American Economic Journal: Microeconomics 2012, 4(3): 57–95
http://dx.doi.org/10.1257/mic.4.3.57

Public Disagreement†

By Rajiv Sethi and Muhamet Yildiz*

We develop a model of deliberation under heterogeneous beliefs and incomplete information, and use it to explore questions concerning the aggregation of distributed information and the consequences of social integration. We show that when priors are correlated, all private information is eventually aggregated and public beliefs are identical to those arising under observable priors. When priors are independently distributed, however, some private information is never revealed, and communication breaks down entirely in large groups. Interpreting integration in terms of the observability of priors, we show how increases in social integration lead to less divergent public beliefs on average. (*JEL* D82, D83, Z13)

Members of different social groups often hold widely divergent beliefs regarding the world in which they live, even when the existence of such disagreement is itself public knowledge (see the next section for examples). Such persistent belief disparities can impede communication and interaction across social and ethnic boundaries and undermine the effectiveness of government policies. Public disagreement also appears to conflict with the standard common-prior assumption in economic theory, which implies that beliefs that are commonly known must also be identical (Aumann 1976) and that the repeated communication of beliefs eventually leads to their convergence (Geanakoplos and Polemarchakis 1982).

In this paper we develop a framework that allows for public disagreement and use it to explore questions concerning the aggregation of distributed information and the consequences of social integration. We consider a finite population of individuals who differ with respect to both their priors and their information about the state of the world. All priors and signals are assumed to be normally distributed; priors may or may not be correlated, and signals are independent. Given their priors and their information, individuals form beliefs and these beliefs are publicly and truthfully announced. The announcements are informative, and individuals update their beliefs based on them. This results in a further round of announcements, which may also be informative. The sequence of announcements continues until no further belief

* Sethi: Department of Economics, Barnard College, Columbia University, 3009 Broadway, New York, NY 10027, and the Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501 (e-mail: rs328@columbia.edu); Yildiz: Department of Economics, Massachusetts Institute of Technology, 50 Memorial Dr, Cambridge, MA 02142 (e-mail: myildiz@mit.edu). We thank the Institute for Advanced Study for financial support and hospitality, and Danielle Allen, Roland Benabou, Sylvain Chassang, Glenn Loury, Eric Maskin, Debraj Ray, Joel Sobel, anonymous referees, and seminar participants at ASSET Meeting, Brown, ISI Delhi, Johns Hopkins, New York University, Princeton, Rutgers, and the University of Illinois at Urbana-Champaign for helpful comments.

† To comment on this article in the online discussion forum, or to view additional materials, visit the article page at http://dx.doi.org/10.1257/mic.4.3.57.

revision occurs. At the end of this process, all beliefs become public information; we call these *public beliefs*. We are interested in whether or not all distributed information is incorporated into public beliefs through the process of communication, and the manner in which the extent of disagreement in public beliefs is affected by patterns of social integration.

We compare two benchmark cases that reflect the extent of social integration: observable and unobservable priors. The case of observable priors may be interpreted as a situation in which individuals understand the thought processes and perspectives of others, even if they do not share them. Such understanding could arise through social integration and mutual understanding that goes beyond the mere announcement of posterior beliefs. Since signals can be deduced from announcements when priors are observable, this case may also be interpreted as a situation in which information (rather than beliefs) can be communicated directly. The alternative case of unobservable priors corresponds to a situation in which individuals are uncertain about the manner in which others process information and form opinions, and cannot directly communicate their information. They observe beliefs but cannot immediately deduce signals from announcements. We take this to represent a less integrated society.

Given the heterogeneity of priors, public beliefs would involve some level of disagreement even if priors were observable and all relevant information aggregated. We show that unobservability of priors may inhibit the communication of some information, resulting in different levels of disagreement relative to the case of observable priors. This happens because unobservability of priors gives rise to a natural signal-jamming problem. An individual's first announcement is a convex combination of his prior and his signal. Since other individuals observe neither the prior nor the signal, they can only extract partial information about each of these from the announcement. At the end of the first round of communication, therefore, beliefs do not reflect all distributed information. We show that when priors are uncorrelated, none of the subsequent announcements has any informational value. As a result, some distributed information remains uncommunicated, despite potentially unlimited rounds of communication. Public disagreement now arises not only because of the heterogeneity of prior beliefs, but also because of informational differences induced by the fact that priors are privately observed.

Although public beliefs differ depending on whether priors are observable or unobservable, it need not be the case that unobservability of priors results in greater public disagreement. That is, there exist realizations of priors and signals such that disagreement is greater when priors are observable than when they are not. In fact, one can easily construct examples in which beliefs converge completely when priors are unobservable but remain apart under observable priors. We show, however, that the *expected value* of public disagreement must be smaller when priors are observable than when they are not. This problem becomes especially acute when the number of communicating individuals is large. When a fixed amount of information is distributed among a large number of individuals, unobservability of priors leads to a complete breakdown in communication: the difference between the public beliefs of any two individuals is approximately equal to the difference in their prior beliefs, as though no information had been received and communicated. Hence, in a large society, public disagreement is greater under unobservable priors than under observable priors at almost all realizations of priors and signals.

With correlated priors the situation is more complex. As long as each individual's prior is correlated with that of at least one other individual, we show that (subject to a regularity condition that is generically satisfied) all distributed information is fully incorporated into public beliefs *even if priors are unobservable*. While individuals may agree to disagree, their eventual beliefs are precisely what they would have been if they had observed each other's signals. This happens because the manner in which an individual responds to the announced beliefs of others reveals his beliefs about their priors, which in turn reveals his own prior. As a consequence, public beliefs in the case of unobserved (but correlated) priors are identical to those resulting from observable priors. However, convergence to public beliefs requires a larger number of rounds of communication when priors are unobserved, and involves levels of statistical sophistication that far exceed those required for convergence under observable priors. And although limiting beliefs are invariant to the manner in which information is distributed in society, beliefs held before convergence has been attained exhibit all of the properties of public beliefs under independently distributed priors.

Taking observability of priors as a proxy for social integration, we investigate the relationship between social integration and public disagreement further. We do so by exploring a variant of the model with uncorrelated priors, two social groups and three possible information structures. We say that society is *fragmented* if no priors are observable, *segregated* if each individual observes only the priors of those within his own social group, and *integrated* if all priors are observed. Our earlier results imply that expected disagreement is greater under fragmentation than under integration. A segregated society with uncorrelated priors behaves in a manner similar to a fragmented society with correlated priors: all distributed information is eventually aggregated. Before such aggregation is complete, however, the expected magnitude of public disagreement is greater under segregation than under integration, and greater under fragmentation than under segregation.

When the population size is large, the dynamics of beliefs under segregation exhibit a number of intriguing characteristics. First, differences in priors can become *amplified* through communication under segregation. In fact, even if there is no ex ante difference in prior beliefs, there will be disagreement after the first round of communication. Second, if the groups are of unequal size, then individuals belonging to the smaller group face a disadvantage under segregation even though all individuals receive equally precise signals and have access to the same belief announcements. The disadvantage arises in the *interpretation* of public announcements. Since minorities (by definition) observe the priors of a smaller segment of the total population, the beliefs of majority group members are more closely aligned with reality (interpreted as the true state) than are the beliefs of minority group members. Finally, we show that when both groups are composed of ex ante identical individuals, realized belief differences under segregation are greater than such differences under either integration or fragmentation. Segregation tends to homogenize within-group beliefs at the expense of amplifying the divergence in between-group beliefs.

The remainder of the paper is structured as follows. In the next section we discuss the existing literature and place of our own contribution within it. We introduce the model in Section II, and explore a special two-person case in Section III. When there are just two individuals, correlated priors result in the same limiting beliefs (and hence the same levels of expected disagreement) as commonly known priors. The general case is examined in Section V, where it is shown that this irrelevance of observability result continues to hold as long as the primitives of the model satisfy a genericity condition. The case of uncorrelated priors (which fails this condition) is explored earlier in Section IV, where we identify conditions under which observability of priors lowers expected disagreement relative to unobservability. Section VI uses our results to explore the relationship between social integration and public disagreement, and Section VII concludes.

I. Related Literature

Examples of Public Disagreement.—There is considerable evidence establishing that the members of different social groups have divergent beliefs on a variety of issues. Here we provide some examples based on the racial divide in the United States. A 1990 survey by the *New York Times* and WCBS found that 29 percent of black respondents (as compared with 5 percent of whites) considered it to be true or possibly true that the AIDS virus was "deliberately created in a laboratory in order to infect black people." Almost 60 percent of blacks believed that it was true or possibly true that the government "deliberately makes sure that drugs are easily available in poor black neighborhoods," and 77 percent gave credence to the claim that "the government deliberately singles out and investigates black elected officials in order to discredit them in a way it doesn't do with white officials." The corresponding numbers for white respondents were 16 percent and 34 percent respectively. These differences cannot be attributed to differences in socioeconomic status or demographic characteristics (Crocker et al. 1999).

More recently, a July 2009 poll by *Research 2000* found that 93 percent of Democrats but only 47 percent of Republicans agreed with the statement that "Barack Obama was born in the United States of America." Based on unpublished poll internals, Weigel (2009) estimated that 97 percent of black respondents but less than 30 percent of Southern whites agreed with the statement that Obama was born in the US. Along similar lines, a June 2008 survey found that while 5 percent of black respondents believed that Barack Obama was a Muslim, the corresponding figure was 12 percent for white respondents, and 19 percent for white evangelical protestants (Pew Research Center 2008). And in a poll conducted just a few days after the 2008 presidential election, 38 percent of black respondents but only 8 percent of whites stated that racial discrimination against blacks in the United States continues to be "a very serious problem" (CNN/Opinion Research 2008).

All of these differences in beliefs are a matter of public record, and appear to persist even when the public nature of the disagreement becomes inescapable. An especially dramatic example of this arose on October 3, 1995, when a nation transfixed by the criminal trial of O.J. Simpson tuned in to hear the announcement of the verdict. The following report describes the scene in New York's Times Square (Allen, O'Shaughnessy, and Chang 1995):

> *"In the moments before the O.J. Simpson verdict was announced, the crowd moved as one, heads all tilted upwards, eyes trained on the giant video screen. But when the verdict was delivered, the crowd split into two distinct camps one predominantly black, the other white and each with a vastly different response. Many blacks… reacted with jubilation. Many whites wore faces of shock and anger directed not only at the verdict, but at the reaction from blacks…Throughout the country, the scene was similar. In Wall Street offices, college campuses, stores, train stations and outside the Los Angeles County Courthouse, the Simpson verdict drew reactions that split along racial lines."*

Differences in reaction to the verdict reflected substantial racial differences in beliefs regarding the likelihood that Simpson was guilty. Brigham and Wasserman (1999) tracked such beliefs over the course of a year, starting with the period of jury selection in 1994 and ending three weeks after the announcement of the verdict. During jury selection 54 percent of whites and 10 percent of blacks in their sample thought that Simpson was "guilty" or "probably guilty." By the time closing arguments were concluded these numbers had risen to 70 percent for whites and 12 percent for blacks, reflecting an even larger racial gap. The final round of the survey, taken several days after the verdict and initial reaction had been made public, showed modest convergence but a significant remaining disparity, with 63 percent of whites and 15 percent of blacks declaring a belief in probable or certain guilt.

Social Impact of Public Disagreement.—Belief disparities can have significant welfare consequences. As Crocker et al. (1999) note, blacks and whites "exist in very different subjective worlds" and "a chasm remains… in the ways they understand and think about racial issues and events." Such differences in beliefs can make "communication and interaction across racial lines painful and difficult," as blacks find "their construal of reality flatly denied" and whites feel hurt or outraged that blacks give credence to conspiracy theories that they find bizarre or outlandish. In addition, beliefs affect responses to government policies such as public health initiatives aimed at reducing the spread of communicable diseases or the promotion of birth control. Most fundamentally, differences in beliefs about the fairness of the justice system or the extent of racial discrimination in daily life can have corrosive effects on the functioning of a democracy and erode confidence and participation in the political process. While a serious analysis of such welfare effects is beyond the scope of this paper, our analysis is motivated in part by the sense that persistent public disagreement can be welfare reducing in subtle but substantial respects.

Public Disagreement and the Common-Prior Assumption.—The persistence of public disagreement appears to conflict with the standard hypothesis in economic theory that differences across individuals in beliefs are due solely to differences in information. If this view were correct, then disagreement itself would be informative and lead to revised beliefs and eventual convergence (Geanakoplos and Polemarchakis 1982). This is the insight underlying Aumann's (1976) theorem, which states that two Bayesian individuals with a common prior must have identical posterior beliefs if these beliefs are themselves common knowledge, no matter how different their information may be. As suggested by Aumann (1976), the widespread public disagreement that one observes in practice can be attributed either to departures from the common prior assumption, or to violations of the hypothesis that Bayesian rationality itself is common knowledge.[1]

Communication and Learning with Heterogeneous Priors.—Our work contributes to a growing literature that allows for heterogeneity in prior beliefs.[2] In particular, Banerjee and Somanathan (2001), Van den Steen (2010), and Che and Kartik (2009) explore strategic communication under observable heterogeneous priors. Since heterogeneous priors lead to heterogeneous preferences, some information cannot be communicated (as in Crawford and Sobel 1982). Our work differs in allowing priors not only to be heterogeneous, but also to be *unobserved*. Furthermore, communication in our model is truthful, nonstrategic and two-sided. We consider nonstrategic communication in order to focus on the role of unobservability of priors in communication.[3] (Moreover, in the applications we have in mind, individuals do not face strong incentives to misrepresent their opinions.) In this we follow Geanakoplos and Polemarchakis (1982), who show how the agreement predicted by Aumann (1976) could arise through a sequence of truthful belief announcements. We adopt the same model of sequential announcement introduced there, but apply it to the case of heterogeneous and possibly unobserved priors.

Another strand of literature on heterogeneous priors focuses on the complementary problem of learning from external sources rather than from communication. Within that paradigm, it has been established that belief differences between Bayesian individuals may increase after they observe a public signal (Dixit and Weibull 2007), and their beliefs may even diverge asymptotically as they observe an infinite sequence of informative signals. Asymptotic divergence can occur when there are infinitely many signal values (Freedman 1965), or when individuals are uncertain about the informativeness of signals (Acemoglu, Chernozhukov, and Yildiz 2009), or when they have bounded memory (Wilson 2003). Yet another literature studies belief divergence (Andreoni and Mylanov 2011) and the formation of approximate common knowledge (Cripps et al. 2008) when players privately learn under a common prior. In this environment beliefs necessarily converge when players communicate their opinions. Interestingly, in their experiments, Andreoni and Mylanov (2011) find that subjects put lower weight on the informative actions of others than they do on their own, as predicted by our model.

[1] Since any updating rule with a mild convexity assumption can be modeled using Bayesian rationality (Shmaya and Yariv 2008), all such violations can be modeled within the general framework in which the individuals' heterogenous priors (i.e., their updating rules) are not known.

[2] Heterogeneous priors play a role in many applications, including work on asset pricing (Harrison and Kreps 1978; Morris 1996; Scheinkman and Xiong 2003), political economy (Harrington 1993), bargaining (Yildiz 2003, 2004), organizational performance (Van den Steen 2005), political polarization (Dixit and Weibull 2007) and mechanism design (Morris 1994; Eliaz and Spiegler 2007; Adrian and Westerfield 2009).

[3] Under broad conditions, Ostrovsky (forthcoming) shows that without heterogenous priors dynamic markets eventually aggregate all information despite strategic behavior by market participants. Taken together with our results, this suggests that information aggregation in public beliefs depends on whether players have *unobservable* heterogenous priors, rather than on whether they behave strategically.

In deliberations, bounded rationality may also lead to asymptotically biased beliefs and public disagreement. For example, DeMarzo, Vayanos, and Zwiebel (2003) analyze a model in which boundedly rational individuals double count some information by ignoring the fact that this information is incorporated into more than one opinion. In such a model, although beliefs converge eventually, the limiting beliefs are biased due to double counting. Similarly, Hafer and Landa (2007) analyze the deliberation of individuals who do not know the logical implications of their information. Such individuals are self-selected to deliberate with people who have a similar bias, and deliberation leads to more extreme beliefs. Our paper differs from this literature by focusing on the informational barriers to communication between rational individuals with differing priors, and on the role of social structure in such communication.

Alternative Causes of Belief Divergence.—Our focus here is on the nature of communication with heterogeneous beliefs under alternative information structures. The heterogeneity itself is a primitive of the model and we do not consider the psychological processes that might give rise to it in practice. A variety of such mechanisms have previously been explored. For instance, there is an extensive literature in psychology on confirmatory bias, which induces individuals to disregard evidence that disconfims previously held views while embracing evidence that is consistent with such views (see Rabin and Schrag 1999, and the references cited therein). Similarly, information may be processed selectively by individuals seeking to maintain a high self-image, as in Benabou and Tirole (2002), Benabou (2009), and Gottlieb (2010). Such selective information processing may lead to divergent beliefs and thereby inhibit communication. Indeed, the nature of the disagreement described in our motivating examples suggests such a mechanism. Regardless of the source of belief heterogeneity, however, it is worth exploring the question of the manner in which beliefs are affected by communication under different information structures, which is our main concern here.

II. The Model

There are n individuals $i \in N = \{1, 2, \ldots, n\}$ and an unknown real-valued parameter θ, which we call the state of the world. Individuals differ with respect to both their prior beliefs and their private information about the state of the world. Before the receipt of any information, individual i believes that θ is normally distributed with mean μ_i and unit variance:[4]

$$\theta \ \sim_i \ N(\mu_i, 1).$$

[4] We use the subscript i to denote the belief of i. For example, E_i and $E_i[\cdot \mid \cdot]$ denote the ex ante and the conditional-expectation operators under i's beliefs. We omit the subscript when all individuals agree. For example, $X \sim N(0,1)$ means that all individuals agree that X has the standard normal distribution. Likewise, E denotes the expectation operator when all individuals agree; e.g., $E[X]$ means that $E_i[X] = E_j[X] = E[X]$ for all $i, j \in N$.

Given these (possibly heterogeneous and privately observed) prior beliefs, each individual i observes a private signal x_i that is informative about θ with additive idiosyncratic noise ε_i:

$$x_i = \theta + \varepsilon_i.$$

All individuals agree that $\theta, \varepsilon_1, \ldots, \varepsilon_n$ are independently distributed, and that

$$\varepsilon_i \sim N(0, \tau^2).$$

Observing x_i, individual i updates[5] his belief about θ to a normal distribution with mean

$$A_{i,1} = \alpha\mu_i + (1 - \alpha)x_i \tag{2}$$

and variance

$$\alpha = \frac{\tau^2}{1 + \tau^2}. \tag{3}$$

Hence, one can think of μ_i as the manner in which individual i processes his information x_i, about which other individuals are uncertain. One can also think of x_i as the component of the belief of i that is perceived to be informative about θ by other individuals, and μ_i as the residual component, which is perceived by others to contain no information about θ. We refer to the pair (μ_i, x_i) as i's *type*, assuming that (μ_i, x_i) is privately known by i unless we explicitly specify that μ_i is observable, in which case μ_i will be common knowledge.

The priors $(\mu_1, \ldots, \mu_n)$ are distributed normally with mean $(\overline{\mu}_1, \ldots, \overline{\mu}_n)$ and variance-covariance matrix $\boldsymbol{\Sigma}$ with entries σ_{ij} for $i,j \in N$. A crucial assumption is that conditional on μ_i, individual i believes that the state θ, the others' priors $\boldsymbol{\mu}_{-i} = (\mu_j)_{j \neq i}$, and the noise terms ε_j, $j \in N$, are all stochastically independent. That is, player i thinks that there is some uncertainty about how each individual j processes his information x_j, but does not think that the manner in which j updates his beliefs reflects any information about θ.

Within this framework, we consider a model of deliberation involving truthful communication of beliefs in a sequence of stages, as in Geanakoplos and Polemarchakis (1982). Once signals are received, beliefs are made public in period 1 by simultaneous (and truthful) announcements $A_{i,1}$, $i \in N$, where $A_{i,1}$ denotes player i's expectation of θ conditional on the prior μ_i and the signal x_i. After observing all announcements, individuals update their beliefs and simultaneously

[5] Throughout the paper, we use the following well-known formula. If $\theta \sim N(\mu, \sigma^2)$ and $\varepsilon \sim N(0, v^2)$, then conditional on signal $s = \theta + \varepsilon$, θ is normally distributed with mean

$$E[\theta | s] = \frac{v^2}{\sigma^2 + v^2}\mu + \frac{\sigma^2}{\sigma^2 + v^2}s \tag{1}$$

and variance $\sigma^2 v^2 / (\sigma^2 + v^2)$.

announce these updated beliefs $A_{i,2}$, $i \in N$, in period 2. Here $A_{i,2}$ denotes i's expectation of θ conditional on his own prior μ_i, his own signal x_i, and the others' initial announcements $\mathbf{A}_{-i,1} = (A_{j,1})_{j \neq i}$. Individuals continue to update and announce their beliefs indefinitely. The limiting values of the sequence of announcements is denoted $A_{i,\infty}$ for $i \in N$. We call $A_{i,\infty}$ the *public belief of i*, emphasizing the fact that this belief becomes public information (i.e., common knowledge) at the end of the communication process. We assume that everything we have described to this point is common knowledge.

REMARK 1: *Since* $(\mu_1, \ldots, \mu_n)$ *may be correlated, i may think that* μ_i *is correlated with both* $\boldsymbol{\mu}_{-i}$ *and* θ*, but* $\boldsymbol{\mu}_{-i}$ *and* θ *are independent conditional on* μ_i*. Such seemingly inconsistent beliefs arise naturally as follows. Suppose that all potentially relevant historical facts are represented by a family* $\{X_m\}_{m \in M}$ *of random variables. Each individual i considers a set* $\{X_m \mid m \in R_i\}$ *of random variables to be relevant for understanding* θ *for some* $R_i \subset M$*; he considers the remaining random variables* X_m *with* $m \notin R_i$ *irrelevant. His conditional expectation of* θ *given* $\{X_m \mid m \in R_i\}$ *is* μ_i*, which is all the relevant information about* θ *in* $\{X_m\}_{m \in M}$ *according to i. Consequently, conditional on* μ_i*,* $\boldsymbol{\mu}_{-i}$ *does not affect his beliefs about* θ*; i.e., he considers* $\boldsymbol{\mu}_{-i}$ *and* θ *to be independent. On the other hand, at the ex ante stage, if i assigns positive probability to* $R_i \cap R_j \neq \emptyset$ *for some* $j \neq i$*, then i considers* μ_i *and* μ_j *to be stochastically dependent.*

REMARK 2: *The assumption that* $\boldsymbol{\mu}_{-i}$ *and* θ *are independent conditional on* μ_i *is without loss of generality: the posterior of j under the belief of player i can be decomposed into two parts, one correlated with* θ*, which i considers the relevant information contained in the belief of j, and one independent from* θ*. We also assume that* x_i *and* x_j *are independent conditional on* θ*. This independence assumption is made only for simplicity and should not affect the qualitative results.*

We conclude this section by describing the two environments that we will investigate. We say that *priors are observable* if $(\mu_1, \ldots, \mu_n)$ is common knowledge (although drawn from an ex ante distribution). We say that *priors are unobservable* if μ_i is privately known by i for each i. We use superscripts ck and u to denote variables in the observable and unobservable priors cases, respectively. For example, we write $A_{i,k}^{ck}$ or $A_{i,k}^{u}$ for the announcement of i at round k, depending on whether priors are observable or unobservable, respectively.

Under observable priors, public beliefs can be easily computed. Each individual i can deduce the signal x_j of any other individual j from the first round announcements. (Specifically, from (2), we have $x_j = (1 + \tau^2)A_{j,1} - \tau^2 \mu_j$.) Hence, individuals extract the entire relevant signal

$$(x_1 + \cdots + x_n)/n = \theta + (\varepsilon_1 + \cdots + \varepsilon_n)/n,$$

where the noise has variance

$$\bar{\tau}^2 = \tau^2/n. \tag{4}$$

Using this signal, they form their public beliefs as follows:

$$(5) \qquad A_{i,\infty}^{ck} = A_{i,2}^{ck} = \frac{\tau^2}{n + \tau^2}\mu_i + \frac{n}{n + \tau^2}\sum_{j=1}^{n}\frac{x_j}{n}.$$

Here, the expression for $A_{i,2}^{ck}$ follows from (1). Since all the available information is revealed by the first announcements, the updating stops at round 2. The difference between the public beliefs of any two individuals $i,j \in N$ is therefore simply

$$(6) \qquad A_{i,\infty}^{ck} - A_{j,\infty}^{ck} = \frac{\tau^2}{n + \tau^2}(\mu_i - \mu_j) = \frac{\overline{\tau}^2}{1 + \overline{\tau}^2}(\mu_i - \mu_j).$$

Holding constant $\overline{\tau}^2$, this difference in beliefs is independent of n. That is, under observable priors, differences in public beliefs between any pair of individuals are due only to differences in priors, which are scaled down according to the precision $1/\overline{\tau}^2$ of the distributed information. This difference in beliefs serves as the benchmark against which we measure belief differences under unobservable priors.[6]

REMARK 3: *As demonstrated above, the case of observable priors is equivalent to the case in which individuals communicate their information directly. Hence, information may be aggregated either if it is transmitted directly, or through knowledge of the manner in which others process information (i.e., their priors). These possibilities are most likely to be feasible in relatively small and well integrated groups. In contrast, in a large, fragmented society, it may not be possible for individuals to communicate their information directly, or to understand the manner in which others incorporate their information into their beliefs. Information is a complex object consisting of many small bits and pieces, and the manner in which these are incorporated into one's final opinion is itself a complex process that involves interpretation in light of one's upbringing and experience. Nevertheless, beliefs may still be communicated through opinion polls in large, fragmented societies, and this allows some inferences to be made. In our subsequent analysis, we compare information aggregation through direct communication (with observable priors) to information aggregation through indirect communication (with unobservable priors).*

III. The Two-Person Case

Before proceeding to more general results, we consider the case of two individuals. We assume without loss of generality that $\mu_i \geq \mu_j$.

[6] Note that if individuals were to receive an infinite sequence of independent signals, then they would each learn the true state even in the absence of communication and there would be no scope for public disagreement even under heterogeneous and unobservable priors.

A. *Observable Priors*

Consider the case in which the priors μ_i and μ_j are common knowledge. Since $n = 2$, (6) implies that the difference in public beliefs is

$$A^{ck}_{i,\infty} - A^{ck}_{j,\infty} = \frac{\tau^2}{2 + \tau^2}(\mu_i - \mu_j). \tag{7}$$

Note that although each individual's public belief depends on the other's initial announcement, the difference in beliefs is independent of both initial announcements, and the individuals agree on the distribution of this difference.

B. *Unobservable Independent Priors*

Next consider the case in which the priors μ_i and μ_j are not observed, and are independently distributed, each with variance σ^2. First round beliefs and announcements are exactly as in the case of observable priors:

$$A^u_{i,1} = \alpha\mu_i + (1 - \alpha)x_i.$$

Observing $A^u_{j,1}$, all i can infer is that $\alpha\mu_j + (1 - \alpha)x_j$ is equal to $A^u_{j,1}$, and cannot know the specific values of each variable. Hence, he attributes some of the variation in $A^u_{j,1}$ to variation in μ_j and some to variation in x_j. More precisely, he observes an additional signal

$$(1 + \tau^2)A^u_{j,1} - \tau^2\overline{\mu}_j = \theta + \tau^2(\mu_j - \overline{\mu}_j) + \varepsilon_j \tag{8}$$

with additive noise $\tau^2(\mu_j - \overline{\mu}_j) + \varepsilon_j$. The noise term has mean 0 and variance $\sigma^2\tau^4 + \tau^2$. He then updates his beliefs to a normal distribution with mean

$$\begin{aligned} A^u_{i,2} &= \frac{\sigma^2\tau^4 + \tau^2}{\alpha + \sigma^2\tau^4 + \tau^2}A^u_{i,1} + \frac{\alpha}{\alpha + \sigma^2\tau^4 + \tau^2}\left((1 + \tau^2)A^u_{j,1} - \tau^2\overline{\mu}_j\right) \\ &= \frac{1}{\gamma}\left((1 + \sigma^2\tau^2)(1 + \tau^2)A^u_{i,1} + (1 + \tau^2)A^u_{j,1} - \tau^2\overline{\mu}_j\right), \end{aligned} \tag{9}$$

where $\gamma = (1 + \tau^2)(1 + \tau^2\sigma^2) + 1$. Here, the first equality is obtained by updating according to (1) starting from $\theta \sim N(A^u_{i,1}, \alpha)$ and using the signal in (8), and the second equality is by (3). Note that (unlike the case of commonly known priors) i puts greater weight on his own announcement than on that of j. This is because i does not know j's prior. When j announces a higher expectation $A^u_{j,1}$, i believes that with some probability j has obtained a higher value of the signal x_j, motivating i to increase his own expectation of θ too. He also thinks that, with some probability, the high announcement may be due to a bias towards higher values (i.e., larger μ_j), in which case i would not want to increase his expectation of θ. Consequently, each player's beliefs become less sensitive to the other's announcement than in the case of commonly known priors.

Even after the second round announcements, i does not know x_j, so there remains some relevant asymmetric information. In other words, the distributed information is not aggregated at the first round.[7] One might hope that further announcements communicate more private information, resulting in the aggregation of the remaining distributed information. This is not the case, however. Since $A^u_{i,1}$ and $A^u_{j,1}$ are sufficient statistics for $A^u_{i,2}$ and $A^u_{j,2}$, the second round announcements provide no additional information, and

$$A^u_{i,2} = A^u_{i,3} = \ldots = A^u_{i,\infty}.$$

The difference in public beliefs is

$$(10) \qquad A^u_{i,\infty} - A^u_{j,\infty} = \frac{1}{\gamma}\big(\sigma^2\tau^2(1 + \tau^2)(A_{i,1} - A_{j,1}) + \tau^2(\overline{\mu}_i - \overline{\mu}_j)\big)$$
$$= \frac{\tau^2}{\gamma}\big((\overline{\mu}_i - \overline{\mu}_j) + \sigma^2\tau^2(\mu_i - \mu_j) + \sigma^2(\varepsilon_i - \varepsilon_j)\big).$$

The difference of opinion has three sources: the difference in the means of the distributions from which priors are drawn $(\overline{\mu}_i - \overline{\mu}_j)$, the difference in the realized values of the priors $(\mu_i - \mu_j)$, and the difference in information $(\varepsilon_i - \varepsilon_j)$. Since communication never completely eliminates informational differences, these differences affect public beliefs. Communication does, however, decrease the role of differential information as the coefficient of $(A_{i,1} - A_{j,1})$ is strictly less than 1. That is, differences in information play a larger role in affecting initial announcements than in affecting public beliefs. As in the common knowledge case, all individuals agree on the distribution of the difference in public beliefs.

Note from (10) that the two individuals will generally agree to disagree even if they have *identical* priors $(\mu_i = \mu_j)$, since they cannot deduce from the announcements that their priors are in fact identical. This makes transparent the obvious but sometimes overlooked fact that the standard common prior assumption requires not only that the players have the same prior, but also that this fact is itself commonly known. Furthermore, even if both individuals have identical priors and receive identical signals $(\varepsilon_i = \varepsilon_j)$, they may disagree once their beliefs have been communicated, provided that the priors themselves are not drawn from identical distributions. The following numerical example illustrates.

Example 1: *Suppose that* $\overline{\mu}_i = 0$, $\overline{\mu}_j = 2$, *and* $\mu_i = \mu_j = x_i = x_j = 1$. *Then,* $A_{i,1} = A_{j,1} = 1$, *while* $A_{i,2} = 0.8$ *and* $A_{j,2} = 1.2$.

In this example, both individuals have identical priors and signals, and make identical initial announcements. But since their priors are not observable, and are drawn from different distributions, they interpret each others announcements in different ways, resulting in a divergence of beliefs over time. Communication can therefore

[7] We say that all distributed information is aggregated if all private signals become known to all individuals; a formal definition is provided in Section V.

lead to increased polarization when priors are unobserved even when individuals receive exactly the same information.

In conclusion, uncertainty about the manner in which other individuals process information hinders the communication of relevant private information through the announcement of beliefs. Consequently, individuals hold different beliefs both because they have (possibly) different priors and because of different information.

C. *A Comparison of Belief Differences*

Note that $A_{i,\infty} - A_{j,\infty}$ measures the amount that i overestimates θ relative to j at the end of the process of deliberation. Hence, we call $A_{i,\infty} - A_{j,\infty}$ the *public bias* of i relative to j. Since uncertainty regarding priors leads to less communication of information, one may think that it also leads to greater public bias. This is not the case. It may so happen that the individuals have very different priors, and knowledge of this may lead to a very large difference of opinion. Indeed, when the priors are not observed, by (10), any amount of public bias is possible, including no bias at all. In contrast, when the priors are common knowledge, by (7), the amount of public bias is constant, depending only on the difference in realized priors.

Figure 1 plots the values of public bias under observed and unobserved priors, respectively, for a set of randomly drawn type realizations.[8] Here, for each realization, the horizontal coordinate is $A^{ck}_{i,\infty} - A^{ck}_{j,\infty}$ and the vertical coordinate is $A^{u}_{i,\infty} - A^{u}_{j,\infty}$. In the realizations that lie below the diagonal, public beliefs differ more when priors are observable. Hence the figure demonstrates that making priors observable may lead to greater disagreement in many cases.

While observability of priors can result in greater public bias for particular type realizations, observability always lowers the ex ante expected value of public bias, $E[A_{i,\infty} - A_{j,\infty}]$. To see this, note that when priors are observable, by (7), the expected bias in public beliefs is

$$E\left[A^{ck}_{i,\infty} - A^{ck}_{j,\infty}\right] = \frac{\tau^2}{2 + \tau^2}(\overline{\mu}_i - \overline{\mu}_j).$$

On the other hand, when the priors are not observable, by (10), the expected public bias is

$$E\left[A^{u}_{i,\infty} - A^{u}_{j,\infty}\right] = \frac{\tau^2(1 + \sigma^2\tau^2)}{1 + (1 + \sigma^2\tau^2)(1 + \tau^2)}(\overline{\mu}_i - \overline{\mu}_j).$$

If $\overline{\mu}_i = \overline{\mu}_j$ then the expected public bias is zero in both cases. If $\overline{\mu}_i > \overline{\mu}_j$, however, then $\sigma^2 > 0$ implies

$$E\left[A^{u}_{i,\infty} - A^{u}_{j,\infty}\right] > E\left[A^{ck}_{i,\infty} - A^{ck}_{j,\infty}\right] > 0.$$

That is, the expected public bias is higher when priors are not observable than when they are observable. This is intuitive because unobservability of priors

[8]The figure is based on 500 realizations of type profiles for parameter values $\sigma^2 = \tau^2 = 1$, $\overline{\mu}_i = 3$, and $\overline{\mu}_j = 0$.

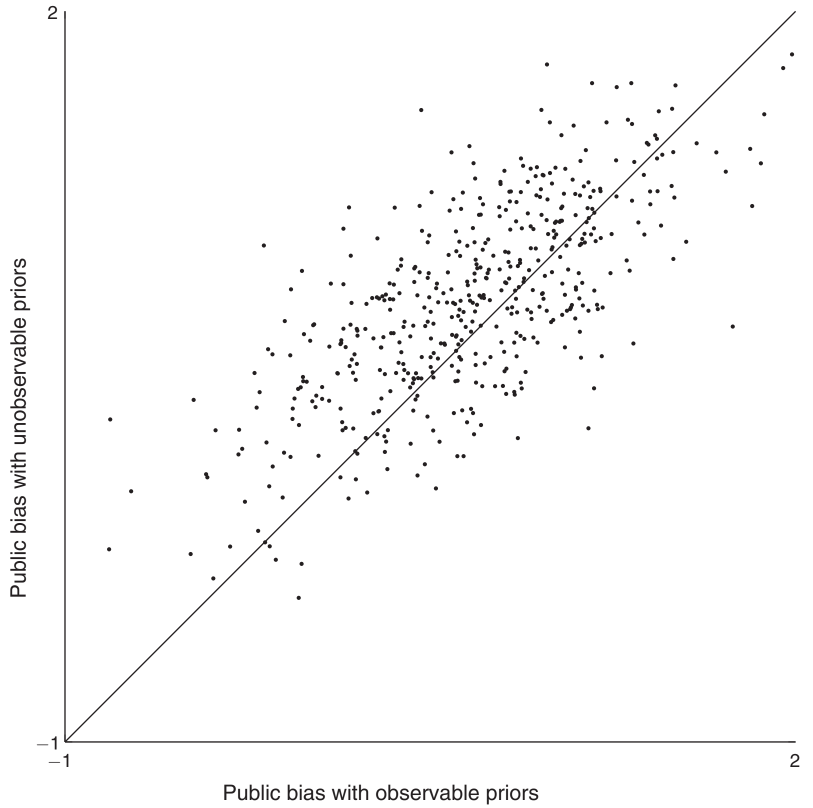

Figure 1. Public Bias with Observable and Unobservable Priors

impedes the full aggregation of the distributed information through deliberation. This result is useful in comparing the difference between the average opinions of various groups. For example, it implies that differences across groups in beliefs about the incidence of police brutality or racial profiling would narrow on average if members of each group were to observe each other's priors and therefore understand how their information is incorporated into beliefs. We return to this point in Section VI.

D. *Unobservable Correlated Priors*

Under the assumption that priors are uncorrelated, we have so far illustrated that unobservability of priors may impede the aggregation of distributed information through deliberation and affect the amount of public disagreement. We now show that when priors are correlated, all distributed information is aggregated and hence the observability of priors has no effect on public beliefs.

Assume that μ_i and μ_j are correlated:

$$\begin{pmatrix} \mu_i \\ \mu_j \end{pmatrix} \sim N\left(\begin{pmatrix} \overline{\mu}_i \\ \overline{\mu}_j \end{pmatrix}, \ \sigma^2 \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}\right),$$

where $\rho \neq 0$. Observing μ_i, i believes that μ_j is distributed normally with mean

$$E_i[\mu_j \mid \mu_i] = \overline{\mu}_j + \rho(\mu_i - \overline{\mu}_i)$$

and variance

$$\mathrm{Var}_i(\mu_j \mid \mu_i) = \sigma^2(1 - \rho^2).$$

That is, $E_i[\mu_j \mid \mu_i]$ is a one-to-one function of μ_i. As before, we have $A^u_{i,1} = A_{i,1}$ and $A^u_{j,1} = A_{j,1}$. Now, for i, the announcement $A^u_{j,1}$ of j in the first round yields an additional noisy signal

$$(1 + \tau^2)A^u_{j,1} - \tau^2 E_i[\mu_j \mid \mu_i] = \tau^2(\mu_j - E_i[\mu_j \mid \mu_i]) + x_j \tag{11}$$
$$= \theta + \tau^2(\mu_j - E_i[\mu_j \mid \mu_i]) + \varepsilon_j.$$

The additive noise $\tau^2(\mu_j - E_i[\mu_j \mid \mu_i]) + \varepsilon_j$ has mean 0 and variance $\sigma^2(1 - \rho^2)\tau^4 + \tau^2$. Updating his belief, in the second round i announces

$$A^u_{i,2} = KA^u_{i,1} + LA^u_{j,1} - \alpha L E_i[\mu_j \mid \mu_i],$$

where K and L are known strictly positive constants.[9] The crucial observation here is that $A^u_{i,2}$ is strictly decreasing in $E_i[\mu_j \mid \mu_i]$, which is i's expectation of j's prior once i has observed his own prior. Player j, having observed $A^u_{i,1}$ and $A^u_{j,1}$ from the previous round, can therefore use $A^u_{i,2}$ to deduce that

$$E_i[\mu_j \mid \mu_i] = (\alpha L)^{-1}(KA^u_{i,1} + LA^u_{j,1} - A^u_{i,2}).$$

Moreover, since $E_i[\mu_j \mid \mu_i] = \overline{\mu}_j + \rho(\mu_i - \overline{\mu}_i)$ and $\rho \neq 0$, there is a one-to-one mapping between μ_i and $E_i[\mu_j \mid \mu_i]$. Hence j correctly infers that

$$\mu_i = \overline{\mu}_i + \rho^{-1}\left((\alpha L)^{-1}(KA^u_{i,1} + LA^u_{j,1} - A^u_{i,2}) - \overline{\mu}_j\right).$$

That is, at the end of second round, all prior beliefs are revealed, and all signals can be inferred. The announcements in all subsequent rounds are therefore precisely the same as in the common knowledge case:

$$A^u_{i,3} = \cdots = A^u_{i,\infty} = A^{ck}_{i,\infty} = \frac{1 + \tau^2}{2 + \tau^2}(A_{i,1} + A_{j,1}) - \frac{\tau^2}{2 + \tau^2}\mu_j.$$

[9] One applies (1), starting from $\theta \sim N(A^u_{i,1}, \alpha)$ and using the signal in (11), to obtain

$$A^u_{i,2} = \frac{\sigma^2(1 - \rho^2)\tau^4 + \tau^2}{\alpha + \sigma^2(1 - \rho^2)\tau^4 + \tau^2}A^u_{i,1} + \frac{\alpha}{\alpha + \sigma^2(1 - \rho^2)\tau^4 + \tau^2}\left((1 + \tau^2)A^u_{j,1} - \tau^2 E_i[\mu_j \mid \mu_i]\right).$$

The desired equation is obtained by letting K and L respectively denote the coefficients of $A^u_{i,1}$ and $A^u_{j,1}$.

Accordingly, when priors are correlated, both individuals can infer each other's prior beliefs from the manner in which they *react* to the initial announcements. All distributed information is therefore aggregated through communication, and the resulting public bias is fully attributable to differences in prior beliefs:

$$A^u_{i,\infty} - A^u_{j,\infty} = \frac{\tau^2}{2 + \tau^2}(\mu_i - \mu_j).$$

We show in Section V that this is true under broad conditions. First, however, we consider the case of uncorrelated priors.

IV. Public Biases

In this section, we explore the impact of observability of priors on the degree of bias in public beliefs under the assumption that priors are independently and identically distributed.

ASSUMPTION 1: *The variance-covariance matrix for priors is* $\mathbf{\Sigma} = \sigma^2\mathbf{I}$.

That is, for all distinct pairs i and j, the priors μ_i and μ_j are independent (i.e., $\sigma_{ij} = 0$) and the variances of priors are equal (i.e., $\sigma_{ii} = \sigma^2$ for all i).

Consider two individuals, i and j. At the end of deliberation, j thinks that the expected value of θ is $A_{j,\infty}$. He also knows that i thinks that the expected value of θ is $A_{i,\infty}$. Therefore, j thinks that i overestimates θ by an amount $A_{i,\infty} - A_{j,\infty}$. This leads to our notion of public bias.

DEFINITION 1: *For any* $i,j \in N$, *the public bias of* i *relative to* j *is* $A_{i,\infty} - A_{j,\infty}$.

Similarly, the *ex ante bias* of i relative to j is $\overline{\mu}_i - \overline{\mu}_j$. The bias after i and j have observed their own priors but before they observe any information is $\mu_i - \mu_j$, which we call the *prior bias* of i relative to j. Note that the ex ante bias is known to all players, and the public bias comes to be known through communication, but the prior bias may never be revealed.

We know from (6) that when priors are common knowledge, the only source of public bias is the difference in realized priors, $\mu_i - \mu_j$, which is scaled down through communication. The following lemma identifies the amount of public bias when priors are unobservable, generalizing the analysis of Section III to n individuals.

LEMMA 1: *Under Assumption* 1, *for any* i *and* j, *the public bias of* i *relative to* j *under unobservable priors is*,

$$(12) \quad A^u_{i,\infty} - A^u_{j,\infty} = \frac{\tau^2}{\gamma_n}\big((\overline{\mu}_i - \overline{\mu}_j) + \tau^2\sigma^2(\mu_i - \mu_j) + \sigma^2(\varepsilon_i - \varepsilon_j)\big).$$

where $\gamma_n = (1 + \tau^2)(1 + \tau^2\sigma^2) + n - 1$.

Under unobservable priors, public bias has three sources: ex ante bias $(\overline{\mu}_i - \overline{\mu}_j)$, prior bias $(\mu_i - \mu_j)$, and informational difference $(\varepsilon_i - \varepsilon_j)$. The informational difference contributes to public bias because unobservability of priors impedes the full aggregation of information. Ex ante bias affects public bias because, without full aggregation, individuals use ex ante information on priors to estimate the information of others.

By Lemma 1, the magnitude of public biases does not depend on θ. Hence all individuals agree on the distribution of these biases (although they disagree on the distribution of public beliefs). Our next result establishes that, if the priors are drawn from distinct distributions, the expected bias is necessarily larger under unobservable priors. (The expected bias is always zero when priors are drawn from the same distribution.)

PROPOSITION 1: *Under Assumption* 1, *if* $\overline{\mu}_i > \overline{\mu}_j$, *then*

$$E\left[A_{i,\infty}^{u} - A_{j,\infty}^{u}\right] > E\left[A_{i,\infty}^{ck} - A_{j,\infty}^{ck}\right] > 0.$$

Consider two individuals i and j. Suppose that $\overline{\mu}_i > \overline{\mu}_j$ so that at the ex ante stage i overestimates θ relative to j, although the actual prior μ_i of i may or may not turn out to be larger than μ_j. After each player k forms his prior and receives his information, all individuals deliberate, communicating their beliefs. At the end of this process, their beliefs become public. Proposition 1 establishes that all individuals expect that, at the end of the process of deliberation, i overestimates θ less vis-à-vis j when priors are observable. That is, $E_k\left[A_{i,\infty}^{ck} - A_{j,\infty}^{ck}\right] < E_k\left[A_{i,\infty}^{u} - A_{j,\infty}^{u}\right]$ according to each $k \in N$. Therefore, making priors observable decreases public biases on average. This suggests that social integration, interpreted as an increased understanding of the manner in which other people think, should result in lower levels of public disagreement on average. We explore these issues further in Section VI.

We conclude this section with a discussion of the manner in which increases in population size affect the aggregation of distributed information. When information is distributed among a large number of individuals, unobservability of priors becomes detrimental for communication, so much so that the bias at the end of the deliberation process is approximately the same as the bias before deliberation begins. Towards establishing this, recall from (4) that the distributed information in society has variance $\overline{\tau}^2 = \tau^2/n$. If one fixes τ and varies n, as n gets large, the distributed information becomes very precise. Consequently, the individuals approximately learn θ from each other. In order to disentangle the effect of group size n from the effect of the information available to the group, we now fix the precision of the distributed information and let n vary.

In particular, consider a family of models $(\tau_n^2, \sigma^2, \overline{\mu}_n, n)$, indexed by the number of individuals n, where $\overline{\mu}_n = (\overline{\mu}_1, \ldots, \overline{\mu}_n)$ is the vector of means for the priors; $\mu_i \sim N(\overline{\mu}_i, \sigma^2)$ for each $i \leq n$. We assume that the variance τ_n^2/n approaches some positive value $\overline{\tau}^2$ as $n \to \infty$. (A special case of this arises if the variance of distributed information is independent of n, i.e., the total amount of information is fixed. In that case, $\tau_n^2 = n\overline{\tau}^2$ for some fixed $\overline{\tau} > 0$.) For any distinct individuals i and j, and any pair of realized priors μ_i and μ_j, this family of models defines a sequence of random variables $(A_{i,\infty}^{u} - A_{j,\infty}^{u})_n$. Our next result shows that under unobservable priors, as the number of individuals n becomes large, this sequence of random variables converges in distribution to $\mu_i - \mu_j$.

PROPOSITION 2: *Under Assumption* 1, *for any family* $(\tau_n^2, \sigma^2, \overline{\mu}_n, n)$ *of models*, *any distinct individuals i and j*, *and any realized priors* (μ_i, μ_j),

$$(A_{i,\infty}^u - A_{j,\infty}^u)_n \xrightarrow{\mathcal{D}} \mu_i - \mu_j .$$

PROOF:

By Lemma 1,

$$E\left[(A_{i,\infty}^u - A_{j,\infty}^u)_n \,|\, \mu_i, \mu_j\right] = \tau_n^2 \sigma^2 \eta (\mu_i - \mu_j) + \eta (\overline{\mu}_i - \overline{\mu}_j),$$

where

$$\eta = \frac{\tau_n^2/n}{n(\tau_n^2/n)^2 \sigma^2 + (\tau_n^2/n)(1 + \sigma^2) + 1}.$$

As $n \to \infty$ and $\tau_n^2/n \to \overline{\tau}^2 > 0$, η goes to 0, while $\tau_n^2 \sigma^2 \eta$ goes to 1. Hence

$$\lim_{n \to \infty} E\left[(A_{i,\infty}^u - A_{j,\infty}^u)_n \,|\, \mu_i, \mu_j\right] = \mu_i - \mu_j.$$

To complete the proof we need to show that the variance of $(A_{i,\infty}^u - A_{j,\infty}^u)_n$ goes to 0 as $n \to \infty$. Since μ_i and μ_j are given, from Lemma 1 we have

$$\text{Var}\left[(A_{i,\infty}^u - A_{j,\infty}^u)_n \,|\, \mu_i, \mu_j\right] = 2\tau_n^2 \sigma^4 \eta^2 = 2\tau_n^2 \sigma^2 \eta (\sigma^2 \eta).$$

Since η goes to 0 and $\tau_n^2 \sigma^2 \eta$ goes to 1 as $n \to \infty$,

$$\lim_{n \to \infty} \text{Var}\left[(A_{i,\infty}^u - A_{j,\infty}^u)_n \,|\, \mu_i, \mu_j\right] = 0.$$

Hence, when the number n of individuals is large, the public bias of i relative to j is approximately equal to the prior bias of i relative to j. In the limit, all distributed information, no matter how precise, is entirely dissipated.[10]

REMARK 4: *For simplicity*, *Proposition* 2 *assumes that there is some residual uncertainty in the limit even if one can aggregate all the information* (*i.e.*, $\lim \tau_n^2/n > 0$). *Although there are many cases in which such uncertainty remains* (*such as religious disagreements*), *some may question the validity of this assumption in other cases. This assumption is not necessary*, *and the proposition illustrates a broader fact. To see this*, *assume that* $\tau_n^2/n \to 0$ *but* $\tau_n^4/n \to \infty$. *That is*, *while the total distributed information becomes completely informative in the limit*, *the individuals' signals get less informative as we distribute the information among a larger group. Then*, *one can easily check in the proof of proposition that* $\eta \to 0$ *and* $\tau_n^2 \sigma^2 \eta \to 1$, *showing that* $(A_{i,\infty}^u - A_{j,\infty}^u)_n \xrightarrow{\mathcal{D}} \mu_i - \mu_j$, *as in the proposition.*

[10] Note that since $\mu_i - \mu_j$ is a constant (conditional on the realized priors μ_i and μ_j), convergence in distribution implies convergence in probability. Hence we also have plim $A_{i,\infty}^u - A_{j,\infty}^u = \mu_i - \mu_j$.

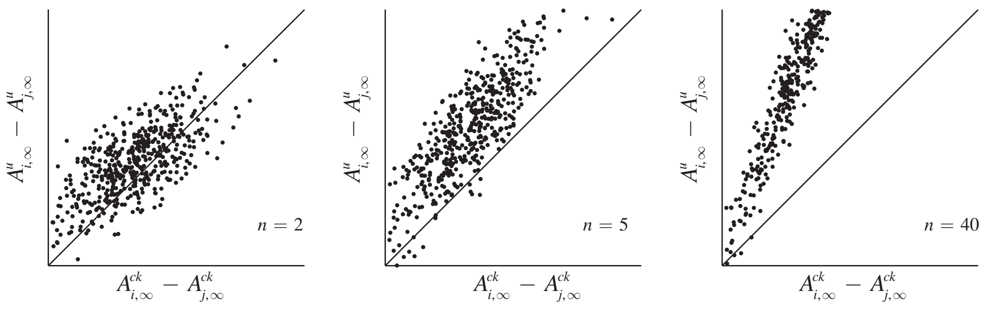

Figure 2. Public Bias with Observable and Unobservable Priors for Various n

Under observable priors, individuals use all distributed information efficiently. Hence, their public beliefs and the bias in those beliefs do not depend on how information is distributed. When priors are unobservable, however, even if individuals have very precise information as a group and announce their beliefs sincerely, they cannot communicate any significant information: at the end of the deliberation process, their beliefs are as they were at the outset. The intuition is that each individual has such a small amount of information that their announcements reveal little more than their priors. Recall from (6) that

$$A^{ck}_{i,\infty} - A^{ck}_{j,\infty} = \frac{\overline{\tau}^2}{1 + \overline{\tau}^2}(\mu_i - \mu_j),$$

so the difference in beliefs under observable priors is independent of n holding $\overline{\tau}$ fixed. An immediate implication of Proposition 2 is therefore the following: as the population size becomes large, so that a given amount of information is distributed among an increasingly large number of individuals, public bias under unobservability is greater *not only in expectation but also for almost all type realizations*. This is illustrated in Figure 2, which repeats the exercise depicted in Figure 1 but for three different values of group size. As n gets large, type realizations for which observability results in greater public bias (which lie below the diagonal) become increasingly rare.[11]

Note that for large n, Proposition 2 and (6) imply that

$$A^u_{i,\infty} - A^u_{j,\infty} = \frac{1 + \overline{\tau}^2}{\overline{\tau}^2}(A^{ck}_{i,\infty} - A^{ck}_{j,\infty}).$$

Hence, the public bias under unobservability is a linear function of the public bias under observability, with slope greater than 1. This is illustrated for the case of $n = 40$ in the right panel of Figure 2.

[11] Each plot is based on 500 realizations of type profiles for parameter values $\sigma^2 = 1$, $\overline{\mu}_i = 3$, $\overline{\mu}_j = 0$, and $\tau_n^2/n = \overline{\tau}^2 = ½$ for all n. Only type realizations at which $\mu_i \geq \mu_j$ (and public biases lie in the positive quadrant) are shown. For realizations at which $\mu_i < \mu_j$ our results imply that as n gets large, public biases will lie below the diagonal in the negative quadrant.

V. Aggregation of Distributed Information

We now turn to the general model with unobservable priors, and provide a near-characterization of the cases in which the private information of an individual is revealed through deliberation. We show that, roughly speaking, if an individual's prior is correlated with the prior of any other individual, his private information is revealed by the end of the second round; otherwise his information is never revealed. Hence, except for certain knife-edge cases (as in the example of independent priors), the process of sequential belief announcements leads to the aggregation of all distributed information.

The idea that an individual's private information is revealed through communication is formalized as follows.

DEFINITION 2: *We say that the private information of individual i is* revealed *by (the end of) round k if* (μ_i, x_i) *is measurable with respect to* $\{A_{j,m}\}_{j \in N, m \leq k}$. *If the private information of individual i is not revealed by round k for any k, we say that his private information is* never revealed.

That is, the private information of i is revealed by the end of round k if, by observing all announcements up to and including those in round k, one can compute his prior belief μ_i and signal x_i. In that case, his private information will be common knowledge at any round $m > k$:

$$A_{j,m} = E_j\left[\theta \mid \mu_i, x_i, \mu_j, x_j, \{A_{i',l}\}_{i' \in N \setminus \{i,j\}, l \leq m}\right] \qquad (\forall j \in N).$$

To present our characterization, we introduce the following notation. For any $i \in N$, we define column vectors $\boldsymbol{\mu}_{-i} = (\mu_j)_{j \neq i}$ and $\boldsymbol{\sigma}_{-i,i} = (\sigma_{j,i})_{j \neq i}$ and write $\boldsymbol{\Sigma}_{-i,-i} = (\sigma_{j,k})_{j \neq i, k \neq i}$ for the variance covariance matrix of $\boldsymbol{\mu}_{-i}$. We write $\mathbf{1}_{k \times l}$ for the $k \times l$-dimensional matrix with entries 1 and $\mathbf{I}$ for the identity matrix. Finally, we define the row vector $\mathbf{M}_i$ as follows:

$$\mathbf{M}_i = \mathbf{1}_{1 \times n-1}\left(\alpha \mathbf{1}_{n-1 \times n-1} + \tau^2 \mathbf{I} + \tau^4 \left(\boldsymbol{\Sigma}_{-i,-i} - \sigma_{ii}^{-1} \boldsymbol{\sigma}_{-i,i} \boldsymbol{\sigma}'_{-i,i}\right)\right)^{-1}.$$

Note that $\mathbf{M}_i$ depends only on the primitives of the model and is therefore independent of all type realizations. The next definition provides the terminology of the characterization.

DEFINITION 3: *We say that i is* isolated *if* $\boldsymbol{\sigma}_{-i,i} = 0$. *We say that i is* regular *under* $(\tau^2, \boldsymbol{\Sigma})$ *if* $\mathbf{M}_i \boldsymbol{\sigma}_{-i,i} \neq 0$. *We say that* $(\tau^2, \boldsymbol{\Sigma})$ *is* regular *if every i is regular under* $(\tau^2, \boldsymbol{\Sigma})$.

Note that i is isolated if and only if μ_i is independent of all other priors μ_j. In this case, i cannot infer any information about the priors of others from his own prior. Consequently, others cannot learn about i's prior from the way he reacts to their announcements, and it is not possible to uncover all of his private information.

The regularity condition $\mathbf{M}_i \boldsymbol{\sigma}_{-i,i} \neq 0$ corresponds to the case that i's second-round announcement contains some new information. It only rules out some knife-edge cases, such as isolation. Indeed, $\mathbf{M}_i \boldsymbol{\sigma}_{-i,i} = 0$ is a nontrivial linear equality restriction on the variances (τ^2, Σ), and hence is satisfied only on a lower-dimensional subspace of the space of all variances (τ^2, Σ). In particular, the set of regular parameters (τ^2, Σ) has full Lebesgue measure and is open and dense.

Our characterization establishes that whether the private information of an individual is revealed depends on whether he is regular or isolated.

PROPOSITION 3: *Assume that priors are not observable. If i is regular, then his private information is revealed by the end of round* 2. *Conversely, if i is isolated, then his private information is never revealed.*

An immediate implication of this is:

COROLLARY 1: *If* (τ^2, Σ) *is regular, then all private information is revealed by the end of round* 2.

This result establishes the *irrelevance of observability*: public beliefs under unobservable priors are identical to public beliefs under common knowledge of priors as long as (τ^2, Σ) is regular. All information is aggregated no matter how little individuals know about each other's way of thinking. Moreover, as in the two person case, this process requires just two rounds of communication.

In order to prove Proposition 3, in the Appendix, we compute the announcements (see Lemma 2). After the first round, the announcement of an individual i is an affine function of the first round announcements of all individuals, the priors of the individuals whose information has been revealed, and the prior μ_i of i himself. In the second round announcement, the coefficient of μ_i is proportional to $\mathbf{M}_i \boldsymbol{\sigma}_{-i,i}$. Hence, when $\mathbf{M}_i \boldsymbol{\sigma}_{-i,i} \neq 0$, other individuals can compute μ_i using the publicly available information and $A_{i,2}$. In that case, the private information of i is revealed by the end of the second round. Moreover, in any round after the first, the coefficient of μ_i is proportional to $\boldsymbol{\sigma}_{-i,i}$. When $\boldsymbol{\sigma}_{-i,i} = 0$, the announcement of i does not contain any new information because it is a function of publicly available information, namely the first round announcements and the priors that have already been revealed. In that case, i's private information is never revealed.

Since all private information is revealed by the end of the second round when (τ^2, Σ) is regular, the difference in public beliefs with unobservable priors is identical to the difference with observable priors: $A_{i,k}^u = A_{i,k}^{ck}$ for all i and $k \geq 3$. Using (6), we therefore have

$$A_{i,\infty}^u - A_{j,\infty}^u = A_{i,\infty}^{ck} - A_{j,\infty}^{ck} = \frac{\tau^2}{n + \tau^2}(\mu_i - \mu_j) = \frac{\overline{\tau}^2}{1 + \overline{\tau}^2}(\mu_i - \mu_j).$$

That is, under the regularity assumption, regardless of whether priors are observable or unobservable, differences in public beliefs are due only to differences in priors, scaled down by a factor that depends on the precision $1/\overline{\tau}^2$ of the distributed information.

The regularity assumption is weaker than genericity and contains many interesting "non-generic" cases, as the following example illustrates.

EXAMPLE 2: *Take $N = B \cup W$ consisting of two groups $B = \{1,2\}$ and $W = \{3,4\}$. For each i, $\sigma_{ii} = \sigma^2$ and for all distinct individuals i and j, $\sigma_{ij} > 0$ if i and j are in the same group and $\sigma_{ij} = 0$ otherwise. That is, from his own prior, an individual can learn about the other individual's prior in his own group, but he cannot learn anything about the other group. Nevertheless, (τ^2, Σ) is regular. One can check that, for any $i \in N$,*

$$\mathbf{M}_i \boldsymbol{\sigma}_{-i,i} \propto 1 + \tau^2 + \sigma^2\tau^2 + \sigma^2\tau^4 + \sigma^2\tau^2\rho + \sigma^2\tau^4\rho \neq 0.$$

This example illustrates that even in a segregated society with no correlation in priors across groups, all distributed information is incorporated into public beliefs.

Consider a society composed of several subgroups such that priors are correlated within groups but independently distributed across groups. In light of Propositions 2 and 3, one might be tempted to conclude that information in such a society would be aggregated within but not across groups. Example 2 demonstrates that this claim would be false. As long as each group is composed of multiple individuals, all individuals can infer the priors and information of all others by observing the manner in which they respond to the announcements of those within their respective groups.[12]

Proposition 3 implies that public beliefs are discontinuous with respect to the correlation of priors. When the priors are correlated, no matter how small the correlation may be, public beliefs incorporate all private information. When priors are independent, however, a substantial amount of private information remains private. This is true even for the third round announcements. The discontinuity in finite rounds stems from our assumption that individuals can communicate their beliefs precisely, and understand the correlation structure perfectly. In reality, individuals have noisy information about the beliefs of others, and imprecise estimates of the correlation structure, which could lead to a continuous relationship between disagreement and correlation, with a substantial amount of private information remaining uncommunicated at each round. In that case, the beliefs at any given round would be a continuous function of parameters such as the correlation coefficients. The public beliefs $A_{i,\infty}$ may remain discontinuous with respect to these parameters. Accordingly, we consider full aggregation after only two rounds to be an artifact of our simplifying assumptions.

Full aggregation after only two rounds is also an artifact of the two-dimensional model we use for tractability. In general, communication of information through announcements can take arbitrarily many rounds, depending on the level at which the model is "closed". More, precisely, in any type space, every type has a belief

[12] In particular, a large population consisting of a small number of subgroups (with correlated priors within but not across groups) does not behave like a small population with independent priors.

about θ (i.e., first-order beliefs), beliefs about θ and the other individuals' first-order beliefs (i.e., second-order beliefs) and so on. Typically, a type space is closed at some order k in the sense that there is only one type with a given first k orders of beliefs. In such a model, one needs to knows the first k orders of beliefs to fully learn the individuals' private information, and this requires k rounds of communications in general. An individual's first-round announcement provides information about his first-order belief. By observing the first-round announcements, one then updates his beliefs using his second-order belief, a joint probability distribution of θ and the first-order beliefs. The second-round announcements then give information only on the second-order beliefs. In the third round then one uses his newly acquired information about the other parties' first and second order beliefs and his own third-order belief to update his belief about θ. In this way, the announcements in the first m rounds reveal only the information contained in first m orders of beliefs. If the model is closed in k orders, one then needs k rounds of communication to learn all of the private information, while some types' information can be revealed in earlier rounds. Although the models are not exactly same, an example of this can be seen in Geanakoplos and Polemarchakis (1982), who show that even under the common prior assumption beliefs may take arbitrarily long to stabilize.[13] Hence the complete aggregation of distributed information could take an arbitrarily large number of communication rounds in a more general setting. Accordingly, we view Proposition 3 to be demonstrating that all information is aggregated under correlated priors once belief announcements cease to be informative.

In our model, beliefs are represented by expectations, and the first and the second-order beliefs of a type (μ_i, x_i) are $A_{i,1} = \alpha\mu_i + (1 - \alpha)x_i$ and $E_i[(\theta, A_{j,1})] = (A_{i,1}, \alpha E[\mu_j \mid \mu_i] + (1 - \alpha)A_{i,1})$, respectively. Under independence, $E[\mu_j \mid \mu_i] = \overline{\mu}_j$, and the second-order beliefs are already determined by the first-order beliefs, i.e., the model is closed at the first order. In fact, informative communication stops after the first round. With correlation, $E[\mu_j \mid \mu_i]$ varies with μ_i, and the second-order beliefs can vary for a fixed first-order belief $A_{i,1}$. Nonetheless, from the second-order beliefs, one can solve for μ_i and x_i, figuring out the entire hierarchy of beliefs. In this case, the model is closed at the second round, and there are two rounds of informative communication.

Complete aggregation of distributed information in the limit relies on the assumption that all individuals have high levels of statistical sophistication. Not only are they able to make rational inferences based on the initial beliefs of others, they are also able to make rational inferences based on the manner in which others adjust their beliefs after hearing each successive round of announcements. This requires that individuals assume that beliefs are as described in the model, and assume that all individuals assume that beliefs are as described in the model, and update their beliefs accordingly … up to high orders. When such strong assumptions fail, individuals may fail to aggregate distributed information fully, and behavior may resemble the case of independent priors, where individuals do not make inferences based on the manner in which others react to information.

[13] Within a special class of models that are closed at the first order, Geanokoplos and Polimarchakis (1982) show also that the beliefs are aggregated at one round "generically."

VI. Social Structure

As an illustration of the theory developed in the previous sections, we now analyze the amount of bias between two groups under three alternative social structures, which we call fragmentation, integration, and segregation. Fragmentation corresponds to a structure in which no individual observes the prior of any other. Under integration, each individual observes the prior of every other individual. And under segregation, each individual observes the priors of all those belonging to the same group, but none of the priors of those in the other group.

More formally, let $N = B \cup W$, where B and W are disjoint sets with $n_b \geq 2$ and $n_w \geq 2$ members, respectively. We maintain the assumption that $\mathbf{\Sigma} = \sigma^2 \mathbf{I}$, so priors are independently distributed, and we assume that for some $\overline{\mu}_b > \overline{\mu}_w$,

$$\overline{\mu}_i = \overline{\mu}_b \quad \text{and} \quad \overline{\mu}_j = \overline{\mu}_w \qquad (\forall i \in B, j \in W).$$

That is, ex ante, members of B overestimate θ relative to members of W.[14] An individual member of B, of course, may turn out to have a higher expectation than an individual member of W once each observes his own prior. We assume that opinions are communicated by successive belief announcement as before and that all announcements are observable.[15] Define the average opinion within each group in period k as follows:

$$\hat{A}_{b,k} = \frac{1}{n_b} \sum_{i \in B} A_{i,k} \quad \text{and} \quad \hat{A}_{w,k} = \frac{1}{n_w} \sum_{j \in W} A_{j,k}.$$

We use the same superscript to denote other within-group averages as well, so $\hat{\mu}_b = \frac{1}{n_b} \sum_{i \in B} \mu_i$, $\hat{\varepsilon}_w = \frac{1}{n_w} \sum_{j \in W} \varepsilon_j$, etc. We are interested in the extent to which average opinion in B exceeds that in W at any given round k, defined as follows:

$$\beta_k \equiv \hat{A}_{b,k} - \hat{A}_{w,k}.$$

We let β_k^F, β_k^I and β_k^S denote the values of this difference under fragmentation, integration and segregation respectively.

We shall refer to beliefs at $k = 2$ as intermediate beliefs, and those at $k = 3$ as limiting beliefs. From the previous section, recall that in both fragmented and integrated societies, $A_{i,k} = A_{i,2}$ for all $k \geq 2$, and the distinction between intermediate and limiting beliefs is not meaningful. However, as we show below, the distinction is important under segregation, since individuals behave as in the correlated priors case (although the priors are in fact independent).

[14] This assumption is without loss of generality even if the groups are of unequal size, because if $\overline{\mu}_b < \overline{\mu}_w$, then we can simply reverse the order on θ by considering $-\theta$. Simply put, we are measuring the biases in the direction that, ex ante, members of B overestimate with respect to the members of W.

[15] The results in this section hold without modification even if only the average announcement in each group is publicly observable. To see this, note from (5) that when priors are observable, the public belief of i depends only on his own prior and the aggregate signal. Similarly, when priors are unobservable, the public belief of i depends only on his own initial announcement and the aggregate announcement in the group from (B2) in the proof of Lemma 1.

A. *Fragmentation*

In a fragmented society, individuals obtain information, form beliefs, communicate these beliefs to pollsters, and observe the aggregate belief distribution. No individual observes the prior belief of any other individual. Instead, he uses his prior belief about the thinking of the others in order to extract the information revealed in the polls. This is the case of unobservable priors.

From Lemma 1, for any round $k \geq 2$, the difference in average opinions across groups is

$$\beta_k^F = \frac{\tau^2}{\gamma_n}(\overline{\mu}_b - \overline{\mu}_w) + \frac{\tau^4\sigma^2}{\gamma_n}(\hat{\mu}_b - \hat{\mu}_w) + \frac{\tau^2\sigma^2}{\gamma_n}(\hat{\varepsilon}_b - \hat{\varepsilon}_w). \tag{13}$$

Hence, the bias has three sources: the ex ante bias between groups $(\overline{\mu}_b - \overline{\mu}_w)$, the average prior bias between groups $(\hat{\mu}_b - \hat{\mu}_w)$, and the average informational difference between groups $(\hat{\varepsilon}_b - \hat{\varepsilon}_w)$. Recalling the definition of γ_n from Lemma 1, the expected value of between-group bias is therefore

$$E\left[\beta_k^F\right] = \frac{\tau^2(1 + \tau^2\sigma^2)}{(1 + \tau^2)(1 + \tau^2\sigma^2) + n - 1}(\overline{\mu}_b - \overline{\mu}_w). \tag{14}$$

B. *Integration*

In an integrated society, each individual observes the priors of every other individual. They communicate directly, understanding the manner in which information is incorporated into beliefs. This is the case of observable priors.

From (6), for any round $k \geq 2$, the difference in average opinions across groups is

$$\beta_k^I = \frac{\tau^2}{\tau^2 + n}(\hat{\mu}_b - \hat{\mu}_w). \tag{15}$$

Hence, the difference across groups in average opinion is the difference between their respective average priors, scaled down by a factor that uses all of the distributed information efficiently. The expected value of this is

$$E\left[\beta_k^I\right] = \frac{\tau^2}{\tau^2 + n}(\overline{\mu}_b - \overline{\mu}_w) \tag{16}$$

and hence, from Proposition 1,

$$E\left[\beta_k^F\right] > E\left[\beta_k^I\right].$$

C. *Segregation*

Now we consider a segregated society partitioned into two components, one for each group. Each component is like an integrated society that is closed to members of the other component; individuals in different groups receive information about each other only through opinion polls. Formally, we assume that the prior of an individual is observable to the members of his own group and unobservable to the members of other group. That is, for each $i \in B$ and $j \in W$, μ_i is common knowledge among B and μ_j is common knowledge among W.

Now, when any $i \in B$ observes the first round announcements of his own group, he extracts all of the relevant information that other members of B have, concluding correctly that

$$\hat{x}_b \equiv \frac{1}{n_b} \sum_{i \in B} x_i = (1 + \tau^2)\hat{A}_{b,1} - \tau^2 \hat{\mu}_b. \tag{17}$$

On the other hand, he can extract only limited information from the announcements of the other group. The only relevant information for him is $(1 + \tau^2)\hat{A}_{w,1} = \hat{x}_w + \tau^2 \hat{\mu}_w$, where he knows neither $\hat{x}_w$ nor $\hat{\mu}_w$. Combining these two pieces of information, he updates his belief, and in the second round, he announces

$$A^S_{i,2} = c_b\big(\alpha_b \mu_i + (1 - \alpha_b)\hat{x}_b\big) + (1 - c_b)\big((1 + \tau^2)\hat{A}_{w,1} - \tau^2 \overline{\mu}_w\big) \quad (i \in B)$$

where

$$\alpha_b = \frac{\tau^2}{\tau^2 + n_b} \tag{18}$$

and

$$c_b = \frac{(\tau^2 + n_b)(1 + \tau^2\sigma^2)}{(1 + \tau^2\sigma^2)(\tau^2 + n_b) + n_w}. \tag{19}$$

Hence the average opinion in B at this stage is

$$\hat{A}^S_{b,2} = c_b\big(\alpha_b \hat{\mu}_b + (1 - \alpha_b)\hat{x}_b\big) + (1 - c_b)\big((1 + \tau^2)\hat{A}_{w,1} - \tau^2 \overline{\mu}_w\big). \tag{20}$$

It turns out that, together with the first round announcements, the second round announcements reveal all relevant information. To see this, consider any $j \in W$. From the average first round announcements of the other group, j deduces that $(1 + \tau^2)\hat{A}_{b,1} = \hat{x}_b + \tau^2\hat{\mu}_b$, and in the second round deduces (20). Since $n_b > 1$, j can solve these two independent linear equations, thereby computing $\hat{x}_b$ and $\hat{\mu}_b$. That is, j does not need to know how members of B think: knowing that members of B know how each other thinks, j can infer all relevant information from the manner in which members of B react to each others' announcements. As a result, by the end of the second round, all distributed information is aggregated, and in the limit, segregated and integrated societies are identical.

PROPOSITION 4: *For each $i \in N$ and $k \geq 3$, $A^S_{i,k} = A^I_{i,k}$ and $\beta^S_k = \beta^I_k$.*

This illustrates the power of the argument behind Proposition 3. When some individuals have information about other individuals (through correlation in Proposition 3 and observation here), third parties can extract that information from the manner in which these individuals react to each other's announcements.

We now turn to intermediate beliefs in a segregated society. From (17) and (20), we obtain

$$(21)\qquad \hat{A}^S_{b,2} = (1 - \alpha_b c_b)\theta + \alpha_b c_b \hat{\mu}_b + (1 - c_b)\tau^2(\hat{\mu}_w - \overline{\mu}_w) + (1 - \alpha_b)c_b\hat{\varepsilon}_b + (1 - c_b)\hat{\varepsilon}_w.$$

Similarly,

$$(22)\qquad \hat{A}^S_{w,2} = (1 - \alpha_w c_w)\theta + \alpha_w c_w \hat{\mu}_w + (1 - c_w)\tau^2(\hat{\mu}_b - \overline{\mu}_b) + (1 - \alpha_w)c_w\hat{\varepsilon}_w + (1 - c_w)\hat{\varepsilon}_b,$$

where α_w and c_w are defined analogously to (18) and (19).

If $n_b = n_w$ then $\alpha_b c_b = \alpha_w c_w$. In that case, intermediate bias, $\beta^S_2 = \hat{A}^S_{b,2} - \hat{A}^S_{w,2}$, does not depend on θ, and all individuals have the same expectation:

$$E\left[\beta^S_2\right] = \frac{\tau^2(1 + \tau^2\sigma^2)}{(1 + \tau^2\sigma^2)(\tau^2 + n/2) + n/2}(\overline{\mu}_b - \overline{\mu}_w).$$

It is easily verified that for any $n > 2$,

$$E\left[\beta^I_2\right] < E\left[\beta^S_2\right] < E\left[\beta^F_2\right].$$

That is, when groups are of equal size, they agree about the value of intermediate bias under all three information structures, and the bias is greatest under fragmentation, least under integration, and intermediate under segregation.

When groups are of unequal size, however, the intermediate bias does depend on θ, and hence the members of different groups will have different expectations about it. Our next result establishes that, in a segregated society, ex ante, members of a minority group will expect a smaller intermediate bias than the members of a majority group. Despite this, it further establishes that they all agree that the expected intermediate bias under segregation is higher than that under integration, and lower than that under fragmentation.

PROPOSITION 5: *If $n_b < n_w$, then, for all $i \in B$ and $j \in W$,*

$$E\left[\beta^I_2\right] < E_i\left[\beta^S_2\right] < E_j\left[\beta^S_2\right] < E\left[\beta^F_2\right].$$

To gain some intuition for the finding that minorities expect lower levels of intermediate bias, consider a highly skewed population with a very large majority group and a very small minority. Then the minority expects the majority to approximately learn θ, and hence to converge to a belief that is close to the minority group prior. The minority therefore expects the initial bias to diminish substantially. In contrast, the majority expects the minority to learn very little, and hence to maintain beliefs that are distant from majority group priors, with little narrowing of the initial bias. Roughly speaking, the minority expects the majority to come around to their way of thinking, while the majority expects no such convergence.

In summary, expected biases are always highest under fragmentation. Expected biases are higher under segregation than under integration with respect to intermediate beliefs, but the two social structures are identical in the limit. This is intuitive, since individuals have the least ability to process information under fragmentation and the greatest ability to process information under integration.

D. *Large Societies*

We have so far compared the expected value of biases under three social structures for arbitrary values of the population size n. In large societies idiosyncratic differences cancel each other out and we can compare the magnitudes of *actual* biases under various social structures state by state. Doing so reveals that our analysis of expectations misses an interesting and potentially disturbing fact about intermediate beliefs: Segregation puts minorities at a disadvantage in processing public information and consequently results in biases even when groups are formed from ex ante identical individuals.

In order to compare biases in large societies (as in Section IV), we consider a family of models indexed by n, such that

$$\text{as} \quad n \to \infty, \quad \tau^2/n \to \overline{\tau}^2 \qquad \text{and} \qquad n_b/n \to r \tag{23}$$

for some $\overline{\tau}^2 > 0$ and $r \in (0, ½)$. That is, we adopt the convention that B is the minority group. In a large fragmented society, by (13), the bias is approximately as great as the ex ante bias:

$$\lim_{n\to\infty} \beta_k^F = \overline{\mu}_b - \overline{\mu}_w \qquad \text{almost surely, for all } k \geq 2. \tag{24}$$

By (15), in a large integrated society, the bias is smaller, to a degree that depends on the precision of the distributed information:

$$\lim_{n\to\infty} \beta_k^I = \frac{\overline{\tau}^2}{\overline{\tau}^2 + 1} (\overline{\mu}_b - \overline{\mu}_w) \qquad \text{almost surely, for all } k \geq 2. \tag{25}$$

In a large segregated society, the bias is identical to that under integration in the limit, as we have seen above:

$$\lim_{n\to\infty} \beta_k^S = \frac{\overline{\tau}^2}{\overline{\tau}^2 + 1} (\overline{\mu}_b - \overline{\mu}_w) \qquad \text{almost surely, for all } k \geq 3.$$

In the limit, both segregated and integrated societies use all available information efficiently.

At the intermediate stage, under segregation, information is not fully aggregated. This does not, however, mean that the magnitude of the bias lies strictly between the corresponding magnitudes under fragmentation and integration respectively. To see this, note from (21) and (22) that average group beliefs at the intermediate stage are given by:

$$\lim_{n\to\infty} \hat{A}^S_{b,2} = \frac{\overline{\tau}^2}{\overline{\tau}^2 + r}\overline{\mu}_b + \frac{r}{\overline{\tau}^2 + r}\theta \tag{26}$$

$$\lim_{n\to\infty} \hat{A}^S_{w,2} = \frac{\overline{\tau}^2}{\overline{\tau}^2 + 1 - r}\overline{\mu}_w + \frac{1 - r}{\overline{\tau}^2 + 1 - r}\theta. \tag{27}$$

Notice that neither group processes information as efficiently as in an integrated society. In effect, a representative member of the minority group faces a noisy signal with variance $\overline{\tau}^2/r$, and a representative member of the majority group faces a noisy signal with variance $\overline{\tau}^2/(1 - r)$. Under integration, each individual obtains a noisy signal with variance $\overline{\tau}$, which is clearly smaller than both $\overline{\tau}^2/r$ and $\overline{\tau}^2/(1 - r)$. Furthermore, under segregation, minorities are disadvantaged in processing public information, since $\overline{\tau}^2/r > \overline{\tau}^2/(1 - r)$. As a result, at the intermediate stage, the majority belief puts greater weight on the true state (and less weight on the prior) when compared with the minority group belief. This disadvantage becomes more pronounced as group sizes become more unequal.[16]

Note that the intermediate stage bias under segregation depends on θ:

$$\lim_{n\to\infty} \beta^S_2 = \frac{\overline{\tau}^2}{\overline{\tau}^2 + r}\overline{\mu}_b - \frac{\overline{\tau}^2}{\overline{\tau}^2 + 1 - r}\overline{\mu}_w - \left(\frac{\overline{\tau}^2}{\overline{\tau}^2 + r} - \frac{\overline{\tau}^2}{\overline{\tau}^2 + 1 - r}\right)\theta \tag{28}$$

almost surely. Because of this dependence, the bias can take any value. In particular, at the intermediate stage, the difference in beliefs under segregation may *increase* (relative to the ex ante belief difference) and therefore exceed the difference under fragmentation. This will occur if θ turns out to be very different from ex ante expectations of it.

An interesting special case arises when the groups have identical ex ante beliefs: $\overline{\mu}_b = \overline{\mu}_w = \overline{\mu} \neq \theta$ for some $\overline{\mu}$. That is, the two groups start out with identical priors, and the true state happens to be different from the priors. Then, from (24) and (25), the intermediate and limiting biases are both negligible under fragmentation and integration: $\lim_{n\to\infty}\beta^F_k = \lim_{n\to\infty}\beta^I_k = 0$ almost surely for $k \geq 2$. However, from (28), the intermediate stage bias under segregation is strictly positive:

$$\lim_{n\to\infty} \beta^S_2 = \left(\frac{\overline{\tau}^2}{\overline{\tau}^2 + r} - \frac{\overline{\tau}^2}{\overline{\tau}^2 + 1 - r}\right)(\overline{\mu} - \theta) \qquad \text{almost surely.}$$

[16] Individuals belonging to a minority within any population tend to have a smaller number of affiliates in friendship networks (Currarini, Jackson, and Pin 2009), which should reinforce this effect. On the other hand, segregation itself tends to be endogenously increasing in the size of the minority group (Sethi and Somanathan 2004), which suggests that the extent of public disagreement may not vary monotonically with the size of the minority group.

Furthermore, from (26–27), the majority group belief at the intermediate stage is closer to the true state. These two facts may be stated as follows.

PROPOSITION 6: *Suppose that $\overline{\mu}_b = \overline{\mu}_w \neq \theta$. Then,*

$$\lim_{n\to\infty} |\hat{A}^S_{b,2} - \theta| > \lim_{n\to\infty} |\hat{A}^S_{w,2} - \theta|$$

and

$$\lim_{n\to\infty} |\beta^S_2| > \lim_{n\to\infty} |\beta^F_k| = \lim_{n\to\infty} |\beta^I_k| = 0.$$

To summarize, when the two groups are composed of ex ante identical individuals, the beliefs of the majority group are more closely aligned with reality than are the beliefs of the minority group at the intermediate stage. Also, the level of bias in intermediate stage beliefs is greater under segregation than under either integration or fragmentation. The former result arises directly from the fact that majority group members have an advantage in the interpretation of public information. The latter result arises because segregation tends to homogenize beliefs *within* groups, which has the effect of creating belief heterogeneity *across* groups. This effect does not arise under either fragmentation or integration.

VII. Conclusions

If a group of individuals share a common prior and are commonly known to be Bayesian (in the sense that each member of the group forms beliefs using Bayes' rule according to the common prior) then public disagreement cannot arise. Accounting for such disagreement therefore requires a departure from one or both of these hypotheses. We have chosen here to explore the implications of heterogeneous priors, while maintaining stringent assumptions regarding Bayesian rationality. Two main results follow from this. First, we find that for generic values of the model's primitives, the extent of public disagreement is independent of whether or not priors are observable, and public beliefs involve the aggregation of all distributed information in the limit. Second, we find that when priors are uncorrelated, the expected value of public bias is lower in an integrated society than in a fragmented one. For large societies, a stronger result holds: public bias is greater in a fragmented society relative to an integrated one under almost all realizations of priors and information. This suggests that social integration (in the sense of better understanding of the priors of others) should result in diminished public disagreement, especially in large populations.

Our results depend on the ability of individuals to make highly sophisticated statistical inferences, based not only on the initial beliefs of others but also on the manner in which these beliefs are adjusted over time on the basis of earlier announcements. If cognitive limitations prevent individuals from making inferences based on the manner in which one person responds to another's announcement, then our intermediate stage analysis applies, and the expected value of bias across social groups depends systematically on the extent of social integration. Expected bias is smallest in integrated societies (where priors are observable both within and between social groups) and largest in fragmented societies (where priors are unobservable even within social groups). Intermediate levels of expected bias arise under segregation, when priors are observable within but not across groups. Hence integration both within and across social groups tends to reduce expected levels of public bias.

In large populations, realized biases may be greater under segregation than under either fragmentation or integration, as belief differences are compressed within groups but amplified across groups. Communication in segregated societies can cause initial biases to be amplified, and new biases to emerge where none previously existed. Despite the fact that all announcements are public and all signals equally precise, members of a minority group face a disadvantage in the interpretation of public information that results in beliefs that are less closely aligned with the true state. If majority group members (or outside observers) fail to appreciate this effect, they may regard the views of minorities as "bizarre" or "outlandish," attributing them to failures in reasoning rather than to structural factors such as the demographic composition and constraints on information exchange induced by the heterogeneity and unobservability of prior beliefs.

Mathematical Appendix

A. *Aggregation of Distributed Information*

In this subsection, we prove Proposition 3. The proof requires the use of the following well-known formula. For any two random vectors $\mathbf{X}$ and $\mathbf{Y}$, if

$$\begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} \sim N\left(\begin{pmatrix} \boldsymbol{\mu}_X \\ \boldsymbol{\mu}_Y \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Sigma}_X & \boldsymbol{\Sigma}_{X,Y} \\ \boldsymbol{\Sigma}_{Y,X} & \boldsymbol{\Sigma}_Y \end{pmatrix}\right),$$

then conditional on $\mathbf{Y}$, $\mathbf{X}$ is distributed with $N(E[\mathbf{X}\,|\,\mathbf{Y}], \mathrm{Var}(\mathbf{X}\,|\,\mathbf{Y}))$ where

$$E[\mathbf{X}\,|\,\mathbf{Y}] = \boldsymbol{\mu}_X + \boldsymbol{\Sigma}_{X,Y}\boldsymbol{\Sigma}_Y^{-1}(\mathbf{Y} - \boldsymbol{\mu}_Y) \tag{A1}$$

$$\mathrm{Var}(\mathbf{X}\,|\,\mathbf{Y}) = \boldsymbol{\Sigma}_X - \boldsymbol{\Sigma}_{X,Y}\boldsymbol{\Sigma}_Y^{-1}\boldsymbol{\Sigma}_{Y,X}.$$

We also need to introduce some more notation. For any subset $N' \subset N$, we use subscript N' to denote the column vector obtained by stacking up all the values for $j \in N'$. For example, we write $\boldsymbol{\mu}_{N'} = (\mu_j)_{j\in N'}$, $\mathbf{A}_{N',k} = (A_{j,k})_{j\in N'}$, and $\boldsymbol{\sigma}_{N',i} = (\sigma_{j,i})_{j\in N'}$. For any subsets N' and N'' of N and any matrix $\mathbf{X} = (x_{i,j})_{i,j\in N}$, we write $\mathbf{X}_{N',N''}$ for the submatrix with entries from N' and N'', i.e., $\mathbf{X}_{N',N''} = (x_{i,j})_{i\in N',j\in N''}$. We use subscript $-i$ instead of $N\backslash\{i\}$, e.g., $\boldsymbol{\mu}_{-i} = (\mu_j)_{j\neq i}$ and $\boldsymbol{\Sigma}_{-i,-i} = (\sigma_{j,k})_{j\neq i,k\neq i}$. We write $\mathbf{1}_{k\times l}$ for the $k \times l$-dimensional matrix with entries 1 and $\mathbf{I}$ for the identity matrix. We write

$$\tilde{\boldsymbol{\mu}}_{-i} \equiv E[\boldsymbol{\mu}_{-i}\,|\,\mu_i] = \overline{\boldsymbol{\mu}}_{-i} + \sigma_{ii}^{-1}\boldsymbol{\sigma}_{-i,i}(\mu_i - \overline{\mu}_i) \tag{A2}$$

$$\tilde{\boldsymbol{\Sigma}}_{-i,-i} \equiv \mathrm{Var}(\boldsymbol{\mu}_{-i}\,|\,\mu_i) = \boldsymbol{\Sigma}_{-i,-i} - \sigma_{ii}^{-1}\boldsymbol{\sigma}_{-i,i}\boldsymbol{\sigma}'_{-i,i}.$$

Using the definitions of R and H in Lemma 2 below, we also write

$$\text{(A3)}\quad \hat{v} = \tau^2/(\tau^2 + 1 + |R|),$$

$$\alpha = \tau^2/(\tau^2 + 1),$$

$$\mathbf{M}_R = 1_{1\times|H|}\left(\hat{v}\mathbf{1}_{|H|\times|H|} + \tau^2\mathbf{I} + \tau^4\tilde{\mathbf{\Sigma}}_{H,H} - \tau^4\tilde{\mathbf{\Sigma}}_{H,R}\tilde{\mathbf{\Sigma}}_{R,R}^{-1}\tilde{\mathbf{\Sigma}}_{R,H}\right)^{-1}$$

$$\mathbf{M}_i = \mathbf{1}_{1\times n-1}\left(\alpha\mathbf{1}_{n-1\times n-1} + \tau^2\mathbf{I} + \tau^4\mathbf{\Sigma}_{-i,-i} - \tau^4\sigma_{ii}^{-1}\boldsymbol{\sigma}_{-i,i}\boldsymbol{\sigma}'_{-i,i}\right)^{-1}.$$

We compute the announcements in the following lemma.

LEMMA 2: *Assume that the priors are not observable. For any $i \in N$ and any round k, let $R \subseteq N\backslash\{i\}$ be the set of other individuals whose private information is revealed by the end of round $k - 1$, and let $H = N\backslash(R \cup \{i\})$. Then,*

$$\text{(A4)}\quad A^u_{i,k} = \frac{\tau^2 + 1}{\tau^2 + 1 + |R|}(1 - \hat{v}\mathbf{M}_R\mathbf{1}_{|H|\times 1})\sum_{j\in R\cup\{i\}} A_{j,1} + (1 + \tau^2)\hat{v}\mathbf{M}_R A_{H,1}$$

$$- \frac{\tau^2}{\tau^2 + 1 + |R|}\mathbf{1}_{1\times|R|}\mu_R - \tau^2\hat{v}\mathbf{M}_R\left(\overline{\mu}_H - \tilde{\mathbf{\Sigma}}_{H,R}\tilde{\mathbf{\Sigma}}_{R,R}^{-1}(\mu_R - \overline{\mu}_R)\right)$$

$$- \tau^2\sigma_{ii}^{-1}\hat{v}\mathbf{M}_R\left(\sigma_{H,i} - \tilde{\mathbf{\Sigma}}_{H,R}\tilde{\mathbf{\Sigma}}_{R,R}^{-1}\sigma_{R,i}\right)(\mu_i - \overline{\mu}_i)$$

when $R \neq \emptyset$ and

$$\text{(A5)}\quad A^u_{i,k} = (1 - \alpha\mathbf{M}_i\mathbf{1}_{n-1\times 1})A_{i,1} + \tau^2\mathbf{M}_i\mathbf{A}_{-i,1} - \tau^2\alpha\mathbf{M}_i\overline{\boldsymbol{\mu}}_{-i}$$

$$- \tau^2\sigma_{ii}^{-1}\alpha\mathbf{M}_i\boldsymbol{\sigma}_{-i,i}(\mu_i - \overline{\mu}_i)$$

when $R = \emptyset$.

PROOF:

We will use mathematical induction on k. We first compute $A^u_{i,2}$, showing that the statement is true for $k = 2$. For each j, since $A_{j,1} = \alpha\mu_j + (1 - \alpha)x_j$,

$$(1 + \tau^2)A_{j,1} = \theta + \varepsilon_j + \tau^2\mu_j.$$

Hence,

$$E_i[(1 + \tau^2)A_{j,1} \,|\, \mu_i, x_i] = A_{i,1} + \tau^2 E_i[\mu_j \,|\, \mu_i].$$

Substituting (A2) in this equality, we obtain

$$\text{(A6)}\quad E_i[(1+\tau^2)\mathbf{A}_{-i,1}\,|\,\mu_i,x_i] = \mathbf{1}_{n-1\times 1}A_{i,1} + \tau^2\overline{\boldsymbol{\mu}}_{-i} + \tau^2\sigma_{ii}^{-1}\boldsymbol{\sigma}_{-i,i}(\mu_i - \overline{\mu}_i).$$

Now, the first round of announcements provides i a new vector $(1+\tau^2)\mathbf{A}_{-i,1} = \theta\mathbf{1}_{n-1\times 1} + \boldsymbol{\varepsilon}_{-i} + \tau^2\boldsymbol{\mu}_{-i}$ of signals with additive normal noise. Notice that, conditional on (x_i,μ_i), the variance of $\theta\mathbf{1}_{n-1\times 1} + \boldsymbol{\varepsilon}_{-i} + \tau^2\boldsymbol{\mu}_{-i}$ is

$$\alpha\mathbf{1}_{n-1\times n-1} + \tau^2\mathbf{I} + \tau^4(\boldsymbol{\Sigma}_{-i} - \sigma_{ii}^{-1}\boldsymbol{\sigma}_{-i,i}\boldsymbol{\sigma}'_{-i,i}).$$

Hence, updating his belief according to (A1), in the second round i announces

$$\begin{aligned}\text{(A7)}\qquad A^u_{i,2} &= E_i[\theta\,|\,\mu_i,x_i,(1+\tau^2)\mathbf{A}_{-i,1}]\\ &= A_{i,1} + \alpha\mathbf{M}_i\big((1+\tau^2)\mathbf{A}_{-i,1} - E_i[(1+\tau^2)\mathbf{A}_{-i,1}\,|\,\mu_i,x_i]\big)\\ &= (1-\alpha\mathbf{M}_i\mathbf{1}_{n-1\times 1})A_{i,1} + \tau^2\mathbf{M}_i\mathbf{A}_{-i,1} - \alpha\tau^2\mathbf{M}_i\overline{\boldsymbol{\mu}}_{-i}\\ &\quad - \tau^2\sigma_{ii}^{-1}\alpha\mathbf{M}_i\boldsymbol{\sigma}_{-i,i}(\mu_i - \overline{\mu}_i),\end{aligned}$$

where the second equality is by (A1) and the definition of $\mathbf{M}_i$, and the last equality is by (A6). Now suppose that the proposition is true for rounds $k' \leq k-1$ and for all j. Then, if

$$\mathbf{M}_R\big(\sigma_{H,j} - \tilde{\boldsymbol{\Sigma}}_{H,R}\tilde{\boldsymbol{\Sigma}}_{R,R}^{-1}\sigma_{R,j}\big) = 0$$

for R defined for k' and j, no new information is revealed by the announcement $A^u_{j,k'}$ because it is measurable with respect to the public information at the end of round $k'-1$. On the other hand, if

$$\mathbf{M}_R(\sigma_{H,j} - \tilde{\boldsymbol{\Sigma}}_{H,R}\tilde{\boldsymbol{\Sigma}}_{R,R}^{-1}\sigma_{R,j}) \neq 0,$$

then we can solve for μ_j from (A4) for k' and j. That is, either the private information of j is revealed by the end of round $k-1$, i.e., $j \in R$, or i knows only that $A_{j,1} = \alpha\mu_j + (1-\alpha)x_j$. Now, if $R = \emptyset$, i has not learned any new information after the first round. In that case, $A^u_{i,k} = A^u_{i,2}$, and (A5) is equivalent to (A7). Now suppose that $R \neq \emptyset$. Individual i knows (μ_i,x_i), (μ_j,x_j) for $j \in R$ and that $A_{j,1} = \alpha\mu_j + (1-\alpha)x_j$ for $j \notin R$. We compute conditional distributions sequentially, first conditioning on (μ_i,x_i), then on (μ_R,x_R), and finally on $A_{H,1} = \alpha\mu_H + (1-\alpha)x_H$, i.e.,

$$\text{(A8)}\qquad (1+\tau^2)A_{H,1} = \mathbf{1}_{|H|\times 1}\theta + \varepsilon_H + \tau^2\mu_H.$$

Conditional on (μ_i,x_i), $(\theta,\boldsymbol{\mu}_{-i},\boldsymbol{\varepsilon}_{-i})$ are independently and normally distributed with $\theta \sim N(A_{i,1},\alpha)$, $\boldsymbol{\mu}_{-i} \sim N(\tilde{\boldsymbol{\mu}}_{-i},\tilde{\boldsymbol{\Sigma}}_{-i,-i})$, and $\boldsymbol{\varepsilon}_{-i} \sim N(0,\tau^2\mathbf{I})$. Then, from (μ_R,x_R), he obtains a new signal $x_R = \mathbf{1}_{|R|\times 1}\theta + \varepsilon_R$ about θ and also potentially new information about μ_H from μ_R. Conditioning on $x_R = \mathbf{1}_{|R|\times 1}\theta + \varepsilon_R$, he updates his belief about θ to $N(\hat{\mu}_i,\hat{v})$ where

$$\hat{\mu}_i = \frac{\tau^2 + 1}{\tau^2 + 1 + |R|} A_{i,1} + \frac{1}{\tau^2 + 1 + |R|} \mathbf{1}_{1\times|R|} x_R$$

$$= \frac{\tau^2 + 1}{\tau^2 + 1 + |R|} \sum_{j \in R \cup \{i\}} A_{j,1} - \frac{\tau^2}{\tau^2 + 1 + |R|} \mathbf{1}_{1\times|R|} \mu_R$$

$$\hat{v} = \frac{\tau^2}{\tau^2 + 1 + |R|}.$$

Conditioning on μ_R, he updates his belief about μ_H to $N(\hat{\mu}_H, \hat{\boldsymbol{\Sigma}}_H)$ where

$$\hat{\mu}_H = \tilde{\mu}_H + \tilde{\boldsymbol{\Sigma}}_{H,R} \tilde{\boldsymbol{\Sigma}}_{R,R}^{-1} (\mu_R - \tilde{\mu}_R)$$

$$\hat{\boldsymbol{\Sigma}}_H = \tilde{\boldsymbol{\Sigma}}_{H,H} - \tilde{\boldsymbol{\Sigma}}_{H,R} \tilde{\boldsymbol{\Sigma}}_{R,R}^{-1} \tilde{\boldsymbol{\Sigma}}_{R,H}.$$

Now, i conditions on (A8) starting from $\theta \sim N(\hat{\mu}_i, \hat{v})$. Given the conditionings so far, by (A8),

$$(1 + \tau^2) A_{H,1} \sim N(\hat{\mu}_i \mathbf{1}_{|H|\times 1} + \tau^2 \hat{\mu}_H, \hat{v}\mathbf{1}_{|H|\times|H|} + \tau^4 \hat{\boldsymbol{\Sigma}}_H + \tau^2 \mathbf{I}).$$

Using (A1), he therefore obtains

$$A_{i,k} = E\left[\theta \mid \mu_i, x_i, \mu_R, x_R, (1 + \tau^2) A_{H,1} = \mathbf{1}_{|H|\times 1}\theta + \varepsilon_H + \tau^2 \mu_H\right]$$

$$= \hat{\mu}_i + \hat{v}\mathbf{1}_{1\times|H|}(\hat{v}\mathbf{1}_{|H|\times|H|} + \tau^4 \hat{\boldsymbol{\Sigma}}_H + \tau^2 \mathbf{I})^{-1}$$

$$\times \left((1 + \tau^2) A_{H,1} - \hat{\mu}_i \mathbf{1}_{|H|\times 1} - \tau^2 \hat{\mu}_H\right)$$

$$= (1 - \hat{v}\mathbf{M}_R \mathbf{1}_{|H|\times 1})\hat{\mu}_i + (1 + \tau^2)\hat{v}\mathbf{M}_R A_{H,1} - \tau^2 \hat{v}\mathbf{M}_R \hat{\mu}_H$$

$$= \frac{\tau^2 + 1}{\tau^2 + 1 + |R|}(1 - \hat{v}\mathbf{M}_R \mathbf{1}_{|H|\times 1}) \sum_{j \in R \cup \{i\}} A_{j,1}$$

$$- \frac{\tau^2}{\tau^2 + 1 + |R|}(1 - \hat{v}\mathbf{M}_R \mathbf{1}_{|H|\times 1}) \mathbf{1}_{1\times|R|} \mu_R$$

$$+ (1 + \tau^2)\hat{v}\mathbf{M}_R A_{H,1}$$

$$- \tau^2 \hat{v}\mathbf{M}_R \big(\overline{\mu}_H + \sigma_{ii}^{-1}\sigma_{H,i}(\mu_i - \overline{\mu}_i) + \tilde{\boldsymbol{\Sigma}}_{H,R}\tilde{\boldsymbol{\Sigma}}_{R,R}^{-1}$$

$$\times \left(\mu_R - \overline{\mu}_R - \sigma_{ii}^{-1}\sigma_{R,i} \times (\mu_i - \overline{\mu}_i)\right)\big),$$

where the second equality is by (A1); the third is by arrangement of terms using the definition of $\mathbf{M}_R$, and the last by substituting the values of $\hat{\mu}_i$ and $\hat{\mu}_H$. By rearranging terms, we obtain the equality in the proposition.

Using Lemma 2, we can now prove Proposition 3.

PROOF OF PROPOSITION 3:

Assume first that i is regular, i.e., $\mathbf{M}_i \boldsymbol{\sigma}_{-i,i} \neq 0$. Then, since no individual's private information is revealed by the end of round 2, by (A5),

$$\mu_i = \overline{\mu}_i + \frac{(1 - \alpha \mathbf{M}_i \mathbf{1}_{n-1\times 1}) A_{i,1} + \tau^2 \mathbf{M}_i \mathbf{A}_{-i,1} - \tau^2 \alpha \mathbf{M}_i \overline{\boldsymbol{\mu}}_{-i} - A_{i,2}}{\alpha \tau^2 \sigma_{ii}^{-1} \mathbf{M}_i \boldsymbol{\sigma}_{-i,i}},$$

i.e., μ_i is measurable with respect to $A_{i,1}$, $\mathbf{A}_{-i,1}$, and $A_{i,2}$. Moreover, since $A_{i,1} = \alpha \mu_i + (1 - \alpha_i) x_i$, we can further compute that

$$x_i = (1 + \tau^2) A_{i,1} - \tau^2 \times \left(\overline{\mu}_i + \frac{(1 - \alpha \mathbf{M}_i \mathbf{1}_{n-1\times 1}) A_{i,1} + \tau^2 \mathbf{M}_i \mathbf{A}_{-i,1} - \tau^2 \alpha \mathbf{M}_i \overline{\boldsymbol{\mu}}_{-i} - A_{i,2}}{\alpha \tau^2 \sigma_{ii}^{-1} \mathbf{M}_i \boldsymbol{\sigma}_{-i,i}} \right),$$

showing that x_i is measurable with respect to $A_{i,1}$, $\mathbf{A}_{-i,1}$, and $A_{i,2}$. Therefore, the private information of i is revealed by round 2. Conversely, suppose that i is isolated, i.e., $\boldsymbol{\sigma}_{-i,i} = 0$. (Note that, in that case, $\overline{\boldsymbol{\mu}}_{-i} = \boldsymbol{\mu}_{-i}$ and $\tilde{\boldsymbol{\Sigma}}_{-i,-i} = \boldsymbol{\Sigma}_{-i,-i}$). Hence, by Lemma 2, for any $k > 1$, if $R \neq \emptyset$, then the coefficient of μ_i is

$$-\tau^2 \sigma_{ii}^{-1} \hat{v} \mathbf{M}_R (\sigma_{H,i} - \boldsymbol{\Sigma}_{H,R} \boldsymbol{\Sigma}_{R,R}^{-1} \sigma_{R,i}) = 0$$

because $\sigma_{H,i} = 0$ and $\sigma_{R,i} = 0$. If $R = \emptyset$, the coefficient is again $\tau^2 \sigma_{ii}^{-1} \alpha \mathbf{M}_i \boldsymbol{\sigma}_{-i,i} = 0$. Thus, $A_{i,k}$ is measurable with respect to the information at the end of round $k - 1$, revealing no new information. On the other hand, since $\boldsymbol{\sigma}_{-i,i} = 0$, (x_R, μ_R) does not provide any information about μ_i, either. It only reduces the variance of x_i without revealing it. Hence, the private information of i is not revealed at any round.

B. *Public Bias*

PROOF OF LEMMA 1:

By Proposition 3, since the priors are independent, no information is revealed. Hence, by Lemma 2, $A^u_{i,\infty} = A^u_{i,2}$, and $A^u_{i,2}$ satisfies (A5). To compute $A^u_{i,2}$ from (A5), first define

$$\varphi = (1 + \tau^2)(1 + \tau^2 \sigma^2)$$

and note that

$$\text{(B1)} \qquad \begin{aligned} \mathbf{M}_i &= \mathbf{1}_{1\times n-1}\big(\alpha \mathbf{1}_{n-1\times n-1} + (\tau^2 + \tau^4\sigma^2)\mathbf{I}\big)^{-1} \\ &= \alpha^{-1}\mathbf{1}_{1\times n-1}(\mathbf{1}_{n-1\times n-1} + \varphi \mathbf{I})^{-1} \\ &= \frac{1}{\alpha\varphi(\varphi + n - 1)}\mathbf{1}_{1\times n-1}\big((\varphi + n - 1)\mathbf{I} - \mathbf{1}_{n-1\times n-1}\big) \\ &= \frac{1}{\alpha(\varphi + n - 1)}\mathbf{1}_{1\times n-1}. \end{aligned}$$

Here, the first equality is obtained by substituting $\boldsymbol{\Sigma} = \sigma^2\mathbf{I}$ in (A3), and the second equality is by simple algebra. In the third equality, we invert the matrix $\mathbf{1}_{n-1\times n-1} + \varphi\mathbf{I}$. It can be easily verified that

$$(\mathbf{1}_{n-1\times n-1} + \varphi\mathbf{I})^{-1} = \frac{1}{\varphi(\varphi + n - 1)}\big((\varphi + n - 1) - \mathbf{1}_{n-1\times n-1}\mathbf{I}\big),$$

yielding the third line. Finally, by adding up the rows of the matrix $((\varphi + n - 1)\mathbf{I} - \mathbf{1}_{n-1\times n-1})$, we obtain (A9). Substituting (A9) in (A5), we then obtain

$$\begin{aligned} A^u_{i,2} &= (1 - \alpha\mathbf{M}_i\mathbf{1}_{n-1\times 1})A_{i,1} + \tau^2\mathbf{M}_i\mathbf{A}_{-i,1} - \alpha\tau^2\mathbf{M}_i\overline{\boldsymbol{\mu}}_{-i} \\ &= \Big(1 - \frac{\mathbf{1}_{1\times n-1}\mathbf{1}_{n-1\times 1}}{\varphi + n - 1}\Big)A_{i,1} + \frac{\tau^2}{\alpha(\varphi + n - 1)}\mathbf{1}_{1\times n-1}\mathbf{A}_{-i,1} \\ &\quad - \frac{\tau^2}{\varphi + n - 1}\mathbf{1}_{1\times n-1}\overline{\boldsymbol{\mu}}_{-i} \\ &= \frac{1}{\varphi + n - 1}A_{i,1} + \frac{1 + \tau^2}{\varphi + n - 1}\sum_{j\neq i} A_{j,1} \\ &\quad - \frac{\tau^2}{\varphi + n - 1}\sum_{j\neq i} A_{j,1}\overline{\mu}_j. \end{aligned}$$

Here, the first equality is simply (A5) for $\boldsymbol{\sigma}_{-i,i} = 0$, and the second equality is just by the substitution of the value of $\mathbf{M}_i$ from (B1). The last equality is by straightforward algebra. By adding and subtracting new terms with $A_{i,1}$ and μ_i, we obtain

$$\text{(B2)} \qquad \begin{aligned} A^u_{i,\infty} = A^u_{i,2} &= \frac{\varphi - (1 + \tau^2)}{\varphi + n - 1}A_{i,1} + \frac{1 + \tau^2}{\varphi + n - 1}\sum_{j=1}^{n} A_{j,1} \\ &\quad + \frac{\tau^2}{\varphi + n - 1}\overline{\mu}_i - \frac{\tau^2}{\varphi + n - 1}\sum_{j=1}^{n}\overline{\mu}_j. \end{aligned}$$

Terms with summations do not depend on i, and hence are cancelled out in the difference, yielding

$$\begin{aligned} A^u_{i,\infty} - A^u_{j,\infty} &= \frac{\varphi - (1 + \tau^2)}{\varphi + n - 1}(A_{i,1} - A_{j,1}) + \frac{\tau^2}{\varphi + n - 1}(\overline{\mu}_i - \overline{\mu}_j) \\ &= \frac{\tau^2\sigma^2(1 + \tau^2)}{\gamma_n}(A_{i,1} - A_{j,1}) + \frac{\tau^2}{\gamma_n}(\overline{\mu}_i - \overline{\mu}_j) \\ &= \frac{\tau^2\sigma^2}{\gamma_n}(\tau^2(\mu_i - \mu_j) + (x_i - x_j)) + \frac{\tau^2}{\gamma_n}(\overline{\mu}_i - \overline{\mu}_j) \\ &= \frac{\tau^4\sigma^2}{\gamma_n}(\mu_i - \mu_j) + \frac{\tau^2\sigma^2}{\gamma_n}(\varepsilon_i - \varepsilon_j) + \frac{\tau^2}{\gamma_n}(\overline{\mu}_i - \overline{\mu}_j). \end{aligned}$$

Here the second equality is by substitution of the definitions $\varphi = (1 + \tau^2) \times (1 + \tau^2\sigma^2)$ and $\gamma_n = \varphi + n - 1$; the third equality is by $(1 + \tau^2)A_{i,1} = \tau^2\mu_i + x_i$ and the last is by $x_i - x_j = \varepsilon_i - \varepsilon_j$.

PROOF OF PROPOSITION 1:

Note from (6) and (12) that

$$E\left[A^{ck}_{i,\infty} - A^{ck}_{j,\infty}\right] = \frac{\tau^2}{\tau^2 + n}(\overline{\mu}_i - \overline{\mu}_j)$$

$$E[A^u_{i,\infty} - A^u_{j,\infty}] = \frac{\tau^2(1 + \tau^2\sigma^2)}{(1 + \tau^2)(1 + \tau^2\sigma^2) + n - 1}(\overline{\mu}_i - \overline{\mu}_j).$$

$E\left[A^{ck}_{i,\infty} - A^{ck}_{j,\infty}\right]$ is independent of σ^2 while $E[A^u_{i,\infty} - A^u_{j,\infty}]$ is increasing in σ^2. Since $E[A^u_{i,\infty} - A^u_{j,\infty}] = E\left[A^{ck}_{i,\infty} - A^{ck}_{j,\infty}\right]$ for $\sigma^2 = 0$, we have $E[A^u_{i,\infty} - A^u_{j,\infty}] > E\left[A^{ck}_{i,\infty} - A^{ck}_{j,\infty}\right]$ for $\sigma^2 > 0$.

C. *Social Groups*

PROOF OF PROPOSITION 5:

For any $i \in B$, $E_i[\hat{\mu}_b] = E_i[\theta] = \overline{\mu}_b$ and $E_i[\hat{\mu}_w] = \overline{\mu}_w$. Hence, by (21) and (22),

$$\text{(C1)} \qquad E_i\left[\beta^S_2\right] = \alpha_w c_w(\overline{\mu}_b - \overline{\mu}_w).$$

Similarly, for any $j \in W$,

$$\text{(C2)} \qquad E_j\left[\beta^S_2\right] = \alpha_b c_b(\overline{\mu}_b - \overline{\mu}_w).$$

Now,

$$\alpha_b c_b = \frac{\tau^2(1 + \tau^2\sigma^2)}{\tau^2(1 + \tau^2\sigma^2) + n_b\tau^2\sigma^2 + n}$$

$$\alpha_w c_w = \frac{\tau^2(1 + \tau^2\sigma^2)}{\tau^2(1 + \tau^2\sigma^2) + n_w\tau^2\sigma^2 + n}.$$

Since $n_b < n_w$, we have $c_b\alpha_b > c_w\alpha_w$, showing that $E_i[\beta_2^S] < E_j[\beta_2^S]$. To see that $E[\beta_2^I] < E_i[\beta_2^S]$, observe that (16) can be obtained from (C1) by setting $\sigma^2 = 0$, and $\alpha_w c_w$ is increasing in σ^2. To see that $E_j[\beta_2^S] < E[\beta_2^I]$, observe that (14) can be obtained from (C2) by setting $n_b = 1$, and $\alpha_b c_b$ is decreasing in n_b.

REFERENCES

Acemoglu, Daron, Victor Chernozhukov, and Muhamet Yildiz. 2009. "Fragility of Asymptotic Agreement under Bayesian Learning." Massachusetts Institute of Technology (MIT) Department of Economics Working Paper 08-09.

Adrian, Tobias, and Mark M. Westerfield. 2009. "Disagreement and Learning in a Dynamic Contracting Model." *Review of Financial Studies* 22 (10): 3873–3906.

Allen, Michael O., Patrice O'Shaughnessy, and Dean Chang. 1995. "Tremor Hits Racial Divide." *New York Daily News*. http://www.nydailynews.com/archives/news/tremor-hits-racial-divide-article-1.693379. October 8.

Andreoni, James, and Tymofiy Mylovanov. 2012. "Diverging Opinions." *American Economic Journal: Microeconomics* 4 (1): 209–32.

Aumann, Robert J. 1976. "Agreeing to Disagree." *Annals of Statistics* 4 (6): 1236–39.

Benabou, Roland. 2009. "Groupthink: Collective Delusions in Organizations and Markets." National Bureau of Economic Research Working Paper 14764.

Benabou, Roland, and Jean Tirole. 2002. "Self-Confidence and Personal Motivation." *Quarterly Journal of Economics* 117 (3): 871–915.

Banerjee, Abhijit, and Rohini Somanathan. 2001. "A Simple Model of Voice." *Quarterly Journal of Economics* 116 (1): 189–227.

Brigham, John C., and Adina W. Wasserman. 1999. "The Impact of Race, Racial Attitude, and Gender on Reactions to the Criminal Trial of O. J. Simpson." *Journal of Applied Social Psychology* 29 (7): 1333–70.

Che, Yeon-Koo, and Navin Kartik. 2009. "Opinions as Incentives." *Journal of Political Economy* 117 (5): 815–60.

CNN/Opinion Research Inc. 2008. "Post-Election Race Relations." http://www.cnn.com/2008/POLITICS/11/11/obama.poll/.

Crawford, Vincent P., and Joel Sobel. 1982. "Strategic Information Transmission." *Econometrica* 50 (6): 1431–51.

Cripps, Martin W., Jeffrey C. Ely, George J. Mailath, and Larry Samuelson. 2008. "Common Learning." *Econometrica* 76 (4): 909–33.

Crocker, Jennifer, Riia Luhtanen, Stephanie Broadnax, and Bruce Evan Blaine. 1999. "Belief in US Government Conspiracies Against Blacks Among Black and White College Students: Powerlessness or System Blame?" *Personality and Social Psychology Bulletin* 25 (8): 941–53.

Currarini, Sergio, Matthew O. Jackson, and Paolo Pin. 2009. "An Economic Model of Friendship: Homophily, Minorities, and Segregation." *Econometrica* 77 (4): 1003–46.

DeMarzo, Peter M., Dimitri Vayanos, and Jeffrey Zwiebel. 2003. "Persuasion Bias, Social Influence, and Unidimensional Opinions." *Quarterly Journal of Economics* 118 (3): 909–68.

DeParle, Jason. 1990. "Talk of Government Being Out to Get Blacks Falls on More Attentive Ears." *New York Times*. http://www.nytimes.com/1990/10/29/nyregion/talk-of-government-being-out-to-get-blacks-falls-on-more-attentive-ears.html?pagewanted=all&src=pm. October 29.

Dixit, Avinash K., and Jörgen W. Weibull. 2007. "Political Polarization." *Proceedings of the National Academy of Sciences* 104 (18): 7351–56.

Eliaz, Kfir, and Ran Spiegler. 2007. "A Mechanism-Design Approach to Speculative Trade." *Econometrica* 75 (3): 875–84.
Freedman, David. 1965. "On the Asymptotic Behavior of Bayes Estimates in the Discrete Case II." *Annals of Mathematical Statistics* 36 (2): 454–56.
Geanakoplos, John D., and Heraklis M. Polemarchakis. 1982. "We Can't Disagree Forever." *Journal of Economic Theory* 28 (1): 192–200.
Gottlieb, Daniel. 2010. "Will You Never Learn? Self Deception and Biases in Information Processing." Unpublished.
Hafer, Catherine, and Dimitri Landa. 2007. "Deliberation as Self-Discovery and Institutions for Political Speech." *Journal of Theoretical Politics* 19 (3): 329–60.
Harrington, Joseph E., Jr. 1993. "Economic Policy, Economic Performance, and Elections." *American Economic Review* 83 (1): 27–42.
Harrison, J. Michael, and David M. Kreps. 1978. "Speculative Investor Behavior in a Stock Market with Heterogeneous Expectations." *Quarterly Journal of Economics* 92 (2): 323–36.
Morris, Stephen. 1994. "Trade with Heterogeneous Prior Beliefs and Asymmetric Information." *Econometrica* 62 (6): 1327–47.
Morris, Stephen. 1996. "Speculative Investor Behavior and Learning." *Quarterly Journal of Economics* 111 (4): 1111–33.
Ostrovsky, Michael. Forthcoming. "Information Aggregation in Dynamic Markets with Strategic Traders." *Econometrica*.
Pew Research Center for the People & the Press. 2008. "Twelve Percent Still Think Obama is Muslim." http://pewresearch.org/databank/dailynumber/?NumberID=509. (accessed October 27, 2008).
Rabin, Matthew, and Joel L. Schrag. 1999. "First Impressions Matter: A Model of Confirmatory Bias." *Quarterly Journal of Economics* 114 (1): 37–82.
Scheinkman, Jose A., and Wei Xiong. 2003. "Overconfidence and Speculative Bubbles." *Journal of Political Economy* 111 (6): 1183–1219.
Sethi, Rajiv, and Rohini Somanathan. 2004. "Inequality and Segregation." *Journal of Political Economy* 112 (6): 1296–1321.
Shmaya, Eran, and Leeat Yariv. 2008. "Experiments on Decisions Under Uncertainty: A Theoretical Framework." Unpublished.
Van den Steen, Eric. 2005. "Organizational Beliefs and Managerial Vision." *Journal of Law, Economics, and Organization* 21 (1): 256–83.
Van den Steen, Eric. 2010. "Culture Clash: The Costs and Benefits of Homogeneity." *Management Science* 56 (10): 1718–38.
Weigel, David. 2009. "How Many Southern Whites Believe Obama Was Born in America?" *Washington Independent*. http://washingtonindependent.com/53396/how-many-southern-whites-believe-obama-was-born-in-america.
Wilson, Andrea. 2003. "Bounded Memory and Biases in Information Processing." Unpublished.
Yildiz, Muhamet. 2003. "Bargaining without a Common Prior: An Immediate Agreement Theorem." *Econometrica* 71 (3): 793–811.
Yildiz, Muhamet. 2004. "Waiting to Persuade." *Quarterly Journal of Economics* 119 (1): 223–48.