

Social Norms and Social Change

Ethan Bueno de Mesquita¹ and Mehdi Shadmehr²

¹Harris School of Public Policy, University of Chicago, USA;
bdm@uchicago.edu

²Department of Public Policy, University of North Carolina at Chapel Hill,
USA; mshadmeh@gmail.com

ABSTRACT

We study how social norms affect social change in a setting where people have both internal motivations and a desire to conform. We distinguish two kinds of internal motivations: *common values* in which people wish to behave consistent with some evolving, uncertain ground truth and *private values* in which individuals genuinely disagree about proper behavior for non-informational reasons. In both settings aggregate behavior changes more slowly than beliefs about proper behavior, and increased information reduces such inertia. Inertia is a more severe problem, but information is more effective, when values are common rather than private. In this common-values setting, we identify conditions under which increased information leads to a normative improvement. Finally, we elucidate empirical implications for the relationships between measures of attitudes, behavior, and descriptive norms. The average perceived descriptive norm is lower than the average action which is lower than the average belief about the right action (injunctive norm). Thus, behavioral forecasts based on survey answers about perceived descriptive or injunctive norms are under- and over-estimates, respectively.

Keywords: Social norms; descriptive norms; injunctive norms; higher-order beliefs; coordination

In many settings, people's preferences depend on both internal motivations to do the right thing and on a desire to conform to the behavior of others. An important literature shows that in such settings, when people are also

MS submitted on 18 August 2021; final version received 8 February 2022

ISSN 1554-0626; DOI 10.1561/100.00021153

© 2023 E. Bueno de Mesquita and M. Shadmehr

uncertain about one another's internal motivations, public information shapes social norms. In particular, individual decision-makers overweight public information because it plays a coordinating role, helping people conform to aggregate behavior (Morris and Shin, 2002).

We study a dynamic version of the canonical model to explore how such norm-based behavior affects the possibility of social change. In our model, people have both internal and conformist motivations.¹ They are uncertain about others' internal motivations. And internal motivations change over time, creating the possibility of desirable social change.

We distinguish two kinds of internal motivations. In *common values* settings, there is a ground truth that individuals would all like their behavior to match, but individuals have different information about what that ground truth is. In *private values* settings there is no ground truth — individuals genuinely disagree about proper behavior for non-informational reasons and are uncertain what others think. We analytically distinguish these two settings and ask how patterns of social change are similar and different across them, always maintaining the assumption that, in addition to having internal motivations, people also desire for their behavior to conform to the behavior of others.

Social distancing or mask-wearing during a global pandemic is an example of a situation that might correspond to our common values model. Individuals are motivated to socially distance or mask in a way that benefits society. (Jordan *et al.* (2020) provide evidence that people have pro-social preferences concerning social distancing.) But the social optimum is shifting over time with changes in the disease and treatment, creating uncertainty about both proper behavior and how others will behave. And, in addition to wanting to do the right thing, people also want to conform to others' behavior to avoid social awkwardness — it is uncomfortable to be the only person who either is or is not wearing a mask in any given setting.²

At least in certain philosophical views, many purely moral decisions might correspond to our private values model. For instance, there might not be a fundamental ground truth regarding when exactly life starts — different people simply disagree. So, instead of wanting their behavior to match some ground truth, people considering abortion may be motivated by some combination of their personal moral views and the desire to conform to the behavior of others in their community.³ Similarly, there might not be a ground truth regarding whether any particular individual has a moral obligation to vote.

¹These are similar to the concepts of *attitudes* and *descriptive norms* in psychology (Ajzen, 2001; Armitage and Conner, 2001; Ravis and Sheeran, 2003).

²In the context of vaccination, based on their large-scale online randomized experiment, Moehring *et al.* (2021) show that people are more likely to express intention to take the vaccine when information about descriptive norms raises their belief about whether others will take the vaccine.

³See Hanschmidt *et al.* (2016) for a review of the literature on social stigma and abortion.

But people have individual views on their voting duties and, as Gerber and Rogers (2009) argue, are also more motivated to vote when they believe others in their community are likely to vote — for similar results in the context of charitable contributions, see Frey and Meier (2004).

Of course, there are gray areas. Later we discuss how our results relate to Bursztyn *et al.*'s (2020) study of the decision by Saudi men about whether to allow their female relatives to work outside of the home. One may or may not believe that there is a fundamental ground truth as to whether god or Islamic law forbids women to work. Where one comes down on this question will determine whether one views the situation as one of common or private values.

Several of our results show that behavior in common and private values settings share important characteristics.

First, Propositions 1 and 2 show that aggregate behavior exhibits inertia in both settings. Individual's actions overweight their common knowledge about past behavior. This is because adhering to past norms helps to coordinate behavior, facilitating conformity with the behavior of others. As a result of such social norms, changes in average social behavior lag behind changes in average internal motivations, whether those internal motivations are driven by a desire to match a ground truth or by shifting moral views.

So, for instance, in the case of COVID-19, this result implies that, following a large positive shock to the optimal level of social distancing or masking (e.g., due to the outbreak of a contagious virus), aggregate social distancing will be below the social optimum, even if individuals' information accurately reflects this shock on average. That is, people will be too slow to adopt social distancing or masking at the outset of the pandemic because of concerns about failing to conform to others' behavior. And following a large negative shock to the optimal level of social distancing or masking (e.g., due to widespread vaccination), aggregate behavior will be above the social optimum. That is, people will be too slow to return to social activities or remove their masks because of concerns about failing to conform to others' behavior. The same is true even in settings without a changing ground truth. As a society's views on moral questions change, behavior will lag convictions.

Second, leadership can improve this situation. Our Propositions 4 and 5 show that public information can reduce inertia by generating new common knowledge that enables people to coordinate their actions around new, more appropriate norms. Such information is more effective, the more precise it is and the more people care about conformity. Moreover, increased precision is itself more impactful the more people care about conformity.

Should such information be provided publicly or privately? In the case where there is a ground truth so that it is unambiguous what desirable social behavior means in the model, such as for COVID-19, our Proposition 8 suggests that public information is better than private information when optimal social behavior is highly correlated over-time and when individuals are poorly

informed, so that the overweighting of prior behavioral norms is more severe. This result reflects that public information creates new common knowledge. As such, people concerned with conforming to social norms overweight the new public information. This overweighting can be beneficial because it helps offset the overweighting of past behavioral norms.

While social norms cause inertia and public messages can reduce that inertia, whether values are common or private, the two settings are not identical. Comparing the model with common and private values, we establish several results.

Proposition 3 shows that the inertia problem is more severe in settings with common values. That is, changes in social behavior lag further behind changes in social convictions when those social convictions reflect an actual ground truth rather than just personal moral views. This is because there is an extra force for conformity with common values. With private values, past behavior is informative about what other people think the right action is and, thus, informative about future aggregate behavior. This creates inertia because of the desire to conform. This force is present with common values as well. And, in addition, with common values, others' past behavior is informative about what the actual right action is. Thus, even a person who did not care about conformity would pay attention to past aggregate actions when there are common values.

Second, Proposition 6 shows that public information is more effective at overcoming inertia in settings with common values than in settings with private values. This is because, in common values settings, public information does not just recalibrate expectations about the behavior of others. It also communicates information about the actual ground truth.

In addition to offering descriptive and normative insight, our model also has several implications for the growing, interdisciplinary, empirical literature on norms and behavior.

First, our results are consistent with the empirical literature in a variety of fields documenting the effect of both injunctive and descriptive norms on behavior (Armitage and Conner, 2001; Bernhardt *et al.*, 2018; Bursztyn *et al.*, 2020; Chen *et al.*, 2010; Cialdini *et al.*, 2006; Field *et al.*, 2021; Fishbein and Ajzen, 2011; Frey and Meier, 2004; Gerber and Rogers, 2009; Goldstein *et al.*, 2008; Reno *et al.*, 1993; Ravis and Sheeran, 2003). Broadly, that literature shows that people's behavior is responsive to information about others' views on correct behavior (injunctive norms) or about others' actual behavior (descriptive norms). In our model, such responsiveness is due to a combination of a desire to conform to the behavior of others and, in the case of common values, the fact that others' views convey information about the right action.

Second, extensive psychological and behavioral literatures measure *attitudes* (what do I think is right), perceived *descriptive norms* (what do I think others

will do), and perceived *injunctive norms* (what do I think others think is right) in order to adjudicate their relative importance as determinants of behavior. (See, for example, Armitage and Conner (2001) and Rivas and Sheeran (2003) for reviews and meta-analyses.) Our model suggests a different kind of empirical implication regarding a comparison of actual behavior to these measures. In particular, consider a situation in which there is a positive shock to either the underlying ground truth (in the common values setting) or the average moral view (in the private values setting). Proposition 9 predicts that the average perceived descriptive norm is lower than the average action (i.e., the true descriptive norm) which is lower than the average attitude (i.e., the true injunctive norm). Put differently, if we forecast behavior by asking people what they think others will do, we will under-estimate the actual behavior. And if we forecast behavior by asking people what they think one ought to do, we will over-estimate behavior. (The relationships are reversed if the shock is negative rather than positive.)

Related Literature

We adopt a dynamic beauty contest model (Morris and Shin, 2002) with a continuum of agents, in which the state of the world evolves according to a random walk. To focus on social norms, the state of the world is revealed at the end of each period (as in Angeletos and La'O (2010) and Angeletos and Lian (2016, p. 76)),⁴ so that we abstract from learning studied by others (Angeletos *et al.*, 2007). Morris and Shin (2002) showed that agents over-react to public information in beauty contest settings. Subsequent literature explored the welfare consequences of this insight and its implications for optimal communication (Morris and Shin, 2007), and its extensions to dynamic settings (Angeletos and La'O, 2010; Huo and Pedroni, 2020; Morris and Shin, 2006). This framework has been applied to study a variety of topics, including leadership in party conferences (Dewan and Myatt, 2008), organizations (Bolton *et al.*, 2013), monetary policy (Lorenzoni, 2010), and judiciaries (Shadmehr *et al.*, 2022).

Our paper provides an application of this framework to the relationship between social norms and social change. As such, it is largely a synthesis of existing theoretical work applied to a new substantive area. Thus, some of the results exist in different contexts in the literature. For example, the private values counterpart of Lemma 1, and its consequences for inertia appear in Angeletos and Lian's (2016, Propositions 27 and 28) comprehensive review. Our comparisons between common and private values settings, some of our normative findings and comparative statics results (e.g., the complementarity

⁴See also Angeletos and La'O (2013), which does not have aggregate uncertainty.

between the conformity weight and the public signal precision), and applications to the empirical literature on attitudes, injunctive, and descriptive norms have not been highlighted in the literature.

Our paper is related to the game-theoretic literature on social norms. Much of this literature interprets social norms as a set of equilibrium expectations and behaviors, so that different social norms correspond to different equilibria of a game with multiple equilibria (Acemoglu and Jackson, 2015; Myerson, 1991, p. 113–114; Postlewaite, 2011). Chwe (2013) argues that many social rituals exist to create common knowledge that coordinates behavior on some social norm. Our paper explores the consequences of that common knowledge for social change. We adopt Acemoglu and Jackson’s (2017) definition of social norms as “the distribution of anticipated payoff-relevant behavior” (p. 246). Our model has a unique equilibrium, and social norms in our setting refer to the average behavior of the population. When agents are heterogeneously informed, the anticipated average behavior depends on the individual’s information. We model people’s desire to conform to social norms as a desire to do what is expected of them, as reflected in other people’s behavior. This is consistent with Elster’s (2011) distinction between the mechanisms underlying social norms and those underlying what he calls moral norms. It is also consistent with the psychological concept of descriptive norms (Armitage and Conner, 2001; Ravis and Sheeran, 2003).

Acemoglu and Jackson (2015) studied the norms of cooperation in a dynamic setting where an agent in each period plays a complete information coordination game with agents of the immediate past and future. Their model features behavioral types who never cooperate and past actions are observed with noise, so some rational agents do not cooperate even in the best equilibrium. When the cooperative action of a generation becomes public for future generations, their expectations about cooperative behavior improves, facilitating cooperation — see also Acemoglu and Wolitzky (2014) and Acemoglu and Jackson (2017). In contrast, agents in our setting are uncertain about one another’s private motivations and about one another’s information in a changing world. While Acemoglu and Jackson (2015, 2017) focus on patterns of cooperation and compliance, our focus is on inertia and over-reactions to public information in a changing world and on their policy implications.

Model

We model a continuum of agents indexed by $i \in [0, 1]$, interacting over time, indexed by $t = 0, 1, \dots$. In each period t , each agent must take an action $a_{it} \in \mathbb{R}$. So, for instance, in the COVID-19 application, a higher action corresponds to more social distancing by the agent or wearing a mask more

frequently. And in the voting application, a higher action corresponds to a greater willingness to vote.

In the common values setting, absent concerns for conformity, all agents' preferred action is θ_t . In the COVID-19 application, we interpret θ_t as corresponding to the socially optimal level of social distancing.

But, in each period, each agent cares about this ideal action *and* about conforming to the average action that others take in that period, $A_t = \int a_{it} di$. This generates a complementarity: if an agent believes that others do little social distancing, this raises that agent's incentive also to do less social distancing. This captures, among other things, social pressure and the cost of deviating from the norms of behavior in the society.

So, in the common values setting, an agent's payoff in period t is:

$$-(1 - \alpha)(a_{it} - \theta_t)^2 - \alpha(a_{it} - A_t)^2, \tag{1}$$

where $\alpha \in (0, 1)$ is the agent's relative weight on conformity.

The ideal action, θ_t , follows a random walk: $\theta_t = \theta_{t-1} + u_t$, where $u_t \sim iidN(0, \sigma_u)$. Individuals do not observe θ_t , but each agent observes a signal: $x_{it} = \theta_t + \epsilon_{it}$, where $\epsilon_{it} \sim iidN(0, \sigma_\epsilon)$. Throughout, we assume that the noise and fundamentals are independent from each other in the standard manner. So people are uncertain about one another's beliefs about the target action.

Individual i observes x_{it} in period t , and θ_{t-1} becomes public in period t .⁵ Individuals discount future payoffs by δ , and each agent maximizes the expected sum of discounted period payoffs.

The private values setting is identical, except that an agent's payoff in period t is:

$$-(1 - \alpha)(a_{it} - x_{it})^2 - \alpha(a_{it} - A_t)^2. \tag{2}$$

That is, in this setting, people do not wish to conform to some ideal action, θ_t . Rather they wish to match their personal convictions, x_{it} , but also remain concerned about conforming their behavior to the behavior of others. Here, we interpret θ_t not as the ideal action, but simply as the mean of the distribution of personal convictions.⁶

⁵Even if θ_{t-1} is not observed in the current period, agents will infer it in equilibrium if they observe the last period's aggregate behavior A_{t-1} .

⁶Individuals in the real world have difference preferences for conformity. To introduce such heterogeneity suppose agents have heterogeneous weights on conformity, so that agent i 's weight on conformity is $\alpha_i \sim iidF$ and α_i s are independent from other variables. The same results in Propositions 1–6 and 9 are obtained, but with the average α_i instead of α . Our normative analysis of Section “Leadership through Information”, however, will also depend on the second moment of α_i introduced through the quadratic distance in the normative measure.

Inertia in Social Change

In this section, we characterize equilibrium in both settings and show that there is inertia in social change. We start with the common values setting.

Because there is a continuum of agents, an individual agent’s action does not affect the aggregate outcome, either in the current or in future periods. Thus, the only link between periods is information. From Equation (1), agent i chooses the following action:

$$a_{it} = (1 - \alpha)\mathbb{E}_{it}[\theta_t] + \alpha \mathbb{E}_{it}[A_t], \tag{3}$$

where $\mathbb{E}_{it}[\cdot]$ is the expectation of i in period t given his information.

Define $\bar{\mathbb{E}}^h$ recursively as follows. $\bar{\mathbb{E}}^0[X] = X$, $\bar{\mathbb{E}}^1[X] = \bar{\mathbb{E}}[\bar{\mathbb{E}}^0[X]] = \int \mathbb{E}_i[X]di$, $\bar{\mathbb{E}}^h[X] = \bar{\mathbb{E}}[\bar{\mathbb{E}}^{h-1}[X]] = \int \mathbb{E}_i[\bar{\mathbb{E}}^{h-1}[X]]di$. That is, $\bar{\mathbb{E}}^1[X]$ is the average expectation of the random variable X in the population; $\bar{\mathbb{E}}^2[X]$ is the average expectation in the population about the average expectation in the population, and so on. Lemma 1 shows that the aggregate action in the population depends on all such higher-order expectations in the population about the target action. All proofs are in the appendix.

Lemma 1. *In the common values setting, the aggregate action in each period depends on all average higher order beliefs in the population about the target action, with lower weights on higher orders⁷:*

$$A_t = \sum_{h=1}^{\infty} (1 - \alpha) \alpha^{h-1} \bar{\mathbb{E}}_t^h[\theta_t]. \tag{4}$$

Now, combining the properties of Normal distribution and Lemma 1 yields:

Proposition 1. *In the common values setting, conformity generates inertia. In particular,*

- $A_t = \theta_{t-1} + \phi u_t$, where $0 < \phi < \beta < 1$, $\phi = \frac{(1-\alpha)\beta}{1-\alpha\beta}$, and $\beta = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_\epsilon^2}$.
- ϕ is decreasing in α , with $\lim_{\alpha \rightarrow 0} \phi(\alpha) = \beta$ and $\lim_{\alpha \rightarrow 1} \phi(\alpha) = 0$.

Because agents care about coordinating their actions, they put extra weight on their common knowledge of the past, which facilitates coordination — reminiscent of the logic of focal points. Individuals have common knowledge that, on average, today’s target action is yesterday’s target action (i.e., $\theta_t \sim$

⁷For convergence, we also need $\lim_{h \rightarrow \infty} \alpha^h \bar{\mathbb{E}}_t^h[A_t] = 0$, which is satisfied for commonly used, linear strategies, and within the more general class of “nonexplosive strategies” defined in Dewan and Myatt (2008, p. 365), or in a variation of the model with a bounded action space.

$N(\theta_{t-1}, \sigma_u)$) and hence over-weight this fact. As a result, today’s aggregate action is biased in the direction of yesterday’s target action.

In our COVID-19 application, Proposition 1 implies that, following a large positive shock to the socially optimal amount of social distancing (like the outbreak of COVID-19), aggregate social distancing will be far lower than is socially optimal.

The analysis and results are qualitatively similar in the private values setting. Except now the optimal action for agent i is:

$$a_{it} = (1 - \alpha)x_{it} + \alpha \mathbb{E}_{it}[A_t]. \tag{5}$$

Again, concern about social conformity creates inertia because people’s information about past behavior coordinates their current behavior.

Proposition 2. *In the private values setting, conformity generates inertia. In particular,*

- $A_t = \theta_{t-1} + \phi^{Pr} u_t$, where $0 < \phi^{Pr} < 1$, $\phi^{Pr} = \frac{1-\alpha}{1-\alpha\beta}$, and $\beta = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_\epsilon^2}$.
- ϕ^{Pr} is decreasing in α , with $\lim_{\alpha \rightarrow 0} \phi^{Pr}(\alpha) = 1$ and $\lim_{\alpha \rightarrow 1} \phi^{Pr}(\alpha) = 0$.

The results in this section provide a way of thinking about empirical findings like those in Bursztyn *et al.*’s (2020) study of whether men allow their female relatives to work in Saudi Arabia. They found that men’s behavior lags well behind their attitudes on average. That is, there are many more men who believe it is acceptable to allow their female relatives to work outside the home than there are men who actually allow their female relatives to do so. And the reason is uncertainty about what other men believe. In particular, men appear to systematically underestimate other men’s acceptance of women working outside the home. Bursztyn *et al.* show that when men are given a private signal about the true beliefs of other men, they significantly change their behavior — becoming much more likely to allow their own female relatives to work.

This is consistent with our results on inertia. Suppose we think of the action in our model as corresponding to a man’s willingness to let his female relatives work outside the home with higher actions corresponding to higher degrees of willingness. As social attitudes become more progressive, aggregate actions will lag behind aggregate beliefs: when men remain uncertain about each other’s beliefs, the fear of being an outlier (the force for conformity to social norms) causes them to underweight their recent, more progressive beliefs about the right thing to do in favor of the common knowledge of familiar, past views. If we give an agent a private signal about the value of θ_t , and that signals turn out to be higher than the agent expected, this improves the agent’s beliefs about θ_t . In both settings, this leads the agent’s action

to increase. In the private values setting, this is because it makes the agent believe others are more likely to take a higher action. And in the common values setting, it has this effect and also directly increases the agent's beliefs about the ideal action. Thus, consistent with the findings of Burszty *et al.*'s experiment, our model predicts that actions will lag beliefs and that providing private information that social beliefs have changed more than had previously been believed will increase actions.

We have seen that social norms inhibit social change in both settings for similar reasons — past behavior is informative about future behavior and people are motivated to conform. The next result shows that this inertia is greater in settings with common values. The reason is that there is an extra force in the common values setting. In particular, past behavior is informative about the value of θ_t . And, in the common values setting, people care directly about matching their behavior to the true θ_t . Hence, even if people had no desire to conform to other's behavior (i.e., if α went to zero), there would still be inertia in the common values setting, because people would use past behavior to help estimate θ_t . Propositions 1 and 2 imply:

Proposition 3. *Inertia is higher in the common values setting than in the private values setting: $\phi = \beta\phi^{Pr}$.*

When prior common knowledge disappears ($\sigma_u \rightarrow \infty$), the two settings converge: $\beta \approx 1$, so that $\phi \approx \phi^{Pr}$. The reason is that the prior common knowledge about θ_t becomes uninformative and is discarded, so that an agent i 's estimate of θ_t in the common values setting becomes their signal: $E[\theta_t | \theta_{t-1}, x_{it}] \approx x_{it}$, as in the private values setting. Indeed, without this prior common knowledge inertia disappears: $\lim_{\sigma_u \rightarrow \infty} \phi = \lim_{\sigma_u \rightarrow \infty} \phi^{Pr} = 1$.

Leadership through Information

Increased information can mitigate the problem of social inertia. To see this, suppose in each period t , in addition to their private signals x_{it} , agents also receive a public signal $p_t = \theta_t + \eta_t$, with $\eta_t \sim iidN(0, \sigma_\eta)$. Such a public signal might be the result of information conveyed by leaders. Lemma 1 still holds because it does not depend on the details of available information. However, the presence of public signals changes the degree of inertia in aggregate behavior.

Proposition 4. *In the setting with common values:*

1. *Public signals reduce inertia. Averaging over the public signal noise, the expected aggregate action is: $E[A_t | \theta_{t-1}, u_t] = \theta_{t-1} + \phi_p u_t$, where $\phi < \phi_p < 1$.*

2. *The effect of the public signal is increasing in its precision and in the weight on conformity. Moreover, the marginal effect of higher precision is increasing in the weight on conformity. That is:*

$$\frac{\partial(\phi_p - \phi)}{\partial(1/\sigma_\eta)}, \quad \frac{\partial(\phi_p - \phi)}{\partial\alpha}, \quad \frac{\partial}{\partial\alpha} \frac{\partial(\phi_p - \phi)}{\partial(1/\sigma_\eta)} > 0.$$

Qualitatively similar results hold in private values settings.

Proposition 5. *In the setting with private values:*

1. *Public signals reduce inertia. Averaging over the public signal noise, the expected aggregate action is: $E[A_t|\theta_{t-1}, u_t] = \theta_{t-1} + \phi_p^{Pr} u_t$, where $\phi_p^{Pr} < \phi_p^{Pr} < 1$.*
2. *The effect of the public signal is increasing in its precision and in the weight on conformity. Moreover, the marginal effect of higher precision is increasing in the weight on conformity. That is:*

$$\frac{\partial(\phi_p^{Pr} - \phi^{Pr})}{\partial(1/\sigma_\eta)}, \quad \frac{\partial(\phi_p^{Pr} - \phi^{Pr})}{\partial\alpha}, \quad \frac{\partial}{\partial\alpha} \frac{\partial(\phi_p^{Pr} - \phi^{Pr})}{\partial(1/\sigma_\eta)} > 0.$$

Any particular public signal may be far from the ideal action due to noise. However, the noise cancels out in expectation, and on average public signals reduce inertia. Individuals put more weight on more informative public signals, raising their effect in reducing inertia. Indeed, when public signals are uninformative, their effect vanishes; and when they become very precise, they will reveal the ideal action and eliminate inertia: $\lim_{\sigma_\eta \rightarrow \infty} (\phi_p(\sigma_\eta) - \phi) = \lim_{\sigma_\eta \rightarrow \infty} (\phi_p^{Pr}(\sigma_\eta) - \phi^{Pr}) = 0$, and $\lim_{\sigma_\eta \rightarrow 0} \phi_p(\sigma_\eta) = \lim_{\sigma_\eta \rightarrow 0} \phi_p^{Pr}(\sigma_\eta) = 1$.

Propositions 4 and 5 also show that the marginal effect of more precise public signals is higher when the society is more concerned with conformity and social norms. Combining these results shows that public information reduces inertia, more so when that information is more precise or when social norms are more entrenched in the sense that agents value conformity more. Moreover, the marginal impact of more precise information is greater exactly when the population is more concerned with adhering to social norms.

We have seen that public information can reduce inertia due to social norms in both settings. The next result shows that such information is more effective in the common values setting. The reason is, again, the presence of an extra force in that setting. In particular, the public signal is informative about the value of θ_t . In both settings, this public signal thus helps to coordinate the behavior of the conformity-minded agents. And, in addition, in the common values setting, it is directly informative about the correct action. Hence, agents' behavior is more responsive to the public signal in the common values setting.

Proposition 6. *The marginal effect of public signals is higher in the common versus private values setting: $\phi_p^{Pr} - \phi^{Pr} = \alpha(\phi_p - \phi)$.*

The above logic implies that when agents do not care about targeting the ideal action θ_t , the extra force in the common values setting disappears. Indeed, Proposition 6 shows that when agents almost only care about conformity (i.e., as α approaches 1), the difference in the marginal effect of public signals in both private and common values settings vanishes.

Achieving Normative Improvements with Common Values

In the common values setting, where there is a ground truth that describes ideal behavior, there is a clear way of thinking about normative improvement — is aggregate behavior closer to ideal behavior? That is not true in the private values setting, where people genuinely disagree about proper behavior. Hence, in this section, we focus on the common values setting. In that setting, we will say that there is a *normative improvement* if $E[\int (a_{it} - \theta_t)^2 di]$ decreases, so that aggregate behavior is closer to ideal behavior.

Propositions 4 shows that, following a shock, an informed leader can send a public signal about the ideal action that helps set public expectations, thereby reducing inertia driven by the desire to conform. The clearer that message (i.e., the lower σ_η), the more this will reduce inertia. The next result characterizes when such an intervention leads to a normative improvement.

Proposition 7. *Improving the clarity of the public signal causes a normative improvement (i.e., $E[\int (a_{it} - \theta_t)^2 di]$ is increasing in σ_η) if: (i) $\alpha \leq 1/2$ or (ii) σ_η is sufficiently small.*

In a common values setting, public messages are a normative improvement if people do not put too much weight on conformity ($\alpha \leq 1/2$) or the public signal is sufficiently informative (σ_η small). Why these conditions? Because agents value conformity, they put excessive weight on all public signals relative to a Bayesian agent who only cares about choosing an action that reflects the best estimate of θ_t (this was the same logic that drove inertia in the first place). Because this distortion is smaller when α is smaller, new public information about the optimal social distancing is always beneficial when citizens put relatively less weight on conformity ($\alpha \leq 1/2$). In the other extreme, when agents almost only care about conformity ($\alpha \approx 1$), they put almost no weight on their private signals. Now, although agents over-weight new public information (p_t), this over-reaction to the new public information helps counter-act their over-reaction to past experience (i.e., $\theta_t \sim (\theta_{t-1}, \sigma_u^2)$), and the overall effect is again beneficial. In between, when $\alpha \in (1/2, 1)$, these effects compete and the overall effect of raising the precision of new public

information may be negative unless it is sufficiently informative (σ_η small) to offset the over-reaction.⁸

For social distancing in the presence of a dangerous infectious disease, we believe the relevant parameter space is $\alpha \leq 1/2$. It is unlikely that people care so much about conformity that over-reaction to new public information trumps its value. Hence, for cases like COVID-19, Proposition 7 suggests that clear and consistent public messages from a leader are likely to be socially beneficial.

Given the overreaction by agents to public messages described above, one may wonder whether there is a better way to deliver information in a common values setting. Would it be better for agents to receive the same level of information, but privately rather than publicly? For instance, perhaps employers or local governments could provide private information to agents, rather than them all observing the same public information in a presidential speech or press conference.

To consider this possibility, contrast the public signal case with a setting where, instead of receiving private and public signals $x_{it} \sim N(\theta_t, \sigma_\epsilon^2)$ and $p_t \sim N(\theta_t, \sigma_\eta^2)$, citizens receive a single private signal x'_{it} with the same amount of information about the ideal action θ_t as the public and private signals combined. In particular, let $x'_{it} = \theta_t + \epsilon'_{it}$, with $\epsilon'_{it} \sim N(0, \sigma_{\epsilon'}^2 = \frac{\sigma_\epsilon^2 \sigma_\eta^2}{\sigma_\epsilon^2 + \sigma_\eta^2})$.

Proposition 8. *The setting with the combination of private and public signals (x_{it}, p_t) is a normative improvement over the setting with more precise private signals x'_{it} when σ_u is sufficiently small or σ_ϵ is sufficiently large.*

When agents believe the past is highly informative about the present (σ_u small) or that they are privately poorly informed (σ_ϵ large), agents put too much weight on their past experience. In such circumstances, it is better for the government to provide public rather than private information. Individuals over-react to public messages. But that will help to counter-act their over-reaction to their past experience. By contrast, when agents believe the past is relatively uninformative (σ_u large) or that they are privately well-informed (σ_ϵ small), private communication is preferred.

Standard accounts frame the problem of social distancing as a public goods problem with the familiar externalities. For example, individuals do not internalize that social distancing has positive health spillovers on others, so there will be under-provision of social distancing. In this setting, the fundamental problem is not informational: even fully informed citizens under-provide social distancing absent some more heavy-handed policy that directly changes citizen incentives — e.g., forced downtown closure or fines for public gatherings. Moreover, enforcing behavioral changes for actions that are largely

⁸Equation (24) in the proof of Proposition 7 shows the necessary and sufficient conditions for when reducing σ_η is a normative improvement.

taken out of the public eye, such as hand washing, handshakes, or private gatherings, is virtually impossible. This aspect of the social distancing challenge has been the focus of public and academic discussions. For example, Allcott *et al.* (2020) study the interaction between risk perception and such externalities in the United States. Dube and Baicker (2020) discuss the importance of individuals sacrificing their interests for the greater good and, drawing on Christensen *et al.*'s (2021) study of the Ebola crisis in Sierra Leon, emphasize the importance of trust in local leaders and institutions.

In contrast, our analysis highlights the role of social norms and strategic uncertainty as an under-appreciated source of counter-productive inertia in aggregate social distancing. The policy implications are also sharply different. While information alone cannot resolve the standard externalities problem, clear and consistent *public* information can dramatically improve social distancing by reducing strategic uncertainty and enabling citizens to coordinate on new norms. Moreover, in contrast to accounts that emphasize the role of the local community in providing trusted information, we highlight the advantage of information provided by national over local leaders, especially in countries where there is trust in the expertise of governmental health organizations. National coverage generates more common knowledge, enabling citizens to better coordinate on new optimal norms of social distancing.

Attitudes, Descriptive Norms, and Injunctive Norms

Extensive literatures across social sciences examine the importance of attitudes, perceived descriptive norms, and perceived injunctive norms for explaining behavior (Armitage and Conner, 2001; Bernhardt *et al.*, 2018; Borsari and Carey, 2003; Bursztyn *et al.*, 2020; Chen *et al.*, 2010; Cialdini *et al.*, 2006; Field *et al.*, 2021; Fishbein and Ajzen, 2011; Frey and Meier, 2004; Gerber and Rogers, 2009; Goldstein *et al.*, 2008; Reno *et al.*, 1993; Ravis and Sheeran, 2003). Loosely, attitudes are understood to reflect an individual's beliefs about right actions, perceived descriptive norms are understood to reflect an individual's beliefs about others' behavior, and perceived injunctive norms are understood to reflect a person's beliefs about what others think they ought to do. The experimental and survey literatures devote considerable effort to measuring and manipulating these concepts, primarily with the goal of adjudicating their relative importance for explaining behavior.

These concepts map naturally to the model. Therefore, our theoretical framework may be useful to organize the existing empirical results and inform future empirical strategies.

In the model, in both common and private values setting, the average/aggregate ideal/right action/behavior in period t is θ_t . In the common values setting, the ideal for everyone is θ_t , so the average is the same. In

the private values setting, the ideal for agent i is x_{it} , so its expectation is $E[x_{it}] = \int x_{it} di = \theta_t$.

We might think of an agent's beliefs about the right/ideal action/behavior as that agent's *attitude*: $E_{it}[\theta_t]$ with common values, x_{it} with private values. In the literature, a person's perception of the injunctive norm is their belief about what others think they ought to do. Thus, it is natural to define the society's *injunctive norm* as the average attitude: $\bar{E}_t[\theta_t] = \int E_{it}[\theta_t] di$ in common values settings, and $\bar{E}_t[x_{it}] = \int x_{it} di$ in private values settings. Then, the *average perception of the injunctive norm* is $\bar{E}_t^2[\theta_t] = \bar{E}_t[\bar{E}_t[\theta_t]]$ in common values settings, and $\bar{E}_t^2[x_{it}] = \bar{E}_t[\bar{E}_t[x_{it}]]$ in private values settings. These perceptions involve higher-order beliefs. As the earlier analysis (e.g., Lemma 1) suggests, our theoretical framework allows for characterization of such higher-order beliefs. Similarly, in the literature, a person's perception of the descriptive norm is their belief about what others will do. Thus, it is natural to define the society's *descriptive norm* as the average/aggregate action/behavior, $A_t = \int a_{it} di$. Then, the *average perception of the descriptive norm* is $\bar{E}_t[A_t]$. Proposition 9 characterizes the relationship between these measures.

Proposition 9. *When the shock to the average ideal action is positive, the average perception of the descriptive norm is lower than the descriptive norm (average action), which is lower than the injunctive norm. When the shock to the average ideal action is negative, these relationships are reversed. Formally, suppose $\theta_{t-1} < \theta_t$. In the common values setting,*

$$\bar{E}_t[A_t] < A_t < \bar{E}_t[\theta_t].$$

In the private values setting,

$$\bar{E}_t[A_t] < A_t < \bar{E}_t[x_{it}].$$

Moreover, the average perception of the injunctive norm (i.e., $\bar{E}_t^2[\theta_t]$ or $\bar{E}_t^2[x_{it}]$) is lower than the injunctive norm, and if and only if $\alpha(1 + \beta) > 1$, the average perception of the injunctive norm is higher than the descriptive norm. Inequalities are reversed if $\theta_{t-1} > \theta_t$, and become equalities if $\theta_{t-1} = \theta_t$. Moreover, the differences are smaller in common values than in private values settings.

To see an immediate implication, consider a situation in which there is a positive shock to either the underlying ground truth (in the common values setting) or the average moral view (in the private values setting). Proposition 9 predicts that if we estimate aggregate behavior by asking people what they think others will do (perceived descriptive norms), we tend to under-estimate actual behavior. And if we estimate behavior by asking people what they think is the right thing to do (injunctive norms), we tend to over-estimate behavior.

We can provide some intuition for these results. We do so for the common values setting (the private values intuitions are analogous).

The average belief about θ_t is an under-estimate when u_t is positive. Why is this the case? If u_t is positive, then the average person gets a signal that is above the prior mean belief, θ_{t-1} . The average person's posterior mean belief is then a weighted average of that signal and the prior mean — hence, it is an under-estimate of the true θ_t . This explains why, when u_t is positive, perceived descriptive norms are below true descriptive norms. Since aggregate actions (descriptive norms) and θ_t are one-to-one, the fact that the average person's beliefs about θ_t are an under-estimate means that the average person's beliefs about A_t are also an under-estimate, which is why $\bar{\mathbb{E}}_t[A_t] < A_t$. In addition to reflecting this informational inertia, aggregate action also exhibits inertia due to the desire to conform. This is why descriptive norms are lower than injunctive norms. And, of course, if u_t is negative, these relationships are reversed.

Conclusion

Social norms can disrupt social change. In particular, aggregate behavior over-reacts to past experience when people are motivated to socially conform and are uncertain about one another's beliefs. This is because the common knowledge created by past experience creates social norms that facilitate coordination. The result is undesirable inertia — behavior tends to fall back toward old norms. This is true in both common- and private-values settings but is most severe in common-values settings, where it is arguably also most problematic.

Communication by leaders can help to mitigate this damaging inertia. If messages are public, the common knowledge they create engenders an over-reaction analogous to that created by past experience. This over-reaction can help off-set the inertia resulting from the over-weighting of past norms. The power of information to overcome norms-based inertia is greater in common values settings than in private values settings.

In common-values settings, where normative improvements are well-defined, providing public information is preferable to providing private information if agents' information is noisy and the socially optimal behavior is sticky (e.g., a rare pandemic shock that requires significantly more social distancing). In the case of COVID-19, both of these conditions are likely to hold. As such, clear and consistent public statements by a national leader are expected to be more effective than similarly informative, but more private, statements (e.g., by local governments or employers).

Importantly, the model provides a framework to study the relationships between individual actions, injunctive and descriptive norms, and their

perceptions in the society, which have been increasingly the focus of empirical studies across social sciences. For example, following a positive shock, we expect that a forecast of aggregate behavior that comes from asking people what they think others will do tends to under-estimate actual behavior. And we expect a forecast of aggregate behavior that comes from asking people what they think is the right thing to do tends to over-estimate actual behavior.

Appendix: Proofs

Proof of Lemma 1: From Equation (3),

$$\begin{aligned}
 A_t &= \int a_{it} di \\
 &= \int ((1 - \alpha)\mathbb{E}_{it}[\theta_t] + \alpha \mathbb{E}_{it}[A_t]) di = (1 - \alpha)\bar{\mathbb{E}}_t[\theta_t] + \alpha \bar{\mathbb{E}}_t[A_t].
 \end{aligned}$$

Iterating yields:

$$A_t = (1 - \alpha)\bar{\mathbb{E}}_t[\theta_t] + (1 - \alpha)\alpha \bar{\mathbb{E}}_t^2[\theta_t] + \alpha^2\bar{\mathbb{E}}_t^2[A_t].$$

Repeated iteration yields the result. □

Proof of Proposition 1: We calculate $\bar{\mathbb{E}}_t^h[\theta_t]$ and use Lemma 1. Because $\mathbb{E}_{it}[\theta_t] = \mathbb{E}_{it}[u_t] + \theta_{t-1}$, we have:

$$\bar{\mathbb{E}}_t^h[\theta_t] = \bar{\mathbb{E}}_t^h[u_t] + \theta_{t-1}. \tag{6}$$

Note that $x_{it} = \theta_t + \epsilon_{it} = \theta_{t-1} + u_t + \epsilon_{it}$. Thus, letting $\beta = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_\epsilon^2}$

$$\mathbb{E}_{it}[u_t] = \beta(x_{it} - \theta_{t-1}) = \beta(u_t + \epsilon_{it}) \Rightarrow \bar{\mathbb{E}}_t[u_t] = \beta u_t. \tag{7}$$

Iterating yields:

$$\bar{\mathbb{E}}_t^h[u_t] = \beta^h u_t. \tag{8}$$

Substituting from Equation (8) into Equation (6) yields:

$$\bar{\mathbb{E}}_t^h[\theta_t] = \beta^h u_t + \theta_{t-1}. \tag{9}$$

Now, substituting from Equation (9) into Equation (4) in Lemma 1 yields:

$$A_t = \sum_{h=1}^{\infty} (1 - \alpha) \alpha^{h-1} (\beta^h u_t + \theta_{t-1}) = \theta_{t-1} + \frac{\beta(1 - \alpha)}{1 - \alpha\beta} u_t. \tag{10}$$

In Equation (10), let $\phi = \frac{(1-\alpha)\beta}{1-\alpha\beta}$. □

Proof of Proposition 2: Mirroring Lemma 1,

$$A_t = \sum_{h=0}^{\infty} (1 - \alpha) \alpha^h \bar{\mathbb{E}}_t^h[\theta_t]. \tag{11}$$

Now, substituting from Equation (9) into Equation (11) yields:

$$A_t = \sum_{h=0}^{\infty} (1 - \alpha) \alpha^h (\beta^h u_t + \theta_{t-1}) = \theta_{t-1} + \frac{1 - \alpha}{1 - \alpha\beta} u_t. \tag{12}$$

In Equation (12), let $\phi^{Pr} = \frac{1-\alpha}{1-\alpha\beta}$. □

Proof of Proposition 3: Follows from Propositions 1 and 2. □

Proof of Proposition 4: With the public signal p_t , $\mathbb{E}_{it}[u_t] = E[u_t|x_{it}, p_t]$. Thus,

$$\begin{aligned} \mathbb{E}_{it}[u_t] &= \frac{\sigma_u^2 \sigma_\eta^2 (x_{it} - \theta_{t-1}) + \sigma_u^2 \sigma_\epsilon^2 (p_t - \theta_{t-1})}{\sigma_u^2 \sigma_\eta^2 + \sigma_u^2 \sigma_\epsilon^2 + \sigma_\epsilon^2 \sigma_\eta^2} \\ &= \frac{\sigma_u^2 \sigma_\eta^2 (u_t + \epsilon_{it}) + \sigma_u^2 \sigma_\epsilon^2 (p_t - \theta_{t-1})}{\sigma_u^2 \sigma_\eta^2 + \sigma_u^2 \sigma_\epsilon^2 + \sigma_\epsilon^2 \sigma_\eta^2}. \end{aligned} \tag{13}$$

Thus,

$$\bar{\mathbb{E}}_t[u_t] = \frac{\sigma_u^2 \sigma_\eta^2 u_t + \sigma_u^2 \sigma_\epsilon^2 (p_t - \theta_{t-1})}{\sigma_u^2 \sigma_\eta^2 + \sigma_u^2 \sigma_\epsilon^2 + \sigma_\epsilon^2 \sigma_\eta^2} = A_u u_t + A_p (p_t - \theta_{t-1}), \tag{14}$$

where

$$A_u = \frac{\sigma_u^2 \sigma_\eta^2}{\sigma_u^2 \sigma_\eta^2 + \sigma_u^2 \sigma_\epsilon^2 + \sigma_\epsilon^2 \sigma_\eta^2} \quad \text{and} \quad A_p = \frac{\sigma_u^2 \sigma_\epsilon^2}{\sigma_u^2 \sigma_\eta^2 + \sigma_u^2 \sigma_\epsilon^2 + \sigma_\epsilon^2 \sigma_\eta^2}, \tag{15}$$

with

$$\lim_{\sigma_\eta \rightarrow \infty} A_p = 0, \quad \lim_{\sigma_\eta \rightarrow \infty} A_u = \beta, \quad \lim_{\sigma_\eta \rightarrow 0} A_p = 1, \quad \text{and} \quad \lim_{\sigma_\eta \rightarrow 0} A_u = 0. \tag{16}$$

Iterating on Equation (14) yields

$$\bar{\mathbb{E}}_t^h[u_t] = (A_u)^h u_t + (1 + \dots + A_u^{h-1}) A_p (p_t - \theta_{t-1}). \tag{17}$$

Substituting from Equation (17) into Equation (6) yields:

$$\begin{aligned} \bar{\mathbb{E}}_t^h[\theta_t] &= (A_u)^h u_t + (1 + \dots + A_u^{h-1}) A_p (p_t - \theta_{t-1}) + \theta_{t-1} \\ &= (A_u)^h u_t + \frac{1 - A_u^h}{1 - A_u} A_p (p_t - \theta_{t-1}) + \theta_{t-1}. \end{aligned} \tag{18}$$

Now, substituting from Equation (18) into Equation (4) in Lemma 1 yields:

$$\begin{aligned}
 A_t &= \sum_{h=1}^{\infty} (1-\alpha) \alpha^{h-1} \left((A_u)^h u_t + \frac{1-A_u^h}{1-A_u} A_p (p_t - \theta_{t-1}) + \theta_{t-1} \right) \\
 &= \theta_{t-1} + \frac{(1-\alpha)A_u}{1-\alpha A_u} u_t + (1-\alpha) \frac{A_p}{1-A_u} \left(\frac{1}{1-\alpha} - \frac{A_u}{1-\alpha A_u} \right) \\
 &\quad \times (p_t - \theta_{t-1}) \\
 &= \theta_{t-1} + \frac{(1-\alpha)A_u}{1-\alpha A_u} u_t + \frac{A_p}{1-\alpha A_u} (p_t - \theta_{t-1}) \\
 &= \theta_{t-1} + \frac{(1-\alpha)A_u u_t + A_p (p_t - \theta_{t-1})}{1-\alpha A_u}.
 \end{aligned} \tag{19}$$

Note that, using Equation (16), if $\sigma_\eta \rightarrow \infty$, Equation (19) simplifies to Equation (10).

For given θ_{t-1} and u_t , aggregate action A_t takes different values for different values of the public signal p_t , depending on the idiosyncratic error term η_t in the public signal. The average public signal, for given θ_{t-1} and u_t , is $E[p_t|u_t, \theta_{t-1}] = \theta_{t-1} + u_t$. Then, averaging over the public signal noise, Equation (19) becomes:

$$E[A_t|u_t, \theta_{t-1}] = \theta_{t-1} + \phi_p u_t, \text{ where } \phi_p = \frac{(1-\alpha)A_u + A_p}{1-\alpha A_u}.$$

From Equation (16), $\lim_{\sigma_\eta \rightarrow 0} \phi_p = 1$ and $\lim_{\sigma_\eta \rightarrow \infty} \phi_p = \phi$. Comparing with ϕ in Proposition 1 yields:

$$\begin{aligned}
 \phi_p - \phi &= \frac{(1-\alpha)A_u + A_p}{1-\alpha A_u} - \frac{(1-\alpha)\beta}{1-\alpha\beta} \\
 &= \frac{\sigma_\epsilon^4 \sigma_u^2}{(\sigma_\epsilon^2 + (1-\alpha)\sigma_u^2)(\sigma_\epsilon^2 \sigma_u^2 + \sigma_\eta^2(\sigma_\epsilon^2 + (1-\alpha)\sigma_u^2))}.
 \end{aligned} \tag{20}$$

Thus, $\phi_p - \phi > 0$ and $\phi_p - \phi$ is increasing in α . Moreover,

$$\frac{d\phi_p}{d\sigma_\eta^2} = -\frac{\sigma_\epsilon^4 \sigma_u^2}{\sigma_\epsilon^2 \sigma_u^2 + \sigma_\eta^2 (\sigma_\epsilon^2 + (1-\alpha)\sigma_u^2)^2} < 0.$$

Thus, reducing the noise in the public signal (less σ_η^2) raises ϕ_p . Further, this derivative is decreasing in α . □

Proof of Proposition 5: Substituting from Equation (18) into Equation (11) in the proof of Proposition 2 yields:

$$\begin{aligned}
 A_t &= \sum_{h=0}^{\infty} (1-\alpha) \alpha^h \left((A_u)^h u_t + \frac{1-A_u^h}{1-A_u} A_p (p_t - \theta_{t-1}) + \theta_{t-1} \right) \\
 &= \theta_{t-1} + \frac{1-\alpha}{1-\alpha A_u} u_t + (1-\alpha) \frac{A_p}{1-A_u} \left(\frac{1}{1-\alpha} - \frac{1}{1-\alpha A_u} \right) \\
 &\quad \times (p_t - \theta_{t-1}) \\
 &= \theta_{t-1} + \frac{1-\alpha}{1-\alpha A_u} u_t + \frac{\alpha A_p}{1-\alpha A_u} (p_t - \theta_{t-1}) \\
 &= \theta_{t-1} + \frac{(1-\alpha)u_t + \alpha A_p (p_t - \theta_{t-1})}{1-\alpha A_u}. \tag{21}
 \end{aligned}$$

For given θ_{t-1} and u_t , aggregate action A_t takes different values for different values of the public signal p_t , depending on the idiosyncratic error term η_t in the public signal. The average public signal, for given θ_{t-1} and u_t , is $E[p_t|u_t, \theta_{t-1}] = \theta_{t-1} + u_t$. Then, averaging over the public signal noise, Equation (21) becomes:

$$E[A_t|u_t, \theta_{t-1}] = \theta_{t-1} + \phi_p^{Pr} u_t, \quad \text{where } \phi_p^{Pr} = \frac{(1-\alpha) + \alpha A_p}{1-\alpha A_u}.$$

From Equation (16), $\lim_{\sigma_\eta \rightarrow 0} \phi_p^{Pr} = 1$ and $\lim_{\sigma_\eta \rightarrow \infty} \phi_p^{Pr} = \phi^{Pr}$. Comparing with ϕ^{Pr} in Proposition 2 yields:

$$\begin{aligned}
 \phi_p^{Pr} - \phi^{Pr} &= \frac{(1-\alpha) + \alpha A_p}{1-\alpha A_u} - \frac{1-\alpha}{1-\alpha\beta} \\
 &= \frac{\alpha\sigma_\epsilon^4\sigma_u^2}{(\sigma_\epsilon^2 + (1-\alpha)\sigma_u^2)(\sigma_\epsilon^2\sigma_u^2 + \sigma_\eta^2(\sigma_\epsilon^2 + (1-\alpha)\sigma_u^2))}. \tag{22}
 \end{aligned}$$

Thus, $\phi_p^{Pr} - \phi^{Pr} > 0$ and $\phi_p^{Pr} - \phi^{Pr}$ is increasing in α . Moreover,

$$\frac{\partial \phi_p^{Pr}}{\partial \sigma_\eta^2} = -\frac{\alpha\sigma_\epsilon^4\sigma_u^2}{\sigma_\epsilon^2\sigma_u^2 + \sigma_\eta^2(\sigma_\epsilon^2 + (1-\alpha)\sigma_u^2)^2} < 0.$$

Thus, reducing the noise in the public signal (less σ_η^2) raises ϕ_p^{Pr} . Further, this derivative is decreasing in α . □

Proof of Proposition 6: From Equations (20) and (22), $\phi_p^{Pr} - \phi^{Pr} = \alpha(\phi_p - \phi)$. □

Proof of Proposition 7: From Equation (3),

$$\begin{aligned}
 a_{it} &= (1 - \alpha)\mathbb{E}_{it}[\theta_t] + \alpha \mathbb{E}_{it}[A_t] \\
 &= (1 - \alpha)\mathbb{E}_{it}[\theta_{t-1} + u_t] \\
 &\quad + \alpha \mathbb{E}_{it} \left[\theta_{t-1} + \frac{(1 - \alpha)A_u u_t + A_p(p_t - \theta_{t-1})}{1 - \alpha A_u} \right] \text{ (from (19))} \\
 &= \theta_{t-1} + (1 - \alpha)\mathbb{E}_{it}[u_t] + \alpha \frac{(1 - \alpha)A_u \mathbb{E}_{it}[u_t] + A_p(p_t - \theta_{t-1})}{1 - \alpha A_u} \\
 &= \theta_{t-1} + \frac{1 - \alpha}{1 - \alpha A_u} \mathbb{E}_{it}[u_t] + \alpha \frac{A_p(p_t - \theta_{t-1})}{1 - \alpha A_u} \\
 &= \theta_{t-1} + \frac{1 - \alpha}{1 - \alpha A_u} (A_u(u_t + \epsilon_{it}) + A_p(p_t - \theta_{t-1})) \\
 &\quad + \alpha \frac{A_p}{1 - \alpha A_u} (p_t - \theta_{t-1}) \text{ (from Equations (13) and (15))} \\
 &= \theta_{t-1} + \frac{(1 - \alpha)A_u}{1 - \alpha A_u} u_t + \frac{(1 - \alpha)A_u}{1 - \alpha A_u} \epsilon_{it} + \frac{A_p}{1 - \alpha A_u} (p_t - \theta_{t-1}).
 \end{aligned}$$

Thus,

$$\begin{aligned}
 a_{it} - \theta_t &= \frac{(1-\alpha)A_u}{1-\alpha A_u} u_t + \frac{(1-\alpha)A_u}{1-\alpha A_u} \epsilon_{it} + \frac{A_p}{1-\alpha A_u} (p_t - \theta_{t-1}) - u_t \\
 &= \frac{(A_p + A_u - 1)u_t + A_p \eta_t + (1-\alpha)A_u \epsilon_{it}}{1-\alpha A_u} \text{ where we used } p_t - \theta_{t-1} = u_t + \eta_t.
 \end{aligned}$$

Thus,

$$\begin{aligned}
 (a_{it} - \theta_t)^2 &= \frac{(A_p + A_u - 1)^2 u_t^2 + A_p^2 \eta_t^2 + (1 - \alpha)^2 A_u^2 \epsilon_{it}^2}{(1 - \alpha A_u)^2} \\
 &\quad + \frac{2(A_p + A_u - 1)u_t A_p \eta_t + 2(A_p + A_u - 1)u_t (1 - \alpha)A_u \epsilon_{it} + 2A_p \eta_t (1 - \alpha)A_u \epsilon_{it}}{(1 - \alpha A_u)^2}.
 \end{aligned}$$

Thus,

$$\int (a_{it} - \theta_t)^2 di = \frac{(A_p + A_u - 1)^2 \sigma_u^2 + A_p^2 \sigma_\eta^2 + (1 - \alpha)^2 A_u^2 \sigma_\epsilon^2 + 2(A_p + A_u - 1)A_p u_t \eta_t}{(1 - \alpha A_u)^2}.$$

Thus,

$$E \left[\int (a_{it} - \theta_t)^2 di \right] = \frac{(A_p + A_u - 1)^2 \sigma_u^2 + A_p^2 \sigma_\eta^2 + (1 - \alpha)^2 A_u^2 \sigma_\epsilon^2}{(1 - \alpha A_u)^2}, \tag{23}$$

where we recognize that if $\alpha = 0$, Equation (23) simplifies to $\frac{\sigma_u^2 \sigma_\eta^2 \sigma_\epsilon^2}{\sigma_u^2 \sigma_\eta^2 + \sigma_u^2 \sigma_\epsilon^2 + \sigma_\eta^2 \sigma_\epsilon^2}$, which is the variance of $\theta_t | \theta_{t-1}, p_t, x_{it}$. Differentiating with respect to σ_η^2 yields:

$$\frac{dE[\int (a_{it} - \theta_t)^2 di]}{d\sigma_\eta^2} = \frac{\sigma_\epsilon^4 \sigma_u^4 (\sigma_\epsilon^2 \sigma_u^2 + \sigma_\epsilon^2 \sigma_\eta^2 + (1 - \alpha)(1 - 2\alpha)\sigma_u^2 \sigma_\eta^2)}{(\sigma_\epsilon^2 \sigma_u^2 + \sigma_\eta^2 \sigma_\epsilon^2 + (1 - \alpha)\sigma_\eta^2 \sigma_u^2)^3}. \tag{24}$$

Thus, if $\sigma_\epsilon^2 + (1 - \alpha)(1 - 2\alpha)\sigma_u^2 \geq 0$ (in particular, if $\alpha \leq 1/2$), the above derivative is strictly positive. If, instead, $\sigma_\epsilon^2 + (1 - \alpha)(1 - 2\alpha)\sigma_u^2 < 0$, the above derivative is strictly positive if and only if σ_η^2 is sufficiently small. \square

Proof of Proposition 8: To obtain $E[\int (a_{it} - \theta_t)^2 di]$ with only x'_{it} , first let $\sigma_\eta \rightarrow \infty$ in Equation (23), and then substitute σ_ϵ^2 with $\sigma_{\epsilon'}^2 = \frac{\sigma_\epsilon^2 \sigma_\eta^2}{\sigma_\epsilon^2 + \sigma_\eta^2}$. Using Equation (16), and recognizing that $\lim_{\sigma_\eta \rightarrow \infty} A_p \sigma_\eta^2 = 0$, the first step yields:

$$\lim_{\sigma_\eta \rightarrow \infty} \frac{(\beta - 1)^2 \sigma_u^2 + (1 - \alpha)^2 \beta^2 \sigma_\epsilon^2}{(1 - \alpha \beta)^2} = \frac{\sigma_\epsilon^2 \sigma_u^2 (\sigma_\epsilon^2 + (1 - \alpha)^2 \sigma_u^2)}{(\sigma_\epsilon^2 + (1 - \alpha)\sigma_u^2)^2}.$$

Substituting σ_ϵ^2 with $\sigma_{\epsilon'}^2$, yields:

$$\frac{\sigma_\epsilon^2 \sigma_u^2 (\sigma_{\epsilon'}^2 + (1 - \alpha)^2 \sigma_u^2)}{(\sigma_{\epsilon'}^2 + (1 - \alpha)\sigma_u^2)^2}. \tag{25}$$

Now, subtracting Equation (23) from Equation (25) yields:

$$\begin{aligned} \Delta &= E \left[\int (a_{it} - \theta_t)^2 di \right]_{(x'_{it})} - E \left[\int (a_{it} - \theta_t)^2 di \right]_{(x_{it}, p_t)} \\ &= \frac{\alpha^2 \sigma_\eta^4 \sigma_\epsilon^4 \sigma_u^4}{(\sigma_\eta^2 \sigma_\epsilon^2 + (1 - \alpha)(\sigma_\eta^2 + \sigma_\epsilon^2)\sigma_u^2)(\sigma_\eta^2 \sigma_\epsilon^2 + \sigma_\eta^2(\sigma_\epsilon^2 + (1 - \alpha)\sigma_u^2))^2} \\ &\quad \times (\sigma_\epsilon^4 (\sigma_\eta^2 + \sigma_u^2) - (1 - \alpha)^2 (\sigma_\eta^2 + \sigma_\epsilon^2)\sigma_u^4). \end{aligned}$$

As expected, $\lim_{\alpha \rightarrow 0} \Delta = 0$, because only the amount of information matter; and $\lim_{\alpha \rightarrow 1} \Delta > 0$, because then citizens put a lot of weight on the pre-existing public information from the previous period, which needs to be countered by new public information about θ_t . Moreover, for any $\alpha > 0$, the setting with both public and private signals (x_{it}, p_t) is a normative improvement over the setting with only private signals (x'_{it}) if and only if $\Delta > 0$, i.e., if and only if

$$\left(\frac{\sigma_\epsilon^2}{\sigma_u^2} \right)^2 > (1 - \alpha)^2 \frac{\sigma_\eta^2 + \sigma_\epsilon^2}{\sigma_\eta^2 + \sigma_u^2}.$$

The result follows from inspection of this inequality. \square

Proof of Proposition 9: In common values settings, from (9), $\bar{\mathbb{E}}_t[\theta_t] = \theta_{t-1} + \beta u_t$, and $\bar{\mathbb{E}}_t^2[\theta_t] = \theta_{t-1} + \beta^2 u_t$. From Proposition 1, $A_t = \theta_{t-1} + \phi u_t$, and hence from Equation (8), $\bar{\mathbb{E}}_t[A_t] = \theta_{t-1} + \phi \beta u_t$. Thus, (a) $\bar{\mathbb{E}}_t[\theta_t] - A_t = (\beta - \phi)u_t > 0$ (from Proposition 1) and $\bar{\mathbb{E}}_t[\theta_t] - \bar{\mathbb{E}}_t^2[\theta_t] = \beta(1 - \beta)u_t$, (b) $\bar{\mathbb{E}}_t^2[\theta_t] - A_t = (\beta^2 - \phi)u_t = \beta(\beta - \phi/\beta)u_t = \beta(\beta - \phi^{Pr})u_t > 0$, where the third equality follows from Proposition 3, and the inequality holds if and only if $\alpha(1 + \beta) > 1$, and (c) $A_t - \bar{\mathbb{E}}_t[A_t] = \phi(1 - \beta)u_t = \beta\phi^{Pr}(1 - \beta)u_t$.

In private values settings, mirroring the above calculations, $E_{it}[x_{it}] = x_{it}$, $\bar{\mathbb{E}}_t[x_{it}] = \theta_t$, $\bar{\mathbb{E}}_t^2[x_{it}] = \theta_{t-1} + \beta u_t$, $A_t = \theta_{t-1} + \phi^{Pr} u_t$, and $\bar{\mathbb{E}}_t[A_t] = \theta_{t-1} + \phi^{Pr} \beta u_t$. Thus, (ap) $\bar{\mathbb{E}}_t[x_{it}] - A_t = (1 - \phi^{Pr})u_t > 0$ (from Proposition 2) and $\bar{\mathbb{E}}_t[\theta_t] - \bar{\mathbb{E}}_t^2[\theta_t] = (1 - \beta)u_t$, (bp) $\bar{\mathbb{E}}_t^2[\theta_t] - A_t = (\beta - \phi^{Pr})u_t > 0$ if and only if $\alpha(1 + \beta) > 1$, and (cp) $A_t - \bar{\mathbb{E}}_t[A_t] = \phi^{Pr}(1 - \beta)u_t$. Comparing (a)–(c) with (ap)–(cp) shows that the differences are smaller in common values settings than in private values settings by a factor of β . \square

References

- Acemoglu, D. and M. O. Jackson. 2015. “History, Expectations, and Leadership in the Evolution of Social Norms”. *Review of Economic Studies*. 82: 1–34.
- Acemoglu, D. and M. O. Jackson. 2017. “Social Norms and the Enforcement of Laws”. *Journal of the European Economic Association*. 12: 245–95.
- Acemoglu, D. and A. Wolitzky. 2014. “Cycles of Conflict: An Economic Model”. *American Economic Review*. 104: 1350–67.
- Ajzen, I. 2001. “Nature and Operation of Attitudes”. *Annual Review of Psychology*. 52(1): 27–58.
- Allcott, H., L. Boxell, J. Conway, M. Gentzkow, M. Thaler, and D. Yang. 2020. “Polarization and Public Health: Partisan Differences in Social Distancing during the Coronavirus Pandemic”. *Journal of Public Economics*. 191: 1–11.
- Angeletos, G.-M., C. Hellwig, and A. Pavan. 2007. “Dynamic Global Games of Regime Change: Learning, Multiplicity, and the Timing of Attacks”. *Econometrica*. 75: 711–56.
- Angeletos, G.-M. and J. La’O. 2010. “Noisy Business Cycles”. In: *NBER Macroeconomics Annual 2009, Volume 24*. University of Chicago Press, 319–78.
- Angeletos, G.-M. and J. La’O. 2013. “Sentiments”. *Econometrica*. 81: 739–79.
- Angeletos, G.-M. and C. Lian. 2016. “Incomplete Information in Macroeconomics: Accommodating Frictions in Coordination”. *Handbook of Macroeconomics*. 2: 1065–240.
- Armitage, C. J. and M. Conner. 2001. “Efficacy of the Theory of Planned Behaviour: A Meta-Analytic Review”. *British Journal of Social Psychology*. 40(4): 471–99.

- Bernhardt, A., E. Field, R. Pande, N. Rigol, S. Schaner, and C. Troyer-Moore. 2018. "Male Social Status and Women's Work". *AEA Papers and Proceedings*. 108: 363–67.
- Bolton, P., M. K. Brunnermeier, and L. Veldkamp. 2013. "Leadership, Coordination and Corporate Culture". *Review of Economic Studies*. 80: 512–37.
- Borsari, B. and K. B. Carey. 2003. "Descriptive and Injunctive Norms in College Drinking: A Meta-Analytic Integration". *Journal of Studies on Alcohol*. 64(3): 331–41.
- Bursztyn, L., A. L. González, and D. Yanagizawa-Drott. 2020. "Misperceived Social Norms: Women Working Outside the Home in Saudi Arabia". *American Economic Review*. 100(4): 1358–98.
- Chen, Y., M. Harper, J. Konstan, and S. X. Li. 2010. "Social Comparisons and Contributions to Online Communities: A Field Experiment on MovieLens". *American Economic Review*. 110(10): 2997–3029.
- Christensen, D., O. Dube, J. Haushofer, B. Siddiqi, and M. Voors. 2021. "Building Resilient Health Systems: Experimental Evidence from Sierra Leone and the 2014 Ebola Outbreak". *Quarterly Journal of Economics*. 136(2): 1145–98.
- Chwe, M. S.-Y. 2013. *Rational Ritual: Culture, Coordination, and Common Knowledge*. Princeton University Press.
- Cialdini, R. R., L. J. Demaine, B. J. Sagarin, D. W. Barrett, K. Rhoads, and P. L. Winters. 2006. "Activating and Aligning Social Norms for Persuasive Impact". *Social Influence*. 1(1): 3–15.
- Dewan, T. and D. P. Myatt. 2008. "The Qualities of Leadership: Direction, Communication, and Obfuscation". *American Political Science Review*. 102: 351–68.
- Dube, O. and K. Baicker. 2020. "How You Can Protect Your Community, Not Just Your Own Health". *New York Times*, March 26, 2020.
- Elster, J. 2011. "Norms". In: *The Oxford Handbook of Analytical Sociology*. Ed. P. Bearman and P. Hedström. Oxford University Press.
- Field, E., R. Pande, N. Rigol, S. Schaner, and C. T. Moore. 2021. "On Her Own Account: How Strengthening Women's Financial Control Impacts Labor Supply and Gender Norms". *American Economic Review*. 111(7): 2342–75.
- Fishbein, M. and I. Ajzen. 2011. *Predicting and Changing Behavior: The Reasoned Action Approach*. New York: Taylor & Francis.
- Frey, B. S. and S. Meier. 2004. "Social Comparisons and Pro-social Behavior: Testing 'Conditional Cooperation' in a Field Experiment". *American Economic Review*. 94(5): 1717–22.
- Gerber, A. S. and T. Rogers. 2009. "Descriptive Social Norms and Motivation to Vote: Everybody's Voting and So Should You". *Journal of Politics*. 71(1): 178–91.

- Goldstein, N., R. B. Cialdini, and V. Griskevicius. 2008. "A Room with a Viewpoint: Using Norm-Based Appeals to Motivate Conservation Behaviors in a Hotel Setting". *Journal of Consumer Research*. 35(3): 472–82.
- Hanschmidt, F., K. Linde, A. Hilbert, S. G. Riedel-Heller, and A. Kersting. 2016. "Abortion Stigma: A Systematic Review". *Perspectives on Sexual and Reproductive Health*. 48(4): 169–77.
- Huo, Z. and M. Pedroni. 2020. "A Single-Judge Solution to Beauty Contests". *American Economic Review*. 110: 526–68.
- Jordan, J. J., E. Yoeli, and D. G. Rand. 2020. "Don't Get it or Don't Spread it? Comparing Self-interested versus Prosocially Framed COVID-19 Prevention Messaging". PsyArXiv, 3 April. <https://doi.org/10.31234/osf.io/yuq7x>.
- Lorenzoni, G. 2010. "Optimal Monetary Policy with Uncertain Fundamentals and Dispersed Information". *Review of Economic Studies*. 77: 305–38.
- Moehring, A., A. Collis, K. Garimella, A. Rahimian, S. Aral, and D. Eckles. 2021. "Surfacing Norms to Increase Vaccine Acceptance". Available at SSRN: <https://ssrn.com/abstract=3782082>.
- Morris, S. and H. S. Shin. 2002. "The Social Value of Public Information". *American Economic Review*. 92: 1521–34.
- Morris, S. and H. S. Shin. 2006. "Inertia of Forward-Looking Expectations". *AEA Papers & Proceedings*. 96: 152–7.
- Morris, S. and H. S. Shin. 2007. "Optimal Communication". *Journal of the European Economic Association*. 5: 594–602.
- Myerson, R. 1991. *Game Theory: Analysis of Conflict*. Boston, MA: Harvard University Press.
- Postlewaite, A. 2011. "Social Norms and Preferences". In: *Handbook of Social Economics, Volume 1A*. Ed. J. Benhabib, A. Bisin, and M. Jackson. Amsterdam: North-Holland.
- Reno, R. R., R. B. Cialdini, and C. A. Kallgren. 1993. "The Transsituational Influence of Social Norms". *Journal of Personality and Social Psychology*. 64(1): 104–12.
- Rivis, A. and P. Sheeran. 2003. "Descriptive Norms as an Additional Predictor in the Theory of Planned Behaviour: A Meta-Analysis". *Current Psychology*. 22(3): 218–33.
- Shadmehr, M., S. Shahshahani, and C. Cameron. 2022. "Coordination and Innovation in Judiciaries: Correct Law vs. Consistent Law". *Quarterly Journal of Political Science*. 17: 61–89.