Penn Institute for Economic Research
Department of Economics
University of Pennsylvania
3718 Locust Walk
Philadelphia, PA 19104-6297
pier@econ.upenn.edu
http://economics.sas.upenn.edu/pier

PIER Working Paper 13-058

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

by

Yonghong An and Xun Tang

http://ssrn.com/abstract=2340104

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

Yonghong AN
University of Connecticut

Xun TANG
University of Pennsylvania

October 7, 2013

Abstract

We study the nonparametric identification and estimation of a structural model for committee decisions. Members of a committee share a common information set, but differ in ideological bias while processing multiple information sources and in individual tastes while weighing multiple objectives. We consider two cases of the model where committee members have or don't have strategic incentives for making recommendations that conform with the committee decision. For both cases, pure-strategy Bayesian Nash equilibria exist, and we show how to use variations in the common information set to recover the distribution of members' private types from individual recommendation patterns. 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 recommendations from external committee members are less distorted by strategic incentives than internal members. There is also evidence that MPC members differ more in their tastes for multiple objectives than in ideological bias.

Keywords: Committee decisions, nonparametric identification, MPC at the Bank of England
JEL: C14, D71

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 various factors in the decision-making. In addition, members may also have strategic concerns such as whether their individual recommendations will conform to committee decisions. As a result, committee members often end up with distinct individual recommendations. Prominent examples include company boards, 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 an important empirical question in its own right, because of the prevalence of committee decisions in social-economic contexts. Besides, inference of members' idiosyncratic preference could shed lights on policy questions such as predicting committee decisions under counterfactual circumstances, e.g., anonymous voting in committees.

To this end, we set up a structural model of committee decisions that rationalizes dissenting recommendations across committee members through their individual heterogeneities in two dimensions. 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 its group deliberation, each member may decide to weigh these sources differently in their individual perceptions about the consequence of each alternative. We refer to these two heterogeneities as "individual tastes for multiple objectives" and "ideological bias towards various information sources" respectively.

We focus on committees that make binary decisions and aggregate individual recommendations from its members through a majority rule. All committee members share a common information set that includes the states of the world and an announced target consisting of multiple objectives, as well as several sources of information that predicts the stochastic outcome from either alternative. The common goal for the committee is to choose an action that could minimize ex ante deviations from the announced target. We analyze the identification and estimation of the model under two cases: one that incorporates members' strategic concerns about conformity to the committee decision (a.k.a. "strategic recommendations"); and one that does not (a.k.a. "expressive recommendations"). In both cases, pure-strategy Bayesian-Nash equilibria (PBNE) exist, and the identification of ideological bias and private tastes are obtained by exploiting how the patterns of individual recommendations, or the conditional choice probabilities, change with the common information set 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 set are both reported in data. Without any strategic concern, the individual choice probabilities in equilibrium take the form of a mixture of taste distributions, where the mixing weights correspond to the probability masses of the ideological bias. Committee members with heterogeneous tastes and ideological bias react differently to the same changes 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 committee (a cross-sectional unit) makes several decisions in multiple episodes. 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 condition on specific bias) depend on endogenous patterns of recommendation by other members in equilibria.

We identify the model with strategic recommendations through sequential steps: First, we apply results from Hu and Schennach (2008) and Kasahara and Shimotsu (2009) to recover individual choice probabilities conditional on the ideological bias. The identifying power comes from observed variations in states, targets and various sources of information that members use to formulate their perceptions. The argument exploits identifying restrictions in the lower-dimensional submodels. Second, we show that, under mild conditions that have clear economic interpretations, 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 as an implication of the structural model. Lastly, we recover the distribution of individual tastes using continuous variation in the component choice probabilities due to changes in common information, by imposing a minimum set of semiparametric restrictions.

This paper contributes to the literature on structural analyses of committee decisions in several ways. First off, we model committee members' private types and their effects on individual recommendations in the presence of a common information set; whereas the existing literature mostly focus on the aggregation of private information among members, e.g., Iaryczower and Shum (2012b). Our model also differs fundamentally from structural voting / election models, which involve numerous decision-makers with heterogenous 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 perceptions. Instead, we rationalize differences in individual choices through idiosyncratic bias and tastes while processing the common information. Besides, our nonparametric approach for identifying the model is innovative. The proposed method is instrumental for understanding the process of committee decisions. Furthermore, recovery of strategic incentives in the model provides a useful framework that can be employed to address mechanism design questions, such as efficiency of open versus anonymous committees.

Based on our identification arguments, we estimate a structural model of interest rate decisions by the Monetary Policy Committee at Bank of England, allowing for strategic concerns among its members. We find that all MPC members tend to put more weights on the forecasts of the Bank than on the forecasts of outsiders 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 two types of committee members, external and internal, differ more in their tastes for multiple objectives than in their ideological bias.

The existing literature on committee decision-making is mainly theoretical and focus on information aggregation.[1] Our paper is closely related to an emerging literature on econometric and empirical analysis of collective decision-making. Iaryczower, Lewis, and Shum (2013) and Iaryczower and Shum (2012b,a) study decision-making of the US supreme court where justices have incomplete information and common values. In a similar framework, Iaryczower, Shi, and Shum (2012) analyze the effects of deliberation among justices using an approach of partial identification. Merlo and De Paula (2010) and Kawai and Watanabe (2013) consider nonparametric identification and estimation of ideological voters' preferences and partial identification of a strategic voting model respectively. As explained above, their models differ qualitatively from our model of committee decisions. Our paper is also related to a recent literature on non-classical measurement error and finite mixture models. Hu and Schennach (2008) provides a general identification result for models with nonclassical measurement error.[2] Hall and Zhou (2003), Kasahara and Shimotsu (2009), and Henry, Kitamura, and Salanié (2013) consider nonparametric identification of finite mixture models using the identification power of covariates.

The rest of the paper is planned as follows. In Sections 2 and 3 we show identification of models for expressive and strategic recommendations respectively. In Section 4 we apply our model to analyze policy decisions by MPC at the Bank of England. Section 5 concludes.

2 Expressive Recommendations

Some committees are ad hoc in that they are organized for a short length of time and only make few (or even a single) decisions. For instance, civil or criminal courts put together one-time juries by summoning randomly chosen eligible citizens. Other committees make

[1]Please see Hao and Suen (2009) for a complete summary.

[2]See Chen, Hong, and Nekipelov (2011) for a survey of recent development in measurement error models.

multiple decisions over its lifetime, but its members are nonetheless not subject to strategic considerations such as reputation for good judgement, or conformity to group choices. One example is the board of directors / partners in charge of major business operations in corporations. These members are at the top of company hierarchy and therefore are hardly constrained by concerns for promotion, etc. In this section, we analyze a structural model where committee members make expressive recommendations, i.e. to propose actions that minimize ex ante deviation from targets based on their individual perceptions.

2.1 The Model

Consider a cross-sectional data containing independent episodes of group decisions by various committees, each of which may consist of a different set of members. In each episode, all members observe some state S drawn from a distribution with a finite support $\mathcal{S}$. Each member formulates an individual perception about how a binary action $d \in \{0,1\}$ could affect a stochastic outcome $Y \in \mathbb{R}^K$ under that state. (We use upper cases to denote random variables and lower cases for their realized values.) The outcome space $\mathcal{Y} \subseteq \mathbb{R}^K$ is finite, with cardinality Q and generic elements denoted by y^q.

In each episode with a state s, the committee can access two sources of information about how both actions could impact the distribution of Y under state s. Denote these two sources by $G(s) \equiv (G_1(s), G_0(s))$ and $H(s) \equiv (H_1(s), H_0(s))$, where $G_d(s)$ and $H_d(s)$ summarize the probability masses of Y given state s and action d according to the two sources respectively. Specifically, the q-th coordinate in $G_d(s)$, denoted $G_{q,d}(s)$, is the probability that "$Y = y^q$ given s and d" according to $G(s)$. Likewise for $H_{q,d}(s)$. Across all decision episodes with states s, both $G_d(s), H_d(s)$ are independent draws from two distinct distributions with their supports being subsets of a connected set $\mathcal{H} \equiv \{v \in \mathbb{R}^Q_+ : \sum_q v_q = 1\}$. Such a specification captures the uncertainty in the information available even under the same states. Both $G(s)$ and $H(s)$ are common knowledge among committee members. To facilitate exposition hereinafter, we refer to G, H respectively as "initial perception" and "empirical evidence". In addition, the committee is informed of a target $\tilde{y} \in \mathcal{Y}$ in each decision episode, which is allowed to vary across episodes. The information set available to committee members is thus summarized by $\mathcal{I} \equiv \{s, \tilde{y}, G(s), H(s)\}$.

Example 1 (*Corporate Board Decisions*) *A corporate board of directors tries to decide on a proposal to merge with another company. Each board member makes a binary recommendation to approve or deny the proposal. All members agree the merger could affect the company's stock price and the rate of return on assets (ROA), i.e.,* $K = 2$. (*Note these two dimensions of outcome do not move in the same direction because changes in stock prices are affected by the post-merger debt structure while the changes in ROA are determined by* *the synergy between the production or sales forces following the merger.) The board members agree on a common target of stock prices and ROA, denoted by* $\tilde{y} \equiv (\tilde{y}_1, \tilde{y}_2)$*.*

While deliberating on the final decision, the board takes into account the industry and market environment as well as status of both parties involved in the merger (s)*. They also have access to two sources of information. These include external evaluations of possible outcomes from an independent management consulting firm, i.e.* $G(s)$*; and predictions of consequences of merger (or non-merger) based on due-diligence by the company staff and historical evidence from past mergers with similar features, i.e.* $H(s)$*. Despite the common goal* $\tilde{y}$ *and information from* $G(s), H(s)$*, the board members hold different views about how the proposal could affect the uncertain outcome of the new merger's stock prices and ROA. They also disagree on relative weights that should be assigned to these two dimensions in the final decision.*

A member i's individual perception about the probability that "the outcome is y^q under s and d" is formulated as:

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

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", for it captures the members' willingness to adjust their initial perception in response to empirical evidence or to balance information from the two sources.

In each episode, every member i independently draws a vector of weights W_i from some distribution with support $\mathcal{W} \subseteq \mathbb{R}^K_{++}$. Member i recommends:

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

where the committee decision D^* aggregates individual recommendations through a majority rule. The expectation in (2) is taken with respect to the stochastic outcome Y and the group 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 the multiple dimensions in the outcome. These weights are heterogeneous and capture discrepancies between members after deliberations. 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 in data be fixed at an odd number I, and maintain the following assumptions throughout the section.[3]

[3]Our method 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) Across decision episodes and members in each committee, the individual bias α_i are independent draws from a distribution F_α over a known discrete support $\mathcal{A} \equiv \{\alpha^1, ., \alpha^J\} \in (0,1)^J$; and the tastes W_i, are independent draws from a continuous distribution F_W with positive densities 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 $\mathcal{I} \equiv \{S, \tilde{Y}, G(S), H(S)\}$.*

The exogeneity of the information set $\mathcal{I}$, and in particular the empirical evidence $H(S)$, is instrumental for our identification methods. This condition can be satisfied even when the empirical evidence potentially depends on the history of past states and committee decisions. As in the example of corporate board decisions, $H(S)$ is based on the accumulated evidence up to the date of decisions. Thus it is subject to 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 plausible in a panel data, where committees are observed to make multiple decisions throughout its tenure.

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, further 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 reported in data. Nevertheless, such observed heterogeneities do not pose any conceptual challenge to our identification exercise in that our method below extends once conditional on characteristics of committee members reported in data.

2.2 The Data and Equilibrium

Consider a data set that records individual recommendations $\{d_i : i = 1, .., I\}$ from many independent episodes of committee decisions. Each committee consists of I members and makes a single or multiple decisions throughout the data. In the later case, the data has some panel structure. However, there is no strategic dependence between decisions made by the same committee across the episodes, and members' private types are independent draws from the same distribution, so that Assumption 1 still holds.[4]

[4]Our method in this section does *not* require the data to have a panel structure. The approach is proposed in an environment where each committee is observed to make a single decision in just one episode.

The states s and the announced targets $\tilde{y}$ are reported in each episode. The empirical evidence $H(\cdot)$, is also reported in data. We also assume the data allows researchers to observe individual and committee choice patterns conditional on the initial perception $G(\cdot)$, at least for some realizations of state s. This holds, for instance, if there exist states of the world s where the initial perception $G(s)$ is known a priori, either as a result of common sense or institutional environment. For example, potential jurors' attitude toward a series of legal phrases are surveyed and analyzed in Kadane (1983). Costanzo, Shaked-Schroer, and Vinson (2010) provide a survey of jury-eligible men and women for their beliefs about police interrogations, false confessions, and expert testimony.

This model admits a unique PBNE where all players adopt dominant pure strategies. Fix a common information set $\mathcal{I} \equiv \{s, \tilde{y}, G(s), H(s)\}$ and define:

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

In words, $\delta_{G,k}$ is the difference in ex ante deviation from target $\tilde{y}$ under the two alternatives for a member who cares only about the k-th dimension of outcome and who puts all the weights on the initial perception $G(s)$ in his individual perception. Define $\delta_{H,k}$ in a similar manner, only with G in (3) replaced by H. That is, a member who cares only about the k-th dimension of outcome and who puts all weight on the empirical evidence would choose $d_i = 1$ if and only if $\delta_{H,k}$ is negative.

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

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

The existence of such a dominant pure-strategy BNE is due to two facts. First, with expressive recommendations, a member i's objective function depends on his recommendation d_i only through ex ante deviation from the target under the stochastic committee decision. Such an ex ante deviation depends on d_i only through its impact on the distribution of the committee decision D^*. Second, by construction, the probability for "$D^* = 1$ conditional on $d_i = 1$" exceeds that for "$D^* = 1$ given $d_i = 0$".

2.3 Identification of F_{α_i} and F_{W_i}

We maintain that individual choices reported in data are generated as members adopt dominant strategies in (4). Lemma 1 has a powerful implication for the empirical task of

On the other hand, if the data does report multiple decisions by the same committee, then by our arguments the model is over-identified. In fact, one can then exploit the panel structure to identify a richer model with strategic interactions, provided the empirical evidence and W_i are random draws for the same member across different episodes. Section 3 elaborates on this idea and identifies a model with strategic interaction.

inferring members' type distributions: Each individual members' decision observed from data could be viewed as independent realizations of the decision rule in (4) with (α_i, W_i) being i.i.d. draws from some unknown distribution.

We maintain the following assumptions on the data-generating process.

Assumption 2 *There exists some* $\{s, \tilde{y}, G(s)\}$ *such that* (*a*) *sign* $(\sum_k w_{i,k}\delta_{G,k}(\mathcal{I}))$ *is known and remains the same for all* $w_i \in \mathcal{W}$*; and* (*b*) *there exists* (H^a, H^b) *such that* $\Pr\{D_i = 1|s, \tilde{y}, G(s), H\} = 0$ *for all* H *in an open neighborhood around* H^a *and* $\Pr\{D_i = 1|s, \tilde{y}, G(s), H\} = 1$ *for all* H *in an open neighborhood around* H^b*.*

Part (a) requires that under certain circumstances, individual recommendations would be degenerate regardless of tastes W_i if based on the initial perception G only. In other words, there are states of the world under which initial perceptions are unequivocal about which of the two actions should be taken so as to achieve the announced target. For example, consider a corporate board of a manufactory that deliberates over a proposed vertical merger with its suppliers, while the announced target is to boost its stock prices while keeping the ROA at the current level. Suppose the initial perception (e.g. an evaluation by an external management consulting firm) suggests that under contemporary market and industry conditions there should be a high cost synergy from the merger. Then, without further investigating empirical evidence $H(s)$, the board members would reach a consensus to recommend the merger given the target, regardless of their heterogeneous tastes.

Part (b) is a joint restriction on $\{s, \tilde{y}, G(s)\}$ and the support of empirical evidence. It requires there be "*extreme evidence*" $H(s)$ under which individuals recommendations are degenerate given the triple of states, targets and initial perceptions. This condition can be verified using the distributions of individual choices and empirical evidence in data.

The rest of Section 1 identifies the distributions of α_i and W_i, conditional on a triple $\{s, \tilde{y}, G(s)\}$ that satisfies Assumption 2. We suppress the dependence on this triple to simplify notations. Without loss of generality, we consider the case where $(s, \tilde{y})$ is such that a member always recommends $d_i = 0$ if based on the initial perception only (i.e. $\sum_k W_{i,k}\delta_{G,k} > 0$). Then from (4)

$$\Pr\{D_i = 1|\mathcal{I}\} = \Pr\left\{c(W_i; \mathcal{I}) \le -\tfrac{1-\alpha_i}{\alpha_i}\right\}, \text{ where} \tag{5}$$

$$c(W_i; \mathcal{I}) \equiv \left(\textstyle\sum_k W_{i,k}\delta_{H,k}(\mathcal{I})\right) / \left(\textstyle\sum_k W_{i,k}\delta_{G,k}(\mathcal{I})\right). \tag{6}$$

2.3.1 Visualizing individual choices

To get an overview of our main methods, it helps to visualize how private types affect individual decisions. Suppose the outcome space is two-dimensional ($K = 2$), which is common in a lot of applications including the corporate board decision example above and the monetary policy decisions at the Bank of England below (where decisions are discretized into binary actions as to whether to increase the current interest rate or not). We sketch our main arguments for $J \equiv |\mathcal{A}| = 2$ and $Q \equiv |\mathcal{Y}| = 3$, i.e. with two possible types of ideological bias and three outcome scenarios. Proofs for general cases with $J > 2$ and $Q > 3$ do not pose new conceptual challenges and require more tedious algebra.

We begin by simplifying notations so as to facilitate our visualization. First, set $W_{i,1} = 1$ as a scale normalization. Drop the second subscript from $W_{i,2}$ and denote its normalized support by $\mathcal{W} \equiv [\underline{w}, \overline{w}]$. Second, without loss of generality, let $\tilde{y} = y^3 \in \mathcal{Y}$ so that the summation in the definition of $\delta_{G,k}, \delta_{H,k}$ is reduced to $\sum_{q=1,2}$. Then 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 $G(s)$. Thus, with a slight abuse of notation we write the common information set $\mathcal{I}$ in a different but more succinct way as $\mathcal{I} \equiv \{s, \tilde{y}, g(s), h(s)\}$ where $g \equiv (g_1, g_2)$ and $h \equiv (h_1, h_2)$ hereinafter. By construction, the support of h is $\{h \in [-1,1]^2 : h_1 + h_2 \in [-1,1]\}$. Third, let $R_i \equiv -(1-\alpha_i)/\alpha_i$ and denote its support $\mathcal{R} \equiv (r^1, r^2)$. W.L.O.G., let $r^1 < r^2$, or equivalently, $\alpha^1 < \alpha^2$. Finally, suppress the dependence of $c(W_i; \mathcal{I})$ and $\Pr\{D_i = 1|\mathcal{I}\}$ on the triple $\{s, \tilde{y}, g(s)\}$ that is conditioned on and write them as $c(W_i; h)$ and $\Pr\{D_i = 1|h\}$, where s is also suppressed from $h(s)$. We also refer to h as the empirical evidence hereinafter.

For a member with $(\alpha_i, W_i) = (\alpha^j, w)$, define an "indifference hyperplane" $\{h : c(w; h) = r^j\}$ (These are lines in $\mathbb{R}^2$ when $Q = 3$). Each hyperplane partitions the space of empirical evidence into two parts: one in which the dominant-strategy for a member with (α^j, w) is to choose 1 and the other 0. While the intercepts of the hyperplanes depend on ideological bias, their slopes depend on tastes only according to (5) and (6). For instance, given any empirical evidence h on the lower-left side of a hyperplane associated with (α^j, w), the dominant choice for a member with (α^j, w) is 1.

We now summarize some properties of the hyperplanes that are useful for the identification exercise. For any w, the two hyperplanes associated with (α^j, w) for $j = 1, 2$ are parallel by construction. For any α^j, the hyperplanes (lines) associated with different weights w intersect at the same point, denoted by $h^j \equiv (h_1^j, h_2^j)$, for all $w \in \mathcal{W}$. Slopes of these hyperplanes are negative because the support of tastes W_i is non-negative. These slopes are the rates of substitution between the two dimensions of the empirical evidence that is required to keep a type-(α^j, w) member indifferent between two alternatives. Also, conditioning on $\{s, \tilde{y}, g(s)\}$, the rate of substitution is independent from the empirical evidence h and the ideological bias α_i, and is monotonic in the taste for outcomes W_i.[5] The direction of monotonicity of the slopes as W_i increases is determined by the target $\tilde{y}$ and the outcome space $\mathcal{Y}$, which is available in data.

For the rest of Section 2.3, we present our method for the case where the hyperplane for type-$(\alpha^j, \underline{w})$ has a greater slope (i.e. is "less steep") than that for type-$(\alpha^j, \overline{w})$. This is

[5]The slope of the hyperplane given taste w (and any α^j) is $\xi(w) \equiv -(a_{1,1} + wa_{1,2})/(a_{2,1} + wa_{2,2})$, where $a_{q,k} \equiv (y_k^q - \tilde{y}_k)^2$. The sign of $\partial\xi(\tilde{y}; w)/\partial w$ depends on the vector of $a_{q,k}$'s but not on w or h.

without loss of generality, since the ranking between these two rates is identifiable from data as argued above. Identification under the other case follows from symmetric arguments.

The six panels in Figure 1 enumerate all possibilities regarding the positions of a generic pair of extreme evidence h^a, h^b relative to the indifference hyperplanes. The "flatter" and the "steeper" hyperplanes are associated with $\underline{w}$ and $\overline{w}$ respectively while dashed and solid hyperplanes are associated with α^1 and α^2, respectively. To lay out our argument, it is convenient to index the set of convex combinations of h^a, h^b, denoted $\mathcal{H}(h^a, h^b)$, by their weights in front of h^b. That is, for any $h \in \mathcal{H}(h^a, h^b)$,

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

For any $\lambda \in [0,1]$, let $h(\lambda)$ be a shorthand for $\lambda h^b + (1-\lambda)h^a$. Also define:

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

In words, $\underline{\lambda}_j$ 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. Hereinafter we refer to them as "indifference thresholds" on $\mathcal{H}(h^a, h^b)$. Thus the positions of extreme evidence relative to the hyperplanes are fully characterized by the order of these thresholds over $\mathcal{H}(h^a, h^b)$.

2.3.2 Finding out the order of indifference thresholds

To reiterate, our identification arguments condition on a triple $\{s, \tilde{y}, g(s)\}$ and a pair of extreme evidence (h^a, h^b) that satisfy Assumptions 2. The first step in our method is to order the indifference thresholds over the set of convex combinations of h^a and h^b, i.e. $\mathcal{H}(h^a, h^b)$. This matters for the subsequent identification of F_{α_i}, F_{W_i}, because it determines how variations in the empirical evidence over $\mathcal{H}(h^a, h^b)$ affect individual choice probabilities through the distributions of α_i and W_i.

To see how the order of thresholds matters for the identification exercise, first consider the case in panel (i) of Figure 1. Under h^b, all members choose 1 regardless of types α_i, W_i; under h^a, all choose 0. For any $h \in \mathcal{H}(h^a, h^b)$ to the right of the indifference hyperplane for $(\alpha^2, \overline{w})$ (i.e. $\lambda(h) < \overline{\lambda}_2$), the individual conditional choice probability (CCP) $\Pr\{D_i = 1|h\}$ is zero. That is, all members would choose 0, regardless of their types (α_i, W_i) whenever h is sufficiently close to h^a. As the evidence moves toward h^b on $\mathcal{H}(h^a, h^b)$ and $\lambda(h)$ crosses $\overline{\lambda}_2$, members with α^2 and w_i close to $\overline{w}$ switch to choose 1 so that $\Pr\{D_i = 1|h\}$ becomes positive. As h moves further towards h^b and $\lambda(h)$ crosses $\overline{\lambda}_1$, members with α^1 and w_i close enough to $\overline{w}$ also start to choose 1. When the evidence moves further toward h^b and beyond $\underline{\lambda}_2$, all members with bias α^2 choose 1 regardless of tastes W_i, while those with α^1 and sufficiently low w_i still choose 0. The CCP becomes degenerate once the evidence moves beyond $\underline{\lambda}_1$ and is sufficiently close to h^b. Hence to recover F_{α_i}, F_{W_i} from the CCPs, one needs to deal with a finite mixture of two non-degenerate distributions at least for certain range of empirical evidence.

Next, consider panel (ii) in Figure 1. There the CCPs can not change as h vary over a range of empirical evidence indexed between $\overline{\lambda}_2$ and $\overline{\lambda}_1$. This is because, conditional on any such evidence h, all members with bias α^2 choose 1 while all those with bias α^1 choose 0, *regardless of their tastes* w_i *respectively.* In comparison, the identification of F_{α_i}, F_{W_i} does not require dealing with finite mixtures of non-degenerate distributions, but the main challenge is to relate the indifference hyperplanes for different tastes w_i to empirical evidence over the set of convex combinations $\mathcal{H}(h^a, h^b)$.

To fully recover the order of the indifference thresholds, we exploit variations in the extreme evidence. In particular, as the pair extreme evidence h^a, h^b move, the relative positions of indifference thresholds will register different patterns of change, due to the difference in the rates of substitution over the indifference hyperplanes. This can be visualized in Figure 1 as follows. Fix h^b and vary h^a vertically (i.e. in the dimension of h_2^a alone), the direction of changes in the distance between the ordered indifference thresholds differ 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 and h_1^a, h^b are fixed; in contrast, such a distance would diminish in case (iv) under the same movements of h_2^a. The following lemma formalizes this argument by providing details about how the comparative statistics in distances between identified thresholds differ across the six scenarios.

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

2.3.3 Recovering F_{α_i} and F_{W_i}: Cases (ii), (iii), (v) and (vi)

With the order in $\{\underline{\lambda}_j, \overline{\lambda}_j\}_{j=1,2}$, we now show how to recover the distributions of α_i and W in each one of the six cases. We first consider the cases where the tripe $\{s, \tilde{y}, g(s)\}$ conditioned on and the pair of extreme evidence (h^a, h^b) used are such that the order of indifference thresholds are as shown in panels (ii), (iii), (v) and (vi) of Figure 1. These are relatively easier cases, because the CCPs are never mixtures of non-degenerate distributions as evidence varies over the convex combination $\mathcal{H}(h^a, h^b)$.

Assumption 3 (*a*) *For the target* $\tilde{y}$*, the 2-by-2 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 at state* s.

Part (a) of this assumption is a mild condition on the announced goal for the committee. Among other things, it rules out uninteresting pathological situations where individual tastes do not matter for calculating ex ante deviations 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. Also note 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. Under the other maintained assumptions, it is sufficient for implying the monotonicity of c in w for almost all evidence. 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, h^b still holds. Also, it is worth noting that part (b) can in principle be verified using data. This is because, as shown below, the initial perceptions are identifiable from CCPs even without part (b) under other maintained assumptions.

Proposition 1 *Suppose Assumption 1 holds, and there exists $\{s, \tilde{y}, g(s)\}$ and a pair of extreme evidence (h^a, h^b) satisfying Assumptions 2 and 3. If the order of indifference thresholds $\{\underline{\lambda}_j, \overline{\lambda}_j\}_{j=1,2}$ is as shown in panels (ii), (iii), (v) or (vi) in Figure 1, then F_{α_i} and F_{W_i} are identified.*

The intuition for this result can also be visualized using the panels in Figure 1. Consider case (ii). By varying the empirical evidence from h^a to h^b, one can recover the second indifference threshold as the infimum of the subinterval of $\mathcal{H}(h^a, h^b)$ over which the individual CCPs is in the interior of $(0,1)$ but invariant. In addition, the invariant CCPs over the interval then identify probability mass function for α^1. The full-rank condition in part (a) of Assumption 3 then implies that the initial perception $g(s)$ can be recovered using knowledge of support of bias $\mathcal{A}$. Consequently, the equations characterizing the indifference hyperplanes are also identified. With the monotonicity of c in tastes induced by part (b) of Assumption 3, we can invert the CCPs to recover (over-identify) the distribution of idiosyncratic tastes.

2.3.4 Recovering F_{α_i} and F_{W_i}: Cases (i) and (iv)

Consider case (i). Define:

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

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 (and r^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

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

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 $\varphi(h) \equiv \Pr(D_i = 1|h)$ 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 4 $\overline{w}_1 < \underline{w}_1$.

This is a joint restriction on the triple $\{s, \tilde{y}, g(s)\}$ and the pair of extreme evidence (h^a, h^b) conditioned on. 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 the two bias types. 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 Assumption 1 holds, and there exists* $\{s, \tilde{y}, g(s)\}$ *and extreme evidence* (h^a, h^b) *satisfying Assumptions 2, 3 and 4. If the order of indifference thresholds* $\{\underline{\lambda}_j, \overline{\lambda}_j\}_{j=1,2}$ *is as shown in panels (i) or (iv) in Figure 1, then* F_{α_i} *and* F_{W_i} *are identified.*

Identification can also be established for the more general case where Assumption 4 fails. Proof under this scenario requires more tedious algebra and is presented in the Appendix B for the sake of completeness.

2.3.5 Discussions about estimation

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 $\mathcal{P} \equiv \{p \in [0,1]^{|\mathcal{A}|} : \sum_{j \le |\mathcal{A}|} p_j = 1\}$ 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

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

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 "$\alpha^j \sum_k W_{i,k} \delta_{H,k}(\mathcal{I}) + (1 - \alpha^j) \sum_k W_{i,k} \delta_{G,k}(\mathcal{I}) \le 0$" 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 Recommendations

Committees are often scheduled to make multiple decisions, and the members could be motivated by career concerns (e.g. promotion, re-election, or reputation, etc). Thus members may care about how likely their recommendations conform with the final committee decision. Such a strategic incentive could lead them to deviate from what would be the decision under expressive recommendation. The strategic interaction between members depends on how the committee aggregates individual recommendations into a group decision. In what follows, we focus on identification when committees adopt a simple majority rule. Our method can be applied under alternative rules for aggregating individual decisions, as long as these rules are known to researchers.

3.1 The Model

Consider a panel data with each cross-sectional unit being an independent committee ($\mathcal{C}$) that makes decisions in several episodes $l = 1, ..., L$. (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 $|\mathcal{C}|$ are fixed throughout the data.[6]

Members' payoffs are similar to that in Section 2.1, except for an additional incentive to conform with the final 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 set in episode l. In a Bayesian Nash Equilibrium, a member i chooses $D_{i,l}(\alpha_i, w_{i,l}; \mathcal{I}_l) \in \{0,1\}$ in episode l as:

$$\arg \min_{d_{i,l} \in \{0,1\}} \left\{ \mathbb{E}_{D^*} \left[1(D_l^* \neq d_{i,l}) | d_{i,l}; \alpha_i, \mathcal{I}_l \right] + \mathbb{E}_{Y,D^*} \left[\sum_k w_{i,l,k} (Y_{l,k} - \tilde{y}_{l,k})^2 | d_{i,l}; \alpha_i, \mathcal{I}_l \right] \right\}, \qquad (7)$$

where $w_{i,l,k}$ is the weight i puts on the k-th objective in episode l. The expectation in (7) 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 (7) captures strategic incentives, or "career concerns". We have normalized the weight on this term to 1, so that $w_{i,l,k}$ are interpreted as weights on the outcome relative to the strategic incentive.[7] Similar specifications were used in Levy (2007) and Bergemann and Morris (2013).

We need to refine the assumptions on the information available to the committee and individuals in the current context of panel data structure.

[6]If data report different number of members across committees, our method applies after conditioning on sizes of committees.

[7]This is an innocuous normalization because by construction the scale of unobserved weights cannot be identified, just as in the model of expressive recommendations.

Assumption 5 (*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 densities over a known support $\mathcal{W} \subseteq \mathbb{R}_+^K$. (b) The tenure of each committee is partitioned into several intervals, each of which consists of m 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} that is 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) is motivated by the reality that individual tastes $W_{i,l}$ are affected by group deliberations in each episode, but often remain idiosyncratic regardless of the pooled information. We allow such idiosyncrasies to be correlated with the common information set $\mathcal{I}_l$. Part (b) captures the fact that members' bias is more persistent than tastes for different objectives. One reason for this is that members need time to get experienced before changing the weights that they apply to two sources of information. For instance, one of the sources of information is updated every m periods. Another reason is that the length of service for each member is in unit of m episodes, and thus personnel changes always occur at the end of an interval with m episodes. (See Example 2 for details.)

Assumption 6 (*a*) *Members in the same committee share the same initial perception $G(S_l)$, which is fixed within each interval of m episodes, but across intervals are i.i.d. draws from some distribution over $\mathcal{H}$ conditional on realizations of 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 realizations of S_l.*

As with expressive recommendations, our identification method can be extended to accommodate committee- and member-level heterogeneities that are reported in data. Also, our identification method below can be extended to allow for heterogeneity in members' initial perceptions, as long as such heterogeneity can be conditioned on in 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 (7), we can rewrite the minimization problem in (7) equivalently as:

$$\max_{d_{i,l}\in\{0,1\}} \Pr(D_l^* = d_{i,l}|d_{i,l};\alpha_i,\mathcal{I}_l) - \sum_{k,q} w_{i,l,k}\left[a_{q,k}(\tilde{y}_l)\Big(\sum_{d\in\{0,1\}} \mathcal{F}_{l,q,d}(s_l;\alpha_i)\Pr(D_l^* = d|d_{i,l},\mathcal{I}_l)\Big)\right]. \quad (8)$$

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

Example 2 (*Monetary Policy at the Bank of England*) *The Monetary Policy Committee (MPC) at the Bank of England meets monthly to set an interest rate they judge will minimize* *ex ante deviation of future outcomes from the targeted inflation and GDP. That is, for each episode (month), the target is two dimensional, i.e.* $\widetilde{y}_l = \{\widetilde{p}_l, \widetilde{\pi}_l\}$ *with* $\widetilde{p}_l$ *being the targeted GDP and* $\widetilde{\pi}_l$ *the targeted inflation rate. (See for example Besley, Meads, and Surico (2008).) The weights that* i *assigns to GDP and inflation, relative to strategic incentives, are given by* $w_{i,l,1}$ *and* $w_{i,l,2}$ *respectively.*

Each member formulates an updated perception $\mathcal{F}_{l,q,d}(s;\alpha_i)$ *about how actions affect the stochastic outcome of GDP and inflation rate. Such a perception is a weighted average of two most relevant sources of information: two sets of forecasts of inflation and output under various interest rates, one by MPC and the other by outsiders (i.e. non-MPC professionals in the private financial sector). The forecasts by outsiders are reported quarterly, while the forecasts for MPC members are adjusted through their monthly deliberations prior to decisions. The length of service in the committee varies across members.*

We now establish the existence of symmetric Bayesian Nash equilibria in the model. As before, we condition on the common information $\mathcal{I}_l$ and suppress it from notations. A pure strategy profile is defined as $\sigma \equiv (\sigma_i)_{i\in\mathcal{C}}$, where $\sigma_i : \mathcal{A} \otimes \mathcal{W} \to \{0,1\}$. 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 PBNE at* $\mathcal{I}_l$ *in the model of committee decisions with strategic recommendations if for all* i *and* $(\alpha_i, w_{i,l})$,

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

where $\sigma_{-i} \equiv (\sigma_j)_{j\neq i}$. *A PBNE is symmetric if* $\sigma_i = \sigma_j$ *for all* i, j.

We first characterize best response functions for committee members. 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:

$$\sigma_i(\alpha_i, w_{i,l}; \sigma_{-i}, \mathcal{I}_l) \equiv 1\left\{\alpha_i \sum\nolimits_k w_{i,l,k}\delta_{H,k}(\mathcal{I}_l)) + (1-\alpha_i)\sum\nolimits_k w_{i,l,k}\delta_{G,k}(\mathcal{I}_l)) \leq \tfrac{p_{1,1}(\sigma_{-i}) - p_{0,0}(\sigma_{-i})}{p_{1,1}(\sigma_{-i}) - p_{1,0}(\sigma_{-i})}\right\}, \tag{9}$$

where $\delta_{H,k}$ and $\delta_{G,k}$ are functions of $\mathcal{I}_l$ defined as in (3).[8] Note the L.H.S. of the inequality in (9) summarizes i's perception of the difference in ex ante deviations under the two alternatives, in the absence of any strategic concerns.

[8]To see this, 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 (8) that depends on d_i as:

$$p_{d_i,d_i}(\sigma_{-i}) - p_{1,d_i}(\sigma_{-i}) \sum\nolimits_{k,q} w_{i,l,k} a_{q,k}(\tilde{y}_l) \Delta\mathcal{F}_{l,q}(s_l;\alpha_i)$$

where $\Delta\mathcal{F}_{l,q}(s;\alpha_i) \equiv \mathcal{F}_{l,q,1}(s;\alpha_i) - \mathcal{F}_{l,q,0}(s;\alpha_i)$.

Let $\phi(\cdot\,;\mathcal{I}_l) \equiv \Pr\{\alpha_i \sum_k w_{i,l,k}\delta_{H,k}(\mathcal{I}_l) + (1-\alpha_i)\sum_k w_{i,l,k}\delta_{G,k}(\mathcal{I}_l) \leq \cdot \mid \mathcal{I}_l\}$ and $\phi'(\cdot\,;\mathcal{I}_l)$ be its derivative with respect to the first argument. Define a mapping $\varphi : (0,1) \to (-\infty,+\infty)$ such that $\varphi(\tau)$ equals the R.H.S. of the inequality in (9) when all others $j \neq i$ follow the same pure strategy which leads to $\Pr(D_{j,l} = 1|\sigma_j) = \tau$. The form of φ depends on the size of the committee. (We present a close form for φ when the number of members $|\mathcal{C}| = 3$ in Appendix C.) Nevertheless, $\varphi(\lambda) \to -\infty$ as $\lambda \to 0$ and $\varphi(\lambda) \to +\infty$ as $\lambda \to 1$ by construction, regardless of the committee size.[9] Let $\varphi \circ \phi$ denote the composition of the two functions.

Lemma 3 *Under Assumptions 5 and 6, the model of committee decisions with strategic recommendations has a symmetric PBNE at* $\mathcal{I}_l$ *if either* (*a*) $\mathcal{W}$ *is bounded; or* (*b*) *there exists* $\eta > 0$ *and a pair* $(\kappa_0, \kappa_1) \subset \mathbb{R}^1$ *such that* $\frac{d\varphi\circ\phi(\kappa;\mathcal{I}_l)}{dt} > 1 + \eta$ *for all* $\kappa > \kappa_1$ *or* $\kappa < \kappa_0$.

We sketch the proof of the lemma in the text. By construction, a symmetric PBNE at $\mathcal{I}_l$ is characterized by some $\kappa^* \in \mathbb{R}$ that solves:

$$\varphi \circ \phi(\kappa^*;\mathcal{I}_l) - \kappa^* = 0. \tag{10}$$

The composite function $\varphi\circ\phi$ is continuous given $\mathcal{I}_l$ due to our maintained assumptions. Also recall that $\varphi(\lambda) \to -\infty$ as $\lambda \to 0$; and $\varphi(\lambda) \to +\infty$ as $\lambda \to 1$. If the support $\mathcal{W}$ is bounded, then the L.H.S. of (10) must be negative if evaluated at some κ sufficiently small, and must be positive at some κ sufficiently large. The Intermediate Value Theorem implies there exists κ^* that solves (10). On the other hand, if $\mathcal{W}$ is unbounded, then the L.H.S. of the inequality in (9) has unbounded support. Condition (b) in Lemma 3 ensures that the derivative of $\varphi \circ \phi$ eventually remains sufficiently greater than 1 in both tails. This guarantees a fixed point exists in (10) as the absolute value of κ becomes sufficiently large. (Condition (b) is a restriction on the tail behavior of the distribution of the index on the L.H.S. of (9). In Appendix C, we provide an example of sufficient conditions that imply (b).)

In general, the model of strategic recommendations admits multiple PBNE because (10) could well admit multiple solutions for a given $\mathcal{I}_l$. For the rest of the paper, 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 PBNE. That is, a single equilibrium is being played across all committees (games) indexed by the same information $\mathcal{I}_l$.

For the rest of Section 3, we maintain that committee members' recommendations in data are rationalized by the symmetric PBNE defined in Definition 1. Our goal is to recover the distributions of ideological bias α_i and tastes $W_{i,l}$ from the distributions of individual recommendations $D_{i,l}$ and committee decisions D^* as well as the information $\mathcal{I}_l$. This is

[9] To see this, suppose $|\mathcal{C}| = 2n+1$. Then $p_{1,1} - p_{1,0} = \binom{2n}{n}\tau^{n+1}(1-\tau)^{n-1}$, which converges to 0 as τ approaches 0 or 1. On the other hand, $p_{1,1} - p_{0,0} = p_{1,1} + p_{1,0} - 1$, which converges to 1 as $\tau \to 1$ and converges to -1 as $\tau \to 0$.

done in three steps. First, recover type-specific CCPs $\Pr(D_{i,l} = d_{i,l}|\alpha_i;\mathcal{I}_l)$ and the probability masses for α_i as the components and weights of a finite mixture respectively, up to unknown values of α_i. Second, show the "type-specific" CCPs are monotonic in α_i. Together with the first step, this implies the distribution of α_i and the type-specific CCPs are fully recovered. Third, identify the distribution of $W_{i,l}$ using variations in type-specific CCPs $\Pr(D_{i,l} = d_{i,l}|\mathcal{I}_l;\alpha_i)$ due to continuous changes in $\mathcal{I}_l$.

3.2 Recovering type-specific CCPs

The first step of identification exploits the panel structure of the data. Under Assumptions 5 and 6, the private types $(\alpha_i, W_{i,l})$ and the common information $\mathcal{I}_l$ are both drawn every m 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 $L = m$ episodes. For the rest of this section, we maintain that, for a known s_l, the realizations of $G(s_l)$ can be effectively conditioned on in 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 $\mathcal{I}_l \equiv \{S_l,\ \tilde{Y}_l,\ G(S_l),\ H_l(S_l)\}$ (that is, $\mathcal{X} \equiv \mathcal{S} \times \mathcal{Y} \times \mathcal{H} \times \mathcal{H}$ where $\mathcal{S}$ and $\mathcal{Y}$ are finite and $\mathcal{H}$ is infinite). Let the lower case $x_l \equiv \{s_l, \tilde{y}_l, G(s_l), H_l(s_l)\} \in \mathcal{X}$ denote a realization of $\mathcal{I}_l$. Let $\mathbf{d}_i^t \equiv (d_{i,l})_{1\leq l\leq t}$, for all $1 \leq t \leq L$. That is, $\mathbf{d}_i^t$ denotes i's decisions up to the t-th episode in the cross-sectional unit. Likewise, let $\mathbf{s}^t$, $\tilde{\mathbf{y}}^t$, $\mathbf{G}^t(\mathbf{s}^t) \equiv (G(s_l))_{1\leq l\leq t}$ and $\mathbf{H}^t(\mathbf{s}^t) \equiv (H_l(s_l))_{1\leq l\leq t}$ denote the history of states, targets, and empirical evidence, respectively up to the t-th episode for all $1 \leq t \leq L$. Note there is no subscript l for $G(\cdot)$ as a function of states s_l, because the initial perception is fixed across episodes l under Assumption 6 (a). Let $\mathbf{x}^t$ denote the history of the common information $\{\mathbf{s}^t, \tilde{\mathbf{y}}^t, \mathbf{G}^t(\mathbf{s}^t), \mathbf{H}^t(\mathbf{s}^t)\}$ for $1 \leq t \leq L$.

Assumption 7 *(a) For all $l \leq L$ and all $\mathbf{d}_i^{l-1}, \mathbf{x}^l$ and α_i, the transition function of x_l satisfies $\Pr(x_l|\mathbf{d}_i^{l-1}, \mathbf{x}^{l-1};\alpha_i) = \Pr(x_l|d_{i,l-1}, x_{l-1})$; and (b) For all x_{l-1} and $d_{i,l-1}$, $\Pr(x_l|d_{i,l-1}, x_{l-1}) > 0$ for all x_l on the support of $\mathcal{I}_l$.*

Part (a) of Assumption 7 requires the transition of the common information set (states, targets, initial perception and empirical evidence) to follow a first-order Markov process that is stationary (time-homogenous).[10] In the example of MPC at the Bank of England, such

[10] This condition is in fact not necessary for the main result in the current subsection (Lemma 4): When transitions between $s, \tilde{y}, G, H$ are episode-dependent, arguments in Proposition 4 of Kasahara and Shimotsu (2009) can be applied to identify the probability mass function of ideological bias and the conditional choice

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 combinations of past state, target, decision and empirical evidence, any state and empirical evidence are reachable in the subsequent episode with positive probability. This assumption is empirically verifiable from the data.

It follows from Assumption 7 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 is stationary (time-homogenous). To see this, recall by construction 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 multiple episodes and with $W_{i,l}$ independent across l and orthogonal to the history $\mathcal{I}^{l-1}$, the type-specific CCPs $\Pr(D_{i,l} = 1|\mathbf{x}^l, \mathbf{d}_i^{l-1}; \alpha_i)$ must be a function of x_l and α_i only. Its time-homogeneity then follows from the assumption that $W_{i,l}$ are i.i.d. draws from the same distribution across episodes.

Thus by the law of total probability and Assumption 7, the joint distribution $\Pr(\mathbf{d}_i^L, \mathbf{x}^L)$ can be written as:

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

where ρ_{α_i} denotes the probability mass at α_i.

Recent development of nonparametric identification of similar models of finite mixture can be found, for example, in Hu and Schennach (2008) and Kasahara and Shimotsu (2009). Hence decisions of committee members, when heterogeneous ideological bias α_i is fixed in multiple episodes, are analogous to dynamic discrete choices with unobserved time-variant individual heterogeneity. Therefore this step of identification applies the method from Kasahara and Shimotsu (2009) to recover the type-specific CCPs $\Pr(d_{i,l}|x_l;\alpha_i)$ and ρ_{α_i}. Similar to Kasahara and Shimotsu (2009), define:

$$\widetilde{\Pr}(\mathbf{d}_i^L, \mathbf{x}^L) \equiv \frac{\Pr(\mathbf{d}_i^L, \mathbf{x}^L)}{\prod_{l=2,.,L} \Pr(x_l|d_{i,l-1}, x_{l-1})} = \sum_{\alpha_i} \rho_{\alpha_i} \Pr(d_{i,1}, x_1 \mid \alpha_i) \prod_{l=2,.,L} \Pr(d_{i,l}|x_l,\alpha_i) \tag{11}$$

where the L.H.S. is directly identifiable from data. As in Kasahara and Shimotsu (2009), integrating out subvectors in $(\mathbf{d}_i^L, \mathbf{x}^L)$ leads to submodels. For instance, integrating out all in $(\mathbf{d}_i^L, \mathbf{x}^L)$ but $(d_{i,1}, x_1)$ leads to $\widetilde{\Pr}(d_{i,1}, x_1) = \sum_{\alpha_i} \rho_{\alpha_i} \Pr(d_{i,1}, x_1|\alpha_i)$; while integrating out all in $(\mathbf{d}_i^L, \mathbf{x}^L)$ but $(d_{i,2}, x_2)$ leads to $\widetilde{\Pr}(d_{i,2}, x_2) = \sum_{\alpha_i} \rho_{\alpha_i} \Pr(d_{i,2}|x_2,\alpha_i)$. Using the first two periods of observations, we obtain $\widetilde{\Pr}(d_{i,1}, x_1, d_{i,2}, x_2) = \sum_{\alpha_i} \rho_{\alpha_i} \Pr(d_{i,1}, x_1 \mid \alpha_i) \Pr(d_{i,2}|x_2,\alpha_i)$. Without loss of generality, let $\{1, 2, \cdots, B\}$ denote a set of realized values for $\mathcal{I}_l$. Denote the type-specific CCPs $\Pr(D_{i,1} = 1, \mathcal{I}_1 = x| \alpha_i = \alpha^j)$ by ξ_x^j, for $j \in \{1, 2, \cdots, |\mathcal{A}|\}$ and $x \in \{1, 2, \cdots, B\}$; and denote $\Pr(D_{i,l} = 1|\mathcal{I}_l = x, \alpha_i = \alpha^j)$ by ζ_x^j. The observed joint probabilities of choice, state

probabilities up to unknown values of α_i but under additional assumptions similar to Assumption 8. Nevertheless, this condition simplifies follow-up arguments in subsequent steps in Sections 3.3 and 3.4.

and empirical evidence at period 1, period 2, and the first two periods are denoted by $\chi_x \equiv \widetilde{\Pr}(D_{i,1} = 1, \mathcal{I}_1 = x)$, $\chi^*_x = \widetilde{\Pr}(D_{i,2} = 1, \mathcal{I}_2 = x)$, and $\chi_{x,x'} = \widetilde{\Pr}(D_{i,1} = 1, \mathcal{I}_1 = x, D_{i,2} = 1, \mathcal{I}_2 = x')$, respectively. Furthermore, define:

$$U \equiv \begin{pmatrix} 1 & \xi_1^1 & \dots & \xi_B^1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \xi_1^{|\mathcal{A}|} & \dots & \xi_B^{|\mathcal{A}|} \end{pmatrix}, V \equiv \begin{pmatrix} 1 & \zeta_1^1 & \dots & \zeta_B^1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \zeta_1^{|\mathcal{A}|} & \dots & \zeta_B^{|\mathcal{A}|} \end{pmatrix}, Q \equiv \begin{pmatrix} 1 & \chi_1^* & \cdots & \chi_B^* \\ \chi_1 & \chi_{1,1} & \cdots & \chi_{1,B} \\ \vdots & \vdots & \ddots & \vdots \\ \chi_B & \chi_{1,B} & \cdots & \chi_{B,B} \end{pmatrix}.$$

Assumption 8 (a) *There exist a set of values for $\mathcal{I}_l$, denoted by $\{1, 2, \cdots, B\}$, such that $B > |\mathcal{A}|$ and both U and V are full rank; and* (b) *there exists some $x^* \in \{1, 2, \cdots, B\}$ such that $\zeta^j_{x^*} > 0$ and $\zeta^j_{x^*} \neq \zeta^{j'}_{x^*}$ for all $j, j' \in \{1, 2, ..., |\mathcal{A}|\}$ with $j \neq j'$.*

The full-rank condition Assumption 8(a) is analogous to that imposed in Kasahara and Shimotsu (2009) for identification. It requires $x_l \equiv \{s_l, \tilde{y}_l, G(s_l), H_l(s_l)\}$ to vary sufficiently in order to achieve identification. With x_l being continuous, Assumption 8(a) only requires there exists a finite set of values on its support $\mathcal{X}$ where the full-rank condition holds. This assumption is in line with Assumption 2 that requires sufficient variations of empirical evidence. Assumption 8 (b) is a condition of non-degeneracy: there exists some states, targeted outcomes, and empirical evidence under which committee members of different types make different decisions. It also implies that the initial perception and the empirical evidence cannot be too close to each other. To see this, consider the extreme case where $G(s_l) = H_l(s_l)$ for all $l \leq L$ Then the updated perception, as a weighted average of the two, would be the same for members of all types. For MPC in the Bank of England, G and H_l are forecasts by outsiders and MPC, respectively, and the difference between them is observed to be apparent.

As shown in Kasahara and Shimotsu (2009), the variation in $\mathcal{I}_l$ helps to recover the distribution of α_i, the cardinality of $\mathcal{A}$, and the choice probability $\Pr(D_{i,l} = d_{i,l} | \mathcal{I}_l; \alpha_i)$. The results are summarized in the following lemma.

Lemma 4 *(Application of Propositions 1 and 3 in Kasahara and Shimotsu (2009)) Suppose Assumptions 5, 6, 7 and 8 hold and $L \geq 3$. Then $|\mathcal{A}| = Rank(Q)$, and the probability mass function ρ_{α_i} and the stationary choice patterns $\Pr(D_{i,l} = 1 | x_l; \alpha_i)$ are identified for all $x_l \in \mathcal{X}$ up to unknown values of α_i.*

Proof of this lemma follows from the same arguments for Proposition 1 and 3 in Kasahara and Shimotsu (2009). A subtle difference is that in our setting we do not require members' ideological bias α_i to be fixed during their tenure in the committee, as stated in Assumption 5. Instead, we only require that α_i is known to be fixed within small intervals that partition the tenure of a committee member. This is essential for analyzing committees when the member compositions vary frequently (e.g., external members in MPC at the Bank of England often serve two to three years). Consequently, members tend to adjust their bias in processing multiple sources of information, due to deliberations with new members, etc.

The identifying power is from variations in the common information set, which imply different restrictions in various lower-dimensional submodels. For example, given the cardinality of the set of values considered in $\{1, 2, ., B\}$ and the number of episodes L, we would derive B^L restrictions in the submodels by integrating out subvectors in the history of common information.

Finally, it is worth noting that the full-rank and non-degeneracy in 7 are only required to hold for some realizations of $\mathcal{I}_l$ while the identification results apply to all realizations of the common information set on the support $\mathcal{X}$.

3.3 Ordering type-specific CCPs

It remains to order the type-specific CCPs according to realizations of α_i on the support $\mathcal{A}$). Our approach in this step differs qualitatively from that used in earlier papers that nonparametrically identify finite mixture models (e.g., in An, Hu, and Shum (2010) and An (2010)). In those papers, the conditions for ordering are either given exogenously by the underlying theoretical model or directly imposed as an assumption on component probabilities. In contrast, in this subsection, we provide conditions on model primitives which imply the monotonicity of type-specific CCPs in α_i.

Assumption 9 *(a) $W_{i,l}$ is independent from $\mathcal{I}_l$ with support $\mathcal{W} \equiv (0, \infty)^K$. (b) There exists $x^* \in \mathcal{X}$ such that there exists no $w \in \mathcal{W}$ so that members with $W_{i,l} = w$ are indifferent between the two alternatives regardless of α_i.*

Part (a) in Assumption 9 further strengthens part (a) in Assumption 5. Thus the private types α_i and $W_{i,l}$ are independent from each other, and are jointly independent from the common information set $\mathcal{I}_l$. Assumption 9 (b) requires that decisions made by the committee members of different ideological bias diverge enough. It is implied by more primitive restrictions on the model structure. (See Appendix C for details.)

Lemma 5 *Suppose Assumption 5 (c) holds. Let $\mathcal{I}_l = x^*$ satisfy Assumption 8 (b) and Assumption 9. Then $\Pr(D_{i,l} = 1|\mathcal{I}_l = x^*, \alpha_i)$ is monotonic in α_i, and the direction of monotonicity is identified.*

The main insight underlying this lemma is that the set of tastes $w_{i,l}$ causing a member to choose 1 changes monotonically as individual bias α_i increases. Independence between $W_{i,l}$, α_i and $\mathcal{I}_l$, together with the positive densities of $W_{i,l}$, imply type-specific CCPs must be monotonic in α_i. Furthermore, the direction of monotonicity is determined by the common information available in data.

We now sketch the intuition for Lemma 5 for the case with a two dimension outcome $K = 2$. It helps to visualize individual choices on the support of tastes $\mathcal{W}$. Below, 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. 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:

$$\Pr(D_i = 1|\mathcal{I} = x; \alpha_i) = \Pr\left\{\sum_k W_{i,k} C_k(x; \alpha_i) \le t(x)\right\} \tag{12}$$

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.

For any x, $\Pr(D_i = 1|\mathcal{I} = x; \alpha_i)$ is visualized in Figure 2 as the probability mass over the half-space $\{w_i : \sum_k w_{i,k} C_k(x; \alpha_i) \le t(x)\}$ in $\mathcal{W}$. With individual tastes independent from bias and the common information, the ordering of $\Pr(D_i = 1|\mathcal{I} = x; \alpha_i)$ in α_i depends on the slopes and intercepts of indifference hyperplanes $\sum_k W_{i,k} C_k(x; \alpha_i) = t(x)$. (These hyperplanes are collections of all members with α_i who are indifferent between the two alternatives under x.) By construction, these hyperplanes, indexed by α_i, either have the same slope or all intersect at the same point.

Part (b) of Assumption 8 requires that, given $\mathcal{I} = x^*$, changes in α_i affect how the support $\mathcal{W}$ is partitioned by the hyperplane (or a line in $\mathbb{R}^2$). Part (b) of Assumption 9 rules out uninteresting cases where some members with certain tastes (W_i) are indifferent between both alternatives *regardless of bias*. In other words, part (b) of Assumption 9 requires the intersection of bias-specific hyperplanes, if exists, must be outside the support of tastes $\mathcal{W}$. These two conditions, together with the independence between W_i and $\alpha_i, \mathcal{I}_l$, ensure changes in α_i always result in monotonic changes in the set of tastes in favor of an action, which in turn leads to monotonic changes in $\Pr(D_i = 1|\mathcal{I} = x^*; \alpha_i)$.

It remains to find out the direction of monotonicity in type-specific CCPs, which depends on how intercepts and slopes of indifference hyperplanes vary with α_i. The latter in turn are determined by ex ante deviations due to each one of the two sources of information alone (i.e. $C_k(x; 1)$ and $C_k(x; 0)$). The direction of monotonicity is identified since the intercepts and slopes are directly observed from the data.

To sum up, $\Pr(D_i = 1|\mathcal{I} = x^*; \alpha_i)$ is identified from Lemma 4 and is shown to be ordered in α_i by Lemma 5. If the specific values for the elements on the support $\mathcal{A}$ are known, it would then follow from Lemma 4 that the probability masses ρ_{α_i}, or weights in the finite mixture (11), are identified.

3.4 Identifying the distribution of W_i

Having identified the type-specific CCPs $\Pr(D_i = d_i | \mathcal{I} = x; \alpha_i)$ for all x and bias α_i, we next recover the joint distribution of W_i. As before, we drop subscripts l from $\mathcal{I}_l$ and $W_{i,l,k}$, and focus our analysis on the case with $K = 2$ for simplicity. To attain identification, we need to introduce additional structural and technical conditions. For any x on the support of $\mathcal{I}$ and any $\gamma \in \mathbb{R}^1_+$ and $v \in \mathbb{R}^1$, define:

$$\Lambda(x, v, \gamma; \alpha_i) \equiv \tfrac{t(x) - C_2(x;\alpha_i)v - C_1(x;\alpha_i)\gamma - C_2(x;\alpha_i)\gamma}{C_1(x;\alpha_i)}.$$

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. $f_{\eta_k}(\cdot)$. (b) Define:*

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

where $\mathbb{I}(\gamma, v) \equiv \mathbb{I}\left\{\tfrac{t(x) - C_2(x;\alpha_i)v - C_1\overline{\eta}_1}{C_1(x;\alpha_i) + C_2(x;\alpha_i)} \leq \gamma \leq \tfrac{t(x) - C_2(x;\alpha_i)v - C_1\underline{\eta}_1}{C_1(x;\alpha_i) + C_2(x;\alpha_i)}\right\}$. Then for some $\alpha \in \mathcal{A}$, the function $K(\cdot, \cdot; \alpha)$ has the following property: "If δ is a bounded function with domain $\mathcal{W}_0$ and $\int_{\mathcal{W}_0} K(w, x; \alpha)\delta(w) dw = 0$ for all x on the support of $\mathcal{I}$, then $\delta(w) = 0$ almost everywhere over $\mathcal{W}_0$." (c) The support for ideological bias $\mathcal{A}$ is known.

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 bounded 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 common condition in nonparametric identification of structural models, e.g., see Andrews (2011). Recently, Hu and Shiu (2012) provide sufficient conditions under which a conditional density is complete. Part (c) is necessary for subsequent identification arguments. Since the cardinality of $\mathcal{A}$ is identified above, part (c) could be replaced by some weaker shape restrictions, such as that require elements in $\mathcal{A}$ to be equally spaced. With knowledge of $\mathcal{A}$, the results from the previous two subsections imply the probability mass function of α_i is fully identified.

Proposition 3 *Suppose Assumptions 5, 6, 7, 8, 9, and 10 hold. The distribution of $W_{i,0}$ is identified.*

The intuition for this result is as follows. By changing variables between $\eta_{i,1}$ and the argument for the C.D.F. of $W_{i,0}$ in the definition of the type-specific CCPs of (12), we can recast this definition in the form of an integral equation:

$$\varphi(x; \alpha_i) = \int_{-\infty}^{\infty} F_{W_0}(\gamma) K(\gamma, x; \alpha_i) d\gamma, \tag{13}$$

where $\varphi(x;\alpha_i) \equiv -\frac{\Pr(D_i=1|\mathcal{I}=x;\alpha_i)C_1(x;\alpha_i)}{C_1(x;\alpha_i)+C_2(x;\alpha_i)}$ and F_{W_0} is the C.D.F. for $W_{i,0}$ (see Appendix A for details). The completeness condition in part (b) then guarantees that there exists a unique solution of $F_{W_0}(\cdot)$ to (13).

3.5 Discussions about estimation

Estimation of the distributions of α_i and $W_{i,0}$ in the model with strategic recommendations takes several steps. We first estimate the cardinality of $\mathcal{A}$ as the rank of the matrix Q defined in Section 3.2. We then estimate the probability masses of α_i and the distribution $F_{W_0}(\cdot)$ using a sieve maximum likelihood estimator (MLE).

Lemma 4 states $|\mathcal{A}| = rank(Q)$. This rank can be estimated using a sequential estimator (Robin and Smith (2000)), which is constructed from a sequential test of the null hypotheses H_r: rank(Q)= r against the alternative hypothesis H_r': rank(Q)> r, $r = 0, 1, \cdots, |\mathcal{A}| - 1$. The estimator $\hat{r}$ can be explicitly defined as $\hat{r} \equiv \min_{r\in\{0,1,\cdots,r-1\}}\{r : H_i \text{ is rejected}, i = 0, 1, \cdots, r - 1; H_r' \text{ is not rejected}\}$. 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 sample size appropriately, it can be shown (Theorem 5.2 in Robin and Smith (2000)) that the rank of Q can be consistently estimated.

Having estimated the cardinality of $\mathcal{A}$, we employ a two-step procedure to estimate the probability masses of α_i and the distribution of $W_{i,0}$. First, estimate the equilibrium probability that the committee decision is d given committee i's decision is d_i, i.e. p^*_{d,d_i}. Then estimate the distributions of α_i and W_0 by a sieve MLE. For any $x \in \mathcal{X}$, the objective $p^*_{1,1}$ is estimated as:

$$\widehat{p}_{1,1}(x) \equiv \widehat{p}(d_l = 1|d_{i,l} = 1, x) = \frac{\widehat{g}(d_l = 1, d_{i,l} = 1, x)}{\widehat{g}(d_{i,l} = 1, x)}.$$

$\widehat{g}(d_l = 1, d_{i,l} = 1, x)$ and $\widehat{g}(d_{i,l} = 1, x)$ are kernel estimators. Also, $p^*_{0,0}(x)$ is estimated by $\widehat{p}_{0,0}(x)$ in a similar fashion. The uniform consistency of $\widehat{p}_{d,d_i}(x)$ under some regularity conditions follows from standard arguments, such as in Fan and Yao (2005). Consequently, $t(x)$ and $C_k(x), k = 1, 2$ can also be consistently estimated using $\widehat{p}_{d,d_i}(x)$. Second, estimate the distribution of α_i and $W_{i,0}$ jointly using sieves MLE and the first-stage estimates. Let $|\mathcal{C}|$ be the number of members in a committee $\mathcal{C}$ (which, W.L.O.G., is fixed across cross-sectional units), $\mathcal{P} \equiv \{p \in [0,1]^{|\mathcal{A}|} : \sum_{j\le|\mathcal{A}|} p_j = 1\}$ be the parameter space for the probability mass function for ideological bias, and 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

$$\widehat{\theta} \equiv (\widehat{p}, \widehat{F}) = \arg \max_{p\in\mathcal{P}, F\in\mathcal{F}_n} \sum_{n=1}^{N} \sum_{i=1}^{|\mathcal{C}|} \log \widehat{\mathcal{L}}_{i,n}(p, F) \tag{14}$$

where i indicates committee member, n indexes the cross-sectional units of independent committees and N is the sample size; and $\widehat{\mathcal{L}}_{i,n}$ is the estimated likelihood based on the conditional distribution of individual and committee decisions given the common information throughout L episodes observed for committee n.

$$\widehat{\mathcal{L}}_{i,n}(p,F) = \sum_{j=1}^{|\mathcal{A}|} p_j \left\{ \prod_{l=1}^{L} \left[\int_0^\infty F(\gamma)\, \widetilde{K}(\gamma, x_{n,l}; \alpha^j) d\gamma \right]^{d_{i,l}} \left[1 - \int_0^\infty F(\gamma)\, \widetilde{K}(\gamma, x_{n,l}; \alpha^j) d\gamma \right]^{1-d_{i,l}} \right\},$$

where

$$\widetilde{K}(\gamma, x_{n,l}; \alpha^j) \equiv -\frac{\widehat{C}_1(x_{n,l};\alpha^j) + \widehat{C}_2(x_{n,l};\alpha^j)}{\widehat{C}_1(x_{n,l};\alpha^j)} \widehat{K}(\gamma, x_{n,l}; \alpha^j),$$

and $\widehat{K}(\cdot)$ is the estimator of $K(\cdot)$ for given known form of $f_{\eta_k}(\cdot), k = 1,2$ as well as the estimates $\widehat{t}(\cdot)$ and $\widehat{C}_k(\cdot)$ from the first step.

Similar to the estimation of expressive recommendations, conditions for consistency of the 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). Our estimator $\widehat{\theta}$ is consistent, as long as the objective function in (14), despite preliminary estimation errors in $\widehat{t}(\cdot)$ and $\widehat{C}_k(\cdot)$, satisfies the conditions listed in those references. We leave a formal proof and the derivation of additional technical conditions needed for consistency to future research. Instead, we adopt a parametric approach for estimation in the empirical section below. Finally, note our estimator can be generalized to be conditional on individual heterogeneities that are reported in data.

4 Empirical Application

We illustrate our methodology by analyzing the decisions of the Bank of England's Monetary Policy Committee (MPC). The committee is made up of nine members. Five of them are internal members, who hold full-time executive positions in the Bank; and the other four are external who have no executive responsibilities and mostly work part time. The committee meets every month to vote for an interest rate. Individual recommendations by members of MPC are aggregated using a simple majority rule.

A growing literature on MPC focus on explaining the discrepancies in the recommendations from different members by their observed characteristics. However, it has been shown that these observed individual characteristics of the members are not sufficient for explaining difference in their recommendation patterns.[11]

We estimate a structural model of MPC decisions, which rationalizes heterogenous recommendations within MPC by the members' idiosyncratic bias and tastes, as well as strategic

[11] Besley, Meads, and Surico (2008) and Harris, Levine, and Spencer (2011) demonstrate that neither the type of membership nor background characteristics fully explain the difference in voting pattern.

career concerns. To the best of our knowledge, this marks the first effort to analyze MPC decisions by explicitly modeling its members' unobserved tastes and ideological bias. That said, the main goal of this section is to provide an illustrating example for our method, rather than a thorough analysis of MPC decisions as such.

4.1 The data and empirical specification

The data we use are compiled from several sources: (1) Individual committee members' voting records are collected from the publicly available "Minutes of MPC Meetings" from the Bank of England, which contain 187 monthly meetings from June 1997 to June 2013 involving 32 committee members (13 internal and 19 external). Each committee member is observed with his votes and membership types (external or internal). In most of the voting periods, committee members vote for two different rates (with only 8 exceptions out of 187). Therefore we transform members' choices to a binary variable $d_{i,l} \in \{0, 1\}$, with 0 being "choose a lower interest rate" and 1 being "vote for a higher interest rate". (2) The quarterly forecasts of MPC and outsiders on inflation and output, which are published in the Bank of England's quarterly *Inflation Report* starting from August 1997. The Bank of England carries out a survey in each quarter just before an *Inflation Report* which presents the summary of the survey. The survey asks a group of external forecasters (other than MPC members) from financial institutions mainly based in the City of London for their prediction of inflation, output growth, etc., under various interest rates. Because the institutions that participate in the survey are prominent and the sample size is fairly large, the outsiders' forecasts can be taken as a good measure of conventional wisdom. Following the literature, we use the forecasts formulated on an assumption of a constant interest rate. (3) Historical data of inflation rates (monthly) and GDP (quarterly) are from Office for National Statistics (ONS).

Our sample contains the data from the three sources above between August 1998 and June 2013,[12] which accounts for 60 quarters or 180 months (include the emergency MPC meeting in September 2001). Table 1 presents summary statistics for the voting records and membership of committee members.

The purpose of the monthly MPC meeting is to set an interest rate to achieve a two-dimensional target: inflation and GDP, i.e., $(\widetilde{y}_l, \widetilde{\pi}_l)$, where $\widetilde{y}_l$ is the targeted GDP growth rate while $\widetilde{\pi}_l$ is the targeted inflation rate. The GDP target $\widetilde{y}_l$ is measured as the potential GDP growth rate using Hodrick-Prescott filter (with a smoothing parameter set to be 1600) from the Bank of England's vintage data of GDP.[13] The inflation target $\widetilde{\pi}_l$ is 2.5% up to December 2003 and 2.0% after that.[14] Similarly, the state s_l contains the current inflation rate

[12] August 1998 is the first month that the complete forecasts of both MPC and outsiders are available.

[13] See Robert and Prescott (1980) for details of the method.

[14] The targeted inflation rate was moved from 2.5 to 2 percent in January 2004 when the targeted inflation

and growth rate of GDP. The empirical evidence $H_l(y,\pi|s,d)$ and MPC's initial perception $G(y,\pi|s,d)$ are constructed from the forecasts of MPC and outsiders, 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 targets $(\tilde{y},\tilde{\pi})$, state $s=(y,\pi)$, empirical evidence H and initial perception G, a member chooses an interest rate by solving the problem in equation (8) and his decision is determined by the individual taste $W_{i,1},W_{i,2}$ and ideological bias α_i, where $W_{i,1}$ and $W_{i,2}$ capture the weights the member imposes on GDP and inflation relative to career concerns, respectively. We assume that a member's ideological bias is unchanged in a quarter and independent across quarters, while individual tastes are i.i.d. across months. Let $W_{i,k}=W_{i,0}+\eta_{i,k}$ for $k=1,2$, where $\eta_{i,k}$ are i.i.d. draws from truncated standard normal distribution between $[-2,2]$ for $k=1,2$. The distribution of $W_{i,0}$ depends on the type of membership (E_i=1 if the member is external and $E_i=0$ otherwise) and is specified as gamma distribution with parameter (a^e,b^e) for $e\in\{0,1\}$. The support of individual bias α_i, $\mathcal{A}$ is specified as $\{\alpha^1,\alpha^2,\alpha^3\}=\{1/4,1/2,3/4\}$ to capture how members balance MPC and outsiders' forecasts. The distribution of α_i also depends on E_i, and its probability masses are denoted as $\Pr(\alpha_i=\alpha^j|E_i=e)=p_{j,e}$ for $1\leq j\leq 3$ and $e=0,1$.

These parametric specifications satisfy some main identifying conditions in Section 3. Assumptions 5 holds because the ideological bias is assumed to be fixed in a quarter and independent across quarters, and the tastes are i.i.d. across individuals and months. Assumption 6 holds due to the institutional fact that forecasts by outsiders are available quarterly while those by MPC members are updated monthly. We verify Assumption 9 in the context of parametric specification in the next subsection.

4.2 Estimation and results

We maintain that committee members' recommendations are rationalized as solutions to (8) in each observation.[15] Our objective is to estimate the parameters in the distribution of committee members' heterogenous tastes and ideological bias. The estimation is based on equation (13) and completed in two steps: first, the difference of probabilities $t(x)\equiv p^*_{1,1}(x)-p^*_{0,0}(x)$ is parametrized and estimated, then the probability masses of α_i and the distribution F_{W_0} are estimated in the second step.

We take a different approach from the estimation strategies in previous section to estimate $t(x)$. For this purpose, we first express $p^*_{d,d_i}(x)$ as a function of individual choice probabilities for both external and internal members, denoted by $\lambda_{Int}(x)\equiv\Pr(d_i=1|E_i=0,x)$ and

index was changed from the retail price index excluding mortgage interest payments (RPIX) to the index of consumer prices (CPI).

[15]Our model of MPC's objective function is in line with the literature. For example, in Geraats (2009), policymakers' objective is to minimize the loss function $W_{i,t}=-\frac{\alpha_i}{2}(\pi_t-\pi^*)^2-\frac{1-\alpha_i}{2}(y_t-y^*)^2$, where α_i is an individual's weight put on inflation.

$\lambda_{Ext}(x) \equiv \Pr(d_i = 1|E_i = 1, x)$, respectively.[16] We adopt a logit form for the individual choice probabilities for both external and internal members with parameters θ_j and $\vartheta_j (j = 0, 1, 2, 3, 4)$:

$$\begin{aligned} \lambda_{Ext}(x) &\equiv \Pr(d_i = 1|E_i = 1, x) = \frac{\exp\{\theta_0 + \sum_{j=1}^{4} \theta_j \omega_j\}}{1 + \exp\{\theta_0 + \sum_{j=1}^{4} \theta_j \omega_j\}}, \\ \lambda_{Int}(x) &\equiv \Pr(d_i = 1|E_i = 0, x) = \frac{\exp\{\vartheta_0 + \sum_{j=1}^{4} \vartheta_j \omega_j\}}{1 + \exp\{\vartheta_0 + \sum_{j=1}^{4} \vartheta_j \omega_j\}}, \end{aligned}$$

where $\omega_j (j = 1, 2, 3, 4)$ is directly recovered from data and it describes how committee members aggregate the two sources of information H and G, targets, and possible outcomes to solve their optimization problem.

$$\omega_j \equiv \begin{cases} \sum_{\{q: y^q \neq \tilde{y}\}} a_{q,k}(\tilde{y}) \big[H_{q,1}(s) - H_{q,0}(s) \big] \text{ for } j = k, \\ \sum_{\{q: y^q \neq \tilde{y}\}} a_{q,k}(\tilde{y}) \big[G_{q,1}(s) - G_{q,0}(s) \big] \text{ for } j = k + 2, k = 1, 2. \end{cases}$$

The validity of Assumption 9 is verified by our data. To do so, we estimate $C_k(\cdot, \cdot)$ and $t(\cdot)$ using data on voting records and states as above. Then for all common information set x in data, we solve for the solutions $(w_{i,1}, w_{i,2})$ for the system of equations: $w_{i,k} C_k(x; \alpha_i) = t(x)$ for all $\alpha_i \in \mathcal{A}$. The solutions, which are based on preliminary estimates of $C_k(\cdot, \cdot)$ and $t(\cdot)$, are sufficiently bounded away from the first quadrant in $\mathbb{R}^2$ with absolute values large relative to the standard errors. As is shown in Figure 2 in the appendix, this suggests our data satisfies Assumption 9.

Table 2 summarizes the estimated parameters, with the standard errors obtained through bootstrap resampling. (We draw $B = 200$ bootstrap samples from the data, and the standard errors in the table are reported as the empirical standard deviation from the 200 estimates calculated from the bootstrap samples.) The results reveal the pattern of members' using information available for their decisions. Overall, there is evidence that external and internal members process various sources of information in a similar manner. Especially, the outsiders' forecasts of GDP play a less important role in decision-making for both types of members (θ_4 and ϑ_4 are both insignificant). On the other hand, external members care more about inflation than external members (ϑ_1 is significantly larger than θ_1).

Based on the results of $t(x)$, we estimate the parameters of committee members' tastes and ideological bias in the second step using MLE. Let $\theta \equiv \{(a^e, b^e)_e, (p_{j,e}, p_{j,e})_{j,e}\}$, $e = 0, 1; j = 1, 2, 3$, $x_t = \{s_t, \tilde{y}_t, G_t(s_t), H_t(s_t)\}$ denote the information set available at episode t as before. The log-likelihood function is given by

$$\mathcal{L} = \sum_{i=1}^{9} \sum_{\tilde{t}=1}^{60} \log L_{i,\tilde{t}}(\theta),$$

[16] For example, for each internal member ($E_i = 0$),

$$p_{1,1}^{Int}(x) \equiv \Pr(D^* = 1|D_i = 1, E_i = 0, x) = \Pr\{\text{at least 4 out of eight others choose 1} \mid E_i = 0, x\},$$

and this probability can be easily expressed as a function of individual choice probabilities.

where the summation indexed by i and $\tilde{t}$ are w.r.t. individuals and quarters, respectively. The function $L_{i,\tilde{t}}(\cdot)$ is defined as

$$L_{i,\tilde{t}}(\theta) = \sum_j p_{j,e_i} \left\{ \prod_{t=3(\tilde{t}-1)+1}^{3\tilde{t}} \left[\int_0^\infty \Phi_{j,e_i}(\gamma, x_t;\theta) d\gamma \right]^{d_{i,t}} \left\{ 1 - \left[\int_0^\infty \Phi_{j,e_i}(\gamma, x_t;\theta) d\gamma \right] \right\}^{1-d_{i,t}} \right\},$$

where

$$\Phi_{j,e_i}(\gamma, x_t;\theta) \equiv F_{W_0}(\gamma; a^{e_i}, b^{e_i})\, \widetilde{K}_{e_i}(\gamma, x_t; \alpha^j)$$

with $\widetilde{K}_{e_i}(\gamma, x; \alpha^j) \equiv -\frac{C_1(x;\alpha_i)+C_2(x;\alpha_i)}{C_1(x;\alpha_i)} K_{e_i}(\gamma, x; \alpha^j)$, and $K_{e_i}(\gamma, x; \alpha^j)$ is defined as in the previous subsection and depends on membership. A computational difficulty arises because it is impractical to evaluate $L_{i,\tilde{t}}(\theta)$ analytically for a given θ. Hence we first use Monte Carlo simulation to get an estimate of $\Phi_{j,e_i}(\gamma, x_t;\theta)$,

$$\widehat{\Phi}_{j,e_i}(x_t;\theta) \equiv \frac{1}{M} \sum_{\tilde{m}=1}^{M} \Phi_{j,e_i}(\gamma_{\tilde{m}}, x_t;\theta),$$

where $(\gamma_{\tilde{m}})_{\tilde{m}\leq M}$ are chosen grid points on the support of f_{η_1} and f_{η_2}. Our estimator is thus expressed as

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

where

$$\widehat{L}_{i,\tilde{t}}(\theta) \equiv \sum_j p_{j,e_i} \left\{ \prod_{t=3(\tilde{t}-1)+1}^{3\tilde{t}} \left\{ d_{i,t} \left[\widehat{\Phi}_{j,e_i}(x_t;\theta) \right] + (1-d_{i,t}) \left[1 - \widehat{\Phi}_{j,e_i}(x_t;\theta) \right] \right\} \right\}. \tag{15}$$

Tables 3 and 4 summarize the bootstrap estimates of the sampling distribution of our sieve MLE for parameters in the distributions of ideological bias and individual tastes. They are both based on the empirical distribution of sieve MLE estimates from $B = 200$ bootstrap samples. Table 3 reports the estimated sampling distribution of our estimator for the probability mass of α_i. With higher probabilities, both types of members put more weights on MPC forecasts than on outsiders' forecasts. That is, $\Pr(\alpha = 75\%) > \max\{\Pr(\alpha = 25\%), \Pr(\alpha = 50\%)\}$ for both types of members. Besides, there is little chance that the members weigh both evidence equally (i.e. $\Pr(\alpha = 50\%)$ is not statistically significant for both types of members). In addition, the results imply that internal members focus slightly more on MPC's forecasts than external members and this is consistent with the estimates of the logit model in the first step.

Table 4 reports the estimated sampling distribution of our estimator for the parameters in the distribution of F_{W_i}. To better interpret the estimated distribution of committee members' tastes, we plot the point estimate for the CDF of tastes for both types using the estimated parameters in Figure 3. Recall that $W_{i,k}, k = 1, 2$ describes the importance of GDP and inflation relative to career concerns in members' decisions. The point estimates in Figure 3 are consistent with a hypothesis that the external members' recommendations are less distorted by strategic incentives than internal members. In other words, internal members care more about the chance that their recommendations conform to the final committee decisions. This new finding in MPC's decisions illustrates that the existence of external members helps to alleviate possible distortions in committee decisions due to strategic incentives for conformity. The results also explain that the heterogeneous recommendation patterns of different types of members are mostly ascribed to the difference in how they weigh multiple dimensions in the target, rather than the difference in how they process various sources of information.

5 Concluding Remarks

We study the identification and estimation of a structural model 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 dimensions in the target outcome (tastes). A motivation for the model is the need to explain why committee members who share the same target and information could end up making different recommendations.

Our model also allows for interactions between members with strategic concerns, such as conformity to the group decision. We show how to nonparametrically recover the distributions of members' private types from members' decisions and sources of common information used in decisions. The identification arguments differ qualitatively for cases with and without strategic concerns. An empirical analysis of MPC decisions at the Bank of England suggests the heterogeneous patterns of recommendations among different types of committee members are ascribed more to heterogeneous tastes for multi-dimensional objectives than to the difference in bias towards various sources of information.

There are a few directions for future research. First, we investigate the models with expressive and strategic recommendations, while leaving out a test of one model against the other. It will be interesting to test which model better describes committee members' behavior. Second, decision makers in our model have a deterministic objective function, which describes their view on how their decisions affect the outcomes, while allowing committee members have multiple views enables us to further investigate committee decision-making in the framework of "robust decision-making". Third, our empirical analysis of MPC focus on the explanation of heterogeneous voting patterns, while more detailed studies will provide useful policy implications. For example, our framework allows us to investigate the effects of committee size, composition of the committee, whether the committee is transparent or secretive, etc.

References

Ai, C., and X. Chen (2003): "Efficient estimation of models with conditional moment restrictions containing unknown functions," *Econometrica*, 71(6), 1795–1843. 2.3.5, 3.5

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

An, Y., Y. Hu, and M. Shum (2010): "Estimating First-Price Auctions with an Unknown Number of Bidders: A Misclassification Approach," *Journal of Econometrics*, 157, 328–341. 3.3

Andrews, D. (2011): "Examples of L2-Complete and Boundedly-Complete Distributions," . 3.4

Bajari, P., H. Hong, J. Krainer, and D. Nekipelov (2010): "Estimating static models of strategic interactions," *Journal of Business & Economic Statistics*, 28(4). 3.1

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

Besley, T., N. Meads, and P. Surico (2008): "Insiders versus outsiders in monetary policymaking," *American Economic Review*, 98(2), 218–223. 2, 11

Chen, X., H. Hong, and D. Nekipelov (2011): "Nonlinear models of measurement errors," *Journal of Economic Literature*, 49(4), 901–937. 2

Chen, X., and X. Shen (1998): "Sieve Extremum Estimates for Weakly Dependent Data," *Econometrica*, 66, 289–314. 2.3.5, 3.5

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

Fan, J., and Q. Yao (2005): *Nonlinear Time Series: Nonparametric and Parametric Methods*. Springer. 3.5

Geraats, P. M. (2009): "Trends in Monetary Policy Transparency," *International Finance*, 12(2), 235–268. 15

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

Hao, L., and W. Suen (2009): "Viewpoint: Decision-making in committees," *Canadian Journal of Economics/Revue canadienne d'économique*, 42(2), 359–392. 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. 11

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

Hu, Y., and S. M. Schennach (2008): "Instrumental variable treatment of nonclassical measurement error models," *Econometrica*, 76(1), 195–216. 1, 3.2

Hu, Y., and J.-L. Shiu (2012): "Nonparametric Identification using instrumental variables: sufficient conditions for completeness," Johns Hopkins Working Paper 581. 3.4

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, 10, 3.2, 3.2, 3.2, 4, 3.2

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. 3.1

Merlo, A., and A. De Paula (2010): "Identification and estimation of preference distributions when voters are ideological," PIER Working Paper. 1

Robert, H., and E. Prescott (1980): "Post-War US Business Cycles: An Empirical Investigation," *Journal of Money, Credit and Banking*, 29(1), 1–16. 13

Robin, J.-M., and R. J. Smith (2000): "Tests of rank," *Econometric Theory*, 16(02), 151–175. 3.5

Shen, X. (1997): "On methods of sieves and penalization," *Annals of Statistics*, 25(6), 2555–2591. 2.3.5, 3.5

Appendix

A Proofs

Proof of Lemma 1. First, we characterize the best response of a member i who has types (α_i, w_i) when 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 notations throughout the proof. Let p_{d',d_i} be i's belief about the chance that his decision is adopted by the committee. That is,

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

where $D^*(\mathbf{d}) \equiv \arg\max_{d^* \in \{0,1\}} \sum_i 1(d_i = d^*)$ denotes the majority rule. Then i's objective function is:

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

where $\mathcal{F}_{q,d_i}(\alpha_i) \equiv \alpha_i H_{q,d_i} + (1-\alpha_i) G_{q,d_i}$, and the dependence on s is suppressed in G_{q,d_i} and H_{q,d_i}. By removing the terms that do not depend on d_i, member i's optimization problem is equivalent to:

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

Thus member i would choose $d_i = 1$ if and only if

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

But note that $p_{1,1} > p_{1,0}$ regardless of other members strategies σ_{-i}, which follows from the definition of $p_{d',d_i}(\sigma_{-i})$ in (A.1) and the fact that $\{\mathbf{d}_{-i} : D^*(1, \mathbf{d}_{-i}) = 1\}$ must be a strict subset of $\{\mathbf{d}_{-i} : D^*(0, \mathbf{d}_{-i}) = 1\}$. Hence i would choose $d_i = 1$ 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}. □

Proof of Lemma 2. 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:

$$\begin{aligned} \lambda_1 &\equiv \sup\{\lambda : \Pr\{D_i = 1 \mid h = (1-\lambda)h^a + \lambda h^b\} = 0\}; \text{ and} \\ \lambda_4 &\equiv \inf\{\lambda : \Pr\{D_i = 1 \mid h = (1-\lambda)h^a + \lambda h^b\} = 1\}. \end{aligned} \quad \text{(A.2)}$$

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 2, $\lambda_1 = \underline{\lambda}_2$ and $\lambda_4 = \underline{\lambda}_1$.

First off, discussions in the second and the third paragraphs in Section 2.3.2 suggest immediately that cases (i), (iv) can be distinguished from the cases (ii), (iii), (v) and (vi). To see why, note 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 (ii), (iii), (v) and (vi), 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.2); 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 panel (ii), $\lambda_2 = \overline{\lambda}_2$ and $\lambda_3 = \overline{\lambda}_1$; in panel (iii), $\lambda_2 = \underline{\lambda}_2$ and $\lambda_3 = \overline{\lambda}_1$.

The unmatched thresholds as $\{\lambda_j(h^a, h^b) : 1 \le j \le 4\}$ are functions of extreme evidence (h^a, h^b) conditioned on. Let $\delta_{j,j'}(h^a, h^b)$ denote the distance between $\lambda_j(h^a, h^b)$ and $\lambda_{j'}(h^a, h^b)$, 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. In what follows, we suppress dependence of $\delta_{j,j'}$ and $\delta'_{j,j'}$ on (h^a, h^b) for simplification. It can be shown from Figure 1 that: in case (i), $\delta'_{1,4} > 0$; in case (ii), $\delta'_{2,3} = 0$ while $\delta'_{1,2} < 0$; in case (iii), $\delta'_{2,3} < 0$; in case (iv), $\delta'_{1,4} < 0$; in case (v), $\delta'_{2,3} = 0$ while $\delta'_{1,2} > 0$; in case (vi), $\delta'_{2,3} > 0$. Hence all six cases can be distinguished from each other using these identifiable partial derivatives. □

Proof of Proposition 1. Fix the triple $\{s, \tilde{y}, G(s)\}$ and the pair of evidence (h^a, h^b). 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, $c(\underline{w}; h(\underline{\lambda}_2)) = r^2$. 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 λ_2, or the infimum of the strict sub-interval of $[\lambda_1, \lambda_4]$ over which CCPs remain constant. 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, $\Pr\{D_i = 1|h\} = \sum_{j=1,2} \Pr\{c(W, h) \le r^j\} \Pr\{R_i = r^j\}$ for all h. In case (ii), $\Pr\{c(W; h(\lambda_2)) \le r^1\} = 0$ and $\Pr\{c(W; h(\lambda_2)) \le r^2\} = 1$. Hence $\Pr\{\alpha_i = \alpha^2\}$, or equivalently $\Pr\{R_i = r^2\}$, is identified as $\Pr\{D_i = 1|\ h(\lambda_2)\}$.

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 conditioned state s. The full-rank condition in Assumption 3 implies $\frac{a_{1,1}}{a_{2,1}} \ne \frac{a_{1,2}}{a_{2,2}}$ with both sides nonzero, which in turn implies $\frac{a_{1,1}+\underline{w}a_{1,2}}{a_{2,1}+\underline{w}a_{2,2}} \ne \frac{a_{1,1}+\overline{w}a_{1,2}}{a_{2,1}+\overline{w}a_{2,2}}$ 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:

$$\Pr\{c(W, h(\lambda)) \le r^2\} = \frac{\Pr\{D_i = 1 | h(\lambda)\}}{\Pr\{\alpha_i = \alpha^2\}}. \tag{A.3}$$

Under Assumption 3, it can be shown that 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

$$\frac{a_{1,1} h_1(\lambda) + a_{2,1} h_2(\lambda)}{a_{1,1} g_1 + a_{2,1} g_2} - \frac{a_{1,2} h_1(\lambda) + a_{2,2} h_2(\lambda)}{a_{1,2} g_1 + a_{2,2} g_2} \tag{A.4}$$

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.3) is either $\Pr\{W_i \le c^{-1}(r^2, h(\lambda))\}$ or $\Pr\{W \ge c^{-1}(r^2, h(\lambda))\}$, depending on the sign of the difference in (A.4), 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. □

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, (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 c.d.f.s 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

$$\begin{aligned} &\tilde{F}(w) = \tfrac{q}{p} F(w), \ \ \forall w \in [\overline{w}_1, \overline{w}]; \text{ and} \\ &1 - \tilde{F}(w) = \tfrac{1-q}{1-p}[1 - F(w)], \ \ \forall w \in [\underline{w}, \underline{w}_1]. \end{aligned} \tag{A.5}$$

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.5) 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.5) 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. □

Proof of Lemma 5. 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 in the text. 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 8 (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.[17] 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 8 (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 9 (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 8(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 8(b) and 9(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. □

[17] 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 9(b) already.

Proof of Equation (13). 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 that the expression of choice probability of alternative $D_i = 1$ in (12) is

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

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

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

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 only consider the case where $C_1(x;\alpha_i) + C_2(x;\alpha_i) > 0$ and $C_1(x;\alpha_i) > 0$ (as the other cases can be analyzed by symmetric argument). In this case,

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

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

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

$$\Lambda(x,v,\gamma;\alpha_i) \equiv \frac{t(x) - C_2(x;\alpha_i)v - C_1(x;\alpha_i)\gamma - C_2(x;\alpha_i)\gamma}{C_1(x;\alpha_i)}.$$

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

$$\Phi(x;\alpha_i) = -\left(\tfrac{C_1(x;\alpha_i)+C_2(x;\alpha_i)}{C_1(x;\alpha_i)}\right)\int_{-\infty}^{\infty} F_{W_0}(\gamma)\left[\int_{\underline{\eta}_2}^{\overline{\eta}_2} \mathbb{I}(\gamma,v) f_{\eta_1}(\Lambda(x,v,\gamma;\alpha_i)) f_{\eta_2}(v)dv\right]d\gamma.$$

The integral equation above can be succinctly summarized as:

$$\varphi(x;\alpha_i) = \int_{-\infty}^{\infty} F_{W_0}(\gamma)\,K(\gamma,x;\alpha_i)d\gamma \tag{A.6}$$

where

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

The completeness of $K(\gamma, x; \alpha_i)$ guarantees that $F_{W_0}(\cdot)$ is the unique solution to the the integral equation (13), i.e., $F_{W_0}(\cdot)$ is nonparametrically identified. □

B Generalization of Proposition 2

We extend the proof of Proposition 2 by relaxing Assumption 4. To do this, we first define

$$\underline{w}_t \equiv \phi_1^{-1}(\phi_2(\underline{w}_{t-1})) \text{ and } \overline{w}_t \equiv \phi_2^{-1}(\phi_1(\overline{w}_{t-1})).$$

We assume alternatively,

Assumption B.1. *For $\{s, \tilde{y}, g(s)\}$ conditioned on and (h^a, h^b) considered, $\underline{w}_1 < \overline{w}_1 < \underline{w}_2$.*

It is sufficient to prove the identification results for the two cases (i) and (iv) in Figure 1 since the proof for the four cases (ii), (iii), (v) and (vi) is unchanged. Here we follow the similar procedure to seek contradictions as in the proof of Proposition 2. The equivalence of the two sets of primitives (q, F) and $(p, \tilde{F})$ implies the following relationship:

$$\begin{aligned} qF(w_1) &= p\tilde{F}(w_1), w_1 \in [\overline{w}_1, \overline{w}_1], \\ qF(w_2) + (1-q)F(w_1) &= p\tilde{F}(w_2) + (1-p)\tilde{F}(w_1), w_2 \in [\underline{w}_1, \underline{w}_2], w_1 \in [\overline{w}_1, \overline{w}_1], \\ q + (1-q)F(w_2) &= p + (1-p)\tilde{F}(w_2), w_2 \in [\underline{w}_1, \underline{w}_2], \end{aligned}$$

where the second equality holds due to Assumption B.1 which guarantees the existence of a non-degenerate and non-empty open interval $(\overline{w}_1, \underline{w}_2)$. The above observational equivalence can be further simplified as,

$$\begin{aligned} \tilde{F}(w_1) &= \frac{q}{p}F(w_1), w_1 \in [\overline{w}_1, \overline{w}_1], \\ \tilde{F}(w_2) &= \frac{q}{p}F(w_2) + \frac{F(w_1) - \tilde{F}(w_1)}{p}, w_2 \in [\underline{w}_1, \underline{w}_2], \\ \overline{\tilde{F}}(w_2) &= \frac{1-q}{1-p}\overline{F}(w_2), w_2 \in [\underline{w}_1, \underline{w}_2], \end{aligned}$$

where the second equality is derived using the relationship $qF(w_1) - p\tilde{F}(w_1) = 0$ for all $w_1 \in [\overline{w}_1, \overline{w}_1]$. Notice that $\underline{w_1}$ changes monotonically as we move w_2, therefore we denote w_1 by $w_1(w_2)$. Both $\tilde{F}(w_2) + \overline{\tilde{F}}(w_2) = 1$ and $F(w_2) + \overline{F}(w_2) = 1$ need to hold on the interval $(\overline{w}_1, \underline{w}_2)$ in order for $\tilde{F}$ and F to be well-defined C.D.Fs.

We express explicitly the restriction imposed by $\tilde{F}(w_2) + \overline{\tilde{F}}(w_2) = 1$:

$$\begin{aligned} 1 &= \tilde{F}(w_2) + \overline{\tilde{F}}(w_2) \\ &= \frac{q}{p}F(w_2) + \frac{F(w_1(w_2)) - \tilde{F}(w_1(w_2))}{p} + \frac{1-q}{1-p}\overline{F}(w_2) \\ &= \frac{q}{p}F(w_2) + \frac{p-q}{p^2}F(w_1(w_2)) + \frac{1-q}{1-p}[1 - F(w_2)]. \end{aligned}$$

Suppressing the argument in $F(\cdot)$ and denote $F(w_1(w_2))$ and $F(w_2)$ by F_1 and F_2, respectively. Then the relationship above provides a quadratic equation of p,

$$(p-q)\left[p^2-(F_1+F_2)p+F_1\right]=0.$$

The solution $p=q$ restricts $F=\tilde{F}$ for all $w\in\mathcal{W}$. Therefore we focus on the quadratic equation of p,

$$p^2-(F_1+F_2)p+F_1=0.$$

It is sufficient to consider the possibilities of two real roots on $(0,1)$.

(a) There are two equal real roots, $p_1=p_2=p^*$. Straightforwardly, we obtain in this case $p^*=\frac{F_1+F_2}{2}$. If we let w_2 increase continuously on the interval $(\overline{w}_1,\underline{w_2})$, both $F_1\equiv F(w_1(w_2))$ and $F_2\equiv F(w_2)$ increase continuously and this contradicts our assumption that p is a constant in $(0,1)$.

(b) There are two distinct real roots, $p_1\neq p_2, 0<p_j<1, j=1,2$. Vieta's theorem implies that $p_1+p_2=F_1+F_2$ and $p_1p_2=F_1$, or equivalently,

$$\frac{p_1}{1+p_1/p_2}=\frac{F_1}{F_1+F_2}.$$

The L.H.S. of the equation above is bounded for p_1 and p_2 are two constants on $(0,1)$. However, the R.H.S. changes with w_2. The relationship $w_1\equiv w_1(w_2)$ permits us to vary w_2 continuously such that $w_1(w_2)\to\underline{w}$. Meanwhile, $w_2>\underline{w}_1$ always holds and $F_2>F(\underline{w}_1)$ is bounded always from zero. As a consequence, the R.H.S. goes to zero as we move w_2 continuously while the L.H.S. keeps a constant. This leads to a contradiction and completes our proof. □

C Existence of PBNE in Section 3

This part of the appendix provides sufficient conditions for condition (b) in Lemma 3 when $|\mathcal{C}|=3$. In a symmetric PBNE, members follow the same pure strategy σ^* which can be 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 $p(\omega^*)\equiv\Pr\{(\alpha_i,w_i)\in\omega^*)$. With $|\mathcal{C}|=3$,

$$p_{1,1}\equiv 1-(1-p(\omega^*))^2;\ p_{0,0}\equiv 1-p(\omega^*)^2;\ p_{1,0}=p(\omega^*)^2,$$

where p_{d,d_i} are defined as in the text. Thus $\varphi(p(\omega^*))\equiv\frac{-1+2p(\omega^*)}{2p(\omega^*)(1-p(\omega^*))}$. Let ϕ be defined as in the text.

TC *(Tail Conditions). $\phi'(\kappa;\mathcal{I}_l)\phi(\kappa;\mathcal{I}_l)^{-2}\to+\infty$ as $\kappa\to-\infty$ and $\phi'(\kappa;\mathcal{I}_l)[1-\phi(\kappa;\mathcal{I}_l)]^{-2}\to+\infty$ as $\kappa\to+\infty$.*

TC is a tail restriction conditional on the information set $\mathcal{I}_l$, which we suppress in notations for simplicity. 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 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. Hence $\varphi \circ \phi(\kappa) = 0$ for some $\kappa \in (-\infty, +\infty)$.

To see why this implies equilibrium existence with unbounded $\mathcal{W}$, 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 $\phi \to 1$. As a result, $\varphi'(\phi(\kappa))\phi'(\kappa)$ 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.

D Sufficient Primitive Conditions for Monotonicity

We now analyze what primitive conditions on the quadruple $(s, \tilde{y}, H, G)$ that are sufficient for Lemma 5. For this purpose, we relax Assumptions 8(b) and 9(b). To fix ideas, consider the case with $K = 2$ and let the support of $W_{i,k}$ be $[0, +\infty)$ for $k = 1, 2$.[18] We follow the notations from the proof of Lemma 5 but make one more simplification by letting $\Delta \equiv p_{1,1} - p_{0,0} = t(x)$. Then the choice probability of the alternative $D_i = 1$ is

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

We now provide two sufficient conditions on model primitives for the CCP $\Pr(D_i = 1|\mathcal{I} = x; \alpha_i)$ to be strictly monotonic in α_i. They are corresponding to the cases that $\sum_k w_{i,k} C_k(x; \alpha_i) = \Delta$ are parallel and intersect, respectively. To make our problem nontrivial, we exclude the case where $C_k(x; 1) = C_k(x; 0)$, $k = 1, 2$ since in such a scenario all the committee members are homogenous in making decisions, thus the unobserved types play no role.

Condition D.1. (i) $C_1(x;1)/C_2(x;1) = C_1(x;0)/C_2(x;0)$. (ii) Either $-\infty < C_1(x;h)/C_2(x;h) < 0$ or sign$[C_1(x;h)]$ =sign$[C_2(x;h)]$=sign(Δ), $h = 0, 1$.

Condition D.1(i) implies that the lines $\sum_k w_{i,k} C_k(x; \alpha_i) = \Delta$ are parallel for all $\alpha_i \in [0, 1]$. There are two subcases which correspond to panel (i) and (ii) in Figure 2 we need to consider: first, the lines have positive slope on (w_{1i}, w_{2i})-plane ($-\infty < C_1(x;h)/C_2(x;h) < 0$), then they

[18] We can also accommodate the case with bounded support of $W_{i,k}$, which only involves more tedious algebra but does not pose any additional conceptual challenge.

are distinguishable without any additional condition as we argued in the proof of Lemma 5. Second, if the slope is negative, then sign$[C_1(x;h)]$ =sign$[C_2(x;h)]$=sign(Δ) ensures that all the lines intersect with the positive $w_{1i}-$ and $w_{2i}-$ axis. Again, the argument of intercepts in proof of Lemma 5 applies and the ordering of the lines is identified. Notice that $C_k(x;h), k=1,2, h=0,1$ are observable, hence Condition D.1 is empirically testable. Hence by imposing condition D.1, we relax the high-level assumption of non-degenerate choice probabilities and substitute it by the assumption on primitives.

Before present condition D.2, we express explicitly the intersection point defined in the proof of Lemma 5 using model primitives. Recall that (w_{1i}^*, w_{2i}^*) satisfy

$$\begin{aligned}\textstyle\sum_k w_{i,k}^* C_k(x;0) = t(x),\\ \textstyle\sum_k w_{i,k}^* C_k(x;1) = t(x).\end{aligned}$$

In matrix form,

$$\begin{pmatrix} C_1(x;0) & C_2(x;0) \\ C_1(x;1) & C_2(x;1) \end{pmatrix}\begin{pmatrix} w_{1i}^* \\ w_{2i}^* \end{pmatrix} = \begin{pmatrix} t(x) \\ t(x) \end{pmatrix},$$

with the solution

$$\begin{pmatrix} w_{1i}^* \\ w_{2i}^* \end{pmatrix} = \frac{t(x)}{|C|}\begin{pmatrix} C_2(x;1)-C_2(x;0) \\ C_1(x;1)-C_1(x;0) \end{pmatrix} \equiv \frac{\Delta}{|C|}\begin{pmatrix} \Delta C_2 \\ \Delta C_1 \end{pmatrix},$$

where $|C| \equiv C_1(x;1)C_2(x;0) - C_2(x;1)C_1(x;0)$ and $\Delta C_j = C_j(x;1) - C_j(x;0), j=1,2$.

Condition D.2. (i) $C_1(x;1)/C_2(x;1) \neq C_1(x;0)/C_2(x;0)$. (ii) One of the following three conditions holds, for $h=0,1$ (a) $-\infty < C_1(x;h)/C_2(x;h) < 0$, and w_{1i}^* or w_{2i}^* is negative. (b) $0 < C_1(x;h)/C_2(x;h) < \infty$, sign$[C_1(x;h)]$ =sign$[C_2(x;h)]$=sign(Δ), and $w_{1i}^* \times w_{2i}^* < 0$. (c) sign$[C_1(x;1)/C_2(x;1)] \neq$ sign$[C_1(x;0)/C_2(x;0)]$, $w_{1i}^* < 0$ and $w_{2i}^* > 0$. Furthermore, conditions (ii)-(a) and (ii)-(b) can be combined as a more intuitive condition (d) sign$[C_1(x;1)/C_2(x;1)]$ = sign$[C_1(x;0)/C_2(x;0)]$, decision of any member i with ideological bias $\alpha_i \in (0,1)$ based on prior belief differs from that on on empirical evidence, and sign$[C_1(x;h)]$ =sign$[C_2(x;h)]$=sign(Δ) if $0 < C_1(x;h)/C_2(x;h) < \infty$.

Condition D.2 is specified for the case that all the lines corresponding to different α_i intersect at (w_{1i}^*, w_{2i}^*). According to the slopes of the lines, all three possible scenarios are summarized in panel (iii), (iv), and (v) of Figure 2. Panel (iii) illustrates the case that all the intersected lines have positive slopes as specified in (ii)-(a), $-\infty < C_1(x;h)/C_2(x;h) < 0$. The ordering of the lines in this case requires that (w_{1i}^*, w_{2i}^*) is not in the first quadrant. Otherwise we may not order the lines according to the event $\sum_k w_{i,k} C_k(x;\alpha_i) = \Delta$. If all the lines have negative slopes, i.e., $0 < C_1(x;h)/C_2(x;h) < \infty$ as shown in panel (iv), the assumption sign$[C_1(x;h)]$ =sign$[C_2(x;h)]$=sign(Δ) ensures that there are no degenerating types such that they all choose $D_i = 1$ with probability zero. Again, we need to exclude the possibility that the intersection point is in the first quadrant as argued above, this requirement is satisfied by imposing the restriction $w^*_{1i} \times w^*_{2i} < 0$. This is because sign$[C_1(x;h)]$ =sign$[C_2(x;h)]$=sign(Δ) implies that the intersection point can only be in the first, the second and the fourth quadrant while $w^*_{1i} \times w^*_{2i} < 0$ restricts it to the second and the fourth quadrant. To investigate the more intuitive condition (ii)-(d), which is equivalent to (ii)-(a) plus (ii)-(b), we first notice that a committee member changes his decision if he switches his information from prior belief to empirical evidence implies that the intersection point cannot be in the first quadrant. The remaining analysis is similar to previous one thus we omit it for brevity. Panel (v) depicts the last case: some of the lines have positive slopes while others have negative slopes, i.e., sign$[C_1(x;1)/C_2(x;1)] \neq$ sign$[C_1(x;0)/C_2(x;0)]$. If the intersection point $(w^*_{1i},\ w^*_{2i})$ is in the second or the third quadrant, it is easy to check that non-degenerate choice probabilities can not be guaranteed. This is due to the fact that for all the α_i such that the line $\sum_k w_{i,k}C_k(x;\alpha_i) = \Delta$ has a negative slope, the choice probability $\Pr(D_i = 1|\mathcal{I} = x;\alpha_i) = 0$ since the line will not pass the first quadrant. Employing the previous argument, we cannot allow the intersection point to be in the first quadrant either. Therefore we impose the condition (ii)-(c) $w^*_{1i} < 0$ and $w^*_{2i} > 0$ to ensure the intersection point to be in the second quadrant.

Both Condition D.1 and Condition D.2 can be similarly imposed if W_{1i} and W_{2i} are distributed on a bounded support in $[0,\infty) \times [0,\infty)$. The proof does not pose any conceptual challenge but the algebra will be more tedious since we have to analyze the lines $\sum_k w_{i,k}C_k(x;\alpha_i) = \Delta$ in a bounded area. We thus omit the analysis in this paper. $\square$

Figure 1: Illustration of the model with expressive recommendations

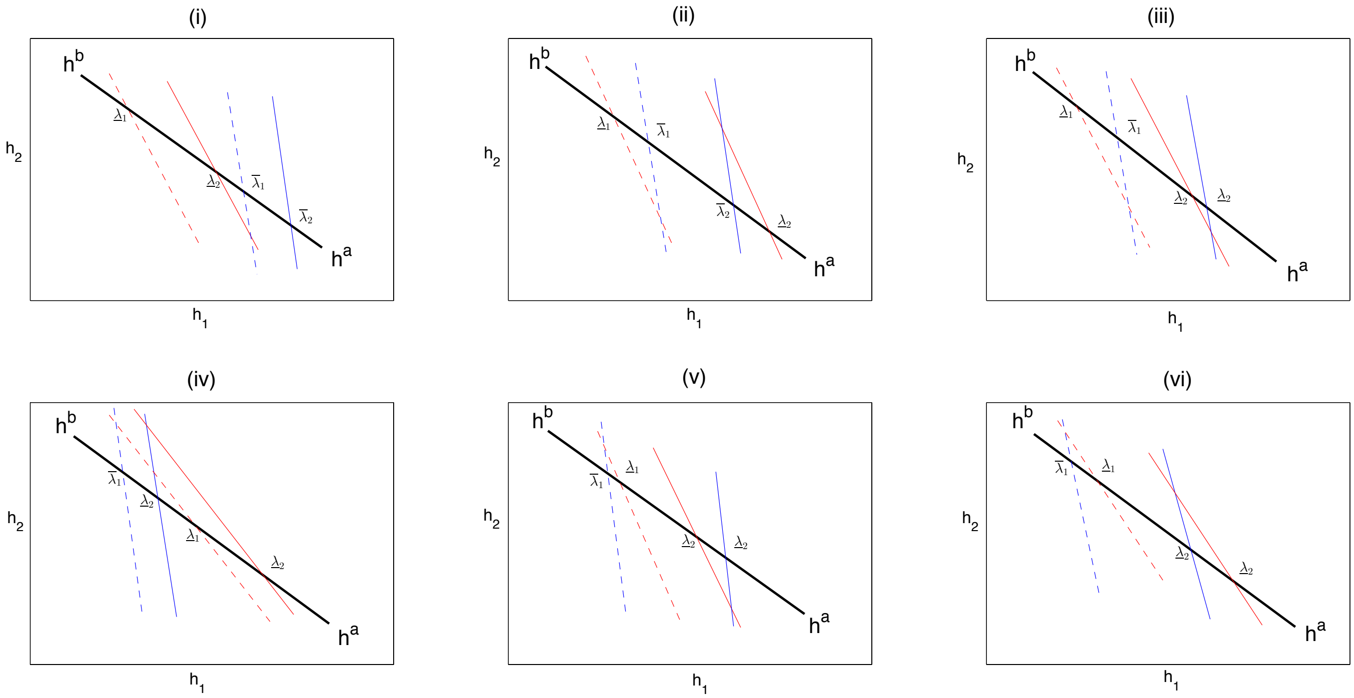

Table 1: Summary of votes

	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: Estimates of individual choice probabilities

External members			Internal members		
parameters	mean	std.	parameters	mean	std.
θ_0	-0.899^{**}	0.092	ϑ_0	-0.525^{**}	0.079
θ_1	0.268^{*}	0.106	ϑ_1	0.446^{**}	0.094
θ_2	1.134^{**}	0.364	ϑ_2	1.233^{**}	0.318
θ_3	-0.424^{*}	0.174	ϑ_3	-0.384^{**}	0.144
θ_4	-0.250	0.781	ϑ_4	0.536	0.671

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

Figure 2: Illustration of the model with strategic recommendations

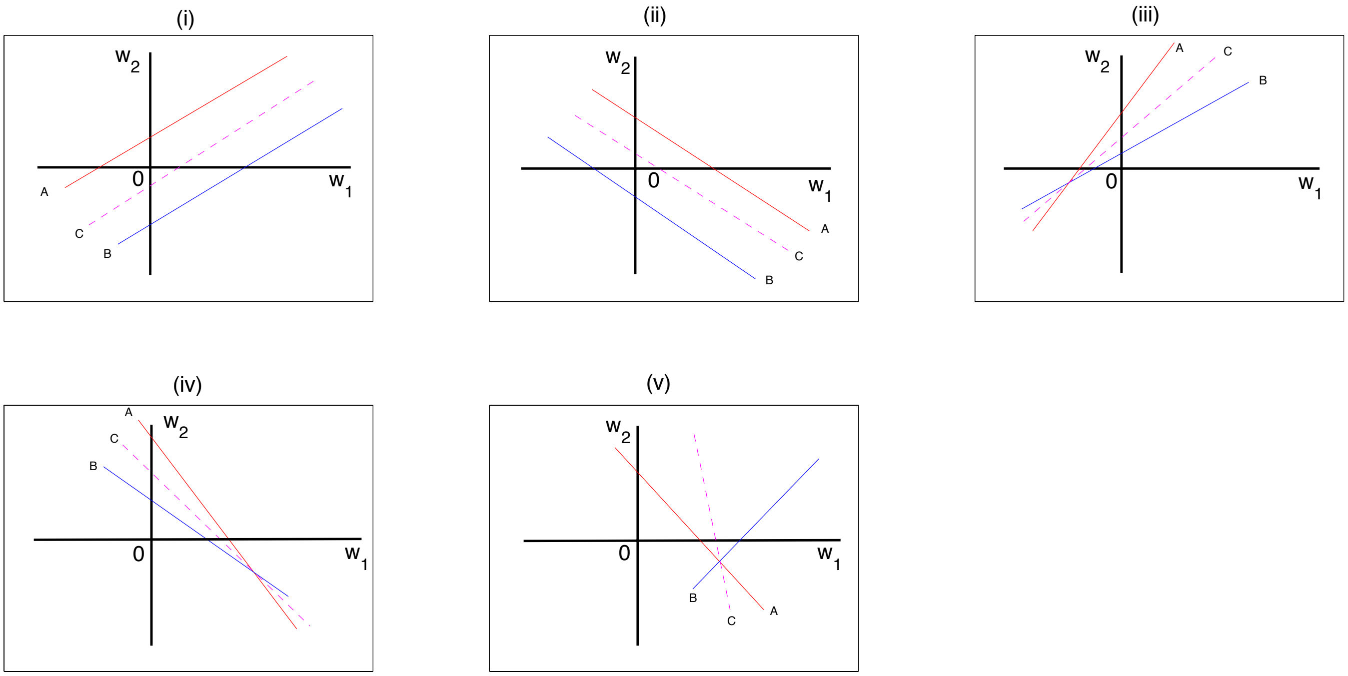

Figure 3: Estimated Distribution for the Tastes of External and Internal Members

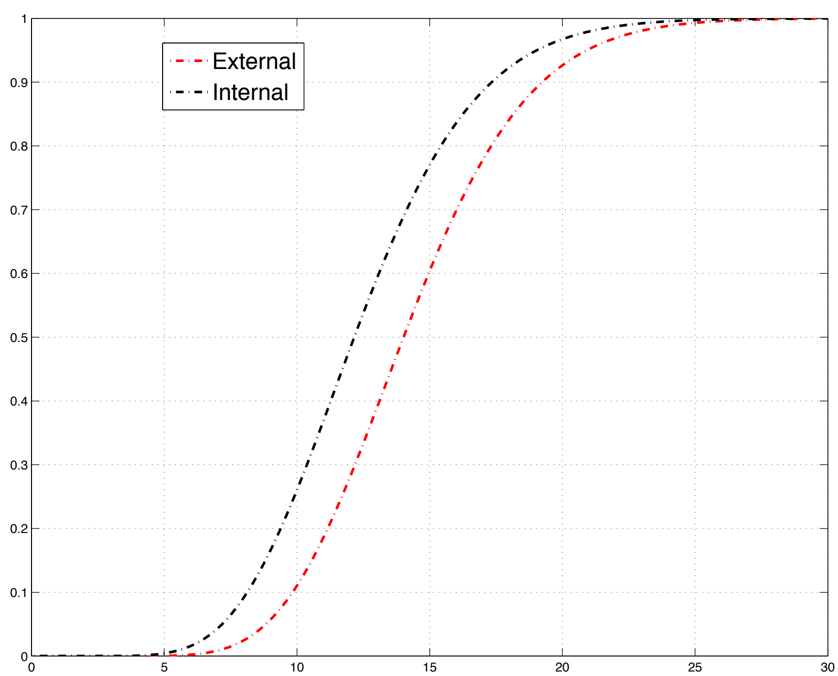

Table 3: Estimates for the Distribution of Ideological Bias[a]

	External members		Internal members	
	mean	std.	mean	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

[a] This table summarizes the empirical distribution of our estimates for the probability masses of α_i, using $B = 200$ bootstrap samples.

Table 4: Estimates for the Distribution of Individual Tastes[a]

	External members				Internal members			
	mean	Q1	median	Q3	mean	Q1	median	Q3
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] Q1: 1^{st} quartile; Q3: 3^{rd} quartile. This table summarizes the empirical distribution of our estimates for (a, b) using $B = 200$ bootstrap samples.