



Management Science

Publication details, including instructions for authors and subscription information:
<http://pubsonline.informs.org>

Information Sampling, Belief Synchronization, and Collective Illusions

Jerker Denrell, Gaël Le Mens

To cite this article:

Jerker Denrell, Gaël Le Mens (2017) Information Sampling, Belief Synchronization, and Collective Illusions. Management Science 63(2):528-547. <http://dx.doi.org/10.1287/mnsc.2015.2354>

Full terms and conditions of use: <http://pubsonline.informs.org/page/terms-and-conditions>

This article may be used only for the purposes of research, teaching, and/or private study. Commercial use or systematic downloading (by robots or other automatic processes) is prohibited without explicit Publisher approval, unless otherwise noted. For more information, contact permissions@informs.org.

The Publisher does not warrant or guarantee the article's accuracy, completeness, merchantability, fitness for a particular purpose, or non-infringement. Descriptions of, or references to, products or publications, or inclusion of an advertisement in this article, neither constitutes nor implies a guarantee, endorsement, or support of claims made of that product, publication, or service.

Copyright © 2016, INFORMS

Please scroll down for article—it is on subsequent pages



INFORMS is the largest professional society in the world for professionals in the fields of operations research, management science, and analytics.

For more information on INFORMS, its publications, membership, or meetings visit <http://www.informs.org>

Information Sampling, Belief Synchronization, and Collective Illusions

Jerker Denrell,^a Gaël Le Mens^b

^a Warwick Business School, University of Warwick, Coventry CV4 7AL, United Kingdom; ^b Department of Economics and Business, Barcelona School of Management and Barcelona Graduate School of Economics, Universitat Pompeu Fabra, 08005 Barcelona, Spain

Contact: jerker.denrell@wbs.ac.uk (JD); gael.le-mens@upf.edu (GLM)

Received: August 3, 2013

Accepted: August 4, 2015

Published Online in Articles in Advance:
March 22, 2016

<https://doi.org/10.1287/mnsc.2015.2354>

Copyright: © 2017 INFORMS

Abstract. We demonstrate that a sampling-based mechanism can offer an alternative explanation for belief synchronization in social groups and the persistence of collective illusions. Our model assumes that people are more likely to sample popular alternatives than unpopular alternatives. We show that this mechanism is sufficient to explain belief synchronization: a strong majority of opinions will likely emerge in favor of one alternative. The reason is that the group is unlikely to move away from a state in which one alternative is very unpopular. If by chance most people come to dislike alternative A, they are all unlikely to sample it again and their opinions of A remain the same. When A is in fact the best alternative, a collective illusion has emerged because people mistakenly believe that a suboptimal alternative is the best. Our model implies that such a collective illusion is persistent. The model thus offers an existence proof that a collective illusion can occur even in settings where people do not infer that popular alternatives are better. The model also casts new light on the effect of online recommendation systems on attitude homogenization and the effect of majority voting on beliefs and attitudes.

History: Accepted by Yuval Rottenstreich, judgment and decision making.

Funding: G. Le Mens benefited from financial support from the Barcelona School of Management, a Juan de la Cierva Fellowship from the Spanish Ministry of Science and Innovation, and the Spanish Ministry of Science and Innovation [Grants ECO2010-17145 and PSI2013-41909-P].

Keywords: status quo • change • learning • social influence • popularity • sampling

1. Introduction

Explaining why collective mistakes emerge and persist is central to understanding social change, immobility, and differences in welfare (Bowles 2004, Elster 1978, North 1990). In developing countries, people keep on using poor domestic hygiene practices even though simple changes would save many lives (Curtis et al. 2000), and farmers fail to use fertilizer despite their potential for large increases in productivity (Duflo et al. 2011). In Western countries, firms persist in making abundant use of temporary workforces despite the inefficiency of this arrangement (Pfeffer 1998), and medical practices and treatments tend to cluster geographically, sometimes in favor of suboptimal alternatives (Coleman et al. 1966, O'Connor et al. 1999, Krumholz et al. 2003). How come groups of people could persist in using inefficient technologies, practices, or institutions when better courses of actions are available?

Prior literature has shown that such mistakes can occur when there are network externalities (Arthur 1989, Elster 1978, North 1990, Pierson 2004) or when a collective illusion emerges in which individuals mistakenly infer that the most popular but suboptimal alternative is the best (Banerjee 1992, Bikhchandani

et al. 1992, Cialdini and Goldstein 2004, Festinger 1957). In the first case, people are likely to herd, i.e., select the popular alternative, even if they know it is inferior to an unpopular alternative because it is difficult to coordinate a switch to the superior alternative. In the second case, people are assumed to make inferences about quality on the basis of popularity. More precisely, if a suboptimal alternative has, by chance, become the most popular one, people who have limited information about the qualities of the alternatives will mistakenly infer that the suboptimal but popular alternative is superior to other, less popular alternatives. This leads to herd behavior where most people chose the same alternative. Such inference can be rational, as in information cascade models (Banerjee 1992, Bikhchandani et al. 1992), but it can also be the product of *belief adjustments*, whereby people tend to process information about popular alternatives in a systematically different fashion from how they process information about unpopular alternatives (Bem 1972, Elster 1983, Festinger 1957, Heider 1958, Samuelson and Zeckhauser 1988).

Here, we demonstrate theoretically that an alternative explanation for collective illusions is possible, an explanation that builds on the sampling perspective

on belief formation (Denrell 2005, Fiedler and Juslin 2006). We consider settings where people learn about the qualities of the alternatives from experience. In contrast to prior work, we assume that popularity does not impact inferences about qualities, but only impacts sampling. In particular, we assume that due to conformity pressures or network externalities, people are more likely to sample alternatives the majority believes to be superior. We demonstrate theoretically that this mechanism, by itself, is sufficient to explain the emergence and persistence of collective illusions. Specifically, we show that if people tend to select popular alternatives, beliefs tend to become synchronized: a strong majority of people will tend to believe that one or the other alternative is the best, even if the qualities of the alternatives are the same. When a strong majority emerges in favor of the inferior alternative, a *collective illusion* has emerged. Our model shows that such a collective illusion is persistent in the sense that it tends to last for a long time.

Importantly, our theoretical results hold even if people do not engage in belief adjustments of the sort described above nor make inferences about quality on the basis of popularity. Our assumptions are thus very different from the assumptions of information cascade models. These models assume that people select the alternative that the majority of others has selected *because* of the inferences they make on the basis of popularity (Bikhchandani et al. 1992). Our model thus provides an existence proof that beliefs can become synchronized even in settings where people are reluctant to infer that what is popular is best for them, such as when they have different tastes or skills.

The intuition behind our results is that when popularity impacts sampling, a population is unlikely to move away from a state in which one alternative is very unpopular compared to other alternatives. If by chance most people come to dislike alternative A, they are all unlikely to sample it again and their opinions of A remain the same. When A is in fact the best alternative, a collective illusion has emerged because people mistakenly believe that a suboptimal alternative is the best. Such a collective illusion is persistent because people avoid alternative A and do not correct their mistaken negative assessment of alternative A.

Our model builds upon prior work which has demonstrated how sampling can systematically impact belief formation when people learn from experience (Denrell 2005, Denrell and Le Mens 2007, Fiedler 2000, Fazio et al. 2004, Le Mens and Denrell 2011). The key mechanism in much of this literature is the asymmetry in experiential learning generated by adaptive sampling. To ensure favorable future outcomes, it makes sense to choose alternatives with favorable past outcomes and avoid alternatives with unfavorable past

outcomes. Such adaptive sampling implies that negative attitudes are more likely than positive attitudes to remain stable. If I dislike A, I am unlikely to sample A again, and my negative attitude toward A does not change. Prior literature has used this mechanism to offer alternative explanations of a range of social psychological phenomena. For example, Denrell (2005) shows how in-group bias could be explained by adaptive sampling. The intuition is that attitudes will be more positive toward people you are likely to meet irrespective of your attitude toward them. If I dislike B, but B is a friend of my friends, I might meet B again. As a result, my negative attitude toward B can change. This would not happen if B were not part of my social network. The key mechanism is that attitudes will become more positive for alternatives the decision maker is “forced” to sample.

In this paper we consider a population of individuals who are all more likely to select alternatives that many others like. As in prior sampling models, sampling behavior is determined by outcomes in prior periods. Unlike prior models, however, there is no inbuilt bias toward sampling a particular set of alternatives (such as people belonging to the in-group). Rather, the sampling bias emerges endogenously (it is determined by what people sample and what outcomes they receive). We ask whether and when such an endogenously emerging sampling bias can generate a persistent collective illusion. Our model thus differs from prior models both conceptually (no inbuilt sampling bias in favor of a specific alternative) and in the focus on a collective (rather than individuals). Analyzing a model with an endogenously emerging sampling bias also requires mathematical techniques different from what was used in prior literature.

In the following, we first explain the rationale behind our assumption that people tend to select alternatives that many others like while not using popularity as a signal of quality nor engaging in belief adjustments. In Section 3, we analyze a simple learning model that gives rise to an information bias that leads to synchronized beliefs and to the emergence of collective illusions. In Section 4, we discuss some extensions to more realistic assumptions regarding choices and belief updating mechanisms. In Section 5 we discuss how our theory relates to prior research on the emergence of collective illusions. In Section 6, we show how our model can be adapted to explain how majority voting can lead to attitude homogenization within groups. Finally, in Section 7, we discuss the difficulties associated with empirically disentangling the sampling effect at the core of our approach and the effect of popularity on quality estimates that is mediated by popularity-based inferences like those at the core of herding models.

2. Majority Influence and Information Sampling

Much prior research on why people might come to agree with a mistaken majority assumes that people change their beliefs to conform to the majority opinion. Theories of social learning in economics and of majority influence in psychology, for example, assume that people use popularity as an indication of quality: if most others like alternative A better than alternative B, people infer that alternative A is likely superior (Banerjee 1992, Bikhchandani et al. 1992, Darley and Latané 1968, Deutsch and Gerard 1955, Prentice and Miller 1993). Such inference about quality on the basis of popularity is often sensible: if most others like and go to restaurant A and few others like and go to restaurant B, this suggests that A offers superior food and service. An abundant literature in economics has shown that, at the population level, such inferences can lead to *information cascades* in which most agents come to mistakenly believe that a suboptimal alternative is optimal (a collective illusion has emerged). A key insight of this literature is that such information cascades, associated collective illusions and resulting herd behavior can be the outcome of rational behavior (Banerjee 1992, Bikhchandani et al. 1992; for a review, see Chamley 2004).

Although popularity-based inferences are often sensible, here we wish to demonstrate that such inferences are *not necessary* to explain the synchronization of opinions in a population. For this purpose, our model deliberately excludes such inferences. Specifically, we do not assume that people change their beliefs when hearing about the behaviors of others: in our model, people's quality estimates are based *only* on the information they obtain by sampling the alternatives. We assume, however, that people are subject to majority influence in the sense that they are more likely to *select and thus sample* an activity that is popular, i.e., believed to be of superior quality by most other people. Thus, our model assumes majority influence, but it does not assume a direct effect of popularity on beliefs (beliefs depend only on one's information sample).

Why assume that the majority only impacts sampling behavior and not beliefs? In cases where popularity-based inferences are important, this assumption may not be realistic since popularity then likely impacts both behavior and beliefs. For theoretical purposes, however, it is useful to examine the case where popularity only influences behavior. We will show that this assumption, combined with individual learning from experience, is sufficient to explain emergence of belief synchronization. This strategy that consists in assuming away intrapsychic belief adjustments is standard in the research tradition in which our paper fits (Fiedler and Juslin 2006, Juslin et al. 2007). Like other information sampling models, our theory

emphasizes the role of the environment and how it can sometimes produce systematically biased samples of information that affect judgments (for a review, see Fiedler and Juslin 2006; for recent publications that adopt this perspective, see Denrell and Le Mens 2011; Eder et al. 2011; Feiler et al. 2013; Fiedler 2008, 2011, 2012; Henriksson et al. 2010; Le Mens and Denrell 2011; Smith and Collins 2009).

There are also many realistic situations where one might expect to see the majority impact sampling behavior although people do not engage in popularity-based inferences. Prior studies have shown that people might publicly adhere to the local norm, by stating an opinion in agreement with the majority or choosing the action chosen by most others, while privately disagreeing or having different preferences. In the Asch (1956) conformity experiments, for example, many participants did not believe that the majority was correct, but nevertheless stated an opinion in agreement with the majority opinion to avoid being seen as a deviant. More generally, studies have shown that people who deviate from the majority opinion are less liked (Levine 1980) and also expect to be disliked (Gerard and Rotter 1961), so it is hardly surprising that people conform publicly even if they privately are of a different opinion. People may also go along with the majority because there is "safety in numbers" in the sense that it can be costly to choose an activity few others will choose (Granovetter 1978, Schelling 1978). For example, a research student will find it easier to get help if she chooses a research method used by many faculty and students in her department. The advantage of having help around may motivate her to choose this method even if she does not believe that it is overall superior. People might also choose the popular alternative because "it is better for reputation to fail conventionally than to succeed unconventionally" (Keynes 1936, p. 158), i.e., because of the adverse reputation effects of receiving a poor outcome with an unusual alternative (Scharfstein and Stein 1990, Zwiebel 1995). Finally, people might select popular alternatives because they are more easily available or visible. For example, people are more likely to be aware of popular songs because they are played more often on radio, etc. Popular alternatives are often also more prominently displayed on websites (such as Google, Spotify, etc.).

Besides, although popularity-based inferences are often sensible, there are settings in which it is doubtful whether the majority knows best or knows what is good for you. People may have unique tastes (Berger and Heath 2007): if others prefer opera to jazz, this may not be enough to convince you that opera is superior. But you might nevertheless go to the opera because you want to be able to talk about it with your friends. Individuals may also suspect that the experiences of others are not relevant because their

circumstances are different. For example, because of wide variations in soil conditions, rice growers during the Indian green revolution were hesitant to adopt techniques simply because they had worked well at neighboring plots (Munshi 2004). Rather than imitating others, they relied on their own experience. Such hesitance to rely on the experience of others is likely common and also sensible especially when the experiences of others contradict your own experience: the fact that the experiences of others differ from yours indicates that your circumstances may differ substantially. The possibility of different circumstances was also one of the explanations that participants in the Asch (1956) experiments came up with to make sense of the behavior of the majority: the fact that the majority disagreed with them indicated that the majority interpreted the question differently or viewed the lines from a different angle. Individuals may also rely more than they should on their own experience simply because it is salient (Simonsohn et al. 2008).

These arguments suggest that there are settings in which popularity-based inferences are unlikely to be important, but people might nevertheless select the popular alternative to avoid being seen as deviant, to coordinate their choices with others, or because popular alternatives are more easily available. In these cases, popularity will influence sampling but not beliefs once sampling is controlled for. A good illustration is the MusicLab experiment (Salganik et al. 2006). In this experiment, thousands of participants could listen to songs by streaming them and then decide to download (or not) the songs to which they had listened. Songs that had been downloaded by many participants (i.e., the popular songs) were more accessible in that they were displayed at the top of the computer screen (Experiment 2). Recent reanalyses of the experimental data suggest that participants made their listening (i.e., sampling) decisions based on popularity. But their evaluations of the songs were not driven by social influence: there was no effect of popularity once sampling was controlled for (Hendricks et al. 2012, Krumme et al. 2012).

3. Model

We consider a population of n people who learn about the qualities of two uncertain alternatives from the private information they obtain by sampling the alternatives. In each period, one individual is randomly picked in the population. This individual selects one of the two alternatives, observes its payoff, and updates her estimate of the quality of this alternative accordingly. We implement the effect of network externalities by assuming that the individual is more likely to sample the alternative that is believed by most to be the superior one. We focus on the essential features of the

learning and choice processes that can give rise to collective illusions. Therefore, we keep our model as simple as possible in this section. In Section 4 we discuss the assumptions of the model and what happens if we make different assumptions.

3.1. Model Specification

Our model is based on the following assumptions.

Payoffs of the Alternatives. The two alternatives potentially differ in “quality” or “efficiency.” The two alternatives have payoffs that follow Bernoulli distributions: the possible payoffs are 0 (failure) and 1 (success). Let p_1 be the probability of success of alternative 1 and p_2 the probability of success of alternative 2. We will say that alternative 1 is superior (inferior) when $p_1 > p_2$ ($p_1 < p_2$). An important feature of the model is that the success probabilities, p_1 and p_2 , are *unknown* to the decision makers.

Quality Estimates. Estimates of the qualities of the alternatives are binary: an individual believes an alternative is either “good” or “bad” (and so she can believe that both are “good,” both are “bad,” or one is “good” while the other is “bad”). The estimate of alternative 1 (alternative 2) for individual i at the beginning of period t will be denoted by $\hat{Q}_{1,t}^i$ ($\hat{Q}_{2,t}^i$). If the estimate of alternative 1 (alternative 2) is “good,” we will write $\hat{Q}_{1,t}^i = 1$ ($\hat{Q}_{2,t}^i = 1$), and if the estimate is “bad,” we will write $\hat{Q}_{1,t}^i = 0$ ($\hat{Q}_{2,t}^i = 0$). Initial estimates are random draws from the payoff distributions. More precisely, for a given individual, the likelihood that the initial estimate of alternative 1 is “good” is p_1 . The likelihood that the initial estimate is “bad” is $1 - p_1$. If $p_1 = 0.5$, then an individual initially believes that alternative 1 is “good” or “bad” with equal probability.

Sampling Rule. In each period, one individual is randomly picked in the population (each individual is equally likely to be picked). This individual makes a choice between the two available alternatives. The crucial assumption is the following: we assume that the individual is *more likely* to sample the popular alternative than the unpopular alternative. Specifically, let $N_{1,t}$ be the number of individuals who have a “good” estimate for alternative 1 at the beginning of period t , and let $N_{2,t}$ be the number of individuals who have a “good” estimate for alternative 2 at the beginning of period t .¹ We assume that the likelihood that an individual chooses the first alternative is

$$pC_1(n_{1,t}, n_{2,t}) = \frac{1}{1 + e^{-s(n_{1,t} - n_{2,t})}}, \quad (1)$$

where $s > 0$ is the sensitivity to popularity. We set $pC_2(n_{1,t}, n_{2,t}) = 1 - pC_1(n_{1,t}, n_{2,t})$. This logistic choice rule is often used to model choice under uncertainty (Erev and Barron 2005, Sutton and Barto 1998) and provides good fit to experimental data on sequential

choices (Bussemeyer and Myung 1992, Denrell 2005). It also fits data on the effect of the number of influence sources on conformity behavior (MacCoun 2012).²

Estimate Updating Rule. We assume that quality estimates are based on the most recent observation. If individual i samples alternative k and gets a success, her estimate becomes “good” (e.g., for alternative 1, $\hat{Q}_{1,t}^i = 1$). If she gets a failure, her estimate becomes “bad” (e.g., for alternative 1, $\hat{Q}_{1,t}^i = 0$). We make this assumption to keep the model as transparent as possible, but the basic results continue to hold if individuals use a weighted average estimate updating rule, as shown in Section 4.1. Although this assumption that only the most recent observation matters might seem simplistic, recent experimental evidence on sequential choice between uncertain alternatives with binary outcomes shows that choices are well predicted by a model that focuses on the most recent outcome only, like ours (Avraami and Kareev 2010).

3.2. Results

In what follows, we first present an analysis of the asymptotic behavior of the model because the such an analysis leads to easily interpretable formulas that make it easier to understand the pattern of behavior implied by the model. Then we demonstrate, using numerical computations, that our main results emerge after a reasonable number of periods.

3.2.1. Asymptotic Analysis. To demonstrate the synchronization of quality estimates, we computed the stationary distribution of the popularities of the two alternatives. Let N_1 and N_2 denote the random variables toward which $N_{1,t}$ and $N_{2,t}$ converge, and let $\tilde{f}_t(n_1, n_2)$ denote the joint stationary distribution of the popularities of N_1 and N_2 .

Proposition 1 (Joint Stationary Distribution of Popularities). *When sampling follows the logistic choice rule (Equation (1)), the stationary distribution of the popularities of the two alternatives, N_1 and N_2 , is*

$$\tilde{f}(n_1, n_2) = K(e^{-sn_1} + e^{-sn_2})(p_1)^{n_1}(1-p_1)^{n-n_1} \cdot (p_2)^{n_2}(1-p_2)^{n-n_2} \binom{n}{n_1} \binom{n}{n_2}, \quad (2)$$

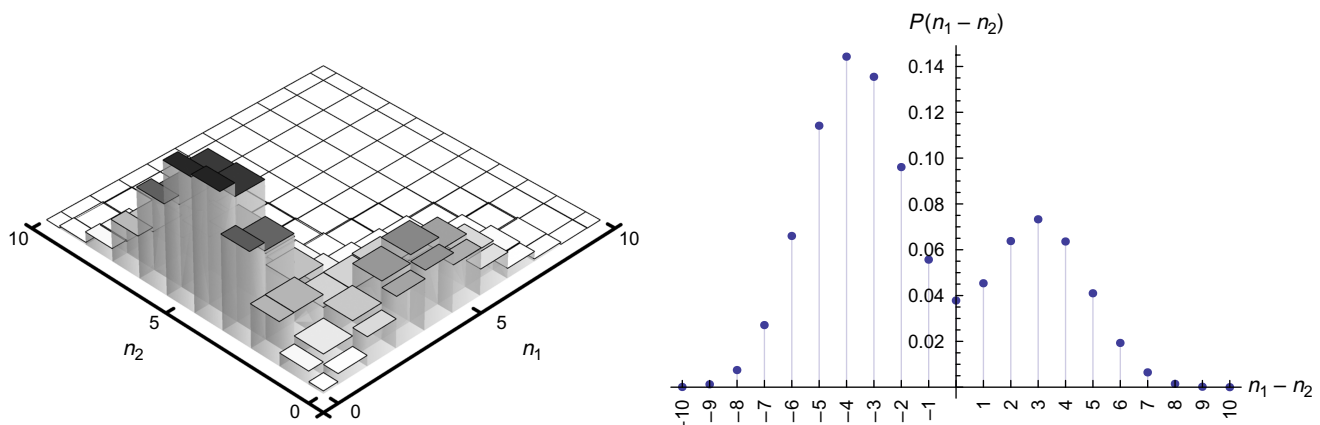
where $K^{-1} = (1-p_1+p_1e^{-s})^n + (1-p_2+p_2e^{-s})^n$.

Proof. The proofs are in the appendix.

This distribution implies that a strong majority favoring one or the other alternatives is likely to emerge (see Figure 1, left panel). To understand why, note that this is a *weighted bivariate binomial distribution*, where the weight is $K(e^{-sn_1} + e^{-sn_2})$. The weighting function puts additional weights on the extremes (where $N_k = 0$) and implies that the joint distribution of popularities will be bi-modal if s is high enough. The first mode is close to $(N_1 = p_1n, N_2 = 0)$: in this case, the proportion of people who like alternative 1 is p_1 , whereas the proportion of those who like alternative 2 is small. The other mode is located at $(N_1 = 0, N_2 = p_2n)$: in this case the proportion of people who like alternative 1 is small, whereas the proportion who like alternative 2 is p_2 .³

The right panel of Figure 1 shows the distribution of the difference in popularities, $N_1 - N_2$, which is also bimodal. One mode is above 0; this corresponds to cases where alternative 1 is the more popular. The other mode is below 0; this corresponds to cases where alternative 1 is the less popular. In the setting of the graph, alternative 1 is inferior to alternative 2 ($p_1 = 0.4$ and $p_2 = 0.45$). Unsurprisingly, the mode above 0 is lower than the mode below 0; alternative 1 is less likely than alternative 2 to become the more popular. But the probability that alternative 1 becomes more popular

Figure 1. (Color online) Features of the Stationary Distribution



Notes. Graphs were obtained with $n = 10$, $p_1 = 0.40$, $p_2 = 0.45$, and $s = 2$. Left panel: The stationary joint distribution of the popularities of alternative 1 and alternative 2 is bimodal. Right panel: The marginal distribution of $N_1 - N_2$ is also bimodal.

than alternative 2 is still substantial (the probability is 0.29) despite the fact that alternative 1 is the inferior alternative. When alternative 1 is more popular despite being inferior, a *collective illusion* has emerged: most people mistakenly believe alternative 1 to be superior to alternative 2. We formalize the definition of *collective illusion* as follows:

Definition 1 (Collective Illusion). We will say that the system is in a collective illusion state when either of the following two situations arises:

- $p_1 < p_2$ and $N_1 > N_2$,
- $p_1 > p_2$ and $N_1 < N_2$.

The intuition for Proposition 1 is that when popularity impacts sampling the system is unlikely to move away from a state in which one alternative is unpopular. This is related to the bimodal character of the asymptotic joint distribution of popularities. To explain this, suppose $p_1 < p_2$ and $N_1 = np_1$ while N_2 is low (i.e., $N_1 > N_2$). Thus, alternative 1 is the most popular, and alternative 2 is believed by many to be “bad.” The system can only exit from such a collective illusion if N_2 increases or N_1 decreases. These events are unlikely to happen, however.

To see why, note first that when alternative 2 is unpopular, it is unlikely to be selected. So, N_2 is unlikely to increase. Next, note that whereas alternative 1 is likely to be selected (because it is popular), N_1 is unlikely to decrease to a value much below np_1 . When alternative 1 is selected, the probability that N_1 decreases by one is $(N_1/n)(1 - p_1)$ whereas the probability that N_1 increases by one is $(1 - N_1/n)p_1$. An increase is more likely than a decrease whenever $N_1 < np_1$, a decrease is more likely if $N_1 > np_1$, and a decrease is as likely as an increase when $N_1 = np_1$. In other words, when alternative 1 is selected, N_1 tends to regress to its mean, which equals np_1 . Overall, this implies that the

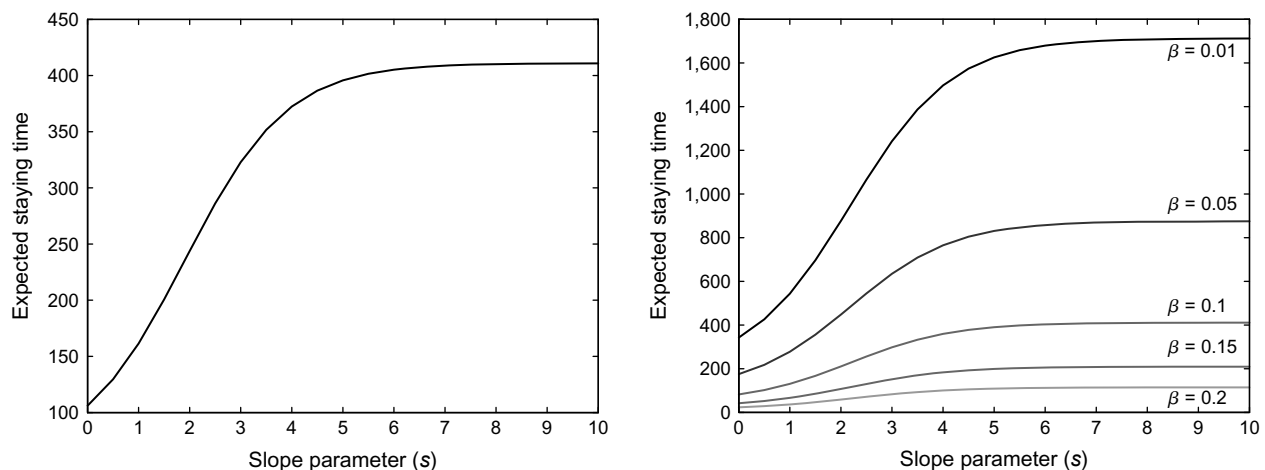
system is very slow to move away from the collective illusion in which $N_1 = np_1$ and N_2 is low.

This reasoning also shows that a collective illusion is a not temporary chance occurrence that will be quickly corrected. Rather, collective illusions produced by our model tend to be self-sustaining: once a collective illusion has emerged, the system is very slow to exit this state. To illustrate the persistence of collective illusions, we computed the expected time the system stays in the collective illusion state once a collective illusion has emerged. In what follows, we assume that alternative 2 is superior to alternative 1 ($p_1 < p_2$).⁴

The left panel of Figure 2 reports the results of these computations. The graph displays the expected number periods the collective illusion persists, given that we know that there is a collective illusion, as a function of the slope parameter of the logistic sampling rule (s). The graph shows that collective illusions tend to persist longer when sampling strongly depends on popularity (s is high). It is useful to note that even when choices are independent ($s = 0$), it still takes some time for the system to exit the collective illusion (because only one individual can sample and change beliefs in each period). But the crucial feature of this figure is that collective illusions persist much longer when people tend to sample popular alternatives.

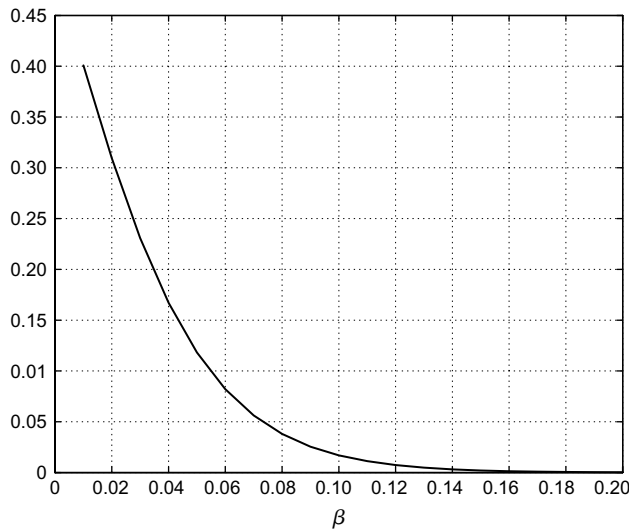
The right panel of Figure 2 shows how this pattern changes with the difference between the qualities of the two alternatives. Intuition suggests that when the alternatives have very different qualities, collective illusions are unlikely to persist. The reason is that, in this case, it is very easy for the agents to learn from their experience that the alternatives are different. We find that this is the case. The graphs of Figure 2 are plotted by assuming that alternative 1 is inferior to alternative 2. More precisely, $p_1 = 0.5 - \beta$ and $p_2 = 0.5 + \beta$, where β is a positive parameter that regulates the difference

Figure 2. Expected Number of Periods the System Stays in the “Collective Illusion” State Once It Has Reached It



Notes. Left panel: The graph was obtained with $n = 10$, $p_1 = 0.40$, $p_2 = 0.45$, and $s = 2$. Right panel: Graphs were obtained with $n = 10$, $p_1 = 0.5 - \beta$, and $p_2 = 0.5 + \beta$.

Figure 3. Proportion of Time the System Is in a Collective Illusion State ($N_1 > N_2$) as a Function of the Difference in True Qualities



Note. The graph was obtained with $n = 10$, $p_1 = 0.5 - \beta$, $p_2 = 0.5 + \beta$, and $s = 2$.

between the qualities of the alternatives. When β is small, the two alternatives have very similar qualities, and when β is larger, alternative 2 is much better than alternative 1. Figure 2 shows that when β is small, collective illusions are persistent. But when β is large, collective illusions do not persist, even if people have a strong tendency to select the popular alternative (i.e., s is high).

We have just seen that the difference between the qualities affects the *persistence* of collective illusions once they occur. But how frequent are collective illusions? Figure 3 shows the proportion of time the

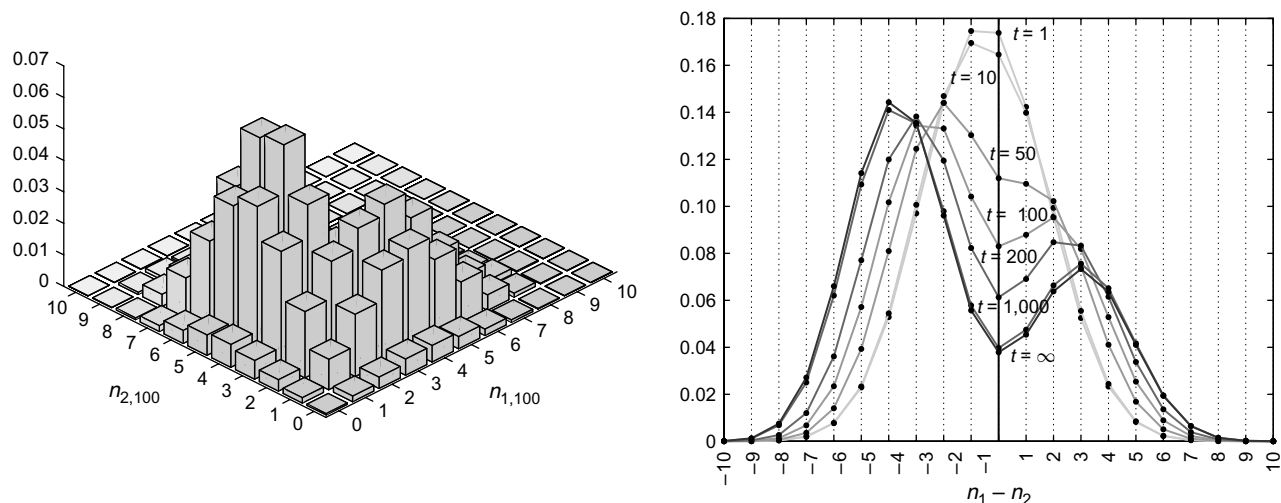
system is in a collective illusion state, as implied by the asymptotic distribution of popularities. Collective illusions are less likely to occur when the alternatives have very different qualities and more likely to occur when the qualities are similar. The intuition for this result is again that it is easier for agents to learn which alternative is the best when the qualities are very different because each observation is more diagnostic than when the alternatives have similar qualities.

Overall, this discussion implies that collective illusions are more likely to emerge and persist when the difference between the qualities of the alternatives is small. Our model shares this feature with some of the existing models that predict the emergence and persistence of collective illusions. In particular, the graph of Figure 3 is very similar to Figure 1 in Bikhchandani et al. (1992), which plots the likelihood of an incorrect cascade as a function of the accuracy of the signal received by the agents. Our model, however, relies on a different mechanism and makes predictions that are distinct from the predictions of information cascade models (see Section 5.3).

3.2.2. Dynamics of the Model: Numerical Computations.

So far, we have discussed properties of the asymptotic distribution of popularities because it has an easily interpretable formulation as a weighted bimodal distribution. Although it does not seem possible to derive a similarly interpretable formula for the distribution of popularities in any given period, it is easy to compute the distribution numerically (by computing successive powers of the transition matrix implied by the above assumptions). The left panel of Figure 4 shows the joint distribution of popularities at the beginning of period 100, assuming that all agents

Figure 4. Features of the Stationary Distribution After a Finite Number of Periods



Notes. Graphs were obtained with $n = 10$, $p_1 = 0.4$, $p_2 = 0.45$, and $s = 2$. Left panel: Joint distribution of the popularities after 100 periods. The stationary joint distribution of the popularities of alternative 1 and alternative 2 is bimodal. Right panel: The marginal distribution of $N_1 - N_2$ is also bimodal as soon as the number of periods is large enough.

in the model start with independent beliefs. The joint distribution of popularities is similar to the asymptotic distribution of Figure 1. It is bimodal. One mode is such that N_1 is close to np_1 and N_2 is low. The other mode is such that N_1 is low and N_2 is close to np_2 .

We can illustrate how the effect unfolds over time by plotting the distribution of $N_{1,t} - N_{2,t}$ for various values of t . The right panel of Figure 4 shows that the distribution of the difference is initially unimodal, and it becomes bimodal after a number of periods somewhere between 50 and 100. This means that after about 10 decisions per agent, if a collective illusion emerges, it will have a self-sustaining character. The self-sustaining character of the collective illusion becomes stronger as the distribution becomes more strongly bimodal. The bimodal character of the distribution becomes more pronounced as the number of period grows. After about 1,000 periods, the distribution has essentially reached its asymptotic state.

3.2.3. Positive Dependence Between Quality Estimates. The synchronization of estimates across individuals is most clearly illustrated by analyzing the correlation between quality estimates of pairs of individuals. The model implies that quality estimates for a given alternative become *positively* correlated. This occurs despite the fact that the individuals in the model make *independent* observations of the payoffs of the alternatives.⁵ Conversely, the quality estimates for competing alternatives become negatively correlated.

Proposition 2 (Correlation Between Quality Estimates of Two Individuals). (i) *The pairwise correlations between stationary quality estimates for a given alternative are positive. For all $i \neq j$, and for $k \in \{1, 2\}$,*

$$\text{corr}(\hat{Q}_k^i, \hat{Q}_k^j) > 0.$$

(ii) *The pairwise correlations between stationary quality estimates for competing alternatives are negative. For all $i \neq j$,*

$$\text{corr}(\hat{Q}_1^i, \hat{Q}_2^j) < 0.$$

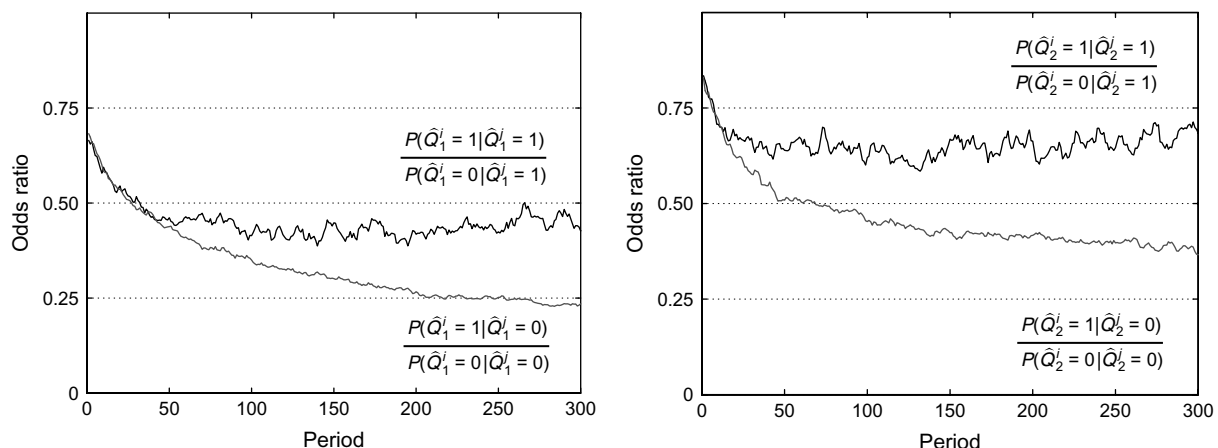
Although these results about correlation between quality estimates are elegantly formulated, our assumption that the payoff distribution is binary implies that correlations are almost surely very low. As such, the correlations do not provide a good intuitive sense about the size of the effect. The following corollary provides an alternate formulation of the same idea in terms of odds ratios.

Corollary 1 (Dependency Between Quality Estimates). *The odds that a given individual believes an alternative is “good” as opposed to “bad” is higher when another individual believes so as well. For all $i \neq j$, and for $k \in \{1, 2\}$,*

$$\frac{P(\hat{Q}_k^i = 1 | \hat{Q}_k^j = 1)}{P(\hat{Q}_k^i = 0 | \hat{Q}_k^j = 1)} > \frac{P(\hat{Q}_k^i = 1 | \hat{Q}_k^j = 0)}{P(\hat{Q}_k^i = 0 | \hat{Q}_k^j = 0)}.$$

Figure 5 displays how these odds ratios evolve over time. The difference in odds ratios predicted by the inequality of Corollary 1 emerges rather quickly, after about 50 periods, and it increases over time. After about 50 periods, the odds that an individual i believes that alternative 1 is “good” rather than “bad” when another individual j also believes alternative 1 is “good” are higher than the odds that i believes that alternative 1 is “good” rather than “bad” when j believes alternative 1 is “bad” (see the left panel). The same holds for alternative 2 (see the right panel). In summary, the quality estimates of the individuals in the group become positively dependent on each other after some time.

Figure 5. Odds Ratios That a Given Individual Believes an Alternative Is “Good” as Opposed to “Bad” When Another Individual Believes That the Same Alternative Is “Good” or “Bad”



Notes. Graphs were obtained on the basis of 10,000 simulations with $n = 10$, $p_1 = 0.4$, $p_2 = 0.45$, and $s = 2$. Left panel: Alternative 1. Right panel: Alternative 2.

These results about the emergence of a positive dependence between estimates are important because they illustrate a possible limitation of the wisdom of crowds (Galton 1907, Surowiecki 2004). The wisdom of crowds refers to the observation that the belief of the majority is often correct. The theoretical explanation of this phenomenon relies on the Condorcet (1785) jury theorem, which assumes that beliefs are formed independently (Grofman et al. 1983). It is well known that if individuals base their beliefs on the beliefs of others, i.e., if social learning operates, then the majority beliefs may not contain much more information than the belief of a single individual (e.g., Lorenz et al. 2011). This is because the beliefs of the members of the group are based on redundant information. Our theory implies that even if beliefs are based on independent pieces of evidence, beliefs will still become dependent if sampling behavior is affected by popularity.

3.3. Relation to Prior Sampling Models

Our formal model is based on but considerably extends prior sampling models. As in prior sampling models (e.g., Denrell 2005), we examine how beliefs evolve when prior beliefs influence sampling. Unlike prior models, we consider a population of individuals, who make their own sampling decisions although influenced by majority beliefs. The most closely related prior model is the model of interdependent sampling by Denrell and Le Mens (2007). Like the current model, the model of interdependent sampling examines how the attitudes of several individuals evolve over time when they influence each others' sampling processes. Unlike the current model, however, the model of interdependent sampling assumes that individuals sample simultaneously (for example, two individuals go together to the same restaurant). This assumption can be realistic for dyads, but is unrealistic for large populations. Our current model examines a population of individuals who sample alternatives at different times, but whose sampling behavior is influenced by the opinions of others. The models thus differ both conceptually and formally. Moreover, whereas the focus of Denrell and Le Mens (2007) was on when joint sampling could lead to positive or negatively correlated attitudes, this paper focuses on how a majority influence on sampling behavior can lead to the emergence of a collective illusion.

4. Discussion of the Assumptions of the Model

Our basic results also hold when there are more than two alternatives and under different assumptions about the payoff distributions and updating rule. The payoff distributions could be Normal instead of Bernoulli, and the quality estimates could be weighted averages of all past observations instead of being based

only on the most recent observation. In this section, we provide a complete description of this version of our model because it leads to formulas that allow for an easy interpretation of the effects of the various model parameters. We also discuss other influences on sampling and belief updating.

4.1. A Model with Normally Distributed Payoffs and Fractional Updating

4.1.1. Model Specification. The structure of the model is the same as before, but the assumptions regarding the payoff distributions and estimate updating rules differ.⁶

Payoffs of the Alternatives. The payoffs of alternative k ($k \in \{1, 2\}$) are independent draws from a normal distribution with mean μ_k and variance σ^2 (both alternatives have the same variance). In what follows, μ_k refers to the "quality" of alternative k .

Quality Estimates. Individual i 's quality estimate for alternative 1 (alternative 2) at the beginning of period t is random variable $\hat{Q}_{1,t}^i$ ($\hat{Q}_{2,t}^i$). The initial estimate of individual i for the quality of alternative k is a random draw from a normal distribution with mean μ_k and variance σ^2 . Let, for all i , $\hat{Z}_t^i = (\hat{Q}_{1,t}^i, \hat{Q}_{2,t}^i)$ and $\hat{\mathbf{Z}}_t = (\hat{Z}_t^1, \dots, \hat{Z}_t^n)$. Similarly, for all i , $\hat{z}_t^i = (\hat{q}_{1,t}^i, \hat{q}_{2,t}^i)$ and $\hat{\mathbf{z}}_t = (\hat{z}_t^1, \dots, \hat{z}_t^n)$.

Sampling Rule. The probability that an individual selects the first alternative is given by a logistic choice rule, as in the model of the previous section. The choice likelihood is based on the average quality estimates of the two alternatives, computed across all the people in the population:

$$pC_{1,t} = pC_1(\hat{\mathbf{z}}_t) = \frac{e^{s(1/n) \sum_{i=1}^n \hat{q}_{1,t}^i}}{e^{s(1/n) \sum_{i=1}^n \hat{q}_{1,t}^i} + e^{s(1/n) \sum_{i=1}^n \hat{q}_{2,t}^i}}, \quad (3)$$

where s is a parameter that regulates the sensitivity of the choice probability to the average quality estimate. We assume $s > 0$. This equation implies that the individual is likely to select the more popular alternative, in the sense that people in the population estimate it more positively. The probability of selecting the second alternative is defined similarly and is denoted by $pC_{2,t} = pC_2(\hat{\mathbf{z}}_t) = 1 - pC_{1,t}$.

Estimate Updating Rule. The revised estimate is a weighted average of the past estimate and the new observation (Bussemeyer and Myung 1992, Denrell 2005). When individual i selects alternative 1, her quality estimate is updated as follows: $\hat{Q}_{1,t+1}^i = \hat{Q}_{1,t}^i(1-b) + bO_{1,t}^i$, where $O_{1,t}^i$ is the observation of the payoff of alternative 1 in period t ($O_{1,t}^i \sim \mathcal{N}(\mu_1, \sigma^2)$) and b is such that $0 < b < 1$. When an alternative is not sampled, its quality estimate does not change. The estimate updating rule is the same for alternative 2.

4.1.2. Analysis. In this setting, it is possible to derive easily interpretable closed-form expressions for the correlations between the quality estimates of any two individuals.⁷ We denote by \hat{Q}_1^i (\hat{Q}_2^i) individual i 's stationary quality estimate of the first (second) alternative.

Proposition 3 (Correlation Between Quality Estimates of Two Individuals). (i) *The pairwise correlations between stationary quality estimates for a given alternative are positive. For all $i \neq j$, and for $k \in \{1, 2\}$,*

$$\text{corr}(\hat{Q}_k^i, \hat{Q}_k^j) = \frac{1}{1 + (2 + e^{-s(\mu_1 - \mu_2)} + e^{s(\mu_1 - \mu_2)})n^2(2 - b)/(s^2\sigma^2b)} > 0. \quad (4)$$

(ii) *The pairwise correlations between stationary quality estimates for competing alternatives are negative. For all $i \neq j$,*

$$\text{corr}(\hat{Q}_1^i, \hat{Q}_2^j) = -\text{corr}(\hat{Q}_k^i, \hat{Q}_k^j) < 0.$$

The sizes of the correlation increases with the weight of new observations b , the standard deviation of the payoff distribution σ , and s , the sensitivity of the choice likelihood to quality estimates. When b is close to 1, the effect is stronger because quality estimates reflect few observations (past observations are heavily discounted). With $n = 5$, $\mu_1 = -0.25$, $\mu_2 = 0$, $\sigma = 2$, $s = 5$, the correlation goes from 0 with $b = 0$ to 0.41 with $b = 1$. It is equal to 0.15 when $b = 0.4$, which is a typical value for b in several experiments on sequential choice under uncertainty (see Table 1 in Denrell 2005). If σ is high, each observation is a “noisy” signal of the quality of the alternative. The correlation goes from 0 with $\sigma = 0$ to 1 when σ is very large.

The correlation is maximal when the two alternatives have the same quality ($\mu_1 = \mu_2$) and it decreases with the difference between the qualities of the alternatives ($|\mu_1 - \mu_2|$). This is not surprising, because when the qualities of the alternatives are very different, each observation is more diagnostic of the ranking of the alternatives. Finally, the correlation decreases with n , the size of the population. The intuition is that when the population is larger, the influence of each member on the sampling behavior of the focal individual is diluted. The reason is that the focal individual's sampling likelihood is assumed to increase with the average estimate of *all* the members in the population.

4.2. The Sampling Rule and the Role of Popularity

Our assumption of a logistic sampling rule (see Equations (1) and (3)) is consistent with much prior research on sequential choice under uncertainty. It has been used in models with a reinforcement learning structure (e.g., Denrell and Le Mens 2011, Erev and Barron 2005, Sutton and Barto 1998) and it fits well data

on repeated choices between uncertain alternatives (e.g., Busemeyer and Stout 2002). A version of the logistic choice rule was also recently proposed by MacCoun (2012) as a model for the effect of the number of influence sources on conformity behavior. MacCoun (2012) shows that the model fits well data from several sources including data from Asch's (1956) experiment on line length judgments and the Milgram et al. (1969) study on the effect of the number of people looking up at a sixth floor window in a street of New York city.

In the logistic choice model, the parameter s regulates how sensitive people are to the opinions of others. A high value of s implies that people are very likely to choose the more popular alternative. As MacCoun (2012) suggests, this parameter can be interpreted as an index of norm clarity. The magnitude of our results—the extent to which quality estimates are influenced by popularity—depends critically on the value of s . The effect is large when s is high and small when s is low. Thus, according to our model, popularity cannot influence quality estimates much if choices are relatively insensitive to popularity. (Equation (4) shows that the correlation between estimates increases with s .) The same holds for many other models of the effect of popularity and conformity on belief formation, such as social impact theory (Latané 1981, Latane and Wolf 1981).

Our model assumes that all individuals in the population, including the focal actor, matter equally. Formally, sampling only depends on N_1 and N_2 . Hence, an individual is not assumed to weight his or her own opinion more than that of others. This assumption was made for simplicity and mathematical tractability (assuming symmetry seems necessary to derive an analytically tractable formula for the joint distribution of the estimates). Nevertheless, computer simulations show that our results still hold when this assumption is relaxed. For example, if the weight of the estimate of the self is equal to 2, whereas the weight of the estimates of others is equal to 0.5, the joint distribution of popularities is still bimodal after 100 periods.⁸

Our model assumes that people are more likely to choose popular alternatives, but this assumption is not always correct. If an alternative has become popular, individuals may wish to choose a less popular alternative to restore their freedom of choice (Brehm 1966) or to distinguish themselves from others (Brewer 1991). In particular, Brewer's (1991) theory of optimal distinctiveness suggests that individuals shy away from alternatives that are too popular. What are the implication of such alternative assumptions for our results? On the one hand, if the preference for distinctiveness is strong, our basic result is unlikely to hold. This suggests an important scope condition for our model (as well as for other models, including herding models, that assume

that people choose popular alternatives). On the other hand, our basic result can still hold if people care about distinctiveness in addition to popularity.

To verify this, consider the following version of our model with binary payoffs. Suppose that the attractiveness of alternative 1 increases linearly with the number of others who have a “good” estimate of that alternative until $N_1 = \delta$, and then decreases linearly until 0 when $N_1 = n$. The same is assumed for alternative 2.⁹ Simulations of this model (with the same parameters as for Figure 1) show that when δ is low (e.g., $\delta = 2$ and $n = 10$), people avoid popular alternatives, and the distribution of the difference in popularity is unimodal with its mode close to 0 (after 100 periods); that is, when people have a strong preference for distinctiveness ($\delta = 2$ and $n = 10$), our main result does not hold. But if the preference for distinctiveness is weaker (δ is equal to 4 or higher and $n = 10$), the distribution of the differences in popularity is bimodal: our main result holds in this case.

4.3. Belief Updating and Popularity

Our model assumes that popularity only influences sampling and has no direct effect on belief updating. In reality, popularity can impact both sampling and belief updating. Most importantly, people may infer that the popular alternative is better. We examine this possibility in the next section when we compare our model to models in the information cascade tradition (see Section 5.3). Several other interactions between popularity and belief updating are possible. First, if people feel pressured to select the popular alternative—and experience this as a threat to their freedom—they might discount or ignore the information sampled under such pressure (Brehm 1966). Second, those who consciously select an unpopular alternative—and thus demonstrate agency—may be motivated to assess it positively (Aronson 2007). Experiences with the unpopular alternative may for this reason be interpreted differently than experiences with the popular alternative. Finally, confirmation bias and contrast effects imply that prior expectations impact how new experiences are interpreted. When going to a popular restaurant, people may have positive expectations and pay attention to experiences confirming their expectations (Nickerson 1998). On the other hand, a negative experience with an alternative for which the decision maker had positive expectations can be particularly striking, and might carry a high weight in the belief updating process. In settings where any of these effects has a large impact on belief updating, our results may not hold. This suggests that our model is most applicable to settings where the influence of popularity is somewhat subtle, such as when people are nudged to select more popular alternatives because these are more available.

5. Relations to Other Explanations

5.1. Coordination Failures Due to Network Externalities

Our theory differs from explanations of collective mistakes that attribute them to a coordination failure due to network externalities (Arthur 1989, Elster 1978, North 1990, Pierson 2004). These explanations assume that people know that they are selecting a suboptimal alternative, but do not switch because it is difficult (or impossible) to coordinate the transition to the superior alternative. Our theory suggests that if network externalities lead people to select and thus sample popular alternatives, most people may not only choose the inferior alternative, but, when this happens, most people will also come to believe that the inferior alternative is of superior quality. Thus, in a vote about whether to switch to another alternative, most people will favor sticking to the status quo even if it is actually inferior. By contrast, explanations that rely on a coordination failure predict that people will switch if a vote can be organized and switching costs are low.

5.2. Belief Adjustments

Our theory also differs from existing mechanisms that explain why people might agree with the beliefs of the majority even when the majority is wrong. Existing explanations for this synchronization of beliefs across members of a social group generally rely on *belief adjustments* (Cialdini and Goldstein 2004, Kunda 1990, Wood 2000). One class of explanations proposes that people might adopt the behavior of the majority and then adjust their (private) beliefs and attitudes with their own behavior (Bem 1972, Festinger 1957). Another class of explanations proposes that people might adjust their (private) beliefs and attitudes to conform with the behavior of the majority because they identify with the majority (Cialdini and Goldstein 2004, Festinger et al. 1950, Hogg and Turner 1987, Turner 1991, Wood 2000). Our explanation does not challenge the experimental evidence underlying those explanations. Rather, it suggests a *complementary* explanation that operates at a *different level of analysis*: these explanations invoke properties of the mind’s information processing mechanisms, whereas our explanation focuses on properties of the sample of information on which these mechanisms operate. To analyze the role of information sampling biases, we have assumed that people are good processors of information, but that they are naïve with respect to the nature of the sample they use to form their beliefs. This strategy is central to the research program on sampling explanations of judgment biases in which our theory fits (Fiedler 2012, Fiedler and Juslin 2006, Juslin et al. 2007). Just like other contributors to this research program, we do *not* claim that people do not engage in such intrapsychic belief adjustments. Empirically analyzing how such

belief adjustments interact with information sampling mechanisms is a necessary next step in this research program.

5.3. Information Cascades

How do our results relate to herding models that rely on information cascades, which also explain collective failures to identify the best alternative (Banerjee 1992, Bikhchandani et al. 1992)? At an abstract level, our model may seem quite similar to herding models. Like herding models, our model assumes that choices between alternatives are determined by the opinions of others. The crucial difference, however, is that our model leaves out one of the key components of information cascade models: the assumption that people rely on popularity to make inferences about the quality of alternatives.

To clarify the difference between our sampling model and herding and information cascade models, suppose that a majority of individuals come to believe that A is superior to B. Herding models assume that information about such majority opinions directly impacts quality estimates: people infer that A is likely superior to B. Because information about the majority opinion influence quality estimates (people infer A is superior), it also has an indirect effect on choice behavior (people choose A because they believe it is superior). Our model does not assume that information about the majority opinion directly influences quality estimates. In our model, the majority opinion influences behavior (people likely sample A). Quality estimates are only indirectly influenced by majority opinions through their effect on sampling (because people sample A their negative opinions about A could change). Stated differently, in our model, popularity only has an indirect effect on beliefs, an effect that is mediated by the effect on popularity on sampling.

Because our model does not assume that information about the majority opinion directly influences quality estimates, it may be more suitable for settings in which such an assumption is not warranted. First, individuals may have different tastes or skills and may not believe that others know best (e.g., Banerjee et al. 2013, Hendricks et al. 2012, Munshi 2004). In situations where tastes and skills differ between people, it makes little sense for an individual to make an inference about how good an alternative is for her on the basis of popularity. For example, a vegetarian should probably not pay much attention to the popularity of restaurants among nonvegetarians. Although it does not make sense, in these situations, for the decision maker to infer how good an alternative will be for her on the basis of popularity, our model can still apply. To see how, note that even when tastes or skills differ, people might still have a tendency to select popular alternatives. Vegetarians might be motivated to go

to a restaurant liked by nonvegetarians to spend time with friends. In such settings, the payoff distributions of the alternatives systematically differ across people, but they still have a tendency to select popular alternatives. To see how we can analyze such situations, consider a variant of our model with binary payoffs, where the only modification is that we now suppose that the success probabilities of the alternatives are specific to each individual. Let p_k^i be the success probability of alternative k for individual i . The proofs of our main results (Propositions 1 and 2) can be easily adapted to this case (at the cost of some complications in the notations). In other words, our theory predicts that quality estimates will tend to synchronize even when there is taste and skill heterogeneity. This is in contrast to the predictions of information cascade accounts of conformity. Since it does not make sense to use popularity to infer quality in these situations, these theories would not predict that beliefs would become synchronized.

Second, our theory also fits settings where information about more popular alternatives is more immediately available and nudges people into selecting them without people forming conscious impressions about the quality of different alternatives. This frequently happens on online recommendation systems embedded in review websites (e.g., TripAdvisor, Yelp, Zagat) or media streaming platforms (e.g., Youtube, Spotify) where items liked by many are often displayed at the top of search results. Readers who browse the Web by clicking on hyperlinks are thus likely to sample again opinions that are already popular in their (online) social circles (Sunstein 2009). Our model suggests that such subtle influences on sampling behavior have the potential to create a synchronization of opinions, even if people do not make inferences about the worthiness of opinions on the basis of their popularities. In fact, our model does not require that people be aware of the fact that some alternatives are more popular than others. All that matters is that they have an increased propensity to sample again alternatives liked by others.

6. Application to Social Choice

The crucial assumption of our model is that people are more likely to sample popular alternatives than unpopular alternatives. An important reason for why this often happens is *social choice*, such as majority voting. When groups select alternatives by using majority voting (Davis 1973, Hastie and Kameda 2005, Stasser et al. 1980), every member will be exposed to the more popular alternative, even those who were in favor of another alternative. Our theory predicts that such a social choice rule will contribute to an homogenization of opinions.

To see how, consider members of a university department who vote about new appointments on a periodic

basis. When a new faculty member joins, all the members of the department get exposed to the research tradition the new member represents. Based on this experience, members can update their attitudes toward this research tradition. Our theory implies that the outcome of this process of learning and voting is an homogenization of attitudes: most people will come to prefer the same research tradition. Although this prediction would not be surprising in settings where all the group members made the same observations, our theory also predicts that voting will lead to an homogenization of attitudes even if the observations of the group members are independent from each others, i.e., are independent realizations of the underlying random variable. This implies, in particular, that attitudes will tend to become more homogenous even if group members have different tastes.

To demonstrate this formally, consider a variant of our model with binary payoffs. All the assumptions remain unchanged except for the sampling rule: we now assume that, in every period, *all the individuals* sample the most popular alternative. If $N_1 > N_2$ ($N_1 < N_2$), everybody selects alternative 1 (alternative 2). If $N_1 = N_2$, one alternative is selected randomly and everybody samples it. The observations of all the individuals are *independent* realizations of the underlying payoff variable. With these assumptions, the stationary distribution of the popularities is in fact very similar to that of the model with binary payoffs (Proposition 1).¹⁰ Thus, quality estimates will become synchronized: most individuals will agree that one alternative is superior to the other one, but the favored alternative does not have to be the superior one.

In summary, majority voting combined with individual learning can lead to an homogenization of opinions, possibly in favor of a less-than-optimal alternative. Unlike prior literature, our model does not need to assume that group members identify with the group (e.g., Mackie 1986). As such, our theory casts new light on the conditions where group membership contributes to attitude and opinion homogeneity.

7. Discussion and Conclusion

Our model offers an existence proof that a sampling-based account could explain phenomena usually attributed to either rational popularity-based inferences or to irrational cognitive biases. Our approach does not require that people engage in sophisticated inferences based on the choices of others or that people engage in motivated cognition. Rather, it emphasizes basic learning processes in an environment where sampling is driven by the opinions of the majority. Sampling influenced by the majority offers a simple alternative explanation for how public conformity in behavior can lead to private acceptance and a possible collective illusion.

How can the sampling-based account proposed in this paper be empirically distinguished from the inference-based mechanism? To address this question, it is useful to consider a setting where fine-grained data about quality estimates are available and popularity affects both quality inferences and sampling behavior. Suppose that at the beginning of each period, people report their quality estimates for the two alternatives A and B (\hat{q}_t^A and \hat{q}_t^B), then learn about the popularity of the alternatives (N_t^A and N_t^B), then report their quality estimates given what they know about the popularities (let us call these $\hat{q}_{t^*}^A$ and $\hat{q}_{t^*}^B$ —these quantities are not part of our formal model), then sample one of the alternatives, then report their quality estimates again (\hat{q}_{t+1}^A and \hat{q}_{t+1}^B). Consider alternative A. In this context, the effect of popularity-based inferences is the effect of popularity on $\hat{q}_{t^*}^A - \hat{q}_t^A$. The effect of popularity through sampling is the effect of popularity on $\hat{q}_{t+1}^A - \hat{q}_{t^*}^A$. The total effect of popularity on quality estimates is the effect of popularity on $\hat{q}_{t+1}^A - \hat{q}_t^A$. If there are no popularity-based quality inferences, we have $\hat{q}_{t^*}^A = \hat{q}_t^A$. In this case, the effect of popularity would be entirely driven by its effect on sampling behavior, as in our formal model. In the remainder of this discussion, we refer to this case as the “direct sampling account.”

This example shows that the “direct sampling account” has empirical implications distinct from inference-based accounts. To cleanly distinguish between these accounts, however, researchers need data about quality estimates *after* people learn about the popularity and *before* they make their sampling decisions (i.e., researchers need $\hat{q}_{t^*}^A$ and $\hat{q}_{t^*}^B$). In many naturally occurring environments, such detailed data on quality estimates are not available. Researchers may have access to data on quality estimates and sampling behavior in each of several periods, but may not know if quality estimates do reflect, or not, the information agents have about the popularities of the alternatives. With such data, researchers cannot distinguish between the direct effect of popularity on sampling behavior and its indirect effect via quality estimates; that is, suppose a researcher finds that popularity in period $t - 1$ affects sampling behavior in period t . Such an effect may occur because popularity directly influences sampling (as in the “direct sampling account”) or because popularity indirectly influence sampling via popularity-based quality inferences.

Our sampling model can be useful even when it is not possible to empirically disentangle the two ways in which popularity has an effect on sampling behavior. The sampling model shows that the effect of popularity on sampling behavior—whether mediated by quality estimates or not—must be taken into account to explain the association between popularity and quality

estimates. Researchers who ignore the effect of popularity on sampling will overestimate the effect of popularity on quality estimates through popularity-based inferences. In addition to any direct effect of popularity on estimates—due to inferences about quality from popularity—our model shows that in settings where people learn about the qualities of the alternatives partly from their own experiences, there is also an indirect effect of popularity on beliefs that operates via sampling behavior.

This paper has focused on collective illusions, but a sampling account could be applied to other phenomena in information economics where prior theoretical accounts have emphasized inferences from choices. Consider signaling (Spence 1973). Models of costly signaling assume that audiences make inferences about individual traits on the basis of choices of activities. A sampling account suggests that these choices can also impact preferences over activities: people choosing education sample information about this activity and learn that they enjoy education. More generally, economic incentives that influence choice will likely have systematic effects on evaluations and beliefs because—as the empirical literature on sampling effects shows—sampling biases have systematic effects on belief formation (Fiedler 2000, Fazio et al. 2004, Fetchenhauer and Dunning 2010).

Acknowledgments

Both authors contributed equally to this work. The authors are grateful for discussions with and comments by Judith Avrahami, Manel Baucells, Jonathan Bendor, Nick Chater, Peter Glynn, Michael Hannan, Robin Hogarth, Greta Hsu, Yaakov Kareev, James Kitts, Balázs Kovács, Thorbjørn Knudsen, Tomás Lejarraga, Ming Leung, Jim March, Johannes Müller-Trede, László Pólos, Saurabh Prakash, Karl Schlag, Hrvoje Stojic, Helena Szrek, and Ezra Zuckerman. The authors benefited from comments by participants at the Annual meeting of the Academy of Management 2008, the Annual Meeting of the American Sociological Association 2008, the Barcelona Conference on Analytical Sociology and Social Mechanisms 2010, the third Meeting of the Theoretical Organization Modeling (TOM) Society in Barcelona in 2012, and seminars at INSEAD in 2010, Tilburg University in 2012, the Chicago Booth School of Business in 2012, Universidad Carlos III de Madrid in 2012, Stanford Graduate School of Business in 2013, Universitat de les Illes Balears in 2014, the University of Lugano in 2014, the University of Southern Denmark in 2014, Heidelberg University in 2014, and the University of California, Davis in 2014.

Appendix A. Proofs for the Model with Binary Payoffs

A.1. Proof of Proposition 1

Note that N_1 and N_2 are the random variables $N_{1,t}$ and $N_{2,t}$ converge to. The state space is of size $(2 \cdot 2)^n$. The Markov process, $\{N_{1,t}, N_{2,t}, t = 1, 2, \dots\}$, is irreducible and aperiodic. Therefore, the joint distribution of (N_1, N_2) is uniquely

defined (Lemma 5.14 in Ross and Pekoz 2007). To demonstrate that a distribution is the asymptotic distribution, it is therefore enough to show that it satisfies a stability equation.

To formulate the stability equation, we begin by writing the joint distribution of estimates at the beginning of period t as a function of the joint distribution of estimates at the beginning of period $t - 1$ and the transition probabilities. Let f_t denote the joint distribution of $(\hat{Q}_{1,t}^1, \hat{Q}_{2,t}^1, \dots, \hat{Q}_{1,t}^n, \hat{Q}_{2,t}^n)$. We can write¹¹

$$f_t(\hat{q}_{1,t}^1, \hat{q}_{2,t}^1, \dots, \hat{q}_{1,t}^j, \hat{q}_{2,t}^j, \dots, \hat{q}_{1,t}^n, \hat{q}_{2,t}^n) = \sum_{k=1}^4 B_{k,t}, \quad (A1)$$

with

$$B_{1,t} = \sum_{j=1}^{t-1} \frac{1}{n} f_{t-1}(\hat{q}_{1,t-1}^1, \hat{q}_{2,t-1}^1, \dots, \hat{q}_{1,t-1}^j, \hat{q}_{2,t-1}^j, \dots, \hat{q}_{1,t-1}^n, \hat{q}_{2,t-1}^n) \cdot p C_1 p_1^{\hat{q}_{1,t-1}^j} (1 - p_1)^{1 - \hat{q}_{1,t-1}^j}, \quad (A2)$$

$$B_{2,t} = \sum_{j=1}^n \frac{1}{n} f_{t-1}(\hat{q}_{1,t-1}^1, \hat{q}_{2,t-1}^1, \dots, \hat{q}_{1,t-1}^j, \hat{q}_{2,t-1}^j, \dots, \hat{q}_{1,t-1}^n, \hat{q}_{2,t-1}^n) \cdot p C_1 p_1^{1 - \hat{q}_{1,t-1}^j} (1 - p_1)^{\hat{q}_{1,t-1}^j}, \quad (A3)$$

$$B_{3,t} = \sum_{j=1}^n \frac{1}{n} f_{t-1}(\hat{q}_{1,t-1}^1, \hat{q}_{2,t-1}^1, \dots, \hat{q}_{1,t-1}^j, \hat{q}_{2,t-1}^j, \dots, \hat{q}_{1,t-1}^n, \hat{q}_{2,t-1}^n) \cdot (1 - p C_1) p_2^{\hat{q}_{2,t-1}^j} (1 - p_2)^{1 - \hat{q}_{2,t-1}^j}, \quad (A4)$$

$$B_{4,t} = \sum_{j=1}^n \frac{1}{n} f_{t-1}(\hat{q}_{1,t-1}^1, \hat{q}_{2,t-1}^1, \dots, \hat{q}_{1,t-1}^j, \hat{q}_{2,t-1}^j, \dots, \hat{q}_{1,t-1}^n, \hat{q}_{2,t-1}^n) \cdot (1 - p C_1) p_2^{1 - \hat{q}_{2,t-1}^j} (1 - p_2)^{\hat{q}_{2,t-1}^j}. \quad (A5)$$

To explain this, consider the summand of B_1 . This is the probability that, in period $t - 1$, individual j gets selected ($1/n$), that the quality estimates equal $\hat{q}_{1,t-1}^1, \hat{q}_{2,t-1}^1, \dots, \hat{q}_{1,t-1}^j, \hat{q}_{2,t-1}^j, \dots, \hat{q}_{1,t-1}^n, \hat{q}_{2,t-1}^n$ (which occurs with probability $f_{t-1}(\hat{q}_{1,t-1}^1, \hat{q}_{2,t-1}^1, \dots, \hat{q}_{1,t-1}^j, \hat{q}_{2,t-1}^j, \dots, \hat{q}_{1,t-1}^n, \hat{q}_{2,t-1}^n)$), that individual j samples alternative 1 (occurs with probability $p C_1$), and that her estimate does not change. When $\hat{q}_{1,t-1}^j = 1$, her estimate does not change when j gets a success (this occurs with probability p_1); when $\hat{q}_{1,t-1}^j = 0$, her estimate does not change when j gets a failure (this occurs with probability $1 - p_1$). In this scenario, no quality estimate changes. We can express B_1 in terms of the quality estimates at the beginning of period t (instead of $t - 1$) as follows:

$$B_{1,t} = \sum_{j=1}^t \frac{1}{n} f_{t-1}(\hat{q}_{1,t}^1, \hat{q}_{2,t}^1, \dots, \hat{q}_{1,t}^j, \hat{q}_{2,t}^j, \dots, \hat{q}_{1,t}^n, \hat{q}_{2,t}^n) \cdot p C_1 p_1^{\hat{q}_{1,t}^j} (1 - p_1)^{1 - \hat{q}_{1,t}^j}.$$

The summand of B_2 is the probability that, in the previous period, individual j gets selected, that the quality estimates equal $\hat{q}_{1,t-1}^1, \hat{q}_{2,t-1}^1, \dots, \hat{q}_{1,t-1}^j, \hat{q}_{2,t-1}^j, \dots, \hat{q}_{1,t-1}^n, \hat{q}_{2,t-1}^n$, that she samples alternative 1 (this occurs with probability $p C_1$), and that her estimate changes. When $\hat{q}_{1,t-1}^j = 1$, agent j believes alternative 1 to be “good,” and thus a change happens when j gets a failure (this occurs with probability $1 - p_1$); when $\hat{q}_{1,t-1}^j = 0$, agent j believe alternative 1 to be “bad,” and thus a change happens when j gets a success (this occurs with probability p_1). In this scenario, only $\hat{q}_{1,t-1}^j$ changes, from $\hat{q}_{1,t-1}^j$

to $1 - \hat{q}_{1,t}^j$. We can hence write B_2 in terms of the quality estimates at the beginning of period t as follows:

$$B_{2,t} = \sum_{j=1}^n \frac{1}{n} f_{t-1}(\hat{q}_{1,t}^1, \hat{q}_{2,t}^1, \dots, 1 - \hat{q}_{1,t}^j, \hat{q}_{2,t}^j, \dots, \hat{q}_{1,t}^n, \hat{q}_{2,t}^n) \cdot pC_1 p_1^{\hat{q}_{1,t}^j} (1 - p_1)^{1 - \hat{q}_{1,t}^j}. \quad (A6)$$

Similarly, we can rewrite

$$B_{3,t} = \sum_{j=1}^n \frac{1}{n} f_{t-1}(\hat{q}_{1,t}^1, \hat{q}_{2,t}^1, \dots, \hat{q}_{1,t}^j, \hat{q}_{2,t}^j, \dots, \hat{q}_{1,t}^n, \hat{q}_{2,t}^n) \cdot (1 - pC_1) p_2^{\hat{q}_{2,t}^j} (1 - p_2)^{1 - \hat{q}_{2,t}^j}, \quad (A7)$$

$$B_{4,t} = \sum_{j=1}^n \frac{1}{n} f_{t-1}(\hat{q}_{1,t}^1, \hat{q}_{2,t}^1, \dots, \hat{q}_{1,t}^j, 1 - \hat{q}_{2,t}^j, \dots, \hat{q}_{1,t}^n, \hat{q}_{2,t}^n) \cdot (1 - pC_1) p_2^{\hat{q}_{2,t}^j} (1 - p_2)^{1 - \hat{q}_{2,t}^j}. \quad (A8)$$

Now that we have described the dynamics of the system in a convenient form, we can express the stationary joint distribution of estimates of all the individuals in the model. Then we use this distribution to derive the stationary distribution of the popularities. Let $(\hat{Q}_1^1, \hat{Q}_2^1, \dots, \hat{Q}_1^n, \hat{Q}_2^n)$ be the random variables toward which the quality estimates $(\hat{Q}_{1,t}^1, \hat{Q}_{2,t}^1, \dots, \hat{Q}_{1,t}^n, \hat{Q}_{2,t}^n)$ converge.

Lemma 1. *The stationary joint density of the estimates of all the individuals (the \hat{q}_1^j 's and \hat{q}_2^j 's can be equal to one or zero) is equal to*

$$f(\hat{q}_1^1, \hat{q}_2^1, \dots, \hat{q}_1^n, \hat{q}_2^n) = K(e^{-s n_1} + e^{-s n_2}) p_1^{n_1} (1 - p_1)^{n - n_1} p_2^{n_2} (1 - p_2)^{n - n_2}, \quad (A9)$$

where $n_1 = \sum_{j=1}^n \hat{q}_1^j$ and $n_2 = \sum_{j=1}^n \hat{q}_2^j$ and $K^{-1} = (1 - p_1 + p_1 e^{-s})^n + (1 - p_2 + p_2 e^{-s})^n$.

Proof. The Markov chain has a unique stationary distribution. To prove that the distribution defined in Equation (A9) is the stationary distribution, it is thus enough to prove that it satisfies the following stability equation that is obtained by letting t converge to infinity in Equation (A1):¹²

$$f(\hat{q}_1^1, \hat{q}_2^1, \dots, \hat{q}_1^j, \hat{q}_2^j, \dots, \hat{q}_1^n, \hat{q}_2^n) = \sum_{k=1}^4 B_k, \quad (A10)$$

with

$$B_1 = \sum_{j=1}^n \frac{1}{n} f(\hat{q}_1^1, \hat{q}_2^1, \dots, \hat{q}_1^j, \hat{q}_2^j, \dots, \hat{q}_1^n, \hat{q}_2^n) \cdot pC_1 p_1^{\hat{q}_{1,t}^j} (1 - p_1)^{1 - \hat{q}_{1,t}^j}, \quad (A11)$$

$$B_2 = \sum_{j=1}^n \frac{1}{n} f(\hat{q}_1^1, \hat{q}_2^1, \dots, 1 - \hat{q}_{1,t}^j, \hat{q}_{2,t}^j, \dots, \hat{q}_1^n, \hat{q}_2^n) \cdot pC_1 p_1^{\hat{q}_{1,t}^j} (1 - p_1)^{1 - \hat{q}_{1,t}^j}, \quad (A12)$$

$$B_3 = \sum_{j=1}^n \frac{1}{n} f(\hat{q}_1^1, \hat{q}_2^1, \dots, \hat{q}_1^j, \hat{q}_2^j, \dots, \hat{q}_1^n, \hat{q}_2^n) \cdot (1 - pC_1) p_2^{\hat{q}_{2,t}^j} (1 - p_2)^{1 - \hat{q}_{2,t}^j}, \quad (A13)$$

$$B_4 = \sum_{j=1}^n \frac{1}{n} f(\hat{q}_1^1, \hat{q}_2^1, \dots, \hat{q}_1^j, 1 - \hat{q}_{2,t}^j, \dots, \hat{q}_1^n, \hat{q}_2^n) \cdot (1 - pC_1) p_2^{\hat{q}_{2,t}^j} (1 - p_2)^{1 - \hat{q}_{2,t}^j}. \quad (A14)$$

Now, we verify that the formula for the joint distribution in Equation (A9) satisfies the above stability equation. Using the

expression for $f(\hat{q}_1^1, \hat{q}_2^1, \dots, \hat{q}_1^j, \hat{q}_2^j, \dots, \hat{q}_1^n, \hat{q}_2^n)$ given by Equation (A9), and noting that $pC_1(n_1, n_2) = 1/(1 + e^{-s(n_1 - n_2)}) = e^{-s n_2}/(e^{-s n_2} + e^{-s n_1})$, the term which is summed up in B_1 can be expressed as

$$K e^{-s n_2} p_1^{n_1 + \hat{q}_1^j} (1 - p_1)^{n - n_1 + 1 - \hat{q}_1^j} p_2^{n_2} (1 - p_2)^{n - n_2}.$$

Note that when $n_1 = \sum_{i \neq j} \hat{q}_1^i + \hat{q}_1^j$, we can write $\sum_{i \neq j} \hat{q}_1^i + 1 - \hat{q}_1^j = n_1 + 1 - 2\hat{q}_1^j$. Hence, the term that is summed up in B_2 can be expressed as

$$K e^{-s n_2} p_1^{n_1 + 1 - \hat{q}_1^j} (1 - p_1)^{n - n_1 + \hat{q}_1^j} p_2^{n_2} (1 - p_2)^{n - n_2}.$$

Then we have

$$B_1 + B_2 = K \sum_{j=1}^n \frac{1}{n} e^{-s n_2} p_1^{n_1} (1 - p_1)^{n - n_1} p_2^{n_2} (1 - p_2)^{n - n_2} \cdot [p_1^{\hat{q}_1^j} (1 - p_1)^{1 - \hat{q}_1^j} + p_1^{1 - \hat{q}_1^j} (1 - p_1)^{\hat{q}_1^j}].$$

However, the term within the bracket always equals 1, for if $\hat{q}_1^j = 1$, then

$$p_1^{\hat{q}_1^j} (1 - p_1)^{1 - \hat{q}_1^j} + p_1^{1 - \hat{q}_1^j} (1 - p_1)^{\hat{q}_1^j} = p_1 + 1 - p_1 = 1. \quad (A15)$$

And if $\hat{q}_1^j = 0$, then

$$p_1^{\hat{q}_1^j} (1 - p_1)^{1 - \hat{q}_1^j} + p_1^{1 - \hat{q}_1^j} (1 - p_1)^{\hat{q}_1^j} = 1 - p_1 + p_1 = 1. \quad (A16)$$

So,

$$B_1 + B_2 = K \sum_{j=1}^n \frac{1}{n} e^{-s n_2} p_1^{n_1} (1 - p_1)^{n - n_1} p_2^{n_2} (1 - p_2)^{n - n_2}. \quad (A17)$$

A similar reasoning gives

$$B_3 + B_4 = K \sum_{j=1}^n \frac{1}{n} e^{-s n_1} p_1^{n_1} (1 - p_1)^{n - n_1} p_2^{n_2} (1 - p_2)^{n - n_2}. \quad (A18)$$

Finally,

$$\sum_{k=1}^4 B_k = K \sum_{j=1}^n \frac{1}{n} (e^{-s n_1} + e^{-s n_2}) p_1^{n_1} (1 - p_1)^{n - n_1} p_2^{n_2} (1 - p_2)^{n - n_2} = K(e^{-s n_1} + e^{-s n_2}) p_1^{n_1} (1 - p_1)^{n - n_1} p_2^{n_2} (1 - p_2)^{n - n_2}, \quad (A19)$$

which is the expression for the joint estimate density in Equation (A9). Also, it is easy to verify that with the proposed formula for K , the expression in Equation (A9) sums to 1. This concludes the proof.

Given the joint distributions of estimates, it is easy to get the joint distribution of the sums, $\bar{f}(n_1, n_2)$. First, what is the probability of one way of assigning beliefs so that the sums are n_1 and n_2 ? From Equation (A9) we know that it is given by

$$f(\hat{q}_1^1, \hat{q}_2^1, \dots, \hat{q}_1^n, \hat{q}_2^n) = K(e^{-s n_1} + e^{-s n_2}) p_1^{n_1} (1 - p_1)^{n - n_1} p_2^{n_2} (1 - p_2)^{n - n_2}. \quad (A20)$$

Second, the number of ways there are to have $\sum_{j=1}^n \hat{q}_1^j = n_1$ and simultaneously $\sum_{j=1}^n \hat{q}_2^j = n_2$ is $\binom{n}{n_1} \binom{n}{n_2}$. By multiplying, we get $\bar{f}(n_1, n_2)$.

A.2. Proof of Proposition 2

(i) Suppose, without loss of generality, that $k = 1$, $i = 1$, and $j = 2$. To show that $\text{corr}(\hat{Q}_1^1, \hat{Q}_2^1) > 0$, it is enough to show that the joint distribution of estimates, f , is totally positive of order 2 with respect to its first and third arguments (TP_2 , see Karlin and Rinott 1980, p. 470–471, Eqs. 1.7–1.9). The density f is TP_2 with respect to its first and third arguments if and only if, for all $\hat{q}_1^1, \hat{q}_1^{1*}, \hat{q}_1^2, \hat{q}_1^{2*}, \hat{q}_2^1, \hat{q}_2^{1*}, \hat{q}_2^2, \hat{q}_2^{2*}, \dots, \hat{q}_1^n, \hat{q}_2^n$ such that $\hat{q}_1^1 < \hat{q}_1^{1*}$ and $\hat{q}_2^1 < \hat{q}_2^{1*}$, we have

$$\begin{aligned} & f(\hat{q}_1^1, \hat{q}_1^{1*}, \hat{q}_2^1, \hat{q}_2^{1*}, \hat{q}_2^2, \hat{q}_2^{2*}, \dots, \hat{q}_1^n, \hat{q}_2^n) \\ & \cdot f(\hat{q}_1^1, \hat{q}_1^2, \hat{q}_1^{1*}, \hat{q}_1^{2*}, \hat{q}_2^1, \hat{q}_2^{1*}, \hat{q}_2^2, \hat{q}_2^{2*}, \dots, \hat{q}_1^n, \hat{q}_2^n) \\ & > f(\hat{q}_1^1, \hat{q}_1^2, \hat{q}_1^{1*}, \hat{q}_1^{2*}, \hat{q}_2^1, \hat{q}_2^{1*}, \hat{q}_2^2, \hat{q}_2^{2*}, \dots, \hat{q}_1^n, \hat{q}_2^n) \\ & \cdot f(\hat{q}_1^1, \hat{q}_1^2, \hat{q}_1^{1*}, \hat{q}_1^{2*}, \hat{q}_2^1, \hat{q}_2^{1*}, \hat{q}_2^2, \hat{q}_2^{2*}, \dots, \hat{q}_1^n, \hat{q}_2^n). \end{aligned} \quad (A21)$$

Let $\hat{q}_1^1, \hat{q}_1^{1*}, \hat{q}_2^1, \hat{q}_2^{1*}$ be such that $\hat{q}_1^1 < \hat{q}_1^{1*}$ and $\hat{q}_2^1 < \hat{q}_2^{1*}$. Given the restrictions on the quality estimates in this model, we have $\hat{q}_1^1 = 0$, $\hat{q}_1^{1*} = 1$, $\hat{q}_2^1 = 0$, and $\hat{q}_2^{1*} = 1$. To simplify the formulas, let denote $\sum_{j=3}^n \hat{q}_j^i$ by $n_1^{3 \rightarrow n}$. Applied to the asymptotic joint density, Equation (A21) becomes

$$\begin{aligned} & K(e^{-s(1+n_1^{3 \rightarrow n})} + e^{-sn_2})p_1^{1+n_1^{3 \rightarrow n}}(1-p_1)^{n-(1+n_1^{3 \rightarrow n})} \\ & \cdot p_2^{n_2}(1-p_2)^{n-n_2}K(e^{-s(0+n_1^{3 \rightarrow n})} + e^{-sn_2})p_1^{0+n_1^{3 \rightarrow n}} \\ & \cdot (1-p_1)^{n-(0+n_1^{3 \rightarrow n})}p_2^{n_2}(1-p_2)^{n-n_2} \\ & > K(e^{-s(1+0+n_1^{3 \rightarrow n})} + e^{-sn_2})p_1^{1+0+n_1^{3 \rightarrow n}}(1-p_1)^{n-(1+0+n_1^{3 \rightarrow n})} \\ & \cdot p_2^{n_2}(1-p_2)^{n-n_2}K(e^{-s(0+1+n_1^{3 \rightarrow n})} + e^{-sn_2})p_1^{0+1+n_1^{3 \rightarrow n}} \\ & \cdot (1-p_1)^{n-(0+1+n_1^{3 \rightarrow n})}p_2^{n_2}(1-p_2)^{n-n_2}. \end{aligned}$$

Simple algebraic manipulations show that this inequality is equivalent to

$$(e^{-s(2+n_1^{3 \rightarrow n})} + e^{-sn_2})(e^{-sn_1^{3 \rightarrow n}} + e^{-sn_2}) > (e^{-s(1+n_1^{3 \rightarrow n})} + e^{-sn_2})(e^{-s(1+n_1^{3 \rightarrow n})} + e^{-sn_2}). \quad (A22)$$

Further manipulations show that this inequality holds true iff

$$e^{-2s} + 1 - 2e^{-s} > 0. \quad (A23)$$

The left-hand side (LHS) of this equation can be rewritten as $(1 - e^{-s})^2$. This quantity is positive because $s > 0$. This implies that the inequality in Equation (A22) is true. In conclusion, f is TP_2 with respect to its first and third arguments. This implies that $\text{corr}(\hat{Q}_1^1, \hat{Q}_2^1) > 0$.

(ii) This proposition follows by a reasoning similar to that used for the proof of (i).

A.3. Proof of Corollary 1

Without loss of generality, consider $k = 1$, $i = 1$, and $j = 2$. Because the density, f , has a multiplicative form, Equation (A21) implies

$$\frac{P(\hat{Q}_1^1 = 1, \hat{Q}_2^1 = 1)}{P(\hat{Q}_1^1 = 0, \hat{Q}_2^1 = 1)} > \frac{P(\hat{Q}_1^1 = 1, \hat{Q}_2^2 = 0)}{P(\hat{Q}_1^1 = 0, \hat{Q}_2^2 = 0)}. \quad (A24)$$

By dividing the numerator and denominator of the LHS by $P(\hat{Q}_1^2 = 1)$, we obtain the LHS of the inequality of the statement of the corollary. By dividing the numerator and denominator of the right-hand side (RHS), by $P(\hat{Q}_1^2 = 0)$, we obtain the RHS of the statement of the corollary.

Appendix B. Proofs for the Model with Normally Distributed Payoffs

We will use the following notation: $\hat{\mathbf{Q}}_1 = (\hat{Q}_1^1, \dots, \hat{Q}_1^n)$, $\hat{\mathbf{Q}}_2 = (\hat{Q}_2^1, \dots, \hat{Q}_2^n)$, $\hat{\mathbf{Z}}_i = (\hat{Q}_i^1, \hat{Q}_i^2)$ and $\hat{\mathbf{Z}} = (\hat{\mathbf{Z}}_1, \dots, \hat{\mathbf{Z}}_n)$ ($\hat{\mathbf{Z}}$ is the random variable $\hat{\mathbf{Z}}_t$ converges to). We first compute the joint distribution of estimates. It is formalized in the following lemma:

Lemma 2. The joint stationary distribution of the quality estimates of all individuals in the population for the two alternatives is

$$h(\hat{\mathbf{z}}) = K(e^{-s(1/n)\sum_{i=1}^n \hat{q}_1^i} + e^{-s\frac{1}{n}\sum_{i=1}^n \hat{q}_2^i}) \prod_{i=1}^n g_1(\hat{q}_1^i)g_2(\hat{q}_2^i), \quad (B1)$$

where

$$K^{-1} = e^{(s^2\sigma^2b)/(2n(2-b))}(e^{-s\mu_1} + e^{-s\mu_2}) \quad (B2)$$

is a normalizing constant, and $g_j(\cdot)$ is the density of the estimate that an individual would converge to if she always sampled alternative j .

In this model, the “popularity” of an alternative is the sum of the quality estimates of that alternative in the population. Let $\bar{Q}_1 = \sum_i \hat{Q}_1^i$ and $\bar{Q}_2 = \sum_i \hat{Q}_2^i$, and let $\bar{h}(\cdot, \cdot)$ denote the joint distribution of \bar{Q}_1 and \bar{Q}_2 . The joint distribution of the popularities is given by the following lemma:

Lemma 3 (Joint Stationary Distribution of Sums of Estimates). The stationary distribution of the sums of quality estimates, \bar{Q}_1 and \bar{Q}_2 is

$$\bar{h}(\bar{q}_1, \bar{q}_2) = K(e^{-s(\bar{q}_1/n)} + e^{-s(\bar{q}_2/n)})\bar{h}_1(\bar{q}_1)\bar{h}_2(\bar{q}_2), \quad (B3)$$

where, for $j \in \{1, 2\}$, $\bar{h}_j(\cdot)$ is the density of a normal distribution with mean $n\mu_j$ and variance $n\sigma^2b/(2-b)$, and K is a normalizing constant.

The joint distribution is a weighted bivariate Normal distribution where the weight is $K(e^{-s(\bar{q}_1/n)} + e^{-s(\bar{q}_2/n)})$. As in the above model, the weighting function puts additional weight on the extremes where the alternatives are unpopular. And thus the joint distribution of popularities is bimodal if s is high enough. This implies that a collective illusion might emerge in which the suboptimal alternative is evaluated more positively than the superior alternative.

B.1. Proof of Lemma 2

We first need to show that the stationary distribution of estimates exists. Because the model defines a Markov process in general state space (noncountable), this is a rather involved proof. Fortunately, we can use some recent developments on the analysis of the stability of Harris chains to get to our objective.

B.1.1. Convergence of the Markov Process. With the above specification, the (random) vector of estimates, $\hat{\mathbf{Z}}_t$, defines a Harris chain that is positive recurrent. This implies that $\hat{\mathbf{Z}}_t$ has a unique stationary distribution h (e.g., Meyn and Tweedie 1993). To prove this and the convergence, we use the following theorem (Glynn 2011).

Theorem 1. Consider a Markov chain $R = (R_t : t \geq 1)$ with state space S . We denote by $E_r(\cdot)$ conditional expectations given that $R_1 = r$, and by $P_r(\cdot)$ conditional probabilities given that $R_1 = r$. Suppose there exists $A \subseteq S$, $\gamma > 0$, $m \geq 1$, and a distribution φ on S , $\epsilon > 0$ and a function $g : S \rightarrow [0, \infty]$ such that

$$(i) \quad E_r g(R_2) \leq g(r) - \epsilon, \quad r \in A^c,$$

- (ii) $\sup_{r \in A} E_r g(R_2) < \infty$,
 (iii) $P_r(R_m \in \cdot) \geq \gamma \varphi(\cdot)$, $r \in A$.

Then, R possesses a unique stationary distribution h , and, for each nonnegative f ,

$$\frac{1}{n} \sum_{i=1}^n f(R_i) \rightarrow \int_S f(r) h(dr) \text{ a.s.}$$

We now verify that the Markov process $(\hat{\mathbf{Z}}_t : t \geq 1)$ satisfies the conditions of the theorem. In our setting, the state space is \mathbb{R}^{2n} .

(i) Let

$$g(\hat{\mathbf{z}}_t) = \sum_{i=1}^n (|\hat{q}_{1,t}^i - \mu_1| + |\hat{q}_{2,t}^i - \mu_2|)$$

and $A = ([\mu_1 - 1, \mu_1 + 1] \times [\mu_2 - 1, \mu_2 + 1])^n$. Let $\hat{\mathbf{z}}_1 \in A^c$. We have

$$\begin{aligned} E_{\hat{\mathbf{z}}_1} g(\hat{\mathbf{Z}}_2) &= \sum_{i=1}^n \frac{1}{n} p C_1(\hat{\mathbf{z}}_1) \left[(1-b) |\hat{q}_{1,1}^i - \mu_1| + b E[o_{1,1}^i - \mu_1] \right. \\ &\quad \left. + \sum_{j \neq i} |\hat{q}_{1,1}^j - \mu_1| + \sum_j |\hat{q}_{2,1}^j - \mu_2| \right] \\ &\quad + \sum_{i=1}^n \frac{1}{n} p C_2(\hat{\mathbf{z}}_1) \left[\sum_j |\hat{q}_{1,1}^j - \mu_1| + (1-b) |\hat{q}_{2,1}^i - \mu_2| \right. \\ &\quad \left. + b E[o_{2,1}^i - \mu_2] + \sum_{j \neq i} |\hat{q}_{2,1}^j - \mu_2| \right]. \end{aligned} \quad (B4)$$

Noting that for all i , $E[o_{1,1}^i - \mu_1] = 0$ and $E[o_{2,1}^i - \mu_2] = 0$, we have

$$\begin{aligned} E_{\hat{\mathbf{z}}_1} g(\hat{\mathbf{Z}}_2) &= \sum_{i=1}^n \frac{1}{n} \left[\sum_j |\hat{q}_{1,1}^j - \mu_1| + \sum_j |\hat{q}_{2,1}^j - \mu_2| \right. \\ &\quad \left. + p C_1(\hat{\mathbf{z}}_1) (-b) |\hat{q}_{1,1}^i - \mu_1| \right. \\ &\quad \left. + p C_2(\hat{\mathbf{z}}_1) (-b) |\hat{q}_{2,1}^i - \mu_2| \right], \end{aligned} \quad (B5)$$

$$\begin{aligned} E_{\hat{\mathbf{z}}_1} g(\hat{\mathbf{Z}}_2) &= \sum_{i=1}^n [|\hat{q}_{1,1}^i - \mu_1| + |\hat{q}_{2,1}^i - \mu_2|] \\ &\quad - \frac{b}{n} \sum_{i=1}^n [p C_1(\hat{\mathbf{z}}_1) |\hat{q}_{1,1}^i - \mu_1| \\ &\quad + p C_2(\hat{\mathbf{z}}_1) |\hat{q}_{2,1}^i - \mu_2|], \end{aligned} \quad (B6)$$

$$\begin{aligned} E_{\hat{\mathbf{z}}_1} g(\hat{\mathbf{Z}}_2) &= g(\hat{\mathbf{z}}_1) - \frac{b}{n} \sum_{i=1}^n [p C_1(\hat{\mathbf{z}}_1) |\hat{q}_{1,1}^i - \mu_1| \\ &\quad + p C_2(\hat{\mathbf{z}}_1) |\hat{q}_{2,1}^i - \mu_2|]. \end{aligned} \quad (B7)$$

The fact that $\hat{\mathbf{z}}_1 \in A^c$ and the definition of A imply that, for all i , $|\hat{q}_{1,1}^i - \mu_1| > 1$ and $|\hat{q}_{2,1}^i - \mu_2| > 1$. Therefore,

$$\begin{aligned} \sum_{i=1}^n [p C_1(\hat{\mathbf{z}}_1) |\hat{q}_{1,1}^i - \mu_1| + p C_2(\hat{\mathbf{z}}_1) |\hat{q}_{2,1}^i - \mu_2|] \\ > \sum_{i=1}^n [p C_1(\hat{\mathbf{z}}_1) + p C_2(\hat{\mathbf{z}}_1)] = n. \end{aligned} \quad (B8)$$

Finally, $E_{\hat{\mathbf{z}}_1} g(\hat{\mathbf{Z}}_2) < g(\hat{\mathbf{z}}_1) - b$. Setting $\epsilon = b > 0$ ensures that condition i is satisfied.

(ii) Let $\hat{\mathbf{z}}_1 \in A$. We have

$$\begin{aligned} E_{\hat{\mathbf{z}}_1} g(\hat{\mathbf{Z}}_2) &= g(\hat{\mathbf{z}}_1) - \frac{b}{n} \sum_{i=1}^n [p C_1(\hat{\mathbf{z}}_1) |\hat{q}_{1,1}^i - \mu_1| + p C_2(\hat{\mathbf{z}}_1) |\hat{q}_{2,1}^i - \mu_2|] \\ &\leq g(\hat{\mathbf{z}}_1). \end{aligned} \quad (B9)$$

Then, $\sup_{\hat{\mathbf{z}}_1 \in A} E_{\hat{\mathbf{z}}_1} g(\hat{\mathbf{Z}}_2) \leq \sup_{\hat{\mathbf{z}}_1 \in A} g(\hat{\mathbf{z}}_1) < \infty$, the last inequality being a consequence of the fact that A is a bounded space.

(iii) Let $B \subseteq \mathbb{R}^{2n}$. In $2n$ transitions, the process can reach any point in the state space \mathbb{R}^{2n} from any starting point. This is the result of a combination of two facts: First, there is a positive probability that all individuals will have been selected and will have each chosen once the first alternative and the second alternative, and thus there is a positive probability that all estimates will have been updated. Second, because updated estimates are weighted averages of previous estimates and normally distributed observations, the support of the joint distribution of updated estimates is the whole state space \mathbb{R}^{2n} . This implies that if we write

$$P_{\hat{\mathbf{z}}_1}(\hat{\mathbf{Z}}_{2n+1} \in B) = \int_B p_{2n}(\hat{\mathbf{z}}_1, \mathbf{r}) d\mathbf{r}, \quad (B10)$$

then the transition density $p_{2n}(\cdot, \mathbf{r})$ is positive. It is also continuous, as a weighted sum of normal densities.

We can then define $\phi(\mathbf{r}) = \inf_{\hat{\mathbf{z}}_1 \in A} P_{\hat{\mathbf{z}}_1}(\hat{\mathbf{Z}}_{2n+1} \in B)$. The facts that $p_{2n}: \mathbb{R}^{2n} \times \mathbb{R}^{2n} \rightarrow \mathbb{R}$ is continuous and that A is a compact set imply that ϕ is continuous. Let $\varphi(B) = \int_B \phi(\mathbf{r}) d\mathbf{r}$, and set $\gamma = 1$.

With these definitions, for all $\hat{\mathbf{z}}_1 \in A$, $P_{\hat{\mathbf{z}}_1}(\hat{\mathbf{Z}}_{2n+1} \in \cdot) \geq \gamma \varphi(\cdot)$, which proves that condition (iii) is satisfied. Q.E.D.

In conclusion, the above theorem implies that there exists a unique stationary joint distribution for the estimates.

B.1.2. The Stationary Distribution of the Quality Estimates.

We begin by introducing some notations. Let $k \in \{1, 2\}$. Let $f_k(\cdot)$ denote the density of the payoff distribution of alternative k . By construction, it is a normal density with mean μ_k and standard deviation σ . We denote by $p_k(r, x)$ the probability density of the event that the estimate of individual i transitions from r in period t to x in period $t + 1$ given that individual i samples alternative k in period t . Under the above assumptions, $p_k(r, x)$ equals $(1/b) f_j((1/b)(x - (1-b)r))$. Let $g_k(\cdot)$ be a normal density with mean μ_k and variance $\sigma^2 b/(2-b)$. It is easy to verify that g_k is stable by application of the transition kernel $p_k(\cdot, \cdot)$:

$$\int_r g_k(r) p_k(r, x) dr = g_k(x). \quad (B11)$$

We can now formulate the stability equation for the joint probability of estimates:

$$\begin{aligned} h(\hat{\mathbf{z}}) &= \sum_{i=1}^n \frac{1}{n} \int_{r_i} h(\hat{q}_1^1, \hat{q}_2^1, \dots, r_i, \hat{q}_2^i, \dots, \hat{q}_1^n, \hat{q}_2^n) \\ &\quad \cdot p C_1(\hat{q}_1^1, \hat{q}_2^1, \dots, r_i, \hat{q}_2^i, \dots, \hat{q}_1^n, \hat{q}_2^n) p_1(r_i, \hat{q}_1^i) dr_i \\ &\quad + \sum_{i=1}^n \frac{1}{n} \int_{r_i} h(\hat{q}_1^1, \hat{q}_2^1, \dots, \hat{q}_1^i, r_i, \dots, \hat{q}_1^n, \hat{q}_2^n) \\ &\quad \cdot p C_2(\hat{q}_1^1, \hat{q}_2^1, \dots, \hat{q}_1^i, r_i, \dots, \hat{q}_1^n, \hat{q}_2^n) p_2(r_i, \hat{q}_2^i) dr_i. \end{aligned} \quad (B12)$$

To explain this, consider the first term in the first sum:

$$\begin{aligned} \frac{1}{n} \int_{r_i} h(\hat{q}_1^1, \hat{q}_2^1, \dots, r_i, \hat{q}_2^i, \dots, \hat{q}_1^n, \hat{q}_2^n) \\ \cdot p C_1(\hat{q}_1^1, \hat{q}_2^1, \dots, r_i, \hat{q}_2^i, \dots, \hat{q}_1^n, \hat{q}_2^n) p_1(r_i, \hat{q}_1^i) dr_i. \end{aligned}$$

This is the probability that individual i gets selected, that all other individuals have estimates $\hat{q}_1^1, \hat{q}_2^1, \dots, \hat{q}_1^n, \hat{q}_2^n$, that individual i has estimates equal to r_i and \hat{q}_2^i , that she samples

the first alternative, and that her estimate goes from r_i to \hat{q}_1^i , which happens with probability $p_1(r_i, \hat{q}_1^i)$. We integrate over all possible estimates about the first alternative. The other terms have a similar explanation. To show that the above joint density (Equation (B1)) does satisfy the stability equation (Equation (B12)), note first that $pC_1(\mathbf{z})$ can be written as follows:

$$pC_1(\hat{\mathbf{z}}) = \frac{e^{-s(1/n)\sum_{j=1}^n \hat{q}_2^j}}{e^{-s(1/n)\sum_{j=1}^n \hat{q}_1^j} + e^{-s(1/n)\sum_{j=1}^n \hat{q}_2^j}}. \quad (\text{B13})$$

We insert this expression and the stationary joint density into the first term of the first sum:

$$\begin{aligned} & \frac{1}{n} \int_{r_i} h(\hat{q}_1^1, \hat{q}_2^1, \dots, r_i, \hat{q}_2^i, \dots, \hat{q}_1^n, \hat{q}_2^n) \\ & \cdot pC_1(\hat{q}_1^1, \hat{q}_2^1, \dots, r_i, \hat{q}_2^i, \dots, \hat{q}_1^n, \hat{q}_2^n) p_1(r_i, \hat{q}_1^i) dr_i \\ & = K \frac{1}{n} e^{-s(1/n)\sum_{j=1}^n \hat{q}_2^j} \prod_{j \neq i} g_1(\hat{q}_1^j) \prod_j g_2(\hat{q}_2^j) \int_{r_i} g_1(r_i) p_1(r_i, \hat{q}_1^i) dr_i \\ & = K \frac{1}{n} e^{-s(1/n)\sum_{j=1}^n \hat{q}_2^j} \prod_{j \neq i} g_1(\hat{q}_1^j) \prod_j g_2(\hat{q}_2^j). \end{aligned} \quad (\text{B14})$$

The last equality is a consequence of the fact that $\int_{r_i} g_1(r_i) p_1(r_i, \hat{q}_1^i) dr_i = g_1(\hat{q}_1^i)$.

Similarly, the first term in the second sum is equal to

$$K \frac{1}{n} e^{-s(1/n)\sum_{j=1}^n \hat{q}_1^j} \prod_{j \neq i} g_1(\hat{q}_1^j) \prod_j g_2(\hat{q}_2^j). \quad (\text{B15})$$

Summing across all terms, the right-hand side of Equation (B12) becomes equal to

$$K(e^{-s(1/n)\sum_{j=1}^n \hat{q}_1^j} + e^{-s(1/n)\sum_{j=1}^n \hat{q}_2^j}) \prod_j g_1(\hat{q}_1^j) g_2(\hat{q}_2^j), \quad (\text{B16})$$

which is $h(\hat{q}_1^1, \hat{q}_2^1, \dots, \hat{q}_1^n, \hat{q}_2^n)$, the stationary joint density.

Noting that the integral of the density should be equal to 1, we obtain the formula for K . We use it to write an explicit formula for h :

$$h(\hat{\mathbf{z}}) = \left(\frac{1}{\sqrt{2\pi(b/(2-b))\sigma^2}} \right)^{2n} \frac{e^{-s(1/n)\sum_{i=1}^n \hat{q}_1^i} + e^{-s(1/n)\sum_{i=1}^n \hat{q}_2^i}}{e^{s^2\sigma^2b/(2n(2-b))}(e^{-s\mu_1} + e^{-s\mu_2})} \cdot e^{-\sum_i ((\hat{q}_1^i - \mu_1)^2 + (\hat{q}_2^i - \mu_2)^2)/(2(b/(2-b))\sigma^2)}. \quad (\text{B17})$$

B.2. Proof of Lemma 3: Joint Distribution of the Sum of Estimates

To get the joint distribution of the sum, we need to perform a complicated integration of the joint density (Equation (B1)) over the region in which the sums equal some constants such as $\sum_{i=1}^n \hat{Q}_1^i = \bar{q}_1$ and $\sum_{i=1}^n \hat{Q}_2^i = \bar{q}_2$. However, note that if the sums equal \bar{q}_1 and \bar{q}_2 , the first part in the density always equals $(e^{-s(\bar{q}_1/n)} + e^{-s(\bar{q}_2/n)})$. Thus, we can ignore the first part.

The second part is then simple—what remains is a sum of independent normal distributions! The distribution of the sum of the estimates is thus a normal distribution with mean equal to the sum of the means and variance equal to the sum of the variances.

B.3. Proof of Proposition 3

(i) The formula follows by calculating $E(\hat{Q}_k^i)$, $E((\hat{Q}_k^i)^2)$, and $E(\hat{Q}_k^i \hat{Q}_k^j)$ by doing the appropriate integrations on the basis of Equation (B1). The calculations are tedious but straightforward; we therefore provide only the key steps. We have

$$\begin{aligned} E(\hat{Q}_k^i) &= \mu_k - \frac{s\sigma^2b}{n(2-b)} \frac{e^{-s\mu_k}}{e^{-s\mu_1} + e^{-s\mu_2}}, \\ E(\hat{Q}_k^i \hat{Q}_k^j) &= (\mu_k)^2 - \frac{s\sigma^2b}{n(2-b)} \frac{e^{-s\mu_k}}{e^{-s\mu_1} + e^{-s\mu_2}} \left(2\mu_k - \frac{s\sigma^2b}{n(2-b)} \right), \\ E((\hat{Q}_k^i)^2) &= (\mu_k)^2 + \frac{\sigma^2b}{2-b} + \frac{e^{-s\mu_k}}{e^{-s\mu_1} + e^{-s\mu_2}} \\ & \cdot \left(\left(\frac{s\sigma^2b}{n(2-b)} \right)^2 - 2 \frac{s\sigma^2b}{n(2-b)} \mu_k \right). \end{aligned}$$

The proof follows by noting that $\text{cov}(\hat{Q}_k^i, \hat{Q}_k^j) = E(\hat{Q}_k^i \hat{Q}_k^j) - E(\hat{Q}_k^i)E(\hat{Q}_k^j)$, $\text{var}(\hat{Q}_k^i) = E((\hat{Q}_k^i)^2) - (E(\hat{Q}_k^i))^2$, and $\text{corr}(\hat{Q}_k^i, \hat{Q}_k^j) = \text{cov}(\hat{Q}_k^i, \hat{Q}_k^j)/\sqrt{\text{var}(\hat{Q}_k^i)\text{var}(\hat{Q}_k^j)}$ and performing the appropriate substitutions and simplifications.

(ii) We have

$$E(\hat{Q}_1^i \hat{Q}_2^j) = \mu_1 \mu_2 - \frac{s\sigma^2b}{n(2-b)} \frac{\mu_2 e^{-s\mu_1} + \mu_1 e^{-s\mu_2}}{e^{-s\mu_1} + e^{-s\mu_2}}.$$

The proof follows from computations similar to those in the proof of (i).

Endnotes

¹ Unless noted otherwise, capital letters refer to random variables. Corresponding lowercase letters denote specific instances of the random variables. We use lowercase letters to refer to nonrandom model parameters.

² See Section 4.2 for further discussion of the logistic choice rule.

³ With the parameters used for constructing the graphs of Figure 1, the first mode is located at (4, 1), and the second mode is located at (0, 4).

⁴ Formally, suppose that we are at the beginning of period t_0 . Suppose that t_0 is large such that we can assume that N_{1,t_0} and N_{2,t_0} follow the asymptotic joint distribution of popularities given by Equation (2). Let T denote the first time, after t_0 , at which the system enters a state such that $N_{1,t} > N_{2,t}$. Using numerical computations on the basis of the transition matrix of the Markov chain defined by our model, we computed the expected “hitting” time of the region of the state space where $N_1 > N_2$. This is $E[T - t_0 \mid N_{1,t_0} < N_{2,t_0}]$. In our numerical computations, we use $t_0 = 10^6$.

⁵ These results are consistent with prior work on the emergence of correlated estimates and attitudes in a dyad as a result of interdependent sampling (Denrell and Le Mens 2007). In fact, Proposition 2 extends Theorem 1 in Denrell and Le Mens (2007) and shows that similar results apply to settings where there are more than two individuals and one alternative and where people do not sample simultaneously.

⁶ For ease of exposition, our model includes just two alternatives. But our results are easily adapted to the case where there are $M > 2$ alternatives.

⁷ In this case, the joint distribution of the popularities has a form very similar to the joint distribution of popularities in the model analyzed above. See Lemma 3 in Appendix B for details.

⁸ This applies for the model with binary payoffs, with $n = 10$, $p_1 = 0.4$, $p_2 = 0.45$ and $s = 2$.

⁹ Formally, we assume $pC_1(n_{1,t}, n_{2,t}) = \alpha_\delta(n_{1,t})/(\alpha_\delta(n_{1,t}) + \alpha_\delta(n_{2,t}))$ with $\alpha_\delta(m) = e^{s(mI_{m \leq \delta} + \delta((n-m)/(N-\delta))I_{m > \delta})}$ and that $I_{m \leq \delta}$ is an indicator variable equal to 1 when $m \leq \delta$; otherwise equal to 0.

¹⁰The omitted proof is almost identical to the proof of Proposition 1. The joint stationary distribution of the sum is similar to Equation (A9), by taking $s = \infty$:

$$\tilde{f}(n_1, n_2) = K(I_{n_1=0} + I_{n_2=0})(p_1)^{n_1}(1-p_1)^{n-n_1} \cdot (p_2)^{n_2}(1-p_2)^{n-n_2} \binom{n}{n_1} \binom{n}{n_2},$$

where $I_{n_j=0}$ is an indicator function equal to 1 when $n_j = 0$ and equal to 0 otherwise.

¹¹This is assuming that $0 < n_{1,t}, n_{2,t} < n$. What follows can easily be adapted to the cases where $n_{1,t}$ or $n_{2,t}$ is equal to 0 or n .

¹²This is assuming that $0 < n_1, n_2 < n$. What follows can easily be adapted to the cases where n_1 or n_2 is equal to 0 or n .

References

- Aronson E (2007) *The Social Animal* (Worth Publishers, New York).
- Arthur WB (1989) Competing technologies, increasing returns, and lock-in by historical events. *Econom. J.* 99(394):116–131.
- Asch SE (1956) Studies of independence and conformity: A minority of one against a unanimous majority. *Psych. Monographs* 70(9): 1–70.
- Avrahami J, Kareev Y (2010) The role of impulses in shaping decisions. *J. Behavioral Decision Making* 24(5):515–529.
- Banerjee AV (1992) A simple model of herd behavior. *Quart. J. Econom.* 107(3):797–817.
- Banerjee A, Chandrasekhar AG, Duflo E, Jackson MO (2013) The diffusion of microfinance. *Science* 341(6144): DOI: 10.1126/science.1236498.
- Bem DJ (1972) Self-perception theory. Berkowitz L, ed. *Advances in Experimental Social Psychology*, Vol. 6 (Academic Press, New York), 1–62.
- Berger J, Heath C (2007) Where consumers diverge from others: Identity signaling and product domains. *J. Consumer Res.* 34(2): 121–134.
- Bikhchandani S, Hirshleifer D, Welch I (1992) A theory of fads, fashion, custom, and cultural change as informational cascades. *J. Political Econom.* 100(5):992–1026.
- Bowles S (2004) *Microeconomics: Behavior, Institutions, and Evolution* (Princeton University Press, Princeton, NJ).
- Brehm JW (1966) *A Theory of Psychological Reactance* (Academic Press, New York).
- Brewer MB (1991) The social self: On being the same and different at the same time. *Personality Soc. Psych. Bull.* 17(5):475–482.
- Bussemeyer JR, Myung IJ (1992) An adaptive approach to human decision making: Learning theory decision theory and human performance. *J. Experiment. Psych.: General* 121(2):177–194.
- Bussemeyer JR, Stout JC (2002) A contribution of cognitive decision models to clinical assessment: Decomposing performance on the Bechara gambling task. *Psych. Assessment* 14(3):253–262.
- Chamley C (2004) *Rational Herds: Economic Models of Social Learning* (Cambridge University Press, New York).
- Cialdini RB, Goldstein NJ (2004) Social influence: Conformity and compliance. *Annual Rev. Psychol.* 55:591–621.
- Coleman JS, Katz E, Menzel H (1966) *Medical Innovation: A Diffusion Study* (Bobbs-Merrill, Indianapolis).
- Condorcet A (1785) *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix* (Imprimerie Royale, Paris), Discours préliminaire, i–cxci; Essai, 1–304.
- Curtis V, Cairncross S, Yonli R (2000) Domestic hygiene and diarrhoea—Pinpointing the problem. *Tropical Medicine Internat. Health* 5(1):22–32.
- Darley JM, Latané B (1968) Bystander intervention in emergencies: Diffusion of responsibility. *J. Personality Soc. Psych.* 8(4):377–383.
- Davis JH (1973) Group decision and social interaction: A theory of social decision schemes. *Psych. Rev.* 80(2):97–125.
- Denrell J (2005) Why most people disapprove of me: Experience sampling in impression formation. *Psych. Rev.* 112(4):951–978.
- Denrell J, Le Mens G (2007) Interdependent sampling and social influence. *Psych. Rev.* 114(2):398–422.
- Denrell J, Le Mens G (2011) Seeking positive experiences can produce illusory correlations. *Cognition* 119(3):313–324.
- Deutsch M, Gerard HB (1955) A study of normative and informational social influences upon individual judgment. *J. Abnormal Soc. Psych.* 51(3):629–636.
- Duflo E, Kremer M, Robinson J (2011) Nudging farmers to use fertilizer: Theory and experimental evidence from Kenya. *Amer. Econom. Rev.* 101(6):2350–2390.
- Eder A, Fiedler K, Hamm-Eder S (2011) Illusory correlations revisited: The role of pseudocontingencies and working-memory capacity. *Quart. J. Experiment. Psych.* 64(3):517–532.
- Elster J (1978) *Logic and Society: Contradictions and Possible Worlds* (Wiley, New York).
- Elster J (1983) *Sour Grapes* (Cambridge University Press, Cambridge, UK).
- Erev I, Barron G (2005) On adaptation maximization and reinforcement learning among cognitive strategies. *Psych. Rev.* 112(4): 912–931.
- Fazio RH, Eiser JR, Shook NJ (2004) Attitude formation through exploration: Valence asymmetries. *J. Personality Soc. Psych.* 87(3):293–311.
- Feiler DC, Tong JD, Larrick RP (2013) Biased judgment in censored environments. *Management Sci.* 59(3):573–591.
- Festinger L (1957) *A Theory of Cognitive Dissonance* (Stanford University Press, Stanford, CA).
- Festinger L, Schachter S, Back KW (1950) *Social Pressures in Informal Groups: A Study of Human Factors in Housing* (Stanford University Press, Stanford, CA).
- Fetchenhauer D, Dunning D (2010) Why so cynical? Asymmetric feedback underlies misguided skepticism regarding the trustworthiness of others. *Psych. Sci.* 21(2):189–193.
- Fiedler K (2000) Beware of samples! A cognitive-ecological sampling approach to judgment biases. *Psych. Rev.* 107(4):659–676.
- Fiedler K (2008) The ultimate sampling dilemma in experience-based decision making. *J. Experiment. Psych.: Learn., Memory, Cognition* 34(1):186–203.
- Fiedler K (2011) Voodoo correlations are everywhere—Not only in neuroscience. *Perspect. Psych. Sci.* 6(2):163–171.
- Fiedler K (2012) Meta-cognitive myopia and the dilemmas of inductive-statistical inference. Ross BH, ed. *The Psychology of Learning and Motivation*, Vol. 57 (Academic Press, London, UK), 1–46.
- Fiedler K, Juslin P (2006) Taking the interface between mind and environment seriously. Fiedler K, Juslin P, eds. *Information Sampling and Adaptive Cognition* (Cambridge University Press, Cambridge, UK), 3–29.
- Galton F (1907) Vox populi. *Nature* 75(1949):450–51.
- Gerard HB, Rotter GS (1961) Time perspective, consistency of attitude and social influence. *J. Abnormal Soc. Psych.* 62(3):565–572.
- Glynn P (2011) Lecture notes for course MS&E321: Stochastic systems. Stanford School of Engineering, Stanford, CA.
- Granovetter MS (1978) Threshold models of collective behavior. *Amer. J. Sociol.* 83(6):1420–1443.
- Grofman B, Owen G, Feld SL (1983) Thirteen theorems in search of the truth. *Theory Decision* 15(3):261–278.
- Hastie R, Kameda T (2005) The robust beauty of majority rules in group decisions. *Psych. Rev.* 112(2):494–508.
- Heider F (1958) *The Psychology of Interpersonal Relations* (Lawrence Erlbaum Associates, Hillsdale, NJ).
- Hendricks K, Sorensen A, Wiseman T (2012) Observational learning and demand for search goods. *Amer. Econom. J.: Microeconomics* 4(1):1–31.
- Henriksson MP, Elwin E, Juslin P (2010) What is coded into memory in the absence of outcome feedback? *J. Experiment. Psych.: Learn., Memory, and Cognition* 36(1):1–16.
- Hogg MA, Turner JC (1987) Social identity and conformity: A theory of referent informational influence. Doise W, Moscovici S, eds. *Current issues in European Social Psychology*, Vol. 2 (Cambridge University Press, New York), 139–182.

- Juslin P, Winman A, Hansson P (2007) The naïve intuitive statistician: A naïve sampling model of intuitive confidence intervals. *Psych. Rev.* 114(3):678–703.
- Karlin S, Rinott Y (1980) Classes of orderings of measures and related correlation inequalities I. Multivariate totally positive distributions. *J. Multivariate Anal.* 10(4):467–498.
- Keynes JM (1936) *General Theory of Employment Interest and Money* (Macmillan, London, UK).
- Krumholz HM, Chen J, Rathore SS, Wang Y, Radford MJ (2003) Regional variation in the treatment and outcomes of myocardial infarction: Investigating New England's advantage. *Amer. Heart J.* 146(2):242–249.
- Krumme C, Cebrian M, Pickard G, Pentland S (2012) Quantifying social influence in an online cultural market. *PLoS ONE*, 7(5):e33785.
- Kunda Z (1990) The case for motivated reasoning. *Psych. Bull.* 108(3):480–498.
- Latané B (1981) The psychology of social impact. *Amer. Psych.* 36(4):343–356.
- Latané B, Wolf S (1981) The social impact of majorities and minorities. *Psych. Rev.* 88(5):438–453.
- Le Mens G, Denrell J (2011) Rational learning and information sampling: On the naïvety assumption in sampling explanations of judgment biases. *Psych. Rev.* 118(2):379–392.
- Le Mens G, Denrell J, Kovács B (2015) Information sampling and the environment: Application to the effect of popularity on evaluations. Working paper.
- Levine JM (1980) Reaction to opinion deviance in small groups. Paulus PB, ed. *Psychology of Group Influence* (Lawrence Erlbaum Associates, Hillsdale, NJ), 375–429.
- Lorenz J, Rauhut H, Schweitzer F, Helbing D (2011) How social influence can undermine the wisdom of crowd effect. *Proc. Natl. Acad. Sci. USA* 108(22):9020–9025.
- MacCoun RJ (2012) The burden of social proof: Shared thresholds and social influence. *Psych. Rev.* 119(2):345–372.
- Mackie DM (1986) Social identification effects in group polarization. *J. Personality Soc. Psych.* 50(4):720–728.
- Meyn SP, Tweedie RL (1993) *Markov Chains and Stochastic Stability* (Springer-Verlag, New York).
- Milgram S, Bickman L, Berkowitz L (1969) Note on the drawing power of crowds of different size. *J. Personality Soc. Psych.* 13(2):79–82.
- Munshi K (2004) Social learning in a heterogeneous population: Technology diffusion in the Indian Green Revolution. *J. Development Econom.* 73(1):185–213.
- Nickerson R (1998) Confirmation bias: A ubiquitous phenomenon in many guises. *Rev. General Psych.* 2(2):175–220.
- North D (1990) *Institutions, Institutional Change, and Economic Performance* (Cambridge University Press, Cambridge, UK).
- O'Connor GT, Quinton HB, Traven ND, Ramunno LD, Dodds TA, Marciniak TA, Wennberg JE (1999) Geographic variation in the treatment of acute myocardial infarction: The cooperative cardiovascular project. *J. Amer. Medical Assoc.* 281(7):627–633.
- Pfeffer J (1998) *The Human Equation* (Harvard Business School Press, Cambridge MA).
- Pierson P (2004) *Politics in Time: History, Institutions, and Social Analysis* (Princeton University Press, Princeton, NJ).
- Prentice DA, Miller DT (1993) Pluralistic ignorance and alcohol use on campus: Some consequences of misperceiving the social norm. *J. Personality Soc. Psych.* 64(2):243–256.
- Ross S, Pekoz EA (2007) *A Second Course in Probability* (www.ProbabilityBookstore.com, Boston, MA).
- Salganik MJ, Dodds PS, Watts DJ (2006) Experimental study of inequality and unpredictability in an artificial cultural market. *Science* 311(5762):854–856.
- Samuelson W, Zeckhauser R (1988) Status quo bias in decision making. *J. Risk Uncertainty* 1(1):7–59.
- Scharfstein DS, Stein JC (1990) Herd behavior and investment. *Amer. Econom. Rev.* 80(3):465–479.
- Schelling TC (1978) *Micromotives and Macrobehaviour* (Norton, New York).
- Simonsohn U, Karlsson N, Loewenstein G, Ariely D (2008) The tree of experience in the forest of information: Overweighing experienced relative to observed information. *Games Econom. Behav.* 62(1):263–286.
- Smith ER, Collins EC (2009) Contextualizing person perception: Distributed social cognition. *Psych. Rev.* 116(2):343–364.
- Spence M (1973) Job market signaling. *Quart. J. Econom.* 87(3):355–374.
- Stasser G, Kerr NL, Davis JH (1980) Influence processes in decision-making groups: A modeling approach. Paulus PB, ed. *Psychology of Group Influence* (Lawrence Erlbaum Associates, Hillsdale, NJ), 432–477.
- Sunstein CR (2009) *Republic.com 2.0* (Princeton University Press, Princeton, NJ).
- Surowiecki J (2004) *The Wisdom of Crowds* (Double Day, New York).
- Sutton R, Barto AG (1998) *Reinforcement Learning* (MIT Press, Cambridge, MA).
- Turner JC (1991) *Social Influence* (Brooks/Cole, Pacific Grove, CA).
- Wood W (2000) Attitude change: Persuasion and social influence. *Annual Rev. Psych.* 51:539–570.
- Zwiebel J (1995) Corporate conservatism and relative compensation. *J. Political Econom.* 103(1):1–25.