

# Supplementary Material

## Why Echo Chambers are Useful

Ole Jann

Nuffield College, University of Oxford

Christoph Schottmüller

University of Cologne and TILEC

August 28, 2018

### 1. Bipolar Polarization

Here we consider the bipolar case, i.e. the case where a player's bias  $b_i$  is either 0 or  $b > 0$ . We denote by  $n_0$  ( $n_b$ ) the number of people with bias 0 ( $b$ ) and without loss of generality we let  $n_0 \geq n_b$ . Recall that a player's expected payoff in a given room configuration can be written as

$$U_i = -\alpha \sum_{j \neq i} \{(b_j - b_i)^2\} - 1/4 [n + \alpha(n-1)n] + (1/4 - p(1-p)) \left[ \zeta_i + \alpha \sum_{j \neq i} \zeta_j \right]. \quad (1)$$

where  $\zeta_i$  is the number of pieces of information player  $i$  has access to (through truthful messages or by observing his signal).

#### 1.1. Segregation and full information as equilibrium

Now we ask the question when segregation, i.e. all players with bias  $b_i = 0$  choosing room 1 and all players with bias  $b_i = b$  choosing room 2, is an equilibrium. Clearly, every player is truthtelling in the most informative messaging equilibrium in this case and player  $i$ 's expected payoff is

$$U_i^{fs} = -\alpha n_{-i} b^2 - 1/4 [n + \alpha(n-1)n] + (1/4 - p(1-p)) [(1-\alpha)n_i + \alpha n_i^2 + \alpha n_{-i}^2] \quad (2)$$

where  $n_i$  is the number of players with the same bias as player  $i$  and  $n_{-i}$  is the number of players with the other bias.

To analyze when segregation is an equilibrium consider the incentives of player  $i$  to unilaterally deviate in his room choice. There are three relevant cases: (i) after the deviation everyone including  $i$  still sends truthful messages, (ii) after the deviation  $i$  will babble but the players in the room he is switching to still send truthful messages and

(iii)  $i$ 's room switch leads to babbling by all players in the room he is switching to. We analyze  $i$ 's incentives to switch as well as conditions for each of these cases in turn.

First, truthful messages after switch. This occurs if and only if  $b \leq p - 1/2$ . If  $i$ 's bias group is the (weakly) smaller group, then  $i$  will benefit from a room switch in this case, see equation (2). If  $i$ 's group is the larger group, he will not. Hence, segregation does not exist as equilibrium if  $b$  is less than  $p - 1/2$ .

Second,  $i$  babbles after the switch but everyone else remains truthtelling. This happens if and only if  $p - 1/2 < b \leq n_{-i}(p - 1/2)$ . In this case,  $i$ 's deviation payoff is

$$U_i^d = -\alpha n_{-i} b^2 - 1/4 [n + \alpha(n - 1)n] + (1/4 - p(1 - p)) [1 + n_{-i} + \alpha n_{-i}^2 + \alpha(n_i - 1)^2]$$

which is higher than  $U_i^{fs}$  if and only if  $n_{-i} > (1 + \alpha)(n_i - 1)$ . Clearly, the members of the smaller group are the ones for who this constraint is more stringent. That is, a segregation equilibrium exists given  $p - 1/2 < b \leq n_{-i}(p - 1/2)$  if and only if the size of the bigger group is at most  $1 + \alpha$  times the size of the smaller group minus 1. Intuitively, members of the smaller group have a lot of information to gain from a switch if the larger group is very large. Hence, a segregation equilibrium only exists if the larger group is not too large.

Third, the switch leads to complete babbling in the room switched to. This will occur if and only if  $b > n_{-i}(p - 1/2)$ . In this case, it is straightforward to see that the switch is unprofitable and a segregation equilibrium exists. This yields the following result.

**Proposition 13.** *Without loss of generality assume  $n_0 \geq n_b$ . A segregation equilibrium exists if and only if one of the two following condition is met: (i)  $b > n_0(p - 1/2)$ , (ii)  $p - 1/2 < b \leq n_0(p - 1/2)$  and  $n_0 \leq (1 + \alpha)(n_b - 1)$ .*

Next we check for which parameter values a full information equilibrium, i.e. all players choosing the same room and report their signal truthfully, exists. Recall that  $n_0 \geq n_b$ . Then, truthful reporting is an equilibrium of the messaging game if and only if  $b(1 - (n_b - 1)/(n - 1)) \leq p - 1/2$ . In the full information equilibrium, each player's expected payoff is

$$U_i^{fi} = -\alpha \sum_{j \neq i} \{(b_j - b_i)^2\} - 1/4 [n + \alpha(n - 1)n] + (1/4 - p(1 - p)) [(1 - \alpha)n + \alpha n^2].$$

Unilateral deviation by player  $i$  in his room choice can lead to two scenarios: either everyone still reports their message truthfully in the room  $i$  is deviating from (but  $i$  himself is no longer able to receive any messages) or not. In either case,  $i$ 's payoff is clearly lower after the deviation than before. Hence, a full information equilibrium exists if and only if  $b(1 - (n_b - 1)/(n - 1)) \leq p - 1/2$  which is equivalent to  $bn_0/(n - 1) \leq p - 1/2$ . This is summarized in figure 1.<sup>1</sup>

<sup>1</sup>Note that  $(n - 1)/n_0 \leq n_0$  as  $n_0 \geq n/2$  and  $n \geq 2$ .

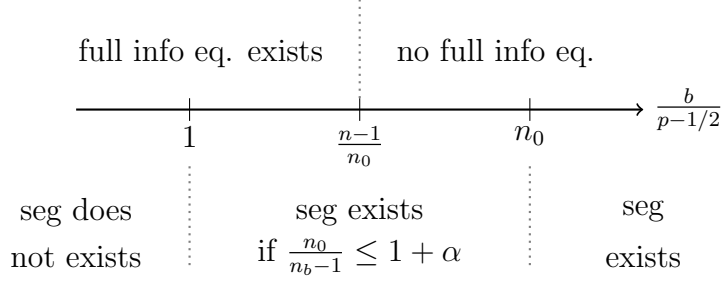


Figure 1: Segregation and full information as room choice equilibria

## 1.2. Welfare optimal room allocation

We will now consider how a planner would assign players to rooms in order to maximize the sum of players' payoffs. The planner's only tool is room assignment knowing the players' biases, i.e. the planner does not observe signals and cannot influence the messages sent or actions taken by the players.

From equation (1), it is clear that the planner's objective is equivalent to maximizing  $\sum_i \zeta_i$ , i.e. the total number of pieces of information by all players. We proceed in a number of lemmas. To avoid case distinction, we use in the following the convention that a player who is alone in a room will send a truthful message.

**Lemma 4.** *There is a welfare optimal room assignment without a room in which all players babble.*

**Proof of lemma 4:** In such a room there would be at least one player of each bias type. Splitting the room into two according to bias type would lead to weakly more transmitted pieces of information.  $\square$

**Lemma 5.** *Assume there are two rooms  $R_1$  and  $R_2$  such that in equilibrium players with bias  $x \in \{0, b\}$  send truthful messages in both rooms. Then players with bias  $x$  will also send truthful messages in a merged room  $R_1 \cup R_2$ .*

**Proof of lemma 5:** As players send truthful messages by assumption in room  $R_i$  for  $i \in \{1, 2\}$ ,

$$\frac{b}{p-1/2} n_{R_i, -x} \leq n_{R_i} - 1$$

has to hold where  $n_{R_i, -x}$  is the number of players in room  $R_i$  who do not have bias  $x$ . Summing this inequality over the two rooms  $i = 1, 2$  yields

$$\frac{b}{p-1/2} (n_{R_1, -x} + n_{R_2, -x}) \leq n_{R_1} + n_{R_2} - 2$$

which is sufficient for

$$\frac{b}{p-1/2} (n_{R_1, -x} + n_{R_2, -x}) \leq n_{R_1} + n_{R_2} - 1.$$

The latter inequality states that it is optimal for players with bias  $x$  to send truthful messages in a room  $R_1 \cup R_2$ .  $\square$

**Corollary 1.** *The welfare maximizing room assignment will not include rooms  $R_1$  and  $R_2$  such that either*

- *both  $R_1$  and  $R_2$  are both populated exclusively by players with bias  $x \in \{0, b\}$ , or*
- *both  $R_1$  and  $R_2$  are populated by players with both biases and all players send truthful messages, or*
- *$R_1$  and  $R_2$  are populated by players with both biases and only the players with bias  $x \in \{0, b\}$  send truthful messages, or*
- *$R_1$  is exclusively populated by players with bias  $x \in \{0, b\}$  and  $R_2$  is populated by players with both biases but only players with bias  $x$  send truthful messages in  $R_2$ .*

**Proof of corollary 1:** In each of these cases, merging the two rooms maintains truthtelling incentives for those that originally sent truthful messages by lemma 5. As the truthful messages are received by more players, merging clearly increases the planner's objective.  $\square$

**Lemma 6.** *The welfare maximizing room assignment will not include rooms  $R_1$  and  $R_2$  such that both rooms contain players of each bias and all players in  $R_1$  send truthful messages while only players of bias  $x \in \{0, b\}$  send truthful messages in  $R_2$ .*

**Proof of lemma 6:** First, we consider the case where more players are in room  $R_1$ , i.e.  $n_{R_1,x} + n_{R_1,-x} \geq n_{R_2,x} + n_{R_2,-x}$ . Note that the number of pieces of information generated in those two rooms is  $n_{R_1}^2 + n_{R_2,x}(n_{R_2,x} + n_{R_2,-x})$ . Note that  $n_{R_2,x} > n_{R_2,-x}$  as otherwise  $x$  would not be truthtelling in  $R_2$  while  $-x$  is not. Consider now an alternative room assignment that differs from the original one in the way that  $n_{R_2,-x}$  players of each bias are moved from  $R_2$  to  $R_1$ . Denote everything after the change using  $\tilde{\cdot}$ . That is,  $\tilde{R}_2$  will contain  $n_{R_2,x} - n_{R_2,-x}$  players of bias  $x$  and none of bias  $-x$  while  $\tilde{R}_1$  will contain  $n_{R_1,x} + n_{R_2,-x}$  players of bias  $x$  and  $n_{R_1,-x} + n_{R_2,-x}$  of type  $-x$ . The crucial result is that all players in room  $\tilde{R}_1$  find truthtelling optimal: As truthtelling was optimal for players of both biases by assumption in  $R_1$ ,

$$\frac{b}{p - 1/2} \leq \frac{n_{R_1} - 1}{\max\{n_{R_1,x}, n_{R_1,-x}\}}.$$

Now note that

$$\frac{n_{R_1} - 1}{\max\{n_{R_1,x}, n_{R_1,-x}\}} \leq \frac{n_{R_1} - 1 + 2n_{R_2,-x}}{\max\{n_{R_1,x} + n_{R_2,-x}, n_{R_1,-x} + n_{R_2,-x}\}}$$

as the latter fraction is increasing in  $n_{R_2,-x}$  and therefore truthtelling is still optimal in  $\tilde{R}_1$ . Clearly, truthtelling is also optimal in  $\tilde{R}_2$  as only players with bias  $x$  are left there. Consequently, the total number of pieces of information in  $\tilde{R}_1$  and  $\tilde{R}_2$  is  $(n_{R_1} + 2n_{R_2,-x})^2 + (n_{R_2,x} - n_{R_2,-x})^2$  which, by  $n_{R_1} \geq n_{R_2} > n_{R_2,x}$ , is strictly greater than the number of pieces of information in rooms  $R_1$  and  $R_2$ . Hence, the planner prefers the room assignment  $\tilde{R}_1, \tilde{R}_2$  over the room assignment  $R_1$  and  $R_2$  (keeping room assignment for players not in those rooms fixed).

Second, consider the case where more players are in room  $R_2$ , i.e.  $n_{R_1,x} + n_{R_1,-x} < n_{R_2,x} + n_{R_2,-x}$ . In this case, we argue that merging the two rooms to  $R_1 \cup R_2$  will yield more information than keeping them separate. By lemma 5, players with bias  $x$  will still be truthtelling in room  $R_1 \cup R_2$ . Assume that players with bias  $-x$  will not tell the truth in  $R_1 \cup R_2$  (otherwise merging the two rooms is clearly optimal). Note that this implies  $n_{R_1,x} + n_{R_2,x} > n_{R_1,-x}, n_{R_2,-x}$  as players with bias  $x$  tell the truth in  $R_1 \cup R_2$  and players with bias  $-x$  do not. The number of pieces of information generated in  $R_1 \cup R_2$  is then  $(n_{R_1,x} + n_{R_2,x})(n_{R_1,x} + n_{R_2,x} + n_{R_1,-x} + n_{R_2,-x})$ . This is greater than the number of pieces of information generated in  $R_1$  and  $R_2$  separately as

$$\begin{aligned} & n_{R_1,x}^2 + 2n_{R_1,x}n_{R_2,x} + n_{R_1,x}n_{R_1,-x} + n_{R_1,x}n_{R_2,-x} + n_{R_2,x}^2 + n_{R_2,x}n_{R_1,-x} + n_{R_2,x}n_{R_2,-x} \\ & > n_{R_1,x}^2 + 2n_{R_1,x}n_{R_1,-x} + n_{R_1,-x}^2 + n_{R_2,x}^2 + n_{R_2,x}n_{R_2,-x} \\ & \Leftrightarrow 2n_{R_1,x}n_{R_2,x} - n_{R_1,x}n_{R_1,-x} + n_{R_1,x}n_{R_2,-x} + n_{R_2,x}n_{R_1,-x} - n_{R_1,-x}^2 > 0 \\ & \Leftrightarrow -n_{R_1,-x}n_{R_1} + n_{R_1,x}n_{R_2} + n_{R_2,x}n_{R_1} > 0 \end{aligned}$$

which holds true by  $n_{R_1} < n_{R_2}$  and  $n_{R_1,x} + n_{R_2,x} > n_{R_1,-x}, n_{R_2,-x}$ . Consequently, the planner would prefer  $R_1 \cup R_2$  to  $R_1$  and  $R_2$  separately.  $\square$

**Lemma 7.** *In a welfare maximizing room allocation, the following cannot occur: There are three rooms  $R_0$  populated only of players with bias 0,  $R_b$  populated only with players of bias  $b$  and  $R_m$  populated with players of both biases.*

**Proof of lemma 7:** If  $R_m$  induces babbling, it is clearly better to assign the bias 0 ( $b$ ) players in there to  $R_0$  ( $R_b$ ) instead. If only players with bias  $x \in \{0, b\}$  send truthful messages in  $R_m$ , then it is clearly better to merge this room with  $R_x$  which maintains truthtelling incentives for players with bias  $x$ , see lemma 5.

Hence, we only have to consider the case where players with both biases send truthful messages in  $R_m$ . For concreteness assume there are more players in  $R_0$  than in  $R_b$  denoted by  $n_0 \geq n_b$  (the reverse case is analyzed analogously). We consider two cases in turn. First, consider  $n_0 \geq n_m + n_b$ . We claim that merging  $R_m$  and  $R_0$  will then lead to more information than keeping these rooms separate. By lemma 5, player with bias 0 will still be truthtelling in  $R_0 \cup R_m$  and the number of pieces of information generated by rooms  $R_b$

and  $R_0 \cup R_m$  is  $n_b^2 + (n_0 + n_{m,0})(n_m + n_0)$  which is strictly higher than the number of pieces of information generated by  $R_0$ ,  $R_b$  and  $R_m$ , i.e.  $n_b^2 + n_0^2 + n_m^2$ , as  $n_0 \geq n_m + n_b > n_m$ .

Second, consider  $n_0 < n_m + n_b$ . The following change in the room allocation creates more information: Move  $n_b$  players from  $R_b$  and  $n_b$  players from  $R_0$  to  $R_m$ . This leaves no one in  $R_b$ ,  $n_m + 2n_b$  in  $R_m$  and  $n_0 - n_b$  players in  $R_0$ . As in the proof of lemma 6, the move maintains truthtelling incentives for players in  $R_m$ . The number of pieces of information generated after the move is  $(n_m + 2n_b)^2 + (n_0 - n_b)^2 = n_m^2 + 4n_b n_m + 5n_b^2 + n_0^2 - 2n_0 n_b$  which is higher than the number of pieces of information generated by  $R_0$ ,  $R_b$  and  $R_m$  without the move, i.e.  $n_0^2 + n_m^2 + n_b^2$ , by  $n_0 < n_b + n_m$ .  $\square$

**Corollary 2.** *The welfare optimal room assignment consists of at most two rooms and at most one room in which players of both biases are present.*

**Proof of corollary 2:** This follows from the combination of corollary 1 and lemmas 4, 6 and 7.  $\square$

Corollary 2 (and the preceding lemmas) leave the following possibilities for welfare optimal room assignment:

- segregation: each bias group has its own exclusive room.
- full integration: one room for all players
- mix: one room exclusively with players of bias  $x$  and one room with both bias types in which either
  - only players of type  $-x$  send truthful messages
  - all players send truthful messages.

The idea behind the mix situation is the following: If there are more of players with bias  $-x$  than players with bias  $x$  and assigning all player to one room would lead to babbling, then it can be optimal not to separate completely but to assign some players of the minority  $x$  to the majority room  $-x$ . The  $x$  players will babble there but they receive a lot of information (truthful messages of all the  $-x$  players). If such a situation is optimal, then clearly it must be the case that assigning one more minority player to the mixed room would lead to babbling. This is the first mixed assignment possibility.

The second mixed assignment possibility refers to a situation where the minority bias group would babble if all players were in one room while the majority would send truthful messages. In this case, taking some players of the majority to a separate room can restore truthtelling by both groups in the mixed room and might be optimal.

As should be clear from the discussion above, the mixed scenario is only optimal if the group sizes differ. With equal group size either segregation or full integration is optimal. Which of the two is optimal depends on whether the bias difference  $b$  is sufficiently small to obtain truthtelling in a fully integrated room.

**Proposition 14.** *Let the number of players  $n$  be strictly greater than 2 and let the two bias groups be of equal size, i.e.  $n_b = n_0 = n/2$ . Then assigning all players to the same room is welfare optimal if  $b/(p-1/2) \leq 2(n-1)/n$ . Otherwise, segregation, i.e. assigning all players with bias 0 to a room  $R_1$  and all players with bias  $b$  to a room  $R_2$  is welfare maximizing.*

**Proof of proposition 14:** If  $b/(p-1/2) \leq 2(n-1)/n$ , then a fully integrated room leads to truthtelling and is therefore clearly optimal. Therefore, consider  $b/(p-1/2) > 2(n-1)/n$ . One fully integrated room leads then to babbling by all players and full integration is therefore worse than segregation. Note furthermore that the alternatives with a mixed room are not feasible with  $n_b = n_0$  and  $b/(p-1/2) > 2(n-1)/n$ : A mixed room containing all players of bias  $x$  will lead to babbling of players with bias  $-x$  by  $b/(p-1/2) > 2(n-1)/n$ . Hence, the scenario of a mixed room in which all players tell the truth is not feasible. A mixed room containing all players of bias  $x$  and babbling of those players in the mixed room with bias  $-x$  might be feasible is, however, clearly worse than segregation as it creates less pieces of information:  $n/2(n/2+n_{m,-x})+(n/2-n_{m,-x}) < 2(n/2)^2$  for  $n_{m,-x} > 0$ . All other room assignments cannot be optimal by corollary 2.  $\square$

Taking the previous proposition and the earlier results on the existence of segregation and full integration equilibrium together yields the following result.

**Proposition 15.** *Let  $n_0 = n_b = n/2$  and  $n > 2$ . Then the welfare optimal room assignment is an equilibrium if  $\alpha \geq 2/(n-2)$ . If  $\alpha < 2/(n-2)$ , then the welfare optimal room assignment is still an equilibrium if  $b/(p-1/2) \notin (2(n-1)/n, n/2)$ .*

Now turn to the case where  $n_0 > n_b$ . If  $b/(p-1/2) \leq (n_0 + n_b - 1)/n_0$ , then full information, i.e. all players in one room and every player sends a truthful message, is feasible and a single room is obviously welfare optimal.

If  $(n_0 + n_b - 1)/n_0 < b/(p-1/2) \leq (n_0 + n_b - 1)/n_b$ , then players with bias  $b$  would babble in a single room. It is important to notice that segregation cannot be optimal in this situation: segregation leads to  $n_b^2 + n_0^2$  pieces of information while in one single room there would still be  $n_0 * (n_0 + n_b)$  pieces of information which is greater than  $n_b^2 + n_0^2$  by  $n_0 > n_b$ .<sup>2</sup> Consequently, there are two options for the welfare maximal room assignment. Either all players are in one single room and babbling by players with bias  $b$  is tolerated or some players with bias 0 are assigned to a separate room in order to balance the mixed room and restore truthtelling incentives for the bias  $b$  players. The maximal number of bias 0 players that can remain in the mixed room without inducing babbling by bias  $b$  players is  $n_{m0} = \lfloor (n_b - 1)/(b/(p-1/2) - 1) \rfloor$ . Consequently, the number of pieces of information in the scenario with  $n_{m0}$  bias 0 players and all bias  $b$  players in one room and  $n_0 - n_{m0}$  bias 0 players in a separate room is  $(n_b + n_{m0})^2 + (n_0 - n_{m0})^2$ . Whether this is higher or

---

<sup>2</sup>Similarly, it is not optimal to put only some bias  $b$  players into their own room: As  $n_b \leq n_0$ , they would have less information there than they would have if they were babbling in one single room.

lower than the number of pieces in a fully intergrated room, i.e.  $n_0(n_0 + n_b) + n_b$ , depends on the parameters. In particular, if  $b/(p - 1/2)$  is close to the lower boundary (and  $n_b$  is not too small), two rooms will be optimal as  $n_{m0}$  will be relatively high. If  $b/(p - 1/2)$  is close to the upper boundary, however,  $n_{m0}$  will be low and one room with babbling will be optimal.

Finally, we consider  $b/(p - 1/2) > (n_0 + n_b - 1)/n_b$ . In this case, even the bias 0 players will babble if all players are in one room. This implies that putting some bias zero players in an own separate room will no longer help: All the remaining bias zero players would have even higher incentives to babble and it would be more informative to fully separate the two bias groups. Consequently, only two options remain: Either the just mentioned segregation or enough bias  $b$  players are assigned to an own separate room to restore truthtelling incentives for the bias 0 players in the mixed room. The maximum number of bias  $b$  players that can remain in the mixed room without destroying truthtelling by bias 0 players is  $n_{mb} = \lfloor (n_0 - 1)/(b/(p - 1/2) - 1) \rfloor$ . This yields  $n_0(n_0 + n_{mb}) + n_{mb} + (n_b - n_{mb})^2$  pieces of information which can, depending on the parameters, be higher or lower than the  $n_0^2 + n_b^2$  pieces of information created by segregation. What is important to notice is that for  $b/(p - 1/2)$  high enough  $n_{bm} = 0$  and in this case segregation is always welfare optimal.

We summarize the preceding discussion in the following proposition.

**Proposition 16.** *Let  $n_0 \geq n_b$ . The welfare maximizing room allocation is as follows:*

1. *For  $b/(p - 1/2) \leq (n - 1)/n_0$ , all players are in one room (where their messages are truthtelling).*
2. *For  $(n - 1)/n_0 < b/(p - 1/2) \leq (n - 1)/n_b$ , let  $n_{m0} = \lfloor (n_b - 1)/(b/(p - 1/2) - 1) \rfloor$ . Then,*
  - (a) *all players in one room (where those with  $b_i = 0$  are truthtelling and those with  $b_i = b$  are babbling) is optimal if  $(n_b + n_{m0})^2 + (n_0 - n_{m0})^2 \leq n_0(n_0 + n_b) + n_b$ ,*
  - (b) *one room with  $n_0 - n_{m0}$  players with  $b_i = 0$  and another room with all other players (in which everyone is truthtelling) is optimal.*
3. *For  $b/(p - 1/2) > (n - 1)/n_b$ , let  $n_{mb} = \lfloor (n_0 - 1)/(b/(p - 1/2) - 1) \rfloor$ . Then,*
  - (a) *one room with  $n_b - n_{mb}$  players with  $b_i = b$  and another room with all other players is optimal if  $n_0(n_0 + n_{mb}) + n_{mb} + (n_b - n_{mb})^2 \geq n_0^2 + n_b^2$*
  - (b) *full segregation is optimal otherwise.*

Note that an increase in  $b/(p - 1/2)$  will decrease both  $n_{m0}$  and  $n_{mb}$ . This implies that we can step monotonically through the different cases of proposition 16 as  $b/(p - 1/2)$  increases. However, depending on parameter values, in particular  $n_0$  and  $n_b$ , some of the



cases may be skipped. However, the first and last case will never be skipped as they always occur for sufficiently low, respectively high, values of  $b/(p - 1/2)$ .

### 1.3. Welfare optimal room assignment as equilibrium

It remains to check when the welfare optimal room assignments correspond to equilibria of the game. If full information is feasible, i.e. if  $b/(p - 1/2) \leq (n_0 + n_b - 1)/n_0$ , then it is clearly an equilibrium. For  $(n_0 + n_b - 1)/n_0 < b/(p - 1/2) \leq (n_0 + n_b - 1)/n_b$ , the welfare optimal room assignment is definitely an equilibrium if it is optimal to keep all players in the same room (and have the smaller group babbling): The point is that this can only be optimal if isolating one player of the larger group would not lead to truthtelling in a room with all other players. (If this was the case, then isolating one player of the larger group would be optimal.) But this implies that any deviation in room choice from the situation with one big mixed room will lead to less information for the deviating player and less (or the same) information for the other players. Following (1), such a deviation will therefore reduce the deviating player's expected payoff. Hence, the welfare maximizing room assignment is an equilibrium if  $(n_0 + n_b - 1)/n_0 < b/(p - 1/2) \leq (n_0 + n_b - 1)/n_b$  and  $(n_b + n_{m0})^2 + (n_0 - n_{m0})^2 \leq n_0(n_0 + n_b)$ . For the case  $(n_b + n_{m0})^2 + (n_0 - n_{m0})^2 > n_0(n_0 + n_b)$  where two rooms are optimal – one mixed room with truthtelling by all players and one with only players of the larger group – the relevant question is whether the  $n_0 - n_{m0}$  players in the room only for bias 0 players would want to deviate to the mixed room. Note that this deviation will destroy truthtelling by the bias  $b$  players. However, a unilaterally deviating player will obtain information from  $n_{m0}$  other players which is more information than the  $n_0 - n_{m0} - 1$  truthful messages he obtains when not deviating.<sup>3</sup> It is therefore unsurprising that the welfare optimal room assignment will only be an equilibrium if  $\alpha$  is sufficiently high. Using (1), the deviation will not be profitable if and only if it decreases  $\zeta_i + \alpha \sum_{j \neq i} \zeta_j$ . This is the case if and only if

$$\begin{aligned} n_{m0} + 1 + \alpha \left( (n_0 - n_{m0} - 1)^2 + (n_{m0} + 1)(n_b + n_{m0}) \right) \\ \leq n_0 - n_{m0} + \alpha \left( (n_0 - n_{m0} - 1)(n_0 - n_{m0}) + (n_{m0} + n_b)^2 \right). \\ \Leftrightarrow \alpha \geq \frac{2n_{m0} - n_0 + 1}{n_0 - 2n_{m0} - 1 + n_b^2 + n_b(n_{m0} - 1)}. \end{aligned} \quad (3)$$

Hence, if  $(n_0 + n_b - 1)/n_0 < b/(p - 1/2) \leq (n_0 + n_b - 1)/n_b$  and  $(n_b + n_{m0})^2 + (n_0 - n_{m0})^2 > n_0(n_0 + n_b)$ , the welfare optimal room assignment is an equilibrium if and only if (3) holds.

For  $b/(p - 1/2) > (n_0 + n_b - 1)/n_b$ , the welfare optimal room assignment can be either

---

<sup>3</sup>To see this, first note that  $n_{m0} \geq n_b$  as no truthtelling mixed room exists if bias  $b$  players were not truthtelling in a perfectly balanced, i.e. same number of players from each bias group, room. Then note that  $(n_b + n_{m0})^2 + (n_0 - n_{m0})^2 > n_0(n_0 + n_b)$  is more demanding for higher  $n_0$  (keeping other parameters fixed). It is straightforward to calculate that the inequality cannot be satisfied (given  $n_b \leq n_{m0}$ ) for  $n_0 \geq 5n_{m0}/3$ . Hence,  $n_{m0} \geq 4n_0/5$  whenever this room assignment is welfare maximizing which implies  $n_{m0} > n_0 - n_{m0} - 1$ .

a mixed room combined with a room exclusively for the smaller group or segregation. In the first case it is straightforward to see that this is an equilibrium. Recall that moving one more player from the separate room to the mixed room would lead to babbling by all players in the mixed room. Clearly, such a deviation is not profitable. In the second case, we already showed before that total separation is an equilibrium if either  $b/(p-1/2) \geq n_0$  or  $\alpha \geq (1+n_0-n_b)/(n_b-1)$ .

We summarize the prior discussion in the following proposition.

**Proposition 17.** *Let  $n_0 \geq n_b$ . The welfare optimal room allocation is an equilibrium in cases (1), (2a) and (3a) of proposition 16. The welfare optimal room allocation is also an equilibrium*

- in case (2b) if

$$\alpha \geq \frac{2n_{m0} - n_0 + 1}{n_0 - 2n_{m0} - 1 + n_b^2 + n_b(n_{m0} - 1)}$$

- in case (3b) if  $\alpha \geq (1+n_0-n_b)/(n_b-1)$ .

If the welfare optimal room allocation is not an equilibrium, then the welfare optimal equilibrium features too little segregation: In case (2b), the equilibrium has all players in a single room while the welfare optimal allocation has two rooms. In case (3b), the welfare optimal room allocation is full segregation which is not an equilibrium.

## 2. Symmetrically single Peaked bias distribution

We now move to symmetrically, single peaked distribution of biases: Assume that biases are on an equally spaced grid  $0, d, 2d, \dots, Kd$  for some  $d > 0$  and  $K \in \mathbb{N}$ . The number of players with bias  $b_i = kd$  is increasing up to  $Kd/2$  and decreasing thereafter. Furthermore, we assume that the number of players with bias  $kd$  equals the number of players with bias  $(K-k)d$  for  $k = 0, 1, \dots, \lfloor K/2 \rfloor$ .

We are going to show a result similar to the one for uniformly distributed biases in the paper. Namely, that a fully integrated room is welfare optimal and an equilibrium under certain conditions.

To state the proposition we need the following notation: Let  $\underline{k}$  be the lowest  $k$  such that  $kd \geq Kd/2 - (p-1/2)(n-1)/n$  and let  $\bar{k}$  be the highest  $k$  such that  $kd \leq Kd/2 + (p-1/2)(n-1)/n$ . Note that due to the discreteness of the grid and following theorem 1, the truthtelling interval in a fully integrated room will cover all players with  $b_i \in [\underline{k}d, \bar{k}d]$ .

**Proposition 18.** *With a symmetric, single peaked distribution of biases, one room containing all players is welfare optimal and also an equilibrium if*

$$\bar{k}d - \underline{k}d + d > (2p-1) \frac{n-2}{n-1}. \quad (4)$$

**Proof of proposition 18:** Theorem 1 states that in the most informative equilibrium of the messaging subgame players in room  $R$  will tell the truth if and only if  $b_i \in \left[ \bar{b} - \frac{n_R-1}{n_R}(p - \frac{1}{2}), \bar{b} + \frac{n_R-1}{n_R}(p - \frac{1}{2}) \right]$ . If this interval covers  $[0, Kd]$ , then one room leads to truthtelling by all players and one single room is clearly optimal. In the remainder of this proof, we therefore assume that this is not the case. Note that – holding  $\bar{b}$  fixed – the length of the interval is increasing in  $n_R$ . If we turn to the case of one fully integrated room, then the truthtelling interval is  $\left[ Kd/2 - \frac{n-1}{n}(p - \frac{1}{2}), Kd/2 + \frac{n-1}{n}(p - \frac{1}{2}) \right]$  as  $\bar{b} = Kd/2$ . We will first show the result under a condition slightly stronger than (4), namely under the condition

$$\bar{k}d - \underline{k}d + d > (2p - 1) \frac{n - 1}{n}. \quad (5)$$

Condition (5) states that the length of the truthtelling interval is less than  $\bar{k} - \underline{k} + d$ . (Note that the length of the truthtelling interval is weakly larger than  $\bar{k}d - \underline{k}d$  due to the discrete grid on which biases are distributed.) This implies that the truthtelling interval would not cover more grid points if it was moved up or down while keeping its length constant. As the truthtelling interval is shorter for any other room (because of  $n_R < n$ ) and the distribution of biases is single-peaked, this implies that there is no room in which more players are truthtelling than in the fully integrated room.

The same conclusion follows if (4) holds instead of (5): (4) states that the length of the truthtelling interval in any room different from the fully integrated room (which therefore contains at most  $n - 1$  players) is less than  $\bar{k}d - \underline{k}d + d$  which again implies that the truthtelling interval of such a room cannot cover more grid points than the fully integrated room and by single peakedness it can therefore also not contain more truthtelling players.

Let  $t^*$  be the maximal number of truthtelling players in any possible room. From the above,  $t^*$  is attained by the fully integrated room if (4) holds. In this case, the number of pieces of information generated in the fully integrated room is  $t^*n + n - t^*$ . We will show that no other room configuration generates more pieces of information: The total number of pieces of information in  $r$  rooms is:  $\sum_R t_R n_R + n_R - t_R = \sum_R t_R (n_R - 1) + n_R \leq \sum_R t^* (n_R - 1) + n_R = t^*(n - r) + n \leq t^*n + n - t^*$ . By proposition 1, one big room with all players is therefore welfare optimal if (4) holds.

In case a single fully integrated room is welfare maximal it is also an equilibrium: Unilateral seöf-isolation would lead to less information for the deviating player and also – by welfare optimality of the fully integrated room – to less information over all. By 1, the deviation is therefore unprofitable.  $\square$

### 3. Alternative signal technologies

In this section, we consider two variations of the model in the paper. The first is a straightforward extensions in which we allow for more signals than just the binary signal structure considered in the paper. (We can also allow for more states but this is relatively immaterial in our setting.) The second changes the signal structure such that no longer each player receives a signal about “his state”  $\theta_i$  as in the main text but instead all players receive a noisy signal about the same one-dimensional state  $\theta$ . For both variations we show that our main result that integration is optimal and an equilibrium if there is little polarization while segregation is optimal and an equilibrium if there is a lot of polarization continue to hold. The main shortcoming of the two variations is that for intermediate values of polarization it is no longer possible to determine the most informative equilibrium of the messaging game as we can no longer rule out that this equilibrium involves mixed strategies. This makes each variation less tractable than the model of the main paper.

#### 3.1. Larger signal and state space

Now allow for an arbitrary finite number of states, biases and signals. We keep the assumption that states and signals of different players are independent and that player  $i$  receives a signal that is partially informative about state  $\theta_i$  (but independent about all other states). We also keep the utility function, i.e. the additive structure. The message space equals the signal space and we assume that lower signals lead to a lower expected value of  $\theta_i$ . For notational simplicity let the signal be the posterior it leads to, i.e.  $\sigma_i = \mathbb{E}[\theta_i|\sigma_i]$ .

Following similar steps as in the main text, we can derive the expected utility difference between sending two messages labeled as high ( $h$ ) and low ( $l$ ). Let  $\mu_{ji}^h$  denote the expected value that  $j$  assigns to  $\theta_i$  upon receiving message  $h$  (given some equilibrium messaging strategy by  $i$ ). The expected utility difference can then, similarly to above, be derived as

$$\begin{aligned} \Delta U_i(\sigma_i) &= \sum_{j \in R_i, j \neq i} (\mu_{ji}^l)^2 - (\mu_{ji}^h)^2 + 2(\mu_{ji}^l - \mu_{ji}^h)(b_j - b_i - \mathbb{E}[\theta_i|\sigma_i]) \\ &= \sum_{j \in R_i, j \neq i} (\mu_{ji}^l - \mu_{ji}^h) [(\mu_{ji}^l + \mu_{ji}^h) + 2(b_j - b_i - \sigma_i)] \\ &= -2(n_{R_i} - 1) (\mu_{ji}^h - \mu_{ji}^l) \left[ \frac{\mu_{ji}^l + \mu_{ji}^h}{2} + \sum_{k \in R_i, k \neq i} \left\{ \frac{b_k}{n_{R_i} - 1} \right\} - b_i - \sigma_i \right] \end{aligned}$$

This expression implies that a truth-telling equilibrium exists if and only if for every player  $i$  and every  $\sigma_i^l < \sigma_i^h$

$$\sigma_i^l \leq \frac{\sigma_i^l + \sigma_i^h}{2} + \sum_{k \in R_i, k \neq i} \left\{ \frac{b_k}{n_{R_i} - 1} \right\} - b_i \leq \sigma_i^h$$

$$\Leftrightarrow \left| \sum_{k \in R_i, k \neq i} \left\{ \frac{b_k}{n_{R_i} - 1} \right\} - b_i \right| \leq \frac{\sigma^h - \sigma^l}{2}.$$

If we assume that all players have the same signal space, this condition is tightest for the player whose bias  $b_i$  is furthest away from the other players' average bias,  $\sum_{j \in R_i, j \neq i} b_j / (n_{R_i} - 1)$ , and for the two signals that are closest together.

It is immediate from the expression above that (i) truthtelling is impossible if bias differences are too high, (ii) adding moderates can establish truthtelling as it can move the average of the other players closer to each player's bias (e.g. consider a room with 2 people with differing biases, then adding a player with the average bias can only help). To state this formally, consider first the expected payoff of player  $i$  when choosing room  $R_i$  and expecting a given (e.g. equilibrium) room allocation:

$$\begin{aligned} & - \mathbb{E} \left[ \left( \sum_{j \in R_i^{truth}, j \neq i} (\mu_{ij} - \theta_j) + \sum_{j \notin R_i, j \in R_i^{bab}} (\bar{\mu}_j - \theta_j) \right)^2 \right. \\ & \quad + \alpha \sum_{j \in R_i, j \neq i} \left( b_j - b_i + \sum_{k \in R_i^{truth} \cup \{j\}} (\mu_{jk} - \theta_k) + \sum_{k \notin R_i, k \in R_i^{bab} \setminus \{j\}} (\bar{\mu}_k - \theta_k) \right)^2 \\ & \quad \left. + \alpha \sum_{j \notin R_i} \left( b_j - b_i + \sum_{k \in R_j^{truth} \cup \{j\}} (\mu_{jk} - \theta_k) + \sum_{k \notin R_i, k \in R_j^{bab} \setminus \{j\}} (\bar{\mu}_k - \theta_k) \right)^2 \right] \end{aligned}$$

where we denote  $\mathbb{E}[\theta_j]$  as  $\bar{\mu}_j$ , the set of players babbling in room  $R_j$  in the messaging equilibrium of the given room allocation as  $R_j^{bab}$  and the set of players sending truthful messages in room  $R_j$  in the messaging equilibrium of the given room allocation as  $R_j^{truth}$ . Note that most of the terms drop out in the expression above as signals are assumed to be independent and therefore  $\mathbb{E}[\mu_{ij} - \theta_j] = 0$  and also  $\mathbb{E}[(\mu_{ij} - \theta_j)(\mu_{ik} - \theta_k)] = 0$ . Consequently, the expression above can be rewritten as

$$\begin{aligned} & - \sum_{j \in R_i^{truth}, j \neq i} \mathbb{E} [(\mu_{ij} - \theta_j)^2] - \sum_{j \notin R_i, j \in R_i^{bab}} \mathbb{E} [(\bar{\mu}_j - \theta_j)^2] \\ & - \alpha \sum_{j \in R_i, j \neq i} (b_j - b_i)^2 - \alpha \sum_{j \in R_i, j \neq i} \sum_{k \in R_i^{truth} \cup \{j\}} \mathbb{E} [(\mu_{jk} - \theta_k)^2] - \alpha \sum_{j \in R_i, j \neq i} \sum_{k \notin R_i, k \in R_i^{bab} \setminus \{j\}} \mathbb{E} [(\bar{\mu}_k - \theta_k)^2] \\ & - \alpha \sum_{j \notin R_i} (b_j - b_i)^2 - \alpha \sum_{j \notin R_i} \sum_{k \in R_j^{truth} \cup \{j\}} \mathbb{E} [(\mu_{jk} - \theta_k)^2] - \alpha \sum_{j \notin R_i} \sum_{k \notin R_i, k \in R_j^{bab} \setminus \{j\}} \mathbb{E} [(\bar{\mu}_k - \theta_k)^2] \end{aligned}$$

As we cannot rule out mixed strategies, this expression will not simplify as neatly as in the main text. However, we can already see from here that a player's payoff is higher if another player is truthtelling than when he is babbling or mixing. This observation will be enough for our purposes.

To state our results we first introduce some notation. Let  $\underline{\sigma} = \min\{|\sigma^j - \sigma^k| : j \neq k, \sigma^j, \sigma^k \in \Sigma\}$  and  $\bar{\sigma} = \max\{|\sigma^j - \sigma^k| : j \neq k, \sigma^j, \sigma^k \in \Sigma\}$  and furthermore,  $\bar{b} = \max_i\{|nb_i - \sum_j b_j|\}$ . We will denote by  $\mathcal{B}_\eta$  the set of biases scaled by  $\eta$ ; that is, it contains all the elements  $\eta b_i$ . We will use this to talk about more spread out biases. If the set of biases is  $\mathcal{B}_\eta$  with  $\eta > 1$ , then biases are more spread out.

**Proposition 19.** *If  $\underline{\sigma} \geq 2\bar{b}/(n-1)$ , then a single room in which all players are truthtelling is both welfare maximizing and an equilibrium.*

*Let the set of biases be  $\mathcal{B}_\eta$  and fix all parameter values apart from  $\eta$ . Generically, full separation is welfare maximizing and an equilibrium if  $\eta$  is sufficiently high.*

**Proof of proposition 19:** Recall that a truthtelling equilibrium exists if and only if for all palyers  $i$   $\left| \sum_{k \neq i} \{b_k/(n-1)\} - b_i \right| \leq (\sigma^h - \sigma^l)/2$  for every  $\sigma^h > \sigma^l$  in  $\Sigma$ . This can be rewritten as  $|\sum_k \{b_k\} - nb_i|/(n-1) \leq (\sigma^h - \sigma^l)/2$ . The condition in the proposition ensures that this inequality holds for all players and all signals. Clearly, having all players in one room and telling the truth is welfare optimal whenever it is feasible.

If  $\left| \sum_{k \in R_i, k \neq i} \{b_k/(n-1)\} - b_i \right| > (\sigma^h - \sigma^l)/2$ , then  $i$  will not be truthful when receiving either signal  $\sigma^l$  or  $\sigma^h$ . Generically,  $\left| \sum_{k \in R_i, k \neq i} \{b_k/(n-1)\} - b_i \right| \neq 0$  for any room configuration containing players from more than one bias group. (This follows from the finiteness of players which obviously implies that the number of such room configurations is finite.) Now observe that the left hand side of the non-truthtelling inequality is scaled by  $\eta$  while the right hand side is not. That is, for  $\eta$  sufficiently high player  $i$  will report the highest (lowest) signal in  $\Sigma$  in all rooms in which  $\sum_{k \in R_i, k \neq j} b_k < n_{R_i} b_i$  ( $\sum_{k \in R_i, k \neq j} b_k > n_{R_i} b_i$ ). Put differently, any room that contains one or more players of a bias not equal to  $b_i$  will lead to totally uninformative messages by  $i$  if  $\eta$  is sufficiently high. For high enough  $\eta$ , this holds true for all players and it is then obvious that full separation is both welfare maximizing and an equilibrium.  $\square$

### 3.2. Single state

In this variation, a state of the world  $\theta \in \Theta$  is distributed according to distribution  $F$ . The state is unobserved but each player  $i$  out of  $n$  players receives a noisy signal  $\sigma_i \in \Sigma$  of the state where  $\sigma_i$  is conditional on  $\theta$  distributed according to  $G_\theta$ . The signals are private and – conditional on the state – independent across players. After observing his signal, a player can access one of  $K \geq 2$  “rooms” and send a message  $m_i \in \mathcal{M}$ . The message is received by all players in the same room. Afterwards each player takes an action  $a_i$ .

The payoff of player  $i$  is  $u(a, b_i, \theta) = -(a_i - b_i - \theta)^2 - \alpha \sum_{j \neq i} (a_j - b_i - \theta)^2$  where  $a$  denotes the vector of actions of all players and  $b_i \in \mathcal{B}$  is a commonly known “bias” of player  $i$ . That is, player  $i$  would like that all players choose the action  $b_i + \theta$ . The parameter  $\alpha$  measures the relative weight players assign to other players’ behavior. Players are assumed to maximize expected utility.

The solution concept used is perfect Bayesian Nash equilibrium.

For simplicity, let  $\Theta = \{\theta^h, \theta^l\}$  and  $\Sigma = \{\sigma^l, \sigma^h\}$  and the signal structure is such that  $\text{prob}(\sigma^j|\theta^j) = p > 1/2$ . We let the message space be binary as well:  $\mathcal{M} = \{h, l\}$ . Furthermore, we let  $\mathcal{B} = \{0, b\}$  and assume that there is at least one player with each of the two biases.

**Action choice** Denote the belief of player  $i$  that the state of the world is high by  $\mu_i$  (after observing his signal and listening to all the messages in his room). The expected utility of player  $i$  can then be written as

$$\begin{aligned} U(a, \mu_i) &= -a_i^2 - \mathbb{E}[(b_i + \theta)^2] + 2a_i(b_i + \mathbb{E}[\theta]) - \alpha \sum_{j \neq i} \mathbb{E}[(a_j - b_i - \theta)^2] \quad (6) \\ &= -a_i^2 - \mu_i(b_i + \theta^h)^2 - (1 - \mu_i)(b_i + \theta^l)^2 + 2a_i(b_i + \mu_i\theta^h + (1 - \mu_i)\theta^l) \\ &\quad - \alpha \sum_{j \neq i} [\mu_i(a_j - b_i - \theta^h)^2 + (1 - \mu_i)(a_j - b_i - \theta^l)^2] \end{aligned}$$

The optimal action choice of player  $i$  is then

$$a_i^* = b_i + \mathbb{E}[\theta] = b_i + \theta^l + \mu_i(\theta^h - \theta^l). \quad (7)$$

**Cheap talk** The cheap talk game can – as usual – have several equilibria. There is always a babbling equilibrium where the message is independent of the observed signal and therefore nothing about the state of the world is learned, e.g.  $m_i(\sigma_i) = \sigma^h$  for all  $\sigma_i \in \Sigma$  and  $\mu_i = p$  ( $\mu_i = 1 - p$ ) if  $\sigma_i = \sigma^h$  ( $\sigma_i = \sigma^l$ ). We will focus on most informative equilibria, that is equilibria where  $m_i(\sigma_i) = \sigma_i$  with as high probability as possible.

Truthful communication is an equilibrium for a given room if all players in this room have the same  $b_i$ . To see this, suppose player  $i$  could maximize his expected utility (1) not only over  $a_i$  but also over the  $a_j$  of all the players in his room. Clearly, he would choose the same action for everyone namely  $b_i + \theta^l + \mu_i(\theta^h - \theta^l)$ . Deviating from the truthful strategy is not profitable because by adhering to truthfulness player  $i$  ensures that all other players in the room choose precisely the action he would have chosen for them (while deviating changes the other players' beliefs and therefore their optimal action). Note that this argument depends on all players having the same bias and truthful communication is normally not an equilibrium if players in a given room have different biases. We state this result for future reference in the following lemma.

**Lemma 8.** *If all players in a given room have the same bias, truthful communication in this room is the most informative equilibrium of the cheap talk game (taking room choice as given).*

We will now analyze the cheap talk problem in rooms in which players with both types of biases are present. In particular, we will be interested in the case of strong differences in opinion, i.e. the case where  $b$  is sufficiently large.

**Lemma 9.** *Let  $n_0 \geq 1$  players with bias  $b_i = 0$  and  $n_b \geq 1$  players with bias  $b_i = b$  be in a room. There exists a  $\bar{b}$  such that for  $b \geq \bar{b}$  babbling is the only equilibrium of the cheap talk game.*

**Proof of lemma 9:** Suppose that there is a non-babbling equilibrium, i.e. an equilibrium where belief  $\mu_j$  depends on the messages of players  $i \neq j$ . Let  $i$  be a player affecting  $j$ 's belief. Without loss of generality, say  $\mu_j$  is lower if  $i$  sends the message  $l$  and higher if  $i$  sends the message  $h$ . By Bayesian updating and independence of the signals,  $\mu_k$  will then be lower when  $i$  sends message  $l$  than when he sends message  $h$  for all  $k \neq i$ . (Moreover two players that observe the same signal themselves and are in the same room will have the same belief because of Bayesian updating and independence of signals.) Hence it is without loss of generality to assume that  $b_i \neq b_j$ . For concreteness, let  $b_i = 0$  and  $b_j = b$  (the proof for the opposite case is analogous).

Now suppose  $i$  observes signal  $\sigma^h$ . We will show that it is optimal for  $i$  to send message  $l$  if  $b$  is sufficiently high. To see this, denote the change in  $i$ 's expected utility (6) when sending message  $l$  instead of message  $h$  as  $\Delta U_i$ <sup>4</sup>

$$\begin{aligned}
\Delta U_i &= -\alpha \sum_{j \neq i} \mathbb{E} [a_j(l)^2 - a_j(h)^2 - 2\theta(a_j(l) - a_j(h))] \\
&= -\alpha \sum_{j \neq i} \mathbb{E} \left[ (\mu_j(l)^2 - \mu_j(h)^2) (\theta^h - \theta^l)^2 + 2(\mu_j(l) - \mu_j(h))(\theta^h - \theta^l)(b_j + \theta^l - \theta) \right] \\
&= \alpha \sum_{j \neq i} \mathbb{E} [(\mu_j(h) - \mu_j(l)) (\theta^h - \theta^l) * ((\mu_j(h) + \mu_j(l))(\theta^h - \theta^l) + 2(b_j + \theta^l - \theta))] \\
&= \alpha (\theta^h - \theta^l) n_b \mathbb{E} [(\mu_j(h) - \mu_j(l)) * ((\mu_j(h) + \mu_j(l))(\theta^h - \theta^l) + 2(b + \theta^l - \theta))] \\
&\quad + \alpha (\theta^h - \theta^l) (n_0 - 1) \mathbb{E} [(\mu_j(h) - \mu_j(l)) * ((\mu_j(h) + \mu_j(l))(\theta^h - \theta^l) + 2(\theta^l - \theta))] \\
&= \alpha (\theta^h - \theta^l) (n_b + n_0 - 1) \\
&\quad \mathbb{E} \left[ (\mu_j(h) - \mu_j(l)) * \left( (\mu_j(h) + \mu_j(l))(\theta^h - \theta^l) + 2 \left( \frac{n_b b}{n_b + n_0 - 1} \theta^l - \theta \right) \right) \right].
\end{aligned}$$

If  $b \geq \theta^h(n_b + n_0 - 1)/(\theta^l n_b)$ , the term inside the expectation is positive for any  $\theta$  and therefore  $\Delta U_i$  is definitely strictly positive. Hence,  $i$  strictly prefers sending message  $l$  to message  $h$  and  $i$  receives signal  $\sigma^h$ . This would imply that  $i$  sends message  $l$  with probability 1 if the signal is  $\sigma^h$  in this equilibrium. But this contradicts that  $\mu_j(l) > \mu_j(h)$ . Hence, choosing  $\bar{b} = \theta^h N / \theta^l$  where  $N$  is the total number of agents implies  $\bar{b} \geq \theta^h(n_b + n_0 - 1)/(\theta^l n_b)$  and gives the result.  $\square$

<sup>4</sup>For a more general proof, one could already go from the first line to  $\alpha \sum_{j \neq i} \mathbb{E} [(a_j(h) - a_j(l)) * (a_j(h) + a_j(l) - 2\theta)]$  and then note that for  $b$  high enough even  $a_j(l) > \theta^h$ .



Lemma 9 implies that – given a finite number of players – the only way allowing meaningful communication if differences in opinion is high is to have only players with the same bias in a room.

If the differences in opinion are minimal, i.e.  $b$  is very low, truthful communication is an equilibrium for any room composition. The reason is the coarseness of the signal structure: Lying in the message game leads – in a truthful equilibrium – to a discrete reaction of all other players in the room. If the difference in bias is very small, this discrete reaction is “too high”, i.e. even those players with a (slightly) different bias react more than the deviating player would wish for. The following lemma formalizes this generalization of lemma 8.

**Lemma 10.** *Let there be  $n_0 \leq n$  players with  $b_i = 0$  and  $n_b \leq n - n_0$  players with  $b_i = b$  in a room. There exists a  $\underline{b} > 0$  such that for  $b \leq \underline{b}$  truthful communication is an equilibrium.*

**Proof of lemma 10:** For  $b = 0$ , truth-telling is strictly better than lying (given that all other players tell the truth). Note that  $i$ 's expected utility is continuous in  $a_j$  and  $a_j^*$  is continuous in  $b_j$ , see (7). Hence,  $U_i$  is continuous in  $b_j$ . However,  $\mu_j$  and therefore  $a_j^*$  reacts discretely to lying. Consequently, truth-telling is still a best response to truth-telling for  $b_j > 0$  sufficiently small.  $\square$

From lemma 9 and lemma 10 we know that for  $b$  low the most informative equilibrium in a room with a given configuration is truth-telling and for  $b$  sufficiently high the “most informative” equilibrium is babbling if players with different biases are present. It seems most likely that  $\bar{b} > \underline{b}$ . In this case, there are mixed strategy equilibria for  $b \in (\underline{b}, \bar{b})$ .

**Room choice equilibria** We claim that separation is an equilibrium if differences in opinion, i.e. the parameter  $b$ , are sufficiently high.

**Proposition 20.** *If  $b \geq \bar{b}$ , the following strategies constitute an equilibrium:*

1. *Players with bias 0 ( $b$ ) go to room 0 (1).*
2. *A player sends truthful messages if only players of the same type are in his room and babbles otherwise.*
3. *Actions are taken according to (7) and beliefs  $\mu_i$  are formed using Bayes' rule (given the equilibrium strategies in 1 and 2).*

*This equilibrium is the most informative equilibrium in the sense that no player has more precise information about the state  $\theta$  in any other equilibrium.*

**Proof of proposition 20:** Given lemma 9, unilateral deviations to other rooms are not profitable: Any such deviation would either lead to being alone in a room or babbling. In either case, the deviating player does not have any information beyond his own signal

about the state of the world. This reduces his expected utility directly. Furthermore, deviations lead to less information for other players which again lowers the deviating player's payoff: Less information for players with the same bias as player  $i$  implies that their actions are further away from  $b_i + \theta$  in expectation. Furthermore, the players with  $b_j \neq b_i$  choose actions further away to  $b_j + \theta$  if they have less information, i.e. variance of their choice is increased while the expected value stays the same. Given the strictly concave loss function, player  $i$  loses from this as well.

Lemmas 8 and 9 imply that no profitable deviation in the cheap talk stage exists. As (7) gives the optimal action (given one's beliefs), no deviation in choosing one's action is profitable either.

By lemma 9, a given player  $i$  cannot observe more "non-babbling" messages than in the suggested equilibrium in any other equilibrium. Given that communication is truthful in the suggested equilibrium, player  $i$  can therefore not have more precise information about  $\theta$  in any other equilibrium.  $\square$

For  $b \leq \underline{b}$ , the most informative equilibrium is clearly that every player goes to the same room and truthfully reports his signal.

**Welfare optimal room allocation** Suppose a social planner could allocate players to rooms. After being assigned a room, players play the same game as above; that is, the planner has no influence on messages or actions. We claim that for  $b \geq \bar{b}$  the welfare optimal allocation is to assign everyone with bias 0 in one room and everyone with bias  $b$  in another room, i.e. the equilibrium described in proposition 20 is welfare optimal. The idea is the following: For  $b \geq \bar{b}$ , the cheap talk game in a room where players with both bias types are present will only have a babbling equilibrium by lemma 9. Consequently, any room allocation that assigns players with different biases to the same room will lead to completely uninformative messages and is therefore equivalent to putting every player to a separate room. By assigning players with the same bias to the same room, the planner achieves the most informative equilibrium. That is, truthful communication is possible in each room. The additional information ensures that player with the same bias as player  $i$  choose actions closer to  $b_i + \theta$ . Furthermore, the players with  $b_j \neq b_i$  choose actions closer to  $b_j + \theta$ , i.e. the variance is reduced while the expected value stays the same. Given the strictly concave loss function, player  $i$  gains from this as well. Note that the welfare notion can be chosen quite strict in the sense that the described allocation maximizes the welfare of every agent. That is, if agent  $i$  could dictatorially decide the room allocation (without having any influence on the messages or actions taken by other players), the same allocation would result.

Similarly, the most informative equilibrium is welfare optimal in the strong sense established above if  $b \leq \underline{b}$ .