Persistence and Reversal in Herd Behavior: Theory and Application to the Decision to Go Public

Lee Nelson
New York University

We model rational herd behavior when the underlying value changes over time, with payoffs that are either dependent or independent of the underlying value. We show that herding does not last forever and is not monotone in signal quality. High correlation among agents' actions does not necessarily imply herding. This suggests alternative empirical methods are needed to detect herding. The model has many applications, including the IPO decision in which payoffs are state dependent. The model implies that the decision to go public is more likely associated with herding than the decision to delay an IPO.

One part of the decision to go public is timing. Should the firm file with the SEC now or wait? Alternatively, having filed and set a target initial public offering (IPO) time frame, should the firm go ahead with the IPO or postpone it? This decision highly depends on expected market conditions several months after the filing.[1] The decision of a firm to file may therefore reveal information about its perception, at the time of filing, of future market conditions. This information is potentially valuable to other firms pondering the same choice. On the other hand, the popular press often regards firms' IPO decisions as blind herd behavior.[2] In this case, no useful information can be derived from observing the decision to file or to postpone an IPO. For example, after several cancellation decisions in August 1998, theglobe.com decided to cancel its IPO. Was this based on its information or was it simply

I would like to thank Avraham Beja, Amir Dembo, Darrell Duffie, Zsuzsanna Fluck, Steve Grenadier, Paul Pfleiderer, Manju Puri, Sven Rady, Bill Sharpe, Steve Tadelis, Ingrid Werner, Bob Wilson, and especially Anat Admati, Moshe Molcho, Dimitri Vayanos, and Jeff Zwiebel for helpful discussions and comments. This article is an extension of one of my PhD thesis chapters, done at the Stanford GSB. Address correspondence to Lee Nelson, TDA Capital Partners, 15 Valley Dr., Greenwich CT 06831, e-mail: leenelsonlee@yahoo.com.

[1] For our purposes we need only assume that firms are sensitive to the same market factors. What these factors are precisely is outside the scope of this article (but see also note 5). The importance of going public in favorable market conditions is often analyzed in the popular press. See, for example "Take Internet, Add Timing, Hit Paydirt," *Wall Street Journal,* April 19, 1999; "IPO Outlook," *Wall Street Journal,* August 30, 1999; "The Secret World of IPOs," *Inc*., June 1997; "Asian IPOs Heat Up on Internet Mania but no Sure Thing," *Investment Dealers Digest*, January 1999; and many others. See also Welch's http://linux.agsm.ucla.edu/ipo/ipopage.html#7 for a general discussion. For particular offerings that did not go as well as expected, it is often said that had market conditions been different the IPO would have done better.

[2] For example, "Activity May Heat Up in IPO Sector This Week," *Wall Street Journal* June 22, 1999, talks of a "frenzy of IPO filings"; "Blasting Off," *Computer Reseller News*, July 1999, uses the term "stampede" and quotes analysts saying it is a "frenzy" and warning of "saturation." "The Secret World of IPOs," *Inc*., June 1997, says ". . . many blamed the [IPO aftermarket] malaise on a host of still-green companies that had been rushed to the market at any price to take advantage of its insatiable appetite for equities."

The Review of Financial Studies Spring 2002 Vol. 15, No. 1, pp. 65–95

herding after previous firms? Similar informational inferences can be drawn in the context of other decisions as well. These decisions include choosing to actively look for M&A partners, adopting financial or technological innovation, etc. For example, seeing that several brick and mortar firms choose to open online divisions may prompt other brick and mortar firms to conclude that conditions are ripe for a brick-to-click migration. This in turn may prompt even more firms to adopt this technology, in some cases inducing firms that should not have adopted this policy to do so anyway (e.g., the *n*th online pet supply store).

In this article we present a general, dynamic model of rational herd behavior, with a changing underlying state (market conditions). While this model applies to many decisions beyond those mentioned above, such as analyst recommendations, for exposition purposes we will concentrate on the application to the IPO choice. The model is a generalization of the herding model in Bikhchandani, Hirshleifer, Welch (1992; hereafter BHW).[3]

In their model there is a constant underlying value that a sequence of identical agents are trying to guess. Each agent in his turn gets a private (noisy) signal about the underlying value and then decides on an observable action. BHW show that in many cases it is optimal for an agent to ignore his information and "follow the crowd" or herd. This creates an informational cascade, since all agents are identical and thus once one agent has herded, all subsequent agents will also ignore their information and the herding will continue indefinitely. The cascade that occurs may be wrong (e.g., all subsequent agents choose *up* while the underlying value is *down*).

A crucial assumption in the BHW model is that the underlying value is static. Other models with a static underlying value include Banerjee (1992), Welch (1992), Chamley and Gale (1994), and Hirshleifer and Welch (1998).[4] This BHW assumption leads to the striking result that once herding starts it never stops. In many cases both this assumption and this result seem inappropriate. That is, the correct decision may change over time, and when herding occurs it does not go on indefinitely. A natural example is the decision to go ahead with an IPO, as mentioned above. "Market conditions" with respect to this decision are far from static.[5] Thus inferring information about market conditions from previous decisions must also take into account the possibility

---

[3] For an excellent and concise review of the rational herding literature see Devenow and Welch (1996).

[4] Hirshleifer and Welch (1998) extend their basic model to allow the possibility that the underlying changes at a specific, predetermined time period.

[5] While we do not give an exact interpretation of the term "market conditions," we have in mind global and industry factors about which firms, through their underwriters, might get information. For recent popular press quotes see, for example, "IPO Outlook," *Wall Street Journal,* August 30, 1999; "Riding the IPO Rollercoaster," *Electronic Business*, September 1998; and "Offerings in the Offing: Good, not Great," *Barron's*, January 1998. "Net IPO Offerings Still Hot Despite Slowing Market," *Information Week*, June 1998; and "IPO Outlook: IPO Delays, Cancellations Are on the Rise," *Wall Street Journal*, April 19, 1999, show the importance of looking at market conditions for IPOs in the firm's particular industry, as opposed to the general IPO market conditions.

that conditions have changed in the interim period. The firm may get a private contemporary signal from its research (with the aid of its underwriter), but this signal is bound to be noisy. Therefore the inferred information is valuable.

Another example, where our model applies, is analysts' stock buy/hold recommendations (analysts' earnings forecasts can also be thought of in a similar manner).[6] The analyst's private research on the firm generates a private signal, and the final recommendation (or action) is public knowledge. An analyst may infer information about whether the stock price was over- or under valued at the time other analysts made their recommendations. However, this underlying value (the valuation of the company relative to the current stock price) may change over time, thus causing the correct recommendation to change over time.[7] Hence the current analyst's research (or signal) is more valuable than a previous analyst's signal.

In this article we extend BHW in two steps. The first step is to allow the underlying value (e.g., market conditions or the relative valuation of a company) to change over time, in a Markov fashion.[8] This allows us to address decisions in a changing environment such as the IPO decision, analysts recommendations, the choice of a restaurant to dine in, etc. Another assumption in BHW is that a correct decision leads to the same payoff, regardless of the state (underlying value). In the first step we maintain this assumption. We therefore call this model the *symmetric* model. The second step we take in this article is to allow different payoffs in different states for an agent (firm) that chooses correctly. We will call this the *asymmetric* model.

In the symmetric model (with a changing underlying), we show that herding waves are of finite, rather than infinite, length. That is, the sequences of firms that ignore their private information are finite. Herding occurs because some information about the current underlying value can be inferred from the previous firms' actions, even though meanwhile the underlying value may have changed. In particular, the aggregate information from many previous actions may be more compelling than the current private (and noisy) signal. This would induce a firm to herd. On the other hand, because of the changing underlying value, the information content of previous nonherding actions becomes stale as time passes. This induces the finiteness of herding waves.[9] The model therefore predicts phenomena such as bursts of activity

[6] Additional applications can be found in Section 3.2.

[7] BHW consider an example where the underlying value may change at a specific period preknown to all. As a result, they get a possible reversal of a cascade in progress at that period.

[8] Independently, Moscarini and Ottaviani, and Smith (1998) incorporated a changing underlying value as in the first-step model in this article. Their analysis focuses on showing that herding waves are of finite length and considering the upper bound on the length of a herd wave. Therefore the overlap with this article is in our Theorem 1 and Lemma 1, but not beyond.

[9] Avery and Zemsky (1998) present a model where herding on the wrong choice is eventually reversed. In their model (with a stationary underlying value) this is achieved by allowing a market maker to update prices and thus any *incorrect* valuation by the market is eventually reversed.

in the IPO market, the M&A market, or the adoption of new technologies. Gul and Lundholm (1995) introduce a similar concept of clustering in which agents' actions are similar, but not all the clustering agents' information is lost. They present cases in which there is no herding but clustering does occur. In their model some of each agent's information is revealed, while in ours there are agents whose information is totally lost (because they herd).

We find that, in the symmetric model, the length of every herding wave, once started, decreases in both signal quality and rate of change of the underlying. The frequency of herding decreases monotonically in the rate of change. We show that within certain parameter bounds there will be herding with positive probability, while outside these bounds no herding will occur. However, although one might think that higher signal quality will induce less herding (in the limit, when the signal is perfect, no other information is needed and no herding occurs), we discover that the frequency of herding is *not* monotone in signal quality. On the one hand, the increase in quality increases the value of the firm's private information and thus makes herd behavior less likely. On the other hand, when all firms have the same quality of signals, the quality of other firms' signals also increases and thus their actions convey more information, increasing the likelihood of herding. Some simulations demonstrating this dual effect are presented. In Section 3 we sketch a model in which these two phenomena can be separated. Using simulations, we also demonstrate the added efficiency from considering previous agents' actions, as opposed to just looking at the current private signal.

It is generally believed that herding induces additional correlation among firms' actions, beyond the correlation induced by the common underlying process (this is clearly true in BHW's model where, once a cascade starts, the correlation is 1). We examine this by comparing the herding model with a model where the signals are public and thus there is no herding.[10] We run simulations on both models to calculate the probability that, given that one firm chose one action, the next firm chooses the same action. Surprisingly we find that for most parameter values there is in fact a *lower* conditional probability (and thus correlation) in the herding model in comparison with the public signal model, particularly under the symmetric model. This happens because in the herding model, when firms are not herding, they are basing their decision on less information than firms in the public signal model (due to the loss of the herding firms' information). Therefore outside of a herding

[10] Perktold (1996) presents a model with repeated private signals in a macroeconomic setting where two infinitely lived agents act concurrently and repeatedly. Agents receive signals in each period and the underlying is a binary Markov chain. His model gives recurring informational cascades. This model is similar to partitioning agents into two groups that share information within the group (i.e., semi-public signal). The break in the cascades occurs partly because each agent collects private information, even in herding periods. With repeated signals, herding would break in the BHW setting as well. BHW give an example to illustrate how public release of information may break a cascade, because it provides additional information to all agents. In contrast, when a cascade is in progress, all the agents' private signals are lost. This lack of additional information is what keeps the cascade going.

wave there can be a high variability in the decisions of consecutive firms, based on their signals. This causes the overall correlation to be low even if the correlation within a herding wave is very high. In contrast, in the public signal model, agents have a relatively steady, large information base for their decisions since all signals are observable. This analysis suggests alternative ways for empirically testing for herding.

Allowing market conditions to change over time brings the model closer to a setting appropriate for examining the decision to go public. However, assuming that the firm gets the same payoff for correctly choosing to go ahead with an IPO, and for correctly choosing to wait, does not seem plausible. Therefore we next consider the asymmetric model where payoffs differ according to the state, when the firm chooses correctly. If the asymmetry is large, infinite cascades on the high payoff (go ahead with the IPO) state are possible. These cascades are induced mostly by the payoff asymmetry, but are also somewhat helped by information inferences. When the asymmetry is not too large, most of the qualitative results mentioned above carry through to the asymmetric model. The one notable exception is the result pertaining to the conditional probability that the current firm chose to IPO, given that the previous firm made that choice. This conditional probability is higher in the herding model than in the public signal model for many more parameter values (when compared with the symmetric model). However, if we consider the conditional probability of choosing to wait, the change is not as pronounced. Thus we can conclude that a burst of IPO activity may indeed be a symptom of herding, but a burst of firms deciding to delay their IPO is less likely to be a symptom of herding and more likely to be based on adverse information (see Section 2 for the intuition behind this result). In addition, the analysis presented in this article suggests characteristics that need to be demonstrated before an empirical finding of high correlation can be suggested as an indication of herding.

The remainder of this article is organized as follows: In Section 1 the basic dynamic, but symmetric, model is presented and analyzed. Although the model is general, it is set in the context of the decision to go public. We also discuss the long-run probability of herding and correlations between consecutive actions within the symmetric model. In Section 2 we present a model with asymmetric payoffs that corresponds more closely to the IPO application.[11] Section 3, contains a brief discussion of a possible extension of the model to incorporate different signal qualities across firms and contains a detailed discussion of applications of the model and its specific implications in these cases. Section 4 concludes and gives some possible future research directions.

[11] Appendix A derives an explicit (asymmetric) payoff structure for the IPO decision from a simple model of consequences of this decision.

1. The Symmetric Model

In this section we present and analyze a model of herd behavior when market conditions can change over time. Below are the model's characteristics. In Section 1.1 the firm's decision process is analyzed, using a representative statistic. We then consider comparative statics on the length and likelihood of herding. In Section 1.2 we use simulations to generate the long-run distribution of the representative statistic. Using this information we can demonstrate additional comparative statics as well as the relative correctness of firms' decision to herd. In addition, we consider a measure of the correlation among firms' actions and relate it to herding.

The model below is general, but it is presented with the terminology of the IPO decision in mind. In order to think of the decision to go public within the framework of this model, we need to keep in mind two issues. The first is that market conditions influence the firm's payoff from an IPO. All other things being equal, a firm is better off going public when market conditions are favorable. Second, it is not obvious from the outcome of previous IPOs whether market conditions will be favorable or not for the next IPO. This is both because market conditions change and because the observing firm does not know how much money the previous firm could raise under different conditions. The previous firm, which has already gone public, does know this, however. Therefore the decision of the previous firm to go ahead with an IPO *does* convey meaningful information to the observing firm. See Section 3.2 and the appendix for details on this application.

The *symmetric* model has the following main characteristics:

- The underlying true value of market conditions, $\langle M_t \rangle$, is a simple Markov process that in each period is either g (should go ahead) or w (should wait). The initial true market conditions, denoted by random variable M_1, are g with a prior probability of $0 < p < 1$. This probability, in essence, embodies all the known information about market conditions before the first firm needs to decide. The transition probabilities are as follows: if $M_t = g$, it has a small probability, $0 < q < \frac{1}{2}$, of changing to $M_{t+1} = w$. Similarly, $\Pr(M_{t+1} = g | M_t = w) = q$. Thus, q is the probability that the conditions will change in the current period.[12]
- In each period a different firm is active. The firm that is active in period t gets a private signal, $S_t \in \{g, w\}$. It then selects an action, A_t, which is either g (go ahead) or w (wait), trying to correctly infer the current market conditions, M_t. The firm's selection is public, thus firm t can take the actions of firms $1, \ldots, t-1$ into account (in addition to its private signal). The firm's signal, S_t, is the same as the true value, M_t, with probability $s > \frac{1}{2}$, that is, $\Pr(S_t = g | M_t = g) = s$ and

[12] If we allow $q > \frac{1}{2}$ then it is more likely that the state changes than not. This makes the actions in a herding wave toggle (i.e., g, w, g, w, etc.). The qualitative results do not change if we consider $1 - q$ instead of q.

$\Pr(S_t = w|M_t = w) = s$.[13] If firms are indifferent between the two possibilities, for didactic reasons, we assume that they choose to follow their private signal (this seems like the rule least conducive to herding). However, this tiebreaker is arbitrary, and any other tiebreaker can be chosen with the exact same results since indifference, unlike in BHW, is generically a zero probability event.

- The utility from "guessing" correctly (i.e., if $A_t = M_t$) is strictly higher than the utility from guessing incorrectly. In general we assume that the utility from guessing incorrectly is 0, the utility from guessing correctly in bad market conditions (i.e., correctly choosing "w") is 1, and the utility from choosing correctly in good market conditions (i.e., a correct choice of "g") is $c \geq 1$. In this section we will consider the benchmark model where payoffs are symmetric, that is, $c = 1$.[14] For simplicity we assume that firms are risk neutral.

These assumptions ensure that the firms need only maximize the conditional probability that $M_t = A_t$. In Section 2 we present an asymmetric model where c can be greater than 1.

1.1 Analysis

In this section a sufficient statistic for the firm's history set is presented and manipulated. This statistic will allow the firm to make its decision, based solely on its signal, and the statistic. Using this statistic, we examine the firms' choices and herding patterns, and find comparative statics on this statistic with respect to the model's parameters.

If a firm chooses differently for different signals, we say that it is *informative*, since its signal can be inferred from its action. If a firm decides on the same value regardless of its signal, we say that it is *herding*, since the information in its signal is lost. All firms are identical and actions are observed, therefore each firm can figure out which of the previous firms herded and which were informative. Also, since informative firms follow their signals and thus reveal them, firm n can deduce the signal of any informative firm before it. The current observable history, H_n, includes the actions A_i for $i < n$, as well as the set of all informative firms up to but not including firm n. Let y_n denote the probability that M_n is g given the current history; that is, $y_n = \Pr(M_n = g|H_n)$. Notice that $y_1 = p$.

The firm maximizes the conditional probability of being correct, that is, it decides to "go" if and only if $\Pr(M_n = g|S_n, H_n) > \frac{1}{2}$. If this probability is greater than half, regardless of the value of S_n then the firm will herd on

[13] If $s < \frac{1}{2}$ then we could consider the complement of the signal as the actual signal and thus be back in this setting.

[14] The appendix derives an appropriate c for the decision to go ahead with an IPO. The general asymmetric model, for any c including the one derived in the appendix, is considered in Section 2.

"go." Plugging in the expression for the conditional probability we get the following condition:

$$\Pr(M_n = g|S_n = w, H_n) = \frac{(1-s)y_n}{(1-s)y_n + s(1-y_n)} > \frac{1}{2}. \tag{1}$$

This can be simplified to the condition $y_n > s$.[15] When $y_n > s$, the firm will ignore a w signal and choose g regardless of its signal. This is because the information the firm infers from previous actions is so strong that one signal of quality s cannot change its decision.

Symmetrically, if it is more likely that the underlying state is bad, regardless of the signal, the firm will herd on "wait." Formally, this occurs when

$$\Pr(M_n = g|S_n = g, H_n) = \frac{sy_n}{sy_n + (1-s)(1-y_n)} < \frac{1}{2}. \tag{2}$$

This condition can be simplified to $y_n < 1-s$. This translates to herding on "go" when $y_n > s$ and herding on "wait" when $y_n < 1-s$. Therefore the nonherding (informative) region is $(1-s) < y_n < s$. In this region the firm follows its signal.

Firm $n+1$ can recursively find y_{n+1} from y_n. The firm considers the two possibilities that lead to the current market conditions being "go" now: the conditions were "go" last period and did not change, or the conditions were "wait" last period but they changed to "go" this period. To recursively get y_{n+1} from y_n we must first check whether firm n herded. If it did then

$$y_{n+1} = y_n(1-q) + (1-y_n)q. \tag{3}$$

Here the case where M_{n+1} was g and no switch occurred corresponds to $y_n(1-q)$, and the case where M_n was w and a switch occurred corresponds to $(1-y_n)q$. Notice that since firm n herded, its action adds no additional information for future firms. If firm n was informative, its signal (which is the same as its action) must be taken into account. The formula for y_{n+1} will then be different for $S_n = g$ and $S_n = w$, but will follow the same construction as Equation (3). For $S_n = g$ this construct is $y_{n+1} = q \cdot \Pr(M_n = w|S_n = g, H_n) + (1-q)\Pr(M_n = g|S_n = g, H_n)$. Substituting the expression from Equation (2) for $\Pr(M_n = g|S_n = g, H_n)$, and its corresponding complement, into the formula we get

$$y_{n+1} = \frac{sy_n(1-q) + (1-s)(1-y_n)q}{sy_n + (1-s)(1-y_n)}. \tag{4}$$

[15] Notice that the conditional probability of conditions being good given a g signal is always higher than the conditional probability given a w signal. Therefore if Equation 1 holds, then the former is also greater than half.

Similarly, for $S_n = w$, we use the construct $y_{n+1} = q \cdot \Pr(M_n = w | S_n = w, H_n) + (1-q)\Pr(M_n = g | S_n = w, H_n)$ and plug in the expression in Equation (1) and its complement. We get

$$y_{n+1} = \frac{(1-s)y_n(1-q) + s(1-y_n)q}{(1-s)y_n + s(1-y_n)}. \tag{5}$$

In these calculations, firm $n+1$ adjusts for firm n's signal and the probability of receiving that signal given the market conditions in period n. The probability of getting $S_n = g$, given the history H_n, is the denominator of the right hand-side of Equation (4), while the probability of getting $S_n = w$, given the history, is the corresponding denominator in Equation (5). The numerator is the probability, given H_n (*not* H_{n+1}), that market conditions at time $n+1$ are good ($M_{n+1} = g$) and S_n is as given [g for Equation (4), and w for Equation (5)].

Notice that the borders between the informative region and the two herding regions vary only with s, not q. Also notice that although the updated y_{n+1} in the herding regions is independent of s, it is s that determines how soon y_n exits the herding regions when herding occurs. To visualize the process y_n goes through, notice that each informative g signal gives y_n a step up (in the region $[0, 1]$) and each informative w signal gives it a step down. See Figure 1 for an illustration of this process.

Theorem 1 characterizes a necessary and sufficient condition for herding to occur with positive probability in this model.

Theorem 1. *For $q > s(1-s)$ there is no herding. For $q < s(1-s)$ herding will occur with positive probability, but if $q > 0$ all herding waves are of finite length.*

Proof. To find the no herding condition we check when $1-s < y_{n+1} < s$ always holds (for $1-s < y_n < s$); that is, the "go" update of Equation (4) must give less than s and the "wait" update of Equation (5) gives more than $1-s$. Since the formulas for y_{n+1} are increasing in y_n we need only check at the corresponding endpoints. For a g signal we check that the condition $y_{n+1} = \frac{s^2(1-q)+(1-s)^2 q}{s^2+(1-s)^2} < s$ holds for $y_n = s$. Since $s > \frac{1}{2}$ this simplifies to the condition $q > s(1-s)$. Symmetrically, for a w signal we check at $y_n = 1-s$, that is, that $\frac{(1-s)^2(1-q)+s^2 q}{(1-s)^2+s^2} > (1-s)$ holds. This simplifies to the same restriction, namely $q > s(1-s)$. Therefore we have shown that for $q > s(1-s)$ there is no herding.

In order to show that for $q < s(1-s)$ there will be herding with positive probability, consider Equation (4). This update function is increasing in y_n for $q < s(1-s)$ and for $y_n \in [s, 1-s]$. Subtracting y_n from this formula

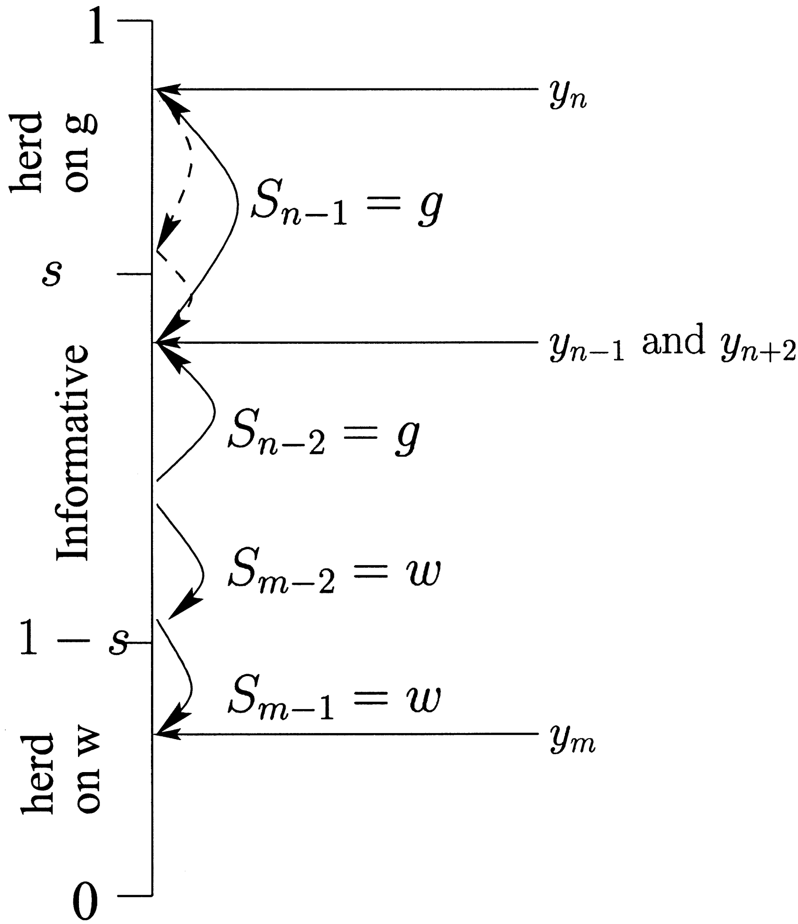

Figure 1
Updating steps, given private signal, for y_n, the conditional probability that the underlying market conditions are good: $M_n = g$. The dashed arrows represent herding firms (signal irrelevant).

we get

$$\frac{q(1-s)+y_n[(1-q)s-(q+1)(1-s)]+y_n{}^2(1-2s)}{sy_n+(1-s)(1-y_n)} \geq \frac{(2s-1)(s(1-s)-q)}{s} > 0. \tag{6}$$

This means that every informative g signal increases y_n by a step whose size in bounded away from zero. Thus given enough g signals in a row we will get into the go-herd range ($y_n > s$); that is, with positive probability we will herd on go (for $q < s(1-s)$). A similar argument can be used for the "wait" update function of Equation (5).

From Equation (3), we see that when there is herding y_n reverts to $\frac{1}{2}$. Since the update during herding is of step size greater or equal to $(2s-1)q > 0$, we get that every herding wave ends in a finite number of steps. This shows the last part of Theorem 1. ■

The intuition of this theorem is that if the rate of change is large, then older actions (signals) bear less information about the current market conditions. Also, the better the quality of the signal the smaller the rate of change that is needed to prevent herding altogether. The reason all herding waves are finite is the possible change in market conditions. When there is a herding wave, as time passes, the informative firms' actions become less and less relevant for the current decision. Eventually they are no longer compelling enough to cause herding and firms will consider their signals and become informative.

We now turn to some comparative statics derived from the updating equations [Equations (3), (4), and (5)], beginning with q. From Equation (3) we see that when there is herding, y_n reverts to $\frac{1}{2}$. In fact, notice that the larger is q the faster Equation (3) reverts to $\frac{1}{2}$ and the faster the herding phase will be broken. Due to the discreteness of the steps it is possible, however, to increase q by so little that the number of steps does not change. Therefore the above statement is only true with a weak inequality. In addition, starting at the same point in the informative region, and everything else being equal, for a lower q we will jump deeper into a herding region than we would with a higher q.[16] Using similar arguments, we can get that the length of the shortest possible informative phase (i.e., the case where all signals are the same) is increasing in q. These results are summarized in Lemma 1.

Lemma 1. *Holding all other parameters fixed, and starting from the same point y_i, the larger is q, the shorter (weakly) the herding phase, and the longer (weakly) the shortest possible informative phase.*

The intuition for the lemma is very similar to that of Theorem 1, namely the higher the rate of change the more relevant the current signal is, compared to all previous signals, for determining current market conditions.

Since the update step for $n+1$, when n has herded, is independent of s, and s decreases the size of the herding regions as it increases, we get that the higher the s the faster we get out of the herding regions. Notice, however, that this does not necessarily mean that firms will herd less often since the informative updating rules, Equations (4) and (5), do depend on s.

[16] Taking derivatives with respect to q of the update equations [Equations (4) (g signal) and (5) (w signal)] we get $\frac{1-s-y_n}{1-s-y_n+2sy_n} < 0$ and $\frac{s-y_n}{s+y_n-2sy_n} > 0$, respectively, where the inequalities hold for $1-s < y_n < s$. This means that starting at the same point in the informative region, and everything else being equal, for a lower q we will jump deeper into a herding region.

Lemma 2. *Holding all other parameters fixed, and starting at a y_i in a herding region, the larger is s, the less (weakly) the herding persists.*

In the next section we will consider the long-run distribution of y_n. It will demonstrate that even though the length of each individual herding wave, once started, may be shorter the higher the s (all other things being equal), the probability of herding in the long run does not change monotonically in s.

1.2 Long-run behavior

In this section we consider three issues for which we need the long-run distribution of y_n. The first is how this long-run distribution (and thus the long-run probability of herding) changes as the parameters change. The second is the efficiency gain that observable actions (and the herding that goes with it) induce. Finally, the third issue we consider is the correlation among firms' actions. This correlation is important because of the general perception that herding induces higher correlation among agents' actions. Given that one firm chose "go," we examine the probability that the next one will go ahead with an IPO as well. This conditional probability is computed for the herding model as well as for a benchmark model where signals are public. The two models are compared.

Ideally these issues would be examined analytically, but, as mentioned above, the long-run distribution of y_n is crucial in both cases. Figures 2 and 3 give the long-run distribution of y_n for two pairs of parameter values (as

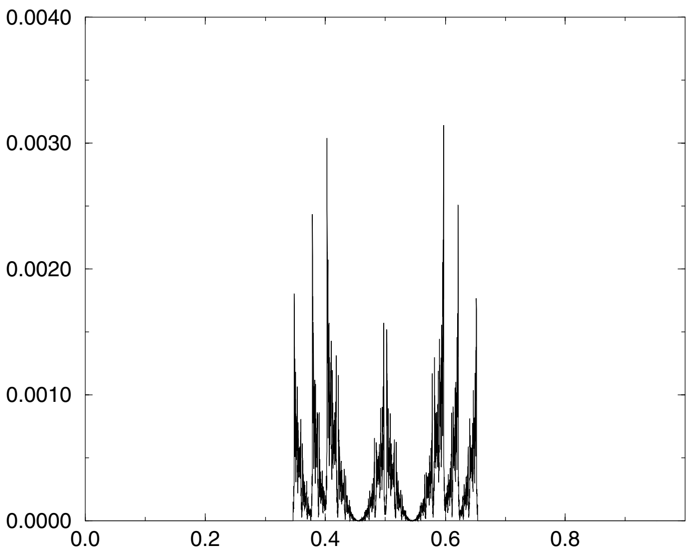

Figure 2
Distribution of $y_n = \Pr$ (underlying is $g|H_n$) for $q = 0.1$ and $s = 0.6$.

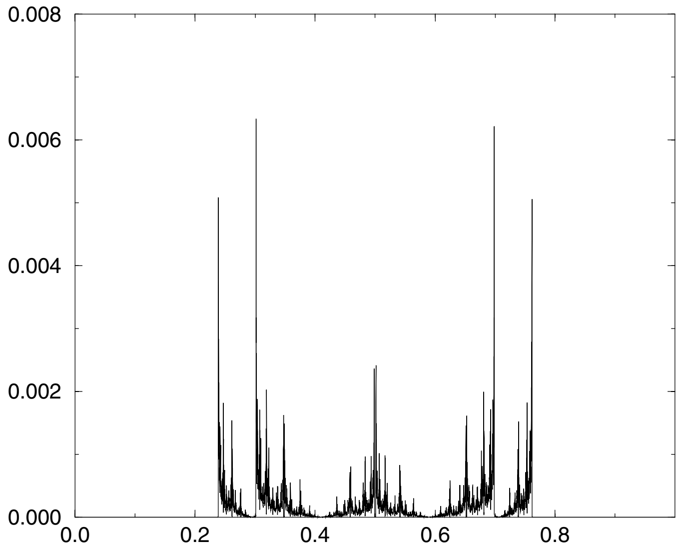

Figure 3
Distribution of $y_n = \Pr$ (underlying is $g|H_n$) for $q = 0.12$ and $s = 0.7$.

found from simulations). As can be seen, the long-run distribution of y_n is rather erratic. This stems mainly from cycles of various sizes that happen with varying frequency. For example, in Figure 2, the highest three peaks on the right form a cycle among y_n values 0.597, 0.651, and 0.621, where the first firm is informative and got a g signal and the next two herd after it and go ahead with the IPO as well. The next firm (the fourth one) then has $y_n = 0.597$, the same as the first firm had. This cycle is similar to the one illustrated in the top part of Figure 1. If the fourth firm gets a g signal another identical cycle commences, otherwise the cycle breaks. There are many such cycles of varying lengths (and herd/informative combinations) possible and the system can move from one cycle to the other depending on the private signals. This configuration makes the long-run distribution analytically intractable. Therefore we use simulations of the long-run distribution to examine the frequency of herding, the fraction of firms that choose correctly, and the correlation among firms' actions.

Table 1 presents simulation results of the probability of herding (either on "go" or on "wait") for various values of s and q.[17] The probability is

[17] The simulations presented here were run on the model starting with $p = 0.5$ and for various q and s values. For each set of values the system was run for 1,000,000 firms. The same simulations were run for other

Table 1
Portion of agents that herded for various signal qualities and rates of change—symmetric model

	s							
q	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95
0.01	0.917	0.916	0.913	0.911	0.904	0.891	0.861	0.767
0.02	0.841	0.840	0.838	0.832	0.820	0.800	0.734	0.618
0.04	0.710	0.707	0.704	0.690	0.670	0.635	0.571	0.438
0.06	0.598	0.599	0.583	0.572	0.528	0.526	0.398	0.0
0.08	0.511	0.513	0.489	0.467	0.446	0.360	0.366	0.0
0.1	0.412	0.421	0.420	0.375	0.328	0.348	0.0	0.0
0.12	0.360	0.344	0.285	0.302	0.315	0.192	0.0	0.0
0.14	0.258	0.267	0.278	0.286	0.188	0.0	0.0	0.0
0.16	0.247	0.252	0.240	0.169	0.0	0.0	0.0	0.0
0.18	0.213	0.193	0.147	0.089	0.0	0.0	0.0	0.0
0.2	0.126	0.115	0.075	0.0	0.0	0.0	0.0	0.0
0.22	0.083	0.048	0.0	0.0	0.0	0.0	0.0	0.0

zero for values for which $q > s(1-s)$, as implied by Theorem 1. Notice that although as s increases (fixing q) the probability of herding decreases, this is not true for all q and for all values of s. For example, for q between 0.07 and 0.1 there is an increase when going from the second highest to the highest applicable s in the table (where herding may occur). There are also increases in the probability of herding for middle values of s. For instance, for $q = 0.12$, going from $s = 0.7$ to $s = 0.8$ there is an increase in the probability of herding. In other simulations, not presented here, we checked the average length of the herding and informative phases for various s and q values. We found that the average length of a herding phase is decreasing in both q and s, while the average length of an informative phase seems to decrease in q, but does not change monotonically in s. This is consistent with Table 1, as well as with Lemmas 1 and 2.

Intuitively there are two opposite effects at work here. On the one hand, for a higher s, the current signal is more accurate (and it is the most recent and hence most relevant single signal). On the other hand, the predecessors' signals are also more accurate, making their combined value higher. If the increase in the aggregate value is stronger, we will see more herding for a higher s, otherwise the better current signal dominates and we will see less herding. The model sketched in Section 3 gives a way to separate these two effects by allowing different firms to have different signal qualities.

In another set of simulations we compared our symmetric model to a model where the signals are public rather than private. The rest of the public signal model's characteristics are the same as those listed in the beginning of this section.[18] In the public signal model the representative statistic is

p values and other (large) numbers of firms and, as expected, there was no difference in the results. The differences for the values we got across simulations (for the same parameters) were less than ± 0.001, or $\pm 0.1\%$.

[18] This model is similar to the benchmark used in Hirshleifer and Welch (1998) with a static underlying. In their benchmark it is the same manager that gets all the signals and makes all the decisions.

$x_n = \Pr(M_n = g|S_1, \ldots, S_n)$. Notice that unlike the representative statistic, y_n, for the herding model, x_n conditions on firm n's signal as well as all previous signals. The decision rule is: choose g i.f.f. $x_n > \frac{1}{2}$. Firm $n+1$ updates its belief according to x_n and its *own* signal, S_{n+1}, in the following manner:

- If $S_{n+1} = g$, then

$$x_{n+1} = \frac{s(x_n(1-q)+q(1-x_n))}{s(x_n(1-q)+q(1-xn))+(1-s)(x_nq+(1-q)(1-x_n))}.$$

- If $S_{n+1} = w$, then

$$x_n = \frac{(1-s)(x_n(1-q)+q(1-x_n))}{(1-s)(x_n(1-q)+q(1-x_n))+s(x_nq+(1-q)(1-x_n))}.$$

Again, the denominators are the probability of receiving the observed S_{n+1}, and the numerators are the probability of receiving the observed S_{n+1} and having $M_{n+1} = g$.

The difference between the herding model, presented earlier in this section, and this model is that in the public signal model the information of firms who ignore their signal is not lost. The loss of information is a socially undesirable effect of herding and reduces the efficiency of future choices. In Table 2 we compare the fraction of firms that chose the correct action in both models, for various parameter values (for which there is herding). Notice that in the extreme case, where only the firm's signal is observable, a fraction s of the firms will choose correctly. We will denote by $\kappa_h(q,s)$, $\kappa_p(q,s)$, and $\kappa_1(q,s)$ the fraction of correct firms for the herding, public signal, and private signal only cases, respectively ($\kappa_1(q,s) = s$).

Table 2
Fraction of firms that have chosen correctly for the herding model, $\kappa_h(q,s)$, and for the public information model, $\kappa_p(q,s)$—symmetric payoffs

		s							
q	Model	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95
0.01	Herd	0.637	0.697	0.759	0.810	0.856	0.898	0.935	0.967
	Public	0.798	0.865	0.908	0.935	0.953	0.967	0.975	0.986
0.05	Herd	0.622	0.681	0.735	0.785	0.831	0.876	0.915	
	Public	0.671	0.740	0.796	0.839	0.876	0.908	0.931	
0.09	Herd	0.615	0.669	0.720	0.766	0.817	0.859		
	Public	0.639	0.697	0.751	0.798	0.838	0.873		
0.13	Herd	0.606	0.662	0.711	0.757	0.803			
	Public	0.618	0.676	0.726	0.772	0.813			
0.17	Herd	0.602	0.653	0.702	0.751				
	Public	0.608	0.661	0.709	0.754				
0.21	Herd	0.601	0.651						
	Public	0.603	0.652						

For parameter values where the models differ, i.e., where there is herding with positive probability.

As expected, both the herding model and the public signal model do better than the private signal only case, for parameter values where there is herding. When there is no herding, each firm follows its signal and thus all three models are equivalent and perform equally well. In such environments (without herding) there is no added value from publicly revealing to the firm additional information beyond its private signal. However, if signals are private, and the rate of change is not too large, there is a social advantage to revealing actions (e.g., by mandating disclosure of actions). By comparing Table 2 with Table 1 we see that a larger advantage of herding (i.e., a bigger $\kappa_h(q, s) - \kappa_1(q, s) = \kappa_h(q, s) - s$, where $\kappa_h(q, s)$ is from Table 2) corresponds to a larger fraction of firms that herd (Table 1). This is intuitively appealing. Note that, as Lemma 1 suggests, when the rate of change, q, increases, the difference $\kappa_h(q, s) - \kappa_p(q, s)$ between the herding and public signal models decreases. However, as q increases both $\kappa_h(q, s)$ and $\kappa_p(q, s)$ get closer to the private signal only case, $\kappa_1(q, s) = s$.

Next, we use simulations to check how the information loss affects the correlation among actions. This is particularly interesting since, in the empirical literature, high correlation has often been equated with herding. We will check this by computing a conditional probability. Specifically, we are interested in Pr (firm chooses "go" | predecessor chose "go").[19] Analysis of the BHW model, with the static underlying, shows that this conditional probability will be higher in their herding model than in the public information model. In fact, the conditional probability will go to one when a cascade occurs.

In our dynamic model, a second effect is introduced. Namely, because the herding firm's decision is based on less information than its parallel in the public signal model, a bad signal right after the end of a herding wave may sway the firms temporarily in the wrong direction.[20] In contrast, in the public signal model a larger number of bad signals are needed to sway the firms in the wrong direction. This, in general, induces higher volatility in the actions of the informative firms in the herding model, and thus less correlation between consecutive actions. Running simulations allows us to see which of these two opposing effects (long herding waves versus volatility outside of the herding wave) dominates for various parameter values.

In Table 3 we see the conditional probability above for various parameter values where herding occurs.[21] For small rates of change and low signal

[19] Since the model is perfectly symmetric, this is the same probability as Pr(firm chooses "wait" | predecessor chose "wait").

[20] This is similar to the notion of fragility presented in BHW. They show that even a small amount of information can break a cascade in progress. In fact, they show that even a small probability of new information coming can break the cascade with a higher probability than the probability of getting the information at all.

[21] The simulations presented here were run on the model starting with $p = 0.5$ and for various q and s values. For each set of values the system was run for 10,000,000 firms. The same simulations were run for other p values and other (large) numbers of firms and, as expected, there was no difference in the results. The differences for the values we got across simulations (with the same parameter values) were less than ± 0.001.

Table 3
Pr(firm chooses 'go' | predecessor chooses 'go') for *s* and *q* where there is herding. In the blank boxes there is no herding and the two models are equivalent

		s							
q	Model	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95
0.01	Herd	**0.959**	**0.960**	**0.962**	0.964	0.965	0.966	0.968	0.969
	Public	0.940	0.952	0.961	0.967	0.967	0.979	0.968	0.984
0.05	Herd	0.832	0.837	0.842	0.850	0.853	0.857	0.860	
	Public	0.834	0.845	0.861	0.872	0.880	0.905	0.918	
0.09	Herd	0.739	0.741	0.752	0.768	0.751	0.790		
	Public	0.764	0.774	0.782	0.813	0.835	0.852		
0.13	Herd	0.664	0.661	0.678	0.703	0.725			
	Public	0.705	0.720	0.736	0.759	0.777			
0.17	Herd	0.624	0.632	0.637	0.636				
	Public	0.659	0.670	0.682	0.686				
0.21	Herd	0.564	0.568						
	Public	0.601	0.601						

Cases were there is a higher conditional probability for the herding model are marked in bold.

quality we indeed get that there is more correlation in the herding model when compared with the public signal model (the long herding waves dominate). However, for most values of q and s, the second effect of volatility outside of a herding wave dominates and we in fact get a *lower* conditional probability in the herding model than in the model with public information (e.g., for $q = 0.1$ and $s = 0.7$ the herding model gives 73.9% while the public signal model gives 77.2%). These findings imply that correlation may not be a good indication for herding when the underlying value changes over time.[22] Notice that, in general, as our intuition would suggest, the higher the rate of change (q) the lower the correlation (for both models). On the other hand, as we saw in Table 1 for the herding model, the effect of an increase in s is not monotone in the public signal model as well. This happens in both models since the two effects (the current signal is better versus previous signals are also better) are present in both models. There is some difference across models in the parameter values where one effect dominates the other.

We have demonstrated that action correlation may not be a good proxy for herding. This presents the question of what is a better way of testing for herding. The model presented above suggests considering a different direction. When looking at Figure 1, we notice that the probability that a firm is correct in its decision declines as a herding wave progresses; that is, the nth firm in a herding wave has a higher probability than the $n+1$st firm of being correct. On the other hand, on average, the quality of information is constant in the public information model. Therefore the model suggests that the ex post success rate within sequences of firms choosing the same action

[22] In other simulations we computed lagged conditional probabilities and found similar patterns, although the values at which there was higher correlation in the herding model changed somewhat.

is more variable in a herding scenario than in a public signal scenario.[23] We believe that this criterion deserves closer empirical scrutiny and may give insight into the robustness of the existing empirical herding literature.

The analysis above suggests that uniformly increasing the quality of the information available to firms will not necessarily decrease herding. On the other hand, as demonstrated in Section 3.1, holding the quality of all other signals constant, an increase in the current firm's information quality *will* reduce the probability of that firm herding. The results above would also indicate that if payoffs were symmetric for the decision to go public, then correlation between the actions of firms may be due to public information rather than herding. However, assuming that the payoff of going public in good market conditions and that of waiting in bad market conditions is the same is not plausible. In the next section we will expand the model above to allow asymmetric payoffs. We then check which of the above results are robust to this adjustment.

The symmetric model can be applied to many problems, such as BHW's classic restaurant choice problem. The decisions for which the model is appropriate are those where the correct answer may change over time. In the restaurant example, it is possible that the quality of food and service has changed (e.g., through change of ownership or chef). Therefore the model presented in this section would be more applicable to this situation than the static BHW model. Additional choices for which this model is applicable range from the choice of technology to use (e.g., operating system), through the use of certain financial instruments (e.g., convertible debt), to labor market decisions (e.g., the decision to give an offer to a candidate, as well as the candidate's decision to go to one of several similar firms). Some of these applications were discussed in BHW and in Devenow and Welch's (1996) review of the herding literature. Further details on the applicability of this model are discussed in Section 3.2.

2. The Asymmetric Model

The asymmetric model is identical to the model presented in Section 1, except for the payoff structure. Rather than assuming state-independent payoffs, we allow one state to have a higher payoff than the other, when the firm chooses correctly. We assume that the payoff for an incorrect choice is 0. For simplicity, we normalize the payoff for a correct choice when market conditions are bad (i.e., a correct choice of "wait") to 1. The payoff for a "go" choice in good market conditions is denoted by c, where $c \geq 1$.[24] In

[23] Simulations of the models presented here show that for almost all parameter values this is the case. The exception is when signal quality is relatively low and at most 1 firm herds in any herding wave. In these cases there is relatively little herding, and it is hard to distinguish between the two models in terms of variability of success.

[24] Using payoffs a and $c \cdot a$, for any constant a, works equally well.

particular, these payoffs can be derived from a simple model of the consequences of the decision to go public (see Appendix A for this derivation).

Most of the intuitions from Section 1 regarding the symmetric model follow through to this model. In particular the recursive updating rules for y_n, Equations (3)–(5), are unchanged since the different payoffs do not affect the conditional probability that market conditions are favorable. What does change is the firm's decision rule. Previously the firm would ignore its signal and "blindly" decide on g if $y_n > s$ (see Figure 1), now the boundary is different because of the different payoffs. The firm will decide g even when it gets a w signal if, given its w signal, its expected payoff for deciding g is higher than its expected payoff for deciding w. In other words, $c \cdot \Pr(M_{n+1} = g | y_n, H_n, S_{n+1} = w) > \Pr(M_{n+1} = w | y_n, H_n, S_{n+1} = w)$. This translates to the condition that $y_n > \frac{s}{(1-s)c+s}$. Similarly the firm chooses w regardless of a g signal if $c \cdot \Pr(M_{n+1} = g | y_n, H_n, S_{n+1} = g) < \Pr(M_{n+1} = w | y_n, H_n, S_{n+1} = g)$, which is equivalent to the condition $y_n < \frac{1-s}{sc+1-s}$. In the interim region ($\frac{1-s}{sc+1-s} \leq y_n \leq \frac{s}{(1-s)c+s}$), the firm follows its signal.

The decision space for the asymmetric model, in terms of y_n, is depicted in Figure 4. As can be seen in the figure, there seems to be much more "room" for herding on "go" as opposed to herding on "wait." In fact, if $c > \frac{s}{1-s}$ then $\frac{s}{(1-s)c+s} < \frac{1}{2}$ and the "go" herding region includes the point $\frac{1}{2}$, which is the point of convergence of the herding update rule [Equation (3)] (see right illustration). In this case, once a firm herds on "go" a cascade starts and all firms after it will also herd on "go." This is not true of herding on "wait," since $\frac{1-s}{sc+1-s} < \frac{1}{2}$ always holds. Notice that the two illustrations in Figure 4 will typically have different s values, and hence different-size jumps for the informative g signal. On the other hand, since the updating in the herding regions is independent of s, the same size of jumps applies (assuming we hold q fixed). These results are summarized in the following theorem.

Theorem 2. *For $c \geq \frac{s}{1-s}$, with positive probability, firms will be trapped in a "go" cascade forever, but if $q > 0$ every "wait" herding wave will be of finite length. For $c < \frac{s}{1-s}$ and $q > 0$, every herding wave is of finite length.*

The intuition for Theorem 2 is a mixture of the intuition for Theorem 1, and the effect of an especially large payoff under good market conditions. Namely, on the one hand, if there is a positive rate of change then old information becomes stale and may induce the firm to break a herding wave and consider its private signal. This causes herding waves to be of finite length. On the other hand, if the payoff for one state is much higher than the payoff for the other, then for any conditional probability above $\frac{1}{2}$ (and for some below $\frac{1}{2}$) the firm will want to choose "go" regardless of its signal. Notice that the conditional probability y_n will never dip below $\frac{1}{2}$ without the help of other firms' signals, but since these firms are herding their private information is lost. Therefore the conditional probability that market conditions are

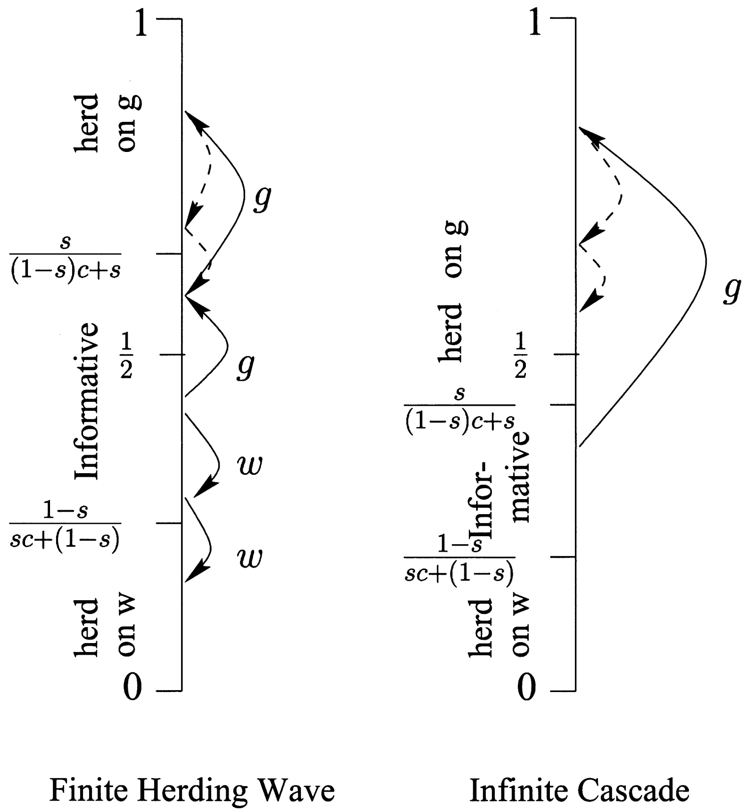

Figure 4
Changes in y_n for the asymmetric model, when $\frac{s}{(1-s)c+s}$ is less than or greater than $\frac{1}{2}$. The solid lines are updates due to informative agents, and the dashed lines are updates due to herding agents. Notice that either the s value or the c value (or both) should be different for the two illustrations.

good will remain high.[25] This brings us back to a situation similar to the one in BHW's original model and an infinite cascade will form. For payoffs that are bounded from above ($c < \frac{s}{1-s}$), many of the intuitions of the symmetric model carry through to the asymmetric model. In particular, Lemma 1 holds as stated.

Corollary 3. *In the asymmetric model, if $c < \frac{s}{1-s}$, then holding all other parameters fixed, and starting from the same point y_i, the larger is q, the shorter (weakly) the herding phase, and the longer (weakly) the shortest possible informative phase.*

[25] If the conditional probability that market conditions are good is $\frac{1}{2}$, and there is no additional information, then the probability that market conditions are good next period is also $\frac{1}{2}$. Thus, without additional information, all future y_ts will be $\frac{1}{2}$.

This means that, in general, the more likely conditions are to change the less herding will occur, provided that the discrepancy in payoffs is not too large.

The payoff c affects the decision of the firm only through the boundaries between the three different regions (herd on "go," informative, and herd on "wait"). Both boundaries decrease (toward 0) as c increases, but the upper boundary moves faster than the lower boundary, increasing the area of the "go" herd region faster than the area of the "wait" herd region shrinks. Since there is also a larger probability of ending up in the "go" herd region to begin with (because of the higher payoff) this means that the higher the c the more herding on "go" we expect to observe. When $c \geq \frac{s}{1-s}$ herding on "go" becomes a cascade. This result is summarized in the following lemma.

Lemma 3. *The higher the discrepancy in payoffs, c, the more likely firms are to herd on "go" and the less likely firms are to herd on "wait."*

Simulations show that the net effect is that there is more herding (in aggregate) for a higher c. In particular, as stated in Theorem 2, when c is large enough, the "go" herd wave becomes an infinite cascade.

Additional simulations demonstrate that the comparative statics with respect to signal quality, s, are a muted version of those for the symmetric model, that is, there is no monotone relationship between signal quality and the probability of herding. This is because the two opposing effects discussed in Section 1.2 (the current firm's signal being better versus everyone else's signals also being better) are present in the asymmetric model as well. Simulation results for two specific values of c are given in Tables 4 and 5. Consider Table 4 ($c = 1.2$). Although for most values the probability of herding decreases as s increases, there are values where there is an increase in herding as s increases (e.g., for $q = 0.08$ when s changes from 0.75 to 0.8). Notice that there are fewer such instances in Tables 4 and 5 than in Table 1, the corresponding table for the symmetric model. The fact that other firms' signals become more valuable when s increases has less influence in the asymmetric model. This is because herding on "go" is partially attributable to the informational inferences (as in the symmetric model) but also partially due to the higher payoff. The signal quality only affects the former consideration, therefore its net effect is relatively smaller.

In order to examine the effect of herding on the correlation among consecutive actions, we apply the procedure outlined in Section 1.2 to the asymmetric model. To this end, we define a public signal model that is identical to the model outlined in the beginning of this section, except that all signals are public. We ran simulations on both models and Tables 6 and 7 report the conditional probability Pr (firm chooses "go" | predecessor chose "go") for both models for various values of s and q. In Table 6 we used $c = 1.2$ and in Table 7 we used $c = 1.51$ (paralleling Tables 4 and 5, respectively). Notice that, compared to the symmetric model, there are far fewer pairs of s and q for which the public signal model has a higher conditional probability than

Table 4
Portion of firms that herded in the asymmetric model for $c = 1.2$

	s							
q	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95
0.02	**0.861**	**0.848**	**0.838**	**0.832**	**0.817**	**0.789**	**0.740**	**0.584**
0.04	**0.745**	**0.724**	**0.707**	**0.691**	**0.664**	**0.619**	**0.494**	0.225
0.06	**0.649**	**0.621**	**0.611**	**0.571**	**0.553**	**0.455**	**0.399**	0.0
0.08	**0.568**	**0.532**	**0.503**	**0.479**	**0.422**	**0.401**	0.202	0.0
0.10	**0.498**	**0.447**	**0.403**	**0.394**	**0.376**	**0.276**	0.117	0.0
0.12	**0.428**	**0.365**	**0.366**	**0.334**	**0.254**	0.181	0.0	0.0
0.14	**0.356**	**0.358**	**0.284**	**0.231**	0.170	0.099	0.0	0.0
0.16	**0.316**	**0.258**	**0.209**	0.162	0.149	0.0	0.0	0.0
0.18	**0.268**	**0.187**	0.156	0.140	0.075	0.0	0.0	0.0
0.20	0.189	0.161	0.141	0.081	0.0	0.0	0.0	0.0
0.22	0.189	0.151	0.086	0.0	0.0	0.0	0.0	0.0
0.24	0.187	0.093	0.043	0.0	0.0	0.0	0.0	0.0

Bold entries denote those parameter values for which there was herding both on "go" and on "wait," other positive entries denote herding only on "go."

the herding model. For example, in Table 7, for $s \leq 0.8$, in the vast majority of table entries there is a higher conditional probability in the herding model than in the public signal model. The effect becomes even stronger when c is larger. These observations imply that in the case of asymmetric payoffs (especially when the asymmetry is relatively pronounced), a higher tendency to choose the high-payoff action may indeed be a symptom of herding.

Unlike in the symmetric model, the discrepancy in payoffs creates an asymmetry in the conditional probabilities of choosing "go" or choosing "wait." The conditional probability of choosing w is presented in Tables 8 and 9 (for $c = 1.2$ and $c = 1.51$, respectively). Comparing Tables 6 and 8, for example, demonstrates that for a choice of "go" there are many more parameter values for which the conditional probability is higher in the herding

Table 5
Portion of firms that herded in the asymmetric model for $c = 1.51$

	s							
q	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95
0.02	1.000	**0.879**	**0.852**	**0.838**	**0.818**	**0.788**	**0.751**	**0.582**
0.04	1.000	**0.781**	**0.733**	**0.709**	**0.670**	**0.645**	**0.542**	0.224
0.06	1.000	**0.704**	**0.640**	**0.605**	**0.525**	**0.514**	**0.411**	0.218
0.08	1.000	**0.624**	**0.543**	**0.512**	**0.485**	**0.383**	0.203	0.0
0.10	1.000	**0.558**	**0.502**	**0.453**	**0.358**	0.236	0.198	0.0
0.12	1.000	**0.499**	**0.410**	**0.336**	0.258	0.187	0.144	0.0
0.14	1.000	**0.451**	**0.313**	0.279	0.186	0.181	0.0	0.0
0.16	1.000	0.415	0.298	0.190	0.180	0.164	0.0	0.0
0.18	1.000	0.352	0.260	0.185	0.175	0.065	0.0	0.0
0.20	1.000	0.335	0.204	0.181	0.128	0.0	0.0	0.0
0.22	1.000	0.326	0.203	0.168	0.080	0.0	0.0	0.0
0.24	1.000	0.256	0.203	0.099	0.0	0.0	0.0	0.0

Bold entries denote those parameter values for which there was herding both on "go" and on "wait," other positive entries denote herding only on "go."

Table 6
Pr(firm chooses 'go' | predecessor chooses 'go') for various *s* and *q* and for *c* = 1.2—asymmetric model (for parameter values where there is herding)

		s							
q	Model	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95
0.01	Herd	**0.975**	**0.970**	**0.969**	**0.969**	0.969	0.969	0.971	0.972
	Public	0.944	0.953	0.962	0.968	0.971	0.977	0.978	0.985
0.05	Herd	**0.890**	**0.872**	**0.869**	0.869	0.870	0.873	0.878	0.891
	Public	0.855	0.857	0.868	0.874	0.876	0.907	0.924	0.911
0.09	Herd	**0.825**	0.798	0.797	0.780	0.794	0.792	0.811	
	Public	0.805	0.801	0.804	0.812	0.834	0.850	0.842	
0.13	Herd	**0.779**	0.742	0.729	0.708	0.730	0.747		
	Public	0.769	0.750	0.758	0.772	0.776	0.783		
0.17	Herd	0.731	0.670	0.670	0.686	0.671			
	Public	0.735	0.718	0.707	0.712	0.714			
0.19	Herd	0.695	0.656	0.661	0.644				
	Public	0.722	0.689	0.686	0.696				
0.21	Herd	0.664	0.650	0.629	0.616				
	Public	0.715	0.676	0.673	0.652				

Cases were there is a higher conditional probability for the herding model are marked in bold.

model, when compared with the choice of "wait." This is because there are longer herding waves (or even infinite cascades) on g than on w. Curiously there are more entries in Table 8 (choosing "wait" for $c = 1.2$ in the asymmetric model), than in Table 3 (the symmetric model), where the conditional probability is higher in the herding model. This is because, in the asymmetric model due to the "lure" of the higher payoff, the choice of "wait," the lower payoff action, is based on more substantial information than in the symmetric model. This decrease in fragility results in less volatility in the actions

Table 7
Pr(firm chooses 'go' | predecessor chooses 'go') for various *s* and *q* and for *c* = 1.51—asymmetric model (for parameter values where there is herding)

		s							
q	Model	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95
0.02	Herd	**1.000**	**0.966**	**0.957**	**0.951**	**0.950**	0.949	0.950	0.953
	Public	0.921	0.924	0.932	0.942	0.949	0.952	0.958	0.973
0.06	Herd	**1.000**	**0.910**	**0.886**	**0.870**	**0.865**	0.866	0.875	0.882
	Public	0.872	0.859	0.859	0.869	0.862	0.888	0.908	0.902
0.10	Herd	**1.000**	**0.857**	**0.828**	**0.809**	0.800	0.806	0.807	
	Public	0.853	0.823	0.815	0.808	0.824	0.815	0.837	
0.14	Herd	**1.000**	**0.821**	0.765	**0.767**	0.732	0.750		
	Public	0.838	0.796	0.766	0.762	0.768	0.783		
0.18	Herd	**1.000**	**0.768**	0.725	0.695	0.709	0.687		
	Public	0.838	0.764	0.755	0.733	0.742	0.727		
0.22	Herd	**1.000**	0.755	0.687	0.678	0.648			
	Public	0.840	0.762	0.732	0.712	0.687			

Cases were there is a higher conditional probability for the herding model are marked in bold.

Table 8
Pr(firm chooses 'wait' | predecessor chooses 'wait') for various s and q and for $c = 1.2$—asymmetric model (for parameter values where there is herding)

		s							
q	Model	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95
0.01	Herd	**0.954**	**0.960**	**0.962**	0.965	0.967	0.969	0.970	0.971
	Public	0.937	0.949	0.959	0.967	0.971	0.978	0.977	0.985
0.05	Herd	0.799	0.820	0.840	0.847	0.857	0.863	0.869	0.880
	Public	0.808	0.832	0.853	0.856	0.873	0.906	0.922	0.895
0.09	Herd	0.705	0.728	0.748	0.756	0.762	0.790	0.783	
	Public	0.719	0.746	0.772	0.794	0.819	0.843	0.800	
0.13	Herd	0.623	0.654	0.680	0.701	0.714	0.698		
	Public	0.645	0.678	0.712	0.737	0.746	0.710		
0.17	Herd	0.561	0.586	0.610	**0.626**	0.626			
	Public	0.568	0.606	0.622	0.600	0.634			
0.19	Herd	**0.534**	**0.562**	**0.577**	0.583				
	Public	0.516	0.536	0.560	0.589				
0.21	Herd	0.513	0.529	0.551	0.575				
	Public	0.515	0.531	0.554	0.578				

Cases were there is a higher conditional probability for the herding model are marked in bold.

outside of a herding wave following a choice of "wait" (as compared with following a choice of "go").

As was demonstrated in this section, the symmetric and asymmetric models have similar characteristics. The notable exception is that when payoffs are asymmetric the correlation in the herding model is higher than that of the corresponding public signal model for many more parameter values. Therefore, in cases where the market conditions may change over time, simply demonstrating correlation among consecutive actions is not tantamount to demonstrating herding. Instead, one must either argue that the payoffs are highly asymmetric, or show that the rate of change is extremely low.[26] Without demonstrating these characteristics, the correlation may be simply due to the fact that this is the correct answer (as in the public signal model). In addition, Tables 6–9 demonstrate that the high correlation is more likely to be due to herding when the firms are all choosing to go ahead with an IPO (herding on the high payoff state). If many firms decide to postpone their IPOs (i.e., "wait," the low payoff choice), then it is more likely that this is because market conditions are indeed unfavorable.

3. Extensions and Implications

In Section 3.1 we propose an extension of the symmetric model that allows some firms to have better signal quality than others (e.g., due to a superior underwriter). This allows us to consider the role of experts in and out of herding waves. In Section 3.2, applications of the symmetric and asymmetric

[26] Even after demonstrating extremely asymmetric payoffs the herding does not immediately follow. One must also argue that the quality of the signal is not too high. In the case of predicting market conditions for an IPO, this is a reasonable assumption.

Table 9
Pr(firm chooses 'wait' | predecessor chooses 'wait') for various s and q and for $c = 1.51$—asymmetric model (for parameter values where there is herding)

		s							
q	Model	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95
0.02	Herd	N/A	**0.916**	**0.926**	0.933	0.937	0.941	0.946	0.951
	Public	0.884	0.905	0.921	0.936	0.947	0.947	0.958	0.973
0.06	Herd	N/A	0.782	**0.812**	0.821	0.832	0.845	0.863	0.871
	Public	0.742	0.783	0.806	0.841	0.851	0.880	0.900	0.885
0.10	Herd	N/A	0.667	0.699	0.729	0.758	**0.785**	0.775	
	Public	0.636	0.691	0.731	0.763	0.785	0.739	0.791	
0.14	Herd	N/A	**0.612**	**0.646**	**0.671**	**0.679**	0.693		
	Public	0.536	0.596	0.619	0.620	0.661	0.705		
0.18	Herd	N/A	0.538	0.562	0.592	0.627	0.662		
	Public	0.521	0.542	0.569	0.599	0.634	0.669		
0.22	Herd	N/A	0.531	0.552	0.577	0.606			
	Public	0.517	0.533	0.557	0.582	0.611			

Cases were there is a higher conditional probability for the herding model are marked in bold. In the spaces marked "N/A" no firm in the simulation chose 'wait' (cascade on 'go').

models to several decisions, including the decision to go public, are discussed in more detail. In addition, implications of the models are drawn.

3.1 Varying signal quality

In this section we briefly examine an extension of the symmetric model, introduced in Section 1, where the signal quality, s, may vary among firms (the asymmetric model can be extended in a similar way). This is intuitively appealing since we may think that some firms are better informed than others, and indeed analysis of this case produces intuitive results. For example, firms with superior information will tend to trust it (and not herd) more than firms with inferior information. This means, for example, that we would expect to see firms with an inferior underwriter herding after firms with a superior underwriter, but not usually vice versa. Applying this model to analysts' recommendations, rather than the IPO decision, produces the implication that less-expert analysts will herd after expert analysts [see Stickel (1990)].

Suppose that firm $n-1$ has signal quality s_{n-1}, while firm n has quality s_n. Firm n's signal quality enters its own decision since it determines the boundaries of the herding regions. The distance between the upper and lower bounds of the informative region from half increases in s_n. Thus, ceteris paribus, a firm with a higher s_n is less likely to start a herd wave and is more likely to break a herding wave (become informative).

In the case of many different s values, if $q > \underline{s}(1-\underline{s})$, where $\underline{s}$ is the minimal s, then no herding is possible. Also, following similar reasoning to Theorem 1, if $q < \overline{s}(1-\overline{s})$, where $\overline{s}$ is the maximal s, then herding occurs with positive probability. For $\overline{s}(1-\overline{s}) < q < \underline{s}(1-\underline{s})$, it is not clear whether there will be herding or not. In particular, the existence of herding may depend on the ordering of the experts (firms with high s_n) in the sequence of firms.

Firm n is influenced by the quality of firm $n-1$'s signal through the informative updating rules [see Equations (4) and (5).[27] Taking derivatives with respect to s of these rules shows that S_{n-1} influences firm n's decision more the higher s_{n-1} is (for $S_{n-1} = g$ we get a higher y_n and the converse holds for $S_{n-1} = w$). This corresponds to preferring to follow an expert rather than a novice.

3.2 Applications and implications

Both the symmetric model, presented in Section 1, and the asymmetric model, presented in Section 2, can be applied to a variety of decision problems. In this section we will discuss some of these applications and the model's implications in their context.

For exposition purposes, we presented the model in the context of a firm's decision to go public, considering its beliefs on market conditions. The firm observes the decisions of previous firms (delay the IPO or go ahead). The firm also receives a private recommendation (signal) from its underwriter concerning their expectations for market conditions. The firm then updates its beliefs on expected market conditions and, considering the possible payoffs, makes its decision. The model implies that while there may be herd behavior on the decision to go ahead with an IPO, it is less likely that firms herd in their decision to postpone an IPO. This latter decision is more likely based on unfavorable information. In addition, the model implies that these results are stronger the larger the discrepancy in payoffs between the choices. This discrepancy may change from industry to industry. Our analysis also suggests that in industries for which IPO market conditions are rapidly changing, no herding will occur. In such industries a frenzy of IPOs is likely to be due to favorable conditions.

The asymmetric model applies to a variety of other decisions in which payoffs are asymmetric. Examples of such decisions are the adoption of new financial products and instruments such as LBOs or asset-backed securities. Executing a timely successful LBO has a large payoff, while the relative payoff for a correct decision that no LBO should be undertaken probably has a smaller payoff. Another such decision is the choice of technology when one technology is superior to another in terms of possible payoffs, but it is not always the superior technology that is embraced by the market (e.g., VHS versus Beta).[28] The model predicts that firms' decisions to adopt the

[27] Other previous firms' signal quality also influences firm n, but this was already taken into consideration by firm $n-1$ in forming y_{n-1}. In addition, there is a decay in the relevance of previous signals the more periods have passed since the signal was current.

[28] Alternatively, the early adoption (or rejection) of a new technology can also be an application of the asymmetric model. Correctly adopting a new technology (e.g., e-business) has a much higher payoff than correctly choosing to pass.

low payoff technology is more likely to stem from information about market conditions, as opposed to herding. The model also implies that there could be a cascade on the high payoff choice even if market conditions are no longer favorable (and thus neither the high nor the low payoff are realized).

As mentioned in Section 1, the symmetric model can be applied to those decisions in which the underlying state changes over time but payoffs are symmetric [see BHW (1992) and Devenow and Welch (1996)]. For example, it is plausible that even though restaurant A starts off being better than restaurant B, the relative ranking could reverse over time (e.g., as the owner or the chef changes). The symmetric model implies that herding on a restaurant will eventually break, unlike in the BHW model. In addition, it implies that a higher rate of change causes shorter lengths of herding waves. This seems to be anecdotally true of "in" restaurants.

The symmetric model also naturally applies to analysts' buy/hold recommendations. The correct recommendation changes over time, and the payoff for choosing correctly is independent of the choice.[29] The model implies that a high correlation among analysts' recommendations is not necessarily a sign of herding.[30] The model also suggests that, if there is herding, analysts will tend to herd after experts. Similar implications would hold for analysts' earnings forecasts. Stickel (1990) looked at analysts' earnings forecasts and found that change in (previous) mean analyst forecasts had strong predictive power for the current analyst forecast. He interprets this as "following the crowd."[31] However, as suggested in Section 1.2, the correlation result alone does not necessarily imply "following the crowd."

The symmetric model can be applied to a job applicant's choice of employer. The applicant can use the interviewing process to assess the merits of the companies (e.g., the value of option packages, work conditions, benefits, etc). This is a private assessment (signal). It is plausible that prior decisions by other applicants convey additional information to the current applicant about these issues. It is also natural to assume that the relative desirability of two companies may change over time as the firms evolve. The symmetric model applies best when the two firms have very similar prospects. The model predicts that following a few successful hires for one company it will be easier

[29] The symmetric model applies more to the buy/hold recommendation as opposed to the buy/sell decision. There is anecdotal evidence that a sell recommendation has negative implications for the institution issuing the warning. In this case the asymmetric model applies and we would conclude that the choice of a sell recommendation is based on information, as opposed to herding. This is not necessarily the case for a buy recommendation.

[30] Welch (2000) looks at analysts' buy/sell recommendations. He finds that recommendations that are bunched closer together in time show more correlation. As Welch suggests, this could be explained by short-term changes that occur around the bunching time. His analysis is consistent with the discussion in Section 1.2. Welch also finds evidence that analysts tend to herd after the consensus. He checks herding by considering the ex post correctness of the consensus choice, as opposed to relying only on correlations.

[31] Stickel also finds that higher quality, "all-American" analysts are less likely to follow the crowd when compared with other analysts. Stickel (1992) checks that analysts on Institutional Investor's "All-American Research Team" do indeed give better predictions, on average, than other analysts. These findings are consistent with the implications of the model with varying signal qualities, discussed in Section 3.1.

for that company, for a while, to hire. On the other hand, the model implies that the firm will not be able to benefit from the inertia (herding) in its hiring success indefinitely.

Both models have some general implications that pertain to all applications, including those above. First, in general, the larger the rate of change the less herding we expect. Second, Section 3.1 implies that firms (agents) will be more likely to herd after experts than after novices. Third, Table 3 and Tables 6–9 demonstrate that high correlation does not immediately imply herding. In the symmetric model, correlation is higher in the herding model than in a public signal model only for very low rates of change and relatively low signal qualities. In the asymmetric model, however, high correlation is predicted mostly following a choice that may bring a high payoff. Even in this case, the correlation is not uniformly higher; that is, there are many parameter values for which the public signal model exhibits higher correlation than the herding model.

4. Concluding Remarks

In this article we examined a general model that is based on Bikhchandini, Hirshleifer, and Welch (1992) but allows the underlying value (e.g., market conditions) to change over time. Initially we preserve BHW's assumption of uniform payoffs for the right choice, independent of the underlying state. In lieu of infinite cascades, as in BHW, we get recurring herding waves, each of finite length. The length of each specific herding wave decreases in the signal quality and in the rate of change. However, we found that the frequency of herding, as well as the average length of the informative (nonherding) waves, does not change monotonically in signal quality. We also examined the correlation between two consecutive firms' actions and found that it is generally lower than in the benchmark case where all signals are public. This stems from the fragility of beliefs in the herding model which causes action choices to vary outside a herding wave.

Next, we considered a generalization of the model above in which payoffs *are* state dependent. This asymmetric model readily applies to many decisions, including the decision to go public. Deciding to go public when market conditions are favorable (corresponding to the "go" state/decision in the general model) clearly has a higher payoff than deciding to wait in adverse market conditions (corresponding to the "wait" state/decision).[32] Our analysis shows that for a high enough asymmetry in payoffs, there are infinite cascades on "go." If the difference in payoffs is not too large, all herding waves are finite. In both cases, firms are much more likely to herd on "go" than to herd on "wait" (for some parameter values there is no herding

[32] See Appendix A for a derivation of this payoff discrepancy in a simple model.

on "wait" but there *is* herding on "go"). In addition, the conditional probability that a firm chooses "go," given that its predecessor choses "go," is higher in the herding model (as opposed to the public signal model) for quite a few parameter values. In particular, this is true for many more parameter values than in the symmetric model. In contrast, the conditional probability of choosing "wait" (given that the firm's predecessor chose "wait") is lower in the herding model for most parameter values. Applying these findings to the decision to IPO, for example, implies that while firms may be (rationally) herding in their decision to IPO, they are less likely to herd when deciding to postpone. Postponement is more likely based on the firm's private signal about market conditions.[33] Lastly, we presented an extended model where different firms are allowed to have different signal qualities.

An interesting direction for future research is the effect of the ordering of firms on the total percentage of firms that go public. For example, could NASDAQ create a wave of IPOs in a certain area (such as internet applications) by encouraging certain companies to go public at certain times. If these decisions created a herding wave many more companies will follow and go public.[34] Controlling the critical positions which can create a herding wave affects the actions of other firms, the amount of publicly available information, and the social desirability of the outcomes.

Appendix A: Deriving IPO Payoffs

In Sections 1 and 2 we presented and analyzed a model of herd behavior when conditions can change over time. In this section we present a framework that can be superimposed on the general herd behavior models. This framework addresses the decision of going ahead with an IPO or delaying. In particular, we develop in this section a payoff structure that is applicable to this decision. Section 2 uses just such a payoff structure.

For simplicity we assume that the firm's payoff from an IPO can have one of two values, depending on market conditions.[35] In order to derive the exact payoffs for the firm's actions and simplify them to the structure we use in Section 2, let us consider the current firm. Assume that this firm has filed with the SEC for an IPO offering and now faces a two-period game: in the first period it can decide whether to go ahead with the IPO or delay, and in the second

[33] See, for example, "IPO Outlook: IPO Delays, Cancellations Are on the Rise," *Wall Street Journal*, April 19, 1999, for anecdotal evidence.

[34] Welch (1992) alludes to this possibility in regard to IPO subscriptions within a model with a stationary underlying. By having the first few customers subscribe, a cascade is created, thus inducing many more to subscribe and the IPO to succeed. Welch stresses that these initial subscribers are critical. The criticality of first movers is a characteristic of BHW as well, and is criticized by Shiller (1995). In our models there are positions, determined by the parameters and the signal realizations, which have more influence on other firms' decisions than others. However, these positions are interspersed in the sequence of firms and, if the payoff discrepancy is not too large, none of them have an irrevocable influence on the outcomes, since every herding wave eventually breaks of its own. On the other hand, if the high payoff is much larger than the low payoff, an infinite cascade can be started by strategically placed firms.

[35] Implicitly this means that we're assuming market conditions affect firms' IPO payoffs in a similar way. However, we do not explicitly model how or why market conditions affect the IPO outcome. These conditions can include market expectation for global and industry factors that influence valuations, such as interest rates (expectations) or beliefs on market growth of a sector.

period, if it delayed, it will go public for sure.[36] There is a discount factor between the two periods, r. As mentioned above, the payoff from going public when market conditions are good is $V_t + H$ and when market conditions are bad the payoff is $V_t + L$, where V_t is firm specific and $H >> L$. We will assume that $V_t = 0$ in order to simplify the discussion, but V_t can take on any value, since it is, by definition, independent of market conditions and thus does not affect the firm's decision. As in the rest of this article, market conditions are Markov (with switching probability q). We will denote true market conditions with capital letters to differentiate them from the chosen actions. The expected payoffs $\pi(a, m)$ for any pair (a, m), where $m \in \{G, W\}$ is (true) market conditions and $a \in \{g, w\}$ is the period 1 action the firm chose, are as follows:

$$\pi(g, G) = H; \;\; \pi(g, W) = L \tag{7}$$

$$\pi(w, G) = \frac{(1-q)H + qL}{1+r} \tag{8}$$

$$\pi(w, W) = \frac{(1-q)L + qH}{1+r}. \tag{9}$$

The payoffs for $a = w$ are just expected discounted cash flows. For example, consider Equation (8) where market conditions are G. In this case, with probability $1 - q$, conditions will remain G in the second period and the payoff will be high (H). With probability q, conditions will switch to W and the payoff will be low (L). These payoffs are discounted by one period to get their expected present value. Similarly for $\pi(w, W)$ [Equation (9)]. We can subtract the constant $\pi(w, G) - \pi(g, W)$ from all payoffs in state W and then normalize all the payoffs so that the payoffs for the pairs (g, W) and (w, G) are 0.[37]

Denote the normalized values by $f(a, m)$. Since $H >> L$ we get that $f(g, G) > f(w, W) > f(g, W) = f(w, G) = 0$. Following this normalization, $f(g, G)$ is higher than $f(w, W)$ by a factor of

$$\frac{(q+r)H + qL}{qH - (q+r)L}.$$

This factor will be denoted by the constant c in Section 2. As we have demonstrated, this derived, normalized, payoff structure is equivalent to a richer payoff structure that is directly derived from the decision to go ahead with an IPO. In Section 2 we use the simple payoff structure to check the robustness of the results obtained in Section 1.

References

Avery, C., and P. Zemsky, 1998, "Multi-Dimensional Uncertainty and Herd Behavior in Financial Markets," *American Economic Review*, 88, 724–748.

Banerjee, A., 1992, "A Simple Model of Herd Behavior," *Quarterly Journal of Economics*, 107, 797–818.

Bikhchandani, S., D. Hirshleifer, and I. Welch, 1992, "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades," *Journal of Political Economy,* 100, 992–1027.

[36] This game is a simplification, of course, but it can easily be justified if there is a large penalty for delaying an IPO too long (e.g., due to the market's inferences from such a move).

[37] This can be done since firms are risk neutral and therefore merely maximize the expected value of $\pi(a, m)$. The firm will choose g (go ahead) if

$$y\pi(g, G) + (1-y)\pi(g, W) > y\pi(w, G) + (1-y)\pi(w, W),$$

where y is the probability that market conditions are good, given all of the firm's information. It follows from this equation that adding a fixed value to both $\pi(g, W)$ and $\pi(w, W)$ will not change the firm's decision. Further, normalizing all $\pi(\cdot, \cdot)$ payoffs by adding/subtracting a constant also does not change the firm's decision.

Chamley, C., and D. Gale, 1994, "Information Revelation and Strategic Delay in a Model of Investment," *Econometrica,* 62, 1065–1085.

Devenow, A., and I. Welch, 1996, "Rational Herding in Financial Economics," *European Economic Review,* 40, 603–615.

Gul, A., and R. Lundholm, 1995, "Endogenous Timing and the Clustering of Agents' Decisions," *Journal of Political Economy,* 103, 1039–1066.

Hirshleifer, D., and I. Welch, 1998, "A Rational Economic Approach to the Psychology of Change: Amnesia, Inertia, and Impulsiveness," working paper, UCLA.

Moscarini, M., M. Ottaviani, and L. Smith, 1998, "Social Learning in a Changing World," *Economic Theory,* 11, 657–665.

Perktold, J., 1996, "Recurring Informational Cascades," working paper, University of Chicago.

Shiller, R., 1995, "Conversation, Information, and Herd Behavior," *American Economic Review,* 85, 181–185.

Stickel, S., 1990, "Predicting Individual Analyst Earnings Forecasts," *Journal of Accounting Research,* 28, 409–417.

Stickel, S., 1992, "Reputation and Performance Among Security Analysts," *Journal of Finance,* 47, 1811–1836.

Welch, I., 1992, "Sequential Sales, Learning, and Cascades," *Journal of Finance,* 47, 695–732.

Welch, I., 2000, "Herding Among Security Analysts," *Journal of Financial Economics,* 58, 369–396.