Interacting Cascades: An Experiment on Between-Group Information Spillovers*

James C. D. Fisher† and John Wooders‡

July 31, 2021

Abstract

This paper reports the results of an experiment on information spillovers between groups. We find that a player who is a member of multiple groups aggregates information and serves as a conduit through which information from one group spills over to another. We also find that such players are more influential than similarly-situated players who are members of only a single group. We introduce a novel measure of the magnitude of information spillovers, and we identify the behavioral characteristics of the players that determines these magnitudes. While between-group spillovers occur, we find that their magnitudes are smaller than social learning theory predicts.

JEL Codes: C7, C9, D7, and D8.

Key Words: experiment, information spillovers, social learning, networks.

*The authors are grateful for comments from participants at the 2019 Meetings of the Asia Pacific ESA, the 2019 NYUAD WESSI Workshop, the 2020 Meetings of the SEA, the 2021 Meetings of the China Econometric Society, the 2021 ESA Global Conference, and seminar participants at Henan University, the US Federal Housing Finance Administration, the US Patent and Trademark Office, and Wuhan University. The authors are also grateful to Olivier Bochet and Marie Claire Villeval for helpful comments and to Richard Kiser for excellent programming assistance.

†School of Economics, Henan University (jamescdfisher@gmail.com).

‡Division of Social Sciences, New York University Abu Dhabi, and the Center for Behavioral Institutional Design (john.wooders@nyu.edu). Wooders gratefully acknowledges financial support from the Australian Research Council's Discovery Projects funding scheme (project number DP140103566) and from Tamkeen under the NYU Abu Dhabi Research Institute Award CG005.

1 Introduction

Consumers are embedded in social networks. When choosing a cell phone, restaurant, or investment product, a consumer's decision reveals information about product quality to her network neighbors (e.g., friends, family, co-workers, etc.). Her neighbors use the information revealed and their own private information when making their own decisions, and those decisions in turn reveal information to their neighbors, including consumers who are strangers to the original consumer. In this way, information spills over from one consumer to another, both within the original consumer's own social group and, through common links, to other groups in which the original consumer is not a member. Such information spillovers are intuitive and a common feature of much of the theory of social learning on networks, which has characterized their profound implications for advertising, marketing, product demand, and pricing. While within-group spillovers have been studied extensively experimentally, to our knowledge between-group information spillovers have not been examined experimentally.

This paper assesses the existence, strength, and behavioral determinants of between-group information spillovers using two experimental games. The first game, introduced in Bikhchandani, Hirshleifer, and Welch [8], henceforth BHW, captures within-group information spillovers and serves as the benchmark for our analysis. In the game, players sequentially guess the true (but unknown) state. Prior to guessing the state, each player observes the guesses of her predecessors and an informative private signal. BHW shows that equilibrium is characterized by the potential emergence of an information cascade in which players rationally ignore their private signals and instead follow the guesses of predecessors. Henceforth we refer to this seminal game as the "Basic Cascade" game.

The second game is the simplest extension of the Basic Cascade that can accommodate information spillovers between groups. In the extension there are two groups of players and, just as in the Basic Cascade, each player observes the guesses of her predecessors *in her own group* and an informative private signal. There is, in addition, a single distinguished player who is a member of both groups and, as such, observes the guesses of her predecessors in both groups. The guess of this player – called the "common player" – is observed by her successors in both groups, and thus the common player is a conduit through which the guesses of players in one group can influence the guesses of the players in the other. The presence of a common player thus enables the information embodied in guesses to spill over from one group to another. We refer to this extension as the "Interacting Cascade" game.

In this paper we test whether information spills over between groups, we measure the magnitudes of information spillovers, and we identify the behavioral characteristics of the players that determine these magnitudes. We measure the magnitude of an information spillover from a player i to a player j (possibly in the other group) as the change in the probability that j guesses the state correctly when i guesses it correctly rather than incorrectly. Equilibrium play identifies the theoretical values of these marginal effects and delivers four testable hypotheses:

Hypothesis 1. Information spills over within groups, i.e., the marginal effect of any player in a group on each of her own-group successors is strictly positive.

Hypothesis 2. The common player aggregates information across groups, i.e., every predecessor of the common player has a strictly positive marginal effect on the common player.

Hypothesis 3. Information spills over between groups, i.e., any predecessor of the common player in one group has a strictly positive marginal effect on every successor of the common player in the other group.

Hypothesis 4. The common player is more influential in the Interacting Cascade than in the Basic Cascade, i.e., the marginal effect of the common player on each of her successor(s) is larger in the Interacting Cascade than that of the similarly-situated player in the Basic Cascade.

As noted earlier, in both games a player observes the guesses of her predecessors in her own group and a private signal whose distribution depends on the state. There are two equally-likely states, "mostly blue" and "mostly red," and there are two signals, "Blue" and "Red." In the mostly blue state the probability the private signal is Blue is $\frac{2}{3}$ and the probability the signal is Red is $\frac{1}{3}$. These probabilities are reversed in the mostly red state. After a player observes her private signal, she guesses the state. We say that a player "follows their signal" if they guess "mostly blue" following a Blue signal and "mostly red" following a Red signal.

Detecting spillovers is not as simple as identifying correlation in guesses because guesses may be statistically correlated even in the absence of information spillovers. An appropriate measure of information spillovers must account for the possibility that guesses may be statistically correlated even in the absence of information spillovers because the distribution of signals depends on the true state. To illustrate, suppose that players 1 and 2 each follows their signals, and player 2 observes player 1's guess. There can be no information spillover since 2's guess does not depend on 1's guess. Nonetheless, 2's guess will be correlated with 1's guess: player 2 guesses Blue with probability $\frac{5}{9}$ if player 1 guesses Blue, but guesses Blue with only probability $\frac{4}{9}$ if player 1 guesses Red. Thus, when examining correlation in guesses one must *condition on the state*.[1]

Chi-square tests of independence shows that, conditional on the state, the null hypothesis that guesses are statistically independent is rejected at the 10% (or smaller) significance level for: (i) the guesses of players in the same group, (ii) the guesses of the common player and her predecessors, and (iii) the guesses of players moving before the common player in one group and moving after the common player in the other group. This simple non-parametric analysis of the data provides strong evidence of information spillovers and Hypotheses 1 to 3.

Comparison of the empirical marginal effects to the equilibrium marginal effects reveals that information spillovers are smaller than predicted by the theory. For example, the empirical marginal effect of the guess of a predecessor of the common player in one group on the guess of a successor of the common player in the other group is around 50% of its predicted value.

To understand the behavioral characteristics of the players that drive the smaller-than-predicted marginal effects, we estimate logistic quantal response equilibrium (QRE) models (see McKelvey and Palfrey [38]) for the experiment games. QRE models capture that players imperfectly maximize payoffs, choosing actions with higher expected payoffs with higher probability. At the same time, they are equilibrium models in that each player's belief, determined by Bayes' Rule, incorporates that her co-players' decisions are noisy. Our QRE models include an additional parameter that captures the possibility that players are not fully Bayesian and suffer from the base rate fallacy, tending to overweight the information obtained from their private signal when forming their beliefs.

We estimate the structural QRE models via maximum likelihood. The results indicate that the degree to which decision-making is noisy depends on a player's position. The estimates support the conclusion that players suffer from the base rate fallacy. We find that the marginal effects implied by the estimated QRE models, which we call the "structural marginal effects," are in line with the empirical marginal effects. Thus, noisy decision-making and the base rate fallacy explain well the observed discrepancy between the empirical and equilibrium marginal effects.

The structural model enables a test of Hypothesis 4 that the common player is more influential in the Interacting Cascade than the Basic Cascade. We consider the null hypothesis that the structural marginal effect of the common player on each of her successors is *smaller* in the Interacting Cascade than the Basic Cascade, against the alternative that it is strictly

[1]In this example, the players' guesses are statistically independent conditional on the state. Conditional on the state being mostly blue, for instance, player 2 guesses Blue with probability $\frac{2}{3}$ regardless of the guess of player 1.

larger. A Wald test rejects this null at the 1.4% significance level, thereby providing strong support for Hypothesis 4.[2]

The structural model also enables a decomposition of the difference between the equilibrium and structural marginal effects to determine the portion of the difference that is due to noisy decision-making and the portion due to the base rate fallacy. For the marginal effects measuring between-group information spillovers, noisy decision-making accounts for about 70% of the difference and the base rate fallacy accounts the rest. For the marginal effects measuring within-group information spillovers, the importance of noisy decision-making and the base rate fallacy depends on player position. To elaborate, consider the difference between the equilibrium and structural marginal effects of player i on player j. For j early in the sequence, the base rate fallacy accounts for the majority of the difference. However, for j late in the sequence, noisy decision-making accounts for the majority of the difference. Across all the marginal effects measuring within-group information spillovers, noisy decision-making accounts for the majority of the difference.

The balance of this section discusses the related literature. Section 2 develops the experimental games and hypotheses. Section 3 describes the experimental procedures and presents preliminary non-parametric results. Section 4 develops the structural models, structural marginal effects, and structural hypothesis tests. Section 5 gives concluding remarks. Appendices provide supplemental information about equilibrium, the structural models and methods, the instructions, and the data. The Online Archive contains the data and computer code; it is available here.

Related Literature

Our work is grounded in the theoretical literature on observational learning. Beginning with Banerjee [6] and BHW [8], this broad literature studies games where players are networked (in multiple interconnected groups) and observe their neighbors before making decisions. A common equilibrium feature of these games is that information spills over both between and within groups (as a consequence of Bayesian updating) – e.g., Acemoglu, Dahleh, Lobel, and Ozdaglar [1], Fisher and Wooders [27], and Lobel and Sadler [36], to name but a few. Bikhchandani, Hirshleifer, Tamuz, and Welch [7] provide a recent survey of these studies.

There is a rich experimental literature on observational learning that documents within-group information spillovers.[3] This literature began with Anderson and Holt [3], and includes

[2] Wald tests likewise provide support for Hypotheses 1 to 3 for both the Interacting and Basic Cascade, consistent with our non-parametric tests.

[3] Field experiments and empirical studies have also found evidence of within-group information spillovers, e.g., Cai, Chen, and Fang [9], Cheng, Rai, Tian, and Xu [14], Cui, Zhang, and Bassamboo [18], Da and Huang [19], Lee, Hosanagar, and Tan [35], Gillingham and Bollinger [29], and Zhang and Liu [42], to name

Agranov, Lopez-Moctezuma, Strack, and Tamuz [2], Angrisani, Guarino, Jehiel, Kitagawa [5], Celen and Hyndman [10], Celen, Kariv, and Schotter [11], De Filippis, Guarino, Jehiel, and Kitagawa [21], Duffy, Hopkins, Kornienko, and Ma [22], Evdokimov and Garfagnini [23], Eyster, Rabin, and Weizsacker [24], Fahr and Irlenbusch [25], Goeree, Palfrey, Rogers, and McKelvey [30], and March and Ziegelmeyer [37].[4] Anderson and Holt [4] provide a survey of these experiments.

In contrast, (to our knowledge) there are no experiments that study information flows between different groups. Our first contribution to the literature is thus a simple experimental design that allows us to cleanly assesses the existence, strength, and behavioral determinants of between-group information spillovers. Using this design, we establish that between-group information spillovers occur, but are smaller than theory predicts principally due to noisy decision-making. In the process, we also (i) confirm within-group information spillovers occur and (ii) establish that they are smaller than theory predicts primarily due to noisy decision-making. The latter comprises our second contribution to the literature since (to our knowledge) there are no experiments that directly assess the strength of within-group information spillovers. Overall, our results both validate the theory's key prediction of between-group and within-group information spillover and illuminate gaps in the theory regarding the intensity of information spillovers. (We discuss the broader implications of these gaps at the conclusion of this paper.)

The observational learning literature is part of a broader literature on social learning, which studies games that allow for networked players to learn through repeated observation of neighbors' decisions.[5] Recent experiments in this literature feature multiple, interconnected groups – e.g., Chandrasekhar, Larreguy, and Xandri [12] and Grimm and Mengel [33].[6] They do not, however, study between-group information spillovers, which is the focus of the present paper. Instead, they focus on the identification of players' belief updating rules.

2 Experimental Games and Hypotheses

We describe the experimental games and develop the hypotheses in this section.

but a few.

[4]In related work, Frey and Van De Rijt [28] experimentally examine sequential voting procedures based on the Basic Cascade and find that they provide more accurate outcomes than simultaneous voting procedures because they allow for the spillover of voters' private information.

[5]Choi, Gallo, and Kariv [15] and Golub [31], respectively, provide overviews of recent experiments and theory in this broad literature.

[6]In a related study, Davis, Gaur, and Kim [20] examine the impact of different types of observational information – e.g., on others' decisions or on others' payoffs – on product demand, in a simple design where players repeatedly visit different retailers while observing all their co-players.

THE EXPERIMENTAL GAMES

We study two games, a "Basic Cascade" and an "Interacting Cascade." The Basic Cascade has a single group of players, while the Interacting Cascade has two groups of players with one member in common. In the experiment, (i) the Basic Cascade has four players – labeled 1, 2, 3, and 4 – and (ii) the Interacting Cascade has seven players – labeled 1^A, 1^B, 2^A, 2^B, 3, 4^A, and 4^B – who are members of two groups – labeled A and B – that share player 3 in common. Given her role in the Interacting Cascade, we refer to player 3 as the "common player." An Interacting Cascade can be seen as two Basic Cascades that share a single member in common, is one of the simplest games with the potential for between-group information spillovers. A stylized representation of both games is given in Figure 1.

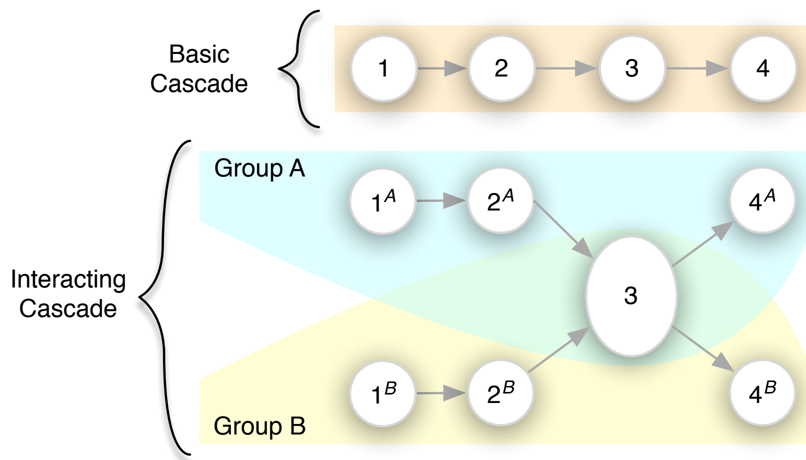


Figure 1: Basic and Interacting Cascades

In both games, Nature moves first and randomly selects, with equal probability, one of two states, which are named "mostly blue," denoted $\mathcal{B}$, and "mostly red," denoted $\mathcal{R}$. Let $c \in \{\mathcal{B}, \mathcal{R}\}$ denote the state. The players do not observe c. The players then move sequentially, according to their labels, each guessing whether the state is $\mathcal{B}$ or $\mathcal{R}$. Before making her guess, each player i observes (i) a private signal and observes either (ii.a) the guesses of her predecessors in the Basic Cascade or (ii.b) the guesses of her predecessors who are *in her group* in the Interacting Cascade.[7] Each private signal is based on the state: it is either "Blue" or "Red" and, in the $\mathcal{B}$ state, the probability it is Blue is $\frac{2}{3}$ and the probability it is Red is $\frac{1}{3}$; these probabilities are reversed in the $\mathcal{R}$ state. Importantly, player 3 in the Interacting Cascade, as a member of both groups, observes the guesses of her predecessors in *both* groups, in addition to her private signal. Player 3 thus provides a conduit through

[7] In the Interacting Cascade, players do not observe the guesses of out-of-group predecessors. For instance, player 2^A observes the guess of player 1^A, but not the guess of player 1^B.

which information may move between groups. A player's payoff is \$1 if she correctly guesses the state and is \$0 otherwise.

We use $\mathcal{B}$ to denote a blue signal or guess and $\mathcal{R}$ to denote a red signal or guess. For the Basic Cascade, write i for the i-th player to move, $s_i \in \{\mathcal{B}, \mathcal{R}\}$ for her signal, and $g_i \in \{\mathcal{B}, \mathcal{R}\}$ for her guess. For the Interacting Cascade, write i^γ for the i-th player to move in group $\gamma \in \{A, B\}$, write $s_i^\gamma \in \{\mathcal{B}, \mathcal{R}\}$ for her signal, and write g_i^γ for her guess. It will be convenient to write g_3, g_3^A, or g_3^B interchangeably for the guess of the common player, and likewise for her signal. For all but the common player, write $\sigma_i^\gamma(g_i^\gamma | g_1^\gamma, \ldots, g_{i-1}^\gamma, s_i^\gamma)$ for the perfect Bayesian equilibrium strategy of player i^γ, i.e., the equilibrium probability that player i^γ guesses g_i^γ given signal s_i^γ and guesses $(g_1^\gamma, \ldots, g_{i-1}^\gamma)$ by her predecessors. We denote the common player's perfect Bayesian equilibrium strategy by $\sigma_3(g_3 | g_1^A, g_2^A, g_1^B, g_2^B, s_3)$.

Equilibrium

In both the Basic and Interacting Cascades there is a unique perfect Bayesian equilibrium under the standard tie-breaking rule that players guess $\mathcal{B}$ or $\mathcal{R}$ with equal probability when indifferent. We illustrate the argument for the first two players of the Interacting Cascade.

Consider player 1^A, who (simultaneously with player 1^B) moves first in the Interacting Cascade. She observes s_1^A and, by Bayes' rule, her (perfect Bayesian) equilibrium belief that the state is $\mathcal{B}$ is

$$\mu_1^A(s_1^A) = \frac{\pi(s_1^A|\mathcal{B})\frac{1}{2}}{\pi(s_1^A|\mathcal{B})\frac{1}{2} + \pi(s_1^A|\mathcal{R})\frac{1}{2}} = \begin{cases} \frac{2}{3} & \text{if } s_1^A = \mathcal{B} \\ \frac{1}{3} & \text{if } s_1^A = \mathcal{R}, \end{cases}$$

where $\pi(s|c)$ is the probability of signal s when the state is c. (We have that $\pi(s|c) = \frac{2}{3}$ if $s = c$ and $\pi(s|c) = \frac{1}{3}$ if $s \neq c$.) Since she earns \$1 for a correct guess and \$0 otherwise, her expected payoff to guessing $\mathcal{B}$ is $\mu_1^A(s_1^A)$ and to guessing $\mathcal{R}$ is $1 - \mu_1^A(s_1^A)$. Consequently, $\mathcal{B}$ is an optimal guess when $\mu_1^A(s_1^A) \geq \frac{1}{2}$ and $\mathcal{R}$ is optimal when $\mu_1^A(s_1^A) \leq \frac{1}{2}$. Hence, the equilibrium strategy of 1^A is to guess the same color as her signal, i.e.,

$$\sigma_1^A(g_1^A|s_1^A) = \begin{cases} 0 & \text{if } g_1^A \neq s_1^A \\ 1 & \text{if } g_1^A = s_1^A. \end{cases}$$

Her guess thus generates information spillovers for her successors in her own group and (as we will see) for player 4^B as well, via the common player.

Write $P(g_1^A|c) = \sum_{s\in\{\mathcal{B},\mathcal{R}\}} \sigma_1^A(g_1^A|s)\pi(s|c)$ for the equilibrium probability that player 1^A guesses g_1^A when the state is c. After 1^A makes her guess, then player 2^A observes (g_1^A, s_2^A).

Given (g_1^A, s_2^A) and σ_1^A, the equilibrium belief of 2^A that the state is $\mathcal{B}$ is

$$\mu_2^A(g_1^A, s_2^A) = \frac{P(g_1^A|\mathcal{B})\pi(s_2^A|\mathcal{B})\frac{1}{2}}{P(g_1^A|\mathcal{B})\pi(s_2^A|\mathcal{B})\frac{1}{2} + P(g_1^A|\mathcal{R})\pi(s_2^A|\mathcal{R})\frac{1}{2}} = \begin{cases} \frac{1}{5} & \text{if } g_1^A = s_2^A = \mathcal{R} \\ \frac{1}{2} & \text{if } g_1^A \neq s_2^A \\ \frac{4}{5} & \text{if } g_1^A = s_2^A = \mathcal{B}. \end{cases}$$

Thus, the equilibrium strategy of 2^A is to guess $\mathcal{B}$ when $g_1^A = s_2^A = \mathcal{B}$ and to guess $\mathcal{R}$ when $g_1^A = s_2^A = \mathcal{R}$. When $g_1^A \neq s_2^A$, then $\mu_2^A(g_1^A, s_1^A) = \frac{1}{2}$ and 2^A guesses $\mathcal{B}$ or $\mathcal{R}$ with equal probability (by the tie-breaking assumption). Formally, her equilibrium strategy is

$$\sigma_2^A(g_2^A|g_1^A, s_2^A) = \begin{cases} 0 & \text{if } g_2^A \neq g_1^A = s_2^A \\ \frac{1}{2} & \text{if } g_1^A \neq s_2^A \\ 1 & \text{if } g_2^A = g_1^A = s_2^A. \end{cases}$$

Notice that the equilibrium strategies and beliefs of 1^A and 2^A are uniquely determined. The analysis for players 1^B and 2^B is symmetric.

The characterization of equilibrium for player 3 and player 4^A (and player 4^B by symmetry) is straightforward; Appendix A provides the details. The characterization of the unique perfect Bayesian equilibrium of a Basic Cascade is analogous; see BHW [8] for details. We utilize these characterizations to compute the marginal effects and develop our core hypotheses concerning information spillovers.

Marginal Effects and Hypotheses

An information spillover occurs when the guess of one player affects the guess of another. We measure an information spillover from player i^γ to player $j^{\gamma'}$ as the change in the probability of a correct guess by player $j^{\gamma'}$ that results from player i^γ guessing correctly rather than incorrectly. To illustrate, consider players 1^A and 2^A in the Interacting Cascade. Write $P(g_2^A|g_1^A, c) = \sum_{s\in\{\mathcal{B},\mathcal{R}\}} \sigma_2^A(g_2^A|g_1^A, s)\pi(s|c)$ for the equilibrium probability that 2^A guesses g_2^A when 1^A guesses g_1^A and when the state is c. The **equilibrium marginal effect of player1$^{\boldsymbol{A}}$ on player 2$^{\boldsymbol{A}}$** is defined as

$$\Delta(1^A, 2^A) = P(g_2^A = \mathcal{B}|g_1^A = \mathcal{B}, c = \mathcal{B}) - P(g_2^A = \mathcal{B}|g_1^A = \mathcal{R}, c = \mathcal{B}),$$

i.e., it is the change in the probability that player 2^A correctly guesses $\mathcal{B}$ when 1^A correctly guesses $\mathcal{B}$ rather than incorrectly guesses $\mathcal{R}$. It is straightforward to verify that $\Delta(1^A, 2^A) = 1/2$. Since the payoff to guessing correctly is 1, the marginal effect can also be interpreted as the equilibrium expected payoff gain to 2^A when 1^A guesses correctly rather than incorrectly.

In general, for $i < j$ and $\gamma, \gamma' \in \{A, B\}$, the **equilibrium marginal effect of player i^γ on player $j^{\gamma'}$** is defined by

$$\Delta(i^\gamma, j^{\gamma'}) = P(g_j^{\gamma'} = \mathcal{B} | g_i^\gamma = \mathcal{B}, c = \mathcal{B}) - P(g_j^{\gamma'} = \mathcal{B} | g_i^\gamma = \mathcal{R}, c = \mathcal{B}),$$

where $P(g_j^{\gamma'} | g_i^\gamma, c)$ is the equilibrium probability player $j^{\gamma'}$ guesses $g_j^{\gamma'}$ given player i^γ guesses g_i^γ and given the state is c.[8]

When defining marginal effects, there are two reasons to take the state as given. First, by doing so, the marginal effect measures information spillovers in terms of the payoff consequences to a player when a predecessor guesses correctly rather than incorrectly. Second, taking the state as given removes the correlation in guesses that results simply from the dependency of players signals on the state, and instead measures the dependency in guesses arising from observation alone. To illustrate this issue, let $P(g_1^B | g_1^A)$ be the equilibrium probability that player 1^B guesses $g_1^{\mathcal{B}}$ given player 1^A guesses g_1^A. It is easy to compute that

$$P(g_1^B = \mathcal{B} | g_1^A = \mathcal{B}) - P(g_2^A = \mathcal{B} | g_1^A = \mathcal{R}) = \frac{5}{9} - \frac{4}{9} > 0.$$

In other words, 1^B guesses $\mathcal{B}$ with higher probability when 1^A guesses $\mathcal{B}$ than when she guesses $\mathcal{R}$. Clearly, however, 1^A's guess cannot influence 1^B's guess as 1^B does not observe it. By contrast, the equilibrium marginal effect $\Delta(i^\gamma, j^{\gamma'})$ defined above removes this spurious correlation – the equilibrium marginal effect of 1^A on 1^B is $\Delta(1^A, 1^B) = 0$.

Table 1 provides the equilibrium marginal effects for the Basic Cascade and the Interacting Cascade. It shows the theoretical magnitudes of information spillovers between and within groups, and thus forms the basis for our hypotheses and forms the benchmark against which our experimental results are measured. The table is read "row-to-column" – e.g., the equilibrium marginal effect of player 2^A on player 4^B is stored in the row labeled "$\Delta(2^A, j^{\gamma'})$" (under the "Player 2^A" heading) at the column labeled "Player 4^B," its value is 0.346.

Proposition 1. Equilibrium Marginal Effects
Table 1 gives the equilibrium marginal effects for the Basic and Interacting Cascades.

Proof. The proof is given in Appendix A. □

Table 1 gives that a correct guess by a player raises the (equilibrium) probability that each of her successors in her group guesses the state correctly. For instance, in the Interacting Cascade, a correct guess by player 1^γ increases the likelihood of a correct guess by players

[8] Since the states are symmetric, it would be equivalent to define the equilibrium marginal effect as $P(g_j^{\gamma'} = \mathcal{R} | g_i^\gamma = \mathcal{R}, c = \mathcal{R}) - P(g_j^{\gamma'} = \mathcal{R} | g_i^\gamma = \mathcal{B}, c = \mathcal{R})$.

	Interacting Cascade					Basic Cascade		
	Player ($j^{Y'}=$) 2^A	Player 2^B	Player 3	Player 4^A	Player 4^B	Player 2	Player 3	Player 4
Player 1^A								
$P(g_i^{Y}=\mathcal{B} \mid g_1^{A}=\mathcal{B}, c=\mathcal{B})$	0.833	.	0.889	0.909	0.868	0.833	0.944	0.944
$P(g_i^{Y}=\mathcal{B} \mid g_1^{A}=\mathcal{R}, c=\mathcal{B})$	0.333	.	0.481	0.440	0.523	0.333	0.222	0.222
$\Delta(1^A, j^{Y'})$	**0.500**	.	**0.407**	**0.469**	**0.346**	**0.500**	**0.722**	**0.722**
Player 1^B								
$P(g_i^{Y}=\mathcal{B} \mid g_1^{B}=\mathcal{B}, c=\mathcal{B})$	.	0.833	0.889	0.868	0.909	.	.	.
$P(g_i^{Y}=\mathcal{B} \mid g_1^{B}=\mathcal{R}, c=\mathcal{B})$	.	0.333	0.481	0.523	0.440	.	.	.
$\Delta(1^B, j^{Y'})$	.	**0.500**	**0.407**	**0.346**	**0.469**	.	.	.
Player 2^A								
$P(g_i^{Y}=\mathcal{B} \mid g_2^{A}=\mathcal{B}, c=\mathcal{B})$	.	.	0.889	0.909	0.868	.	0.944	0.944
$P(g_i^{Y}=\mathcal{B} \mid g_2^{A}=\mathcal{R}, c=\mathcal{B})$	.	.	0.481	0.440	0.523	.	0.222	0.222
$\Delta(2^A, j^{Y'})$	.	.	**0.407**	**0.469**	**0.346**	.	**0.722**	**0.722**
Player 2^B								
$P(g_i^{Y}=\mathcal{B} \mid g_2^{B}=\mathcal{B}, c=\mathcal{B})$	.	.	0.889	0.868	0.909	.	.	.
$P(g_i^{Y}=\mathcal{B} \mid g_2^{B}=\mathcal{R}, c=\mathcal{B})$	.	.	0.481	0.523	0.440	.	.	.
$\Delta(2^B, j^{Y'})$	.	.	**0.407**	**0.346**	**0.469**	.	.	.
Player 3								
$P(g_i^{Y}=\mathcal{B} \mid g_3=\mathcal{B}, c=\mathcal{B})$	.	.	.	0.982	0.982	.	.	0.965
$P(g_i^{Y}=\mathcal{B} \mid g_3=\mathcal{R}, c=\mathcal{B})$	.	.	.	0.056	0.056	.	.	0.083
$\Delta(3, j^{Y'})$	.	.	.	**0.926**	**0.926**	.	.	**0.882**

Cells either give (i) the equilibrium probability that the column player guesses correctly given the specified guess of the row player and given the state is mostly blue or (ii) the equilibrium marginal effect of the row player on the column player (**bolded**). Basic Cascade results are reported in the group A rows for expositional simplicity.

Table 1: Equilibrium Marginal Effects

2^γ, 3, and 4^γ by 0.5, 0.407, and 0.346 respectively for $\gamma \in \{A, B\}$. Similarly, in the Basic Cascade, a correct guess by player 1 increases the likelihood of a correct guess by players 2, 3, and 4 by at least 0.5. Intuitively, this occurs because a player's guess reveals information about her signal to those who directly observe it – namely, her successors in her group – that they incorporate into their beliefs and guesses. In other words, information spills over within groups.

Hypothesis 1. Information spills over within groups, i.e., the (equilibrium) marginal effect of any player on each of her successors her group is strictly positive.

Table 1 also gives that a correct guess by each of players 1^A, 1^B, 2^A, and 2^B increases the probability of a correct guess by player 3 by 0.407. Intuitively, this occurs because the common player infers the information contained in each of her predecessors' guesses across both groups and incorporates it into her belief and guess. In other words, the common player aggregates information across both groups.

Hypothesis 2. The common player aggregates information across groups, i.e., the marginal effects of players 1^A, 1^B, 2^A, and 2^B on the common player are all strictly positive.[9]

Table 1 further shows that (i) a correct guess by player 1^A or player 2^A increases the probability of a correct guess by player 4^B by 0.346 and (ii) a correct guess by 1^B and 2^B increases the likelihood of a correct guess by player 4^A by the same amount. Thus, the information embodied in the guesses of the member of one group influences the guesses of the members of the other group, despite the lack of direct observation of the guesses of the members of the original group. Intuitively, this occurs because the common player serves as a conduit for information to move between groups: since the common player's guess embeds information aggregated from both groups, her successors in one group use it to draw inferences about the signals of the players in the other group, which they incorporate into their beliefs and guesses.

Hypothesis 3. Information spills over between groups, i.e., the marginal effects of players 1^A or player 2^A on player 4^B are strictly positive, as are those of players 1^B and 2^B on player 4^A.

Our last hypothesis concerns the magnitudes of the (equilibrium) marginal effects in the Basic and Interacting Cascades. Table 1 shows that in the Interacting Cascade the marginal effect of player 3's guess on player 4^A and player 4^B is greater than the marginal effect of player 3's guess on player 4 in the Basic Cascade, i.e., $\Delta(3, 4^A) = \Delta(3, 4^B) = .926 > \Delta(3, 4) = .882$. This result is intuitive since player 3's guess in the Interacting Cascade

[9] We evaluate Hypothesis 2 separately from Hypothesis 1 because it is necessary for Hypothesis 3.

reflects the information embodied in the guesses of four predecessors, rather than just two in the Basic Cascade. Consequently, players 4^A and 4^B shift their beliefs more in response to a change in player 3's guess than does player 4, resulting in larger changes to their probabilities of a $\mathcal{B}$ guess. In this sense, player 3 is more influential in the Interacting Cascade than the Basic Cascade.

Hypothesis 4. The common player is more influential in the Interacting Cascade than in the Basic Cascade. In particular, the marginal effect of player 3 on her successors is greater in the Interacting Cascade than in the Basic Cascade.

3 Experimental Procedures and Preliminary Analysis

We describe the experimental procedures in this section. We then study information spillovers via a non-parametric analysis of the data, which first focuses on the marginal effects and then on tests of independence. Appendix D provides a summary of the data.[10]

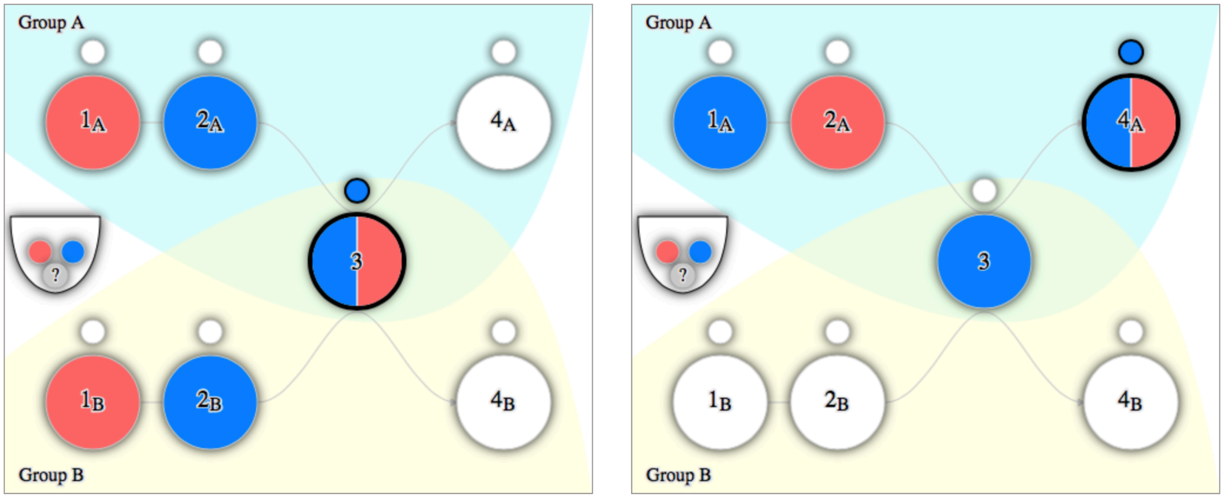
Experimental Procedures

Our experiment consisted of 12 sessions of each of the Basic and Interacting Cascades, which were conducted at the University of Arizona's Economic Sciences Laboratory (ESL). For each session of the Interacting Cascade, seven subjects were recruited and randomly assigned to one of the seven roles $(1^A, 1^B, \ldots, 4^A, 4^B)$. Each subject played 15 rounds of the game in the same fixed role. Likewise, for each session of the Basic Cascade, four subjects were recruited, they were randomly assigned to one of the four roles $(1, 2, 3, 4)$, and they played 15 rounds of the game in the same fixed role. Thus, a total of 132 subjects participated, with each making 15 guesses. Subjects were paid \$1 for each correct guess, in addition to \$10 for participating. Subjects were recruited from the ESL's database, which primarily draws from the university's student body, and each subject participated in only one session.

In the experiment, there were two identical and opaque containers, holding blue and red poker chips: the $\mathcal{B}$ container had two blue chips and one red chip, and the $\mathcal{R}$ container had two red chips and one blue chip. At the start of each round (of a session), one of the two containers was randomly and privately selected.[11] Each subject, when it was her turn to

[10]The Online Archive contains the raw data.

[11]Before the first round of a session, instructions were read aloud to the subjects, who then completed a computerized demonstration that was designed to familiarize them with the software, the procedures, and the experimental game. In the demonstration, each subject played multiple rounds of the game with simulated co-players, while taking on a different player's role in each round. The instructions for the Interacting Cascade treatment are reproduced in Appendix C; all instructions and procedures are given in the Online Archive.



(a) When Player 3 Guesses (b) When Player 4^A Guesses

Figure 2: Computer Software Screenshots

guess, drew a chip at random from the selected container, privately observed its color, and then returned the chip to the container.[12] The experimenter also observed the color of the drawn chip and entered it into a smartphone so that the color appeared on the subject's computer screen and was recorded in a database. Each subject then observed her group's history of guesses on her computer screen and then made her own guess via her computer screen.

Figure 2a shows the screen of player 3 in the Interacting Cascade after observing her signal, but prior to making her guess. The players' positions are represented by large circles (a player's own position is bolded), and the color of a large circle displays the player's guess. Signals are represented by small circles. In the figure, the small circle for player 3 shows that she drew a $\mathcal{B}$ chip. White circles represent signals and guesses that player 3 does not observe. At player 3's turn, she would guess $\mathcal{B}$ by clicking on the blue half-circle and guess $\mathcal{R}$ by clicking on the red half-circle. After player 3 guessed, it became players 4^A and 4^B' turns. Figure 2b show player 4^A's screen when her own signal was $\mathcal{B}$ and when player 3 guessed $\mathcal{B}$.

At the end of each round: (i) the experimenter revealed whether the container was $\mathcal{B}$ or $\mathcal{R}$ by emptying its contents in front of the subjects and entered the result into a smartphone for recording in a database; and (ii) each subject's computer screen revealed whether the

[12]The use of physical poker chips and containers made the signal-generating process transparent to subjects; a pilot experiment indicated subjects were "suspicious" of computer-generated signals. While it is straightforward to fully implement a "pencil and paper" experiment with the Basic Cascade (e.g., Anderson and Holt [3]), such an approach is impractical for the Interacting Cascade since not all subjects observe all guesses as they are made.

subject's guess was correct. The Basic Cascade experiment proceeded analogously.

Preliminary Analysis

We begin the analysis of the data by introducing empirical measures of the equilibrium marginal effects, which we term the "empirical" marginal effects. Since behavior may differ across states, we distinguish between the empirical marginal effects for the $\mathcal{B}$ and $\mathcal{R}$ states. In the Interacting Cascade, for $i < j$ and $\gamma, \gamma' \in \{A, B\}$, the **empirical marginal effect of player i^γ on player $j^{\gamma'}$ for the $\kappa \in \{\mathcal{B}, \mathcal{R}\}$ state** is

$$\hat{\Delta}(i^\gamma, j^{\gamma'}, \kappa) = \hat{P}(g_j^{\gamma'} = \kappa | g_i^\gamma = \kappa, c = \kappa) - \hat{P}(g_j^{\gamma'} = \kappa | g_i^\gamma \neq \kappa, c = \kappa),$$

where $\hat{P}(g_2^A | g_1^A, c)$ is the empirical frequency that player $i^{\gamma'}$ guesses $g_j^{\gamma'}$ given player i^γ guesses g_i^γ and given the state is c.[13,14] Empirical marginal effects for the Basic Cascade are defined analogously.

Table 2 reports the empirical marginal effects. The table shows, for instance, that $\hat{\Delta}(1^A, 3, \mathcal{B}) = 0.479$ and $\hat{\Delta}(1^A, 3, \mathcal{R}) = 0.218$, i.e., a correct guess by player 1^A increases the frequency of a correct guess by player 3 by 0.479 and 0.218 when the states are $\mathcal{B}$ and $\mathcal{R}$ respectively. Comparing Tables 1 and 2 shows that the empirical marginal effects are, on average, smaller than their equilibrium counterparts in both the Basic and Interacting Cascades. For instance, the empirical marginal effects that measure between-group information spillovers – i.e., those identified by Hypothesis 3 – average about 47% of their equilibrium values. We will return to these differences in Section 4, wherein we investigate the behavioral determinants of the differences.

Table 2 provides preliminary evidence in support of our hypotheses. First, it shows that the empirical marginal effects of each player on her successors in her group are always positive in both the Basic and Interacting Cascades, suggesting that information spills over within groups. Second, it shows that the empirical marginal effects of players 1^A, 1^B, 2^A, and 2^B on player 3 are positive, suggesting that the common player aggregates information across groups. Third, it shows that the empirical marginal effects of players 1^A and 2^A on player 4^B, as well as those of players 1^B and 2^B on player 4^A, are positive, suggesting that information spills over between groups. Fourth, it shows that player 3 has a greater empirical marginal effect in the Interacting Cascade than the Basic Cascade, indicating common player is more influential in the Interacting Cascade than the similarly-situated player in the Basic Cascade.

Since information spillovers from one player to another theoretically correlate the pair's

[13] We label subjects by their roles; so, when we discuss the guesses of player i^γ, we mean the guesses of all subjects who played the role of player i^γ during the experiment.

[14] If subjects follow equilibrium, then it is readily verified that the empirical marginal effects are equal in expectation to the equilibrium marginal effects.

	Interacting Cascade					Basic Cascade		
	Player ($j^{Y'}=$) 2^A	Player 2^B	Player 3	Player 4^A	Player 4^B	Player 2	Player 3	Player 4
Player 1^A								
$\hat{\Delta}(1^A, j^{Y'}, \mathcal{B})$	0.373	.	0.479	0.672	0.145	0.222	0.365	0.459
$\hat{\Delta}(1^A, j^{Y'}, \mathcal{R})$	0.214	.	0.218	0.409	0.077	0.128	0.283	0.387
Player 1^B								
$\hat{\Delta}(1^B, j^{Y'}, \mathcal{B})$	.	0.215	0.394	0.253	0.577	.	.	.
$\hat{\Delta}(1^B, j^{Y'}, \mathcal{R})$	.	0.414	0.466	0.215	0.486	.	.	.
Player 2^A								
$\hat{\Delta}(2^A, j^{Y'}, \mathcal{B})$	.	.	0.297	0.368	0.088	.	0.259	0.349
$\hat{\Delta}(2^A, j^{Y'}, \mathcal{R})$	.	.	0.499	0.460	0.247	.	0.321	0.292
Player 2^B								
$\hat{\Delta}(2^B, j^{Y'}, \mathcal{B})$	.	.	0.431	0.151	0.334	.	.	.
$\hat{\Delta}(2^B, j^{Y'}, \mathcal{R})$	.	.	0.377	0.240	0.566	.	.	.
Player 3								
$\hat{\Delta}(3, j^{Y'}, \mathcal{B})$	.	.	.	0.577	0.382	.	.	0.348
$\hat{\Delta}(3, j^{Y'}, \mathcal{R})$	.	.	.	0.694	0.573	.	.	0.487

Each cell gives the empirical marginal effect of the row player on the column player for the specified state. Basic Cascade results are reported in the group A rows for expositional simplicity.

Table 2: Empirical Marginal Effects

	Interacting Cascade					Basic Cascade		
	Player 2^A	Player 2^B	Player 3	Player 4^A	Player 4^B	Player 2	Player 3	Player 4
Player 1^A								
Independence Test p -Values	0.000***	.	0.000***	0.000***	0.253	0.071*	0.000***	0.000***
Player 1^B								
Independence Test p -Values	.	0.000***	0.000***	0.008***	0.000***	.	.	.
Player 2^A								
Independence Test p -Values	.	.	0.000***	0.000***	0.036**	.	0.000***	0.000***
Player 2^B								
Independence Test p -Values	.	.	0.000***	0.040**	0.000***	.	.	.
Player 3								
Independence Test p -Values	.	.	.	0.000***	0.000***	.	.	0.000***

Each cell gives the p -value of a test of independence between the row player's guess and the column player's guess. The null is that their guesses are independent for both states and the alternative is that their guesses are not independent for at least one state. Asterisks denote statistical significance: * indicates a rejection of the null at least at the 10% level; ** indicates a rejection of the null at least at the 5% level; and *** indicates a rejection of the null at least at the 1% level. Basic Cascade results are reported in the group A rows for expositional simplicity.

Table 3: Independence Tests

guesses (holding the state constant), χ^2-Tests of Independence are a simple way to non-parametrically test for the presence of information spillovers. Specifically, we employ a joint χ^2-Test of Independence that evaluates the null that that a pair's guesses are (i) independent when the state is $\mathcal{B}$ and are (ii) independent when the state is $\mathcal{R}$, against the alternative that the pairs' guesses are not independent – i.e., are correlated – for at least one state.[15] Table 3 reports the p-values from this test. The table shows, for instance, that the p-value from the test for players 2^A and 4^B is 0.036.

Table 3 provides support for Hypotheses 1 to 3. First, regarding Hypothesis 1, the table shows that information spills over within a group because the guess of any player is correlated with the guess of each of her successors in the same group. Specifically, for each such pair, the table shows that the null of the test is rejected at least at the 7% significance level in both the Basic and Interacting Cascades, leading us to conclude correlation of the pair's

[15]To elaborate, let $\mathcal{T}_\mathcal{B}$ ($\mathcal{T}_\mathcal{R}$) be the test statistic of a χ^2-Tests of Independence that the pair's guesses are independent when the state is $\mathcal{B}$ ($\mathcal{R}$) against the alternative that they are not independent when the state is $\mathcal{B}$ ($\mathcal{R}$); see Mood, Graybill, and Boes (MGB) [39] for details. Let $\chi^2_k(\cdot)$ denote the Chi-Square cumulative distribution function with k-degrees of freedom. MGB show that, under their respective nulls, $\mathcal{T}_\mathcal{B}$ and $\mathcal{T}_\mathcal{R}$ converge to $\chi^2_1(\cdot)$ as the sample size grows large, with respective p-values of $1-\chi^2_1(\mathcal{T}_\mathcal{B})$ and $1-\chi^2_1(\mathcal{T}_\mathcal{R})$. Hence, under the null of our test, $\mathcal{T}_\mathcal{B}+\mathcal{T}_\mathcal{R}$ converges to $\chi^2_2(\cdot)$. We thus take $\mathcal{T}_\mathcal{B}+\mathcal{T}_\mathcal{R}$ to be our test statistic and observe that it has a p-value of $1-\chi^2_2(\mathcal{T}_\mathcal{B}+\mathcal{T}_\mathcal{R})$.

guesses (for at least one state). Second, regarding Hypothesis 2, the table shows that the common player aggregates information because the guess of player 3 is correlated with each of the guesses of players 1^A, 1^B, 2^A, and 2^B. Specifically, for each such pair, the table shows that null of the test is rejected at least at the 1% significance level, leading us to conclude the correlation of the pair's guesses. Third, regarding Hypothesis 3, the table shows that information spills over between groups because the guesses of players 2^A and 4^B, players 1^B and 4^A, and players 2^B and 4^A are correlated. Specifically, for each such pair, the table shows that null of the test is rejected at least at the 4% significance level, leading us to conclude the correlation of the pair's guesses.

The tests of independence, however, (i) do not account for the direction of information spillovers (i.e., from a player to her successors) and (ii) cannot assess the magnitudes of information spillovers and thus Hypothesis 4. Additionally, non-parametric methods do not generally admit the types of counterfactual analyses that are key to understanding the behavioral determinants of information spillovers. With these factors in mind, we turn to structural analysis for the balance of the paper.

4 Structural Analysis

We describe the structural analysis of the Basic and Interacting Cascades in this section: we first build and estimate structural models, and then test and decompose the "structural" marginal effects that emerge under them.

Structural Models

Our baseline structural model for the Interacting Cascade is a variant of McKelvey and Palfrey's [38] logistic quantile response equilibrium, in which the strategy of each player i^γ (for $\gamma \in \{A, B\}$), after observing her group's history of guesses and her own signal $(g_1^\gamma, \ldots, g_{i-1}^\gamma, s_i^\gamma)$, is

$$\tilde{\sigma}_i^\gamma(g_i^\gamma | g_1^\gamma, \ldots, g_{i-1}^\gamma, s_i^\gamma) = \frac{\mathbb{I}(g_i^\gamma = \mathcal{B}) + \mathbb{I}(g_i^\gamma = \mathcal{R}) e^{\lambda_i(1-2\tilde{\mu}_i^\gamma(g_1^\gamma, \ldots, g_{i-1}^\gamma, s_i^\gamma))}}{1 + e^{\lambda_i(1-2\tilde{\mu}_i^\gamma(g_1^\gamma, \ldots, g_{i-1}^\gamma, s_i^\gamma))}}, \tag{1}$$

where $\lambda_i \geq 0$ is a rationality parameter, $\tilde{\mu}_i^\gamma(g_1^\gamma, \ldots, g_{i-1}^\gamma, s_i^\gamma)$ is the Bayesian belief of i^γ that the state is $\mathcal{B}$, and $\mathbb{I}(\cdot)$ is the indicator function. The model has four parameters $(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$. Players with the same index, but different groups – i.e., i^A and i^B for $i \in \{1, 2, 4\}$ – share the same parameter λ_i because they have symmetric information sets.[16] (For

[16] We previously estimated richer structural models that allowed each player to have her own rationality parameter and found results analogous to those presented in this paper; the details are omitted for expositional simplicity.

expositional simplicity, we typically suppress the dependency of beliefs and strategies on the parameter vector.) Appendix B provides a detailed development of the model (including beliefs) and a discussion of its key properties.

The randomness in a player's guess is driven by noise in her decision-making process, which may be rooted in her bounded rationality, behavioral heuristics, idiosyncratic preference shocks, or other traits. Regardless of the reason for the noise, a player is most likely to guess the state with the highest subjective expected payoff under equation (1). Thus, for each player i^γ and each $(g_1^\gamma, \ldots, g_{i-1}^\gamma, s_i^\gamma)$, we have that (i) $\tilde{\sigma}_i^\gamma(\cdot) \to \sigma_i^\gamma(\cdot)$ as $(\lambda_1, \lambda_2, \lambda_3, \lambda_4) \to \boldsymbol{\infty}$ and that (ii) $\tilde{\sigma}_i^\gamma(\cdot) \to \frac{1}{2}$ as $(\lambda_1, \lambda_2, \lambda_3, \lambda_4) \to \mathbf{0}$; the formal argument is given in Appendix B. In other words, players' strategies converge to equilibrium as the rationality parameters grow large and converge to uniform randomness as the rationality parameters shrink to zero. Thus, higher values of $(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ correspond to lower decision-making noise and greater rationality. An important feature of the model is that players' account for their predecessors' rationality parameters when forming beliefs. Thus, when the model is estimated, the rationality parameters capture information about both (i) a player's own rationality and (ii) her successors' assessments of her rationality.

It will be useful to extend the baseline model to capture the "base rate fallacy." This fallacy, which was first documented by Anderson and Holt [3] in the Basic Cascade and was later formalized by Goeree, Palfrey, Rogers, and McKelvey (GPRM) [30], refers to an individual's tendency to overweight their private information relative to the information provided by their predecessors' guesses. We follow GPRM and extend the baseline model by introducing a new parameter α as an exponent on the contribution of each player's signal to her Bayesian belief, i.e., by replacing each instance of $\pi(s_i^\gamma|c)$ in the belief of player i^γ with $\pi(s_i^\gamma|c)^\alpha$. We refer to α as the "base rate fallacy parameter." Thus, the baseline model obtains when $\alpha = 1$ and the base rate fallacy obtains when $\alpha > 1$. We call this extension the "base rate fallacy extension" and observe that it has five parameters $(\lambda_1, \lambda_2, \lambda_3, \lambda_4, \alpha)$. Appendix B provides a detailed development of this extension.

The baseline structural model for the Basic Cascade and its base rate fallacy extension are analogously defined. Details are given in Appendix B.

The parameters of the structural models are estimated from the data via Maximum Likelihood Estimation (MLE). To elaborate, focus on the baseline structural model for the Interacting Cascade. For the k-th round of play, our data consist of players' signals $(s_1^A, s_1^B, s_2^A, s_2^B, s_3, s_4^A, s_4^B)$ and guesses $(g_1^A, g_1^B, g_2^A, g_2^B, g_3, g_4^A, g_4^B)$, as well as the state c. In light of equation (1), the likelihood of this round is

$$L_k = \frac{1}{2}\tilde{\sigma}_3(g_3|g_1^A, \ldots, g_2^B, s_3)\pi(s_3|c) \prod_{(i,\gamma)\in\{1,2,4\}\times\{A,B\}} \tilde{\sigma}_i^\gamma(g_i^\gamma|g_1^\gamma, \ldots, g_{i-1}^\gamma, s_i^\gamma)\pi(s_i^\gamma|c).$$

Since there are data for 180 rounds of play, the log-likelihood is $\sum_{k=1}^{180}\ln(L_k)$. This sum is maximized via a modified Newton-Raphson algorithm via selection of the parameters to return the MLE estimates $(\hat{\lambda}_1, \hat{\lambda}_2, \hat{\lambda}_3, \hat{\lambda}_4)$. The other structural models are estimated in an analogous fashion. Details are given in Appendix B.

Table 4 reports the estimated parameters for the four structural models, as well as alternative flavors of these models where all players are constrained to have the same rationality parameter λ. The table's rows report the estimates and standard errors for each parameter in columns corresponding to each model, while the top (bottom) sub-table gives the constrained (unconstrained) estimates. The table shows, for instance, that the rationality parameters for the Interacting Cascade structural models range between approximately 5 and 8. This indicates that players are neither fully rational nor irrational, but rather are somewhere in-between and so make decisions with an intermediate amount of noise. The table also shows that the base rate fallacy parameter is approximately 1.2 in the same models, indicating players overweight their own information. The table further shows that all parameters are statistically different from zero at the 1% level of significance.

Table 4 also presents the results of two Wald tests. The first, label "Test of Lambda Equality," is a test of the null that $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4$ against the two-sided alternative that the rationality parameters are unequal. The second, labeled "Test of Base Rate Fallacy," is a test of the null that $\alpha = 1$ against the two-sided alternative that $\alpha \neq 1$. The table shows that both nulls are rejected at least at the 10% significance level in every specification. In light of this, we conclude that the best structural models for both the Basic and Interacting Cascades are the unrestricted, base rate fallacy extensions, which we refer to henceforth as the "preferred structural models." We work exclusively with the preferred structural models in the balance of the paper.

Structural Marginal Effects

We next introduce the analogue of the equilibrium marginal effects for the structural models, which we term the "structural" marginal effects. For the preferred structural model for the Interacting Cascade, let $P(g_j^{\gamma'}|g_i^\gamma, c, \boldsymbol{\lambda}, \alpha)$ be the probability in the model, when its parameters are $(\boldsymbol{\lambda}, \alpha) = (\lambda_1, \lambda_2, \lambda_3, \lambda_4, \alpha)$, that player $j^{\gamma'}$ guesses $g_j^{\gamma'}$ given player i^γ guesses g_i^γ and given the state is c, where $i < j$ and $\gamma, \gamma' \in \{A, B\}$. Let

$$\Delta(i^\gamma, j^{\gamma'}, \boldsymbol{\lambda}, \alpha) = P(g_j^{\gamma'} = \mathcal{B}|g_i^\gamma = \mathcal{B}, c = \mathcal{B}, \boldsymbol{\lambda}, \alpha) - P(g_j^{\gamma'} = \mathcal{B}|g_i^\gamma = \mathcal{R}, c = \mathcal{B}, \boldsymbol{\lambda}, \alpha)$$

Parameters	Interacting Cascade		Basic Cascade	
	Baseline	Base Rate Fallacy Extension	Baseline	Base Rate Fallacy Extension
Constrained Structural Models				
$\hat{\lambda}$	6.647***	6.252***	7.122***	6.324***
	(0.307)	(0.337)	(0.463)	(0.510)
$\hat{\alpha}$	.	1.213***	.	1.358***
	.	(0.071)	.	(0.108)
Number Obs.	180	180	180	180
Log-Likelihood	-365.898	-360.876	-201.291	-194.548
Test of Base Rate Fallacy *p* -Value	.	0.003***	.	0.001***
Unconstrained Structural Models				
$\hat{\lambda}_1$	8.064***	7.275***	9.507***	8.261***
	(0.630)	(0.596)	(1.117)	(1.110)
$\hat{\lambda}_2$	5.793***	5.650***	38.121**	13.119***
	(0.658)	(0.619)	(15.186)	(5.029)
$\hat{\lambda}_3$	6.878***	6.906***	4.813***	4.616***
	(1.161)	(1.150)	(0.614)	(0.616)
$\hat{\lambda}_4$	5.382***	5.073***	6.510***	6.152***
	(0.572)	(0.551)	(0.962)	(0.956)
$\hat{\alpha}$	.	1.211***	.	1.207***
	.	(0.075)	.	(0.121)
Number Obs.	180	180	180	180
Log-Likelihood	-360.32	-356.053	-188.215	-186.681
Test of Lambda Equality *p* -Value	0.015**	0.027**	0.003***	0.030**
Test of Base Rate Fallacy *p* -Value	.	0.005***		0.087*

Standard errors are reported in parentheses and asterisks denote statistical significance: * indicates a statistical difference from zero at least at the 10% level of significance; ** indicates a statistical difference from zero at least at the 5% level of significance; and *** indicates a statistical difference from zero at least at the 1% level of significance. The "Test of Lambda Equality" is a test of the null that all rationality parameters are the same. The "Test of Base Rate Fallacy" is a test of the null that that α = 1.

Table 4: Structural Model Coefficient Estimates

	Interacting Cascade					Basic Cascade		
	Player ($j^{\gamma'}=$) 2^A	Player 2^B	Player 3	Player 4^A	Player 4^B	Player 2	Player 3	Player 4
Player 1^A								
$\Delta(1^A, j^{\gamma'}, \hat{\boldsymbol{\lambda}}, \hat{\alpha})$	0.318***	.	0.298***	0.315***	0.127***	0.215***	0.339***	0.357***
Player 1^B								
$\Delta(1^B, j^{\gamma'}, \hat{\boldsymbol{\lambda}}, \hat{\alpha})$	.	0.318***	0.298***	0.127***	0.315***	.	.	.
Player 2^A								
$\Delta(2^A, j^{\gamma'}, \hat{\boldsymbol{\lambda}}, \hat{\alpha})$	.	.	0.304***	0.321***	0.130***	.	0.361***	0.383***
Player 2^B								
$\Delta(2^B, j^{\gamma'}, \hat{\boldsymbol{\lambda}}, \hat{\alpha})$	.	.	0.304***	0.130***	0.321***	.	.	.
Player 3								
$\Delta(3, j^{\gamma'}, \hat{\boldsymbol{\lambda}}, \hat{\alpha})$	.	.	.	0.529***	0.529***	.	.	0.406***

Each cell gives the structural marginal effect of the row player on the column player for the preferred structural models at their estimated parameters. The structural marginal effects are tested for equality to zero and asterisks denote statistical significance: * indicates a statistical difference from zero at least at the 10% level of significance; ** indicates a statistical difference from zero at least at the 5% level of significance; and *** indicates a statistical difference from zero at least at the 1% level of significance. Basic Cascade results are reported in the group A rows for expositional simplicity.

Table 5: Structural Marginal Effects

be the change in the probability that $j^{\gamma'}$ correctly guesses $\mathcal{B}$ when i^γ correctly guesses $\mathcal{B}$ rather than incorrectly guesses $\mathcal{R}$ when the parameters are $(\boldsymbol{\lambda}, \alpha)$.[17] Let $(\hat{\boldsymbol{\lambda}}, \hat{\alpha})$ be the estimated parameters (from Table 4). We define the **structural marginal effect of player i^γ on player $j^{\gamma'}$** as $\Delta(i^\gamma, j^{\gamma'}, \hat{\boldsymbol{\lambda}}, \hat{\alpha})$, which is the displayed equation evaluated at the estimated parameters.[18] The structural marginal effects of the Basic Cascade preferred structural model are analogously defined. Appendix B provides a formal development of the structural marginal effects.

Table 5 presents the structural marginal effects for the Basic and Interacting Cascades. The table shows, for instance, that $\Delta(1^A, 4^B, \hat{\boldsymbol{\lambda}}, \hat{\alpha}) = 0.127$, i.e., the structural marginal effect of player 1^A on player 4^B is 0.127. Since players with the same index have the same rationality parameter, it is readily established that the structural marginal effects are symmetric across groups, i.e., $\Delta(i^A, j^A, \hat{\boldsymbol{\lambda}}, \hat{\alpha}) = \Delta(i^B, j^B, \hat{\boldsymbol{\lambda}}, \hat{\alpha})$ and $\Delta(i^A, j^B, \hat{\boldsymbol{\lambda}}, \hat{\alpha}) = \Delta(i^B, j^A, \hat{\boldsymbol{\lambda}}, \hat{\alpha})$ for any $i < j$; this is evident in the table. Since a player's $\mathcal{B}$ guess weakly increases her successor's belief that the state is $\mathcal{B}$, it is easily shown that the structural marginal effects are non-negative; this is also evident in the table.

[17] It is readily verified that $\lim_{\boldsymbol{\lambda}\to\boldsymbol{\infty},\alpha\to 1} \Delta(i^\gamma, j^{\gamma'}, \boldsymbol{\lambda}, \alpha) = \Delta(i^\gamma, j^{\gamma'})$, i.e., the structural marginal effects converge to the equilibrium marginal effects as decision-making noise and the base rate fallacy vanish.

[18] By symmetry, it would be equivalent to define the structural marginal effect as $P(g_j^{\gamma'} = \mathcal{R} | g_i^\gamma = \mathcal{R}, c = \mathcal{R}, \hat{\boldsymbol{\lambda}}, \hat{\alpha}) - P(g_j^{\gamma'} = \mathcal{R} | g_i^\gamma = \mathcal{B}, c = \mathcal{R}, \hat{\boldsymbol{\lambda}}, \hat{\alpha})$.

Comparing Tables 2 and 5 shows that the structural marginal effects are in-line with their empirical counterparts when the latter are averaged across states. However, comparing Tables 1 and 5 shows that the structural marginal effects are smaller than their equilibrium counterparts, indicating that information spills over with a lower intensity than predicted by equilibrium theory.[19] Specifically, the structural marginal effects that measure between-group information spillovers – i.e., those identified by Hypothesis 3 – are about four-tenths of their equilibrium values. Additionally, the structural marginal effects that measure within-group information spillovers – i.e., those identified by Hypothesis 1 – range between four-tenths and three-quarters of their equilibrium values. We provide a detailed analysis of these differences at the close of this section, wherein we assess the relative importance of noisy decision-making and the base rate fallacy in driving the departures from equilibrium.

Table 5 provides strong support for Hypotheses 1 to 3, confirming the results of the independence tests reported in Table 3. Specifically, for each structural marginal effect, Table 5 reports the results of a Wald test of the null that the structural marginal effect equals zero against the alternative that it does not equal zero; Appendix B contains the details of the test. The table shows that all structural marginal effects are statistically different from zero at the 1% significance level. Consequently, we have evidence that: (i) information spills over within a group because the structural marginal effect of any player on each of her successors in her group is strictly positive and statistically different from zero, i.e., Hypothesis 1 holds; (ii) the common player aggregates information because the structural marginal effect between each of her predecessors (i.e., players 1^A, 1^B, 2^A, and 2^B) and herself is strictly positive and statistically different from zero, i.e., Hypothesis 2 holds; and (iii) information spills over between groups because the structural marginal effects of 1^A and 2^A on player 4^B are strictly positive and statistically different from zero, as are those of 1^B and 2^B on player 4^A, i.e., Hypothesis 3 holds.

The tests in Table 5, however, provide no evidence for or against Hypothesis 4 since they do not compare structural marginal effects across the Basic and Interacting Cascades. Further, the tests are pairwise in nature. Yet, Hypothesis 1 to 3 are joint in nature. For example, Hypothesis 1 is the structural marginal effect of each player on every one of her successors is strictly positive. We thus wish to test the joint null that at least one of these marginal effects is zero against the alternative that all of them are strictly positive. In this way, rejecting the null provides support for Hypothesis 1.

Structural Hypothesis Tests

To directly assess Hypotheses 1 to 4, we formulate Wald tests of each hypothesis using

[19] Wald tests reject the null that the structural marginal effects equal their equilibrium counterparts at least at the 5% significance level; detailed results are omitted for expositional simplicity.

the preferred structural models of the Basic and Interacting Cascades. Table 6 presents the results. Each row of the table corresponds to a hypothesis and describes (i) the null and alternative (via an intermediate variable T) and (ii) the test's p-value. We overview the tests and their results below; additional details are given in Appendix B.

Hypothesis 1 is that information spills over within groups. Specifically, $\Delta(i^\gamma, j^\gamma, \hat{\boldsymbol{\lambda}}, \hat{\alpha}) > 0$ for all $\gamma \in \{A, B\}$, $i \in \{1, \ldots, 4\}$, and $j \in \{i, \ldots 4\}$ in the Interacting Cascade. Equivalently, $\prod_{\gamma \in \{A,B\}} \prod_{i \in \{1,\ldots,4\}} \prod_{j \in \{i,\ldots,4\}} \Delta(i^\gamma, j^\gamma, \hat{\boldsymbol{\lambda}}, \hat{\alpha}) \neq 0$ since the structural marginal effects are non-negative, which simplifies to

$$\prod_{i \in \{1,\ldots,4\}} \prod_{j \in \{i,\ldots,4\}} \Delta(i^A, j^A, \hat{\boldsymbol{\lambda}}, \hat{\alpha}) \neq 0$$

since the structural marginal effects are symmetric. In the first row of Table 6 the null is that the left-hand-side of the displayed equation equals zero and the alternative that it does not equal zero; the row also reports the analogous test for the Basic Cascade. The p-value is 8.4% for the Basic Cascade and 1% for the Interacting Cascade, thereby establishing that information spills over within groups.

Hypothesis 2 is that the common player aggregates information across groups, i.e., $\Delta(i^\gamma, 3, \hat{\boldsymbol{\lambda}}, \hat{\alpha}) > 0$ for all $\gamma \in \{A, B\}$ and $i \in \{1, 2\}$. Equivalently, $\prod_{\gamma \in \{A,B\}} \prod_{i \in \{1,2\}} \Delta(i^\gamma, 3, \hat{\boldsymbol{\lambda}}, \hat{\alpha}) \neq 0$ since the structural marginal effects are non-negative, which simplifies to

$$\prod_{i \in \{1,2\}} \Delta(i^A, 3, \hat{\boldsymbol{\lambda}}, \hat{\alpha}) \neq 0$$

since the structural marginal effects are symmetric. In the second row of Table 6 the null is that the left-hand-side of the displayed equation equals zero and the alternative that it does not equal zero. The p-value is 1%, thereby establishing that the common player aggregates information.

Hypothesis 3 is that information spills over between groups, i.e.., if $\Delta(i^\gamma, 4^{\neg\gamma}, \hat{\boldsymbol{\lambda}}, \hat{\alpha}) > 0$ for all $i < 3$ and $\gamma \in \{A, B\}$, where $\neg\gamma$ denotes the group that is not γ. Equivalently, $\prod_{\gamma \in \{A,B\}} \prod_{i \in \{1,2\}} \Delta(i^\gamma, 4^{\neg\gamma}, \hat{\boldsymbol{\lambda}}, \hat{\alpha}) \neq 0$ since the structural marginal effects are non-negative, which simplifies to

$$\prod_{i \in \{1,2\}} \Delta(i^A, 4^B, \hat{\boldsymbol{\lambda}}, \hat{\alpha}) \neq 0$$

since the structural marginal effects are symmetric. In the third row of Table 6 the null is that the left-hand-side of the displayed equation equals zero and the alternative is that it does not equal zero. The p-value is 1%, thereby establishing that information spills over between groups.

Test Description	Test Information: T	Null	Alternative	p-Value
Hypothesis 1. Information Spills Over Within Groups	$T = \Delta(1^A, 2^A, \hat{\boldsymbol{\lambda}}, \hat{\alpha})\ \Delta(1^A, 3, \hat{\boldsymbol{\lambda}}, \hat{\alpha}) \times \Delta(1^A, 4^A, \hat{\boldsymbol{\lambda}}, \hat{\alpha})\ \Delta(2^A, 3, \hat{\boldsymbol{\lambda}}, \hat{\alpha}) \times \Delta(2^A, 4^A, \hat{\boldsymbol{\lambda}}, \hat{\alpha})\ \Delta(3, 4^A, \hat{\boldsymbol{\lambda}}, \hat{\alpha})$	T = 0	T ≠ 0	Basic Cascade: 0.084* Interacting Cascade: 0.001***
Hypothesis 2. Common Player Aggregates Information Across Groups	$T = \Delta(1^A, 3, \hat{\boldsymbol{\lambda}}, \hat{\alpha})\ \Delta(2^A, 3, \hat{\boldsymbol{\lambda}}, \hat{\alpha})$	T = 0	T ≠ 0	0.000***
Hypothesis 3. Information Spills Over Between Groups	$T = \Delta(1^A, 4^B, \hat{\boldsymbol{\lambda}}, \hat{\alpha})\ \Delta(2^A, 4^B, \hat{\boldsymbol{\lambda}}, \hat{\alpha})$	T = 0	T ≠ 0	0.000***
Hypothesis 4. Common Player is More Influential in the Interacting Cascade than the Basic Cascade	$T = \Delta_I(3, 4^A, \hat{\boldsymbol{\lambda}}, \hat{\alpha}) - \Delta_B(3, 4, \hat{\boldsymbol{\lambda}}, \hat{\alpha})$	T ≤ 0	T > 0	0.014**

Where necessary, we denote the structural marginal effects for the Basic and Interacting Cascades as $\Delta_B(\cdot)$ and $\Delta_I(\cdot)$ respectively. Asterisks denote statistical significance: * indicates a rejection of the null at least at the 10% level; ** indicates a rejection of the null at least at the 5% level; and *** indicates a rejection of the null at least at the 1% level.

Table 6: Structural Hypothesis Tests

Hypothesis 4 is that the common player is more influential in the Interacting Cascade than the similarly-situated player in the Basic Cascade, i.e.,

$$\min\{\Delta_I(3, 4^A, \hat{\boldsymbol{\lambda}}_I, \hat{\alpha}_I), \Delta_I(3, 4^B, \hat{\boldsymbol{\lambda}}_I, \hat{\alpha}_I)\} > \Delta_B(3, 4, \hat{\boldsymbol{\lambda}}_B, \hat{\alpha}_B),$$

where $\Delta_B(\cdot, \hat{\boldsymbol{\lambda}}_B, \hat{\alpha}_B)$ and $\Delta_I(\cdot, \hat{\boldsymbol{\lambda}}_I, \hat{\alpha}_I)$ are respectively the Basic and Interacting Cascade structural marginal effects and where $(\hat{\boldsymbol{\lambda}}_B, \hat{\alpha}_B)$ and $(\hat{\boldsymbol{\lambda}}_I, \hat{\alpha}_I)$ are respectively the estimated parameters for the Basic and Interacting Cascades. Equivalently,

$$\Delta_I(3, 4^A, \hat{\boldsymbol{\lambda}}_I, \hat{\alpha}_I) - \Delta_B(3, 4, \hat{\boldsymbol{\lambda}}_B, \hat{\alpha}_B) > 0$$

since the structural marginal effects are symmetric. In the fourth row of Table 6 the null is that the left-hand-side of the displayed equation is no greater than zero and the alternative is that it is strictly positive. The p-value is 1.4%, thereby establishing that the common player is more influential in the Interacting Cascade.

Decomposition of the Differences Between the Equilibrium and Structural Marginal Effects

Recall from Tables 2 and 5 that the equilibrium marginal effects are larger than the structural marginal effects. We decompose the difference into two parts, the portion due to noisy decision-making and the portion due to the base rate fallacy.

To elaborate on the decomposition for the Interacting Cascade, for a pair of players i^γ and $j^{\gamma'}$, with $i < j$ and $\gamma, \gamma' \in \{A, B\}$, write the difference between their equilibrium and structural marginal effects in terms of the structural model, i.e.,

$$\Delta(i^\gamma, j^{\gamma'}) - \Delta(i^\gamma, j^{\gamma'}, \hat{\boldsymbol{\lambda}}, \hat{\alpha}) = \Delta(i^\gamma, j^{\gamma'}, \boldsymbol{\infty}, 1) - \Delta(i^\gamma, j^{\gamma'}, \hat{\boldsymbol{\lambda}}, \hat{\alpha}),$$

where $\Delta(i^\gamma, j^{\gamma'}, \boldsymbol{\infty}, \alpha) = \lim_{\boldsymbol{\lambda}\to\boldsymbol{\infty}} \Delta(i^\gamma, j^{\gamma'}, \boldsymbol{\lambda}, \alpha)$ and where the equality follows from the fact that $\Delta(i^\gamma, j^{\gamma'}) = \Delta(i^\gamma, j^{\gamma'}, \boldsymbol{\infty}, 1)$ since the structural model converges to equilibrium as $\boldsymbol{\lambda} \to \boldsymbol{\infty}$. It is thus clear that the difference of the equilibrium and structural marginal effects is determined by the change in parameters from $(\boldsymbol{\infty}, 1)$ to $(\hat{\boldsymbol{\lambda}}, \hat{\alpha})$. Hence, the goal of the decomposition is to determine the change in $\Delta(i^\gamma, j^{\gamma'}, \cdot)$ due to: (i) the shift in the rationality parameters from $\boldsymbol{\infty}$ to $\hat{\boldsymbol{\lambda}}$, while holding all else constant, which reflects the change due to noisy decision-making; and (ii) the shift in the base rate fallacy parameter from 1 to $\hat{\alpha}$, while holding all else constant, which reflects the change due to base rate fallacy. Mathematically, both $\Delta(i^\gamma, j^{\gamma'}, \boldsymbol{\infty}, 1) - \Delta(i^\gamma, j^{\gamma'}, \hat{\boldsymbol{\lambda}}, 1)$ and $\Delta(i^\gamma, j^{\gamma'}, \boldsymbol{\infty}, \hat{\alpha}) - \Delta(i^\gamma, j^{\gamma'}, \hat{\boldsymbol{\lambda}}, \hat{\alpha})$ and describe the change in $\Delta(i^\gamma, j^{\gamma'}, \cdot)$ due to the shift in the rationality parameters, while holding all other parameters constant. Likewise, both $\Delta(i^\gamma, j^{\gamma'}, \hat{\boldsymbol{\lambda}}, 1) - \Delta(i^\gamma, j^{\gamma'}, \hat{\boldsymbol{\lambda}}, \hat{\alpha})$ and

$\Delta(i^\gamma, j^{\gamma'}, \boldsymbol{\infty}, 1) - \Delta(i^\gamma, j^{\gamma'}, \boldsymbol{\infty}, \hat{\alpha})$ describe the change in $\Delta(i^\gamma, j^{\gamma'}, \cdot)$ due to the shift in the base rate fallacy parameter, while holding all other parameters constant. We thus take averages and ascribe

$$\psi_{\text{NDM}} = \frac{1}{2}(\Delta(i^\gamma, j^{\gamma'}, \boldsymbol{\infty}, 1) - \Delta(i^\gamma, j^{\gamma'}, \hat{\boldsymbol{\lambda}}, 1)) + \frac{1}{2}(\Delta(i^\gamma, j^{\gamma'}, \boldsymbol{\infty}, \hat{\alpha}) - \Delta(i^\gamma, j^{\gamma'}, \hat{\boldsymbol{\lambda}}, \hat{\alpha}))$$

to the shift in the rationality parameters and ascribe

$$\psi_{\text{BRF}} = \frac{1}{2}(\Delta(i^\gamma, j^{\gamma'}, \hat{\boldsymbol{\lambda}}, 1) - \Delta(i^\gamma, j^{\gamma'}, \hat{\boldsymbol{\lambda}}, \hat{\alpha})) + \frac{1}{2}(\Delta(i^\gamma, j^{\gamma'}, \boldsymbol{\infty}, 1) - \Delta(i^\gamma, j^{\gamma'}, \boldsymbol{\infty}, \hat{\alpha}))$$

to the shift in the base rate fallacy parameters. A bit of algebra shows that

$$\Delta(i^\gamma, j^{\gamma'}, \boldsymbol{\infty}, 1) - \Delta(i^\gamma, j^{\gamma'}, \hat{\boldsymbol{\lambda}}, \hat{\alpha}) = \psi_{\text{NDM}} + \psi_{\text{BRF}}.$$

The decompositions for the Basic Cascade are analogous.[20]

Table 7 reports the results of the decompositions for the Basic Cascade and for group A players in the Interacting Cascade; the decompositions for the group B players are symmetric. The table shows, for instance, that for players 1^A and 4^B: (i) the difference between their equilibrium and structural marginal effect is 0.219; (ii) that $\psi_{\text{NDM}} = 0.158$, i.e., 0.158 of the difference is attributed to noisy decision-making; and (iii) that $\psi_{\text{BRF}} = 0.06$, i.e., 0.06 of the difference is attributed to the base rate fallacy. Thus, the base rate fallacy accounts for 28% (= 0.06/0.219) of the difference, as is reported in the table; noisy decision-making accounts for the remaining 72%. Similar percentages obtain for the marginal effect of 2^A on 4^B and thus for all marginal effects that measure between-group information spillovers.

For the marginal effects that measure within-group information spillovers, the table shows that the importance of the base rate fallacy varies – e.g., in the Interacting Cascade it accounts for over 100% of the difference for players 1^A and 2^A, but only 47% of the difference for players 2^A and 4^A.[21] In particular, as j increases, the base rate fallacy accounts for a decreasing portion of the difference between the equilibrium and structural marginal effects of player i^γ on j^γ in the Interacting Cascade, where $i < j$ and $\gamma \in \{A, B\}$. This occurs

[20] Our decomposition is tightly linked to the Shapley Value since both ψ_{BRF} and ψ_{NDM} equal the Shapley Values for players in an associated cooperative game; the details are omitted for expositional simplicity. Studies in empirical economics (e.g., Shorrocks [41]) and machine learning (e.g., Cohen, Ruppin, and Dror [16]) increasingly leverage Shapley Value-linked decompositions, like ours.

[21] The base rate fallacy accounts for over 100% of the difference for players 1^A and 2^A precisely because 2^A only observes the guess of a single predecessor. Simply, the increase in the base rate fallacy parameter from 1 to $\hat{\alpha}$ overweights 2^A's signal to such an extent that a change in 1^A's guess has a minimal effect on her belief and thus, when her rationality parameter is large, her guess. Hence, $\Delta(1^A, 2^A, \boldsymbol{\infty}, \hat{\alpha})$ is close to zero, so $\Delta(1^A, 2^A, \boldsymbol{\infty}, 1) - \Delta(1^A, 2^A, \boldsymbol{\infty}, \hat{\alpha})$ and ψ_{BRF} are both large. Similar phenomena do not emerge for other pairs of players because their members have more information from the guesses of their predecessors.

	Interacting Cascade				Basic Cascade		
	Player ($j^{\gamma'}=$) 2^A	Player 3	Player 4^A	Player 4^B	Player 2	Player 3	Player 4
Player 1^A							
$\Delta(1^A, j^{\gamma'})$ - $\Delta(1^A, j^{\gamma'}, \hat{\boldsymbol{\lambda}}, \hat{\alpha})$	0.182	0.110	0.154	0.219	0.285	0.383	0.366
ψ_{NDM}	-0.097	0.044	0.087	0.158	-0.036	0.229	0.209
ψ_{BRF}	0.279	0.066	0.068	0.060	0.322	0.154	0.157
Base Rate Fallacy %	**153%**	**60%**	**44%**	**28%**	**113%**	**40%**	**43%**
Player 2^A							
$\Delta(2^A, j^{\gamma'})$ - $\Delta(2^A, j^{\gamma'}, \hat{\boldsymbol{\lambda}}, \hat{\alpha})$	.	0.103	0.148	0.216	.	0.361	0.339
ψ_{NDM}	.	0.035	0.079	0.155	.	0.200	0.175
ψ_{BRF}	.	0.068	0.070	0.061	.	0.161	0.165
Base Rate Fallacy %	.	**66%**	**47%**	**28%**	.	**45%**	**48%**
Player 3							
$\Delta(3, j^{\gamma'})$ - $\Delta(3, j^{\gamma'}, \hat{\boldsymbol{\lambda}}, \hat{\alpha})$	.	.	0.398	0.398	.	.	0.475
ψ_{NDM}	.	.	0.383	0.383	.	.	0.263
ψ_{BRF}	.	.	0.015	0.015	.	.	0.212
Base Rate Fallacy %	.	.	**4%**	**4%**	.	.	**45%**

For each pair of row and column players, cells give (i) the difference between the pair's equilibrium marginal effects and structural marginal effects, (ii) the amount of the difference due to noisy decision-making, (iii) the amount of the difference due to the base rate fallacy, and (iv) the portion of the difference due to the base rate fallacy (**bolded**). Group B decompositions are identical to the those reported for group A and are thus omitted. Basic Cascade results are reported in the group A rows for expositional simplicity.

Table 7: Decompositions of Differences Between Equilibrium and Structural Marginal Effects

because players who guess later in the queue have more information from the guesses of their predecessors, so increasing the weight on their own signals has less of an impact on how they guess and thus on their marginal effects, all else equal. The table shows that the portion of the difference attributed to the base rate fallacy stabilizes at an average of 44% for $j > 2$ in the Basic Cascade, and shrinks to a comparable amount for $j = 4$ in the Interacting Cascade.

Both noisy decision-making and the base rate fallacy are economically important drivers of the reduction in the intensity of information spillovers relative to equilibrium. For within-group information spillovers, however, noisy decision-making accounts for the majority of the reduction. As to within-group information spillovers, noisy decision-making accounts for the majority of reduction in eight of the thirteen cases reported in Table 7, while the base rate fallacy accounts for the majority of the reduction in the remaining five cases.

5 Conclusion and Discussion

This paper reports the results of an experiment on information spillovers in networks. A key element of the theory of social learning is that information spills over both between and within groups. Our experimental results establish that, while these spillovers occur, the intensity is smaller than predicted by the theory. Specifically, noisy decision-making is the primary driver of the reduction, with non-Bayesian updating due to players' overweighting of their own information playing a secondary role.

We close by discussing the broader implications of our results. The theory of social learning has been applied to study a range of topics, including advertising and marketing (e.g., Joshi and Musalem [34] and Zhang, Liu, and Chen [43]), learning (e.g., Acemoglu, Dahleh, Lobel, and Ozdaglar [1]), pricing (e.g., Chen and Papanastasiou [17, 13], Crapis, Ifrach, Maglaras, and Scarsini [17], and Papanastasiou and Savva [40]), and product design (e.g., Feldman, Papanastasiou, and Segev [26]), to name but a few recent examples. In each of these domains, equilibrium information spillovers have been found to play a central role in the behavior of economic actors. For instance, in pricing, Papanastasiou and Savva show that equilibrium information spillovers from early adopters increase the expected value of a product to successors, enabling a monopolist to significantly raise prices over time. It is not, however, the mere presence of information spillovers that drives these results; rather, it is the fact that equilibrium information spillovers are strong. Our finding that information spills over at a reduced strength relative to equilibrium thus has deep implications for the theory of social learning. For instance, in pricing, it changes the nature of the monopolist's optimal pricing path, as the following numerical example illustrates.

Example. Dynamic Pricing.

Consider an extension of the Interacting Cascade where each player i^γ, with $i \in \{1, 2, 3, 4\}$ and $\gamma \in \{A, B\}$, chooses whether to purchase a monopolist's product at price p_i. Purchasing the product is analogous to guessing $\mathcal{B}$ and not purchasing is analogous to guessing $\mathcal{R}$. The state is product quality. With probability $\frac{1}{2}$ the product is high-quality and purchasing it gives i^γ a payoff of $1 - p_i$, and with probability $\frac{1}{2}$ the product is low-quality and purchasing it gives i^γ a payoff of $-1-p_i$; not purchasing gives i^γ a payoff of 0. A high-quality product is analogous to the $\mathcal{B}$ state and a low-quality product is analogous to the $\mathcal{R}$ state. Signals correctly indicate the state with probability $\frac{2}{3}$. Players sequentially make purchase decisions and observe their predecessors' decisions as described in Section 2. Before Nature moves, the monopolist posts an index-specific price vector (p_1, p_2, p_3, p_4), with $p_i \in \{0, 0.1, 0.2, \ldots, 0.9, 1\}$ for expositional simplicity. The monopolist receives p_i whenever i^γ makes a purchase.

There is a unique perfect Bayesian equilibrium, where players randomize uniformly when indifferent.[22] In the subgame where all prices are zero, equilibrium is identical to that described in Section 2. However, the monopolist always charges a positive price in equilibrium. In equilibrium, a purchase by a player indicates, to her successors, that a high-quality product is more likely. It thus increases the expected value of the product to her successors, which allows the monopolist to charge them higher prices. The result is an optimal pricing path that begins low and increases substantially, specifically $(p_1, p_2, p_3, p_4) = (0.3, 0.5, 0.6, 0.6)$.

The results of our experiments indicate, however, that equilibrium overstates the intensity of information spillovers. To capture this, we employ a logistic quantile response equilibrium model where (i) all players, save for the monopolist, make noisy decisions and overweight their own information and (ii) the monopolist takes players' behavioral characteristics as given when choosing prices. We parameterize the model using the estimates from Table 4 for the preferred Interacting Cascade structural model.

At these parameters, a purchase by a player increases the expected value of the product to her successors by a smaller amount than in equilibrium (due to the reduced intensity of information spillovers). This limits the monopolist's ability to charge these successors higher prices. The result is a lower and flatter optimal pricing path, specifically $(\tilde{p}_1, \tilde{p}_2, \tilde{p}_3, \tilde{p}_4) = (0.3, 0.4, 0.4, 0.5)$. $\triangle$

Our experimental results also establish that between-group information spillovers occur, showing that players who belong to multiple groups serve to aggregate and transmit information between groups. The players who observe their decisions recognize their centrality as evidenced by their larger marginal effects in the Interacting Cascade.

[22] For expositional simplicity, we omit a detailed characterization of both the extension's equilibrium and logistic quantile response equilibrium; the Online Archive contains code that computes these objects.

A Appendix: Equilibrium

This Appendix characterizes the perfect Bayesian equilibrium of the Interacting Cascade and shows how to compute its equilibrium marginal effects. The equilibrium and marginal effects of the Basic Cascade are analogous and thus omitted.

Equilibrium Characterization

We begin with a characterization of the unique perfect Bayesian equilibrium of the Interacting Cascade, where players randomize with equal probability between $\mathcal{B}$ and $\mathcal{R}$ when they are indifferent. (BHW [8] provide an analogous characterization for the Basic Cascade.)

When player i^γ ($\neq 3$), for $\gamma \in \{A, B\}$, moves she observes her group's history and her own signal $(g_1^\gamma, \ldots, g_{i-1}^\gamma, s_i^\gamma)$, which her equilibrium strategy σ_i^γ maps into a distribution over guesses. Specifically, $\sigma_i^\gamma(g_i^\gamma | g_1^\gamma, \ldots, g_{i-1}^\gamma, s_i^\gamma)$ is the equilibrium probability that i^γ makes guess g_i^γ after observing $(g_1^\gamma, \ldots, g_{i-1}^\gamma, s_i^\gamma)$, where

$$\sigma_i^\gamma(\mathcal{B}|g_1^\gamma, \ldots, g_{i-1}^\gamma, s_i^\gamma) = \begin{cases} 0 & \text{if } \mu_i^\gamma(g_1^\gamma, \ldots, g_{i-1}^\gamma, s_i^\gamma) < \frac{1}{2} \\ \frac{1}{2} & \text{if } \mu_i^\gamma(g_1^\gamma, \ldots, g_{i-1}^\gamma, s_i^\gamma) = \frac{1}{2} \\ 1 & \text{if } \mu_i^\gamma(g_1^\gamma, \ldots, g_{i-1}^\gamma, s_i^\gamma) > \frac{1}{2}, \end{cases} \tag{2}$$

$\sigma_i^\gamma(\mathcal{R}|\cdot) = 1 - \sigma_i^\gamma(\mathcal{B}|\cdot)$, and $\mu_i^\gamma(g_1^\gamma, \ldots, g_{i-1}^\gamma, s_i^\gamma)$ is her Bayesian belief about the probability of the $\mathcal{B}$ state. Player 3's equilibrium strategy $\sigma_3(g_3|g_1^A, g_1^B, g_2^A, g_2^B, s_3)$ is defined analogously because she observes $(g_1^A, g_1^B, g_2^A, g_2^B, s_3)$.

The belief of player i^γ (including the common player) is defined recursively,

$$\mu_i^\gamma(g_1^\gamma, \ldots, g_{i-1}^\gamma, s_i^\gamma) = \begin{cases} \frac{\pi(s_i^\gamma|\mathcal{B}) \prod_{j=1}^{i-1} P(g_j^\gamma|g_1^\gamma,\ldots,g_{j-1}^\gamma,\mathcal{B})}{\sum_{c\in\{\mathcal{B},\mathcal{R}\}} \pi(s_i^\gamma|c) \prod_{j=1}^{i-1} P(g_j^\gamma|g_1^\gamma,\ldots,g_{j-1}^\gamma,c)} & \text{if } i \neq 3 \\ \frac{\pi(s_i^\gamma|\mathcal{B}) \prod_{\gamma'\in\{A,B\}} P(g_2^{\gamma'}|g_1^{\gamma'},\mathcal{B})P(g_1^{\gamma'}|\mathcal{B})}{\sum_{c\in\{\mathcal{B},\mathcal{R}\}} \pi(s_i^\gamma|c) \prod_{\gamma'\in\{A,B\}} P(g_2^{\gamma'}|g_1^{\gamma'},c)P(g_1^{\gamma'}|c)} & \text{if } i = 3, \end{cases} \tag{3}$$

with

$$P(g_j^\gamma|g_1^\gamma, \ldots, g_{j-1}^\gamma, c) = \begin{cases} \sum_{s_j^\gamma\in\{\mathcal{B},\mathcal{R}\}} \sigma_j^\gamma(g_j^\gamma|g_1^\gamma, \ldots, g_{j-1}^\gamma, s_j^\gamma)\pi(s_j^\gamma|c) & \text{if } j \neq 3 \\ \sum_{(g_1^{\gamma'}, g_2^{\gamma'}, s_j)\in\{\mathcal{B},\mathcal{R}\}^3} \sigma_3(g_j^\gamma|g_1^\gamma, g_2^\gamma, g_1^{\gamma'}, g_2^{\gamma'}, s_j^\gamma)P(g_2^{\gamma'}|g_1^{\gamma'}, c)P(g_1^{\gamma'}|c)\pi(s_j|c) & \text{if } j = 3. \end{cases} \tag{4}$$

Since players 1^A and 1^B have no predecessors, these recursions are well-defined.

As discussed in the main text, player 1^γ, for $\gamma \in \{A, B\}$, follows her own signal in equilibrium, choosing $\mathcal{B}$ when her signal is $\mathcal{B}$ and $\mathcal{R}$ when her signal is $\mathcal{R}$. Player 2^γ follows

Cascades		3's Signal	Action
Group A	Group B		
g	g	s_3	g
g	none	s_3	g
g	$\neg g$	s_3	s_3
none	none	s_3	s_3

Table 8: Interacting Cascade – Common Player's Equilibrium Behavior

Cascade in γ	3's Guess	4^γ's Signal	4_g's Guess
g	g	s_4	g
g	$\neg g$	$\neg g$	$\neg g$
g	$\neg g$	g	$\frac{1}{2}\mathcal{B} + \frac{1}{2}\mathcal{R}$
none	g	s_4^γ	g

Table 9: Interacting Cascade – Player 4^γ's Equilibrium Behavior

her own signal if it matches player 1^γ's guess; otherwise, she follows $\frac{1}{2}\mathcal{B} + \frac{1}{2}\mathcal{R}$, randomizing uniformly between $\mathcal{B}$ and $\mathcal{R}$.

For a guess $g \in \{\mathcal{B}, \mathcal{R}\}$, we say that there is a **cascade on $\boldsymbol{g}$ in group $\boldsymbol{\gamma}$** if players 1^γ and 2^γ both guess g.[23] Player 3's equilibrium behavior is described by Table 8; we $\neg g$ for a cascade on not g and we write "none" for no cascade (i.e., when 1^γ and 2^γ make different guesses). The table shows that if there are cascades on different guesses in each group (third row) or if there is no cascade in either group (fourth row), then player 3 follows her signal s_3. Otherwise, if there is a cascade on guess g in at least one group (and no opposing cascade in the other group), then player 3 guesses g.

Player 4^γ's equilibrium behavior is described by Table 9. The table shows that player 4^γ only randomizes if there is a cascade on g in her own group and player 3 guesses $\neg g$ (see row three). Player 4^γ then infers that the other group had a cascade on $\neg g$ and player 3's signal was $\neg g$. Thus, 4^γ is indifferent between g and $\neg g$ when her own signal is g. However, when her own signal is $\neg g$, then she guesses $\neg g$. The table further shows (i) that 4^γ guesses g when her predecessors in her group all guess g and (ii) that 4^γ makes the same guess as the common player when there is no cascade in her group.

Marginal Effects

We describe the computation of the equilibrium marginal effects reported in Table 1 of Proposition 1 for the Interacting Cascade.

Given signals $\boldsymbol{s} = (s_1^A, s_1^B, s_2^A, s_2^B, s_3, s_4^A, s_4^B) \in \{\mathcal{B}, \mathcal{R}\}^7$ and the state c, the equilibrium

[23] Fisher and Wooders [27] provide context for this definition.

probability of guesses $\boldsymbol{g} = (g_1^A, g_1^B, g_2^A, g_2^B, g_3, g_4^A, g_4^B) \in \{\mathcal{B}, \mathcal{R}\}^7$ is

$$P(\boldsymbol{g}|\boldsymbol{s}, c) = \sigma_3(g_3|g_1^A, g_1^B, g_2^A, g_2^B, s_3) \times \prod_{\gamma \in \{A,B\}} \prod_{i \in \{1,2,4\}} \sigma_i^\gamma(g_i^\gamma|g_1^\gamma, \ldots, g_{i-1}^\gamma, s_i^\gamma) \tag{5}$$

since players randomize independently at their information sets and do not observe the state. Since the probability of $\boldsymbol{s}$ given c is

$$P(\boldsymbol{s}|c) = \pi(s_3|c) \times \prod_{\gamma \in \{A,B\}} \prod_{i \in \{1,2,4\}} \pi(s_i^\gamma|c), \tag{6}$$

the joint probability of $(\boldsymbol{g}, \boldsymbol{s})$ given c is

$$P(\boldsymbol{g}, \boldsymbol{s}|c) = P(\boldsymbol{g}|\boldsymbol{s}, c)P(\boldsymbol{s}|c). \tag{7}$$

For $\gamma, \gamma' \in \{A, B\}$ and $i < j$, let $P(g_j^{\gamma'}|g_i^\gamma, c)$ give the probability that $j^{\gamma'}$ guesses $g_j^{\gamma'}$ given the state is c and given i^γ guesses g_i^γ, i.e.,

$$P(g_j^{\gamma'}|g_i^\gamma, c) = \frac{\sum_{(\boldsymbol{g}^\dagger, \boldsymbol{s}^\dagger) \in \Theta_1} P(\boldsymbol{g}^\dagger, \boldsymbol{s}^\dagger|c)}{\sum_{(\boldsymbol{g}^\dagger, \boldsymbol{s}^\dagger) \in \Theta_2} P(\boldsymbol{g}^\dagger, \boldsymbol{s}^\dagger|c)}, \tag{8}$$

where

$$\Theta_1 = \{(g_1^{\gamma\dagger}, ..., g_i^{\gamma\dagger}, \ldots, g_j^{\gamma'\dagger}, \ldots, g_4^{\gamma'\dagger}, \boldsymbol{s}^\dagger)|g_i^{\gamma\dagger} = g_i^\gamma \text{ and } g_j^{\gamma'\dagger} = g_j^{\gamma'}\}$$

is the set of joint guesses and signals where i^γ and $j^{\gamma'}$ respectively guess g_i^γ and $g_i^{\gamma'}$ and

$$\Theta_2 = \{(g_1^{\gamma\dagger}, ..., g_i^{\gamma\dagger}, \ldots, g_j^{\gamma'\dagger}, \ldots, g_4^{\gamma'\dagger}, \boldsymbol{s}^\dagger)|g_i^{\gamma\dagger} = g_i^\gamma\}$$

is the set of joint guess and signal where i^γ guesses g_i^γ.

Formally, for $\gamma, \gamma' \in \{A, B\}$ and $i < j$, the **equilibrium marginal effect of player $\boldsymbol{i^\gamma}$ on player $\boldsymbol{j^{\gamma'}}$** is

$$\Delta(i^\gamma, j^{\gamma'}) = P(g_j^{\gamma'} = \mathcal{B}|g_i^\gamma = \mathcal{B}, c = \mathcal{B}) - P(g_j^{\gamma'} = \mathcal{B}|g_i^\gamma = \mathcal{R}, c = \mathcal{B}). \tag{9}$$

In words, the marginal effect is the change in the probability that $j^{\gamma'}$ guesses $\mathcal{B}$, given the $\mathcal{B}$ state, which results from a change in the guess of i^γ from $\mathcal{B}$ to $\mathcal{R}$. This formula is used to compute the equilibrium marginal effects in Proposition 1. The corresponding code is given in the Online Archive, as is the code for the Basic Cascade's equilibrium marginal effects.

B Appendix: Structural Models

This Appendix describes the details of the Interacting Cascade structural model that was introduced in the main text, including its estimation, structural marginal effects, and structural hypothesis testing. The structural models for the Basic Cascade are analogous and thus omitted.

Structural Models

In the baseline Interacting Cascade structural model, the probability player i^γ, for $\gamma \in \{A, B\}$, guesses g_i^γ after observing her group's history and signal $(g_1^\gamma, \dots, g_{i-1}^\gamma, s_i^\gamma)$ for $i \neq 3$ (or $(g_1^A, g_1^B, g_2^A, g_2^B, s_i)$ for $i = 3$) is

$$\tilde{\sigma}_i^\gamma(g_i^\gamma | g_1^\gamma, \dots, g_{i-1}^\gamma, s_i^\gamma) = \frac{\mathbb{I}(g_i^\gamma = \mathcal{B}) + \mathbb{I}(g_i^\gamma = \mathcal{R}) e^{\lambda_i (1 - 2\tilde{\mu}_i^\gamma(g_1^\gamma, \dots, g_{i-1}^\gamma, s_i^\gamma))}}{1 + e^{\lambda_i (1 - 2\tilde{\mu}_i^\gamma(g_1^\gamma, \dots, g_{i-1}^\gamma, s_i^\gamma))}}, \tag{10}$$

where $\lambda_i \geq 0$ is her rationality parameter and $\tilde{\mu}_i^\gamma(g_1^\gamma, \dots, g_{i-1}^\gamma, s_i^\gamma)$ is her Bayesian belief of the probability of the $\mathcal{B}$ state. The belief is defined recursively using equations (3) and (4), with σ_j^γ replaced by $\tilde{\sigma}_j^\gamma$ and so on.

Let $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ denote the vector of rationality parameters. Observe that when $\boldsymbol{\lambda} = \mathbf{0}$, all players guess $\mathcal{B}$ or $\mathcal{R}$ with equal probability. As $\boldsymbol{\lambda} \to \boldsymbol{\infty}$, their guessing behavior converges to equilibrium.

Lemma 2. Convergence.
Let $\boldsymbol{\lambda} = (\lambda, \dots, \lambda)$, then for each player i^γ, with $\gamma \in \{A, B\}$, and each $(g_1^\gamma, \dots, g_{i-1}^\gamma, s_i^\gamma) \in \{\mathcal{B}, \mathcal{R}\}^i$, we have

$$\lim_{\lambda \to \infty} \tilde{\sigma}_i^\gamma(\mathcal{B} | g_1^\gamma, \dots, g_{i-1}^\gamma, s_i^\gamma) = \sigma_i^\gamma(\mathcal{B} | g_1^\gamma, \dots, g_{i-1}^\gamma, s_i^\gamma)$$

and each $(g_1^A, g_1^B, g_2^A, g_2^B, s_i^\gamma) \in \{\mathcal{B}, \mathcal{R}\}^5$ for $i = 3$, we have

$$\lim_{\lambda \to \infty} \tilde{\sigma}_3(\mathcal{B} | g_1^A, g_1^B, g_2^A, g_2^B, s_i^\gamma) = \sigma_3(\mathcal{B} | g_1^A, g_1^B, g_2^A, g_2^B, s_i^\gamma).$$

Proof. This follows from the recursive nature of the game and the fact that

$$\lim_{\lambda \to \infty} \frac{1}{1 + e^{\lambda(1-2\mu)}} = \begin{cases} 0 & \text{if } \mu < \frac{1}{2} \\ \frac{1}{2} & \text{if } \mu = \frac{1}{2} \\ 1 & \text{if } \mu > \frac{1}{2}. \end{cases} \tag{11}$$

To illustrate, consider player 1^A. Since her belief only depends on her own signal, it is the same in equilibrium and in the structural model. Thus, equation (11) implies $\lim_{\lambda_1^A\to\infty}\tilde{\sigma}_1^A = \sigma_1^A$; likewise for player 1^B. Since $\tilde{\mu}_2^A$ is a continuous function of $\tilde{\sigma}_1^A$, it follows that $\lim_{\lambda_1^A\to\infty}\tilde{\mu}_2^A(g_1^A, s_2^A) = \mu_2^A(g_1^A, s_2^A)$ (point-wise) and thus that $\lim_{(\lambda_1^A,\lambda_2^A)\to\infty}\tilde{\sigma}_2^A \to \sigma_2^A$ by equation (11); likewise for player 2^B. Continuing in this manner for players 3, 4^A, and 4^B establishes the lemma. □

In the base rate fallacy extension, a player's belief is modified by adding an exponential weight $\alpha > 0$ to the contribution of her own signal. Specifically, the belief of player i^γ is recursively defined by

$$\tilde{\mu}_i^\gamma(g_1^\gamma, \dots, g_{i-1}^\gamma, s_i^\gamma) = \begin{cases} \frac{\pi(s_i^\gamma|\mathcal{B})^\alpha \prod_{j=1}^{i-1} P(g_j^\gamma|g_1^\gamma,\dots,g_{j-1}^\gamma,\mathcal{B})}{\sum_{c\in\{\mathcal{B},\mathcal{R}\}} \pi(s_i^\gamma|c)^\alpha \prod_{j=1}^{i-1} P(g_j^\gamma|g_1^\gamma,\dots,g_{j-1}^\gamma,c)} & \text{if } i \neq 3 \\ \frac{\pi(s_i^\gamma|\mathcal{B})^\alpha \prod_{\gamma'\in\{A,B\}} P(g_2^{\gamma'}|g_1^{\gamma'},\mathcal{B})P(g_1^{\gamma'}|\mathcal{B})}{\sum_{c\in\{\mathcal{B},\mathcal{R}\}} \pi(s_i^\gamma|c)^\alpha \prod_{\gamma'\in\{A,B\}} P(g_2^{\gamma'}|g_1^{\gamma'},c)P(g_1^{\gamma'}|c)} & \text{if } i = 3 \end{cases}$$

and equation (4) with σ_j^γ replaced by $\tilde{\sigma}_j^\gamma$. In all other respects, the model is unchanged. The base rate fallacy extension reduces to the baseline model when $\alpha = 1$.

As part of the Online Archive, we include computer code that computes players' beliefs, strategies, and so on for each of the Interacting Cascade structural models considered in the paper, as well as their Basic Cascade analogues.

Maximum Likelihood Estimation

Let $\dot{\boldsymbol{\lambda}}$ be the true parameters for the population for the baseline Interacting Cascade structural model. These parameters are estimated from the data via Maximum Likelihood Estimation (MLE) under the assumption that rounds are independent. For any $\boldsymbol{\lambda}$, equations (5) to (7) – with σ_i^γ replaced by $\tilde{\sigma}_i^\gamma$ – imply that the single-observation likelihood of guesses $\boldsymbol{g} = (g_1^A, g_1^B, \dots, g_4^B)$, signals $\boldsymbol{s} = (s_1^A, s_1^B, \dots, s_4^B)$, and state c is $P(\boldsymbol{g}|\boldsymbol{s}, c)P(\boldsymbol{s}|c)\frac{1}{2}$. Since the Interacting Cascade experiment consists of 180 rounds, indexed $k = 1, \dots, 180$, the log-likelihood is

$$\sum_{k=1}^{180} \ln(P(\boldsymbol{g}_k|\boldsymbol{s}_k, c_k)) + \sum_{k=1}^{180} \ln(P(\boldsymbol{s}_k|c_k)) + \sum_{k=1}^{180} \ln(\frac{1}{2}),$$

where $(\boldsymbol{g}_k, \boldsymbol{s}_k, c_k)$ denote the k-th round's guesses, signals, and state. Since the second and third sums do not depend on $\boldsymbol{\lambda}$, we maximize the first sum by choice of $\boldsymbol{\lambda} \geq 0$ via a modified Newton-Raphson algorithm in Stata 17. The result is the MLE estimator $\hat{\boldsymbol{\lambda}}$. The asymptotic variance-covariance matrix $\hat{\boldsymbol{V}}$ is estimated using the inverse of the negative Hessian of the first sum.

Under the weak condition that $\dot{\boldsymbol{\lambda}}$ is contained in some convex-domain $(0, M)^4$, where $M < \infty$, the model is identified and $\hat{\boldsymbol{\lambda}}$ is asymptotically normal with variance-covariance matrix (estimated by) $\hat{\boldsymbol{V}}$; the details are discussed in the Online Archive. For each player

i^γ, the difference between $\tilde{\sigma}_i^\gamma$ and σ_i^γ rapidly converges to zero as $\boldsymbol{\lambda}$ grows large – e.g., $|\tilde{\sigma}_i^\gamma - \sigma_i^\gamma| \leq 3.0 \times 10^{-14}$ at $\boldsymbol{\lambda} = \mathbf{500}$. Thus, taking M to be large – e.g., $M \geq 500$ – results in no loss because the structural model can capture behavior that is (very) close to equilibrium. Further, prior work – e.g., Anderson and Holt [4] – and our own non-structural analysis indicates that individuals do make guesses with error.

The MLE of the base rate fallacy extension is analogous. The Online Archive contains code that implements these estimators, as well as their Basic Cascade analogues.

Structural Marginal Effects

For the baseline Interacting Cascade structural model, $\gamma, \gamma' \in \{A, B\}$, and $i < j$, let $\Delta(i^\gamma, j^{\gamma'}, \boldsymbol{\lambda})$ be given by equations (5) to (9) – with σ_i^γ replaced by $\tilde{\sigma}_i^\gamma$, etc. – evaluated at $\boldsymbol{\lambda}$. It is readily verified that $\Delta(i^\gamma, j^{\gamma'}, \boldsymbol{\lambda})$ is (i) a continuous and differentiable function of $\boldsymbol{\lambda}$ and (ii) is non-negative. The **structural marginal effect of player $\boldsymbol{i^\gamma}$ on player $\boldsymbol{j^{\gamma'}}$** is $\Delta(i^\gamma, j^{\gamma'}, \hat{\boldsymbol{\lambda}})$.

The structural marginal effects for the base rate fallacy extension are computed analogously. The Online Archive contains computer code that compute these structural marginal effects, as well their Basic Cascade analogues.

Structural Hypothesis Testing

The hypothesis tests employed in Tables 5 and 6 are based on the (non-linear) Wald framework for MLE.[24] We describe these tests in the context of the baseline Interacting Cascade structural model, since the tests for the base rate fallacy extension and the Basic Cascade structural models are analogous. The Online Archive contains code that implements these tests.

Table 5, for each pair of players $(i^\gamma, j^{\gamma'})$ in the table, reports the result of a test of the null that $\Delta(i^\gamma, j^{\gamma'}, \dot{\boldsymbol{\lambda}}) = 0$ against the alternative that $\Delta(i^\gamma, j^{\gamma'}, \dot{\boldsymbol{\lambda}}) \neq 0$. The corresponding Wald test statistic is

$$\mathcal{T}(i^\gamma, j^{\gamma'}, \hat{\boldsymbol{\lambda}}) = \frac{(\Delta(i^\gamma, j^{\gamma'}, \hat{\boldsymbol{\lambda}}))^2}{\nabla(\Delta(i^\gamma, j^{\gamma'}, \hat{\boldsymbol{\lambda}}))\, \hat{\boldsymbol{V}}\, \nabla(\Delta(i^\gamma, j^{\gamma'}, \hat{\boldsymbol{\lambda}}))'},$$

where $\nabla(\Delta(i^\gamma, j^{\gamma'}, \hat{\boldsymbol{\lambda}}))$ is the gradient of $\Delta(i^\gamma, j^{\gamma'}, \hat{\boldsymbol{\lambda}})$ evaluated at $\hat{\boldsymbol{\lambda}}$. Per Greene [32], $\mathcal{T}(i^\gamma, j^{\gamma'}, \hat{\boldsymbol{\lambda}})$ converges in distribution to the χ_1^2-distribution, with associated p-value $1 - \chi_1^2(\mathcal{T}(i^\gamma, j^{\gamma'}, \hat{\boldsymbol{\lambda}}))$, where $\chi_1^2(\cdot)$ is the cumulative distribution function for the Chi-Square with one degree of freedom.

Each row of Table 6, save for the last, reports the result of a test of the null that $T(\dot{\boldsymbol{\lambda}}) = 0$ against the alternative that $T(\dot{\boldsymbol{\lambda}}) \neq 0$, where $T(\cdot)$ is given in the row's "T" column. Since

[24] Greene [32] provides a formal development of Wald testing in the context of MLE.

$T(\cdot)$ is continuously differentiable in the parameters, the corresponding Wald test statistic is

$$\mathcal{T}(\hat{\boldsymbol{\lambda}}) = \frac{(T(\hat{\boldsymbol{\lambda}}))^2}{\nabla(T(\hat{\boldsymbol{\lambda}}))\,\hat{\boldsymbol{V}}\,\nabla(T(\hat{\boldsymbol{\lambda}}))'},$$

where $\nabla(T(\hat{\boldsymbol{\lambda}}))$ is the gradient of $T(\hat{\boldsymbol{\lambda}})$ evaluated at $\hat{\boldsymbol{\lambda}}$. Per Greene [32], $\mathcal{T}(\hat{\boldsymbol{\lambda}})$ converges in distribution to a χ^2_1-distribution, with associated p-value $1 - \chi^2_1(\mathcal{T}(\hat{\boldsymbol{\lambda}}))$.

To describe the test in the last row of Table 6, let $\hat{\boldsymbol{\lambda}}_B$, $\hat{\boldsymbol{V}}_B$, and $\dot{\boldsymbol{\lambda}}_B$ ($\hat{\boldsymbol{\lambda}}_I$, $\hat{\boldsymbol{V}}_I$, and $\dot{\boldsymbol{\lambda}}_I$) be the MLE estimates, associated asymptotic variance-covariance matrix, and true parameters for the structural model for the Basic (Interacting) Cascade. For $\gamma, \gamma' \in \{A, B\}$ and $i < j$, let $\Delta_B(i, j, \hat{\boldsymbol{\lambda}}_B)$ ($\Delta_I(i^\gamma, j^{\gamma'}, \hat{\boldsymbol{\lambda}}_I)$) be the structural marginal effects of player i (i^γ) on player j ($j^{\gamma'}$) in the structural model for the Basic (Interacting) Cascade evaluated at $\hat{\boldsymbol{\lambda}}_B$ ($\hat{\boldsymbol{\lambda}}_I$).

The goal of the test in the last row of Table 6 is to examine the null that $T(\dot{\boldsymbol{\lambda}}_B, \dot{\boldsymbol{\lambda}}_I) = \Delta_I(3, 4^A, \dot{\boldsymbol{\lambda}}_I) - \Delta_B(3, 4, \dot{\boldsymbol{\lambda}}_B) \leq 0$ against the alternative that $T(\dot{\boldsymbol{\lambda}}_B, \dot{\boldsymbol{\lambda}}_I) > 0$. We employ the test statistic

$$\mathcal{T}(\hat{\boldsymbol{\lambda}}_B, \hat{\boldsymbol{\lambda}}_I) = \frac{T(\hat{\boldsymbol{\lambda}}_B, \hat{\boldsymbol{\lambda}}_I)}{\sqrt{\nabla(T(\hat{\boldsymbol{\lambda}}_B, \hat{\boldsymbol{\lambda}}_I))\,\hat{\boldsymbol{V}}_J\,\nabla(T(\hat{\boldsymbol{\lambda}}_B, \hat{\boldsymbol{\lambda}}_I))'}},$$

where $\nabla(T(\hat{\boldsymbol{\lambda}}_B, \hat{\boldsymbol{\lambda}}_I))$ is the gradient of $T(\hat{\boldsymbol{\lambda}}_B, \hat{\boldsymbol{\lambda}}_I)$ evaluated at $(\hat{\boldsymbol{\lambda}}_B, \hat{\boldsymbol{\lambda}}_I)$ and where

$$\hat{\boldsymbol{V}}_J = \begin{bmatrix} \hat{\boldsymbol{V}}_B & \mathbf{0} \\ \mathbf{0} & \hat{\boldsymbol{V}}_I \end{bmatrix}.$$

We will show (i) that $\mathcal{T}(\hat{\boldsymbol{\lambda}}_B, \hat{\boldsymbol{\lambda}}_I)$ is asymptotically Normal with mean $T(\dot{\boldsymbol{\lambda}}_B, \dot{\boldsymbol{\lambda}}_I)$ and variance one and (ii) that, for a given significance α, the probability that $\mathcal{T}(\hat{\boldsymbol{\lambda}}_B, \hat{\boldsymbol{\lambda}}_I)$ exceeds $\Phi^{-1}(1-\alpha)$ is bounded-above by α under the null, where $\Phi(\cdot)$ is the standard Normal cumulative distribution function and $\Phi^{-1}(\cdot)$ is its inverse. Hence, the test's p-value is $1 - \Phi(\mathcal{T}(\hat{\boldsymbol{\lambda}}_B, \hat{\boldsymbol{\lambda}}_I))$.

We now establish (i) and (ii). Consider the joint vector $(\hat{\boldsymbol{\lambda}}_B, \hat{\boldsymbol{\lambda}}_I)$. Since both parameter estimates are independent (by independence of the sessions), this vector is asymptotically Normal with mean $(\dot{\boldsymbol{\lambda}}_B, \dot{\boldsymbol{\lambda}}_I)$ and variance-covariance matrix estimated by $\hat{\boldsymbol{V}}_J$. Since $T(\cdot)$ is continuously differentiable, Theorem D.21A in Greene [32] implies

$$\frac{T(\hat{\boldsymbol{\lambda}}_B, \hat{\boldsymbol{\lambda}}_I) - T(\dot{\boldsymbol{\lambda}}_B, \dot{\boldsymbol{\lambda}}_I)}{\sqrt{\nabla(T(\hat{\boldsymbol{\lambda}}_B, \hat{\boldsymbol{\lambda}}_I))\,\hat{\boldsymbol{V}}_J\,\nabla(T(\hat{\boldsymbol{\lambda}}_B, \hat{\boldsymbol{\lambda}}_I))'}}$$

is asymptotically standard Normal. Under the null, either $T(\dot{\boldsymbol{\lambda}}_B, \dot{\boldsymbol{\lambda}}_I) = 0$ or $T(\dot{\boldsymbol{\lambda}}_B, \dot{\boldsymbol{\lambda}}_I) < 0$. If $T(\dot{\boldsymbol{\lambda}}_B, \dot{\boldsymbol{\lambda}}_I) = 0$, then the probability $\mathcal{T}(\hat{\boldsymbol{\lambda}}_B, \hat{\boldsymbol{\lambda}}_I) \geq \Phi^{-1}(1-\alpha)$ is α. If $T(\dot{\boldsymbol{\lambda}}_B, \dot{\boldsymbol{\lambda}}_I) < 0$, then the asymptotic distribution of $\mathcal{T}(\hat{\boldsymbol{\lambda}}_B, \hat{\boldsymbol{\lambda}}_I)$ is left-shifted. Hence, the probability $\mathcal{T}(\hat{\boldsymbol{\lambda}}_B, \hat{\boldsymbol{\lambda}}_I) \geq$

$\Phi^{-1}(1-\alpha)$ is strictly smaller than α.

C Appendix: Instructions

The two pages of instructions for the Interacting Cascade are given in Figures 3a and 3b respectively; whitespace is omitted.[25]

D Online Appendix: Detailed Summary of Experimental Outcomes

Per Appendix A, there are four types of cascades: (i) no cascade in either group; (ii) a cascade in exactly one group; (iii) the same cascade in both groups; and (iv) different cascades in both groups. Table 10 summarizes players' behavior in the data by the type of cascade, their signals, and their guesses; it also compares these to equilibrium. Note that the table aggregates across groups – e.g., players 1^A and 1^B are aggregated into position 1, players 2^A and 2^B are aggregated into position 2, and so on – for simplicity. The table shows, for example, that (players in) position 2 guessed the color of their signal 86.1% of the time in the Interacting Cascade and 89.4% of the time in the Basic Cascade. It also shows that both proportions are close to the equilibrium prediction of 78%.

[25] The figures stats that players have "...20 chances to guess... ." The first five chances consist of an individual decision task and the next 15 chances consist of plays of the experimental game. An analysis of the data from the decision task found that it is not informative about play of the experimental game; it is thus omitted for expositional simplicity.

Instructions

Welcome. This is an experiment in decision-making. In this experiment, you'll play a Guessing Game. At the end of the experiment, you'll be paid a dollar for each correct guess you made. You'll have 20 chances to guess; so, if you're attentive and lucky, you can earn up to $20, in addition to your $10 show-up payment.

The experiment takes place through a computer. It is important that you don't talk or otherwise try to communicate with other participants. If you communicate with another participant, you'll be asked to leave and forfeit your earnings.

If you have any questions, raise your hand and I'll come over and help you.

Timing:

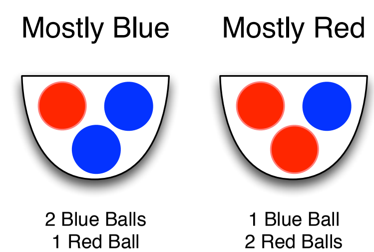
1. Instructions. Read these instructions. Once everyone's read these instructions, a lab assistant will ask you if you have any questions and will distribute an ID code to you.
2. Enter the lab. Please sit at the computer with your position number on it, place your belongings under the desk, enter your ID code into the appropriate field, and click "Login." (Hang onto your ID code since it's how we'll know how much to pay you.)
3. Demo. Please click "Start Demo." The demo will familiarize you with the structure of the Guessing Game and the computer program. Please read through the slides and make sure that you understand what's happening; the information in the demo will help you make the best guesses possible. (You won't be paid for guesses in the demo.)
4. Experiment. Once everyone's completed the demo, the Guessing Game starts.
5. Get Paid. Once you've completed the experiment, please wait quietly. We'll come to you, pay you your earnings, and then you can go.

Please see the reverse for a summary of useful information from the Demo.

(a) Page 1

<u>Useful Information from Demo:</u>

At the start of each experimental round, one of the two containers at right is selected. Both containers have an *equal* chance of being selected.



Your private signal is the ball/token you choose from the selected container.

If the container is Mostly Red, there is a 2/3rd chance the ball will be red and 1/3rd chance it will be blue.

If the container is Mostly Blue, then there is a 1/3rd chance the ball will be red and 2/3rd chance it will be blue.

In addition, every other participant (including you) only observes the history of decisions in their group. That is,

1^A observes no previous decisions, and 1^B observes no previous decisions.

2^A observes the decision of 1^A, while 2^B observes the decision of 1^A.

3 observes the decisions of 1^A, 2^A, 1^B, and 2^B (as 3 is a member of both groups).

4^A observes the decisions of 1^A, 2^A, and 3, while 4^B observes the decisions of 1^B, 2^B, and 3.

(b) Page 2

Figure 3: Interacting Cascade Instructions

Subject Position and Cascade Type	Experiment							Equilibrium						
	Interacting Cascade				Basic Cascade			Interacting Cascade				Basic Cascade		
	Frequency of Event (Count)	Fraction of Row's Guesses Agreeing with Private Signal	Fraction of Row's Guesses Agreeing with Cascade	Fraction of Row's Guesses Agreeing with Player 3	Frequency of Event (Count)	Fraction of Row's Guesses Agreeing with Private Signal	Fraction of Row's Guesses Agreeing with Cascade	Frequency of Event	Fraction of Row's Guesses Agreeing with Private Signal	Fraction of Row's Guesses Agreeing with Cascade	Fraction of Row's Guesses Agreeing with Player 3	Frequency of Event	Fraction of Row's Guesses Agreeing with Private Signal	Fraction of Row's Guesses Agreeing with Cascade
Position 1	1.000 (360)	0.944	.	.	1.000 (180)	0.961	.	1.000	1.000	.	.	1.000	1.000	.
Position 2	1.000 (360)	0.861	.	.	1.000 (180)	0.894	.	1.000	0.780	.	.	1.000	0.780	.
Position 3														
No Cascade in Either Group	0.078 (14)	0.929	.	.	0.367 (66)	0.879	.	0.049	1.000	.	.	0.222	1.000	.
Cascade in Exactly One Group	0.478 (86)	0.616	0.953	.	0.633 (114)	0.772	0.833	0.346	0.570	1.000	.	0.778	0.570	1.000
Same Cascade in Both Groups	0.306 (55)	0.673	0.945	.		.	.	0.358	0.620	1.000	.		.	.
Different Cascades in Both Groups	0.139 (25)	1.000	.	.		.	.	0.247	1.000	.	.		.	.
(Total Number of Guesses)	(180)				(180)			.				.		
Position 4														
No Cascade	0.317 (114)	0.763	.	0.667	0.367 (66)	0.939	.	0.222	0.568	.	1.000	0.222	0.780	.
Cascade	0.683 (246)	0.691	0.898	0.854	0.633 (114)	0.649	0.877	0.778	0.624	0.877	0.965	0.778	0.570	1.000
Cascade and Position 3's Guess Agree	0.586 (211)	0.682	0.938	0.938		.	.	0.654	0.595	1.000	1.000		.	.
Cascade and Position 3's Guess Disagree	0.097 (35)	0.743	0.657	0.343		.	.	0.124	0.778	0.222	0.778		.	.
(Total Number of Guesses)	(360)				(180)			.				.		

Subject behavior is aggregated by position across groups A and B in the interacting cascade for simplicity. Equilibrium proportions are computed via Monte Carlo methods.

Table 10: Detailed Summary of Experimental Outcomes

References

[1] D. Acemoglu, M. Dahleh, I. Lobel, and A. Ozdaglar. Bayesian Learning in Social Networks. *Review of Economic Studies*, 78:1201–1236, 2011.

[2] M. Agranov, G. Lopez-Moctezuma, P. Strack, and O. Tamuz. Social Learning in Groups. *Working Paper, California Institute of Technology*, 2021.

[3] L. Anderson and C. Holt. Information Cascades in the Laboratory. *American Economic Review*, 87(5):847–862, 1997.

[4] L. Anderson and C. Holt. Chapter 39: Information Cascade Experiments. *Handbook of Experimental Economics Results*, 1:335–343, 2008.

[5] M. Angrisani, A. Guarino, P. Jehiel, and T. Kitagawa. Information Redundancy Neglect versus Overconfidence: A Social Learning Experiment. *American Economic Journal: Microeconomics*, Forthcoming, 2021.

[6] A. Banerjee. A Simple Model of Herd Behavior. *Quarterly Journal of Economics*, 107(3):797–817, 1992.

[7] S. Bikhchandani, D. Hirshleifer, O. Tamuz, and I. Welch. Information Cascades and Social Learning. *NBER Working Paper*, 2021.

[8] S. Bikhchandani, D. Hirshleifer, and I. Welch. A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades. *The Journal of Political Economy*, 100(5):992–1026, 1992.

[9] H. Cai, Y. Chen, and H. Fang. Observational Learning: Evidence from a Randomized Natural Field Experiment. *The American Economic Review*, 99(3):864–882, 2009.

[10] B. Celen and K. Hyndman. Social Learning Through Endogenous Information Acquisition: An Experiment. *Management Science*, 58(8), 2012.

[11] B. Celen, S. Kariv, and A. Schotter. An Experimental Test of Advice and Social Learning. *Management Science*, 56(10):1687–1701, 2010.

[12] A. Chandrasekhar, H. Larreguy, and J. Xandri. Testing Models of Social Learning on Networks: Evidence from Two Experiments. *Econometrica*, 88(1):1–32, 2020.

[13] L. Chen and Y. Papanastasiou. Seeding the Herd: Pricing and Welfare Effects of Social Learning Manipulation. *Management Science*, Forthcoming.

[14] Z. Cheng, A. Rai, F. Tian, and S. Xu. Social Learning in Information Technology Investment: The Role of Board Interlocks. *Management Science*, 67(1), 2021.

[15] S. Choi, E. Gallo, and S. Kariv. Chapter 17: Networks in the Laboratory. *The Oxford Handbook of the Economics of Networks*, 2016.

[16] S. Cohen, E. Ruppin, and G. Dror. Feature Selection Based on the Shapley Value. In

IJCAI'05: Proceedings of the 19th International Joint Conference on Artificial Intelligence, pages 665–670. Association for Computing Machinery, 2005.

[17] D. Crapis, B. Ifrach, C. Maglaras, and M. Scarsini. Monopoly Pricing in the Presence of Social Learning. *Management Science*, 63(11):3531–3997, 2017.

[18] R. Cui, D. Zhang, and A. Bassamboo. Learning from Inventory Availability Information: Evidence from Field Experiments on Amazon. *Management Science*, 65(3):955–1453, 2019.

[19] Z. Da and X. Huang. Harnessing the Wisdom of Crowds. *Management Science*, 66(5):1783–2290, 2020.

[20] A. Davis, V. Gaur, and D. Kim. Consumer Learning from Own Experience and Social Information: An Experimental Study. *Management Science*, 67(5):2657–3320, 2021.

[21] R. De Filippis, A. Guarino, P. Jehiel, and T. Kitagawa. Non-Bayesian Updating in a Social Learning Experiment. *Journal of Economic Theory*, Forthcoming, 2021.

[22] J. Duffy, E. Hopkins, T. Kornienko, and M. Ma. Information Choice in a Social Learning Experiment. *Games and Economic Behavior*, 118:295–315, 2019.

[23] P. Evdokimov and U. Garfagnini. Individual vs. Social Learning: An Experiment. *Working Paper, Higher School of Economics*, 2020.

[24] E. Eyster, M. Rabin, and G. Weizsacker. An Experiment on Social Mislearning. *Working Paper, Harvard University*, 2018.

[25] R. Fahr and B. Irlenbusch. Who Follows the Crowd – Groups or Individuals? *Journal of Economic Behavior & Organization*, 80(1):200–209, 2011.

[26] P. Feldman, Y. Papanastasiou, and E. Segev. Social Learning and the Design of New Experience Goods. *Management Science*, 65(4), 2019.

[27] J. Fisher and J. Wooders. Interacting Information Cascades: On the Movement of Conventions Between Groups. *Economic Theory*, 63(1):211–231, 2017.

[28] V. Frey and A. Van De Rijt. Social Influence Undermines the Wisdom of the Crowd in Sequential Decision Making. *Management Science*, Forthcoming, 2021.

[29] K. Gillingham and B. Bollinger. Social Learning and Solar Photovoltaic Adoption. *Management Science*, Forthcoming.

[30] J. Goeree, T. Palfrey, B. Rogers, and R. McKelvey. Self-Correcting Information Cascades. *Review of Economic Studies*, 74(3):733–762, 2007.

[31] B. Golub and E. Sadler. Chapter 19: Learning in Social Networks. *The Oxford Handbook of the Economics of Networks*, 2016.

[32] W. Greene. *Econometric Analysis, 6th Edition.* Pearson, 2008.

[33] V. Grimm and F. Mengel. Experiments on Belief Formation in Networks. *Journal of the European Economic Association*, 18(1):49–82, 2020.

[34] Y. Joshi and A. Musalem. When Consumers Learn, Money Burns: Signaling Quality via Advertising with Observational Learning and Word of Mouth. *Marketing Science*, 40(1), 2021.

[35] Y. Lee, K. Hosanagar, and Y. Tan. Do I Follow My Friends or the Crowd? Information Cascades in Online Movie Ratings. *Management Science*, 61(9):2013–2280, 2015.

[36] I. Lobel and E. Sadler. Preferences, Homophily, and Social Learning. *Operations Research*, 64(3):564–584, 2016.

[37] C. March and A. Ziegelmeyer. Altruistic Observational Learning. *Journal of Economic Theory*, Forthcoming, 2021.

[38] R. McKelvey and T. Palfrey. Quantal Response Equilibria for Extensive Form Games. *Experimental Economics*, 1:9–41, 1998.

[39] A. Mood, F. Graybill, and D. Boes. *Introduction to the Theory of Statistics, 3rd Edition*. McGraw-Hill, 1974.

[40] Y. Papanastasiou and N. Savva. Dynamic Pricing in the Presence of Social Learning and Strategic Consumers. *Management Science*, 63(4), 2017.

[41] A. Shorrocks. Decomposition Procedures for Distributional Analysis: A Unified Framework Based on the Shapley Value. *Journal of Economic Inequality*, 11(99-126), 2013.

[42] J. Zhang and P. Liu. Rational Herding in Microloan Markets. *Management Science*, 58(5), 2012.

[43] J. Zhang, Y. Liu, and Y. Chen. Social Learning in Networks of Friends versus Strangers. *Marketing Science*, 34(4):473–626, 2015.