Dynamic Equilibrium Bunching

Tao Wang
Queen's University

Department of Economics
Queen's University
94 University Avenue
Kingston, Ontario, Canada
K7L 3N6

11-2011

Dynamic Equilibrium Bunching

Tao Wang*

Queen's University

November 2011

Abstract

In this paper, we analyze the asymmetric pure strategy equilibria in a dynamic game of pure information externality. Each player receives a private signal and chooses whether and when to invest. In some of the periods, only a subgroup of the players make decisions, which we call bunching, while the rest of the players do not invest regardless of their signals. Bunching is different from herding; it occurs in the first period and recursively until herding takes place or the game runs out of undecided players. We find that any asymmetric pure strategy equilibrium is more efficient than the symmetric mixed strategy equilibrium. When players become patient enough, herding of investment disappears in the most efficient asymmetric pure strategy equilibrium, while the least efficient asymmetric pure strategy equilibrium resembles those in a fixed timing model, producing an exact match when the discount factor is equal to 1. Bunch sizes are shown to be independent of the total number of players; adding more players to the game need not change early players' behavior. All these are unique properties of the asymmetric pure strategy equilibria. We also show that the asymmetric pure strategy equilibria can accommodate small heterogeneities of the players in costs of acquiring signals, discount factors, or degree of risk aversion. In any of these environments, there exists a unique welfare maximizing equilibrium which provides a natural way for the players to coordinate.

Keywords: bunching, herding, endogenous timing, asymmetric equilibrium, information externality

JEL Classification numbers: C73, D82, G01

1 Introduction

Frequently in life, firms or individuals need to make Yes or No decisions under uncertainty. For example, one needs to decide on whether or not to purchase a newly-released cell phone, whether or not to go to the cinema to watch a new movie, or whether or not to make an investment. These decisions involve risk, since the outcomes are uncertain.

In this kind of environment, an agent could benefit from the observations of other agents' actions, as those actions could reveal information which is useful to the agent. The existing literature has focused on herd behavior; when later investors observe many investments (or no-investments) made by earlier investors, they ignore their own private information and follow suit. An important setup in a typical herding model is that players make their choices according to a fixed order; they must act in their exogenously determined time slot, no sooner and no later. We categorize this type of models as the fixed timing models.[1]

Chamley and Gale (1994) are the first to employ an endogenous timing model to investigate the optimal investment issue. They examine a model where an unknown number of investment options are held by a certain number of players, and the expected payoff of each individual depends on the total number of options. They focus on a symmetric mixed strategy equilibrium of that model.

In a fixed timing model, agents make decisions one by one. In a symmetric pure or mixed strategy equilibrium of an endogenous timing model, all agents use the same strategy and act simultaneously. The intermediate case, where only a subgroup of agents optimally make decisions at one time, has not been studied.

We investigate this phenomenon in this paper. In our endogenous timing model of pure informational externality, players choose whether and when to invest. A player possesses private information (good signal or bad signal) which could be revealed to other players through his action. Therefore, a player must balance the effect of discounting and the incentive to wait for more information to be revealed. We find that in the asymmetric pure strategy equilibria of our model, players make decisions in subgroups, until herding occurs.

More specifically, in these asymmetric equilibria, starting from the first period, only a subgroup of players make decisions. The rest of the players (who have not yet invested) just wait and observe the actions of these players. We term this behavior "bunching", and the subgroup of players making decisions in a given period a "bunch". When players bunch, they make their decisions at the same time (but their decisions may not be the same). When players herd, they make the same decision regardless of their private information.

If enough of the players in a given bunch invest, revealing that they have received good

---

[1]See Banerjee (1992) and Bikhchandani, Hirshleifer and Welch (1992) for the pioneer work on the fixed timing models.

signals, then herding of investment occurs, meaning that the rest of the players will invest in the next period. If enough players in the bunch do not invest, then herding of no-investment occurs, meaning that the rest of the players will never invest in any future periods. Lastly and importantly, if only an intermediate number of players invest, then another bunch will form in the next period. This pattern will continue recursively until there are no undecided players left in the game.

These asymmetric pure strategy equilibria have many unique properties. First, bunching occurs in all of these asymmetric equilibria. It occurs at least in the first period. Bunch sizes are uniquely and endogenously determined. A larger bunch reveals too much information and even a bunch member would like to wait. A smaller bunch reveals too little information and a player outside the bunch does not want to wait. These incentives uniquely pin down the equilibrium bunch size in any given period.

Second, each of these asymmetric equilibria is more efficient than the symmetric mixed strategy equilibrium.[2] Players using mixed strategies can only achieve the payoffs of investing in the period they randomize. This means that the benefits of investing later (from more revealed information) are completely canceled by the discounting. This is not the case in the asymmetric equilibria. Players outside a bunch can benefit from the information revealed in the bunch and earn higher payoffs in the equilibrium.

Third, when players are patient enough, the least efficient asymmetric equilibria produces outcomes resembling those in a fixed timing model. In the equilibrium, only one player makes a decision in each period (i.e., the bunch size is one in every period) until herding occurs. This one-per-period decision making is an equilibrium result in our model, while it is assumed in a fixed timing model. A symmetric mixed strategy equilibrium cannot produce such outcomes because of the randomness in the outcomes.[3]

Fourth, in the most efficient equilibrium, bunching occurs most often and herding occurs least often. In this equilibrium, as long as herding on no-investment does not occur, players make their investment decisions one by one according to their private signal, and herding on investment completely disappears. The information revelation is maximized in this way, and the equilibrium outcome is more efficient than the one obtained from the fixed-timing model, where herding occurs most often. These results are specific to the asymmetric pure strategy equilibria. A symmetric mixed strategy equilibrium cannot be easily linked to the outcomes of a fixed-timing model.

Finally, the asymmetric pure strategies in the equilibrium for all players (except the last bunch if undersized) are upward independent of the number of players. Consider the asymmetric equilibrium strategies of the players when there are n players in the game. Then

---

[2]We need to forbid investment from the players with bad signals to make our model comparable to Chamley and Gale (1994) in this comparison.

[3]The symmetric mixed strategy equilibrium we refer to in this paper is the one characterized in Chamley and Gale (1994), unless otherwise specified.

the strategies of these (old) players (except the last bunch if undersized) can be used to construct a new asymmetric pure strategy equilibrium in a game with more than n players (when additional players are added in), with the old players' strategies and information updating exactly the same as before. This property allows us to apply the asymmetric pure strategies to situations where new players may arrive in later periods in the game. Characterizing a symmetric mixed strategy equilibrium would be an immense task in such situations.

It is interesting to note how information is revealed in a bunch. Knowing the bunch size, one can deduce from the actions of the players in the bunch how many of them received good signals and how many did not, since players with different signals will act differently in a bunch. Seeing someone invest in a bunch is the same as seeing a good signal, while seeing someone not invest in a bunch is the same as seeing a bad signal. Players not in a particular bunch do not invest in the given period regardless of their signals. Therefore, seeing just a few players invest in a period does not necessarily mean that there are many bad signals among the players. For example, suppose there are 1000 players in total and the size of the first bunch is 5. Then if 5 out of the 1000 players invest in the first period, it is very good news to other players. Note that the players do not need to know the identities of the players in each bunch. As long as they know the size of the bunch, they can infer the numbers of good signals and bad signals in a given bunch by just knowing the number of investments in that given period.

We compare the social welfare in our equilibria to the socially optimal level. Analysis of the general model is unmanageably complicated, but in a two-period version of our model, we show that the bunching sizes are smaller in the asymmetric equilibria compared to the socially optimal level.

In this paper, we characterize both asymmetric and symmetric pure strategy equilibria, even though we focus on the former. We are able to provide closed form solutions for the equilibrium strategies. When the number of players is small enough or when players are impatient enough, a unique symmetric pure strategy equilibrium exists. In this equilibrium, every player with a good signal invests right away. In all other cases, asymmetric pure strategy equilibria exist.

Our analysis is robust to small perturbations in signal acquisition costs. The equilibria characterized in the paper are still valid if private signals are costly and players have the choice of not acquiring their signals. The equilibrium strategies in our model are also robust to small perturbations in players' discounting factors, as well as to introducing some degree of player risk aversion. We can rank the asymmetric equilibria according to the social welfare they generate in these cases, and there is a unique asymmetric equilibrium that maximizes the social welfare. This equilibrium serves as a natural coordination for the players.

The results in our paper have a wide range of applications. Consider a new product of unknown quality. Some consumers may decide early on whether or not to buy it (i.e.,

bunching), while other consumers wait and see. When the sales figures are very good (relative to the equilibrium bunch size), a huge wave of consumers may subsequently buy it. On the contrary, if the sales figures are really bad (again, relative to the equilibrium bunch size), no one would buy it anymore. If the sales figures are just moderate, then another group of consumers may decide whether or not to buy the new product (i.e., bunching again), and so on and so forth. Another application is in investment decisions. When the economy is uncertain, investors need to see many signals to be reasonably sure about the state of the economy. In this case, we often observe some early decision makers, big investment banks, who decide on whether or not to invest early, depending on the information they receive at the time. Other investors just stay on the sideline. Then herding may occur, if many investors from this first bunch invest (or not invest). When neither of the sides in the bunch dominates (in a sense to be made clear in the analysis), another bunch of investors forms, etc.

The literature on herding in economics was started by Banerjee (1992) and Bikhchandani, Hirshleifer and Welch (1992), with the latter referring to herd behavior as information cascade, as people no longer reveal their own private information. In a herding equilibrium, the equilibrium outcome depends crucially on the private information of the first few individuals, and only these first few players' private signals are revealed in the equilibrium, leading to socially inefficient outcomes.[4] Note that these are fixed timing models.

Rosenberg, Solan and Vieille (2007) and Zhang (2009) have different models of endogenous timing. In Rosenberg, Solan and Vieille (2007), players play a one-armed bandit and continuously receive private signals. Zhang (2009) generalizes Chamley and Gale (1994) to both continuous signals and multiple discrete signals.[5] Zhang (1997) presents another endogenous timing model where players with different precision of signals choose their optimal timing to act in continuous time. His paper shows that the player with the highest precision invests first and then there is an investment surge leading to an information cascade. Similar results are also found in Aoyagi (1998) in a model with a different setup. All of these papers focus on symmetric equilibria. In this paper, we focus on asymmetric pure strategy equilibria and highlight their different properties.

This paper is organized as follows. In Section 2, we describe the model and its assumptions. In Section 3, we perform some preliminary analysis, including defining the myopic

---

[4]Herding and information cascades have been investigated by many researchers. Cao, Han and Hirshleifer (2011) show that even if players can communicate with each other regarding their payoffs from past choices, herding can still occur, but cannot last forever. Khanna and Mathews (2011) study a model in which players' actions may affect each other's payoffs. They find that a leader may spend more to get more precise information to avoid the adverse effect of actions being copied, and this information gain can dominate the information loss from herding.

[5]Herd behavior in models of endogenous timing has been confirmed in experiments. (See Sgroi (2003) for a review.) In an empirical study, Moretti (2011) examines the behavior of movie goers and finds that peer effect (in the form of information externality) plays an important role in their movie watching decisions.

best response function, defining the patterns of equilibrium behavior, and determining the bunch sizes. In Section 4, we characterize the symmetric pure strategy perfect Bayesian equilibria (PBE) and the asymmetric pure strategy PBE. In Section 5, we perform comparative static analysis regarding the discount factor, the probability of a good state and the number of players. In Section 6, we perform a welfare analysis. We compare our asymmetric pure strategy equilibria to the symmetric mixed strategy equilibrium in Chamley and Gale (1994). We analyze the socially optimal strategies in a two-period version of our model. We also analyze the socially optimal strategies when there are heterogeneous players. In Section 7, we discuss some of the applications of our model. In Section 8, we conclude. In the appendix, we analyze the equilibria when signals are costly. We also characterize the equilibria when the assumptions in the paper are violated. Finally, the appendix contains lengthy proofs.

2 The Model

Consider a model with n players. Each player decides whether to make an investment. The action space for each player is given by $A = \{0, 1\}$, where 0 stands for no investment, and 1 stands for investment. The outcome of the investment depends on the state of the economy. The state space is denoted as $S \in \{G, B\}$. The payoff from investment is π_G when the state is G and is π_B when the state is B, where $\pi_G > 0 > \pi_B$. Anyone who never invests earns 0. With prior probability q the state is G, and with prior probability $1 - q$ the state is B.

There are T periods starting from period 1 in this game. Assume that $T \geq n + 1$ as the game will end in at most $n+1$ periods.[6] Each player can invest only once. If a player invests in a period, he gets his payoff in that period, and has no further decisions to make. So in any period t, only those players who have not invested in an earlier period make investment decisions. These players can observe other players' actions (but not the payoffs) in previous periods (from period 1 to $t - 1$). Players are impatient and their payoffs are discounted by a common discount factor $\delta \in [0, 1]$.

The state of the economy remains the same throughout the game. Before period 1, each player receives a private signal $s \in \{g, b\}$, which is correlated with the state of the economy. A player's action may reveal his private signal, which is useful to other players in predicting the state. However, one player's action does not affect another player's payoff directly. Therefore, this is a game of pure information externality.

A player's signal is distributed independently and identically across all the players, with the following probabilities conditional on the underlying state:

---

[6]As we will see in the analysis, in each period at least one player will reveal his signal. So we need at most n periods for the signal revelation, plus one more period for any player who has not yet invested to make a final investment decision.

$$\Pr(s = g | S = G) \equiv \Pr(g|G) = p > \frac{1}{2},$$

$$\Pr(s = b | S = G) \equiv \Pr(b|G) = 1 - p,$$

$$\Pr(s = b | S = B) \equiv \Pr(b|B) = p,$$

$$\Pr(s = g | S = B) \equiv \Pr(g|B) = 1 - p.$$

Here, we assume that the signal precision (i.e., p) is the same in state G or B for simplicity. Because of the symmetry in signal precision, we have

$$\Pr(S|g, b) = \frac{\Pr(g, b|S)\Pr(S)}{\Pr(g, b|S)\Pr(S) + \Pr(g, b|S')\Pr(S')} = \frac{p(1-p)\Pr(S)}{p(1-p)\Pr(S) + p(1-p)\Pr(S')}$$

$$= \Pr(S),$$

where $S = G, B$ and $S' = B, G$; that is, a good signal and a bad signal cancel out each other's effect. Similarly, it is straight-forward to show that the probability of a good state conditional on m_1 good signals and m_2 bad signals can be simplified as

$$\Pr(S|m_1 g, m_2 b) = \begin{cases} \Pr(S|(m_1 - m_2)g), & \text{if } m_1 > m_2, \\ \Pr(S), & \text{if } m_1 = m_2, \\ \Pr(S|(m_2 - m_1)b), & \text{if } m_1 < m_2. \end{cases} \tag{1}$$

The above expressions state that the probability of a state S given m_1 good signals and m_2 bad signals is equivalent to $m_1 - m_2$ net good signals if $m_1 > m_2$, etc. Each good signal cancels out with a bad signal, because a good signal has the same precision as a bad signal.[7]

We focus on pure strategy equilibria in the analysis. Obviously, the actions of the players may reveal their signals. For example, if a player's strategy in a certain period is to invest when his signal is g and not to invest when his signal is b, then other players can completely infer his signal from whether or not he has invested in that period. Of course, if his strategy is to invest in that period regardless of his signal, then other players gain no information from his action. Therefore, a player's action either reveals his signal completely or reveals nothing at all. Note that this kind of updating is appropriate only for pure strategy equilibria.

In the analysis, the total number of good signals and the total number of bad signals known by a player at the beginning of period t, denoted by $(N_{g,t}, N_{b,t})$, play an important role in deciding that player's action. This pair of numbers includes both the public information at the beginning of period t (i.e., the total number of good and bad signals inferred from the actions of other players in periods 1 to $t-1$) and the player's private signal. Note also that this pair of numbers is different from player to player, since each player may have different private signal. We shall omit the index for the player since it will not cause any confusion

---

[7]When the precisions are different, the Bayesian updating is more complicated, but the analysis is similar.

in later analysis. In what follows, we use both $\Pr(S|N_{g,t}, N_{b,t})$ and $\Pr(S|N_{g,t}g, N_{b,t}b)$ interchangeably (depending on which one is clearer in the particular context) to refer to the probability of state S conditional on the player's information at the beginning of period t. The expected payoff for a player playing action $a \in A$ in period t given $(N_{g,t}, N_{b,t})$ is written as $\pi(a|N_{g,t}, N_{b,t})$.

In this paper, we will focus on the following assumptions:

$$\pi(1|g) = \Pr(G|g)\pi_G + \Pr(B|g)\pi_B > 0, \tag{A1}$$

and

$$\pi(1|b) = \Pr(G|b)\pi_G + \Pr(B|b)\pi_B < 0. \tag{A2}$$

With these assumptions, without learning any other information, it is profitable for a player to invest if his private signal is g and not to invest if his private signal is b.

We further assume that

$$\pi(1) = \Pr(G)\pi_G + \Pr(B)\pi_B < 0; \tag{A3}$$

that is, unconditional on any signal, it is not profitable to invest. If $\pi(1) \geq 0$, the results are comparable and will be stated in a separate theorem in the appendix. Note that (A3) implies (A2).

We will discuss the cases when these assumptions do not hold in the appendix. It is shown that the equilibrium strategies are either parallel or trivial.

3   Preliminary Analysis

3.1   Myopic Best Response

Define $BR(N_{g,t}, N_{b,t}) \in A = \{0, 1\}$ as the myopic best response function for player i in period t. Assuming that player i has not yet invested in previous periods, given his information $(N_{g,t}, N_{b,t})$, then restricting attention to the current period t, the best action for him is

$$BR(N_{g,t}, N_{b,t}) = \begin{cases} 1, & \text{if } N_{g,t} - N_{b,t} \geq 1, \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

Of course, a player may or may not always follow this myopic best response function, as not investing in period t (and thus getting payoff 0 in period t) does not mean that he will not invest in a later period. However, we will show in the proof of Theorem 2 in the appendix, this myopic best response is in fact optimal for a deviating player.

Given player i's information $(N_{g,t}, N_{b,t})$ at the beginning of period t, if there are k_t of new signals to be revealed in period t, then the total number of revealed signals is $k_t + N_{g,t} + N_{b,t}$ at the beginning of period $t+1$. Then player i's expected payoff from using the myopic best response function in period $t+1$ is

$$R(k_t|N_{g,t}, N_{b,t}) \;=\; \sum_{r=0}^{k_t} \pi\left((BR\left(N_{g,t} + r, N_{b,t} + k_t - r\right)|N_{g,t} + r, N_{b,t} + k_t - r\right)$$

$$\cdot \Pr\left(N_{g,t} + r, N_{b,t} + k_t - r|N_{g,t}, N_{b,t}\right),$$

where r is the number of good signals among the k_t new signals. This R function captures player i's expected payoff when he waits one period (expecting k_t additional signals to be revealed) and then uses his myopic best response. We have the following lemma.

Lemma 1 *Given* $(N_{g,t}, N_{b,t})$, $R(k_t|N_{g,t}, N_{b,t})$ *is strictly increasing in* k_t *and converges to* $R(\infty|N_{g,t}, N_{b,t}) \equiv \Pr(G|N_{g,t}, N_{b,t})\pi_G$.

Proof See appendix. ■

This lemma illustrates that more signals are better for a player. This is because more signals will help a player to make a more precise decision. When there are infinitely many signals, by the Law of Large Numbers, the proportions of good signals and bad signals will be, respectively, exactly the same as the probabilities of receiving a good signal and a bad signal conditional on the true state. Therefore, by examining the proportions of good signals and bad signals, a player can tell exactly which state is the true underlying state. If the underlining state is revealed to be G, the player invests, and if the underlining state is revealed to be B, the player does not invest.

3.2   Patterns of Equilibrium Behavior

As will become clear later, only the following four patterns of behavior will occur in a given period in a pure strategy PBE. Specifically, in a given period, if only a subgroup of players make decisions on whether or not to invest, while all other players do not invest regardless of their signals, then we call it bunching, and the subgroup a bunch. When bunching, every player in the bunch makes a decision in the *same period*. In equilibrium, players could also herd. When herding, every player who herds makes the *same decision* regardless of his signal.[8]

---

[8]Different from the herd behavior that leads to information cascades in our model, Scharfstein and Stein (1990) examine reputational herding, where smart types receive perfectly correlated signals and dumb types receive random signals. Players herd so as to build a reputation for being a smart type. Gul and Lundholm

Definition 1 *In a given period,* **Herding of Investment** *occurs when every player (regardless of his signal) who has not invested invests now.*

Definition 2 *In a given period,* **Herding of No-Investment** *occurs when every player (regardless of his signal) who has not invested does not invest. No further investment will occur after this period.*

Definition 3 *In a given period,* **Bunching of Size** k *occurs when a pre-specified group (i.e., bunch) of k players selected from those who did not belong to a previous bunch make decisions simultaneously according to their own signals – invest if the signal is g and not invest if it is b.*

Definition 4 *In a given period,* **the Last Bunching** *occurs when there are not enough players left to form a bunch of the desired size. All players who did not belong to previous bunches make decisions in this period.*

In the definition for the last bunching above, the decisions made by the bunch members could be different from those made by a player in a regular bunch. More specifically, a player with a bad signal does not invest in a regular bunch. But he may invest in the last bunch if it is profitable to do so. (In this case, it is both herding and bunching.) Meanwhile, a player with signal g always invests in a bunch, last period or not. (Refer to Lemma 7 for their behavior.)

In the asymmetric equilibria we will characterize, either Herding of Investment, Herding of No-Investment, or Bunching may occur in a particular period, and Last Bunching may occur only at the end of the game. Bunching will definitely occur in period 1, and may occur again in later periods if bunching occurred in all previous periods.

3.3   Bunch Size Determination

In any period, given any history of the game, a player who has not yet invested needs to make an investment decision. He trades off the benefit of investing right away with the benefit of waiting. Bunching occurs frequently in our equilibria and the size of bunching is the focus of this subsection. Recall that in a given period, bunching of size k occurs when k players

---

(1995) define the concept of clustering, a behavior that appears similar to an information cascade, but is in fact different. When clustering, players make similar decisions at almost (but not exactly) the same time, but players' private information is completely revealed. The definition for herding is slightly different in financial markets literature as it involves price movements. It is used by Park and Sabourian (forth.) to investigate informational herding and contrarianism in financial markets.

selected from those who did not belong to a previous bunch make decisions simultaneously according to their own signals – invest if the signal is g and not invest if it is b.

At the beginning of a given period, let d denote the difference between the number of good signals and the number of bad signals a player learned from other players' actions in previous periods. This d is common to every player, and is public information at the beginning of a given period. When d is negative, bad signals outnumber good signals, and no player will invest. In this case, further bunching can never occur. Therefore, we will focus on non-negative d in the following analysis, and we interpret a non-negative d as the degree of public optimism.

Suppose that a bunch of size k_d^* is to be formed in a given period. Then k_d^* must satisfy two conditions. The first condition is that a player inside the bunch (with either signal) does not want to be a non-bunch person. This means that if the player has signal g, he prefers to invest now instead of waiting one period; if the player has signal b, he prefers to wait one period instead of investing now. Equivalently, we have

$$\pi(1|d+1,0) \geq \delta R\left(k_d^* - 1|d+1,0\right), \tag{IC1}$$

$$\pi(1|d,1) \leq \delta R\left(k_d^* - 1|d,1\right). \tag{IC2}$$

The second condition is that a player outside the bunch (who has not previously invested, with either signal) does not want to deviate. This implies that regardless of his signal the player prefers to wait one period instead of investing now. Equivalently, we have

$$\pi(1|d+1,0) \leq \delta R\left(k_d^*|d+1,0\right), \tag{IC3}$$

$$\pi(1|d,1) \leq \delta R\left(k_d^*|d,1\right). \tag{IC4}$$

The following lemmas will help us to establish the existence and the properties of k_d^*.

Lemma 2 *For any given k, $\pi(1|d,0) - \delta R\left(k|d,0\right)$ is increasing in d.*

Proof See appendix. ∎

This lemma implies that the marginal benefit of additional k signals is decreasing in the number of net good signals already obtained. When there are more net good signals, the additional k signals become less valuable. Therefore, if a player with more net good signals finds it optimal to wait for k additional signals, it must be optimal for a player with fewer net good signals to wait as well.

The following two lemmas consolidate the incentive compatibility constraints.

Lemma 3 *(IC3) implies (IC4).*

From Lemma 2, this lemma is immediate.

Lemma 4 *(IC1) and (IC3) imply (IC2).*

Proof By Lemma 2 and (IC3), we have

$$0 \geq \pi(1|d+1,0) - \delta R\left(k_d^*|d+1,0\right) > \pi(1|d,0) - \delta R\left(k_d^*|d,0\right).$$

Also,

$$\pi(1|d,0) = \pi(1|d,1)\Pr(b|d,0) + \pi(1|d+1,0)\Pr(g|d,0),$$

and

$$R\left(k_d^*|d,0\right) = R\left(k_d^* - 1|d,1\right)\Pr(b|d,0) + R\left(k_d^* - 1|d+1,0\right)\Pr(g|d,0).$$

From (IC1) and $\pi(1|d,0) < \delta R\left(k_d^*|d,0\right)$, it must be that $\pi(1|d,1) < R\left(k_d^* - 1|d,1\right)$. This completes the proof. ■

This lemma means that if it is optimal for a player with signal g to wait for k_d^* signals but not $k_d^* - 1$ signals, then it is optimal for a player with signal b to wait for $k_d^* - 1$ signals.

From the above two lemmas, (IC1), (IC2), (IC3) and (IC4) reduce to (IC1) and (IC3).

Suppose that $\pi(1|d+1,0) \geq \delta R\left(\infty|d+1,0\right) = \delta\Pr\left(G|d+1,0\right)\pi_G$. Then even if there are infinitely many signals to be revealed in this period, a player with signal g does not find it optimal to wait for these signals. In this case, any player with signal g will invest right away. We define $k_d^* = \infty$ in this case. It happens when players discount future payoffs a lot (i.e., δ is very small) or when signals are already very precise (i.e., p is very close to 1).

Now suppose that $\pi(1|d+1,0) < \delta R\left(\infty|d+1,0\right)$. In this case, as we shall show in the following lemma, there exists a unique and finite k_d^* satisfying (IC1) and (IC3). Note that (IC1) and (IC3) cannot be binding at the same time.

Lemma 5 *Suppose that $\pi(1|d+1,0) < \delta\Pr\left(G|d+1,0\right)\pi_G$. Then for each $d \geq 0$, there exists a unique k_d^* satisfying (IC1) and (IC3); that is,*

$$\delta R\left(k_d^* - 1|d+1,0\right) \leq \pi(1|d+1,0) \leq \delta R\left(k_d^*|d+1,0\right). \tag{3}$$

Suppose that $\pi(1|d+1,0) \geq \delta\Pr\left(G|d+1,0\right)\pi_G$. Then there exists no k_d^* satisfying (IC1) and (IC3). In this case, $k_d^* = \infty$.

Proof From Lemma 1, $R\left(k|d+1,0\right)$ is increasing in k. Since $\pi(1|d+1,0) < \delta R\left(\infty|d+1,0\right)$, the curve of $\delta R\left(k|d+1,0\right)$ intersects the line $\pi(1|d+1,0)$ only once and at $k \geq 1$. (c.f. Figure 1.)
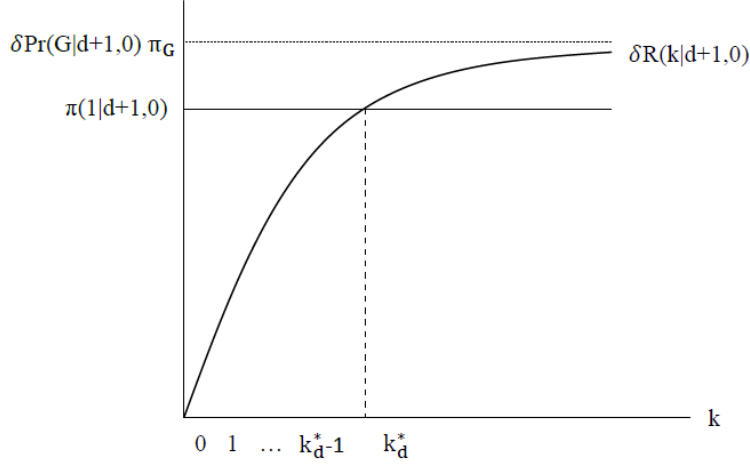
Figure 1: The determination of k_d^*

If $\pi(1|d+1,0) \geq \delta R(\infty|d+1,0) = \delta \Pr(G|d+1,0)\pi_G$, then it is obvious that no k can satisfy (IC1) and (IC3). In this case, a player with signal g would not wait for any finite number of signals to be revealed. In this case, $k_d^* = \infty$. ∎

The k_d^* uniquely determined in the above lemma is increasing in d, the number of net good signals at the time. The following proposition shows that when there are more net good signals, the bunch size has to be larger. We have

Lemma 6 $k_0^* \leq k_1^* \leq k_2^* \leq k_3^* \leq \cdots$.

Proof Since $\pi(1|d,0) - \delta R(k|d,0)$ is decreasing in k, and by Lemma 2, an increase in d shifts the curve $\pi(1|d,0) - \delta R(k|d,0)$ upwards, resulting in a bigger k. Note here the inequalities are not strict because of the discreteness of k. ∎

One way to interpret this lemma is that the situation is more optimistic when there are more net good signals. Therefore, players have more incentive to invest now. In order for the players outside the bunch to wait, the bunch size (and thus the number of signals to be revealed) must be larger than when the number of net good signals is small. From the proof, we can see that when d is larger, (IC1) is not affected; however, (IC3) becomes more difficult to satisfy. When (IC3) is violated, one must increase k to meet the requirement of (IC3).

The above k_d^* is the number of additional signals such that a player inside the bunch does not wait while a player outside the bunch prefers to wait. However, near the end of the game, the number of players left may be greater or smaller than this k_d^*. Let $\tilde{n}_t$ denote at the beginning of period t the number of players who did not belong to any previous bunches

in the game. Let d_t be the difference between the numbers of good signals and bad signals inferred from other players' actions at the beginning of period t (i.e., from period 1 to period $t-1$). The determination of the bunch size above assumes implicitly that $\tilde{n}_t \geq k^*_{d_t}$. When $\tilde{n}_t < k^*_{d_t}$, that is, when the number of players not belonging to any previous bunch is not large enough to form a new bunch of the desired size $k^*_{d_t}$, these remaining players still form a (last) bunch. But the players in this last bunch may behave differently from a regular bunch. We have the following lemma.

Lemma 7 *Suppose that in period t, $\tilde{n}_t < k^*_{d_t}$. Then these $\tilde{n}_t$ players will form the last bunch of the game.*

(a) A player with signal g in the last bunch will invest in period t.

(b) A player with signal b in the last bunch will invest in period t if and only if $\pi(1|d_t, 1) > \delta R(\tilde{n}_t - 1|d_t, 1)$.

(c) A player with signal b in a previous bunch will invest in period t if and only if $\pi(1|d_t, 0) > \delta R(\tilde{n}_t|d_t, 0)$.

The game will end in period t if both groups of players in (b) and (c) invest in period t. Otherwise, those players in (b) and/or (c) will make a final investment decision in period $t+1$ and the game ends.

Note that in the above lemma, if the group in (b) invests, then the group in (c) will definitely invest, since $\pi(1|d_t, 1) > \delta R(\tilde{n}_t - 1|d_t, 1)$ implies $\pi(1|d_t, 0) > \delta R(\tilde{n}_t|d_t, 0)$. Note also that if $d_t = 0$ or 1, the condition in (b) never holds, since $\pi(1|d_t, 0) < 0$ and $\delta R(\tilde{n}_t|d_t, 0) \geq 0$. Therefore, the players in the last bunch still act according to their signals; i.e., a player with signal g invests and a player with signal b does not invest. If $d_t \geq 2$, then it is possible that a player with signal b in the last bunch will invest right away. This is because investing right away brings in positive expected payoff, while waiting one period for just a few extra signals does not bring enough gain to offset the discounting. This happens when $\tilde{n}_t$ is relatively small.

In equilibrium, Herding of Investment occurs when the actions in previous periods indicate that good signals outnumber bad signals by a sufficiently large number. Herding of No-Investment occurs when the opposite happens, i.e., the actions in previous periods indicate that bad signals sufficiently outnumber good signals. In short, herding occurs when the probability of either state (G or B) becomes sufficiently large. Bunching occurs when the situation is not as certain. It is clear that the actions of the players in a bunch (except maybe the last one) will reveal their signals completely. Meanwhile, herding will reveal none of the players' signals.

4 Equilibrium Analysis

4.1 Symmetric Pure Strategy PBE

In this subsection, we shall characterize the conditions under which a unique symmetric pure strategy PBE exists. This equilibrium exists whenever the number of players (n) in the game does not exceed the size of bunching in period 1 (i.e., k_0^*). If the size of bunching is infinite in period 1, i.e., $k_0^* = \infty$, then the game (with any finite number of players) always has a unique symmetric pure strategy PBE. We have the following theorem.

Theorem 1 *Suppose that (A1) and (A3) hold. Then there exists a unique equilibrium if and only if $n \leq k_0^*$, or equivalently,*

$$\pi(1|g) \geq \delta R\left(n - 1|g\right). \tag{4}$$

Furthermore, this equilibrium is in symmetric pure strategies. Bunching of size n occurs in period 1. If investments strictly outnumber no-investments in the bunch, then in period 2, any player who did not previously invest will invest; otherwise, none of those players will invest. No further investment will occur after period 2.

Proof The right hand side of the inequality (4) is the discounted expected payoff for a player with a good signal if he deviates by waiting and using the myopic best response in period 2. The benefit by doing so is that he can then observes $n - 1$ other players' actions in period 1.

The inequality (4) implies that a player with a good signal finds it more profitable to invest in period 1 than waiting until period 2. A player with a bad signal will only invest if the good signals outnumber the bad signals by at least 1, as implied by inequality (A1). ∎

Note that inequality (4) will not be satisfied for $n > k_0^*$. This is because the expected gain from seeing all other $n - 1$ players' actions (and thus signals) is increasing in the total number of players. Therefore, when n becomes large enough, a player with signal g has enough incentive to deviate from the symmetric strategy of investing in period 1 so that he can best respond to other players' signals in period 2.

We have the following two corollaries; their proofs are straight-forward given (4). The first corollary is about the number of players. When there are more players in the game, a player may have enough incentive to wait when all the other players make decisions in period 1.

Corollary 1 *Suppose that (A1) and (A3) hold. Given δ, there exists an n^*, such that the symmetric pure strategy equilibrium exists if and only if $n \leq n^*$.*

Fixing the total number of players, when players become more patient, the incentive for waiting one period is stronger, and thus it is more difficult for a symmetric pure strategy equilibrium to exist.

Corollary 2 *Suppose that (A1) and (A3) hold. Given the total number of players n, the symmetric pure strategy equilibrium exists if and only if $\delta \leq \bar{\delta}(n) \equiv \pi(1|g)/R(n-1|g) < 1$.*

4.2   Asymmetric Pure Strategy PBE

In this section, we characterize the equilibrium when $n > k_0^*$, or equivalently,

$$\pi(1|g) < \delta R(n-1|g). \tag{5}$$

From the last section, we know that in this case a symmetric pure strategy PBE does not exist, so we are looking for the asymmetric pure strategy PBE. In an asymmetric equilibrium, players with the same signal may act differently.

We characterize the asymmetric equilibrium in the following theorem.

Theorem 2 *Suppose (A1), (A3) and (5) hold. There exists an asymmetric pure strategy PBE, which has the following recursive properties. Let $d_1 = 0$. In period t, $t = 1, 2, ...$*

(a) if $d_t > 1$, Herding of Investment occurs, and the game ends;

(b) if $d_t < 0$, Herding of No-Investment occurs, and the game ends;

(c) if $d_t = 0$ and $\tilde{n}_t \geq k_0^*$, then Bunching of size k_0^* occurs;

(d) if $d_t = 1$ and $\tilde{n}_t \geq k_1^*$, then Bunching of size k_1^* occurs;

(e) if $d_t = 0$ and $\tilde{n}_t < k_0^*$, or if $d_t = 1$ and $\tilde{n}_t < k_1^*$, then the Last Bunching of size $\tilde{n}_t$ occurs.

Proof See appendix. ∎

In (c), (d), and (e) of the above theorem where bunching occurs, players with signal b do not invest. In (e), the final bunch is formed, and it is undersized. From this theorem, we know that in the first period, Bunching of size k_0^* forms since $d_1 = 0$. In the second period, if $d_2 > 1$, that is, $d_2 \geq 2$, then even a player with signal b will find it profitable to invest since his belief is given by at least one net good signal. Therefore, by assumption (A1), the expected payoff from investing with (at least) one good signal is positive. If $d_2 < 0$, that is, $d_2 \leq -1$, then even a player with a good signal does not invest since his belief is given by at most zero net good signals. From assumption (A3), a player's payoff is negative with a belief of zero net good signals.

This equilibrium has the following property. In any period, players outside the bunch do not need to know the identities of the players in previous bunches in order to infer the total numbers of good signals and bad signals in those bunches. The total number of players who have previously invested gives them the total number of good signals among the previous bunches; by subtracting this number from the total size of all previous bunches, they can then infer the total number of bad signals. For example, suppose that the size of the first bunch $k_0^* = 5$ and 3 players invested in the first period. Then a player need not know who these 3 players are; he can infer right away that there are 3 good signals and the rest of the bunch (i.e., 2 players) have bad signals. We have the following corollary.

Corollary 3 *In any period of the game, a player need not know the identities of other players; he need not know who has invested and who has not, as long as he knows how many players have invested and whether or not he belongs to a particular bunch.*

Depending on k_0^* being even or odd, either $d_2 = 0$ or $d_2 = 1$ can happen, but not both. There are three different bunching patterns which are described in the following corollaries.

Corollary 4 *Suppose that k_0^* is even. Then the bunching size in any period is k_0^*.*

Since $d_1 = 0$, the bunching size in period 1 is k_0^*. If k_0^* is even, case (d) cannot occur. When cases (a) and (b) occur, the game ends as no further investments will take place in future periods. When case (c) occurs, the bunch size is k_0^* again. Therefore, the same pattern repeats recursively and the bunching size in each period is k_0^*.

Corollary 5 *Suppose that k_0^* is odd and k_1^* is even. Then the bunching size is k_0^* in the first period, and k_1^* in all subsequent periods.*

If k_0^* is odd, then in the second period, case (c) cannot happen. When case (d) happens, $d_2 = 1$ and bunching of size k_1^* forms. If k_1^* is even, the difference between the number of investments and non-investments in the second period must be even. Therefore, in the third period, only cases (a), (b) and (d) can occur. This second-period pattern repeats recursively and the bunching pattern is $k_0^*, k_1^*, k_1^*....$

Corollary 6 *Suppose that both k_0^* and k_1^* are odd. The bunching size is k_0^* in any odd period, and becomes k_1^* in any even period.*

If k_0^* is odd, then again in the second period, case (c) cannot happen and the bunch size becomes k_1^*. Given that the difference between the number of investments and non-investments in the second period must be odd, the cumulative difference between the number

of investments and non-investments becomes even again by the end of the second period. Therefore, in the third period, only cases (a), (b), and (c) can happen. This alternating pattern between k_0^* and k_1^* repeats recursively. The bunching pattern becomes $k_0^*, k_1^*, k_0^*, k_1^*$....

Theorem 2 and its corollaries imply that the information being revealed in any one period cannot be too large: either herding of all players (and thus no information is revealed) or bunching of fixed sizes (and thus only the players in the bunch reveal their signals) occurs. If too many players reveal their signals in one period, then any one of them will have a strong incentive to wait for the signal revelation before deciding whether or not to invest in the next period. At the same time, the information revelation cannot be too small either; otherwise no player would wait for the revelation. To summarize, the information revealed in any period should be just the right amount to sustain the described equilibrium. Note that the equilibrium bunching sizes k_0^* and k_1^* are independent of the total number of players n.

4.3   Other Asymmetric Pure Strategy PBE

The equilibria characterized in Theorem 2 are not the only asymmetric pure strategy equilibria in the game. There are other equilibria, which take similar forms but with more bunching. It turns out that when the difference between the numbers of good signals and bad signals revealed before period t is at least 2, i.e., $d_t \geq 2$, herding is not the unique equilibrium behavior; bunching may also occur in this situation. The equilibrium characterized in Theorem 2 has the least bunching; all other equilibria we will characterize in this subsection will have more bunching. We first have the following lemma.

Lemma 8 *For any fixed $\delta \in [0, 1)$, there exists a $\bar{d} > 0$, such that k_d^* is finite when $d_t \leq \bar{d}$, and k_d^* is infinite when $d_t > \bar{d}$.*

Proof   Given δ, k_d^* is finite only if

$$\pi(1|d_t + 1, 0) < \delta R\left(\infty|d_t + 1, 0\right),$$

that is, when

$$\Pr\left(G|d_t + 1, 0\right)\pi_G + \Pr\left(B|d_t + 1, 0\right)\pi_B < \delta \Pr\left(G|d_t + 1, 0\right)\pi_G.$$

Rearranging, we have

$$\pi_G + \frac{\Pr\left(B|d_t + 1, 0\right)}{\Pr\left(G|d_t + 1, 0\right)}\pi_B < \delta\pi_G.$$

Since $\frac{\Pr(B|d_t+1,0)}{\Pr(G|d_t+1,0)}$ is increasing in d_t and converging to 0, there exists an upper bound for d_t for the above inequality to hold. Define this upper bound as $\bar{d}$.

17

When the above inequality does not hold, $k_d^* = \infty$. (c.f. Lemma 5.) This concludes the proof of this lemma. ∎

We can now characterize the alternative asymmetric equilibria in the following theorem.

Theorem 3 *Suppose (A1), (A3) and (5) hold. Then all asymmetric pure strategy PBE have the following recursive properties. Let $d_1 = 0$. In period t, $t = 1, 2, ...$,*

(a) if $d_t < 0$, Herding of No-Investment occurs, and the game ends;

(b) if $d_t = 0, 1$, and $\tilde{n}_t \geq k_{d_t}^*$, then Bunching of size $k_{d_t}^*$ occurs;

(c) if $d_t = 2, ..., \bar{d}$, and $\tilde{n}_t \geq k_{d_t}^*$, then Bunching of size $k_{d_t}^*$ is one equilibrium behavior, and Herding of Investment is another equilibrium behavior;

(d) if $d_t = 0, 1$, and $\tilde{n}_t < k_{d_t}^*$, then the Last Bunching of size $\tilde{n}_t$ occurs;

(e) if $d_t \geq 2$, and $\tilde{n}_t < k_{d_t}^*$, then the Last Bunching of size $\tilde{n}_t$ is one equilibrium, and Herding of Investment is another equilibrium.

When $d_t < 2$, the strategies in this equilibrium are the same as those in Theorem 2. When $d_t \geq 2$, there are multiple equilibria in this "subgame". Bunching is one equilibrium, while herding is another equilibrium. When bunching occurs, it reveals the private signals of the players in the bunch. Meanwhile, herding reveals no private information. Therefore, the outcomes of the equilibria in Theorem 3 are more informative than those in Theorem 2.

5 Comparative Static Analysis

In this section, we will perform a comparative static analysis on the discount factor, the probability of having an underlying state G, and the number of players.

5.1 The Discount Factor

In this subsection, we investigate the effect of δ on the asymmetric equilibrium strategies. We compare our asymmetric equilibria with the equilibrium in the following fixed timing model. In this fixed timing model, only one pre-specified player makes a decision on investment in any one period. If a player does not invest in his specified period, he will have no other chance to invest.

We show that when players become infinitely patient (i.e., $\delta = 1$), the outcome in one of our equilibria resembles that in the fixed timing model. In the equilibrium, only one player makes a decision in each period (i.e., the bunch size is one); herding of investment occurs

when there are two or more net good signals, and herding of no-investment occurs when there are zero or negative number of net good signals. We also find that this is the least efficient equilibrium among all asymmetric pure strategy equilibria in this case (i.e., $\delta = 1$). In the most efficient equilibrium, players no longer herd on investment. Finally, we show that most of the above results can be obtained when δ is close to 1.

We first have the following lemma.

Lemma 9 *Suppose (A1), (A3) and (5) hold. For a given d, if δ increases, the bunching size k_d^* decreases.*

Proof From Figure 1, the curve of $\delta R\left(k|d+1,0\right)$ shifts upward as δ increases, intersecting $\pi(1|d+1,0)$ at a smaller k. ∎

An increase in δ means that the players become more patient and waiting is less costly. So in order for a more patient player in the bunch to invest right away (instead of waiting one period for more signals to be revealed), the number of signals to be revealed needs to be smaller. This can be seen from (IC1). Furthermore, (IC3) becomes easier to satisfy when δ increases.

Now we discuss the special case of no discounting, i.e., $\delta = 1$.

Lemma 10 *Suppose (A1), and (A3) hold. If $\delta = 1$, then $k_d^* = 1$, for every $d \geq 0$. In this case, bunching occurs for every $d \geq 0$ is an equilibrium.*

The proof for this lemma is straight-forward: the only value satisfying the bunch size determination inequalities (IC1) and (IC3) is $k_d^* = 1$. When $\delta = 1$, there is no cost in waiting one period for more revealed signals. In any given period, it is not an equilibrium that every player waits (i.e., not invest). It is also not an equilibrium that more than one player are in the bunch (to make the invest/not invest decision), since one of the bunch members would wait. The only equilibrium behavior is for the players to reveal their signals one per period until herding occurs. Nevertheless, there are still multiple asymmetric pure strategy equilibria. This is because when $d \geq 2$ in a given period, bunching with a bunch size of 1 is one equilibrium behavior, and herding is another equilibrium behavior. Theorem 2 and Theorem 3 characterize all of the asymmetric pure strategy equilibria.

In the equilibrium with the least bunching, its outcome nests the equilibrium outcome in the fixed timing model. Since $d_1 = 0$, bunching of size $k_0^* = 1$ occurs in period 1. If the bunch member does not invest and reveals that his signal is b, then $d_2 = -1$ and herding of no-investment occurs and the game ends. If he invests and reveals that his signal is g, then $d_2 = 1$. In this case, $k_1^* = 1$ and bunching of size 1 occurs in period 2. In period 2, if the revealed signal is b, then $d_3 = 0$ and bunching of size 1 occurs again in period 3. If

the revealed signal is g, then $d_3 = 2$ and herding occurs in this equilibrium. Since $\delta = 1$, players do not discount their future payoffs, players in herding can invest one player per period. This outcome corresponds exactly to the fixed timing model as in Banerjee (1992) or BHW (1992). In a fixed timing model, players make decisions sequentially according to a fixed order, and only one player can invest in each period. Therefore, our asymmetric pure strategy equilibria naturally include the equilibrium from the fixed timing model. A similar link from the symmetric mixed strategy equilibrium to the fixed timing model is not possible.

The above equilibrium has the least bunching (and thus the most herding) among all co-existing asymmetric pure strategy equilibria. Because the information revealed is the least, this equilibrium is the least efficient. We now consider the equilibrium with the most bunching. When $d_t \leq 1$, players' behavior is exactly the same as in the above equilibrium. However, when $d_t \geq 2$, bunching of size 1 occurs (instead of herding) and private signals are still being revealed. Herding of investment never occurs. The total amount of information revealed by this equilibrium is the most. Theorem 3 above characterizes this most efficient equilibrium as well as other less efficient equilibria.

The above results are summarized in the following proposition.

Proposition 1 *Suppose (A1), (A3) and (5) hold. If $\delta = 1$, we have the following results.*

(1) The asymmetric pure strategy equilibria described in Theorem 2 coincide with the equilibrium in the fixed timing model. In these equilibria, one player makes a decision in each period.

(2) The asymmetric equilibria described in Theorem 3 have less herding and are more efficient than the equilibrium described in Theorem 2. Therefore, all pure strategy asymmetric equilibria are at least as efficient as the equilibrium in the fixed timing model.

(3) The most efficient equilibrium has bunching (of size 1) in every period whenever $d \geq 0$. In this case, herding of investment never occurs.

The above properties are not specific to the case of $\delta = 1$. In fact, most properties go through as long as δ is close enough to 1. Define $\hat{\delta} \equiv \frac{\pi(1|n-2,0)}{R(1|n-2,0)} < 1$. As long as $\delta \geq \hat{\delta}$, the bunch size for any $d \geq 0$ is always 1. This is because $\delta \geq \hat{\delta}$ guarantees $k_{n-2}^* = 1$, and k_d^* is increasing in d. We have the following proposition.

Proposition 2 *Suppose (A1), (A3) and (5) hold. Suppose also that $\delta \in [\hat{\delta}, 1)$. Then $k_d^* = 1$ for any $d \geq 0$. In this case, (2) and (3) in Proposition 1 remain valid. In (1), the players' investment decisions and the resulting equilibrium outcome are still the same as in the fixed timing model. But in the asymmetric pure strategy equilibria in our model, the investment*

decisions are made earlier than in the fixed timing model where players have to wait for their pre-specified periods to make their decisions.

When $\delta < 1$, an equilibrium corresponding exactly to the fixed timing model cannot be generated, since players are no longer indifferent to investing in later periods. Herding of investment must now take place in the same period. Nevertheless, the least efficient equilibrium reveals exactly the same amount of information as the fixed timing model. Since players are not forced to wait for their turn to invest in our model, our asymmetric equilibria are more efficient than the equilibrium in the fixed timing model.

In the next subsection, we investigate the effect of q, the prior belief on the underlying state being G, on players' equilibrium behavior.

5.2   The Probability of a Good State

We first show that the bunch size is increasing in the probability of a good state, q.

Lemma 11 *If q increases, the bunching size k_d^* increases.*

Proof   See appendix.   ■

An increase in q implies that the underlying state is more likely to be good. So in order to induce players outside the bunch but with good signals to wait until the next period as required by the equilibrium condition, the number of signals to be revealed in that period must be larger.

Similarly to the last subsection, when q is close to zero, the bunching size for any $d = 0, ..., n-2$ shrinks to one. In this case, the equilibria in Proposition 2 continue to be valid. Therefore, the implications for efficiency and the comparison with the fixed timing model continue to hold. If q is relatively small or δ relatively large such that $k_0^* = k_1^* = 1$, and $k_2^* \geq n-2$, then the equilibrium outcome in an asymmetric equilibrium coincides with the fixed timing model. This is because when $d = 2$, all remaining (at most $n-2$) players with signal g would invest in either model. We have the following proposition.

Proposition 3 *Suppose (A1), (A3) and (5) hold. Suppose also that $k_d^* = 1$ for every $d = 0, ..., n-2$. Then (2) and (3) in Proposition 1 remain valid. If $k_0^* = k_1^* = 1$ and $k_2^* \geq n-2$, then the equilibrium outcome resembles the fixed timing equilibrium in the sense that any given player with signal g makes the same investment decision in the two models.*

Note that in the above proposition, players with signal b may make different decisions. In the fixed timing model, the first player does not invest in the first period if his signal is b,

and cannot invest later as he is not allowed to. In our model, however, he will invest (later) if it is revealed that the good signals outnumber the bad signals in the game.

5.3 Upward Independence of n

The strategies in an asymmetric pure strategy equilibrium for the players may be invariant to the arrival of new players. Suppose that there are at first n players. We claim that these players' asymmetric equilibrium strategies can be part of the new asymmetric equilibrium when the number of players increases to n'. In the new game with n' players, the old players continue to use their old strategies until the period when the old last bunch was formed. The new players are arranged to fill up the last bunch in the old game if it was originally undersized, and to form new bunches afterwards. This constitutes a new asymmetric equilibrium. In this new equilibrium, players from the old game make decisions first. We have the following proposition.

Proposition 4 *Suppose that there are n players and that (A1), (A3) and (5) hold. Suppose also that these players use the equilibrium strategies described in Theorem 2 (or Theorem 3). Suppose now new players are added and it becomes a game of $n' > n$ players. In this new game, there exists an asymmetric pure strategy PBE as is described by Theorem 2 (or Theorem 3) in which the players from the old game continue to use their old equilibrium strategies until the period the last bunch in the old game was formed.*

First note that if (5) holds for a smaller n, it will hold for a larger n. Second, in the new equilibrium, unless herding occurs (in which case the new players will herd with the old players), the new players do not make investment decisions until the period the last bunch in the old game is formed. Last, the strategies for the old players during the period when the last bunch in the old game was formed vary from case to case. In all cases, the strategies of the old players with signal g remain the same in the new game during the period when the last bunch in the old game is formed. In that period, a player with signal g in the last bunch would invest in the old game; he still finds it optimal to invest in the new game, even though the bunch size may have grown from an undersized one to a full size one. He does not want to wait (i.e., not invest) in that period, as the size of the bunch provides the right incentive for him to invest. The strategies for the old players with signal b in the last bunch, however, may change from investing to not investing if they follow the strategies in Theorem 3. Again, the full bunch size now provides incentive for them not to invest. (They will invest in the old game because the bunch was undersized and could not provide enough incentive for these players to wait.)

This property of the asymmetric pure strategy equilibria provides an easy way to handle situations where players may arrive later in the game. Earlier players are not affected by

these players until the old players are used up in the bunch forming process, since in the new equilibrium the new players do not make decisions until they become part of a bunch. (They do act if herding occurs, but the game is over then.) This property is valid for the asymmetric pure strategy equilibria only. It is obviously not valid for a symmetric mixed strategy equilibrium, since the probability of investment in the randomization would depend on the total number of players in the game.

Note that there are other equilibria in the new game. The new players could make decisions first, followed by the old players. Mixing the players in any order would also be an equilibrium.

We have also investigated the effect of signal precision p on the bunching sizes and found that the relationship is ambiguous. An increase in p makes additional signals more valuable and thus increases the incentive for a player to wait. At the same time, it makes existing signals more precise, and thus decreases the incentive to wait. The trade-off of these two effects is ambiguous.

6 Welfare Analysis

6.1 Comparison with Chamley and Gale (1994)

In this subsection, we will compare the welfare in a symmetric mixed strategy equilibrium in Chamley and Gale (1994) with an asymmetric pure strategy equilibrium in this paper. In Chamley and Gale (1994), a player cannot invest if he does not have an investment option. To facilitate a proper comparison, in this subsection only, we modify our model and do not allow any player with signal b to invest.

The welfare is defined as the sum of the total expected payoffs of all the players. Since a player with signal b cannot invest, his expected payoff is zero. Therefore, we only need to compare the total expected payoffs for players with signal g.

In a symmetric mixed strategy equilibrium, players with signal g are to mix between investing in period 1 and waiting. Therefore, their expected payoffs must be exactly equal to the payoff from investing in period 1, which is equal to $\pi(1|g)$. Therefore, the total ex ante expected payoffs, defined by W^{sm}, can be simply expressed as

$$W^{sm} = n \Pr(g) \pi(1|g). \tag{6}$$

Now consider the asymmetric pure strategy equilibria in our paper. The equilibria in Theorem 2 and Theorem 3 still apply when players with signal b are not allowed to invest. This is because these players never invest before all players with signal g have invested and

23

not allowing them to invest would not affect the information updating process for those players with signal g.

In this modified game, there are multiple equilibria, but they all start with bunching of size k_0^* in period 1. In this period, players outside the first bunch strictly prefer waiting to investing, implying that they have higher expected payoffs waiting for the k_0^* signals to be revealed. Investing in period 1 yields $\pi(1|g)$ for each player with signal g in the bunch. Waiting (i.e., not investing) for each player with signal g outside the bunch yields at least $\delta R(k_0^*|g)$, which can be obtained by using the myopic best response function. This myopic best response function is used by the bunch members in period 2 if bunching occurs. Again, a player with signal g outside the second period bunch would have strictly higher payoff. Following this logic, we can see that a player with signal g receives higher payoff if he is in a later bunch. The sum of the total expected payoffs in an asymmetric pure strategy equilibrium, denoted by $\underline{W}^{ap}$, satisfies the following:

$$\underline{W}^{ap} > n \Pr(g) \left[\frac{k_0^*}{n} \pi(1|g) + \frac{n - k_0^*}{n} \delta R(k_0^*|g) \right],$$

or,

$$\underline{W}^{ap} > \Pr(g) \left[k_0^* \pi(1|g) + (n - k_0^*) \delta R(k_0^*|g) \right]. \tag{7}$$

We have the following proposition.

Proposition 5 *Any asymmetric pure strategy equilibrium Pareto dominates the symmetric mixed strategy equilibrium.*

Proof As implied by (IC3), when $d = 0$, we have $\pi(1|g) \leq \delta R(k_0^*|g)$. Comparing (6) with (7), it is immediate that $\underline{W}^{ap} > W^{sm}$. Players in the first bunch have the same payoff as in the symmetric mixed strategy equilibrium. Players outside the first bunch have strictly higher payoff as in the symmetric mixed strategy equilibrium. This completes the proof. ■

Note that the right hand side of (7) is the sum of payoffs when the players outside the first bunch use the myopic best response function in the second period. As we mentioned earlier, if bunching occurs in period 2, some players would receive higher payoffs than $R(k_0^*|g)$, the payoff from the myopic best response. Whenever bunching occurs, players outside that bunch would get higher payoffs than players inside the bunch. The equilibrium characterized in Theorem 2 is the worst asymmetric pure strategy equilibrium in terms of efficiency and discounted payoffs, since it has the least bunching. Nevertheless, it is still more efficient than the symmetric pure strategy equilibrium.

6.2 Welfare Analysis for $T = 2$

In the analysis until now, we assume that the time horizon is long enough, i.e., $T \geq n + 1$. In Theorem 1, players adopt the symmetric pure strategies, and the game ends in at most 2

periods. In Theorem 2 and Theorem 3, players adopt asymmetric pure strategies, and the game ends in at most $n + 1$ periods.

In this subsection, we will perform the welfare analysis in a game with $T = 2$. The analysis for the general game is much more complicated, but the results should remain qualitatively the same.

For the symmetric pure strategy equilibrium in Theorem 1, restricting attention to two periods is not binding, as the game will end in two periods anyway. For the asymmetric pure strategy equilibrium of Theorem 2, the restriction changes the equilibrium strategies of the players not belonging to the first bunch. Now only one bunch can be formed (in period 1), and all players who have not yet invested in period 1 must decide whether or not to invest in period 2. The equilibrium bunching size (in period 1) is again determined by (3), and is also independent of the total number of players n.

Define the social welfare as the sum of total payoffs for all players, written as $W(k, n)$, where k is the bunch size (i.e., the number of players making decisions) in period 1. The social welfare function is then given by

$$W(k, n) = \Pr(g)\left[k\pi(1|g) + (n - k)\delta R(k|g)\right] \tag{8}$$

$$+ \Pr(b)\delta\left[kR(k - 1|b) + (n - k)R(k|b)\right].$$

This social welfare function is based on the fact that there are k players in the bunch in period 1 and the rest of $n - k$ players will not invest in that period. A player in the bunch will invest in period 1 if his signal is g, and will not invest if his signal is b. In the latter case, he will make his investment decision in period 2 after observing the actions (and thus inferring the signals) of the rest $(k - 1)$ bunch members. A player outside the bunch will make his investment decision in period 2 after observing the actions (and thus inferring the signals) of all the players in the bunch.

Bunching in period 1 has three effects. The first effect is that a player in the bunch has only his own private signal to make his investment decision, and thus the decision is less informed. The larger the bunch size is, the more aggregate efficiency (on these bunch members) the society loses. The second effect is that a bunch member provides information to the players outside the bunch through his action. In period 2, players who have not invested can observe the actions of these bunch members and infer their signals. The larger the bunch size, the more informed these players become, and thus the society gains more aggregate efficiency (on non-bunch-members). The third effect is discounting. Investing early avoids the effect of discounting in payoffs. Trading off these three effects, there exists a socially optimal bunching size, which is denoted by $k^{so}(n)$. In what follows, we will compare this bunching size with the equilibrium bunching size. Note that the equilibrium bunching size in this two-period model is exactly the same as the equilibrium bunching size k_0^* in the many period model. This is because the trade-off for a player inside and outside the bunch is exactly the same.

In the following analysis, we will first consider the case of finitely many players, and then consider the limiting case when the number of players goes to infinity.

6.2.1 Finite Number of Players

Our first result is that the socially optimal bunching size is weakly greater than the equilibrium bunching size.

Proposition 6 $k_0^* \leq k^{so}(n)$.

Proof See appendix. ∎

This proposition is intuitive. In an equilibrium, players in the bunch consider their own payoffs only; they do not consider the (positive) externality effect of their actions. Starting with the equilibrium bunch size, a social planner would like to add weakly more players into the bunch. This is because one more player in the bunch now provides extra information to other players (and thus improves the payoffs of those players). But this positive externality was not considered in the equilibrium bunch size determination.

Note that the equilibrium bunch size k_0^* is independent of the total number of players n, while the socially optimal bunch size $k^{so}(n)$ depends on n. In the following proposition, we show that when the number of players increases, social welfare maximization would require the bunching size to be larger.

Proposition 7 $k^{so}(n)$ *is an increasing function of* n.

Proof See appendix. ∎

The intuition for this proposition is as follows. First fix the bunch size in the first period. When the number of players becomes larger, more players depend on the information revealed in the first period. Therefore, it is socially desirable to reveal more information by increasing the bunch size.

6.2.2 Infinite Number of Players

In this subsubsection, we will examine the socially optimal bunching size when the number of players goes to infinity.

Proposition 8 $\lim_{n\to\infty} k^{so}(n) = \infty$.

Proof See appendix. ∎

The above proposition shows that the bunching size must go to infinity when the number of players goes to infinity. The intuition is that when there are a large number of players making decisions in period 2, relying on the signals from the bunch in period 1, the socially optimal bunch size must also be large so that the information provided by the bunch is sufficiently precise. A small increase in the precision will benefit a large number of players in period 2.

Even though the bunch size goes to infinity when n goes to infinity, the following proposition shows that the percentage of players in the socially optimal bunch goes to zero as n goes to infinity.

Proposition 9 *Suppose that* $\pi(1|g) < \delta R(\infty|g)$. *Then* $\lim_{n\to\infty} \frac{k^{so}(n)}{n} = 0$.

Proof See appendix. ∎

This proposition states that the social planner needs only a small percentage of the players to be in the bunch in period 1 when the total number of players is large. The logic behind this is as follows. Even though the percentage of players in the bunch is small, the total number is still large and going to infinity as n goes to infinity. The players in the bunch can provide very precise information for the players outside the bunch as the bunch size is large. Meanwhile, the players in the bunch take up only a small proportion of the total population of players, and their payoffs have little effect when doing the averaging. The socially optimal welfare is thus achieved.

6.3   Welfare Analysis for Heterogeneous Players

In this subsection, we consider three variations of the original model by introducing three types of heterogeneities for the players – heterogeneous discount factors, heterogeneous costs of acquiring signals, and heterogeneous risk preferences. In what follows, we discuss these three cases one by one. We argue that when the heterogeneities are small, the asymmetric pure strategy equilibria characterized in Theorem 2 and Theorem 3 are still valid. When players are heterogeneous, there is one unique welfare maximizing asymmetric pure strategy equilibrium. Even though other asymmetric equilibria remain valid, this welfare maximizing equilibrium could become a natural way for the players to coordinate on their strategies.

6.3.1   Heterogeneous Discount Factors

Suppose that the discount factors are slightly different across the players, and can be ranked as $\delta_1 \leq \delta_2 \leq ... \leq \delta_n$. Suppose that the differentiations are small, such that for each d, there

is a unique k_d^* satisfying (3) for all $i = 1, ..., n$:

$$\delta_i R\left(k_d^* - 1 | d + 1, 0\right) \leq \pi(1 | d + 1, 0) \leq \delta_i R\left(k_d^* | d + 1, 0\right).$$

Then Theorem 2 and Theorem 3 are still valid.

To maximize the social welfare, a social planner would coordinate the players in such a way that the least patient players make decisions first. For example, the social planner would choose the most impatient k_0^* players, i.e., players with δ_1 to $\delta_{k_0^*}$, to form the bunch in the first period. If bunching occurs again in period 2, then the remaining most impatient players will be chosen, etc. If every player's discount factor is different, this coordination is unique and maximizes the welfare.

6.3.2   Heterogeneous Costs of Acquiring Signals

In the appendix, we examine a variation of the original model by introducing a cost of acquiring the private signals by the players. We show that our pure strategy equilibria are robust to such a cost when it is small.

Suppose that the costs of acquiring signals are heterogeneous across the players, and can be ranked as $c_1 \leq c_2 \leq ... \leq c_n$. Suppose that all these costs are small so that the asymmetric equilibria still hold, that is, the strategies in Theorem 2 and Theorem 3 are still valid. In this case, the social planner would choose the players with the lowest costs to form the first bunch, the players with the next lowest costs to form the second bunch, and so on. Again, if the costs are all different, then the welfare maximizing equilibrium is unique. This is also a natural way to coordinate the players in the asymmetric equilibria.

6.3.3   Heterogeneous Risk Preferences

In the original model, the players are assumed to be risk neutral. If they are slightly risk averse, Theorem 2 and Theorem 3 are still valid. The bunch size k_d^* is still determined by (3):

$$\delta R_i\left(k_d^* - 1 | d + 1, 0\right) \leq \pi_i(1 | d + 1, 0) \leq \delta R_i\left(k_d^* | d + 1, 0\right),$$

where π_i and R_i now represent the expected payoffs of player i.

Suppose that the players' degree of risk aversion can be ranked. The players who make decisions early have less information, and thus incur higher risk. This is because the underlying states are more uncertain when there is less information. This can be seen from $Var\left(Y\right) \geq E_x Var\left(Y | x\right).$[9] This implies that the expected risk (as measured by the vari-

---

[9] This is because $Var\left(Y\right) = E(Y^2) - (EY)^2$, $E_x Var\left(Y | x\right) = E_x\left[E\left(Y^2 | x\right) - \left(E\left(Y | x\right)\right)^2\right] = E(Y^2) - E_x\left(E\left(Y | x\right)\right)^2$. By Jensen's inequality, we know $(EY)^2 = \left(E_x\left(E\left(Y | x\right)\right)\right)^2 \leq E_x\left(E\left(Y | x\right)\right)^2$. Therefore, $Var\left(Y\right) \geq E_x Var\left(Y | x\right)$.

ance) is lower when more signals are revealed. When there are more signals, the players can predict the states more precisely, and the optimal decisions by the players involve less risk. Therefore, when players are slightly risk averse, the social planner would choose those who are less risk averse to bunch earlier to maximize welfare. This would also be a natural way for the players to coordinate in the asymmetric equilibrium.

7 Applications

Our model can be applied to the adoption of new products. New products, such as new iPhones, new albums, or new movies, are of unknown quality. Each consumer may obtain some private information from various kinds of sources, such as advertisements, word of mouth, or personal investigation. Some of these consumers may decide early on whether or not to purchase these new products. (These consumers can be regarded as the first bunch.) Meanwhile, other consumers may decide to wait regardless of their private information; they wait for the sales figures, etc., and some of them make their decisions then. When the sales figures are very good relative to people's expectation – for example, the new iPhones were out of stock everywhere on the first day of sale, Michael Jackson's album "thriller" reached 100 million copies, or the box office of "Avatar" was record high during the opening weekend – then a huge wave of consumers may follow suit (i.e., herding). On the contrary, if the sales figures are really bad, then few may buy the new product, which may then disappear from the market soon (herding of no-investment). If the sales figures are just moderate, then another group of consumers may decide whether or not to buy the new product (i.e., bunching again), and so on and so forth.

In the above application, the sales record reveals how many consumers have purchased the new product. Each consumer has some prior anticipation about how many people should be in the first bunch (who make decisions first). The actual number of units sold (or, in the case of supply shortage, the number of people who want to buy) then represents the number of players with good signals in the bunch. If this number is relatively large, e.g., if we see many people queuing outside the Apple Store at 3:00am in the morning, then we can infer that the number of good signals outweighs the number of bad signals by a significant margin in the bunch. In this case, the consumers may herd and buy. However, if the consumers see relatively few people buying the new product, they infer that there are many more bad signals than good signals in the bunch. In this case, they herd and will not buy the product. Meanwhile, if the sales figure is mediocre, then it is likely that another subgroup of consumers will make their purchase decisions.

In another application, we consider the investment decisions when the economy is highly uncertain, such as when we are near the end of a recession. In that situation, investors are highly unsure about the state of the economy, and a signal (either good or bad) is very valuable. We usually observe that a group of investors, such as big investment banks,

decide early; they either invest or do not invest, depending on the information they receive. Other investors just stay on the sideline. Then herding may occur, if many investors from this first bunch invest (or not invest). When neither of the sides in the bunch dominates, another bunch of investors forms, etc. An important implication from our analysis is that it is not the amount of investments relative to the total amount of possible investments that matters. It is the amount of investments relative to the equilibrium bunch size that matters. Financial analysts have long incorporated the supposed trading volumes versus the actual trading volumes (in addition to the market prices) in their analysis. In IPO offerings, the expected subscription rate versus the actual subscription rate plays an important role in the analysis. In a sense, this practice reflects some of the intuitions we obtain in this paper.

The recent popularity of group discount websites, especially Groupon.com, provides a possible venue to test the equilibrium behavior of the buyers. On these websites, the number of units sold is posted and instantly updated. Buyers are sometimes unsure about how good a deal a coupon offers. They may have some private information. But from the sales figure, they can see the actions of other buyers and infer useful information. Here, the total number of potential buyers could be difficult to obtain, but the bunch sizes could be calculated if we know other parameters of the model. Asymmetric equilibria are useful when we do not have precise information on the total number of players.

8   Conclusion

In this paper, we characterize the pure strategy equilibria when players with private information decide not only whether or not to make an investment in a risky environment, but also when to make the investment. We show that a unique symmetric pure strategy equilibrium exists when players are impatient, or when the number of players is small. Whenever these conditions are not satisfied, asymmetric pure strategy equilibria exist. In these equilibria, players either herd or bunch in any given period. Bunching can occur only when there was bunching in all previous periods. When herding occurs, the game ends. In these equilibria, some players make decisions early and some later, regardless of their signals. Later players learn information from the early decision makers and can make more informed decisions, but their payoffs are discounted.

We focus on the asymmetric pure strategy equilibria in this paper. These asymmetric equilibria exhibit the property of bunching. In bunching, a bunch of players make decisions while the rest of the players do nothing regardless of their signals. When players are very patient, the asymmetric equilibria enclose those in a comparable fixed timing model from the literature. When the discount factor is equal to 1, the least efficient equilibrium is the one resembling the fixed timing model, while in the most efficient equilibrium, the bunch size is always equal to 1 and herding of investment disappears. Bunch sizes are shown to be independent of the total number of players. The equilibrium asymmetric pure strategies are

invariant to adding more players to the game. All these properties are unique to asymmetric pure strategy equilibria. These equilibria are shown to be more efficient than symmetric mixed strategy equilibria.

The analysis in this paper has many applications. It applies to the adoption of new products, the issuing of IPOs, investments in the financial markets, etc. When people make decisions in a dynamic setting, the asymmetric equilibria in our model predict that they are likely to make decisions in groups. This is also true when players do not initially have private information and must pay a small fee if they want to acquire a private signal. Thus, the equilibria characterized in this paper are robust to small fees for information acquisition.

In this paper, the players interact with each other through the revelation of their private signals; the actions they take do not affect each other's payoff directly. When a player's payoff is affected directly by other players' actions, the analysis would depend on whether their actions are complements or substitutes. This will be explored in our next project.

9 Appendix

9.1 Costly Signal Acquisition

In this subsection, we investigate whether the equilibrium strategies analyzed in the main text are robust to a small cost of acquiring information. Instead of giving each player a signal at the beginning of the game for free, now suppose that each player has to pay a small signal acquisition fee $c > 0$ to get his signal. Players can choose whether and when to acquire this signal, and then decide whether and when to invest. Since future payoffs are discounted, it is obvious that players will not acquire their signals until the period in which they are about to make their investment decisions. That means only those players in the bunch of the current period acquire their signals in that period.

In what follows, we shall argue that the equilibrium described in Theorem 2 continues to be an equilibrium when c is small enough. The bunch sizes remain the same as in Theorem 2. Apart from the fact that the players now pay an information acquisition fee (in the period that their bunch is about to act), the equilibrium strategies for the players are exactly the same as in Theorem 2.

Consider a player in the first bunch of k_0 players. If he acquires the signal in period 1, his expected payoff is

$$\Pr(g)\,\pi(1|g) + \Pr(b)\,\delta R\,(k_0 - 1|b) - c, \tag{9}$$

that is, with probability $\Pr(g)$, he gets signal g and invests, and his expected payoff is $\pi(1|g)$; with probability $\Pr(b)$, he gets signal b and does not invest, and his expected payoff is $\delta R\,(k_0 - 1|b)$, since in the next period he uses the one-period best response to the signals

of the other $k_0 - 1$ players in his bunch.

Now suppose that he deviates by not acquiring the signal in period 1. Then in period 1, he does not invest. In period 2, if the difference between the numbers of good signals and bad signals in the first bunch $d_2 > 1$, herding of investment occurs. The player will invest with all remaining players regardless of the signal he may receive by paying the fee. Therefore, he will not acquire his signal in this case. Similarly, if $d_2 < 0$, herding of no-investment occurs, and he will not pay the fee to acquire his signal. If $d_2 = 0$ or $d_2 = 1$, herding does not occur. In these cases, he will pay the cost to acquire his signal as his optimal action depends on it. He will invest if his signal is g, and not invest if his signal is b.

To summarize, if in the first period this player does not acquire his signal, then in the second period he will best respond to the $k_0 - 1$ signals from the first bunch. (Note that when herding occurs, his best response remains the same regardless of his signal.) Therefore, we can express this player's expected payoff as

$$\Pr(g)\,\delta R(k_0 - 1|g) + \Pr(b)\,\delta R(k_0 - 1|b) - c\Pr(d_2 = 0 \text{ or } d_2 = 1). \tag{10}$$

It is clear that when c is sufficiently small, the requirement that the player's payoff in (9) is greater than the payoff in (10) is equivalent to (IC1), because those terms with c go to zero as c goes to zero. When this condition holds, it is optimal for the player (who is in the first bunch) to acquire the signal in period 1.

Similarly, for a player not in the first bunch, it is optimal not to invest and not to acquire the signal in period 1. This is because his expected payoff from not acquiring the signal in period 1 is

$$\Pr(g)\,\delta R(k_0|g) + \Pr(b)\,\delta R(k_0 - 1|b) - c\Pr(d_2 = 0 \text{ or } d_2 = 1), \tag{11}$$

and his expected payoff if he deviates (and thus acquires his signal in period 1) is

$$\Pr(g)\,\pi(1|g) + \Pr(b)\,\delta R(k_0|b) - c. \tag{12}$$

When c is sufficiently small, the condition that the player's payoff in (11) be greater than the payoff in (12) is equivalent to (IC3). When this condition holds, it is optimal for the player (who is not in the first bunch) not to acquire the signal in period 1.

We thus have the following theorem when c is sufficiently small, which is parallel to Theorem 2. It shows that the equilibria in our model are robust to introducing a small cost of acquiring signals.

Theorem 4 *Suppose (A1), (A3) and (5) hold and c is sufficiently small. There exists an asymmetric pure strategy PBE, which has the following recursive properties. Let $d_1 = 0$. In period t, $t = 1, 2, ...$*

(1) if $d_t > 1$, Herding of Investment occurs, and the game ends;

(2) if $d_t < 0$, Herding of No-Investment occurs, and the game ends;

(3) if $d_t = 0$, and $\tilde{n}_t \geq k_0^*$, then Bunching of size k_0^* occurs;

(4) if $d_t = 1$, and $\tilde{n}_t \geq k_1^*$, then Bunching of size k_1^* occurs;

(5) if $d_t = 0$, and $\tilde{n}_t < k_0^*$, or if $d_t = 1$, and $\tilde{n}_t < k_1^*$, then the Last Bunching of size $\tilde{n}_t$ occurs.

9.2 Equilibrium Analysis when (A1) and (A3) Fail

In this subsection, we shall consider the situations where the assumptions in Theorem 2 do not hold. We first consider the case when (A3) does not hold. Let

$$\pi(1) = \Pr(G)\pi_G + \Pr(B)\pi_B \geq 0. \tag{A3'}$$

This implies that unconditional on any signal, it is profitable to invest. Observe that (A3') implies (A1). The results when (A2) and (A3') hold are similar to Theorem 2 and are summarized in the following theorem.

Theorem 5 *Suppose (A2), (A3') and (5) hold. There exists an asymmetric pure strategy PBE, which has the following recursive properties. Let $d_1 = 0$. In period t, $t = 1, 2, ...$*

(a) if $d_t > 0$, Herding of Investment occurs, and the game ends;

(b) if $d_t < -1$, Herding of No-Investment occurs, and the game ends;

(c) if $d_t = -1$, and $\tilde{n}_t \geq k_{-1}^*$, then Bunching of size k_{-1}^* occurs;

(d) if $d_t = 0$, and $\tilde{n}_t \geq k_0^*$, then Bunching of size k_0^* occurs;

(e) if $d_t = -1$, and $\tilde{n}_t < k_{-1}^*$, or if $d_t = 0$, and $\tilde{n}_t < k_0^*$, then the Last Bunching of size $\tilde{n}_t$ occurs.

In the above theorem, we have similar bunching patterns as in Theorem 2, except that now unconditional on any signal, players have non-negative investment payoffs. Basically, in this case, we need one fewer good signal than in Theorem 2. Other aspects of the equilibrium strategies are the same.

We now consider the case where (A2) holds, i.e., $\pi(1|b)$ is negative, but (A1) does not hold, i.e., $\pi(1|g)$ is non-positive. The implication is trivial: no one will ever invest, and Herding of No-Investment is the only outcome.

Finally, we consider the case where (A1) holds, i.e., $\pi(1|g)$ is positive, but (A2) does not hold, i.e., $\pi(1|b)$ is non-negative. In this case, Herding of Investment occurs in $t = 1$.

9.3 Proofs

Proof for Lemma 1.

In this proof, we abuse the notations by using $(N_{g,t+1}, N_{b,t+1})$ to replace $(N_{g,t} + r, N_{b,t} + k_t - r)$, in order to make the equations short.

$$R(k_t|N_{g,t}, N_{b,t})$$

$$= \sum_{r=0}^{k_t} \pi\left((BR\left(N_{g,t+1}, N_{b,t+1}\right)|N_{g,t+1}, N_{b,t+1}\right) \Pr\left(N_{g,t+1}, N_{b,t+1}|N_{g,t}, N_{b,t}\right)$$

$$= \sum_{s=g,b}\sum_{r=0}^{k_t} \pi\left((BR\left(N_{g,t+1}, N_{b,t+1}\right)|N_{g,t+1}, N_{b,t+1}, s\right) \Pr\left(N_{g,t+1}, N_{b,t+1}, s|N_{g,t}, N_{b,t}\right)$$

$$< \sum_{s=g,b}\sum_{r=0}^{k_t} \pi\left((BR\left(N_{g,t+1}, N_{b,t+1}, s\right)|N_{g,t+1}, N_{b,t+1}, s\right) \Pr\left(N_{g,t+1}, N_{b,t+1}, s|N_{g,t}, N_{b,t}\right)$$

$$= R(k_t + 1|N_{g,t}, N_{b,t})$$

This completes the proof for the first part of the lemma.

When k_t goes to infinity, there will be infinite number of signals to be revealed in period t. Given the underlying state, the signals are i.i.d. According to the Law of Large Numbers, the proportion of good signals converges to its expected value with probability one, that is, it converges to p in state G, and $1 - p$ in state B. Note that $p > \frac{1}{2}$. Therefore, in period $t + 1$, when the state is G, the total number of good signals will exceed the total number of bad signals by any given finite number with probability one. Seeing this, the best response will always be $a = 1$, that is, to invest. Likewise, in period $t + 1$, when the state is B, the total number of bad signals will exceed the total number of good signals by any given finite number with probability one. Seeing this, the best response will always be $a = 0$, that is, not to invest.

Therefore, the expected payoff is $\Pr(G|N_{g,t}, N_{b,t})\pi_G + \Pr(B|N_{g,t}, N_{b,t})0$. This completes the proof for the second part of the lemma.

Proof for Lemma 2.

We have

$$\pi(1|d,0) - R(k|d,0)$$

$$= \left(\sum_{r=0}^{k} - \sum_{r+d>k-r} \right) \pi(1|d+r,k-r)\Pr(d+r,k-r|d,0)$$

$$= \sum_{r+d\leq k-r} \pi(1|d+r,k-r)\Pr(d+r,k-r|d,0).$$

Since $\pi(1|d+r,k-r)$ is increasing in d and each $\pi(1|d+r,k-r)$ is negative when $r+d \leq k-r$. Also,

$$\Pr(d+r,k-r|d,0)$$

$$= \Pr(r,k-r|G)\Pr(G|d,0) + \Pr(r,k-r|B)\Pr(B|d,0)$$

$$= \left[p^{k-r}(1-p)^r - p^r(1-p)^{k-r} \right]\Pr(B|d,0) + p^r(1-p)^{k-r}.$$

When $r+d \leq k-r$ and $p > \frac{1}{2}$, $p^{k-r}(1-p)^r - p^r(1-p)^{k-r} > 0$. And recall that $\Pr(B|d,0)$ is decreasing in d. Thus, $\Pr(d+r,k-r|d,0)$ is decreasing in d and it is positive.

Thus, each term after the summation sign is negative and decreasing in d in absolute value. Also, the range for r such that $r+d \leq k-r$ is shrinking as d increases. Therefore, $\pi(1|d,0) - R(k|d,0)$ is increasing in d.

Since

$$\pi(1|d,0) - \delta R(k|d,0) = (1-\delta)\pi(1|d,0) + \delta[\pi(1|d,0) - R(k|d,0)],$$

and $\pi(1|d,0)$ is increasing in d, we conclude that $\pi(1|d,0) - \delta R(k|d,0)$ is increasing in d.

Proof for Theorem 2.

Denote the players in the bunch in period t as H_t, $t = 1,..,T$. We call a player with signal $s = g,b$ a type s player in the proof.

In period 1, a type b player $i \in H_1$ does not have incentive to deviate by investing in period 1, according to assumption (A2). For the same reason, a type b player $i \in \overline{H_1}$ does not have incentive to deviate by investing in the period 1. By (IC3), A type g player $i \in \overline{H_1}$ does not have incentive to deviate by investing in period 1.

Now we will prove that a type g player $i \in H_1$ does not want to deviate by waiting. By (IC1), it is not optimal for this player to wait until period 2 and use his myopic best response then. In the following, we prove that the myopic best response in period 2 is overall optimal for him, that is, he does not want to wait further. The myopic best response function defined

in (2) means that if it gives positive expected payoff, the player should invest right away, instead of waiting further. To show that it is overall optimal for him to use his myopic best response in period 2, we need to consider what happens in period 2. There are 3 cases in period 2, corresponding to the three cases described in the Corollaries 4, 5 and 6.

We consider the bunching pattern in Corollary 4 first. Recall that only cases (a), (b) and (c) in Theorem **2** can occur.

If a type g player $i \in H_1$ deviates, other players will believe he has a bad signal instead of a good one. That makes his own update different from the public. To be precise, he knows that there are two additional good signals than what the other players believe (inferred from the actions of the first bunch).

If $d_1 > 1$ happens, it is actually $\tilde{d}_1 > 3$ for the deviated player. Since all the other players will herd by investing, no further information is to be revealed. Thus, he has no incentive to wait further. It is overall optimal for him to invest right away, which is also his myopic best response.

If $d_1 < 0$ happens, it is actually $\tilde{d}_1 < 2$ for the deviated player. Since all the other players will herd by not investing, no further information is to be revealed as well. Thus, he has no incentive to wait further. It is overall optimal for him to use his myopic best response.

If $d_1 = 0$ happens, it is actually $\tilde{d}_1 = 2$ for the deviated player. If

$$\pi(1|2g) \geq \delta R\left(k_0^*|2g\right) \tag{13}$$

holds, then instead of waiting until the third period (the RHS above) after seeing another k_0^* signals from H_2 in this case, he will invest right away (the LHS above), which is his myopic best response.

We now show that (13) holds. The LHS of (13) is

$$\begin{aligned}
\pi(1|2g) &= \pi(1|3g)\Pr(g|2g) + \pi(1|2g, b)\Pr(b|2g) \\
&= \pi(1|3g)\Pr(g|2g) + \pi(1|g)\Pr(b|2g).
\end{aligned}$$

And the RHS of (13) is

$$\begin{aligned}
& R\left(k_0^*|2g\right) \\
&= R\left(k_0^* - 1|3g\right)\Pr\left(g|2g\right) + R\left(k_0^* - 1|2g, b\right)\Pr\left(b|2g\right) \\
&= R\left(k_0^* - 1|3g\right)\Pr\left(g|2g\right) + R\left(k_0^* - 1|g\right)\Pr\left(b|2g\right).
\end{aligned}$$

By (IC1), we have $\pi(1|g) \geq \delta R\left(k_0^* - 1|g\right)$. Together with Lemma 2, we have $\pi(1|3g) > \delta R\left(k_0^* - 1|3g\right)$. Thus, it is clear that (13) holds. Therefore, to use his myopic best response in period 2 rather than wait further is optimal for this type g player $i \in H_1$. And in all later

periods, the incentive is the same as the players in the second bunch. This completes the proof for the case in Corollary 4.

Now consider the bunching pattern in Corollary 5. Recall that only cases (a), (b) and (d) in Theorem 2 can occur. The argument is the same as in the above when herding, i.e., cases (a) and (b) occur. In case (d), a second bunch H_2 of size k_1^* should occur in the second period.

If $d_1 = 1$ happens, it is actually $\tilde{d}_1 = 3$ for the deviated player. If

$$\pi(1|3g) \geq \delta R\left(k_1^*|3g\right) \tag{14}$$

holds, then instead of waiting until the third period (the RHS of the above), he will invest right away (the LHS of the above), which is his myopic best response. Similar to the proof for (13), the LHS of (14) is

$$\pi(1|3g) = \pi(1|4g)\Pr(g|3g) + \pi(1|2g)\Pr(b|3g),$$

and the RHS of (13) is

$$R\left(k_1^*|3g\right) = R\left(k_1^* - 1|4g\right)\Pr\left(g|3g\right) + R\left(k_0^* - 1|2g\right)\Pr\left(b|3g\right).$$

By (IC1), we have $\pi(1|2g) \geq \delta R\left(k_1^* - 1|2g\right)$. Together with Lemma 2, we have $\pi(1|4g) > \delta R\left(k_1^* - 1|4g\right)$.

Thus, it is clear that (14) holds. Therefore, to use the myopic best response function in period 2 rather than to wait further is overall optimal for this type g player $i \in H_1$. And in all later periods, the incentive is the same as the players in the second bunch. This completes the proof for the case in Corollary 5.

Last consider the bunching pattern in Corollary 6. The proof for this part is similar to the proof for the case in Corollary 5, except for a type g player $i' \in H_3$. Since then, $\tilde{d}_2 = 2$ and a fourth bunch of size k_0^* is followed. The condition for this player i' not to deviate by waiting further is again $\pi(1|2g) \geq \delta R\left(k_0^*|2g\right)$, the same as (13).

This completes the proof for Theorem 2.

Proof for Lemma 11.

We have

$$\pi(1|d,0) - R\left(k|d,0\right)$$

$$= \Pr\left(G|d,0\right) \sum_{r+d \leq k-r} \binom{k}{r} p^r \left(1-p\right)^{k-r}\pi_G + \Pr\left(B|d,0\right) \sum_{r+d \leq k-r} \binom{k}{r}\left(1-p\right)^r p^{k-r}\pi_B$$

$$= \sum_{r+d \leq k-r} \binom{k}{r}\left[p^r\left(1-p\right)^{k-r}\Pr\left(G|d,0\right)\pi_G + \left(1-p\right)^r p^{k-r}\Pr\left(B|d,0\right)\pi_B\right]. \tag{15}$$

37

By Bayes' Rule, we have $\Pr\left(G|d,0\right) = \frac{p^d q}{p^d q + (1-p)^d (1-q)} = \frac{p^d}{p^d + (1-p)^d (1-q)/q}$, which is an increasing function of q, and so is the first group of terms inside the brackets of (15). In addition, $\Pr\left(B|d,0\right) = 1 - \Pr\left(G|d,0\right)$, which is a decreasing function of q. Since $\pi_B < 0$, the second group of terms inside the brackets of (15) is then increasing in q.

Similarly, we know $\pi(1|d,0)$ is an increasing function of q.

Since

$$\pi(1|d,0) - \delta R\left(k|d,0\right) = (1-\delta)\,\pi(1|d,0) + \delta\left(\pi(1|d,0) - R\left(k|d,0\right)\right),$$

$\pi(1|d,0) - \delta R\left(k|d,0\right)$ is increasing in q as well.

Since $\pi(1|d+1,0) - \delta R\left(k|d+1,0\right)$ is decreasing in k, and the bunching size k_d^* is the biggest integer satisfying $\pi(1|d+1,0) - \delta R\left(k|d+1,0\right) \geq 0$. Therefore, an increase in q will increase k_d^*.

Proof for Proposition 6.

Suppose not. Then $k_0^* > k^{so}\left(n\right)$. We shall ignore the argument n in $k^{so}\left(n\right)$ in the proof when it does not cause confusion.

Recall that

$$\delta R\left(k_0^* - 1|g\right) \leq \pi(1|g) \leq \delta R\left(k_0^*|g\right)$$

and that $R\left(k|g\right)$ is a strictly increasing function of k. Since $k_0^* > k^{so}$, it must be that $k^{so} \leq k_0^* - 1$. From the above inequality, we know that $\delta R\left(k^{so}|g\right) \leq \pi(1|g)$. Thus,

$$\begin{aligned}
&k_0^* \pi(1|g) + (n - k_0^*)\,\delta R\left(k_0^*|g\right) \\
=\;& k^{so}\pi(1|g) + (k^* - k^{so})\,\pi(1|g) + (n - k^*)\,\delta R\left(k^*|g\right) \\
>\;& k^{so}\pi(1|g) + (k^* - k^{so})\,\delta R\left(k^{so}|g\right) + (n - k^*)\,\delta R\left(k^{so}|g\right) \\
=\;& k^{so}\pi(1|g) + (n - k^{so})\,\delta R\left(k^{so}|g\right),
\end{aligned}$$

and

$$\begin{aligned}
&k_0^* R\left(k_0^* - 1|b\right) + (n - k_0^*)\,R\left(k_0^*|b\right) \\
>\;& k^{so}R\left(k^{so} - 1|b\right) + (k_0^* - k^{so})\,R\left(k^{so}|b\right) + (n - k_0^*)\,R\left(k^{so}|b\right) \\
=\;& k^{so}R\left(k^{so} - 1|b\right) + (n - k^{so})\,R\left(k^{so}|b\right).
\end{aligned}$$

Combining the above two inequalities, we have

$$W(k_0^*, n) > W(k^{so}, n).$$

Thus, $k^{so}(n)$ is not the socially optimal bunching size. This is a contradiction. Therefore, we conclude that $k_0^* \leq k^{so}(n)$.

Proof for Proposition 7.

Suppose not. Then for some $n_1 < n_2$, we have $k^{so}(n_1) > k^{so}(n_2)$. For simplicity, we write $k^{so}(n_i)$ as k_i^{so}, $i = 1, 2$. Since

$$\Pr(g)\left[k_1^{so}\pi(1|g) + (n_1 - k_1^{so})R(k_1^{so}|g)\right] + \Pr(b)\left[k_1^{so}R(k_1^{so} - 1|b) + (n_1 - k_1^{so})R(k_1^{so}|b)\right]$$

$$\geq \Pr(g)\left[k_2^{so}\pi(1|g) + (n_1 - k_2^{so})R(k_2^{so}|g)\right] + \Pr(b)\left[k_2^{so}R(k_2^{so} - 1|b) + (n_1 - k_2^{so})R(k_2^{so}|b)\right],$$

then

$$\Pr(g)\left[n_1\left(R(k_1^{so}|g) - R(k_2^{so}|g)\right) + (k_1^{so} - k_2^{so})\pi(1|g) + k_2^{so}R(k_2^{so}|g) - k_1^{so}R(k_1^{so}|g)\right]$$

$$+ \Pr(b)\delta[n_1\left(R(k_1^{so}|b) - R(k_2^{so}|b)\right) + k_1^{so}\left(R(k_1^{so} - 1|b) - R(k_1^{so}|b)\right)$$

$$-k_2^{so}\left(R(k_2^{so} - 1|b) - R(k_2^{so}|b)\right)]$$

$$\geq 0.$$

Since $k_1^{so} > k_2^{so}$, we have $R(k_1^{so}|\cdot) - R(k_2^{so}|\cdot) > 0$. Thus, the above expression is strictly greater than zero if we replace n_1 by a larger number n_2. Rearranging, we obtain

$$W(k_1^{so}, n_2) > W(k_2^{so}, n_2).$$

Therefore, k_2^{so} is not the social optimal bunching size for n_2. This is a contradiction. So $k^{so}(n)$ must be an increasing function for every n.

Proof for Proposition 8.

We first consider the case where $\pi(1|g) \geq \delta R(\infty|g) = \delta \Pr(G|g)\pi_G$. In this case, all players should be in the bunch in period 1 in equilibrium. Since the socially optimal bunching size is at least the equilibrium bunching size, the socially optimal bunching size is also n. Therefore, $\lim_{n \to \infty} k^{so}(n) = \infty$.

We now consider the case where $\pi(1|g) < \delta R(\infty|g) = \delta \Pr(G|g)\pi_G$. We shall prove the proposition by contradiction. Suppose that $\lim_{n \to \infty} k^{so}(n) \neq \infty$. Then there exist a subsequence $\{k^{so}(n_i)\}_{i=1,\ldots,\infty}$ of $\{k^{so}(n)\}$, such that $k^{so}(n_i) \to \bar{k}$, as $i \to \infty$. From (8), for a given k, we have

$$\frac{W(k,n)}{n} = \Pr(g)\left[\frac{k}{n}\pi(1|g) + \left(1 - \frac{k}{n}\right)\delta R(k|g)\right] \tag{16}$$

$$+ \Pr(b)\delta\left[\frac{k}{n}R(k-1|b) + \left(1 - \frac{k}{n}\right)R(k|b)\right].$$

39

Then
$$\lim_{i \to \infty} \frac{W\left(k^{so}\left(n_i\right), n_i\right)}{n_i} = \Pr\left(g\right)\delta R\left(\overline{k}|g\right) + \Pr\left(b\right)\delta R\left(\overline{k}|b\right).$$

Meanwhile, for any given fixed k, we have
$$\lim_{n \to \infty} \frac{W(k, n)}{n} = \Pr\left(g\right)\delta R\left(k|g\right) + \Pr\left(b\right)\delta R\left(k|b\right).$$

Since $R\left(k|\cdot\right)$ is strictly increasing in k, for any $k > \overline{k}$, we can conclude that $\lim_{n \to \infty} \frac{W(k,n)}{n} > \lim_{n \to \infty} \frac{W\left(k^{so}\left(n_i\right), n\right)}{n}$. That is, the subsequence $\{k_i^{so}\left(n\right)\}_{i=1,\dots,\infty}$ does not maximize the social welfare. This is a contradiction. Thus, we have $\lim_{n \to \infty} k^{so}\left(n\right) = \infty$ in this case as well.

Proof for Proposition 9.

For a given converging sequence $\{\frac{k(n)}{n}\}_{n=1,\dots,\infty}$, with $\lim_{n \to \infty} k\left(n\right) = \infty$ and $\lim_{n \to \infty} \frac{k(n)}{n} = \alpha$, we have

$$\lim_{n \to \infty} \frac{W\left(k\left(n\right), n\right)}{n}$$

$$= \lim_{n \to \infty} \left\{\Pr\left(g\right)\left[\frac{k\left(n\right)}{n}\pi(1|g) + \left(1 - \frac{k\left(n\right)}{n}\right)\delta R\left(k\left(n\right)|g\right)\right] + \right.$$

$$\left. \Pr\left(b\right)\delta\left[\frac{k\left(n\right)}{n}R\left(k\left(n\right) - 1|b\right) + \left(1 - \frac{k\left(n\right)}{n}\right)R\left(k\left(n\right)|b\right)\right]\right\}$$

$$= \Pr\left(g\right)\left[\alpha\pi(1|g) + (1 - \alpha)\delta R\left(\infty|g\right)\right] + \Pr\left(b\right)\delta R\left(\infty|b\right).$$

The per player social welfare $\frac{W(k(n),n)}{n}$ is maximized in the limit if $\alpha = 0$. If $\lim_{n \to \infty} \frac{k^{so}(n)}{n} \neq 0$, then we can find a sequence of $k\left(n\right)$, with $\lim_{n \to \infty} k\left(n\right) = \infty$ and $\lim_{n \to \infty} \frac{k(n)}{n} = 0$, (e.g., $k\left(n\right) = \sqrt{n}$) such that the per player welfare from $k\left(n\right)$ is larger than from $k^{so}\left(n\right)$, when n is large enough. Therefore, we conclude that $\lim_{n \to \infty} \frac{k^{so}(n)}{n} = 0$.

References

[1] Aoyagi, Masaki "Equilibrium Delay in a Simple Dynamic Model of Investment." Economic Theory, Vol. 12, No. 1, pp. 123-146, 1998.

[2] Banerjee, Abhijit V. "A Simple Model of Herd Behavior." Quarterly Journal of Economics, Vol. 107, No. 3, pp. 797-817, 1992.

[3] Bikhchandani, Sushil, Hirshleifer, David, and Welch, Ivo. "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades." Journal of Political Economy, Vol. 100, No.5, pp. 992-1026, 1992.

[4] Cao, H. Henry, Han, Bing, and Hirshleifer, David "Taking the Road Less Traveled By: Does Conversation Eradicate Pernicious Cascades?" Journal of Economic Theory, Vol 146, pp. 1418–1436, 2011.

[5] Chamley, Christophe, and Gale, Douglas "Information Revelation and Strategic Delay in a Model of Investment." Econometrica, Vol. 62, No. 5, pp. 1065-1085, 1994.

[6] Gul, Faruk, and Lundholm, Russell "Endogenous Timing and the Clustering of Agents' Decisions." Journal of Political Economy, Vol. 103, No. 5, pp. 1039-1066, 1995.

[7] Khanna, Naveen and Mathews, Richmond D. "Can Herding Improve Investment Decisions?" RAND Journal of Economics, Vol. 42, No. 1, pp. 150–174, 2011.

[8] Moretti, Enrico "Social Learning and Peer Effects in Consumption: Evidence from Movie Sales." Review of Economic Studies, Vol. 78, pp. 356-393.

[9] Park, Andreas, and Sabourian, Hamid "Herding and Contrarian Behaviour in Financial Markets." Econometrica, forthcoming.

[10] Rosenberg, Dinah, Solan, Eilon, and Vieille, Nicolas "Social Learning in One-arm Bandit Problems." Econometrica, Vol. 75, No. 6, pp. 1591–1611, 2007.

[11] Scharfstein, David, and Stein, Jeremy C. "Herd Behavior and Investment." American Economic Review, Vol. 80, No.3, pp. 465-479, 1990.

[12] Sgroi, Daniel "The Right Choice at the Right Time: A Herding Experiment in Endogenous Time." Experimental Economics, Vol. 6, No. 2, pp. 159-180, 2003.

[13] Zhang, Jianbo "Strategic Delay and the Onset of Investment Cascades." RAND Journal of Economics, Vol. 28, No. 1, pp. 188-205, 1997.

[14] Zhang, Yi "Robust Information Cascade with Endogenous Ordering." Working Paper, Singapore Management University, 2009.