative noise (α) can counterbalance this.

Finally, it is worth reiterating some features of our setup. While signals are noisy, policymakers are assumed to know the distributions from which shocks are drawn, and to understand whether a period of higher uncertainty (as we define it here) is driven by a higher volatility of the state or the noise around signals. Moreover, it assumes that policymakers weight signals optimally. It does not account for the possibility that policymakers might shift the relative weights, including α, strategically. For example, if it is better to be conventionally wrong than unconventionally right, policymakers could downweight their own view by more than they optimally would (Baddeley (2010), Sibert (2006)). On the other hand, reputational considerations could lead to policymakers 'anti-herding', or taking more extreme positions than might be expected (Rulke and Tillmann (2011)).

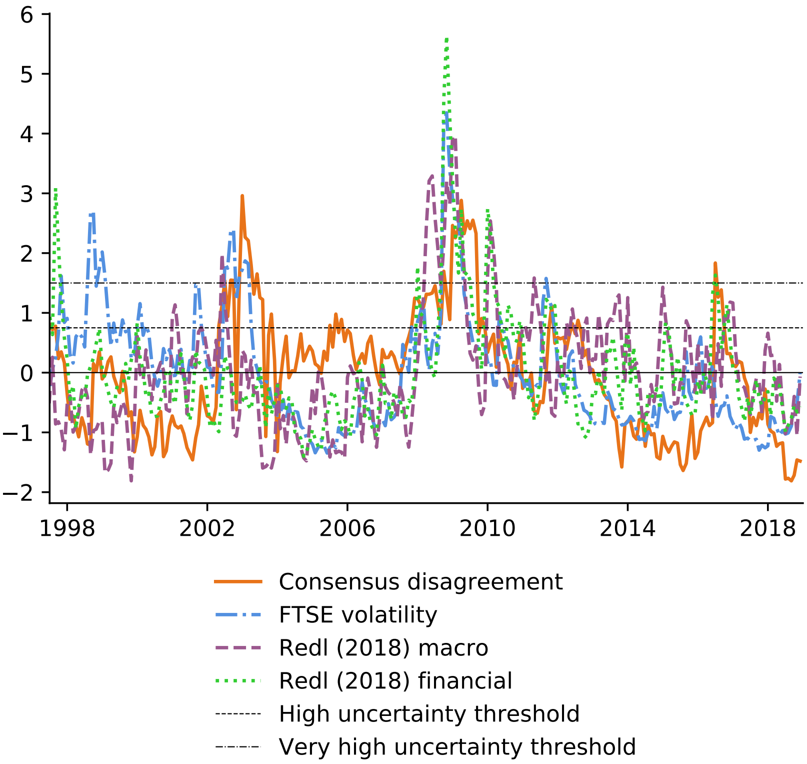
3 Data and Measurement Approach

The model described above offers predictions for the patterns we might expect between individual and committee economic assessments and voting, in periods of high and low uncertainty. Yet there are empirical challenges to testing these predictions. Uncertainty is a difficult concept to measure, and private economic assessments are unobservable. In this section, we describe the datasets we use and our approach to preparing them for our econometric exercises.

3.1 Measuring uncertainty

Measuring economic uncertainty is notoriously difficult. Since the financial crisis, there has been renewed interest in the concept, and a number of metrics for changes in the second-moment have been identified. Redl (2019) splits the approaches into observable proxies (such as news citations, realised stock volatilities, or disagreement between professional forecasters), and those derived from econometric techniques such as Jurado et al. (2015). The latter have the advantage of better controlling for the first-moment effects of shocks, while the former have the benefit of simplicity, real-time observability and ubiquity in the literature.

Figure 2: Measures of macroeconomic uncertainty



Notes: All series have been standardised by their mean and standard deviation. Consensus disagreement (orange) is measured as the standard deviation of forecasts for GDP growth by the Consensus panel of forecasters, averaged over the one and two year horizons (Consensus Economics (2018)). FTSE volatility (blue) measures the option-implied volatility of the FTSE All Share index. Redl (2019) macro and financial uncertainty are Jurado et al. (2015)-style measures of uncertainty, constructed as the unforecastable component common to a large dataset of macroeconomic and financial variables. The high (very high) uncertainty threshold is defined as a value 0.75 (1.5) standard deviations above the series means.

Figure 2 plots a range of uncertainty metrics for the UK. Consensus disagreement (orange) is measured as the standard deviation of forecasts for GDP growth by the Consensus panel of forecasters, averaged over the one and two year horizons (Consensus Economics (2018)). Forecast disagreement such as this is a commonly-used proxy for uncertainty. However, conceptually, disagreement and uncertainty are different: differences in the mean forecasts of individuals do not necessarily coincide well with the distributions that individuals perceive. Whether disagreement is well-correlated with uncertainty empirically is a matter of debate, as evidenced from the conflicting evidence from Boero et al. (2008), Boero et al. (2015). FTSE volatility (blue dash-dot) measures the option-implied volatility of the FTSE All Share index, which is a commonly-used proxy popularised by Bloom (2009). Again, this measure correlates well with other metrics, but is likely to be driven by changes in risk premia as well as changes in uncertainty. Alongside these we show a macro uncertainty measure (purple dashed) and a financial uncertainty measure (green dotted) from Redl (2019). These are constructed, in line with Jurado et al. (2015), as the unforecastable component common to a large dataset of macroeconomic and financial variables.

These series are all positively correlated, but to differing degrees. They offer slightly different perspectives on the timings of high uncertainty in the UK. All of the measures register very high uncertainty during the financial crisis (the peak of all four series is either in 2008 or 2009). Most of the series register heightened uncertainty around the Dot-Com bust of 2002-2004, and again around the UK's Referendum on EU membership in 2016. The period of lowest uncertainty on most of the measures occurred between around 2004 and 2007. There is some difference in the extent to which the measures identify heightened uncertainty around the euro-area crisis of 2011-12 and in the period immediately after the MPC's inception (with the Asian crisis and the LTCM default).

For our analysis we construct dummy variables for periods of high and very high uncertainty, rather than using the raw series. Given the variation in episodes identified by the uncertainty measures, we opt to combine the signal from two commonly-used measures in the literature, FTSE volatility and Consensus disagreement. We register uncertainty as high (very high) when either measure sits more than 0.75 (1.5) standard deviations above its mean – the horizontal lines in Figure 2. These definitions give a reasonable proportion of high and very high uncertainty episodes (around 30% and 10% respectively of the sample). In Section 4 we demonstrate robustness of our analysis to alternative uncertainty measures.

3.2 Measuring votes

The Bank of England's nine-member MPC was established in 1997 to set monetary policy for the UK.[14] MPC member votes are made public and attributed in the minutes that accompany each policy meeting. Since the financial crisis, where multiple monetary policy instruments have been in operation, the MPC has held multiple votes at each policy meeting.

We create a composite vote variable ($V_{i,t}$) from these public records. This indicates whether the vote made by member i at time t was looser than their vote at the previous meeting, unchanged or tighter, coded as -1, 0 and 1 respectively. This could be delivered by a vote for a different level of Bank rate ($i_{i,t}$) relative to their preference at the previous meeting ($i_{i,t-1}$), or a change in the size of the Asset Purchase Facility ($A_{i,t} - A_{i,t-1}$):[15]

$$V_{i,t} = \begin{cases} -1, & i_{i,t} < i_{i,t-1} \text{ or } A_{i,t} > A_{i,t-1}, \\ 0, & i_{i,t} = i_{i,t-1} \text{ and } A_{i,t} = A_{i,t-1}, \\ 1, & i_{i,t} > i_{i,t-1} \text{ or } A_{i,t} < A_{i,t-1}. \end{cases}$$

We also construct an indicator for when a member has made a dissent (a vote against the majority), $D_{i,t}$:

[14] At times, the committee has been smaller, for example at the committee's inception, and at points since between the appointment of new members.

[15] While technically possible when multiple policy tools are in use, no MPC member has voted for a tightening using one policy instrument and a loosening with another over our sample.

$$D_{i,t} = \begin{cases} 0, & i_{i,t} = i_t \text{ and } A_{i,t} = A_t, \\ 1, & i_{i,t} \neq i_t \text{ and/or } A_{i,t} \neq A_t. \end{cases}$$

3.3 Constructing measures of MPC views from text

To measure the views of MPC members individually and committee-wide, we analyse two main communication vehicles: the minutes of committee meetings and each individual's speeches.[16] The minutes summarise the policy meetings, including the MPC's collective assessment of the latest data, views expressed in the meeting (unattributed) and the outcome of the policy vote.[17] Speeches are made by individual MPC members, and may focus directly on current monetary policy or span broader topics.[18] Table 3 summarises these different objectives, formats and linguistic styles.

Table 3: Characteristics of MPC minutes and speeches

	Minutes	**Speeches**
Purpose	Explanation of the outlook, policy issues and decision	Varied, some explanation of view/votes
Topics	Consistent, always monetary policy	Varied
Individuality	Some mention, unattributed	Wholly attributed
Structure	Consistent, standard sections	Varied
Language	Formal, consistent	More informal, inconsistent
Frequency	8 times per year (from 12)	c.2-5/member/year (rising)
Timing	Regular timing, tied to vote	Irregular
Count (Jun 97 - Dec 18)	251	755

We analyse the period from the inception of the MPC in June 1997 to December 2018. The MPC met monthly from 1997 to 2016, and 8 times per year since then (excepting a special meeting in 2001), yielding 251 minuted meetings and 755 speeches made by 40 MPC members. The frequency of speeches has increased over the committee's existence, with an increasing share made by MPC members other than the Governor, so making them more representative of committee views over time (Appendix A reports speeches per year and per MPC member).

3.3.1 Approach

We aim to extract measures of policymakers' inclination towards contractionary/expansionary policy: hawkishness/dovishness. To locate hawkish and dovish sentiment in minutes and speeches, we identify the parts of the speech that focus on monetary policy topics, and then assess the tone in which they are discussed. In particular, we aim to identify *references to changes in economic conditions* that would occasion a change in voting pattern.

We take an automated approach, creating a central bank policy-specific dictionary similar to that of Apel and Blix Grimaldi (2014).[19] In contrast to studies that search for individual words to indicate tone, our dictionary comprises two-word combinations – *bigrams*. By combining a salient monetary policy topic and an indicator of direction or sentiment (approximately, a noun and an adjective), we obtain more precise context for the sentiment expressed. The

[16] A third is the quarterly *Inflation Report*, now *Monetary Policy Report*, and economic projections published as part of that.

[17] Following the Warsh review of MPC transparency (Warsh (2014)), transcripts of policy meetings will be published with an 8-year lag, beginning with the March 2015 meeting. (Bank of England (2014))

[18] Such as Japan's Great Recession (Posen (2010)) or robots and job automation (Haldane (2015))

[19] Historically, as described in Bholat et al. (2015), text analysis has used manual classification (e.g. Ehrmann and Fratzscher (2007)). More recently, studies have utilised automated text analysis techniques to count appearances of keywords, or to automatically extract topics of interest (including Tetlock (2007), Apel and Blix Grimaldi (2014), Nyman et al. (2015), Baker et al. (2016), Malmendier et al. (2017)).

nouns represent key concepts for policymakers: inflation, unemployment, demand, output, spare capacity and so on.[20] To these we add adjectival words to provide direction: e.g. higher, gain, reduction, drop. In both cases, words are stemmed to allow us to find variations: inflationary, inflating; reduced, reducing; speed up, sped up; and so on. The full list of nouns and adjectives is given in Tables 7 and 8 in Appendix B.

Each bigram is categorised as hawkish (associated with contractionary policy) or dovish (expansionary). Thus 'higher inflation', 'lower unemployment', and 'pick up (in) production' are classified as hawkish; 'lower gdp', 'declining consumption', 'increased spare capacity' are classified as dovish. Having constructed our categorised dictionary of bigrams, we apply it to each sentence of each committee minute and speech. Counting the number of hawkish and dovish bigrams per document, we calculate an index of net change in sentiment for each minute (S_t^{MPC}) and speech (S_{it}):

$$S = \frac{N_{hawk} - N_{dove}}{N_{hawk} + N_{dove}}$$

The index can range from -1 (entirely dovish) to 1 (entirely hawkish). Zero indicates an even balance between hawkish and dovish sentiment.

3.3.2 Considerations in constructing and applying the dictionary

The approach is very similar to Apel and Blix Grimaldi (2014), however, we use a larger list of both nouns and adjectives, because MPC members' speeches use a much more varied vocabulary than the more formal minutes. In constructing the dictionary we extract the words co-occurring with our chosen policy-relevant nouns, and from those we select 'adjectives' giving a direction of change. Monetary policy 'nouns' are included only where their direction in a policy rule is unambiguous.[21]

We select directional 'adjectives' which represent changes rather than levels: 'increase', 'fall', 'higher' rather than 'high', 'low'. As described in Section 2, focussing on changes allows us to combine votes over multiple policy instruments, is consistent with the approach in the literature, and deals with the large proportion of votes that are for 'no change'. Moreover, the speeches in our sample have many more references to changes than levels: 490 speeches had more than ten change-bigram hits; only 225 speeches had more than ten levels-bigram hits. Working in changes increases the confidence in our sentiment measures.

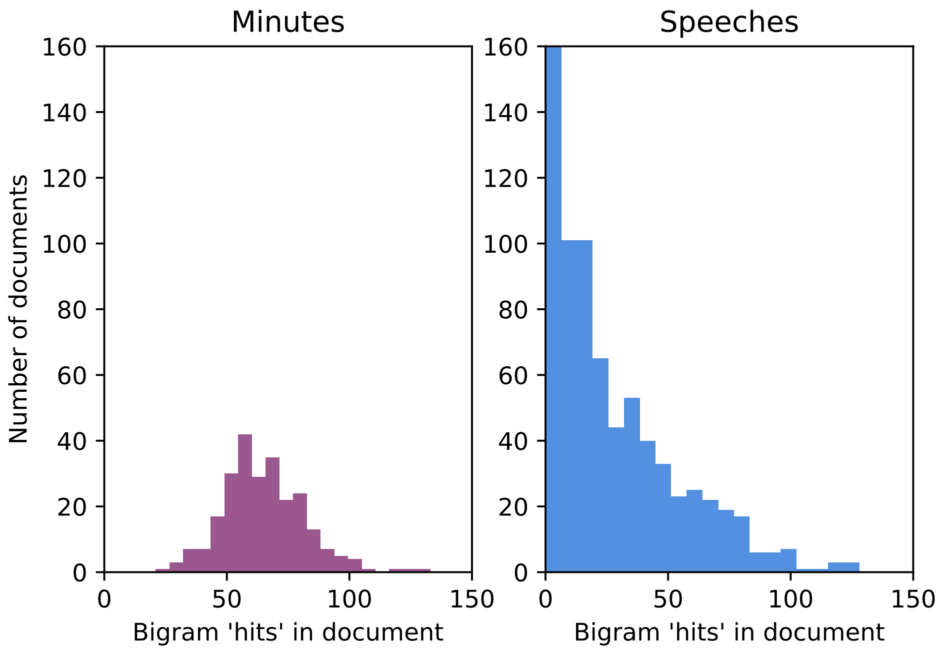
Speeches that do not discuss monetary policy issues are not relevant to our analysis. We see from Figure 3 (right-hand panel) that a significant tail of speeches contain very few instances ('hits') of the monetary policy bigrams. Contrast this with the hits for the minutes (left-hand panel), all of which we know to be addressing monetary policy. These low hit-rate speeches include those on financial stability or prudential regulation topics, where MPC members have additional responsibilities in those areas, plus remarks and speeches addressing members' wider interests. Therefore we eliminate any speech containing ten or fewer bigrams from the analysis to exclude non-monetary policy speeches.[22]

[20]Two-word phrases are treated as a single noun. Examples include 'claimant count', 'world trade', 'excess demand'.

[21]For this reason, asset prices and other financial market indicators are omitted. Additionally, 'growth' is omitted as it may represent an adjective or a noun: "reduced (economic) growth" vs "reduced growth in prices"; the same applies to 'recovery' and 'development'.

[22]To inform our threshold, we conducted a simulation exercise to approximate the process by which MPC members randomly draw a sample of bigrams (either hawkish or dovish) to represent their true view. Where the sample is very small, e.g. fewer than three bigrams, the standard deviation of the sampled statistics can be up to nine or ten times the asymptotic value (if an MPC member could use thousands of bigrams in their speeches), but it falls back sharply. Drawing 10 bigrams delivers a standard deviation of the test statistic of around three times the asymptotic value, offering a good balance between signal-to-noise ratio and maintaining sample size.

Figure 3: Bigram counts in minutes and speeches



In order to locate non-contiguous bigrams, e.g. "*unemployment* is *rising*", "the *rising* level of *unemployment*', we remove non-affecting stopwords which might come between our pairs, and allow for both noun-adjective and adjective-noun configurations. This lets us pair non-contiguous words if the filling is not significant (and if it does not cross sentence/clause boundaries), while minimising the chance of erroneously pairing an adjective and a noun which do not belong together.[23] Finally, negations in sentences are especially difficult to interpret correctly, so we exclude from the sample all sentences which contain negations. More detail on all text preparation steps is provided in Appendix B.

3.3.3 Features of sentiment measures

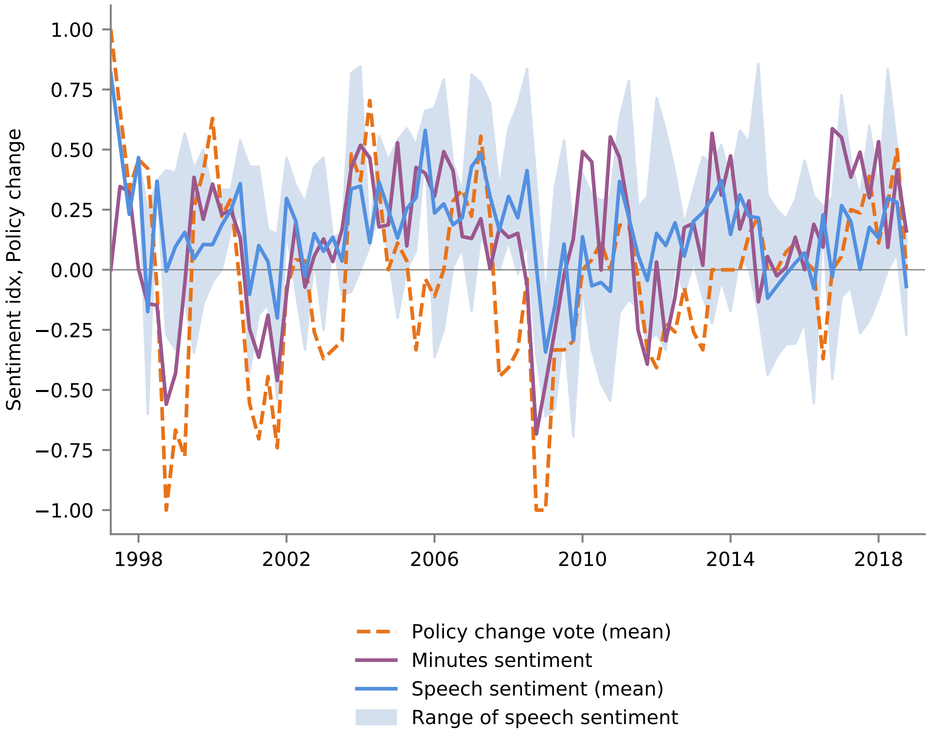
Figure 4 charts the sentiment measured in the minutes and speeches against a measure of the mean policy vote, aggregated to a quarterly frequency. It reveals a comovement between the mean of the speech sentiment in a quarter (blue line) and the minutes sentiment (purple line), as well as the average policy vote (dashed orange line). This is reassuring, since it indicates that in periods when the MPC were voting to tighten policy, our measure shows greater hawkishness in their minutes and speeches.

We note a few features of the measures:

- On average over the sample, the sentiment measure from the minutes is slightly positive (at 0.09) compared to an average policy vote that is a little negative (-0.05).
- Minutes sentiment is also fairly persistent, despite capturing changes in policy votes and changes in the view of economic conditions: the auto-correlation for the monthly minutes sentiment measure made around these speeches is 0.62. However this is also true of the change in the monthly policy vote, with an autoregressive coefficient 0.54.
- While speeches are given infrequently (and so we cannot look at the persistence of any individual member's measure), there is again some persistence in the average speech senti-

[23] We do not want a sentence like "Data collection frequency is higher but unemployment figures are unaffected" to yield a match against the 'higher unemployment' bigram. This would occur with a simple n-words-either-side pattern.

Figure 4: Sentiment from MPC minutes and members' speeches



Notes: Only speeches with more than ten bigrams are reported. The minutes sentiment is the mean from the two or three meetings on the quarter. The speech sentiment is the mean / minimum / maximum over the quarter as a whole.

ment measure. There is a higher correlation between the minutes and the policy vote than the average speech sentiment and the policy vote (Figure 7 in the Annex). The minutes are also better correlated with leads of the policy vote.

- There is dispersion in the sentiment of speeches across the whole sample, visible in the range in Figure 4. This can reflect both the spread of views about economic conditions and the noise in the speeches' signal of views. The number of bigrams in each speech is typically lower than for the minutes, as speeches offer a less focussed discussion of economic conditions than the minutes.

To validate our text strategy and understand the likely magnitude of measurement error, we check the sentiment in speeches made close together by the same policymaker. In Figure 8 of the Appendix, we report the sentiment in speeches made within 10 days and 30 days of one another. The correlations are 0.61 and 0.43 respectively, offering reassurance that there is a fair common signal in the speeches.

4 Empirical Analysis of MPC Sentiment and Voting

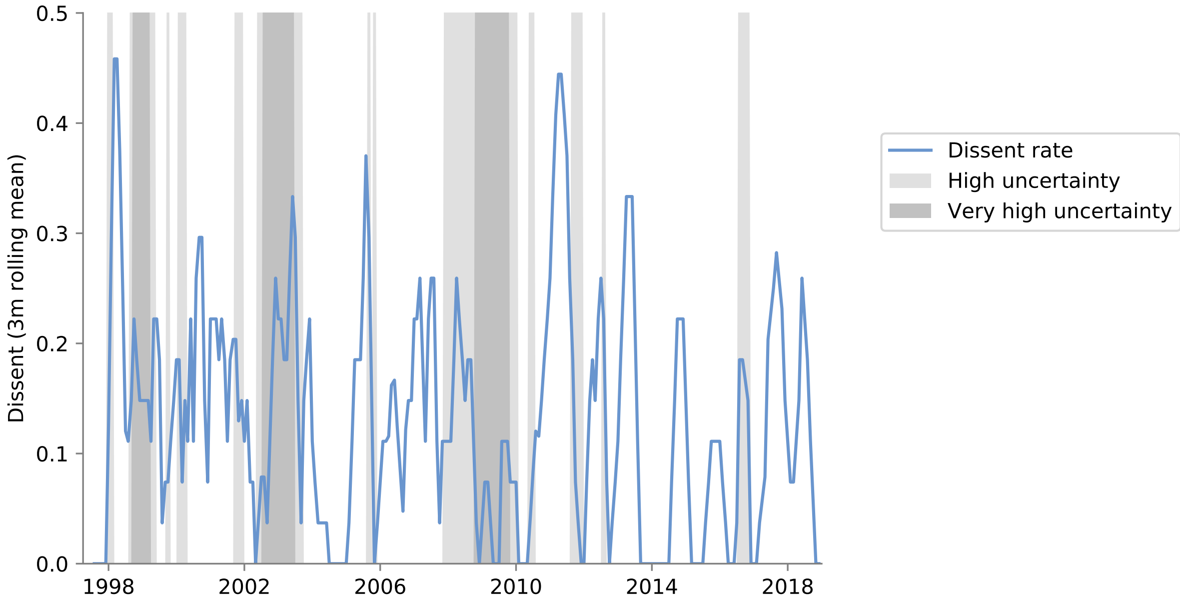
We now turn to an empirical analysis of MPC voting. First, we describe the properties of our sentiment measures and votes. Second, we describe our empirical strategy for associating the two. Finally, we describe our results.

4.1 Properties of sentiment and voting in periods of high uncertainty

In Section 2 we reported the properties of dissent and estimates of the state from *simulated* data where higher uncertainty is proxied by different calibrations of shock processes (Figure 1 and Table 1). Here, we construct the equivalent statistics for the actual MPC voting record and our measures of economic sentiment.

Figure 5 reports the share of dissenting votes on the Bank of England's MPC and periods of high (or very high) macroeconomic uncertainty. There is no clear relationship between the two. The average dissent rates in periods of high (12.2%) and very high macroeconomic uncertainty (12.8%) are not statistically significantly different from those periods with low uncertainty (13.1%). Observing specific episodes as examples, there were few dissents in the 2008-9 very high uncertainty period, while earlier very high uncertainty periods (such as 2003-4) had roughly average dissent rates.

Figure 5: Share of dissenting votes and periods of high macroeconomic uncertainty, 1997-2018



Notes: 'Dissent rate' indicates the share of MPC votes that were different from the majority in the meeting, with the normal maximum being 4 votes out of 9, or 0.44. In 1998, the MPC had 8 members temporarily which made a 50:50 split possible, with the Governor having the casting vote. High (very high) uncertainty is defined as a month in which *either* the consensus measure of uncertainty *or* FTSE volatility was more than 0.75 (1.5) standard deviations above its mean.

We use our sentiment measures from text as indicators of individual and committee views. $\Delta S_{i,t}$ and ΔS_t^{MPC} are, respectively, the change in an individual MPC member's economic assessment from a speech, and the change in the MPC's latest economic assessment from the previous minutes. We construct a measure of the difference between an individual's economic views and those of the remainder of the committee as $\Delta S_{i,t} - \Delta S_t^{MPC}$.

The first row of Table 4 reports the mean absolute difference between an individual member's sentiment relative to that of the rest of the committee, in normal and high uncertainty times. This measure indicates no significant change in the dispersion of MPC member sentiment in periods of higher uncertainty. The second and third rows show the volatility of the individual and committee sentiment measures in different periods. These do have a modest increase in their standard deviation (significant at the 10% level).

Taking these results together, we observe some increase in the volatility of estimates of the state of the economy in periods of uncertainty but no increase in the dispersion of estimates, nor in the dissent rate. These patterns could either be consistent with calibration A or calibration C described in Section 2. Under calibration A the volatility of estimates increases to reflect the increased volatility of the true state, but there is no shift in the informativeness of different information sources and so the dispersion remains unchanged. Under calibration C, there is an increase in the noise around estimates which would alone make estimates more dispersed, but

Table 4: Differences in individual sentiment by uncertainty

	Uncertainty level		Significant difference?
	'Normal'	High	
Mean $abs\left(\Delta S_{i,t} - \Delta S_t^{MPC}\right)$	0.270	0.269	No
Std. dev. $(\Delta S_{i,t})$	0.277	0.311	Yes*
Std. dev. $\left(\Delta S_t^{MPC}\right)$	0.228	0.259	Yes*

Notes: *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels respectively. High uncertainty where ≥0.75 s.d. above mean.

is offset by a fall in the perceived noise around the views of others. Under the information that we can observe, these two explanations are observationally equivalent.

4.2 Regression specification

Next, we conduct ordered probit regressions of changes in policy votes, repeating the exercise on simulated data. We follow the same specification as set out in equation 5, but with additional controls:

$$V_{i,t} = \alpha_i + \rho_1 V_{i,t-1} + \left(\beta_1 + \beta_2 U_t^H\right) * \left(\Delta S_{i,t-1} - \Delta S_{t-1}^{MPC}\right) + \left(\beta_3 + \beta_4 U_t^H\right) * \Delta S_{t-1}^{MPC} + \beta_5 U_t^H + X_t + \eta_{i,t} \quad (7)$$

As described above, $\Delta S_{i,t-1}$ and ΔS_{t-1}^{MPC} are, respectively, the change in an individual MPC member's economic assessment from a speech made in the run-up to the vote, and the change in the MPC's latest economic assessment from the previous minutes. For us to estimate this specification, there needs to have been a speech by MPC member i in the run-up to the policy vote at time t. As some votes have a speech ahead of them while others do not, we have an unbalanced panel.[24] α_i allows for a member-specific fixed effect that could account for the role of preferences in votes as well as other time-invariant unobservable characteristics. To estimate the member-specific fixed effect we need an adequate number of observations per committee member, so we exclude MPC members for whom we have fewer than three monetary policy speeches (an assumption to which we show robustness).[25] To control for economic conditions, we include the change in the MPC's latest *Inflation Report* forecasts for inflation (at the two year horizon) and GDP growth (at the one year horizon), consistent with the timing assumptions of Besley et al. (2008) given lags in the monetary transmission mechanism. We also include controls for member i's lagged policy vote, and a number of factors that have been found to be important in explaining policy votes summarised in X_t: lagged policy decision, a dummy variable for whether the MPC member is an internal MPC member, and a dummy variable indicating the Governor at the time.

We again include interactions between a dummy variable for periods of high uncertainty and MPC and individual sentiment. The role of individual sentiment $\left(\beta_1 + \beta_2 U_t^H\right)$ thus depends on whether uncertainty is high or low, and likewise for MPC sentiment $\left(\beta_3 + \beta_4 U_t^H\right)$.

Finally, in computing standard errors we cluster by MPC member and meeting.

4.3 Regression results

Table 5 reports the results of our regression analysis. The first column shows the baseline specification, where the policy vote depends on the MPC's assessment of the change in economic conditions (S_{t-1}^{MPC}) and the individual MPC member's personal assessment relative to

[24] A potential concern for our exercise would be a correlation between the timing of MPC member speeches and the errors. In the Appendix we report the proportion of speeches that were preceded or followed by a dissent or vote for tightening/loosening and show that these are close to sample averages (Table 9).

[25] We do not include meeting-specific fixed effects due to the relatively small number of observations per meeting.

Table 5: Ordered probit regression of policy vote on economic sentiment and uncertainty

	Baseline	High uncertainty	Very high uncertainty	Excluding crisis	Tighter restrictions	Low GDP growth	Not at ELB
Member i sentiment	0.73***	1.33***	1.33***	0.80***	0.99**	1.29***	0.48
rel. to MPC	(0.24)	(0.30)	(0.30)	(0.27)	(0.34)	(0.31)	(0.44)
MPC sentiment	2.71***	3.15***	3.16***	2.55***	2.63***	3.18***	2.74***
	(0.39)	(0.53)	(0.53)	(0.52)	(0.59)	(0.54)	(0.81)
High uncertainty		-0.65***	-0.58**	-0.60**	-0.95***	-0.67***	-0.74**
(HU)		(0.24)	(0.26)	(0.29)	(0.25)	(0.25)	(0.32)
Very high uncert.			-0.25				
(VHU)			(0.39)				
Low GDP growth						-0.07	
(LGDP)						(0.22)	
HU * member i		-2.25***	-2.89***	-2.20***	-2.00***	-2.56***	-1.19
sent. rel. MPC		(0.53)	(0.69)	(0.79)	(0.67)	(0.65)	(0.87)
VHU * member i			1.32				
sent. rel. MPC			(0.82)				
LGDP * member i						0.71	
sent. rel. MPC						(0.66)	
HU * MPC sent.		-1.47*	-1.72*	-0.70	-0.65	-1.51	-0.87
		(0.79)	(0.96)	(0.97)	(1.00)	(0.92)	(1.14)
VHU * MPC			-0.05				
sentiment			(1.53)				
LGDP * MPC						-0.20	
sentiment						(0.91)	
Lagged vote,	-0.92***	-1.00***	-1.03***	-1.03***	-1.07***	-1.01***	-0.87***
member i	(0.24)	(0.25)	(0.26)	(0.30)	(0.29)	(0.25)	(0.31)
Lagged policy	1.47***	1.37***	1.41***	1.48***	1.41***	1.39***	1.12***
change	(0.26)	(0.28)	(0.29)	(0.33)	(0.31)	(0.28)	(0.36)
Lower cut-off	-1.05	-1.43*	-1.36*	-1.15	-0.68	-1.39*	-1.05*
	(0.80)	(0.79)	(0.82)	(0.73)	(0.87)	(0.81)	(0.59)
Upper cut-off	2.14***	2.04***	2.15***	2.48***	2.91***	2.10***	1.84***
	(0.80)	(0.79)	(0.82)	(0.75)	(0.91)	(0.80)	(0.59)
Member FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes
IR forecasts	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Governor dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Internal dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample	473	473	473	412	324	473	176

Notes: *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels respectively. Regression excluding crisis removes speeches made between 2008 and 2010 from the sample. Tighter restrictions limit the sample to speeches including more than 20 bigrams and MPC members who have given five or more eligible speeches. Low GDP growth indicates a one-year ahead GDP growth forecast that is more than one percentage point below trend. The trend is assumed to be the average annual GDP growth rate since the 1950s until the quarter of the forecast being made, i.e. it varies over the sample, but is typically around 2.6%. Standard errors clustered by MPC member and meeting.

that $(S_{i,t-1} - S_{t-1}^{MPC})$. Both are found to be positively associated with the policy vote, and significant at the 1% level.[26] This is reassuring since it suggests that our sentiment measure is capturing the relative hawkishness or dovishness of the MPC, both collectively and individually. It also confirms an important role for both in determining individual monetary policy votes, as predicted in our simulation exercise. The regression specification includes member-specific fixed

[26]Note that the ordered probit specification means that the coefficient cannot be interpreted as the marginal effect of the variable on the vote. Rather it is the effect on the latent variable which, when a threshold is exceeded, leads to a vote for a loosening, no change or tightening. However, the sign of the coefficient is the same – a higher value of the latent variable indicates a higher chance of a tightening.

effects, but these are not important in explaining the *change* in an individual's policy vote (only two of the thirty one fixed effects are significant at the 10% level). Again, this is consistent with our simulation exercise, and the fact that the dependent variable is policy rate *changes*: rather, we might expect individual factors to have a greater influence on the *level*.[27]

The second column introduces the high uncertainty dummy, as well as interactions with the two sentiment measures. Higher uncertainty is associated with a lower policy vote. This is beyond the effects that high uncertainty might have via the economic assessment of the committee and individual members, and the *Inflation Report* forecasts.[28] The interaction of high uncertainty and the two sentiment measures are also both negative, again consistent with the results from our simulated data regressions. A policymaker places less weight on information about which they are less confident in voting. The reduction in the coefficient is greater for individual views. This lower weight on the individual's own relative assessment indicates that individuals set less store by their private assessments at times of higher uncertainty.

Combined with Table 4, which showed that there is no significant difference in the average spread of views in periods of high uncertainty, the downward weighting of private assessments would tend to point to smaller differences in votes when uncertainty is high. The fact that we did not observe a fall in the dissent rate in periods of uncertainty perhaps reflects the scale of the differences – with discontinuities in voting, the effects may not always be large enough to influence dissents.

4.3.1 Robustness

The remaining columns set out five robustness exercises.

- Column three considers whether very high uncertainty (a level greater than 1.5 standard deviations above the mean) has a different effect from high uncertainty (0.75 standard deviations above the mean). The very high uncertainty dummy variable is insignificant. The interaction terms between very high uncertainty and individual and MPC sentiment are both insignificant.

- Column four ('Excluding crisis') shows that the results remain when we exclude the financial crisis period (taken to be 2008-2010), although the interaction between high uncertainty and MPC sentiment becomes insignificant.

- In the fifth column we tighten the restrictions on the speeches that we include. In particular, we exclude speeches with fewer than 20 bigrams (rather than 10 bigrams in our baseline specifications), and only consider MPC members for whom we have five or more observations. The downweighting of MPC views in periods of high uncertainty becomes insignificant.

- Next, given that high uncertainty tends to be correlated with weak GDP growth, column six considers a variant where we also include a low GDP growth dummy (when it is forecast to be more than 1 percentage point below trend) and interactions with the sentiment measures. The sign on the interaction between high uncertainty and individual sentiment measures continue to be negative and significant, but the MPC sentiment interaction loses its significance. Overall, the result that individual views are downweighted in periods of high uncertainty appears to be more robust than that of MPC views.

- Finally, we may be concerned that the use of unconventional policy tools in the post-crisis period is important. For example, there could be additional frictions to expanding QE

[27] For example, as in the specification of Equation 5.

[28] This is unlike our simulated regressions where there is no feedback from uncertainty to the state of the economy.

compared to a cut in Bank Rate (given that purchases take time to complete), or greater uncertainty about its effects, that affects the probability that a decline in sentiment feeds into a looser policy vote. The final column restricts the sample to those periods in which Bank Rate was not at its effective lower bound (before the cut to 0.5% in 2008, and after the rise back to 0.5% from 0.25% in 2017). This restricts the sample considerably. While the signs are retained, many of our results lose their significance.

Table 11 of the Appendix demonstrates that our choice of uncertainty indicator is not key to our results: the weight on individual sentiment is consistently negative across a range of alternative uncertainty measures. We also consider the results when uncertainty enters as a level rather than a dummy variable. The coefficients are consistently negative, and significant for all but the consensus measure of uncertainty.

Brooks et al. (2012) argue that the prevalence of 'no change' votes warrants a specification that takes account of the greater probability of that choice: a 'zero inflated' probit. In Appendix C we follow their exercise. Table 10 reports the results when we allow for a splitting equation that includes factors that might be associated with a higher probability of 'no change': *Inflation Report*, dummy variables for internals and the Governor, and in some specifications the high uncertainty dummy variable. Some of these variables prove significant in explaining the propensity to vote for a no change, and there is some change in estimated coefficients. However, our results regarding the importance of individual and committee sentiment in periods of normal and high uncertainty are unchanged.

5 Conclusions

Monetary policy committees face the challenging task of interpreting a wide range of incoming data under considerable uncertainty. A prominent hypothesis to explain differences in views of individuals on a committee is differences in the assessments of economic conditions that they form when analysing those data. But a key obstacle in understanding the role of these assessments in voting is that they are typically unobservable.

Our novel text-based measure of committee member views extracted from speeches allows us to test the predictions of a small model of voting, in which policymakers need to estimate the state of the economy from noisy signals. We corroborate a number of key predictions from the model. First, we show that both individual and committee economic sentiment measures positively and significantly affect votes. Second, the weights on both of these estimates are lower in periods of higher uncertainty, consistent with committee members placing less weight on views that they are less sure about. The weights on individual views fall by relatively more, and shows greater robustness to alternative specificaions. Finally, we show that there has been no change in dissent rates in periods of high uncertainty, and that the dispersion in committee assessments is also unchanged.

Public commentators often look to votes as an indicator of the degree of uncertainty around policy decisions, and as an indicator of whether the committee is indeed working as intended, with individuals debating diverse views to arrive at robust policy decisions. Our paper demonstrates that the links between uncertainty, estimates of the state of the economy and votes are nuanced, meaning that such inferences from the vote alone are unlikely to be robust. And while a 'cacophony of voices' critique may be levelled at individualistic communications on a committee (Blinder (2004)), we show the value of individual members' communications for conveying differences in views which votes may hide. This suggests that individualistic communications are important for explaining policy decisions and convey accountability and transparency benefits.

Bibliography

Apel, M. and Blix Grimaldi, M. (2014). How Informative Are Central Bank Minutes? *Review of Economics*, 65(1):53–76.

Apel, M., Claussen, C. A., and Lennartsdotter, P. (2010). Picking the Brains of MPC Members. Working Paper Series 237, Sveriges Riksbank (Central Bank of Sweden).

Baddeley, M. (2010). Herding, social influence and economic decision-making: socio-psychological and neuroscientific analyses. Philiosophical Transactions of the Royal Society 365, Royal Society.

Baker, S. R., Bloom, N., and Davis, S. J. (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics*, 131(4):1593–1636.

Bank of England (2014). Transparency and Accountability at the Bank of England.

Berk, J. M., Bierut, B., and Meade, E. (2010). The Dynamic Voting Patterns of the Bank of England's MPC. DNB Working Papers 261, Netherlands Central Bank, Research Department.

Besley, T., Meads, N., and Surico, P. (2008). Insiders versus Outsiders in Monetary Policymaking. *American Economic Review*, 98(2):218–223.

Bhattacharjee, A. and Holly, S. (2005). Inflation Targeting, Committee Decision Making and Uncertainty: The case of the Bank of England's MPC. Cambridge Working Papers in Economics 0530, Faculty of Economics, University of Cambridge.

Bholat, D. M., Hansen, S., Santos, P. M., and Schonhardt-Bailey, C. (2015). Text mining for central banks.

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

Bloom, N. (2009). The Impact of Uncertainty Shocks. *Econometrica*, 77(3):623–685.

Boero, G., Smith, J., and Wallis, K. (2008). Uncertainty and Disagreement in Economic Prediction: The Bank of England Survey of External Forecasters. *Economic Journal*, 118(530):1107–1127.

Boero, G., Smith, J., and Wallis, K. F. (2015). The Measurement and Characteristics of Professional Forecasters' Uncertainty. *Journal of Applied Econometrics*, 30(7):1029–1046.

Brooks, R., Harris, M. N., and Spencer, C. (2012). Inflated ordered outcomes. *Economics Letters*, 117(3):683–686.

Consensus Economics (2018). Consensus Forecasts – G-7 Countries & Western Europe — Consensus Economics. `http://www.consensuseconomics.com/download/G7_Economic_Forecasts.htm`. [Online; accessed 30-April-2018].

Ehrmann, M. and Fratzscher, M. (2007). Communication by central bank committee members: Different strategies, same effectiveness? *Journal of Money, Credit and Banking*, 39(2-3):509–541.

Gerlach-Kristen, P. (2006). Monetary policy committees and interest rate setting. *European Economic Review*, 50(2):487–507.

Gerlach-Kristen, P. (2008). The Role of the Chairman in Setting Monetary Policy: Individualistic vs. Autocratically Collegial MPCs. *International Journal of Central Banking*, 4(3):119–143.

Haldane, A. G. (2015). Labour's Share. Speech given at the Trades Union Congress, 12 November 2015.

Hansen, S., McMahon, M., and Velasco Rivera, C. (2014). Preferences or private assessments on a monetary policy committee? *Journal of Monetary Economics*, 67(C):16–32.

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

Jurado, K., Ludvigson, S. C., and Ng, S. (2015). Measuring Uncertainty. *American Economic Review*, 105(3):1177–1216.

King, M. (2007). The MPC ten years on. Lecture to the Society of Business Economists, London.

Malmendier, U., Nagel, S., and Yan, Z. (2017). The Making of Hawks and Doves: Inflation Experiences on the FOMC. NBER Working Papers 23228, National Bureau of Economic Research, Inc.

Nyman, R., Kapadia, S., Tuckett, D., Gregory, D., Ormerod, P., and Smith, R. (2015). News and narratives in financial systems: Exploiting big data for systemic risk assessment. *Bank of England Working Paper*, (704).

Picault, M. and Renault, T. (2017). Words are not all created equal: A new measure of ECB communication. *Journal of International Money and Finance*, 79(C):136–156.

Posen, A. S. (2010). The realities and relevance of Japan's Great Recession: Neither "Ran" nor "Rashomon". STICERD Public Lecture, London School of Economics.

Redl, C. (2019). Uncertainty Matters: Evidence from Close Elections. In *NBER International Seminar on Macroeconomics 2019*, NBER Chapters. National Bureau of Economic Research, Inc.

Riboni, A. and Ruge-Murcia, F. (2014). Dissent in monetary policy decisions. *Journal of Monetary Economics*, 66(C):137–154.

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

Rulke, J.-C. and Tillmann, P. (2011). Do FOMC members herd? *Economics Letters*, 113(2):176–179.

Sibert, A. (2006). Central Banking by Committee. *International Finance*, 9(2):145–168.

Stockton, D. (2012). Review of the Monetary Policy Committee's Forecasting Capability. Technical report.

Surico, P. (2003). US Monetary Policy Rules: the Case for Asymmetric Preferences. Royal Economic Society Annual Conference 2003 199, Royal Economic Society.

Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *Journal of Finance*, 62:1139–1168.

Vissing-Jorgensen, A. (2019). Central Banking with Many Voices: The Communications Arms Race.

Warsh, K. (2014). Transparency and the Bank of England's Monetary Policy Committee.

Weale, M. (2015). Independence and Responsibility: Observations from and External Member of the Bank of England's Monetary Policy Committee. Speech at the Thomas Guggenheim Conference, 28th June 2015.

A Speech counts

Figure 6: Annual count of MPC member speeches, June 1997 – December 2018

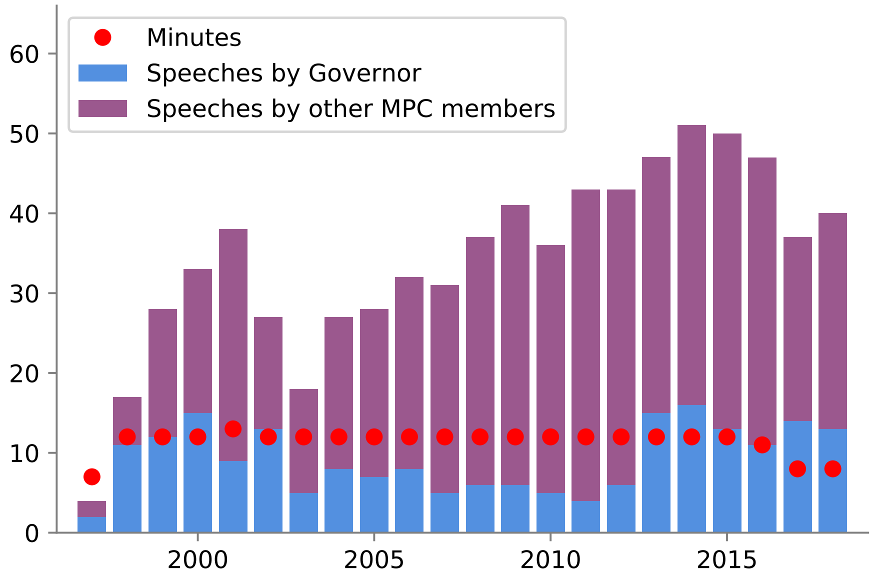


Table 6: Count of speeches by MPC member in the period June 1997 to Dec 2018

Member	**Speeches**	**Member**	**Speeches**	**Member**	**Speeches**
Adam Posen	15	DeAnne Julius	5	Mervyn King	78
Alan Budd	1	Eddie George	66	Michael Saunders	7
Andrew Haldane	34	Gertjan Vlieghe	7	Nemat Shafik	13
Andrew Large	9	*Howard Davies*	*0*	Paul Fisher	17
Andrew Sentance	27	Ian McCafferty	15	Paul Tucker	53
Ben Broadbent	24	Ian Plenderleith	7	Rachel Lomax	12
Charles Bean	43	John Gieve	19	Richard Lambert	5
Charles Goodhart	1	John Vickers	4	Silvana Tenreyro	2
Christopher Allsopp	2	Jon Cunliffe	26	Spencer Dale	17
Dave Ramsden	6	Kate Barker	16	Stephen Nickell	20
David Blanchflower	12	Kristin Forbes	14	Sushil Wadhwani	14
David Clementi	19	Marian Bell	4	Tim Besley	8
David Miles	24	Mark Carney	78	Willem Buiter	2
David Walton	3	Martin Weale	26		

B Text Preparation

Dictionary bigrams and the text of the minutes and speeches are reduced to their base stem using a Porter stemmer to allow matches across different phrasings and conjugations ('increase', 'increased', 'increasing' all stem to 'increas'). Numbers and punctuation are removed, and documents split into sentences for bigram search. To more cleanly identify bigrams of economic term and direction, we remove stopwords which may interpose between the pair of words being sought, where they can be safely removed without a change of meaning:

- Common pronouns, conjunctions, articles, etc.: 'the', 'it', 'of', 'with', 'is', ... but not contradictory connectors ('but', 'except', etc.)
- Dates and real names: 'q1', 'November', 'Carney'...
- Technical terms which do not change the tone: 'level', 'aggregate', 'actual', ...
- Modifiers which do not change the direction of the tone: 'generally', 'sharply', ...

We remove all sentences containing a negating word: 'not', 'cannot', 'no', 'nobody', 'none', 'nothing', 'nowhere', 'never', 'neither', 'nor', '-n't' suffix.

After stemming and removing the initial set of stopwords, we identify trigrams – sets of three words – for which our bigrams are the outer pair of words ("higher *something* inflation"), and assess the most common words in between. Where appropriate, these middle words – the filling in the bigram sandwich – are added to the stopword list. Manual review is required as some filling words will change the meaning or correctly separate the bigrams, while other can happily be dropped. For example, we do not want to identify bigrams across parts and clauses of a sentence: 'This inflation resulted in increased tension in financial markets' should not yield the bigram 'inflation increased': 'resulted' here correctly separates the parts of the sentence. Likewise, '...had no effect on output. Rising tensions...' should not yield [output rising] as a bigram: we should respect the sentence boundary. Each sentence is considered individually to avoid erroneous identification of bigrams across sentence breaks.

B.1 Wordlists for bigram construction

The nouns (Table 7) are combined with the adjectives / directional modifiers (Table 8) to create bigrams:

Table 7: Nouns used to categorise text

Concept	**Noun (stemmed)**
Inflation & Prices	inflation (-ary pressure), (commodity/consumer/energy/house/import/input/oil/retail) price, cost (pressure), CPI, RPI, RPIX, wage (pressure/settlement)
Exchange rates	ERI*, exchange rate*, pound*, sterling*
Labour	labour market/demand, job, income, hours, earnings, employment, unemployment* (rate), claimant count*
Activity	activity, (household) consumption, construction, (aggregate/consumer/domestic/excess/external) demand, expenditure, export, external environment, GDP (business) investment, manufacturing, output (gap), pmi, (industrial) production, productivity, (retail) sales, services, slack*, spare capacity*, (consumer/household) spending, (world) trade, world economy
Confidence	(business/consumer) confidence, business expectations

Notes: * indicates that the accompanying adjectives are inverted, i.e. positive adjectives for unemployment are associated with a lower inflation outlook. Where the noun is itself a two-word phrase ("claimant count"), these are identified in the text as a single unit.

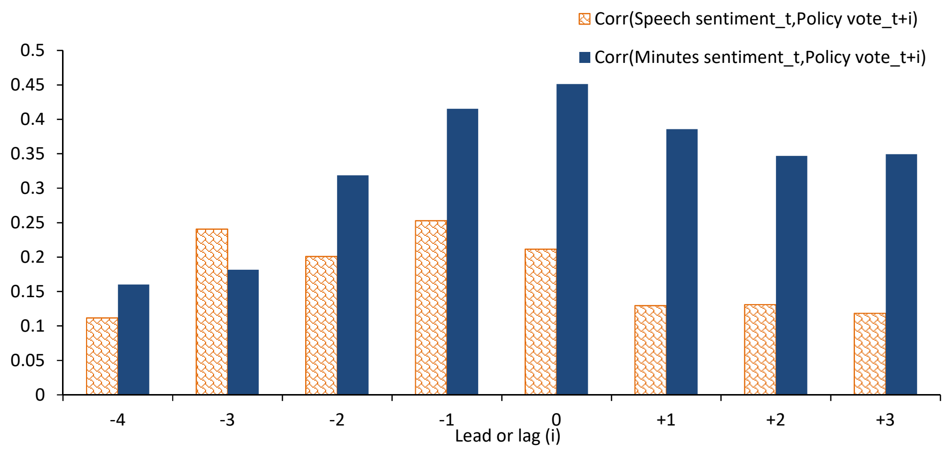
Table 8: Adjectives / Direction modifiers used to categorise text

Positive	**Negative**	**Positive**	**Negative**	**Positive**	**Negative**
bigger	collapse	overshoot	edge down	surge	recede
boost	contract	overshot	edged down	tick up	recession
edge up	cut	pick up	edging down	ticked up	reduce
edged up	dampen	picked up	fall	ticking up	reduction
edging up	decelerate	picking up	fallen	up	shrink
expand	decline	picks up	fell	upside	slow down
expansion	decrease	pickup	less	upside news	slowed down
faster	depress	push up	lessen	upside risk	slowing down
gain	deteriorate	pushed up	lose	upswing	smaller
greater	diminish	pushing up	loss	upturn	soft
grow	downside	quicken	lost	upward	soften
grown	downside news	quicker	lower	widen	tick down
harden	downside risk	raise	narrow	wider	ticked down
higher	downswing	rise	narrower		ticking down
hike	downturn	risen	normalise		undershoot
increase	downward	rose	push down		undershot
larger	drop	strengthen	pushed down		weaken
more	ease	stronger	pushing down		weaker

Notes: Where the adjective is itself a two-word phrase ("pick up"), these are identified in the text as a single unit. In each case these words are stemmed to match with different phrasing (e.g. speed up/sped up; reduce/reduction)

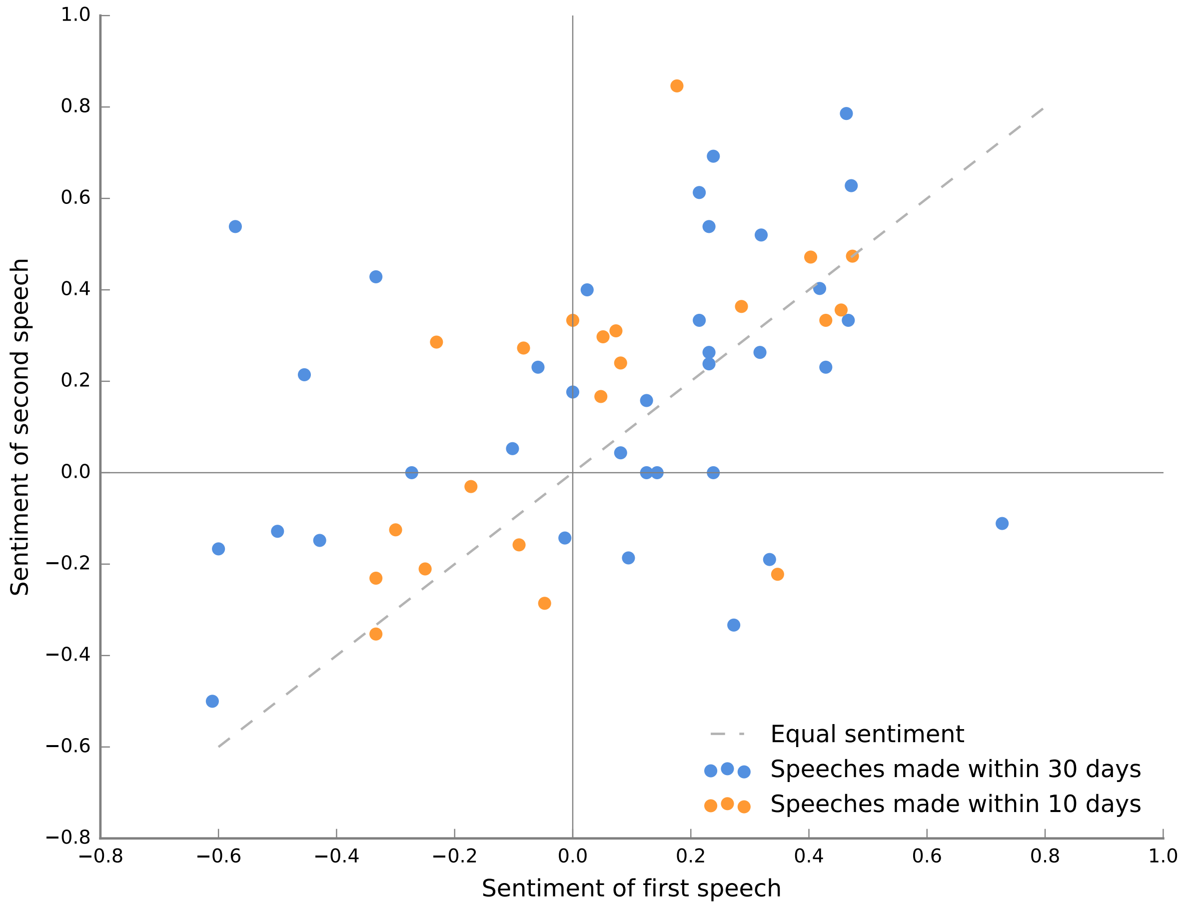
C Robustness

Figure 7: Correlogram for policy votes, speech sentiment and minutes sentiment



Notes: Correlation between lags and leads of policy votes and speech and minutes sentiment. Given that speeches are not made on the day of policy decisions, each speech is associated with the vote immediately after. As such, the information available at the time of a speech is lower than the minutes (which aligns perfectly with the time t vote).

Figure 8: Sentiment of speeches made close together



Notes: The days between speeches are measured as calendar days (as opposed to working days). In our sample, there are 21 instances of two speeches made by the same MPC member within 10 days of each other and a further 35 made within 30 days.

Table 9: Proportion of votes that were preceded or followed by a speech

	Share preceded by a speech	Share followed by a speech
All votes	28.0%	28.5%
Looser votes	25.9%	27.0%
No change votes	29.7%	30.2%
Tighter votes	25.6%	26.8%
Dissenting votes	27.6%	25.8%

Notes: Reports the share of votes of different types that were preceded by a speech (since the previous meeting) or followed by a speech (before the next meeting).

Table 10: Zero-inflated ordered probit regression of policy vote on economic sentiment and uncertainty

	Baseline	High uncertainty
Member i sentiment relative to MPC	0.60**	0.97***
	(0.25)	(0.37)
MPC sentiment	2.79***	2.78***
	(0.42)	(0.81)
High uncertainty (HU)		-0.68
		(0.69)
HU * member i sentiment rel. MPC		-2.01***
		(0.63)
HU * MPC sentiment		-0.15
		(2.60)
Lagged vote, member i	-0.92***	-1.19
	(0.30)	(1.16)
Lagged policy change	1.67***	1.64*
	(0.31)	(0.84)
Lower cut-off	-1.34*	-1.94
	(0.81)	(1.54)
Upper cut-off	2.33***	2.05**
	(0.81)	(1.03)
Member FEs	Yes	Yes
IR forecasts	Yes	Yes
Governor dummies	Yes	Yes
Internal dummy	Yes	Yes
Splitting equation		
IR dummy	4.07***	1.40
	(1.17)	(4.46)
High uncertainty		-0.73
		(0.97)
Internal dummy	Yes	Yes
Governor dummies	Yes	Yes
Constant	Yes	Yes
Sample	473	473

Notes: *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels respectively. Standard errors clustered by MPC member and meeting.

Table 11: Ordered Probit regression of policy vote, V_{it} under alternative uncertainty measures

	Consensus disagreement			FTSE volatility			Redl macroeconomic uncertainty		
	Continuous	High dummy	Very high dummy	Continuous	High dummy	Very high dummy	Continuous	High dummy	Very high dummy
$\left(\Delta S_{i,t} - \Delta S_t^{MPC}\right)$	0.74***	1.06***	1.05***	0.43*	1.16***	1.15***	0.67***	0.86***	0.86***
	(0.25)	(0.29)	(0.29)	(0.25)	(0.27)	(0.28)	(0.24)	(0.29)	(0.29)
ΔS_t^{MPC}	2.78***	3.41***	3.41***	2.06***	3.00***	3.04***	2.97***	3.11***	3.11***
	(0.40)	(0.51)	(0.51)	(0.45)	(0.48)	(0.49)	(0.44)	(0.52)	(0.52)
U_t	-0.42***			-0.49***			-0.32***		
	(0.09)			(0.11)			(0.09)		
U_t^H		-0.68***	-0.67***		-0.59**	-0.36		-0.38**	-0.37*
		(0.21)	(0.25)		(0.28)	(0.29)		(0.18)	(0.21)
U_t^{VH}			0.06			-0.73			-0.05
			(0.38)			(0.48)			(0.37)
$U_t * \left(\Delta S_{i,t} - \Delta S_t^{MPC}\right)$	-0.03			-0.50**			-0.17		
	(0.18)			(0.24)			(0.19)		
$U_t^H * \left(\Delta S_{i,t} - \Delta S_t^{MPC}\right)$		-1.27**	-1.71**		-2.32***	-3.37***		-0.67	-0.65
		(0.65)	(0.98)		(0.62)	(0.80)		(0.51)	(0.54)
$U_t^{VH} * \left(\Delta S_{i,t} - \Delta S_t^{MPC}\right)$			1.23			2.26**			-0.03
			(0.91)			(1.03)			(0.79)
$U_t * \Delta S_t^{MPC}$	-0.90***			-0.42			-040*		
	(0.33)			(0.32)			(0.30)		
$U_t^H * \Delta S_t^{MPC}$		-3.13***	-3.22***		-1.94**	-2.81**		-1.07*	-1.06***
		(0.87)	(1.04)		(0.96)	(1.23)		(0.65)	(0.75)
$U_t^{VH} * \Delta S_t^{MPC}$			0.36			0.88			-0.06
			(1.57)			(1.89)			(0.97)
$i_{i,t-1}$	-0.97***	-0.95***	-0.96***	-1.08***	-0.93***	-0.98***	-0.92***	-0.92***	-0.92***
	(0.26)	(0.26)	(0.27)	(0.26)	(0.25)	(0.25)	(0.24)	(0.24)	(0.24)
i_{t-1}	1.48***	1.47***	1.47***	1.39***	1.41***	1.42***	1.27***	1.38***	1.37***
	(0.28)	(0.28)	(0.28)	(0.28)	(0.28)	(0.28)	(0.27)	(0.26)	(0.27)
Lower cut-off	-1.58*	-1.34*	-1.22	-1.48**	-1.19	-1.22	-1.48*	-1.21	-1.23
	(0.81)	(0.74)	(0.81)	(0.89)	(0.84)	(0.84)	(0.78)	(0.77)	(0.78)
Upper cut-off	1.87**	2.12***	2.25***	1.87**	2.18***	2.20***	1.86**	2.05***	2.03***
	(0.81)	(0.74)	(0.81)	(0.89)	(0.84)	(0.84)	(0.77)	(0.77)	(0.78)
Member FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
IR forecasts	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Governor dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Internal dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample	473	473	473	473	473	473	473	473	473

Notes: *, **, and *** indicate statistical significance at the 10%, 5% and 1% levels respectively. Continuous uncertainty measures are standardised relative to their mean and standard deviation. A high (very high) uncertainty dummy indicates a value of more than 0.75 (1.5) standard deviations above its mean.