Econ Theory (2008) 37:203–241
DOI 10.1007/s00199-007-0313-9

RESEARCH ARTICLE

Monopoly pricing in the binary herding model

Subir Bose · Gerhard Orosel · Marco Ottaviani · Lise Vesterlund

Received: 28 March 2007 / Accepted: 30 October 2007 / Published online: 23 November 2007

Abstract How should a monopolist price when selling to buyers who learn from each other's decisions? Focusing on the case in which the common value of the good is

This paper generalizes the model and the results contained in "Monopoly pricing with social learning" by Ottaviani (1996) and "Optimal pricing and endogenous herding" by Bose, Orosel, and Vesterlund (2001). Ottaviani (1996) first formulated the problem and derived implications for learning and welfare. Independently, Bose, Orosel, and Vesterlund (2001) formulated a similar model but with a different focus on the dependence of the solution on the model's parameters. Bose, Orosel, and Vesterlund's team and Ottaviani then joined efforts to partially characterize the solution for the general case with a finite number of signal realizations (Bose et al. 2006), and to provide a full characterization of the equilibrium for the case with symmetric binary signals, which is done in the present paper. Vesterlund thanks the NSF for financial support.

S. Bose
Department of Economics, University of Illinois at Urbana-Champaign, Champaign, IL 61820, USA
e-mail: bose@uiuc.edu
URL: http://netfiles.uiuc.edu/bose/www

G. Orosel
Department of Economics, University of Vienna, Hohenstaufengasse 9, A-1010 Vienna, Austria
e-mail: gerhard.orosel@univie.ac.at
URL: http://homepage.univie.ac.at/gerhard.orosel

M. Ottaviani (✉)
Economics Subject Area, London Business School, Sussex Place, London, NW1 4SA, England
e-mail: mottaviani@london.edu
URL: http://faculty.london.edu/mottaviani

L. Vesterlund
Department of Economics, University of Pittsburgh, 4916 WW Posvar Hall,
Pittsburgh, PA 15260, USA
e-mail: vester@pitt.edu
URL: http://www.pitt.edu/∼vester

binary and each buyer receives a binary private signal about that value, we completely answer this question for all values of the production cost, the precision of the buyers' signals, and the seller's discount factor. Unexpectedly, we find that there is a region of parameters for which learning stops at intermediate and at extreme beliefs, but not at beliefs that lie between those intermediate and extreme beliefs.

Keywords Monopoly · Public information · Social learning · Herd behavior · Informational cascade · Binary signal

JEL Classification D83 · L12 · L15

1 Introduction

This paper analyzes a simple model of dynamic monopoly pricing with buyers who learn from each other's decisions. For the case with *fixed prices*, Banerjee (1992) and Bikhchandani et al. (1992) have shown that herd behavior results when buyers have private signals of bounded informativeness. Eventually, the public information inferred from prior decisions swamps the private information of an individual buyer. An *informational cascade* then arises from the first time a buyer's private information has no effect on his decision. From that point onward, buyers imitate the prevailing behavior of the initial buyers, and, so, their private information is lost. In this paper, we embed Bikhchandani et al.'s (1992) social learning model in a market framework in which a monopolist adjusts prices dynamically in response to past purchase decisions.

In our setting with variable prices, the current price not only reflects the product's quality (as currently perceived), but also affects the social learning process and, thus, the perception of quality by future buyers. This intertemporal linkage is relevant for understanding pricing of new products, as well as career paths and wages demanded by workers to influence their future employers, who observe the employment history. Will the seller (or, the worker) be more expensive (or, choosier) in the early stages after product introduction (or, early in their career) in order to build a strong track record (or, a strong vita), even at the risk of not selling (or, remaining unemployed) for a long time? When will the seller want to sell out by inducing an informational cascade in which all potential buyers purchase regardless of their private information? Could it be that learning ceases for intermediate beliefs? How do monopoly prices evolve over time?

This paper addresses these questions by providing a complete characterization of the solution of a simplified version of the model formulated in Bose et al. (2006), henceforth called BOOV. The general model features an exogenous sequence of potential buyers, one in each period. Each buyer has a unit demand for a product with a common value that is either high or low. The seller and the buyers do not know the true common value, but each buyer observes a partially informative private signal about the value. In each period, the prices posted in the past and purchase decisions made by past buyers are publicly observed by the buyers as well as the seller. Having observed this information the seller posts a price, and the current buyer decides whether to accept or reject the offer. The game then proceeds to the next period.

For the general case in which each buyer has access to a signal with many possible realizations, BOOV shows that, in the long run, the monopolist generally (i.e., with the exception of one specific parameter configuration) settles on a fixed price and all subsequent buyers either purchase the good (in a *purchase cascade*) or do not purchase it (in an *exit cascade*). Thus, typically, all buyers herd on the same decision eventually and no information is revealed thereafter. However, the general model is not sufficiently tractable to address the four characterization questions that are the focus of this paper: (1) whether the monopolist finds higher prices more attractive in the initial periods, (2) when it is optimal for the seller to induce informational cascades, (3) whether an informational cascade can occur at intermediate beliefs, and (4) how prices evolve over time. These questions have no obvious answer and without a rigorous analysis it is not clear what to expect.

In this paper we simplify the model by restricting attention to the case in which buyers have access to symmetric *binary* signals about the product's unknown value. By assuming that the private signal is either "high" or "low", we are able to determine the seller's posted price for every possible public belief and thus obtain a complete characterization of the dynamic pricing policy. In particular, we derive sharp and somewhat unexpected predictions regarding the *short-run* price dynamics and the *long-run* occurrence of informational cascades.

The major advantage of the binary signal structure is that, at any public belief, the seller who wishes not to exit the market needs to consider only two possible selling prices: a low *pooling* price at which the respective buyer purchases the good irrespective of the private signal, and a high *separating* price at which only a buyer with a high signal purchases the good. Whenever the seller charges the separating price, the current buyer's private signal is perfectly revealed by the purchase decision. On the other hand, when the seller charges the pooling price, both buyer types purchase the good, so that no information is revealed. Therefore, the price determines whether the purchase decision reveals either all or none of the current buyer's private information. This dichotomy of all or nothing makes the model analytically tractable.

Since the seller benefits from revealing public information (see Ottaviani and Prat 2001 and BOOV), the seller's expected future profits are maximized at the separating price. Thus, the expected immediate profits from the pooling and the separating prices, respectively, are crucial for determining which price is charged. Specifically, the separating price is uniquely optimal when it also maximizes expected immediate profits, and the pooling price can only (but need not) be optimal when its expected immediate profits exceeds those of the separating price.

The properties of the equilibrium depend on the interaction between the three main parameters of the model: the precision of the signal, the seller's discount rate, and the unit cost of production. We distinguish two cases. First, when the signal is sufficiently precise, it is uniquely optimal for the seller to charge the separating price if and only if the public belief belongs to a single connected (non-empty) interval. What defines a sufficiently precise signal is independent of the seller's discount rate but depends on the cost of production. Whenever the seller charges the separating price, the respective buyer's action reveals the signal, allowing the seller and future buyers to update their beliefs accordingly. The seller then continues to charge the separating price as long as the public belief about the good's value stays within this interval, and demands the pooling price or exits the market when the public belief hits the boundaries of this interval.

Second, when instead the signal is sufficiently noisy, the properties of the equilibrium depend crucially on how the cost of production compares to the low value of the good. When the cost of production is below the low value of the good, the separating price is optimal in a connected (though possibly empty) set of beliefs, whereas the pooling price is optimal for beliefs that are either more optimistic or more pessimistic. The solution is more complicated when instead the production cost exceeds the low value of the good. In this situation, it is possible that the separating and the pooling price are each uniquely optimal within two disconnected non-empty intervals of public beliefs. As the public belief increases, exit is initially optimal for the seller, then the separating price is optimal, then the pooling price is optimal, then the separating price is optimal once more, and finally the pooling price becomes again optimal for sufficiently optimistic public beliefs.

From a technical point of view, the characterization obtained in this paper relies on the combined application of dynamic programming techniques and the Perron–Frobenius theory of nonnegative matrices. This combination is the keystone of the proof of Lemma 5, which constitutes a central building block of our analysis.

The rest of the paper is organized as follows. After discussing the related literature in Sect. 2, Sect. 3 introduces the model, Sect. 4 explains the main trade off, Sect. 5 presents the optimal pricing policy, Sect. 6 derives the stochastic properties of the price sequence, and Sect. 7 concludes. Appendices A.1 and A.2 contain respectively the proofs of the propaedeutic results and of the propositions.

2 Related work

The first paper to analyze *monopoly pricing* in the presence of herding is Welch (1992). He considers a similar setup to ours where buyers have binary signals, but investigates the problem of *static* monopoly pricing in which the seller cannot adjust the price depending on the purchase history. He shows that in this case it is optimal for the seller to charge a low price and immediately trigger herding. Note that expected monopoly profits are clearly higher when prices are allowed to depend on the history of previous purchases, as in our model.

Avery and Zemsky (1998) are the first to study the effect of the *dynamic* adjustment of prices on the occurrence of informational cascades. They focus on a competitive financial market in which informed agents can choose to either buy or sell. In their as well as in our setting, prices adjust to reflect the information revealed from past trades. But in addition our monopolist sets prices so as to control the learning process.

Caminal and Vives (1996) consider a two-period model with a continuum of buyers privately informed about the quality of two competing products. Since in their model second-period buyers observe first-period quantities but not first-period prices, the sellers have an incentive to set low first-period prices to boost sales in an attempt to convince buyers that their quality is high. By assuming instead that past prices are observed, in our model we find that the seller initially posts high prices in order to induce social learning.

As in Bergemann and Välimäki (1996), in our model information about quality becomes publicly available over time. In their setting, the product price affects the information that is publicly revealed only through the identity of the product purchased (and therefore experienced) by the buyer. When instead buyers decide also on the basis of pre-existing *private* information, as in our model, the price affects the amount of public information depending on the induced probabilities (conditional on the product's true quality) that the buyer purchases the product.

Building on the first incarnation of this paper (Ottaviani 1996), Moscarini and Ottaviani (1997) analyze a two-period version of the model presented here. Chamley (2004, Sect. 4.5.2) briefly discusses a version of this model in which the seller's cost exceeds the low value of the good. As explained below, our complete characterization of the seller's optimal strategy reveals a number of somewhat unexpected properties of the solution.[1]

More broadly, we share with Sgroi (2002) the interest in studying the effect of policies aimed at influencing the social learning process. In a model with fixed prices, Sgroi considers the effect of allowing a group of buyers to make simultaneous decisions before the other buyers make sequential decisions. In his context, the trade off is between the cost of experimentation from less informed decisions made by the guinea pigs and the value of the public information their decisions reveal. Similarly, in our context the monopolist tends to subsidize the learning process by charging separating prices in the initial periods.[2]

In our model, equilibrium prices are high and decline on average, similar to what happens in the separating equilibrium of Bagwell and Riordan's (1991) signaling game. In their model, the high quality seller needs to charge a higher price when there is a smaller fraction of informed buyers; as buyers become better informed over time, the price charged by the high quality seller decreases. The two models provide different explanations for the same phenomenon.

Finally, in Bar-Isaac (2003) a seller privately informed about product quality supplies a sequence of (uninformed) buyers, whose noisy satisfaction with the product is revealed publicly after purchase. In his model prices are assumed to be fixed, so that signaling takes place through the seller's decision to remain in the market or exit. Our model eliminates instead the possibility of signaling by assuming that the seller has no private information.[3]

[1] Chamley presents diagrammatically a numerical solution for particular parameter values at page 79 and discusses the model at pages 81–83. We prove here that the properties of that specific illustration (seller's exit at sufficiently pessimistic public beliefs, purchase cascade at sufficiently optimistic public beliefs, and continuing learning at all intermediate public beliefs) do not hold in general.

[2] See also Gill and Sgroi (2005) for a model in which a monopolist asks a reviewer to evaluate the quality of the product before launching it. As in our model, buyers have symmetric binary signals and the monopolist controls the product's price. In their model however, as in Welch (1992), the monopolist cannot change the price over time depending on the history of past sales.

[3] If the seller had some imperfect private information about the good's value, for some parameter configurations the seller might be able to signal this information through prices, given that buyers are also privately informed (compare Judd and Riordan 1994). As we focus on the aggregation of the information possessed by buyers, we improve tractability by assuming that the seller has no private information.

3 Model

A risk-neutral monopolist (or *seller*) offers identical goods to a sequence of risk-neutral potential *buyers* with quasi-linear preferences and unit demand. In each period $t \in \{1, 2, ...\}$, a different buyer arrives to the market, indexed by the arrival time.

The payoff of buyer t is $(v - p_t)\, a_t$, where v is the value of the good, p_t is its price, action $a_t = 1$ indicates purchase of one unit of the good, and $a_t = 0$ no purchase. The good's common value is either low or high, $v \in \{L, H\}$, with $0 \leq L < H$. Without loss of generality we choose the monetary unit such that $H - L = 1$. The good's value is unknown to the seller and the buyers. The initial prior belief that the value is high, $v = H$, is commonly known to be equal to $\lambda_1 \in (0, 1)$.

Buyer t *privately* observes a random signal about the good's value v, denoted by $\tilde{s}_t \in \mathcal{S}$ with realization s_t (abbreviated as s if time does not matter), where $\mathcal{S} = \{l, h\}$. Conditional on the true value v, buyers' signals are independent and identically distributed for all buyers, and they are imperfectly informative. The signals are correct with probability $\alpha \in (1/2, 1)$ and incorrect with probability $1 - \alpha$. That is, $\Pr(s_t = h|H) = \Pr(s_t = l|L) = \alpha$ and $\Pr(s_t = h|L) = \Pr(s_t = l|H) = 1 - \alpha$. A high signal realization h (respectively a low signal l) then indicates that the high value H is more (respectively less) likely than according to the prior.

The seller has a constant marginal cost c per unit sold, with $0 \leq c < H$.[4] This cost is incurred only if the good is sold. In the analysis, we distinguish three cases, depending on whether the unit cost c is equal to, below, or above L. The seller's payoff is equal to the discounted sum of profits, $\sum_{t=1}^{\infty} \delta^{t-1} (p_t - c)\, a_t$, with discount factor $\delta \in [0, 1)$. For tractability, we further assume that the seller has no private information about the good's value.

The key feature of the model is that the buyers' actions are publicly observed. This allows future buyers as well as the seller to possibly learn the signals those buyers had. The sequence of events in each period t is as follows:

1. The seller and buyer t observe the purchase decisions taken by previous buyers, as well as the prices posted in the past.[5] This public history at time t is denoted by $h_t = (p_1, a_1, ..., p_{t-1}, a_{t-1})$, with $h_1 = \varnothing$. The set of all possible histories is denoted by $\mathcal{H}$.
2. The seller makes a take-it-or-leave-it price offer p_t for a unit of the good to buyer t. A pure strategy for the seller is a function $p : \mathcal{H} \to (L, L + 1)$ that maps every history h_t into a price p_t, $t \in \{1, 2, ...\}$.[6]

[4] See Neeman and Orosel (1999) and Taylor (1999) on dynamic bidding or pricing, respectively, for a *single* good.

[5] If instead buyers observed only purchase decisions but not prices, the seller would have an incentive to mislead buyers by making them believe that an observable sale has occurred at the high separating price, rather than at the actual low pooling price. If buyers understand this incentive, they cannot be misled in equilibrium. However, because the seller cannot commit not to lower the unobservable price if the high price is expected, the equilibrium in pure Markov strategies results in the seller triggering an informational cascade immediately. For details see Bose et al. (2001).

[6] Since it is common knowledge that the seller can always sell at some price above L and is unable to sell at a price at or above H, we restrict attention to $p_t \in (L, H) = (L, L + 1)$.

3. Buyer t observes signal $\tilde{s}_t$ about the good's value v and takes action a_t. Buyer t purchases the good if and only if its expected value $E(v|s_t, h_t)$ exceeds the price. For technical reasons, we make the (innocent) tie-breaking assumption that a buyer purchases the good when indifferent between purchasing and not.

This model is a dynamic game between a long-run player (the seller) and a sequence of short-run players (the buyers). Following any history $h_t = (p_1, a_1, ..., p_{t-1}, a_{t-1})$ of past prices and actions, past play induces a common prior belief for period t that is shared between the seller and the current buyer, $\lambda_t \equiv \Pr(v = H \mid h_t)$. Since buyers decide only once, it is immediate to characterize their behavior as a function of this *public belief*, the current price charged, and the realization of their private signal. In each period and for any public belief, the seller maximizes expected profits by anticipating how the buyers react to the prices posted. The public belief is then the state variable in the seller's dynamic optimization problem. The perfect Bayesian equilibrium (PBE) of this game coincides with its Markov perfect equilibrium and is derived directly from the seller's optimal strategy.

When focusing on a single period, we often drop the time subscript and treat the public belief λ as the key parameter for comparative statics. Since signals, if revealed, lead to discrete jumps in λ_t, any λ_t must be an element of a countable set which for any prior $\lambda_1 \in (0, 1)$ is defined as the set of beliefs that can be attained starting from λ_1,

$$\Lambda(\lambda_1) \equiv \left\{ \lambda \;\middle|\; \begin{array}{l} \text{there exists an integer } T \text{ and a sequence of signal realizations} \\ (s_1, ..., s_T) \in \mathcal{S}^T \text{ such that } \lambda = \Pr(H|\lambda_1; s_1, ..., s_T) \end{array} \right\} \subset (0, 1). \quad (1)$$

For any given $\lambda^j \in \Lambda(\lambda_1)$, we define for all $k \in \{1, 2, ...\}$, $\lambda^{j+k} \equiv \Pr\left(H|\lambda^j, k \text{ signals } s = h\right)$ and $\lambda^{j-k} \equiv \Pr\left(H|\lambda^j, k \text{ signals } s = l\right)$, where $s \in \{l, h\}$ denotes the signal realization. We denote the updated beliefs that the value is high conditional on a high and low signal, respectively, as $\lambda^+ \equiv \Pr(H|\lambda, s = h)$ and $\lambda^- \equiv \Pr(H|\lambda, s = l)$.

4 Main trade off

Buyer t's optimal strategy is simply to buy if and only if the posted price is (weakly) lower than the expected value conditional on the privately observed signal:

$$\begin{aligned} p_t \le E(v|\lambda_t, s_t) &= \Pr(H|\lambda_t, s_t) H + [1 - \Pr(H|\lambda_t, s_t)] L \\ = \Pr(H|\lambda_t, s_t)(H - L) + L &= \Pr(H \mid \lambda_t, s_t) + L. \end{aligned}$$

When wishing to sell with positive probability at any given public λ_t, the seller need only consider two possible prices, either the separating price $p^H(\lambda_t) \equiv \lambda_t^+ + L = \frac{\lambda_t \alpha}{\lambda_t \alpha + (1-\lambda_t)(1-\alpha)} + L$, which is the highest price at which type h buyer purchases the good (type l declines to buy at this price), or the pooling price $p^L(\lambda_t) \equiv \lambda_t^- + L = \frac{\lambda_t(1-\alpha)}{\lambda_t(1-\alpha) + (1-\lambda_t)\alpha} + L$, which is the highest price at which both type h and type l buyer purchase. Clearly, any price $p \in \left(p^L(\lambda_t), p^H(\lambda_t)\right)$ is suboptimal. The H and L labels derive from the fact that the separating price is higher than the pooling price for any given belief λ_t.[7]

In addition, the seller can always charge a price strictly higher than $p^H(\lambda_t)$ at which no buyer purchases. Although all prices strictly larger than $p^H(\lambda_t)$ are non-selling prices, the multiplicity is clearly inconsequential, since all these prices result in the same outcomes and profits. Hence, we treat all such prices as one *exit price,* which we denote by $p^E(\lambda_t)$. For the results in Sect. 6 we further assume that $p^E(\lambda) = p^H(\lambda) + \varepsilon$, with $\varepsilon > 0$ arbitrarily small.

Given the buyers' strategies, it is clear that whenever either $p^H(\lambda_t)$ or $p^L(\lambda_t)$ is uniquely optimal, λ_t determines the seller's optimal price p_t in period t. If $p^H(\lambda_t)$ and $p^L(\lambda_t)$ are both optimal for some $\lambda_t \in \Lambda(\lambda_1)$, the seller may condition the price p_t on aspects of the history h_t that are not reflected in λ_t. However, in all three cases the maximum of the seller's expected discounted profits from period t onwards is uniquely determined by λ_t. For any $t \in \{1, 2, ...\}$ and $\lambda_t \in \Lambda(\lambda_1)$, we denote this payoff by $V(\lambda_t)$. That is, $V : \Lambda(\lambda_1) \to \mathbb{R}$ is the seller's value function.

In choosing between the pooling and separating price the seller must take into account the (updated) probability that a buyer has received the high or the low signal. This probability is determined by λ_t and given by $\Pr(s_t = h|h_t) = \lambda_t \alpha + (1 - \lambda_t)(1 - \alpha) \equiv \varphi(\lambda_t)$ and $\Pr(s_t = l|h_t) = \lambda_t(1 - \alpha) + (1 - \lambda_t)\alpha = 1 - \varphi(\lambda_t)$, respectively.

Following the pooling price in period t, buyer t purchases regardless of the private signal, so that $\lambda_{t+1} = \lambda_t$. Thus, if $p^L(\lambda_t)$ is uniquely optimal at λ_t, it continues to be optimal forever after. In accordance with Bikhchandani et al. (1992), we then have an informational cascade.[8]

Definition An *informational cascade* occurs at time T, if all buyers $t \geq T$ make the same purchase decisions regardless of their signal realizations.

If $p^L(\lambda_t)$ is optimal at λ_t, then $V(\lambda_t) = p^L(\lambda_t) - c + \delta V(\lambda_{t+1}) = p^L(\lambda_t) - c + \delta V(\lambda_t)$, and thus

$$V(\lambda_t) = \frac{p^L(\lambda_t) - c}{1 - \delta}.$$

Whenever the pooling price $p^L(\lambda_t)$ is uniquely optimal for some λ_t, a *purchase cascade* is triggered or continued in period t. Similarly, whenever an exit price $p^E(\lambda_t)$ is uniquely optimal for some λ_t, we have an *exit cascade*.

Clearly, the stochastic process of the updated probabilities $\{\lambda_t\}_{t=1}^{\infty}$ is a martingale. If the seller charges the pooling price $p^L(\lambda_t)$ at some t or stays out of the market, no information is revealed so that $\lambda_{t+1} = \lambda_t$. If the seller demands the separating price

7 Because these prices are increasing functions of $\lambda_t \in (0, 1)$, the separating price for a high belief can be lower than the pooling price corresponding to a lower belief.

8 Since in this model there is a finite signal space as in Bikhchandani et al. (1992), an informational cascade is equivalent to herding. See Smith and Sørensen (2000).

$p^H(\lambda_t)$, buyer t 's action reveals the signal realization s_t so that the expected belief in period $t+1$ is

$$E(\lambda_{t+1}|\lambda_t) = \varphi(\lambda_t)\lambda_t^+ + [1-\varphi(\lambda_t)]\lambda_t^- = \lambda_t, \tag{2}$$

verifying the martingale property of beliefs.

In this setting, the current price charged in any given period serves two roles. First, the price determines the amount of rent that the seller can extract from the current buyer. Second, the price affects the amount of the current buyer's information that is transmitted to future buyers. Since expected future profits depend on this information, there is a trade-off between current and future rent extraction. To determine the seller's optimal price it is therefore useful to break down the seller's expected payoff into the expected *immediate* and *future profits*.

4.1 Expected future profits

The seller's expected *future profits* are the sum of the expected discounted profits from the next period onwards conditional on the present information λ_t.

If the seller charges the pooling price $p^L(\lambda_t)$ in period t and triggers a purchase cascade, then the future profits are $\delta\frac{p^L(\lambda_t)-c}{1-\delta}$. Denote $p^L(\lambda_t)-c$ by $R_L(\lambda_t)$ and note that $R_L(\lambda_t) = \lambda_t^- + L - c = \frac{\lambda_t(1-\alpha)}{\lambda_t(1-\alpha)+(1-\lambda_t)\alpha} + L - c$, which is a strictly convex function for $\lambda \in [0,1]$. As Lemma 1 shows, strict convexity of $R_L(\lambda)$ implies that the expected future profits from charging the separating price in period t and the optimal price thereafter always exceed the future profits from charging the pooling price from period t onwards. Thus, in expectation the seller's profits increase with public information.

Lemma 1 *For all $\lambda_t \in \Lambda(\lambda_1)$ the expected future profits to the seller from charging the separating price $p^H(\lambda_t)$ in period t (and charging the conditionally optimal price thereafter), strictly exceed the expected future profits from charging the pooling price $p^L(\lambda_t)$ in period t and thereafter.*

This result can also be derived as a corollary of a more general result proved by Ottaviani and Prat (2001) and generalized by Saak (2007). Ottaviani and Prat show that a sufficient condition for the monopolist to benefit from the revelation of public information is that this information is affiliated to the private information of the buyer—this assumption is automatically satisfied when the state is binary, as in the model considered here. In a general model with finitely many signal realizations, BOOV's Proposition 4 establishes the corresponding result that an information-revealing price maximizes future profits. In our setting with binary signals, no information is revealed when the pooling price or an exit price is posted, so that the separating price is the only price that reveals information. The advantage of focusing on the case with binary signals is that the amount of information revealed at different prices can be Blackwell ranked.

Since by Lemma 1 the separating price generates higher expected future profits than the pooling price, for the separating price to be optimal it is sufficient that the separating price generates higher immediate profits. For the pooling price to be optimal instead, it is necessary (but not sufficient) that its immediate profits exceeds those associated to the separating price.

4.2 Expected immediate profits

We now turn to the seller's expected *immediate profits* in period t. If the seller charges the pooling price, immediate profits are $R_L(\lambda_t) = p^L(\lambda_t) - c$, a convex function of λ with $R_L(0) = L - c$ and $R_L(1) = 1 + L - c = H - c$. When instead the seller charges the separating price $p^H(\lambda_t) = \lambda_t^+ + L$, the probability of a sale is $\varphi(\lambda_t) \equiv \Pr(s_t = h|\lambda_t)$ and expected immediate profits are $R_H(\lambda_t) = \left[p^H(\lambda_t) - c\right]\varphi(\lambda_t) = [\alpha + (2\alpha - 1)(L - c)]\lambda_t + (1 - \alpha)(L - c)$, a linear function of λ_t with $R_H(0) = (1 - \alpha)(L - c)$ and $R_H(1) = \alpha(1 + L - c) = \alpha(H - c)$.

We now compare the expected immediate profits from the pooling and separating price. Since $R_L(1) = (H - c) > \alpha(H - c) = R_H(1)$ and profits are continuous in λ, for sufficiently optimistic beliefs the immediate profits from the pooling price always exceed those from the separating price. More generally, $R_L(\lambda) - R_H(\lambda)$ depends crucially on how the cost of production compares to the low value L of the good. We distinguish three cases:

1. If $c = L$, either the two functions have a unique intersection for $\lambda > 0$ when α is large enough, or $R_H(\lambda) < R_L(\lambda)$ for all $\lambda > 0$, as shown in Fig. 1.a.1 and 1.a.2, respectively.
2. If $c < L$, we have $0 < R_H(0) < R_L(0)$ so that the two functions have either two intersections, no intersection, or a point of tangency depending on the level of α, as illustrated in Fig. 1.b.1–1.b.3.
3. If $c > L$, we have $0 > R_H(0) > R_L(0)$ so that $R_H(\lambda)$ intersects with $R_L(\lambda)$ exactly once. This case is illustrated in Fig. 1.c.1.

While the same price does not generally maximize both expected immediate and future profits, in some situations we need not consider both. When the belief is sufficiently optimistic, the monopolist benefits little by revealing information to future buyers, hence the myopically optimal price is dynamically optimal. When λ_t is high, the separating and the pooling price are almost identical, thus the potential increase in the seller's expected future profits from an increase of λ_t, even to its upper limit of 1, is small. With $\alpha < 1$ the probability of selling at the separating price is significantly lower, hence $R_L(\lambda_t) > R_H(\lambda_t)$ for high λ_t. According to the following lemma, the seller triggers a purchase cascade whenever the belief is sufficiently optimistic.

Lemma 2 *For every discount factor $\delta \in (0, 1)$ the seller chooses the pooling price $p^L(\lambda)$ whenever λ is sufficiently high, i.e., there exists an $\epsilon_\delta > 0$ such that $p_t = p^L(\lambda_t)$ whenever $\lambda_t \in (1 - \epsilon_\delta, 1)$.*

Thus, along the equilibrium path the belief λ_t is bounded away from 1 and buyers cannot learn asymptotically that the good's true value is high. Not surprisingly, the belief at which herding is triggered depends on the seller's discount factor. In fact, given any public belief $\lambda_t \in (0, 1)$ it is optimal for a sufficiently patient seller to charge the separating price $p^H(\lambda_t)$ or to exit the market. (see Lemma 4, Appendix A.1).

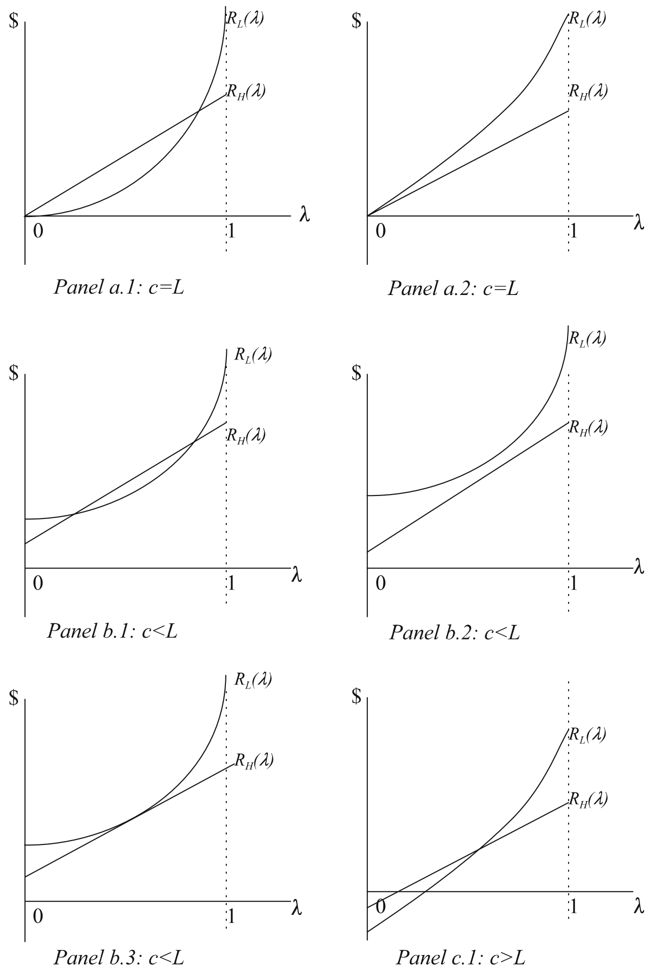

Fig. 1 In these graphs, R_H (λ) represents the expected immediate profit from the separating price p^H (λ), and R_L (λ) the expected immediate profit from the pooling price p^L (λ)

As suggested by the expected immediate profits (Fig. 1), the seller's optimal strategy depends on how the cost c relates to the low value of the object L. We distinguish three qualitatively different cases:

1. *Borderline case* ($c = L$). The borderline case is particularly instructive to demonstrate the mechanics of the solution. Even if this case may be regarded as non-generic, there are plausible situations in which $c = L$. For example, patent holders often have zero marginal cost from licensing ($c = 0$) and patents can be completely worthless to licensees ($L = 0$). In the borderline case the seller's (positive) profits per sale converge to zero for $\lambda \to 0$.
2. *Case with socially worthless information* ($c < L$). In this case, the buyers' private information has no social value because the socially optimal allocation requires

that all buyers purchase the good regardless of the realizations of their private signals. However, information has private value for the monopolist, since it affects the buyer's willingness to pay. In this case, the seller's profits per sale are bounded away from zero for all probabilities $\lambda \in [0, 1]$.

3. *Case with socially valuable information* ($c > L$). In this case, the socially optimal decision (purchase or no purchase) depends on the quality of the good (respectively, $v = H$ or $v = L$). Since the buyers collectively know this quality, their information is privately as well as socially valuable. The seller's option to exit the market matters in this case, but not in the first two cases.

5 Optimal pricing

To illustrate the seller's optimization problem and the resulting equilibrium, we begin in Sect. 5.1 by considering the borderline case with $c = L$. The analysis of this case provides the basic insights on how the equilibrium depends on the signal precision α and the seller's discount factor δ. When $c = L$ as well as when $c < L$ (analyzed in Sect. 5.2), we need compare only two prices since exit is never optimal and a purchase cascade is the only occurrence of herding. When instead $c > L$, exit becomes relevant (Sect. 5.3).

5.1 Borderline case ($c = L$)

Without loss of generality, we normalize $L = 0$, so that $p^H(\lambda_t) = \lambda_t^+$ and $p^L(\lambda_t) = \lambda_t^-$. Since the expected immediate profits from the separating price is $R_H(\lambda_t) = \varphi(\lambda_t)p^H(\lambda_t) = \alpha\lambda_t$, the difference between the expected immediate profits from the separating and the pooling price is

$$R_H(\lambda_t) - R_L(\lambda_t) = \alpha\lambda_t - p^L(\lambda_t) = \frac{\lambda_t\left[\alpha^2(1-\lambda_t) - (1-\alpha)(1-\alpha\lambda_t)\right]}{\lambda_t(1-\alpha) + (1-\lambda_t)\alpha}.$$

Figure 1.a.1 and 1.a.2 make clear that we need to distinguish the case in which $R_L(\lambda_t) > R_H(\lambda_t)$ for all $\lambda_t \in (0, 1)$ from the case in which $R_H(\lambda_t) > R_L(\lambda_t)$ for some $\lambda_t \in (0, 1)$:

Lemma 3 *If $\alpha^2 > 1-\alpha$, there exists a $\bar{\lambda}_\alpha \in (0, 1)$ such that the expected immediate profits from the separating price $p^H(\lambda_t)$ are identical to those from the pooling price $p^L(\lambda_t)$ for $\lambda_t = \bar{\lambda}_\alpha$, whereas they are larger for $\lambda_t < \bar{\lambda}_\alpha$ and smaller for $\lambda_t > \bar{\lambda}_\alpha$. Furthermore, $\bar{\lambda}_\alpha \to 1$ for $\alpha \to 1$. If $\alpha^2 \leq 1-\alpha$, the expected immediate profits from the pooling price $p^L(\lambda_t)$ exceed those of the separating price $p^H(\lambda_t)$ for all $\lambda_t \in \Lambda(\lambda_1)$.*

Figure 1.a.1 and 1.a.2 illustrate the cases with $\alpha^2 > 1-\alpha$ and $\alpha^2 \leq 1-\alpha$, respectively. The precision of the buyers' signal is critical for the seller's optimal strategy. For example, if the signal is almost perfect (α close to 1) even an extremely impatient seller demands the separating price unless λ_t is close to 1. This follows directly from Lemmas 1 and 3 because $\bar{\lambda}_\alpha \to 1$ for $\alpha \to 1$.

We consider the three sub-cases : (i) $\alpha^2 > 1-\alpha$ or, equivalently, $\alpha > \frac{1}{2}(\sqrt{5}-1) \simeq 0.618$; (ii) $\alpha^2 = 1-\alpha$, or $\alpha = \frac{1}{2}(\sqrt{5}-1)$; and (iii) $\alpha^2 < 1-\alpha$, or $\alpha < \frac{1}{2}(\sqrt{5}-1)$. We say that the signal is *precise* in case (i) and *noisy* in case (iii), while (ii) is the *threshold* case in which the likelihood ratio $\frac{1-\alpha}{\alpha}$ is equal to the probability α that the signal is correct.[9] Propositions 1, 2, and 3 below provide a full characterization of the seller's optimal strategy.

Consider first the seller's optimal pricing strategy when the signal is precise.

Proposition 1 *If* $c = L$ *and* $\alpha^2 > 1-\alpha$, *there exists a critical belief* $\lambda^{**} \geq \bar{\lambda}_\alpha$, $\lambda^{**} \in \Lambda(\lambda_1)$, *that depends on the seller's discount factor* δ, *such that it is uniquely optimal for the seller to demand*

$$p_t = \begin{cases} p^H(\lambda_t) & \text{whenever} \quad \lambda_t < \lambda^{**} \\ p^L(\lambda_t) & \text{whenever} \quad \lambda_t > \lambda^{**}. \end{cases}$$

For $\lambda_t = \lambda^{**}$, $p^L(\lambda_t)$ *is optimal, but* $p^H(\lambda_t)$ *may be optimal as well. Moreover,* $\lambda^{**} > \bar{\lambda}_\alpha$ *for* $\delta > 0$, *and* $\lambda^{**} = \min_{\lambda \in \Lambda(\lambda_1) \cap [\bar{\lambda}_\alpha, 1]} \lambda$ *for* $\delta = 0$, *so that if the seller is completely impatient,* λ^{**} *is the smallest* $\lambda \in \Lambda(\lambda_1)$ *such that* $\lambda \geq \bar{\lambda}_\alpha$. *For* $\delta \to 1$, $\lambda^{**} \to 1$, *so that as the seller becomes infinitely patient the separating price* $p^H(\lambda)$ *becomes uniquely optimal for the seller for all* $\lambda \in (0, 1)$.[10]

Proposition 1 shows that the set of attainable beliefs can be partitioned such that for low beliefs the separating price is optimal and for high beliefs the pooling price is optimal for the seller. As argued earlier, the intuition is that at high beliefs there is little to gain and much to lose from the separating price (Lemma 2). At low beliefs the converse holds for precise signals. Although at low beliefs there is a high probability of no sale at the separating price, the pooling price is low and thus the opportunity cost of not selling at the pooling price is small as well. With precise signals, the separating price is sufficiently large, relative to the pooling price, to more than compensate for the expected loss of no sale. The separating price then maximizes expected immediate profits and thus, by Lemma 1, the seller's expected payoff.

When the good's true value is low, λ_t may never hit λ^{**} and consequently herding may never arise. If herding does not occur, the belief will converge to zero due to the martingale convergence theorem. Thus, with precise signals it may asymptotically be revealed that the good's true value is low.

9 Note that the likelihood ratio $\frac{1-\alpha}{\alpha}$ determines the function $p^L(\lambda_t)$, i.e., the immediate return of the pooling price as a function of λ_t. The probability α determines the expected immediate return $\alpha\lambda_t$ of the separating price as a function of λ_t. For $\lambda_t = 0$ the respective derivatives are $\frac{dp^L(\lambda_t)}{d\lambda_t} = \frac{1-\alpha}{\alpha}$ and $\frac{d(\alpha\lambda_t)}{d\lambda_t} = \alpha$; thus these two slopes are identical in the threshold case with $\alpha^2 = 1-\alpha$. Incidentally, $\alpha^2 = 1-\alpha$ is the equation for the golden section.

10 In general $p^L(\lambda^{**})$ will be *uniquely* optimal at $\lambda_t = \lambda^{**}$ because $\Lambda(\lambda_1)$ is a discrete set and thus the λ where both $p^L(\lambda)$ and $p^H(\lambda)$ are optimal (and which would imply $\lambda = \lambda^{**}$) is generically not attainable. However, it is possible that this λ is attainable and hence at $\lambda_t = \lambda^{**}$ the separating price is also optimal. This holds for the Propositions 2–6 as well.

Next consider the seller's optimal strategy in the threshold case ($\alpha^2 = 1 - \alpha$). In this case, the seller's patience plays a more important role in determining the optimal pricing strategy.

Proposition 2 *If $c = L$ and $\alpha^2 = 1 - \alpha$, there exists a discount factor $\delta^* \in (0, 1)$ such that for all $\delta \in [0, \delta^*]$ the uniquely optimal prices are given by $p_t = p^L(\lambda_1)$ for all $t \in \{1, 2, ...\}$. For each discount factor $\delta \in (\delta^*, 1)$ there exists a critical belief $\lambda^{**} \in \Lambda(\lambda_1)$ that depends on δ, such that it is uniquely optimal for the seller to demand*

$$p_t = \begin{cases} p^H(\lambda_t) & \text{whenever} \quad \lambda_t < \lambda^{**} \\ p^L(\lambda_t) & \text{whenever} \quad \lambda_t > \lambda^{**}. \end{cases}$$

For $\lambda_t = \lambda^{**}$, $p^L(\lambda_t)$ is optimal, but $p^H(\lambda_t)$ may be optimal as well. Finally, $\lambda^{**} \to 0$ for $\delta \to \delta^*$, and $\lambda^{**} \to 1$ for $\delta \to 1$, i.e., whereas a sufficiently impatient seller always chooses the pooling price and triggers herding immediately, for a seller that becomes infinitely patient the separating price $p^H(\lambda)$ becomes uniquely optimal for all $\lambda \in (0, 1)$.

In contrast to the case with precise signals, in the threshold case the separating price generates lower expected immediate profits than the pooling price even for small λ_t (Fig. 1.a.2). Thus, for an impatient seller the pooling price $p^L(\lambda_1)$ is always uniquely optimal and herding arises immediately. For a patient seller the threshold case is similar to the precise signal case, and herding may, but need not, arise. The difference is that the belief $\bar{\lambda}_\alpha$ at which $R_H(\lambda) = R_L(\lambda)$ is equal to zero in the threshold case, but is positive in the case of precise signals, as discussed in Lemma 3.

Finally, consider the seller's optimal price when the signal is noisy, $\alpha^2 < 1 - \alpha$. As in the threshold case, with noisy signals the pooling price $p^L(\lambda_1)$ generates higher expected immediate profits (see Lemma 3 and Fig. 1.a.2), and hence an impatient seller chooses the pooling price and triggers herding immediately. For a patient seller herding does not occur immediately for priors λ_1 that lie within some range $(\lambda^*, \lambda^{**})$, but in contrast to the two previous cases herding arises eventually.[11] These results are shown in the following proposition, along with the possibility that sometimes the pooling and the separating price may both be optimal.

Proposition 3 *If $c = L$ and $\alpha^2 < 1 - \alpha$, there exist discount factors $\delta^{**} \in (0, 1)$ and $\delta^{***} \in [\delta^{**}, 1)$ such that:*

1. *For all $\delta \in [0, \delta^{**})$, the uniquely optimal prices are given by $p_t = p^L(\lambda_1)$ for all $t \in \{1, 2, ...\}$;*
2. *For $\delta \in [\delta^{**}, \delta^{***}]$, $p_t = p^L(\lambda_t)$ is optimal for all $\lambda_t \in \Lambda(\lambda_1)$, but $p_t = p^H(\lambda_t)$ is optimal as well for at least one $\lambda_t \in \Lambda(\lambda_1)$; and*
3. *For each $\delta \in (\delta^{***}, 1)$ there exist critical beliefs λ^* and λ^{**} (that depend on δ) in $\Lambda(\lambda_1)$, where $0 < \lambda^* < \lambda^{**} < 1$, such that it is uniquely optimal for the seller*

[11] Notice that for any $\lambda_1 \in (0, 1)$, $(\lambda^*, \lambda^{**}) \cap \Lambda(\lambda_1) \neq \emptyset$ for sufficiently large δ because, as Proposition 3 shows, $\lambda^* \to 0$ and $\lambda^{**} \to 1$ for $\delta \to 1$.

to demand

$$p_t = \begin{cases} p^H(\lambda_t) & \text{whenever} \quad \lambda_t \in (\lambda^*, \lambda^{**}) \\ p^L(\lambda_t) & \text{whenever} \quad \lambda_t \in (0, \lambda^*) \text{ or } \lambda_t \in (\lambda^{**}, 1)\,; \end{cases}$$

for $\lambda_t = \lambda^*$ *and* $\lambda_t = \lambda^{**}$, $p^L(\lambda_t)$ *is optimal, but* $p^H(\lambda_t)$ *may be optimal as well. For* $\delta \to 1$, $\lambda^* \to 0$ *and* $\lambda^{**} \to 1$, *i.e., as the seller becomes infinitely patient the separating price* $p^H(\lambda)$ *becomes uniquely optimal for the seller for all* $\lambda \in (0, 1)$.

In contrast to the case with precise signals and the threshold case, the seller triggers herding also for sufficiently low public beliefs, with the consequence that eventually herding occurs (almost surely). To understand this difference, consider the expected payoff achieved by the seller when always charging the separating price. Since in any period t the expected immediate profit from the separating price is $\alpha\lambda_t$ and λ_t is a martingale, the expected payoff is $\frac{\alpha\lambda_t}{1-\delta}$. If the separating price is optimal for all sufficiently small λ_t, it must be that for $\lambda_t \to 0$ the seller's expected payoff $V(\lambda_t)$ converges to $\frac{\alpha\lambda_t}{1-\delta}$.[12] On the other hand, the seller's payoff from triggering herding is $\frac{p^L(\lambda_t)}{1-\delta}$, which for $\lambda_t \to 0$ converges to $\frac{1}{1-\delta}\frac{1-\alpha}{\alpha}\lambda_t$ since $p^L(\lambda_t) = \frac{\lambda_t(1-\alpha)}{\lambda_t(1-\alpha)+(1-\lambda_t)\alpha}$. Consequently, if $\frac{\alpha}{1-\delta}\lambda_t > \frac{1}{1-\delta}\frac{1-\alpha}{\alpha}\lambda_t$, or $\alpha^2 > 1-\alpha$, the separating price is optimal for all sufficiently small λ_t, whereas if $\frac{\alpha}{1-\delta}\lambda_t < \frac{1}{1-\delta}\frac{1-\alpha}{\alpha}\lambda_t$, or $\alpha^2 < 1-\alpha$, the pooling price is optimal for all sufficiently small λ_t. This explains why λ^* is positive when the signal is noisy (Proposition 3) but zero when it is precise (Proposition 1). Since $\frac{\alpha}{1-\delta}\lambda_t = \frac{1}{1-\delta}\frac{1-\alpha}{\alpha}\lambda_t$ in the threshold case, the probability to reach the upper threshold λ^{**} gives the separating price the edge and implies $\lambda^* = 0$ (Proposition 2).

To summarize Propositions 1–3, herding may but need not arise in the case $c = L$. Depending on the parameters, the seller either initiates herding immediately or starts with a separating price. In the latter case, the buyer purchases the good following the observation of a high signal, which results in positive updating by the seller and the future buyers. The seller then continues to charge the separating price as long as the updated public belief is within a certain interval. An informational cascade arises as soon as the belief hits the lower barrier λ^* or the upper barrier λ^{**}.

The absorbing barriers are optimally chosen by the seller and thus the seller's problem can also be seen as one of optimal stopping. As soon as the price exceeds a critical level in some period t and the buyer actually buys at this price, the seller reduces the price somewhat (from $p_t = p^H(\lambda_t) = \lambda_t^+$ to $p_{t+1} = p^L(\lambda_{t+1}) = \lambda_{t+1}^- = \lambda_t < \lambda_t^+$ since $\lambda_{t+1} = \lambda_t^+$ because of the sale at t) and triggers herding. However, the price may never hit this critical level. Instead, the price may converge to zero and thereby reveal that the common value of the object is low. Although one might expect the seller to trigger a purchase cascade in order to prevent buyers from asymptotically learning

[12] This follows because $V(\lambda_t)$ cannot be below $\frac{\alpha\lambda_t}{1-\delta}$ and exceeds it only because the seller will switch to the pooling price if and when λ_t attains the (upper) threshold λ^{**}. But for any small λ_t the threshold λ^{**} will be reached only with a minute probability and, if at all, only after a long time. Because of this and discounting, the effect on $V(\lambda_t)$ becomes vanishingly small for $\lambda_t \to 0$, and thus if for all sufficiently small λ_t the separating price is optimal, $V(\lambda_t)$ converges to $\frac{\alpha}{1-\delta}\lambda_t$.

that the true value is low, the seller will not do so, except when the precision of the signal is poor.

If the signal is precise (or in the threshold case if the seller is patient) and $\lambda_1 < \lambda^{**}$, either a purchase cascade occurs in the long run or it is learnt asymptotically that the value is low. As shown in Proposition 1, in this case the seller does not trigger a purchase cascade as long as $\lambda_t < \lambda^{**}$, and $\Pr(\lambda_t < \lambda^{**} \text{ for all } t \in \{1, 2, ...\} | L) > 0$. Hence, with positive probability an informational cascade does not arise conditional on the true value being low. Conditional instead on the good's quality being high, an informational cascade arises with probability one.

If the signal is precise ($\alpha^2 > 1 - \alpha$), there is a range of λ, i.e., the interval $\left(0, \bar{\lambda}_\alpha\right)$, where the seller's degree of patience δ plays no role. For λ in this interval, regardless of δ the seller charges the separating price and herding will not occur in this range. Nevertheless, the seller's degree of patience matters because, as Proposition 1 shows, the separating price is optimal not only for probabilities $\lambda \in \left(0, \bar{\lambda}_\alpha\right)$ but also for probabilities $\lambda \in \left[\bar{\lambda}_\alpha, \lambda^{**}\right)$, and λ^{**} does depend on δ.

Depending on the signal's precision and the seller's patience, there may, but need not, be public beliefs such that the separating price is optimal. However, if the separating price is optimal for some public belief, then it is optimal for a connected set of public beliefs. As shown in Sect. 5.2, the optimal strategy retains this simple property also in the case in which information is socially worthless. If instead the information is socially valuable, we show in Sect. 5.3 that there are situations in which the pooling price is uniquely optimal for beliefs surrounded by beliefs at which the separating price is optimal.

5.2 Case with socially worthless information ($c < L$)

When $c < L$ the seller also triggers a purchase cascade when the belief is sufficiently pessimistic, regardless of the signal's precision. The intuition is as follows. When the belief λ approaches zero, the separating and the pooling price both converge to $L > c$. If the seller charges the pooling price, she gets the approximate profit $L - c$ for sure, whereas if she demands the separating price, she gets it only with a probability close to $1 - \alpha < 1/2$. Thus, as seen in Fig. 1.b.1–b.3, the difference in the expected immediate profit is bounded away from zero. On the other hand, the difference in the expected future profits converges to zero when the belief approaches zero. Consequently, for a small λ the difference in the expected immediate profits exceeds the difference in the expected future profits and the pooling price is optimal. Because of this and Lemma 2, the seller's optimal strategy is to trigger herding at low as well as at high public beliefs, analogously to the case $c = L$ with noisy signals.

Depending on the signal's precision, the seller's patience and the cost c, there may but need not be intermediate public beliefs such that the separating price is optimal. However, as the following proposition shows, if the separating price is optimal for some public belief, then it is optimal for *all* intermediate public beliefs between the optimistic and pessimistic public beliefs, respectively, where the pooling price is optimal.

Proposition 4 *If $c < L$, the seller's optimal strategy is characterized by the following properties:*

1. *If the discount factor δ is sufficiently close to 1 (i.e., if the seller is sufficiently patient), there exist probabilities λ^* and λ^{**} (that depend on δ) in $\Lambda(\lambda_1)$, where $0 < \lambda^* < \lambda^{**} < 1$, such that the separating price $p^H(\lambda)$ is uniquely optimal for the seller for all $\lambda \in (\lambda^*, \lambda^{**})$, whereas the pooling price $p^L(\lambda)$ is uniquely optimal for the seller for all $\lambda \in (0, \lambda^*)$ and all $\lambda \in (\lambda^{**}, 1)$. For $\lambda = \lambda^*$ and $\lambda = \lambda^{**}$ the pooling price $p^L(\lambda)$ is optimal for the seller, but the separating price $p^H(\lambda)$ may be optimal as well. For $\delta \to 1$ it holds that $\lambda^* \to 0$ and $\lambda^{**} \to 1$, hence as the discount factor converges to 1 the separating price $p^H(\lambda)$ becomes uniquely optimal for all $\lambda \in (0, 1)$.*
2. *If $\alpha^2 \leq 1 - \alpha$, there exists a $\delta^{**} > 0$ such that the pooling price $p^L(\lambda)$ is uniquely optimal for all $\lambda \in (0, 1)$ whenever $\delta < \delta^{**}$; that is, if $\alpha^2 \leq 1 - \alpha$ and the discount factor δ is sufficiently low, the pooling price $p^L(\lambda)$ is uniquely optimal for all $\lambda \in (0, 1)$.*
3. *If the signal precision α is sufficiently high, there exist probabilities $\mu^* \in \Lambda(\lambda_1)$ and $\mu^{**} \in \Lambda(\lambda_1)$, where $0 < \mu^* < \mu^{**} < 1$, that depend only on α and the difference $L - c$ between the low value and the unit cost, such that for all $\lambda \in [\mu^*, \mu^{**}]$ the separating price $p^H(\lambda)$ is uniquely optimal for the seller regardless of the discount factor δ.*

Part 2 of the proposition corresponds to Fig. 1.b.2, where $R_L(\lambda) > R_H(\lambda)$ for all λ, and part 3 corresponds to Fig. 1.b.1, where $R_H(\lambda) > R_L(\lambda)$ for some intermediate values of λ. Fig. 1.b.3 depicts the threshold case. Proposition4 implies that whenever $c < L$ herding will occur with probability 1. We saw in Proposition 1 that when $c = L$ and signals are precise, $p^H(\lambda)$ is optimal for $\lambda \in (0, \bar{\lambda}_\alpha)$ regardless of the discount factor δ. Thus, part 3 of Proposition 4 is analogous to Proposition 1. Similarly, part 2 of Proposition 4 is analogous to Propositions 2 and 3, respectively. Notice that, as before, the degree of patience matters more when the signal is noisy. When the signal is precise, there is a range of public beliefs where it is uniquely optimal to charge the separating price regardless of δ.

5.3 Case with socially valuable information ($c > L$)

As illustrated in Fig. 1.c.1, when $c > L$ the seller's option to exit the market becomes relevant. Whenever the belief is sufficiently pessimistic, the separating price (and a fortiori the pooling price) is below the cost c. The seller may still stay in the market and demand the separating price because there is a positive probability that there is a sequence of high signals, revealed by sales, that lead the price to rise above c. However, if λ is sufficiently low, the probability of ever reaching the profitable range of prices is small and, moreover, even if all future buyers receive high signal realizations, the seller has nevertheless to incur a loss for a long sequence of periods until at least the separating price exceeds the unit cost. In addition, due to discounting the present value of potential future profits is low. Consequently, whenever λ is sufficiently low, it is not worthwhile for the seller to incur these losses, and an exit cascade results.

If for some $\lambda \in \Lambda(\lambda_1)$ it holds that $p^L(\lambda) \leq c < p^H(\lambda)$, it is uniquely optimal for the seller to stay in the market and demand the separating price $p^H(\lambda)$ because the expected immediate profit from demanding the separating price, $\varphi(\lambda)\left[p^H(\lambda) - c\right]$, is positive (and thus the seller's expected payoff is positive), whereas the payoff from charging the pooling price and triggering herding, $\frac{1}{1-\delta}\left[p^L(\lambda) - c\right]$, is non-positive. From $c > L$ it follows that there are exactly two attainable probabilities $\lambda \in \Lambda(\lambda_1)$ that satisfy $p^L(\lambda) \leq c < p^H(\lambda)$. One, call it $\underline{\lambda}$, is the largest $\lambda \in \Lambda(\lambda_1)$ such that $p^L(\lambda) \leq c$. Since $p^H\left(\underline{\lambda}\right) > p^L\left(\underline{\lambda}^+\right) > c$, it must be that $c < p^H\left(\underline{\lambda}\right)$. In addition, $p^H\left(\underline{\lambda}^-\right) = p^L\left(\underline{\lambda}^+\right)$ implies $p^L\left(\underline{\lambda}^-\right) < c < p^H\left(\underline{\lambda}^-\right)$, whereas for all $\lambda < \underline{\lambda}^-$, $\lambda \in \Lambda(\lambda_1)$, it holds that $p^L(\lambda) < p^H(\lambda) < c$, and for all $\lambda \geq \underline{\lambda}^+$, $\lambda \in \Lambda(\lambda_1)$, it holds that $c < p^L(\lambda) < p^H(\lambda)$.

Recall that the seller's expected immediate profit is linear in λ_t for the separating price $p^H(\lambda_t)$, and strictly convex in λ_t for the pooling price $p^L(\lambda_t)$. Moreover, if λ_t is sufficiently large, the expected immediate profit from demanding the separating price is strictly below the immediate profit from charging the pooling price since $R_H(1) = \alpha(1 + L - c) < 1 + L - c = R_L(1)$, as illustrated by Fig. 1.c.1. In combination with the analysis of the preceding paragraph, this implies that there is exactly one $\lambda \in (0, 1)$, which we denote by $\bar{\lambda}_\alpha$, such that the expected immediate profit from demanding the separating price is positive and equal to the immediate profit from charging the pooling price, i.e., $R_H\left(\bar{\lambda}_\alpha\right) = R_L\left(\bar{\lambda}_\alpha\right) > 0$.[13] Furthermore, $R_H(0) = -(1-\alpha)(c - L) > -(c - L) = R_L(0)$ implies $R_H(\lambda) > R_L(\lambda)$ for all $\lambda < \bar{\lambda}_\alpha$. Consequently, at all public beliefs below $\bar{\lambda}_\alpha$ either the separating price or exit is optimal for the seller.

The seller's optimal strategy depends again on the precision of the signal α and the discount factor δ. The solution's qualitative characteristics depend critically on whether $\alpha \geq \left(1 - \alpha^2\right)(H - c)$ or not.[14] First, if signals are sufficiently precise, i.e., if $\alpha \geq \left(1 - \alpha^2\right)(H - c)$, Proposition 5 below shows that the separating price is again optimal only in one single connected interval.[15] Second, if instead the signal precision is sufficiently low, $\alpha < \left(1 - \alpha^2\right)(H - c)$, the separating price may be optimal in two disconnected intervals.

First, for the case with $\alpha \geq \left(1 - \alpha^2\right)(H - c)$ we derive a result analogous to Proposition 1 , modified for exit. Provided the seller is active in the market, the seller demands the separating price if λ is below a certain threshold and the pooling price

[13] Note that for $c \to L$, $\bar{\lambda}_\alpha$ converges to the $\bar{\lambda}_\alpha$ that we have defined above for the case $c = L$. In general, $\bar{\lambda}_\alpha$ will not be an attainable λ, i.e., will not be an element of $\Lambda(\lambda_1)$.

[14] Without the normalization $H - L = 1$ the condition would read $\alpha \geq \left(1 - \alpha^2\right)\frac{H-c}{H-L}$.

[15] The condition $\alpha \geq \left(1 - \alpha^2\right)(H - c)$ is equivalent to

$$\frac{1-\alpha}{\alpha} - \alpha \leq \frac{c - L}{H - c},$$

which is independent of the normalization $H - L = 1$. For $c = L$ this condition reduces to $\alpha^2 \geq 1 - \alpha$ (i.e., the signal must not be noisy in the sense of Sect. 5.1), which is consistent with our analysis of the borderline case in Sect. 5.1. However, for $c > L$ the condition $\alpha^2 \geq 1 - \alpha$ is sufficient but not necessary for $\alpha \geq \left(1 - \alpha^2\right)(H - c)$ to hold, since $\frac{c-L}{H-c} > 0$.

if λ is above that threshold. This corresponds to what we have found for the case with $c = L$ and precise signals. However, in contrast to Proposition 1 exit is now optimal whenever λ is sufficiently small. As noted in footnote 15, precise or threshold case signals in the sense of Sect. 5.1 are sufficient but not necessary for the condition $\alpha \geq \left(1-\alpha^2\right)(H-c)$ to hold.

Proposition 5 *If $c > L$ and $\alpha \geq \left(1-\alpha^2\right)(H-c)$, the optimal strategy of the seller is characterized by the following properties. For every discount factor $\delta \in [0,1)$ there exist probabilities λ^E and λ^{**} (that depend on δ) in $\Lambda(\lambda_1)$, with $0 < \lambda^E < \bar{\lambda}_\alpha \leq \lambda^{**} < 1$, such that for all $\lambda \in \left(0, \lambda^E\right)$ it is uniquely optimal for the seller to exit the market, whereas for all $\lambda \in \left(\lambda^E, 1\right)$ it is uniquely optimal for the seller to stay in the market. Moreover, for all $\lambda \in \left(\lambda^E, \lambda^{**}\right)$ the separating price $p^H(\lambda)$ is uniquely optimal for the seller, and the interval $\left(\lambda^E, \lambda^{**}\right)$ contains at least two attainable values of λ (i.e., $\left(\lambda^E, \lambda^{**}\right) \cap \Lambda(\lambda_1)$ contains at least two elements). For all $\lambda \in (\lambda^{**}, 1)$ the pooling price $p^L(\lambda)$ is uniquely optimal for the seller. For $\lambda = \lambda^E$ exit is optimal, but staying in the market and charging the separating price $p^H(\lambda)$ may be optimal as well; and for $\lambda = \lambda^{**}$ the pooling price $p^L(\lambda)$ is optimal, but the separating price $p^H(\lambda)$ may be optimal as well. For $\delta = 0$, $\lambda^{**} = \min_{\lambda \in \Lambda(\lambda_1) \cap [\bar{\lambda}_\alpha, 1]} \lambda$, i.e., if the seller is completely impatient, λ^{**} is the smallest $\lambda \in \Lambda(\lambda_1)$ such that $\lambda \geq \bar{\lambda}_\alpha$. For $\delta \to 1$, $\lambda^{**} \to 1$, i.e., as the seller becomes infinitely patient the separating price $p^H(\lambda)$ becomes uniquely optimal for the seller for all $\lambda \in \left(\lambda^E, 1\right)$ and thus the seller either charges the separating price or exits the market.*

Second, for the case with $\alpha < \left(1-\alpha^2\right)(H-c)$ we find the possibility that an informational cascade arises for interior beliefs—a key finding of this paper.

Proposition 6 *If $c > L$ and $\alpha < \left(1-\alpha^2\right)(H-c)$, the optimal strategy of the seller is characterized by the following properties. For every discount factor $\delta \in [0,1)$ there exist probabilities $\lambda^E, \hat{\lambda}', \hat{\lambda}''$ and λ^{**}(that depend on δ) in $\Lambda(\lambda_1)$, where $0 < \lambda^E < \bar{\lambda}_\alpha \leq \hat{\lambda}' \leq \hat{\lambda}'' \leq \lambda^{**} < 1$, such that for all $\lambda \in \left(0, \lambda^E\right)$ it is uniquely optimal for the seller to exit the market, whereas for all $\lambda \in \left(\lambda^E, 1\right)$ it is uniquely optimal for the seller to stay in the market. Moreover, for all $\lambda \in \left(\lambda^E, \hat{\lambda}'\right)$ and for all $\lambda \in \left(\hat{\lambda}'', \lambda^{**}\right)$ it is uniquely optimal for the seller to charge the separating price $p^H(\lambda)$, whereas for all $\lambda \in \left(\hat{\lambda}', \hat{\lambda}''\right)$ and for all $\lambda \in \left(\lambda^{**}, 1\right)$ it is uniquely optimal for the seller to charge the pooling price $p^L\left(\lambda\right)$. For $\lambda = \lambda^E$ exit is optimal, but staying in the market and charging the separating price $p^H(\lambda)$ may be optimal as well; and for $\lambda \in \left\{\hat{\lambda}', \hat{\lambda}'', \lambda^{**}\right\}$ the pooling price $p^L\left(\lambda\right)$ is optimal, but the separating price $p^H\left(\lambda\right)$ may be optimal as well. For every $c > L$ there exists a discount factor $\delta' \in [0,1)$ such that $\hat{\lambda}' = \hat{\lambda}'' = \lambda^{**}$ (and thus the interval $\left(\hat{\lambda}', \hat{\lambda}''\right)$ is empty) for all $\delta \in \left(\delta', 1\right)$. There also exist parameter values of $\delta \in \left(0, 1\right)$ and $c > L$ such that $\hat{\lambda}' < \hat{\lambda}'' < \lambda^{**}$ and $\left(\hat{\lambda}', \hat{\lambda}''\right)$ contains at least one attainable λ (i.e., $\left(\hat{\lambda}', \hat{\lambda}''\right) \cap \Lambda\left(\lambda_1\right) \neq \emptyset$). When the seller becomes arbitrarily patient, $\delta \to 1$, we have $\lambda^{**} \to 1$ (and $\hat{\lambda}' = \hat{\lambda}'' = \lambda^{**}$), so that the separating price $p^H\left(\lambda\right)$ becomes uniquely optimal for the seller for all $\lambda \in \left(\lambda^E, 1\right)$ and thus the seller either charges the separating price or exits the market.*

An important and unexpected conclusion from our analysis is that there are cases in which the separating price is uniquely optimal in two disconnected intervals. This proves that the property of the specific illustration in Chamley (2004, pp. 81–83)—according to which the seller exits the market for sufficiently pessimistic public beliefs, charges separating prices at which learning takes place for all public beliefs in an intermediate interval, and induces a purchase cascade at sufficiently optimistic beliefs—does not hold in general. For parameter values that satisfy $\alpha < \left(1-\alpha^2\right)(H-c)$ it is possible that when we increase λ, starting with $\lambda = 0$, exit is optimal for the seller for extremely low values of λ, then the separating price is optimal, then the pooling price, then the separating price once more, and finally the pooling price is again optimal for the seller. The proof of Proposition 6 (and specifically of Lemma 7 in the Appendix) implies that the possibility that the separating price may be optimal in two disconnected intervals is not due to the seller's exit option. If exit were not feasible, the separating price would be charged at *all* sufficiently low public beliefs rather than only for those where exit is not optimal, but when we increase λ above that level it would still be the case that first the pooling price becomes optimal, then the separating price once more, and finally the pooling price again.[16]

To gain some intuitive understanding of Proposition 6, the following continuity argument is useful. Consider a situation corresponding to Proposition 3, in which $c = L$, signals are noisy ($\alpha^2 < 1-\alpha$), and the pooling price is optimal for $\lambda \in (0, \lambda^*)$ and $\lambda \in (\lambda^{**}, 1)$, with $0 < \lambda^* < \lambda^{**} < 1$, whereas the separating price is optimal for $\lambda \in (\lambda^*, \lambda^{**})$. That is, the pooling price is optimal for sufficiently pessimistic and optimistic public beliefs, whereas the separating price is optimal for intermediate public beliefs. What is the effect of increasing the cost c by a small amount so that $c-L$ is positive but almost zero? We know that with $c > L$ for extremely low values of λ first exit and then the separating price is optimal, but for the remaining (higher) levels of λ we expect a situation that is similar to the case $c = L$, provided $c - L$ is sufficiently small. That is, for pessimistic (though not extremely pessimistic) public beliefs the pooling price is optimal, for intermediate public beliefs the separating price, and for optimistic public beliefs again the pooling price is optimal. Thus, when we increase λ from zero to one, exit is optimal for the seller for extremely low values of λ, then the separating price is optimal, then the pooling price, then the separating price once more, and finally the pooling price is again optimal for the seller.[17]

To further understand the unexpected result of Proposition 6, consider how the public belief λ affects the difference in the seller's payoff between the separating and

[16] If exit is not feasible, the separating price is optimal for all sufficiently low public beliefs λ since whenever $p^L(\lambda) \leq c$ the separating price generates a strictly larger expected immediate profit than the pooling price. This follows because the separating price exceeds the pooling price and deters buyers who have observed a low signal realization (notice that for $p^L(\lambda) < c$ the seller makes a loss at the pooling price, and for $p^H(\lambda) < c$ this holds even at the separating price).

[17] In contrast to the case with $c = L$, the separating price is optimal at the lowest public beliefs such that the seller does not exit the market, and this makes the separating price more attractive to the seller also at higher public beliefs (because it reduces the negative effect on the seller's expected payoff of a non-sale). Therefore, the condition for the occurrence of the separating price at low beliefs, $\alpha < \left(1-\alpha^2\right)(H-c)$, is more restrictive (and only necessary) than the respective (necessary and sufficient) condition for the case with $c = L$, i.e., $\alpha^2 < 1-\alpha$.

the pooling price. Why is it possible for the separating price to be optimal for λ in an interval, but then no longer optimal for a higher λ, and finally optimal again for even larger λs? First, the difference in the expected immediate profit between the separating and the pooling price decreases in λ to the right of the point where $R_L(\lambda)$ and $R_H(\lambda)$ intersect, as also illustrated in Fig. 1.c.1.[18] Second, the informativeness of the signal is maximized for $\lambda = 1/2$. For λ close to zero the signal provides little information, but becomes more informative as λ increases towards $1/2$, and then again becomes less informative for λ beyond $1/2$. Therefore, for moderate values of λ the (positive) difference in the expected *future* profits between the separating and the pooling price increases, whereas the difference in the expected *immediate* profit between the separating and the pooling price decreases. Because of these two counteracting forces, the difference in the seller's *total* expected profit between the separating and the pooling price need not be monotonic in λ and may first decrease from a positive to a negative value (implying a switch from the separating to the pooling price), then increase to a positive value (making the separating price optimal yet another time) and finally decrease again to a negative value (and then the pooling price is optimal anew). This intuitive argument sheds some light on Proposition 6. More precisely, the proof of Lemma 7 in the Appendix implies (i) that for $\alpha \geq \left(1-\alpha^2\right)(H-c)$ a function related to the difference in the total expected profit between the separating and the pooling price decreases monotonically in λ and thus $\alpha < \left(1-\alpha^2\right)(H-c)$ is a necessary condition for the "recurrence" of the separating price; and (ii) that there can be at most two disconnected intervals in which the separating price is optimal.

6 Price sequence

We now derive the stochastic properties of the price sequence. To this end, we need to define more precisely the prices that result in an exit cascade. An exit cascade results at any price strictly above the separating price, since no buyer purchases the good at such a price. We make the natural assumption that the exit price is $p^E(\lambda) = p^H(\lambda) + \varepsilon$, with $\varepsilon > 0$ arbitrarily small. Under this assumption, we show that the price decreases on average.

Proposition 7 *The price sequence is a supermartingale,* $E\left[p_{t+1}|\lambda_t\right] \leq p_t$.

To translate this result into an empirical statement, the average should be taken over a number of (ex-ante identical) markets. Once prices are averaged in this way, they should display a downward trend.[19]

---

[18] Under the condition of Proposition 6, $\alpha < \left(1-\alpha^2\right)(H-c)$, the difference in expected immediate profits between the separating and the pooling price decreases in λ everywhere. The slope of $R_H(\lambda)$ is given by $R'_H(\lambda) = \alpha - (2\alpha - 1)(c - L)$ and for the slope of $R_L(\lambda)$ it holds that $R'_L(\lambda) \geq R'_L(0) = \frac{1-\alpha}{\alpha}$. Since $\alpha < \left(1-\alpha^2\right)(H-c)$ and $c > L$ imply (cf. footnote 15) $\frac{1-\alpha}{\alpha} > \alpha > \alpha - (2\alpha - 1)(c - L)$, it follows that $\frac{d}{d\lambda}\left[R_L(\lambda) - R_H(\lambda)\right] > 0$. However, for the argument it is actually sufficient that this result holds for all λs that satisfy $R_L(\lambda) > R_H(\lambda)$, and this is true whenever $c > L$.

[19] However, it is easy to show that for a fixed good of high quality the price tends to increase on average in the learning phase.

This result is driven by the following three properties. First, the separating price $p^H(\lambda)$ is a concave function of the belief λ. This first property, combined with the fact that the belief is a martingale (see Eq. 2), implies by Jensen's inequality that prices decrease on average when separating prices are charged in period $t+1$ as well as in period t. Second, the pooling price is lower than the separating price, $p^L(\lambda) < p^H(\lambda)$. This second property implies that the price decreases when a purchase cascade is started. Third, the exit price $p^E = p^H(\lambda) + \varepsilon$ is arbitrarily close to the separating price provided that ε is sufficiently small, and thus the argument given above for the case in which the separating price is charged in period $t+1$ extends to the case in which exit takes place following no purchase. This third property implies that the price decreases also when an exit cascade is started.

This result on the stochastic properties of the price sequence is special to our binary signal setting. We have constructed a simple two-period example with three signal realizations in which the optimal price increases on average over time in the learning phase. With more than two signal realizations, there are cases in which the price increases when a purchase cascade is induced.[20]

7 Conclusion

This paper analyzes monopoly pricing with an exogenous sequence of buyers who learn from each other's purchases. Specifically, the goods sold are non-durable and have a common but unknown value, which is either high or low. In contrast to herding models with a fixed price, the monopolist is free to post a new price in every period. Consequently, the monopolist can strategically influence learning dynamically rather than only once—as it is the case if the price is set at the outset once and for all—or not at all—as it is the case if the price is exogenously given.[21] Focusing on the

[20] When a purchase cascade is induced after a purchase at a non-pooling price, there are two effects on the price. On the one hand, the purchase results in an updated belief at which all prices are higher. On the other hand, for any given belief the pooling price charged in a purchase cascade is lower than any price at which some buyers do not purchase. This second effect tends to result in a price reduction. With two signal realizations the second effect dominates, while there are three-signal examples in which the first effect dominates.

[21] If the price is exogenously fixed, the standard model applies and, depending on the price and the prior, a cascade occurs either immediately, after one signal revelation, or as soon as two consecutive identical signals are realized (see e.g., Chamley 2004, Sect. 4.1, pp. 62–67). If the monopolist could choose the price but would not be able to change it thereafter, only one out of four prices can be optimal. These four prices are the pooling price $p^L(\lambda_1)$, the separating price $p^H(\lambda_1)$, the exit price $p^E(\lambda_1)$, and the "prior price" $p^P(\lambda_1) \equiv E(v|\lambda_1) = \lambda_1 + L$. The "prior price" did not appear so far, but, in contrast to the case with flexible prices, this price may be optimal if the monopolist can choose the price only once and for all. The reason is that if at the prior price $p^P(\lambda_1)$ there is no sale and thus the public belief drops to λ_1^-, the seller has a "second chance" in the sense that the second buyer will buy at price $p^P(\lambda_1) = p^H(\lambda_1^-)$ when observing a high signal (and this will bring back the public belief to λ_1). This is not true for a higher price. Moreover, a price below the prior price but above the pooling price cannot be optimal because it does not increase the probability of a sale in any period. A similar argument shows that prices below the separating but above the prior price cannot be optimal. Consequently, depending on the parameters the optimal fixed price can only be one of the four prices $\{p^L(\lambda_1), p^P(\lambda_1), p^H(\lambda_1), p^E(\lambda_1)\}$.

case with symmetric binary signals, we are able to provide a complete characterization of the equilibrium and answer the four questions posed at the outset of the paper.

First, a (patient) monopolist charges a relatively high price in the initial periods in order to build a good track record, even at the risk of failing to sell for a long time. When the buyer has a binary signal and past prices are observable, a high (separating) price allows future buyers to infer the information of the current buyer while a low (pooling) price induces the buyer to purchase regardless of the signal realization. The monopolist benefits from the social learning process and so charges the separating price.

Second, we determine when the seller induces an informational cascade in which all potential buyers purchase regardless of their private information. Unless infinitely patient, the monopolist always induces a purchase cascade after a sufficiently long stream of successful sales at relatively high prices. If the value of the good is known to exceed its cost of production, the monopolist never exits the market and eventually induces a purchase cascade, even after many buyers refuse to purchase at separating prices. As in the bandit literature, the intuitive reason is that the seller finds it too costly to achieve complete learning. However, when the signal is sufficiently precise and the cost of production is exactly equal to the low value of the good, the monopolist demands the separating price for all sufficiently pessimistic beliefs and hence allows the buyers to learn the true value of the good conditional on this value being low.

Third, we show that there is a region of parameters for which learning stops at intermediate beliefs—a result we certainly could not foretell. When the signal is sufficiently uninformative and the production cost exceeds the low value of the good by a sufficiently small amount, the separating price is optimal in two disjoint intervals of public beliefs, whereas the pooling price is optimal for beliefs that lie, respectively, between or above these two intervals. In this case, a purchase cascade may arise for both high and intermediate public beliefs. With the exception of this case (as well as the trivial case in which they are triggered for all public beliefs), informational cascades are triggered only for extreme (i.e., sufficiently optimistic or pessimistic) but not for intermediate public beliefs.

Fourth, we show that monopoly prices on average decrease over time. In the initial periods when the monopolist charges separating prices, there is active learning. Since the expected belief tomorrow is equal to the belief today and the separating price is a concave function of the belief, prices decrease on average in the learning phase.[22] In addition, the price is reduced deterministically at the onset of a purchase cascade. While this fourth set of results is special to our binary signal setting, the other three results reflect more general properties.

[22] In turn, concavity of the separating price in the prior belief is driven by concavity of the posterior belief following a high signal. Intuitively, the marginal impact of an increase in the prior belief on this posterior belief is lower for higher prior beliefs.

Appendix

Notation

Let $F(\lambda)$ denote the difference in the seller's expected payoff obtained by (i) charging the separating price now and the pooling price (associated with the updated belief) from the next period onwards and (ii) charging the pooling price now and forever, i.e., $F(\lambda) \equiv [\alpha + (2\alpha - 1)(L - c)]\lambda + (1 - \alpha)(L - c) + \delta[\varphi(\lambda)\frac{p^L(\lambda^+)-c}{1-\delta} + [1 - \varphi(\lambda)] \times \frac{p^L(\lambda^-)-c}{1-\delta}] - \frac{p^L(\lambda)-c}{1-\delta}$. By $\frac{1}{1-\delta} = 1 + \frac{\delta}{1-\delta}$, we have

$$F(\lambda) = \alpha c + (1 - \alpha) L + [\alpha + (2\alpha - 1)(L - c)]\lambda - p^L(\lambda) + \frac{\delta}{1 - \delta}[\varphi(\lambda) p^L(\lambda^+) + (1 - \varphi(\lambda)) p^L(\lambda^-) - p^L(\lambda)]. \tag{3}$$

If $F(\lambda) > 0$, the separating price $p^H(\lambda)$ is the uniquely optimal price. If $F(\lambda) \leq 0$, the separating or the pooling price (or both) may be optimal.

Given λ_1, let $\lambda^j \in \Lambda(\lambda_1)$ be an arbitrary element of the set defined in (1). Define $\varphi_j \equiv \varphi\left(\lambda^j\right) \equiv \Pr\left(s = h|\lambda^j\right)$.

Overview

Section A.1 collects the proofs of all the lemmas. Lemmas 5, 6, 8, and Corollary 1 are key for the characterization of the optimal price in the subintervals in which $F(\lambda) \leq 0$. Essentially, we show that whenever $F(\lambda) \leq 0$ the separating price is uniquely optimal only if it is linked by a chain of separating prices (each optimal at its respective λ) to a public belief λ at which $F(\lambda) > 0$. Lemma 7 derives, for different parameter configurations, the subintervals of the set (0, 1) of possible public beliefs λ for which $F(\lambda)$ is positive, zero, or negative. These results are used repeatedly in the proofs of the propositions.

Section A.1 also contains the proofs of Lemmas 1–3 whose statements appear in the main text, as well as the statement and proof of Lemma 4 which is a result of independent interest mentioned in Sect. 4.2. The proofs of Lemmas 1, 2, and 4 are included to make the paper self contained, but these results can also be established as corollaries of BOOV's Lemmas 1, Proposition 6, and Proposition 5, respectively.

Section A.2 collects the proofs of all the propositions stated in the text.

A.1 Proofs of the propaedeutic results

Proof of Lemma 1 The function $p^L(\lambda) = \frac{1-\alpha}{1-\alpha+(1-\lambda)\alpha/\lambda} + L$, where $\lambda \in [0, 1]$, is strictly convex. Moreover, $V(\lambda) \geq \frac{1}{1-\delta}\left[p^L(\lambda) - c\right]$ for all $\lambda \in \Lambda(\lambda_1)$. If the seller charges the separating price $p^H(\lambda_t)$ in t and the optimal price thereafter, the expected future payoff is $\delta\varphi(\lambda_t)\ V\left(\lambda_t^+\right) + \delta[1 - \varphi(\lambda_t)]\ V\left(\lambda_t^-\right) \geq \delta\varphi(\lambda_t)\ \frac{p^L(\lambda_t^+)-c}{1-\delta} +$

$\delta[1-\varphi(\lambda_t)]\frac{p^L(\lambda_t^-)-c}{1-\delta} > \delta\frac{p^L(\lambda_t)-c}{1-\delta}$ because $\lambda_t = \varphi(\lambda_t)\lambda_t^+ + [1-\varphi(\lambda_t)]\lambda_t^-$ and $p^L(\lambda)$ is strictly convex. □

Proof of Lemma 2 The seller's payoff from charging $p^L(\lambda_t)$ from period t onwards is $\frac{p^L(\lambda_t)-c}{1-\delta} = (\lambda_t^- + L - c) + \frac{\delta}{1-\delta}\left[p^L(\lambda_t) - c\right]$. Since $p^H(\lambda) < H$ for all $\lambda \in (0, 1)$, the expected payoff from charging $p^H(\lambda_t)$ is less than $[\alpha + (2\alpha - 1)(L-c)]\lambda_t + (1-\alpha)(L-c) + \frac{\delta}{1-\delta}(H-c)$. The difference, $\lambda_t^- - [\alpha + (2\alpha-1)(L-c)]\lambda_t + \alpha(L-c) + \frac{\delta}{1-\delta}\left[p^L(\lambda_t) - H\right]$, converges to $1-\alpha+(1-\alpha)(L-c) = (1-\alpha)(1+L-c) = (1-\alpha)(H-c) > 0$ for $\lambda_t \to 1$. Thus, $p^L(\lambda_t)$ generates a higher expected payoff than $p^H(\lambda_t)$ whenever λ_t is sufficiently large. □

Proof of Lemma 3 Assume $\alpha^2 > 1-\alpha$. Then, $\alpha^2(1-\lambda) - (1-\alpha)(1-\alpha\lambda)$ is equal to $\alpha^2 - (1-\alpha) > 0$ for $\lambda = 0$ and to $-(1-\alpha)^2 < 0$ for $\lambda = 1$. The term $\alpha^2(1-\lambda) - (1-\alpha)(1-\alpha\lambda)$ is continuous and decreasing in $\lambda \in [0, 1]$ since $\alpha > 1/2$. Hence there is a unique $\bar{\lambda}_\alpha \in (0, 1)$ such that $\alpha^2(1-\bar{\lambda}_\alpha) - (1-\alpha)(1-\alpha\bar{\lambda}_\alpha) = 0$, and $\alpha\lambda_t - p^L(\lambda_t) \gtreqless 0$ for $\lambda_t \lesseqgtr \bar{\lambda}_\alpha$. Since $\alpha^2(1-\lambda) - (1-\alpha)(1-\alpha\lambda)$ increases in α, $\bar{\lambda}_\alpha \to 1$ for $\alpha \to 1$. When $\alpha^2 \leq 1-\alpha$, we have $\alpha^2(1-\lambda_t) - (1-\alpha)(1-\alpha\lambda_t) \leq -(1-\alpha)^2\lambda_t < 0$, hence $\alpha\lambda_t - p^L(\lambda_t) < 0$ for all $\lambda_t \in (0, 1)$. □

Lemma 4 *Given any* $\lambda \in (0, 1)$*, if the seller is sufficiently patient, then charging the pooling price* $p^L(\lambda)$ *is not an optimal strategy. Either the separating price or exit is optimal.*

Proof of Lemma 4 It is enough to show that for each $\lambda \in (0, 1)$, $F(\lambda) > 0$ if δ is sufficiently large. Since p^L is strictly convex, $[\varphi(\lambda)p^L(\lambda^+) + (1-\varphi(\lambda))p^L(\lambda^-) - p^L(\lambda)] > 0$. For $\delta \to 1$, the right hand side of Eq. (3) becomes arbitrarily large. Hence $F(\lambda) > 0$ for δ sufficiently large. □

Lemma 5 *Let* $\left\{\lambda^j, \lambda^{j+1} ..., \lambda^{j+K}\right\}$, $K \geqq 2$, *where* $\lambda^{j+k} \equiv \Pr\left(H \mid \lambda^j, k \text{ signals } s = h\right)$*, be a sequence of* λ*'s such that* $F\left(\lambda^{j+k}\right) \leqq 0$ *for all* $k \in \{0, ..., K\}$*. If* $V\left(\lambda^{j+k}\right) > \frac{p^L(\lambda^{j+k})-c}{1-\delta}$ *for all* $k \in \{1, ..., K-1\}$*, then either* $V\left(\lambda^j\right) > \frac{p^L(\lambda^j)-c}{1-\delta}$ *or* $V\left(\lambda^{j+K}\right) > \frac{p^L(\lambda^{j+K})-c}{1-\delta}$ *(or both).*

Proof of Lemma 5 The proof is by contradiction. Assume $V\left(\lambda^j\right) = \frac{p^L(\lambda^j)-c}{1-\delta}$ and $V\left(\lambda^{j+K}\right) = \frac{p^L(\lambda^{j+K})-c}{1-\delta}$. We simplify the notation by $V_{j+k} \equiv V\left(\lambda^{j+k}\right)$, $k \in \{0, 1 \ldots, K\}$. Recall that $p^L\left(\lambda^{j+k}\right) = \lambda^{j+k-1} + L$. We have

$$V_{j+k} = \begin{cases} \lambda^{j-1} + L - c + \delta V_j & \text{for } k = 0 \\ [\alpha + (2\alpha-1)(L-c)]\lambda^{j+k} + (1-\alpha)(L-c) & \\ \quad +\delta\left(1-\varphi_{j+k}\right)V_{j+k-1} + \delta\varphi_{j+k}V_{j+k+1} & \text{for } k \in \{1 \ldots, K-1\} \\ \lambda^{j+K-1} + L - c + \delta V_{j+K} & \text{for } k = K. \end{cases} \tag{4}$$

Defining

$$V \equiv \begin{bmatrix} V_j \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ V_{j+K} \end{bmatrix}, \quad J \equiv \begin{bmatrix} \lambda^{j-1} + L - c \\ [\alpha + (2\alpha - 1)(L - c)]\,\lambda^{j+1} + (1 - \alpha)(L - c) \\ [\alpha + (2\alpha - 1)(L - c)]\,\lambda^{j+2} + (1 - \alpha)(L - c) \\ \cdot \\ \cdot \\ \cdot \\ [\alpha + (2\alpha - 1)(L - c)]\,\lambda^{j+K-1} + (1 - \alpha)(L - c) \\ \lambda^{j+K-1} + L - c \end{bmatrix},$$

and

$$A \equiv \begin{bmatrix} 1 & 0 & 0 & 0 & . \; . \; . & 0 & 0 & 0 \\ 1 - \varphi_{j+1} & 0 & \varphi_{j+1} & 0 & . \; . \; . & 0 & 0 & 0 \\ 0 & 1 - \varphi_{j+2} & 0 & \varphi_{j+2} & . \; . \; . & 0 & 0 & 0 \\ - & - & - & - & - - - - - & - & - & - \\ 0 & 0 & 0 & 0 & . \; . \; . & 1 - \varphi_{j+K-1} & 0 & \varphi_{j+K-1} \\ 0 & 0 & 0 & 0 & . \; . \; . & 0 & 0 & 1 \end{bmatrix},$$

(4) can be rewritten as $V = J + \delta A V$. Notice that A is a nonnegative square matrix and that each row sum is 1. Therefore, A has a Frobenius root of 1, since the Frobenius root must lie between the lowest and the highest row sum of A (see, e.g., Takayama 1974; Theorem 4.C.11, p. 388). Hence, the inverse $(I - \delta A)^{-1} = \frac{1}{\delta}\left(\frac{1}{\delta} I - A\right)^{-1}$, where I denotes the identity matrix, exists and is nonnegative for $\delta \in (0, 1)$, since for any nonnegative square matrix A and real number ρ the matrix $B \equiv (\rho I - A)$ has a nonnegative inverse $B^{-1} \geq 0$, if and only if ρ exceeds the Frobenius root of A (see e.g., Takayama 1974; Theorem 4.D.2, p. 392). Obviously, $\delta = 0$ implies $(I - \delta A)^{-1} = I$. Therefore, for $\delta \in [0, 1)$ we have

$$V = (I - \delta A)^{-1} J. \tag{5}$$

Define $P_{j+k} \equiv p^L\left(\lambda^{j+k}\right) - c = \lambda^{j+k-1} + L - c$ for $k \in \{0, 1, ..., K\}$, and

$$P \equiv \begin{bmatrix} P_j \\ P_{j+1} \\ \cdot \\ \cdot \\ \cdot \\ P_{j+K} \end{bmatrix} = \begin{bmatrix} \lambda^{j-1} + L - c \\ \lambda^{j} + L - c \\ \cdot \\ \cdot \\ \cdot \\ \lambda^{j+K-1} + L - c \end{bmatrix}.$$

By construction $V \geqq \frac{1}{1-\delta} P$, and at least one inequality is strict because by assumption $V\left(\lambda^{j+k}\right) > \frac{p^L\left(\lambda^{j+k}\right) - c}{1-\delta}$ for all $k \in \{1, ..., K - 1\}$ and $K \geqq 2$. Thus,

$$V \nleqq \frac{1}{1 - \delta} P. \tag{6}$$

Note that

$$AP = \begin{bmatrix} \lambda^{j-1} + L - c \\ (1-\varphi_{j+1}) \left[\lambda^{j-1} + L - c\right] + \varphi_{j+1} \left[\lambda^{j+1} + L - c\right] \\ \vdots \\ (1-\varphi_{j+k}) \left[\lambda^{j+k-2} + L - c\right] + \varphi_{j+k} \left[\lambda^{j+k} + L - c\right] \\ \vdots \\ (1-\varphi_{j+K-1}) \left[\lambda^{j+K-3} + L - c\right] + \varphi_{j+K-1} \left[\lambda^{j+K-1} + L - c\right] \\ \lambda^{j+K-1} + L - c \end{bmatrix}$$

and that for $\lambda = \lambda^{j+k},\ k \in \{1, ..., K-1\}$,

$$\varphi(\lambda) \left[p^L(\lambda^+) - c\right] + [1-\varphi(\lambda)] \left[p^L(\lambda^-) - c\right] = \left(1-\varphi_{j+k}\right) \left[\lambda^{j+k-2} + L - c\right] \\ +\varphi_{j+k} \left[\lambda^{j+k} + L - c\right].$$

This and the assumption that $F\left(\lambda^{j+k}\right) \leqq 0$ for all $k \in \{0, 1, ..., K\}$ imply $J + \frac{\delta}{1-\delta} AP - \frac{1}{1-\delta} P \leqq 0$, or $J \leqq \frac{1}{1-\delta} (I - \delta A) P$. Since $(I-\delta A)^{-1}$ is semipositive, $(I-\delta A)^{-1} J \leqq \frac{1}{1-\delta} P$. Together with (5) this gives $V \leqq \frac{1}{1-\delta} P$, which contradicts (6). This contradiction proves the lemma. □

Corollary 1 *Let* $\left\{\lambda^j, \ldots, \lambda^{j+K}\right\}$, $K \geqq 2$, *where* $\lambda^{j+k} \equiv \Pr\left(H \mid \lambda^j, k \text{ signals } s = h\right)$, *be a set of* λ*'s such that* $F\left(\lambda^{j+k}\right) \leqq 0$ *for all* $k \in \{0, ..., K\}$. *If* $V\left(\lambda^j\right) = \frac{p^L(\lambda^j)-c}{1-\delta}$ *and* $V\left(\lambda^{j+K}\right) = \frac{p^L(\lambda^{j+K})-c}{1-\delta}$, *then* $V\left(\lambda^{j+k}\right) = \frac{p^L(\lambda^{j+k})-c}{1-\delta}$ *for all* $k \in \{1, ..., K-1\}$.

Lemma 6 *Let* $\left\{\lambda^j, \lambda^{j+1}, \lambda^{j+2}\right\}$, *where* $\lambda^{j+k} \equiv \Pr\left(H|\lambda^j, k \text{ signals } s = h\right)$, *be a sequence of* λ*'s such that* $F\left(\lambda^{j+1}\right) < 0$. *If* $V\left(\lambda^{j+k}\right) = \frac{p^L(\lambda^{j+k})-c}{1-\delta}$ *for all* $k \in \{0, 2\}$, *then* $p^H\left(\lambda^{j+1}\right)$ *cannot be optimal, i.e., the price* $p\left(\lambda^{j+1}\right) = p^L\left(\lambda^{j+1}\right)$ *is uniquely optimal.*

Proof of Lemma 6 By assumption $F\left(\lambda^{j+1}\right) < 0$, and $V\left(\lambda^{j+k}\right) = \frac{p^L(\lambda^{j+k})-c}{1-\delta}$ for $k \in \{0, 2\}$. With the notation $V_{j+k} \equiv V\left(\lambda^{j+k}\right)$, $k \in \{0, 2\}$, the expected payoff from the price $p^H\left(\lambda^{j+1}\right)$ equals $[\alpha + (2\alpha - 1)(L-c)]\lambda^{j+1} + (1-\alpha)(L-c) + \delta\left(1-\varphi_{j+1}\right)V_j + \delta\varphi_{j+1}V_{j+2} = F\left(\lambda^{j+1}\right) + \frac{p^L(\lambda^{j+1})-c}{1-\delta} < \frac{p^L(\lambda^{j+1})-c}{1-\delta}$, since $F\left(\lambda^{j+1}\right) < 0$. Hence, $p^H\left(\lambda^{j+1}\right)$ is not optimal. □

Lemma 7 *If either* $c > L$ *and* $\alpha \geq \left(1-\alpha^2\right)(H-c)$, *or* $c = L$ *and* $\alpha > 1-\alpha^2$, *there exists a* $\lambda' \in \left[\bar{\lambda}_\alpha, 1\right)$ *such that*

$$F(\lambda) \begin{cases} > 0 & \text{for all } \lambda \in \left(0, \lambda'\right) \\ = 0 & \text{for } \lambda = \lambda' \\ < 0 & \text{for all } \lambda \in \left(\lambda', 1\right). \end{cases}$$

The number λ' *is strictly increasing in* δ, $\lambda' = \bar{\lambda}_\alpha$ *for* $\delta = 0$, *and* $\lambda' \to 1$ *for* $\delta \to 1$.

If $c > L$ *and* $\alpha < \left(1-\alpha^2\right)(H-c)$, *the equation* $F(\lambda) = 0$, *has at least one and at most three solutions* $\lambda \in (0,1)$*; and whenever it has exactly one solution, the results of the previous paragraph hold for* $\alpha < \left(1-\alpha^2\right)(H-c)$ *as well. For every* $c > L$ *there exists a* $\delta' \in [0,1)$ *such that* $F(\lambda) = 0$*, has exactly one solution* $\lambda \in (0,1)$ *for all* $\delta \in \left(\delta', 1\right)$.

If $c = L$ *and* $\alpha^2 = 1-\alpha$, *there exists a* $\delta^* \in (0,1)$ *such that for all* $\delta \in \left[0, \delta^*\right]$, $F(\lambda) < 0$ *for all* $\lambda \in (0,1)$. *For each* $\delta \in (\delta^*, 1)$ *there exists a* $\lambda' \in (0,1)$ *such that*

$$F(\lambda) \begin{cases} > 0 & \text{for all } \lambda \in \left(0, \lambda'\right) \\ = 0 & \text{for } \lambda = \lambda' \\ < 0 & \text{for all } \lambda \in \left(\lambda', 1\right). \end{cases}$$

The number λ' *is strictly increasing in* δ, $\lambda' \to 0$ *for* $\delta \to \delta^*$, *and* $\lambda' \to 1$ *for* $\delta \to 1$.

If either $c = L$ *and* $\alpha^2 < 1-\alpha$ *or* $c < L$, *there exists a* $\delta^* \in (0,1)$ *such that for all* $\delta \in \left[0, \delta^*\right)$, $F(\lambda) < 0$ *for all* $\lambda \in (0,1)$. *For each* $\delta \in \left[\delta^*, 1\right)$, *there exist a* $\lambda' \in (0,1)$ *and a* $\lambda'' \in \left[\lambda', 1\right)$ *such that*

$$F(\lambda) \begin{cases} > 0 & \text{for all } \lambda \in \left(\lambda', \lambda''\right) \\ = 0 & \text{for } \lambda \in \left\{\lambda', \lambda''\right\} \\ < 0 & \text{for all } \lambda \in \left(0, \lambda'\right) \cup \left(\lambda'', 1\right). \end{cases}$$

Whereas λ' *is strictly decreasing in* $\delta \in \left[\delta^*, 1\right)$, λ'' *is strictly increasing in* $\delta \in \left[\delta^*, 1\right)$. *For all* $\delta \in (\delta^*, 1)$, $\lambda' < \lambda''$; *and for* $\delta = \delta^*$, $\lambda' = \lambda''$. *For* $\delta \to 1$, $\lambda' \to 0$ *and* $\lambda'' \to 1$.

Proof of Lemma 7 For notational convenience we define $\gamma \equiv c - L$. Simple but tedious calculations show that $F(\lambda) = 0$ for $\lambda \in (0,1)$ if and only if

$$\begin{aligned} (2\alpha-1)^3 \frac{\delta}{1-\delta} \lambda (1-\lambda) = {} & \alpha^2 \left(1-\alpha-\alpha^2\right)(1-\lambda) + (1-\alpha)^3 \lambda \\ & + \frac{(1-\alpha)^4 \lambda^2}{1-\lambda} (1-\gamma) - \frac{\alpha^4 (1-\lambda)^2}{\lambda} \gamma \\ & + 2\alpha (1-\alpha)(2\alpha-1) \lambda \gamma - \alpha^2 \left(1-\alpha^2\right) \gamma, \end{aligned} \quad (7)$$

and $F(\lambda) < 0$ if and only if the left-hand side of (7) is less than the right-hand side of (7). The left-hand side of (7) is strictly concave in λ. Since $1-\gamma = H - L - [c-L] = H - c > 0$ and the second derivatives $\frac{d^2}{d\lambda^2}\left(\frac{\lambda^2}{1-\lambda}\right) = \frac{2}{(1-\lambda)^3}$ and $\frac{d^2}{d\lambda^2}\left(\frac{(1-\lambda)^2}{\lambda}\right) = \frac{2}{\lambda^3}$ are both positive, the right-hand side of (7) is strictly convex in λ if $c \le L$ and thus $\gamma \le 0$. Consequently, for $c \le L$ there are at most two solutions of (7) and thus of $F(\lambda) = 0$.

Consider the case $c > L$, i.e., $\gamma > 0$. Dividing both sides of (7) by $\lambda > 0$ gives

$$\begin{aligned}(2\alpha-1)^3\ \frac{\delta}{1-\delta}\ (1-\lambda) = {} & \frac{\alpha^2}{\lambda}\left[1-\alpha-\alpha^2-\left(1-\alpha^2\right)\gamma\right] \\ & -\alpha^2\left[1-\alpha-\alpha^2\right]+(1-\alpha)^3+\frac{(1-\alpha)^4\lambda}{1-\lambda}(1-\gamma) \\ & -\ \frac{\alpha^4(1-\lambda)^2}{\lambda^2}\gamma+2\alpha(1-\alpha)(2\alpha-1)\gamma. \end{aligned} \tag{8}$$

The left-hand side of (8) is positive and strictly decreasing in $\lambda \in (0, 1)$. For $\alpha \geq \left(1-\alpha^2\right)(H-c)$ the right-hand side of (8) is strictly increasing in $\lambda \in (0, 1)$ because $1-\alpha-\alpha^2-\left(1-\alpha^2\right)\gamma = \left(1-\alpha^2\right)(H-c)-\alpha \leq 0$. Moreover, for λ sufficiently close to zero it is negative, and for $\lambda \to 1$ it diverges to ∞. Consequently, if $c < L$ and $\alpha \geq \left(1-\alpha^2\right)(H-c)$, Eq. (8) and thus $F(\lambda) = 0$ has exactly one solution. For the same reason $F(\lambda) = 0$ has exactly one solution if $c = L$ and $\alpha > 1-\alpha^2$ because then $1-\alpha-\alpha^2-\left(1-\alpha^2\right)\gamma < 0$. Hence, whenever either $c < L$ and $\alpha \geq \left(1-\alpha^2\right)(H-c)$ or $c = L$ and $\alpha > 1-\alpha^2$, there exists a $\lambda' \in \left[\bar{\lambda}_\alpha, 1\right)$ such that

$$F(\lambda)\begin{cases} > 0 & \text{for all } \lambda \in \left(0, \lambda'\right) \\ = 0 & \text{for } \lambda = \lambda' \\ < 0 & \text{for all } \lambda \in \left(\lambda', 1\right). \end{cases}$$

For $\lambda \in (0, 1)$ the left-hand side of (8) is strictly increasing in δ and arbitrary large for $\delta \to 1$, whereas the right-hand side of (8) is independent of δ. Therefore, λ' is strictly increasing in δ and $\lambda' \to 1$ for $\delta \to 1$. For $\delta = 0$ we get $F(\lambda) = \varphi(\lambda)\left[p^H(\lambda)-c\right]-p^L(\lambda)-c$ and thus $\lambda' = \bar{\lambda}_\alpha$ by definition of $\bar{\lambda}_\alpha$. This proves the first paragraph of Lemma 7.

If $c = L$ and $\alpha^2 = 1-\alpha$, the right-hand side of (8) reduces to $(1-\alpha)^3+\frac{(1-\alpha)^4\lambda}{1-\lambda}$, which is positive and strictly increasing in $\lambda \in (0, 1)$. Define $\delta^* \in (0, 1)$ as the (unique) solution of $(2\alpha-1)^3\ \frac{\delta}{1-\delta} = (1-\alpha)^3$. Then $F(\cdot) < 0$ for all $\delta \in \left[0, \delta^*\right]$. For all $\delta \in (\delta^*, 1)$ the equation $F(\lambda) = 0$ has a unique solution λ', which is strictly increasing in δ, and $\lambda' \to 1$ for $\delta \to 1$. This implies the third paragraph of Lemma 7.

Next, we show that if $c > L$ and $\alpha < \left(1-\alpha^2\right)(H-c)$, Eq. (8) and thus $F(\lambda) = 0$ has at least one and at most three solutions. With the definition

$$\begin{aligned} G(\lambda) \equiv {} & (2\alpha-1)^3\ \frac{\delta}{1-\delta}\lambda+\frac{\alpha^2}{\lambda}\left[\left(1-\alpha^2\right)(H-c)-\alpha\right]-\alpha^2\left[(1-\alpha)-\alpha^2\right] \\ & +(1-\alpha)^3+\frac{(1-\alpha)^4\lambda}{1-\lambda}(1-\gamma)-\frac{\alpha^4(1-\lambda)^2}{\lambda^2}\gamma+2\alpha(1-\alpha)(2\alpha-1)\gamma, \end{aligned}$$

(8) is equivalent to $G(\lambda) = (2\alpha-1)^3\ \frac{\delta}{1-\delta}$. We prove that $G(\lambda)$ can be partitioned into a strictly concave and a strictly convex part, which implies that $G(\lambda) = (2\alpha-1)^3$

$\frac{\delta}{1-\delta}$ can have at most three solutions. Differentiating $G(\lambda)$ gives

$$G'(\lambda) = (2\alpha-1)^3 \frac{\delta}{1-\delta} - \frac{\alpha^2}{\lambda^2}\left[\left(1-\alpha^2\right)(H-c)-\alpha\right] + \frac{(1-\alpha)^4(1-\gamma)}{(1-\lambda)^2} + \frac{2\alpha^4\gamma(1-\lambda)}{\lambda^3}$$

and

$$\begin{aligned} G''(\lambda) &= \frac{2\alpha^2}{\lambda^3}\left[\left(1-\alpha^2\right)(H-c)-\alpha\right] + \frac{2(1-\alpha)^4(1-\gamma)}{(1-\lambda)^3} - \frac{2\alpha^4\gamma(3-2\lambda)}{\lambda^4} \\ &= \frac{2\alpha^2}{\lambda^3}\left[\left(1-\alpha^2\right)(H-c)-\alpha-\frac{\alpha^2\gamma(3-2\lambda)}{\lambda}\right] + \frac{2(1-\alpha)^4(1-\gamma)}{(1-\lambda)^3}. \end{aligned}$$

Since $2\lambda < 3$ and $\gamma < 1$, $G''(\lambda)$ increases whenever $\left(1-\alpha^2\right)(H-c)-\alpha-\frac{\alpha^2\gamma(3-2\lambda)}{\lambda} < 0$. Moreover, $G''(0) = -\infty$ and $G''(1) = \infty$. Furthermore, $G''(\lambda) \leq 0$ implies $\left(1-\alpha^2\right)(H-c)-\alpha-\frac{\alpha^2\gamma(3-2\lambda)}{\lambda} < 0$. Consequently, $G''(\lambda)$ has a positive slope for any λ that satisfies $G''(\lambda) = 0$, and therefore there exists some $\bar{\lambda} \in (0,1)$ such that $G''\left(\bar{\lambda}\right) = 0$, $G''(\lambda) < 0$ for $\lambda < \bar{\lambda}$ and $G''(\lambda) > 0$ for $\lambda > \bar{\lambda}$. That is, $G(\lambda)$ is strictly concave for $\lambda < \bar{\lambda}$ and strictly convex for $\lambda > \bar{\lambda}$. It follows that the equation $G(\lambda) = (2\alpha-1)^3 \frac{\delta}{1-\delta}$, i.e. (8), and thus $F(\lambda) = 0$ can have at most three solutions. Moreover, since $G(0) = -\infty$ and $G(1) = \infty$, there is at least one solution. If $F(\lambda) = 0$ has exactly one solution $\lambda \in (0,1)$, the proof of the first paragraph of Lemma 7 applies as well. Finally, the right-hand side of (8) increases in λ if λ is sufficiently large (whereas the left-hand side decreases in λ); and $(2\alpha-1)^3(1-\lambda)\frac{\delta}{1-\delta} \to \infty$ for $\delta \to 1$, whereas the right-hand side of (8) is independent of δ. Therefore, if δ is sufficiently close to 1, any solution of (8) will be in the range where the right-hand side of (8) increases in λ. Consequently, $F(\lambda) = 0$ has exactly one solution whenever δ is sufficiently close to 1. This proves the second paragraph of Lemma 7.

For the proof of the last paragraph of Lemma 7 recall that we have already shown that for $c \leq L$ there are at most two solutions of (7) and thus of $F(\lambda) = 0$. The (strictly concave) left-hand side of (7) is zero for $\lambda = 0$ and for $\lambda = 1$. For $\lambda = 0$ the (strictly convex) right-hand side of (7) is infinite if $\gamma = c - L < 0$, and $\alpha^2\left(1-\alpha-\alpha^2\right) > 0$ if $c = L$ and $\alpha^2 < 1-\alpha$. For $\lambda = 1$ the right-hand side of (7) is infinite, since $\gamma < 1$. Consequently, $F(\lambda) < 0$ for all $\lambda \in (0,1)$ if δ is sufficiently small, and there is exactly one δ, denoted by δ^*, such that $F(\lambda) = 0$ has exactly one solution. For all $\delta > \delta^*$, $F(\lambda) = 0$ has two different solutions, $\lambda' \in (0,1)$ and $\lambda'' \in \left(\lambda', 1\right)$. Since the left-hand side of (7) increases in δ (whereas the the right-hand side of (7) is independent of δ), λ' is strictly decreasing in $\delta \in \left[\delta^*, 1\right)$ and λ'' is strictly increasing in $\delta \in \left[\delta^*, 1\right)$. Moreover, for $\delta \to 1$ the left-hand side of (7) diverges and thus $\lambda' \to 0$ and $\lambda'' \to 1$ for $\delta \to 1$. □

Lemma 8 *Let* $\{\lambda^j, \lambda^{j-1}, ...\}$, *where* $\lambda^{j-k} \equiv \Pr\left(H|\lambda^j, k \text{ signals } s = l\right)$, *be a sequence of* λ*'s such that* $\lambda^j \in (0,1)$ *and* $F\left(\lambda^{j-k}\right) \leqq 0$ *for all* $k \in \{0, 1, ...\}$. *If* $V\left(\lambda^j\right) = \frac{p^L(\lambda^j)-c}{1-\delta}$, *then* $V\left(\lambda^{j-k}\right) = \frac{p^L(\lambda^{j-k})-c}{1-\delta}$ *for all* $k \in \{1, 2, ...\}$.

Proof of Lemma 8 Lemma 5 excludes the case where $V\left(\lambda^{j-k}\right) > \frac{p^L(\lambda^{j-k})-c}{1-\delta}$ for some, but not all $k \in \{1, 2, ...\}$. Therefore, we only have to show that it is not possible that $V\left(\lambda^{j-k}\right) > \frac{p^L(\lambda^{j-k})-c}{1-\delta}$ for all $k \in \{1, 2, ...\}$. Assume the contrary, i.e., $V\left(\lambda^{j-k}\right) > \frac{p^L(\lambda^{j-k})-c}{1-\delta}$ for all $k \in \{1, 2, ...\}$ and $V\left(\lambda^j\right) = \frac{p^L(\lambda^j)-c}{1-\delta}$. With the notation $V_{j-k} \equiv V\left(\lambda^{j-k}\right)$, $k \in \{0, 1, ...\}$, we have

$$V_{j-k} = \begin{cases} \lambda^{j-1} + L - c + \delta V_j & \text{for } k = 0 \\ [\alpha + (2\alpha - 1)(L - c)]\,\lambda^{j-k} & \\ \quad + (1-\alpha)(L-c) & \\ \quad + \delta\varphi_{j-k} V_{j-k+1} + \delta\left(1 - \varphi_{j-k}\right) V_{j-k-1} & \text{for } k \in \{1, 2, ...\} \end{cases} \tag{9}$$

Let the infinite matrix C be defined by

$$C \equiv \begin{bmatrix} 1 & 0 & 0 & 0 & . & . & & & 0 & 0 & 0 & & . & . & . \\ \varphi_{j-1} & 0 & 1-\varphi_{j-1} & 0 & . & . & & & 0 & 0 & 0 & & . & . & . \\ 0 & \varphi_{j-2} & 0 & 1-\varphi_{j-2} & . & . & & & 0 & 0 & 0 & & . & . & . \\ & & & & & & & & & & & & & & \\ - & - & - & - & - & - & - & - & - & - & - & & - & - & - \\ 0 & 0 & 0 & 0 & & . & . & 0 & \varphi_{j-k} & 0 & 1-\varphi_{j-k} & & 0 & . & . \\ - & - & - & - & - & - & - & - & - & - & - & & - & - & - \\ . & . & . & . & . & . & . & . & . & . & . & & . & . & . \\ . & . & . & . & . & . & . & . & . & . & . & & . & . & . \\ . & . & . & . & . & . & . & . & . & . & . & & . & . & . \end{bmatrix}.$$

With I denoting the infinite identity matrix, this gives

$$I - \delta C \equiv \begin{bmatrix} 1-\delta & 0 & 0 & 0 & 0 & 0\ldots \\ -\delta\varphi_{j-1} & 1 & -\delta\left(1-\varphi_{j-1}\right) & 0 & 0 & 0\ldots \\ 0 & -\delta\varphi_{j-2} & 1 & -\delta\left(1-\varphi_{j-2}\right) & 0 & 0\ldots \\ 0 & 0 & -\delta\varphi_{j-3} & 1 & -\delta\left(1-\varphi_{j-3}\right) & 0\ldots \\ . & . & . & . & . & \ldots \\ . & . & . & . & . & \ldots \\ . & . & . & . & . & \ldots \end{bmatrix}.$$

Define

$$V \equiv \begin{bmatrix} V_j \\ V_{j-1} \\ . \\ . \\ . \\ . \end{bmatrix}, \quad J \equiv \begin{bmatrix} \lambda^{j-1} + L - c \\ [\alpha + (2\alpha-1)(L-c)]\,\lambda^{j-1} + (1-\alpha)(L-c) \\ [\alpha + (2\alpha-1)(L-c)]\,\lambda^{j-2} + (1-\alpha)(L-c) \\ . \\ . \\ . \end{bmatrix},$$

$$P \equiv \begin{bmatrix} P_j \\ P_{j-1} \\ . \\ . \\ . \\ . \end{bmatrix} \equiv \begin{bmatrix} \lambda^{j-1} + L - c \\ \lambda^{j-2} + L - c \\ . \\ . \\ . \\ . \end{bmatrix}.$$

With these definitions and (9) we get $V = J + \delta C V$, or

$$J = (I - \delta C) V. \tag{10}$$

The assumption $F\left(\lambda^{j-k}\right) \leqq 0$ for all $k \in \{1, 2, ...\}$ implies $\frac{\delta}{1-\delta} CP + J \leqq \frac{1}{1-\delta} P$, or

$$J \leqq \frac{1}{1-\delta} (I - \delta C) P. \tag{11}$$

From (10) and (11) we get

$$(I - \delta C) V \leqq \frac{1}{1-\delta} (I - \delta C) P. \tag{12}$$

Define the infinite vector $x = (x_1, x_2, ...) \gg 0$ by $x_1 = \frac{1}{1-\delta}$ and $x_k = 1$ for $k \in \{2, 3, ...\}$. Then the infinite vector $z \equiv x (I - \delta C)$ has the elements $z_1 = 1 - \delta\varphi_{j-1} \in (0, 1]$, $z_2 = 1 - \delta\varphi_{j-2} \in (0, 1]$, $z_k = 1 - \delta\left(1 - \varphi_{j-k+1} + \varphi_{j-k-1}\right) \in (0, 1]$ for $k \in \{3, 4, ...\}$, where $1 - \delta\left(1 - \varphi_{j-k+1} + \varphi_{j-k-1}\right) \in (0, 1]$ follows from $\varphi_{j-k+1} - \varphi_{j-k-1} \in (0, 1)$ and $\delta \in [0, 1)$. With the help of some calculations,

$$\frac{\lambda^{j-k-1}}{\lambda^{j-k}} = 1 - \frac{(2\alpha - 1)\left(1 - \lambda^{j-k}\right)}{(1-\alpha)\lambda^{j-k} + \alpha\left(1 - \lambda^{j-k}\right)} < 1 - \frac{(2\alpha - 1)\left(1 - \lambda^{j}\right)}{\alpha} \in (0, 1),$$

and thus $\sum_{k=0}^{\infty} \lambda^{j-k-1}$ converges. Since $zP = \sum_{k=0}^{\infty} z_k \lambda^{j-k-1}$ and $z_k \in (0, 1]$ for all $k \in \{1, 2, ...\}$, $0 < zP \leq \sum_{k=0}^{\infty} \lambda^{j-k-1} < \infty$. Multiplying both sides of (12) by $x \gg 0$ gives

$$zV \leqq \frac{1}{1-\delta} zP < \infty. \tag{13}$$

However, the proof's assumption that $V\left(\lambda^{j-k}\right) > \frac{p^L\left(\lambda^{j-k}\right) - c}{1-\delta}$ for all $k \in \{1, 2, ...\}$, together with $V\left(\lambda^{j}\right) = \frac{p^L\left(\lambda^{j}\right) - c}{1-\delta}$ and $z \gg 0$, implies $zV > \frac{1}{1-\delta} zP$. This contradiction proves the lemma. □

A.2 Proof of the propositions

Proof of Proposition 1 We know from Lemma 2 that given δ, $V(\lambda) = \frac{1}{1-\delta}\left[p^L(\lambda) - c\right]$, if λ is sufficiently close to 1. Define λ^{**} as the smallest $\mu \in \Lambda(\lambda_1)$ such that for all $\lambda \geq \mu$, $\lambda \in \Lambda(\lambda_1)$, it holds that $V(\lambda) = \frac{1}{1-\delta}\left[p^L(\lambda) - c\right]$. Since $V(\lambda) > \frac{1}{1-\delta}\left[p^L(\lambda) - c\right]$ for all $\lambda < \bar{\lambda}_\alpha$, $\lambda \in \Lambda(\lambda_1)$, λ^{**} exists and $\lambda^{**} \geq \bar{\lambda}_\alpha$. We show by contradiction that $V(\lambda) > \frac{1}{1-\delta}\left[p^L(\lambda) - c\right]$ for all $\lambda < \lambda^{**}$, $\lambda \in \Lambda(\lambda_1)$. Assume that for some $\lambda < \lambda^{**}$, $\lambda \in \Lambda(\lambda_1)$, it holds that $V(\lambda) = \frac{1}{1-\delta}\left[p^L(\lambda) - c\right]$, and let λ^j be the largest such $\lambda < \lambda^{**}$. This implies $F\left(\lambda^j\right) \leq 0$ (otherwise $p^H\left(\lambda^j\right)$ would be uniquely optimal). Hence by Lemma 7, $\lambda' \leq \lambda^j$ and $F(\lambda) < 0$ for all

$\lambda > \lambda^j$, $\lambda \in \Lambda(\lambda_1)$. Since by construction $\lambda^j < \lambda^{**}$ and $V(\lambda) > \frac{1}{1-\delta}\left[p^L(\lambda) - c\right]$ for all $\lambda \in \Lambda(\lambda_1) \cap \left(\lambda^j, \lambda^{**}\right)$, and since the definition of λ^{**} implies $V\left(\lambda^-\right) > \frac{1}{1-\delta}\left[p^L\left(\lambda^-\right) - c\right]$ for $\lambda = \lambda^{**}$, there exists a set $\left\{\lambda^j, ..., \lambda^{j+K}\right\}$, $K \geq 2$, that satisfies the assumptions of Lemma 5 and, in addition, $V\left(\lambda^K\right) = \frac{p^L(\lambda^K)-c}{1-\delta}$. Because of this, Lemma 5 implies $V\left(\lambda^j\right) > \frac{p^L(\lambda^j)-c}{1-\delta}$, whereas by construction $V\left(\lambda^j\right) = \frac{p^L(\lambda^j)-c}{1-\delta}$. This contradiction proves that $V(\lambda) > \frac{1}{1-\delta}\left[p^L(\lambda) - c\right]$ and thus that $p^H(\lambda)$ is uniquely optimal for all $\lambda < \lambda^{**}$, $\lambda \in \Lambda(\lambda_1)$. Moreover, because of Lemma 7 and $\lambda^{**} \geq \lambda'$ (which follows from $F(\lambda^{**}) \leq 0$), Lemma 6 implies that only $p^L(\lambda)$ is optimal for all $\lambda > \lambda^{**}$, $\lambda \in \Lambda(\lambda_1)$. The rest of Proposition 1 follows from the first paragraph of Lemma 7. □

Proof of Proposition 2 From Lemma 2 we know that given δ, the pooling price $p^L(\lambda)$ is uniquely optimal whenever λ is sufficiently close to 1. By Lemma 7, there exists a $\delta^* \in (0, 1)$ such that for all $\delta \in \left[0, \delta^*\right]$, $F(\lambda) < 0$ for all $\lambda \in \Lambda(\lambda_1)$, and thus Lemma 8 implies that $p^L(\lambda)$ is optimal for all $\lambda \in \Lambda(\lambda_1)$. Since for any $\varepsilon > 0$, $(0, \varepsilon) \cap \Lambda(\lambda_1) \neq \emptyset$, the rest of the proof of Proposition 2 is analogous to the proof of Proposition 1 (note that for $\alpha^2 = 1 - \alpha$, $\bar{\lambda}_\alpha = 0$). □

Proof of Proposition 3 From Lemma 2 we know that given δ, the pooling price $p^L(\lambda)$ is uniquely optimal whenever λ is sufficiently close to 1. By Lemma 7, there exists a $\delta^* \in (0, 1)$ such that for all $\delta \in \left[0, \delta^*\right)$, $F(\lambda) < 0$ for all $\lambda \in \Lambda(\lambda_1)$. Therefore Lemma 8 implies that $p^L(\lambda)$ is optimal for all $\delta \in (0, \delta^*)$. Although $F(\lambda) > 0$ for some $\lambda \in (0, 1)$ if $\delta > \delta^*$, these λ's may not be elements of $\Lambda(\lambda_1)$ and thus $F(\lambda) < 0$ for all $\lambda \in \Lambda(\lambda_1)$ may still hold. Hence define $\delta^{**} \geq \delta^* > 0$ as the supremum of δ in the set $\{\delta | F(\lambda) < 0$ for all $\lambda \in \Lambda(\lambda_1)\}$. Together with Lemma 6 the definition of δ^{**} implies that for all $\delta \in (0, \delta^{**})$, $p^L(\lambda)$ is uniquely optimal for all $\lambda \in \Lambda(\lambda_1)$.

Consider now the case $\delta \in (\delta^{**}, 1)$. From Lemma 7 we know that if $\delta \in (\delta^{**}, 1)$ is sufficiently large (and therefore λ' and λ'' of Lemma 7 are sufficiently close to 0 and 1, respectively), there exists a $\lambda \in \Lambda(\lambda_1)$ such that $V(\lambda) > \frac{1}{1-\delta}\left[p^L(\lambda) - c\right]$. Let $\delta^{***} \geq \delta^{**}$ denote the supremum of δ in the set $\left\{\delta \mid V(\lambda) = \frac{1}{1-\delta}\left[p^L(\lambda) - c\right]\right.$ for all $\lambda \in \Lambda(\lambda_1)\}$. Continuity implies $V(\lambda) = \frac{1}{1-\delta}\left[p^L(\lambda) - c\right]$ for all $\lambda \in \Lambda(\lambda_1)$, if $\delta = \delta^{***}$. For any $\delta \in (\delta^{***}, 1)$ define λ^* as the largest $\mu \in \Lambda(\lambda_1)$ such that for all $\lambda \leq \mu, \lambda \in \Lambda(\lambda_1)$, it holds that $V(\lambda) = \frac{1}{1-\delta}\left[p^L(\lambda) - c\right]$. If such a μ does not exist, $\lambda^* \equiv 0$. Define λ^{**} as the smallest $\mu \in \Lambda(\lambda_1)$ such that for all $\lambda \geq \mu$, $\lambda \in \Lambda(\lambda_1)$, it holds that $V(\lambda) = \frac{1}{1-\delta}\left[p^L(\lambda) - c\right]$. These definitions and the definition of δ^{***} imply $\lambda^* < \lambda^{**}$. Since $F(\lambda^*) \leq 0$ if $\lambda^* > 0$, and $F(\lambda^{**}) \leq 0$, Lemma 7 implies $F(\lambda) < 0$ for all $\lambda \in \Lambda(\lambda_1)$ that satisfy either $\lambda < \lambda^*$ or $\lambda > \lambda^{**}$. If we can show that $\lambda^* > 0$, the part of Proposition 3 that relates to $\delta \in (\delta^{***}, 1)$ follows from Corollary 1 and Lemma 8.

Hence, we prove that indeed $\lambda^* > 0$. First, we show that if $\lambda^* = 0$, then $p^H(\lambda)$ is uniquely optimal for all $\lambda \in (0, \lambda^{**})$. Corollary 1 implies that $p^L(\lambda)$ is optimal for $\lambda \in (\lambda^{**}, 1)$. Assume that $p^L\left(\lambda^0\right)$ is optimal for some $\lambda^0 \in (0, \lambda^{**}) \cap \Lambda(\lambda_1)$. This implies $F\left(\lambda^0\right) \leq 0$ and thus $\lambda^0 \leq \lambda'$ because of Corollary 1, Lemma 7, and the definition of λ^{**}. Consequently, $F(\lambda) \leq 0$ for all $\lambda \in \left(0, \lambda^0\right)$. Hence by Lemma 8,

$p^L(\lambda)$ is optimal for all $\lambda \in (0, \lambda^0) \cap \Lambda(\lambda_1)$. It follows that $\lambda^* \geq \lambda^0 > 0$, which contradicts $\lambda^* = 0$. Therefore, $\lambda^* = 0$ implies that $p^H(\lambda)$ is uniquely optimal for all $\lambda \in (0, \lambda^{**}) \cap \Lambda(\lambda_1)$. Thus, if $\lambda^j \in (0, \lambda^{**}) \cap \Lambda(\lambda_1)$ it holds (cf. (9) with $L - c = 0$) that $V_{j-k} = \alpha\lambda^{j-k} + \delta\varphi_{j-k}V_{j-k+1} + \delta\left(1 - \varphi_{j-k}\right)V_{j-k-1}$ for all $k \in \{1, 2, ...\}$. Hence after some calculations,

$$\frac{V_{j-k} - V_{j-k-1}}{\lambda^{j-k} - \lambda^{j-k-1}} = \frac{\alpha}{\theta_{j-k}} + \delta\frac{\varphi_{j-k}}{\theta_{j-k}}\frac{V_{j-k+1} - V_{j-k}}{\lambda^{j-k} - \lambda^{j-k-1}} + \delta\frac{1 - \varphi_{j-k-1}}{\theta_{j-k}}\frac{V_{j-k-1} - V_{j-k-2}}{\lambda^{j-k} - \lambda^{j-k-1}}$$

for $k \in \{1, 2, ...\}$, where θ_{j-k} is given by $\theta_{j-k} \equiv 1 - \delta\left(\varphi_{j-k} - \varphi_{j-k-1}\right)$. If the limit $D_H \equiv \lim_{k\to\infty}\frac{V_{j-k}-V_{j-k-1}}{\lambda^{j-k}-\lambda^{j-k-1}}$ exists, $D_H = \alpha + \delta(1-\alpha)D_H + \alpha\delta D_H$ (since $\lim_{k\to\infty}\varphi_{j-k} = 1-\alpha$), which gives $D_H = \frac{\alpha}{1-\delta}$. We show that the limit D_H does, in fact, exist. First we prove that $\frac{V_{j-k}-V_{j-k-1}}{\lambda^{j-k}-\lambda^{j-k-1}}$ is decreasing in k. This follows because $V_{j-k} = \alpha\lambda^{j-k} + \delta\varphi_{j-k}V_{j-k+1} + \delta\left(1-\varphi_{j-k}\right)V_{j-k-1}$ and $\alpha\lambda^{j-k} < \lambda^{j-k-1}$ (which holds because of $\alpha^2 < 1-\alpha$) imply

$$V_{j-k} - \delta V_{j-k} < \lambda^{j-k-1} + \delta\left[\varphi_{j-k}V_{j-k+1} + \left(1-\varphi_{j-k}\right)V_{j-k-1} - V_{j-k}\right];$$

hence, because $V_{j-k} \geq \frac{p^L\left(\lambda^{j-k}\right)-c}{1-\delta} = \frac{\lambda^{j-k-1}}{1-\delta}$,

$$0 \leq V_{j-k} - \frac{\lambda^{j-k-1}}{1-\delta} < \frac{\delta}{1-\delta}\left[\varphi_{j-k}V_{j-k+1} + \left(1-\varphi_{j-k}\right)V_{j-k-1} - V_{j-k}\right].$$

Therefore, $\varphi_{j-k}V_{j-k+1} + \left(1-\varphi_{j-k}\right)V_{j-k-1} > V_{j-k}$. Since $\varphi_{j-k}\lambda^{j-k+1} + \left(1-\varphi_{j-k}\right)\lambda^{j-k-1} = \lambda^{j-k}$, this implies $\frac{V_{j-k+1}-V_{j-k}}{\lambda^{j-k+1}-\lambda^{j-k}} > \frac{V_{j-k}-V_{j-k-1}}{\lambda^{j-k}-\lambda^{j-k-1}}$. Consequently, $\frac{V_{j-k}-V_{j-k-1}}{\lambda^{j-k}-\lambda^{j-k-1}}$ is decreasing in k. Moreover, it is easy to check that $V_{j-k} \geq V_{j-k-1}$ and thus $\frac{V_{j-k}-V_{j-k-1}}{\lambda^{j-k}-\lambda^{j-k-1}} \geq 0$. As $\frac{V_{j-k}-V_{j-k-1}}{\lambda^{j-k}-\lambda^{j-k-1}}$ is decreasing in k and bounded from below, $D_H \equiv \lim_{k\to\infty}\frac{V_{j-k}-V_{j-k-1}}{\lambda^{j-k}-\lambda^{j-k-1}}$ exists and thus $D_H = \frac{\alpha}{1-\delta}$. Consider now the slopes of the seller's payoffs from herding, i.e.,

$$\frac{1}{1-\delta}\frac{p^L\left(\lambda^{j-k}\right) - p^L\left(\lambda^{j-k-1}\right)}{\lambda^{j-k} - \lambda^{j-k-1}} = \frac{1}{1-\delta}\frac{\lambda^{j-k-1} - \lambda^{j-k-2}}{\lambda^{j-k} - \lambda^{j-k-1}} = \frac{1}{1-\delta}\frac{1 - \frac{1-\alpha}{(1-\alpha)\lambda^{j-k-1}+\alpha\left(1-\lambda^{j-k-1}\right)}}{\frac{\alpha}{\alpha\lambda^{j-k-1}+(1-\alpha)\left(1-\lambda^{j-k-1}\right)} - 1}.$$

The respective limit is $D_L \equiv \lim_{k\to\infty}\left[\frac{1}{1-\delta}\frac{p^L\left(\lambda^{j-k}\right)-p^L\left(\lambda^{j-k-1}\right)}{\lambda^{j-k}-\lambda^{j-k-1}}\right] = \frac{1}{1-\delta}\frac{1-\frac{1-\alpha}{\alpha}}{\frac{\alpha}{1-\alpha}-1} = \frac{1}{1-\delta}\frac{1-\alpha}{\alpha} > \frac{\alpha}{1-\delta} = D_H$ since $1-\alpha > \alpha^2$. This and $\lim_{k\to\infty}\frac{p^L\left(\lambda^{j-k}\right)-c}{1-\delta} = 0 = \lim_{k\to\infty}V_{j-k}$ imply that $V_{j-k} < \frac{p^L\left(\lambda^{j-k}\right)-c}{1-\delta}$ for large k (i.e., small λ^{j-k}), since from

$D_L > D_H$ it follows that $\frac{1}{1-\delta}\left[p^L(\lambda) - c\right]$ is steeper than $V(\lambda)$ for all sufficiently small λ's. This contradicts that the separating price $p^H\left(\lambda^{j-k}\right)$ is uniquely optimal for all $\lambda = \lambda^{j-k} \in (0, \lambda^{**})$, and thus proves $\lambda^* > 0$.

Finally, consider the case $\delta \in \left[\delta^{**}, \delta^{***}\right]$. The definition of δ^{***} implies that for all $\delta \in \left[\delta^{**}, \delta^{***}\right]$, $V(\lambda) = \frac{1}{1-\delta}\left[p^L(\lambda) - c\right]$ for all $\lambda \in \Lambda(\lambda_1)$ and therefore $p^L(\lambda)$ is optimal for all $\lambda \in \Lambda(\lambda_1)$. Moreover, the definitions of δ^{**} and δ^{***}, respectively, imply that $F(\lambda) = 0$ for at least one $\lambda \in \Lambda(\lambda_1)$, and therefore $p^H(\lambda)$ is also optimal for at least one $\lambda \in \Lambda(\lambda_1)$. □

Proof of Proposition 4 First, we show that given any $\delta \in [0, 1)$ the seller always charges the pooling price for sufficiently low λ's. Since $p^L(\lambda) > L > c$ for all $\lambda \in (0, 1)$, the seller never exits the market. In any period τ the seller's expected immediate profit is bounded by

$$\max\left\{\lambda_\tau^- + L - c, \alpha\lambda_\tau + [\alpha\lambda_\tau + (1-\alpha)(1-\lambda_\tau)](L-c)\right\} < \lambda_\tau + L - c.$$

Hence, in any period t the seller's expected future payoff is bounded by $\frac{\delta[\lambda_t + L - c]}{1-\delta}$ since the stochastic process $\{\lambda_\tau\}_{\tau=t}^{\infty}$ is a martingale. Thus, in period t the seller's expected payoff from the separating price is bounded by $\alpha\lambda_t + [\alpha\lambda_t + (1-\alpha)(1-\lambda_t)](L-c) + \frac{\delta}{1-\delta}[\lambda_t + L - c]$, whereas the payoff from the pooling price is $\frac{1}{1-\delta}\left[\lambda_t^- + L - c\right]$. The difference in the seller's expected payoff from the pooling price and the separating price exceeds $\lambda_t^- + L - c - \alpha\lambda_t - [\alpha\lambda_t + (1-\alpha)(1-\lambda_t)](L-c) - \frac{\delta}{1-\delta}\left(\lambda_t - \lambda_t^-\right)$, which converges to $\alpha(L-c) > 0$ for $\lambda_t \to 0$. Therefore, the pooling price is optimal whenever λ is sufficiently low. The rest of the first part of the proposition follows from Lemma 7, analogously to the proof of Proposition 3.

The second part of the proposition follows directly from Lemma 7 (as explained in the proof of Proposition 3, $\delta^{**} \geq \delta^* > 0$). For the proposition's third part notice that the difference in the seller's expected immediate profit from the separating and the pooling price, respectively, is

$$\begin{aligned}\Delta(\alpha, \lambda_t) &\equiv \alpha\lambda_t + [\alpha\lambda_t + (1-\alpha)(1-\lambda_t)](L-c) - \left(\lambda_t^- + L - c\right)\\ &= \alpha\lambda_t - \frac{(1-\alpha)\lambda_t}{(1-\alpha)\lambda_t + \alpha(1-\lambda_t)} - [(1-\alpha)\lambda_t + \alpha(1-\lambda_t)](L-c).\end{aligned}$$

For $\lambda_t = 0$ and $\lambda_t = 1$, respectively, $\Delta(\alpha, \lambda_t) < 0$; and for $\alpha = \lambda_t \in \Lambda(\lambda_1)$ we get $\Delta(\alpha, \lambda_t) = \Delta(\alpha, \alpha) = \alpha^2 - \frac{1}{2} - 2\alpha(1-\alpha)(L-c)$, which is positive for $\alpha < 1$ sufficiently close to 1. Therefore, and since $\Delta(\alpha, \lambda_t)$ is strictly concave in λ_t (because λ_t^- is a strictly convex function of λ_t whereas $\alpha\lambda_t - [1 - \alpha\lambda_t - (1-\alpha)(1-\lambda_t)](L-c)$ is linear in λ_t), there exist μ and $\mu' > \mu$, such that for a sufficiently large fixed α it holds that $\Delta(\alpha, \lambda_t) > 0$ for all $\lambda_t \in \left[\mu, \mu'\right]$. Moreover, if $\Delta(\alpha, \alpha) > 0$ and $\alpha = \lambda_t$ is sufficiently close to 1, $\Delta\left(\alpha, \lambda_t^-\right)$ or $\Delta\left(\alpha, \lambda_t^+\right)$, or both, will be positive as well. Hence for $\lambda_t = \alpha$ at least $\lambda_t^- \in \left[\mu, \mu'\right]$ or $\lambda_t^+ \in \left[\mu, \mu'\right]$. Let μ^* be the the smallest and μ^{**} the largest element of $\left[\mu, \mu'\right] \cap \Lambda(\lambda_1)$. For α sufficiently large, μ and μ' can be chosen such that $\mu^{**} > \mu^*$. Because of $\Delta(\alpha, \lambda_t) > 0$ for all $\lambda_t \in \left[\mu^*, \mu^{**}\right]$, Lemma 1 implies that for all $\lambda \in \left[\mu^*, \mu^{**}\right]$ the separating price is uniquely optimal regardless of δ. This proves that there exist an α and associated $\mu^* \in \Lambda(\lambda_1)$, $\mu^{**} \in \Lambda(\lambda_1)$, $0 < \mu^* < \mu^{**} < 1$, such that regardless of δ the separating price is uniquely optimal for all $\lambda \in [\mu^*, \mu^{**}]$. Note that μ^* and μ^{**} depend on α. Thus, we write $\mu^*(\alpha)$ and $\mu^{**}(\alpha)$ for the following argument. Since $\frac{d\Delta(\alpha,\alpha)}{d\alpha} = 2\alpha + 2(2\alpha - 1)(L - c) > 0$, $\Delta(\alpha, \alpha) > 0$ implies $\Delta(\alpha', \alpha') > 0$ for all $\alpha' > \alpha$. Therefore, the previous argument implies that if for some α the separating price is uniquely optimal regardless of δ for all $\lambda \in [\mu^*(\alpha), \mu^{**}(\alpha)]$, then for all $\alpha' \in (\alpha, 1)$ the separating price is uniquely optimal regardless of δ for all $\lambda \in [\mu^*(\alpha'), \mu^{**}(\alpha')]$, where $0 < \mu^*(\alpha') < \mu^{**}(\alpha') < 1$ and $\mu^*(\alpha') \in \Lambda(\lambda_1)$, $\mu^{**}(\alpha') \in \Lambda(\lambda_1)$. Since above we have shown the existence of such an α, the proposition follows. □

Proof of Proposition 5 First, we show that there exists a $\lambda^E > 0$ such that for all $\lambda \in (0, \lambda^E)$ it is uniquely optimal for the seller to exit the market, whereas for all $\lambda \in (\lambda^E, 1)$ it is uniquely optimal for the seller to stay in the market. In any period τ the seller's expected immediate profit is bounded by

$$\max\left\{\lambda_\tau^- + L - c, \alpha\lambda_\tau + [\alpha\lambda_\tau + (1-\alpha)(1-\lambda_\tau)](L-c)\right\} < \lambda_\tau - [c - L].$$

This implies (because the stochastic process $\{\lambda_\tau\}_{\tau=t}^\infty$ is a martingale) that if the seller never exits the market, her expected payoff conditional on λ_t is bounded by $\frac{\lambda_t - [c-L]}{1-\delta}$. Consequently, whenever $\lambda_t < c - L$ the seller is better off exiting the market than staying in the market forever. Since the seller's expected profits increase in λ, there must exist a critical $\lambda^E \in \Lambda(\lambda_1)$ such that exit is optimal for $\lambda = \lambda^E$ and uniquely optimal for all $\lambda \in (0, \lambda^E)$, whereas for all $\lambda \in (\lambda^E, 1)$ it is uniquely optimal for the seller to stay in the market. As shown in Sect. 5.3, $c > L$ implies that there are exactly two attainable probabilities $\lambda \in \Lambda(\lambda_1)$ that satisfy $p^L(\lambda) \le c < p^H(\lambda)$. From this it follows that $(\lambda^E, \lambda^{**}) \cap \Lambda(\lambda_1)$ contains at least two elements. Since exit is optimal for $\lambda = \lambda^E$, it must hold that $p^L(\lambda^E) < p^H(\lambda^E) \le c$ and thus the pooling price $p^L(\lambda^E)$ cannot be optimal at $\lambda = \lambda^E$. Finally, note that $\lambda^E < \bar{\lambda}_\alpha$ because $p^L(\bar{\lambda}_\alpha) - c > 0$. Because of this, the rest of the proof is analogous to the proof of Proposition 1. □

Proof of Proposition 6 For the parts of the proposition that relate to λ^E and $\lambda^E < \bar{\lambda}_\alpha$ the proof of Proposition 5 applies as well. Because of Lemma 7 and the proof of Proposition 5, there exists a $\delta' \in [0, 1)$ such that the results of Proposition 5 hold for all $\delta \in (\delta', 1)$. In this case, $\hat{\lambda}' = \hat{\lambda}'' = \lambda^{**}$ and the interval $(\hat{\lambda}', \hat{\lambda}'')$ is empty. Moreover, since $F(\lambda) > 0$ whenever λ is sufficiently small and since (because of Lemma 7) there are at most three solutions $\lambda \in (0, 1)$ of $F(\lambda) = 0$, the pooling price can be optimal at most in two separated subintervals of $(0, 1)$. Thus, we only need to prove (a) that the pooling price $p^L(\lambda)$ is optimal for $\lambda \in \{\hat{\lambda}', \hat{\lambda}'', \lambda^{**}\}$, but the separating price $p^H(\lambda)$ may be optimal as well; and (b) that there exist parameter values of $\delta \in (0, 1)$ and $c > L$ such that $\hat{\lambda}' < \hat{\lambda}'' < \lambda^{**}$.

First we show (b). For $c \ge L$ we get $(1 - \alpha^2)(H - c) = (1 - \alpha^2)(1 + L - c) \le 1 - \alpha^2$. Thus, $\alpha < (1 - \alpha^2)(H - c)$ implies $\alpha^2 < 1 - \alpha$. From Proposition 3 we know that if $c = L$ and $\alpha^2 < 1 - \alpha$, then there exist critical probabilities $\lambda^* \in \Lambda(\lambda_1)$ and

$\lambda^{**} \in \Lambda(\lambda_1)$ for each $\delta \in (\delta^{***}, 1)$, $\delta^{***} < 1$, where $0 < \lambda^* < \lambda^{**} < 1$, such that it is uniquely optimal for the seller to demand $p_t = p^H(\lambda_t)$ whenever $\lambda_t \in (\lambda^*, \lambda^{**})$, and $p_t = p^L(\lambda_t)$ whenever $\lambda_t \in (0, \lambda^*) \cup (\lambda^{**}, 1)$. Fix some $\delta \in (\delta^{***}, 1)$, some $\mu' \in (0, \lambda^*) \cap \Lambda(\lambda_1)$, and some $\mu'' \in (\lambda^*, \lambda^{**}) \cap \Lambda(\lambda_1)$. For any unit cost c, let $V(\lambda; c)$ denote the value function of the seller's optimization problem, and let $W(\lambda; c)$ denote the maximum expected profit of the seller in a given period, if $p^H(\lambda)$ is chosen in this period and from the next period onwards the optimal price is chosen, i.e., $W(\lambda; c) \equiv \varphi(\lambda)\left[p^H(\lambda) - c\right] + \delta\left[\varphi(\lambda) V\left(\lambda^+; c\right) + (1 - \varphi(\lambda)) V\left(\lambda^-; c\right)\right]$. Obviously, $V\left(\lambda'; c\right) = W\left(\lambda'; c\right)$ if and only if $p^H(\lambda)$ is optimal for $\lambda = \lambda'$, given c. Since profits change continuously with c, $V(\lambda; c)$ and $W(\lambda; c)$ are continuous in c. If $c = L$, (i) $p^L(\lambda)$ is uniquely optimal for $\lambda = \mu'$ and hence $\frac{p^L(\mu') - c}{1-\delta} = V\left(\mu'; L\right) > W\left(\mu'; L\right) > 0$; and (ii) $p^H(\lambda)$ is uniquely optimal for $\lambda = \mu''$ and thus $V\left(\mu''; L\right) > \frac{p^L(\mu'') - c}{1-\delta}$. Choose now some $\bar{c} = L + \varepsilon$ such that $\varepsilon > 0$ is sufficiently small to imply $\frac{p^L(\mu') - \bar{c}}{1-\delta} > W\left(\mu'; L\right)$ and $V\left(\mu''; \bar{c}\right) > \frac{p^L(\mu'') - \bar{c}}{1-\delta}$. Because from this and $W\left(\mu'; L\right) > 0$ it follows that $p^L\left(\mu'\right) - \bar{c} > 0$, market exit is not optimal for $\lambda = \mu'$. Therefore, $\mu' > \lambda^E$. Since all profits decrease in c, $W\left(\mu'; L\right) \geq W\left(\mu'; \bar{c}\right)$, and thus $\frac{p^L(\mu') - \bar{c}}{1-\delta} > W\left(\mu'; \bar{c}\right)$. Hence for $\lambda = \mu'$ it is uniquely optimal for the seller to stay in the market and charge the pooling price $p^L\left(\mu'\right)$. Define $\lambda^{E+} \equiv \Pr(H \mid \lambda^E, s = h)$. Since exit is optimal for $\lambda = \lambda^E$, it must hold that $p^H\left(\lambda^E\right) \leq c$, and thus $p^L\left(\lambda^{E+}\right)$ cannot be optimal because $p^L\left(\lambda^{E+}\right) < p^H\left(\lambda^E\right)$ implies $p^L\left(\lambda^{E+}\right) < c$. Furthermore, exit is not optimal for $\lambda = \lambda^{E+} \in \left(\lambda^E, 1\right)$. Hence the separating price is uniquely optimal for $\lambda^{E+} \in \left(\lambda^E, \mu'\right)$. Moreover, $V\left(\mu''; \bar{c}\right) > \frac{p^L(\mu'') - \bar{c}}{1-\delta}$ implies that the separating price $p^H\left(\mu''\right)$ is uniquely optimal for $\lambda = \mu'' > \mu'$, whereas the pooling price $p^L\left(\mu'\right)$ is uniquely optimal for $\lambda = \mu'$. This proves point (b).

Finally, we prove point (a), i.e., that the pooling price $p^L(\lambda)$ is optimal for $\lambda \in \{\hat{\lambda}', \hat{\lambda}'', \lambda^{**}\}$, but the separating price $p^H(\lambda)$ may be optimal as well. For λ^{**} this holds because we can choose λ^{**} to be the smallest λ such that staying in the market and charging the pooling price $p^L(\lambda)$ is optimal (though not necessarily uniquely optimal) for all $\lambda \in \left[\lambda^{**}, 1\right) \cap \Lambda(\lambda_1)$. This implies that $\lambda^{**} > \lambda^E$ (since $p^L\left(\lambda^E\right) < p^H\left(\lambda^E\right) \leq c$), and that $\hat{\lambda}'' < \lambda^{**}$ if $\hat{\lambda}' < \lambda^{**}$. If $\hat{\lambda}' < \lambda^{**}$, there must exist $\mu' \in \Lambda(\lambda_1)$ and $\mu'' \in \Lambda(\lambda_1)$ such that $\lambda^E < \mu' < \mu'' < \lambda^{**}$ and the pooling price $p^L\left(\mu'\right)$ is optimal (though not necessarily uniquely optimal) for $\lambda = \mu'$, whereas the separating price $p^H\left(\mu''\right)$ is optimal (though not necessarily uniquely optimal) for $\lambda = \mu''$. This holds because if such a pair of μ' and μ'' does not exist, the interval $\left(\lambda^E, 1\right)$ can be partitioned into subintervals $\left(\lambda^E, \lambda^{**}\right)$, $\left[\lambda^{**}\right]$, and $(\lambda^{**}, 1)$ such that the separating price is uniquely optimal in $\left(\lambda^E, \lambda^{**}\right)$ and the pooling price is uniquely optimal in $(\lambda^{**}, 1)$, and optimal in for $\lambda = \lambda^{**}$, and this implies $\hat{\lambda}' = \hat{\lambda}'' = \lambda^{**}$, contradicting $\hat{\lambda}' < \lambda^{**}$. Since μ' and μ'' exist, we can choose $\hat{\lambda}'$ to be the smallest $\lambda > \lambda^E$ such that the pooling price $p^L(\lambda)$ is optimal, and $\hat{\lambda}''$ to be the largest $\lambda < \mu''$ such that the pooling price $p^L(\lambda)$ is optimal. From this (a) follows. □

Proof of Proposition 7 When the seller exits the market or sells to both buyer types, the purchase behavior of the current buyer does not reveal any information. Once an informational cascade has started, beliefs and prices remains trivially constant, $\lambda_{t+1} = \lambda_t$ and $p_{t+1} = p_t$.

If instead at the current belief the seller charges the separating price, there is active social learning. In this case, the belief in the next period is either $\lambda_{t+1} = \lambda^+ (\lambda_t)$ if the current buyer purchases or $\lambda_{t+1} = \lambda^- (\lambda_t)$ following no purchase. The price in the following period depends on whether at the realized belief λ_{t+1} the seller charges the separating price or induces either a purchase or an exit cascade.

Starting from a belief in the separating region, at the updated belief following a purchase by the current buyer, $\lambda_{t+1} = \lambda^+ (\lambda_t)$, it is optimal for the seller either (A) to charge the separating price, $p_{t+1} = p^H (\lambda_{t+1}) = \lambda^+ (\lambda_{t+1}) + L = \lambda^+ \left(\lambda^+ (\lambda_t)\right) + L$, or (B) to charge the pooling price, $p_{t+1} = p^L (\lambda_{t+1}) = \lambda^- \left(\lambda^+ (\lambda_t)\right) + L = \lambda_t + L$. If instead the current buyer does not purchase, at the updated belief $\lambda_{t+1} = \lambda^- (\lambda_t)$, three prices can be optimal: either (i) the separating price, $p_{t+1} = p^H (\lambda_{t+1}) = \lambda^+ \left(\lambda^- (\lambda_t)\right) + L = \lambda_t + L$, or (ii) the pooling price $p_{t+1} = p^L (\lambda_{t+1}) = \lambda^- \left(\lambda^- (\lambda_t)\right) + L$, or (iii) the exit price, $p_{t+1} = p^H (\lambda_{t+1}) + \varepsilon = \lambda^+ \left(\lambda^- (\lambda_t)\right) + \varepsilon + L$. Overall, there are six possibilities that we cover in turn below.

First, consider case (A) in which following a purchase from the current buyer, $a_t = 1$, the period $t + 1$ price is separating, $p^H \left(\lambda^+ (\lambda_t)\right) = \lambda^+ \left(\lambda^+ (\lambda_t)\right) + L$. In case (A.i), the separating price is also charged in period $t + 1$ following no purchase in period t. Since the separating price $p^H (\lambda) = \lambda^+ (\lambda) + L = \frac{\lambda\alpha}{\lambda\alpha + (1-\lambda)(1-\alpha)} + L$ is a concave function of the belief λ, Jensen's inequality and the martingale property of beliefs (verified in equation (2)) imply that the price is decreasing on average, $E_t[p_{t+1}] < p_t$ where the inequality is strict by the assumption that the signal is informative, $\alpha > 1/2$. In case (A.ii), the argument given for case (A.i) implies a fortiori that $E_t[p_{t+1}] < p_t$ since $p^H \left(\lambda^- (\lambda_t)\right) > p^L \left(\lambda^- (\lambda_t)\right)$. In case (A.iii), again we have $E_t[p_{t+1}] < p_t$ provided that ε is small enough.

Consider next case (B) in which following a purchase from the current buyer, $a_t = 1$, in period $t + 1$ the seller induces a purchase cascade by charging $p^L \left(\lambda^+ (\lambda_t)\right) = \lambda^- \left(\lambda^+ (\lambda_t)\right) + L = \lambda_t + L$. In case (B.i), the price decreases deterministically to $p_{t+1} = \lambda_t + L < p_t = \lambda^+ (\lambda_t) + L$. In case (B.ii), the price decreases both following a purchase and no purchase. In case (B.iii), the price decreases provided that ε is sufficiently small. □

References

Avery, C., Zemsky, P.: Multidimensional uncertainty and herd behavior in financial markets. Am Econ Rev **88**, 724–748 (1998)

Bagwell, K., Riordan, M.: High and declining prices signal product quality. Am Econ Rev **81**, 224–239 (1991)

Bar-Isaac, H.: Reputation and survival: learning in a dynamic signalling model. Rev Econ Stud **70**, 231–251 (2003)

Banerjee, A.V.: A simple model of herd behavior. Q J Econ **107**, 797–817 (1992)

Bergemann, D., Välimäki, J.: Learning and strategic pricing. Econometrica **64**, 1125–1149 (1996)

Bikhchandani, S., Hirshleifer, D., Welch, I.: A theory of fads, fashion, custom, and cultural change as informational cascades. J Polit Econ **100**, 992–1026 (1992)

Bose, S., Orosel, G.O., Vesterlund, L.: Optimal pricing and endogenous herding. Department of Economics, University of Vienna, Working Paper No. 0204 (2001)

Bose, S., Orosel, G.O., Ottaviani, M., Vesterlund, L.: Dynamic monopoly pricing and herding. RAND J Econ **37**, 910–928 (2006)

Caminal, R., Vives, X.: Why market shares matter: an information-based theory. RAND J Econ **27**, 221–239 (1996)

Chamley, C.: Rational Herds: Economic Models of Social Learning. Cambridge: University Press (2004)

Judd, K., Riordan, M.: Price and quantity in a new product monopolist. Rev Econ Stud **61**, 773–789 (1994)

Gill, D., Sgroi, D.: The Optimal Choice of Pre-launch Reviewer. University of Oxford and University of Cambridge mimeo (2005)

Neeman, Z., Orosel, G.O.: Herding and the winner's curse in markets with sequential bids. J Econ Theory **85**, 91–121 (1999)

Moscarini, G., Ottaviani, M.: Economic models of social learning. In: Battigalli, P., Montesano, A., Panunzi, F. (eds.) Decisions, Games and Markets. Boston: Kluwer (1997)

Ottaviani, M.: Monopoly pricing with social learning. In: Social Learning in Markets. Economics Ph.D. Thesis, Massachusetts Institute of Technology (1996)

Ottaviani, M., Prat, A.: The value of public information in monopoly. Econometrica **69**, 1673–1683 (2001)

Saak, A.E.: A note on the value of public information in monopoly. Econ Theor **33**, 369–379 (2007)

Sgroi, D.: Optimizing Information in the Herd: Guinea pigs, profits, and Welfare. Games Econ Behav **39**, 137–166 (2002)

Smith, L., Sørensen, P.: Pathological outcomes of observational herding. Econometrica **68**, 371–398 (2000)

Takayama, A.: Mathematical Economics. Hinsdale: Dryden Press (1974)

Taylor, C.R.: Time-on-the-market as a sign of quality. Rev Econ Stud **66**, 555–578 (1999)

Welch, I.: Sequential sales, learning and cascades. J Financ **47**, 695–732 (1992)