Abstract

The literature on informational cascades and herding theory has for a decade focused on the externality and suboptimal outcomes generated from decision-making when action spaces are coarser than private information spaces. Much of the output has therefore been positive, not normative. This paper redresses this imbalance by detailing several direct applications for marketing and business arising from herding theory. We see that business practices such as encouraging early sales, or selling to groups rather than individual customers, can be justified theoretically by a direct application of herding theory.

Keywords: informational cascade, herding, marketing strategies.
JEL Classification: D82, D83, L15, M30.

CONTROLLING THE HERD: APPLICATIONS OF HERDING THEORY

Daniel Sgroi [1]
Churchill College
University of Cambridge

1. Introduction

The literature on informational cascades and herding theory since Banerjee (1992) has focused on the externality and suboptimal outcomes generated from decision-making when action spaces are coarser than private information spaces. Much of the output has therefore been descriptive: stressing the *herd externality* and the difficulty of those within a herd to avoid the informational pressure to repeat the actions of their predecessors. Very little work has focused on normative suggestions of how to break herds, or better yet, manipulate them to ensure higher sales for firms. In many cases the existence of a herd will be suggestive of a particular business response. This paper redresses the imbalance in the literature by listing several direct applications for marketing and business arising from herding theory. This paper builds on the work in Sgroi (2000), which focuses on the manipulation of ordering in a sequential herd, designed to maximize sales for a firm, or welfare from a social planner's perspective. However this forms just one of the four business practices which include encouraging early sales, manipulating biased consumers and selling to groups rather than individual customers.

1.1. **Overview.** The first section deals with lessons drawn from herding models which focus on sequential decision-making. In particular we examine the role of forced early decision-making and biased consumers when a firm wants to maximize the chance of a herd acting in favor of it's product. The second section switches attention to herding when agents can wait, so the timing of decisions becomes endogenous. Here we look at the choice between individual sales methods and group sales when the firm is aiming

[1]The author would like to thank Huw Dixon, Paul Klemperer, Meg Meyer, Andrew Temple, an anonymous referee for the working paper series and seminar participants in Oxford and Helsinki for many useful comments. Financial support from AEA Technology is gratefully acknowledged. The author can be contacted via email at daniel.sgroi@econ.cam.ac.uk.

to maximize sales, and the worth of gathering more information when the firm is the consumer. In each section on applications remarks are made which might be construed as possible informal policies intended to help firms which face the prospect of informational cascades on or against their products. We finish with some conclusions.

2. Sequential Herding

In sequential herd models the main features are the coarseness of actions spaces relative to signal spaces, the exogenously determined sequential ordering of decisions by consumers, and the externality generated by this information structure. Once a herd begins public information will swamp private information and all later decision-makers will copy the actions of their predecessors, but no new information will be available to later decision-makers so their will be no scope for later agents to "break" the herd. This section sets up a version of the model first used in the seminal herd paper by Bikchandani, Hirshleifer, and Welch (1992), and then goes on to analyze direct business applications of this model, initially by summarizing Sgroi (2000).

2.1. A Simple Model.

Consider a sequence of $N \in \mathbb{N}_{++}$ agents, the ordering of which is exogenous and common knowledge, each deciding whether to purchase (Y) or not purchase (N) some product. Each agent observes the actions (Y or N) of his predecessors. The cost of purchase is $C = \frac{1}{2}$, and results in the gain of V which has prior probability $\frac{1}{2}$ of returning 0 or 1. The agents each receive a conditionally independent signal about V defined as $X_i \in \{H, L\}$ for agent i. The signals are informative in the following sense.

Definition 1. *Signals are informative, but not fully-revealing, in the sense that:*

$$\Pr[X_i = H \mid V = 1] = \Pr[X_i = L \mid V = 0] = p \in (0.5, 1)$$

$$\Pr[X_i = H \mid V = 0] = \Pr[X_i = L \mid V = 1] = 1 - p \in (0, 0.5)$$

Define the history up to agent n as the set of actions of agents 1 to $n-1$ so $H_{n-1} \equiv \{A_1, A_2, ..., A_{n-1}\}$ where $A_i \in \{Y, N\}$. Now define the information set of agent i as $I_i \equiv \{H_{i-1}, X_i\}$. It will be the case that in certain circumstances X_i will be inferable from A_i but this will not always be true. Now define N^{odd} as the set of agents from N indexed by only odd numbers from $\mathbb{N}_{++}$, and equivalently define N^{even}. Define also $\mathbb{N}^{odd}$ as the set of odd numbers in $\mathbb{N}_{++}$, and equivalently define $\mathbb{N}^{even}$. Define $E[\pi_i]$ to be agent i's *ex ante* expected payoff (i.e. his expected payoff before his signal draw). Finally define $\#X_i$ as the number of signals or actions of type X_i drawn or taken up to and including agent i. Now $X_1 = H \Leftrightarrow A_1 = Y$ and $X_1 = L \Leftrightarrow A_1 = N$. Agent 2 can infer agent 1's signal, X_1, from his action, A_1, and so has an information set $I_2 = \{X_1, X_2\}$. If $X_2 = H$ and $A_1 = Y \Rightarrow X_1 = H$ then agent 2 adopts so $A_2 = Y$. If

$X_2 = H$ and $A_1 = N \Rightarrow X_1 = L$ or if $X_2 = L$ and $A_1 = Y \Rightarrow X_1 = H$ agent 2 will have two conflicting signals so we require a tie-breaking rule. We use a simple coin-flipping rule which is known to all agents:

Condition 1. *(Tie-breaking rule) If I_i includes an equal weighting of H and L signals then* $\Pr[A_i = Y] = \Pr[A_i = N] = \frac{1}{2}$. *This rule is common knowledge.*

Consider a possible chain of events. The *first agent* will purchase if $X_1 = H$ and reject if $X_1 = L$. The *second agent* can infer the signal of the first agent from his action. He will then purchase if $X_2 = H$ having observed purchase by the first agent. If he observed rejection but received the signal $X_2 = H$ then he will flip a coin following the tie-breaking rule. If he receives $X_2 = L$ and $A_1 = N$ then he too will choose $A_2 = N$. If the first agent purchased then he would be indifferent and so flip a coin. The *third agent* is the first to face the possibility of a herd. If he observed two purchases, so $H_2 = \{Y, Y\}$ then $A_3 = Y$ for all X_3 since he knows that $X_1 = H$ and the second agent's signal is also more likely to be H than L, so the weight of evidence is in favor of purchase *regardless of* X_3. This initiates a Y *cascade*: the *forth agent* will also adopt as will the fifth, etc. Similarly if the third agent observes that both previous choices were rejections then he too will reject, initiating a N *cascade*. An informational cascade occurs if an individual's action does not depend upon his private information signal. The individual, having observed the actions of those ahead of him in a sequence, who follows the behavior of the preceding individual, without regard to his own information, is said to be in a cascade. A model-specific definition would be:

Definition 2. *Informational Cascades.* *A Y cascade is said to occur if $A_{i-1} = Y \Rightarrow A_i = Y$ for all X_i. A N cascade is said to occur if $A_{i-1} = N \Rightarrow A_i = N$ for all X_i.*

For the sake of clarity define the *initiator* of a herd or cascade as the agent whose decision to go Y or N makes the following agent's signal irrelevant. The cascade *traps* the agent who first faces a deterministic optimal choice regardless of his signal value, and all subsequent agents. So in the case of $H_2 = \{Y, Y\}$ a Y cascade is initiated by agent 2 and agent 3 finds himself trapped in the Y cascade. Note that if $H_2 = \{Y, N\}$ or $H_2 = \{N, Y\}$ then agent 3 will be in the same position, pre-signal draw, as agent 1. Note also that if agent 3 finds himself trapped in a cascade so too will agents $4, 5, 6, ..., N$. We should note that a cascade once started will last forever, even if it is based on an action which would not be chosen if all the agents' signals were common knowledge. Finally, the possibility of convergence to the incorrect outcome through the loss of information contained in later agents' private signals might be phrased in terms of a discernible negative *herd externality* as suggested by Banerjee (1992). A social planner will wish to minimize the impact of this negative externality on consumer welfare. In some cases it will also be in the interests of a firm to work against this externality. As we see later though in some cases a firm will actually use this externality to its own advantage, when it wishes to sell a low quality product.

From the model specifications we can derive the unconditional *ex ante* probabilities of a Y cascade, N cascade, or no cascade after n agents. Define $Y(n)$ to be a Y cascade initiated by agent n and similarly define $N(n)$ for a N cascade and $No(n)$ for no cascade by agent n. For example, $\Pr[Y(2)]$ is simply the probability that the first two agents both choose Y. The following functions conditional on $V = 1$ are fully calculated in part 1 of the appendix. After an even number of n agents we have:

$$\Pr[Y(n) \mid V = 1] = \frac{p(p+1)}{2}\frac{1-\left(p-p^2\right)^{\frac{n}{2}}}{1-(p-p^2)} \tag{2.1}$$

$$\Pr[N(n) \mid V = 1] = \frac{(p-2)(p-1)}{2}\frac{1-\left(p-p^2\right)^{\frac{n}{2}}}{1-(p-p^2)} \tag{2.2}$$

These expressions allow us to make a number of clarifying remarks. Note that from equation 2.1 $\Pr[Y(n) \mid V = 1]$ is increasing in p and n but from equation 2.2 we have that $\Pr[N(n) \mid V = 1]$ is high even for p much higher than $\frac{1}{2}$. Therefore, even when a great majority of the signals are of type H, a product still faces the prospect of a possible herd against its purchase. This is worrying for both consumers and for a firm with a high quality product. The symmetric case where $V = 0$ would apply when the product is of low quality, and the results provide some hope for the manufacturer of such a product, since there is always the chance of a Y cascade. Of course there is no reason for a firm to stay passive in the face of such potential herds.

2.2. **Early Sales and "Guinea Pigs".** Consider a single firm with a product it wishes to sell facing a sequence of consumers as described in section 2.1. Abstracting from profit-maximization, in the context of the current model we will consider the firm's aim to be simply to sell as many units as possible. We will first consider the optimal strategy of a firm which has a "good" product, i.e. for which $V = 1$ and then consider a firm which has a "bad" product, i.e. for which $V = 0$. We allow the firm to force a fixed number of consumers to decide early. This fixed group we will call $M \subset N$. This alters the information structure of decision-making producing a large initial set of decisions $(M + 1)$ followed by the remaining $N - M - 1$ deciding as before in sequence. The key feature is that the initial group will not be able to learn from observing others so must decide on the basis of their own private information; however, later decision-makers will have a larger pool of inferable signals to act upon. We might imagine the firm approaching a sub-set of all consumers and offering some incentive to make a quick decision. This can come in a variety of forms. The firm might send time-limited money-off coupons to certain potential consumers, or offer free products to high profile consumers or members of the press who agree to advertise their experience through writing a review. Perhaps the best example is that of a movie premiere full of high profile celebrities and members of the press whose opinion will be sought. The rest of this section formally examines these ideas. The reader is advised to refer to Sgroi (2000) for a more formal step by step examination of the firm's problem.[2]

2.2.1. *Good Products to Sell.* Assume first that $V = 1$ so the decision to purchase is the right one. Define the number of units sold as $Q_N(M+1) \equiv \#Y_N(M+1)$ which is a function of $M+1$ for a population of agents of size N and simply reads the number of Y decisions made by a population of N agents when there are an additional M guinea pigs chosen to decide with the first agent. In terms of the model the firm's objective is clearly to maximize the number of units sold. In order to do this the firm faces an important trade-off.

1. It wishes to maximize the probability of a Y cascade by choice of M, since this will raise the number of purchases by those outside the initial decision group. For any given choice of M there will only be a remainder population outside the group of guinea pigs of size $N-M-1$, so the population which learns is of size $N-M-1$. Therefore the firm is interested in ensuring that this remainder population opts for a Y cascade, so intuitively it is interested in maximizing $(N-M-1)\Pr[Y(M+1) \mid V=1]$. A Y cascade will be initiated by the group of $M+1$ guinea pigs if $Q_{M+1} \geq \frac{M+1}{2}+1$, a N cascade will be initiated if $Q_{M+1} \leq \frac{M+1}{2}-1$ or alternatively there will be no net public information and no cascade will occur if $Q_{M+1} = \frac{M+1}{2}$. Having noted this it is easy to see that the probability of a Y cascade being initiated by a given number of $M+1$ guinea pigs will be $\Pr\left[Q_{M+1} \geq \frac{M+1}{2}+1 \mid V=1\right]$.

2. It also wishes to sell its product to as many of the guinea pigs as possible. The sales to the first $M+1$ is very simply defined as $p(M+1)$ since there will be no learning within this group.

Furthermore, the firm also knows that even if a Y cascade is not initiated by the initial group of guinea pigs later agents may still initiate a Y cascade. Part 2 of the appendix reduces the firm's problem to:

$$\max_M \left\{ p(M+1) + (N-M-1)(M+1)! \sum_{x=0}^{\frac{M-1}{2}} \frac{p^{M+1-x}(1-p)^x}{x!(M+1-x)!} + \right. \tag{2.3}$$

[2] Many of the calculations from Sgroi (2000) are reproduced in the parts 2 and 3 of the appendix.

$$\left(1-\sum_{x=0}^{\frac{M-1}{2}}\frac{(M+1)![p^{M+1-x}(1-p)^{x}+p^{x}(1-p)^{M+1-x}]}{x!(M+1-x)!}\right)\frac{p(p+1)}{2}\sum_{n=2}^{N-M-1}\frac{\left(N-M-\frac{n}{2}\right)\left(1-\left(p-p^{2}\right)^{\frac{n-M-1}{2}}\right)}{1-(p-p^{2})}\Bigg\}$$

Differentiating this requires the use of the digamma and hypergeometric distributions and produces a fairly complex result. Some comparative statics should provide some intuition. Table I gives the optimal choice of M for various values of p and N.

Table I: Optimal M Values for the Firm when $V = 1$ for Given Values of p and N

p		0.6	2/3	3/4	4/5
N	50	23	19	13	11
	100	37	27	17	15

Table II gives the expected number of units sold for various different choices of M by the firm for a market of size $N = 100$ for different values of p.

Table II: Expected Units Sold for Different Values of M, $N = 100$

M		9	29	49	69	89
	0.51	41	48	51	51	51
p	2/3	77	87	83	77	70
	4/5	95	94	90	86	82

Table III holds p constant at $2/3$ and varies the size of the market, again looking at the impact on the expected number of units sold (with percentage of market size in brackets) of a change in M.

Table III: Expected Units Sold for Different Values of M, $p = \frac{2}{3}$

M		9	29	49	69	89
	100	77 (77)	87 (87)	83 (83)	77 (77)	70 (70)
N	150	117 (78)	135 (90)	132 (88)	126 (84)	120 (80)
	250	195 (78)	230 (92)	231 (92)	226 (90)	220 (88)

Finally, table IV considers the percentage of the market which purchases the product when $p = \frac{2}{3}$ and we vary N and the ration of M/N.

Table IV: Success Rate for Different Percentages of the Market Forced to Decide Early, $p = \frac{2}{3}$

M/N		9%	25%	49%	75%	91%
	100	77%	87%	83%	75%	69%
N	150	83%	89%	83%	75%	69%
	250	91%	91%	83%	75%	70%

Analyzing tables I to IV reveals a number of interesting comparative statics. Firstly the impact of raising M on total number of units purchased is non-monotonic. So we do not expect corner-solutions. Secondly, the impact of M is very dependent on the value of N and p. Thirdly, optimal M is rising in N but falling in p. Finally, switching to percentages reduces the importance of N but does not eliminate it, so the solution cannot be expressed as a fixed percentage of the market for a given p. Some casual observations would add that a figure of around 25% of the market for $p = \frac{2}{3}$ whilst not optimal seems reasonable for an N between 100 and 250, though it is a little high for N approaching 250. So the trade-off gives us a value of M which is nicely in the interior, and not too high a level for a reasonable value of p. As for the impact of N and p we can reason as follows. As p rises the chance of a Y cascade without resort to guinea pigs rises and this seems sufficient to outweigh the similarly beneficial fall in the number of guinea pigs who do not purchase from the firm. Therefore, a rising p value indicates that the number of guinea pigs should be reduced, holding N constant. A rising N value indicates that the number of guinea pigs should rise, though not as a percentage of N. So the firm should raise the absolute number but reduce the percentage of the market acting as guinea pigs. This seems sensible given that market size is decreasingly important for learning in a herding model, since once a herd has started it will not stop, regardless of the number of agents remaining in the sequence.

2.2.2. *Bad Products to Sell.* Now we consider the case when $V = 0$, so the firm now has a "bad" product. Part 3 of the appendix shows that the firm's problem has now changed to become:

$$\max_{M} \left\{ (1-p)(M+1) + (N-M-1)(M+1)! \sum_{x=0}^{\frac{M-1}{2}} \frac{(1-p)^{M+1-x} p^{x}}{x!(M+1-x)!} + \right. \tag{2.4}$$

$$\left. \left(1 - \sum_{x=0}^{\frac{M-1}{2}} \frac{(M+1)![p^{M+1-x}(1-p)^{x} + p^{x}(1-p)^{M+1-x}]}{x!(M+1-x)!} \right) \frac{(p-2)(p-1)}{2} \sum_{n=2}^{N-M-1} \frac{\left(N-M-\frac{n}{2}\right)\left(1-\left(p-p^{2}\right)^{\frac{n-M-1}{2}}\right)}{1-(p-p^{2})} \right\}$$

Now we carry out some of the comparative statics from the previous subsection with the only difference being the move from the $V = 1$ state to the $V = 0$ state. Table V repeats the findings of table I for the new state. Note the collapse in the value of M that would be selected by the firm as we move from state $V = 1$ to $V = 0$. The firm has to carefully balance the desire to initiate a Y cascade by manipulating the number of guinea pigs, by the need to avoid too much information being revealed and a N cascade being initiated.

Table V: Optimal M Values for the Firm when $V = 0$ for Given Values of p and N

p		0.6	2/3	3/4	4/5
N	50	17	13	9	7
	100	27	19	11	9

Table VI, much like table II, gives the expected number of units sold for various different choices of M by the firm for a market of size $N = 100$ for different values of p. The figures for optimal M when $p = 0.51$ are not surprisingly very similar, but moving to a higher figure for p yields very different results with far fewer units being sold especially for higher values of M supporting the findings in table VI.

Table VI: Expected Units Sold for Different Values of M, $N = 100$

M		9	29	49	69	89
	0.51	41	48	50	50	49
p	2/3	74	77	66	53	40
	4/5	89	76	60	44	28

Table VII carries out the same process as table III but for $V = 0$, holding p constant at 2/3 and varying the size of the market, looking at the impact on the expected number of units sold (with percentage of market size in brackets) of a change in M. Finally, table VIII mirrors table IV by evaluating the percentage of the market which purchases the product when $p = \frac{2}{3}$ for various values of N and M/N.

Table VII: Expected Units Sold for Different Values of M, $p = \frac{2}{3}$

M		9	29	49	69	89
	100	74 (74)	77 (77)	66 (66)	53 (53)	40 (40)
N	150	113 (75)	125 (83)	116 (77)	103 (69)	90 (60)
	250	192 (77)	220 (88)	215 (86)	203 (86)	190 (76)

Table VIII: Success Rate for Different Percentages of the Market Forced to Decide Early, $p = \frac{2}{3}$

M/N		9%	25%	49%	75%	91%
	100	74%	77%	66%	53%	40%
N	150	80%	81%	67%	49%	39%
	250	88%	83%	67%	50%	39%

2.2.3. *Summary.* What can we conclude?

Remark 1. *Our first lesson is that regardless of the quality of a firm's product it would do well to consider the dangers of a herd, and to maximize the probabilities of a positive herd on its product through the judicious use of "guinea-pigs" and schemes designed to encourage early purchase by a carefully selected proportion of the population. It is doubly important to notice that this proportion must be carefully calculated as an incorrect number might prove very damaging for sales, for example trying to encourage too high a number of early sales if the product is not a good one.*

2.3. **Consumers with Biases.** Having considered the case when consumers have identical information and preferences, let us now consider the case where some consumers have biases. How should a firm respond to potential customers who have such known differences in preference? Consider a sequential model exactly as in section 2.1, except that we include two special agents with very different signals, a *pessimist* with a negative bias against the firm's product and an *optimist* with a positive bias in favor. We might imagine a customer with a lower or higher cost of adoption, perhaps because he is about to replace an old product, or a customer with a very high cost of adoption because he has just invested in a new product which is especially costly to replace.

We first examine the benchmark case when as is intuitive a firm would like to exploit the existence of consumers with a bias in favor of their product. We then examine the case when a consumer wishes to have access to finer information. This will naturally be costly for the consumer, and will therefore reduce the funds available for purchase, and so will raise *de facto* costs of adoption. This provides a direct justification for the existence of pessimists with higher purchase costs, but also provides motivation for such pessimists to have finer information than other consumers. The result is the counter-intuitive result that such pessimists with higher costs might actually be unambiguously useful to the firm as obvious candidates to be asked to decide first.

2.3.1. *Consumers with Varying Costs of Adoption.*

Definition 3. *Define a pessimist to have a cost of $C = x$, where $x > \frac{1}{2}$ and an optimist to have a cost of $C = 1 - x$. We denote the payoff of the optimist at location i in the sequence as π_i^o and the payoff of the pessimist at location i in the sequence as π_i^p.*

A second useful definition establishes some notion of how informative a signal is.

Definition 4. *Following definition 1, a signal is mildly informative if* $p \in \left[\frac{1}{2}, x\right)$*, where* x *is defined as in definition 3. A signal is very informative if it more informative than a mildly informative signal.*

We will assume in this section that while firms know what the cost of adoption is for each customer, customers only know their own cost of adoption. There is one very important benchmark result we can immediately state about the actions of the biased consumers for a given fixed value of p for all agents in the sequence.

Proposition 1. *If* $p \in \left[\frac{1}{2}, x\right)$ *and* $x > \frac{1}{2}$ *then we can immediately say that a generalized optimist will adopt if asked to decide first regardless of his signal value, while a generalized pessimist will never adopt.*

Proof. First note that for informativeness we have restricted signal accuracy such that $p > \frac{1}{2}$, but signals are mild in the sense of definition 4. The expected payoff of the first decision maker if he is an optimist with a good signal who chooses to adopt is: $E\left[\pi_1^o \mid X_1 = H, A_1 = Y\right] = p - (1-x) > 0$ for $p > 1 - x$ which is trivially true. If the optimist instead receives a bad signal, but still chooses to adopt then: $E\left[\pi_1^o \mid X_1 = L, A_1 = Y\right] = (1-p) - (1-x) > 0$ for $p < x$. Therefore for $p \in \left[\frac{1}{2}, x\right)$ the optimist will always adopt. Now consider the pessimist with a good signal choosing to adopt: $E\left[\pi_1^p \mid X_1 = H, A_1 = Y\right] = p - x > 0$ for $p > x$ which cannot be true. If his signal is bad: $E\left[\pi_1^p \mid X_1 = L, A_1 = Y\right] = (1-p) - x > 0$ for $p < 1 - x$ which also cannot be true. ∎

So if signals are informative, but not too informative then a pessimist will never purchase while an optimist will always purchase. So we have a sensible marketing ploy:

Remark 2. *If a firm knows how optimistic/pessimistic it's customers are, but customers do not know how optimistic/pessimistic their fellow customers are, then the firm should always begin by selling to optimistic customers if information is mildly informative.*

Two caveats are necessary. Firstly, what if information is very informative, so $p \in [x, 1)$? Now an optimist may not purchase and a pessimist may purchase. Nevertheless if other consumers do not know who the optimists/pessimists are then it still seems sensible to approach optimists first as they are still more likely to adopt the new product. Secondly, what if every C is common knowledge? Now consumers can weight the actions of their predecessors by their value of C. We therefore lose the benefit gained from a more likely purchase by an optimist. Combined with common knowledge of a general p there is nothing to gain by changing the order of purchase. We can expand this last point by arguing that even if other consumers do not know who the biased consumers are they may well assume that a firm that is able might approach optimistic consumers first for mildly informative values of p and hence act as if the optimist was approached first.

This result is very intuitive and provides a benchmark for our next result, when we will establish conditions under which it is pessimists rather than optimists which are valuable to the firm.

2.3.2. *Varying the Fineness of Information.* In this section we make two major changes. Firstly we allow customers to know whether their fellow customers are pessimistic or not (though not of course to know the value of any other customer's signals). This is modeled by allowing customers to know the distributions from which their fellow customers draw their signals. This is relevant here because of the second major change: we allow for varying fineness of information, so we allow draws to come from different distributions.

Given the dangers of herding, of which all rational consumers are aware, some might wish to acquire additional information, or more specifically, better information. "Better" in this sense is best described in terms of the fineness of information. A firm which has finer information will be better able to make a good decision, but will have to pay for this increase in information quality, and therefore might be reasonably modeled as having fewer funds to purchase. A simple way of capturing this idea would be to raise the quality of a pessimist's information, while at the same time justifying the higher costs of adoption faced by the pessimist.

Alternatively, imagine that our pessimistic customer, who might have a strong reason not to switch to a new product, is in the position of being asked to consider a new product. Given the pessimism of this customer it might be reasonable to expect him to acquire better quality information on which to base a decision, so we might wonder if a pessimist is more likely to have access to better quality information than an unbiased customer.

Following either motivation, we can ask what would change if the quality of information is not in fact fixed for all agents, but can vary for our pessimistic customers. In particular consider signal draws to now be of the form $p_i \sim U_L\left(0, \frac{1}{2}\right) \cup U_H\left(\frac{1}{2}, 1\right)$. This distribution is a natural extension of $U\left(0,1\right)$ where we retain informativeness so force $P\left[p = 0.5\right] = 0$. Now consider the coarseness of information for our two types of consumer, a pessimist and an unbiased consumer.

The unbiased consumer as before gets basic information about the type of signal, either $X_i = L$ or $X_i = H$ which can now be better described as picks from either $U_L\left(0, \frac{1}{2}\right)$ or $U_H\left(\frac{1}{2}, 1\right)$.[3] So our unbiased consumer i's information is too coarse to specify a value for p_i, but can specify which half of the distribution this signal is taken from. In expectation then he can best approximate signal values as $X_i = L \Rightarrow p_i = \frac{1}{4}$ and $X_i = H \Rightarrow p_i = \frac{3}{4}$.

[3]Hence the use of the distribution $p_i \sim U_L\left(0, \frac{1}{2}\right) \cup U_H\left(\frac{1}{2}, 1\right)$.

The pessimist however has full access to the signal value so knows exactly the value of p_i, and will judge values close to 0 to indicate a strong negative signal about the product, and values close to 1 to indicate a strong positive piece of information. We retain all other elements of the model in 2.1. To summarize:

Definition 5. *Define a pessimist to have a cost of $C = x$, where $x > \frac{1}{2}$. We denote the payoffs, actions and signals of the pessimist at location i with the superscript p. Let $p_i^p \sim U_L\left(0, \frac{1}{2}\right) \cup U_H\left(\frac{1}{2}, 1\right)$. Leave $C = \frac{1}{2}$ as before for all unbiased consumers and let them retain coarse information of the form $X_i = L \Rightarrow p_i = \frac{1}{4}$ and $X_i = H \Rightarrow p_i = \frac{3}{4}$.*

We will in fact focus on pessimistic customers, since as suggested they might have more reason than most to want to consider more information. Consider what might happen if a pessimist decides first in the chain, and decides to adopt. Not only is this a certain sign that $X_1 = H$, but also that $p_1 > \frac{3}{4}$, so $E\left[p_1 \mid A_1^p = N\right] = \frac{7}{8}$. Therefore the decision by the pessimist to adopt is in expectation sufficient to start a Y cascade so long as $p_2 < \frac{7}{8}$, which will of course be the case if the second consumer is unbiased. If we similarly define optimists with higher quality information but allow them to retain lower costs we will of course obtain the symmetric result that an optimist starting the sequence with a decision not to adopt is sufficient to trap an unbiased consumer in a N cascade. Furthermore if a pessimist decides not to purchase this restricts the signal value to be below $\frac{7}{8}$, and so provides a less damaging expected signal than were an unbiased consumer to decide against purchasing in the same position. We can formalize this.

Proposition 2. *In a model in which pessimists have finer information $p_i^p \sim U_L\left(0, \frac{1}{2}\right) \cup U_H\left(\frac{1}{2}, 1\right)$, and customers know the distributions from which their fellow customer's signals are drawn, a pessimist deciding to adopt first, followed by a sequence of unbiased consumers, will initiate a Y cascade but will be irrelevant if he decides not to adopt.*

Proof. First consider the second agent's expectation of the first agent's signal if the first agent is a known pessimist and decides to adopt. $A_1^p = Y \Rightarrow p_1^p > \frac{3}{4} \Rightarrow E\left[p_1 \mid A_1^o = N\right] = \frac{7}{8}$. Now the second agent may have a positive signal and will hence adopt as will all future agents. If the second agent does not have a positive signal his signal strength will be $\frac{1}{4}$ and so the net signal strength will exceed $\frac{1}{2}$ and the agent will still decide to adopt. The pessimist's action is therefore a sufficient statistic for the determination of the actions of future agents. If the pessimist decides not to adopt by a symmetric argument the net signal after the action of the second agent will overcompensate if the second agent decides to adopt, or initiate a herd if the second agent decides not to adopt. The second agent is therefore the initiator of the herd regardless of the pessimist's decision not to adopt. ∎

How does this information help the firm hoping to find their product caught up in a Y herd?

Remark 3. *It is in a firm's best interest to convince a known pessimistic customer to decide first. If the pessimist decides against his decision will be effectively ignored by later movers, but if he decides to purchase he is likely to single-handedly initiate a cascade in favor of the firm's product.*

2.3.3. *Summary.* The two remarks in this section are very different, so a great deal depends upon how variable information quality may be and whether biases are common knowledge. The best way to draw these together is with a final remark.

Remark 4. *If pessimistic customers naturally try to improve their quality of information before making a decision AND are well known as pessimists to other customers then it is a good idea to start with them first. If all customers are likely to have the same quality of information AND do not know how optimistic or pessimistic their fellow customers are then it is a good idea to start with optimistic customers.*

Finally, we can tie this result to remark 1 by noting that a pessimist plays a very similar role here to the guinea pigs in the earlier model. In both cases we have an agent with the ability to signal in a finer way to later decision-makers. In the guinea-pig model we might consider the group of $M+1$ agents to form one super-agent able to act in a less coarse way and therefore provide a less noisy signal for later decision-makers. The super-agent can in effect go beyond binary signalling by choosing a fraction of adoption $\#Y_{M+1}/M+1$ which is a finer signal than a simple decision to adopt or not. The pessimist in a similar way reveals finer information to later decision-makers if he decides to adopt by offering an insight into the exact nature of this signal. In a strong sense this recalls the use of biased contests to provide useful information in Meyer (1991).

3. Endogenous Timing

This section switches the emphasis to herding models where agents can decide to wait to benefit from observing the actions of others. The herd externality applies just as much here as once again action spaces are coarser than signal spaces so actions do not perfectly reveal signals. The result is once again that once a herd begins all agents will follow the actions of their predecessors, and new information will cease to become available. Since consumers can decide at any time in these models a herd will start and end rapidly with all remaining agents deciding simultaneously after a herd has been initiated. The basic model used in this section follows Gale (1996).

3.1. **A Simple Model.** In general we will consider N agents, but initially we will restrict ourselves to $N = 2$.[4] These agents have a decision problem which operates in two dimensions: whether to invest in a project and if so when to invest. The return to this project is the state of the world, w, which is initially assumed fixed at the beginning of time. Time is indexed by $t \in \mathbb{T}_{++}$ and is therefore discrete and strictly positive. Agents do not directly observe w, instead receiving a signal, μ, at $t = 1$. We use superscript to index agents and subscript to index time, so μ_t^i is the signal of agent $i \in \{1, 2\}$ at time t. We will use i and j to denote our two agents; usually i is the agent whose decision problem we are considering and j will be the other agent. The signals μ^i and μ^j are independent and identically drawn from the uniform distribution with support $[-1, 1]$, so $\mu^i \sim U[-1, 1]$ for $i \in \{1, 2\}$. These signals do not change over time, are drawn before the first period, and the state of the world w is set equal to the sum of all signals, $w = \mu^i + \mu^j$. Actions are defined as: $x^i = 1 \Leftrightarrow$ "invest"; and $x^i = 0 \Leftrightarrow$ "do not invest". An agent can observe his own signal, but not the signal of the other agent. In each period actions are made simultaneously, so the two agents cannot observe each others' actions. However, in period 2, the agent will know the action which the other agent performed in period 1, and through the observed choice of action some information about the nature of the other agent's signal may be revealed. Finally we have payoffs, π_t^i, where $t \in \mathbb{T}_{++}$ and $i \in \{1, 2\}$, discounted strictly by $\delta \in (0, 1)$:

$$\pi_t^i = \begin{cases} \delta^{t-1} w & \text{if} \quad x^i = 1 \\ 0 & \text{if} \quad x^i = 0 \end{cases}$$

Consider the problem faced by agent i: whether and if so when to invest. Myopically we could consider the following simple rule: (i) invest (i.e. $x_t^i = 1$) if and only if $E[\pi_t^i] > 0$; (iia) if an investment is to be made, then make it at $t = 1$ if and only if $E[\pi_1^i] > E[\pi_2^i]$, if not then wait. In these rules the profit function explicitly includes discounting. This might seem a sensible rule to adopt, but while we are capturing a notion of the cost of delay since we have an implicit $\delta < 1$ in the second period payoff, we are not capturing the benefit of delay, namely the *option value* of waiting. This option value comes about because of the possibility that for some reason agent i may have invested at time 1 when doing so was foolish given the information available to him at time 2. We will consider the cost and benefit of delay in turn, but first we will define a symmetric signal value $\overline{\mu}$ such that $\mu^i > \overline{\mu} > 0 \Leftrightarrow x^i = 1$. We have not yet said anything about what to do at $t = 2$, but we have defined an alternative decision rule for $t = 1$: (iib) invest at $t = 1$ (i.e. set $x_1^i = 1$) if and only if $\mu^i > \overline{\mu} > 0$.

[4] The endogenous-timing model is very different from the sequential model of section 2.1 so alternative notation will be used to avoid confusion.

Proposition 3. *There is some symmetric $\overline{\mu}$ such that it is optimal for agent i to invest at time $t = 1$ if and only if $\mu^i > \overline{\mu} > 0$.*

Proof. See Appendix, part 4. ∎

Proposition 4. *(i). The game will end by $t = 2$, i.e. if agent i did not invest at time $t = 1$ he will either invest when $t = 2$ or never invest. (ii). Agent i will only invest at $t = 2$ if agent j invested at $t = 1$.*

Proof. See Appendix, part 5. ∎

There are numerous features of this model which are very much in keeping with the herding literature: information is not fully revealed, there is no direct mapping from signal to action which can be inverted to reveal agent's signals; errors are made and private information may be ignored, in particular even if $\mu^i > 0$ for $i = 1, 2$ neither will invest unless $\mu^i > \overline{\mu}(\delta)$ for at least some i; and the errors which lead to incorrect decisions in turn lead to welfare losses, even though there is minimal delay in this model. It has also been shown that the game will effectively end at $t = 2$, beyond this point agents have either invested or will never do so. The addition of further agents would allow the game to continue beyond two periods of interest, but we need at least one agent to invest in a period or investment will stop, as in the two agent case. This is formally shown to be true in the statement and proof of proposition 5 which extends proposition 4 to the multi-agent case.

Proposition 5. *A single period of no investment will end the prospect of any further investment in a model with $N \in \mathbb{N}_{++}$ agents.*

Proof. See Appendix, part 6. ∎

Gale (1996) provides an intuition for this result, pointing out that in a model of this type there must be a possibility of investment collapse as a necessary condition of equilibrium. This comes about because in order to have any delay there must be a positive option value, and this in turn implies a positive probability that agents will never invest.

3.2. **Selling to Groups.** We first investigate the nature of the problem caused by the structure of sales in this model and see that this can be overcome by a simple sales mechanism. Consider the danger of investment breakdown:

Definition 6. ***Full investment breakdown*** *is said to occur when a project has positive value but it is not carried out by any agent because of problems of asymmetric information and uncertainty about the true value of the project.*

Close examination of this definition reveals that full investment breakdown is a form of informational cascade. Consider a situation in which both agents have signal values below the threshold $\overline{\mu}$ and where $w = \mu^i + \mu^j > 0$. Neither agent would invest at $t = 1$, then having observed a period of no investment, they would never invest. The agents are effectively trapped in an informational cascade on the action "do not invest", producing an investment breakdown. Complete revelation effectively bounds profits.

Note also that a problem might also emerge when one agent has a positive signal and one a negative signal but where again, the project is worthwhile in the sense that $w = \mu^i + \mu^j > 0$.

Definition 7. *A* ***partial investment breakdown*** *is said to occur when a project has positive value and one agent invests, but the other agent fails to invest because of problems of asymmetric information and uncertainty about the true value of the project.*

In both cases a worthwhile project is not sufficiently obviously profitable and so delay occurs as both agents wait in an attempt to learn more about the others' signal. Let us go through the decision-process in detail. Both agents begin with positive signals that lie just under $\overline{\mu}$. This results in a unilateral decision to delay in order to gain more information. In the second period both observe a failure to invest by the other agent which results in an updated set of beliefs which suggest to each agent that the others' signal in expectation $(\overline{\mu}-1)/2 < 0$. If an agent had reason to delay when the other agent's signal was in expectation neutral a negative expected signal will reinforce to decision not to invest. Of course there will be no new information forthcoming as neither agent will now ever decide to invest and so we have a reverse cascade.

What is the solution to this problem? An obvious direct suggestion would be to change the nature of sales. Rather than allowing each agent the luxury of delay in order to observe the action of the other an alternative might be tried, for example:

1. All consumers are gathered and asked to anonymously complete a simple questionnaire designed to reveal the strength of signal values.
2. The results are revealed to all consumers simultaneously.
3. They are then asked to publicly and simultaneously vote on whether they would purchase the product.
4. Once again the results are revealed publicly.

This form of sales is easy to implement when the number of potential consumers is low. In practice it is used by some firms to avoid the perennial sales director's nightmare when marketing a new technology of a potential customer replying to a sales pitch "I am of course interested, tell me who else has purchased?"[5]

[5] AEA Technology who sponsor the author of this paper have provided a good example. A perception of some difficulty in sales led the company's rail division to try a similar scheme to sell a patented

To summarize:

Remark 5. *When a new product is released and consumers can delay their decision about adoption in order to observe the actions of their fellow consumers then a form of reverse cascade known as investment breakdown can occur. This can be overcome through a sales mechanism designed to allow consumers to pool their signals and therefore avoid a profit and welfare-damaging herd.*

3.3. **Joint Ventures and Information Gathering.** Finally we consider joint ventures by consumers specifically designed to improve information, perhaps because of a fear of a reverse cascade and the welfare-damaging implications this entails. We might imagine that the consumers here are large enough to be interested in such an arrangement, perhaps they are firms themselves considering taking on a new potentially profitable, but expensive technology. A reasonable reaction would be to pool information especially when there is little strategic interaction between the firms. Alternatively the firm selling the new technology might wish to provide unbiased credible information to consumers which rule out a reverse cascade. In both cases the issue is whether such information revelation which will undoubtedly be expensive and time-consuming is worthwhile.

3.3.1. *Complete Revelation of the True State.* We now examine the role of information revelation in the decision process. In particular, consider the possibility of *complete revelation* of the true state.

Definition 8. ***Complete revelation*** *of the true state of the world $\{w_t, t \in \mathbb{T}_{++}\}$ at some pre-determined point in the future $t = \tau^*$, where $\tau^* \in \mathbb{T}_{++}$ is common knowledge to all agents, is said to occur when the true value of the state of the world w_{τ^*} at time $t = \tau^*$ is revealed to all agents.*

Assume that there exists a benevolent "third party" who can observe the true state of the world at certain points in time.[6] Consider a three period model: the first two periods are as in the previous sections. However, in the third period we allow the agents to know the true state of the world. Therefore for $\delta \in (0,1)$ and $i \in \{1,2\}$ the agents are left with:

$$\pi_3^i = \begin{cases} \delta^2 w_3 & if \quad x_3^i = 1 \\ 0 & if \quad x_3^i = 0 \end{cases}$$

smart-sanding technology. The results were very positive with voting revealing a preference to purchase which resulted in good sales.

[6]The third party here is perhaps best considered to be a joint committee set up by both firms to aggregate their information and therefore produce a totally accurate picture of the value of the project. This will of course take time to implement.

Proposition 6. *If $\tau^* > 2$ then the game will end at time $t = \tau^*$ with a decision to invest or never invest, where τ^* is the time of complete revelation. In particular, if $\tau^* = 3$, then the game will end in period 3 with a decision to invest or never invest.*

Proof. See Appendix, part 7. ∎

Put simply this proposition establishes that the quick decisions of proposition 4 will be slowed by the potential for further information. To examine this formally define the information set in time t of agent i as J_t^i. This will naturally include his own signal as well as the past history of agent j's actions up to and including agent j's action in time $t-1$. Agent i has three potential periods in which it might be optimal to invest, and we will henceforth consider $\tau^* = 3$. The natural way to examine the decision problem is via a backward induction or dynamic programing approach. We will consider what the agent would do in period 3, assuming he is at period 3, then examine decisions in period 2 in the light of actions in period 3. Finally, we will look at period 1 having considered the optimal decision in period 2. This is made feasible by the simple observation that having reached period 3 the agent's best decision is to invest if and only if $\pi_3^i = \delta^2 w_3 > 0 \Rightarrow w_3 > 0$. Therefore we can disregard period 4 and onwards. We will define the threshold signal value used here as $\widehat{\mu}$ to differentiate it from the previous signal value. The precise calculations are made in part 8 of the appendix, but collapse to agent i knowing that his worst possible payoff by period three is $\max\left\{\frac{1}{2}\delta^2\left(\widehat{\mu}+1\right),0\right\}$.[7] In period 2 the agent will therefore invest if and only if:

$$E\left[\pi_2^i \mid J_2^i\right] - \max\left\{\tfrac{1}{2}\delta^2\left(\mu^i+1\right),0\right\} > -\frac{1}{4}\delta\left(1+\widehat{\mu}\right)\left(3\widehat{\mu}-1\right)$$

And therefore the first-period decision defines $\widehat{\mu}$ when solved with equality:

$$\widehat{\mu} - \max\left\{\delta\widehat{\mu}, \tfrac{1}{2}\delta^2\left(\widehat{\mu}+1\right),0\right\} = -\frac{1}{4}\delta\left(1+\widehat{\mu}\right)\left(3\widehat{\mu}-1\right)$$

Part 8 of the appendix completes the agent's decision-making problem resulting in the calculation of a new threshold value:

$$\overline{\overline{\mu}} = \tfrac{1}{6\delta}\left\{2\delta^2 - 2\delta - 4 + \left[\left(4-2\delta^2+2\delta\right)^2 + 12\delta\left(2\delta^2+\delta\right)\right]^{\frac{1}{2}}\right\}$$

[7] Here and throughout payoffs will always be discounted back to time $t = 1$, so the $t = 3$ payoff is expressed as the discounted value at $t = 1$. When comparing time $t = 2$ and $t = 3$ the payoff at time $t = 3$ should only be discounted once by δ to return a payoff discounted back to $t = 2$ but by discounting all payoffs to $t = 1$ there is greater consistency throughout.

Comparing this with our previous value for $\overline{\mu}$, we have $\overline{\mu} < \overline{\overline{\mu}}$ for $\delta \in (0,1]$ and $\overline{\mu} = \overline{\overline{\mu}}$ for $\delta = 0$. The addition of the possibility of complete revelation therefore results in a new threshold value which lies above the old threshold value for $\delta > 0$. The addition of the prospect of new information has increased the chance that agents will delay their decision and wait for this new information.

3.3.2. *The Usefulness of New Information.* So new information will slow decision-making, but presumably it will assist the firms in making the best possible decision. This section will examine precisely this trade-off.

Consider the following example. A project may be worth a strictly positive amount, with $\mu^1 = 0.7$ and $\mu^2 = -0.65$. Agent 1 will invest at time $t = 1$ and agent 2 will delay. Even after observing agent 1's decision to invest, agent 2 will not invest at time $t = 2$ for a wide range of values of $\delta \in (0,1)$. This example can be generalized, so for $i \neq j$ whenever $-\mu^i < \mu^j < -\frac{1}{2}(\overline{\mu}+1)$ there will be a situation of partial breakdown with agent j failing to invest in a worthwhile project.[8] In this case complete revelation will be useful for agent j though not for agent i.

Proposition 7. *Complete revelation will only be of any benefit in the fraction of cases given by:*

$$f(\overline{\mu}(\delta)) \equiv f(\delta) = \frac{2-6\overline{\mu}+32\overline{\mu}^2-10\overline{\mu}^3+\overline{\mu}^4}{50-20\overline{\mu}+2\overline{\mu}^2}$$

Proof. See Appendix, part 9. ∎

It should be noted that $\frac{\partial f(\delta)}{\partial \delta} > 0$ and $\frac{\partial f(\overline{\mu})}{\partial \overline{\mu}} > 0$, so $\max_\delta f(\delta) = f(1)$ which is roughly 0.07. With maximum patience complete revelation will be of use in 7% of cases. For a more reasonable patience level, say 0.5, we get $f(0.5) \approx 0.038$, so gathering extra information is here worthwhile in under 4% of cases. We see that extra information is indeed useful but in only a surprisingly small fraction of cases, and this is when such extra information is assumed to be costlessly obtained.

Proposition 8 gives a necessary condition for undertaking complete revelation when information gathering has a cost.

Proposition 8. *With a cost of gathering information $C_g > 0$ an ex ante (even before signal values are realized) necessary condition for welfare-improving complete revelation*

[8] This figure can be derived in the following way. For a project to be worthwhile it must be the case that $\mu^i + \mu^j > 0$ which implies that $-\mu^i < \mu^j$. However, after observing that agent i has invested, agent j can calculate $E\left[\mu^i \mid x_1^i = 1\right] = \frac{1}{2}(\overline{\mu}+1)$ since it is uniform with support $[\overline{\mu}, 1]$. In order for agent j to fail to invest after observing investment by agent i, it must be the case that $\frac{1}{2}(\overline{\mu}+1) + \mu^j < 0$ which implies that $\mu^j < -\frac{1}{2}(\overline{\mu}+1)$.

when revelation occurs at $t = \tau^*$, *is:*

$$C_g < \delta^{\tau^*} \left(\overline{\mu}^3 + \frac{(3+\overline{\mu})(1-\overline{\mu})^3}{2(5-\overline{\mu})^2} \right)$$

Proof. See Appendix, part 10. ∎

It should be stressed that this is a necessary condition and a very weak one, based on maximum possible signal values throughout, and it will rarely be sufficient. This implies that the information will only be of any use in a small number of cases. Add to this the fact that when the information is of any use this is exactly when the value of the project is likely to be positive but small, and the total value of the complete revelation is seen to be low. To give some idea of the magnitudes involved consider the following example.

Example 1. *All approximations are to three significant figures. For* $\delta = 0.5$, $\tau^* = 3$ *we have* $\overline{\mu}(\delta) \approx 0.155$ *and* $\overline{\overline{\mu}}(\delta) \approx 0.208$. *So the prospect of extra information raises the threshold value by* 34% *of the original value. Furthermore* $f(\delta) \approx 0.038$, *so by proposition 7 complete revelation is only useful in* 3.8% *of cases. Now using proposition 8 we have as a necessary condition that the cost of public information gathering* C_g *must be below* 0.0055. *To put this into a reasonable metric, the maximum possible ex ante project value is* 2, *so a necessary condition for complete revelation in the case when* $\delta = 0.5$ *is that the cost of information gathering not exceed* 0.28% *of the maximum value of a given project, so for a project with a maximum possible value of $1million, the cost of information gathering should not exceed $2800.*

Note that a low value of δ makes the necessary condition even stricter. A final proposition stresses the importance of when the extra information is due to be released.

Proposition 9. *As* $\tau^* \to \infty$ *complete revelation will have a falling effect on the threshold value, with the effect disappearing completely in the limit. Therefore, a necessary condition for complete revelation to be weakly profit-improving as* $\tau^* \to \infty$ *is that it has zero cost. The necessary and sufficient condition with* $\tau^* \to \infty$ *involves the need for negative costs.*

Proof. See Appendix, part 11. ∎

So yet another qualifier for the usefulness of complete revelation is that it should not occur too late.

3.3.3. *Summary.* What we have seen is that new information is useful as it will certainly overcome the dangers of a reverse cascade (investment breakdown), but it will also shift out the threshold value beyond which an immediate decision to adopt is taken. This movement will reduce the amount of swift decisions and so add to discounting and in many cases will actually reduce profits as compared with the situation when information gathering is not entertained. However most damaging at all is the fact that such information gathering will not be costless and when a direct cost is combined with the extra delay produced the net result is that only when the costs are very low will extra information be useful.

Remark 6. *When reverse cascades are a serious danger consumers should not be too quick to attempt to overcome these dangers through further information gathering unless such extra information can be obtained at very low cost (often in the order of below 1% of the total potential value of the project). if extra information is costless then it is unambiguously useful, though even here the implied extra delay diminishes the benefits.*

4. Conclusion

What we have seen is that in both sequential and endogenous-time example herd models it is possible, indeed sensible, to imagine firms and consumers seeking ways to avoid welfare and profit damaging cascades. In sequential models these methods may come in the form of encouraging early entry as in the first model, or using the extra information provided by the actions of biased consumers as in the second. In endogenous-time models similar concerns might lead to an attempt to once again increase the information available to consumers, though methods that work through sales might be more cost-effective than attempting to provide direct extra information as there is always a danger that this will lead to extra delay in decision-making. The simple models used in this paper show how the dangers of herding can provide direct policy advice to firms and consumers and not simply provide sound theoretical underpinnings for otherwise difficult to understand decision-making patterns. In particular many of the methods suggested in this paper are already widely used whether in terms of the group selling techniques of section 3 or methods designed to encourage early decision-making in section 2. In many cases however firms may not realize the full extent of the dangers that arise from the herd externality and so the models examined in this paper might translate as direct advice to such firms. It is hoped that this paper will leave a general message that while simple theoretical models are a useful tool, it is not only theorists that can alter the structure of models; agents can and do make changes to informational structures or timing decisions themselves when such alterations translate into increased profits or utility.

Appendix

1. Derivation of Conditional Herd Probabilities:

Starting with 2 agents we have $\Pr[Y(2) \mid V=1] = p^2 + \frac{(1-p)p}{2}$ & $\Pr[Y(2) \mid V=0] = (1-p)^2 + \frac{p(1-p)}{2}$. Therefore $\Pr[Y(2)] = \frac{1-p+p^2}{2}$. Similarly, we have $\Pr[N(2)] = \frac{1-p+p^2}{2}$.

No cascade by agent 2 will occur with probability $1 - \Pr[Y(n)] - \Pr[N(n)]$, therefore $\Pr[No(2)] = p - p^2$. Note of course that this can be alternatively calculated as the occurrence of HL or LH and a coin flip by agent 2, so $\Pr[No(n)] = \frac{1}{2}p(1-p) + \frac{1}{2}(1-p)p$. Further note that $\Pr[Y(2)]$ and $\Pr[N(2)]$ are not conditional on V since they are fully symmetric so $\Pr[N(n)] = \frac{1}{2}(1 - \Pr[No(n)])$. Now note that $\Pr[Y(4)] = \Pr[Y(2)] + \Pr[No(2)]\Pr[Y(2)]$ and similarly for $\Pr[N(4)]$. Further $\Pr[No(4)] = (\Pr[No(2)])^2$. Using this we can easily deduce the general probabilities after an even number of n agents to be $\Pr[No(n)] = (\Pr[No(2)])^n = (p-p^2)^{\frac{n}{2}}$ for no cascade, and $\Pr[Y(n)] = \Pr[N(n)] = \frac{1}{2}\{1 - \Pr[No(n)]\} = \frac{1}{2}\left[1 - (p-p^2)^{\frac{n}{2}}\right]$ for a Y or N cascade. Now consider the probability of the correct or incorrect cascade occurring:

$$\Pr[Y(2) \mid V=1] = p^2 + \tfrac{1}{2}p(1-p) = \tfrac{1}{2}p(p+1)$$

$$\Pr[No(2) \mid V=1] = \tfrac{1}{2}p(1-p) + \tfrac{1}{2}p(1-p) = p(1-p)$$

$$\Pr[N(2) \mid V=1] = (1-p)^2 + \tfrac{1}{2}p(1-p) = \tfrac{1}{2}(p-2)(p-1)$$

After an even number of n agents we have:

$$\Pr[No(n) \mid V=1] = (\Pr[No(2) \mid V=1])^{\frac{n}{2}} = \left(p-p^2\right)^{\frac{n}{2}}$$

$$\Pr[Y(n) \mid V=1] = \Pr[Y(2) \mid V=1] + \Pr[Y(2) \mid V=1]\Pr[No(2) \mid V=1]$$

$$+ \Pr[Y(2) \mid V=1]\Pr[No(4) \mid V=1] + ... + \Pr[Y(2) \mid V=1]\Pr\left[No\left(\tfrac{n}{2}\right) \mid V=1\right]$$

$$= \Pr[Y(2) \mid V=1]\left[1 + \left(p-p^2\right) + ... + \left(p-p^2\right)^{\frac{n}{2}}\right]$$

Now using the sum of a geometric series we have:

$$\Pr[Y(n) \mid V=1] = \frac{p(p+1)}{2}\frac{1-\left(p-p^2\right)^{\frac{n}{2}}}{1-(p-p^2)}$$

Similarly we can calculate for $\Pr[N(n) \mid V=1]$:

$$\Pr[N(n) \mid V=1] = \frac{(p-2)(p-1)}{2}\frac{1-\left(p-p^2\right)^{\frac{n}{2}}}{1-(p-p^2)}$$

2. Calculation of Expression 2.3:

Expression 2.3 is based on three parts. The first part is simply $p(M+1)$, the number of units purchased within the group of guinea pigs, and is simply the probability of a high signal given $V=1$ multiplied by the size of the initial group. The second part is more complex $(N - M - 1)(M+1)! \sum_{x=0}^{\frac{M-1}{2}} \frac{p^{M+1-x}(1-p)^x}{x!(M+1-x)!}$. This is the size of the remaining population of agents, $N - M - 1$, multiplied by the probability of a Y cascade being induced by the initial group, which is:

$$\Pr\left[Q_{M+1} \geq \tfrac{M+1}{2} + 1 \mid V = 1\right] = (M+1)! \sum_{x=0}^{\frac{M-1}{2}} \tfrac{p^{M+1-x}(1-p)^x}{x!(M+1-x)!}$$

The derivation of the final part of expression 2.3 incorporates the possibility that the initial group failed to initiate a Y cascade. Despite this there is still a good chance of a Y cascade being initiated by later agents. Start with a signal which is on aggregate neutral, being revealed by the guinea pigs, which occurs with probability:

$$\Pr\left[Q_{M+1} = \tfrac{M+1}{2} \mid V = 1\right] = 1 - (M+1)! \sum_{x=0}^{\frac{M-1}{2}} \tfrac{p^{M+1-x}(1-p)^x + p^x(1-p)^{M+1-x}}{x!(M+1-x)!}$$

Which is simply one minus the combined probability of a Y cascade and a N cascade. Now consider the actions of the decisions made by the post-guinea pig agents. If n agents make decisions and initiate a Y cascade this must involve $\frac{1}{2}(n-2)$ agents choosing Y and $\frac{1}{2}(n-2)$ agents choosing N with a crucial 2 agents tipping the balance in favor of a Y cascade. So of the n we have $\frac{1}{2}(n-2)+2$ agents deciding to purchase. Now we have only a population of size $N - M - 1 - n$ remaining. So we have a total of $N - M - \frac{1}{2}n$ who purchase in the event of a Y cascade being initiated by agent n. For example, if $n = 2$ then the Y cascade failed to be initiated by the initial group of guinea pigs, but the $M+2$nd agent and the $M+3$rd agent both decide to purchase initiating a Y cascade which still results in the entire $N - M - 1$ purchasing. If $n = 4$, then from the first 4 after the initial group of $M+1$, 3 will decide to purchase and 1 will decide otherwise, resulting in $N - M - 2$ units being purchased. This all has to be multiplied by the probability of no cascade being initiated by the group of guinea pigs and the probability of a Y cascade being initiated by the group of n immediately following the guinea pig, which is therefore:

$$\sum_{n=2}^{N-M-1} \left[\theta \left(N - M - \tfrac{n}{2}\right) \tfrac{p(p+1)}{2} \tfrac{1-\left(p-p^2\right)^{\frac{n-M-1}{2}}}{1-(p-p^2)} \right]$$

where

$$\theta = \left(1 - (M+1)! \sum_{x=0}^{\frac{M-1}{2}} \tfrac{p^{M+1-x}(1-p)^x + p^x(1-p)^{M+1-x}}{x!(M+1-x)!} \right) \tag{A1}$$

Combining these three parts yields a function giving the total number of units sold by the firm as a function of p, N and M. Of these we allow the firm to vary only M making the final problem to maximize expression 2.3 by choice of M.

3. Calculation of Expression 2.4:

Expression 2.4 is also based on three parts. The first part is now $(1-p)(M+1)$, since the probability of a high signal given $V=0$ has changed to be $1-p$. The second part has also slightly changed to now be: $(N-M-1)(M+1)!\sum_{x=0}^{\frac{M-1}{2}}\frac{(1-p)^{M+1-x}p^x}{x!(M+1-x)!}$. This is the size of the remaining population of agents, $N-M-1$, multiplied by the new probability of a Y cascade being induced, now that $V=0$, which is:

$$\Pr\left[Q_{M+1} \geq \tfrac{M+1}{2}+1 \mid V=0\right] = (M+1)!\sum_{x=0}^{\frac{M-1}{2}}\tfrac{(1-p)^{M+1-x}p^x}{x!(M+1-x)!}$$

The derivation of the final part of the expression is much as in the case when $V=1$ except we now use the probability that a Y cascade occurs given $V=0$. Note that the aggregate neutral signal being revealed by the guinea pigs occurs with the same probability as before, so:

$$\Pr\left[Q_{M+1} = \tfrac{M+1}{2} \mid V=0\right] = 1-(M+1)!\sum_{x=0}^{\frac{M-1}{2}}\tfrac{p^{M+1-x}(1-p)^x+p^x(1-p)^{M+1-x}}{x!(M+1-x)!}$$

The population size is also the same, at $N-M-\frac{n}{2}$. However, the new Y cascade probability changes the final term to:

$$\sum_{n=2}^{N-M-1}\left[\theta\left(N-M-\tfrac{n}{2}\right)\tfrac{(p-2)(p-1)}{2}\tfrac{1-\left(p-p^2\right)^{\frac{n-M-1}{2}}}{1-(p-p^2)}\right]$$

Where θ is defined as in equation A1. Combining these three parts once again yields the function which the firm will maximize by choice of M.

4. Proof of Proposition 3:

First the "if" part will be proven, then the "only if".

($\Rightarrow$) The cost of delay can be seen intuitively as $(1-\delta)\mu^i$. This is simply the expected payoff at time 1 minus the expected payoff at time 2. The difference is in some sense, the cost of delay. Since the unconditional expectation $E\left[\mu^j\right]=0$ which is true for any signal distribution symmetric about zero, such as the uniform $[-1,1]$. Consider the benefit in delay: the option value. Here we need to consider the possibility of regret, where an investment made at time 1 actually seems less sensible when information made available at time 2 is revealed. Information of this sort comes about if it is observed that agent j did not invest at time 1, therefore revealing that $\mu^j < \overline{\mu}$ which provides some evidence that the state of the world is less likely to merit investment.[9] This can be avoided if agent i waits and so provides the option value of waiting which occurs with probability $\Pr\left[\mu^j < \overline{\mu}\right]$. The option value can therefore be defined as the expected loss avoided by agent i by not investing at $t = 1$ in the event that agent j does not invest at $t = 1$:

$$\text{(A2)} \qquad -\delta \Pr\left[\mu^j < \overline{\mu}\right]\left\{\mu^i + E\left[\mu^j \mid \mu^j < \overline{\mu}\right]\right\}$$

We have a condition which leaves the marginal decision-maker indifferent when deciding to invest at time 1: indifference occurs when the option value exactly offsets the delay cost; this is none other than the standard value matching condition for a dynamic programming problem. This condition implicitly defines the value of $\overline{\mu}$ using the properties of the uniform distribution:

$$\text{(A3)} \quad (1-\delta)\,\overline{\mu} = -\delta \Pr\left[\mu^j < \overline{\mu}\right]\left\{\overline{\mu} + E\left[\mu^j \mid \mu^j < \overline{\mu}\right]\right\} \Rightarrow \overline{\mu} = \frac{-(4-2\delta) \pm \left[(4-2\delta)^2 + 12\delta^2\right]^{\frac{1}{2}}}{6\delta}$$

For $\delta \in (0,1)$ and $\overline{\mu} \in [-1,1]$ we can rule out one of these two results, eliminating:

$$\text{(A4)} \qquad \overline{\mu} = \frac{1}{6}\delta^{-1}\left\{-(4-2\delta) - \left[(4-2\delta)^2 + 12\delta^2\right]^{\frac{1}{2}}\right\} \notin [-1,1] \text{ for } \delta \in (0,1)$$

This leaves the value of $\overline{\mu}$ uniquely given as:

$$\text{(A5)} \qquad \overline{\mu} = \frac{1}{3} + \frac{2}{3}\delta^{-1}\left[\left(\delta^2 - \delta + 1\right)^{\frac{1}{2}} - 1\right]$$

Equation A5 is well defined for $\delta \in (0,1)$ and gives a range of values for $\overline{\mu}$ of $\overline{\mu} \in \left(0, \frac{1}{3}\right]$, that can be roughly approximated by the linear function $\overline{\mu} = \frac{1}{3}\delta$ over the relevant range of values of δ. It has been shown that there exists a unique value of $\overline{\mu}$ given in equation A5 such that if $\mu^i > \overline{\mu}$ the cost of delay is strictly offset by the option value of waiting. We have the "$>$" relation since the cost of delay is rising in μ^i (and falling in δ) which therefore defines the optimal decision rule for agent i at time 1. The assumption of a positive option value to delay implies that $\overline{\mu} > 0$.

($\Leftarrow$) Consider what value μ^i must take if agent i has optimally decided to invest at time 1. Optimally deciding to invest implies that the delay cost is strictly offset by the option value:

$$\text{(A6)} \qquad (1-\delta)\,\mu^* < -\delta \Pr\left[\mu^j < \mu^*\right]\left\{\mu^* + E\left[\mu^j \mid \mu^j < \mu^*\right]\right\}$$

[9]There is an assumption of symmetry here - or alternatively the equilibrium decision rules found could be referred to as the symmetric decision rules. The totally symmetric nature of the problem makes this a natural assumption and a natural equilibrium to seek. Gul and Lundholm (1995) make a strong case for the relevance of the symmetric equilibrium in a decision model of this type.

where μ^* implicitly defines the value of μ required for this inequality relation to hold. But this is exactly the value $\overline{\mu}$ we defined in the first part of the proof.

5. Proof of Proposition 4:

We are given that agent i did not invest at time $t = 1$. If this was so we know from proposition 3 that $\mu^i < \overline{\mu}$. Investment will benefit agent i if $E\left[\pi_2^i\right] > 0$. Two rationales for delay at $t = 1$ are possible and are considered in turn.

(a). If $\mu^i \in (-1, 0]$ and therefore $E\left[\pi_2^i\right] < 0$, only if new information suggested a rise in $E\left[\pi_2^i\right]$ would it be rational to decide to invest. Agent i must have observed one of two possible histories: $x_1^j = 1$ or $x_1^j = 0$. Only if he observed $x_1^j = 1$ would he raise his expectation of π_2^i:

$$E\left[\pi_2^i \mid x_1^j = 1\right] = \mu^i + E\left[\mu^j \mid \mu^j > \overline{\mu}\right] = \mu^i + \tfrac{1+\overline{\mu}}{2} > \mu^i = E\left[\pi_1^i\right] \tag{A7}$$

$$E\left[\pi_2^i \mid x_1^j = 0\right] = \mu^i + E\left[\mu^j \mid \mu^j < \overline{\mu}\right] = \mu^i - \tfrac{1-\overline{\mu}}{2} < \mu^i = E\left[\pi_1^i\right] \tag{A8}$$

Since this is a symmetric problem the same is true for agent j if $\mu^j \in (-1, 0]$, therefore if agent j did not invest at $t = 1$ then he too would only raise his expectation if $x_1^i = 1$. If neither invest then no increase in expectation occurs at $t = 2$ and so neither invest at $t = 2$, and hence no rise in expectation occurs at $t = 3$, and so on. Therefore we have shown that if one agent does not observe investment from the other he will not invest and the next period will look much like the second, so the decision not to invest becomes permanent. If either agent invested the other would increase his expectation, but only once (since the other player may not act again) and will therefore raise his expectation so $E\left[\pi_2^i \mid x_1^j = 1\right] > 0$ and invest at $t = 2$ or despite the increase it will be the case that $E\left[\pi_2^i \mid x_1^j = 1\right] < 0$ because his signal was so low, and no investment will take place at $t = 2$ or ever.

(b). If $\mu^i \in (0, \overline{\mu})$ and $E\left[\pi_2^i\right] > 0$ then he was delaying despite expecting positive profit because of the positive option value to delay. This option value has however been expended. If $x_1^j = 1$ then he would have been better off investing at $t = 1$ and would have done so had he realized that agent j would definitely invest. He will invest at $t = 2$ since there will be no further revelations as agent j has *de facto* left the game. Now if $x_1^j = 0$ agent i will lower his payoff expectation as will agent j therefore if it was optimal for them to delay at $t = 1$ it is optimal to delay at $t = 2$ *a fortiori* and so it will be optimal not to invest at $t = 2, 3, 4$, etc. We have shown that it all cases, agent i will either invest at $t = 1$, invest at $t = 2$, or never invest and so proven part (i). Furthermore in all the cases examined it is only optimal to invest at $t = 2$ if agent j invested at $t = 1$ and therefore we have also proven part (ii).

6. Proof of Proposition 5:

We need to show that if there is no investment at an arbitrary time, $t = \tau$, then there will be no investment at time $t = \tau + 1, \tau + 2, \ldots.$ We know from proposition 4 that if there is investment at time $t = \tau$ then agent i will not alter his optimal decision not to invest, and by symmetry this will be the case for all i. The only additional information revealed at time $t = \tau + 1$ lowers expected payoffs so as in proposition 4 agents will either go from a position where $\mu^i \in (-1, 0] \Rightarrow E\left[\pi^i_{\tau+1}\right] < 0$ and will then certainly not invest at time $t = 2$, or $\mu^i \in (0, \overline{\mu}) \Rightarrow E\left[\pi^i_{\tau+1}\right] > 0$ and they will have decided optimally to delay because of a positive option value, and it will remain optimal to delay *a fortiori* just as in the two agent case. At time $t = \tau + 3$ agent i is in an identical position to the position at time $t = \tau + 2$, since no agents have invested once more, so there is no additional information at all being revealed, and this will clearly be the case for $t = \tau + 4, \tau + 5, \tau + 6\ldots.$ Therefore there will be no reason for any agent to change his optimal decision not to invest.

7. Proof of Proposition 6:

There is no option value at time $t = \tau^*$. Since the state of the world is now known with certainty, both w_{τ^*} and δ are known to agent i at time $t = \tau^*$, so there is no longer any need to consider the actions or information of agent j. Therefore a very simple decision rule is optimal: $x^i_{\tau^*} = 1 \Leftrightarrow \pi^i_{\tau^*} = \delta^2 w_{\tau^*} > 0$. Once we factor in further discounting $E\left[\pi^i_4\right] < \pi^i_3$. Therefore if $\pi^i_3 \geq 0$ there is no reason to delay beyond the period in which information revelation takes place. If $\pi^i_3 < 0$ then investment is not profitable now and will be even more unprofitable beyond $t = 3$. In general, we can say that since $E\left[\pi^i_{\tau^*+1}\right] < \pi^i_{\tau^*}$ where $t = \tau^*$ is the time of full information revelation, then $\pi^i_{\tau^*} \geq 0 \Leftrightarrow x^i_{\tau^*} = 1$. Alternatively if $\pi^i_{\tau^*} < 0$ then $E\left[\pi^i_{\tau^*+1}\right] < 0 \Rightarrow E\left[\pi^i_{\tau^*+2}\right] < 0$ etc. therefore $x^i_t = 0$ for all $t \geq \tau^*$. Thus the solution to the decision problem is fully determined by time $t = \tau^*$.

8. Calculation of the New Signal Threshold, when Information Gathering is Undertaken:

In period 3 the agent will observe the true value of the project w_3. So agent i will know whether the project is worth a strictly positive amount or not, i.e. will know whether $\mu^i + \mu^j > 0$. If $\mu^i + \mu^j < 0$ then the agent can obtain $\pi^i_3 = 0$ simply through not investing. If $\mu^i + \mu^j > 0$, then the agent will obtain $E\left[\pi^i_3 \mid \mu^j > -\mu^i\right] = \frac{1}{2}\delta^2\left(\mu^i + 1\right)$. Therefore in period 3 agent i will receive $\max\left\{\frac{1}{2}\delta^2\left(\mu^i + 1\right), 0\right\}$. Now we move back to period 2 where the agent can wait until period 3 to obtain $\max\left\{\frac{1}{2}\delta^2\left(\mu^i + 1\right), 0\right\}$, or invest to obtain an expected $\delta\mu^i$. To roll back the decision to period 1 we now have the expected payoff from delay of $\max\left\{\delta\mu^i, \frac{1}{2}\delta^2\left(\mu^i + 1\right), 0\right\}$ and the expected payoff to investing of μ^i. This defines the cost of delay as $\mu^i - \max\left\{\delta\mu^i, \frac{1}{2}\delta^2\left(\mu^i+1\right), 0\right\}$. Finally we compare this with the benefit of delay, the real option value, of $-\frac{1}{4}\delta\left(1+\widehat{\mu}\right)\left(3\widehat{\mu}-1\right)$. Setting these two terms to equality defines the threshold value for the problem above which investment will take place:

$$\widehat{\mu} - \max\left\{\delta\widehat{\mu}, \tfrac{1}{2}\delta^2\left(\widehat{\mu}+1\right), 0\right\} = -\frac{1}{4}\delta\left(1+\widehat{\mu}\right)\left(3\widehat{\mu}-1\right)$$

Now we know $\widehat{\mu} > 0$, but we need to check whether $\delta\widehat{\mu} < \frac{1}{2}\delta^2\left(\widehat{\mu}+1\right)$. If $\delta\widehat{\mu} > \frac{1}{2}\delta^2\left(\widehat{\mu}+1\right)$ then we return to the same threshold value as in the case without revelation, so $\widehat{\mu} = \overline{\mu}$. However:

$$\delta\widehat{\mu} < \tfrac{1}{2}\delta^2\left(\widehat{\mu}+1\right) \Rightarrow \widehat{\mu} < \tfrac{\delta}{2-\delta}$$

which is always true for $\delta \in (0,1]$. So we can define a new threshold value $\overline{\overline{\mu}}$:

$$\overline{\overline{\mu}} - \tfrac{1}{2}\delta^2\left(\overline{\overline{\mu}}+1\right) = -\frac{1}{4}\delta\left(1+\overline{\overline{\mu}}\right)\left(3\overline{\overline{\mu}}-1\right)$$

$$\Rightarrow \overline{\overline{\mu}} = \tfrac{1}{6\delta}\left\{2\delta^2 - 2\delta - 4 + \left[\left(4-2\delta^2+2\delta\right)^2 + 12\delta\left(2\delta^2+\delta\right)\right]^{\frac{1}{2}}\right\}$$

and comparing this with our previous value for $\overline{\mu}$, we have $\overline{\mu} < \overline{\overline{\mu}}$ for $\delta \in (0,1]$ and $\overline{\mu} = \overline{\overline{\mu}}$ for $\delta = 0$.

9. Proof of Proposition 7:

Complete revelation is only of any use if, in the world before the prospect of revelation, the firms' signal values were such that full or partial investment breakdown would have occurred. This requires that μ^i and μ^j are both in the region $[-1, \overline{\mu}(\delta)]$ and that $w = \mu^i + \mu^j > 0$. The distribution of the value of the project at t, below the threshold value, is the sum of two uniform distributions with support $[-1, \overline{\mu}]$ and is therefore triangular with support $[-2, 2\overline{\mu}]$. Denote the probability of investment breakdown as $g\left(\overline{\mu}(\delta)\right) \equiv g(\delta)$. Using the properties of the triangular distribution $g(\delta)$ is given by:

$$g(\delta) = \Pr\left[w > 0 \mid \mu^i < \overline{\mu} \;\&\; \mu^j < \overline{\mu}\right]\Pr\left[\mu^i < \overline{\mu} \;\&\; \mu^j < \overline{\mu}\right]$$

$$= 2\left(\tfrac{\overline{\mu}}{\overline{\mu}+1}\right)^2\left(\tfrac{\overline{\mu}+1}{2}\right)^2 = \frac{1}{2}\overline{\mu}^2 \tag{A9}$$

Denote the probability of partial investment breakdown as $h\left(\overline{\mu}(\delta)\right) \equiv h(\delta)$.

$$h(\delta) = 2\Pr\left[-\mu^i < \mu^j < -\tfrac{\overline{\mu}+1}{2}\right] = 2\Pr\left[\mu^j < -\tfrac{\overline{\mu}+1}{2}\right]\Pr\left[\mu^i + \mu^j > 0 \mid \mu^j < -\tfrac{\overline{\mu}+1}{2}\right]$$

The second probability is in fact just the probability that a drawing from a triangular distribution with support $\left[-2, \frac{1}{2}(1-\overline{\mu})\right]$ is strictly positive. Using the characteristics of the uniform and triangular distributions this yields:

$$h(\alpha, \delta) = \left(\tfrac{1-\overline{\mu}}{2}\right)\left[2\left(\tfrac{1-\overline{\mu}}{5-\overline{\mu}}\right)^2\right] = \tfrac{(1-\overline{\mu})^3}{(5-\overline{\mu})^2} \tag{A10}$$

Now since $f(\delta) = g(\delta) + h(\delta)$, combining equations A9 and A10 yields:

$$f(\delta) = \frac{1}{2}\overline{\mu}^2 + \tfrac{(1-\overline{\mu})^3}{(5-\overline{\mu})^2} = \tfrac{2-6\overline{\mu}+32\overline{\mu}^2-10\overline{\mu}^3+\overline{\mu}^4}{50-20\overline{\mu}+2\overline{\mu}^2}$$

Which is the required fraction of time when complete revelation is useful.

10. Proof of Proposition 8:

The extra delay in investment causes a small but strictly positive loss in joint payoffs of $\kappa > 0$. From proposition 7 complete revelation will be useful in countering full investment breakdown for the fraction of cases $\frac{1}{2}\overline{\mu}^2$. In these cases the maximum potential gain in profit is the sum of the two highest signal values which still lie in the investment breakdown signal region, i.e. the two highest signals for which $\mu^i \in [0, \overline{\mu}(\delta)]$, $\mu^j \in [0, \overline{\mu}(\delta)]$ and $\mu^i + \mu^j > 0$. This produces the maximum possible combined signal value of $2\overline{\mu}$. Furthermore the return will only occur at the point of complete revelation, therefore the payoffs must be discounted up to that point. Hence we have a maximum possible gain from countering full investment breakdown of $2\delta^{\tau^*}\overline{\mu}\left(\frac{1}{2}\overline{\mu}^2\right)$. Now complete revelation is also useful in countering partial investment breakdown which occurs in the fraction of cases $(5-\overline{\mu})^{-2}(1-\overline{\mu})^3$. In these cases the maximum possible gain is $1 + \frac{1}{2}(\overline{\mu}+1)$, which must again be discounted up to the point of complete revelation, and so we have a maximum possible gain from countering partial investment breakdown of $\delta^{\tau^*}\frac{1}{2}(3+\overline{\mu})(1-\overline{\mu})^3(5-\overline{\mu})^{-2}$. Combining all of this we have a necessary condition on the cost of information gathering:

$$C_g \leq 2\delta^{\tau^*}\overline{\mu}\left(\frac{1}{2}\overline{\mu}^2\right) + \delta^{\tau^*}\frac{1}{2}(3+\overline{\mu})(1-\overline{\mu})^3(5-\overline{\mu})^{-2} - \kappa$$

So we have as a weaker necessary condition:

$$C_g < \delta^{\tau^*}\left(\overline{\mu}^3 + \tfrac{(3+\overline{\mu})(1-\overline{\mu})^3}{2(5-\overline{\mu})^2}\right)$$

11. Proof of Proposition 9:

As $\tau^* \to \infty$ so complete revelation is known to occur a great many periods in the future. Consider the case where there is complete revelation and the threshold signal is

$\overline{\overline{\mu}}$. When $t = \tau^*$ the general value matching indifference condition is:[10]

$$\overline{\overline{\mu}} - \max\left\{\delta\overline{\overline{\mu}}, \tfrac{1}{2}\delta^{\tau^*-1}\left(\widehat{\mu}+1\right), 0\right\} = -\frac{1}{4}\delta\left(1+\widehat{\mu}\right)\left(3\widehat{\mu}-1\right)$$

Now $\tau^* \rightarrow \infty \Rightarrow \tau^* - 1 \rightarrow \infty$. Therefore, since $\delta \in (0,1)$, we have that $\lim_{\tau^*-1\rightarrow\infty} \delta^{\tau^*-1} = 0$. Hence in the limit expression the indifference condition becomes:

$$\overline{\overline{\mu}} - \max\left\{\delta\overline{\overline{\mu}}, 0, 0\right\} = -\frac{1}{4}\left(3\delta\overline{\overline{\mu}}^2 + 2\overline{\overline{\mu}} - 1\right)$$

This is exactly the same expression as when there is no complete revelation. Therefore as $\tau^* \rightarrow \infty$ the threshold value with complete revelation $\overline{\overline{\mu}}$ does indeed tend to the threshold signal value without complete revelation $\overline{\mu}$ as would be expected and their is no effect on the threshold value and therefore decision-making from complete revelation If complete revelation was costly it would therefore reduce firm's profits. Since from the proof of proposition 8, we know that the extra delay in investment causes a small but strictly positive loss in joint payoffs of $\kappa > 0$ the necessary and sufficient condition with $\tau^* \rightarrow \infty$ would involve the need for negative costs to compensate for κ.

References

Banerjee, A. V. (1992): "A Simple Model of Herd Behaviour," *Quarterly Journal of Economics*, 107, 797–817.

Bikchandani, S., D. Hirshleifer, and I. Welch (1992): "A Theory of Fads, Fashion, Custom and Cultural Change as Informational Cascades," *Journal of Political Economy*, 100, 992–1026.

Gale, D. (1996): "What Have We Learned from Social Learning?," *European Economic Review*, 40, 617–628.

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

Meyer, M. A. (1991): "Learning from Coarse Information: Biased Contests and Career Profiles," *Review of Economic Studies*, 58, 15–41.

Sgroi, D. (2000): "Optimizing Information in the Herd: Guinea Pigs, Profits and Welfare," Discussion paper, No. W14-2000. Nuffield College, Oxford. Forthcoming in Games and Economic Behavior.

[10]The general value matching indifference condition is a generalization of the indifference condition for the case when $\tau^* = 3$.