Strategic Polarization in Group Interactions *

Ganesh Iyer
University of California, Berkeley

Hema Yoganarasimhan
University of Washington

September 28, 2020

*We thank Przemek Jeziorski, Yuichiro Kamada, Sridhar Moorthy, Jiwoong Shin, Jidong Zhou and seminar participants at Carnegie-Mellon University, Columbia University, Harvard University, MIT, University of Toronto and Yale University. Authors contributed equally and are listed alphabetically. Email: giyer@haas.berkeley.edu, hemay@uw.edu.

STRATEGIC POLARIZATION IN GROUP INTERACTIONS

ABSTRACT

We study the phenomenon of strategic polarization of actions in groups. Agents with private preferences choose a public action (voice opinions), and the mean of their actions represents the group's realized outcome. They face a trade-off between influencing the group outcome and truth-telling. In a simultaneous move game, agents strategically shade their actions towards the extreme leading to polarization in equilibrium. Agents with more extreme preferences shade even more. The group outcome is also more extreme than the mean preference. The strategic group influence motive can create substantial polarization in actions even when the preferences are relatively moderate. Compared to a simultaneous actions game, randomized (or exogenous) sequential actions lowers polarization when agents' preferences are relatively similar. Sequential actions can even lead to moderation if later agents have preferences which are close to the center. Endogenizing the order of moves (through a first-price sealed bid auction) always increases polarization, but it is also welfare enhancing.

Introduction

Many business and organizational settings involve group decisions. Group interactions and decisions are observed in both formal and informal groups ranging from social media groups, sales teams, corporate personnel committees, academic committees, community groups and political action committees. While a natural premise of group interactions, going back to Rawls (1971), is that they may promote greater alignment of opinions and actions, it is also common for individuals to become more divergent in their opinions and actions, especially when their actions have consequences for the group's outcomes. Indeed, a body of experimental evidence shows that instead of enabling greater alignment, group deliberations often lead to more polarized behavior (Isenberg, 1986).

Polarization can lead to outcomes in firms and in social settings which do not represent the preferences of the agents. This paper examines the phenomenon of polarization of the observed actions of agents who interact on business, organizational, or even socio-political issues: i.e., when members of a group take actions (voice opinions) that are more extreme in the direction of their true preferences. This characterization of polarization of agents' observed actions is consistent with the idea of polarization of group decisions described in the existing literature (Sunstein, 2002). Consider the following examples of group interactions that illustrate relevant aspects of the phenomenon studied in this paper:

- **Brand Perceptions:** After the 2016 elections discussions on social media forums and Twitter groups about the advertising campaigns of brands Nike, Starbucks, and Chobani which have taken positions on socio-political issues have become more polarized. In these online interactions consumers with liberal and conservative political preferences compete to make their views about the brands heard (see Kate Taylor 11/22/2017, *Business Insider*). How do these interactions shape overall brand perceptions?
- **Online Reviews:** Research shows that reviewers on Yelp place a significant weight on the past reviews of others. Furthermore, some reviewers on Yelp seem to have a greater motivation to deviate from the past ratings posted by others and to be more extreme

(see Dai et. al 2018). Reviewers also tend to exaggerate more in the direction of their preference when they see a current review rating which is significantly higher or lower than their preference.

- **Corporate Lobbying and Climate Change:** During the heated climate change debates of 2008 in Congress, the most intensive lobbying efforts in the electric utilities industry were by two companies at the opposite and extreme ends of the environmental performance spectrum. Southern Company one of the highest polluting utilities in the nation spent $14 million in 2008 on climate change lobbying, while at the same time PG&E one of the greenest utilities in terms of carbon emissions had an even more intense campaign and spent an estimated $27 million. Recent research (see Delmas et. al 2016) provides empirical evidence for a systematic U-shaped relationship between greenhouse emissions and lobbying intensity: the companies with the least and most carbon emissions are the most aggressive and extreme voices in the climate debate. What might account for this pattern?
- **Diversity and Inclusion:** Recently, HR groups within firms have struggled with how to balance merit vs. diversity in recruiting decisions. Often, the final decisions look more extreme than the individual preferences of group members. For example, Google's senior executives met and decided that certain hires from the third quarter of 2017 onwards must be "all diverse". That meant all the hires had to be Black, Hispanic, or female. None could be a white or Asian-American man (Eastland, 2018).
- **Marketing Faculty Hiring:** Faculty hiring decisions in marketing departments often bring into play the motivations to influence the group. Consider the potential debate within a department regarding hiring in new emerging research areas. Many marketing departments may consider whether to allocate scarce faculty slots towards expertise in machine learning or field experimentation. Similarly there may be arguments about hiring in sub-areas areas, such as, behavioral versus quantitative or theory versus empirical. Group members may have different private preferences across the research areas to which scarce faculty slots should be allocated to. What mechanisms might be

available to a department chair seeking to reduce polarization and division in the hiring deliberations?

- **Gun Control:** In 2015, the state of Texas passed the *campus carry law*, formally known as Senate Bill 11 (SB 11) (Aguilar, 2016). Following this, the University of Texas assembled a nineteen member working group to discuss how to implement this law into practice. The working group members with disparate views deliberated on implementation dimensions such as where to allow guns on campus, age limits, and the manner in which guns may be carried on campus. The deliberations led to some extreme outcomes such as the conclusion that the committee would not recommend a ban on guns in the classroom (Campus Carry Policy, 2015).[1]

Common Themes

Some common themes emerge from these examples and are the focus of our analysis. First, they represent contexts in which agents' have strongly held preferences. For example, in hiring decisions, agents often have strong preferences on the role of diversity or on the importance of research areas. Similarly, consumers may have strong political preferences that shape their reactions to advertising campaigns take positions on socio-political issues. Many other issues that dominate our business and public policy discussions also fall under this category –the abortion debate, how much immigration to allow, should there be separation of church, and state and the size of government. In such cases, when participating in a group decision-making process, an agent's motivation is to achieve an outcome that closely conforms to her preferences. This is in contrast to and distinct from settings where agents care about some underlying true state of the world and have uncertainty about this state. In that case the agents try to aggregate their information as a group to uncover the uncertainty and condition their decision on it. Thus, the actions/voiced opinions in our paper represent agents' stated preferences rather than beliefs/information about a true state of the world.

[1] See Huitlin (2015) and Armed Campuses (2016) for comprehensive discussions of campus carry laws across different states and the recent developments in this area.

Second, both individual opinions and the eventual outcomes may often end up looking more extreme than one would expect *a priori* based on the initial distribution of the group members' preferences. For example, in the Google case, the decision to make "all diverse" hires in 2017 was more extreme than expected. This pattern of polarized opinions and group outcomes is common in other socio-political spheres and even a cursory reading of current news would suggest that many issues show evidence of polarization despite the presence of moderating influences (Cohn, 2014).

Third, in these examples individuals' actions or voiced opinions are not necessarily a binary/discrete choice, as is common in voting models; rather it is typically a choice of the extent or the magnitude of an action (i.e., continuous choice). In the faculty hiring example, it could be the strength and the number of arguments presented by group members to make a behavioral versus quantitative faculty hire. In the campus carry example, the choice for the members of the group was not a simple "should guns be allowed on campus or not?." Rather it was a nuanced decision on where it would be allowed (classrooms), not allowed (child-care units), and allowed with discretion (single-user offices), where it can be stored (in-person or locked vehicle), who is it allowed for (over 21 years, with license).

Past research in economics mainly attributes polarization to imperfect information aggregation and polarization of agents' beliefs in group settings; see the Related Literature section for details. A parallel literature in psychology attributes polarization of group decisions to behavioral biases stemming from social comparison and persuasive argumentation (Zuber et al., 1992; Baron, 2005). In this paper we present a general rationale for the polarization of actions in groups as distinct from the existing accounts based on either imperfect information aggregation or behavioral biases. Our analysis links the preferences of agents in the group to the resulting polarization in their observed actions. We ask – Can polarization of group decisions stem from strategic interactions between agents, even if the agents are rational and there is no imperfect information on some true state of the world?

Research Agenda and Approach

We propose a theory of group polarization with two related objectives. First, the theory connects the emergence of polarized group outcomes to the strategic motives of individuals for group influence. We develop a model of group decision-making, where agents have heterogeneous preferences over an issue. The basic analysis starts with a group of two agents. Each agent's utility function has two components: First, an agent incurs dis-utility if she chooses an action (or voices an opinion) that is different from her true preference. This is represented as a convex cost, which is increasing in the extent of the misalignment between her action and her true preference. This can be interpreted as a reputational (or even a psychological) cost of misreporting her true preference. Second, an agent cares about the distance of the group's eventual outcome from her true preference. This represents the group influence motive – individuals would like to move the group's eventual outcome towards their true preference. The game consists of each agent privately observing her true preference and choosing a publicly observable action (opinion). The mean of their public actions represents the group's decision or outcome and thus this outcome is influenced by the actions of all the agents. Our model connects the trade-off between the desire for group influence and truth-telling to the polarization of actions at the individual and group level, i.e., where individuals take actions that are more extreme than their true preferences and the aggregate group outcome is more extreme than the mean of the group's true preferences.

Next, we examine some important variables that can influence the existence and the extent of polarization – the size of the group, sub-group interactions, partial knowledge of the other agents' types, and the game structure. With respect to the last variable, we investigate the timing of actions or the order in which agents voice opinions. Each agent may simultaneously chooses an action without observing that of the other. Alternatively, agents may voice their opinions sequentially, in which case those who speak/act later will be able to observe the opinions of those who spoke before. Thus the main difference between these two timings is the "observability" of others' actions. We examine whether group decisions are more polarized if agents speak simultaneously or sequentially. If individuals were to speak sequentially, who has a greater incentive to speak first – those with more extreme preferences or the moderates?

The questions pertaining to timing and observability of actions are important because we see examples of both models in practice. The standard secret ballot models, where each agent submits her/his opinion without observing the actions of others can be interpreted as simultaneous actions model. For example, a department chair can survey everyone's opinion simultaneously (e.g., through online survey tools) and aggregate their opinions to make a decision. On the other hand, on a social media site, the opinions posted by previous members will be observed and can affect the actions of later members.[2] Indeed, with the advent of online ballots and opinion-sharing forums, both models are equally easy to design and implement. However, we do not have good answers to which of these models lead to more polarized decisions.

Further, we ask whether endogenizing the timing of actions by allowing agents to influence the speaking order affects polarization, and if so how? Specifically, we consider a game where agents can participate in a first-stage auction to bid on the right to decide when they speak. The first-stage auction may be interpreted as agents lobbying with a principal (eg., a department chair or a policy maker) to influence the speaking order. There is a long history in economics of modeling lobbying activities as auctions (Che and Gale, 1998), and our endogenous choice model reflects agents' costs to influence the rules of the game through lobbying (Potters and Van Winden, 1992; Harstad et al., 2006). Within this context, we examine which agents will have higher incentives to bid for the right to mandate the speaking order and when will they prefer to speak. Finally, we seek to compare the

[2] A recent study on Twitter finds that when individuals express opinions in sequence, exposure to opposing views leads users to become more entrenched (extreme) in their views compared to their original positions (Bail et al., 2018). On Twitter and other popular online opinion-sharing platforms, individuals with more extreme views often contribute more than those with more moderate views. This may lead to a distribution of opinions that is polarized, with many extreme opinions, and few moderate opinions.

social welfare under different game forms, and derive the relationship between the extent of polarization in the system and the overall welfare.

Results and Contribution

First, we show that in a simultaneous game, agents indulge in *strategic shading* towards the extremes, i.e., take actions (voice opinions) that are more extreme than their true preferences in order to pull the group outcome closer to their preferences. Shading towards the extreme occurs at both individual and group levels. Further, and consistent with the pattern of lobbying in the electrical utility industry example, agents with extreme preferences shade more than moderates, as they anticipate that the equilibrium outcome is likely to be farther away from their preference. Importantly, an agent's incentive to shade and the extent of polarization in the group outcome are independent of the preference distribution. In other words, our model shows that the polarization of observed actions does not necessarily stem from polarization of preferences, but rather from the strategic motivations of the agents.

Extending the analysis to many players we show that the extent of shading goes down as groups get larger, which suggests that group size can be used as a mechanism to control polarization. We also examine the interaction between subgroups in which agents within a subgroup have homogeneous preferences while there is heterogeneity across subgroups to show that smaller sub-groups have the incentive to become even more extreme.

The analysis of the timing actions and the comparison of simultaneous and sequential choice games establishes some of the important results of the paper. With sequential actions, polarization occurs whenever the agent who moves later is relatively more extreme compared to the first agent, whereas moderation occurs if the agent who moves later has less extreme preferences. The second agent's motivation looms larger on the joint outcome because she can condition her action on the observed action of the first agent and pull the outcome closer to her preference. For example, this pattern is often visible in online forums, where agents who come later tend to express progressively more extreme opinions (Bail et al., 2018). Thus the analysis shows that given the group influence motive, the informational benefit of waiting to react to the actions of others is more attractive than moving first and setting the agenda. We then compare the two timing games and show that if the preferences of the agents are relatively similar then the group decision is more moderate in the sequential actions game, whereas with dissimilar preferences it is the simultaneous actions game that results in more moderate group outcomes.

Next, we discuss the findings from the endogenous choice game, where agents can participate in a first-price sealed bid auction for the right to determine the speaking order. We find that agents with more extreme preferences bid more for the right to determine the speaking order, and upon winning all agents regardless of their preferences prefer to wait. More importantly, because the more extreme agents bid more, the group outcome in the endogenous sequential game is always polarized. But the extent of polarization can be higher or lower compared to the simultaneous game: when the players preferences are relatively similar, endogenizing the speaking order leads to relatively less polarization compared to the case of simultaneous actions. The practical implication is that in groups where players are similarly inclined, allowing for endogenous sequential timing helps to mitigate group polarization. Interestingly, we find that the endogeneous sequential choice game has highest total welfare despite the fact that it may lead to higher polarization, because it more efficiently allocates the right to decide the speaking order.

To summarize, our paper contributes a general theory of polarization of group outcomes which is based on the strategic motivation of agents for group influence. Distinct from the existing research which has focused on the polarization of beliefs arising either from imperfect updating or from behavioral biases, we analyze the polarization of observed actions in groups. The presence of the group influence motive can result in substantial polarization of actions even when preferences are relatively moderate. We also contribute by identifying several important mechanisms that govern the extent of polarization in actions. First, we show how the timing or the observability of actions mitigates polarization. The observability of the actions of prior agents mitigates polarization when agents' preferences are relatively similar, but allowing agents to influence the speaking order exacerbates polarization. Second, we identify mechanisms which can be used in practice to moderate the extent of polarization in group decisions. These include the role of group size, the speaking order, and the amount of knowledge agents' have about others, all of which can be controlled by a group coordinator. Finally, from a overall welfare perspective, we show that game formats that lead to more polarized outcomes do not necessarily lead to lower welfare.

Related Research

Research in psychology starting with Stoner (1961) shows evidence that group deliberation can make individuals and the overall group decision to be extreme in the direction of their original proclivities. Evidence and the psychological basis for polarization has been demonstrated in different contexts including jury decisions (Main and Walker, 1973), faculty evaluations and pay, attitudes towards women (Myers, 1975), and judgements of attractiveness (Myers, 1982) to name a few. We identify a strategic rationale for group polarization which can accommodate and explain the different studies in this literature. In other words, the strategic trade-off between the incentive to influence the group outcome and truth-telling provides a general rationalization of polarization which is independent of the contexts that motivate the different studies in psychology.

A more recent stream of literature focuses on providing different economic explanations for polarization of "beliefs" in group interactions. Dixit and Weibull (2007) analyze a model of Bayesian updating by agents with heterogeneous normally distributed priors about a true (policy) state and a common noise. In this set-up while the mean belief of the group may diverge under Bayesian updating after observing the common signal, individual-level polarization does not occur. Nevertheless, Baliga et al. (2013) show that polarization of individual beliefs can occur if there is ambiguity aversion of individuals who observe a common signal. Acemoglu et al. (2009) considers a Bayesian learning problem for agents with different priors about the distribution of signals and show that even a tiny amount of signal uncertainty leads to significant disagreement in asymptotic beliefs.[3] In contrast to this literature, our analysis is about the polarization of observed actions resulting from the strategic incentives of agents. This allows us to establish a rationale for the polarization of group actions even in circumstances where preferences and beliefs of the agents are relatively moderate.

A parallel stream of research examines the role of behavioral biases or non-bayesian updating on polarization. An early paper by Rabin and Schrag (1999) formalizes a model of confirmatory bias where agents ignore signals which do not confirm with their initial impression, and update in the direction of their current beliefs generating polarization.[4] Bénabou (2012) investigates the emergence of collective denial in groups as agents form overoptimistic beliefs by ignoring negative signals. Glaeser and Sunstein (2009) analyze non-Bayesian behavior in which agents fail to account for the common sources of information of others' opinions. We complement this literature by identifying the role of a general group influence motive and how it interacts with the timing and the observability of actions in determining the extent of polarization. Whether the outcome of the group is more or less polarized depends upon strength of the group influence motive as well as whether agents who move later are relatively more or less extreme.

There is also a related literature on strategic communication and cheap talk in persuasion games with multiple senders (experts) who try to influence a decision maker. Early papers by Gilligan and Krehbiel (1989) and Austen-Smith (1990) model debates as cheap talk messages from multiple senders with different interests to show that such debates will only affect the outcome if the agents' preferences are not too dissimilar. Within this stream,

[3] Other papers in this area include Kondor (2012), who shows that belief polarization can be generated when agents see different private signals that are correlated with a common public signal. A similar idea is present in Andreoni and Mylovanov (2012).

[4] This in turn is related to the research stream in social psychology starting from Lord et al. (1979) which provides experimental evidence that groups of individuals who hold differing opinions about socio-political issues use information in a biased manner, by incorporating confirming evidence more readily than disconfirming evidence.

Krishna and Morgan (2001) show that consulting two perfectly informed experts rather than one is beneficial when the experts are biased in opposite directions. A group of extremists do not have informational value in this framework. However, Bhattacharya and Mukherjee (2013) show that if the experts are uncertain about their information then a decision maker may indeed prefer to hear from more extreme experts. Our paper does not deal with strategic information transmission, but rather with the strategic effect of the group influence motive in creating an incentive for polarized actions.

Our paper is also related to the literature in marketing research on group decision making that focuses on linking group behavior to that of individual members. Rao and Steckel (1991) develop an empirical model where group preferences is a weighted linear model of individual preferences. Their model attempts to account for observed group polarization in the data. Eliashberg and Winkler (1981) study group decision making and examine how uncertain group payoffs should be divided among the members in a Pareto optimal manner, given their risk attitudes and preferences. More broadly, our paper also adds to the literature on social effects in marketing. A stream of empirical papers document the existence of social effects; see Hartmann et al. (2008) for an overview and Sun et al. (2019) for a recent documentation of social effects using data from field experiments. A related stream considers the impact of these social effects on firms' strategies: For example, Amaldoss and Jain (2005) analyze the competitive pricing strategies of conspicuous goods when consumers have preferences for uniqueness and conformity and Yoganarasimhan (2012) analyzes a monopolist's advertising decisions in a market where consumers engage in social signaling.

Model

We first present the basic model of group interactions, where the mean of actions of the agents is seen as the group outcome. Consider a group of two agents i and j, where each agent's preference (denoted as x_i and x_j) is independently drawn from a distribution $g(x)$, which is symmetric around zero and with support over the real line $\mathbb{R}$. The cumulative density of the distribution is given by $G(x) = \int_{-\infty}^{x} g(t)dt$ and $G(\infty) = 1$.[5]

Agent i's true preference or type x_i is her private information. Both agents simultaneously choose a publicly observable action, $\{a_i, a_j\} \in \mathbb{R}$. In a group interaction, an agent's action can be interpreted as her voiced opinion and the action is continuous. After both agents have spoken, assume that a neutral third-party or principal implements the mean of their voiced preferences or actions as the group outcome or decision.

The utility of agent i is given by the following convex loss function:

$$u(x_i, a_i, a_j) = -r(x_i - a_i)^2 - (1 - r)(x_i - \bar{a})^2 \quad (1)$$

where $\bar{a} = \frac{a_i + a_j}{2}$, and $r \in (0, 1)$ represents the relative weights that the agents places on the different components of their utility. Agents obtain dis-utility from two sources. First, their utility is decreasing in the distance between their action and their preference, $i.e.$, they prefer to voice opinions close to their true preference. This could stem from a disinclination to misreport their preferences (cost of lying) or a reduced form representation of credibility of actions arising from potential reputational concerns.[6] Second, their utility is decreasing in the distance between the group outcome ($\bar{a}$) and their true preference. The main results of the paper are not dependent to this assumption of the mean as the group outcome. They hold qualitatively for any decision rule that uses a linear combination of agents' preferences and puts non-zero weights on the actions of all players; see Web Appendix §A.1 for details.[7] What is necessary for the results of the paper to qualitatively hold is that the outcome measure is a function of the actions of all the agents in the group. In other words, agents'

[5]In the analysis, to illustrate some of the results we use as an example preferences that are independently drawn from $U[-1, 1]$ (and actions $\{a_i, a_j\} \in \mathbb{R}$).

[6]One could consider a model of reputation in which any misreporting today has consequences for the future. The first term in the utility function can be seen as the reduced form equivalent of this model. This reduced form representation captures the credibility of actions; as we can see, when r approaches 1 the players truly report their preferences.

[7]The idea of using a linear combination of individual preferences to represent the group preferences/outcomes has a long history in consumer research and marketing, and has been shown to have significant empirical validity (Rao and Steckel, 1991; Corfman and Lehmann, 1987; Elisahberg et al., 1986).

have a taste for influencing the group's outcome. A greater value of r represents issues for which agents have stronger relative preference for voicing opinions that are consistent with their true preferences. Overall, the agent's utility function displays the single-peakedness property where each agent has an ideal point (in this case x_i). Actions as well as outcomes away from this point are less than ideal, and strictly monotonically decreasing in both directions.

We consider a game in which nature first draws the preferences x_i and x_j for the agents based on which they choose their publicly observable actions. The actions a_i and a_j may be chosen simultaneously in which case each agent's choice is contingent only on the private information about her preference. Alternatively, the agents may move sequentially in which case the agent who moves second will be able to choose her actions contingent upon her private information as well as the observed actions of the first mover.

Benchmark Cases

Before we analyze of the private information game, it is useful to derive two benchmark cases – i) the first-best socially optimal solution, and ii) the perfect information case. In the first case, a social planner chooses actions to maximize the joint surplus of the two agents:

$$W(x_i, x_j, a_i, a_j) = \sum_{k=i,j} -r(x_k - a_k)^2 - (1-r)(x_k - \bar{a})^2 \tag{2}$$

The welfare maximizing choices are $a_i^* = x_i$ and $a_j^* = x_j$. The socially optimal action for both agents is truth-telling and the joint decision shows no distortion from the preferences. Suppose now that the agents have perfect information on each other's types and move simultaneously. Denoting the agents' equilibrium actions as $\{a_i^p, a_j^p\}$, we can derive: $a_i^p = \frac{3r+1}{4r}x_i - \frac{1-r}{4r}x_j$ and $a_j^p = \frac{3r+1}{4r}x_j - \frac{1-r}{4r}x_i$. While both agents deviate from truth-telling by reporting a weighting of their own preference and the other agent's preference, the mean action, $\bar{a}^p = \frac{x_i+x_j}{2}$, perfectly reflects the mean preferences of the group. Thus, with perfect information too the group's joint decision is not distorted.

Simultaneous Actions

Equilibrium

Consider the game in which agents choose their actions without observing the other agent's type and actions. We proceed to derive the Bayesian Nash equilibrium of this game and focus without loss of generality on agent i. Let $\hat{a}_j$ denote the equilibrium action of j. Because j's preference (x_j) is her private information at the time of choosing the action, i's expected utility from choosing a_i as $EU(x_i, a_i) = \frac{\int_{\mathbb{R}} u(x_i, a_i, \hat{a}_j) g(x_j) \mathrm{d}x_j}{\int_{\mathbb{R}} g(x_j) \mathrm{d}x_j}$. By differentiating $EU(x_i, a_i)$ and setting it equal to zero at i's equilibrium action $a_i = \hat{a}_i$ gives us:

$$\left.\frac{\partial EU(x_i, a_i)}{\partial a_i}\right|_{a_i = \hat{a}_i} = 2r(x_i - \hat{a}_i) - \frac{(1-r)}{2}\left[-(2x_i - \hat{a}_i) + \int_{\mathbb{R}} \hat{a}_j g(x_j) \mathrm{d}x_j\right] = 0 \quad (3)$$

In obtaining the above first order condition, we can set $\frac{\mathrm{d}\hat{a}_j}{\mathrm{d}a_i} = 0$ because in a simultaneous equilibrium, any change in the action of agent i has no impact on the equilibrium action of agent j. Simplifying equation (3) gives us $\hat{a}_i$ as:

$$\hat{a}_i = \frac{2(1+r)}{1+3r} x_i - \frac{(1-r)}{(1+3r)} \int_{\mathbb{R}} \hat{a}_j g(x_j) \mathrm{d}x_j \quad (4)$$

Integrating i's equilibrium action $\hat{a}_i$ over the entire range of x_i gives us:

$$\int_{\mathbb{R}} \hat{a}_i g(x_i) \mathrm{d}x_i = \frac{2(1+r)}{1+3r} \int_{\mathbb{R}} x_i g(x_i) \mathrm{d}x_i - \frac{(1-r)}{(1+3r)} \int_{\mathbb{R}} \hat{a}_j g(x_j) \mathrm{d}x_j \int_{\mathbb{R}} g(x_i) \mathrm{d}x_i \quad (5)$$

Because $\int_{\mathbb{R}} \hat{a}_i g(x_i) \mathrm{d}x_i = \int_{\mathbb{R}} \hat{a}_j g(x_j) \mathrm{d}x_j$, and because $E(x) = 0$ for a symmetric distribution, we can uniquely identify $\int_{\mathbb{R}} \hat{a}_j g(x_j)_j \mathrm{d}x_j = 0$. We thus have $\hat{a}_i = \frac{2(1+r)}{1+3r} x_i$. These results are summarized below in Proposition 1 below:

Proposition 1 *In the simultaneous actions game, there exists a unique Bayesian Nash equilibrium, where an agent i with preference x_i chooses action $\hat{a}_i = \frac{2(1+r)}{1+3r} x_i$.*

An implication of the Proposition is that agents' actions are more extreme than their true preferences (the multiplier $\mu(r) = \frac{2(1+r)}{1+3r} > 1$, for all $r < 1$). Moreover, this shift to extremity is in the direction of their original preference: *i.e.*, those with positive x_i always move right, while those with negative x_i always move towards the left. When picking the optimal action, agent i's calculation of the expected action of the other agent will be $E(\hat{a}_j) = 0$. Consider the trade-off faced by the agent if she chooses to report her true preference and choose $a_i = x_i$: Given this choice she expects the mean of the actions to be $E_i[\bar{a}] = x_i/2$ and the distance between her preference and the mean to be $E_i[x_i - \bar{a}] = x_i/2$. We know that i's utility is decreasing both in the distance between her type and the mean action and in the distance between her type and her action. By reporting $a_i = x_i$ the agent does not incur any cost from misreporting, and her expected loss is purely the cost of the joint outcome being misaligned with her preference, $EU(x_i, x_i) = -\frac{1-r}{4}x_i^2 - \frac{\mu(r)^2}{4}E(x^2)$, where $E(x^2) = \int_{\mathbb{R}} x_j^2 g(x_j)dx_j$. If instead, she exaggerates her opinion by ϵ in the direction away from zero, she successfully moves the mean closer to her own preference x_i. However in doing so, she also incurs an extra cost from lying, which is increasing with ϵ. Overall, her expected utility is $EU(x_i, x_i + \epsilon) = -r\epsilon^2 - \frac{1-r}{4}(x_i - \epsilon)^2 - \frac{\mu(r)^2}{4}E(x^2)$. For small values of ϵ, $EU(x_i, x_i + \epsilon) > EU(x_i, x_i)$ and the converse is true for ϵ large enough. Hence, in equilibrium, i picks the optimal value of a_i that minimizes the loss from the distance between the group's outcome and their own preference, but one that does not inflate the cost of exaggerating.[8]

Next, recall that group polarization is defined as the tendency of the joint outcome to move towards a more extreme point in the direction indicated by the members' original preferences. The equilibrium derived above satisfies this definition. The mean pre-deliberation preference of the group is $\bar{x} = \frac{x_i + x_j}{2}$ while the mean post-deliberation outcome is $\bar{a} = \frac{\hat{a}_i + \hat{a}_j}{2} = \frac{1+r}{1+3r}(x_i + x_j)$. If $\bar{x} > 0$, then $\bar{a} > \bar{x}$; else if $\bar{x} < 0$, then $\bar{a} < \bar{x}$. Hence, if

[8]We have not explicitly included abstention as part of the players' strategy set. Abstention can be seen as equivalent to not voicing any opinion. The analysis above would hold if we assume that not voicing any opinion implies an action that is consistent with true preferences. When agent i abstains, her actions are aligned with her true preferences and so her utility from the first term to $-r(x_i - a_i)^2 = 0$. However, by abstaining, she has no effect on the group's final outcome and $\bar{a} = a_j$. So the utility from abstaining is $-(1-r)(x_i - a_j)^2$ which is strictly lower than the utility of voicing her true preferences (and obtaining $-(1-r)\left(x_i - \frac{x_i + a_j}{2}\right)^2$.) Thus abstaining is always a dominated choice.

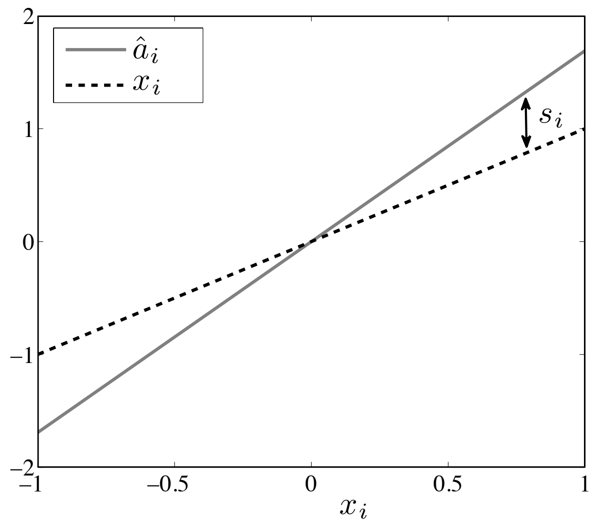

Figure 1: Equilibrium actions and shading in a two-player simultaneous game with $r = 0.1$.

the preferences of the two agents in the group is initially predisposed towards the right, the group decision is even more rightwards. Alternatively, if the group is predisposed towards the left, then its joint decision is even more leftwards.

Comparative Statics

To investigate the comparative statics, we denote the extent to which an agent i shades her opinion in equilibrium as $s_i = |\hat{a}_i - x_i| = \frac{1-r}{1+3r}|x_i|$. Figure 1 depicts an example of the equilibrium actions of agents whose preferences are drawn from $U[-1, 1]$ and their shading as a function of their preference x_i, for $r = 0.1$. We can see that $\frac{\mathrm{d}s_i}{\mathrm{d}r} \leq 0$ and as would be expected agents shade their actions less if the cost associated with lying is higher. Second, $\frac{\mathrm{d}s_i}{\mathrm{d}|x_i|} > 0$ suggests that agents near the extremes shade more than moderates, who are closer to the center. For example, as discussed in the introduction consumers who have strong liberal and conservative political preferences are the ones who are more vocal in their Twitter activity about the advertising campaigns of brands. Similarly this also provides a rationale for the U-shaped relationship described in Delmas et al. (2016) between the greenhouse emissions and lobbying intensity: the more extreme companies with the least and most carbon emissions are also the more extreme voices in the climate debate.

The result also implies that the overall group shift is proportional to the initial tendency of the group. The mean shift of a group is given by $\bar{s} = |\bar{a} - \bar{x}| = \frac{1-r}{1+3r}|\bar{x}|$ and so the shift exhibited by an extreme group of agents is higher than that exhibited by a relatively moderate group. While all groups tend towards the extremes in their decisions, this effect is exacerbated in extreme groups.

An interesting point is that the extent of shading, s_i, is independent of the distribution $g(x)$ as long as it is symmetric. Agents' shade the same amount irrespective of whether they are drawn from a uniform distribution or from a more polarized preference distributions where the masses are near the extrema. This suggests that polarization in this model does not stem from preference polarization. Rather it is a strategic choice made by agents in order to influence the group decisions.

Extensions and Robustness

We now consider several extensions of the main model to illustrate the robustness of the results as well as to develop a comprehensive theoretical account by examining how polarization is influenced by the game structure, group size and characteristics, preference characteristics and the informational endowment of the agents.

A. Group Size Effects ($m > 2$)

We first extend the game to interactions between $m > 2$ agents to investigate the role of the group size. Recall that the preferences of the agents are private information and are independently drawn and all agents simultaneously choose their public action $\{a_{\mathrm{m},i}, a_{\mathrm{m},j}, ...\} \in \mathbb{R}$, with agent i's action denoted as $a_{\mathrm{m},i}$. Agent i's utility is given by $u(x_i, a_{\mathrm{m},i}, a_{\mathrm{m},-i}) = -r(x_i - a_{\mathrm{m},i})^2 - (1-r)\,(x_i - \bar{a}_{\mathrm{m}})^2$, where $a_{\mathrm{m},-i}$ denotes the actions of all agents except i and $\bar{a}_{\mathrm{m}} = \frac{1}{m}\sum_{i=1}^{m} a_{\mathrm{m},i}$.

In the Bayesian Nash equilibrium of this game agent i chooses action $\hat{a}_{\mathrm{m},i} = \frac{rm^2+(1-r)m}{rm^2+(1-r)} x_i$. This shows the robustness of the equilibrium of the two-agent group. Agent i's strategy is linear in her type x_i with the multiplier $\mu_{\mathrm{m}}(r) = \frac{rm^2+(1-r)m}{rm^2+(1-r)}$. The extent of shading by agent i is $s_{\mathrm{m},i} = |\hat{a}_{\mathrm{m},i} - x_i| = \frac{(1-r)(m-1)}{rm^2+(1-r)}|x_i|$ and so as in the basic model, for all m, r, agents near the extreme shade their opinions more than those near the center. The extent of shading is also increasing in r, and independent of $g(\cdot)$.

The m-agent case provides an additional insight – it shows that the level of shading s_i is increasing in m up to $\frac{1}{\sqrt{r}} + 1$, but decreasing after that (see Web Appendix §A.2.2). In other words, after a certain point, as the number of players increases, agents tend to shade less. As $m \to \infty$, shading goes to zero, $i.e.$, players report the truth. Agents exaggerate in order to pull the mean ($\bar{a}$) closer to their preference. But when the number of agents in the group becomes large, the marginal impact of any one agent's action on the group mean outcome becomes negligible. Hence, in very large groups, the incentive to exaggerate is low. This indicates a plausible solution to the polarization problem – a social planner wishing to reduce polarization of actions may do so by picking larger groups. However, while group size can be a potential remedy, it may also have practical limits. For example, involving large groups in decision-making is likely to be costly to implement in some cases and may simply be infeasible in others. For example, faculty groups in many schools are small. Therefore, later in the paper we will analyze the role of timing of the actions and ask whether sequential choices may be a potential solution to the polarization problem.

B. Sub-group Interactions

In many situations interactions occur between sub-groups, where each sub-group consists of many agents having the same preferences over an issue, but different from that of the other sub-group. For example, academic marketing departments consists of quantitative and behavioral sub-groups, political deliberations in the U.S., *e.g.*, in the Senate, occur between multi-agent Democratic and Republican sub-groups. On issues such as abortion rights, gun control, and taxation, conservatives groups have different preferences than liberals, but citizens within each sub-group have similar preferences.

Consider an extension of the basic model with a population of $m > 2$ agents that are divided into two subgroups 1 and 2 of sizes n_1 and n_2 so that $n_1 + n_2 = m$. The

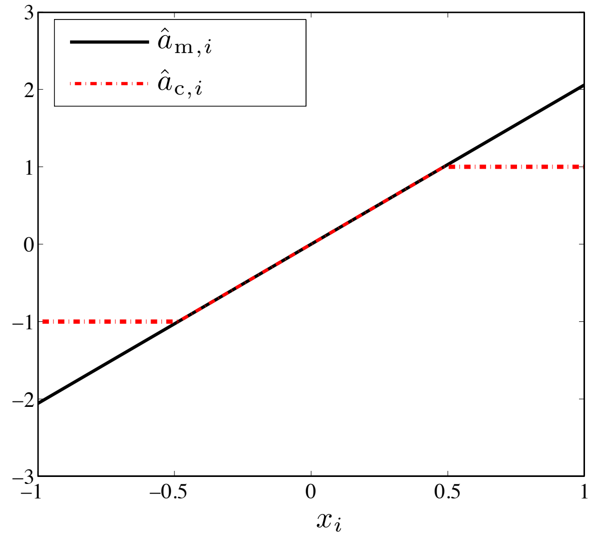

Figure 2: Equilibrium actions in unconstrained and constrained m-player simultaneous choice games with $r = 0.1$ and $m = 5$.

sub-group sizes are common knowledge. The preferences x_1 and x_2 of the sub-groups are independently drawn from $g(x)$. All agents within a sub-group know their individual preferences (and that of the others within their sub-group), but they do not observe the preferences of the other sub-group. We can write the expected utility of an individual agent i from sub-group 1 and agent j from sub-group 2 as:

$$EU_1^i = -r(x_1 - a_1^i)^2 - (1-r)\left(x_1 - \frac{1}{m}\left(a_1^i + \sum_{k_1=1(\neq i)}^{n_1} a_1^{k_1} + \sum_{k_2=1}^{n_2} \int_{\mathbb{R}} a_2^{k_2} g(x_2)dx_2\right)\right)^2 \tag{6a}$$

$$EU_2^j = -r(x_2 - a_2^j)^2 - (1-r)\left(x_2 - \frac{1}{m}\left(a_2^j + \sum_{k_2=1(\neq i)}^{n_2} a_2^{k_2} + \sum_{k_1=1}^{n_1} \int_{\mathbb{R}} a_1^{k_1} g(x_1)dx_1\right)\right)^2 \tag{6b}$$

In Web Appendix A.3, we present the solution for the symmetric (for agents within a sub-group) Bayesian Nash equilibrium and show that the equilibrium actions are $\hat{a}_i(x_i, m, n_i) =$

$x_i \frac{m(mr+(1-r))}{m^2 r+(1-r)n_i}$ and $\hat{a}_j(x_j, m, n_j) = x_j \frac{m(mr+(1-r))}{m^2 r+(1-r)n_j}$. The main results of the two-agent model continue to hold: the sub-group's action is linear in its preferences and sub-groups near the extremes shade more.

The main point of this analysis is to understand the role of the sub-group size on actions. Specifically, for a given population size m, how would a sub-group's size affect actions. It can be seen that for a given m, $\frac{\partial \hat{a}_i}{\partial n_i}$ and $\frac{\partial \hat{a}_j}{\partial n_j}$are both negative. The implication is that smaller sub-groups can become even more extreme. For example, in the academic hiring scenario the department chair may expect to see more extreme opinions and actions if one of the sub-groups is smaller than the other. This result can be seen as being consistent with the role of the Tea party movement in U.S. politics which was associated with pulling the Republican party more to right and with adopting increasingly conservative economic and social positions. For example, the Tea party members in the senate adopted increasingly conservative positions on environment, trade, budget, and immigration (Todd et al., 2014). This has happened even as the percentage of tea party supporters reported by polls diminished from 30% at the beginning of 2011 to 17% in October 2015 (Gallup, 2015).

C. Constrained Choice Model

So far we allowed agents to choose actions beyond the range of types. However, this may not be always feasible. For example, if the public action is supposed to be a revelation of private type, then it is impossible for agents to proffer a type which doesn't exist. Thus, we now consider a model where agents' actions are bounded within a credible range. Specifically, consider a scenario where agents' preferences are drawn from a uniform distribution bounded at -1 and 1, i.e., $U[-1, 1]$, and they are required to choose an action that lies between $[-1, 1]$.[9] As in the earlier models, agents simultaneously choose actions and the mean of their actions is implemented as the group's decision.

In Web Appendix §A.4, we show that in a m-player game constrained-choice, the optimal response function continues to be $\frac{rm^2+(1-r)m}{rm^2+(1-r)} x_i$ till it hits the bounds (i.e, 1 or -1),

[9]The results translate directly to any symmetric bounded preference distribution. We use $U[-1, 1]$ mainly to illustrate the actions pictorially in Figure 2.

and then is constrained to be at the bounds of the distribution. Please see Figure 1 for a pictorial depiction of the equilibrium actions in this bounded setting (denoted by $\hat{a}_{\mathrm{c},i}$).

In sum, expanding the number of players, constraining the choice set, or considering sub-groups does not affect the key results. Hence, moving forward, we retain the two-player, unconstrained choice model and focus our attention on other interesting modifications such as information revelation and sequential choices.

D. Asymmetric Type Distribution

We now consider a situation where the distribution of types, $g(x)$ is asymmetric, i.e., $E(x) \neq 0$. In Web Appendix §A.5, we show that for any general asymmetric distribution $g(x)$, then we can derive $a_i = \frac{2(1+r)}{1+3r} x_i - \frac{1-r}{1+3r} E(x)$. The effect of the asymmetry of the type distribution is intuitive. Suppose $x_i > 0$ and agent i has right leaning preferences, but that the distribution of the agent types is skewed in the opposite direction, i.e., $E(x) < 0$. In this case, agent i will have the incentive to be more extreme than in the symmetric case and to shade her action even more to the right. In contrast, if the distribution is skewed in the same direction as the agent's preference, this works against the tendency to be extreme and may even lead to moderation.

E. Partial Knowledge

In many instances players may have some knowledge about the preferences of their rivals. This may especially be the case in smaller groups such as faculty groups or corporate teams in which members have a history of prior interactions. For example, suppose that each agent i knows whether the other agent j is in $\mathbb{R}_+$ or $\mathbb{R}_-$ but actual locations of (x_i, x_j) are still private information for the respective players. Thus each agent knows which side the other "leans" toward, but not their exact preference.

Given the setup there are two possible cases of partial knowledge: First, the case where each agent knows that the other player's preference is on the opposite side (i.e., opposite leaning). The alternative case is one in which each agent knows that the other is on the same side (i.e., similar leaning). The equilibrium analysis is in Web Appendix §A.6.

Opposite Leaning Agents: Suppose without loss of generality let $x_i \in \mathbb{R}_+$ and $x_j \in \mathbb{R}_-$. In other words, agent i is known to be left leaning and agent j to be right leaning. The equilibrium actions in this case turn out to be:

$$\hat{a}_i = \frac{2(1+r)}{1+3r} x_i - \frac{(1-r^2)}{2r(1+3r)} \int_{\mathbb{R}_-} x_j g(x_j) \mathrm{d}x_j \tag{7}$$

$$\hat{a}_j = \frac{2(1+r)}{1+3r} x_j - \frac{(1-r^2)}{2r(1+3r)} \int_{\mathbb{R}_+} x_i g(x_i) \mathrm{d}x_i \tag{8}$$

Recall from Proposition 1 that in the full private information case each agent's actions are a function of only their own preferences. With partial knowledge the equilibrium actions of agent i responds not only to her own preference x_i, but also to $E(x_j) = \int_{\mathbb{R}_-} x_j g(x_j) \mathrm{d}x_j$ conditional on the knowledge that the other agent is opposite leaning. Thus each agent's actions are now also a function of the knowledge of the rival that they possess.

Note that $\frac{2(1+r)}{1+3r} > 1$ which means that with opposite leaning agents, the actions are still more extreme in response to the own preference. Further, the knowledge that x_j is opposite leaning also adds to the polarization of a_i. Note that $\frac{1-r^2}{2r(1+3r)} < \frac{2(1+r)}{1+3r}$, as long as $r > \frac{1}{5}$. Thus, an agent responds more to her own private type compared to partial information about the rival as long as the truth-telling motive is strong enough. Finally, we have $\bar{a} = \frac{1+r}{1+3r}(x_i + x_j)$ and so similar to Proposition 1 we will have that $\bar{x} > 0$, $\bar{a} > \bar{x}$ and $\bar{x} < 0$, $\bar{a} < \bar{x}$. Thus, the extent of polarization in the group outcome remains the same as in the main model.

Similar Leaning Agents: Next, consider the alternative case in which both agents know that they are on the same side of zero, i.e., they are similar leaning. Without loss of generality let $(x_i, x_j) \in R_+$. As derived in the Appendix the equilibrium actions are:

$$\hat{a}_i = \frac{2(1+r)}{1+3r} x_i - \frac{(1-r)}{(1+3r)} E(x_j) \tag{9}$$

$$\hat{a}_j = \frac{2(1+r)}{1+3r} x_j - \frac{(1-r)}{(1+3r)} E(x_i), \tag{10}$$

where $E(x) = \int_{\mathbb{R}_+} x g(x) \mathrm{d}x$. When the group members are similar leaning and on the same side of zero, the presence of partial knowledge can have a moderating influence and the equilibrium actions will be less extreme. Indeed, when a player's preference is sufficiently small when compared to the expected value of the preference distribution (e.g., x_i is small enough compared to $E(x)$), it is even possible that the player will choose a moderating action away from her direction of preference and towards zero. We can also note that the partial knowledge of the rival's preference (which is expressed through the effect of the expected value of the preference $E(x)$) has a greater effect on actions when agents have opposite as compared to similar preference leanings. Finally, the mean of the group actions is $\bar{a} = \frac{2(1+r)}{1+3r}(\bar{x}) - \frac{(1-r)}{1+3r}E(x)$. Thus the group outcome maybe more extreme than the mean realized preferences but is also moderated by partial knowledge to the extent of the expected preference.

Thus, in general partial knowledge induces agents to consider the available information about the other player and this can be a force for moderation of group actions when agents end up being similar leaning. More generally, suppose that each agent has a noisy (but better than the prior) information signal about the private information of the other agent. Then as the precision of the signal improves, the preferences of the other player will have a greater effect in moderating actions. At the extreme with perfect information we will get the previously discussed the benchmark case in which agents have full information. The availability of information about the private preferences of group members can thus be used as a strategic instrument by the principal to moderate group behavior.

F. Incentive to Disclose Preferences

Suppose agents had the opportunity to verifiably communicate their preference would they have the incentive to do so? Such communication which makes preferences common knowledge has the potential to reduce polarization in actions. Accordingly, consider an ex-ante disclosure game in which each agent i simultaneously chooses whether or not to reveal the private information about x_i prior to the agents choosing their public actions a_i. In Web Appendix §A.7, we solve for the equilibrium of the disclosure game and show that the scenario where both agents choose not to disclose their type is an equilibrium. Thus our basic model with private information emerges as an equilibrium even when players can communicate their preferences.

G. Alternative Preferences

In order to have a better perspective of the role of the group influence motive in our model, we compare it with some alternative preference formulations to understand what type of preferences may counter the polarization of actions.

Consider first an alternative formulation in which each agent's social preference is to minimize the distance between their actions and the true mean preference of the group. This can be seen as a taste for conformity with group preferences. Might this help to reduce the extent of polarization in actions? Specifically, suppose that agent i minimizes the distance between a_i and and average of the group's true preference $\bar{x}$. Thus the alternative utility function for say i would be $U(x_i, a_i, a_j) = -r(x_i - a_i)^2 - (1-r)(a_i - \bar{x})^2$, and equivalently for j. The equilibrium actions, as derived in Web appendix §A.8, are $\hat{a}_i = \frac{(1+r)}{2}x_i$ and $\hat{a}_j = \frac{(1+r)}{2}x_j$. These equilibrium actions show moderation, in that they are closer to the center than the agent's true preferences. This is natural since i has an incentive to move her action closer to $\frac{x_i+E(x_j)}{2}$. Since $E(x) = 0$, this naturally leads to the agent moving closer to zero, or moderating her action. Further, the equilibrium outcome is $\bar{a} = \frac{(1+r)}{4}(x_i + x_j)$, and $|\bar{a}| < |\bar{x}|$. Thus the group outcome is more moderate than the average preferences.

As another alternative formulation, consider the case when agents care about minimizing the distance between their public actions and the group's outcome. In other words, the agent cares about how closely their actions or voiced opinions conform to the mean outcome of the group. Specifically, suppose that the agent i's utility function was $U(x_i, a_i, a_j) = -r(x_i - a_i)^2 - (1-r)(a_i - \bar{a})^2$, and equivalently for j. We derive the equilibrium in the appendix and can show that for a symmetric distribution of agent types, the equilibrium actions are $\hat{a}_i = \frac{4r}{1+3r}x_i$. This implies that $|\hat{a}_i| < |x_i|$; so the agents moderate their actions to be closer to center, and the group outcome is also moderate compared to the mean of preferences, i.e., $|\bar{a}| < |\bar{x}|$. Overall, if people care about conforming with their peer's actions, we get moderation in actions and group decisions, similar to the main findings in (Bernheim, 1994).

Together, these two extensions with alternative preferences highlight the importance and role of the group influence motive (i.e., the agents' desire to move the group's outcome closer to her true preferences) in driving polarization.

Sequential Actions

We now consider the case in which players may voice their opinions in sequence. On Yelp consumers see the past reviews while providing their ratings. Similarly on social media groups individuals may express opinions in sequence. A department meeting the chair may mandate the order in which different faculty members may speak. Indeed in many institutional settings members typically take turns to speak. In the Federal Open Market Committee (FOMC) meetings the members of the committee express their preferred policy position sequentially in an order that varies across meetings. The committee chairman summarizes these positions into an overall group directive on the federal short term interest rates. Similarly, in juries and legislative bodies the order of speaking is often pre-determined by the institutional rules.

Accordingly, we consider a two-period model in which one of the agents is randomly picked to speak in the first period and the other follows in the second period upon observing the action taken by the first. We refer to this model as the "exogenous" sequential choice model where the order of agent actions is exogenously determined and is uncorrelated to the agents' preferences. The speaking order can be interpreted as being either determined by institutional rules or by a third-party. Then in a subsequent section we will consider the case when the agents bid to endogenously choose the speaking order.

Equilibrium in the Sequential Game

Let $a_{\mathrm{xt},i}$ denote agent i's action in period t, in this e**<u>x</u>**ogenous sequential actions game. Without loss of generality, suppose that agent i speaks in the first period and j in the second period. We solve for the Perfect Bayesian equilibrium (PBE) for this game, and derive the equilibrium actions of both players starting with the second player. A PBE consists of strategy profile (and associated beliefs) for the two agents that specify their optimal actions given their beliefs and the strategies of the other agent. Further, the beliefs of each agent are consistent with the strategy profile and are determined by Bayes rule where possible. In this game, the first agent i's strategy $a_{\mathrm{x}1,i}(x_i, \hat{a}_{\mathrm{x}2,j})$ is a function of her type x_i and her (consistent) beliefs about the optimal actions of j in period 2, whereas the second agent j's strategy $a_{\mathrm{x}2,j}(x_j, a_{\mathrm{x}1,i})$ is a function of her type and the action of player i that she observes.

Period 2 – The utility of the second player j when she chooses action $a_{\mathrm{x}2,j}$ in response to the first player's observed action $a_{\mathrm{x}1,i}$ is $u\left(x_j, a_{\mathrm{x}2,j}, a_{\mathrm{x}1,i}\right) = -r(x_j - a_{\mathrm{x}2,j})^2 - (1 - r)\left(x_j - \bar{a}_x\right)^2$ where $\bar{a}_{\mathrm{x}} = \frac{a_{\mathrm{x}1,i}+a_{\mathrm{x}2,j}}{2}$. The optimal choice of agent j given the first period choice of i can be derived as $\hat{a}_{\mathrm{x}2,j} = \frac{2(1+r)}{1+3r}x_j - \frac{(1-r)}{1+3r}a_{\mathrm{x}1,i}$.

Period 1 – We can now solve for i's first period choice. While i doesn't know the second player's type, her belief will be that j will choose an optimal action $\hat{a}_{\mathrm{x}2,j}$ in response to her action. So her expected utility from choosing action $a_{\mathrm{x}1,i}$ is obtained by taking the expectation of $u\left(x_i, a_{\mathrm{x}1,i}, \hat{a}_{\mathrm{x}2,j}\right)$ over the full range of x_j which gives us:

$$EU_{\mathrm{x}1}\left(x_i, a_{\mathrm{x}1,i}\right) = -r\left(x_i - a_{\mathrm{x}1,i}\right)^2 - (1-r)\left[\left(x_i - \frac{2r}{1+3r}a_{\mathrm{x}1,i}\right)^2 + \left(\frac{1+r}{1+3r}\right)^2 \int_{\mathbb{R}} x_j^2 g(x_j)\mathrm{d}x_j\right] \tag{11}$$

Taking the F.O.C of equation (11) and following a similar analysis to that in the simultaneous equilibrium case gives us the equilibrium action of i as $\hat{a}_{\mathrm{x}1,i} = \frac{(1+3r)(3+r)}{(1+3r)^2+4r(1-r)}x_i$. Therefore, the first player i has a unique optimal response that is both linear in her type x_i and is symmetric around zero.

Characterizing the Group Outcome

Having derived the individual equilibrium actions, we now proceed to characterize the mean equilibrium outcome. For a given x_i and x_j, the mean equilibrium outcome is $\bar{a}_{\mathrm{x}} = \frac{\hat{a}_{\mathrm{x}1,i}+\hat{a}_{\mathrm{x}2,j}}{2} = \frac{2r(3+r)}{(1+3r)^2+4r(1-r)}x_i + \frac{1+r}{1+3r}x_j$. Before stating the results, we define the following relationships:

- *Polarization*: $|\bar{a}_{\mathrm{x}}| > |\bar{x}|$ and $\bar{a}_{\mathrm{x}}\bar{x} > 0$. Polarization is said to have occurred if the mean of equilibrium actions ($\bar{a}_{\mathrm{x}}$) is more extreme (farther away from zero) than the mean of preferences $\bar{x}$, and this shift is in the direction of the group's initial tendency $\bar{x}$.

- *Reverse Polarization*: $|\bar{a}_{\mathrm{x}}| > |\bar{x}|$ and $\bar{a}_{\mathrm{x}}\bar{x} \leq 0$. Reverse Polarization is the case where the mean equilibrium outcome ($\bar{a}_{\mathrm{x}}$) is more extreme than the mean of preferences $\bar{x}$, and the shift is in the direction *opposite* to the group's initial tendency $\bar{x}$.

- *Moderation*: $|\bar{a}_{\mathrm{x}}| \leq |\bar{x}|$. Moderation refers to the case where the mean of the equilibrium actions ($\bar{a}_{\mathrm{x}}$) lies closer to zero than the mean of preferences ($\bar{x}$).

The following proposition summarizes the equilibrium extent of shading as measured by the relationship between the mean actions and preferences:

Proposition 2 *Let* $k_1(r) = \frac{(1+3r)(1-r)}{(1+3r)^2+4r(1-r)}$, $k_2(r) = \frac{(1+3r)\left[(1+3r)^2+16r\right]}{[(1+3r)^2+4r(1-r)][2(1+r)+(1+3r)]}$, *and* $k_3(r) = \frac{(1+3r)^3+8r(1-r)(1+r)}{2(1+r)[(1+3r)^2+4r(1-r)]}$, *and without loss of generality, let* $x_i \geq 0$. *Comparison of the mean equilibrium outcome* ($\bar{a}_{\mathrm{x}}$) *with the mean of preferences* ($\bar{x}$)*:*

- *Polarization occurs if* $x_j > k_1(r)x_i$ *or* $x_j < -x_i$.
- *Reverse Polarization occurs if* $-x_i \leq x_j < -k_2(r)x_i$.
- *Moderation occurs if* $-k_2(r)x_i \leq x_j \leq -k_1(r)x_i$

Proof: See Appendix □

The effect of sequential actions on group polarization is summarized in Figure 3. Polarization occurs whenever the second player's preference is relatively extreme or comparable to that of the first player $i.e.$, $x_j > k_1(r)x_i$ or $x_j < -k_2(r)x_i$. In a sequential game, the second player can condition her action on that of the first player and is therefore always able to pull the mean outcome $\bar{a}_{\mathrm{x}}$ close to her own preference. When j is extreme, she pulls the mean outcome to the extreme too, thereby leading to polarization. Note that within this region when $-x_i \leq x_j \leq -k_2(r)x_i$, the polarization is reverse in the sense that it is in the direction opposite to that implied by $\bar{x}$. This happens when x_i and x_j lie on

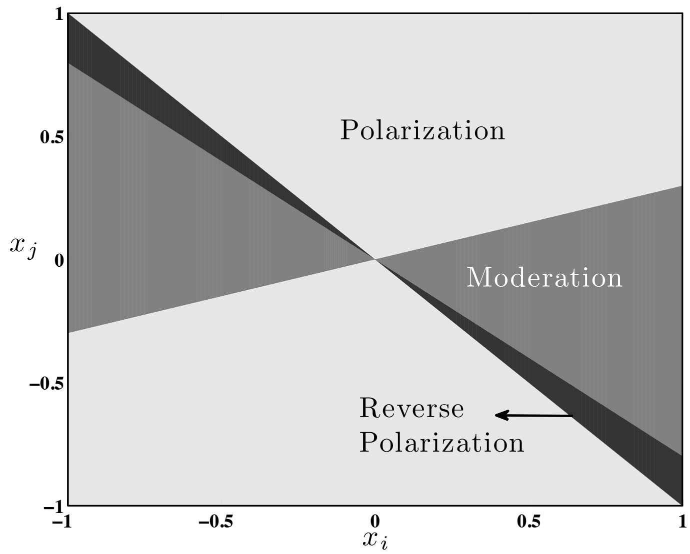

Figure 3: Regions of polarization, reverse polarization, and moderation in an exogenous sequential choice game; shown for $\{x_i, x_j\}$ drawn from $U[-1, 1]$.

opposite sides of zero, and $|x_i|$ is slightly greater than $|x_j|$. In other words, while the agents have preferences that are on opposite side of the issue, the first mover's preference is only slightly more intense. This implies that the mean group preference $\bar{x}$ lies on the same side of zero as the first mover i. However, in the second period, agent j's optimal action is able to ensure that the mean action $\bar{a}_{\text{x}}$ is closer to her than to i, $i.e.$, lies on the same side of zero as her own preference x_j – opposite to that of x_i and $\bar{x}$. Therefore, in this region, the group's mean action or outcome can be seen as being polarized but in the reverse direction.

In contrast, when the second mover j's preference is closer to zero compared to i, then she can choose her second period action so as to bring the group's outcome closer to her preference. This provides a moderating influence, and the overall outcome is closer to zero than the mean preferences. Thus both polarization and moderation are possible in this exogenous sequential game with the actual outcome depending upon on the relative preferences of both players and leans in the direction of the second player. So if the second player is relatively extreme, the outcome is also extreme; however if she is moderate, the outcome is moderate too.

Comparing Simultaneous and Sequential Games

We compare the simultaneous and sequential action games to understand how the extent of group polarization is affected by the timing of actions.

Proposition 3 *Let $\{\hat{a}_i, \hat{a}_j\}$ and $\{\hat{a}_{\mathrm{x}1,i}, \hat{a}_{\mathrm{x}2,j}\}$ denote the equilibrium actions of i and j in the simultaneous choice and exogenous sequential choice games, respectively. Without loss of generality, let $x_j > 0$. Then:*

a) $\hat{a}_{\mathrm{x}2,j} \geq \hat{a}_j$ *if* $x_i \leq 0$ *and* $\hat{a}_{\mathrm{x}2,j} < \hat{a}_j$ *if* $x_i > 0$.

b) $|\hat{a}_{\mathrm{x}1,i}| \geq |\hat{a}_i|$ *and* $\frac{\mathrm{d}|\hat{a}_{\mathrm{x}1,i}|}{\mathrm{d}r} < 0$

Proof: See Appendix $\square$

Consider first the action of the second player j in the sequential game $\hat{a}_{\mathrm{x}2,j} = \frac{2(1+r)}{1+3r}x_j - \frac{(1-r)}{1+3r}\hat{a}_{\mathrm{x}1,i}$. Part (a) of the Proposition shows that if the players lie on opposite sides of zero, then j becomes more extreme in the sequential actions game as compared to the simultaneous game. In contrast, when the two players preferences on the same side of zero, then in the second period j is less extreme, in response to i's action. That is, when the first player i chooses an action close to j's own preference, then she is more moderate in comparison to the simultaneous case. This is because exaggeration in the simultaneous case is driven by the anticipation of the other players' opinion. But in the sequential case player j already observes an action that shows that player i is not from the opposite camp and so the incentive to exaggerate decreases.

In contrast, the first player i's action in the sequential game is always more extreme than that in the simultaneous case ($i.e.$, $|\hat{a}_{\mathrm{x}1,i}| \geq |\hat{a}_i|$) because i knows that the second player j can compensate for her action in either direction. This is not an issue when j's preferences are similar to her own. But if the preferences happen to be very different, then by virtue of speaking second, j can nullify the effect of i's actions. Because this effect does not exist in the simultaneous game, i's action in the sequential game is more extreme.

Next, we compare the equilibrium outcomes in the two game formats.

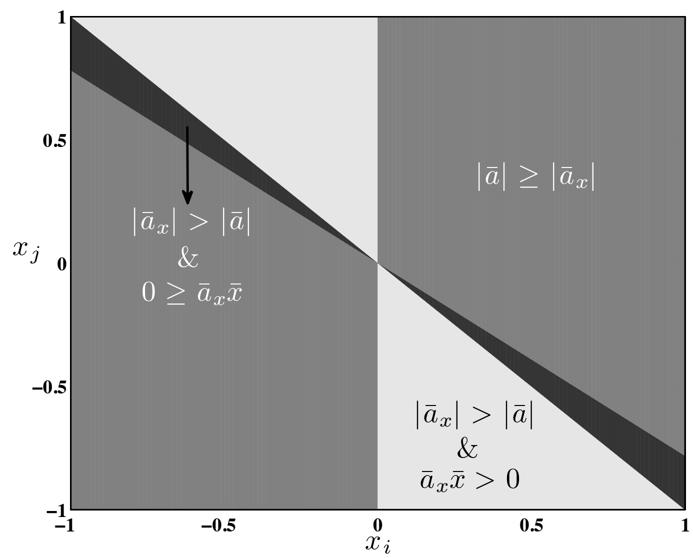

Figure 4: Comparison of mean outcome in the exogenous sequential game with that in the simultaneous game; shown for $\{x_i, x_j\}$ drawn from $U[-1, 1]$.

Proposition 4 *Comparison of the mean equilibrium outcome in the exogenous sequential game $\bar{a}_{\mathrm{x}}$ with that from the simultaneous game ($\bar{a}$):*

- $|\bar{a}_{\mathrm{x}}| > |\bar{a}|$ *and* $\bar{a}_{\mathrm{x}}\bar{x} > 0$ *if* $x_j < -x_i$, *i.e., the mean outcome in the exogenous sequential game is more extreme than that in the simultaneous game, in the direction of the initial tendency* $\bar{x}$.
- $|\bar{a}_{\mathrm{x}}| > |\bar{a}|$ *and* $\bar{a}_{\mathrm{x}}\bar{x} \leq 0$ *if* $-x_i \leq x_j < -k_3(r)x_i$, *i.e., the mean outcome in the exogenous sequential game is more extreme than that in the simultaneous game, but in the "opposite" direction of the initial tendency* $\bar{x}$.
- $|\bar{a}_{\mathrm{x}}| \leq |\bar{a}|$ *if* $x_j \geq -k_3(r)x_i$, *i.e., the mean outcome in the exogenous sequential game is moderate compared to the mean outcome in the simultaneous game,*

Proof: see Appendix $\square$

If i and j are relatively similar, then the mean outcome is more moderate in the sequential game ($i.e.$, $|\bar{a}_{\mathrm{x}}| < |\bar{a}|$). On the other hand, if i and j lie on opposite sides of zero (are relatively different) and j is relatively extreme, then the mean outcome in the sequential game is more extreme (see Figure 4). Overall, the propositions above reveal the insight that sequential actions may make agents more polarized than in the simultaneous case if they have divergent preferences. But if the individuals have relatively similar preferences, sequential actions has the potential to lead to less polarization of actions. These results highlight how the timing of the game may be exploited to combat polarization. For example, a planner or a group coordinator with an objective to reduce polarization can do so by assigning speaking orders if she expects players to be similar. However, if she expects them to be dissimilar, she may instead opt for a simultaneous choice format.

Value of the Speaking Order

We now analyze the relative value of the speaking order for the players by comparing the ex-ante expected utilities from speaking in the first and second periods. Agent i's a priori expected utility from speaking first and choosing action $a_{\mathrm{x},i}$ is given by equation (11). In equilibrium, i optimally chooses action $\hat{a}_{\mathrm{x}1,i}$. Substituting this into equation (11) gives us:

$$EU_{\mathrm{x}1}\left(x_i, \hat{a}_{\mathrm{x}1,i}\right) = -\frac{(1-r)(1+r)^2}{(1+3r)^2 + 4r(1-r)} x_i^2 - \frac{(1-r)(1+r)^2}{(1+3r)^2} \int_{\mathbb{R}} x^2 g(x) \mathrm{d}x \qquad (12)$$

Similarly, we can calculate i's a priori expected utility from speaking second as follows:

$$\begin{aligned} EU_{\mathrm{x}2}\left(x_i, \hat{a}_{\mathrm{x}2,i}\right) &= \frac{\int_{\mathbb{R}} u\left(x_i, \hat{a}_{\mathrm{x}2,i}, \hat{a}_{\mathrm{x}1,j}\right) g(x_j) \mathrm{d}x_j}{\int_{\mathbb{R}} g(x_j) \mathrm{d}x_j} \\ &= -\frac{r(1-r)}{1+3r} x_i^2 - \frac{r(1-r)(1+3r)(3+r)^2}{\left[(1+3r)^2 + 4r(1-r)\right]^2} \int_{\mathbb{R}} x^2 g(x) \mathrm{d}x \qquad (13) \end{aligned}$$

The following proposition compares the equilibrium expected utilities of speaking in the first and second periods, for a player i of type x_i.

Proposition 5 *Let $D_{\mathrm{x}}(x_i) = EU_{\mathrm{x}1}\left(x_i, \hat{a}_{\mathrm{x}1,i}\right) - EU_{\mathrm{x}2}\left(x_i, \hat{a}_{\mathrm{x}2,i}\right)$ denote the difference between the equilibrium expected utilities of agent i from speaking in the first and second periods.*

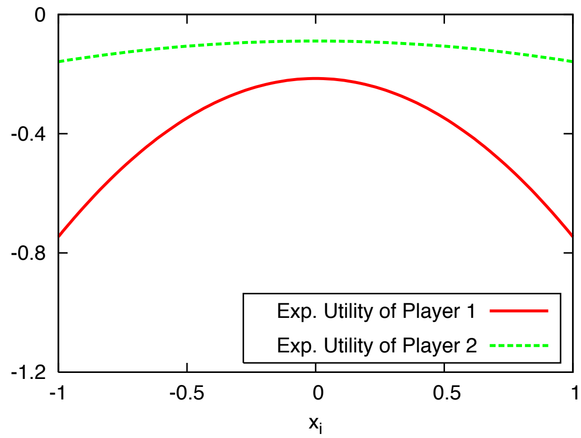

Figure 5: Players' expected utilities in the exogenous sequential choice game. Shown for $\{x_i, x_j\}$ drawn from $U[-1, 1]$ for $r = 0.1$.

a) $D_{\mathrm{x}}(x_i) \leq 0 \Rightarrow$ *for any agent* i*, the expected utility from speaking in the second period is greater than or equal to that from speaking in the first period.*

b) $\frac{\mathrm{d}D_{\mathrm{x}}(x_i)}{\mathrm{d}|x_i|} \leq 0 \Rightarrow$ *the difference in expected utilities of speaking in the first and second periods is increasing in* $|x_i|$.

Proof: See Appendix $\square$

The proposition highlights an important trade-off in incentives: Moving first allows an agent to "set the agenda" by committing to an observable action, whereas moving second allows the agent the flexibility to optimally adjust to the first period actions. The analysis indicates a rationale for why agents would wait to react to the actions of other agents, rather than to act first and set the group's agenda. The general point is that, irrespective of the type of the agent, the social influence motive makes the value of flexibility that comes from speaking second to be higher than the commitment value of speaking first and setting the agenda. A player who speaks second observes the first player's action and has the opportunity, if needed, to compensate for it. This works to the second player's advantage irrespective of whether or not the first player's action was close to or not from her preference. If the first player chose an action very different from the second player's preference, then she can compensate by picking a more extreme action in the opposite direction. But if the first player were to choose an action which is already close to her own preference, then the second player can also choose an action close to her preference and thereby not incur the cost of exaggerating.

Because the second player can can adjust her action based on the observed actions of the first player, the first player in anticipation of this behavior has the incentive to be more extreme, which in turn reduces her utility even more. Thus when agents care about influencing the overall group outcome towards their true preference, they prefer if given the choice to wait and delay their actions. The benefit of speaking second is higher for players who are more extreme – moderates have less to lose from speaking first. In general, moderates suffer less from decisions which are away from the middle. However, a player whose preference is more extreme on the right suffers a lot if the final outcome is more extreme to the left (and vice-versa). In sum, the analysis suggests that there are inherent advantages to waiting, especially for agents with more extreme preferences.

Endogenous Sequential Actions: Bidding to Speak

As described above the trade-off faced between truth-telling and group influence leads to a preference among agents to wait and speak in the second period and further such a strategy is more beneficial for players with extreme preferences. The natural question is what would happen in a group where the speaking order is endogenous. Given that speaking second is the dominant choice, there would exist a market for the order of speaking which may be characterized by allowing agents to endogenously bid for the right to determine the speaking order. In reality such an endogenous choice game implies the idea that group members may be willing to take costly actions to determine whether they are able to speak in the most favorable position.

Consider then an extension to the game where in a prior period 1 both agents participate in a first-price sealed bid auction for the opportunity to decide the speaking order. This first stage auction can also be seen agents lobbying to influence the speaking order. A neutral organizer/auctioneer receives agents' bids and announces the winner: if $b_i > b_j$, then i is the winner and in the event of a tie the winner is randomly chosen. The organizer announces the winner (but not the bid amounts) and so each player's beliefs will be based on inferences about the other's type depending upon who won the auction. In period 2, the winner chooses the preferred speaking order. In period 3 the players act based on the speaking-order determined by the winner. Players' have the same utility as in equation (1), except now the winner of the auction (say i) also pays her bid b_i to the organizer in period 1 for the right to choose the speaking order.

We derive the symmetric equilibrium bidding strategies of this game where the equilibrium bidding functions $\beta(x)$ are symmetric around zero. Unlike in the standard auction models where the bidder valuations are exogenously specified, the challenge in deriving the equilibrium strategies in this model stems from the fact that a bidder's valuation for the speaking order is endogenous to the outcome of the auction itself.

Proposition 6 *In the game where the agents participate in a first-price sealed-bid auction to decide on the right to determine the speaking order, there exists a unique symmetric PBE in which, agent i*

- *has a bidding strategy $\beta(x_i) = f(r)\frac{\int_0^{x_i} x^2 g(x)\mathrm{d}x}{\int_0^{x_i} g(x)\mathrm{d}x}$ and chooses to speak second if she wins the auction. The function $f(r)$ is defined in Web Appendix §F and has the property that $f(r) > 0 \, \forall \, r > 0$.*

- *If agent i speaks first, then she chooses $\hat{a}_{\mathrm{n}1,i} = \frac{(1+3r)(3+r)}{(1+3r)^2+4r(1-r)}x_i$, whereas if she speaks second, she chooses $\hat{a}_{\mathrm{n}2,i} = \frac{2(1+r)}{1+3r}x_i - \frac{(1-r)}{1+3r}a_{\mathrm{n}1,j}$.*

Proof: See Web Appendix F. □

Note that the equilibrium actions of the agents in this endogenous game end up being the same as that in the exogenous sequential game. Clearly, the agent who moves second faces the same game as the agent in the exogenous sequential game because she always chooses her response $a_{\mathrm{n}2,i}$ in response to the first player's action $a_{\mathrm{n}1,j}$. The incentives facing the first player is more subtle. If she is speaking first; this could either be because she won the auction and chose to go first, or because she lost and was asked to go first by the other player (the former case turns out to be off the equilibrium path). Regardless, i does not know j's type because j has not yet spoken. Therefore, i's actions will depend on her beliefs about j's type which will be based on the observed outcome of the auction and j's choice of speaking order (if j had the opportunity to decide it).

In a symmetric PBE, the region, say, W to which i can expect j to belong to is symmetric around zero, irrespective of the exact scenario under which i is speaking first. Hence, i's expected utility from speaking first is obtained by taking the expectation of $u\left(x_i, a_{\mathrm{n}1,i}, \hat{a}_{\mathrm{n}2,j}\right)$ over $x_j \in W$, *i.e.*, $EU_{\mathrm{n}1}\left(x_i, a_{\mathrm{n}1,i}\right) = \frac{\int_W u(x_i, a_{\mathrm{n}1,i}, \hat{a}_{\mathrm{n}2,j}) g(x_j) \mathrm{d}x_j}{\int_W g(x_j)\mathrm{d}x_j}$. This can be simplified by substituting for $\hat{a}_{\mathrm{n}2,i}$:

$$\begin{aligned} EU_{\mathrm{n}1}\left(x_i, a_{\mathrm{n}1,i}\right) = & - \quad r(x_i - a_{\mathrm{n}1,i})^2 - (1-r)\left[\left(x_i - \frac{2r}{1+3r}a_{\mathrm{n}1,i}\right)^2 + \left(\frac{1+r}{1+3r}\right)^2 \frac{\int_W x_j^2 g(x_j)\mathrm{d}x_j}{\int_W g(x_j)\mathrm{d}x_j}\right] \\ & - \quad 2\frac{1+r}{1+3r}\left(x_i - \frac{2r}{1+3r}a_{\mathrm{n}1,i}\right)\frac{\int_W x_j g(x_j)\mathrm{d}x_j}{\int_W g(x_j)\mathrm{d}x_j} \end{aligned} \tag{14}$$

The last term vanishes because W is symmetric around zero. By setting $\left.\frac{\mathrm{d}EU_{\mathrm{n}1}(x_i, a_{\mathrm{n}1,i})}{\mathrm{d}a_{\mathrm{n}1,i}}\right|_{a_{\mathrm{n}1,i} = \hat{a}_{\mathrm{n}1,i}} = 0$, we can solve for $\hat{a}_{\mathrm{n}1,i} = \frac{(1+3r)(3+r)}{(1+3r)^2 + 4r(1-r)}x_i$ which turns out to be the same as in the exogenous sequential case.

Players' equilibrium beliefs are that upon losing, they will forfeit the right to decide the speaking order, and the right to move second. Given this, the equilibrium bidding strategy can be specified. In deriving the equilibrium bidding strategy, note that a players' value from winning the auction is endogenous, unlike a traditional first-price sealed-bid auction, where players' valuations are exogenously given. The approach to deriving the equilibrium is to show that equilibrium bidding strategies are increasing strictly monotonically in $|x_i|$. The equilibrium bidding strategy is derived in the Appendix to be $\beta(x_i) = f(r)\frac{\int_0^{x_i} x^2 g(x)\mathrm{d}x}{\int_0^{x_i} g(x)\mathrm{d}x}$ and it is monotonically increasing in $|x_i|$. Thus agents with more extreme preferences have higher value for choosing the speaking order and will accordingly bid higher. The multiplier, $f(r)$, of the equilibrium bidding function is monotonically decreasing in r, i.e., as players' need to pull the final outcome ($\bar{a}_{\mathrm{n}}$) closer to own preference increases (as r decreases), their bid increases. At $r = 0$, the bidding strategy simplifies to $\beta(x_i)|_{r=0} = 2\frac{\int_0^{x_i} x^2 g(x)\mathrm{d}x}{\int_0^{x_i} g(x)\mathrm{d}x}$, which is the highest, whereas at $r = 1$, the bidding strategy devolves to $\beta(x_i)|_{r=1} = 0$.

Consider now the mean equilibrium outcome of the endogenous sequential actions game. For a given x_i and x_j suppose $|x_i| < |x_j|$, without loss of generality. Then j wins the auction and chooses to speak second, and the mean equilibrium outcome will be $\bar{a}_{\mathrm{n}} = \frac{\hat{a}_{\mathrm{n}1,i}+\hat{a}_{\mathrm{n}2,j}}{2} = \frac{2r(3+r)}{(1+3r)^2+4r(1-r)}x_i + \frac{1+r}{1+3r}x_j \quad \forall \quad |x_i| < |x_j|$. Proposition 7 compares the mean outcome, $\bar{a}_{\mathrm{n}}$, with the mean of the preferences $\bar{x}$ and the mean outcome in the simultaneous choice game $\bar{a}$.

Proposition 7 *In the equilibrium of the endogenous sequential actions game polarization always occurs ($|\bar{a}_{\mathrm{n}}| > |\bar{x}|$ and $\bar{a}_{\mathrm{n}}\bar{x} > 0$). Comparison of the extent of polarization across the different games yields:*

- *If $x_i \cdot x_j < 0$, then $|\bar{a}_{\mathrm{n}}| > |\bar{a}|$.*
- *If $x_i \cdot x_j > 0$, then $|\bar{a}_{\mathrm{n}}| \leq |\bar{a}|$*

Allowing the agents to compete for the right to speak always leads to the polarization of actions. When the speaking order is endogenous, the agent who wins the right to speak always prefers to wait. Further, it is the agents with more extreme preferences who have the incentive to bid more for the right to determine the speaking order. This leads to the important point that if agents were to bid for the right to speak then the actions of the agents and the group outcome is always polarized. *Unlike in the exogenous sequential game where moderation is a possibility, allowing agents to choose the speaking order always leads to polarization.*

It is also useful to compare the outcome of the endogenous sequential actions game with that of the simultaneous game. If the two players lie on opposite sides of zero, then endogenous sequential actions produces more polarization than the simultaneous actions game. On the other hand, if both players lie on the same side of zero, then the outcome is less polarized than that in the simultaneous game. The basic mechanism at play is that the second player in the sequential actions game can condition her action to that of the first player and accordingly pull the group outcome closer to her own preference. Recall that the first player's action in the sequential case is always more extreme than in the simultaneous case because of a compensation effect: i.e., she knows that the second player can observe and compensate for her action. In the endogenous sequential actions game, it is the more extreme player who ends up winning the right to be the second player. Given this, the player who loses the auction and speaks first can infer that the other player has more extreme preferences. This inference induces her to be even more extreme. When the two players' preferences are on opposite sides of zero then not only does the second player have the incentive to be more extreme (after observing the first player's actions) in order to pull the joint outcome towards her preference, but the inference effect also induces the first player to be more extreme. Consequently, the group becomes more polarized than in the simultaneous actions game. In contrast, when the players' preferences are on the same side of zero, the second player's knowledge of first's actions implies that she does not need to shade and take too extreme actions. The implication is that when the players are similarly inclined, endogenizing the speaking order can help to reduce polarization.

Welfare Comparisons

We start with the planner's problem to understand how a principal would design the group interaction to maximize social welfare. The welfare in the two player system for any x_i and x_j is given by Equation (2). Note that any pecuniary transfers (such as bids) are canceled out since they remain within the system and hence have no impact on the total welfare. In discussion forums, the speaking formats (or the timing game-forms) are design decisions and are often chosen before the agents' preferences are drawn. Therefore we can consider the expected welfare for a given game-form across the distribution of player types as a relevant measure for making welfare comparisons, i.e., $EW = \frac{\int_{x_i}\int_{x_j} W(x_i,x_j)g(x_i)g(x_j)\mathrm{d}x_i\mathrm{d}x_j}{\int_{x_i}\int_{x_j} g(x_i)g(x_j)\mathrm{d}x_i\mathrm{d}x_j}$. Denote the expected welfare for the first-best case to be EW_{FB}, the simultaneous case by

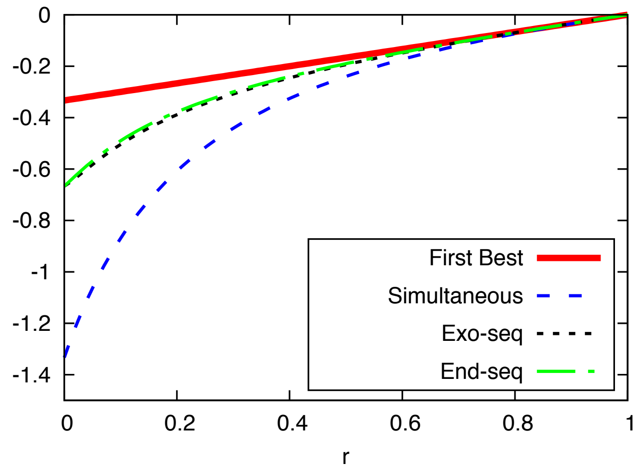

Figure 6: Expected social welfare for the four game forms. Shown for $U[-1, 1]$.

EW_s, the exogenous sequential by EW_x and the endogenous sequential by EW_n.

Figure 6 shows the relationship between expected welfare functions as a function of r for the case of $U[-1, 1]$. It can be seen that $EW_n > EW_x > EW_s$. Please see the Appendix for a detailed derivation of this inequality. Within the sequential action formats allowing agents to endogenously bid for the speaking order increases expected welfare as compared to the exogenous assignment of speaking order across the agents. The market clearing for the speaking order through the first-price auction mechanism improves efficiency. Further, it can be seen that the expected welfare under the sequential game (irrespective of whether it is exogenous or endogenous) is higher than that for the simultaneous game. It is interesting that even though the endogenous sequential choice game produces higher polarization, it also increases the welfare by allowing those with more extreme preferences to obtain the outcomes they desire.

Conclusion

Many formal and informal forums in business, organizational and socio-political settings facilitate group interactions shape views and outcomes on important issues ranging from brand choices, faculty hiring, diversity, and gun control. One might expect such interactions to help in the exchange of information and in better coordination. However, one only need to look the current landscape in the U.S. to observe that deliberations on many of these issues, if anything, lead to more polarization. Understanding the mechanisms that make actions and opinions more polarized is important, because they may create conflict and impede effective policy making.

We develop a theory that links the polarization of actions to a fundamental trade-off faced by agents between influencing others in a deliberation and representing their true preferences. When agents take actions with the objective to have group influence, deliberations can lead to polarization. Further, it is the more extreme agents who end up becoming more polarized in their actions. We analyze the role of simultaneous versus sequential timing of the actions of agents. With sequential actions polarization occurs whenever the agent who moves later is more extreme compared to the first agent. With sequential actions we also highlight the tradeoff between the commitment value of moving first versus the value of flexibility to optimally adjust to the first period actions that comes from moving later. The group influence incentive makes flexibility valuable and induces agents to wait.

The comparison of the different timings shows that if the preferences of the agents are dissimilar, then the sequential actions game produces less polarization as compared to simultaneous actions, whereas if the agents have similar preferences it is the simultaneous actions game that leads to less polarization. When the timing of sequential actions is endogenous and agents bid for the right to choose the order of actions, agents with more extreme preferences bid more, and the winning agents regardless of their preferences prefer to wait. We also examine the role of the group size and the presence of sub-groups to show that larger groups show less polarization, while smaller sub-groups tend to go more extreme. Further we investigate alternative preferences and information endowments to expand and clarify the role of the group influence motive in creating polarization.

The investigation of the incentives of firms (or principals) interested in managing polarization seems to be rich area for analysis. For example a firm might be interested in managing consumer opinions on social media and how they affect brand perceptions. Similarly platform design may have to deal with how groups in online forums interact. Another possible future direction would be to use data on group decisions and to identify the different sources of polarization in groups (e.g., strategic incentives, polarization of beliefs, and behavioral biases). Lab and/or field experiments or observational data with exogenous variation on these different sources can help further our understanding of how these factors contribute. It will also be usrful to emprically investigate the extent and the nature of the divergence between the polarization of actions and preferences/beliefs. It will also be interesting to test our theoretical predictions on how different mechanisms help moderate or exacerbate polarization, such as sequential vs. simultaneous model, group size, and observability of others' actions. Finally, it would be useful to tie polarized group decisions/outcomes to broader firm-level or societal outcomes, e.g., consumer engagement and social welfare.

References

D. Acemoglu, V. Chernozhukov, M. Yildiz, et al. Fragility of Asymptotic Agreement Under Bayesian Learning. Unpublished paper, Department of Economics, Massachusetts Institute of Technology, 2009.

J. Aguilar. At Shooting Range, Abbott Signs "Open Carry" Bill. The Texas Tribune, 2016. Online, posted 13-June-2015, retrieved 27-March-2016.

W. Amaldoss and S. Jain. Pricing of Conspicuous Goods: A Competitive Analysis of Social Effects. *Journal of Marketing Research*, 42(1):30–42, 2005.

J. Andreoni and T. Mylovanov. Diverging Opinions. *American Economic Journal: Microeconomics*, 4(1):209–232, 2012.

Armed Campuses, 2016. URL `http://www.armedcampuses.org/`. Online, acessed 27-March-2016.

D. Austen-Smith. Information transmission in debate. *American Journal of political science*, pages 124–152, 1990.

C. A. Bail, L. P. Argyle, T. W. Brown, J. P. Bumpus, H. Chen, M. B. F. Hunzaker, J. Lee, M. Mann, F. Merhout, and A. Volfovsky. Exposure to Opposing Views on Social Media can Increase Political Polarization. *Proceedings of the National Academy of Sciences*, 115(37):9216–9221, 2018. ISSN 0027-8424. doi: 10.1073/pnas.1804840115. URL `https://www.pnas.org/content/115/37/9216`.

S. Baliga, E. Hanany, and P. Klibanoff. Polarization and Ambiguity. *The American Economic Review*, 103(7):3071–3083, 2013.

R. S. Baron. So Right It's Wrong: Groupthink and the Ubiquitous Nature of Polarized Group Decision Making. *Advances in Experimental Social Psychology*, 37:219–253, 2005.

R. Bénabou. Groupthink: Collective Delusions in Organizations and Markets. *Review of Economic Studies*, 80(2):429–462, 2012.

B. D. Bernheim. A Theory of Conformity. *Journal of political Economy*, 102(5):841–877, 1994.

S. Bhattacharya and A. Mukherjee. Strategic information revelation when experts compete to influence. *The RAND Journal of Economics*, 44(3):522–544, 2013.

Campus Carry Policy. Campus Carry Policy Working Group – Final Report. Final Report, University of Texas at Austin, 2015. Online, retrieved 27-March-2016.

Y.-K. Che and I. L. Gale. Caps on political lobbying. *The American Economic Review*, 88(3): 643–651, 1998.

N. Cohn. Polarization Is Dividing American Society, Not Just Politics. New York Times, 2014. Online, posted 12-June-2014, retrieved 27-March-2016.

K. Corfman and D. Lehmann. Models of Cooperative Group Decision-Making and Relative Influence: An Experimental Investigation of Family Purchase Decisions. *Journal of Consumer Research*, 14(1):–13, 1987.

A. K. Dixit and J. W. Weibull. Political Polarization. *Proceedings of the National Academy of Sciences*, 104(18):7351–7356, 2007.

T. Eastland. Google in the Dock. Washington Examiner, 2018. Online, retrieved 1-Aug-2019.

J. Eliashberg and R. Winkler. Risk Sharing and Group Decision Making. *Management Science*, 27 (11):1221–1235, 1981.

J. Elisahberg, S. LaTour, A. Rangaswamy, and S. Louis. Assessing the Predictive Accuracy of Two Utility-Based Theories in a Marketing Channel Negotiation Context. *Journal of Marketing Research*, 23(2):101–110, 1986.

D. Fudenberg and J. Tirole. *Game Theory*. MIT Press, Cambridge, MA, 1991.

Gallup. Tea Party Movement, 2015. URL `http://www.gallup.com/poll/147635/tea-party-movement` Online, retrieved 27-March-2016.

T. W. Gilligan and K. Krehbiel. Asymmetric information and legislative rules with a heterogeneous committee. *American journal of political science*, pages 459–490, 1989.

E. L. Glaeser and C. R. Sunstein. Extremism and Social Learning. *Journal of Legal Analysis*, 1(1): 263–324, 2009.

B. Harstad, J. Svensson, et al. *Bribes, Lobbying, and Development*. 2006.

W. R. Hartmann, P. Manchanda, H. Nair, M. Bothner, P. Dodds, D. Godes, K. Hosanagar, and C. Tucker. Modeling Social Interactions: Identification, Empirical Methods and Policy Implications. *Marketing letters*, 19(3-4):287–304, 2008.

S. Huitlin. Guns on Campus: An Overview, 2015. URL `http://www.ncsl.org/research/education/guns-on-campus-overview.aspx`. Online, posted 5-October-2015, retrieved 27-March-2016.

D. J. Isenberg. Group Polarization: A Critical Review and Meta-Analysis. *Journal of Personality and Social Psychology*, 50(6):1141, 1986.

P. Kondor. The More We Know About The Fundamental, The Less We Agree On The Price. *The Review of Economic Studies*, 79(3):1175–1207, 2012.

V. Krishna and J. Morgan. A Model of Expertise. *The Quarterly Journal of Economics*, 116(2): 747–775, 2001.

C. G. Lord, L. Ross, and M. R. Lepper. Biased Assimilation and Attitude Polarization: The Effects of Prior Theories on Subsequently Considered Evidence. *Journal of Personality and Social Psychology*, 37(11):2098, 1979.

E. C. Main and T. G. Walker. Choice Shifts and Extreme Behavior: Judicial Review in the Federal Courts. *The Journal of Social Psychology*, 91(2):215–221, 1973.

D. G. Myers. Discussion-induced Attitude Polarization. *Human Relations*, 28(8):699–714, 1975.

D. G. Myers. Polarizing Effects of Social Interaction. *Group Decision Making*, pages 125–161, 1982.

J. Potters and F. Van Winden. Lobbying and Asymmetric Information. *Public Choice*, 74(3):269–292, 1992.

M. Rabin and J. L. Schrag. First Impressions Matter: A Model of Confirmatory Bias. *Quarterly journal of Economics*, pages 37–82, 1999.

V. Rao and J. Steckel. A Polarization Model for Describing Group Preferences. *Journal of Consumer Research*, pages 108–118, 1991.

J. Rawls. A Theory of Justice. *Bamforth N, Malik M and O'Cinneide C*, 1971.

J. A. F. Stoner. *A Comparison of Individual and Group Decisions Involving Risk*. PhD thesis, Massachusetts Institute of Technology, 1961.

M. Sun, X. Zhang, and F. Zhu. U-shaped Conformity in Online Social Networks. *Marketing Science*, 38(3):461–480, 2019.

C. R. Sunstein. The Law of Group Polarization. *Journal of Political Philosophy*, 10(2):175–195, 2002.

C. Todd, M. Murray, and C. Dann. There's The Tea Party – And Then There's Everyone Else. NBCNews, 2014. URL `http://www.nbcnews.com/politics/first-read/tea-party-stands-alone-key-issues-n`

Online, published = 19-June-2014, retrived = 27-March-2016.

H. Yoganarasimhan. Cloak or Flaunt? The Fashion Dilemma. *Marketing Science*, 31(1):74–95, 2012.

J. A. Zuber, H. W. Crott, and J. Werner. Choice Shift and Group Polarization: An Analysis of the Status of Arguments and Social Decision Schemes. *Journal of Personality and Social Psychology*, 62(1):50, 1992.

Web Appendix

A Extensions of the Simultaneous Action Game

A.1 General Linear Combination

We now consider a case where the group outcome is a linear combination of all agents' revealed preferences such that $\bar{a} = \alpha a_i + (1-\alpha)a_j$, where $0 < \alpha < 1$. We can then write the utility as:

$$u_\alpha(x_i, a_i, a_j) = -r(x_i - a_i)^2 - (1-r)(x_i - \alpha a_i - (1-\alpha)a_j)^2 \tag{A1}$$

Because j's preference (x_j) is her private information at the time of choosing the action, i's expected utility from choosing a_i as $EU_\alpha(x_i, a_i) = \frac{\int_{\mathbb{R}} u_\alpha(x_i, a_i, \hat{a}_j) g(x_j) \mathrm{d}x_j}{\int_{\mathbb{R}} g(x_j) \mathrm{d}x_j}$. By differentiating $EU(x_i, a_i)$ and setting it equal to zero at i's equilibrium action $a_i = \hat{a}_i$ gives us:

$$\left.\frac{\partial EU_\alpha(x_i, a_i)}{\partial a_i}\right|_{a_i = \hat{a}_i} = 2r(x_i - \hat{a}_i) + 2(1-r)\alpha \left[x_i - \alpha \hat{a}_i) + (1-\alpha) \int_{\mathbb{R}} \hat{a}_j g(x_j) \mathrm{d}x_j\right] = 0. \tag{A2}$$

We can then write the equilibrium action of i as:

$$\hat{a}_i = \frac{r + (1-r)\alpha}{r + (1-r)\alpha^2} x_i - \frac{(1-r)\alpha(1-\alpha)}{r + (1-r)\alpha^2} \int_{\mathbb{R}} \hat{a}_j g(x_j) \mathrm{d}x_j \tag{A3}$$

As in the main model, we can easily show that $\int_{\mathbb{R}} \hat{a}_j g(x_j) \mathrm{d}x_j = 0$. Therefore, we have:

$$\hat{a}_i = \frac{r + (1-r)\alpha}{r + (1-r)\alpha^2} x_i \tag{A4}$$

We can easily show that $|\hat{a}_i| > x_i$ and $|\alpha \hat{a}_i + (1-\alpha)\hat{a}_j| > \bar{x}$. Thus, we see polarization in individual agents' actions as well as the group outcome.

A.2 Analysis of m-Player Game

A.2.1 Equilibrium of m-Player Game

To derive the Bayesian Nash equilibrium for a m-player simultaneous game, we can write the utility of a player i as:

$$u(x_i, a_{\mathrm{m},i}, a_{\mathrm{m},-i}) = -r(x_i - a_{\mathrm{m},i})^2 - (1-r)\left(x_i - \frac{a_{\mathrm{m},i}}{m} - \frac{\sum_{j=1}^{m-1} a_{\mathrm{m},j}}{m}\right)^2 \tag{A5}$$

To obtain i's expected utility for any $a_{\mathrm{m},i}$ in equilibrium, we can take the expectation of equation (A5) over all the other $m-1$ agents' equilibrium actions:

$$EU_{\mathrm{m}}(x_i, a_{\mathrm{m},i}) = \frac{\int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \int_{\mathbb{R}} \left(u(x_i, a_{\mathrm{m},i}, \hat{a}_{\mathrm{m},-i})\right) g(x_1) g(x_2) \ldots g(x_{m-1}) \mathrm{d}x_1 \mathrm{d}x_2 \ldots \mathrm{d}x_{m-1}}{\int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \int_{\mathbb{R}} g(x_1) g(x_2) \ldots g(x_{m-1}) \mathrm{d}x_1 \mathrm{d}x_2 \ldots \mathrm{d}x_{m-1}} \tag{A6}$$

Substituting for $u(x_i, a_{\mathrm{m},i}, a_{\mathrm{m},-i})$ from equation (A5) and simplifying, we have:

$$\begin{aligned} & EU_{\mathrm{m}}\left(x_i, a_{\mathrm{m},i}\right) = -r(x_i - a_{\mathrm{m},i})^2 \qquad \text{(A7)} \\ + \quad & \frac{2(1-r)}{m}\left(x_i - \frac{a_{\mathrm{m},i}}{m}\right) \frac{\int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \int_{\mathbb{R}} \left(\sum_{j=1}^{m-1} a_{\mathrm{m},j}\right) g(x_1)g(x_2)...g(x_{m-1})\mathrm{d}x_1\mathrm{d}x_2...\mathrm{d}x_{m-1}}{\int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \int_{\mathbb{R}} g(x_1)g(x_2)...g(x_{m-1})\mathrm{d}x_1\mathrm{d}x_2...\mathrm{d}x_{m-1}} \\ - \quad & (1-r)\left[\left(x_i - \frac{a_{\mathrm{m},i}}{m}\right)^2 + \frac{\int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \int_{\mathbb{R}} \left(\sum_{j=1}^{m-1} a_{\mathrm{m},j}\right)^2 g(x_1)g(x_2)...g(x_{m-1})\mathrm{d}x_1\mathrm{d}x_2...\mathrm{d}x_{m-1}}{m^2 \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \int_{\mathbb{R}} g(x_1)g(x_2)...g(x_{m-1})\mathrm{d}x_1\mathrm{d}x_2...\mathrm{d}x_{m-1}}\right] \end{aligned}$$

We know that $\int_{\mathbb{R}} g(x)dx = 1$. Also, from i's perspective, all the other agents' types are i.i.d from the distribution $g(x)$. So we can simplify equation (A7) to:

$$\begin{aligned} EU_{\mathrm{m}}\left(x_i, a_{\mathrm{m},i}\right) = & - \quad r(x_i - a_{\mathrm{m},i})^2 + 2(1-r)\frac{m-1}{m}\left(x_i - \frac{a_{\mathrm{m},i}}{m}\right)\int_{\mathbb{R}} a_{\mathrm{m},j} g(x_j)\mathrm{d}x_j \qquad \text{(A8)} \\ & - \quad (1-r)\left[\left(x_i - \frac{a_{\mathrm{m},i}}{m}\right)^2 + \int_{\mathbb{R}} \cdots \int_{\mathbb{R}} \int_{\mathbb{R}} \left(\sum_{j=1}^{m-1} a_{\mathrm{m},j}\right)^2 g(x_1)g(x_2)...g(x_{m-1})\mathrm{d}x_1\mathrm{d}x_2...\mathrm{d}x_{m-1}\right] \end{aligned}$$

Taking the F.O.C of Equation (A8) w.r.t $a_{\mathrm{m},i}$ at $\hat{a}_{\mathrm{m},i}$ gives us:

$$\left.\frac{\mathrm{d}EU_{\mathrm{m}}\left(x_i, a_{\mathrm{m},i}\right)}{\mathrm{d}a_{\mathrm{m},i}}\right|_{a_{\mathrm{m},i}=\hat{a}_{\mathrm{m},i}} = 2r(x_i - \hat{a}_{\mathrm{m},i}) - \frac{(1-r)(m-1)}{m^2}\int_{\mathbb{R}} \hat{a}_{\mathrm{m},j} g(x_j)\mathrm{d}x_j + \frac{1-r}{m}\left(x_i - \frac{\hat{a}_{\mathrm{m},i}}{m}\right) = 0 \qquad \text{(A9)}$$

This gives us $\hat{a}_{\mathrm{m},i}$ as:

$$\hat{a}_{\mathrm{m},i} = \frac{\left[rm^2 + (1-r)m\right]x_i - (1-r)(m-1)\int_{\mathbb{R}} \hat{a}_{\mathrm{m},j} g(x_j)\mathrm{d}x_j}{rm^2 + (1-r)} \qquad \text{(A10)}$$

Integrating $\hat{a}_{\mathrm{m},i}$ w.r.t x_i, we have:

$$\int_{\mathbb{R}} \hat{a}_{\mathrm{m},i} g(x_i)\mathrm{d}x_i = \frac{\left[rm^2 + (1-r)m\right]\int_{\mathbb{R}} x_i g(x_i)\mathrm{d}x_i - (1-r)(m-1)\int_{\mathbb{R}} \hat{a}_{\mathrm{m},j} g(x_j)\mathrm{d}x_j \int_{\mathbb{R}} g(x_i)\mathrm{d}x_i}{rm^2 + (1-r)} \qquad \text{(A11)}$$

Since $\int_{\mathbb{R}} g(x_i)\mathrm{d}x_i = 1$ and $\int_{\mathbb{R}} \hat{a}_{\mathrm{m},i} g(x_i)\mathrm{d}x_i = \int_{\mathbb{R}} \hat{a}_{\mathrm{m},j} g(x_j)\mathrm{d}x_j = E(a_m)$ this simplifies to $E(a_m) = E(x)$. Since $E(x) = 0$ in a symmetric distribution, we have:

$$\hat{a}_{\mathrm{m},i} = \frac{rm^2 + (1-r)m}{rm^2 + (1-r)} x_i \qquad \text{(A12)}$$

A.2.2 Comparative Statics in m-Player Game

We now derive the comparative statics regarding the extent of shading $s_{\mathrm{m},i}$.

i) $\frac{\mathrm{d}s_{\mathrm{m},i}}{\mathrm{d}r} = (m-1)\frac{-(1-r)(m^2-1)-[rm^2+(1-r)]}{[rm^2+(1-r)]^2}|x_i|$. This simplifies to $\frac{\mathrm{d}s_{\mathrm{m},i}}{\mathrm{d}r} = -\frac{(m-1)m^2}{[rm^2+(1-r)]^2}|x_i|$. We know that $m, m-1 > 0$, because by definition $m \geq 2$. Also, it is clear that $[rm^2 + (1-r)]^2 > 0$ and that $|x_i| \geq 0$. It therefore follows that $\frac{\mathrm{d}s_{\mathrm{m},i}}{\mathrm{d}r} \geq 0$.

ii) $\frac{\mathrm{d}s_{\mathrm{m},i}}{\mathrm{d}|x_i|} = \frac{(1-r)(m-1)}{rm^2+(1-r)}$. Since $0 < r < 1$ and $m \geq 2$, this value is always positive. So $\frac{\mathrm{d}s_{\mathrm{m},i}}{\mathrm{d}|x_i|} > 0$.

iii) $\frac{\mathrm{d}s_{\mathrm{m},i}}{\mathrm{d}m} = (1-r)\frac{1-r(m-1)^2}{[rm^2+(1-r)]^2}$. If $m < \frac{1}{\sqrt{r}}+1$, then $1-r(m-1)^2 > 0$; so in this range $\frac{\mathrm{d}s_{\mathrm{m},i}}{\mathrm{d}m} > 0$. Else if $m > \frac{1}{\sqrt{r}} + 1$, then $1 - r(m-1)^2 < 0$; so in this range $\frac{\mathrm{d}s_{\mathrm{m},i}}{\mathrm{d}m} < 0$.

A.3 Equilibrium for Sub-group Interactions

The expected utility of agent i from sub-group 1 and agent j in sub-group 2 are given in Equations (6). The first-order condition for agent i is $\frac{\mathrm{d}EU_1^i}{\mathrm{d}a_1^i} = 0$ which can be calculated to be:

$$r(x_1 - a_1^i) + \frac{1-r}{m}\left(x_1 - \frac{a_1^i + \sum\limits_{k_1=1(\neq i)}^{n_1} a_1^{k_1} + \sum\limits_{k_2=1}^{n_2} \int_{\mathbb{R}} a_2^{k_2} dx_2}{m}\right) = 0 \tag{A13}$$

As we are looking for a symmetric in actions with sub-group equilibrium, we can set $a_1^i = a_1^k = \hat{a}_1$, and $a_2^k = \hat{a}_2$, and simplify (A13) to:

$$x_1(r + \frac{1-r}{m}) - \hat{a}_1(r + \frac{n_1(1-r)}{m^2}) - \frac{(1-r)n_2}{m^2}\int_{\mathbb{R}} \hat{a}_2 g(x_2)dx_2 = 0 \tag{A14}$$

Thus the equilibrium action $\hat{a}_1$ for agents from sub-group 1 is:

$$\hat{a}_1 = x_1\left(\frac{m(mr + (1-r))}{m^2r + n_1(1-r)}\right) - \frac{(1-r)n_2}{m^2r + n_1(1-r)}\int_{\mathbb{R}} \hat{a}_2 g(x_2)dx_2 \tag{A15}$$

Integrating the equilibrium action over the entire distribution of x_1 we can get:

$$\int_{\mathbb{R}} \hat{a}_1 g(x_1)dx_1 = E(x_1)\left(\frac{m(mr + (1-r))}{m^2r + n_1(1-r)}\right) - \frac{(1-r)n_2}{m^2r + n_1(1-r)}\int_{\mathbb{R}} \hat{a}_2 g(x_2)dx_2 \tag{A16}$$

By noting that $E(x_1) = 0$ for a symmetric distribution, we have that $\int_{\mathbb{R}} \hat{a}_1 g(x_1)dx_1 = \int_{\mathbb{R}} \hat{a}_2 g(x_2)dx_2$. Therefore the equilibrium $\hat{a}_1$ is $\hat{a}_1 = x_1\left(\frac{m(mr+(1-r))}{m^2r+n_1(1-r)}\right)$. Using similar analysis for sub-group 2 we can also derive $\hat{a}_2 = x_2\left(\frac{m(mr+(1-r))}{m^2r+n_2(1-r)}\right)$.

A.4 Analysis of Constrained Choice Game

For a m-player simultaneous choice where preferences are drawn from $U[-1,1]$ and actions are curtailed between $[-1,1]$, there exists a unique Perfect Bayesian equilibrium, where player i chooses action $\hat{a}_{\mathrm{c},i} = min\{\frac{rm^2+(1-r)m}{rm^2+(1-r)}x_i, 1\}$ if $x_i \geq 0$, and $\hat{a}_{\mathrm{c},i} = max\{\frac{rm^2+(1-r)m}{rm^2+(1-r)}x_i, -1\}$ if $x_i < 0$.

We now present the detailed proof for the above statement. Let $\hat{a}_{\mathrm{c},i}$ be i's equilibrium action, and let $A_{\mathrm{c},i}$ be i's optimal action. It therefore follows that $\hat{a}_{\mathrm{c},i} = min\{A_{\mathrm{c},i}, 1\}$ if $A_{\mathrm{c},i} \geq 0$ and $\hat{a}_{\mathrm{c},i} = max\{A_{\mathrm{c},i}, -1\}$ if $A_{\mathrm{c},i} < 0$. Player i's optimal action $A_{\mathrm{c},i}$ can be derived using the same steps those used in Web Appendix A.2.1. We therefore have:

$$A_{\mathrm{c},i} = \frac{\left[2rm^2 + 2(1-r)m\right] x_i - (m-1)\int_{-1}^{1} \hat{a}_{\mathrm{c},j} \mathrm{d}x_j}{2rm^2 + 2(1-r)} \tag{A17}$$

Note that $A_{\mathrm{c},i}$ is monotonically increasing in x_i. Hence, if at some point $x_i = q_2 > 0$, $A_{\mathrm{c},i} = 1$, then for all $x_i > q_2$, $A_{\mathrm{c},i} > 1$, which implies that $\hat{a}_{\mathrm{c},i} = 1$. Similarly, if at some point $x_i = q_1 < 0$, $A_{\mathrm{c},i} = -1$, then for all $x_i < q_2$, $A_{\mathrm{c},i} < -1$, which implies that $\hat{a}_{\mathrm{c},i} = -1$. Therefore, we can express $\int_{-1}^{1} \hat{a}_{\mathrm{c},i} \mathrm{d}x_i$ as follows:

$$\int_{-1}^{1} \hat{a}_{\mathrm{c},i} \mathrm{d}x_i = \int_{-1}^{q_1} -1 \mathrm{d}x_i + \int_{q_1}^{q_2} A_{\mathrm{c},i} \mathrm{d}x_i + \int_{q_2}^{1} 1 \mathrm{d}x_i \tag{A18}$$

Substituting for $A_{\mathrm{c},i}$ from Equation (A17) and integrating, we have:

$$\int_{-1}^{1} \hat{a}_{\mathrm{c},i} \mathrm{d}x_i = -(1+q_1) + \frac{1}{2}\frac{rm^2 + (1-r)m}{rm^2 + (1-r)}\left(q_2^2 - q_1^2\right) - (q_2 - q_1)\frac{m-1}{2rm^2 + 2(1-r)}\int_{-1}^{1} \hat{a}_{\mathrm{c},j} + (1-q_2) \tag{A19}$$

We also know that $A_{\mathrm{c},i}$ is just equal to -1 at q_1 and $A_{\mathrm{c},i}$ is just equal to 1 at q_2. So, we have:

$$-1 = \frac{\left[2rm^2 + 2(1-r)m\right] q_1 - (m-1)\int_{-1}^{1} \hat{a}_{\mathrm{c},j} \mathrm{d}x_j}{2rm^2 + 2(1-r)} \tag{A20}$$

$$1 = \frac{\left[2rm^2 + 2(1-r)m\right] q_2 - (m-1)\int_{-1}^{1} \hat{a}_{\mathrm{c},j} \mathrm{d}x_j}{2rm^2 + 2(1-r)} \tag{A21}$$

Next, we multiply Equation (A20) with q_1 and Equation (A21) with q_2 and subtract the latter from the former. This gives us:

$$(q_2 - q_1)\frac{m-1}{2rm^2 + 2(1-r)}\int_{-1}^{1} \hat{a}_{\mathrm{c},j} \mathrm{d}x_j = (q_2^2 - q_1^2)\frac{2rm^2 + 2(1-r)m}{2rm^2 + 2(1-r)} - (q_1 + q_2) \tag{A22}$$

Next, we substitute the L.H.S of Equation (A22) in Equation (A19) and obtain:

$$\int_{-1}^{1} \hat{a}_{\mathrm{c},i} \mathrm{d}x_i = -\frac{1}{2}\frac{rm^2 + (1-r)m}{rm^2 + (1-r)}\left(q_2^2 - q_1^2\right) \tag{A23}$$

Adding Equation (A20) to Equation (A21) gives us:

$$(q_1 + q_2)\frac{rm^2 + (1-r)m}{rm^2 + (1-r)} = \frac{(m-1)}{rm^2 + (1-r)}\int_{-1}^{1} \hat{a}_{\mathrm{c},j} \mathrm{d}x_j \tag{A24}$$

Substituting the L.H.S into Equation (A23) gives us:

$$\int_{-1}^{1} \hat{a}_{\mathrm{c},i} \mathrm{d}x_i = -\frac{q_2 - q_1}{2} \frac{(m-1)}{rm^2 + (1-r)} \int_{-1}^{1} \hat{a}_{\mathrm{c},j} \mathrm{d}x_j \tag{A25}$$

If $q_2 - q_1 = 0$, then we directly have: $\int_{-1}^{1} \hat{a}_{\mathrm{c},j} \mathrm{d}x_j = 0$. Else if $q_2 - q_1 \neq 0$, it still follows that $\int_{-1}^{1} \hat{a}_{\mathrm{c},i} \mathrm{d}x_i = \int_{-1}^{1} \hat{a}_{\mathrm{c},j} \mathrm{d}x_j = 0$. Hence the optimal response of agent i is:

$$A_{\mathrm{c},i} = \frac{rm^2 + (1-r)m}{rm^2 + (1-r)} x_i \tag{A26}$$

It is clear that $\frac{rm^2+(1-r)m}{rm^2+(1-r)} > 0$. So the equilibrium response of agent i is given by $\hat{a}_{\mathrm{c},i} = min\{\frac{rm^2+(1-r)m}{rm^2+(1-r)} x_i, 1\}$ if $x_i \geq 0$ and $\hat{a}_{\mathrm{c},i} = max\{\frac{rm^2+(1-r)m}{rm^2+(1-r)} x_i, -1\}$ if $x_i < 0$.

A.5 Asymmetric Distribution

We now consider a game where the distribution of types, $g(x)$, is asymmetric, i.e., $E(x) \neq 0$. To solve for the equilibrium in this case, we start with the general equation considered in Equation (4):

$$\hat{a}_i = \frac{2(1+r)}{1+3r} x_i - \frac{(1-r)}{(1+3r)} \int_{\mathbb{R}} \hat{a}_j g(x_j) \mathrm{d}x_j$$

As usual, integrating i's equilibrium action $\hat{a}_i$ over the entire range of x_i gives us:

$$\int_{\mathbb{R}} \hat{a}_i g(x_i) \mathrm{d}x_i = \frac{2(1+r)}{1+3r} \int_{\mathbb{R}} x_i g(x_i) \mathrm{d}x_i - \frac{(1-r)}{(1+3r)} \int_{\mathbb{R}} \hat{a}_j g(x_j) \mathrm{d}x_j \int_{\mathbb{R}} g(x_i) \mathrm{d}x_i \tag{A27}$$

Unlike the symmetric distribution case, here we know that $E(x_i) = \int_{\mathbb{R}} x_i g(x_i) \mathrm{d}x_i \neq 0$. But we can re-arrange the terms to write $\int_{\mathbb{R}} \hat{a}_i g(x_i) \mathrm{d}x_i$ as follows:

$$\int_{\mathbb{R}} \hat{a}_i g(x_i) \mathrm{d}x_i = E(x_i) \tag{A28}$$

We can now substitute the above equivalence in Equation (4) to derive the a_i as follows:

$$\hat{a}_i = \frac{2(1+r)}{1+3r} x_i - \frac{(1-r)}{1+3r} E(x_i) \tag{A29}$$

A.6 Partial Knowledge

A.6.1 Same-side Leaning Agents

First, we consider the case where both agents i and j lean on the same side. Without loss of generality, let both agents be drawn from $\mathbb{R}_+$. Then, we have:

$$\hat{a}_i = \frac{2(1+r)}{1+3r} x_i - \frac{(1-r)}{(1+3r)} \int_{\mathbb{R}_+} \hat{a}_j g(x_j) \mathrm{d}x_j \tag{A30}$$

Integrating i's equilibrium action $\hat{a}_i$ over the entire range of x_i gives us:

$$\int_{\mathbb{R}_+} \hat{a}_i g(x_i)\mathrm{d}x_i = \frac{2(1+r)}{1+3r}\int_{\mathbb{R}_+} x_i g(x_i)\mathrm{d}x_i - \frac{(1-r)}{(1+3r)}\int_{\mathbb{R}_+} \hat{a}_j g(x_j)\mathrm{d}x_j \int_{\mathbb{R}_+} g(x_i)\mathrm{d}x_i \quad \text{(A31)}$$

Since we know that $x_i \in \mathbb{R}_+$, $\int_{\mathbb{R}_+} g(x_i)\mathrm{d}x_i = 1$. This gives us:

$$\int_{\mathbb{R}_+} \hat{a}_i g(x_i)\mathrm{d}x_i = \int_{\mathbb{R}_+} x_i g(x_i)\mathrm{d}x_i \quad \text{(A32)}$$

Substituting the above equivalence into Equation (A33), we have:

$$\hat{a}_i = \frac{2(1+r)}{1+3r}x_i - \frac{(1-r)}{(1+3r)}\int_{\mathbb{R}_+} x_j g(x_j)\mathrm{d}x_j \quad \text{(A33)}$$

A.6.2 Opposite Leaning Agents

Next, we consider the case where i knows that j is on the opposite side (and vice-versa). Without loss of generality, let i be drawn from $\mathbb{R}_+$ and j be drawn from $\mathbb{R}_-$. Then, we have:

$$\hat{a}_i = \frac{2(1+r)}{1+3r}x_i - \frac{(1-r)}{(1+3r)}\int_{\mathbb{R}_-} \hat{a}_j g(x_j)\mathrm{d}x_j \quad \text{(A34)}$$

Integrating i's equilibrium action $\hat{a}_i$ over the entire range of x_i gives us:

$$\int_{\mathbb{R}_+} \hat{a}_i g(x_i)\mathrm{d}x_i = \frac{2(1+r)}{1+3r}\int_{\mathbb{R}_+} x_i g(x_i)\mathrm{d}x_i - \frac{(1-r)}{(1+3r)}\int_{\mathbb{R}_-} \hat{a}_j g(x_j)\mathrm{d}x_j \int_{\mathbb{R}_+} g(x_i)\mathrm{d}x_i \quad \text{(A35)}$$

Since we know that $x_i \in \mathbb{R}_+$, $\int_{\mathbb{R}_+} g(x_i)\mathrm{d}x_i = 1$. This gives us:

$$\int_{\mathbb{R}_+} \hat{a}_i g(x_i)\mathrm{d}x_i = \frac{2(1+r)}{1+3r}\int_{\mathbb{R}_+} x_i g(x_i)\mathrm{d}x_i - \frac{(1-r)}{(1+3r)}\int_{\mathbb{R}_-} \hat{a}_j g(x_j)\mathrm{d}x_j \quad \text{(A36)}$$

Similarly, for j, we can write:

$$\int_{\mathbb{R}_-} \hat{a}_j g(x_j)\mathrm{d}x_j = \frac{2(1+r)}{1+3r}\int_{\mathbb{R}_-} x_j g(x_j)\mathrm{d}x_j - \frac{(1-r)}{(1+3r)}\int_{\mathbb{R}_+} \hat{a}_i g(x_i)\mathrm{d}x_i \quad \text{(A37)}$$

Because the distribution $g(x)$ is symmetric, we know that $\int_{\mathbb{R}_+} x_i g(x_i) = -\int_{\mathbb{R}_-} x_j g(x_j)$. Therefore, adding up Equations (A36) and (A37), we have:

$$\int_{\mathbb{R}_+} \hat{a}_i g(x_i)\mathrm{d}x_i = -\int_{\mathbb{R}_-} \hat{a}_j g(x_j)\mathrm{d}x_j \quad \text{(A38)}$$

We can thus write $\int_{\mathbb{R}_-} \hat{a}_j g(x_j)\mathrm{d}x_j$ as follows:

$$\int_{\mathbb{R}_-} \hat{a}_j g(x_j)\mathrm{d}x_j = \frac{(1+r)}{2r}\int_{\mathbb{R}_-} x_j g(x_j)\mathrm{d}x_j \quad \text{(A39)}$$

Substituting this in Equation (A34), we have:

$$\hat{a}_i = \frac{2(1+r)}{1+3r}x_i - \frac{(1-r)(1+r)}{2r(1+3r)}\int_{\mathbb{R}_-} x_j g(x_j)\mathrm{d}x_j \quad \text{(A40)}$$

A.7 Game with Disclosure

Consider a two-stage game such that in Stage 1, agents have the opportunity to simultaneously reveal their type (which is verifiable). Then in Stage 2, the players simultaneously choose an action. There are four possible equilibria of this game: (D,D), (ND, ND), (D, ND), and (ND, ND), where D stands for Disclosure and ND stands for Non-Disclosure. The (ND, ND) equilibrium is equivalent to the baseline game considered in the main analysis, where agents' types are private information.

We first consider a (ND, ND) equilibrium and see whether agent i has the incentive to deviate. Suppose that i deviates are reveals her type and j plays the equilibrium strategy of not disclosing her type x_j. Then we can derive the optimal actions of j in this case as:

$$\hat{a}_j = \frac{2(1+r)}{1+3r}x_j - \frac{(1-r)}{(1+3r)}\hat{a}_i \tag{A41}$$

Note that there is no integral over a_i since j does not have uncertainty on i's type anymore. Next, we can derive a_i as:

$$\hat{a}_i = \frac{2(1+r)}{1+3r}x_i - \frac{(1-r)}{(1+3r)}\int_{\mathbb{R}} \hat{a}_j g(x_j)\mathrm{d}x_j \tag{A42}$$

Substituting a_j from Equation (A41) into the above equation, we have:

$$\hat{a}_i = \frac{1+3r}{4r}x_i \tag{A43}$$

Now we can write the expected utility of i when she plays the equilibrium strategy as:

$$EU_{ND,ND}(x_i, a_i) = -\frac{r(1-r)^2}{(1+3r)^2}x_i^2 - \frac{(1-r)4r^2}{(1+3r)^2}x_i^2 - \frac{(1-r)(1+r)^2}{(1+3r)^2}\int_{\mathbb{R}} x_j^2 g(x_j)\mathrm{d}x_j. \tag{A44}$$

Similarly, we can write the expected utility of i when she deviates as:

$$EU_{D,ND}(x_i, a_i) = -\frac{(1-r)^2}{16r}x_i^2 - \frac{(1-r)}{4}x_i^2 - \frac{(1-r)(1+r)^2}{(1+3r)^2}\int_{\mathbb{R}} x_j^2 g(x_j)\mathrm{d}x_j. \tag{A45}$$

Comparing Equations (A44) and (A45), we can see that $EU_{ND,ND}(x_i, a_i) > EU_{D,ND}(x_i, a_i)$ if:

$$\begin{aligned}
& -\frac{r(1-r)^2}{(1+3r)^2}x_i^2 - \frac{(1-r)4r^2}{(1+3r)^2}x_i^2 > \frac{(1-r)^2}{16r}x_i^2 - \frac{(1-r)}{4}x_i^2 \\
\Rightarrow \quad & -\frac{r(1-r)}{(1+3r)^2} - \frac{4r^2}{(1+3r)^2} > -\frac{(1-r)}{16r} - \frac{1}{4} \\
\Rightarrow \quad & -\frac{r}{(1+3r)} > -\frac{1+3r}{16r} \\
\Rightarrow \quad & -16r^2 + (1+3r)^2 > 0 \\
\Rightarrow \quad & (1+7r)(1-r) > 0
\end{aligned} \tag{A46}$$

The above inequality is always true. Hence, we know that agent i has no incentive to deviate from the (ND, ND) equilibrium. Based on the same inequality, we also know that both (D, ND) and (ND, D) cannot be equilibrium outcomes since the agent who is disclosing her type will always benefit from deviating and choosing to not disclose her type.

A.8 Alternative Preferences

We first consider the case where agents care about choosing an action that is close to the mean preferences of the group. That is, suppose that the agent's utility can be expressed as:

$$u_1(x_i, a_i, a_j) = -r(x_i - a_i)^2 - (1-r)(x_i - \bar{x})^2 \tag{A47}$$

Taking the expectation, i's expected utility from choosing a_i in this setting can be written as $EU_1(x_i, a_i) = \frac{\int_{\mathbb{R}} u_1(x_i, a_i, \hat{a}_j) g(x_j) \mathrm{d}x_j}{\int_{\mathbb{R}} g(x_j) \mathrm{d}x_j}$. By differentiating $EU_1(x_i, a_i)$ and setting it equal to zero at i's equilibrium action $a_i = \hat{a}_i$ gives us:

$$\left.\frac{\partial EU_1(x_i, a_i)}{\partial a_i}\right|_{a_i = \hat{a}_i} = 2r(x_i - \hat{a}_i) - 2(1-r)\left[\hat{a}_i - \frac{x_i}{2} - \frac{1}{2}\int_{\mathbb{R}} x_j g(x_j) \mathrm{d}x_j\right] = 0 \tag{A48}$$

This in turn simplifies to:

$$\hat{a}_i = \frac{(1+r)}{2} x_i \tag{A49}$$

Next, consider the second alternative preference structure, where agents care about choosing an action that is close to the group's decision ($\bar{a}$). We can write agent i's utility in this case as:

$$u_2(x_i, a_i, a_j) = -r(x_i - a_i)^2 - (1-r)(a_i - \bar{a})^2 \tag{A50}$$

Taking the expectation, i's expected utility from choosing a_i in this setting can be written as $EU_2(x_i, a_i) = \frac{\int_{\mathbb{R}} u_2(x_i, a_i, \hat{a}_j) g(x_j) \mathrm{d}x_j}{\int_{\mathbb{R}} g(x_j) \mathrm{d}x_j}$. By differentiating $EU_2(x_i, a_i)$ and setting it equal to zero at i's action $\hat{a}_i$, we have:

$$\left.\frac{\partial EU_2(x_i, a_i)}{\partial a_i}\right|_{a_i = \hat{a}_i} = 2r(x_i - \hat{a}_i) - \frac{(1-r)}{2}\left[\hat{a}_i - \int_{\mathbb{R}} \hat{a}_j g(x_j) \mathrm{d}x_j\right] = 0 \tag{A51}$$

Re-arranging the terms, we have:

$$\hat{a}_i = \frac{4r}{(1+3r)} x_i + \frac{(1-r)}{(1+3r)} \int_{\mathbb{R}} \hat{a}_j g(x_j) \mathrm{d}x_j \tag{A52}$$

Taking the integral of the above equation with respect to the distribution of x_i, we have:

$$\int_{\mathbb{R}} \hat{a}_i g(x_i) \mathrm{d}x_i = \frac{4r}{(1+3r)} \int_{\mathbb{R}} x_i g(x_i) \mathrm{d}x_i + \frac{(1-r)}{(1+3r)} \int_{\mathbb{R}} \hat{a}_j g(x_j) \mathrm{d}x_j \tag{A53}$$

Since $\int_{\mathbb{R}} x_i g(x_i)\mathrm{d}x_i = 0$, we have $\int_{\mathbb{R}} \hat{a}_i g(x_i)\mathrm{d}x_i = 0$. This gives us the equilibrium action of i as:

$$\hat{a}_i = \frac{4r}{(1+3r)} x_i \tag{A54}$$

B Equilibrium of the Exogenous Sequential Game

Let i and j be the first and second players, respectively. The second player j's equilibrium action has already been derived in the main text (see section **??**), and is equal to:

$$\hat{a}_{\mathrm{x}2,j} = \frac{2(1+r)}{1+3r} x_j - \frac{(1-r)}{1+3r} a_{\mathrm{x}1,i} \tag{A55}$$

So we now derive the first player's optimal action $\hat{a}_{\mathrm{x}1,i}$. In equilibrium, i's utility from choosing action $a_{\mathrm{x}1,i}$ is given by:

$$u\left(x_i, a_{\mathrm{x}1,i}, \hat{a}_{\mathrm{x}2,j}\right) = -r(x_i - a_{\mathrm{x}1,i})^2 - (1-r)\left(x_i - \bar{a}_{\mathrm{x}}\right)^2 \tag{A56}$$

where $\bar{a}_{\mathrm{x}} = \frac{a_{\mathrm{x}1,i} + \hat{a}_{\mathrm{x}2,j}}{2}$. We know that when i chooses $a_{\mathrm{x}1,i}$, in response, j chooses $\hat{a}_{\mathrm{x}2,j} = \frac{2(1+r)}{1+3r} x_j - \frac{(1-r)}{1+3r} a_{\mathrm{x}1,i}$. This in turn gives us $\bar{a}_{\mathrm{x}}$ as:

$$\bar{a}_{\mathrm{x}} = \frac{(1+r)x_j + 2r a_{\mathrm{x}1,i}}{1+3r} \tag{A57}$$

Substituting this value of $\bar{a}_{\mathrm{x}}$ into Equation (A56), we have:

$$\begin{aligned} u\left(x_i, a_{\mathrm{x}1,i}, \hat{a}_{\mathrm{x}2,j}\right) = -r(x_i - a_{\mathrm{x}1,i})^2 \quad &- \quad (1-r)\left[\left(x_i - \frac{2r}{1+3r} a_{\mathrm{x}1,i}\right)^2 + \left(\frac{1+r}{1+3r} x_j\right)^2\right] \\ &+ \quad 2(1-r)\left[\left(x_i - \frac{2r}{1+3r} a_{\mathrm{x}1,i}\right)\left(\frac{1+r}{1+3r} x_j\right)\right] \end{aligned} \tag{A58}$$

Hence, the expected utility of agent i in from choosing action $a_{\mathrm{x}1,i}$ in period 1 is:

$$EU_{\mathrm{x}1}\left(x_i, a_{\mathrm{x}1,i}\right) = \frac{\int_{\mathbb{R}} u\left(x_i, a_{\mathrm{x}1,i}, \hat{a}_{\mathrm{x}2,j}\right) g(x_j)\mathrm{d}x_j}{\int_{\mathbb{R}} g(x_j)\mathrm{d}x_j} \tag{A59}$$

We can simplify the above as follows:

$$\begin{aligned} EU_{\mathrm{x}1}\left(x_i, a_{\mathrm{x}1,i}\right) = \quad &- \quad r(x_i - a_{\mathrm{x}1,i})^2 + 2(1-r)\left(x_i - \frac{2r}{1+3r} a_{\mathrm{x}1,i}\right)\left(\frac{1+r}{1+3r}\right)\int_{\mathbb{R}} x_j g(x_j)\mathrm{d}x_j \\ &- \quad (1-r)\left[\left(x_i - \frac{2r}{1+3r} a_{\mathrm{x}1,i}\right)^2 + \left(\frac{1+r}{1+3r}\right)^2 \int_{\mathbb{R}} x_j^2 g(x_j)\mathrm{d}x_j\right] \end{aligned} \tag{A60}$$

Computing the integrals in Equation (A60), we have:

$$EU_{\mathrm{x}1}\left(x_i, a_{\mathrm{x}1,i}\right) = -r\left(x_i - a_{\mathrm{x}1,i}\right)^2 - (1-r)\left[\left(x_i - \frac{2r}{1+3r}a_{\mathrm{x}1,i}\right)^2 + \left(\frac{1+r}{1+3r}\right)^2 \int_{\mathbb{R}} x_j^2 g(x_j)\mathrm{d}x_j\right] \tag{A61}$$

Solving the first-order condition for agent i's action gives us the equilibrium choice $\hat{a}_{\mathrm{x}1,i}$:

$$\hat{a}_{\mathrm{x}1,i} = \frac{(1+3r)(3+r)}{(1+3r)^2 + 4r(1-r)}x_i \tag{A62}$$

Thus, given x_i and x_j, there exists a unique exogenous sequential choice equilibrium, where $\hat{a}_{\mathrm{x}1,i} = \frac{(1+3r)(3+r)}{(1+3r)^2+4r(1-r)}x_i$ and $\hat{a}_{\mathrm{x}2,j} = \frac{2(1+r)x_j-(1-r)a_{\mathrm{x}1,i}}{1+3r}$.

C Proof of Proposition 2

Recall that the mean of the equilibrium outcome for the exogenous sequential game can be expressed as:

$$\bar{a}_{\mathrm{x}} = \frac{\hat{a}_{\mathrm{x}1,i} + \hat{a}_{\mathrm{x}2,j}}{2} = \frac{2ra_{\mathrm{x}1,i} + (1+r)x_j}{1+3r} = \frac{2r(3+r)}{(1+3r)^2 + 4r(1-r)}x_i + \frac{1+r}{1+3r}x_j \tag{A63}$$

Next we show that $0 < k_1(r), k_2(r), k_3(r) < 1$.

- $k_1(r) < 1$ if $\frac{(1+3r)(1-r)}{(1+3r)^2+4r(1-r)} < 1 \Rightarrow (1+3r)(1-r) < (1+3r)^2 + 4r(1-r) \Rightarrow -4r(1+3r) < 4r(1-r)$, which is always true if $0 < r < 1$. Hence $k_1(r) < 1$.
- $k_2(r) < 1$ if $\frac{(1+3r)\left[(1+3r)^2+16r\right]}{[(1+3r)^2+4r(1-r)][2(1+r)+(1+3r)]} < 1$, $\Rightarrow 4r(1+3r)(3+r) < 2(1+3r)^2(1+r) + 4r(1-r)(1+r) \Rightarrow -(1+3r)(1-r)^2 < 4r(1-r)(1+r)$, which is always true since $0 < r < 1$.
- $k_3(r) < 1$ if $\frac{(1+3r)^3+8r(1-r)(1+r)}{2(1+r)[(1+3r)^2+4r(1-r)]} < 1 \Rightarrow (1+3r)^3 < 2(1+r)(1+3r)^2$, $\Rightarrow r < 1$, which we know is always true.

Further, $k_1(r), k_2(r), k_3(r) > 0$ since all the terms in their numerators and denominators are positive. Thus $0 < k_1(r), k_2(r), k_3(r) < 1$. Next, we present the proofs of Proposition 3a and 3b.

a) First, we compare $\bar{a}_{\mathrm{x}}$ with $\bar{x}$. Without loss of generality, let $x_i \geq 0$.

- Polarization – $|\bar{a}_{\mathrm{x}}| > |\bar{x}|$ and $\bar{a}_{\mathrm{x}}\bar{x} > 0$.
  - First, consider the case where $x_j > k_1(r)x_i$. In this case, $\bar{a}_{\mathrm{x}}\bar{x} > 0$ because both $\bar{a}_{\mathrm{x}} > 0$ and

$\bar{x} > 0$. The condition $|\bar{a}_{\mathrm{x}}| > |\bar{x}|$ therefore simplifies to $\bar{a}_{\mathrm{x}} > \bar{x}$. This can be expressed as:

$$\begin{aligned} &\frac{2r(3+r)}{(1+3r)^2+4r(1-r)}x_i + \frac{1+r}{1+3r}x_j > \frac{1+r}{1+3r}x_i + \frac{1+r}{1+3r}x_j \\ \Rightarrow \quad & x_j > \frac{(1+3r)(1-r)}{(1+3r)^2+4r(1-r)}x_i \end{aligned} \tag{A64}$$

where the multiplier of x_i is $k_1(r) = \frac{(1+3r)(1-r)}{(1+3r)^2+4r(1-r)}$. By definition,(A64) is satisfied.

- Second, consider the case where $-x_i < x_j \leq k_1(r)x_i$. Here, $\bar{x} > 0$. So the condition $|\bar{a}_{\mathrm{x}}| > |\bar{x}|$ simplifies to $\bar{a}_{\mathrm{x}} < \bar{x}$. From the previous case, we know that $\bar{a}_{\mathrm{x}} > \bar{x}$ if and only if $x_j > k_1(r)x_i$. This is not possible here since by definition $x_j \leq k_1(r)x_i$.
- Third, consider the case where $x_j = -x_i$. Here $\bar{x} = 0$. So, we cannot have $\bar{a}_{\mathrm{x}}\bar{x} > 0$.
- Fourth, consider the case where $x_j < -x_i$. Here, $\bar{x} < 0$. Since we require $\bar{a}_{\mathrm{x}}\bar{x}$ to be greater than zero, it follows that $\bar{a}_{\mathrm{x}} < 0$. So the condition $|\bar{a}_{\mathrm{x}}| > |\bar{x}|$ simplifies to $\bar{a}_{\mathrm{x}} < \bar{x} \Rightarrow x_j < k_1(r)x_i$, which is always true since by definition $x_j < -x_i$.

Hence, polarization occurs in the first and third cases, $i.e.$, when $x_j > k_1(r)x_i$ or when $x_j < -x_i$.

- Reverse Polarization – $|\bar{a}_{\mathrm{x}}| > |\bar{x}|$ and $\bar{a}_{\mathrm{x}}\bar{x} \leq 0$.
  - First, consider the case where $x_j > -k_2(r)x_i$. Here both $\bar{a}_{\mathrm{x}} > 0$ and $\bar{x} > 0$ since $k_2(r) < 1$. So it cannot be that $\bar{a}_{\mathrm{x}}\bar{x} \leq 0$. Therefore, this case is ruled out.
  - Second, consider the case where $-x_i \leq x_j < -k_2(r)x_i$. Here $\bar{x} \geq 0$. So the condition $|\bar{a}_{\mathrm{x}}| \geq |\bar{x}|$ simplifies to $\bar{a}_{\mathrm{x}} \leq -\bar{x}$. This can be expressed as:

$$\begin{aligned} &\frac{2r(3+r)}{(1+3r)^2+4r(1-r)}x_i + \frac{1+r}{1+3r}x_j < -\frac{1+r}{1+3r}x_i - \frac{1+r}{1+3r}x_j \\ \Rightarrow \quad & x_j < \frac{(1+3r)\left[(1+3r)^2+16r\right]}{\left[(1+3r)^2+4r(1-r)\right]\left[2(1+r)+(1+3r)\right]}x_i \end{aligned} \tag{A65}$$

    where the multiplier of x_i is labeled $k_2(r) = \frac{(1+3r)\left[(1+3r)^2+16r\right]}{\left[(1+3r)^2+4r(1-r)\right]\left[2(1+r)+(1+3r)\right]}$. Since by definition $x_j < -k_2(r)x_i$, condition (A65) is always satisfied.
  - Third, consider the case where $x_j < -x_i$. Here $\bar{x} < 0$. So the condition $|\bar{a}_{\mathrm{x}}| \geq |\bar{x}|$ simplifies to $\bar{a}_{\mathrm{x}} \geq \bar{x}$, which we know is the same as $x_j \geq k_1(r)x_i > 0$. However, this is not possible since by definition $x_j < -x_i$.

Hence, Reverse Polarization only occurs when $-x_i \leq x_j < -k_2(r)x_i$.

- Moderation – $|\bar{a}_{\mathrm{x}}| \leq |\bar{x}|$.
  - First, consider the case where $x_j < -x_i$. Here, $\bar{x} \leq 0$. So the condition $|\bar{a}_{\mathrm{x}}| \leq |\bar{x}|$ simplifies to $-\bar{x} \leq \bar{a}_{\mathrm{x}} \leq \bar{x}$. If $-\bar{x} \leq \bar{a}_{\mathrm{x}}$, it then follows that $x_j \geq k_2(r)x_i \geq 0$, which is impossible since by definition $x_j < -x_i$.
  - Second, consider the case where $-x_i \leq x_j < -k_2(r)x_i$. Here, $\bar{x} \geq 0$. So the condition $|\bar{a}_{\mathrm{x}}| \leq |\bar{x}|$ simplifies to $-\bar{x} \leq \bar{a}_{\mathrm{x}} \leq \bar{x}$. We know that this condition can be expressed as $-k_2(r)x_i \leq x_j \leq k_1(r)x_i$. This is not possible, since by definition $x_j < -k_2(r)x_i$.
  - Third, consider the case where $-k_2(r)x_i \leq x_j \leq k_1(r)x_i$. Here, $\bar{x} > 0$. So the condition $|\bar{a}_{\mathrm{x}}| \leq |\bar{x}|$ simplifies to $\bar{x} \leq \bar{a}_{\mathrm{x}} \leq -\bar{x}$. This in turn can be expressed as $-k_2(r)x_i \leq x_j \leq k_1(r)x_i$, which we know is true by definition.
  - Fourth, consider the case where $x_j > k_1(r)x_i$. Here, $\bar{x} > 0$. So the condition $|\bar{a}_{\mathrm{x}}| \leq |\bar{x}|$ simplifies to $\bar{x} \leq \bar{a}_{\mathrm{x}} \leq -\bar{x}$. This in turn can be expressed as $-k_2(r)x_i \leq x_j \leq k_1(r)x_i$, which cannot be true, since by definition $x_j > k_1(r)x_i$.

  Hence, Moderation occurs only when $-k_2(r)x_i \leq x_j \leq k_1(r)x_i$.

Proof of Proposition 3

Required to show that, for $x_j > 0$, $\hat{a}_{\mathrm{x}2,j} \geq \hat{a}_j$ if $x_i \leq 0$, and $\hat{a}_{\mathrm{x}2,j} < \hat{a}_j$ if $x_i > 0$.

i) $x_i \leq 0$

$\hat{a}_{\mathrm{x}2,j} - \hat{a}_j$ can be simplified to $\hat{a}_{\mathrm{x}2,j} - \hat{a}_j = -\frac{(1-r)}{1+3r}\hat{a}_{\mathrm{x}1,i}$. We know that $\hat{a}_{\mathrm{x}1,i} = \mu_x(r)x_i \leq 0$ because $\mu_x(r) > 0$, $x_i \leq 0$. It therefore follows that $-\frac{(1-r)}{1+3r}\hat{a}_{\mathrm{x}1,i} \geq 0 \Rightarrow \hat{a}_{\mathrm{x}2,j} \geq \hat{a}_j$.

ii) $x_i > 0$

As before $\hat{a}_{\mathrm{x}2,j} - \hat{a}_j = -\frac{(1-r)}{1+3r}\hat{a}_{\mathrm{x}1,i}$. However, here $\hat{a}_{\mathrm{x}1,i} = \mu_x(r)x_i > 0$ because $\mu_x(r), x_i > 0$. Hence it follows that $\hat{a}_{\mathrm{x}2,j} - \hat{a}_j < 0 \Rightarrow \hat{a}_{\mathrm{x}2,j} < \hat{a}_j$.

b) The derivative of $\hat{a}_{\mathrm{x}1,i}$ w.r.t r can be calculated and simplified to $\frac{\mathrm{d}\hat{a}_{\mathrm{x}1,i}}{\mathrm{d}r} = -\frac{4(5r^2+6r+5)}{[(1+3r)^2+4r(1-r)]^2}x_i$.

Since $r > 0$, it follows that $\frac{\mathrm{d}\hat{a}_{\mathrm{x}1,i}}{\mathrm{d}r} < 0$

c) To show that $|\hat{a}_{\mathrm{x}1,i}| \geq |\hat{a}_i|$, we need to show that $\mu_{\mathrm{x}}(r) \geq \mu(r)$.

$$\begin{aligned}\mu_{\mathrm{x}}(r) - \mu(r) &= \frac{(1+3r)(3+r)}{(1+3r)^2+4r(1-r)} - \frac{2(1+r)}{1+3r} \\ &= \frac{(1-r)^3}{(1+3r)[(1+3r)^2+4r(1-r)]} > 0 \text{ if } r > 0 \end{aligned} \tag{A66}$$

Therefore, $\mu_{\mathrm{x}}(r) \geq \mu(r) \Rightarrow |\hat{a}_{\mathrm{x}1,i}| \geq |\hat{a}_i|$. □

D Proof of Proposition 4

We now compare $\bar{a}_{\mathrm{x}}$ with $\bar{a}$. Without loss of generality, let $x_i \geq 0$.

- $|\bar{a}_{\mathrm{x}}| > |\bar{a}|$ and $\bar{a}_{\mathrm{x}}\bar{x} > 0$.
  - First, consider the case where $x_j \geq -x_i$. Here, $\bar{x} \geq 0$. So the condition $|\bar{a}_{\mathrm{x}}| > |\bar{a}|$ simplifies to $\bar{a}_{\mathrm{x}} > \bar{a} \Rightarrow \frac{2r(3+r)}{(1+3r)^2+4r(1-r)}x_i > \frac{1+r}{1+3r}x_i$, which is impossible since $0 < r < 1$ and $x_i \geq 0$.
  - Second, consider the case where $-x_j < x_i$. Here $\bar{x} < 0$. So the condition $|\bar{a}_{\mathrm{x}}| > |\bar{a}|$ simplifies to $\bar{a}_{\mathrm{x}} < \bar{a} \Rightarrow$, $\frac{2r(3+r)}{(1+3r)^2+4r(1-r)}x_i < \frac{1+r}{1+3r}x_i$, which is always true for $0 < r < 1$, $x_i \geq 0$.

  Hence, for $-x_j < x_i$, the mean outcome in the exogenous sequential game is more polarized than that in the simultaneous game, and this polarization is in the same direction as $\bar{x}$.

- $|\bar{a}_{\mathrm{x}}| > |\bar{a}|$ and $\bar{a}_{\mathrm{x}}\bar{x} \leq 0$.
  - First, consider the case where $x_j < -x_i$. Then $\bar{a}, \bar{x} < 0$. So for $\bar{a}_{\mathrm{x}}\bar{x} \leq 0$ to be true, we require $\bar{a}_{\mathrm{x}} \leq 0$, which is not possible since $\bar{a}_{\mathrm{x}} < \bar{x} < 0$.
  - Second, consider the case where $-x_i \leq x_j < -k_3(r)x_i$. Here, $\bar{x} \geq 0$ and the condition $|\bar{a}_{\mathrm{x}}| > |\bar{a}|$ simplifies to $\bar{a}_{\mathrm{x}} > -\bar{a}$. This can be expressed as:

$$\frac{2r(3+r)}{(1+3r)^2+4r(1-r)}x_i < -\frac{1+r}{1+3r}x_i$$
$$\Rightarrow \quad x_j < \frac{(1+3r)^3+8r(1-r)(1+r)}{2(1+r)\left[(1+3r)^2+4r(1-r)\right]}x_i \tag{A67}$$

    where the multiplier of x_i is labeled $k_3(r) = \frac{(1+3r)^3+8r(1-r)(1+r)}{2(1+r)[(1+3r)^2+4r(1-r)]}x_i$. Since by definition, $x_j < k_3(r)x_i$, condition (A67) is always satisfied.
  - Third, consider the case where $x_j \geq k_3(r)x_i$. Then, $\bar{x} \geq 0$ and the $|\bar{a}_{\mathrm{x}}| > |\bar{a}|$ simplifies to $\bar{a}_{\mathrm{x}} < -\bar{a}$. However, from the second case, we know that this condition can only be satisfied when $x_j < k_3(r)x_i$, which cannot hold here, since by definition $x_j \geq k_3(r)x_i$.

  Hence, for $-x_i \leq x_j < -k_3(r)x_i$, the mean outcome in the exogenous sequential game is more extreme than that in the simultaneous game, but in the direction opposite to that indicated by the mean preference $\bar{x}$.

- $|\bar{a}_{\mathrm{x}}| \leq |\bar{a}|$.
  - First, consider the case where $x_j \geq -k_3(r)x_i$. Here, $\bar{a}_{\mathrm{x}}, \bar{x}, \bar{a} \geq 0$. So the condition $|\bar{a}_{\mathrm{x}}| \leq$

$|\bar{a}|$ simplifies to $\bar{a}_{\mathrm{x}} \leq \bar{a} \Rightarrow$, $\frac{2r(3+r)}{(1+3r)^2+4r(1-r)} x_i \leq \frac{1+r}{1+3r} x_i$, which is always true for $0 < r < 1$, $x_i \geq 0$.

- Second, consider the case where $-x_i \leq x_j < k_3(r)x_i$. Here also $\bar{x}, \bar{a} \geq 0$. So the condition $|\bar{a}_{\mathrm{x}}| \leq |\bar{a}|$ simplifies to $-\bar{a} \leq \bar{a}_{\mathrm{x}} \leq \bar{a}$. From the case before, we know that $\bar{a}_{\mathrm{x}} \leq \bar{a}$ for $x_i \geq 0$. However, to ensure that $-\bar{a} \leq \bar{a}_{\mathrm{x}}$, we need $x_j \geq k_3(r)x_i$, which cannot be true since by definition $-x_i \leq x_j < k_3(r)x_i$.
- Third, consider the case where $x_j < -x_i$. Here, $\bar{x}, \bar{a} < 0$. So the condition $|\bar{a}_{\mathrm{x}}| \leq |\bar{a}|$ simplifies to $\bar{a} \leq \bar{a}_{\mathrm{x}} \leq -\bar{a}$. The condition $\bar{a}_{\mathrm{x}} \geq \bar{a}$ reduces to $x_j \geq -k_3(r)x_i$, which we know is not possible since $x_j < -x_i$ and $0 < k_3(r) < 1$.

Hence, for $x_j \geq -k_3(r)x_i$, the mean outcome in the exogenous sequential game is less extreme (moderate) compared to that in the simultaneous game.

E Proof of Proposition 5

The expected utility of the first player i in equilibrium is given by (11). Substituting for $\hat{a}_{\mathrm{x}1,i}$ gives us:

$$\begin{aligned} EU_{\mathrm{x}1}\left(x_i, \hat{a}_{\mathrm{x}1,i}\right) = & - r\left(x_i - \frac{(1+3r)(3+r)}{(1+3r)^2+4r(1-r)} x_i\right)^2 \\ & - (1-r)\left[\left(x_i - \frac{2r(3+r)}{(1+3r)^2+4r(1-r)} x_i\right)^2 + \frac{(1+r)^2}{3(1+3r)^2}\int_{\mathbb{R}} x_j^2 g(x_j)\mathrm{d}x_j\right] \end{aligned} \tag{A68}$$

This in turn simplifies to:

$$EU_{\mathrm{x}1}\left(x_i, \hat{a}_{\mathrm{x}1,i}\right) = -\frac{(1-r)(1+r)^2}{(1+3r)^2+4r(1-r)} x_i^2 - \frac{(1-r)(1+r)^2}{(1+3r)^2}\int_{\mathbb{R}} x_j^2 g(x_j)\mathrm{d}x_j \tag{A69}$$

Next, consider the a priori expected utility of player j, in equilibrium. It is obtained by integrating the utility of the second player over the range of x_i. That is, $EU_{\mathrm{x}2}\left(x_j, \hat{a}_{\mathrm{x}2,j}\right) = \int_{\mathbb{R}} u\left(x_j, \hat{a}_{\mathrm{x}2,j}, \hat{a}_{\mathrm{x}1,i}\right) g(x_i)\mathrm{d}x_i$. Substituting for $\hat{a}_{\mathrm{x}2,j}$ as $\frac{2(1+r)}{1+3r} x_j - \frac{(1-r)}{1+3r} a_{\mathrm{x}1,i}$, we have:

$$\begin{aligned} EU_{\mathrm{x}2}\left(x_j, \hat{a}_{\mathrm{x}2,j}\right) = & - r\int_{\mathbb{R}}\left(x_j - \frac{2(1+r)}{1+3r} x_j - \frac{1-r}{1+3r}\hat{a}_{\mathrm{x}1,i}\right)^2 g(x_i)\mathrm{d}x_i \\ & - (1-r)\int_{\mathbb{R}}\left(x_j - \frac{1+r}{1+3r} x_j - \frac{2r}{1+3r}\hat{a}_{\mathrm{x}1,i}\right)^2 g(x_i)\mathrm{d}x_i \end{aligned} \tag{A70}$$

Substituting for $\hat{a}_{\mathrm{x}1,i}$ and integrating, we have:

$$EU_{\mathrm{x}2}\left(x_j, \hat{a}_{\mathrm{x}2,j}\right) = -\frac{r(1-r)}{1+3r} x_i^2 - \frac{r(1-r)(1+3r)(3+r)^2}{\left[(1+3r)^2+4r(1-r)\right]^2}\int_{\mathbb{R}} x^2 g(x)\mathrm{d}x \tag{A71}$$

Next, we compare the difference in the expected utilities for a specific player i, where $D_{\mathrm{x}}(x_i) = EU_{\mathrm{x}1}(x_i, \hat{a}_{\mathrm{x}1,i}) - EU_{\mathrm{x}2}(x_i, \hat{a}_{\mathrm{x}2,i})$. We can show that $D_{\mathrm{x}}(x_i) \leq 0$ for all i if the following two conditions are satisfied:

$$\frac{r(1-r)(1+3r)(3+r)^2}{[(1+3r)^2+4r(1-r)]^2} > -\frac{(1-r)(1+r)^2}{(1+3r)^2} \tag{A72}$$

and

$$-\frac{r(1-r)}{1+3r} \geq -\frac{(1-r)(1+r)^2}{(1+3r)^2+4r(1-r)} \tag{A73}$$

First, consider the inequality (A72), which can be simplified to:

$$(1+r)^2\left[(1+3r)^2+4r(1-r)\right]^2 > r(3+r)^2(1+3r)^3 \tag{A74}$$

This in turn simplifies to:

$$\begin{aligned} & (1+3r)^3\left[(1+r)^2(1+3r)-r(3+r)^2\right]+(1+r)^2\left[16r^2(1-r)^2+8r(1-r)(1+3r)^2\right] > 0 \\ \Rightarrow\quad & 16r^2(1+r)^2(1-r)^2+(1+3r)^2(1-r)\left[2r(1+r)^2+(1+3r)(1-r)\right] > 0 \end{aligned} \tag{A75}$$

Since both the terms in the R.H.S of inequality (A75) are non-negative, the inequality is always true. Therefore, (A72) is always true. Next, consider the inequality (A73), which can be expressed as:

$$\frac{(1-r)(1+r)^2}{(1+3r)^2+4r(1-r)}x_i^2 \geq \frac{r(1-r)}{1+3r}x_i^2 \tag{A76}$$

This is true if:

$$\begin{aligned} & (1-r)\left[(1+r)^2(1+3r)-r[(1+3r)^2+4r(1-r)]\right]x_i^2 \geq 0 \\ \Rightarrow\quad & (1-r)^2\left[2r^2+5r+1\right]x_i^2 \geq 0 \end{aligned} \tag{A77}$$

We know that $x_i^2 \geq 0$ and that both $(1-r)^2$ and $2r^2+5r+1$ are positive for $0 < r < 1$. Hence this inequality is always true too. Further, since both (A72) and (A73) are always true, it follows that $EU_{\mathrm{x}2}(x_i, \hat{a}_{\mathrm{x}2,i}) > EU_{\mathrm{x}1}(x_i, \hat{a}_{\mathrm{x}1,i})$.

b) Now we prove the second part of the Proposition. Let $x_i > 0$, then:

$$\frac{\mathrm{d}D_{\mathrm{x}}(x_i)}{\mathrm{d}x_i} = 2x_i(1-r)\left[\frac{r}{1+3r}-\frac{(1+r)^2}{(1+3r)^2+4r(1-r)}\right] \tag{A78}$$

Since we have already shown that (A73) is true, we know that $(1-r)\left[\frac{r}{1+3r}-\frac{(1+r)^2}{(1+3r)^2+4r(1-r)}\right] > 0$. It therefore follows that $2x_i(1-r)\left[\frac{r}{1+3r}-\frac{(1+r)^2}{(1+3r)^2+4r(1-r)}\right] \leq 0$. Hence, $\frac{\mathrm{d}D_{\mathrm{x}}(x_i)}{\mathrm{d}x_i} \leq 0$. Similar proof applies for $x_i < 0$. □

F Proof of Proposition 6

The solutions for the optimal actions for periods 3 and 4 are analogous to that in the exogenous sequential choice game and are outlined in the main text. Below, we derive the players' optimal action for the first two periods.

Period 2 – A player who has lost the auction makes no decisions in period 2. So we only consider the actions of a player who won the auction in period 1. Suppose player j bids according to the symmetric bidding function $\beta(\cdot)$, then player i belief upon winning is that player j must belong to a some symmetric region W for her to have won the auction. In that case, i's expected utility from speaking first is:

$$\begin{aligned} EU_{\text{n1}}\left(x_i, a_{\text{n1},i}\right) = & - \quad r(x_i - a_{\text{n1},i})^2 - (1-r)\left[\left(x_i - \frac{2r}{1+3r}a_{\text{n1},i}\right)^2 + \left(\frac{1+r}{1+3r}\right)^2 \frac{\int_W x_j^2 g(x_j)\mathrm{d}x_j}{\int_W g(x_j)\mathrm{d}x_j}\right] \\ & - \quad 2\frac{1+r}{1+3r}\left(x_i - \frac{2r}{1+3r}a_{\text{n1},i}\right)\frac{\int_W x_j g(x_j)\mathrm{d}x_j}{\int_W g(x_j)\mathrm{d}x_j} \end{aligned} \tag{A79}$$

The last term vanishes because W is symmetric around zero. So:

$$\begin{aligned} EU_{\text{n1}}\left(x_i, a_{\text{n1},i}\right) &= -r(x_i - a_{\text{n1},i})^2 - (1-r)\left[\left(x_i - \frac{2r}{1+3r}a_{\text{n1},i}\right)^2 + \left(\frac{1+r}{1+3r}\right)^2 \frac{\int_W x_j^2 g(x_j)\mathrm{d}x_j}{\int_W g(x_j)\mathrm{d}x_j}\right] \\ &= -\frac{(1-r)(1+r)^2}{(1+3r)^2 + 4r(1-r)}x_i^2 - \frac{(1-r)(1+r)^2}{(1+3r)^2}\frac{\int_W x_j^2 g(x_j)\mathrm{d}x_j}{\int_W g(x_j)\mathrm{d}x_j} \end{aligned} \tag{A80}$$

Similarly, simplifying i's expected utility from speaking second gives us:

$$EU_{\text{n2}}\left(x_i, a_{\text{n2},i}\right) = -\frac{r(1-r)}{(1+3r)}\left[x_i^2 + \left(\frac{(1+3r)(3+r)}{(1+3r)^2 + 4r(1-r)}\right)^2 \int_W g(x_j)x_j^2\mathrm{d}x_j\right] \tag{A81}$$

i prefers to speak second if:

$$EU_{\text{n2}}\left(x_i, a_{\text{n2},i}\right) > EU_{\text{n1}}\left(x_i, a_{\text{n1},i}\right) \tag{A82}$$

Let $A(r) = \frac{(1-r)(1+r)^2}{(1+3r)^2+4r(1-r)}$, $B(r) = \frac{(1-r)(1+r)^2}{(1+3r)^2}$, $C(r) = \frac{r(1-r)}{(1+3r)}$, and $D(r) = \frac{r(1-r)}{(1+3r)}\left(\frac{(1+3r)(3+r)}{(1+3r)^2+4r(1-r)}\right)^2$. It is trivial to show that the multiplier of x_i^2 in $EU_{\text{n2}}\left(x_i, a_{\text{n2},i}\right)$ is always greater than that in $EU_{\text{n1}}\left(x_i, a_{\text{n1},i}\right) \Rightarrow A(r) > C(r)$. Similarly, we can show that the multiplier of $\int_W x_j^2 g(x_j)\mathrm{d}x_j$ in $EU_{\text{n2}}\left(x_i, a_{\text{n2},i}\right)$ is also greater than that in $EU_{\text{n1}}\left(x_i, a_{\text{n1},i}\right) \Rightarrow B(r) > D(r)$. Therefore, in a symmetric bidding equilibrium, for all x_is, the expected value from speaking second is higher than that from speaking first. Therefore, upon winning the auction, all types will choose to speak second.

Period 1 At the beginning of period 1, before placing her bid, player i knows that if she wins, she will choose to go second and if player j wins, she will go first. Hence, her expected utility from choosing a bid b_i and then choosing the optimal action in the subsequent periods is

$$EU(x_i, b_i) = \frac{\int_{L_i} u(x_i, \hat{a}_{\mathrm{n1},i}, \hat{a}_{\mathrm{n2},j}) g(x_j)\mathrm{d}x_j + \int_{W_i} \left[u(x_i, \hat{a}_{\mathrm{n1},j}, \hat{a}_{\mathrm{n2},i}) - b_i\right] g(x_j)\mathrm{d}x_j}{\int_{\mathbb{R}} g(x_j)\mathrm{d}x_j} \tag{A83}$$

where $j \in W_i$ for i to win the auction and $j \in L_i$ for her to lose the auction, if she chooses a bid b_i. This simplifies to:

$$\begin{aligned} 2EU(x_i, b_i) = & - A(r)x_i^2 \int_{L_i} g(x_j)\mathrm{d}x_j - B(r) \int_{L_i} x_j^2 g(x_j)\mathrm{d}x_j \\ & - C(r)x_i^2 \int_{W_i} g(x_j)\mathrm{d}x_j - D(r) \int_{W_i} x_j^2 g(x_j)\mathrm{d}x_j - \int_{W_i} b_i g(x_j)\mathrm{d}x_j \\ = & - 2x_i^2 + [A(r) - C(r)]x_i^2 \int_{W_i} g(x_j)\mathrm{d}x_j + [B(r) - D(r)] \int_{W_i} x_j^2 g(x_j)\mathrm{d}x_j - \int_{W_i} b_i g(x_j)\mathrm{d}x_j \end{aligned} \tag{A84}$$

Now consider any two types x' and x'' and a bidding function $\beta(\cdot)$. In equilibrium, x' can do no better by playing x'''s strategy $\beta(x'')$ over her own strategy $\beta(x')$ and vice-versa. That is:

$$EU(x', \beta(x')) \geq EU(x', \beta(x'')) \tag{A85}$$

$$EU(x'', \beta(x'')) \geq EU(x', \beta(x'')) \tag{A86}$$

Substituting the simplified expressions for the expected utilities into the above inequalities and adding them up gives us:

$$[A(r) - C(r)]\,(x'^2 - x''^2) \left(\int_{W_1} g(x_j)\mathrm{d}x_j - \int_{W_2} g(x_j)\mathrm{d}x_j \right) \geq 0 \tag{A87}$$

We know that $A(r) - C(r) > 0$. So if $x'^2 > x''^2$, then for the above inequality to hold, we require that $\int_{W_1} g(x_j)\mathrm{d}x_j - \int_{W_2} g(x_j)\mathrm{d}x_j \geq 0 \Rightarrow$ the region over which a player wins upon bidding $\beta(x')$ is greater than that over which she wins when she bids $\beta(x'')$. In other words, the equilibrium bidding strategies are monotonically increasing in $|x|$. Further, following the technique as that outlined in Fudenberg and Tirole (1991) p. 217, we can show strict monotonicity, *i.e.*, if $|x'| > |x''|$, then $\beta(x') > \beta(x'')$.

Now that we have shown that the bidding strategies are monotonically increasing in $|x|$, for the specific bid b_i by player i (when the other player uses the bidding function $\beta(\cdot)$), we re-write Equation (A84) as:

$$EU(x_i, b_i) = -2x_i^2 + 2[A(r) - C(r)]x_i^2 \int_0^{\beta^{-1}(b_i)} g(x)\mathrm{d}x + 2[B(r) - D(r)] \int_0^{\beta^{-1}(b_i)} x^2 g(x)\mathrm{d}x - 2\int_0^{\beta^{-1}(b_i)} b_i g(x)\mathrm{d}x \tag{A88}$$

Further, we specify the following expression for the derivatives:

$$\frac{\mathrm{d}\left[\int_0^{\beta^{-1}(b_i)} F(x)\right]}{\mathrm{d}b_i}\mathrm{d}x = F(V)\frac{\mathrm{d}V}{\mathrm{d}b_i} \tag{A89}$$

where $V = \beta^{-1}(b_i)$. To obtain the equilibrium bidding function, we can calculate the F.O.C of Equation (A88) as $\left.\frac{\mathrm{d}EU(x_i,b_i)}{\mathrm{d}b_i}\right|_{b_i=\hat{b}_i} = 0$. This simplifies to:

$$[A(r) - C(r)]\, x_i^2 \frac{g(\beta^{-1}(\hat{b}_i))}{\beta'(\beta^{-1}(\hat{b}_i))} + [B(r) - D(r)] \frac{g(\beta^{-1}(\hat{b}_i)) \left[\beta^{-1}(\hat{b}_i)\right]^2}{\beta'(\beta^{-1}(\hat{b}_i))} - \left[\hat{b}_i \frac{g(\beta^{-1}(\hat{b}_i))}{\beta'(\beta^{-1}(\hat{b}_i))} + \int_0^{\beta^{-1}(\hat{b}_i)} g(x)\mathrm{d}x\right] = 0$$

In equilibrium $\hat{b}_i = \beta(x_i)$ and so $\beta^{-1}(\hat{b}_i) = x_i$. So the above equation simplifies to:

$$[A(r) - C(r) + B(r) - D(r)]\, x_i^2 g(\beta^{-1}(\hat{b}_i)) = \frac{\mathrm{d}\left[\beta(x_i)\int_0^{x_i} g(x)\mathrm{d}x\right]}{\mathrm{d}x_i} \tag{A90}$$

Integrating this from 0 to x_i, we have:

$$\begin{aligned} \hat{\beta} = \beta(x_i) &= [A(r) - C(r) + B(r) - D(r)] \left[\frac{\int_0^{x_i} x^2 g(x)\mathrm{d}x}{\int_0^{x_i} g(x)\mathrm{d}x}\right] \\ &= f(r)\left[\frac{\int_0^{x_i} x^2 g(x)\mathrm{d}x}{\int_0^{x_i} g(x)\mathrm{d}x}\right], \end{aligned} \tag{A91}$$

where $f(r) = A(r) - C(r) + B(r) - D(r)$ is the multiplier of $\left[\frac{\int_0^{x_i} x^2 g(x)\mathrm{d}x}{\int_0^{x_i} g(x)\mathrm{d}x}\right]$. Since $A(r) - C(r) > 0$ and $B(r) - D(r) > 0$, $f(r)$ is positive, which recovers the assumption that the bidding function is symmetric around zero. $\square$

G Social Welfare Derivations and Comparisons

G.1 Welfare Under First-Best Planner's Choice

The social planner's choice involve agents choosing their true preferences as their actions, $i.e.$, $a_i = x_i$ and $a_j = x_j$, and $\bar{a} = \frac{x_i + x_j}{2}$. Then:

$$\begin{aligned}
W_{FB}(x_i, x_j, a_i, a_j) &= W_{FB}(x_i, x_j) \\
&= -r(x_i - a_i)^2 - (1-r)(x_i - \bar{a})^2 - r(x_j - a_j)^2 - (1-r)(x_j - \bar{a})^2 \\
&= -(1-r)\frac{(x_i - x_j)^2}{2} \qquad \text{(A92)}
\end{aligned}$$

Then the expected welfare is given by:

$$\begin{aligned}
EW_P &= \frac{\int_{\mathbb{R}} \int_{\mathbb{R}} W_1(x_i, x_j) g(x_i) g(x_j) \mathrm{d}x_i \mathrm{d}x_j}{\int_{\mathbb{R}} \int_{\mathbb{R}} g(x_i) g(x_j) \mathrm{d}x_i \mathrm{d}x_j} \\
&= -\frac{(1-r)}{2 \int_{\mathbb{R}} \int_{\mathbb{R}} g(x_i) g(x_j) \mathrm{d}x_i \mathrm{d}x_j} \left[\int_{\mathbb{R}} \int_{\mathbb{R}} (x_i^2 + x_j^2 - 2x_i x_j) g(x_i) g(x_j) \mathrm{d}x_i \mathrm{d}x_j \right]
\end{aligned}$$

We know that $\int_{\mathbb{R}} \int_{\mathbb{R}} g(x_i) g(x_j) \mathrm{d}x_i \mathrm{d}x_j = 1$ and $\int_{\mathbb{R}} \int_{\mathbb{R}} x_i x_j g(x_i) g(x_j) \mathrm{d}x_i \mathrm{d}x_j = 0$ because $g(\cdot)$ is a symmetric distribution. Therefore, EW_1 simplifes to:

$$\begin{aligned}
EW_{FB} &= \frac{-(1-r)}{2} \int_{\mathbb{R}} \int_{\mathbb{R}} \left[x_i^2 + x_j^2 \right] g(x_i) g(x_j) \mathrm{d}x_j \mathrm{d}x_j \\
&= -(1-r) \int_{\mathbb{R}} x^2 g(x) \mathrm{d}x \qquad \text{(A93)}
\end{aligned}$$

G.2 Welfare in the Simultaneous Game

In the simultaneous game, the actions are $a_i = \frac{2(1+r)}{1+3r} x_i$, $a_j = \frac{2(1+r)}{1+3r} x_j$, and the mean action is $\bar{a} = \frac{(1+r)}{1+3r}(x_i + x_j)$. Substituting this in the welfare function we can obtain:

$$\begin{aligned}
W_s(x_i, x_j, a_i, a_j) &= W_s(x_i, x_j) \\
&= -\frac{r(1-r)^2}{(1+3r)^2}(x_i^2 + x_j^2) - \frac{1-r}{(1+3r)^2} \left[2rx_i - (1+r)x_j \right]^2 \\
&\quad -\frac{1-r}{(1+3r)^2} \left[2rx_j - (1+r)x_i \right]^2 \qquad \text{(A94)} \\
&= -\frac{r(1-r)^2}{(1+3r)^2}(x_i^2 + x_j^2) \\
&\quad -\frac{1-r}{(1+3r)^2} \left[4r^2(x_i^2 + x_j^2) + (1+r)^2(x_i^2 + x_j^2) - 8r(1+r)x_i x_j \right]
\end{aligned}$$

The expected welfare is then given by:

$$EW_s = \frac{\int_{\mathbb{R}}\int_{\mathbb{R}} W_s(x_i,x_j)g(x_i)g(x_j)\mathrm{d}x_i\mathrm{d}x_j}{\int_{\mathbb{R}}\int_{\mathbb{R}} g(x_i)g(x_j)\mathrm{d}x_i\mathrm{d}x_j}$$

As before, $\int_{\mathbb{R}}\int_{\mathbb{R}} g(x_i)g(x_j)\mathrm{d}x_i\mathrm{d}x_j = 1$ and $\int_{\mathbb{R}}\int_{\mathbb{R}} x_ix_jg(x_i)g(x_j)\mathrm{d}x_i\mathrm{d}x_j = 0$. So:

$$EW_s = -\frac{2(1-r)}{(1+3r)^2}[r(1-r)+4r^2+(1+r)^2]\int_{\mathbb{R}} x^2g(x)\mathrm{d}x \tag{A95}$$

G.3 Welfare in the Exogenous Sequential Choice Game

Without loss of generality, assume that i speaks first and j speaks second. Then, we have the actions of the two agents as $a_i = \frac{(1+3r)(3+r)}{(1+3r)^2+4r(1-r)}x_i$, $a_j = \frac{2(1+r)}{1+3r}x_j - \frac{1-r}{1+3r}a_i$, and the mean action as $\bar{a} = \frac{2r(3+r)}{(1+3r)^2+4r(1-r)}x_i + \frac{1+r}{1+3r}x_j$. Using these, we can further derive the following expressions:

$$\begin{aligned}
x_i - a_i &= -\frac{2(1-r)(1+r)}{(1+3r)^2+4r(1-r)}x_i \\
x_j - a_j &= -\frac{1-r}{1+3r}\left[x_j - \frac{(1+3r)(3+r)}{(1+3r)^2+4r(1-r)}x_i\right] \\
x_i - \bar{a} &= \frac{(1+3r)(1+r)}{(1+3r)^2+4r(1-r)}x_i - \frac{1+r}{1+3r}x_j \\
x_j - \bar{a} &= \frac{2r}{1+3r}x_j - \frac{2r(3+r)}{(1+3r)^2+4r(1-r)}x_i
\end{aligned}$$

Substituting the above terms in the welfare equation, we have:

$$\begin{aligned}
W_x(x_i,x_j,a_i,a_j) &= W_x(x_i,x_j) \\
&= -\frac{4r(1-r)^2(1+r)^2}{((1+3r)^2+4r(1-r))^2}x_i^2 - (1-r)\left[\frac{(1+3r)(1+r)}{(1+3r)^2+4r(1-r)}x_i - \frac{1+r}{1+3r}x_j\right]^2 \\
&\quad -\frac{r(1-r)^2}{(1+3r)^2}\left[x_j - \frac{(1+3r)(3+r)}{(1+3r)^2+4r(1-r)}x_i\right]^2 \\
&\quad -(1-r)\left[\frac{2r}{1+3r}x_j - \frac{2r(3+r)}{(1+3r)^2+4r(1-r)}x_i\right]
\end{aligned} \tag{A96}$$

As before, the expected welfare is given by:

$$EW_x = \frac{\int_{\mathbb{R}}\int_{\mathbb{R}} W_3(x_i,x_j)g(x_i)g(x_j)\mathrm{d}x_i\mathrm{d}x_j}{\int_{\mathbb{R}}\int_{\mathbb{R}} g(x_i)g(x_j)\mathrm{d}x_i\mathrm{d}x_j}$$

As before, this implies that the integrals of the $x_i x_j$ terms are canceled out. Further, and $\int_{\mathbb{R}}\int_{\mathbb{R}} g(x_i)g(x_j)\mathrm{d}x_i\mathrm{d}x_j = 1$.

$$\begin{aligned} EW_x = & - \frac{4r(1-r)^2(1+r)^2}{\left[(1+3r)^2+4r(1-r)\right]^2}\int_{\mathbb{R}}\int_{\mathbb{R}} x_i^2 g(x_i)g(x_j)\mathrm{d}x_i\mathrm{d}x_j \\ & - \frac{r(1-r)^2}{(1+3r)^2}\int_{\mathbb{R}}\int_{\mathbb{R}}\left[x_j^2+\frac{(1+3r)^2(3+r)^2}{\left[(1+3r)^2+4r(1-r)\right]^2}x_i^2\right]g(x_i)g(x_j)\mathrm{d}x_i\mathrm{d}x_j \\ & - (1-r)\int_{\mathbb{R}}\int_{\mathbb{R}}\left[\frac{(1+3r)^2(1+r)^2+4r^2(3+r)^2}{\left[(1+3r)^2+4r(1-r)\right]^2}x_i^2+\frac{(1+r)^2+4r^2}{(1+3r)^2}x_j^2\right]g(x_i)g(x_j)\mathrm{d}x_i\mathrm{d}x_j \end{aligned}$$

This simplifies to:

$$\begin{aligned} EW_x = & -\frac{(1-r)\left[(1+r)^2+4r^2+r(1-r)\right]}{3(1+3r)^2}\int_{\mathbb{R}} x^2 g(x)\mathrm{d}x \\ & - \left[\frac{(1-r)\left[4r(1-r)(1+r)^2+4r^2(3+r)^2+(1+3r)^2(1+r)^2+r(1-r)(3+r)^2\right]}{3\left[(1+3r)^2+4r(1-r)\right]^2}\right]\int_{\mathbb{R}} x^2 g(x)\mathrm{d}x \end{aligned} \tag{A97}$$

G.4 Welfare in the Endogenous Sequential Choice Game

As before, assume that i speaks first and j speaks second. Recall that the players' actions here are the same as that in the exogenous sequential choice game. However, we know that $|x_i| < |x_j|$. So while the welfare equation remains the same, the integrations regions are different. Thus, we have:

$$\begin{aligned} W_n(x_i,x_j,a_i,a_j) &= W_n(x_i,x_j) \\ &= -\frac{4r(1-r)^2(1+r)^2}{((1+3r)^2+4r(1-r))^2}x_i^2-(1-r)\left[\frac{(1+3r)(1+r)}{(1+3r)^2+4r(1-r)}x_i-\frac{1+r}{1+3r}x_j\right]^2 \\ &\quad -\frac{r(1-r)^2}{(1+3r)^2}\left[x_j-\frac{(1+3r)(3+r)}{(1+3r)^2+4r(1-r)}x_i\right]^2 \\ &\quad -(1-r)\left[\frac{2r}{1+3r}x_j-\frac{2r(3+r)}{(1+3r)^2+4r(1-r)}x_i\right] \end{aligned} \tag{A98}$$

and the expected welfare is:

$$EW_n = \frac{\iint_{R1\cup R2} W_4(x_i,x_j)g(x_i)g(x_j)\mathrm{d}x_i\mathrm{d}x_j}{\iint_{R1\cup R2} g(x_i)g(x_j)\mathrm{d}x_i\mathrm{d}x_j} \tag{A99}$$

where the two regions R1 and R2 are defined as follows:

$$\begin{aligned} R_1 &\equiv x_i \in [0,\infty], \quad x_j \in [x_i,\infty]\cup[-\infty,-x_i] \\ R_2 &\equiv x_i \in [-\infty,0], \quad x_j \in [-x_i,\infty]\cup[-\infty,x_i] \end{aligned}$$

As before, we can show that the integral of the $x_i x_j$ terms over $R_1 \cup R_2$ is zero. So we now consider the integrals of the x_j^2 and x_i^2 terms. Since inference on j's type is conditional on i, we first integrate over j's type. Because of the symmetry of the distrbution, it is easy to show that:

$$\iint\limits_{R1 \cup R2} x_i^2 g(x_i) g(x_j) \mathrm{d}x_i \mathrm{d}x_j = 4 \int_0^\infty x^2 g(x)(1 - G(x)) \mathrm{d}x \tag{A100}$$

$$\iint\limits_{R1 \cup R2} g(x_i) g(x_j) \mathrm{d}x_i \mathrm{d}x_j = 4 \int_0^\infty g(x)(1 - G(x)) \mathrm{d}x \tag{A101}$$

Substituting these expressions back in the expected welfare function, we have:

$$\begin{aligned} EW_n = & \left(-\frac{(1-r)\left[(1+r)^2 + 4r^2 + r(1-r)\right]}{(1+3r)^2} \right. \\ - & \left. \frac{(1-r)\left[4r(1-r)(1+r)^2 + 4r^2(3+r)^2 + (1+3r)^2(1+r)^2 + r(1-r)(3+r)^2\right]}{3\left[(1+3r)^2 + 4r(1-r)\right]^2} \right) \\ & \cdot \left(\frac{\int_0^\infty x^2 g(x)(1-G(x))\mathrm{d}x}{\int_0^\infty x^2 g(x)(1-G(x))\mathrm{d}x} \right) \end{aligned} \tag{A102}$$