Identifying Structural Models of Committee Decisions with Heterogeneous Tastes and Ideological Bias*

Yonghong AN
Texas A&M University

Xun TANG
Rice University

August 6, 2015

Abstract

In practice, members of a committee often make different recommendations despite a common goal and shared sources of information. We study the nonparametric identification and estimation of a structural model where such discrepancies are rationalized by the members' unobserved types, which consist of ideological bias while weighing different sources of information, and tastes for multiple objectives announced in the policy target. We consider models with and without strategic incentives for members to make recommendations that conform to the final committee decision. We show that pure-strategy Bayesian Nash equilibria exist in both cases, and that the variation in common information recorded in the data helps us to recover the distribution of private types from the members' choices. Building on the identification result, we estimate a structural model of interest rate decisions by the Monetary Policy Committee (MPC) at the Bank of England. We find some evidence that the external committee members are less affected by strategic incentives for conformity in their recommendations than the internal members. We also find that the difference in ideological bias between external and internal members is statistically insignificant.

Keywords: Committee decisions, nonparametric identification, heterogeneity, mixture model
JEL: C14, D71

*We are grateful to an associate editor and three anonymous referees for their exceptionally helpful comments. We also thank seminar participants at Rice University, University of Toronto, University of Pennsylvania, Renmin University, the Tsinghua International Conference in Econometrics (2013), and the North American Winter Meetings of the Econometric Society (2015) for useful feedbacks. All errors remain our own.

1 Introduction

Public policy or business decisions are often made by committees organized to serve a common cause. Despite the information shared through group deliberations, members of a committee may disagree over the weights assigned to relevant factors in their decisions. In addition, an individual member may have strategic concerns about whether his recommendation conforms to the final decision made by the committee. As a result, members often end up with distinct recommendations. Prominent examples include corporate boards of directors in the private sector, the U.S. Supreme Court judges, and committees in charge of monetary policies at central banks such as Monetary Policy Committee (MPC) at the Bank of England and the Federal Open Market Committee (FOMC) at the U.S. Federal Reserve.

Understanding the mechanism that generates discrepancies in members' recommendations is important, because committee decisions are prevalent in social-economic contexts. Besides, inference of member preference sheds lights on policy questions such as predicting committee decisions under counterfactual circumstances, e.g., anonymous voting in committees.

Preview of the model and identification. To this end, we set up a structural model that rationalizes dissenting recommendations from committee members by their individual heterogeneity, or private types. First, members do not agree on the relative importance of multiple objectives announced as the goal of the committee. Second, as the committee pool multiple sources of information through group deliberations, each member may decide to weigh these sources differently while forming their own perceptions about the consequences of each choice. We refer to these two sources of heterogeneity as "tastes for multiple objectives" and "ideological bias towards various information sources", respectively.

We focus on committees that make a binary decision by aggregating individual recommendations through a majority rule. All committee members share common information, which includes some states of the world, an announced policy target that consists of multiple objectives, and several sources of information that describe the uncertainty in outcome from both alternatives. The goal for the committee is to choose an action that minimizes the ex ante deviation from the target.

We study the identification and estimation of the model under two scenarios: one in which members have strategic concerns about conformity to the committee decision (a.k.a. "strategic recommendations"); and one in which they do not (a.k.a. "expressive recommendations"). In both cases, pure-strategy Bayesian-Nash equilibria (p.s.BNE) exist, and the identification of ideological bias and individual tastes are obtained by exploiting how the patterns of individual recommendations, or the conditional choice probabilities, change with the common information in data.

In the model with expressive recommendations, members follow a simple dominant strategy in equilibrium. We show how to identify this model under a cross-sectional data environment, where each independent committee is observed to make a single decision. Recommendations from individual members and the common information are both recorded in the data. Without any strategic concern, the individual choice probabilities in equilibrium are mixtures of taste distributions, with the mixing weights determined by the probability masses of ideological bias. Committee members with heterogeneous tastes and ideological bias react differently to the same variation in the common information. This allows us to recover the distribution of two-dimensional private types from their recommendation patterns.

Identifying the model with strategic recommendations requires a qualitatively different argument, and a panel structure where each cross-sectional unit (a committee) makes several group decisions. As in the case with expressive recommendations, members' conditional choice probabilities are finite mixtures. However, with strategic concerns, the component probabilities in the mixture (which are conditional on the ideological bias) depend on endogenous patterns of recommendation by other members. We identify this model through sequential steps: First, we apply results from Hu (2008) to recover individual choice probabilities conditional on the ideological bias. The identifying power comes from observed variation in the common information for the committee. Second, we show that, under mild conditions with economic interpretation, the component probabilities are monotonic in ideological bias. This allows us to match the identified component probabilities with specific values of bias. In contrast to most existing literature, our monotonicity result is derived endogenously from the structural model, as opposed to being imposed exogenously. Lastly, we recover the distribution of individual tastes using continuous variation in the component choice probabilities in response to changes in common information. This is done with an additional minimum set of semi-parametric restrictions on the taste distribution.

Relation to existing literature and contribution. Our paper is closely related to an emerging literature on the econometric and empirical analysis of collective decision-making (e.g., Iaryczower, Lewis, and Shum (2013) and Iaryczower and Shum (2012b,a)). A classical approach for modeling committee decisions in this literature is via a common value model with incomplete information. See, for example, the recent work in Iaryczower and Shum (2012b) and Iaryczower, Shi, and Shum (2012). Hansen, McMahon, and Rivera (2014) also analyze the decisions of MPC in a common value framework. In a typical common value model, there exists a correct decision determined by the state of the nature, which is unknown to all members. The members receive correlated signals about the state, and make decisions based on their update beliefs with or without group deliberations. The discrepancy in members' recommendations is thus attributed to the difference in realized private information, as well as their ideology-related payoffs.

We consider our model as a useful alternative to the common value approach in several ways. First, in our model the members share the same sources of information, which are at least partly reported in the data. The differences in members' recommendations are thus rationalized by idiosyncracies in how they process information from different sources, or weigh multiple policy objectives. Second, our model introduces more structural restrictions on how the factors reported in the data (e.g., committees' announced goals and sources of information) affect individual and committee decisions. This added structural link allows us to exploit exogenous variation in these observed factors to recover the distribution of two-dimensional private types, which includes a new source of heterogeneity in how members weigh multiple objectives in the policy target. Third, our model also lends itself to an array of situations where there may not exist any unambiguous interpretation of a correct decision under each state of the nature.

The identification of the model with expressive recommendations is related to the recent literature on the identification of voter preferences in spatial voting models (e.g., Degan and Merlo (2009) and Merlo and de Paula (2013)) and the inference of strategic voting in structural election models (e.g., Kawai and Watanabe (2013)). The setting of our model differs fundamentally from these other voting/election models, which involve numerous decision-makers with heterogeneous information but no group deliberations. In contrast, we allow members to have a common information set that consists of states, shared targets and sources of information that affect individual perception. We rationalize differences in individual choices by their idiosyncratic bias and tastes despite the common information. Our identification argument is also innovative.

Furthermore, the identification of our method uses a new source of exogenous variation that differs qualitatively from those exploited in these other voting/election papers. In a spatial voting model, candidates are represented as points in the domain of voter preferences and a voter supports the candidate closest to his ideal point by certain metric. Thus, variation in the choice sets (candidates) is instrumental to recover voter preferences. For inference of strategic voting in election models, Kawai and Watanabe (2013) uses the variation in the voting outcome among municipalities (in different districts) with similar characteristics versus the variation in the vote shares and characteristics of other municipalities in the same district. In contrast, we use the variation in common information utilized by committee members to trace out the distribution of their private types when the choice set is fixed (binary). That is, the identifying power in our model is derived from members' heterogeneous adjustment of their choice patterns in response to the same changes in common information.

The identification of the model with strategic recommendations draws on the literature of identifying non-classical measurement error models and finite mixture models. Hu (2008) and Hu and Schennach (2008) provides a general identification result for models with non-classical measurement errors. Hall and Zhou (2003), Kasahara and Shimotsu (2009), and Henry, Kitamura, and Salanié (2013) study nonparametric identification of finite mixture models using exogenous variation in covariates. The results of finite mixture models have been recently applied to different contexts. For example, Aguirregabiria and Mira (2013) identify games of incomplete information with multiple equilibria and common unobserved heterogeneity using such results.

Summary of empirical results. We estimate a structural model of interest rate policy decisions by the Monetary Policy Committee (MPC) at the Bank of England, taking into account the strategic interaction among its members due to concerns for conformity to group decisions. We find that MPC members tend to put more weights on the internal forecasts by the Bank of England than on external forecasts by professionals surveyed in the private sector. We also find that the recommendations from external committee members (who have no executive responsibilities at the Bank and only offer committee service on a part-time basis) are less distorted by strategic incentives for conformity than internal members (who hold full-time executive positions at the Bank). A third finding is that the difference in ideological bias between external and internal members is statistically insignificant.

Roadmap. The rest of this article is organized as follows. In Sections 2 and 3 we present identification results for the models with expressive and strategic recommendations, respectively. In Section 4 we conduct an empirical analysis of the monetary policy decisions by MPC at the Bank of England. Section 5 concludes.

2 Expressive Recommendation

Some committees are short-term and make a single or very few decisions (e.g., juries in civil or criminal courts). Other committees make multiple decisions over its lifetime, but members are not subject to strategic considerations such as reputation for good judgement, or have no incentive to conform to group choices. (For example, members in a corporate board of directors/partners are at the top of company hierarchy, and thus may well be free from concerns such as promotion, etc.) In such cases, committee members propose actions that they consider to be optimal based on individual perception. We now define the equilibrium in such a model of expressive recommendation, and establishes its identification.

2.1 The model and equilibrium

Consider a cross-sectional data that report the group decisions made by independent committees. Each of these committees may consist of a different set of members. In each episode, all members observe a state S drawn from a distribution with a finite support $\mathcal{S}$. Each member formulates an individual perception about how a binary action $D \in \{0,1\}$ affects the distribution of the stochastic outcome $Y \in \mathbb{R}^K$ under the current state. A group decision D^* is made by aggregating individual recommendations via a majority rule. (We use lower cases y, d, s to denote the realized values for these random variables.) The outcome space $\mathcal{Y} \subseteq \mathbb{R}^K$ is finite with cardinality Q, and a generic element in $\mathcal{Y}$ is denoted by y^q.

In an episode under the state s, the committee members have two sources of information about how the actions impact the distribution of Y. Denote these two sources by $G(s) \equiv (G_1(s), G_0(s))$ and $H(s) \equiv (H_1(s), H_0(s))$, with $G_d(s)$, $H_d(s)$ being the probability masses of Y if action d is taken under state s according to the two sources respectively. Specifically, $G_d(s) \equiv (G_{q,d}(s) : q = 1, ..., Q)$, where $G_{q,d}(s)$ is the probability for "$Y = y^q$" conditional on the action d and the state s according to the source $G(s)$. The other source $H(s)$ is defined likewise. Across all decision episodes under the state s, $G_d(s)$ and $H_d(s)$ are independent draws from two distinct distributions with support $\mathcal{H} \equiv \{v \in \mathbb{R}^Q_+ : \sum_q v_q = 1\}$. Both $G(s)$ and $H(s)$ are common knowledge among committee members. To simplify exposition, we refer to G, H as the "initial perception" and the "empirical evidence" respectively. In addition, the committee is informed of a target $\tilde{y} \in \mathcal{Y}$ in each decision episode. The information available to committee members under the state s and target $\tilde{y}$ is summarized by $\mathcal{I} \equiv \{s, \tilde{y}, G(s), H(s)\}$.

Example 1 (*Corporate Board Decisions*) *A corporate board of directors deliberates over a proposal to merge with another company. Each member of the board makes a binary recommendation (to merge or not to merge). All members agree that the merger affects the company's stock price and its rate of return on assets (ROA), i.e., $K = 2$. The board members agree on an ideal target of stock prices and ROA, denoted by $\tilde{y} \equiv (\tilde{y}_1, \tilde{y}_2)$. While making their decisions, the members take account of the industry and market environment as well as the state of both firms in the merger (s). In addition, they have access to two sources of information: an external assessment of the merger outcome by an independent management consulting firm $G(s)$; and the company's own prediction of consequences of the merger based on historical evidence from past mergers under the same states $H(s)$. Despite the common goal $\tilde{y}$ and information $G(s)$ and $H(s)$, the board members hold different views about how the proposal could affect the merger's stock prices and ROA. They also disagree on relative weights that should be assigned to these two dimensions in the decision.*

A member i's individual perception about the probability that "the outcome will be $Y = y^q$ if d is chosen under the state s" is

$$\mathcal{F}_{q,d}(s; \alpha_i) \equiv (1 - \alpha_i) G_{q,d}(s) + \alpha_i H_{q,d}(s) \tag{1}$$

for $d \in \{0, 1\}$, where $\alpha_i \in (0, 1)$ is independently drawn from a multinomial distribution for each i. We refer to α_i as an "ideological bias", as it captures the members' willingness to adjust their initial perception in response to empirical evidence or more generally to balance information from two sources.

Every member i independently draws a vector of weights $W_i = w_i$ from some distribution with support $\mathcal{W} \subseteq \mathbb{R}^K_{++}$. Member i recommends:

$$d_i(\alpha_i, w_i; \mathcal{I}) \equiv \arg\min_{d_i \in \{0,1\}} \mathbb{E}\left[\textstyle\sum_k w_{i,k} (Y_k - \tilde{y}_k)^2 \middle| d_i; \alpha_i, \mathcal{I} \right] \tag{2}$$

where the expectation in (2) is taken with respect to the outcome Y and the final committee decision D^* given i's perception formulated in (1). Such an expectation depends on other members' strategies as well as the distributions of (α_i, W_i). (We provide an explicit form of this conditional expectation in the proof of Lemma 1.) We refer to W_i as individual "tastes" for multiple dimensions in the outcome. These heterogeneous weights capture the discrepancy between members that persists even after the group deliberation.

The individual types (α_i, W_i) are private information of each member, but their distribution is common knowledge among all committee members. For the rest of this section, we let the size of committees be fixed at an odd number I, and maintain the following assumptions throughout the section. Our method also applies to the cases with an even number of members as long as the tie-breaking rule is specified and known to econometricians.

Assumption 1 *(i) For each member i, the ideological bias α_i is independently drawn from a distribution F_α over a known discrete support $\mathcal{A} \equiv \{\alpha^j : j = 1,.,J\} \in (0,1)^J$; individual tastes W_i are independent draws from a continuous distribution F_W with positive density over a known support $\mathcal{W} \subseteq \mathbb{R}_+^K$. (ii) α_i and W_i are independent from each other, and jointly independent from $\mathcal{I}$.*

The exogeneity of the information set $\mathcal{I}$, and in particular the empirical evidence H, is instrumental for our identification methods. In committee decisions, the common information $\mathcal{I}$ is usually pooled after group deliberations, and thus reflects members' consensus about the uncertainty under different choices. On the other hand, the private types are meant to capture the idiosyncracies in members' perspectives that persist after group deliberations. Hence the orthogonality conditions in Assumption 1 are plausible in situations where the individual idiosyncracies are due to random noises independent across members and orthogonal to common consensus.

This condition is plausible even when the empirical evidence potentially depends on the history of states and committee decisions in the past. As in the example of corporate board decisions, $H(S)$ is based on the accumulated evidence up to the date of the decision. Thus it is consists of random shocks that vary across episodes, and may well be orthogonal to individual types (α_i, W_i). By the same argument, the exogeneity of accumulated evidence is also in a panel data, where committees are observed to make multiple decisions throughout its tenure.

That said, we also acknowledge Assumption 1 rules out situations where members' tastes for multiple objectives or ideological bias are related through hidden factors not reported in the data on common information. One example for such an unobserved factor would be private consensus between a subset of the members, that is not recorded in pooled information formed after group deliberations. If this is a concern in some specific empirical context, then one would need to use a more complex model that takes account of these unobserved factors. However, it remains an open question whether such a more sophisticated model could be identified under any interpretable conditions given the kind of data scenario we consider here.

That the support $\mathcal{A}$ is finite is relevant in environments where members are known a priori to belong to a small number of distinct groups with varying emphasis on both sources of information. In this case, assuming the elements of $\mathcal{A}$ are known, say, by stating that $\alpha^j \in \mathcal{A}$ are spaced with equal distance over $[0,1]$, serves as an approximation of the actual data generating process (DGP). In other cases where α_i is in fact continuously distributed over $[0,1]$ in DGP, our method below should be interpreted as showing identification for a coarser, discretized version of the model.

Finally, note ideological bias α_i could be related to observed demographics of a committee member. Examples of such demographic variables reported in data include political affiliation, education background and professional experiences of committee members, etc. Likewise, the distribution of tastes W_i may also depend on individual-level variables in the data. Nevertheless, such observed heterogeneities do not pose any conceptual challenge to identification in that our method below extends once conditional on characteristics of committee members reported in data.

This model admits a unique pure-strategy Bayesian Nash equilibrium (p.s.BNE) in which each player adopts a dominant strategy. Fix an information set $\mathcal{I} \equiv \{s, \tilde{y}, G(s), H(s)\}$ and define:

$$\delta_{G,k} \equiv \textstyle\sum_q (y_k^q - \tilde{y}_k)^2 [G_{q,1}(s) - G_{q,0}(s)]. \tag{3}$$

In words, $\delta_{G,k}$ is the difference in the ex ante distance from the target $\tilde{y}$ under the two alternatives for a hypothetical member who *only* cares about the k-th dimension in the outcome and whose perception is based on $G(s)$ *exclusively*. In a similar fashion, define $\delta_{H,k}$ by replacing G in (3) with H. For example, a member i who only cares about the k-th dimension in the outcome and puts all weight on the empirical evidence $H(s)$ recommends $D_i = 1$ if and only if $\delta_{H,k} < 0$. (By construction, $\delta_{G,k}$ and $\delta_{H,k}$ are functions of the common information $\mathcal{I}$, which is suppressed in the notation for simplicity.)

Lemma 1 *Under Assumption 1, the model of committee decisions with expressive recommendations has a unique p.s.BNE where each member i follows a dominant pure strategy*

$$\sigma_i^*(\alpha_i, w_i; \mathcal{I}) \equiv 1\left\{\alpha_i \textstyle\sum_k w_{i,k}\delta_{H,k} + (1-\alpha_i)\textstyle\sum_k w_{i,k}\delta_{G,k} \le 0\right\}. \tag{4}$$

As the lemma shows, in the unique dominant strategy p.s.BNE each member makes an individual decision that is solely based on his private types and the common information set $\mathcal{I}$. Remarkably, for each member i the strategy in (4) is the dominant strategy even after taking account of the uncertainty in the endogenous group decision in equilibrium. Such a dominant strategy p.s.BNE exists as a result of two distinctive features of the model: First, a member i's objective function depends on his choice d_i only through ex ante deviation from the target under the committee decision D^*. Such an ex ante deviation depends on d_i only through its impact on the distribution of D^*. Second, by construction, the probability for "$D^* = 1$ conditional on $d_i = 1$" is greater than the probability for "$D^* = 1$ given $d_i = 0$".

2.2 Identification

We now discuss how to recover the distribution of members' private types α_i and W_i in a typical empirical environment where the data record each member's choices $\{d_i : i = 1, 2, ..., I\}$ from a large number of independent committees, each of which makes a single decision in the sample. The states s, the announced targets $\tilde{y}$ as well as the empirical evidence $H(s)$ are reported for each committee decision. We also assume the data allows researchers to observe individual and committee choice patterns conditional on the initial perception $G(s)$ at least for some state. This holds, for instance, if there exists some state s where the initial perception $G(s)$ is known to the researcher a priori, either due to common sense or institutional environment. For example, jurors' attitude toward a series of legal phrases are surveyed and analyzed in Kadane (1983); and Costanzo, Shaked-Schroer, and Vinson (2010) provide a survey of jury-eligible citizens for their beliefs about police interrogations, false confessions, and expert testimony.

We maintain that individual choices in the data are generated as members adopt dominant strategies in (4). Lemma 1 has a powerful implication for identification: Each member's decision is an independent realization of the decision rule in (4) with (α_i, W_i) drawn randomly from some distribution. We also maintain the following assumption (where $\mathcal{N}_\varepsilon(.)$ denotes neighborhood with a radius $\varepsilon > 0$):

Assumption 2 *There exists* $\{s, \tilde{y}, G(s)\}$ *such that* (*a*) *the sign of* $\sum_k w_{i,k}\delta_{G,k}$ *is fixed and known for all* $w_i \in \mathcal{W}$*; and* (*b*) *there exists* (H^a, H^b) *and* $\varepsilon > 0$ *such that* $\Pr\{D_i = 1 \mid s, \tilde{y}, G(s), H\}$ *is* 0 *for all* $H \in \mathcal{N}_\varepsilon(H^a)$*, and is* 1 *for all* $H \in \mathcal{N}_\varepsilon(H^b)$*.*

Part (a) states that under some state and target, the recommendation by an individual who *only* takes account of the initial perception G would be degenerate, regardless of his tastes for policy objectives (w_i). For example, consider a corporate board of a manufactory that deliberates over a proposed merger with its suppliers. The announced target is to boost its stock prices while keeping the ROA at the current level. Suppose an initial evaluation by a management consulting firm suggests that under the current industry conditions there will be a high cost synergy from the merger. Based on this source of information alone, the board members would reach a consensus to proceed with the merger, regardless of their heterogeneous tastes.

Part (b) is a joint restriction on $\{s, \tilde{y}, G(s)\}$ and the support of the empirical evidence $H(s)$. It requires there be "*extreme evidence*" $H(s)$ under which individuals recommendations are degenerate under the state and targets, holding the initial perception fixed. In principle, this condition can be verified using data from individual choices and empirical evidence.

- **The structural link between members' choices and private types**

For better exposition, we define type-specific "indifference hyperplanes" which partition the support of empirical evidence into subsets according to the dominant strategies prescribed. Fix a triple $\{x, \tilde{y}, G(s)\}$ under Assumption 2. For a member with type (α, w), define the indifference hyperplane as the set of empirical evidence that equates this member's ex ante distance from target under both alternatives. That is, the indifference hyperplane consists of empirical evidence H such that

$$\left(\sum_k w_k \delta_{H,k}\right) / \left(\sum_k w_k \delta_{G,k}\right) = \frac{\alpha - 1}{\alpha}. \tag{5}$$

This hyperplane partitions the support for the empirical evidence into two parts: one in which the dominant strategy for the member is to choose 1 and the other to choose 0. By construction, the indifference hyperplanes associated with different α are parallel for any w.

To present the main idea effectively, we focus on the case with $K = 2$, $J \equiv |\mathcal{A}| = 2$ and $Q \equiv |\mathcal{Y}| = 3$ (the outcome space is two-dimensional with three possible scenarios, and there are two types of ideological bias). Proofs for general cases with $J, K > 2$ and $Q > 3$ require more tedious algebra but do not pose new challenges. The indifference hyperplanes can be visualized as lines in $\mathbb{R}^2$ in this case. To see how, first set $w_1 = 1$ as a scale normalization and drop the subscript from w_2 to simplify notation. Denote the normalized support for W_2 by $\mathcal{W} \equiv [\underline{w}, \overline{w}]$. Without loss of generality, suppose $\tilde{y} = y^3 \in \mathcal{Y}$. The summation in $\delta_{G,k}, \delta_{H,k}$ is thus reduced to $\sum_{q=1,2}$, and individual choices depend on $H(s)$ only through $h_q(s) \equiv H_{q,1}(s) - H_{q,0}(s)$ for $q = 1, 2$. Likewise for $G(s)$. With a slight abuse of notation, we define the relevant common information set $\mathcal{I} \equiv \{s, \tilde{y}, g(s), h(s)\}$, where $g \equiv (g_1, g_2)$, $h \equiv (h_1, h_2)$. The support of h is $\{h \in [-1, 1]^2 : h_1 + h_2 \in [-1, 1]\}$ by construction. We will also refer to h as empirical evidence.

For example, suppose part (a) of Assumption 2 holds with a positive sign (that is, the state and the target are such that a member who only uses the initial perception G would always recommend $d_i = 0$). Given any empirical evidence h on the lower-left side of a hyperplane associated with (α, w), the dominant choice for a member with (α, w) is 1. While the intercepts of these indifference hyperplanes (lines in $\mathbb{R}^2$) depend on ideological bias, their slopes only depend on tastes according to (5).

As the realized empirical evidence moves, the probability for recommending an action changes by an amount that equals the measure of the types whose indifference hyperplanes are crossed by this movement of empirical evidence. We fully exploit this structural link to recover the distribution of members' private types from these conditional choice probabilities (CCPs). Below we provide more intuitive details. The formal proof of identification as well as detailed technical conditions are collected in Appendix A.

To see how our argument works, it is necessary to first summarize some intuitive properties of indifference hyperplanes that are instrumental for our argument. For any α, the hyperplanes (lines) associated with different weights w must intersect at the *same* point $h(\alpha) \equiv (h_1(\alpha), h_2(\alpha))$ and this can be derived from the equation implied by the intersection. Slopes of these hyperplanes are necessarily negative because it reflects the rate of substitution between two dimensions in empirical evidence $h(\alpha)$ that is required to keep a member indifferent between the two alternatives. Also, the rate of substitution $-(a_{1,1}+wa_{1,2})/(a_{2,1}+wa_{2,2})$, where $a_{q,k} \equiv (y_k^q - \tilde{y}_k)^2$, is monotonic in the taste w. The direction of monotonicity in the slopes as w increases is determined by the target $\tilde{y}$ and the outcome space $\mathcal{Y}$. For the rest of Section 2.2, we present our method for the case where the hyperplane for type-$(\alpha, \underline{w})$ has a greater slope (i.e. is "less steep") than that for type-$(\alpha, \overline{w})$. This is without loss of generality, since the ranking between these two rates can be recovered from the data. Identification under the other case follows from symmetric arguments.

- **Recovering the order of indifference thresholds**

For identification, we exploit the variation in the members' CCPs over a set of empirical evidence that form the shortest path between the extreme evidence. This set, denoted $\mathcal{H}(h^a, h^b)$, consists of convex combinations of h^a and h^b (with $h_q^l \equiv H_{q,1}^l - H_{q,0}^l$ for $l \in \{a, b\}$). The first step in our argument is to recover the order in which the members with different types change actions as the realized evidence move over the path $\mathcal{H}(h^a, h^b)$. This is a prerequisite for answering the identification question, because such an order is associated with the relative position of the indifference hyperplanes and determines the direction as well as the magnitude of changes in individual choice probabilities.

Without loss of generality, we consider the case where part (a) of Assumption 2 holds with a positive sign. The six panels in Figure 1 enumerate the possible positions of generic extreme evidence (h^a, h^b) relative to indifference hyperplanes. The "flatter" and the "steeper" lines are associated with $\underline{w}$ and $\overline{w}$ respectively; and the dashed and solid lines are associated with two different values $\alpha^1 < \alpha^2$ respectively. In some cases (shown in panels (i) and (iv) in Figure 1) the CCP is observed to vary continuously over a connected interval of empirical evidence in the path $\mathcal{H}(h^a, h^b)$. In the other cases (shown in the other four panels in Figure 1) there exists a subinterval on the path where the CCP is observed to remain constant at a non-degenerate level (strictly between 0 and 1).

In these latter four cases, the intersections of the path $\mathcal{H}(h^a, h^b)$ with the four hyperplanes associated with $(\alpha^1, \underline{w}), (\alpha^1, \overline{w}), (\alpha^2, \underline{w}), (\alpha^2, \overline{w})$ (hereinafter referred to as the "*indifference* *threshold*") are identified as the infima or suprema of the set of empirical evidence on the path where the CCPs remain constant (either degenerate at 0 or 1, or non-degenerate at some intermediate value). In the other two cases, only two of the indifference thresholds can be identified.

To fully pin down the ordering of indifference thresholds, we have yet to distinguish the first four cases from each other, and distinguish between the other two cases. To do so, we use variation in the location of extreme evidence h^a and h^b. The key idea is that, as the extreme evidence h^a, h^b (and hence the path linking them) move, the relative position of indifference thresholds shows different patterns of changes because of the distinctive rates of substitution over the type-specific indifference hyperplanes. This idea can be better visualized in Figure 1 as follows: Fix h^b and vary h^a vertically (that is, hold h_1^a constant and change h_2^a alone). The sign of changes in the distance between ordered indifference thresholds are different across the six scenarios. For instance, in case (i), the distance between the first and the last threshold becomes larger when h_2^a increases; in contrast, such a distance diminishes in case (iv) under the same movement of h_2^a. Similar arguments allow us to distinguish between the other four cases. We formalizes this argument in Lemma A1 in the appendix.

- **Identifying the distribution of member types**

The identification of the distribution of member types (α, w) in the four cases (ii), (iii), (v) and (vi) are relatively easier, because the CCPs are not mixtures of non-degenerate distributions when the empirical evidence in the conditioning set varies over the path $\mathcal{H}(h^a, h^b)$.

Assumption 3 (*a*) *The matrix* $[a_{1,1}, a_{1,2}; a_{2,1}, a_{2,2}]$ *is full-rank, where* $a_{q,k} \equiv (y_k^q - \tilde{y}_k)^2$ *for* $q = 1, 2$. (*b*) *The initial perception is such that* $a_{1,1}g_1 + a_{2,1}g_2$ *and* $a_{1,2}g_1 + a_{2,2}g_2$ *are both non-zero.*

This puts additional restrictions on the triple (of state, target and initial perception) under Assumption 2. Part (a) is a condition on the announced goal for the committee. It rules out uninteresting pathological situations where individual tastes do not matter in the ex ante distance from the target outcome. Identification of the distribution of W_i would fail without this condition, because individual CCPs would be independent from members' idiosyncratic weights. This condition is verifiable given knowledge of $\tilde{y}$ and the outcome space $\mathcal{Y}$. Part (b) rules out another pathological case where one out of the two dimensions in outcome does not affect individual decisions through initial perception at all. Given the other assumptions, it is sufficient for implying the monotonicity of c in w for almost all evidence. (See proof of Proposition 1 in Appendix A.) In principle part (b) can also be verified from the data. This is because, as shown below, the initial perceptions are identifiable from CCPs even without part (b) under other maintained assumptions. It is possible to attain identification even under the "knife-edge" case when part (b) fails, provided the monotonicity of c in w given h^a and h^b still hold.

Proposition 1 *Suppose Assumptions 1, 2 and 3 hold. Then the marginal distributions of* α_i *and* W_i *are both identified in the cases illustrated in (ii), (iii), (v) and (vi) in Figure 1.*

To see the intuition behind this proposition, consider case (ii). First, by varying the empirical evidence from h^a to h^b, one can recover the second indifference threshold as the infimum of the subinterval on $\mathcal{H}(h^a, h^b)$ over which the CCP remains constant in the interior of $(0,1)$. The non-degenerate CCP over this interval equals the probability mass functions for α^1. Second, the full-rank condition in part (a) of Assumption 3 implies that the initial perception $g(s)$ can be recovered from the location of indifference thresholds (which is identified from the previous step) and the assumed knowledge of bias support $\mathcal{A}$. Thus the left-hand side of (5), as a function of individual tastes w and empirical evidence h that characterizes the indifference hyperplane, is identified. Besides, the choice probabilities conditional on the realized bias are identified once the probability mass function of α is known. Finally, the mild regularity condition in part (b) of Assumption 3 implies the left-hand side of (5) is monotone in w. This enables us to invert the CCPs to recover the distribution of idiosyncratic tastes W_i.

The identification of type distributions F_α, F_W in the other two cases (i) and (iv) requires more involved arguments, because the CCPs becomes a finite mixture of component choice probabilities over some sections on the path of empirical evidence $\mathcal{H}(h^a, h^b)$. Nonetheless, the main idea for identification is still to exploit the impact of the empirical evidence on CCPs. By varying the empirical evidence over the convex path $\mathcal{H}(h^a, h^b)$, we get a system of equations that relate the type distributions to observed CCPs. We show that this provides the basis for identification, as long as there is sufficient variation in the empirical evidence in the following sense: The range of evidence that lead to non-degenerate CCPs are sufficiently apart for members with different types of bias α_i. In other words, the ranges of evidence $\mathcal{H}(h^a, h^b)$ covered by the two sets of indifference hyperplanes associated with different bias types must be sufficiently different from each other.

2.3 Estimation strategy

With the model identified nonparametrically, one can use sieve maximum likelihood estimator (MLE) to jointly estimate the probability mass function for α_i and the distribution of W_i. That is, let $\mathcal{F}_n$ denote an appropriately chosen sequence of sieve spaces for continuous cumulative distribution functions over $\mathcal{W}$ such that as $n \to \infty$, $\mathcal{F}_n$ becomes dense in the parameter space for F_W. Let $\mathcal{P} \equiv \{p \in [0,1]^J : \sum_{j \le J} p_j = 1\}$ denote the parameter space of the probability mass function for ideological bias. Define:

$$(\widehat{p}, \widehat{F}) = \arg \max_{p \in \mathcal{P}, F \in \mathcal{F}_n} \frac{1}{N} \sum_{n=1}^{N} \hat{\mathcal{L}}_n(p, F)$$

where n indexes the cross-sectional units of independent committees and N is the sample size; and

$$\hat{\mathcal{L}}_n(p,F) \equiv \sum_{i=1}^{I} \log \sum_{j\leq J} p_j \psi_j(\mathcal{I}_n, F)^{d_{n,i}}[1-\psi_j(\mathcal{I}_n, F)]^{1-d_{n,i}}$$

where $\mathcal{I}_n$ is the common information in the n-th committee in data; $d_{n,i}$ is the recommendation by member i in committee n; and $\psi_j(\mathcal{I},F)$ is the probability that a member chooses 1 conditional on the common information $\mathcal{I}$ and an ideological bias $\alpha_i = \alpha^j$ when the distribution of individual tastes is F. That is, $\psi_j(\mathcal{I},F)$ is the probability that "$\alpha^j \sum_k W_{i,k}\delta_{H,k}(\mathcal{I}) + (1-\alpha^j)\sum_k W_{i,k}\delta_{G,k}(\mathcal{I}) \leq 0$" when the distribution of W_i is F. Conditions for consistency of sieve MLE, as well as discussions on appropriate choices of the sieve space $\mathcal{F}_n$, are provided in Shen (1997), Chen and Shen (1998) and Ai and Chen (2003).

3 Strategic Recommendation

In some other cases, committees are organized to make multiple decisions and the members have career concerns such as reputation which affects promotion or re-election. Thus a member cares about how likely his recommendation conforms with the final decision adopted by the committee. Such an incentive may cause a member to make a different recommendation than what he would make in the absence of such a strategic concern.

The nature of the strategic interaction between committee members depends on how individual recommendations are aggregated in a committee decision. In what follows, we consider the case of a simple majority rule. Our method for identification can be applied under alternative committee decision rules, as long as they are known to researchers.

3.1 The model

Consider a panel data with its cross-sectional units being independent committees ($\mathcal{C}$) that make decisions in several episodes $l = 1, ..., \overline{L}$. (We suppress indices for committees to simplify notations.) Individual recommendations in each episode are also observed in data. Each committee aggregates individual choices by a majority rule known to all members: That is,

$$D_l^* = \max_{d\in\{0,1\}} \sum_{i\in\mathcal{C}} 1\{D_{i,l} = d\},$$

where D_l^* and $D_{i,l}$ are committee and member i's decisions in episode l respectively. To simplify exposition, suppose the size of committees I are fixed throughout the data, which is not necessarily greater than J, the cardinality of the support for the ideological bias α_i.

Members' payoffs are similar to that in Section 2.1, except for an additional incentive to conform with the committee decision. Let $\mathcal{I}_l \equiv \{S_l, \tilde{Y}_l, G_l(S_l), H_l(S_l)\}$ denote the random common information in episode l. In a Bayesian Nash Equilibrium, a member i chooses

$d_{i,l} \in \{0,1\}$ in episode l to minimize:

$$\mathbb{E}_{D^*}\left[1(D_l^* \neq d_{i,l})|d_{i,l},\alpha_i,\mathcal{I}_l\right] + \mathbb{E}_{Y,D^*}\left[\sum_k w_{i,l,k}(Y_{l,k} - \tilde{y}_{l,k})^2|d_{i,l},\alpha_i,\mathcal{I}_l\right], \tag{6}$$

where $w_{i,l,k}$ is the weight i puts on the k-th objective in episode l. The expectation in (6) is taken with respect to (Y_l, D_l^*) given i's perception in (1). Such an expectation depends on other members' strategies and the joint distribution of types $(\alpha_i, W_i)_{i\in\mathcal{C}}$. The first term in (6) captures the strategic incentive, or "career concerns". By construction the scale of unobserved weights cannot be identified, just as in the model of expressive recommendations. Thus, we normalize the weight on this term to 1, so $w_{i,l,k}$ is interpreted as the relative weights on the outcome (compared with the strategic incentive). Similar specifications were used in Levy (2007) and Bergemann and Morris (2013).

Assumption 4 (*a*) *Across members i and episodes l, the tastes $W_{i,l} \equiv (W_{i,l,k})_{k=1}^K$ are independent draws from a continuous distribution $F_{W_i|\mathcal{I}_l}$ with positive density over a known support $\mathcal{W} \subseteq \mathbb{R}_+^K$.* (*b*) *The tenure of each committee is partitioned into intervals, each of which consists of L consecutive episodes. For each i, α_i is fixed within an interval. Across the intervals, α_i are i.i.d. draws from a multinomial distribution F_{α_i} (with support $\mathcal{A}$) independent from $\mathcal{I}_l$.* (*c*) *α_i and W_i are independent given $\mathcal{I}_l$, with $F_{\alpha_i}, F_{W_i|\mathcal{I}_l}$ being common knowledge among committee members.*

Part (a) states that the individual tastes $W_{i,l}$ are idiosyncratic and i.i.d. across members and decision episodes conditional on the common information following the group deliberations. This assumption allows an individual's tastes to be correlated across the multiple episodes throughout the tenure of the committee. It also allows individual tastes to be correlated with the common information set $\mathcal{I}_l$. It does rule out the cases where individual tastes are affected by unobserved noises that are serially correlated even after controlling for the transition of common information. An implication of part (b) is that a member's ideological bias is more persistent than tastes for different objectives. This is a intuitive condition because tastes for policy may well depend on state of the world which is realized in each every episode; whereas the ideological bias by definition won't be updated at a higher frequency than the information sources.

Part (b) is also empirically motivated. For instance, in the context of monetary policy committees, Wesche (2003) reports some empirical evidence that members may frequently change their understanding of economy and adjust the weights W accordingly. On the other hand, it takes time for the members to change ideological bias towards various sources of information. This would be the case, for example, if one of the sources of information is only updated every L periods. Another reason is that the length of service for each member is in unit of L episodes, and thus personnel changes always occur at the end of an interval with L episodes. Consequently, members tend to adjust their bias in processing multiple sources of information, due to deliberations with new members, etc. Another (technical) reason for this assumption is a practical concern about model tractability and identifiability. If both α_i and W_i are modeled as random draws across all individual members and episodes, then it may not be possible to separately recover the marginal distributions of α_i and W_i respectively from the choice patterns.

Part (c) states that after controlling for the common information $\mathcal{I}_l$ the members' preferences over sources of information are not affected by their tastes for policy targets. In our empirical application of MPC decisions, this means a member who puts higher weights on inflation than GDP is not necessarily inclined to rely more on one of the two sources of information (the predictions of policy consequences by outsiders or the Bank of England). Part (c) does rule out the dependence between α_i (ideological bias) and W_i (tastes for outcomes) through unobserved individual heterogeneity.

Assumption 4 rules out serial correlation of ideological bias as well as its correlation between committee members. In situations where this is a concern due to correlation via unobserved time-varying factors (e.g., members may learn their colleagues' private information or, a member tries to influence his colleagues' beliefs about his types), one would need to use a more complex model that accommodate such correlation. Nevertheless, we maintain Assumption 4 as a first-order approximation/simplification of the data-generating process in the MPC application in Section 4.

Assumption 5 *(a) Members in the same committee share the same initial perception $G(S_l)$, which is fixed within each interval of L episodes, but across intervals are i.i.d. draws from some distribution over $\mathcal{H}$ conditional on S_l. (b) Across all episodes (within and across intervals), the empirical evidence $H_l(S_l)$ are i.i.d. draws from some distribution over $\mathcal{H}$ conditional on S_l.*

Together with Assumption 4, these conditions in Assumption 5 accommodate the practical situation where the two sources of information are updated at different frequencies. (For instance, in the example of monetary policy committees, the forecasts by outside institutions are provided on a quarterly basis while the Bank of England estimates are available on a monthly basis.) This difference in the frequency between information sources plays a key role in helping us to separately back out the two distribution of private types (α_i, W_i). In other contexts where both sources G, H and types α_i, W_i are drawn independently across all episodes, it is not clear whether it is feasible to recover the distributions of α_i, W_i from individual choice patterns alone. As in Section 2, our method for identification can be extended to accommodate heterogeneity in committees, members as well as initial perceptions reported in the data.

Let $a_{q,k}(\tilde{y}_l) \equiv (y_k^q - \tilde{y}_{l,k})^2$ and $\mathcal{F}_{l,q,d}(s;\alpha_i) \equiv \alpha_i H_{l,q,d}(s) + (1-\alpha_i)G_{q,d}(s)$, where $\tilde{y}_l$ is the target announced for episode l; $H_{l,q,d}(s)$ denotes in episode l, the probability that $Y = y^q$

given s and d according to empirical evidence and this is similar to $H_{q,d}$ defined in Section 2. Applying the law of iterated expectations to the second term in the objective function in (6), we can rewrite the minimization problem in (6) equivalently as:

$$\max_{d_{i,l}\in\{0,1\}} \Pr(D_l^* = d_{i,l}|d_{i,l},\alpha_i,\mathcal{I}_l) - \sum_{k,q} w_{i,l,k}\left[a_{q,k}(\tilde{y}_l)\Big(\sum_{d\in\{0,1\}} \mathcal{F}_{l,q,d}(s_l;\alpha_i)\Pr(D_l^* = d|d_{i,l},\mathcal{I}_l)\Big)\right]. \quad (7)$$

where $\Pr(D_l^* = d_{i,l}|d_{i,l};\alpha_i,\mathcal{I}_l)$ depends on strategies adopted by other members than i.

For an empirical example of this model, consider the Monetary Policy Committee (MPC) at the Bank of England. The MPC meets each month, indexed by l, to set an interest rate that they believe will help to achieve the announced policy targets in inflation ($\widetilde{\pi}_l$) and GDP ($\tilde{y}_l$). See for example Besley, Meads, and Surico (2008). Each member weighs the outcomes in GDP and inflation (relative to their strategic incentives) differently; and each has an updated perception $\mathcal{F}_{l,q,d}(s;\alpha_i)$ about how the chosen interest rate will affect prospective outcomes. Such a perception is a weighted average of two forecasts about inflation and output under different interest rates: one by MPC and the other by outside professionals from the private financial sector. The outside forecasts are reported quarterly, whereas the forecasts by MPC are adjusted through monthly deliberations prior to decision. Members in a committee may differ in the lengths of their tenure.

3.2 Equilibrium: definition and existence

In the subsequent discussion about the model equilibrium, we condition the strategies and the equilibrium definition on the common information $\mathcal{I}_l$, and suppress it from the notation when there is no ambiguity. A pure strategy profile is defined as $\sigma \equiv \{\sigma_i\}_{i\in\mathcal{C}}$, where σ_i maps from the support of private types to the binary action. Let $\pi_i(d_i,\alpha_i,w_{i,l};\sigma_{-i})$ denote the ex ante payoff for a member i, given his choice d_i and types $(\alpha_i,w_{i,l})$ and others' strategies $\sigma_{-i} \equiv \{\sigma_j\}_{j\neq i}$.

Definition 1 *A profile σ is a pure-strategy Bayesian Nash equilibrium (p.s.BNE) if for all i and $(\alpha_i,w_{i,l})$,*

$$\sigma_i(\alpha_i,w_{i,l}) = \arg\max_{d_i\in\{0,1\}} \pi_i(d_i,\alpha_i,w_{i,l};\sigma_{-i})$$

where $\sigma_{-i} \equiv \{\sigma_j\}_{j\neq i}$. A p.s.BNE is symmetric if $\sigma_i = \sigma_j$ for all i,j.

We give an intuitive argument for the existence of p.s.BNE in the text, and leave the technical details in Appendix B.1. First, we show that a member i's best response to the others' strategies σ_{-i} takes a simple form: choose $d_i = 1$ if i's ex ante difference between the losses under the two alternatives is lower than a threshold. This thresholds are determined by choices probabilities of other members under their strategies in σ_{-i}. Such an expected difference, denoted by Δ_i, depends on individual types $\alpha_i, W_{i,l}$ and common information $\mathcal{I}_l$. Let $\phi : \mathbb{R} \to [0,1]$ denote the CDF of i's ex ante difference Δ_i.

Next, define $\varphi : [0,1] \to \mathbb{R}$ so that $\varphi(\tau)$ is the afore-mentioned threshold defining i's best response when all other members adopt identical pure strategies that lead to an individual CCP $\tau \in (0,1)$. We show the mapping φ has the following properties in the limit:

$$\varphi(\tau) \to -\infty \text{ as } \tau \to 0 \text{ and } \varphi(\tau) \to +\infty \text{ as } \tau \to 1. \tag{8}$$

This means if the other members' strategies are such that their CCPs diminish to zero, then a member's best response is such that his own CCP also diminishes to zero. Note that we drop the subscript i from both φ and ϕ due to our focus on symmetric equilibria, and we suppress the dependence of ϕ, φ on $\mathcal{I}_l$ for simplicity.

A symmetric p.s.BNE is then characterized by a threshold $\kappa^* \in \mathbb{R}$ such that

$$\varphi \circ \phi(\kappa^*) = \kappa^* \tag{9}$$

where " $\circ$ " denotes the composition of two functions. With $\varphi \circ \phi$ being continuous under maintained assumptions, it suffices to show there exist κ, κ' so that $\varphi \circ \phi(\kappa) < \kappa$ and $\varphi \circ \phi(\kappa') > \kappa'$. Then the Intermediate Value Theorem implies a solution to the fixed-point equation in (9) exists. With the established property of best response in (8), it is easy to show such a pair (κ, κ') exists. For example, if both α_i and W_i have bounded support, then the support of ex ante difference Δ_i is also bounded. Hence $\varphi \circ \phi(\kappa) < \kappa$ for κ sufficiently small and $\varphi \circ \phi(\kappa') > \kappa'$ for κ' sufficiently large. It is also possible to show existence of such a pair κ, κ' when the support of private types is unbounded under some mild technical conditions on the tail of the distribution ϕ. See Appendix B.1 for details.

In general, the model of strategic recommendations admits multiple p.s.BNE because (9) could well admit multiple solutions for a given $\mathcal{I}_l$. We follow the convention of literature on empirical games (e.g. Bajari, Hong, Krainer, and Nekipelov (2010) and Lewbel and Tang (2013)), and assume that data-generating process only involves a single symmetric p.s.BNE. That is, across all committees (games) with identical information $\mathcal{I}_l$, the same p.s.BNE is played. For the rest of Section 3, we maintain that members' recommendations in data are rationalized by the symmetric p.s.BNE defined above.

It is worth mentioning that in principle the argument for equilibrium existence (as well as the identification argument presented below) could be modified to accommodate asymmetric pure-strategy BNE. Nevertheless this generalization comes at the cost of more stringent requirement for the data: The identification argument and consequently the estimation procedure need to be based on choice probabilities conditional on member identities. Given the data constraint, we only focus on the case with symmetric p.s.BNE in our application.

Our goal is to recover the distributions of ideological bias α_i and tastes $W_{i,l}$ from the distributions of individual recommendation $D_{i,l}$ and committee decisions D_l^* given the information in $\mathcal{I}_l$. This is done in three steps. First, we recover type-specific CCPs $\Pr(D_{i,l} = d_{i,l} \mid \mathcal{I}_l, \alpha_i)$ and the probability masses of α_i (as the components and weights of a finite mixture respectively) up to unknown values of α_i. Second, we show the "type-specific" CCPs are monotonic in α_i. Together with the first step, this allows us to pin down the probability mass function of α_i and the type-specific CCPs. Third, we identify the distribution of $W_{i,l}$ using the impact of continuous changes in $\mathcal{I}_l$ on the type-specific CCPs $\Pr(D_{i,l} = d_{i,l} \mid \mathcal{I}_l, \alpha_i)$.

3.3 Recovering type-specific CCPs

The first step of identification exploits the panel structure of the data. Under Assumptions 4 and 5, the private types $(\alpha_i, W_{i,l})$ and the common information $\mathcal{I}_l$ are both drawn every L episodes in the data. Thus for identification purposes, we consider the DGP equivalently as one in which each cross-sectional unit (i.e. an independent committee) is observed to make decisions in L episodes. For the rest of this section, we maintain that, for a known s_l, the realized information $G(s_l)$ can be conditioned on in the data. Among other things, this happens when $G(s_l)$ is observed for a subset of the state space $\mathcal{S}$. For instance, in the application of monetary policy decisions at the Bank of England in Section 4, the initial perception refers to quarterly forecasts of policy outcomes by outsiders, which is reported in data.

Let $\mathcal{X}$ denote the support of the common information set $\mathcal{I}_l$, and let the lower case $x_l \equiv \{s_l, \tilde{y}_l, G(s_l), H_l(s_l)\}$ denote a realized value of $\mathcal{I}_l$. Let $\mathbf{d}_i^t \equiv (d_{i,l})_{1 \le l \le t}$, for all $1 \le t \le L$. That is, $\mathbf{d}_i^t$ denotes i's decisions up to the t-th episode in the cross-sectional unit. Likewise, let $\mathbf{s}^t$, $\tilde{\mathbf{y}}^t$, $\mathbf{G}^t(\mathbf{s}^t)$ and $\mathbf{H}^t(\mathbf{s}^t)$ denote the history of states, targets, initial perception and empirical evidence, respectively up to the t-th episode. Let $\mathbf{x}^t$ denote the history of the common information. For a generic vector A that consists of discrete components A_d and continuous components A_c, we use $\Pr(a_d, A_c = a_c \mid \omega)$, or simply $\Pr(a_d, a_c \mid \omega)$, as a shorthand for $\frac{\partial}{\partial a} \Pr(A_d = a_d, A_c \le a \mid \omega)\big|_{a=a_c}$ for any generic event ω.

Assumption 6 *(a) For all $l \le L$ and all $\mathbf{d}_i^{l-1}, \mathbf{x}^l$ and α_i, the transition function of x_l satisfies $\Pr(x_l \mid \mathbf{d}_i^{l-1}, \mathbf{x}^{l-1}; \alpha_i) = \Pr(x_l \mid d_{i,l-1}, x_{l-1})$. (b) For all x_{l-1} and $d_{i,l-1}$, $\Pr(x_l \mid d_{i,l-1}, x_{l-1}) > 0$ for all x_l on the support of $\mathcal{I}_l$.*

Part (a) requires the transition of the common information to follow a first-order Markov process that is time-homogenous. In the example of MPC at the Bank of England, such stationarity holds when each committee member's choices depend on the current information set of states in the same way across multiple episodes. Part (b) states that starting from any combination of past state, target, decision and empirical evidence, any state and empirical evidence are reachable in the subsequent episode with positive probability.

It follows from Assumption 6 that type-specific CCPs $\Pr(d_{i,l} | \mathbf{d}_i^{l-1}, \mathbf{x}^l, \alpha_i)$ only depend on the contemporary information set x_l and ideological bias α_i, and are time-homogenous. To see this, recall that a member's decision in episode l is a function of $W_{i,l}, \alpha_i$ and $\mathcal{I}_l$. With α_i fixed across L episodes and $W_{i,l}$ independent across l and orthogonal to $\mathcal{I}^{l-1}$, the type-specific CCPs $\Pr(D_{i,l} = 1 \mid \mathbf{x}^l, \mathbf{d}_i^{l-1}, \alpha_i)$ is a function of x_l and α_i only, and is time-homogeneous under the maintained assumptions.

Thus by the law of total probability and Assumption 6, the joint distribution $\Pr(\mathbf{d}_i^L, \mathbf{x}^L)$ is:

$$\sum_{\alpha_i \in \mathcal{A}} \rho(\alpha_i) \Pr(\mathbf{d}_i^L, \mathbf{x}^L \mid \alpha_i) = \sum_{\alpha_i \in \mathcal{A}} \Pr(d_{i,1}, x_1, \alpha_i) \prod_{l=2,.,L} \Pr(x_l \mid d_{i,l-1}, x_{l-1}) \Pr(d_{i,l} \mid x_l, \alpha_i)$$

where $\rho(\cdot)$ is the probability mass function for α_i. When heterogeneous ideological bias α_i is fixed in multiple episodes, decisions of committee members described above are analogous to dynamic discrete choices with unobserved time-variant individual heterogeneity. Hence the recent development of nonparametric identification of similar models in a framework of measurement errors, e.g., Hu (2008) can be applied in this step of identification to recover the type-specific CCPs $\Pr(d_{i,l} \mid x_l, \alpha_i)$ and $\rho(\alpha_i)$. We show below that data of three periods are enough for identification. Without loss of generality, let $l-2, l-1$ and l denote the three periods, we also drop the index of committee member i whenever there is no ambiguity.

Let f denote the joint probability mass and density of individual member choice and information in three consecutive periods, $D_l, \mathcal{I}_l, D_{l-1}, \mathcal{I}_{l-1}$, D_{l-2} and $\mathcal{I}_{l-2}$:

$$\begin{aligned} & f_{D_l,\mathcal{I}_l,D_{l-1},\mathcal{I}_{l-1},D_{l-2},\mathcal{I}_{l-2}} = \sum\nolimits_{j=1}^{J} f_{D_l,\mathcal{I}_l,D_{l-1},\mathcal{I}_{l-1},D_{l-2},\mathcal{I}_{l-2},\alpha=\alpha^j} \\ = & \sum\nolimits_{j=1}^{J} f_{D_l|\mathcal{I}_l,\alpha} f_{\mathcal{I}_l|D_{l-1},\mathcal{I}_{l-1}} f_{D_{l-1}|\mathcal{I}_{l-1},\alpha} f_{\mathcal{I}_{l-1}|D_{l-2},\mathcal{I}_{l-2}} f_{D_{l-2},\mathcal{I}_{l-2},\alpha=\alpha^j} \end{aligned}$$

where the second equality is due to the restrictions imposed on CCP and law of motion in Assumption 6. The relationship holds for all possible values of $D_l, \mathcal{I}_l, D_{l-1}, \mathcal{I}_{l-1}, D_{l-2}, \mathcal{I}_{l-2}$ and α. This allows us to fix $(D_l, D_{l-1}, \mathcal{I}_{l-1}, D_{l-2})$ at any of their realizations $(d_l, d_{l-1}, x_{l-1}, d_{l-2})$ and rewrite the the equation above as:

$$\frac{f_{d_l,\mathcal{I}_l,d_{l-1},x_{l-1},d_{l-2},\mathcal{I}_{l-2}}}{f_{\mathcal{I}_l|d_{l-1},x_{l-1}} f_{x_{l-1}|d_{l-2},\mathcal{I}_{l-2}}} = \sum\nolimits_{j=1}^{J} \Pr(d_l \mid \mathcal{I}_l, \alpha) \Pr(d_{l-1} \mid x_{l-1}, \alpha) f_{d_{l-2},\mathcal{I}_{l-2},\alpha=\alpha^j},$$

where the left-hand side is directly identifiable from data and the components on the right-hand are not directly identifiable and contains the type-specific CCPs $\Pr(d_l \mid \mathcal{I}_l, \alpha)$. Similarly, the joint distribution $f_{D_{l-1},\mathcal{I}_{l-1},D_{l-2},\mathcal{I}_{l-2}}$ at a given pair of realized choices (d_{l-1}, d_{l-2}) can be decomposed into

$$\frac{f_{d_{l-1},\mathcal{I}_{l-1},d_{l-2},\mathcal{I}_{l-2}}}{f_{\mathcal{I}_{l-1}|d_{l-2},\mathcal{I}_{l-2}}} = \sum\nolimits_{j=1}^{J} \Pr(d_{l-1} \mid \mathcal{I}_{l-1}, \alpha) f_{d_{l-2},\mathcal{I}_{l-2},\alpha=\alpha^j}. \tag{10}$$

To apply the arguments of identification in Hu (2008), we partition the support of $\mathcal{I}_l$ of each l into M intervals, and define $\mathcal{J}_l = k$ if $\mathcal{I}_l$ is in the k-th interval. Consequently, the preceding two displays have the following matrix forms:

$$L_{d_l,\mathcal{J}_l,d_{l-1},x_{l-1},d_{l-2},\mathcal{J}_{l-2}} = L_{d_l|\mathcal{J}_l,\alpha}\Lambda_{d_{l-1}|x_{l-1},\alpha}L_{d_{l-2},\alpha,\mathcal{J}_{l-2}} \quad (11)$$

$$L_{d_{l-1},\mathcal{J}_{l-1},d_{l-2},\mathcal{J}_{l-2}} = L_{d_{l-1}|\mathcal{J}_{l-1},\alpha}L_{d_{l-2},\alpha,\mathcal{J}_{l-2}}, \quad (12)$$

where $\Lambda_{d_{l-1}|x_{l-1},\alpha}$ is a $J \times J$ diagonal matrix with its k-th diagonal element $\Pr(d_{l-1} \mid x_{l-1}, \alpha = \alpha^k)$; $L_{d_l,\mathcal{J}_l,d_{l-1},x_{l-1},d_{l-2},\mathcal{J}_{l-2}}$, $L_{d_l|\mathcal{J}_l,\alpha}$, $L_{d_{l-2},\alpha,\mathcal{J}_{l-2}}$, $L_{d_{l-1},\mathcal{J}_{l-1},d_{l-2},\mathcal{J}_{l-2}}$ are of dimensions $M\times M$, $M\times J$, $J \times M$, $M \times M$ respectively, with their entries defined as follows:

$$\begin{aligned}
(L_{d_l,\mathcal{J}_l,d_{l-1},x_{l-1},d_{l-2},\mathcal{J}_{l-2}})_{i,j} &= \frac{\Pr(d_l, \mathcal{J}_l = i, d_{l-1}, x_{l-1}, d_{l-2}, \mathcal{J}_{l-2} = j)}{\Pr(\mathcal{J}_l = i \mid d_{l-1}, x_{l-1})\Pr(x_{l-1} \mid d_{l-2}, \mathcal{J}_{l-2} = j)}, \\
(L_{d_{l-1},\mathcal{J}_{l-1},d_{l-2},\mathcal{J}_{l-2}})_{i,j} &= \frac{\Pr(d_{l-1}, \mathcal{J}_{l-1} = i, d_{l-2}, \mathcal{J}_{l-2} = j)}{\Pr(\mathcal{J}_{l-1} = i \mid d_{l-2}, \mathcal{J}_{l-2} = j)}, \\
\left(L_{d_l|\mathcal{J}_l,\alpha}\right)_{i,k} &= \Pr(d_l|\mathcal{J}_l = i, \alpha = \alpha^k), \\
(L_{d_{l-2},\alpha,\mathcal{J}_{l-2}})_{k,j} &= \Pr(d_{l-2}, \alpha = \alpha^k, \mathcal{J}_{l-2} = j),
\end{aligned}$$

for $i, j = 1, 2, \cdots, M$; $k = 1, 2, \cdots, J$.

Assumption 7 *(a) There exists a known partition of $\mathcal{X}$ into M intervals, with $M \geq J$, such that both $L_{d_{l-1}|\mathcal{J}_{l-1},\alpha}$ and $L_{d_{l-2},\alpha,\mathcal{J}_{l-2}}$ have rank J. (b) There exist J realizations of $\mathcal{I}_{l-1}$, denoted $x^{(1)}, \cdots, x^{(J)}$, such that $\Pr(d_{l-1} \mid \mathcal{I}_{l-1} = x^{(j)}, \alpha) > 0$ for all α; and $\Pr(d_{l-1} \mid \mathcal{I}_{l-1} = x^{(j)}, \alpha) \neq \Pr(d_{l-1} \mid \mathcal{I}_{l-1} = x^{(j)}, \alpha')$ for any $\alpha \neq \alpha'$ in $\mathcal{A}$, $j = 1, 2, \cdots, J$. Moreover, the matrix $(L_{d_{l-1}|\mathcal{I}_{l-1},\alpha})_{i,j} \equiv \Pr(d_{l-1} \mid \mathcal{I}_l = x^{(i)}, \alpha = \alpha^j)$ has full rank.*

Similar full rank conditions to Assumption 7 (a) are widely used for identifying models with unobserved heterogeneity in the literature, e.g., Hu (2008) and An (2010). Assumption 7 (a) implies that the rank of the $M\times M$ matrix $L_{d_{l-1},\mathcal{J}_{l-1},d_{l-2},\mathcal{J}_{l-2}}$ must be J. This is a restriction on the distribution of common information $\mathcal{I}$ and the ideological bias α. The full rankness of $L_{d_{l-1}|\mathcal{J}_{l-1},\alpha}$ requires that the CCPs to vary across members with different ideological bias once conditional on common information. Besides, recall that the information $\mathcal{I}_{l-2}$ is also affected by each member's α_i, because the ideological bias affects the committee decisions in previous episodes which in turn affects the dynamic transition of common information across episodes. Thus, the rank condition on $L_{d_{l-2},\alpha,\mathcal{J}_{l-2}}$ requires that ideological bias have enough impact on the dynamics of common information via committee choices. This is in line with Assumption 2 that requires sufficient variation in the empirical evidence.

Assumption 7 (b) is a non-degeneracy condition. Given the common information, committee members of different types have different choice probabilities. It necessarily requires that the initial perception and the empirical evidence should be sufficiently different from each other. In the extreme case where $G(s_l) = H_l(s_l)$ for all $l \leq L$, the updated perception, as a weighted average of the two sources, would be the same for members of all types. In the context of monetary policy decisions by the Bank of England, G and H_l are forecasts by the outsiders and MPC respectively, and as discussed in our empirical application, it is reasonable to assume that they are sufficiently different in general.

Lemma 2 *Suppose Assumptions 4, 5, 6 and 7 hold and* $L \geq 3$. *Then* $J = Rank(L_{d_l, \mathcal{J}_l, x_{l-1}, \mathcal{J}_{l-2}})$ *and the probability mass function for* α_i *and the stationary choice patterns* $\Pr(D_{i,l} = 1 \mid \mathcal{I}_l = x; \alpha_i)$ *are identified for all* $x \in \mathcal{X}$ *up to unknown values of* α_i.

The proof of this lemma follows from the insight in Hu (2008), and is included in Appendix B.2.

3.4 Ordering type-specific CCPs

To match the type-specific CCPs with realized values of α_i, we introduce mild conditions on the model primitives which imply that the type-specific CCPs are monotone in the ideological bias. In what follows, we drop subscripts l from $x_l \equiv \{s_l, \tilde{y}_l, G(s_l), H_l(s)\}$ and $W_{i,l}$ for simplicity.

Assumption 8 *(a)* W_i *is independent from* $\mathcal{I}$ *with support* $\mathcal{W} \equiv (0, \infty)^K$. *(b) There exists* $x^* \in \mathcal{X}$ *such that conditional on* x^* *there exists no* $w \in \mathcal{W}$ *so that members with* $W_i = w$ *are indifferent between the two alternatives regardless of* α_i.

Part (a) in Assumption 8 strengthens part (a) in Assumption 4. On the one hand, W_i, defined as individual preferences over the multiple dimensions of outcomes, is intended to capture the idiosyncratic tastes of committee members. On the other hand, $\mathcal{I}$ consists of common information that are not affected by such idiosyncratic tastes of committee members. In our empirical application of MPC decisions, this common information includes the policy targets set by the government $\tilde{Y}_l$, the outsiders' perception about policy consequences $G(\cdot)$ and the forecasts by the Bank of England $H_l(\cdot)$. Part (a) also requires individual tastes to be inherent in the sense that they are not responsive to the state of economy S_l. It is not clear how the current identification strategy could be extended to a richer model that relaxes part (a) and allows for general dependence between the members' tastes and the states.

Assumption 8 (b) requires that decisions made by the committee members of different ideological bias must diverge sufficiently. This is a mild condition that guarantees the component choice probabilities (conditional on the unobserved ideological bias) recovered from the previous step must be distinct for different realized values of bias. An and Tang (2013) provides a detailed discussion on how this condition is implied by primitive restrictions on model elements in Appendix D. In other words, part (b) helps to rule out technically pathological cases, and allows us to focus on the structural link between variation in the data and the recoverability of model elements.

Lemma 3 *Suppose Assumption 4 (c) holds. Let* $\mathcal{I} = x^*$ *satisfy Assumption 7 (b) and Assumption 8. Then* $\Pr(D_i = 1 \mid \mathcal{I} = x^*, \alpha_i)$ *is monotone in* α_i*, and the direction of monotonicity is identified.*

The proof is presented in Appendix B.3. The main idea for the lemma is to show that the set of tastes $w_{i,l}$ causing a member to choose 1 changes monotonically as individual bias α_i increases. Then the independence between W_i, α_i and $\mathcal{I}$ implies the type-specific CCPs must be monotonic in α_i. It is also shown that the direction of monotonicity is determined by the common information alone, and therefore is identifiable from the data.

3.5 Identifying the taste distribution

Having fully recovered the type-specific CCPs $\Pr(D_i = d_i \mid \mathcal{I} = x; \alpha_i)$ for all x and bias α_i, we turn to the identification of the distribution of tastes F_{W_i}. Due to the technical nature of the proof, we summarize the main idea for identification in the text, and leave a detailed formal proof in Appendix B.4.

Our identification result builds on an additional parametric structure: $W_{i,k} = W_{i,0} + \eta_{i,k}$ for $k \leq K$, where $W_{i,0}$ and $\eta_{i,k}$ are respectively the common and idiosyncratic factors in the tastes. They are assumed to be continuously distributed over $\mathcal{W}_0 \subset \mathbb{R}_+$ and $[\underline{\eta}_k, \overline{\eta}_k]$ respectively, and are independent from common information and ideological bias. The main idea is to use the variation in the common information to trace out the distribution of $W_{i,0}$ and $\eta_{i,k}$ from the type-specific CCPs.

By construction, we can write type-specific CCPs, which are already identified from Section 3.4, as:

$$\Pr(D_i = 1 \mid \mathcal{I} = x; \alpha_i) = \Pr\left\{\textstyle\sum_k W_{i,k} C_k(x; \alpha_i) \leq t(x)\right\} \tag{13}$$

for some C_k that depends on common information and individual bias and some $t(x)$ that is a function of directly identifiable choice probabilities. (See the first paragraph in Appendix B.3 for the closed forms of $C_k(.)$ and $t(.)$.) Applying the law of total probability, we can write (13) as an integral of the CDF of W_0, denoted by F_{W_0}. By changing variables between $\eta_{i,1}$ and the expression that enters $F_{W_0}(\cdot)$ in the integrand, we can write (13) as:

$$\Pr(D_i = 1 \mid \mathcal{I} = x; \alpha_i) = \int_0^\infty F_{W_0}(s)\, \mathcal{K}(s, x; \alpha_i) ds \tag{14}$$

where the left-hand side is already identified in the preceding step, and $\mathcal{K}(\cdot, x; \alpha_i)$ on the right-hand side is a type-specific function of the common information whose closed form is presented in Appendix B.4. The lower limit of the integral is 0 because W_0 is non-negative. The functional form of $\mathcal{K}(\cdot, x; \alpha_i)$ depend on the distribution function of η_k, which is assumed known. Furthermore, as shown in Section 3.4, the type-specific CCPs and the probability masses of α_i are fully identified. Thus $\mathcal{K}$ is identified over its respective domains. As long as a kernel component in $\mathcal{K}(\cdot,\cdot;\alpha_i)$ satisfies a completeness condition at least for some values of α_i, the integral equation in (14) admits a unique solution in $F_{W_0}(\cdot)$.

3.6 Estimation strategy

We propose a two-step procedure for estimating the probability masses of α_i and the distribution of W_0. First, we estimate the cardinality of $\mathcal{A}$ as the rank of the matrix $L_{d_l,\mathcal{J}_l,x_{l-1},\mathcal{J}_{l-2}}$ (denote the matrix Q) defined in Section 3.3; then estimate the probability masses of α_i and $F_{W_0}(\cdot)$ using a sieve maximum likelihood estimator (MLE).

Lemma 2 states $J = \text{Rank}(Q)$. To estimate the rank of Q, we propose to use the estimator from Robin and Smith (2000). This estimator is based on a sequential test of the null hypotheses H_r: rank$(Q)= r$ against the alternative H_r': rank$(Q)> r$ for $r = 0,1,\cdots,J-1$. Specifically, it is defined as the minimum of r such that H_i is rejected for all $i = 0,1,\cdots,r-1$ and H_r is not rejected. The critical regions are obtained based on the result that the limiting distribution of the test statistic is a weighted average of χ^2-distributions for each step of testing. Allowing the significance level of each step to depend on the sample size appropriately, it can be shown (Theorem 5.2 in Robin and Smith (2000)) that the rank of Q can be consistently estimated.

Next, we estimate p^*_{d,d_i} in the equilibrium (the probability that $D = d$ given $D_i = d_i$) using local (constant, linear or polynomial) kernel regressions. Denote these estimates by $\hat{p}_{d,d_i}$. Standard arguments (such as in Fan and Yao (2005)) can be used to show the uniform consistency of $\widehat{p}_{d,d_i}(x)$ under appropriate regularity conditions. Then plug in $\hat{p}_{d,d_i}$ to estimate $t(\cdot)$ and $C_k(\cdot)$ for $k \leq K$. Denote these estimates by $\hat{t}(\cdot)$ and $\hat{C}_k(\cdot)$.

Let I be the number of members in a committee $\mathcal{C}$, which is fixed across cross-sectional units. Let $\mathcal{W}_0$ denote the support of $W_{i,0}$; let $\mathcal{P} \equiv \{p \in [0,1]^J : p_1 + p_2 + ... + p_J = 1\}$. Let $\mathcal{F}_n$ denote an appropriately chosen sequence of sieves space for continuous distributions over $\mathcal{W}_0$ such that as $n \to \infty$, $\mathcal{F}_n$ becomes dense in the parameter space for F_{W_0}. Then the sieve MLE is given by

$$\widehat{\theta} \equiv (\widehat{p},\widehat{F}) = \arg\max_{p\in\mathcal{P},F\in\mathcal{F}_n} \sum_{n=1}^N \sum_{i=1}^I \log \widehat{\mathcal{L}}_{i,n}(p,F) \tag{15}$$

where i,n are indices for a committee member and a committee respectively, N is the sample size for committees, and $\widehat{\mathcal{L}}_{i,n}$ is the estimated likelihood based on the distribution of individual and committee decisions given common information throughout L decision episodes. Specifically, for each $\theta \equiv (p,F)$, let $\widehat{\Phi}_{j,e_i}(x_t;\theta)$ be an simulation-based estimator for the integral on the right-hand side of (14), using simulated draws in s and a sample analog estimator for $\mathcal{K}$ (which involves $\widehat{t}(\cdot)$ and $\widehat{C}_k(\cdot)$ from the first step as well as the known pdf of η_k). Then, the estimated likelihood is

$$\widehat{\mathcal{L}}_{i,n}(\theta) = \sum_{j=1}^J p_j \left\{ \prod_{l=1}^L \left[\widehat{\Phi}_{j,e_i}(x_t;\theta)\right]^{d_{i,l}} \left[1-\widehat{\Phi}_{j,e_i}(x_t;\theta)\right]^{1-d_{i,l}} \right\}.$$

Conditions for consistency of the sieve MLE, and the choice of the sieve space $\mathcal{F}_n$, should follow the guidelines from Shen (1997), Chen and Shen (1998) and Ai and Chen (2003). We expect $\widehat{\theta}$ to be consistent, as long as the objective function in (15) satisfies the conditions listed in those references after taking account of the estimation errors in $\widehat{t}(\cdot)$ and $\widehat{C}_k(\cdot)$. We leave the formal proof (along with the necessary technical conditions) for consistency to future research. In our empirical section below, we adopt a parametric approach for estimation. Our estimator can also be generalized to include individual heterogeneity reported in data.

4 Empirical Application

In this section, we apply the model of strategic recommendations to analyze the interest rate policy decisions made by the Monetary Policy Committee (MPC) at the Bank of England. The committee consist of nine members. Five of them are internal members, who hold full-time executive positions in the bank; the other four are external with no executive responsibilities within the bank. The committee meets monthly to vote for an interest rate so as to achieve the targeted inflation rate and output (measured by GDP, or gross domestic product). Recommendations from individual members are aggregated using a simple majority rule.

Each MPC member makes monthly decisions throughout his tenure. An external member serves a 3-year term, and may be reappointed for another three years; internal members generally serve longer terms than their external colleagues. A member thus cares about whether his choice conforms to the final committee decision because of concerns about reappointment, and/or reputation for good judgement. Hence we analyze MPC decisions at the Bank of England using the model of strategic recommendations.

Our model specification is motivated by empirical evidence in the literature. First, Berk, Bierut, and Meade (2010) documented that MPC forecasts of inflation and output under various interest rates, and predictions by professionals in the private financial sector are both instrumental for the choice of monetary policies. These sources of information are available to all committee members and recorded in the data. Second, Besley, Meads, and Surico (2008) and Harris, Levine, and Spencer (2011) concluded that the characteristics of MPC members reported in the data are not sufficient for explaining discrepancies in their voting patterns. Because the MPC is organized to collectively process information from various sources through group deliberations, it is plausible that at the end of the day all members of the committee share the same pooled information set. Hence, we only attribute the members' distinct decisions to idiosyncracies in how they weigh multiple policy objectives (heterogeneous tastes) and how they utilize information from several sources (ideological bias). Our empirical analysis also takes into account the strategic interaction between members due to their concerns about conformity to committee decisions.

We estimate the structural model of MPC decisions with strategic career concerns to investigate how members' private types affect their recommendations. This provides some evidence about how the effects differ across external and internal members in the committee. To the best of our knowledge, this marks the first effort to analyze MPC decisions using a structural model involving members' (unobserved) private types.

4.1 Data

The data for the analysis are compiled from several sources. Our first source of data is the publicly available minutes from MPC meetings at the Bank of England. These minutes are from 187 monthly meetings between June 1997 and June 2013, involving 32 committee members (13 internal and 19 external). These minutes include the complete voting records of all committee members as well as their observed types (external or internal). In 179 out of the 187 meetings, the committee members vote for two different rates. Therefore we practically model the members' choices as binary $d_{i,l} \in \{0, 1\}$, with 0 being "vote a lower interest rate" and 1 being "vote for a higher interest rate".

The second source of data is the Bank of England's *Inflation Report*, which has been published quarterly since August 1997. These reports contain MPC's monthly forecasts about the key economic indicators such as inflation rates and gross domestic product (GDP). In addition to *monthly* MPC forecasts, the reports also include summaries of *quarterly* surveys conducted by the Bank of England. The surveys sample external professionals from (mostly London-based) financial institutions in the private sector, and ask them to forecast key economic indicators under various interest rate policies. Because the institutions represented in the survey are prominent and the sample size is fairly large, these outsider forecasts offer a good measure of general sentiments from the private sector, which is also taken into account in MPC decisions. Following the literature, we use the forecasts formulated on an assumption of a constant interest rate.

To identify the model elements, we rely on the exogenous variation in common information (including MPC forecasts) that is independent from the members' private types. We argue such an orthogonality condition is plausible and consistent with the interpretation of MPC forecasts and members' private types in our context. The MPC forecasts are constructed by its members through group deliberations and internal communication. To a large extent, they reflect the members' consensus about the economic outlook, and thus can be interpreted as pooled information after deliberations. In contrast, ideological bias α_i and tastes for multiple objectives W_i are meant to capture the idiosyncracies in members' perspectives that persist after the committee deliberations. Thus it is reasonable to posit these remaining idiosyncracies are orthogonal to whatever consensus the MPC members form.

The third source of the data is the historical record of U.K. monthly inflation rates and quarterly GDP provided by the Office for National Statistics (ONS).

Our sample contains the data from these three sources between August 1998 and June 2013, which last for 60 quarters or 180 months including the emergency MPC meeting in September 2001. (The complete record of both MPC forecasts *and* outsider surveys only dates back to August 1998.) Table 1 presents summary statistics for the voting records and membership of committee members. On average, each internal and external member attends 75 and 37 meetings respectively. Internal members vote for a lower interest rate ($d_{il} = 0$) in 65.3% of the meetings, and the external members in 73.5% of the meetings.

In each *Inflation Report*, the prediction of inflation rates in both MPC forecasts and outsider surveys are reported as probability distributions with the support $\{< 1.5\%, [1.5\%, 2.5\%), [2.5\%, 3.5\%), > 3.5\%\}$. Similarly, the forecasts of GDP growth are reported as distributions supported on $\{< 1\%, [1\%, 2\%), [2\%, 3\%), > 3\%\}$. The data are summarized in Table 2 for both sources. For instance, in outsider surveys the external professionals forecast for the probability that the GDP grows by more than 3% is 0.187 on average, and has a minimum at 0.01 and a maximum at 0.047; whereas MPC forecasts the same probability to be 0.306 on average with the minimum and maximum at 0 and 0.88 respectively. From August 1998 to June 2013, the ONS reports the average quarterly growth rate of GDP is 0.624% and the average inflation rate is 2.14%.

4.2 Econometric specification

The MPC meets monthly to set an interest rate, guided by some target rates of GDP growth and inflation. Denote these targets by $(\widetilde{y}_l, \widetilde{\pi}_l)$, where $\widetilde{y}_l$ is the GDP growth rate and $\widetilde{\pi}_l$ the inflation rate. The GDP target $\widetilde{y}_l$ is measured as the potential GDP growth rate using Hodrick-Prescott filter (Hodrick and Prescott (1980)) with a smoothing parameter set to be 1600 from the Bank of England's vintage data of GDP. Such an approach has been widely used in the literature to estimate potential GDP, e.g., Batini, Jackson, and Nickell (2005).

We verify that the potential GDP obtained from the Hodrick-Prescott filter is consistent with the forecast of inflation using the New Keynesian Phillips Curve (NKPC) $\pi_t = \beta E_t(\pi_{t+1}) + \kappa \Delta y_t$, where β is the discount factor and $E_t(\pi_{t+1})$ is the forecasted inflation rates for period $t+1$ and Δy_t is the GDP gap at period t. In this formulation, $\beta E_t(\pi_{t+1})$ reflects the discounted current expectations of next period's inflation rate. Specifically, we use the historical inflation rates as π_t and Bank of England's forecasts as $E_t(\pi_{t+1})$. The GDP gap is calculated as the difference between the logarithm of historical GDP and logarithm of potential GDP obtained using Hodrick-Prescott filter. We follow the literature (e.g., Galí and Gertler (1999)) and set the discount factor $\beta = 0.95$. We then estimate the parameter κ by a simple linear regression. The estimated κ is 0.01 and this is consistent with the textbook evidence (e.g., Galí (2009)) and some empirical results (e.g., Mavroeidis, Plagborg-Møller, and Stock (2014)), where κ ranges from 0.005-0.05. Considering the complexity of estimating Phillips Curve, our simple analysis here only serves as an auxiliary check of the results for the Hodrick-Prescott filter.

The inflation target of Bank of England $\widetilde{\pi}_l$ is 2.5% up to December 2003 and 2.0% from January 2004. The state s_l contains the current inflation rate and growth rate of GDP. In our context, $H_l(y, \pi \mid s, d)$ and $G(y, \pi \mid s, d)$ are constructed from the MPC forecasts and the outsider surveys, respectively. The support of state $\mathcal{S}$ is discretized as values $\{(y^{\text{high}}, \pi^{\text{high}}), (y^{\text{high}}, \pi^{\text{low}}), (y^{\text{low}}, \pi^{\text{high}}), (y^{\text{low}}, \pi^{\text{low}})\}$.

Given the announced targets $(\tilde{y}, \tilde{\pi})$, the current states $s = (y, \pi)$ and two information sources H and G, a MPC member votes for an interest rate that solves (7). Of course, the recommendation is determined by the member's realized types $W_i \equiv (W_{i,1}, W_{i,2})$ and ideological bias α_i, with $W_{i,1}$, $W_{i,2}$ being the weights this member puts on GDP and inflation respectively relative to his career concerns.

A member's ideological bias reflects the weights applied to MPC and outsider forecasts in his updated perception. Because MPC forecasts are updated monthly while the outsider surveys are collected quarterly, we maintain that weights on these sources α_i are drawn at the same frequency (quarterly) as the outsider forecasts. On the other hand, policy targets are announced prior to each monthly meetings. Thus we allow members to draw their tastes for multiple policy objectives W_i monthly. This is consistent with identifying conditions in Assumptions 4 and 5.

As in Assumption 10, let $W_{i,k} = W_{i,0} + \eta_{i,k}$ for $k = 1, 2$, where $\eta_{i,k}$ are assumed to be i.i.d. draws from a truncated standard normal distribution between $[-2, 2]$ for $k = 1, 2$. The distribution of $W_{i,0}$ depends on the observed type of members ($E_i = 1$ if the member is external and $E_i = 0$ otherwise) and is specified as a gamma distribution with parameter (a^e, b^e) for $e = 0, 1$. The support of individual bias α_i, $\mathcal{A}$ is $\{\alpha^1,\ \alpha^2,\ \alpha^3\} = \{1/4,\ 1/2,\ 3/4\}$ to capture how members weigh MPC forecasts and outsider forecasts from the surveys. The distribution of α_i also depends on E_i, and its probability masses are denoted as $p^{\alpha}_{j,e} \equiv \Pr(\alpha_i = \alpha^j \mid E_i = e)$ for $1 \leq j \leq 3$ and $e = 0, 1$.

4.3 Estimation strategy

We estimate the parameters in the distribution of private types $\theta \equiv \{a^e,\ b^e,\ p^{\alpha}_{1,e},\ p^{\alpha}_{2,e},\ p^{\alpha}_{3,e} : e = 0, 1\}$. Our estimation strategy is based on (15), with model primitives parametrized as above. The first step is to estimate $C_k(x; \alpha_i) \equiv \alpha_i \left(A^H_{1,k} - A^H_{0,k}\right) + (1 - \alpha_i)\left(A^G_{1,k} - A^G_{0,k}\right)$ for $k = 1, 2$; and $t(x) \equiv p^*_{1,1}(x) - p^*_{0,0}(x)$. Both objectives depend on p^*_{d,d_i}, the equilibrium probability that the committee decision is d given a member i's decision d_i. Once p^*_{d,d_i} is estimated, both $C_k(x, \alpha_i)$ and $t(x)$ can be estimated by plugging in the estimates of p^*_{d,d_i}.

We estimate $p^*_{1,1}$ and $p^*_{0,0}$ by expressing them as functions of type-specific individual choice probabilities $\lambda_{Ext} \equiv \Pr(d_i = 1|E_i = 1, x)$ and $\lambda_{Int} \equiv \Pr(d_i = 1|E_i = 0, x)$, which are then estimated via a logit specification with parameters ξ_j and ϑ_j for $0 \leq j \leq 4$:

$$\begin{aligned} \lambda_{Ext}(x) &= \exp(\xi_0 + \textstyle\sum_{j=1}^4 \xi_j\omega_j)/\big[1 + \exp(\xi_0 + \sum_{j=1}^4 \xi_j\omega_j)\big], \\ \lambda_{Int}(x) &= \exp(\vartheta_0 + \textstyle\sum_{j=1}^4 \vartheta_j\omega_j)/\big[1 + \exp(\vartheta_0 + \sum_{j=1}^4 \vartheta_j\omega_j)\big], \end{aligned} \tag{16}$$

where ω_j for $1 \leq j \leq 4$ are quantities that are directly recoverable from the data and that describe how committee members utilize common information in their optimization problem. Specifically, ω_1 (and ω_2) are defined as

$$\textstyle\sum_q a_{q,k}(\tilde{y})\big[H_{q,1}(s) - H_{q,0}(s)\big]$$

where k corresponds to inflation (and to GDP respectively); ω_3 (and ω_4) are defined similarly only with H replaced by G. Let $\hat{\xi}_j$ and $\hat{\vartheta}_j$ denote the logit estimate of ξ and ϑ, respectively. By plugging $\hat{\xi}_j$ and $\hat{\vartheta}_j$ into the definition of $C_k(x, \alpha_i)$ and $t(x)$, we obtain the estimates $\widehat{C}_k(x, \alpha_i)$ and $\hat{t}(x)$.

We verify Assumption 8 using our data under the specification above. For all common information set x in our data, we solve from the solutions $(w_{i,1}, w_{i,2})$ for the system of equations: $w_{i,k}\widehat{C}_k(x; \alpha_i) = \hat{t}(x)$ for all $\alpha_i \in \mathcal{A}$. The solutions are sufficiently bounded away from the first quadrant in $\mathbb{R}^2$ with absolute values large relative to the standard errors. This suggests that our data satisfies Assumption 8 (as shown in Figure 2 and related discussions in Appendix B.3).

Our second step is to estimate the structural parameters in θ. Based on $\hat{t}(x)$ and $\widehat{C}_k(x; \alpha_i)$, we estimate the parameters of committee members' tastes and ideological bias θ in this step using MLE. Let $x_t = \{s_t, \tilde{y}_t, G_t(s_t), H_t(s_t)\}$ denote the information set available at an episode t; let i and $\tilde{t}$ index individuals and quarters in our data (with $i \leq 9$ and $\tilde{t} \leq 60$), respectively. The estimator of θ is

$$\widehat{\theta} = \arg\max_{\theta} \textstyle\sum_{i,\tilde{t}} \log \widehat{L}_{i,\tilde{t}}(\theta) \tag{17}$$

where

$$\widehat{L}_{i,\tilde{t}}(\theta) \equiv \textstyle\sum_j p_{j,e_i} \left\{ \prod_{t=3(\tilde{t}-1)+1}^{3\tilde{t}} \left[\widehat{\Phi}_{j,e_i}(x_t; \theta)\right]^{d_{i,t}} \left[1 - \widehat{\Phi}_{j,e_i}(x_t; \theta)\right]^{1-d_{i,t}} \right\}$$

where $\widehat{\Phi}_{j,e_i}(x_t; \theta)$ is a simulation-based estimator for the integral on the right-hand side of (14). The specific form of $\widehat{\Phi}_{j,e_i}$ involves $\hat{t}$ and $\hat{C}_k$, and is presented in Appendix C. To calculate the standard error of our estimator, we use a bootstrap resampling procedure.

4.4 Results

Table 3 presents the estimates for ξ and ϑ in the first step. In general, external and internal members of MPC process two sources of information in a similar manner. Among the four categories of forecasts, the outsiders' forecasts of GDP play a less important role than other forecasts for both internal and external members (ξ_4 and ϑ_4 are both insignificant). On the other hand, MPC forecasts of inflation and GDP, and outsiders' forecast of inflation all affect the voting patterns significantly for both types of members.

The estimates of ideological bias for the two observed types are summarized in Table 4. The standard errors are calculated using 200 bootstrap iterations. Recall that α is a member's weight on the MPC forecasts H in his updated perception. The estimates demonstrate that both groups of members are more likely to put greater weights on MPC forecasts (55.1% for external members and 61.3% for internal members). There is also significant probability that they put more weights on outsiders' forecasts (G) (about 30% for both groups). Furthermore, it is unlikely that the members apply equal weights to both sources of information (i.e. the probability of $\alpha = 50\%$ is statistically insignificant for both types of members). The point estimates suggest that internal members might focus slightly more on MPC's forecasts than external members, which is consistent with the estimates of the logit model in the first step. Nevertheless, a formal test based on the Wald statistic fails to reject the null that the weights on MPC forecasts are identical for internal and external members even at a 10% level. (Formally, the null is $(p^{\alpha}_{1,1}, p^{\alpha}_{2,1}) = (p^{\alpha}_{1,0}, p^{\alpha}_{2,0})$. The consistent estimator for the covariance matrix in the Wald statistic is constructed via bootstrap resampling. The degree of freedom in the limit distribution of the Wald statistic under the null is two. The test statistic is 0.601 with a p-value of 0.741). These results indicate that ideological bias plays an important role in committee members' decisions, and that the observed types of members do not have a significant impact on the voting patterns through ideological bias. These results reinforce the empirical findings in the literature (e.g., Besley, Meads, and Surico (2008) and Harris, Levine, and Spencer (2011)) that observed member characteristics alone cannot explain members' voting records.

Table 5 reports the point estimates for parameters in the distribution of $W_{i,0}$, i.e., (a^e, b^e) for $e = 0, 1$. Figure 3 plots the distributions of F_{W_0} for external members and internal members based on these point estimates. The estimated distribution for external members is shown to first-order stochastically dominate that for internal members. Recall that $W_{i,0}$ describes the deterministic part of $W_{i,1}, W_{i,2}$, or the weights that a member i applies to GDP and inflation relative to the career concerns. Hence a greater W_0 implies less relative weights on career concerns. Thus Figure 3 implies external members tend to apply less weight to strategic career concerns than internal members. A Wald test statistic for the null $H_0 : (a^1, b^1) = (a^0, b^0)$ is 4.748, which leads to a p-value of 0.093. (Again, the estimator for the covariance matrix in the Wald statistic is calculated via bootstrap resampling.) Thus, we reject the null that the distributions of W_0 is the same for external and internal members at 10% significance level. This can be interpreted as (somewhat weak) evidence that internal MPC members care more about conformity to the group decision than their external colleagues.

To approximate the sampling distribution of our estimators, we report empirical percentiles $(20\%, 50\%, 80\%)$ among 200 estimates based on the bootstrap samples in Table 5. The sign of difference in these estimated percentiles is consistent with the hypothesis that external members are less distorted by strategic incentives than internal members. That is to say, internal members care more about how likely their votes conform to the final committee decisions. This means the presence of external members in MPC could help to alleviate distortions in committee decisions due to strategic incentives for conformity.

To sum up, our empirical evidence shows that (1) MPC members tend to put more weights on the internal forecasts by the Bank of England than on external forecasts by professionals surveyed in the private sector. (2) The recommendations from external committee members are less distorted by strategic incentives for conformity than internal members. (3) The difference in ideological bias between external and internal members is statistically insignificant.

5 Concluding Remarks

We study the identification and estimation of structural models of committee decisions when the members have two-dimensional private types: how they process different sources of information (ideological bias); and how they weigh multiple objectives (tastes). The model is suitable for rationalizing the dissenting recommendations made by committee members, who share the same target and information.

Our models take account of strategic interactions between members due to concerns about conformity to the committee decision. We show how to nonparametrically recover the distributions of members' private types, using their choice patterns and the variation in common information. The identification arguments differ qualitatively for cases with and without strategic concerns for conformity. We apply our model to analyze monetary policy decisions by MPC at the Bank of England. Our analysis provides new insights about MPC's decision process. First, the ideological bias (how members process information from different sources) is important to explain the variation in recommendations, while the observed types of members (external or internal) are not. Second, internal members tend to have more concerns for conformity to the final committee decision than external members do. Finally, the difference in voting patterns of internal and external members can be explained better by difference in tastes for multiple objectives than by difference in ideological bias.

References

Aguirregabiria, V., and P. Mira (2013): "Identification of Games of Incomplete Information with Multiple Equilibria and Common Unobserved Heterogeneity," University of Toronto, Working Paper. 1

Ai, C., and X. Chen (2003): "Efficient estimation of models with conditional moment restrictions containing unknown functions," *Econometrica*, 71(6), 1795–1843. 2.3, 3.6

An, Y. (2010): "Nonparametric Identification and Estimation of Level-k Auctions," Manuscript, Johns Hopkins University. 3.3

An, Y., and X. Tang (2013): "Identifying Structural Models of Committee Decisions with Heterogeneous Tastes and Ideological Bias," . 3.4

Andrews, D. (2011): "Examples of L2-Complete and Boundedly-Complete Distributions," . B.4

Bajari, P., H. Hong, J. Krainer, and D. Nekipelov (2010): "Estimating static models of strategic interactions," *Journal of Business & Economic Statistics*, 28(4). 3.2

Batini, N., B. Jackson, and S. Nickell (2005): "An open-economy new Keynesian Phillips curve for the UK," *Journal of Monetary Economics*, 52(6), 1061–1071. 4.2

Bergemann, D., and S. Morris (2013): "Robust predictions in games with incomplete information," *Econometrica*, 81(4), 1251–1308. 3.1

Berk, J. M., B. K. Bierut, and E. E. Meade (2010): "The dynamic voting patterns of the Bank of England's MPC," . 4

Besley, T., N. Meads, and P. Surico (2008): "Insiders versus outsiders in monetary policymaking," *American Economic Review*, 98(2), 218–223. 3.1, 4, 4.4

Chen, X., and X. Shen (1998): "Sieve Extremum Estimates for Weakly Dependent Data," *Econometrica*, 66, 289–314. 2.3, 3.6

Costanzo, M., N. Shaked-Schroer, and K. Vinson (2010): "Juror Beliefs About Police Interrogations, False Confessions, and Expert Testimony," *Journal of Empirical Legal Studies*, 7(2), 231–247. 2.2

Degan, A., and A. Merlo (2009): "Do voters vote ideologically?," *Journal of Economic Theory*, 144(5), 1868–1894. 1

d'Haultfoeuille, X. (2011): "On the completeness condition in nonparametric instrumental problems," *Econometric Theory*, 27(03), 460–471. B.4

Fan, J., and Q. Yao (2005): *Nonlinear Time Series: Nonparametric and Parametric Methods*. Springer. 3.6

Galí, J. (2009): *Monetary Policy, inflation, and the Business Cycle: An introduction to the new Keynesian Framework*. Princeton University Press. 4.2

Galí, J., and M. Gertler (1999): "Inflation dynamics: A structural econometric analysis," *Journal of monetary Economics*, 44(2), 195–222. 4.2

Hall, P., and X.-H. Zhou (2003): "Nonparametric estimation of component distributions in a multivariate mixture," *Annals of Statistics*, 31(1), 201–224. 1

Hansen, S., M. McMahon, and C. V. Rivera (2014): "Preferences or private assessments on a monetary policy committee?," *Journal of Monetary Economics*, 67, 16–32. 1

Harris, M., P. Levine, and C. Spencer (2011): "A decade of dissent: explaining the dissent voting behavior of Bank of England MPC members," *Public Choice*, 146(3), 413–442. 4, 4.4

Henry, M., Y. Kitamura, and B. Salanié (2013): "Identifying Finite Mixtures in Econometric Models," *Quantitative Economics, Forthcoming*. 1

Hodrick, R., and E. Prescott (1980): "Post-War US Business Cycles: An Empirical Investigation," *Journal of Money, Credit and Banking*, 29(1), 1–16. 4.2

Hu, Y. (2008): "Identification and Estimation of Nonlinear Models with Misclassification Error Using Instrumental Variables: A General Solution," *Journal of Econometrics*, 144, 27–61. 1, 3.3, 3.3, 3.3

Hu, Y., and S. M. Schennach (2008): "Instrumental variable treatment of nonclassical measurement error models," *Econometrica*, 76(1), 195–216. 1

Hu, Y., and J.-L. Shiu (2012): "Nonparametric Identification using instrumental variables: sufficient conditions for completeness," Johns Hopkins Working Paper 581. B.4

Iaryczower, M., G. Lewis, and M. Shum (2013): "To elect or to appoint? Bias, information, and responsiveness of bureaucrats and politicians?," *Journal of Public Economics*, 97, 230–244. 1

Iaryczower, M., X. Shi, and M. Shum (2012): "Words Get in the Way: The Effect of Deliberation in Collective Decision-Making," Working paper: California Institute of Technology. 1

Iaryczower, M., and M. Shum (2012a): "Money in Judicial Politics: Individual Contributions and Collective Decisions," Princeton University. 1

——— (2012b): "The Value of Information in the Court: Get it Right, Keep it Tight," *American Economic Review*, 102(1), 202–237. 1

Kadane, J. B. (1983): "Juries hearing death penalty cases: Statistical analysis of a legal procedure," *Journal of the American Statistical Association*, 78(383), 544–552. 2.2

Kasahara, H., and K. Shimotsu (2009): "Nonparametric identification of finite mixture models of dynamic discrete choices," *Econometrica*, 77(1), 135–175. 1

Kawai, K., and Y. Watanabe (2013): "Inferring strategic voting," *American Economic Review*, 103(2), 624–662. 1

Levy, G. (2007): "Decision Making in Committees: Transparency, Reputation, and Voting Rules," *American Economic Review*, 97(1), 150–168. 3.1

Lewbel, A., and X. Tang (2013): "Identification and Estimation of Games with Incomplete Information Using Excluded Regressors," Working paper: University of Pennsylvania. 3.2

Mavroeidis, S., M. Plagborg-Møller, and J. H. Stock (2014): "Empirical evidence on inflation expectations in the New Keynesian Phillips Curve," *Journal of Economic Literature*, 52(1), 124–188. 4.2

Merlo, A. M., and A. de Paula (2013): "Identification and estimation of preference distributions when voters are ideological," Discussion paper, Centre for Microdata Methods and Practice, Institute for Fiscal Studies. 1

Robin, J.-M., and R. J. Smith (2000): "Tests of rank," *Econometric Theory*, 16(02), 151–175. 3.6

Shen, X. (1997): "On methods of sieves and penalization," *Annals of Statistics*, 25(6), 2555–2591. 2.3, 3.6

Wesche, K. (2003): "Monetary policy in Europe: Evidence from time-varying Taylor rules," Discussion paper, Bonn econ discussion papers. 3.1

Appendix

A Proofs for Section 2

This part of the appendix provides further details regarding the equilibrium in the model of expressive recommendation in Section 2, and the identification of the members' type distributions therein. First, we present the proof of the equilibrium characterization in Lemma 1.

Proof of Lemma 1. First, we characterize the best response of a member i with types (α_i, w_i) when the other members adopt generic strategies $\sigma_{-i} \equiv \{\sigma_j\}_{j \neq i}$, with σ_j mapping from the support of α_j, W_j to binary actions $\{0,1\}$. Suppress $\mathcal{I}$ from notation throughout the proof. For a member i, define:

$$p_{1,d_i}(\sigma_{-i}) \equiv \Pr(D^* = 1 \mid D_i = d_i; \sigma_{-i}) = \sum_{\{\mathbf{d}_{-i}: D^*(d_i, \mathbf{d}_{-i})=1\}} \left(\prod_{j \neq i} \Pr(D_j = d_j \mid \sigma_j) \right) \tag{A.1}$$

where $D^*(\mathbf{d}) \equiv \arg\max_{z \in \{0,1\}} \sum_i 1(d_i = z)$ denotes the majority rule. In a Bayesian Nash equilibrium, a member i with (α_i, w_i) seeks to pick d_i to minimize

$$\sum_k \left\{ w_{i,k} \left[\sum_q (y_k^q - \tilde{y}_k)^2 \left[\mathcal{F}_{q,1}(\alpha_i) p_{1,d_i}(\sigma_{-i}) + \mathcal{F}_{q,0}(\alpha_i)(1 - p_{1,d_i}(\sigma_{-i})) \right] \right] \right\}$$

where $\mathcal{F}_{q,1}(\alpha_i) \equiv \alpha_i H_{q,1} + (1-\alpha_i) G_{q,1}$ and $\mathcal{F}_{q,0}$ is defined likewise. (The dependence on s is suppressed in $G_{q,d}$ and $H_{q,d}$ for $d = 0, 1$.) By removing the terms that do not depend on the individual recommendation d_i, the optimization problem is equivalent to:

$$\arg\min_{d_i \in \{0,1\}} \sum\nolimits_k \{ w_{i,k}[\alpha_i \delta_{H,k} + (1-\alpha_i)\delta_{G,k}] \} \, p_{1,d_i}(\sigma_{-i})$$

where $\delta_{H,k}, \delta_{G,k}$ are defined as in the text. Thus member i chooses $d_i = 1$ if and only if

$$[p_{1,1}(\sigma_{-i}) - p_{1,0}(\sigma_{-i})] \sum\nolimits_k w_{i,k}[\alpha_i \delta_{H,k} + (1-\alpha_i)\delta_{G,k}] \leq 0.$$

Next, note that $\{\mathbf{d}_{-i} : D^*(1, \mathbf{d}_{-i}) = 1\}$ must be a strict subset of $\{\mathbf{d}_{-i} : D^*(0, \mathbf{d}_{-i}) = 1\}$ by construction. Therefore, it follows from the definition of $p_{1,d_i}(\sigma_{-i})$ in (A.1) that $p_{1,1} > p_{1,0}$ regardless of other members' strategies σ_{-i}. Thus in equilibrium, i chooses $d_i = 1$ if and only if $\sum_k w_{i,k}[\alpha_i \delta_{H,k} + (1-\alpha_i)\delta_{G,k}] \leq 0$ for all σ_{-i}. The uniqueness of Bayesian Nash equilibrium follows directly from this characterization. □

A.1 Identifying the order of indifference thresholds

As in the text, we focus on the model with $K = 2$, $J \equiv |\mathcal{A}| = 2$ and $Q \equiv |\mathcal{Y}| = 3$; and consider without loss of generality the scenario where Assumption 2 (a) holds with a positive sign. The rest of Appendix A present arguments conditioning on a triple $\{s, \tilde{y}, g(s)\}$ under Assumption 2. We suppress the dependence on the triple to simplify notations. Define

$$c(w; h) \equiv \left(\sum_k w_k \delta_{H,k}\right) / \left(\sum_k w_k \delta_{G,k}\right) \tag{A.2}$$

so that the indifference hyperplane for type (α, w) is

$$\{h : c(w; h) = (\alpha - 1)/\alpha\}.$$

For any pair of extreme evidence (h^a, h^b), we index the elements in the set of their convex combinations $\mathcal{H}(h^a, h^b)$ by their weights on h^b. That is, an $h \in \mathcal{H}(h^a, h^b)$ is indexed by

$$\lambda(h) \equiv \{\lambda : h = (1 - \lambda)h^a + \lambda h^b\}.$$

For any $\lambda \in [0, 1]$, let $h(\lambda)$ be a shorthand for the inverse mapping, that is $h(\lambda) \equiv \lambda h^b + (1 - \lambda)h^a$. For $j = 1, 2$, define $r^j \equiv (\alpha^j - 1)/\alpha^j$ (recall $\mathcal{A} \equiv \{\alpha^1, \alpha^2\}$) and

$$\underline{\lambda}_j \equiv \{\lambda : c(\underline{w}\ ,\ h(\lambda)) = r^j\} \quad \text{and} \quad \overline{\lambda}_j \equiv \{\lambda : c(\overline{w}\ ,\ h(\lambda)) = r^j\}.$$

That is, $\underline{\lambda}_j$ is the index for an $h \in \mathcal{H}(h^a, h^b)$ under which a member with $(\alpha^j, \underline{w})$ is indifferent between two alternatives and likewise for $\overline{\lambda}_j$. We refer to $\underline{\lambda}_j, \overline{\lambda}_j$ as "indifference thresholds" on $\mathcal{H}(h^a, h^b)$. The positions of extreme evidence relative to the hyperplanes are fully characterized by the order of these thresholds over $\mathcal{H}(h^a, h^b)$.

Lemma A1. Suppose Assumption 1 holds. For any $\{s, \tilde{y}, g(s)\}$ and $\{h^a, h^b\}$ under Assumption 2, the order of $\{\underline{\lambda}_j, \overline{\lambda}_j\}_{j=1,2}$ is identified.

Proof of Lemma A1. Fix the triple $\{s, \tilde{y}, G(s)\}$. For a pair (h^a, h^b) satisfying part (b) of Assumption 2, define the index for the infimum and supremum of the set of evidence that has non-degenerate CCPs:

$$\begin{aligned} \lambda_1 &\equiv \sup\{\lambda : \Pr\{D_i = 1 \mid h = (1 - \lambda)h^a + \lambda h^b\} = 0\}; \text{ and} \\ \lambda_4 &\equiv \inf\{\lambda : \Pr\{D_i = 1 \mid h = (1 - \lambda)h^a + \lambda h^b\} = 1\}. \end{aligned} \tag{A.3}$$

For example, in panel (i) of Figure 1, $\lambda_1 = \overline{\lambda}_2$ and $\lambda_4 = \underline{\lambda}_1$; on the other hand, in panel (ii) of Figure 1, $\lambda_1 = \underline{\lambda}_2$ and $\lambda_4 = \underline{\lambda}_1$. First off, the two cases (i) and (iv) can be distinguished from the other four cases (ii), (iii), (v) and (vi) as follows. In (i) and (iv), the CCPs must be strictly increasing in h over the set of convex combinations indexed between $[\lambda_1, \lambda_4]$ under Assumption 1. In contrast, in each of the other four cases, the CCPs must be invariant over a certain range of evidence over $\mathcal{H}(h^a, h^b)$.

For all six cases (i)-(vi), let λ_1, λ_4 be defined as in (A.3); for cases (ii), (iii), (v), (vi), let λ_2, λ_3 be the infimum and the supremum of a strict sub-interval in $[\lambda_1, \lambda_4]$ over which the CCP conditional on the convex combination of extreme evidence is constant. While $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ are all directly identified from the CCPs in (ii), (iii), (v) and (vi), the matching between them and the four indifference thresholds $\{\underline{\lambda}_j, \overline{\lambda}_j\}_{j=1,2}$ vary across these cases. For instance, in (ii), $\lambda_2 = \overline{\lambda}_2$ and $\lambda_3 = \overline{\lambda}_1$; in (iii), $\lambda_2 = \underline{\lambda}_2$ and $\lambda_3 = \overline{\lambda}_1$.

The unmatched thresholds as $\{\lambda_j : 1 \leq j \leq 4\}$ are functions of extreme evidence (h^a, h^b) conditioned on. Let $\delta_{j,j'}$ denote the distance between λ_j and $\lambda_{j'}$, which by construction is differentiable at (h^a, h^b) in both arguments due to Assumption 2. Let $\delta'_{j,j'}$ denote the partial derivative of $\delta_{j,j'}$ with respect to the second coordinate of h^a for $l \in \{a, b\}$. Recall that in panels (i), (iv), only λ_1, λ_4 are identified while in the other four panels all four indifference thresholds are identified. The following patterns are evident from Figure 1. In (i), $\delta'_{1,4} > 0$; in (ii), $\delta'_{2,3} = 0$ while $\delta'_{1,2} < 0$; in (iii), $\delta'_{2,3} < 0$; in (iv), $\delta'_{1,4} < 0$; in (v), $\delta'_{2,3} = 0$ while $\delta'_{1,2} > 0$; in (vi), $\delta'_{2,3} > 0$. Hence all six cases can be distinguished from each other using these identifiable partial derivatives. □

A.2 Recovering the type distributions

First, we prove the identification of F_α, F_W in the four cases in (ii), (iii), (v), (vi) in Proposition 1.

Proof of Proposition 1. Fix the triple $\{s, \tilde{y}, G(s)\}$ and the pair of evidence (h^a, h^b) under Assumption 2, 3. Consider the case (ii). In this case, $\underline{\lambda}_2$ is identified as $\lambda_1 \equiv \inf\{\lambda : \Pr\{D_i = 1 \mid h(\lambda)\} > 0\}$. By construction, $c(\underline{w}; h(\underline{\lambda}_2)) = r^2$ where c is defined in (A.2). With $\underline{w}$ and (h^a, h^b) known and $\underline{\lambda}_2$ identified, the equation allows us to solve for $\delta_{G,1} + \underline{w}\delta_{G,2}$, where $\delta_{G,k}$ is defined in (3). In addition, $\overline{\lambda}_2$ is identified as the infimum of the strict sub-interval of $[\lambda_1, \lambda_4]$ over which CCPs remain constant, which is denoted by λ_2. (Recall that both λ_1 and λ_4 are defined as in the proof of Lemma A1 above.) Likewise, this allows us to solve for $\delta_{G,1} + \overline{w}\delta_{G,2}$.

First, identify the probability mass function for α_i. By construction, $\Pr\{D_i = 1 \mid h\} = \sum_{j=1,2} \Pr\{c(W, h) \leq r^j\} \Pr\{R_i = r^j\}$ for all h. In case (ii), $\Pr\{c(W; h(\lambda_2)) \leq r^1\} = 0$ and $\Pr\{c(W; h(\lambda_2)) \leq r^2\} = 1$. Hence $\Pr\{\alpha_i = \alpha^2\}$, or equivalently $\Pr\{R_i = r^2\}$, is identified as $\Pr\{D_i = 1 \mid h(\lambda_2)\}$.

To identify the distribution F_{W_i} in case (ii), we need to first recover the initial perception g_q for $q = 1, 2$ at the state s in the conditioning set. The full-rank condition in Assumption 3 implies $\frac{a_{1,1}}{a_{2,1}} \neq \frac{a_{1,2}}{a_{2,2}}$ with both sides nonzero, which in turn implies $\frac{a_{1,1}+\underline{w}a_{1,2}}{a_{2,1}+\underline{w}a_{2,2}} \neq \frac{a_{1,1}+\overline{w}a_{1,2}}{a_{2,1}+\overline{w}a_{2,2}}$ and both sides are nonzero. With $\delta_{G,1} + w\delta_{G,2}$ identified for $w = \underline{w}, \overline{w}$ using arguments above, this inequality implies g_1 and g_2 are identified (at the state s conditioned on) as the unique solution to:

$$\begin{pmatrix} a_{1,1} + \underline{w}a_{1,2} & a_{2,1} + \underline{w}a_{2,2} \\ a_{1,1} + \overline{w}a_{1,2} & a_{2,1} + \overline{w}a_{2,2} \end{pmatrix} \begin{pmatrix} g_1 \\ g_2 \end{pmatrix} = \begin{pmatrix} \delta_{G,1} + \underline{w}\delta_{G,2} \\ \delta_{G,1} + \overline{w}\delta_{G,2} \end{pmatrix}.$$

As a result, the function $c(w; h)$ is identified for any given w, h.

Next, note for all $\lambda \in (\lambda_1, \lambda_2)$, we have:

$$\Pr\{c(W, h(\lambda)) \leq r^2\} = \frac{\Pr\{D_i = 1 | h(\lambda)\}}{\Pr\{\alpha_i = \alpha^2\}}. \tag{A.4}$$

Under Assumption 3, the sign of the derivative of c with respect to w at any $h(\lambda)$ with $\lambda \in [\lambda_1, \lambda_2]$ must equal the sign of

$$\frac{a_{1,1}h_1(\lambda) + a_{2,1}h_2(\lambda)}{a_{1,1}g_1 + a_{2,1}g_2} - \frac{a_{1,2}h_1(\lambda) + a_{2,2}h_2(\lambda)}{a_{1,2}g_1 + a_{2,2}g_2} \tag{A.5}$$

which does not depend on w. Note for any $\{s, \tilde{y}, G(s)\}$ and (h^a, h^b) conditioned on, this sign is non-zero for almost all λ over $[\lambda_1, \lambda_2]$.

Let $c^{-1}(r, h)$ denote the inverse of c at r given h. Then the left-hand side of (A.4) is either $\Pr\{W_i \leq c^{-1}(r^2, h(\lambda))\}$ or $\Pr\{W \geq c^{-1}(r^2, h(\lambda))\}$, depending on the sign of the difference in (A.5), which is known given $\{s, \tilde{y}, g(s)\}$ and (h^a, h^b). Also $c^{-1}(r^2, h(\lambda))$ is continuous in λ with $c^{-1}(r^2, h(\lambda_1)) = \underline{w}$ and $c^{-1}(r^2, h(\lambda_2)) = \overline{w}$ by construction. Therefore, F_{W_i} is identified almost everywhere over its support $\mathcal{W}$. Identification of F_{α_i} and F_{W_i} under the other three cases (iii), (v) and (vi) follows from symmetric arguments. $\square$

Next, we provide the details for identification under the other two cases (i) and (iv) in Figure 1 for the rest of this part of the appendix. Let's consider case (i); and let λ_1, λ_4 be defined as in the proof of Lemma A1. Knowing that $\lambda_4 = \underline{\lambda}_1$ is on the indifference hyperplane for $(\alpha^1, \underline{w})$ allows us to solve for $\delta_{G,1} + \underline{w}\delta_{G,2}$ at $\{s, \tilde{y}, g(s)\}$, using knowledge of α^1. Likewise, knowing $\lambda_1 = \overline{\lambda}_2$ and α^2 allows us to solve for $\delta_{G,1} + \overline{w}\delta_{G,2}$. These in turn allow one to solve for $\overline{\lambda}_1$ and $\underline{\lambda}_2$ using $c(\overline{w}; h(\overline{\lambda}_1)) = r^1$ and $c(\underline{w}; h(\underline{\lambda}_2)) = r^2$. Thus all indifference thresholds $\{\underline{\lambda}_j, \overline{\lambda}_j\}_{j=1,2}$ are identified. Let $\lambda_2 \equiv \overline{\lambda}_1$ and $\lambda_3 \equiv \underline{\lambda}_2$. The pair of evidence $h(\underline{\lambda}_2)$ and $h(\overline{\lambda}_1)$ are of particular importance for the identification question for the following reason: For any h outside the interval between these two, $\varphi(h) \equiv \Pr\{D_i = 1|h\}$ is a product of the marginal distribution of W and the probability mass function for α_i. For any h between these two, the choice pattern $\varphi(h)$ takes the form of a finite mixture.

To characterize the range of tastes that are involved in such a finite mixture, we introduce the following definition and notations. For $j = 1, 2$, define $\phi_j : \mathcal{W} \to [\lambda_1, \lambda_4]$ as

$$\phi_j(w) = \{\lambda \in [\lambda_1, \lambda_4] : c(w, h(\lambda)) = r^j\}$$

where c is defined in (A.2). In words, $\phi_j(\cdot)$ describes how far the empirical evidence needs to move to the direction of h^b in order for an individual with taste w and ideological bias α^j to become different between both alternatives. The image of $\phi_j(\cdot)$ is $[\lambda_1, \lambda_4]$ by construction. For $\lambda \in [\lambda_1, \lambda_4]$, the inverses of $\phi_j(\cdot)$ is defined as :

$$\phi_1^{-1}(\lambda) \equiv \begin{cases} c^{-1}(r^1, h(\lambda)) & \text{for } \lambda \in (\lambda_2, \lambda_4) \\ \overline{w} & \text{for } \lambda \in (\lambda_1, \lambda_2) \end{cases} ;$$

and

$$\phi_2^{-1}(\lambda) \equiv \begin{cases} c^{-1}(r^2, h(\lambda)) & \text{for } \lambda \in (\lambda_1, \lambda_3) \\ \underline{w} & \text{for } \lambda \in (\lambda_3, \lambda_4) \end{cases}.$$

In words, for any given $h(\lambda)$ and α^j, the function $\phi_j^{-1}(\cdot)$ returns a cutoff value in individual taste w beyond which a member with α^j would vote for $D_i = 1$. The reported cutoff is censored at the boundaries of support $\mathcal{W}$ by construction. Let $\underline{w}_0 \equiv \underline{w}$ and $\overline{w}_0 \equiv \overline{w}$ so that $\lambda_3 = \phi_2(\underline{w}_0)$ and $\lambda_2 = \phi_1(\overline{w}_0)$. Define $\underline{w}_1 \equiv \phi_1^{-1}(\lambda_3)$ and $\overline{w}_1 \equiv \phi_2^{-1}(\lambda_2)$. By construction, $\underline{w}_1 > \underline{w}_0$ while $\overline{w}_1 < \overline{w}_0$. Of course, the values for $\overline{w}_1$ and $\underline{w}_1$ depend on the triple $\{s, \tilde{y}, g(s)\}$ conditioned on and the extreme evidence $\{h^a, h^b\}$ considered. Both $\underline{w}_1$ and $\overline{w}_1$ have an intuitive economic interpretation. Recall that the function c under Assumption 3 is monotonic in w for almost all pairs of extreme evidence. Then for any evidence h between $h(\lambda_4)$ and $h(\lambda_3)$, an application of the law of total probability suggests $\varphi(h)$ equals

$$\Pr\{W \le w^*\} \Pr\{\alpha_i = \alpha^1\} + \Pr\{\alpha_i = \alpha^2\}$$

for some w^* located on $[\underline{w}_0, \underline{w}_1]$. Likewise, for any h between $h(\lambda_2)$ and $h(\lambda_1)$, $\varphi(h) \equiv \Pr(D_i = 1|h)$ equals:

$$\Pr\{W \le w'\} \Pr\{\alpha_i = \alpha^2\}$$

for some w' located on $[\overline{w}_1, \overline{w}_0]$. We show identification under the following condition, which is sufficient but not necessary.

Assumption 9 $\overline{w}_1 < \underline{w}_1$.

This is a joint restriction on the triple $\{s, \tilde{y}, g(s)\}$ and the pair of extreme evidence (h^a, h^b) in the conditioning set.[1] Essentially it requires that, as the evidence varies over the set of convex combinations $\mathcal{H}(h^a, h^b)$, the ranges of evidence that lead to non-degenerate CCPs are sufficiently apart for members with different types of bias α_i. In other words, the ranges of evidence $\mathcal{H}(h^a, h^b)$ covered by the two sets of indifference hyperplanes for α^1 and α^2 (i.e. $[\lambda_1, \lambda_3]$ and $[\lambda_2, \lambda_4]$ in panel (i) of Figure 1) must be sufficiently non-overlapping. With the indifference thresholds identified above, this condition is verifiable. It is also worth noting that for our identification method to apply, we only need the support of empirical evidence to contain one such pair.

Proposition 2 *Suppose Assumptions 1, 2, 3 and 9 hold. Then F_{α_i} and F_{W_i} are identified in the cases (i) and (iv) in Figure 1.*

Proof of Proposition 2. Suppose members' decisions can be rationalized as following dominant strategies under two sets of model primitives: $(F_\alpha, F_W) \neq (\tilde{F}_\alpha, \tilde{F}_W)$. That is,

[1] In an earlier version of the paper, An and Tang (2013), we extend the proof of Proposition 2 for the general case where Assumption 9 fails.

(F_α, F_W) are true parameters in the data-generating process and $(\tilde{F}_\alpha, \tilde{F}_W)$ are alternative parameters that are observationally equivalent to (F_α, F_W). To simplify notations, we use q and p to denote the probability that $\alpha_i = \alpha^2$ according to F_α and $\tilde{F}_\alpha$ respectively, and drop the subscripts W in the two CDFs F_W and $\tilde{F}_W$.

For the two sets of primitives to be observationally equivalent, it must be the case that:

$$\begin{aligned} qF(w) &= p\tilde{F}(w) \ \forall w \in [\overline{w}_1, \overline{w}]; \text{ and} \\ q + (1-q)F(w) &= p + (1-p)\tilde{F}(w), \ \forall w \in [\underline{w}, \underline{w}_1]. \end{aligned}$$

According to these restrictions, a necessary condition for $(q, F) \neq (p, \tilde{F})$ is $p \neq q$. Then the observational equivalence of (q, F) and $(p, \tilde{F})$ requires

$$\begin{aligned} &\tilde{F}(w) = \tfrac{q}{p}F(w), \ \ \forall w \in [\overline{w}_1, \overline{w}]; \text{ and} \\ &1 - \tilde{F}(w) = \tfrac{1-q}{1-p}[1 - F(w)], \ \ \forall w \in [\underline{w}, \underline{w}_1]. \end{aligned} \tag{A.6}$$

With $\overline{w}_1 < \underline{w}_1$, the open interval $(\overline{w}_1, \underline{w}_1)$ is non-empty and non-degenerate. It then follows from the two conditions in (A.6) that

$$\tfrac{q}{p}F(w) + \tfrac{1-q}{1-p}[1 - F(w)] = 1 \ \forall w \in (\overline{w}_1, \underline{w}_1),$$

or equivalently

$$p^2 - [q + F(w)]p + qF(w) = 0 \ \forall w \in (\overline{w}_1, \underline{w}_1),$$

The solutions require either $p = q$ or $p = F(w)$ for all w in $(\overline{w}_1, \underline{w}_1)$. If $p = q$, then the restrictions in (A.6) implies $F = \tilde{F}$ over $\mathcal{W}$. Suppose $p \neq q$, then $p = F(w)$. This contradicts the assumption that F_W is strictly monotonic on $\mathcal{W}$. Hence there exists no $(p, \tilde{F})$ that differs from (q, F) but is observationally equivalent to (q, F) at the same time. □

B Proofs for Section 3

B.1 Details in Section 3.1

Let $p_{d,d_i}(\sigma_{-i})$ denote the probability that the committee chooses d given i's decision d_i and others' strategies σ_{-i}. With $(\alpha_i, W_{i,l})$ independent across i, we have $p_{1,1}(\sigma_{-i}) > p_{1,0}(\sigma_{-i})$ regardless of σ_{-i} due to the simple majority rule. It follows from $p_{1,d_i} = 1 - p_{0,d_i}$ that i's best response to σ_{-i} is:

$$\sigma_i(\alpha_i, w_{i,l}; \sigma_{-i}) \equiv 1\left\{\alpha_i \sum_k w_{i,l,k}\delta_{H,k} + (1-\alpha_i)\sum_k w_{i,l,k}\delta_{G,k} \le \tfrac{p_{1,1}(\sigma_{-i}) - p_{0,0}(\sigma_{-i})}{p_{1,1}(\sigma_{-i}) - p_{1,0}(\sigma_{-i})}\right\}, \tag{B.1}$$

where the dependence of $\delta_{H,k}$ and $\delta_{G,k}$, which are defined in Section 2, on $\mathcal{I}_l$ is suppressed in notation. The left-hand side of the inequality in (B.1) is Δ_i, or i's ex ante difference between the losses under the two alternatives. To see why (B.1) defines i's best response to σ_{-i}, note that $p_{1,d_i}(\sigma_{-i}) = 1 - p_{0,d_i}(\sigma_{-i})$ by construction. Using this fact, one can write the part of the objective function in (7) that depends on d_i as:

$$p_{d_i,d_i}(\sigma_{-i}) - p_{1,d_i}(\sigma_{-i}) \sum_{k,q} w_{i,l,k} a_{q,k}(\tilde{y}_l)[\mathcal{F}_{l,q,1}(s;\alpha_i) - \mathcal{F}_{l,q,0}(s;\alpha_i)].$$

The form in (B.1) follows from solving the inequality that defines the best response.

Let $\phi(\cdot) \equiv \Pr\{\alpha_i \sum_k w_{i,l,k}\delta_{H,k} + (1-\alpha_i)\sum_k w_{i,l,k}\delta_{G,k} \leq \cdot \mid \mathcal{I}_l\}$ denote the distribution of Δ_i. As stated in the text, define a mapping $\varphi : (0,1) \to (-\infty, +\infty)$ such that $\varphi(\tau)$ equals the right-hand side of the inequality in (B.1) when all others $j \neq i$ follow the same pure strategy which leads to $\Pr(D_{j,l} = 1 \mid \sigma_j) = \tau$. The form of φ depends on the size of the committee. For instance, consider a symmetric p.s.BNE with $|\mathcal{C}| = I = 3$. The members follow the same pure strategy σ^* characterized by a subset of the support of private types $\omega^*(\sigma^*) \equiv \{(\alpha_i, w_i) \in \mathcal{A} \otimes \mathbb{R}_+^K : \sigma^*(\alpha_i, w_i) = 1\}$. Let $p(\omega^*) \equiv \Pr\{(\alpha_i, w_i) \in \omega^*\}$. Then with three members, $p_{1,1} \equiv 1 - (1-p(\omega^*))^2$; $p_{0,0} \equiv 1 - p(\omega^*)^2$ and $p_{1,0} = p(\omega^*)^2$. These imply $\varphi(p(\omega^*)) \equiv \frac{2p(\omega^*)-1}{2p(\omega^*)[1-p(\omega^*)]}$.

We now show that $\varphi(\tau) \to -\infty$ as $\tau \to 0$ and $\varphi(\tau) \to +\infty$ as $\tau \to 1$, regardless of the committee size. Consider the general case $I = 2n+1$. Then $p_{1,1} - p_{1,0} = \binom{2n}{n}\tau^n(1-\tau)^n$, which converges to 0 as τ approaches 0 or 1. On the other hand, the difference in the numerator is $p_{1,1} - p_{0,0} = p_{1,1} + p_{1,0} - 1$, which converges to 1 as $\tau \to 1$ and converges to -1 as $\tau \to 0$. This implies the limit pattern of φ as $\tau \to 0$ or 1.

Lemma B1. *Suppose Assumptions 4 and 5 hold. There exist symmetric p.s.BNE if there exists κ, κ' such that $\varphi \circ \phi(\kappa) > \kappa$ while $\varphi \circ \phi(\kappa') > \kappa'$.*

Proof of Lemma B1. A symmetric p.s.BNE exists if there exists a constant cutoff $\kappa^* \in \mathbb{R}$ such that $\varphi \circ \phi(\kappa^*) = \kappa^*$. Under maintained assumptions, the composite function $\varphi \circ \phi$ is continuous over $\mathbb{R}$. If there exists κ, κ' with $\varphi \circ \phi(\kappa) - \kappa < 0 < \varphi \circ \phi(\kappa') - \kappa'$, then by the Intermediate Value Theorem there exists κ^* between κ and κ' with $\varphi \circ \phi(\kappa^*) = \kappa^*$. □

As we mention in the text, a pair (κ, κ') exists if the support of private types α_i, W_i is bounded. Alternatively, such a pair exists if $\exists \eta > 0$ such that $d\varphi \circ \phi(\kappa)/d\kappa > 1 + \eta$ for all κ sufficiently large and small on $\mathbb{R}$. In this case, $\varphi \circ \phi(\kappa)$ eventually increases (decreases) faster than κ as κ take extremely large or small values. Hence $\varphi \circ \phi(\kappa) - \kappa$ eventually becomes positive (negative) for κ large (small) enough.

Furthermore, we conclude this subsection with yet a third example where primitive conditions on ϕ alone implies the existence of p.s.BNE when private types have unbounded support. Consider the case with $|\mathcal{C}| = 3$.

TC*(Tail Conditions).* $\lim_{\kappa \to -\infty} \phi'(\kappa)\phi(\kappa)^{-2} > 1$ *and* $\lim_{\kappa \to \infty} \phi'(\kappa)[1-\phi(\kappa)]^{-2} > 1$.

TC is a tail restriction conditional on the common information $\mathcal{I}_l$, which we suppress in notations. TC requires the rate of increase of $[\phi(\kappa)]^{-1}$ as κ decreases (and the rate of increase of $[1-\phi(\kappa)]^{-1}$ as κ increases) is unbounded when κ gets sufficiently small (and large respectively). This implies that as $\kappa \to \pm\infty$, the rate of changes in $\varphi \circ \phi(\kappa)$ (which is a monotonic function in κ over $(-\infty,+\infty)$) eventually exceeds one.

To see why this implies equilibrium existence with unbounded taste support, note $d(\varphi \circ \phi(\kappa) - \kappa)/d\kappa = \varphi'(\phi(\kappa))\phi'(\kappa) - 1$ where $\varphi'(\phi(\kappa))\phi'(\kappa)$ is positive for all $\kappa \in \mathbb{R}$. This is because $\varphi'(\phi) = \frac{2\phi^2-2\phi+1}{2\phi^2(1-\phi)^2} > 0$ with $\phi \in [0,1]$. Besides, $\varphi'(\phi) \equiv \frac{d}{d\phi}\varphi(\phi)$ is $O(\phi^{-2})$ as $\phi \to 0$, and is $O\left([1-\phi]^{-2}\right)$ as $\phi \to 1$. If ϕ' is bounded, this implies $\varphi'(\phi(\kappa))\phi'(\kappa)$ is $O(\phi'(\kappa)\phi(\kappa)^{-2})$ as $\kappa \to -\infty$ and is $O(\phi'(\kappa)[1-\phi(\kappa)]^{-2})$ as $\kappa \to +\infty$. Assumption TC ensures the rate of increase or decrease in $\varphi'(\phi(\kappa))\phi'(\kappa)$ must eventually exceeds that in κ (which is constantly 1 or -1) as κ approaches ∞ or $-\infty$. Thus it follows from continuity of φ and ϕ and the intermediate value theorem that solutions to the fixed point equation $\varphi \circ \phi(\kappa^*) = \kappa^*$ must exist.

B.2 Proof of Lemma 2

Proof of Lemma 2. Under Assumption 7 (a), the cardinality of the support of ideological bias J is identified as the rank of $L_{d_{l-1},\mathcal{J}_{l-1},d_{l-2},\mathcal{J}_{l-2}}$. It also implies there exists a coarser partition of $\mathcal{X}$ into $M = J$ intervals such that both $L_{d_{l-1}|\mathcal{J}_{l-1},\alpha}$ and $L_{d_{l-2},\alpha,\mathcal{J}_{l-2}}$ are non-singular square matrices. Hence we can write the inverse of (12) as

$$L^{-1}_{d_{l-1},\mathcal{J}_{l-1},d_{l-2},\mathcal{J}_{l-2}} = L^{-1}_{d_{l-2},\alpha,\mathcal{J}_{l-2}} L^{-1}_{d_{l-1}|\mathcal{J}_{l-1},\alpha}.$$

This together with (11) imply that

$$L_{d_l,\mathcal{J}_l,d_{l-1},x_{l-1},d_{l-2},\mathcal{I}_{l-2}} L^{-1}_{d_{l-1},\mathcal{J}_{l-1},d_{l-2},\mathcal{J}_{l-2}} = L_{d_l|\mathcal{J}_l,\alpha}\Lambda_{d_{l-1}|x_{l-1},\alpha} L^{-1}_{d_{l-1}|\mathcal{J}_{l-1},\alpha}, \tag{B.2}$$

where the same partition are used in defining $\mathcal{J}_l$ and $\mathcal{J}_{l-1}$. Because of the time-homogeneity of CCPs, $L_{d_{l-1}|\mathcal{J}_{l-1},\alpha} = L_{d_l|\mathcal{J}_l,\alpha}$ for $d_l = d_{l-1}$. Thus the right-hand side of the display above can be written as $L_{d_l|\mathcal{J}_l,\alpha}\Lambda_{d_{l-1}|x_{l-1},\alpha}L^{-1}_{d_l|\mathcal{J}_l,\alpha}$. Without loss of generality, choose $d_l = d_{l-1} = 1$. Under Assumption 7 (a)-(b), for any given $x^{(j)}$ the eigen-decomposition of the left-hand side in (B.2) uniquely determines J eigenvalues $\Pr(D_{l-1} = 1 \mid \mathcal{I}_l = x^{(j)}, \alpha = \alpha^k), k = 1,2,\cdots,J$ but only up to unknown matches with α.

The next step is to recover the probability mass function for α_i, and the CCPs conditional on all elements in $\mathcal{X}$. For any given d_{l-1}, d_{l-2} and x_{l-2}, (10) provides a linear system with J equations:

$$\frac{\Pr(d_{l-1},\mathcal{I}_{l-1} = x^{(j)}, d_{l-2}, x_{l-2})}{\Pr(\mathcal{I}_{l-1} = x^{(j)} \mid d_{l-2}, x_{l-2})} = \sum_{k=1}^{J} \Pr(d_{l-1} \mid \mathcal{I}_{l-1} = x^{(j)}, \alpha = \alpha^k) \Pr(d_{l-2}, x_{l-2}, \alpha = \alpha^k), \tag{B.3}$$

for $j = 1,2,...,J$ where the J unknowns are $\Pr(d_{l-2}, x_{l-2}, \alpha = \alpha^k), k = 1,2,\cdots,J$. Under Assumption 7 (b), the unknowns are uniquely solved from the linear system. Thus the

CCP $\Pr(d_{l-2} \mid x_{l-2}, \alpha)$ is identified as the ratio $\Pr(d_{l-2}, x_{l-2}, \alpha)/\sum_{d_{l-2}\in\{0,1\}} \Pr(d_{l-2}, x_{l-2}, \alpha)$. By solving the linear system for each given x_{l-2} and taking the ratio, $\Pr(d_{l-2} \mid \mathcal{I}_{l-2} = x_{l-2}, \alpha)$ is identified for all $x_{l-2} \in \mathcal{X}$. Integrating out both $d_{l-1} \in \{0,1\}$ and $x_{l-2} \in \mathcal{X}$ from the solution of the linear system, $\Pr(d_{l-2}, x_{l-2}, \alpha = \alpha^k)$, we obtain the probability mass function for α at all α^k, $k = 1, 2, \cdots, J$.

Alternatively, one can first identify $\Pr(d_{l-1} \mid x_{l-1}, \alpha)$ by conducting the eigen-decomposition in (B.2) for a given d_{l-1} and all the $x_{l-1} \in \mathcal{X}$. Then we apply (B.3) to identify the probability mass function. This approach requires more restrictive conditions on the conditional choice probabilities than Assumption 7 (b), i.e., for all $x \in \mathcal{X}$, $\Pr(d_{l-1} \mid \mathcal{I}_{l-1} = x, \alpha) > 0$ for all α; $\Pr(d_{l-1} \mid \mathcal{I}_{l-1} = x, \alpha) \neq \Pr(d_{l-1} \mid \mathcal{I}_{l-1} = x, \alpha')$ for any $\alpha \neq \alpha'$ in $\mathcal{A}$. □

B.3 Proof of Lemma 3

We prove the lemma for the case with $K = 2$. Generalization to the other cases with $K \geq 3$ requires more tedious algebra, but does not pose additional challenge in principle. Let $A^H_{d_i,k}$ be i's ex ante deviation from the k-th dimension of the target conditional on his choice d_i and the empirical evidence H. That is, $A^H_{d_i,k} \equiv \sum_q a_{q,k}(\tilde{y}) \left(\sum_{d\in\{0,1\}} H_{q,d} p^*_{d,d_i}(x)\right)$, where $p^*_{d,d_i}(x)$ is the probability that $D^* = d$ when $D_i = d_i$ and $\mathcal{I} \equiv x$. Define $A^G_{d_i,k}$ similarly. Note p^*_{d,d_i} is directly identifiable from data in equilibrium. Under our assumptions, α_i, W_i and $\mathcal{I}_l$ are mutually independent, and type-specific CCPs in equilibrium are:

$$\Pr(D_i = 1 \mid \mathcal{I} = x; \alpha_i) = \Pr\left\{\sum_k W_{i,k} C_k(x; \alpha_i) \leq t(x)\right\} \tag{B.4}$$

where $C_k(x; \alpha_i) \equiv \alpha_i \left(A^H_{1,k} - A^H_{0,k}\right) + (1-\alpha_i)\left(A^G_{1,k} - A^G_{0,k}\right)$ for $k = 1, 2$; and $t(x) \equiv p^*_{1,1}(x) - p^*_{0,0}(x)$. That is, for a member with bias α_i, $C_k(x; \alpha_i)$ is the difference in ex ante deviation from targets in the k-th dimension between the two alternatives.

Proof of Lemma 3. For any given x, a member i chooses 1 if and only if $\sum_k w_{i,k} C_k(x; \alpha_i) \leq t(x)$, where C_k is defined as above. Figure 2 visualizes this event on the $(W_{i,1}, W_{i,2})$-plane and the five panels are all the cases allowed by the non-degeneracy condition Assumption 7 (b) (we suppress the subscripts l for episodes in the proof and let $W_{i,k}$ denote the k-th coordinate in W_i.) Let $\mathcal{B} \equiv (0, \infty) \times (0, \infty)$. Then

$$\Pr(D_i = d_i | \mathcal{I} = x; \alpha_i) = \iint_{\mathcal{B}} \mathbb{I}(uC_1(X; \alpha_i) + vC_2(X; \alpha_i) \leq t(X)) f_{W_{i,1}, W_{i,2}}(u, v) du dv.$$

Since W_i is independent from X, the ordering of the type-specific CCPs w.r.t. α_i is determined by the relative positions of lines $\sum_k w_{i,k} C_k(x; \alpha_i) = t(x)$ on the $(W_{i,1}, W_{i,2})$-plane. Without loss of generality, let the supremum and the infimum of the support of α_i be 1 and 0 respectively.

The three lines in Figure 2 denote $\sum_k w_{i,k} C_k(x; 1) = t(x)$ (line A) and $\sum_k w_{i,k} C_k(x; 0) = t(x)$ (line B) and a generic $\sum_k w_{i,k} C_k(x; \alpha_i) = t(x)$ (line C) respectively. The event "$D_i = 1$

given α_i and $\mathcal{I}_l$" is shown in the figure as a half-space defined by the line $\sum_k w_{i,k} C_k(x;\alpha_i) = t(x)$.

The slope of the line associated with a generic α_i, i.e. $-C_1(x;\alpha_i)/C_2(x;\alpha_i)$, is either monotonic or invariant in $\alpha_i \in (0,1)$. To see this, write $C_k(x;\alpha_i)$ as:

$$C_k(x;\alpha_i) = \alpha_i C_k(x;1) + (1-\alpha_i) C_k(x;0) \text{ for } k = 1,2,$$

where $C_k(x;1) \equiv A^H_{1,k} - A^H_{0,k}$ and $C_k(x;0) \equiv A^G_{1,k} - A^G_{0,k}$ (with $A^H_{d_i,k}, A^G_{d_i,k}$ defined as in the text). Taking derivative of $-C_1(x;\alpha_i)/C_2(x;\alpha_i)$ w.r.t. α_i, we get

$$-\frac{d}{d\alpha_i}\left(\frac{C_1(x;\alpha_i)}{C_2(x;\alpha_i)}\right) = \frac{C_1(x;1)C_2(x;0) - C_2(x;1)C_1(x;0)}{[\alpha_i C_2(x;1) + (1-\alpha_i)C_2(x;0)]^2},$$

where the numerator does not depend on α_i and is known, and the denominator is positive by construction. Hence a line associate with a generic α_i necessarily lies between the lines A and B and their slopes are either monotonic or invariant in α_i. Also note that by construction, if lines A and B intersect at $(w^*_{i,1}, w^*_{i,2})$, then line C must also intersect with them at the same point. In the following, we prove the lemma under two scenarios above, depending on whether the lines A and B intersect or not.

Case 1. $C_1(x;1)C_2(x;0) = C_2(x;1)C_1(x;0)$. Thus the lines A, B, and C are parallel. Then the position of $\sum_k w_{i,k} C_k(x;\alpha_i) = t(x)$ can be attained by the intercepts on $W_{i,1}$-axis or $W_{i,2}$-axis, being $t(x)/C_1(x;\alpha_i)$ and $t(x)/C_2(x;\alpha_i)$, respectively, which are both monotonic in α_i. For instance, suppose $C_k(x;\alpha_i) > 0$ for $k = 1,2$, the intercept on $W_{i,2}$-axis is $t(x)/[\alpha_i C_1(x;1) + (1-\alpha_i)C_1(x;0)]$, which is monotonic in α_i. The monotonicity violates when $C_1(X;1) = C_1(X;0)$, however in this case we will have $C_2(X;1) = C_2(X;0)$, too. Consequently, all the members behave exactly the same regardless of their ideological bias. This would be ruled out by Assumption 8(b) already. Such a proof is generic no matter whether the slope of the parallel lines is positive (panel (i)) or negative (panel (ii)). In the former case with positive slopes, the proof is straightforward since the parallel lines always intersect with the first quadrant and the intercept on $W_{i,1}$-axis or $W_{i,2}$-axis must be positive, which can be used to rank the lines.

In the latter case of negative slopes, Assumption 7 (b) guarantees that there is at most one realization of $\alpha_i \in [0,1]$ such that the corresponding line has no intersection with the first quadrant since such a line corresponds to the fact that this type of decision makers choose alternative $d = 1$ with probability zero. According to the monotonicity of the slope in α_i, this type of decision makers must have $\alpha_i = 0$ or $\alpha_i = 1$. All the other lines can be ordered by the intercepts again.

Case 2. $C_1(x;1)C_2(x;0) \neq C_2(x;1)C_1(x;0)$. Thus lines A and B intersect at $(w^*_{i,1}, w^*_{i,2})$, and so does line C. Assumption 8 (b) restricts the intersection point $(w^*_{i,1}, w^*_{i,1}) \notin (0,\infty) \times (0,\infty)$. Otherwise there will be some members with tastes $(w^*_{i,1}, w^*_{i,2})$ who are always indifferent between the two alternatives regardless of their ideological bias α_i for the given $\mathcal{I}_l$. This case can be further divided into three subcases as illustrated by panel (iii), (iv), and (v) in Figure 2. In panel (iii), all the lines have positive slopes and they interact with the first quadrant for sure and this permits us to employ the intercept arguments as in Case 1 again to order the lines. As for panel (iv), Assumption 7(b) implies that the intersection point can only be in the second or the fourth quadrant. Otherwise if it is in the third quadrant, there will be more than two types of members whose choice is always $d = 0$ as we argued in Case 1, i.e., more than two lines with negative slope and do not pass the first quadrant. A similar proof can be applied to the subcase depicted in panel (v) where the intersection point can only be in the second quadrant due to Assumptions 7(b) and 8(b). In all three cases, the argument of intercepts which depend on $C_k(x;h), k = 1,2, h = 0,1$ still holds and this provides an ordering of all the lines. Consequently the choice probabilities $\Pr(D_i = d_i|\mathcal{I} = x;\alpha_i)$ are completely ordered according to α_i. □

B.4 Details in Section 3.5

As in Appendix B.3, we focus on the case with $K = 2$ to simplify exposition. Generalization to the other cases with $K \geq 3$ does not pose any additional challenge. Some definitions and assumptions are necessary for presenting the proof of the identification result in Section 3.5.

Definition 2 *Let $\mathcal{H} \equiv L^2(\mathcal{W}_0)$ be a Hilbert space. For some $\alpha \in \mathcal{A}$, the function $K(\gamma, x;\alpha)$ is L^2-complete in $\mathcal{H}$ if for all $x \in \mathcal{X}$, $\int_{\mathcal{W}_0} (K(w,x;\alpha))^2 dw < \infty$ and for any function $\delta \in \mathcal{H}$, $\int_{\mathcal{W}_0} K(w,x;\alpha)\delta(w)dw = 0$ for all $x \in \mathcal{X}$ implies $\delta = 0$ almost surely in $\mathcal{W}_0$.*

The following assumption imposes an additional parametric structure on committee members' taste.

Assumption 10 *(a) $W_{i,1} = W_{i,0} + \eta_{i,1}$ and $W_{i,2} = W_{i,0} + \eta_{i,2}$, where $W_{i,0}$ is continuously distributed over $\mathcal{W}_0 \subset \mathbb{R}_+$ and is independent from α_i and $\mathcal{I}$; and $\eta_{i,k} \in \mathbb{R}^1$ is continuously distributed over support $[\underline{\eta}_k, \overline{\eta}_k]$ for $k = 1,2$ with a known p.d.f. f_{η_k}. (b) For some $\alpha \in \mathcal{A}$, the function $K(\cdot,\cdot;\alpha)$ is L^2-complete.*

Part (a) of Assumption 10 allows members' private tastes on different dimensions to be correlated. Part (b) requires that, at least for some α_i, a L^2-completeness condition is satisfied by $K(\cdot,\cdot;\alpha_i)$. It is a joint restriction on the densities of $\eta_{i,1}$ and $\eta_{i,2}$ as well as model elements involved in $t(x)$, $C_2(x;\alpha_i)$ and $C_1(x;\alpha_i)$.

Completeness is a commonly imposed condition in nonparametric identification of structural models, e.g., see Andrews (2011). Recently, d'Haultfoeuille (2011) provides sufficient conditions for bounded completeness for random vectors. Hu and Shiu (2012) provide some sufficient conditions for L^2-completeness, and the basic idea is that if a function is close to a sequence with its limit being complete then the function is complete. The sufficient conditions in Hu and Shiu (2012) focus on a family of conditional densities. Unfortunately, all these existing sufficient conditions do not apply to the completeness of $K(\cdot,\cdot;\alpha_i)$.

Proposition B1. *Suppose Assumptions 4, 5, 6, 7, 8, and 10 hold. The distribution of* $W_{i,0}$ *is identified.*

Proof of Proposition B1. Under Assumption 10, identification of the joint distribution $F_{W_1,W_2}(\cdot)$ is equivalent to that of the distribution $F_{W_0}(\cdot)$. Recall

$$\Pr(D_i = 1|\mathcal{I} = x;\alpha_i) = \Pr\left\{\sum_k W_{i,k}C_k(x;\alpha_i) \le t(x)\right\}.$$

By the law of total probability and our specification of $W_{i,k}$,

$$\begin{aligned}\Pr(D_i = 1|\mathcal{I} = x;\alpha_i) &= \Pr\left\{\sum_k (W_{i,0}+\eta_{i,k})C_k(x;\alpha_i) \le t(x)\right\} \\ &= \int_{\underline{\eta}_2}^{\overline{\eta}_2}\left[\int_{\underline{\eta}_1}^{\overline{\eta}_1} F_{W_0}\left(\tfrac{t(x)-C_1(x;\alpha_i)u-C_2(x;\alpha_i)v}{C_1(x;\alpha_i)+C_2(x;\alpha_i)}\right) f_{\eta_1}(u)du\right] f_{\eta_2}(v)dv\end{aligned}$$

Denote the choice probability $\Pr(D_i = 1|\mathcal{I} = x;\alpha_i)$ by $\Phi(x;\alpha_i)$. The variation of x provides a linear integral equation of our identification objective $F_{W_0}(\cdot)$. With the support of α_i assumed known, the functions $C_k(x;\alpha_i)$ are also known. W.L.O.G., we consider the case where $C_1(x;\alpha_i) + C_2(x;\alpha_i) > 0$ and $C_1(x;\alpha_i) > 0$. (The other cases can be analyzed by symmetric arguments.) In this case,

$$\begin{aligned}&\Pr(D_i = 1 \mid \mathcal{I} = x;\alpha_i) \\ =\ & \int_{\underline{\eta}_2}^{\overline{\eta}_2}\left[\int_{\underline{\eta}_1}^{\overline{\eta}_1} F_{W_0}\left(\tfrac{t(x)-C_1(x;\alpha_i)u-C_2(x;\alpha_i)v}{C_1(x;\alpha_i)+C_2(x;\alpha_i)}\right) f_{\eta_1}(u)du\right] f_{\eta_2}(v)dv \\ =\ & \int_{\underline{\eta}_2}^{\overline{\eta}_2}\left[\int_{-\infty}^{\infty} \mathbb{I}(\gamma,v)F_{W_0}(\gamma)\, f_{\eta_1}(\Lambda(x,v,\gamma;\alpha_i))\left(-\tfrac{C_1(x;\alpha_i)+C_2(x;\alpha_i)}{C_1(x;\alpha_i)}\right)d\gamma\right] f_{\eta_2}(v)dv \\ =\ & -\left(\tfrac{C_1(x;\alpha_i)+C_2(x;\alpha_i)}{C_1(x;\alpha_i)}\right)\int_{\underline{\eta}_2}^{\overline{\eta}_2}\left[\int_{-\infty}^{\infty} F_{W_0}(\gamma)\,\mathbb{I}(\gamma,v) f_{\eta_1}(\Lambda(x,v,\gamma;\alpha_i))\,d\gamma\right] f_{\eta_2}(v)dv,\end{aligned}$$

where the second equality is due to change of variables, and the indicator function $\mathbb{I}(\cdot,\cdot)$ and $\Lambda(x,v,\gamma;\alpha_i)$ are defined as

$$\begin{aligned}\mathbb{I}(\gamma,v) &\equiv \mathbb{I}\left\{\tfrac{t(x)-C_2(x;\alpha_i)v-c_1\overline{\eta}_1}{C_1(x;\alpha_i)+C_2(x;\alpha_i)} \le \gamma \le \tfrac{t(x)-C_2(x;\alpha_i)v-c_1\underline{\eta}_1}{C_1(x;\alpha_i)+C_2(x;\alpha_i)}\right\}, \\ \Lambda(x,v,\gamma;\alpha_i) &\equiv \frac{t(x)-C_2(x;\alpha_i)v-C_1(x;\alpha_i)\gamma-C_2(x;\alpha_i)\gamma}{C_1(x;\alpha_i)}.\end{aligned}$$

Applying Fubini's theorem to the R.H.S. of the equation above,

$$\begin{aligned}\Pr(D_i = 1 \mid \mathcal{I} = x;\alpha_i) &= -\left(\tfrac{C_1(x;\alpha_i)+C_2(x;\alpha_i)}{C_1(x;\alpha_i)}\right) \\ &\quad \times \int_{-\infty}^{\infty} F_{W_0}(\gamma)\left[\int_{\underline{\eta}_2}^{\overline{\eta}_2} \mathbb{I}(\gamma,v) f_{\eta_1}(\Lambda(x,v,\gamma;\alpha_i))\, f_{\eta_2}(v)dv\right]d\gamma.\end{aligned}$$

Due to the non-negativity of W_0, the integral equation above can be succinctly summarized as:

$$\varphi(x;\alpha_i) = \int_0^\infty F_{W_0}(\gamma)\, K(\gamma, x;\alpha_i) d\gamma$$

where

$$\begin{aligned}\varphi(x;\alpha_i) &\equiv -\frac{\Pr(D_i = 1 \mid \mathcal{I} = x;\alpha_i) C_1(x;\alpha_i)}{C_1(x;\alpha_i) + C_2(x;\alpha_i)}, \\ K(\gamma, x;\alpha_i) &\equiv \int_{\underline{\eta}_2}^{\overline{\eta}_2} \mathbb{I}(\gamma, v) f_{\eta_1}(\Lambda(x, v, \gamma;\alpha_i))\, f_{\eta_2}(v) dv.\end{aligned}$$

Alternatively, the integral equation that link the CCPs and the unknowns can be written as

$$\Pr(D_i = 1 \mid \mathcal{I} = x;\alpha_i) = \int_0^\infty F_{W_0}(s)\, \mathcal{K}(s, x;\alpha_i) ds,$$

where

$$\mathcal{K}(s, x;\alpha_i) \equiv -\tfrac{C_1(x;\alpha_i)+C_2(x;\alpha_i)}{C_1(x;\alpha_i)} K(s, x, \alpha_i)$$

with $C_k(\cdot)$ and $t(\cdot)$ defined as in the first paragraph of Appendix B.3.

If L^2-completeness of $K(\gamma, x;\alpha_i)$ holds for some α_i, then it guarantees that F_{W_0} is the unique solution to the the integral equation (14) indexed by that α_i. □

C Details in Section 4.3

This section provides some further details about the simulated Maximum Likelihood estimation of θ in Section 4.3. Let

$$\widehat{\mathcal{K}}(\gamma, x;\alpha^j) \equiv -\tfrac{\widehat{C}_1(x;\alpha_i)+\widehat{C}_2(x;\alpha_i)}{\widehat{C}_1(x;\alpha_i)} \widehat{K}(\gamma, x;\alpha^j)$$

be the sample analog estimator for $\mathcal{K}$ defined above, where

$$\widehat{K}(\gamma, x;\alpha_i) \equiv \int_{\underline{\eta}_2}^{\overline{\eta}_2} \widehat{\mathbb{I}}(\gamma, v) f_{\eta_1}\left(\widehat{\Lambda}(x, v, \gamma;\alpha_i)\right) f_{\eta_2}(v) dv,$$

with $\widehat{\mathbb{I}}(\gamma, v)$ and $\widehat{\Lambda}(x, v, \gamma;\alpha_i)$ being estimates of $\mathbb{I}(\gamma, v)$ and $\Lambda(x, v, \gamma;\alpha_i)$ respectively by plugging in $\widehat{C}_k(x, \alpha_i)$ and $\hat{t}(x)$.

The simulation-based estimator for $\widehat{\Phi}_{j,e_i}$ in the log-likelihood is defined as:

$$\widehat{\Phi}_{j,e_i}(x_t;\theta) \equiv M^{-1} \textstyle\sum_{m=1}^M F_{W_0}(\gamma_m; a^{e_i}, b^{e_i})\, \widehat{\mathcal{K}}(\gamma_m, x_t;\alpha^j)/g(\gamma_m),$$

where $\gamma_m, 1 \leq m \leq M$ are random draws from a chosen density $g(\cdot)$ over the support $[0, \infty)$. In practice, we propose to use moderately positively skewed densities such as *Gamma*$(5,1)$, *Gamma*$(2,2)$ or *Gamma*$(7.5,1)$.

Figure 1: Illustration of the model with expressive recommendations

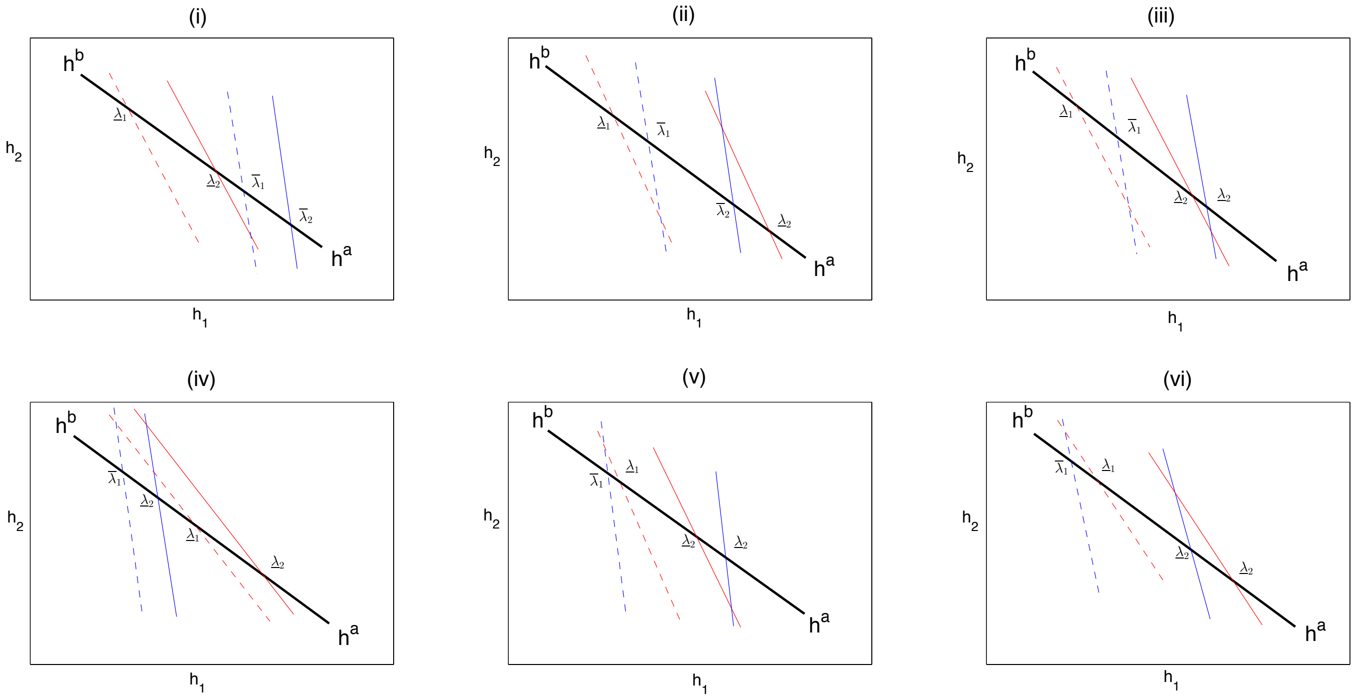

Table 1: Summary of votes

	Voted for higher rate	Voted for lower rate	Total
All members (31)	500	1108	1608
Internal members (12)	312	587	899
External members (19)	188	521	709

Table 2: Summary Statistics of Forecasts

	Forecasts	Outsiders (probabilities)				MPC (probabilities)			
		mean	std.	min	max	mean	std.	min	max
GDP	<1%	0.173	0.146	0.01	0.81	0.140	0.152	0	0.25
	[1%, 2%)	0.276	0.086	0.06	0.67	0.218	0.095	0	0.6
	[2%, 3%)	0.364	0.110	0.01	0.59	0.336	0.157	0.01	0.98
	≥ 3%	0.187	0.094	0.01	0.47	0.306	0.176	0	0.88
Inflation	<1.5%	0.176	0.119	0.01	0.64	0.205	0.189	0	0.97
	[1.5%, 2.5%)	0.538	0.103	0.29	0.83	0.449	0.189	0.03	0.97
	[2.5%, 3.5%)	0.226	0.129	0.02	0.69	0.233	0.167	0	0.82
	> 3.5%	0.060	0.040	0	0.25	0.113	0.132	0	0.54

Figure 2: Illustration of the model with strategic recommendations

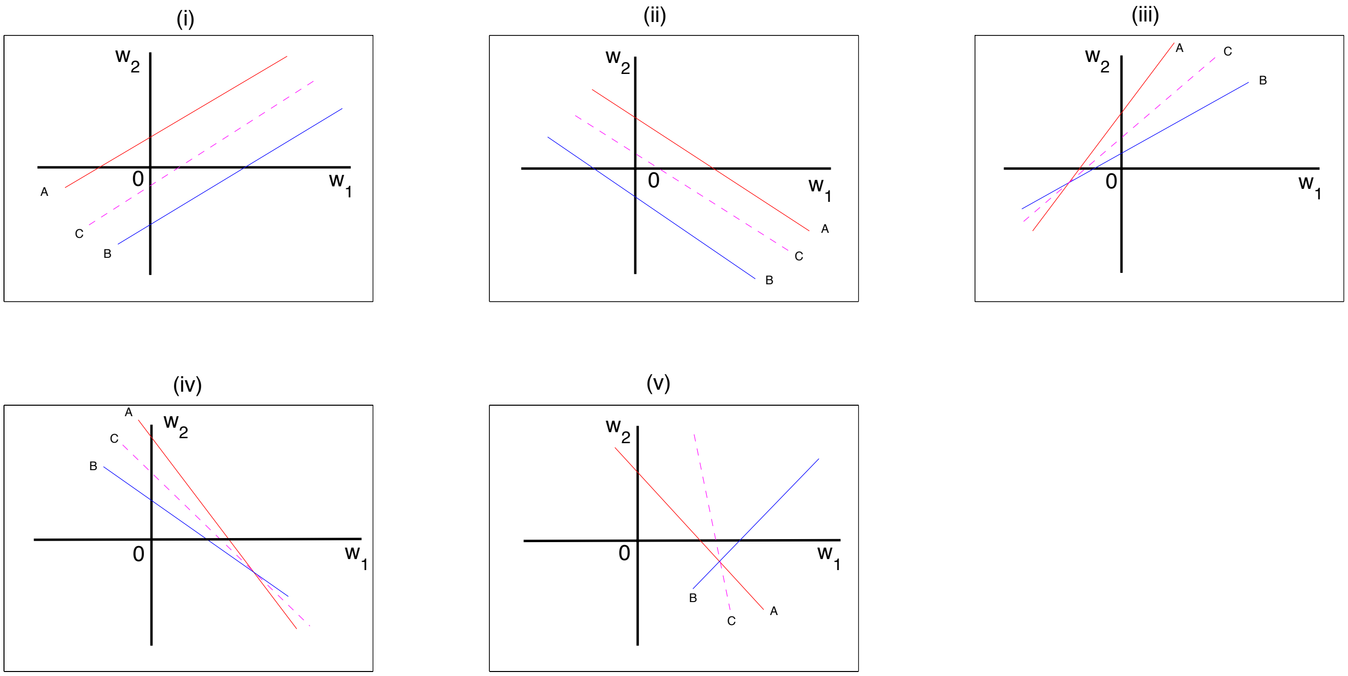

Figure 3: Estimated Distribution for the Tastes of External and Internal Members

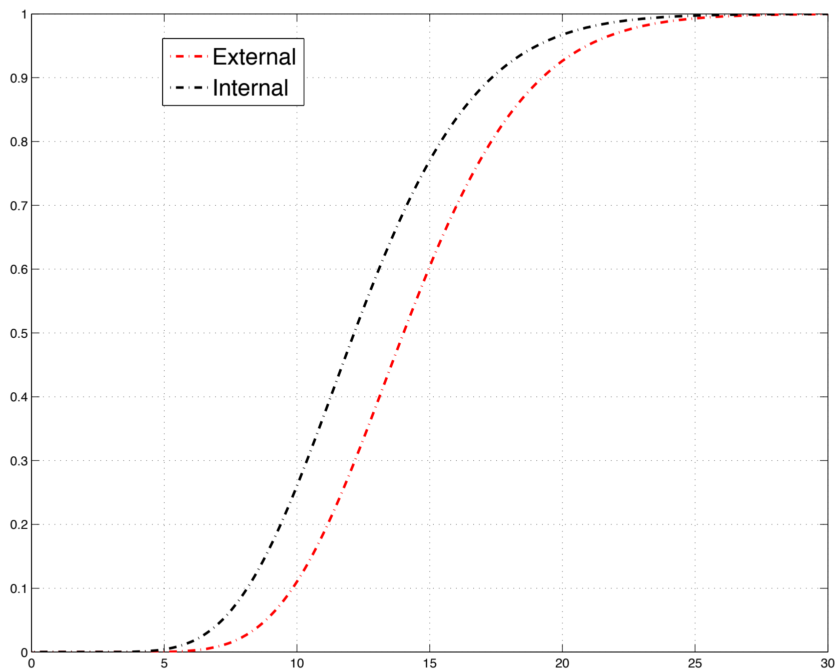

Table 3: Estimates of individual choice probabilities

External members		Internal members	
parameters	estimate	parameters	estimate
ξ_0	-0.899^{**}	ϑ_0	-0.525^{**}
	(0.092)		(0.079)
ξ_1 (Inflation-MPC)	0.268^{*}	ϑ_1 (Inflation-MPC)	0.446^{**}
	(0.106)		(0.094)
ξ_2 (GDP-MPC)	1.134^{**}	ϑ_2 (GDP-MPC)	1.233^{**}
	(0.364)		(0.318)
ξ_3 (Inflation-outsiders)	-0.424^{*}	ϑ_3 (Inflation-outsiders)	-0.384^{**}
	(0.174)		(0.144)
ξ_4 (GDP-outsiders)	-0.250	ϑ_4 (GDP-outsiders)	0.536
	(0.781)		(0.671)

* significant at 2.5 % level, ** significant at 1 % level.

Table 4: Estimates for the Distribution of Ideological Bias

	External members		Internal members	
	estimate	std.	estimate	std.
$\Pr(\alpha = 25\%)$	0.349^{***}	0.098	0.329^{***}	0.090
$\Pr(\alpha = 50\%)$	0.100	0.089	0.058	0.106
$\Pr(\alpha = 75\%)$	0.551^{***}	0.099	0.613^{***}	0.104

*** significant at 1 % level.

Table 5: Estimates for the Distribution of Individual Tastes[a]

	External members				Internal members			
	estimate	20-th pctl.	median	80-th pctl.	estimate	20-th pctl.	median	80-th pctl.
a	14.961	3.077	13.848	21.331	11.834	1.713	12.248	13.563
b	0.958	0.976	1.004	7.448	1.064	0.995	1.011	10.351

[a] This table summarizes the estimate and empirical distribution of our estimates for (a, b) using $B = 200$ bootstrap samples.