Aggregate information cascades*

Antonio Guarino, Heike Harmgart and Steffen Huck†

Abstract

We introduce a new model of aggregate information cascades where only one of two possible actions is observable to others. Agents make a binary decision in sequence. The order is random and agents are not aware of their own position in the sequence. When called upon, they are only informed about the total number of others who have chosen the observable action before them. This informational structure arises naturally in many applications. Our most important result is that only one type of cascade arises in equilibrium, the aggregate cascade on the observable action. A cascade on the unobservable action never arises.

1 Introduction

A hiring committee must make a decision on a job candidate who has just been interviewed. The candidate mentions that three other companies have already made him an offer, information that the committee can verify. On the

*We thank Roberto Burguet, Marco Cipriani, Olivier Compte, Giuseppe Della Corte, Douglas Gale, Nicola Pavoni, Giulia Sestieri and participants in the June 2007 IESE Workshop on Industrial Organization and Finance for useful comments. The revision of the paper owes a lot to the very insightul comments and suggestions of an advisory editor and of a referee. We thank Silvia Camussi and Minyu Chen for excellent research assistance. Errors are ours. We gratefully acknowledge financial support from the ESRC, the ERC and the Leverhulme Trust.

†Guarino: Department of Economics and ELSE, University College London (e-mail: aguarino@ucl.ac.uk); Harmgart: European Bank for Reconstruction and Development (EBRD) (e-mail: harmgarh@ebrd.com); Huck: Department of Economics and ELSE, University College London (e-mail: s.huck@ucl.ac.uk).

other hand, the committee can only speculate on how many rival companies have already rejected the candidate's job application.

A manager of a venture capital firm discusses a project with an inventor who needs capital to develop a new product. The inventor has already secured funds from two other venture capital firms, information that the present manager can verify. The manager will also have some private information about the viability of the project but he can only speculate about how often the inventor was turned down by other rival firms who thought that the project was bad.[1]

A restaurant goer must decide whether or not he wants to dine at a particular restaurant he stands in front of. He has some private information on how good the restaurant is, and he is able to peer through the window to see how many others have already decided to dine there. But he can only speculate about how many others stood before the same door and decided to pass.

What these examples have in common is that agents who have to decide between two options have only aggregate information about one of the two options (offering a job, financing a project, dining in a restaurant), simply because the choice of the other option is not observable. In this paper we study the properties of social learning in this type of environment.

This informational environment arises naturally in many social interactions. Like in the case of the restaurant goer or the venture capital firm, in many circumstances, a decision maker can gather some aggregate information (how many firms have already adopted a new technology, invested in a specific project, etc.), but he can rarely observe all the individual decisions. Clearly, if the decision is binary, knowing the number of agents who have made a certain decision also helps to update on the number of agents who have made the opposite decision. But this is not equivalent to knowing it. And, as we will show in this paper, this makes an important difference for social learning.

The social learning literature so far has focussed on situations where, in

[1] This example could also be extended to the market for syndicated loans where several banks jointly offer funds to a borrowing firm. See for example Sufi (2007) for an empirical analysis of the effect of information provision between several lenders and the borrower on the syndicate structure of the contract.

principle, all available actions are observable. The seminal models of informational cascades (Banerjee, 1992, and Bikhchandani *et al.*, 1992), for instance, contemplates a sequence of binary decisions which are all observable. Agent i knows whether each predecessor in the sequence, from agent 1 to agent $i-1$, decided in favor of one option or the other. Several studies since have relaxed these stringent assumptions, some of which we will discuss briefly below. The question of what happens if some actions are not observable at all, however, has not yet been addressed.

In the sequential models of Banerjee (1992) and Bikhchandani *et al.* (1992) informational cascades arise: at a certain point in the sequence, agents rationally neglect their own private information, that is, they choose the same action independently of the information they receive (and follow the decisions of the predecessors). In particular, different types of cascades can arise. If the decision is binary, say, between investing and not, there can be cascades in which, from a certain point onwards, all decision makers decide to invest, as well as cascades in which, from a certain point onwards, all decisions makers decide not to invest.

At a first glance, one could think that this may be the case in our set up, too. If a restaurant goer sees many people in a restaurant, he could disregard his information and just join the crowd; and if he sees the restaurant empty, he could decide to go somewhere else independently of his private signal.

We will show, on the contrary, that only the first cascade is possible. In equilibrium, cascades on the unobservable action cannot arise and a restaurant about which some people have read good reviews will not remain empty for ever. Intuitively, a cascade on the unobservable action cannot be an equilibrium outcome since if all agents chose the unobservable action (not dining at the restaurant) independently of their signal, then the observation that no one has chosen the observable action (dining at the restaurant) would be completely uninformative. Then it would be optimal for an agent with a signal in favor of the observable action to deviate, that is, to follow his signal and choose that action.

After showing that cascades on the unobservable action cannot occur, we will show that the equilibrium in our model predicts a very simple behavior, with agents following their private signal until a threshold in the aggregate number of the observable action has been reached: after that threshold a cascade occurs, with all agents choosing the observable action independently of their signal. We will illustrate the properties of these cascades, and compare the welfare in our economy with that in the seminal model of Bikhchandani *et al.* (1992). We will see that under some conditions welfare can even be higher in our model, despite agents having access to less information.

Our work contributes to the study of social learning when there is imperfect observability of other agents' actions. The paper most closely related to ours is probably the recent work by Herrera and Hörner (2009). This paper shares the same motivation with ours, in that it focuses on social learning when only one action is observable. While the motivation is the same, the framework is very different. In their continuous time model, agents arrive randomly over time, and only those who invest are observed. Therefore, the lack of observed investments may reflect either the choice of predecessors not to invest or the lack of investment opportunities. Time itself is informative. In terms of results, exactly as it happens in the standard model (Smith and Sørensen, 2000), in their model cascades can occur on both actions if and only if beliefs are bounded while, with unbounded beliefs, learning is asymptotically complete.[2] The probability of an incorrect cascade is higher or lower than in the standard model depending on some properties of the signal distributions.

Other papers in the social learning literature study what happens when we remove the strong assumption that agents can observe the entire history of individual decisions.[3] Çelen and Kariv (2004) extend the standard model of sequential social learning by allowing each agent to observe the decision of his immediate predecessor only. The prediction of these authors is that behavior does not settle on a single action. Long periods of herding can be observed, but switches to the other action occur. As time passes, the periods of herding become longer and longer, and the switches increasingly rare. Larson (2008) analyzes a situation in which agents observe a weighted average of past actions before making a choice in a continuous action space. In

[2] Beliefs are (un)bounded when the support of the loglikelihood ratio—the logarithm of the ratio between the probabilities of a signal conditional on one state relative to the other—is (un)bounded.

[3] For comprehensive surveys of the literature see, among others, Gale (1996), Hirshleifer and Theo (2003), Chamley (2004) and Vives (2007).

contrast to our work, the focus is not on whether a cascade occurs or not, but on the speed of learning (since the continuous action space guarantees that complete learning eventually occurs). An interesting observation of this study is that the speed of learning depends on how effectively the noise coming from early actions is purged. Collander and Hörner (2009) present a model in which an agent observes only the total number of choices of each type (in a binary action setting), rather than the full sequence of actions. They characterize conditions under which later agents optimally imitate the minority, rather than the majority action.

Banerjee and Fudenberg (2004), Smith and Sørensen (2008) and Acemoglu *et al.* (2010) study social learning when agents can only observe a sample of predecessors' actions. Banerjee and Fudenberg (2004) present a model in which, at every time, a continuum of agents choose a binary action after observing a sample of previous decisions (and, possibly, of signals on the outcomes). This can be interpreted as a model of word of mouth communication in large populations. The authors find sufficient conditions (on the sampling rule, etc.) for herding to arise, and conditions for all agents to settle on the correct choice. Smith and Sørensen (2008) study a sequential decision model in which agents can only observe unordered random samples from predecessors' actions (e.g., because of word of mouth communication). They characterize different conditions on the sampling procedure and on the beliefs to have complete or incomplete learning. When the past is not over-sampled, that is, not affect for ever by any one individual, and when beliefs are unbounded, complete learning eventually obtains. Acemoglu *et al.* (2010) analyze a situation in which agents observes the past actions of a stochastically-generated neighborhood of individuals. Differently from Smith and Sørensen (2008), who assume that "samples are unordered" (individuals do not know the identity of the agents they have observed), in Acemoglu *et al.* (2010) individuals know the identity of the agents in their realized neighborhood. In this sense, their work studies social learning in social networks. In this set up, when beliefs are unbounded, there is asymptotic learning (defined as convergence of the actions to the correct one) as long as there is some minimal amount of "expansion in observations". For many common deterministic and stochastic networks, bounded private beliefs are, instead, incompatible with asymptotic learning, as in the standard model. Nevertheless, the authors find conditions under which asymptotic learning obtains even with bounded private beliefs for a large class of stochastic network topologies.

Finally the issue of imperfect observability is also discussed in recent papers by Eyster and Rabin (2008) and Guarino and Jehiel (2009) in contexts in which agents are not fully rational. The imperfect observability can actually alleviate some biases that bounded rationality produces in a classical model of learning with continuous action space similar to that of Lee (1992).

The remainder of the paper is organized as follows. In Section 2 we introduce the formal model. We present its equilibrium analysis in Section 3. Section 4 discusses when informational cascades arise in the sequence of decisions. Section 5 studies the welfare properties of our equilibrium. Section 6 illustrates extensions of our model. Section 7 concludes with a discussion. The Appendix contains some of the proofs.

2 The Model

In our economy there are n agents who have to decide in sequence whether or not to take up a certain option. For convenience, we shall refer to this choice as the decision about whether or not to invest. Time is discrete and indexed by $i = 1, 2, ..., n$. Each agent makes his choice only once in the sequence. Agents are numbered according to their positions: agent i chooses at time i only. An agent's action space is given by $\{0, 1\}$, and his action is denoted by $I_i \in \{0, 1\}$ (where 1 is interpreted as investment). An agent's payoff π_i depends on his choice and on the true state of the world $\omega \in \{0, 1\}$. The prior probability of $\omega = 1$ is $r \geq \frac{1}{2}$.[4] If $\omega = 1$, an agent receives a payoff of 1 if he chooses to invest, and a payoff of 0 otherwise; vice versa if $\omega = 0$, that is,

$$\pi_i = \omega I_i + (1 - \omega)(1 - I_i).$$

The sequence in which agents make their choices is randomly determined. All sequences are equally likely. The agents, however, are not informed about which sequence has been chosen, and do not know their own position in the sequence. Aggregate investments are the only observable variable: when called upon, agent i is informed about the total number of agents who have decided to invest before him. While the aggregate number of investments is observable, each individual decision to invest or not is not publicly known, nor is the

[4] Later we will also discuss the case $r < \frac{1}{2}$.

total number of non-investments. We denote the total number of agents who have invested before agent i by T_i: agent i is informed about $T_i = \sum_{j=1}^{i-1} I_j$.[5] In addition to observing T_i, each agent i receives a symmetric binary signal σ_i distributed as follows:

$$\Pr(\sigma_i = 1 \mid \omega = 1) = \Pr(\sigma_i = 0 \mid \omega = 0) \equiv q.$$

Note that, conditional on the state of the world, the private signals are i.i.d.. We shall refer to $\omega = 1$ as the "good state" and to $\omega = 0$ as the "bad state." A signal pointing in the direction of the good state ($\sigma_i = 1$) shall be called a "good signal" and a signal pointing in the opposite direction ($\sigma_i = 0$) a "bad signal." We assume that $1 > q > r$. This condition ensures that, in the one-agent case, an agent would not invest upon a bad signal while he would invest upon a good signal, which renders the problem interesting. The condition also implies that $q > \frac{1}{2}$, that is, the signal respects the monotone likelihood ratio property. Finally, the signal is not perfectly informative, which makes social learning possible and relevant.

An agent's information set is, therefore, $\{T_i, \sigma_i\}$. An agent's mixed strategy induces a probability with which the agent invests for each $\{T_i, \sigma_i\}$. We denote such a probability by $\mathcal{I}_i(T_i, \sigma_i)$.[6]

To conclude the description of our model, it is useful to introduce the notion of an aggregate information cascade. The definition is similar to the standard definition of information cascade, with the characteristic that histories are summarized by the aggregate statistic T_i.

Definition 1 *An aggregate information cascade occurs when there is a critical value of T_i after which all agents choose the same action independently of their signals. In particular:*

In an aggregate up cascade (AUC) there is a critical value T^{UP} such that if $T_k = T^{UP}$ all agents from k onwards choose to invest regardless of their signals. Consequently, there is some k such that $T_{k+j} = T_k + j$ for all $j = 1, ..., n - k$.

[5] As we have already mentioned, and will write formally below, an agent does not know his index i. The only thing agent i, the ith agent in the sequence, knows about his position is that he is not among the first T_i agents.

[6] We use the same symbol to indicate both a random variable and its realization, unless it creates ambiguities.

In an aggregate down cascade (ADC) there is a critical value T^{DOWN} such that if $T_k = T^{DOWN}$ all agents from k onwards choose not to invest regardless of their signals. Consequently, in an ADC there is some k such that $T_{k+j} = T_k$ for all $j = 1, ..., n-k$.

We are now ready to start analyzing the equilibrium decisions in our economy.

3 Equilibrium Analysis

The ultimate goal of our analysis is to understand the social learning process that occurs in our economy. Each agent can learn about the true state of the world from the aggregate information that he receives about his predecessors' choices. This can lead to better decisions. On the other hand, it may be that in our economy, as in the canonical models of social learning of Banerjee (1992) and Bikhchandani *et al.* (1992), there is room for information cascades. In such a case, the process of information aggregation will not be efficient. We will show that, indeed, AUCs, that is, cascades of investments, are possible even in our set up, just as in the canonical models. In contrast, ADCs, that is, cascades of non-investments, never occur in equilibrium, unlike in the canonical models.

We shall restrict the entire analysis to symmetric Perfect Bayesian equilibria (PBEs). For convenience, we shall drop the qualification and simply speak of an "equilibrium."[7] Since the equilibrium is symmetric, we will drop the subscript and simply write $\mathcal{I}(T_i, \sigma_i)$.

To start our analysis, it is convenient to focus first on the case of $T_i = 0$, in which an agent observes that no one has invested before him. At first glance, the decision problem in such a situation appears to be fairly complicated. If the agent knew that $T_i = 0$ simply because he is the first decision maker, then he should certainly follow his private signal, since that is the only information available. If, instead, he knew that he is not the first decision maker, then he could decide not to invest independently of the signal, as other agents have already chosen the non-investment option. Intuitively, one might think of $T_i = 0$ as bad information if there are many agents. Suppose that n is

[7] Our economy is represented by a symmetric game and there is nothing in the environment that could help agents to coordinate on an asymmetric outcome. Therefore, the restriction to symmetric equilibria is natural.

very large and you observe that nobody has invested before you. But at the same time your own private signal is good. Would you trust your own signal? Of course, the answer to this question would depend on the other agents' strategies. While the problem is made hard due to the fact that the agent does not know his position in the sequence, it is made easier due to the fact that the only thing that matters about other agents' strategies is what these specify for the very same case of $T_i = 0$.

We now prove that in equilibrium, after observing $T_i = 0$, agents do not play independently of their signal or against it.

Lemma 1 *There exists no equilibrium in which an agent, after observing $T_i = 0$, always invests (i.e., for both signal realizations) or never invests, or plays against his signal.*

Proof We prove the lemma by contradiction. Suppose that for $T_i = 0$ agents choose either to invest always or never (independently of their private signals). Consider the latter possibility first, i.e., consider a pure-strategy equilibrium with $\mathcal{I}(0,0) = \mathcal{I}(0,1) = 0$. Then, along the equilibrium path, nobody ever invests and, for any agent $i = 1, ..., n$, $T_i = 0$. Hence, $T_i = 0$ does not reveal any information on the true state of the world and agent i is better off by following his informative signal σ_i. Now, consider the case of investment after $T_i = 0$, i.e., an equilibrium with $\mathcal{I}(0,0) = \mathcal{I}(0,1) = 1$. In this case, along the equilibrium path, only the first agent in the sequence observes that nobody else has invested before, that is, $T_i = 0$ if and only if $i = 1$. Hence, after observing $T_i = 0$ agent i knows that he is the first agent in the sequence and follows his signal. Finally, suppose that for $T_i = 0$ agents choose to play against their private information, i.e., consider an equilibrium with $\mathcal{I}(0, \sigma_i) = 1 - \sigma_i$. Then, along the equilibrium path, after observing $T_i = 0$, agent i knows that he is either the first in the sequence or all other agents before him have received good signals. In both cases, if the agent receives a good signal, he follows it.■

According to the previous lemma, the only remaining possibilities are that an agent observing $T_i = 0$ either follows his own signal, that is, $\mathcal{I}(0, \sigma_i) = \sigma_i$, or mixes (for at least one signal realization). We will show that, in equilibrium, agents do follow their signal. To this aim, we first state a lemma that trivially follows from Bayesian updating.

Lemma 2 *In equilibrium,* $\mathcal{I}(T_i, 1) \geq \mathcal{I}(T_i, 0)$ *for all* T_i*. In particular, if* $0 < \mathcal{I}(T_i, 0) < 1$ *then* $\mathcal{I}(T_i, 1) = 1$*, and if* $0 < \mathcal{I}(T_i, 1) < 1$ *then* $\mathcal{I}(T_i, 0) = 0$*.*

Proof In equilibrium, each agent will infer the same information from observing a particular value of T_i. Whatever the posterior induced by just observing T_i, it follows immediately from Bayes's rule that an agent who has an additional good signal cannot be more pessimistic than an agent with a bad signal. The first part of the lemma results from this consideration and from expected payoff maximization. The second part follows from the same argument and the additional observation that mixing requires the agent being indifferent between the two actions.■

We are now ready to prove the following result:

Lemma 3 *In equilibrium, after observing* $T_i = 0$*, an agent follows his own signal, that is,* $\mathcal{I}(0, \sigma_i) = \sigma_i$*.*

Proof We first prove that it is indeed optimal for an agent i to follow his own good signal after $T_i = 0$ provided that everybody else does the same. Assuming such behavior of others, an agent i who observes $T_i = 0$ and $\sigma_i = 1$ attaches to the good state a posterior of

$$\Pr(\omega = 1 \mid T_i = 0, \sigma_i = 1) = \frac{rq\sum_{j=1}^{n}(1-q)^{j-1}}{rq\sum_{j=1}^{n}(1-q)^{j-1} + (1-r)(1-q)\sum_{j=1}^{n} q^{j-1}}.$$

He will follow his good signal if this posterior is greater than 1/2, that is, if

$$rq\sum_{j=1}^{n}(1-q)^{j-1} > (1-r)(1-q)\sum_{j=1}^{n} q^{j-1}.$$

Solving for the sums and rearranging the terms, we obtain the condition $r > \frac{1-q^n}{2-(1-q)^n-q^n}$, which is always satisfied for $r \geq \frac{1}{2}$. Now we show that an agent i who assumes that the others play according to their signals and observes $T_i = 0$ and $\sigma_i = 0$ does not invest, that is, we need that

$$\Pr(\omega = 1 \mid T_i = 0, \sigma_i = 0) =$$
$$\frac{r(1-q)\sum_{j=1}^{n}(1-q)^{j-1}}{r(1-q)\sum_{j=1}^{n}(1-q)^{j-1} + (1-r)q\sum_{j=1}^{n} q^{j-1}} < \frac{1}{2},$$

or

$$r(1-q)\sum_{j=1}^{n}(1-q)^{j-1} < (1-r)q\sum_{j=1}^{n} q^{j-1},$$

which can be written as

$$\frac{r}{(1-r)} < \frac{q^2}{(1-q)^2}\frac{1-q^n}{1-(1-q)^n}.$$

Since $r < q$ we also have $\frac{r}{1-r} < \frac{q}{1-q}$. Hence, the above inequality holds if $\frac{q}{1-q}\frac{1-q^n}{1-(1-q)^n} > 1$. This can be rewritten as $2q > 1 + q^{n+1} - (1-q)^{n+1}$ which is true for $q > 1/2$.

Finally, to complete the proof, we have to show that a mixed strategy equilibrium does not exist. By the previous lemma we know that mixing cannot occur for both signals. We now rule out the case in which $0 < \mathcal{I}(0,0) < 1$ and $\mathcal{I}(0,1) = 1$. For an agent to be indifferent between investing and not investing after observing $T_i = 0$ and $\sigma_i = 0$ it must be that $\Pr(\omega = 1 \mid T_i = 0, \sigma_i = 0) = 1/2$. Using Bayes's rule, this can be re-written as

$$r \Pr(T_i = 0, \sigma_i = 0 \mid \omega = 1) = (1-r)\Pr(T_i = 0, \sigma_i = 0 \mid \omega = 0),$$

or

$$r\sum_{j=1}^{n}(1-q)^j(1-p)^{j-1} = (1-r)\sum_{j=1}^{n} q^j(1-p)^{j-1},$$

where p denotes the probability with which all other agents who see $T_i = 0$ and $\sigma_i = 0$ invest. Rewriting this as

$$\sum_{j=1}^{n}\left[\left(r(1-q)^j - (1-r)q^j\right)(1-p)^{j-1}\right] = 0$$

makes it obvious that there is no $p > 0$ that solves the equation: since $q > r$, the left-hand side is strictly negative for any positive p. To conclude, we rule out the case of mixing after observing the good signal. Agent i's indifference between investing and not investing after observing $T_i = 0$ and $\sigma_i = 1$ requires $\Pr(\omega = 1 \mid T_i = 0, \sigma_i = 1) = 1/2$. If all other agents $j \neq i$ behave according to $I(0,0) = 0$ and $I(0,1) = p$, after applying Bayes's rule and some algebraic manipulation, this equality becomes

$$rq \sum_{j=1}^{n} (1 - pq)^{j-1} = (1 - r)(1 - q) \sum_{j=1}^{n} (1 - p(1 - q))^{j-1}.$$

For $r \geq \frac{1}{2}$ the left-hand side of is strictly greater than the right-hand side for any value of p, which ends the proof.■

Our analysis essentially shows that, when facing a situation with no previous investments, an agent follows his signal.[8] This clearly indicates that an ADC in which all agents choose not to invest, never occurs in equilibrium. In other words, to go back to one of our examples, a restaurant will not stay empty forever only because it is empty when it opens.

While this puts already a lot of structure on the equilibrium solution of our game, we still have to investigate what happens for different values of the aggregate investment T_i. Before we start this analysis, it is worth mentioning that, as we will see, in some cases there are multiple equilibria in our model: an equilibrium in which a cascade starts after T investments can coexist with one without a cascade after T. In these cases, agents who coordinate on the AUC equilibrium at T could revert to the other equilibrium at $T+1$. We rule out this trivial issue, by making the following assumption:

Assumption Once agents have coordinated on an AUC at T^{UP}, they will remain coordinated on that cascade for any $T > T^{UP}$.

We now start our analysis by establishing an intuitive monotonicity result: a higher value of T_i is always good news, that is, when an agent observes a higher number of investments made before him, he cannot be less willing to invest himself. Once this monotonicity lemma is established, we will be able to show two fundamental results about aggregate cascades.

[8] It is worth pointing out that in the proof of the previous lemma we use the condition $r \geq \frac{1}{2}$. For $r < \frac{1}{2}$, the same results obtain under some conditions on the parameters r and q. When these conditions are not satisfied, there can be a mixed strategy equilibrium in which an agent does not invest upon receiving a bad signal and mixes upon receiving a good signal. We refer the reader to the working paper version of the paper for all the details. Apart from this aspect, the entire analysis for $r < \frac{1}{2}$ is identical to that presented in the main text.

Lemma 4 *In equilibrium, if $T_i' < T_i''$ then $\mathcal{I}(T_i', \sigma_i) \leq \mathcal{I}(T_i'', \sigma_i)$ for both $\sigma_i = 0$ and $\sigma_i = 1$. If $0 < \mathcal{I}(T_i, \sigma_i) < 1$, then $\mathcal{I}(T_i - 1, \sigma_i) = 0$ and $\mathcal{I}(T_i + 1, \sigma_i) = 1$.*

Proof See the Appendix.■

While this lemma seems very intuitive (how could a fuller restaurant be worse news than an emptier one?) it is actually not trivial to prove it. At the core of the proof there is, however, a very simple logic. Essentially, it is the earlier monotonicity result (Lemma 2) that drives this one. Agents with good signals are more likely to invest than agents with bad signals. Good signals are more likely to be generated in the good state than in the bad state. Hence, T_i grows, on average, "faster" in the good state than in the bad state. Therefore, the higher T_i the more confident can the agent be of being in the good state.

Equipped with these lemmata, we are now ready to state a proposition that characterizes which form of cascades will or will not arise. We will show that ADCs never occur in equilibrium, while AUCs are always part of an equilibrium.

Proposition 1 *(i) An equilibrium in pure strategies exists. In any such equilibrium, an agent with a good signal invests with probability independently of T_i (i.e., $\mathcal{I}(T_i, 1) = 1$ for all T_i); an agent with a bad signal does not invest up to a threshold T^{UP}, and invests with probability one after the threshold is reached (i.e., $\mathcal{I}(T_i, 0) = 0$ for all $T_i < T^{UP}$, and $\mathcal{I}(T_i, 0) = 1$ for all $T_i \geq T^{UP}$). For the threshold T^{UP} that triggers an AUC it holds that $0 < T^{UP} \leq \frac{n}{2} + 1$.*

(ii) There can exist an equilibrium in which, for a T^{MIX} ($0 < T^{MIX} < \frac{n}{2} + 1$), an agent with a bad signal mixes (i.e., $0 < \mathcal{I}(T^{MIX}, 0) < 1$). In this equilibrium, for $T_i < T^{MIX}$, an agent follows his bad signal and does not invest and for all $T_i > T^{MIX}$, he invests with probability one. An agent with a good signal invests with probability one for all T_i. In this equilibrium an AUC occurs for $T^{UP} = T^{MIX} + 1$. If an equilibrium with such a T^{MIX} exists, then there exist also an equilibrium in which $\mathcal{I}(T^{MIX}, 0) = 0$ and one in which $\mathcal{I}(T^{MIX}, 0) = 1$.[9]

[9] In social learning models of sequential decision making, typically the equilibrium solution requires a tie-breaking rule for the cases of indifference, and there is no need to invoke mixed strategies. In our case, mixed strategies are needed since the optimal decision of an agent observing T_i depends on what other agents do when facing the same situation.

Proof We will first prove part (i) of the proposition in four steps.

(1) The part of the proposition concerning agents' behavior upon observing good signals follows immediately from Lemma 3 and Lemma 4. According to Lemma 3, $\mathcal{I}(0,1) = 1$, and according to Lemma 4, $\mathcal{I}(T_i,1) \geq \mathcal{I}(0,1) = 1$.

(2) In regard to agents' behavior upon observing bad signals, note that from Lemma 3, $\mathcal{I}(0,0) = 0$, and from Lemma 4 $\mathcal{I}(T_i,0) \geq \mathcal{I}(T_i-1,0)$.

(3) To show that $T^{UP} \leq \frac{n}{2}+1$, consider an agent i who observes $T_i > \frac{n}{2}$ and suppose he knew that he were the last agent in the sequence. This agent knows that there were T_i good signals and $n - T_i - 1$ bad signals. Hence, even if this agent's own signal is bad, he knows that there were altogether more good signals than bad signals and he will decide to invest. Now, agent i cannot be sure that he is the last agent. But if he is not the last agent, then there were fewer bad signals and still T_i good signals, which would be even better news. Hence, an agent who observes $T_i > \frac{n}{2}$ will always invest and an AUC occurs, a contradiction.

(4) To prove the existence, in the Appendix, we show that the belief of an agent receiving a bad signal and observing any $T_i = t$ is higher if all other agents invest after observing the same t than if they do not invest. If in both cases the belief of the agent is lower than $\frac{1}{2}$, then $\mathcal{I}(T_i,0) = 0$ is part of the equilibrium. If in both cases the belief of the agent is higher than $\frac{1}{2}$, then $\mathcal{I}(T_i,0) = 1$ is part of the equilibrium. If one belief is lower and the other is higher, then both pure strategy equilibria with $\mathcal{I}(T_i,0) = 0$ and $\mathcal{I}(T_i,0) = 1$ exist.

As for part *(ii)*, we prove the possibility of such a (partially) mixed strategy equilibrium by example. This is contained in the next section. The same argument as in point (4) (and developed in the Appendix) proves that when such an equilibrium exists, there exist also two pure strategy equilibria. Lemma 3 implies that T^{MIX} must be greater than 0. The same argument as in point (2) above establishes that T^{MIX} must be lower than $\frac{n}{2}+1$. Lemma 4 implies that, along the equilibrium path, mixing can only occur for one value of the aggregate investments.■

The proposition clearly implies that there are no cascades on the unobservable action, since agents with good signals always invest. Incidentally, we note that such a result just comes from an equilibrium argument. One could imagine that, when facing a "low" value of T_i, in order to make his decision, agent i should consider all possible sequences and attach a probability to the event that he is the first in the sequence, or the second, etc. After all, a low number of investments may merely come from the fact that only few agents had the opportunity to invest so far, in which case the low value of T_i should be considered good news. Or it could arise from many agents having had the option of investing but only few using it, in which case the low T_i should be viewed as bad news. This inference process could be quite complicated. Our analysis solves the problems by just invoking some equilibrium arguments.

The proposition also shows that AUCs do arise—and are, in fact, part of any equilibrium. The value $\frac{n}{2}+1$ is, obviously, just an upper bound for the critical value T^{UP} that triggers an AUC. Depending on the parameters' values, AUCs may well be triggered by a lower number of investments. But AUCs are indeed part of all equilibria. Of course, this does not imply that AUCs will necessarily occur in a population of finite size n, since there is always the possibility of sufficiently many bad signals occurring such that T^{UP} is not reached. In the next section we will show how the threshold T^{UP} varies with n and the other parameters of the model. We will also show that in a large population (for n tending to infinity) this threshold is reached almost surely, that is, an AUC occurs with probability one.

Before we move to this analysis, it is worth noticing that, despite the AUC consists in agents investing independently of the signal, agents keep learning even during a cascade. In particular, during an AUC the belief on the good state of the world is increasing in the number of investments.

Proposition 2 *During an AUC, agents revise their beliefs upwards, that is,* $\Pr(\omega = 1 \mid T_i = T^{UP} + j + 1, \sigma_i) > \Pr(\omega = 1 \mid T_i = T^{UP} + j, \sigma_i)$ *for* $j = 0, 1, 2, ...n - T^{UP} - 2$*, and for both* $\sigma_i = 0$ *and* $\sigma_i = 1$*.*

Proof See the Appendix.■

To understand the proposition intuitively, consider the case in which $T_i = n - 1$. In this case agent i learns that he is the last agent in the sequence and that all predecessors have invested. In particular, this implies that the first T^{UP} agents invested, which gives the highest posterior on $\omega = 1$. If, instead, the agent observes $T_i = n - 2$, with some probability he is the last in the sequence and one of the first T^{UP} agents did not invest, which means he received a bad signal. This, of course, implies a lower posterior on $\omega = 1$. Similarly for lower levels of T_i. Put it differently, a T_i higher than T^{UP} reveals the information that other agents have already observed T^{UP}, which is good news, since it increases the probability that the AUC started earlier in the sequence and there were fewer bad signals.

4 Aggregate Up Cascades

We now illustrate our theory, focusing in particular on when an AUC occurs. The case in which there are only three agents will provide a useful starting point.

4.1 An Example with Three Agents

Consider the case in which $n = 3$. From Lemma 3 we know that $\mathcal{I}(0, \sigma_i) = \sigma_i$ and from Proposition 3 we know that $\mathcal{I}(2, \sigma_i) = 1$ and that $\mathcal{I}(1, 1) = 1$. Thus, these results alone immediately give us the equilibrium actions for five out of the six possible contingencies agents can face. The only remaining question is now what agents do after observing $T_i = 1$ and $\sigma_i = 0$. This depends on the values of parameters r and q. Let us first check under which conditions agent i rationally follows his bad signal. Recall that we are analyzing a symmetric equilibrium, therefore suppose each other agent j chooses $\mathcal{I}(1, 0) = 0$. Then it is optimal for agent i to do the same if his posterior for the good state is not bigger than $1/2$, that is, if

$$\frac{r[q(1-q) + 2q(1-q)^2]}{r[q(1-q) + 2q(1-q)^2] + (1-r)[q(1-q) + 2q^2(1-q)]} \leq \frac{1}{2}$$

which is equivalent to

$$q \geq 2r - \frac{1}{2} \equiv \underline{q}.$$

Similarly, $\mathcal{I}(1, 0) = 1$ is optimal if

$$r[q(1-q) + q(1-q)^2] \geq (1-r)[q(1-q) + q^2(1-q)].$$

which is equivalent to

$$q \leq 3r - 1 \equiv \overline{q}$$

Note that $\underline{q} \leq \overline{q}$. Hence, in the case in which $\underline{q} > 0.5$ and $\overline{q} < 1$, we obtain three equilibrium regions. For $q < \underline{q}$ there is a unique pure-strategy equilibrium in which $\mathcal{I}(1,0) = 1$ and an AUC starts with $T_i = 1$. For $q > \overline{q}$ there is a unique pure-strategy equilibrium in which $\mathcal{I}(1,0) = 0$ and an AUC starts only with $T_i = 2$. Finally, for $\underline{q} \leq q \leq \overline{q}$ both the two pure-strategy equilibria exist and there is a mixed-strategy equilibrium as well—with $\mathcal{I}(1,0) = \frac{1+2q-4r}{q-r}$ and and an AUC starts only with $T_i = 2$. Note that, when the two states of the world are equally likely ($r = 0.5$), $\underline{q} = \overline{q} = 0.5$ and, in the unique pure-strategy equilibrium, $\mathcal{I}(1,0) = 0$ and an AUC starts with $T_i = 2$.

4.2 The General Case with n Agents

The logic of the previous example can easily be generalized to any finite number of agents. We focus our attention to equilibria in pure strategies only, since the possibility that agents mix for a particular value of T_i does not modify our analysis substantially.

The first pure strategy equilibrium threshold T^{UP} illustrated in the example can be computed by verifying, for each $T_i = 1, 2, 3, ..,$ that an agent has an incentive to follow his bad signal when facing T_i assuming that for all values up to (and including) T_i other agents do follow their bad signals. The first T_i for which this is not verified is the threshold T^{UP}. In other words, the threshold T^{UP} is the first T_i for which the following equality is not satisfied:

$$\sum_{k=0}^{n-T_i-1} \binom{T_i+k}{T_i} q^{T_i}(1-q)^{k+1} r \leq \sum_{k=0}^{n-T_i-1} \binom{T_i+k}{T_i} (1-q)^{T_i} q^{k+1}(1-r).$$

The left (right) hand side represents the probability of observing T_i investments and a bad signal when $\omega = 1$ ($\omega = 0$), multiplied by the prior probability of the good (bad) state of the world. The agent observing T_i investments could be in position T_i+1, $T_i+2, .., n$. If he is in position T_i+k+1 then any combination of T_i investments (i.e., good signals) and k decisions not to invest (i.e., bad signals) before him in the sequence is possible, which explains the binomial coefficient. As long as the inequality is satisfied, the agent observing T investments will find it optimal to follow his bad signal. The first T_i for which this is not optimal represents the threshold for an AUC.

The second pure strategy equilibrium threshold T^{UP} is computed assuming that agents follow their signals up to (but excluding) a certain $T_i = 1, 2...$

and then invest independently of their signals (i.e., they are in an AUC) for T_i, and then verifying that an agent actually finds it optimal to neglect his bad signal when facing T_i. The first T_i for which this is verified is the threshold T^{UP}. In other words, the threshold T^{UP} is the first T_i for which the following equality is satisfied:

$$\sum_{k=0}^{n-T_i-1} \binom{T_i+k-1}{k} q^{T_i}(1-q)^{k+1} r \geq \sum_{k=0}^{n-T_i-1} \binom{T_i+k-1}{k} (1-q)^{T_i} q^{k+1}(1-r).$$

The interpretation of the inequality is the same as for the previous one, given the different assumption on the behavior upon observing T_i investments. In particular, note that since we start from the assumption of a cascade after T_i investments, the agent's predecessor must have observed $T_i - 1$ investments and invested, so that the agent is the first to observe T_i. This explains the differences in the binomial coefficients between the two expressions. It is clear that, by construction, for any set of parameters, this threshold is (weakly) lower than that illustrated above.

Note that the multiplicity of equilibria, with possibly more than one T^{UP}, arises from a sort of information complementarity: if all other agents invest upon any signal realization after observing a specific T_i, then that T_i is better news (than if agents invested only upon a good signal). To see why, compare the two equilibria above. In the first type of equilibrium, observing any T_i is worse news than in the second. In the first equilibrium the assumption is that T_i does not induce a cascade. Therefore, if agent i observes T_i, it is possible that other agents have observed it too and did not to invest because they received the bad signal. In the other equilibrium, T_i triggers a cascade. Therefore, if agent i observes it, it means he is the first one to observe it, which lowers the probability of predecessors having received the bad signal.

Note also that these two types of equilibria provide the upper bound and the lower bound on the equilibrium T^{UP}. It is easy to verify that when these bounds are not consecutive integers, there exist other equilibria in which the threshold is in between these two values.[10]

[10] To see why, suppose in the first equilibrium $T^{UP} = t'$ and in the second $T^{UP} = t''$, with $t' > t'' + 1$. Now consider $t = t'' + 1$. Let us verify that an equilibrium with $T^{UP} = t$ exists. First, observe that following the signal for $T_i = t''$ can be part of an equilibrium (actually, so is in the first equilibrium). Second, using a monotonicity argument as that illustrated in the text, one can show that if the posterior on $\omega = 1$ is higher than $1/2$ assuming agents

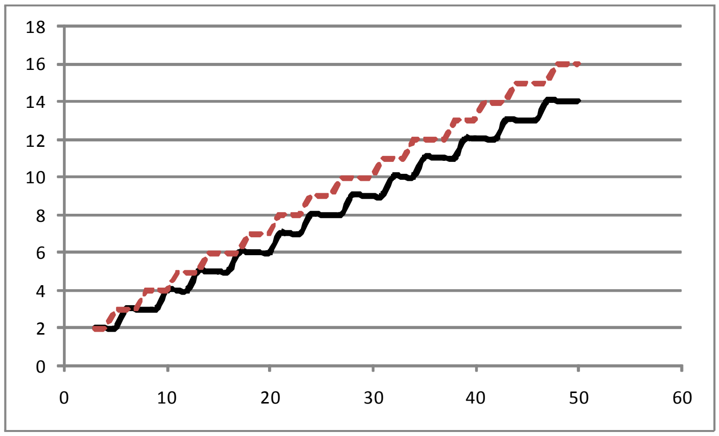

Figure 1: Thresholds for an AUC for different values of the number of agents n in the population (from 3 to 50), r=0.5, q=0.75. The dashed line refers to the first equilibrium and the solid line to the second.

Figure 1 shows the threshold values in the two equilibria above for $r = 0.5$ and $q = 0.75$ and various values of n (the dashed line refers to the first equilibrium; the solid line to the second).[11] When $n = 3$, an AUC occurs after $T^{UP} = 2$ investments. For $n = 10$, it occurs after 4 investments in the first type of equilibrium and after 5 in the second. For $n = 50$, the thresholds become 14 and 16.

Perhaps not surprisingly, the thresholds are (weakly) increasing in n. This is true not only for the specific example of Figure 1 but generally. With a slight abuse of notation, let us define by $T^{UP}(r, q, n)$ the threshold level for a cascade as a function of the parameters r and q and the number of agents in the economy. We have the following result:

Lemma 5 *For any n, $T^{UP}(r, q, n) \leq T^{UP}(r, q, n + 1)$.*

Proof See the Appendix.■

While for finite n the threshold cannot be expressed in an explicit form, things are different in the case of a large population (n tending to infinity).

are in a cascade the first time for $T^{UP} = t''$, then this posterior is higher than 1/2 also for $T^{UP} = t$ in the conjectured equilibrium.

[11] The thresholds have been computed in Matlab.

In the next proposition, we compute the threshold for n going to infinity and show that it is reached almost surely.

Proposition 3 *For n going to infinity, an AUC occurs with probability one at a threshold level $T^{UP}(r,q,n)$ such that* $\lim_{n\to\infty} \frac{T^{UP}(r,q,n)}{n} = \frac{q(q-1)(q-r)}{q^2(2r-1)-r(2q-1)}$. *For* $r = \frac{1}{2}$, $\lim_{n\to\infty} \frac{T^{UP}(\frac{1}{2},q,n)}{n} = q(1-q)$.

Proof See the Appendix.■

In our model, as in the canonical model of Bikhchandani *et al.* (1992), when n goes to infinity, an informational cascade occurs almost surely. Whereas in the canonical model it can occur on either action (and be incorrect, i.e., a cascade can occur on action 1 when the state of the world is 0 and vice versa), in our model it always occurs on the observable action, independently of the state of the world (therefore being incorrect whenever the state of the world is 0).

5 A Welfare Analysis

In our economy agents can make mistakes. This happens whenever a cascade has not occurred and the agent receives the incorrect signal, or whenever an AUC has occurred and the state of the world is $\omega = 0$.

One may wonder what the probability of a mistake is. Recall that *ex ante* all agents are identical, and receive a payoff of 1 for a correct decision and of 0 for an incorrect one. Therefore, the *ex ante* probability of a mistake is also an (inverse) measure of *ex ante* welfare in our economy. More precisely, it is equivalent to the *per capita*, *ex ante* welfare loss.

In the next proposition we illustrate this probability for a large economy:

Proposition 4 *For n going to infinity, the ex ante probability that an agent makes the incorrect decision is $1-q$.*

Proof See the Appendix.■

The *ex ante* probability of a mistake in our economy is equal to $1-q$, which is also the probability of a mistake in the case each agent just decides on the basis of his own private signal. Observing others' decisions is, therefore, not payoff improving for an agent in this large economy. Although surprising at a first glance, the result can be easily understood. A cascade is a situation in which public information swamps private information. As we know from the previous analysis, when the population grows large, the threshold $T^{UP}(r,q,n)$ is such that the information contained in the observable history of decisions is just enough to swamp the information contained in the private signal. Therefore, if the agent acts before the AUC occurs, just following the noisy signal, he makes a mistake with probability $1-q$. And if he acts after the occurrence of an AUC, he makes a mistake with the same probability, since the information contained in $T^{UP}(r,q,n)$ has the same precision as the noisy signal. Hence, the result.

In our economy agents have less information than in the canonical model of Bikhchandani *et al.* (1992), where they can observe the entire sequence of actions. How does less information affects welfare? Intuitively, one may believe that more information cannot harm agents. On the other hand, if more information means a higher probability of cascades and if during cascades agents are more likely to take the wrong decision, information may indeed be harmful.

To attack the problem, we compare the *ex ante* welfare in our model with that in a set-up in which, like in Bikhchandani *et al.* (1992), agents can observe the decision of each predecessor to invest or not, that is, agent i's information set is $\{I_1, I_2, ..., I_{i-1}, \sigma_i\}$. In all other aspects the two models are identical. We restrict the analysis to $n \geqslant 3$, since for $n < 3$ the two models give identical predictions. Moreover, we only present the case in which both states of the world are equally likely ($r = 0.5$).[12]

When agents can observe the entire history of actions, they can be indifferent between the two actions, like when the second agent observes an investment and receives a bad signal. For these cases, one needs to adopt a tie-breaking rule. It turns out that the tie-breaking rule is relevant for our comparison. We will compare the two models under different tie-breaking rules, which will help us to understand the forces that drive the results. The following proposition summarizes our findings:

Proposition 5 *The ex ante probability of making a mistake when agents can only observe* $\{T_i, \sigma_i\}$

[12]The analysis of the case $r > 0.5$ is similar and does not offer more intutitions.

a) is higher than when they can observe $\{I_1, ..., I_{i-1}, \sigma_i\}$ *for any* $n \geqslant 3$ *if the tie-breaking rule is "follow the own signal with probability* 1;*"*

b) is lower than when they can observe $\{I_1, ..., I_{i-1}, \sigma_i\}$ *for any finite* $n \geqslant 3$ *if the tie-breaking rule is "follow the predecessor's decision with probability* 1;*" asymptotically, for* n *going to infinity, the two probabilities are identical and equal to* $1 - q$.

Proof See the Appendix.■

Let us discuss the intuition behind this result. The first tie-breaking rule contemplated in the proposition implies that an agent neglects his signal only when there is a majority of at least two identical actions in the history of actions $\{I_1, ..., I_{i-1}\}$. When an agent acts in a cascade, his decision is, therefore, based on two signals pointing in the same direction. The *ex ante* probability that he makes a mistake under this circumstance is $\frac{(1-q)^2}{q^2+(1-q)^2}$ and lower than when he acts alone deciding on the basis of one signal only, $(1-q)$. In the proof of the proposition, we show that, when agents can only observe the aggregate number of investments, the *ex ante* probability of a mistake, once in a cascade, lies between these two values. Moreover, we prove that cascades are more likely to occur when agents observe the entire history than when they only know aggregate investments. These two observations prove the result.

The second tie-breaking rule in the proposition implies that, when agents can observe the entire history, they simply follow the first decision I_1. Thus, each agent's decision is based on the first signal realization, and the probability of a mistake is equal to $1-q$ for everyone. When agents can only observe the aggregate number of investments, instead, as we said, when a cascade occurs the probability of a mistake is lower than $1 - q$ for any finite $n \geqslant 3$, and is equal to $1 - q$ only asymptotically (as we know from Proposition 4). This proves point b in the proposition.

In simple words, agents are better off in Bikhchandani *et al.* (1992) economy under the first tie-breaking rule, and *ex ante* better off in our economy under the second one. These two tie-breaking rules are at the two extremes. Other rules would require agents to randomize between investing and not. When agents randomize putting a higher weight on their signal and less on the predecessor's decision, the history of decisions becomes more informative and the probability of a mistake, conditional on being in a cascade, decreases. It is, therefore, easy to conjecture that for any n and q there is a threshold such that if agents follow their own signal with a probability lower than this threshold when indifferent, agents are better off in our model.[13]

6 Extensions

We have presented our model in its simplest form, with two states, two signals and two actions. We chose this simple model since we had in mind applications (like the case of an investment in a new technology, or the decision to hire a candidate or not) in which typically there are only two available actions and one is normally non observable (e.g., the decision not to invest, or not to hire).

Nevertheless, one may wonder whether the main message of our study extends to other set ups. In the social learning literature, two main extensions have been considered, one concerning the signal space and one concerning the action space. In this section, we will consider these extensions and show that our main results are not altered. We do not intend to replicate the entire analysis for more general structures of the signal and of the action space, but only briefly illustrate how our main result on the impossibility of ADCs can be generalized.

6.1 Signal Space

Smith and Sørensen (2000) have extended the classical model of Bikhchandani *et al.* (1992) to a general structure of the signal space. Their main result is that an information cascade occurs almost surely if the beliefs are bounded.[14] If, instead, beliefs are unbounded, an information cascade never occurs and beliefs converge to the truth.

Let us consider what happens in our model when we generalize the signal space. Recall that our main result is that an ADC cannot occur, while an AUC occurs almost surely. If we allow for signals of unbounded precision, *a fortiori*, an ADC cannot occur. An AUC cannot occur either, since, whatever the content of the aggregate public information, there is always the possibility that an agent receives a very precise signal and decides not to invest even if

[13] We actually computed this threshold numerically for various parameter values. We do not report them in the interest of space.

[14] Beliefs are said to be bounded if the support of the likelihood ratio is bounded.

many other agents have opted for investment. Since the argument is standard, we do not develop it further and refer the reader to Smith and Sørensen (2000).

The remaining question is whether the impossibility of an ADC is a result of our binary signals or holds more generally. Suppose the signals $\{\sigma_i\}$ are i.i.d. and drawn according to a state-contingent density function $f(\sigma_i|\omega)$ on a support $[e, E]$ (with $e < E$). To avoid trivialities, we assume that, for almost all signal realizations, either $\frac{f(\sigma_i|\omega=1)}{f(\sigma_i|\omega=0)}\frac{r}{1-r} < 1$ (the "bad signals") or $\frac{f(\sigma_i|\omega=1)}{f(\sigma_i|\omega=0)}\frac{r}{1-r} > 1$ (the "good signals"), and that signals respect the standard monotone likelihood ratio property. We now show that the result at the heart of our finding on the impossibility of ADCs holds under this more general structure of private information (keeping all other aspects of the model unchanged).

Proposition 6 *Given this structure of the private signal, there exists no equilibrium in which, if nobody has invested so far (i.e., $T_i = 0$), an agent chooses not to invest for any signal, that is, $\mathcal{I}(0, \sigma_i) = 0$ for all $\sigma_i \in [e, E]$.*

Proof By contradiction, suppose that for $T_i = 0$ agents choose never to invest (independently of their private signals). Then, along the equilibrium path, nobody ever invests and, for any agent $i = 1, ..., n$, $T_i = 0$. Hence, $T_i = 0$ does not reveal any information on the true state of the world. Since the posterior probability (conditional on public information) that $\omega = 1$ is still r, agent i is better off by following his informative signal σ_i.■

According to this proposition, an ADC in which all agents do not invest is, indeed, impossible. To conclude that an ADC cannot occur, one has only to show that also the intuitive monotonicity result (Lemma 4) can be extended to this more general structure. In the interest of space, we do not provide such a proof, but note that our proof of Lemma 4 can easily be extended to this case.[15]

[15] For a given $T_i = t$, one can partition the set of all signal realizations in the set for which the agent invests and that for which he does not. Then, one can prove that for the first set an agent would invest after observing $t + 1$ by following the same steps as in the proof of Lemma 4 (realizations of the first set would play the same role as $\sigma_i = 1$, and realizations of the second set would play the same role as $\sigma_i = 0$).

6.2 Action Space

Let us now consider a more general action space on top of the more general signal space. Suppose each agent i can choose his action a_i among $m+1$ possible alternatives, denoted by $0, 1, 2, ...m$, and let the action 0 be the unobservable one. A simple interpretation could be that agents must choose the scale of an investment project. While the total number of each type (size) of project is observable, the decision not to invest (action 0) is unobservable. Let us say that agents choose the optimal action to maximize the conditional expected value of a standard quadratic payoff function $-(a_i - \omega)$.

Now, the analog of our case $T_i = 0$ is the case in which an agent observes that no investment of any size has yet been chosen. Let us denote such a case by an m-dimensional vector $(0, 0, ...0)$, where each entry represents the number of investments of size $1, ..., m$.

We now prove a proposition similar to the previous one:

Proposition 7 *Given this structure of the private signal and this action space, there exists no equilibrium in which, if nobody has invested so far, that is, agent i observes $(0, 0, .., 0)$, the agent chooses not to invest for any signal, that is, $\mathcal{I}((0, 0, .., 0), \sigma_i) = 0$ for $\sigma_i \in [e, E]$.*

Proof By contradiction, suppose that when observing $(0, 0, ...0)$, agents choose never to invest (independently of their private signals). Then, along the equilibrium path, nobody ever invests and, any agent $i = 1, ..., n$, observes $(0, 0, ...0)$. Hence, $(0, 0, ...0)$ does not reveal any information on the true state of the world. Since the posterior probability (conditional on public information) that $\omega = 1$ is still r, agent i is better off by following his informative signal σ_i.■

Again, according to this result, an ADC in which all agents do not invest is impossible. And, again, to conclude that an ADC is impossible, one has only to extend the monotonicity result, a proof that we omit in the interest of space.[16]

[16] In this set up, the monotonicity result states that if an investment of a certain size is chosen for a particular signal realization and $T_i = t$, then for the same signal and $T_i = t+1$ the agent would choose an investment of the same size or larger. The proof coudl be amended along the lines indicated in the previous footnote.

Finally, an alternative set up considered in the literature is that of a continuous action space (Lee, 1993). The standard result in this case is that learning is efficient and cascades do not occur. This structure, however, is not interesting for our case, since the probability of a specific action would be zero. Of course, one could discretize the action space and assume that actions belonging to an interval of positive measure are unobservable. This would render the model identical to the one just described and the same considerations would apply.

7 Conclusion

We have introduced a new model of information cascades. The crucial difference between our model and those already in the literature is that only one action taken by agents is observable by others. When it is their turn to make the binary decision, agents simply receive aggregate information about how many others before them took the observable action. We have argue that this setup arises naturally in many scenarios: for example, when entrepreneurs seek investors they will typically not inform them about how many others have turned them down before, but, surely, they will mention who else decided previously to invest in their project. This asymmetry in observability significantly affects the equilibria in such games. Most importantly, there can be no information cascades on the unobservable action.

Our result has important implications. In particular, it implies that a new, good project (e.g., a technological innovation, a new product or service, a new medical treatment) will not be neglected for ever simply because there is lack of interest at the beginning. Sooner or later (i.e., as soon as people start receiving good information on it) the new project will start diffusing. A lack of initial interest will not represent a barrier to future adoption because of informational considerations.

Our study has also an interesting consequence for applications where a third party can decide which information to release. Consider the introduction of a new medical treatment. An agency in health policy must decide how to disclose information on the adoption of this treatment: it can release information on how many physicians have already adopted it; or on how many have considered it but have decided not to adopt it; or, finally, it can reveal both figures. Suppose the agency considers as the worst case scenario the situation in which the treatment is widely adopted while ultimately resulting in bad health outcomes, for instance because of potential side effects.[17] As we know from our analysis, the way the information is disclosed makes a big difference for the diffusion of the new treatment. By withholding information on a particular decision the agency can, in fact, guarantee that an informational cascade on that decision does not occur.[18] In the light of this result, the agency can then optimally withhold information on the number of doctors who decided to adopt the treatment.

References

[1] Acemoglu, D., Dahleh, M., Lobel, I., and Ozdaglar, A., 2010. Bayesian Learning in Social Networks. *Mimeo*, MIT.

[2] Banerjee, A., 1992. A Simple Model of Herd Behavior. *Quarterly Journal of Economics* 107, 797-817.

[3] Banerjee, A.V., and D. Fudenberg, 2004. Word-of-Mouth Learning. *Games and Economic Behavior* 46, 1-23.

[4] Bikhchandani, S., D. Hirshleifer, and I. Welch, 1992. A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades. *Journal of Political Economy* 110, 992-1026.

[5] Çelen, B., and S. Kariv, 2004. Observational Learning under Imperfect Information. *Games and Economic Behavior* 47, 72-86.

[6] Chamley, C., 2004. *Rational Herds*. Cambridge University Press.

[7] Collander S., Hörner J., 2009. The Wisdom of the Minority. *Journal of*

[17] This is of course unmodelled in our analysis. In our welfare analysys we studied whether partial or full observability gave the higher ex ante *per capita* payoffs to the decision makers. Here we are considering a case in which, for instance because the doctors' interests are not perfectly aligned with those of patients', the relevant welfare criterium is different.

[18] In their seminal paper on information cascades, Bikhchandani *et al.* (1992) have argued that the adoption of medical procedures is often based on fairly weak information and that in many cases doctors tend to imitate others. As an example they cite the widespread use of tonsillectomies in the sixties and seventies and argue that it was essentially an information cascade. In the sixties and seventies, according to Bikhchandani *et al.* (1992), the sheer fact that the majority of physicians employed the procedure overrode any private information individual doctors might have had against tonsillectomies. And this was a "wrong cascade"—a cascade that generated the worst outcome, since it eventually turned out that tonsillectomies did more harm than good.

Economic Theory 144, 1421-1439.

[8] Gale, D., 1996. What Have We Learned from Social Learning. *European Economic Review* 40, 617-628.

[9] Guarino, A., Harmgart, H. and S. Huck, 2007. When Half the Truth is Better than the Truth: Theory and Experiments on Aggregate Information Cascades. *ELSE Working Paper* 257.

[10] Guarino, A., Jehiel, P., 2009. Social Learning with Coarse Inference. *ELSE Working Paper* 337.

[11] Herrera, H., Hörner, J., 2009. Biased Social Learning. Cowles Foundation Discussion Paper No. 1738.

[12] Eyster, E., Rabin, M., 2008. Naive Herding. *Mimeo*, Berkeley and LSE.

[13] Hirshleifer, D. and S. Teoh, 2003. Herd Behaviour and Cascading in Capital Markets: a Review and Synthesis. *European Financial Management* 9, 25-66.

[14] Larson, N., 2008. In with the New Or Out with the Old? Bottlenecks in Social Learning. *Mimeo,* University of Virginia.

[15] Lee, I. H., 1993. On the Convergence of Informational Cascades. *Journal of Economic Theory* 61, 395-411.

[16] Smith, L., and P. Sørensen, 2008. Rational Social Learning with Random Sampling. *Mimeo*, MIT and University of Copenhagen.

[17] Sufi, A., 2007. Information Asymmetry and Financing Arrangements: Evidence from Syndicated Loans. *Journal of Finance* 62, 629-668.

[18] Vives, X., 2008. *Information and Learning in Markets: The Impact of Market Microstructure.* Princeton University Press.

8 Appendix

A Proof of Lemma 4

The proposition is equivalent to saying that in equilibrium $\mathcal{I}(t,\sigma_i) \leq \mathcal{I}(t+1,\sigma_i)$ for any $t = 0,1,2...$and both $\sigma_i = 0$ and $\sigma_i = 1$ (with strict inequality if $0 < \mathcal{I}(t,\sigma_i) < 1$). Because of expected payoff maximization, this inequality holds if, whenever $\Pr(\omega = 1 \mid T_i = t, \sigma_i) \geq \frac{1}{2}$, we have $\Pr(\omega = 1 \mid T_i = t+1, \sigma_i) > \frac{1}{2}$.

There are three relevant possibilities:

1. $\Pr(\omega = 1 \mid T_i = t, \sigma_i = 0) > \frac{1}{2}$ and $\Pr(\omega = 1 \mid T_i = t, \sigma_i = 1) > \frac{1}{2}$,

2. $\Pr(\omega = 1 \mid T_i = t, \sigma_i = 0) < \frac{1}{2}$ and $\Pr(\omega = 1 \mid T_i = t, \sigma_i = 1) > \frac{1}{2}$,

3. $\Pr(\omega = 1 \mid T_i = t, \sigma_i = 0) = \frac{1}{2}$ and $\Pr(\omega = 1 \mid T_i = t, \sigma_i = 1) > \frac{1}{2}$.

Case 1 is the case of an informational cascade. In such a case, by our Assumption,

$$\mathcal{I}(t, \sigma_i) = \mathcal{I}(t + 1, \sigma_i) = 1$$

for both σ_i, and therefore the proposition obviously holds.

Now let us consider Case 2. In this case we want to show that $\Pr(\omega = 1 \mid T_i = t + 1, \sigma_i = 1) > \frac{1}{2}$ (while nothing must be shown for the case of a bad signal). Suppose not, i.e., suppose $\Pr(\omega = 1 \mid T_i = t + 1, \sigma_i = 1) \leq \frac{1}{2}$. Let us consider, first, the case of the strict inequality.

By Bayes's rule,

$$\Pr(\omega = 1 \mid T_i = t + 1, \sigma_i = 1) =$$
$$= \frac{\Pr(T_i{=}t{+}1|\omega{=}1,\sigma_i{=}1)r}{\Pr(T_i{=}t{+}1|\omega{=}1,\sigma_i{=}1)r+\Pr(T_i{=}t{+}1|\omega{=}0,\sigma_i{=}1)(1-r)}.$$

As we suppose that this is strictly smaller than $\frac{1}{2}$ it follows that

$$\frac{\Pr(T_i = t + 1 \mid \omega = 1, \sigma_i = 1)}{\Pr(T_i = t + 1 \mid \omega = 0, \sigma_i = 1)} < \frac{1 - r}{r}.$$

which is equivalent to

$$\frac{\Pr(T_i = t + 1 \mid \omega = 1)}{\Pr(T_i = t + 1 \mid \omega = 0)} < \frac{1 - r}{r}.$$

By the law of total probabilities,

$$\begin{aligned}
&\Pr(T_i = t + 1 \mid \omega = 1) \\
&= \Pr(T_i = t + 1 \mid \omega = 1, T_{i-1} = t) \Pr(T_{i-1} = t|\omega = 1) \\
&+ \Pr(T_i = t + 1 \mid \omega = 1, T_{i-1} = t + 1) \Pr(T_{i-1} = t + 1|\omega = 1) \\
&= q \Pr(T_{i-1} = t|\omega = 1) + \Pr(T_{i-1} = t + 1|\omega = 1).
\end{aligned}$$

Notice that the last equality comes from the fact that we are analyzing Case 2 and that we are assuming no investment after observing $t + 1$.

Now the decision problem of agent $i - 1$ is identical to the one of agent i.

So, by applying recursively the same law, we obtain:

$$\begin{aligned}
&\Pr(T_i = t+1 \mid \omega = 1) \\
&= q\Pr(T_{i-1} = t|\omega = 1) + \Pr(T_{i-1} = t+1|\omega = 1, \sigma_i = 1) \\
&= q\Pr(T_{i-1} = t|\omega = 1) + [q\Pr(T_{i-2} = t|\omega = 1) + \Pr(T_{i-2} = t+1|\omega = 1)] \\
&+q\Pr(T_{i-1} = t|\omega = 1) + q\Pr(T_{i-2} = t|\omega = 1) + [q\Pr(T_{i-3} = t|\omega = 1) \\
&+\Pr(T_{i-3} = t+1|\omega = 1)] + ... \\
&= q\Pr(T_{i-1} = t|\omega = 1) + q\Pr(T_{i-2} = t|\omega = 1) + q\Pr(T_{i-3} = t|\omega = 1) \\
&+... + q\Pr(T_{i-m} = t|\omega = 1)]
\end{aligned}$$

for some m (note that m depends on the value of i: indeed, for any value of i there is an m such that $\Pr(T_{i-m} = t+1|\omega = 1) = 0$). Similarly, conditioning on $\omega = 0$,

$$\begin{aligned}
&\Pr(T_i = t+1 \mid \omega = 0) \\
&= (1-q)\Pr(T_{i-1} = t|\omega = 0) + (1-q)\Pr(T_{i-2} = t|\omega = 0) \\
&+... + (1-q)\Pr(T_{i-m} = t|\omega = 0).
\end{aligned}$$

Some algebraic computations show that for any pair of terms in the two expressions above, the following inequality holds:

$$\frac{\Pr(T_{i-j} = t|\omega = 1)}{\Pr(T_{i-j} = t|\omega = 0)} \geq \frac{\Pr(T_i = t|\omega = 1)}{\Pr(T_i = t|\omega = 0)}.$$

Since we know that $\Pr(\omega = 1 \mid T_i = t, \sigma_i = 1) > \frac{1}{2}$ and, therefore,

$$\frac{\Pr(T_i = t \mid \omega = 1)}{\Pr(T_i = t \mid \omega = 0)} > \frac{1-r}{r},$$

simple algebra shows that

$$\frac{\Pr(T_i = t+1 \mid \omega = 1)}{\Pr(T_i = t+1 \mid \omega = 0)} > \frac{1-r}{r},$$

a contradiction.

Note that the same proof holds true when, by contradiction, we assume that

$$\Pr(\omega = 1 \mid T_i = t+1, \sigma_i = 1) = \frac{1}{2}.$$

The only difference is that in such a case

$$\Pr(T_i = t+1 \mid \omega = 1)$$
$$= q\Pr(T_{i-1} = t|\omega = 1) + s\Pr(T_{i-1} = t+1|\omega = 1),$$

where s represents the probability by which an agent receiving the good signal decides not to invest. This change does not affect the above inequalities.

Finally, note that the proof for Case 3 (for both the good and the bad signal) is identical to Case 2 just described, with the exception that

$$\Pr(T_i = t+1 \mid \omega = 1, T_{i-1} = t) = q + (1-q)u,$$

and

$$\Pr(T_i = t+1 \mid \omega = 0, T_{i-1} = t) = qu + (1-q),$$

where u is the probability of investment by an agent receiving a bad signal.■

B Proof of Proposition 1

We complete the proof in the text concerning the existence of the equilibrium. We have already proven in the text that $\mathcal{I}(T_i, 1) = 1$ is optimal for any T_i. Therefore, for the existence of a pure strategy (symmetric Perfect Bayesian) equilibrium we must only prove that for an agent observing $(t, 0)$ it is optimal not to invest if all other agents observing $(t, 0)$ do not invest, or it is optimal to invest if all other agents observing $(t, 0)$ invest, or both (in which case, there exists an equilibrium with $\mathcal{I}(t, 1) = 0$ and one with $\mathcal{I}(t, 1) = 1$). With a slight abuse of notation, let us denote by $\Pr(\omega = 1 \mid T_i = t, \sigma_i = 0, u)$ the probability attached to $\omega = 1$ by an agent observing $(t, 0)$ and knowing that all other agents invest with probability $u \in [0, 1]$ when observing the same couple $(t, 0)$. Then, what we want to prove is that

$$\Pr(\omega = 1 \mid T_i = t, \sigma_i = 0, u = 0) < \Pr(\omega = 1 \mid T_i = t, \sigma_i = 0, u = 1).$$

Indeed, if this is verified, then, if

$$\Pr(\omega = 1 \mid T_i = t, \sigma_i = 0, u = 0) < \Pr(\omega = 1 \mid T_i = t, \sigma_i = 0, u = 1) < \frac{1}{2}$$

there exists a pure strategy equilibrium in which $\mathcal{I}(t, 0) = 0$; if, on the other hand,

$$\frac{1}{2} < \Pr(\omega = 1 \mid T_i = t, \sigma_i = 0, u = 0) < \Pr(\omega = 1 \mid T_i = t, \sigma_i = 0, u = 1),$$

there exists a pure strategy equilibrium in which $\mathcal{I}(t,0)=1$; finally, if

$$\Pr(\omega=1 \mid T_i=t,\sigma_i=0,u=0)<\frac{1}{2}<\Pr(\omega=1 \mid T_i=t,\sigma_i=0,u=1),$$

both equilibria with $\mathcal{I}(t,0)=0$ and $\mathcal{I}(t,0)=1$ exist.

To prove the inequality, observe that

$$\frac{\Pr(\omega=1|T_i=t,\sigma_i=0,u=1)}{\Pr(\omega=0|T_i=t,\sigma_i=0,u=1)}=$$
$$\frac{\Pr(T_i=t|\omega=1,\sigma_i=0,u=1)}{\Pr(T_i=t|\omega=0,\sigma_i=0,u=1)}\frac{r(1-q)}{q(1-r)}=$$
$$\frac{q^t\left[1+k_1(1-q)+k_2(1-q)^2+...+k_{n-1-t}(1-q)^{n-t-1}\right]}{(1-q)^t\left[1+k_1q+k_2q^2+...+k_{n-1-t}q^{n-t-1}\right]}\frac{r(1-q)}{q(1-r)},$$

where k_1,k_2...are binomial coefficients. Similarly, one can show that

$$\frac{\Pr(\omega=1|T_i=t,\sigma_i=0,u=0)}{\Pr(\omega=0|T_i=t,\sigma_i=0,u=0)}=$$
$$\frac{\Pr(T_i=t|\omega=1,\sigma_i=0,u=0)}{\Pr(T_i=t|\omega=0,\sigma_i=0,u=0)}\frac{r(1-q)}{q(1-r)}=$$
$$\frac{q^t\left[1+(k_1+1)(1-q)+(k_2+k_1+1)(1-q)^2+...+(k_{n-1-t}+k_{n-2-t}+...+1)(1-q)^{n-t-1}\right]}{(1-q)^t\left[1+(k_1+1)q+(k_2+k_1+1)q^2+...+(k_{n-1-t}+k_{n-2-t}+...+1)q^{n-t-1}\right]}\frac{r(1-q)}{q(1-r)}.$$

The extra terms in this expression come from the fact that, if $u=0$, the number of investments $T_i=t$ can be observed by agent i also after one or more predecessors have already observed it and, after having bad signals, did not invest. The comparison between the two expressions immediately proves the inequality.

Similarly, one can show that if agents invest with a probability $0<u<1$, then

$$\frac{\Pr(\omega=1|T_i=t,\sigma_i=0,u)}{\Pr(\omega=0|T_i=t,\sigma_i=0,u)}=$$
$$\frac{\Pr(T_i=t|\omega=1,\sigma_i=0,u)}{\Pr(T_i=t|\omega=0,\sigma_i=0,u)}\frac{r(1-q)}{q(1-r)}=$$
$$\frac{q^t\left[1+(k_1+(1-u))(1-q)+\left(k_2+k_1(1-u)+(1-u)^2\right)(1-q)^2+...+\left(k_{n-1-t}+k_{n-2-t}(1-u)+...+(1-u)^{n-2-t}\right)(1-q)^{n-t-1}\right]}{(1-q)^t\left[1+(k_1+(1-u))q+(k_2+k_1(1-u)+(1-u)^2)q^2+...+(k_{n-1-t}+k_{n-2-t}(1-u)+...+(1-u)^{n-2-t})q^{n-t-1}\right]}\frac{r(1-q)}{q(1-r)}.$$

It is easy to verify that this expression is lower than the first and higher than the second, which proves part *(ii)* of the proposition given that, if for $u\in(0,1)$, $\Pr(\omega=1|T_i=T^{MIX},\sigma_i=0,0<u<1)=\frac{1}{2}$, then $\Pr(\omega=1|T_i=T^{MIX},\sigma_i=0,u=0)<\frac{1}{2}$ and $\Pr(\omega=1|T_i=T^{MIX},\sigma_i=0,u=1)>\frac{1}{2}$.■

C Proof of Proposition 2

Let us prove the proposition for $j = 1$, that is, let us prove that

$$\Pr(\omega = 1 \mid T_i = T^{UP} + 1, \sigma_i) > \Pr(\omega = 1 \mid T_i = T^{UP}, \sigma_i),$$

which is equivalent to proving that

$$\Pr(\omega = 1 \mid T_i = T^{UP} + 1) > \Pr(\omega = 1 \mid T_i = T^{UP}).$$

First, note that

$$\Pr(\omega = 1 \mid T_i = T^{UP} + 1) =$$

$$\sum_{k=2}^{n-T^{UP}} \Pr(\omega = 1 \mid T_i = T^{UP} + 1, i = T^{UP} + k) \Pr(i = T^{UP} + k \mid T_i = T^{UP} + 1).$$

To simplify the notation, let $\Pr(\omega = 1 \mid T_i = T^{UP} + 1, i = T^{UP} + k) = a_k$ and $\Pr(i = T^{UP} + k \mid T_i = T^{UP} + 1) = p_k$ so that

$$\Pr(\omega = 1 \mid T_i = T^{UP} + 1) = a_2 p_2 + a_3 p_3 + ... + a_n p_n.$$

Similarly,

$$\Pr(\omega = 1 \mid T_i = T^{UP}) =$$

$$\sum_{k=1}^{n-T^{UP}} \Pr(\omega = 1 \mid T_i = T^{UP}, i = T^{UP} + k) \Pr(i = T^{UP} + k \mid T_i = T^{UP}).$$

Again to simplify the notation, let $\Pr(\omega = 1 \mid T_i = T^{UP}, i = T^{UP} + k) = b_k$, and $\Pr(i = k \mid T_i = T^{UP} + 1) = q_k$, so that

$$\Pr(\omega = 1 \mid T_i = T^{UP}) = b_1 q_1 + b_2 q_2 + b_3 q_3 + ... + b_n q_n.$$

Now, it is easy to verify that $a_k > a_{k+1}$, $b_k > b_{k+1}$, and $a_{k+1} = b_k$.

To conclude the proof, we now show that $p_{k+1} > q_k$. To this aim, simply observe that

$$q_k = \Pr(i = T^{UP} + k \mid T_i = T^{UP}) =$$

$$\frac{\Pr(T_i = T^{UP} \mid i = T^{UP} + k) \Pr(i = T^{UP} + k)}{\sum_{k=1}^{n-T^{UP}} \Pr(T_i = T^{UP} \mid i = T^{UP} + k) \Pr(i = T^{UP} + k)} =$$

$$\frac{\Pr(T_i = T^{UP} \mid i = T^{UP} + k)}{\sum_{k=1}^{n-T^{UP}} \Pr(T_i = T^{UP} \mid i = T^{UP} + k)},$$

$$p_{k+1} = \Pr(i = T^{UP} + k + 1 \mid T_i = T^{UP} + k + 1) =$$
$$\frac{\Pr(T_i = T^{UP} + 1 \mid i = T^{UP} + k + 1)\Pr(i = T^{UP} + k + 1)}{\sum_{k=1}^{n-T^{UP}-1} \Pr(T_i = T^{UP} + 1 \mid i = T^{UP} + k + 1)\Pr(i = T^{UP} + k + 1)} =$$
$$\frac{\Pr(T_i = T^{UP} + 1 \mid i = T^{UP} + k + 1)}{\sum_{k=1}^{n-T^{UP}-1} \Pr(T_i = T^{UP} + 1 \mid i = T^{UP} + k + 1)}.$$

and

$$\Pr(T_i = T^{UP} \mid i = T^{UP} + k) = \Pr(T_i = T^{UP} + 1 \mid i = T^{UP} + k + 1).$$

Since the numerators in the expressions for p_{k+1} and q_k are identical and so are the terms in the sums at the denominators (but the denominator in p_{k+1} has one less term), clearly $p_{k+1} > q_k$. This concludes the proof for $j = 1$. An identical analysis proves that the inequality holds for $j = 2, ...n - T^{UP} - 2$.■

D Proof of Lemma 5

With a slight abuse of notation, let us denote by $\Pr(\omega = 1|T_i = t, \sigma_i = 0, n)$ the probability that the state of the world is 1 upon observing t investments and a bad signal in an economy with n agents.

Now, suppose $t = T^{UP}(r, q, n)$. We want to show that

$$\Pr(\omega = 1|T_i = t - 1, \sigma_i = 0, n) > \Pr(\omega = 1|T_i = t - 1, \sigma_i = 0, n + 1),$$

that is,

$$\Pr(\omega = 1|T_i = t - 1, n) > \Pr(\omega = 1|T_i = t - 1, n + 1).$$

By the law of total probabilities,

$$\Pr(\omega = 1|T_i = t - 1, n) =$$
$$\sum_{k=0}^{n-t} \Pr(\omega = 1|T_i = t - 1, i = t + k, n)\Pr(i = t + k|T_i = t - 1, n).$$

Similarly,

$$\Pr(\omega = 1|T_i = t-1, n+1) =$$

$$\sum_{k=0}^{n-t+1} \Pr(\omega = 1|T_i = t-1, i = t+k, n+1) \Pr(i = t+k|T_i = t-1, n+1).$$

Observation 1: $\Pr(\omega = 1|T_i = t-1, i = t+k, n) = \Pr(\omega = 1|T_i = t-1, i = t+k, n+1)$ for any $k = 0, 1, ..., n-t$. This observation means that in the two sums above the first part of each term is the same (but the second sum has one extra term).

Observation 2: $\Pr(\omega = 1|T_i = t-1, i = t+k, n) > \Pr(\omega = 1|T_i = t-1, i = t+k+1, n)$ for any $k = 0, 1, .., n-t-1$. (And similarly for the case of $n+1$ agents.)

Given these two observations, to complete the proof, we only have to show that

$$\Pr(i = t+k|T_i = t-1, n+1) < \Pr(i = t+k|T_i = t-1, n)$$

for any $k = 0, 1, .., n-t$.

To this aim, note that

$$\Pr(i = t+k|T_i = t-1, n) =$$

$$\frac{\Pr(T_i = t-1|i = t+k, n)}{\sum_{l=0}^{n-t} \Pr(T_i = t|i = t+l, n)} = \frac{\binom{(t-1)+k-1}{t-1} q^{t-1}(1-q)^k}{\sum_{l=0}^{n-t} \binom{(t-1)+l-1}{t-1} q^{t-1}(1-q)^l},$$

and, similarly,

$$\Pr(i = t+k|T_i = t-1, n+1) =$$

$$\frac{\binom{(t-1)+k-1}{t-1} q^{t-1}(1-q)^k}{\sum_{l=0}^{n-t+1} \binom{(t-1)+l-1}{t-1} q^{t-1}(1-q)^l}.$$

While the numerator in these two expressions is the same, the denominator differs in that there is an extra term, which proves our result.■

E Proof of Proposition 3

To prove the proposition, let us compute the probability $\Pr(\omega = 1|T_i \geq T^{UP})$. By Bayes's rule,

$$\Pr(\omega = 1|T_i \geq T^{UP}) = \frac{\Pr(T_i \geq T^{UP}|\omega = 1)r}{\Pr(T_i \geq T^{UP}|\omega = 1)r + \Pr(T_i \geq T^{UP}|\omega = 0)(1-r)}.$$

Consider, first, the case $\omega = 1$. For large n, by the Strong Law of Large Numbers, the number of aggregate investments T^{UP} will be reached almost surely after a number M_1 of agents have made a decision such that $qM_1 = T^{UP}$ (since, in equilibrium, agents follow their private signals when observing up to T^{UP} investments). Hence $M_1 = \frac{T^{UP}}{q}$. Similarly, for $\omega = 0$, T^{UP} will be reached almost surely after a number M_0 of agents have made a decision such that $(1-q)M_0 = T^{UP}$. Hence $M_0 = \frac{T^{UP}}{1-q}$. Therefore, for large n, the probability of observing at least T^{UP} investments when $\omega = 1$ is equivalent to the probability of acting after M_1 agents and hence equal to $\Pr(T_i \geq T^{UP}|\omega = 1) = \frac{(n-\frac{T^{UP}}{q})}{n}$. Similarly, $\Pr(T_i \geq T^{UP}|\omega = 0) = \frac{(n-\frac{T^{UP}}{1-q})}{n}$. An agent observing at least T^{UP} investments and receiving a bad signal decides to invest (i.e., he is in an AUC) when

$$\Pr(\omega = 1|T_i \geq T^{UP}) \geq q,$$

that is, when

$$\frac{\frac{(n-\frac{T^{UP}}{q})}{n}r}{\frac{(n-\frac{T^{UP}}{q})}{n}r + \frac{(n-\frac{T^{UP}}{1-q})}{n}(1-r)} \geq q,$$

which can be rearranged as $T^{UP} \geq \frac{q(q-1)(q-r)}{q^2(2r-1)-r(2q-1)}n$. Finally, note that we have found this inequality by noting that agents who observe $T_i < T^{UP}$ follow the signal and agents who observe a higher T_i neglect the signal, which implies that

$$T^{UP} = \frac{q(q-1)(q-r)}{q^2(2r-1)-r(2q-1)}n.$$

For $r = 0.5$, $T^{UP} = q(1-q)n$.■

F Proof of Proposition 4

Let us first consider $\omega = 1$. In this case, an agent makes a mistake only if he acts before the threshold is reached and, additionally, receives the wrong signal. As we know from the previous analysis, the probability of acting before T^{UP} is $\Pr(T_i < T^{UP}|\omega = 1) = \frac{(\frac{T^{UP}}{q})}{n} = \frac{(q-1)(q-r)}{q^2(2r-1)-r(2q-1)}$. Therefore, the probability of a mistake is $\frac{(q-1)(q-r)}{q^2(2r-1)-r(2q-1)}(1-q) = \frac{(r-q)(q-1)^2}{r+2q^2r-2qr-q^2}$. Let us now consider $\omega = 0$. In this case an agent makes a mistake if he acts before the threshold and receives the wrong signal, or if he acts after the threshold.

Since, as we know, $\Pr(T_i < T^{UP}|\omega = 0) = \frac{\frac{T^{UP}}{1-q}}{n} = -\frac{q(q-r)}{q^2(2r-1)-r(2q-1)}$, the probability of making a mistake is now equal to

$$\frac{q(r-q)}{q^2(2r-1)-r(2q-1)}(1-q)+(1-\frac{q(r-q)}{q^2(2r-1)-r(2q-1)}) = \frac{(q-1)\left(q^2+rq-r\right)}{r+2q^2r-2qr-q^2}.$$

Hence, the *ex ante* total probability of a mistake is equal to

$$\frac{(r-q)(q-1)^2}{r+2q^2r-2qr-q^2}r + \frac{(q-1)\left(q^2+rq-r\right)}{r+2q^2r-2qr-q^2}(1-r) = 1-q.$$

■

G Proof of Proposition 5

Let us first prove part a. For the sake of exposition, we will refer to the condition in which agent i's information set is $\{I_1, I_2, ..., I_{t-i}, \sigma_i\}$ as "perfect observability" and to the situation in which agent i's information set is $\{T_i, \sigma_i\}$ as "imperfect observability." Under both conditions, an agent can be in two situations: either he acts on the basis of his own signal or he is in a cascade. Consider first the case of perfect observability. If the agent acts according to his signal, the *ex ante* probability that he makes a mistake is $1-q$. If, instead, he is in a cascade, then his action is based on a majority of two (either good or bad) signals and the *ex ante* probability of a mistake is $\frac{(1-q)^2}{q^2+(1-q)^2}$. Now consider the the case of imperfect observability. If he acts according to his signal, the *ex ante* probability that he makes a mistake is, again, $1-q$. If, instead, he is in a cascade, then he makes a mistake with a probability α. Therefore, under perfect observability, the *ex ante* probability of a mistake is $(1-q)(1-p) + \frac{(1-q)^2}{q^2+(1-q)^2}p$ (where p is the probability that a cascade occurs), and under imperfect observability it is $(1-q)(1-s) + \alpha s$ (where s is the probability that a cascade occurs). We prove our result in two steps. First, we show that $\frac{(1-q)^2}{q^2+(1-q)^2} \leq \alpha$. Second, we show that the probability of a cascade is higher under the perfect observability condition, that is, $p > s$.

Step 1: $\frac{(1-q)^2}{q^2+(1-q)^2} \leq \alpha$.

Let us prove the first inequality. By contradiction, suppose $\alpha < \frac{(1-q)^2}{q^2+(1-q)^2}$. This means that $\Pr(\omega = 0|T_i = T^{UP}) < \frac{(1-q)^2}{q^2+(1-q)^2}$, that is, $\frac{\Pr(\omega=1|T_i=T^{UP})}{\Pr(\omega=0|T_i=T^{UP})} > \frac{q^2}{(1-q)^2}$. Then, consider the situation of an agent observing $T^{UP}-1$ investments. Since, by assumption he is not in a cascade, his decision to invest is based on a good signal. Then, the history with $T^{UP}-1$ investments that he observes is such that $\frac{\Pr(\omega=1|T_i=T^{UP}-1)}{\Pr(\omega=0|T_i=T^{UP}-1)} > \frac{q}{(1-q)}$. (Indeed, if this were not the case, then $\frac{\Pr(\omega=1|T_i=T^{UP})}{\Pr(\omega=0|T_i=T^{UP})}$ could not be greater than $\frac{q^2}{(1-q)^2}$ since even in the best scenario in which after $T^{UP}-1$ there were only one investment (and no bad signals), the likelihood ratio could not increase by more than $\frac{q}{1-q}$). In that case the agent should invest independently of the signal, which contradicts the fact that he is not in a cascade.

Step 2: The probability of a cascade is higher when there is full observability.

We know that in our model an AUC starts not before there are at least two investments (from the example in Section 4.1 and from Lemma 5). This means that, unless signal realizations always alternate, that is, whenever $\sigma_i = 1$, then $\sigma_{i+1} = 0$, and *vice versa*, there cannot exist a sequence of signals such that a cascade occurs under imperfect observability without occurring (at the same time or at an earlier time) under perfect observability. Moreover, note that whenever a cascade occurs under imperfect observability for a sequence of alternating signals $S_1 \equiv \{1, 0, 1, 0, 1, ..., 0, 1\}$, it occurs under perfect observability for the sequence of signals $S_2 \equiv \{1, 1, 1, 0, 1, ..., 0, 1\}$ or for the sequence of signals $S_3 \equiv \{0, 0, 1, 0, 1, ..., 0, 1\}$. If $\omega = 1$, then $\Pr(S_2|\omega = 1) > \Pr(S_1|\omega = 1)$, and if $\omega = 0$, then $\Pr(S_3|\omega = 0) > \Pr(S_1|\omega = 0)$. Therefore, a cascade is more likely under perfect than under imperfect observability in both states of the world.

For part b, observe that, as shown above, the *ex ante* probability of a mistake under imperfect observability is $(1-q)(1-s)+\alpha s$, and, as it is easy to verify, $\alpha < 1-q$ for any finite n. For n tending to infinity, we know from Proposition 4 that $\alpha = 1-q$, which concludes the proof.■