Preference for Similarity and Information Collection in Groups

MariaGiovanna Baccara (Stern, NYU) Leeat Yariv (Caltech)

Current Version: February 15, 2008

Preliminary Draft. Please Do Not Circulate.

Abstract. We study a model in which agents make individual decisions on several dimensions (e.g., life-decisions on consumption and education, meal choices regarding food and wine, etc.). Individuals' tastes differ in terms of the relative weight they put on each dimension. Before making decisions, each individual can collect information on at most one dimension and all information collected is shared within a group of peers. For any group of individual, we characterize the equilibrium choice of information collection.

We then consider the group of peers (or friends) as an object of choice. We characterize the peer group's optimal composition for each individual in the population. We show that, for each individual, there is a large equivalence class of optimal groups, potentially with maximal variance of tastes. We also characterize the stable groups, that is, groups that are optimal for all participating individuals, and illustrate the similarity requirements such stable groupings entail. Our results help explain an array of applications, ranging from friendship formation, to selective exposure.

JEL classification: D82, D85.

Keywords: Homophily, Group Choice, Selective Exposure.

1. Introduction

The empirical literature has identified a wide array of realms in which underlying social networks impact outcomes, ranging from job search, to the spread of epidemics, to crime and addictive behaviour. Understanding the process by which social networks evolve is therefore crucial for both predictions of future outcomes, as well as for the determination of policies directed at affecting those outcomes. One overwhelming empirical correlate with the formation of links is that of similarity, i.e., agents appear to be connected to those who are similar to them (in demographics, political opinions, looks, etc.).[1] Consequently, this phenomenon, known as homophily in the sociology literature, has recently received growing attention from economists, who have been focusing on explaining the strategic origins of homophily, as well as studying its implications on network formation, segregation outcomes, and social welfare.[2]

We provide a simple model in which the group of peers is an object of choice. Our goal is to provide a taxonomy of different environments generating homogeneous or heterogeneous preferences as stable outcomes.

Specifically, we study a model in which agents make individual decisions on several dimensions (e.g., life-decisions on consumption and education, meal choices regarding food and wine, etc.). Individuals' tastes differ in terms of the relative weight they put on each dimension.

Before making decisions, each individual can collect information on at most one dimension. We assume that all information collected is shared within a group of peers. For any group of individual, we characterize the equilibrium choice of information collection.

We then consider the group of peers (or friends) as an object of choice. Depending on her tastes, each individual prefers certain peer groups to others. We characterize the peer group's optimal composition for each individual in the population. We show that, for each individual, there is a large equivalence class of optimal groups, potentially with maximal variance of tastes. We also characterize the stable groups, that is, groups that are optimal for all participating individuals, and illustrate the similarity requirements such stable groupings entail. Thus, our

[1] See McPherson M., Smith-Lovin L. and J. Cook, (2001).

[2] See Currarini, Jackson and Pin (2007) and Peski (2007).

results suggest that in situations in which each member has some leverage in choosing their group partners, global stability occurs when tastes are sufficiently close. Nonetheless, each individual cares only about order-statistic type of attributes of the group.

As it turns out, when information gathering is costly, a free-riding problem arises in groups that are too homogeneous. Introducing group heterogeneity has the benefit of mitigating this problem. In particular, we show that for any positive information gathering cost, for sufficiently large groups stability will entail extreme preference polarization.

We study the interplay between the information collection and the degree of similarity in the groups. We also analyse how the degree of similarity in stable groups is affected by the properties of the information structure.

Our results help explain an array of applications, ranging from friendship formation, to selective exposure.

1.1. Related Literature.

To be completed.

2. THE MODEL

There are two issues at stake: A and B, each taking a value in $\{0,1\}$. Assume issues are determined independently at the outset of the game, with p_I denoting the probability of issue I being 1.

There are n agents in a group. Each agent i needs to make a decision on both dimensions, namely pick an action $v \in V = \{0,1\} \times \{0,1\}$. Utilities are determined individually. Each agent i is characterized by a taste $t_i \in [0,1]$. The utility of agent i from choosing v when the realized states are A and B is given by:

$$u_i(v; A, B) = t_i \mathbf{1}_A(v_1) + (1 - t_i)\, \mathbf{1}_B(v_2).$$

Prior to making a decisions, agents have access to information in two stages. First, each agent selects simultaneously one of two information sources, α or β. Information source α provides the realized state A with probability $q_\alpha > \frac{1}{2}$. That is, upon choosing information source α, the agent observes a signal $s \in \{0, 1, \varnothing\}$, according to:

$$\Pr(s = A) = q_\alpha, \quad \Pr(s = \varnothing) = 1 - q_\alpha.$$

Similarly, information source β provides the realized state B with probability $q_\beta > \frac{1}{2}$.

After all information sources are selected, all signals are realized and made public within the group.

Let us now step backwards and suppose that, given a group size n and an agent of type $t \in [0, 1]$, this agent is given the possibility of choosing the other $n - 1$ agents in the group. After characterizing the optimal group chosen by agent t, $(t_1^*(t), .., t_n^*(t))$ (in which one element is t), we characterize the stable groups-that is

Definition 1. *A group $(t_1, .., t_n)$ is stable if it is optimal for all its components, i.e. $(t_1, .., t_n) = (t_1^*(t_i), .., t_n^*(t_i))$ for all $i = 1, .., n$.*

3. Analysis of Equilibrium

3.1. Information Scources. Let us now consider a fixed group of $n < N$ agents with tastes $t_1, t_2, ..., t_n$, and let us study the equilibrium choice of information sources. Assume first that $t_i \neq t_j$ for all i, j and that a pure equilibrium exists (note that when tastes are different across agents, equilibria are in pure strategies generically).

When there are k_x agents choosing source x, the probability state x is revealed is $1 - (1 - q_x)^{k_x}$, $x = \alpha, \beta$.

Without loss of generality, assume that $t_1 > t_2 > ... > t_n$. An equilibrium is characterized by the profile of chosen sources $(x_1, ..., x_n)$, where $x_i \in \{\alpha, \beta\}$ is the source chosen by agent i.

Lemma If there exist $i > j$ such that $x_i = \beta$ and $x_j = \alpha$, then $(y_1, ..., y_n) \in \{\alpha, \beta\}^n$, where $y_l = x_l$ for all $l \neq i, j$, $y_i = \alpha$, and $y_j = \beta$ constitutes an equilibrium as well.

In order to assess the information content within a particular group, the Lemma suggests that is without loss of generality to concentrate on a (pure) equilibrium of the form $(x_1, ..., x_n)$

such that for some $\kappa = 1, ..., n$, $x_1 = ... = x_k = \alpha$ and $x_{\kappa+1} = ... = x_n = \beta$. In words, we concentrate on equilbria in which any agent choosing the source α cares more about A than any agent choosing the source β.

In fact, the IC constraints now allow us to characterize the equilibrium. Indeed, there are two crucial incentive constraints determining κ: that of agent κ and that of agent $\kappa + 1$.

The IC constraint for agent κ is $U(t_\kappa, \kappa, n-\kappa) \geqslant U(t_\kappa, \kappa - 1, n - \kappa + 1)$.

The IC constraint for agent $\kappa + 1$ is $U(t_{\kappa+1}, \kappa, n-\kappa) > U(t_{\kappa+1}, \kappa + 1, n - \kappa - 1)$.

Note that in the (generic) case of multiple pure equilibria, we consider the equilibrium entailing the most α source signals.

This boils down to looking for the maximum k such that $U(t_k, k, n-k) \geqslant U(t_k, k-1, n-k+1)$, which amounts to (see terms in proof of Lemma 3.1):

$$\kappa = \max_k \left\{ t_k \geqslant \frac{(1-p_\beta)(1-q_\beta)^{n-k} q_\beta}{(1-p_\alpha)(1-q_\alpha)^{k-1} q_\alpha + (1-p_\beta)(1-q_\beta)^{n-k} q_\beta} \right\},$$

where we use the convention here that $\kappa = 0$ if the set above is empty.

Naturally, the threshold is n if $U(t_n, n, 0) \geqslant U(t_n, n-1, 1)$, or

$$t_n \geqslant \frac{(1-p_\beta) q_\beta}{(1-p_\alpha)(1-q_\alpha)^{n-1} q_\alpha + (1-p_\beta) q_\beta}.$$

Similarly, the threshold is 0 if $U(t_1, 0, n) > U(t_1, 1, n-1)$

$$t_1 < \frac{(1-p_\beta)(1-q_\beta)^{n-1} q_\beta}{(1-p_\alpha) q_\alpha + (1-p_\beta)(1-q_\beta)^{n-1} q_\beta}.$$

3.2. Group Composition. We can now turn to the optimal group composition from an individual agent's point of view.

First, note that for each t and group size n, there exists an optimal (from the point of view of type t) allocation of agents who choose source α and source β, k^n and $n - k^n$, defined by:

$$k^n(t) = \max_k \left\{ t \geqslant \frac{(1-p_\beta)(1-q_\beta)^{n-k} q_\beta}{(1-p_\alpha)(1-q_\alpha)^{k-1} q_\alpha + (1-p_\beta)(1-q_\beta)^{n-k} q_\beta} \right\} \tag{1}$$

where, again, if the set over which the maximization is taken is empty, $k^n(t) = 0$. Condition in (1) amounts to require k^n to be the maximum k such that $U(t, k, n-k) \geqslant U(t, k-1, n-k+1)$.

In the generic case in which the agent is indifferent between two allocations, we choose the one entailing more α source choices. Naturally, $k^n(t)$ is weakly increasing in t and in n.

Denote by

$$t^n(k) = \frac{(1-p_\beta)(1-q_\beta)^{n-k} q_\beta}{(1-p_\alpha)(1-q_\alpha)^{k-1} q_\alpha + (1-p_\beta)(1-q_\beta)^{n-k} q_\beta}.$$

This is the type that is indifferent between source α and β when precisely $k-1$ other agents choose source α. Note that $t^n(k)$ is increasing in k. This reflects the fact that types who would like more α source signals are naturally those who care more about the realization of state A. Furthermore, $t^n(k)$ is decreasing in n. Intuitively, when fixing the number of source α signals at k, increasing n implies an increase in the number of source β signals, thereby a reduction in the marginal benefit of an additional source β signal. In particular, the marginal type who is indifferent between the two sources will care less about the realization of state A.

From our analysis thus far, we get the following proposition:

Proposition 1 (Optimal Group Composition)

1. For any type t, the class of optimal groups comprising of tastes $t_1 > ... > t_n$ (one of which is t) entails

$$t_{k^n(t)} \geqslant t^n(k^n(t)) \text{ and } t_{k^n(t)+1} < t^n(k^n(t)+1).$$

2. A group comprising of tastes $t_1 > ... > t_n$ is optimal for all t_i if and only if either $t_i < t(1)$ for all i, or $t_i \geqslant t(n)$ for all i, or there exists $k = 0, ..., n$, such that for all i,

$$t_i \in [t^n(k), t^n(k+1)).$$

The intuition of Proposition 1 is the follwing. First, let us focus on the optimal group composition from an individual agent's point of view. Depending on his type, this agent optimal group will be formed by a set of $k^n(t)$ agents that will choose source α and $n - k^n(t)$ agents that will choose source β. This amounts to pick any k agents (or $k-1$ agents, if this set includes the agent himself) whose type is above the threshold $t^n(k^n(t))$ and $n - k^n(t)$ agents whose type is below the threshold $t^n(k^n(t)+1)$. Note that an optimal group ca always be obtained by picking agents of extreme types 0 and 1 only.

On the other hand, stable groups can be formed only by agents whose types lie in the interval $[t^n(k), t^n(k+1))$. This is because, given a relatively moderate group, an extreme agents added to that group will tend to deviate from it and pick friends with a similar polarization in tastes.

Proposition 1 suggests that in situations in which each member has some leverage in choosing their group partners, global stability occurs when tastes are sufficiently close. Nonetheless, each individual cares only about order-statistic type of attributes of the group.

4. Comparative Statics

In this section we study how the results in Proposition 1 are affected by the type of the agent who chooses the group, the size of the group and the properties of the information structure.

4.1. Individual Group Choice. The Proposition indicates some immediate comparative statics. Indeed, suppose that $t > t'$. It follows that $k^n(t) \geqslant k^n(t')$ and therefore $t^n(k^n(t)) \geqslant t^n(k^n(t')).$ Thus, type t's optimal group consists of more agent who surpass a higher threshold than type t'. Nonetheless, an optimal group for type t may entail a higher variance of tastes than an optimal group for t'.

Suppose agent is type t. How does t relate to the types he chooses to be in his group? Can t be an extreme of the chosen types?

If $t > \frac{(1-p_\beta)q_\beta}{(1-p_\alpha)(1-q_\alpha)^{n-1}q_\alpha+(1-p_\beta)q_\beta}$, then all friends of t at the right of t is one of optimal choices.

To be completed.

4.2. Group Stability. Note that the lenghts of the intervals that characterize stable groups in Proposition 1 provide a proxy for equilibrium homophily. In this section, we study how these lenghts are affected by the size of the group n, by the information structure and by

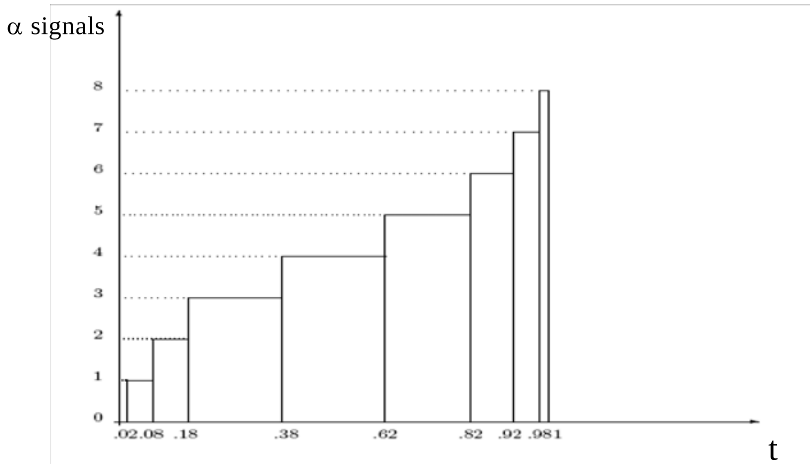

Figure 1:

the proximity of the intervals to the extreme types.

Let us start from analyzing how the lenght of the stable group intervals is affected by how moderate the types that compose the groups are. It is easy to show that the intervals are smaller close to the extreme types, and become wider for moderate types. This suggests that more extreme individuals can form stable groups only with people that are relatively very similar to them. On the other hand, moderate types can for stable groups with people that are relatively more heterogeneous. Figure 1 describes the stable groups intervals in the type space $[0, 1]$.

5. Endogenous Collection of Information

Suppose that observing a signal (of either source) is costly. Namely, observing a signal comes at a cost of $c > 0$. Our problem then becomes a type of a free rider problem, which can be modified in the initial group selection.

Note that now there is a maximal number $n^x_{\max}$ of signals of either type $x = \alpha, \beta$ that can be bought in any group equilibrium, defined as the maximal integer for which

$$(1 - q^x)^{n^x_{\max}-1} - (1 - q^x)^{n^x_{\max}} = (1 - q^x)^{n^x_{\max}-1} q^x \geq c$$

$$\Leftrightarrow n_{\max}^{x} = \left\lfloor \frac{\ln c - \ln q^{x}}{\ln (1-q^{x})} + 1 \right\rfloor .$$

These are the number of signals agents with the most extreme tastes ($t = 0$ or $t = 1$) would be willing to invest in.

5.1. Analysis of Equilibrium. As before, without loss of generality, assume that $t_1 > t_2 > ... > t_n$. An equilibrium is characterized by the profile of chosen sources $(x_1, ..., x_n)$, where $x_i \in \{\alpha, \beta, \varnothing\}$ is the source chosen by agent i (and $\varnothing$ stands for agent i no acquiring information). There is an analogue of our initial lemma in the current setup.

Lemma If there exist $i > j$ such that $x_i \in \{\beta, \varnothing\}$ and $x_j \in \{\alpha, \varnothing\}$ then $(y_1, ..., y_n) \in \{\alpha, \beta, \varnothing\}^n$, where $y_l = x_l$ for all $l \neq i, j$, $y_i = x_j$, and $y_j = x_i$ constitutes an equilibrium as well.

In particular, from an information point of view, we can restrict attention to equilibria of the form $(x_1, ..., x_n)$, where there exists $\kappa^{\alpha} = 0, ..., n$ and $\kappa^{\beta} = 1, ..., n+1$, $\kappa^{\beta} \geqslant \kappa^{\alpha}$, such that for all $1 \leqslant i \leqslant \kappa^{a}$, $x_i = \alpha$, for all $\kappa^{\alpha} < i < \kappa^{\beta}$, $x_i = \varnothing$, and for all $\kappa^{\beta} \leqslant i \leqslant n$, $x_i = \beta$.

Note that now there is pressure to choose highly variant groups in order to get the sufficient incentives to acquire information. In a sense, the cost does not affect the choice of signal source, but only the desire to invest in information.

Proposition 2 For any κ^{a} and $\kappa^{\beta} \geqslant \kappa^{\alpha}$, with $\kappa^{x} < n_{\max}^{x}$ for $x = \alpha, \beta$, there exists $T_{\alpha}(\kappa^{a}, \kappa^{\beta})$, $T_{\beta}(\kappa^{a}, \kappa^{\beta})$ and $\underline{T}_{\varnothing}(\kappa^{a}, \kappa^{\beta})$ and $\overline{T}_{\varnothing}(\kappa^{a}, \kappa^{\beta})$ such that $t_i \geq T_{\alpha}(\kappa^{a}, \kappa^{\beta})$ for any $1 \leqslant i \leqslant \kappa^{a}$, $t_i \in [\underline{T}_{\varnothing}(\kappa^{a}, \kappa^{\beta}), \overline{T}_{\varnothing}(\kappa^{a}, \kappa^{\beta}))$ for any $\kappa^{a} \leq i < \kappa^{\beta}$ and $t_i < T_{\beta}(\kappa^{a}, \kappa^{\beta})$ for any $\kappa^{\beta} \leq i \leq n$.

From an individual agent's point of view, the optimal group consists of κ^{a} agents choosing the α signal (such that $t_i \geq T_{\alpha}(\kappa^{a}, \kappa^{\beta})$), and $n - \kappa^{\beta} + 1$ agents who choose the β signal (such that $t_i < T_{\beta}(\kappa^{a}, \kappa^{\beta})$). In making his choice, the individual has to consider additional constraints with respect to the $c = 0$ case. In particular, if $\kappa^{x} > n_{\max}^{x}$, the constraint $\kappa^{x} \leq n_{\max}^{x}$ is binding. This means that if the agent picks any agent to gather information on source, say α, in addition to $n_{\max}^{\alpha}$ such agent will not bring additional information. Thus, if $\kappa^{\beta} < n_{\max}^{\beta}$, more agents can be added to gather information on source β instead. If both constrais are binding, the optimal group will include agent that do not collect information in equilibrium.

Proposition 3 (Costly Signals)

An optimal group has one of the following structures

1. $\kappa^\alpha \le n^\alpha_{\max}$ and $n - \kappa^\beta + 1 \le n^\beta_{\max}$. In this case, agents $t_i = 1$ for any $1 \leqslant i \leqslant \kappa^a$ and $t_i = 0$ for any $\kappa^\beta \le i \le n$ are an optimal group. There will be κ^α and $n - \kappa^\beta + 1$ signals α and, respectively, β collected.

2. Either $\kappa^\alpha > n^\alpha_{\max}$ or $n - \kappa^\beta + 1 > n^\beta_{\max}$. In this case if $\kappa^\alpha > n$, $(n - \kappa^\beta + 1 > n^\beta_{\max})$ then $t_i = 1$ $(t_i = 0)$ for any $1 \leqslant i \leqslant n^\alpha_{\max}$ $(n - n^\beta_{\max} + 1 \le i \le n)$ and $t_i = 0$ $(t_i = 1)$ for $n^\alpha_{\max} \le i \le n$ $(0 \le i \le n^\beta_{\max} - 1)$ is an optimal group. There will be $n^\alpha_{\max}$ and $n^\beta > n - \kappa^\beta + 1$ signals α and, respectively, β collected.

3. $\kappa^\alpha > n^\alpha_{\max}$ and $n - \kappa^\beta + 1 > n^\beta_{\max}$. In this case, $t_i = 1$ for any $1 \leqslant i \leqslant n^\alpha_{\max}$ and $t_i = 0$ for $n^\alpha_{\max} \le i \le n$ is an optimal group. There will be $n^\alpha_{\max}$ and $n^\beta_{\max}$ signals α and, respectively, β collected.

The intuition of Proposition 3 is the following. If none of the constraints $\kappa^\alpha \le n^\alpha_{\max}$ and $n - \kappa^\beta + 1 \le n^\beta_{\max}$ is binding, the unconstrained optimum for the individual agent can be achieved, and the analysis is similar to the $c = 0$ case. However, if, say, $\kappa^\alpha > n^\alpha_{\max}$, then one of the constraints is binding. Thus, after choosing $n^\alpha_{\max}$ agents at $t_i = 1$ for $1 \leqslant i \leqslant n^\alpha_{\max}$, the agent is better off switching to $t_i = 0$ for $i > n^\alpha_{\max}$. This could lead to more information collection on source β than in the unconstrained solution. Finally, if both $\kappa^\alpha > n^\alpha_{\max}$ and $n - \kappa^\beta + 1 > n^\beta_{\max}$, the agent is constrained to collect $n^\alpha_{\max}$ and $n^\beta_{\max}$ signals α and, respectively, β. This can always be achieved by selecting extreme agents only.

Note that the optimal groups described in points (2) and (3) of Proposition 3 are also stable. Thus, in contrast with the $c = 0$ case, polarized groups can be stable if information is costly.

5.2. Adding Members. Consider now the addition of agents. In particular, suppose a group comprised of tastes $t_1 > ... > t_n$ and equilibrium κ. Suppose agent of taste t^P chooses an additional agent for the group (this can be a metaphor for a Principal who contemplates adding an expert to an already existing team, or a club that considers a new member, a group of friends adding new people to their circle, etc. In particular, the agent who makes the choice may or may not be part of the original group of agents).

The following proposition illustrates the optimal range of tastes for an additional member t^P would choose.

Proposition 4 (Optimal Additional Members)
1. If $t^P > t^{n+1}(\kappa+1)$ and $t_{k+1} < t^{n+1}(\kappa+1)$ then any agent with taste $t < t^{n+1}(\kappa+1)$ would be optimal.
2. If $t^P < t^{n+1}(\kappa+1)$ and $t_{\kappa+1} < t^{n+1}(\kappa+1)$, then any agent with taste $t \geqslant t^{n+1}(\kappa+1)$ would be optimal.
3. In all other cases, any agent would be optimal.

Intuitively, if in the original group precisely κ agents chose source α, then adding a member can alter the number of agents choosing source α to $\kappa' \in \{\kappa, \kappa+1\}$. To understand why that is the case, suppose, for instance, that a new member allows for an equilibrium in which $\kappa+2$ and $(n+1)-(\kappa+2) = n-\kappa-1$ agents choose sources α and β, respectively. This implies, that for some agent i, choosing the $\kappa+2'$nd source α signals when there are only $n-\kappa-1$ source β signals is preferrable, but choosing the $(n-\kappa)'$th source β signal when there are only κ source α signals is preferrable, in clear contradiction.

The Proposition now has a clear interpretation. When, e.g., the choosing agent cares sufficiently about the realization of state A, then, if there are already $\kappa+1$ agents who are willing to choose source α in a group of $n+1$, then the additional member's tastes are inconsequential. Otherwise, a member who cares sufficiently about the realization of state A would be necessary.

Is there a majority within the group that would agree on the taste of one agent? May be interesting to hint at some simple group formation process such as that.

6. References

Jackson M., Currarini S. and Pin P. (2007) "An Economic Model of Friendship: Homophily, Minorities and Segregation," *Mimeo*.

McPherson M., Smith-Lovin L. and J. Cook, (2001) "Birds of a Feather: Homophily in Social Networks", *Annual Review of Sociology*.

Peski M, (2007) "Complementarities, Group Formation and Preferences for Similarity" Mimeo.

Appendix

Proof of Lemma 3.1 Indeed, assume that

$$n_\alpha = |\{x_l = \alpha, l \neq i, j\}| \quad \text{and} \quad n_\beta = |\{x_l = \beta, l \neq i, j\}|$$

Denote by $U(t, k_\alpha, k_\beta)$ the expected utility of an agent with tastes t who is in a group in which k_x agents choose source x, $x = \alpha, \beta$. Under equilibrium x, IC constraint for player i requires that $U(t_i, n_\alpha + 1, n_\beta + 1) \geqslant U(t_i, n_\alpha + 2, n_\beta)$. That is,

$$\begin{aligned} & t_i \left[1 - (1 - q_\alpha)^{n_\alpha + 1} (1 - p_\alpha)\right] + (1 - t_i) \left[1 - (1 - q_\beta)^{n_\beta + 1} (1 - p_\beta)\right] \\ \geqslant \;\; & t_i \left[1 - (1 - q_\alpha)^{n_\alpha + 2} (1 - p_\alpha)\right] + (1 - t_i) \left[1 - (1 - q_\beta)^{n_\beta} (1 - p_\beta)\right]. \end{aligned}$$

$$\Leftrightarrow (1 - t_i) (1 - p_\beta) (1 - q_\beta)^{n_\beta} q_\beta - t_i (1 - p_\alpha) (1 - q_\alpha)^{n_\alpha + 1} q_\alpha \geqslant 0.$$

Similarly, the IC constraing for player j requires that $U(t_j, n_\alpha + 1, n_\beta + 1) \geqslant U(t_j, n_\alpha, n_\beta + 2)$, translating to

$$t_j (1 - p_\alpha) (1 - q_\alpha)^{n_\alpha} q_\alpha - (1 - t_j) (1 - p_\beta) (1 - q_\beta)^{n_\beta + 1} q_\beta \geqslant 0.$$

Clearly,

$$\begin{aligned} & (1-t_j)(1-p_\beta)(1-q_\beta)^{n_\beta} q_\beta - t_j (1-p_\alpha)(1-q_\alpha)^{n_\alpha+1} q_\alpha \\ > \ & (1-t_i)(1-p_\beta)(1-q_\beta)^{n_\beta} q_\beta - t_i (1-p_\alpha)(1-q_\alpha)^{n_\alpha+1} q_\alpha \geqslant 0. \end{aligned}$$

and, similarly,

$$\begin{aligned} & t_i (1-p_\alpha)(1-q_\alpha)^{n_\alpha} q_\alpha - (1-t_i)(1-p_\beta)(1-q_\beta)^{n_\beta+1} q_\beta \\ > \ & t_j (1-p_\alpha)(1-q_\alpha)^{n_\alpha} q_\alpha - (1-t_j)(1-p_\beta)(1-q_\beta)^{n_\beta+1} q_\beta \geqslant 0. \end{aligned}$$

This assures that under the profile $(y_1, ..., y_n)$ the IC constraints for agents i, j are satistfied. That is,

$$\begin{aligned} U(t_i, n_\alpha+1, n_\beta+1) &\geqslant U(t_i, n_\alpha, n_\beta+2), \\ U(t_j, n_\alpha+1, n_\beta+1) &\geqslant U(t_i, n_\alpha+2, n_\beta). \end{aligned}$$

The IC constraints for all other agents remain identical to those under x. ■

Proof of Proposition 1.

1. For any type t, an optimal choice for a group entails choosing $k^n(t)$ and $n-k^n(t)$ types such that the first group chooses source α and the second chooses source β. This is achieved if and only if for $i = 1, .., k^n(t)$

$$U(t_i, k^n(t), n-k^n(t)) \geqslant U(t_i, k^n(t)-1, n-k^n(t)+1)$$

or $t_i \geqslant t^n(k^n(t))$ for $i = 1, .., k^n(t)$. For $i = k^n(t)+1, .., n$, it must be the case

$$U(t_i, k^n(t), n-k^n(t)) \geq U(t_i, k^n(t)+1, n-k^n(t)-1)$$

or $t_i < t^n(k^n(t)+1)$.

2. Suppose the group $t_1 > ... > t_n$ is optimal for all i. Thus, either $t_i < t(1)$ for all i, or

$t_i \geqslant t(n)$ for all i, or it must be the case that there exist a k such that for all t_i

$$U(t_i, k, n-k) \geq U(t_i, k-1, n-k+1)$$

or $t_i \geq t^n(k)$

and

$$U(t_i, k, n-k) \geq U(t_i, k+1, n-k-1)$$

or $t_i \leq t^n(k+1)$.

Suppose now that $t_i < t(1)$ for all i. Then, the group must be optimal for all i since they all prefer all sources to be β. Same for $t_i \geqslant t(n)$ for all i. Suppose now that there exist a k such that $t_i \in [t^n(k), t^n(k+1))$ for all i. Then, by definition, we have $U(t_i, k, n-k) \geq U(t_i, k-1, n-k+1)$ and $U(t_i, k, n-k) \geq U(t_i, k+1, n-k-1)$ for all t_i. This implies that k is the optimal number of signals α for all i. To see this note that the marginal utility of a signal from one source is decreasing. Indeed

$$\begin{aligned}
& U(t_i, k, n-k) - U(t_i, k-1, n-k+1) \\
= \; & t_i \left[1 - (1-q_\alpha)^k (1-p_\alpha)\right] + (1-t_i) \left[1 - (1-q_\beta)^{n-k} (1-p_\beta)\right] - \\
& t_i \left[1 - (1-q_\alpha)^{k-1} (1-p_\alpha)\right] - (1-t_i) \left[1 - (1-q_\beta)^{n-k+1} (1-p_\beta)\right] \\
= \; & t_i(1-p_\alpha) \left[(1-q_\alpha)^{k-1} - (1-q_\alpha)^k\right] - (1-t_i)(1-p_\beta) \left[(1-q_\beta)^{n-k} - (1-q_\beta)^{n-k+1}\right]
\end{aligned}$$

is decreasing in k. ■

Proof of Proposition 2.

Let us look at IC constraints to characterize the equilibrium.

For $1 \leqslant i \leqslant \kappa^a$, it must be the case that

$(IC1_\alpha)$ $U(t_i, \kappa^a, n-\kappa^\beta+1) - c \geqslant U(t_i, \kappa^a - 1, n - \kappa^\beta + 2) - c$

and

$(IC2_\alpha)$ $U(t_i, \kappa^a, n-\kappa^\beta+1) - c \geqslant U(t_i, \kappa^a - 1, n - \kappa^\beta + 1)$

$(IC1_\alpha)$ becomes

$$t_i \geq \frac{(1-q_\beta)^{n-\kappa^\beta+1}(1-p_\beta)q_\beta}{(1-q_\alpha)^{\kappa^a-1}(1-p_\alpha)q_\alpha+(1-q_\beta)^{n-\kappa^\beta+1}(1-p_\beta)q_\beta}$$

$(IC2_\alpha)$ becomes

$$t_i \geq \frac{c}{(1-p_\alpha)(1-q_\alpha)^{\kappa^a-1}q_\alpha}$$

Thus, we have

$$\begin{aligned} t_{\kappa^a} &\geq \max\left\{\frac{(1-q_\beta)^{n-\kappa^\beta+1}(1-p_\beta)q_\beta}{(1-q_\alpha)^{\kappa^a-1}(1-p_\alpha)q_\alpha+(1-q_\beta)^{n-\kappa^\beta+1}(1-p_\beta)q_\beta};\frac{c}{(1-p_\alpha)(1-q_\alpha)^{\kappa^a-1}q_\alpha}\right\} \\ &\equiv T_\alpha(\kappa^a,\kappa^\beta) \end{aligned}$$

Let us now focus on $\kappa^a \leq i < \kappa^\beta$. We have

$(IC1_\varnothing) U(t_i,\kappa^a,n-\kappa^\beta+1) \geqslant U(t_i,\kappa^a+1,n-\kappa^\beta+1)-c$

$(IC2_\varnothing) U(t_i,\kappa^a,n-\kappa^\beta+1) \geqslant U(t_i,\kappa^a,n-\kappa^\beta+2)-c$

$(IC1_\varnothing)$ becomes

$$t_i \leq \frac{c}{(1-q_\alpha)^{\kappa^a}(1-p_\alpha)q_\alpha}$$

and $(IC2_\varnothing)$ becomes

$$t_i \geq 1-\frac{c}{(1-q_\beta)^{n-\kappa^\beta+1}(1-p_\beta)q_\beta}$$

Thus, for $\kappa^a \leq i < \kappa^\beta$, we have

$$t_i \in \left[1-\frac{c}{(1-q_\beta)^{n-\kappa^\beta+1}(1-p_\beta)q_\beta},\frac{c}{(1-q_\alpha)^{\kappa^a}(1-p_\alpha)q_\alpha}\right) \equiv [\underline{T}_\varnothing(\kappa^a,\kappa^\beta),\overline{T}_\varnothing(\kappa^a,\kappa^\beta))$$

Finally, for $\kappa^\beta \leq i \leq n$, we have

$(IC1_\beta)$ $U(t_i,\kappa^a,n-\kappa^\beta+1)-c \geqslant U(t_i,\kappa^a+1,n-\kappa^\beta)-c$

and

$(IC2_\beta)$ $U(t_i, \kappa^a, n-\kappa^\beta+1) - c \geqslant U(t_i, \kappa^a, n-\kappa^\beta)$

$(IC1_\beta)$ becomes

$$t_i \leq \frac{(1-q_\beta)^{n-\kappa^\beta}(1-p_\beta)q_\beta}{(1-q_\beta)^{n-\kappa^\beta}(1-p_\beta)q_\beta + (1-q_\alpha)^{\kappa^a}(1-p_\alpha)q_\alpha}$$

and $(IC2_\beta)$ becomes

$$t_i \leq \frac{(1-q_\beta)^{n-\kappa^\beta}(1-p_\beta)q_\beta - c}{(1-q_\beta)^{n-\kappa^\beta}(1-p_\beta)q_\beta}$$

Thus,

$$\begin{aligned} t_{\kappa^\beta} &\leq \min\left\{\frac{(1-q_\beta)^{n-\kappa^\beta}(1-p_\beta)q_\beta - c}{(1-q_\beta)^{n-\kappa^\beta}(1-p_\beta)q_\beta}; \frac{(1-q_\beta)^{n-\kappa^\beta}(1-p_\beta)q_\beta}{(1-q_\beta)^{n-\kappa^\beta}(1-p_\beta)q_\beta + (1-q_\alpha)^{\kappa^a}(1-p_\alpha)q_\alpha}\right\} \\ &\equiv T_\beta(\kappa^a, \kappa^\beta) \end{aligned}$$

■

Proof of Proposition 4. Suppose an agent of taste t is added, so that the new group consists of tastes $\{t_1, ..., t_n, t\}$. Denote by κ' the new equilibrium. From the equilibrium incentive constraints, it follows directly that $\kappa' \in \{\kappa, \kappa+1\}$.

An agent with tastes t^P strictly prefers κ to $\kappa+1$ source α signals if and only if:$U(t^P, \kappa, n+1-\kappa) > U(t^P, \kappa+1, n-\kappa)$. That is,

$$\begin{aligned} & t^P\left[1-(1-q_\alpha)^\kappa(1-p_\alpha)\right] + (1-t^P)\left[1-(1-q_\beta)^{n+1-\kappa}(1-p_\beta)\right] \\ > \ & t^P\left[1-(1-q_\alpha)^{\kappa+1}(1-p_\alpha)\right] + (1-t^P)\left[1-(1-q_\beta)^{n-\kappa}(1-p_\beta)\right]. \end{aligned}$$

$$\Leftrightarrow \left(1-t^P\right)(1-p_\beta)(1-q_\beta)^{n-\kappa}q_\beta - t^P(1-p_\alpha)(1-q_\alpha)^\kappa q_\alpha > 0.$$

$$\Leftrightarrow t^P < \frac{(1-p_\beta)(1-q_\beta)^{n-\kappa}q_\beta}{(1-p_\alpha)(1-q_\alpha)^\kappa q_\alpha + (1-p_\beta)(1-q_\beta)^{n-\kappa}q_\beta} = t^{n+1}(\kappa+1).$$

Suppose then that $t^P < t^{n+1}(\kappa+1)$ so that the deciding agent prefers κ source α signals and $n+1-\kappa$ source β signals. If $t_{\kappa+1} < t^{n+1}(\kappa+1)$ then, since $t_1, ..., t_\kappa \geqslant t^n(\kappa) > t^{n+1}(\kappa)$, any $t < t^{n+1}(\kappa+1)$ would be optimal. If $t_{\kappa+1} \geqslant t^{n+1}(\kappa+1)$, whatever t is, at least $\kappa+1$ agents (and therefore, exactly $\kappa+1$ agents) choose source α. It follows that any t would be optimal.

Similarly, suppose $t^P > t^{n+1}(\kappa+1)$ so that the deciding agent prefers $\kappa+1$ source α signals and $n-\kappa$ source β signals. If $t_{\kappa+1} \geqslant t^{n+1}(\kappa+1)$ then for any t at least (and, therefore precisely) $\kappa+1$ agents would choose the source α signal and any t would be optimal. If $t_{\kappa+1} < t^{n+1}(\kappa+1)$, then any $t \geqslant t^{n+1}(\kappa+1)$ would be optimal.

When $t^P = t^{n+1}(\kappa+1)$, the deciding agent is indifferent among all potential additional members. ■