Optimizing Information in the Herd *

Daniel Sgroi
Churchill College, University of Cambridge,
Cambridge CB3 0DS, UK

Abstract

Herding arises when an agent's private information is swamped by public information in what Jackson and Kalai (1997) call a recurring game. The agent will fail to reveal his own information and will follow the actions of his predecessor and, as a result, useful information is lost, which might have highlighted a better choice for later decision-makers. This paper evaluates the strategy of forcing a sub-set of agents to make their decision early from the perspective of a social planner, and a firm with a valuable or valueless product. Promotional activity by firms can be explained as an attempt to overcome the herd externality and maximize sales.

J.E.L. classification numbers: D82 (asymmetric and private information), D83 (search, learning and information), L15 (information and product quality), M30 (marketing and advertising).

* The author would like to thank Ernesto Dal Bo, Paul Klemperer, Meg Meyer, John Morgan, Peter Sørensen, Andrew Temple and Lucy White for their useful comments. Financial support from AEA Technology is gratefully acknowledged. Email: daniel.sgroi@econ.cam.ac.uk.

1 Introduction

Should a firm with a new product release it to the entire market on the same day or pre-release to a select set of customers first? Should movie premieres be made available to members of the press and high profile celebrities or should films simply be released worldwide on the same day? Should a firm's product be released on the same day in several markets or released sequentially in different markets? Is it optimal for every country in the European Community to conduct separate drug trials or should they pool and have one drug trial? Is it sensible for the government to slowly release new drug treatments to doctors or should there be a core release to a select group of doctors first?

All of these questions relate to whether a slow sequential release of information is better than a discrete simultaneous release in the first instance followed by a slow sequential release thereafter. The term *better* may relate to consumer welfare (in the case of government planners) or profitability (in the case of firms). In both cases the approach taken in this paper is to model the learning process as a potential herding phenomenon. This allows a quick identification of the crucial trade-off involved between the gains for those late in the sequence who have access to more information if an initial group of "guinea pigs" is used, and the costs for those within the group of guinea pigs. More generally, forcing a sub-set of agents to make decisions quickly is important in any herd context when agents learn by observing each others' actions. Later decision-makers in such a model can gain a great deal from early movers, but they are also liable to become trapped in an information-damaging herd in which it is impossible to derive information from agents' actions, as they are simply copying their predecessors regardless of their own private information. Forced early movement overcomes this problem by providing later movers with genuinely informative actions and easily inferable private information. However, this is clearly damaging for those forced to decide early since they can no longer gain from observational learning themselves. We therefore have a clear trade-off.

This paper begins in section 2 by developing a herding or *informational cascade* model based on the work of Bikhchandani *et al.* (1992) which has been usefully categorized by Jackson and Kalai (1997) as a *recurring game*. The paper then moves on in section 3 by developing individual payoff functions, then a total consumer welfare function. Section 4 then examines the central trade-off from a social planner's perspective, finding that the optimal size of this sub-set of guinea pigs depends on the total number of agents in the sequence and the probability that a signal points in the right direction. This part of the paper shows that decision-making by a sequence of pairs is superior to sequential individual decision-making, and might therefore be an interesting informational justification for marriage! Section 5 considers the problem from a firm's perspective. The firm is naturally not interested in maximizing consumer welfare, but is instead interested in ensuring as many successful sales as possible. The firm can still use guinea pigs through some form of promotional campaign to raise its sales and the paper produces various comparative statics to show how knowledge of the number of consumers and the probability that their private information is correct is joined with the firms' knowledge of the true state to provide a clear optimal policy. A firm with a good product, will make more use of guinea pigs and intervene aggressively, doing so all the more as the number of consumers in the market increases and the quality of consumers' private information falls. There are clear implications for welfare as it is in the firm's interest to assist consumers to make the right choice. With a bad product the firm is indirectly interested in minimizing consumer welfare by convincing the maximum number to purchase when doing so is not optimal, and will attempt to prevent any observational learning about the firm's low quality product.

The findings in this paper provide at least one way to improve on the consumer welfare damaging herds that can start in many circumstances. Using the results in this paper it is possible to look for any potential herd situation in which a herd might move in the wrong direction and calculate an optimal number of guinea pigs, plus the percentage improvement over *laissez-faire*, that an attempt to force early movement could achieve. For example, if we were dealing with a group of 100 individuals, with private information which is correct $\frac{2}{3}$ of the time, it would be optimal to force 22 of these to move early, perhaps through the use of schemes to provide extra incentives for early movement (such as price reductions, or free gifts), limiting the number in this scheme to 22 (through randomization, or via prior selection). This would maximize overall consumer welfare and minimize the chance of an incorrect herd. For a firm we can now find some number of early and perhaps high profile consumers which would provide maximum sales. The firm could seek out members of the press or celebrities and provide them with its product at some reduced rate, and the extra sales certainly justify reducing prices to these early movers. In this way we can justify various observed practices from a card-holder day at a department store in which a select number of customers are urged to try out new stock, to the use of movie premieres in which celebrities and members of the press gather to pass judgement on a new film before the general release of the film. Although there are undoubtedly many reasons for such promotional activity by firms, such practices increase the likelihood of products achieving success in the marketplace, and work against the herd externality.

2 The Model

This section sets up a version of the model first used in the seminal herd paper by Bikhchandani *et al. (1992)* more generally characterized by Jackson and Kalai (1997) as a recurring game. Most early herd papers were primarily concerned with herd probabilities and proving that herds were likely or even certain to occur. The main focus of this paper is rather with payoff functions and the development of a total consumer welfare function. However, in order to derive such functions we cannot avoid looking at herd probabilities. We can motivate this in two ways. Firstly, we will consider the role of a social planner, government or regulatory agency. The agents might simply be consumers of a product and the social planner might simply wish to maximize consumer welfare. Later we will consider a firm which has a different aim. The firm wishes to maximize profit. We proxy this aim by maximizing sales since price is not a variable we consider, so the firm will attempt to induce as many agents to purchase as possible.

2.1 Preliminaries

Consider a sequence of $N \in \mathbb{N}_{++}$ agents, the ordering of which is exogenous and common knowledge, each deciding whether to adopt/purchase (Y) or not (N) some product or technology. Each agent observes the actions (Y or N) of his predecessors. The cost of adoption 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 sense that:

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

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

We assume that signals are identically distributed and note that the restriction on p suffices to produce informative but non-fully revealing signals. 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 $E\left[\pi_i\right]$ 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.*

2.2 Cascades

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 favour 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 behaviour 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 to will agents $4, 5, 6, ..., N$. We should note that:

Proposition 3 *A cascade once started will last forever.*

Proof. See appendix. ■

Consequently, 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.

2.3 Calculating Herd Probabilities

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. The appendix calculates the following functions conditional on $V = 1$. After an even number of n agents we have:

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

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

Note that from eq. 1 $\Pr[Y(n) \mid V = 1]$ is increasing in p and n but from eq. 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 a social planner 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. As we will see, firms and planners can manipulate the probability of a Y cascade, so they need not remain passive in the face of a potential herd.

3 Payoffs and Consumer Welfare

In this section we move away from standard herding concerns and instead focus on the calculation of individual payoffs for agents along the sequence. These calculations are then used to derive an expression for total consumer welfare. Evaluating the payoffs to the potential consumers and the total consumer welfare function, represents the first step in finding a consumer welfare-improving policy for a social planner. We will look at a more active role for the social planner, in which it can improve on the *laissez-faire* outcome in the next section. As we see later, consumer welfare also indirectly plays an important role in finding the optimal policy for a firm.

3.1 Individual Payoffs

We will begin by looking at the *ex ante* expected payoff of the first agent. Note that the two conditional prior probabilities are $\Pr[V=1] = \Pr[V=0] = \frac{1}{2}$ and signal probabilities are $\Pr[X_i = H \mid V=1] = \Pr[X_i = L \mid V=0] = p$. So we have:

$$E[\pi_1] = \frac{E[\pi_1|V=0]}{2} + \frac{E[\pi_1|V=1]}{2} = \frac{2p-1}{4} > 0 \tag{3}$$

Note also that $A_i = N \Rightarrow \pi_i = 0$ so we need only consider $A_i = Y$ when calculating payoffs. Since $p > \frac{1}{2}$ (signals are informative) we have that $E[\pi_1] > 0$. Now consider the second agent:

$$E[\pi_2] = \frac{\left[(1-p)^2 + \frac{p(1-p)}{2} + \frac{(1-p)p}{2}\right]\left(-\frac{1}{2}\right)}{2} + \frac{\left[p^2 + p(1-p)\right]\left(\frac{1}{2}\right)}{2} = \frac{2p-1}{4} = E[\pi_1] \tag{4}$$

So we have the interesting result, that $E[\pi_2] = E[\pi_1]$. In fact in general we can say that:

Proposition 4 *Agent k will have the same ex ante expected payoff as agent $k+1$ where $k \in N^{odd}$.*

Proof. See appendix. ■

It is interesting to note the similarity between this proposition and the role of information played in Meyer (1991). Meyer shows that in certain circumstances an extra draw from an informative distribution will not increase the probability of a decision-maker choosing correctly between two alternatives. She goes on to suggest biasing the second draw in such a way as to improve the probability of success and applies this to show the optimality of biased promotion tournaments.

We have the result that $E[\pi_n] = E[\pi_{n+1}]$ for $n \in N^{odd}$. This allows us to concentrate on agent $n \in N^{odd}$. Now consider what happens by agent $n-1$. There are three possibilities:

1. We have a Y cascade, so agent n will go Y.
2. We have a N cascade, so agent n will go N.
3. We have no cascade, so agent n will go Y if $X_n = Y$ and will opt for N if $X_n = L$.

This allows the calculation of agent n's payoff since we know that $\pi_n = \frac{1}{2}$ if $A_n = Y$ and $V = 1$, and $\pi_n = -\frac{1}{2}$ if $A_n = Y$ and $V = 0$, otherwise there will be no payoff. Therefore:

$$E\left[\pi_n \mid n \in N^{odd}\right] = -\frac{\{\Pr[Y(n-1)|V=0] + (1-p)\Pr[No(n-1)|V=0]\}}{4} + \frac{\{\Pr[Y(n-1)|V=1] + p\Pr[No(n-1)|V=1]\}}{4}$$

This simply evaluates the probability that agent n goes for Y in both states of the world then weights this by the payoffs of $\frac{1}{2}$ and $-\frac{1}{2}$. Note that if there is no cascade we have to weight the chance of agent n going Y by the chance of his observing a H signal in each state of the world, that is $(1-p)$ if $V = 0$ and p if $V = 0$. This can be simplified to:

$$E\left[\pi_n \mid n \in N^{odd}\right] = \frac{2p-1}{4}\frac{1-\left(p-p^2\right)^{\frac{n+1}{2}}}{1-p+p^2} \tag{5}$$

It is also interesting to note that:

$$E\left[\pi_n \mid Y(n)\right] = \frac{1-\left(p-p^2\right)^{\frac{n-1}{2}}}{1-(p-p^2)}E\left[\pi_1\right]$$

So all payoffs are related to the payoff of the first agent. Similarly calculations for agent $n \in N^{even}$ yield:

$$E\left[\pi_n \mid n \in N^{even}\right] = \frac{2p-1}{4}\frac{1-\left(p-p^2\right)^{\frac{n}{2}}}{1-p+p^2} \tag{6}$$

3.2 Aggregate Consumer welfare

Define aggregate consumer welfare Ω as equal to *ex ante* expected payoffs summed over all N agents, so $\Omega = \sum_{n \in N} \pi_n$. Proposition 4 immediately tells us that since $E\left[\pi_n\right] = E\left[\pi_{n+1}\right]$ for $n \in N^{odd}$ we can consider aggregate payoffs to be simply:

$$\sum_{n \in N} \pi_n = \begin{cases} 2\sum_{n \in N^{odd}} E\pi(n) \text{ for } N \in \mathbb{N}^{even} \\ 2\left[\sum_{n \in N^{odd}} E\pi(n)\right] - \pi(N) \text{ for } N \in \mathbb{N}^{odd} \end{cases} \tag{7}$$

Based on eq. 7 we can now calculate total *ex ante* expected consumer welfare since:

$$\sum_{n \in N^{odd}} E\left[\pi_n\right] = \frac{(2p-1)\left[N - \sum_{n \in N^{odd}}\left(p-p^2\right)^{\frac{n+1}{2}}\right]}{4(1-p+p^2)} \tag{8}$$

And so we have:

$$\Omega = \begin{cases} \frac{2p-1}{2(1-p+p^2)} \left[N - \sum_{n \in N^{odd}} \left(p - p^2\right)^{\frac{n+1}{2}}\right] \text{ for } N \in \mathbb{N}^{even} \\ \frac{2p-1}{2(1-p+p^2)} \left[N - \sum_{n \in N^{odd}} \left(p - p^2\right)^{\frac{n+1}{2}}\right] - \frac{1}{4}\left(2p-1\right) \frac{1-\left(p-p^2\right)^{\frac{N+1}{2}}}{1-p+p^2} \text{ for } N \in \mathbb{N}^{odd} \end{cases} \tag{9}$$

Or alternatively:

$$\Omega = \left[N - \sum_{n \in N^{odd}} \left(p - p^2\right)^{\frac{n+1}{2}} - \sum_{n \in N^{even}} \left(p - p^2\right)^{\frac{n}{2}}\right] \tag{10}$$

This then provides our benchmark. Any intervention that aims to strictly improve aggregate consumer welfare will have to raise overall payoffs above the non-intervention level of Ω. This then provides the benchmark for an interventionist social planner.

4 The Social Planner

Now we consider the role of a consumer welfare-maximizing social planner. We allow our social planner to force an additional $M \subset N$ agents to move in the first period. These agents will join with the first agent to give us a set of $M+1$ "guinea pigs". This has two major effects on consumer welfare:

1. Those $M+1$ who move first will not get access to later information and so will have less to use to help them make the optimal decision.
2. Those $N-(M+1)$ who move later will have more information at their disposal and so should have a better chance of making the correct decision.

We can therefore use the consumer welfare equation from the previous section and make two adjustments. We can use the sum for $N-(M+1)$ agents and then add a further $ME\left[\pi_1\right]$ to the total, but first we must adjust the probability of being caught in a cascade to reflect the extra information made available to later movers. For simplicity we will assume that $N \in \mathbb{N}^{even}$, though the results in the section above allow us to look at $N \in \mathbb{N}^{odd}$ in an equivalent way. Now define $\Omega = \Omega\left(M, N, p\right)$ as the general level of consumer welfare. Since $N \in \mathbb{N}^{even}$ we can use the simple form for $M = 0$ (no intervention) given as:

$$\Omega\left(0, N, p\right) = \frac{2p-1}{2(1-p+p^2)} \left[N - \sum_{n \in N^{odd}} \left(p - p^2\right)^{\frac{n+1}{2}}\right] \tag{11}$$

4.1 Polar Examples

Consider what happens when $M = 1$ so we force two agents to act without social learning, rather than just a single agent as in the model in Section 3. We then get 2 initial moves followed by a sequence of $N - 2$ agents. The two start-agents will share the same *ex ante* expected payoff as the first agent in the standard sequence, that is $\frac{1}{4}(2p-1)$. However, the new sequence from the third agent onwards will have access to more noiseless information; 3 signals instead of 2. Therefore the agent deciding after the start-agents may well face the prospect of a cascade before drawing his signal.

Consider the payoff of the third agent. He faces three possibilities, a Y cascade, N cascade or no cascade as before, but the probabilities are slightly different. Consider $V = 0$:

$$E[\pi_3 \mid V = 0, M = 1] = (1-p)^2\left(-\tfrac{1}{2}\right) + 2p(1-p)^2\left(-\tfrac{1}{2}\right)$$

Now consider $V = 1$:

$$E[\pi_3 \mid V = 0, M = 1] = p^2\left(\tfrac{1}{2}\right) + 2p^2(1-p)\left(\tfrac{1}{2}\right)$$

Therefore his unconditional payoff is:

$$E[\pi_3 \mid M = 1] = \tfrac{1}{4}(2p-1)\left(2p - 2p^2 + 1\right) \tag{12}$$

Now note that in the non-intervention case we have:

$$E[\pi_3 \mid M = 0] = \tfrac{1}{4}(2p-1)\left(p - p^2 + 1\right) \tag{13}$$

Now consider which has the higher value. Let us assume that eq. 12 has the higher value than eq. 13 and see if this is true: $\frac{1}{4}(2p-1)\left(2p-2p^2+1\right) > \frac{1}{4}(2p-1)\left(p-p^2+1\right) \Rightarrow p - p^2 > 0$ which is clearly true for $p \in (0, 0.5)$. So the third agent gains slightly from the extra noiseless information, while the first two agents each receive what they would have anyway under non-intervention, expected payoffs of $\frac{1}{4}(2p-1)$. Clearly agent 4 cannot be any worse off, so we have found an unambiguously better situation, hence $\Omega(1, N, p) > \Omega(0, N, p)$.

Now consider $M = N-1$. In this case all the agents that usually move after the first agent are instructed to move immediately without scope for learning. The result here requires no analysis, all will move and gain the same *ex ante* expected payoff of $\frac{1}{4}(2p-1)$, so $\Omega(N-1, N, p) = \frac{1}{4}N(2p-1)$. Now since $\Omega(0, N, p) > \frac{1}{4}N(2p-1)$ this is clearly a worse situation for total consumer welfare than non-intervention. So we see by example that $\Omega(0, N, p)$ is not the highest consumer welfare achievable, but is also not the lowest. We need to find another candidate for optimal consumer welfare and therefore the optimal structure for our problem. This also constitutes a proof that:

Proposition 5 *The optimal level of M (number of guinea pigs) will lie strictly between 0 and $N-1$.*

Proof. As shown above $\Omega(1,N,p) > \Omega(0,N,p)$ and $\Omega(0,N,p) > \Omega(N-1,N,p)$ which immediately proves the proposition. ■

4.2 The Welfare-Maximizing Structure

So far we have $\Omega(1,N,p) > \Omega(0,N,p)$ and $\Omega(0,N,p) > \Omega(N-1,N,p)$. The agent moving immediately after the $M+1$ agents make the initial move, agent $M+2$, will see various possibilities. Divide the set of size $M+1$ into subsets M_Y and M_N which are made from those who choose Y and N respectively. Agent $M+2$ faces three mutually exclusive possibilities:

1. $M_Y - M_N \geq 2$ which will create a Y cascade.
2. $M_N - M_Y \geq 2$ which will create a N cascade.
3. $M_Y - M_N \in (-2, 2)$ which produces no cascade.

Part 4 of the appendix derives the function:

$$\Omega(M,N,p) = \frac{(2p-1)(M+1)}{4} + \frac{\frac{1}{2}(2p-1)(M+1)!\,p^{\frac{M+2}{2}}(1-p)^{\frac{M+2}{2}}}{\left[\left(\frac{M+1}{2}\right)!\right]^2} \sum_{z=1}^{\frac{N-M}{2}} \sum_{s=1}^{\frac{2z-2}{2}} \left(\frac{p^2-p+1}{2}\right)^s$$

$$+\frac{N-M-1}{4}\left\{\frac{(2p-1)(M+1)!\,p^{\frac{M+2}{2}}(1-p)^{\frac{M+2}{2}}}{\left[\left(\frac{M+1}{2}\right)!\right]^2} + \sum_{x=0}^{\frac{M-1}{2}} \frac{(M+1)!\left[p^{M+1-x}(1-p)^x - (1-p)^{M+1-x}p^x\right]}{(M+1-x)!\,x!}\right\}$$

We then need to optimize this with respect to M and a partial further calculation is made in the appendix.

Despite the complexity of the function, a number of features stand out:

1. There is a unique maximum for any value of p and N assured by the concavity of $\Omega(M,N,p)$.

2. The maximum is in the interior of the range of M for all values of p and non-trivial N.

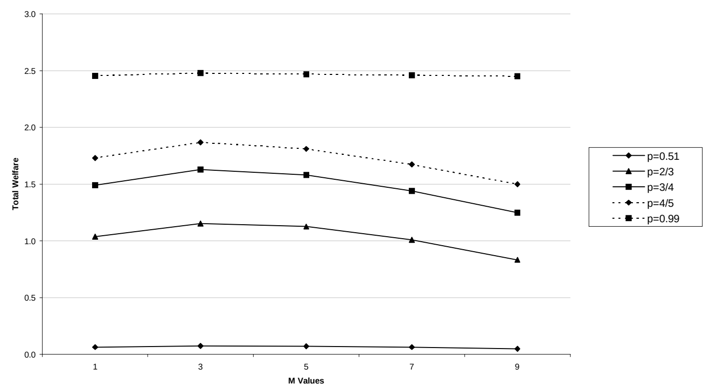

Figure 1: Optimal M Values, $N = 10$

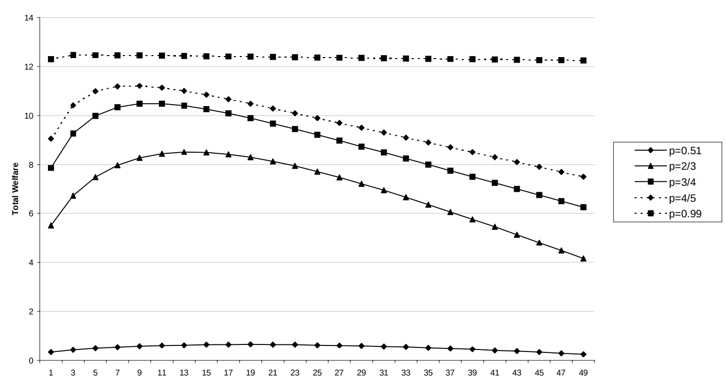

Figure 2: Optimal M Values, $N = 50$

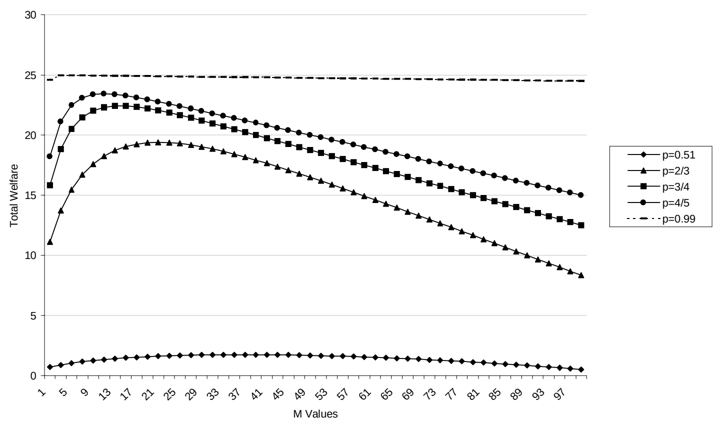

Figure 3: Optimal M Values, $N = 100$

To give some examples figures 1 to 3 show how the total consumer welfare value $\Omega(M, N, p)$ evolves as $M + 1 \in \mathbb{N}^{even}$ increases, with N and p held fixed. Figure 1 fixes $N = 10$, figure 2 fixes $N = 50$ and figure 3 fixes $N = 100$. Figure 2 provides different optimal values for M for different values of p. However, as in figure 1 the optimal value of M is never 0 or N, instead it is given at some interior value. Figure 3 provides a convex shape and a set of interior optima for M. For $N = 10$ the optimal value of M is 4 for the five values of p examined. The existence of a trade-off between the value of guinea pigs and the loss to the guinea pigs' consumer welfare because of their failure to learn from others is captured by the convex shape of the function, which is a general feature, and is clearly visible in figures 2 and 3. Table I summarizes figures 1, 2 and 3 by providing optimal M values for the given values of p and N. It is important to note the restriction $M + 1 \in \mathbb{N}^{even}$.

To give some feeling for the results consider one example. For $N = 100$ and $p = \frac{2}{3}$, we have an optimal value of $M = 21$. This states that total consumer welfare will be maximized if we have a structure in which 21 of 100 agents join agent 1 and move immediately. Then the remaining 78 agents move in sequence. This gives agent 23 access to 23 signals unpolluted by possible herding, and guarantees that all agents will get full access to at least 23% of all signals. This provides significantly more consumer welfare than a standard herd when all act in strict sequence (twice as much) or when there is no social learning so all act independently (over twice as much).

To give some comparative statics note that as p rises the optimal number of guinea pigs falls. This is the case since a rising p increases the probability of a good decision and a correct herd which is good for consumer welfare. The gains from increased guinea pigs are therefore not so great, while the disadvantage of reducing the information available to them is still present. As N rises the number of guinea pigs needs to rise, though the proportion falls.

Table I: Optimal M Values for Consumer Welfare Given Values of p and N

p		0.51	2/3	3/4	4/5	0.99
N	50	19	13	11	9	3
	100	37	21	13	11	5

4.3 A Note about Marriage

Before we move on to consider the firm, we will examine one alternative structure implied by Proposition 4. The proposition reveals that for $k \in N^{odd}$ payoffs will always be identical for agents k and $k+1$, the reasoning revolved around the valueless nature of agent k's information to agent $k+1$. This immediately tells us that it is at least weakly better for welfare to have all decisions "made in pairs".

Remark 6 *Welfare is weakly improved by grouping by going from a system of sequential decision-making with no guinea-pigs to a system where all decision making is done by a sequence of pairs, where each pair-member must decide without knowledge of their partner's decision.*

Consider a structure in which agents 1 and 2 decide simultaneously, then agents 3 and 4 observe agents 1 and 2, and also decide simultaneously, etc. We have eliminated only the useless observation of the direct predecessor when that predecessor can be indexed by an odd number. This will increase the information available to each successive pair in the same way as increasing the number of guinea pigs, but without the associated cost to the pair. The net result suggests that simultaneous decisions by pairs is at least as good as a strict sequence of decision-making. Taken literally this can provide an interesting argument for the informational gain of splitting the population into pairs via formal marriage or a similar informal link, but where each individual decides without reference to the decision made by their partners!

Remark 7 *In a sequential decision-making system with no guinea pigs, deference to one's partner is optimal.*

This remark is simply a corollary to proposition 1. Since agent $k+1 \in N^{even}$ will be behaving optimally by copying the decision made by agent k, we have a theory of deference to one's partner! On informational grounds one possible optimal policy in a sequential decision-making world with no guinea pigs would be for all husbands to defer to their wives' decisions, or *vice versa.*

5 The Firm

Now we move away from the objective of consumer welfare maximization and instead consider the aims of the firm. Consider a single firm with a product it wishes to sell. 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 then compare the optimal actions of a firm with the optimal actions of a social planner.

5.1 Promoting a Good Product

Assume that $V = 1$ so the decision to purchase is the right one. Unfortunately the firm cannot convince all consumers that this is the case. However, the firm can manipulate the structure of the herd in much the same way as the social planner through the use of guinea pigs. In this case 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.

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 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.

3. 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 5 of the appendix reduces the firm's problem to:

$$\max_{M}\left\{p\,(M+1)+\sum_{x=\frac{M+3}{2}}^{M+1}\frac{(N-M-1)(M+1)!p^{x}(1-p)^{M+1-x}}{x!(M+1-x)!}\right. \tag{14}$$

$$\left.+\sum_{x=0}^{\frac{N-M-1}{2}}\frac{(M+1)!p^{\frac{M+3}{2}}(1-p)^{\frac{M+1}{2}}\left(p-p^{2}\right)^{x}[2+(p+1)(N-M-2x)]}{2\left(\frac{M+1}{2}\right)!\left(\frac{M+1}{2}\right)!}\right\}$$

Differentiating this requires the use of the digamma and hypergeometric distributions and produces a fairly complex result. However, some comparative statics should provide some intuition for the result. Table II gives the optimal choice of M for various values of p and N, and provides an interesting comparison with table I.

Table II: 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	15	13	11	7
	100	29	21	13	9

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

M		9	29	49	69	89
	0.51	53	53	53	53	52
p	2/3	86	88	83	77	70
	4/5	97	94	90	86	82

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

M		29	49	69	89
N	100	88 (88)	83 (83)	83 (83)	70 (70)
	150	137 (91)	133 (88)	127 (84)	120 (80)

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

M/N		9%	25%	49%	75%	91%
	50	82%	85%	81%	74%	69%
N	100	86%	89%	83%	75%	69%
	150	89%	90%	83%	75%	69%

Analyzing tables II to V 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. So to give an example a figure of around 20% of the market for $p = \frac{2}{3}$ seems reasonable for an N of 100. 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.

5.2 Promoting a Bad Product

Now we consider the case when $V = 0$. Part 6 of the appendix shows that the firm's problem has now changed to become:

$$\max_{M} \left\{ (1-p)(M+1) + \sum_{x=\frac{M+3}{2}}^{M+1} \frac{(N-M-1)(M+1)!(1-p)^{x} p^{M+1-x}}{x!(M+1-x)!} \right. \tag{15}$$

$$\left. + \sum_{x=0}^{\frac{N-M-1}{2}} \frac{(M+1)!(1-p)^{\frac{M+3}{2}} p^{\frac{M+1}{2}} \left(p-p^{2}\right)^{x} [2+(2-p)(N-M-2x)]}{2\left(\frac{M+1}{2}\right)!\left(\frac{M+1}{2}\right)!} \right\}$$

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 VI repeats the findings of table II for the new state. Note that the best possible policy is now always to ensure that as many consumers as possible choose independently and hence avoid possible observational learning. There is then no real trade off for a firm when the state is $V = 0$, and the unique optimal value of expression 15 for $p \in (0.5, 1)$ is in fact $M = N - 1$. Rather than being informative about the optimal M the following tables instead show how sales are affected by alternative choices of M.

Table VI: 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	49	49	49	49
	100	99	99	99	99

Table VII, much like table III, 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 confirm that a firm with a bad product always does better by restricting observational learning, and forcing all consumers to make decisions independently.

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

M		9	29	49	69	89
	0.51	40	43	45	46	48
p	2/3	12	12	17	23	30
	4/5	3	6	10	14	18

Table VIII carries out the same process as table IV 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 IX 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. The best possible result for the firm when $V = 0$ is to sell to a share equal to $(1 - p)$ of the market and can do this by forcing all consumers to choose independently and hence rely only on their private information. What would ensue is not really the same as a promotional campaign in the sense of a firm wth a good product but rather an on the spot offer of purchase to *every* consumer without allowing them the advantage of observing previous decisions.

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

M		29	49	69	89
N	100	12 (12)	17 (17)	23 (23)	30 (30)
	150	13 (9)	17 (12)	23 (16)	30 (20)

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

M/N		9%	25%	49%	75%	91%
	50	16%	15%	19%	26%	31%
N	100	12%	11%	17%	25%	31%
	150	10%	9%	17%	25%	31%

5.3 A Note about Welfare

The firm's desire to maximize the number of units sold when the state is $V = 1$ is intuitively consumer welfare maximizing, since consumers would wish to purchase given $V = 1$. Therefore the firm's problem is effectively identical to the problem $\max_M \Omega\left(M, N, p \mid V = 1\right)$. However, if we consider $V = 0$ it should be immediately obvious that the firm will not maximize consumer welfare. In fact the firm is interested in *minimizing* consumer welfare. To summarize:

Table X: Consumer Welfare

Agent	Social Planner	Firm, $V = 1$	Firm, $V = 0$
Optimal Policy	Maximize Ω	Maximize Ω	Minimize Ω

5.4 A Note about Revelation

If a social planner has a reasonable knowledge of p and of N it might be possible to deduce whether the firm is promoting a good product ($V = 1$) or a bad product ($V = 0$) based on the firm's choice of M. Since we know that a firm with a bad product is likely to try to force consumers to decide with only their own private signals to inform them, a social planner or indeed consumer might use this fact to help reveal a firm's own preceptions of its product. The firm's own choice of M might therefore become a signal of V. If we follow this avenue of thought we would need to consider ways in which a firm in state $V = 0$ can emulate the firm in the state $V = 1$ and this returns us to more traditional signaling models which might work together with herding concerns to provide an interesting extension to this paper. One way of completely removing this concern is to assume that *only the firm knows the value of* N *and* M, which is in fact quite a reasonable assumption in most markets. If a consumer is approached and asked to be a guinea pig the key question is whether that consumer knows how many others are being asked to decide early and how large is the potential market. If consumers do not know the answer to these questions then backward inference to V becomes impossible and signaling concerns can be safely put to one side.

References

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

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

Jackson, M.O. and Kalai, E. (1997). "Social Learning in Recurring Games," *Games and Economic Behavior* **21**, 102-134.

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

Appendix

1. Proof of Proposition 3. There exist many proofs that an informational cascade will last forever in the literature. The following is therefore short and intuitive. If an agent i is in a cascade, then by definition, regardless of his signal he will follow his predecessor's action. Therefore his action conveys no information and agent $i+1$ can only draw the same inference from all previous actions as agent i. Agent $i+1$'s information set is therefore made up of exactly the same public information plus one private observation. Since agent $i+1$'s signal is drawn from the same distribution as agent i and since agent i's action was no dependent on his signal neither will the action of agent $i+1$ be dependent on his own private signal. By induction, as all agents after agent $i+1$ also have draws from the same distribution they can only draw the same inference from all previous actions as agent i. Therefore all agents after i will simply repeat the action of agent $i-1$.

2. Derivation of Conditional Herd Probabilities. 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. So $1-\Pr[Y(n)]-\Pr[N(n)]=\Pr[No(n)]$ for all n. 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 $\Pr[N(2)]=\frac{1}{2}\left(1-p+p^2\right)$. 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 $\Pr[No(n)]=(\Pr[No(2)])^n=\left(p-p^2\right)^{\frac{n}{2}}$ for no cascade, and $\Pr[Y(n)]=\Pr[N(n)]=\frac{1}{2}\{1-\Pr[No(n)]\}=\frac{1}{2}\left[1-\left(p-p^2\right)^{\frac{n}{2}}\right]$ for a Y or N cascade. Now note that as $p\to 1$ cascades tend to start sooner, so more precise signals raise the probability of histories that lead to the correct cascades where we define correct cascades as a Y cascade if $V=1$ or a N cascade if $V=0$. As Bikhchandani *et al.* (1992) note, the probability of not being in a cascade falls exponentially with the number of agents, for example for a very noisy signal, $p=\frac{1}{2}+\varepsilon$ with $\varepsilon\to 0$ we have $\Pr[No(10)]<0.1$. Now we conclude this part of the

analysis by considering the probability of the correct or incorrect cascade occurring:

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

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

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

After an even number of n agents we have:

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

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

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

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

Now using the sum of a geometric series we have:

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

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

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

3. Proof of Proposition 4. We need to show that agent k will have the same *ex ante* expected payoff as agent $k+1$ where $k \in N^{odd}$. Consider a sequence of N individuals and let $k < N$ be an odd number. Now consider an arbitrary agent k and agent $k+1$. Agent k will either be in a cascade or not. We will consider these possibilities in turn. 1. If agent k is in a cascade then agent $k+1$ will simply duplicate agent k's action and so will obtain the same expected payoff. 2. If agent k is not in a cascade then all the revealed information from agents

1 to $k-1$ must be neutral, i.e. can be ignored. This is only possible if k is odd, as otherwise the set of information arising from agents 1 to $k-1$ would always be biased towards one choice. Agent k will then decide based upon his own signal, choosing $A_K = Y$ if $X_k = H$ and $A_K = N$ if $X_k = L$. The signal of agent k is therefore perfectly inferable to agent $k+1$. Now we examine the return to agent $k+1$. If agent $k+1$ receives the same signal as agent k, so $X_{k+1} = X_k$ then he will choose as agent k did and therefore receive the same expected payoff. If agent k receives a different signal, so $X_{k+1} \neq X_k$, then he will have received one signal suggesting Y and one suggesting N and will therefore be perfectly indifferent between the two. Indifference implies the expected return from choosing either option is the same. Therefore one optimal policy for agent $k+1$ is always to duplicate the choice of agent k regardless of his own signal. Since this is always optimal the expected payoffs of the two agents must be the same. For completeness we should note that as shown earlier payoffs are the same for agents 1 and 2, and since it has been shown true for arbitrary agents k and $k+1$, where k is odd, it will be generally true by induction for $k < N$ where k is an odd number.

4. The Optimal Choice of M by a Social Planner. Note that when $V = 1$, we have a probability of a Y cascade being introduced by our $M+1$ agents of:

$$p^{M+1} + \tfrac{(M+1)!}{M!} p^M (1-p) + \tfrac{(M+1)!}{(M-1)!2!} p^{M-1} (1-p)^2 + ... + \tfrac{(M+1)!}{(M+1-x)!x!} p^{M+1-x} (1-p)^x$$

Where x is the highest whole number less than $\frac{M-1}{2}$. We can simplify this to:

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

For $V = 0$ by a similar calculation we have:

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

Which yields the unconditional probability:

$$\Pr[M_Y - M_N \geq 2] = \tfrac{1}{2} \sum_{x=0}^{\frac{M-1}{2}} \tfrac{(M+1)!}{(M+1-x)!x!} \left[p^{M+1-x} (1-p)^x + (1-p)^{M+1-x} p^x \right] \qquad (16)$$

Now for $M_Y = M_N \in (-2, 2)$ which results in no cascade we have $M_Y = M_N$ in the case when

$M+1 \in \mathbb{N}^{even}$. We will from now on concentrate on this case for simplicity. So we have:

$$\Pr\left[M_Y = M_N\right] = \frac{(M+1)!}{\left[\left(\frac{M+1}{2}\right)!\right]^2} p^{\frac{M+2}{2}} (1-p)^{\frac{M+2}{2}} \tag{17}$$

We would need $X_{M+2} = Y$ to induce our $M+2$nd agent to choose Y after observing $M_Y = M_N$, which will happen with probability p if $V = 1$ and probability $(1-p)$ if $V = 0$ and yield a payoff of $\frac{1}{2}$ or $-\frac{1}{2}$ respectively. Therefore we have a payoff under the no cascade assumption of:

$$E\left[\pi_{M+2} \mid M_Y = M_N\right] = \frac{2p-1}{4} \left\{ \frac{(M+1)!}{\left[\left(\frac{M+1}{2}\right)!\right]^2} p^{\frac{M+2}{2}} (1-p)^{\frac{M+2}{2}} \right\} \tag{18}$$

Now we add in the probability of our agent being caught in a herd on Y, to yield an unconditional expected payoff:

$$E\left[\pi_{M+2}\right] = \frac{1}{4} \left\{ \frac{(2p-1)(M+1)! p^{\frac{M+2}{2}} (1-p)^{\frac{M+2}{2}}}{\left[\left(\frac{M+1}{2}\right)!\right]^2} + \sum_{x=0}^{\frac{M-1}{2}} \frac{(M+1)!\left[p^{M+1-x}(1-p)^x - (1-p)^{M+1-x} p^x\right]}{(M+1-x)!x!} \right\} \tag{19}$$

For example, in the $M = 1$ case, eq. 19 gives us $E\left[\pi_3 \mid M = 1\right] = \frac{1}{4}(2p-1)\left(1 + 2p - 2p^2\right)$ just as in eq. 12. The total consumer welfare expression is:

$$\Omega(M, N, p) = \frac{1}{4}(2p-1)(M+1) + \left\{E\left[\pi_{M+2}\right] + E\left[\pi_{M+3}\right] + ... + E\left[\pi_N\right]\right\}$$

Which we simply need to express fully in terms of M and then optimize with respect to M. We need to note the relationship between $E\left[\pi_{M+2}\right]$ and $E\left[\pi_{M+i}\right]$ for $i \in \{3, 4, ..., N\}$. First note that from proposition 4. Therefore we have:

$$\Omega(M, N, p) = \frac{1}{4}(2p-1)(M+1) + 2 \sum_{i=1}^{\frac{N-M}{2}} E\left[\pi_{M+2i}\right]$$

Comparing agent $M+4$ to agent $M+2$ we see that agent $M+4$ basically gets the same payoff but faces a slightly higher chance of being in a Y cascade which is good when $V = 1$ and bad when $V = 0$. This extra probability is basically just the chance that no cascade was initiated before agent $M+2$, but that agent $M+2$ and agent $M+3$ both went for action Y. This occurs with probability $\Pr\left[No \mid M+2\right]\left\{\frac{1}{2}\left[p^2 + \frac{1}{2}p(1-p)\right] + \frac{1}{2}\left[(1-p)^2 + \frac{1}{2}p(1-p)\right]\right\}$ and the net gain is simply $\frac{1}{4}(2p-1)$. So we simply get:

$$E[\pi_{M+4}] = E[\pi_{M+2}] + \frac{p^2-p+1}{2}\frac{\frac{1}{4}(2p-1)(M+1)!p^{\frac{M+2}{2}}(1-p)^{\frac{M+2}{2}}}{\left[\left(\frac{M+1}{2}\right)!\right]^2}$$

For agent $M+6$ we have a similar calculation except that now we need no cascade before agent $M+4$:

$$E[\pi_{M+6}] = E[\pi_{M+4}] + \left(\frac{p^2-p+1}{2}\right)^2 \frac{\frac{1}{4}(2p-1)(M+1)!p^{\frac{M+2}{2}}(1-p)^{\frac{M+2}{2}}}{\left[\left(\frac{M+1}{2}\right)!\right]^2}$$

And so on, so in general for $z > 1$:

$$E[\pi_{M+2z}] = E[\pi_{M+2z-2}] + \left(\frac{p^2-p+1}{2}\right)^{\frac{2z-2}{2}} \frac{\frac{1}{4}(2p-1)(M+1)!p^{\frac{M+2}{2}}(1-p)^{\frac{M+2}{2}}}{\left[\left(\frac{M+1}{2}\right)!\right]^2}$$

Clearly we have a nested structure and can therefore resolve this as follows:

$$E[\pi_{M+6}] = E[\pi_{M+2}] + \left[\frac{p^2-p+1}{2} + \left(\frac{p^2-p+1}{2}\right)^2\right] \frac{\frac{1}{4}(2p-1)(M+1)!p^{\frac{M+2}{2}}(1-p)^{\frac{M+2}{2}}}{\left[\left(\frac{M+1}{2}\right)!\right]^2}$$

And in general:

$$E[\pi_{M+2z}] = E[\pi_{M+2}] + \frac{\frac{1}{4}(2p-1)(M+1)!p^{\frac{M+2}{2}}(1-p)^{\frac{M+2}{2}}}{\left[\left(\frac{M+1}{2}\right)!\right]^2} \sum_{s=1}^{\frac{2z-2}{2}} \left(\frac{p^2-p+1}{2}\right)^s \tag{20}$$

So for total consumer welfare we have:

$$\Omega(M, N, p) = \frac{(2p-1)(M+1)}{4} + 2\sum_{z=1}^{\frac{N-M}{2}} E[\pi_{M+2z}] \tag{21}$$

$$= E[\pi_{M+2}](N-M-1) + \frac{(2p-1)(M+1)}{4} + \frac{\frac{1}{2}(2p-1)(M+1)!p^{\frac{M+2}{2}}(1-p)^{\frac{M+2}{2}}}{\left[\left(\frac{M+1}{2}\right)!\right]^2} \sum_{z=1}^{\frac{N-M}{2}} \sum_{s=1}^{\frac{2z-2}{2}} \left(\frac{p^2-p+1}{2}\right)^s$$

We can now insert the expression for $E\left[\pi_{M+2}\right]$ to yield:

$$\Omega\left(M,N,p\right) = \frac{(2p-1)(M+1)}{4} + \frac{\frac{1}{2}(2p-1)(M+1)!p^{\frac{M+2}{2}}(1-p)^{\frac{M+2}{2}}}{\left[\left(\frac{M+1}{2}\right)!\right]^2} \sum_{z=1}^{\frac{N-M}{2}} \sum_{s=1}^{\frac{2z-2}{2}} \left(\frac{p^2-p+1}{2}\right)^s$$

$$+\frac{N-M-1}{4}\left\{\frac{(2p-1)(M+1)!p^{\frac{M+2}{2}}(1-p)^{\frac{M+2}{2}}}{\left[\left(\frac{M+1}{2}\right)!\right]^2} + \sum_{x=0}^{\frac{M-1}{2}} \frac{(M+1)!\left[p^{M+1-x}(1-p)^x-(1-p)^{M+1-x}p^x\right]}{(M+1-x)!x!}\right\}$$

Now we differentiate with respect to M using a convenient set of abbreviations:

$$A = \frac{1}{2}\left(N-M+2\right)$$

$$B = \left(\frac{M+1}{2}\right)!$$

$$C = \frac{1}{2}\left(p^2+p+1\right)$$

$$D = -\frac{A\left(p^4-2p^3+p^2-1\right)-4C^A-p^2+2p-3+p^4-2p^3}{\left(p^2-p-1\right)^2}$$

$$E = \left(2p-1\right)\left(M+1\right)!p^{\frac{M+2}{2}}\left(1-p\right)^{\frac{M+2}{2}}$$

$$\Omega^{'}\left(M,N,p\right) = \frac{2p-1}{4} + \frac{E}{B^2}\left(\frac{2D+N-M-1}{2}\frac{2\Psi(M+2)+\ln(p)+\ln(1-p)-2\Psi\left(\frac{M+3}{2}\right)}{4}\right.$$

$$\left.-\frac{2C^A\ln C-\frac{1}{2}p^4+p^3-\frac{1}{2}p^2+\frac{1}{2}}{2(p^2-p-1^2)}-1\right) - \frac{1}{4}\sum_{x=0}^{\frac{M-1}{2}}\frac{(M+1)!\left(p^{(M+1-x)}(1-p)^x-(1-p)^{(M+1-x)}p^x\right)}{(M+1-x)!x!}$$

$$+\frac{N-M-1}{4}\frac{\partial}{\partial M}\sum_{x=0}^{\frac{M-1}{2}}\frac{(M+1)!\left(p^{(M+1-x)}(1-p)^{x}-(1-p)^{(M+1-x)}p^{x}\right)}{(M+1-x)!x!}$$

Where $\Psi(\cdot)$ is the digamma function. This expression can then be set equal to zero to yield an implicit function for the optimal value of M given p and N. Furthermore $\Omega_{MM}(M,N,p)<0$ across the whole range of p and N implying that $\Omega(M,N,p)$ is concave, which is sufficient to provide a unique maximum.

5. The Firm's Problem when V=1. Expression 14 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:

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

The derivation of the final part of expression 14 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:

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

Which is simply the only sequence of signals in the initial M (a precisely identical number o fhigh and low signals) multiplied by the number of possible permutations of such an occurrence. Alternatively it can be calculated as 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. We then see that the agents from $M+2$ onwards are in an identical position to those in a standard sequemtial herdng model excepting of course that their total purchases must be pre-multiplied by the probability that they are not already in a herd induced by the initial $M+1$. By dividing the set of agents into odd and even agents, we see that odd agents can never initiate a herd but even agents can. We first consider the odd agents who are not in a herd. They will purchase with probability p but we must weight their decision by the falling probability that they are not in a herd. This is captured by the expression $p\sum_{x=0}^{\frac{N-M-1}{2}}\left(p-p^2\right)^x$. The expression for the remaining even agents is a little more complex and incorporates the chance that each even agent initiates a cascade which will in turn generate sales from future odd and even agents. This is weighted by the falling size of the remaining population of agents which therefore falls inside a summation of their total impact on the market as follows: $\frac{1}{2}\left(p^2+p\right)\sum_{x=0}^{\frac{N-M}{2}}\left(N-M-2x\right)\left(p-p^2\right)^x$. Combining all four elements yields expression 14 in the main text. The function yielded is a function of p, N and M. Of these we allow the firm to vary only M making the final problem to maximize expression 14 by choice of M.

6. The Firm's Problem when V=0. Expression 15 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 $\Pr\left[Q_{M+1}\geq\frac{M+1}{2}+1\mid V=0\right]$. 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:

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

The population size remaining is also the same. However, the new Y cascade probability changes the final term to a slightly different combination of purchases by odd agents outside herds, now $(1-p)\sum_{x=0}^{\frac{N-M-1}{2}}\left(p-p^2\right)^x$ and the contribution of even agents, which is now $\frac{1}{2}\left(p^2-p\right)\sum_{x=0}^{\frac{N-M}{2}}\left(N-M-2x\right)\left(p-p^2\right)^x$. Combining these three parts once again yields the function which the firm will maximize by choice of M.