A Theory of Demand for Search Goods

Ken Hendricks
University of Texas at Austin

Alan Sorensen
Stanford University & NBER

Tom Wiseman
University of Texas at Austin

April 15, 2008
Preliminary and Incomplete, Please do not quote

Abstract

We examine demand for search goods, where a consumer’s valuation for a product depends on the product’s quality and a consumer-specific preference component. Both components are initially unknown to the consumer but she can learn about her preferences from the purchasing decisions of other consumers and through costly search. We use a variant of the sequential search models of Banerjee (1992), Bikchandani, Hirshleifer, and Welch (1992) and Smith and Sorensen (2000). The search option leads to different market dynamics and outcomes than the standard herding models. The results explain the unpredictability of sales, even conditional on quality, in markets for search goods like books, movies, and music and the highly skewed distribution of sales in these markets. The model yields testable predictions regarding the impact of product quality, search costs, and price on the probability of a “bad herd” in which a high-quality product ends up with low sales. We test the model’s predictions using Salganik, Dodds, and Watts’ (2006) data from an experimental music market.
1 Introduction

In cultural markets such as books, music and movies, consumers face an overwhelmingly large choice set that grows over time as many new products flow into the market each week. A small fraction of the products are profitable and, even among profitable products, the distribution of sales is highly skewed with top products selling many times more than the median product. The skewness may simply be a reflection of the products’ relative qualities. Rosen (1981) develops a theory of demand for talent that attributes the inequality of rewards to differences in observable product quality. He argues that more talented artists produce higher quality products for essentially the same costs. In equilibrium, higher quality products sell more units at higher unit prices. As a result, the market reward function is convex, with small differences in quality leading to large differences in returns. The main problem with applying this explanation to cultural markets is that higher quality products tend to sell for the same prices as lower quality products. For example, Hendricks and Sorensen (2007) compare prices of three kinds of CDs offered for sale by a major online retailer: new releases, catalog titles by artists with new releases, and catalog titles by artists without new releases at the time. They found that not only did hits sell for the same prices as duds, catalog titles tended to sell for the same prices as new releases. The lack of price variation across quality is also true of new releases of hardcover books, paperback books and videos. The source of the skewness in revenues in these markets is mainly variation in quantity, not price.

Salganik and Watts (2007) point to another feature of cultural markets that Rosen’s model fails to explain: the unpredictability of demand. The fact that most products fail to be profitable suggests that experts are unable to predict flops very well. Salganik and Watts cite a number of examples which suggest that experts are also unable to predict which products will be hits\(^1\) and quote industry executives as referring to their work as “a crap game”. Salganik and Watts argue, however, that the lack of predictability is not simply inability to identify product quality. The problem appears to be more fundamental: the mapping from product quality to outcomes is itself stochastic. That is, even if all products have the same intrinsic quality, sales are likely to be highly skewed. Salganik and Watts

\(^1\)The examples include the first book in the Harry Potter series which was rejected by eight publishers before being accepted; Star Wars, which Fox believed was all but certain to be a dud; and Bob Dylan’s big hit “Like a Rolling Stone” which was almost not released.
attribute the unpredictability in sales to social influence: individual consumers want to buy what others buy. As a result, a product’s success or lack of success reinforces itself, leading to unpredictable consequences at the collective level. Salganik, Dodds, and Watts (2006) (hereafter referred to as SDW) investigate this hypothesis experimentally by creating an artificial music market in which thousands of participants listen to previously unknown songs and download the ones they like with knowledge of the download choices of previous participants. The results strongly support their hypothesis.

In this paper, we develop a theory of market demand that can explain why sales in cultural markets are both unpredictable and unequal. The model is a variant of the herding models introduced by Banerjee (1992) and Bikhchandani, Hirshleifer and Welch (1992) and subsequently generalized by Smith and Sorensen (2000). Consumers enter the market sequentially to purchase a product. Their preferences for the product consists of a common component and an idiosyncratic component. The common component is a measure of product quality; the idiosyncratic component represents the heterogeneity in consumer preferences, which is important in cultural markets since even hit products have relatively small market shares. Consumers do not know the quality or the idiosyncratic value of the product. Prior to making their purchasing decision, consumers learn about their preferences by observing the purchasing decisions of prior consumers and by engaging in search. Search consists of acquiring a costly, private signal about preferences. For example, for music, it would involve listening to the songs either on the radio or at the listening post in the store; for books, it would involve reading reviews and checking out the book at the bookstore. Search is assumed to be valuable even when consumers know product quality. Indeed, the idiosyncratic component to preferences is the primary reason why consumers are likely to search before purchasing. However, search is costly, so if consumers know that product quality is low, they would prefer not to search and therefore not to buy.

The option to search leads to different dynamics and outcomes than in the standard herding models. A herd can form on the “search” action only if quality is high. In this case, the market learns the true quality of the product and its market share converges to the correct share. A herd can also form on the “no search” action and is certain to do so if quality is low. In this case, the market does not learn the true quality and market shares converge to 0. Thus, social influence can lead a population of consumers not to buy

---

2 Smith and Sorensen (2000) admit heterogenous preferences but assume that consumers know the idiosyncratic value of the product.
high quality products but search ensures that they realize their correct market shares some of the time and prevents the population (eventually) from buying low quality products. The model yields testable predictions regarding the impact of product quality, search costs and price on the number of consumers who search and on the probability of a “bad” herd occurring on high quality products.

We examine several of the model predictions using data from the SDW experiments. The main challenge in applying the herding model to field data sets is measuring intrinsic product quality. However, SDW conduct one experiment in which consumers have to independently decide which songs to listen and download without knowledge of the downloading decisions of other consumers. The probability of downloading a song conditional on listening to it is a measure of the song’s intrinsic quality. The distinguishing feature of the herding models is that social influence works through the learning process: consumers tend to buy what others buy because they search the products that others buy. By contrast, in social effects models (see Brock and Durlauf (1999), social influence works through preferences: consumers want to buy products that others buy. In the SDW social influence experiments, the herding model predicts that consumers are more likely to listen to songs with higher download rates, but the download information will not affect the likelihood that they will download the song conditional on listening to it. The social effects model predicts that the download information will affect the likelihood that consumer will download a song conditional on listening to it. Conditional on finding support for the herding model, we then test the model’s predictions regarding the impact of song quality on the distribution of the number of consumers who listen to it.

The paper is organized as follows. In Section II, we develop a herding model in which consumers base their decision to search on publicly available information. In Section III, we extend the model by endowing consumers with a private signal about their preferences prior to making their search decision. In Section IV we discuss the empirical implications of the model using data from the SDW experiments. Section V concludes.

2 Herding Model with Search

In this section we present a sequential choice model in which heterogenous consumers arrive randomly and have the option of searching before deciding whether or not to purchase a product of unknown quality. Search involves acquiring a costly, private signal about
preferences for the product. Each consumer’s purchasing decision is observable to later consumers but her search decision is not observable.

Let $X$ denote the unknown quality of the product. There are two quality levels: $X = H$ and $X = L$, where $H > L$. We will refer to $H$ as the high quality state and $L$ as the low quality state. We normalize $L = 0$. There is a common prior belief $\mu$ that assigns a probability to each state. An infinite sequence of consumers indexed by $t = 1, \ldots, T$ enter in exogenous order. Each consumer makes an irreversible decision on whether or not to purchase the product. Consumer $t$’s utility for the product is given by

$$V_t = X + U_t$$

where $X$ denotes the mean utility of the product and $U_t$ is the idiosyncratic component which is identically and independently distributed across consumers with zero mean. Consumer $t$ does not know $U_t$. The price of the product is $p$. Consumers’ utility is quasilinear in wealth, so consumer $t$’s net payoff from purchasing the product is $V_t - p$.

Consumers have two available actions. Buying the product involves risk since the ex post payoff may be negative. She can reduce the likelihood of this event by choosing to Search (S) before making her purchasing decision. Search involves paying a cost $c$ to obtain a private, informative signal about $V_t$, and then purchasing if the expectation of $V_t$ conditional on the signal exceeds $p$ and not purchasing otherwise. For notational simplicity, it will be convenient to assume that the signal is perfectly informative and reveals $V_t$ precisely.\(^3\) Note that search remains a valuable option for consumer $t$ even if she has learned $X$. The other action that she can choose is to Not Search and Not Buy (N). Let $a_t \in \{N, S\}$ denote the action chosen by consumer $t$.

We impose several restrictions on the heterogeneity in preferences and the cost of search. Define the expected value of search in state $X$ as

$$w(p, X) = \int_{X}^{\infty} (X - p + u)dF_U(u)$$

where $F_U$ is the distribution of the idiosyncratic preference shock.

\(A1:\) (a) $w(p, H) - c > 0$ and $w(p, H) - c > H - p$; (b) $w(p, 0) - c < 0$.

Condition (a) of Assumption A1 states that, conditional on $H$, the payoff to $S$ is positive and exceeds the payoff from buying without search.\(^4\) It implies that consumers never purchase

\(^3\)The important restriction is that the signal is informative about $X$ and not just $U$.

\(^4\)We rule out this action for now but intend relax this restriction when we discuss extensions of the model.
without search even when they know that the state is $H$. The second inequality states that the consumer’s payoff to $S$ is negative if she knows that the state is $L$. It implies that the consumer’s optimal action in state $L$ is $N$.

Consumer $t$’s action generates a purchasing outcome $b_t \in \{0, 1\}$. Here $b_t = 0$ is the outcome in which consumer $t$ does not purchase the good and $b_t = 1$ is the outcome in which consumer $t$ purchases the product. Outcome 0 occurs if consumer $t$ chooses $N$ or if she chooses $S$ and obtains a realization of $V_t$ such that her net payoff from purchase is negative. Outcome 1 arises if consumer $t$ chooses $S$ and obtains a realization of $V_t$ such that her net payoff from purchase is positive.

Before taking her action, consumer $t$ observes the purchasing decisions of consumers 1 through $t - 1$. She does not observe which actions they chose or, if they searched, the signals they obtained or, if they purchased, the payoffs they realized. The space of possible $t$-period purchase histories is given by $\Omega_t = \{0, 1\}^{t-1}$ and a particular history is denoted by $\omega_t$. The initial history is defined as $\omega_1 = \emptyset$. The assumption that consumers observe the entire ordered action history $\omega$ is obviously quite strong. In most markets, consumers are likely to know only the number of past purchases. However, as we shall see later, the results do not change if we assume that consumer $t$ knows only her place in the sequence and the number of outcome 1’s, $\#_{t-1}$ so that $\omega_n = (n, \#_{n-1}) \in \{t\} \times \{0, \ldots, t - 1\}$. The important issue is whether consumers observe some statistics of the purchasing history that are informative about $X$. Note that the $t$-period history is not informative about $U_t$.

We now define the expected payoffs to consumer $t$ for each action. Given any history $\omega_t$, consumer $t$ updates her beliefs about $X$ using Bayes rule. Let $\mu_t(\omega_t)$ represent her posterior belief that the state is $H$ conditional on history $\omega_t$. Then her expected payoff to $S$ is

$$\pi_t(S) = \mu_t w(p, H) + (1 - \mu_t) w(p, 0).$$

(1)

Recall that $L$ is normalized to zero. If the consumer chooses $N$, then her payoff is zero.

We look for a Bayesian equilibrium where everyone computes posterior beliefs using Bayes rule, knows the decision rules of all consumers and knows the probability laws determining outcomes under those rules. The posterior belief of consumer $t$ is also the public belief in period $t$. Since consumer $t$’s optimal decision rule maps posterior beliefs into an action, this means that later consumers can infer her action.

Because of the distinction between actions and outcomes, we have to careful in defining a herd. We say that a herd on action $a$ occurs at time $n$ if each consumer $t \geq n$ chooses
action a. Note that while a herd on N implies that all future outcomes are the same (all 1’s and 0’s, respectively), a herd on S does not. The outcome for a consumer who chooses S depends not only on X (which is common across consumers) but also on the realization of the idiosyncratic component U. In fact, a herd on S precludes the event that all future outcomes are the same (almost surely) - if the outcome does not vary with the realization of U, then it is not worthwhile paying c to search.

A related concept is outcome convergence. Let \( \lambda_t \in [0, 1] \) be the fraction of the first \( t - 1 \) consumers whose outcome was 1 (purchase). Outcome convergence is the event that \( \lambda_t \) converges to some limit \( \lambda \in [0, 1] \). A herd implies outcome convergence. A herd on N leads to \( \lambda = 0 \); and a herd on S leads to \( \lambda = 1 - F_U(p - X) \).

How does our model differ from the standard herding model? In the standard herding model, each consumer obtains a free private signal about X and then chooses to either buy or not. In our model, the private signal is costly and each consumer has to choose whether or not to acquire it before making her purchasing decision. As we shall see, the endogeneity of the private signal have important implications for outcomes and for the learning dynamics.

### 2.1 Outcomes and Learning Dynamics

We begin by defining a threshold belief. Let \( \hat{\mu} \) represent the posterior belief at which a consumer is indifferent between S and N. From (1),

\[
\hat{\mu} = \frac{c - w(p, 0)}{w(p, H) - w(p, 0)}.
\]

Assumption A1 implies that \( \hat{\mu} \in (0, 1) \).

Suppose \( \mu_t > \hat{\mu} \). Then consumer \( t \) is certain to search and consumer \( t + 1 \)'s belief that the state is H is

\[
\mu_{t+1}(b_t) = \begin{cases} 
\mu_t(1 - F_U(p - H)) & \text{if } b_t = 1 \\
\frac{\mu_t(1 - F_U(p - H)) + (1 - \mu_t)(1 - F_U(p - H))}{\mu_t F_U(p - H) + (1 - \mu_t)F_U(p)} & \text{if } b_t = 0.
\end{cases}
\]

If \( \mu_t < \hat{\mu} \), then consumer \( t \) chooses N and \( \mu_{t+1} = \mu_t \).

In studying the dynamics of beliefs and actions, we follow Smith and Sorensen (2000) and work with the public likelihood ratio that the state is L versus H rather than public beliefs. Define

\[
l_t = \frac{1 - \mu_t}{\mu_t}
\]
and the likelihood analogue of the threshold belief, \( \hat{l} \). The dynamics of the system are then determined by two first-order difference equations, a transition matrix, and an initial condition. Let \( l_0 \) denote the prior likelihood ratio. Given any \( l_t < \hat{l} \),

\[
\begin{align*}
    l_{t+1}(b_t) = \begin{cases} 
    1 - F_U(p) & \text{if } b_t = 1 \\
    1 - F_U(p - H) & \text{if } b_t = 0 \\
    F_U(p) & \text{if } b_t = 1 \\
    F_U(p - H) & \text{if } b_t = 0.
    \end{cases}
\end{align*}
\]  

(3)

Otherwise, \( l_{t+1} = l_t \). The transition probabilities conditional on \( H \) are \( 1 - F_U(p - H) \) and \( F_U(p - H) \) respectively; the transition probabilities conditional on \( L \) are \( 1 - F_U(p) \) and \( F_U(p) \) respectively.

Smith and Sorensen (2000) show the following:

Lemma 1. Conditional on state \( H \), the public likelihood ratio is a martingale. It converges to a random variable with support in \( [0, \infty) \) so fully wrong learning has probability zero. Conditional on state \( L \), the inverse public likelihood ratio is a martingale. It converges to a random variable with support in \( [0, \infty) \).

The martingale property rules out convergence to nonstationary limit beliefs such as cycles or to incorrect point beliefs. Figure 1 presents the phase diagram for the two difference equations. The fixed point for each equation is zero. It is stable for the “buy” difference equation: a sequence of “buy” outcomes causes the likelihood ratio to decrease in ever smaller increments towards 0. It is unstable for the “no buy” equation: a sequence of “no buy” outcomes causes the likelihood ratio to increase in every larger increments. Thus, conditional on \( H \), there is a positive probability that learning is complete, that is, beliefs converge asymptotically to the truth: \( \lim_{T \to \infty} l_T \to 0 \). When this is the case, a herd on \( S \) almost surely forms in finite time.

When public beliefs exceed \( \hat{l} \), the optimal action switches from \( S \) to \( N \) and learning stops. Hence, public beliefs in the set \{ \( l \geq \hat{l} \) \} are stationary points but these points are not reached asymptotically. There is a positive probability that public beliefs will “jump” into this set after a finite number of periods and a herd on the action \( N \) will form. We will refer to \{ \( l \geq \hat{l} \) \} as the failure set.

We can adapt Smith and Sorensen’s (2000) arguments to derive the following results.

Proposition 2. Suppose \( 0 < l_0 < \hat{l} \). (a) Then outcome convergence occurs almost surely. In state \( H \), \( \lambda = 1 - F_U(p - H) \) with positive probability and \( \lambda = 0 \) with positive probability.
In state \( L \), \( \lambda = 0 \). (b) Beliefs converge to the truth in state \( H \) when \( \lambda = 1 - F_U(p - H) \); otherwise learning is incomplete. (c) Actions in state \( H \) converge almost surely to \( S \) when \( \lambda = 1 - F_U(p - H) \); otherwise they are equal to \( N \) after a finite number of periods.

The proposition demonstrates how search modifies the results of herding models. When consumers cannot search prior to purchasing, the market never learns the true quality of products since eventually everybody either buys or does not buy and both outcomes can occur with positive probability in each state. Thus, the market share of high quality products never converges to their correct market shares, and the market share of low quality products converges to the wrong market share with positive probability. By contrast, in markets where consumers always search before purchasing, the market can learn the true quality of high quality products and outcomes can converge to the correct market shares. They are certain to do so for low quality products and with positive probability for high quality products. High quality products can still fail in our model due to the herding effect: outcomes converge to 0 with positive probability. Thus, our model generates a stochastic mapping between quality and market outcomes that allows market shares to vary with quality but explains why high quality products can still fail.

2.2 Comparative Statics

We turn next the comparative statics of our model. Throughout this section we will assume that the number of consumers, \( T \), is large but finite. Let \( M \) denote the equilibrium number of consumers who search. Since everyone searches as long as public beliefs lies below the threshold, it is a random variable whose realization depends upon whether and when public beliefs jump into the failure set. Let \( F_{M|X} \) denote the distribution of \( M \) conditional on state \( X \). The expected revenues to a high quality product are given by

\[
R(H) = p(1 - F_U(p - H))\overline{M}(H)
\]

where \( \overline{M}(H) \) is the expected value of \( M \) conditional on state \( H \).

The following proposition establishes that expected number of consumers who search fall when search costs increase.

**Proposition 3** \( F_{M|X}(m; c) \leq F_{M|X}(m; c') \) for \( c > c' \), \( X = L, H \).

**Corollary 4** An increase in \( c \) reduces expected sales in both states.
We are able to obtain a comparative static result for costs in this model because the increase in costs has no effect on the consumers’ decision rules. Consumers are certain to search until public beliefs enter the failure set and the search costs are sunk when they make their purchasing decision. Consequently, the impact of search costs on $M$ is determined solely by its impact on $\tilde{l}$.

By contrast, the impact of an increase in quality or price on $M$ (and the probability of an incorrect herd) is ambiguous because these changes cause the purchasing rule to change. An increase in price reduces the probability of purchase in both states and therefore the value of search. Hence, an increase in price causes $\tilde{l}$ to fall. However, this does not imply that $M$ falls or that the probability of a herd on $N$ increases. The reduction in the purchase probability means that, given any sequence of realizations of idiosyncratic shocks, zero outcomes are more likely to occur. But the change in purchasing rule changes the informativeness of the signal generated by the consumer’s purchasing decision. More zero outcomes does not necessarily mean that public beliefs are more pessimistic because consumers will take the into account the change in the decision rule when they update their beliefs about the state. It is not difficult to construct paths in which the number of consumers who search increases with price or to construct examples in which probability of a herd on $N$ increases with price. The same issue arises with an increase in product quality. Pastine and Pastine (2006) obtained similar “perverse” results when they studied the effect of changing the accuracy of signals on the probability of incorrect herds.

3 Herding Model with Private Search

In this section, we extend our model to allow consumers to have free private signals prior to making their search decisions. We will assume that the signals provide information about $X$ but not $U$. The private signals imply that the consumers’ search decisions are also private information. This feature of the model has important implications for outcomes and learning dynamics.

Smith and Soresen (2000) have shown that there is no loss in generality in defining the private signal that a consumer receives as her private belief. Here we denote the signal by $\sigma$ and define it as the probability that the state is $H$. Conditional on the state, the signals are

\footnote{It is not yet clear whether we can generalize the model to allow the signals to be informative about $X$ and $U$.}
i.i.d. across consumers and drawn from a distribution $F_X$, $X = H, L$. We assume that $F_L$ and $F_H$ are continuous and differentiable with densities $f_L$ and $f_H$. Under the assumption that both states are equally likely, the unconditional distribution of $\sigma$ is $F = (F_L + F_H)/2$ with density $f$. Smith and Sorensen (2000) show that defining the private signal in this way implies that the joint distribution of $X$ and $\sigma$ possesses the monotone likelihood ratio property. As is well known, this property implies that the conditional distributions $F_L$ and $F_H$ as well as their hazard and reverse hazard rates are ordered. Private beliefs are bounded if the convex hull of the common support of $F_L$ and $F_H$ consists of an interval $[d, \bar{d}]$ where $d > 0$ and $\bar{d} < 1$.

We begin with the consumer’s decision problem. Given public likelihood ratio $l$ and private signal $\sigma$, a consumer’s posterior belief that the state is $H$ is

$$r(\sigma, l) = \frac{\sigma}{\sigma + (1 - \sigma)l}. \quad (4)$$

The posterior belief at which consumer is indifferent between $S$ and $N$ is given by

$$\hat{\sigma} = \frac{c - w(L)}{w(H) - w(L)}. \quad (5)$$

Using equations (4) and (5), we can define the private signal at which a consumer is indifferent between search and no search as

$$\hat{\sigma}(l) = \frac{(c - w(L))l}{w(H) - c + (c - w(L))l}. \quad (6)$$

Thus, given $l$, the consumer’s optimal action is to choose $S$ if $\sigma \geq \hat{\sigma}$ and to choose $N$ if $\sigma < \hat{\sigma}$. We will refer to $\hat{\sigma}$ as the search threshold.

Next we define the cascade regions. A cascade on action $a \in \{S, N\}$ occurs when a consumer chooses $a$ regardless of the realization of her private signal $\sigma$. Let $\underline{l}$ denote the largest value of the public likelihood ratio such that a consumer is certain to choose $S$. From equation (5), $\underline{l}$ satisfies $\hat{\sigma}(\underline{l}) = \underline{d}$. Solving this equation for $\underline{l}$ yields

$$\underline{l} = \frac{\underline{d}(w(H) - c)}{(1 - \underline{d})(c - w(L))}. \quad (7)$$

Let $\bar{l}$ denote the lowest value of the public likelihood ratio such that a consumer is certain to choose $N$. Once again, using equation (5), $\bar{l}$ satisfies $\hat{\sigma}(\bar{l}) = \bar{d}$. Solving this equation for $\bar{l}$ yields

$$\bar{l} = \frac{\bar{d}(w(H) - c)}{(1 - \bar{d})(c - w(L))}. \quad (8)$$
Thus, we can partition the values of the public likelihood ratio into three intervals. When \( l < l_t \), there is a cascade on \( S \); when \( l_t \leq l \leq \bar{l} \), the consumer searches with probability \( 1 - F(\tilde{\sigma}(l)) \) and does not search with probability \( F(\tilde{\sigma}(l)) \); and when \( l > \bar{l} \), there is a cascade on \( N \).

We now characterize the dynamics of the public likelihood ratio. Suppose \( \underline{l} < l_t < \bar{l} \). Then the probability that consumer \( t \) buys the product in state \( X \) is

\[
\Pr\{b_t = 1|X, l_t\} = (1 - F_X(\tilde{\sigma}(l_t)))(1 - F_U(p - X)).
\]

It is the probability that consumer \( t \) searches in state \( X \) times the probability that she gets a realization of \( U \) that lies above the purchasing threshold \( p - X \). The probability that consumer \( t \) does not buy the product in state \( X \) is

\[
\Pr\{b_t = 0|X, l_t\} = F_X(\tilde{\sigma}(l_t)) + (1 - F_X(\tilde{\sigma}(l_t)))F_U(p - X).
\]

The first term is the probability of the event that consumer \( t \) searches in state \( X \) and gets a value of \( U \) that lies below the purchase threshold; the second term is the probability that consumer \( t \) does not search in state \( X \). Using Bayes’ rule, the public likelihood ratio in period \( t + 1 \) is given by

\[
l_{t+1}(b_t) = \begin{cases} 
    \left[ \frac{(1 - F_L(\tilde{\sigma}(l_t)))(1 - F_U(p))}{(1 - F_H(\tilde{\sigma}(l_t)))(1 - F_U(p - H))} \right] l_t & \text{if } b_t = 1 \\
    \left[ \frac{F_L(\tilde{\sigma}(l_t)) + (1 - F_L(\tilde{\sigma}(l_t)))F_U(p)}{F_H(\tilde{\sigma}(l_t)) + (1 - F_H(\tilde{\sigma}(l_t)))F_U(p - H)} \right] l_t & \text{if } b_t = 0
\end{cases}
\]

When \( l_t < \underline{l} \), \( F_L(\tilde{\sigma}) = F_H(\tilde{\sigma}) = 0 \), and the dynamic system reduces to the one specified in the previous section. When \( l_t > \bar{l} \), \( F_L(\tilde{\sigma}) = F_H(\tilde{\sigma}) = 1 \) and \( l_{t+1} = l_t \).

Following Smith and Sorenson (2000), it is straightforward to show that, conditional on \( H \), the public likelihood process is a martingale. It converges to a random variable with support in \([0, \infty)\) so fully wrong learning has probability zero. A cascade on \( N \) generates a herd on \( N \). However, a cascade on \( S \) does not generate a herd on \( S \) because of the distinction between actions and outcomes. If enough consumers choose not to buy following search, the public likelihood ratio will increase and enter the region where the optimal action depends upon the private signal.

For notational convenience, define

\[
\psi_0(l_t) = \frac{F_L(\tilde{\sigma}(l_t)) + (1 - F_L(\tilde{\sigma}(l_t)))F_U(p)}{F_H(\tilde{\sigma}(l_t)) + (1 - F_H(\tilde{\sigma}(l_t)))F_U(p - H)}
\]
\[
\psi_1(l_t) = \frac{(1 - F_L(\hat{\sigma}(l_t)))(1 - F_U(p))}{(1 - F_H(\hat{\sigma}(l_t)))(1 - F_U(p - H))}
\]

The following lemma follows from our assumptions on the conditional distributions \(F_H\) and \(F_L\).

**Lemma 5** (i) \(\psi_0\) is continuous, \(\psi_0(l) > 1\) on \([0, \overline{l})\), and \(\psi_0(l) = 1\) on \([\overline{l}, \infty)\). (ii) \(\psi_1(l) < 1\) and continuous on \([0, \overline{l})\) with \(\lim_{l_t \uparrow \overline{l}} \psi(l) > 0\).

A particular useful property for comparative static purposes is when the posterior likelihood ratio following each action is monotone increasing in the prior likelihood ratio. Smith and Sorensen (2005) find that this is the case in their model if the density of the private belief log-likelihood ratio is log-concave. Log-concavity is sufficient to establish that the “buy” difference equation (i.e., \(l_{t+1} = \psi_1(l_t)l_t\)) is monotone increasing but it is not sufficient for the “not buy” difference equation (i.e., \(l_{t+1} = \psi_0(l_t)l_t\)). In the former case, there is no distinction between outcome and action: when consumer \(t + 1\) observes the outcome “buy”, she knows that consumer \(t\) chose action \(S\). This is not true when consumer \(t + 1\) observes the outcome “not buy”: it is consistent with consumer \(t\) choosing either action \(N\) or \(S\). As a result, we need to impose an additional restriction on the distribution of \(U\) [and \(\sigma\)] to ensure monotonicity. For any private belief \(\sigma\), let \(\gamma\) denote the natural log of the corresponding likelihood ratio: \(\gamma = \ln \left(\frac{1 - \sigma}{\sigma}\right)\). Denote the unconditional distribution of \(\gamma\) by \(F'\), with density \(f'\).

**A2:** (i) The density of the log likelihood ratio, \(f'\), is log-concave and (ii) \(E[\gamma|\gamma \geq g] - E[\gamma|\gamma < g] > \ln F_U(p) - \ln F_U(p - H)\) for all \(g\) in the support of \(\gamma\).

Part (ii) of Assumption A2 ensures that at any public belief, a decision not to search (after receive the private belief) is more suggestive of state \(L\) than deciding to search but then not buying.

**Lemma 6** Assumption A2 implies that \(\psi_0(l_t)l_t\) and \(\psi_1(l_t)l_t\) are increasing in \(l_t\).

Figure 2 illustrates a dynamic system that satisfies Assumption A2. The cascade set for \(S\) is the interval \([0, \overline{l}]\) and the cascade set for \(N\) is the interval \([\overline{l}, \infty)\). The “no buy” difference equation intersects the diagonal at 0 and at \(\overline{l}\) and lies everywhere above the diagonal in between these two values. It is linear on the cascade set for \(S\) and strictly increasing on the interval \((\overline{l}, \overline{l})\). The “buy” difference equation intersects the diagonal at 0 and lies below the
diagonal for positive values of \( l_t \). It is linear on the cascade set for \( S \) and increasing on the interval \((l_t, \hat{t})\). Active dynamics occur when the prior is such that \( l_0 \in (0, \hat{t}) \).

**Proposition 7** Suppose \( 0 < l_0 < \hat{t} \). (a) Outcome convergence occurs almost surely. In state \( H \), \( \lambda = 1 - F_U(p - H) \) with positive probability and \( \lambda = 0 \) with positive probability; in state \( L \), \( \lambda = 0 \). (b) In state \( H \), beliefs converge to the truth when \( \lambda = 1 - F_U(p - H) \); otherwise, the limit belief converges to \( \bar{t} \) and learning is incomplete. Beliefs can never enter the cascade set for \( N \) from outside. (c) Actions in state \( H \) converge almost surely to \( S \) when \( \lambda = 1 - F_U(p - H) \); otherwise they converge almost surely to \( N \).

The key feature of the dynamics is that public beliefs can never enter the cascade set for \( N \) from outside. For any \( l_0 < \hat{t} \), the dynamics are forever trapped between the stationary points 0 and \( \hat{t} \). A herd always eventually starts. In state \( H \), the herd can form with positive probability on either \( N \) or on \( S \). If it forms on \( S \), then the market learns the true state from the frequency of purchases. In state \( L \), the herd can only form on \( N \). Intuitively, if a herd formed on \( S \), then the purchase frequency reveals the state is \( L \), which contradicts the assumption that it is not optimal to search in state \( L \).

Hendricks and Sorensen (2007) show the release of a hit new album by an artist can substantially raise sales of catalog albums. We can interpret the release of a hit new album as a strongly positive, public signal about the catalog album. What is the effect of such a signal after public beliefs about the catalog albums have converged? Clearly, there is an impact only if beliefs converged to the cascade set on \( N \). In that case, if the signal is strong enough, then public beliefs will move away from \( \bar{t} \) and re-converge, possibly to a herd on \( S \).

**Proposition 8** The introduction of a sufficiently positive, public signal after beliefs have converged (a) raises expected sales if the initial herd is on \( N \); (b) has no effect on sales if the initial herd is on \( S \).

The probability that product sales increase in response to the positive public signal depends upon the quality of the product. The higher the quality of the product, the more likely sales will increase substantially. However, there is a selection effect. A herd on \( N \) is more likely to occur for lower quality products. It is this selection effect that adds skew to the distribution of outcomes.
3.1 Comparative Statics

We examine the impact of product quality, search costs and price on the probability of a “bad” herd occurring in state $H$.

Our earlier result that the distribution of $M$ is stochastically ordered in search costs does not extend to situations in which consumers have free private signals. The problem is that an increase in search costs lowers search thresholds, thereby increasing the likelihood that a consumer does not search. The change in search rule changes the informativeness of the signal generated by the consumer’s purchasing decision, which consumers will take into account when they update their beliefs about the state. As a result, given any path of signal realizations, it is not possible to determine whether the number of consumers who search is larger or smaller at a higher cost level: for some paths, the number will increase and for other paths, it will decrease.

However, we are able to determine the impact of search costs and other model parameters on long-run sales. Since the likelihood ratio process $(l_t)$ is a bounded martingale conditional on state $H$, the $E[l_\infty] = l_0$. This property of martingales was not useful in the previous model because public beliefs could jump into the failure set from outside of the set. Hence, the support of $l_\infty$ in the model without the free private signals consists of 0 and a subset of the failure set. When consumers have free private signals whose distribution satisfy A2, public beliefs cannot enter the cascade set for $N$ from outside. This implies that $\text{supp}l_\infty = \{0, \overline{l}\}$, from which it follows that

$$\Pr\{l_\infty = \overline{l}\} = \frac{l_0}{\overline{l}}.$$  

Thus, the sign of the changes in product quality, search costs, and price on the probability of a “bad” herd (i.e., long-run sales are 0) are determined by the sign of their impact on the value of $\overline{l}$. Differentiating $\overline{l}$ with respect to $H$, $c$ and $p$ yields the following results:

**Proposition 9** Suppose the state is $H$. Then the probability of a limit cascade on $N$ is (a) strictly decreasing in $H$; (b) increasing and convex in $c$; (c) increasing in $p$.

The proposition yields several predictions. An increase in search costs increases the likelihood that long-run sales of a high quality product is zero and hence reduces its expected long-run sales. The impact of the increase is larger at higher cost levels. An increase in price also increases the probability of zero long-run sales but the sign of the second derivative of $\overline{l}$ with respect to $p$ depends upon the distribution of the private signal. This issue is
important because, if $f$ is convex, then the model could explain why search goods like music albums and books sell for the same prices even when the distributors know that the quality of their products vary.

What is the effect of an increase in the precision of the free signal on the probability of a limit cascade on $N$? The answer is somewhat surprising: none, if the support of the distribution does not change. The only properties of $F$ that matter are $\overline{a}$, the upper bound of its support, and continuity. This result has important implications for the kind of information that the market should reveal about consumers’ purchasing decisions.

A number of papers (e.g., [7]) have argued that the decline in search costs due to the Internet has disproportionately increased sales of niche products and reduced the concentration in sales. The next proposition provides support for this claim.

**Proposition 10** Suppose the state is $H$. Then the impact of an increase in $c$ (or an increase in $p$) on the probability of a limit cascade on $N$ is smaller (in absolute value) for higher quality products.

The results follows from differentiating $f$ with respect to $H$ and $c$ (and $H$ and $p$). The Proposition implies that a decrease in search costs has a larger impact on long-run sales of niche products (i.e., medium quality products) than on high quality products.

We need to be cautious in interpreting the above result as implying a reduction in the concentration in sales across products. The distribution of sales across products is the equilibrium to a model in which consumers face a choice set with multiple products of unknown quality. Equilibrium outcomes are likely to depend upon the type of search rule used by consumers in this environment and the number of products they are willing to purchase. If search costs and indirect utility function are additive, then the consumer’s decision problem reduces to one studied in this paper: she searches each product of unknown quality if the expected utility from doing so exceeds the utility of the outside good, which is a constant. But, in practise, search costs are unlikely to be additive and the indirect utility function is surely not additive due to budget constraints, preferences. The implications of non-linearities in search costs and utility on the distribution of sales is an interesting issue that deserves further study.
3.2 Extensions

Discuss extensions such as (i) introducing buy without search option; (ii) signals are informative about $X$ and $U$; (iii) search costs are heterogenous and private.

4 Application

Discuss the empirical content of the herding model; describe the Salganik experiment; present the data and hypotheses to be tested; report results.
References


