SUBSIDIZING LEMONS FOR EFFICIENT INFORMATION AGGREGATION

VADIM MARTYNOV

Abstract. We study how social planning can reduce the inefficiencies of social learning, stemming from herding and informational cascades. A social planner is introduced to the classical sequential social learning model. She can tax or subsidize players' actions in order to maximize social welfare, a discounted sum of agents' utilities. We solve or accurately approximate the expected utility of the social planner and the optimal pricing strategy for various signal distributions. In equilibrium, it is optimal to increase the price for the better action, causing a reduction in current agent's utility, but also a net gain, due to the information this action reveals. The addition of the social planner significantly improves social welfare and the asymptotic speed of learning.

1. Introduction

The aggregation of information in society is a complex, and, at the same time, a very interesting process from an economic point of view. Especially it is interesting to see how it affects individuals' choices. People's decisions often rely on two types of information. The first one is their private knowledge about the choice they face. The second source is information received from society, in particular, what other people did before. Social learning models have been used to analyze how people make decisions based on these types of information. They explain phenomena such as herding [2], informational cascades [4], and asymptotic learning [12].

Usually, these models have huge inefficiencies. For example, people might end up in the wrong cascade. Furthermore, when herding occurs, people's actions convey much less information about their private signals than they do in the beginning. One of the reasons this happens is because people do not take into account how their actions affect future generations' utility. The main aftermath of this is a decrease in

I would like to thank my adviser Omer Tamuz, Federico Echenique, Laura Doval, Luciano Pomatto, Colin Camerer, Alejandro Robinson Cortes, Lucas Nunez, Yimeng Li, Mali Zhang, Theory and Graduate seminars participants at Caltech for helpful comments and discussions.

social welfare and in the asymptotic speed of learning [10]. A natural question arises: how can we improve this social learning system?

In this paper we start with the classical sequential learning model with binary states of nature and corresponding binary actions. Our leading example is the case in which there are two goods/technologies on the market: the old one, whose characteristics are well-known to everybody, and the new one, which is better than the former one in one state of the world and worse in the other. The actions here are represented by buying either the new or the old product. Agents have the same preferences: they want their actions to match the underlying state. Players get private signals that are i.i.d. conditioned on the realized state of the world. The players are exogenously ordered and, in every time period, one agent chooses which product to buy, based on her private signal and the actions of her predecessors.

A social planner would like to maximize the discounted sum of players' utilities. She can choose relative prices, by taxing or subsidizing the goods in every period, that are publicly observed. In other words, she chooses a cost, which can be positive, negative or zero, of the new product, normalizing the price of the old one to be 0. The social planner's choice is based on the public information available at that time: prices in previous periods and corresponding players' actions.

We find the optimal pricing policy, which, perhaps, is counterintuitive: tax the good that is more likely to be better. This results in an expected loss today, because agents whose private information is not strong enough are less likely to buy the better product. But at the same time, their choice will give the social planner more information about their private signals, which in its turn significantly increases the utility of future generations.

As a motivational example, imagine there are two drugs that treat the same disease but have unknown benefits/side effects, as they were tested only on a small sample of population. The only way the government can collect information about the real effects is by observing which one people bought, given their personal and public knowledge. Here the government plays the role of the social planner who can tax or subsidize one drug or the other. Without it there is a significant risk that society will stick with the wrong medicine or that it will learn the truth very slowly. We show how these risks become lower when the social planner is involved in this process.

Our model overcomes a major criticism of the sequential social learning model – the assumption that players have to know the order in which the actions were taken by their predecessors and then make a complicated calculation by Bayes rule to obtain the current public belief. In our setup, the optimal prices are functions of the public belief, which contains all the necessary information about the past. In equilibrium agents can recover the public belief from the price, and do not need to know the actions and specific order of their predecessors. Thus, agents only need to observe the current price and private signals to choose the optimal action.

We start with introducing the model, describing individuals' and the social planner's behaviors and providing an intuitive illustration of how prices change the belief - action relationship. After this, we go to the binary signals case and calculate the expected utility function and its asymptotic characteristics when δ goes to 1. On the way to these results, we establish a few interesting and helpful properties of the binary case. We show that the public belief is a random walk, that depends on the difference between the number of $High$ and Low signals up to this point. Another property is that when an agent takes into account their private signal, their expected utility is equal to the signals' precision.

In the next section we study continuous distributions: bounded and unbounded. We find that there is still a difference in terms of asymptotic learning: when private signals have bounded strength the underlying state is never revealed. However, the beliefs at which learning stops are closer to 0 and 1 than they would be without the social planner. For general signal distributions this problem becomes very complex and we can not find the analytical solution. But we provide a good qualitative description for a specific unbounded signal distribution. This description looks similar to the one in the binary case: when one product seems to be better, it should be taxed, in order to extract more information in the future periods and exploit the convexity of the expected utility function. Moreover, the prices are bounded away from 0 and 1 when the public belief converges to one of the extreme values. This implies that eventually people are going to buy the right product with probabilities close to 1, but at the same time, it is going to be positively priced, to increase the social benefits. Finally, we calculate numerically the expected utility and the optimal pricing function for this signal distribution.

1.1. **Literature review.** Several previous studies consider a similar question: how prices might affect social learning in different scenarios.

Crapis, Ifrach, Maglaras, Scarsini (2016) [8] consider a situation when people with heterogeneous preferences observe not only actions of their predecessors but also their reviews of the product, the outcome, in a non-Baysian framework. Numerical experiments suggest that pricing policies that account for social learning may increase revenues considerably relative to policies that do not.

Papanastasiou, Savva (2016) [11] and Manaswini (2012) [3] allow both buyers and a monopolist to act strategicly over a finite number of time periods. The first one finds that the social learning increases the firm's expected profit and contrary to previous results in the literature, preanounced prices are not beneficial to the firm. The second one shows that prices are no longer sub-martingales, but that for some range of beliefs they can be super-martingales too.

In a more classic setup Bose, Orosel, Ottaviani, and Vesterlund (2006) [5] consider a binary model when a monopolists chooses pricing strategy in order to maximize its revenue and incurs some cost to produce the product. They find some qualitative results. For example the objective is convex, increases in number of periods the game is played and that at some point herding occurs.

There is also some literature on optimal pricing in networks. Candogan, Bimpikis and Ozdaglar (2012) [7] study the optimal pricing strategy of a monopolist in a two period game, where an agent's utility depends not only on her action, but also on decisions of her neighbors. They find that optimal price should depend on Bonacich centrality, on a markup term proportional to the influence that the network exerts and on a term that is independent of the network.

Another paper in this area is by Campbel (2013) [6]. He models a firm's ability to strategically influence the probability individual engages in WOM, word of mouth, through the price. The author derives the comparative static results of connectivity, mean-preserving spread of friendships, and clustering of friends on price.

2. Model

Let $\theta \in \{High, Low\}$ be the true state of the world, where both states a priori are equally likely[1]. At the beginning of the game one state is realized and does not change.

There are countably many rational agents $t \in \{1, 2, \dots\}$, who receive private signals s_t. These signals are i.i.d. conditional on the state of the world: if $\theta = High$, they have cumulative distribution function (CDF) G_H and if $\theta = Low$ - G_L. The corresponding PDFs are g_H and g_L. We assume that signals never completely reveal the true state, which

[1] We make this simplification of a (1/2,1/2) prior to reduce the complexity of the presentation, but all results hold for general priors

is the same as saying that the conditional distributions are absolutely continuous with respect to each other.

Suppose there are two products on the market with different prices in each time period: the new one and the old one. When $\theta = High$ the new product is better than the old one and when $\theta = Low$ - vise versa. Each agent t has a decision to make: whether to buy the new product ($a_t = 1$) and pay the price c_t or to stick with the old one ($a_t = 0$). We normalize the price of the old product to be equal to 0. Utility of each agent from the action a_t is 1, if it matches the state ($a_t = 0$ and $\theta = Low$ or $a_t = 1$ and $\theta = High$)), minus the cost, if she buys the new product, -$\mathbb{1}\{a_t = 1\}c_t$. Players act sequentially and their decisions are based on two types of information: private information from their own private signal s_t and public information (history) h_t. The latter one includes actions that were taken before player t and the sequence of prices of the new good c_t, so $h_t = \{a_1, a_2, \ldots a_{t-1}, c_1, c_2, \ldots c_{t-1}\}$. Before we explain what is the real role of c_t we need some notation.

2.1. **Agent's decision process.** Denote the posterior belief of the agent t that the new good is better by

$$\mu_t = \mathbb{P}(\theta = High | h_t, s_t).$$

We will also refer to it as the *total belief* as it combines public and private information. Also, let us call the corresponding likelihood ratio, $\mu_t/(1 - \mu_t)$ - the *total likelihood ratio.*

Note that μ_t also represents the expected utility of player t for taking action 1, not including the price. As h_t and s_t are independent of each other, the posterior belief has two components: the private belief $\mathbb{P}(\theta = 1|s_t)$, which is known only to player t, and the public belief $p_t = \mathbb{P}(\theta = 1|h_t)$, which is known to everyone who observed history up to time t. Also, denote by l_t the likelihood ratio of the public belief

$$l_t = \frac{p_t}{1 - p_t}.$$

We can see that there is a monotone bijection between the public belief and its likelihood ratio and we are going to use them interchangeably. As $p_t \in [0, 1]$ then $l_t \in [0, \infty]$. Also denote by $F_H(l_t)$, $F_L(l_t)$ the CDF's of l_t conditional on θ.

The posterior belief μ_t captures how confident the player is about buying the new product. We obtain p_t from $p_0 = 1/2$ and h_t using Bayes rule.

Now let us go back to c_t. Suppose that $c_t = 0$. As agents are expected utility maximizers, player t buys the new product if her posterior belief μ_t is greater than $1/2$. Now, if $c_t \neq 0$ then she buys it only if $\mu_t \geq$

$1/2 + c_t/2$: if she buys the new product then her expected utility is $\mu_t - c_t$ which has to be greater or equal to $1 - \mu_t$ - the utility from buying the old product. In other words, these prices reflect how confident you should be in the new product, in comparison to the old one, in order for you to buy the former one.

We can summarize it in the following way

$$\begin{cases} a_t = 0, \text{ utility } 1 - \mu_t & \text{when } \mu_t < \frac{1}{2} + \frac{c_t}{2} \\ a_t = 1, \text{ utility } \mu_t - c_t & \text{when } \mu_t \geq \frac{1}{2} + \frac{c_t}{2} \end{cases}$$

The agent wants to guess the correct state of the world ($a_t = \theta$) in general. But there are some situations when it is more profitable to take the opposite action in order to avoid the cost c_t. Imagine that $\mu_t = 0.57$ and $c_t = 0.15$. Even though the total belief tells us to buy the new product ($\mu_t > 0.5$) we would get less utility by doing this ($0.57 - 0.15 = 0.42$) rather than buying the old one ($1 - 0.57 = 0.43$). Hence, the price forces some people with a not very strong belief to switch to a "non optimal" action, while people with a strong belief are not affected by it.

We can also state this condition in terms of the likelihood ratio: $a_t = 1$ iff[2]

$$\frac{\mu_t}{1 - \mu_t} = \frac{p_t}{1 - p_t} \frac{g_H(s_t)}{g_L(s_t)} \geq k_t,$$

where

$$k_t = \frac{1 + c_t/2}{1 - c_t/2}.$$

If $c_t = 0$ ($k_t = 1$) we get the usual conditions for taking action 1 .We are going to use both c_t (for beliefs) and k_t (for likelihood ratios) as prices but in different settings. Notice that we can rewrite the condition above as follows

$$\frac{p_t}{(1 - p_t)k_t} \cdot \frac{g_H(s_t)}{g_L(s_t)} \geq 1.$$

We can interpret this as if there were no price, $k_t = 1$, but we had a lower public belief, that corresponds to the likelihood ratio $p_t/((1 - p_t)k_t)$.

Definition 1. *Let us call $p_t/((1 - p_t)k_t)$ - modified likelihood ratio. Similarly, we call the public belief that corresponds to the modified likelihood ratio - modified public belief.*

[2]For simplicity, we assume that agents choose action-1 when indifferent. This will have no impact on our results.

Last thing we mention here is how the public belief evolves after another player takes an action. If at time t the public belief is p_t and the price is k_t, then the public belief at time $t+1$ after taking action $a_t = 1$ satisfies the following formula

$$l_{t+1} = \frac{p_{t+1}}{1-p_{t+1}} = \frac{p_t}{1-p_t} \cdot \frac{1-F_H(y)}{1-F_L(y)},$$

where $y = (1-p_t)k_t/p_t$ and F_i - CDFs of the likelihood ratio conditioned on the true state. And if $a_t = 0$

$$l_{t+1} = \frac{p_{t+1}}{1-p_{t+1}} = \frac{p_t}{1-p_t} \cdot \frac{F_H(y)}{F_L(y)}.$$

This is an application of the Bayes rule, given that player t buys (does not buy) the new product if her total likelihood ratio is above (below) k_t. This implies that her private likelihood ratio is above (below) $(1-p_t)k_t/p_t$.

Now let us look at this game from the social planner's perspective, who would like to maximize the discounted sum of the expected utilities of agents with a discount factor δ.

2.2. **Social planner.** We introduce a long run risk neutral social planner who chooses a price, by taxing or subsidizing the new product, in each period and then returns the collected money to players in the following way. If in time period $t-1$ c_{t-1} was collected from player $t-1$ then this money is put in the bank with the interest rate $1/\delta$ and returned to player t in the next period. This mechanism implies two things. First, redistribution does not affect players' decisions and they still act according to the previous subsection. This is true, because player's t choice does not change how much money she gets back, it is already decided by the previous player. Second, the total amount of money that is taken from the players is 0, budget-balanced. Thus, the discounted sum of players' payoff is equal to the discounted sum of $u_t(\mu_t, c_t) = \mathbb{1}\{a_t(\mu_t, c_t) = 1\}\mu_t + \mathbb{1}\{a_t(\mu_t, c_t) = 0\}(1-\mu_t)$. Therefore, the utility function of the social planner, who implements a pricing strategy $\{c_t\}_{t=1}^{\infty}$ is

$$u_{\{k_t\}}(p) = \sum_{t=1}^{\infty} u_t(\mu_t(p, h_{t-1}, s_t), k_t),$$

where $\mu_t(p, h_{t-1}, s_t)$ means that it is path dependent.

Due to the money redistribution defined above, if $\mu_t > 1/2$ and player t buys the new product we treat her utility from society's prospective as just μ_t instead of subtracting the price, as it is returned later. But from the players perspective they still take into account the price.

The optimal pricing strategy plays another crucial role in this model. Imagine that the social planner does not do anything ($\forall t\ k_t = 1$) then when the herd occurs, the difference between public beliefs in two consequitive periods converges to 0 extremely quickly. This is shown in [10]. Without loss of generality assume that they are buying the new product. But if we set a high price for the new product and observe that people are still buying it, then the posterior belief would grow faster than before. In other words, the asymptotic speed of learning would increase.

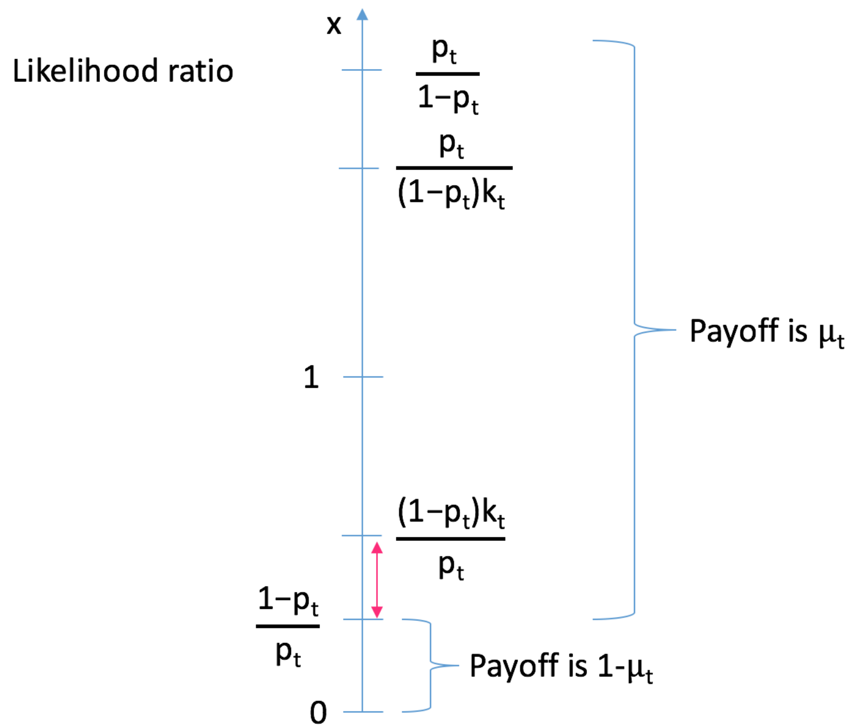
Let us denote by $u(p_t)$ the expected utility of the social planner with the public belief p_t when we use the optimal prices k_t^*. Due to the stationarity of the social planner's problem the expected utility depends on the history only through the public belief. Then it should satisfy the following Bellman equation:

(1)
$$u(p_t) = \max_{k_t}\{(1-\delta)(\text{expected gain}(p_t) - \text{expected loss}(k_t, p_t)) + \delta\, \mathbb{E}\, u(p_{t+1})\}.$$

The expected gain calculates the expected utility in this period if there were no price, $\mathbb{E}\, u_t(\mu_t, 1)$. First, we calculate the expected μ_t and then the expected payoff is equal to $\max(\mu_t, 1-\mu_t)$. If we receive a private signal with a likelihood ratio x then the total likelihood ratio is equal to $y = p_t x/(1-p_t)$ and therefore, $\mu_t = y/(y+1)$. Depending on whether the corresponding likelihood ratio is above or below 1, the expected utility is either equal to μ_t or $1-\mu_t$, so we divide this into two cases: $0 \le x \le (1-p_t)/p_t$ and $x > (1-p_t)/p_t$. Therefore, the expected gain(p_t) is equal to

$$\int_0^{\frac{1-p_t}{p_t}} \left(1 - \frac{x\frac{p_t}{1-p_t}}{x\frac{p_t}{1-p_t}+1}\right)(p_t f_H(x) + (1-p_t) f_L(x))dx +$$
$$+ \int_{\frac{1-p_t}{p_t}}^{\infty} \left(\frac{x\frac{p_t}{1-p_t}}{x\frac{p_t}{1-p_t}+1}\right)(p_t f_H(x) + (1-p_t) f_L(x))dx.$$

The second term, the expected loss(k_t, p_t), calculates utility that agent loses due to a non optimal action. Notice that if we apply price k_t, then the only loss that can occur is when the likelihood ratio of the private signal is strong enough to make the total likelihood ratio less than 1 for the modified likelihood ratio $(1-p_t)k_t/p_t$, but not for the initial one, $(1-p_t)/p_t$. When this happens, the likelihood ratio of the private signal x can be between $(1-p_t)/p_t$ and $(1-p_t)k_t/p_t$. If the total belief is μ_t then the loss that player bears is $\mu_t - (1-\mu_t) = 2\mu_t - 1$. Therefore, the expected loss(k_t, p_t) is equal to



Red area corresponds to the private likelihood ratios when people take non optimal actions. Here $\mu_t = (x\frac{p_t}{1-p_t})/(x\frac{p_t}{1-p_t}+1)$.

FIGURE 1. Private likelihood ratios and today payoff.

$$\int_a^b \left(\frac{2x\frac{p_t}{1-p_t}}{x\frac{p_t}{1-p_t}+1} - 1 \right) (p_t f_H(x) + (1-p_t) f_L(x)) dx,$$

where $a = (1-p_t)/p_t$, $b = (1-p_t)k_t/p_t$.

The high complexity of this problem is due to the intricacy of the random walk of the public belief. Equation (1) shows the trade-off between losing some utility today due to the fact that some people (whose posterior likelihood is between 1 and k_t) take a non optimal action (and paying $(2\mu_t - 1)$ for this) and gaining utility through more a disperse belief tomorrow. The latter one occurs due to the convexity of u which we prove a bit later.

In other words, strategic pricing can help to aggregate information more efficiently and increase the social welfare as well as speed up the asymptotic learning.

In the subsequent sections we are going to consider both discrete and continuous, bounded and unbounded private signals and analyze how it affects properties of the optimal pricing policy and the corresponding solution. But before doing this we are going to state a general property of the expected utility function.

Proposition 2. *The expected utility function $u(p)$ is convex, $u(1) = u(0) = 1$ and is symmetric around $1/2$, i.e $u(p) = u(1-p)$.*

Most of the time we are going to assume, without loss of generality, that $p_t \geq 0.5$, so the new product is a priory better, and will try to find the optimal price, or at least a price that is better than no price at all. For $p_t < 0.5$, a symmetric analysis applies.

3. Binary signals

In this section we are going to calculate the utility and the optimal strategy of the social planner when private signals are Bernoulli distributed. This means that each agent is going to be told whether the state is *High* or *Low*, as her private signal, and this information is going to be correct with probability $q > 1/2$.

Definition 3. *Define a* ***learning period****, LP_t to be all periods t' up to time t, conditioned on h_t, such that a player t' took into account her private signal when she chose an action $a_{t'}$.*

We abuse notation a bit and ignore the subscript t as it is going to be clear which period we have in mind. Vadim: If people disregard their private signals when they take actions, then the public belief does not change, so in this sense, we *do not learn* in these periods. It is helpful to keep in mind this distinction between LP and not LP.

There are a few nice characteristics of the binary distribution. The first one is that if t is in the learning period then agent t's action reveals her private signal. Indeed, if we know that people are going to take into account their private signals then the only possibility is that they act according to them.

Lemma 4. *If $t \in LP$ then player t's action, a_t, reveals her private signal s_t.*

Thus, during the learning period actions not only convey but also reveal the information.

The second one is that there are only two nontrivial multipliers by which we update the likelihood ratio: either we observe the $High$ signal and multiply l_t by

$$\frac{\mathbb{P}(\theta = High|s_t = High)}{\mathbb{P}(\theta = Low|s_t = High)} = \frac{q}{1-q},$$

or we observe the Low signal and then multiply l_t by

$$\frac{\mathbb{P}(\theta = High|s_t = Low)}{\mathbb{P}(\theta = Low|s_t = Low)} = \frac{1-q}{q}.$$

When player t ignores s_t we do not update the public belief, so l_t is multiplied by 1 and from now on all claims about the public belief assume that actions depend on the private signals. Thus, if player t's action is informative then l_t is multiplied by either $q/(1-q)$ or $(1-q)/q$ depending on her action.

The third, and the best characteristic of the binary case is summarized in the following lemma.

Lemma 5. *In the binary case, the public belief is a random walk and its position depends on the difference in the number of times we observed High and Low signals. Moreover, if we observed n High signals and k Low ones between periods t and $t+n+k$ then the likelihood ratio at time $t+n+k$ is*

$$l_{n+k} = l_t \cdot \left(\frac{q}{1-q}\right)^{n-k}$$

Lemma 5 tells us that in order to find l_t we just need to calculate the difference in number of times agents took actions 1 and 0 during the LP and take $q/(1-q)$ to the corresponding power.

Definition 6. *We say the public belief (LR)* ***goes up*** *if we observed the $High$ signal, during LP. Analogously, the public belief (LR)* ***goes down*** *if we observed the Low one.*

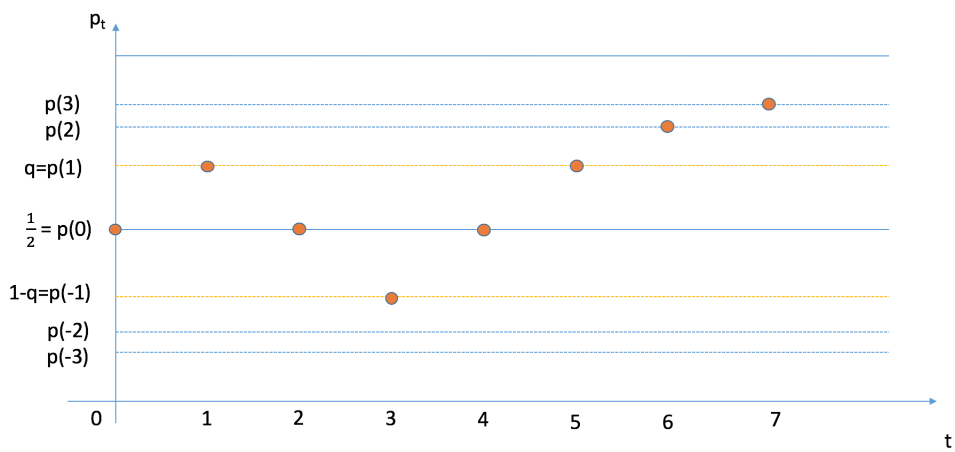
Notice that the public belief increases (decreases) when it goes up (down) as $q > 1/2$.

If at time t we have $p_t = 1/2$ and we go n times up then by Bayes rule the public belief is

$$p_{t+n} = \frac{q^n}{q^n + (1-q)^n}$$

and if we go down n times from $p_t = 1/2$ then

$$p_{t+n} = \frac{(1-q)^n}{q^n + (1-q)^n}.$$



Where $p(k) = \frac{q^k}{q^k+(1-q)^k}$ and $p(-k) = \frac{(1-q)^k}{q^k+(1-q)^k}$ as in (2).

FIGURE 2. Random walk of the public belief with binary signals.

Definition 7. *For $k \geq 0$ define **level** k to be the public belief if it observed k more High signals that Low ones and denote it by $p(k)$. Similarly, define **level** $(-k)$ to be the public belief if we observed k more Low singals than High ones and denote it by $p(-k)$*

$$
\begin{cases} p(k) = \dfrac{q^k}{q^k + (1-q)^k} \\ p(-k) = \dfrac{(1-q)^k}{q^k + (1-q)^k} \end{cases} \tag{2}
$$

where $k > 0$.

For example, public belief 1/2 corresponds to level 0.

The last property we mention is that if $k_t = 1$ then agents start ignoring their private signals (learning period stops) (a cascade[3] occurs) after one or two people take the same action. If we start with $p_t = 1/2$ and someone takes action 1, p_{t+1} will be equal to q. Even if the next player gets *Low* private signal her posterior belief is going to be 1/2, hence, she would take action 1 disregarding her private signal. This means, we stop aggregating private information extremely quickly

[3]This is an event in the classical model when people disregard their private signals and take the same action, thus public belief does not update after this. Once it has started it does not stop. Although, in our model we are able to stop it by introducing a price, which brings the public belief back to the region where agents act according to their private signals

and therefore unable to get to a high public belief, which hurts social welfare.

To improve this situation, by maximizing the social welfare, we are going to implement the optimal pricing scheme. The following lemma will help in our analysis.

Lemma 8. *There exists an optimal strategy of the following form: the social planner* 1) *picks* $N \in \mathbb{N}$, 2) *chooses prices such that actions reveal the private signals until the public belief reaches either level* N *or* $-N$*. After that there are no prices,* $k_t = 1$*.*

Notice that from the social planner's perspective this game is stationary. Thus, in this binary situation, there is only one way how she can affect the outcome: she can choose how long we are going to distinguish signals, so how long does the learning period continue. As we said above, due to the stationarity, there is no incentive to choose a non zero price c_t once the learning period has ended. Notice, that since we are not biased towards one or the other state these two levels should be symmetric around 1/2: upper bound - level N, lower bound - $-N$.

Thus, we pick $N \in \mathbb{N}$ at the beginning and then in every period, until p_t hits either level N or $-N$, choose a price which allows us to separate $High$ and Low private signals. This is the pricing scheme.

How exactly do these prices look like? Suppose at time t the likelihood ratio $l_t > 1$ (p_t is greater than 1/2), which we assume is above $q/(1-q)$ (otherwise do not need any price, we are still going to figure out the private signal from the action) then the social planner chooses a price $k_t > 1$ such that

$$\frac{l_t}{k_t} = \frac{q}{1-q} - \varepsilon > \frac{1}{2},$$

for a small, positive ε.

Now we go back to the initial problem (1). Recall that one of our goals is to maximize the social welfare. To do this, first we need to calculate our expected utility from starting at the public belief 1/2, $u(0.5)$, when private information is acquired until p_t hits one of the barriers that are at distance N from the initial belief 1/2. And we are interested in N which maximizes $u(0.5)$.

Let us start with a few observations.

Lemma 9. *The expected utility that the social planner gets today in the learning period is* q*.*

There is a simple intuition behind this. As our action follows our signal in LP we take the right action only with probability q. If the true state of the world is $High$ we receive the corresponding signal only with probability q. Hence, we are going to make the right action with the same probability, which equals to our expected utility today.

Now, let us go back to our Bellman equation and rewrite it for this case

(3)
$$u(p(n)) = (1-\delta)q + \delta\left(\mathbb{P}(\text{signal } High)u(p(n+1)) + \mathbb{P}(\text{signal } Low)u(p(n-1))\right),$$

where $\mathbb{P}(\text{signal } High) = (p_t q + (1-p_t)(1-q))$ and $\mathbb{P}(\text{signal } Low) = p_t(1-q) + (1-p_t)q$. We choose N in order to maximize utility at $1/2$, $p(0)$.

This is better than the general form (1) but still complicated, as the probability of going up or down depends on the current public belief.

Fortunately, in order to calculate $u(p(N))$ the analysis can be simplified. Recall that the only thing that the social planner controls is how far away are the absorbing boundaries from $1/2$, in other words she chooses N. After it is fixed, with probability $1/2$ we are going to be in the $High$ state and the probability of going up is just q instead of $p_t q + (1-p_t)(1-q)$ and the probability of going down is $(1-q)$ instead of $(1-p_t)q + p_t(1-q)$. Also, with probability $1/2$ we are in the Low state and again can simplify recurrence relation (3). Therefore, our utility at $1/2$ is going to be the average of $u_H(p(N))$ and $u_L(p(N))$, where $u_i(p(N))$ are defined by the following recurrence problems: in the $High$ state

(4)
$$\begin{cases} u_H(p(k)) = (1-\delta)q + \delta(qu_H(p(k+1)) + (1-q)u_H(p(k-1))) \\ u_H(p(2N)) = p(N) \ = \frac{q^N}{q^N+(1-q)^N} \\ u_H(p(0)) = p(-N) \ = \frac{(1-q)^N}{q^N+(1-q)^N} \end{cases}$$

and in the Low state

(5)
$$\begin{cases} u_L(p(k)) = (1-\delta)q + \delta((1-q)u_L(p(k+1)) + qu_L(p(k-1))) \\ u_L(p(N)) = 1-p(N) \ = \frac{(1-q)^N}{q^N+(1-q)^N} \\ u_L(p(-N)) = 1-p(-N) \ = \frac{q^N}{q^N+(1-q)^N} \end{cases}$$

Here the boundary conditions come from the fact that when we reach levels N or $-N$ we stop learning and in every period just receive expected utility that is equal to the belief. To sum up the paragraph above, if we can solve two problems (4) and (5), we can get $u(p(0))$.

Theorem 10. *The expected utility of public belief* $1/2$ *when we stop learning upon arrival at levels* $2N$ *or* 0 *has the following form*

$$u\left(\frac{1}{2}\right) = \frac{\left(\frac{q^N}{q^N+(1-q)^N} - q\right) - \left(-\frac{(1-q)^N}{(1-q)^N+q^N} + q\right)\left(\frac{1-q}{q}\right)^N}{a_1^N + a_2^N} + q,$$

and

$$a_i = \frac{1}{2}\left(\frac{1}{\delta q} \pm \sqrt{\frac{1}{\delta^2 q^2} - \frac{4}{q} + 4}\right).$$

To prove this theorem we use recurrence and linear equations techniques. To get the utility function for some other initial level k (with belief $p(k)$) the same technique can be applied.

Even though this may not look very friendly, the numerator has a nice interpretation. The first term is a difference between the public belief at level N and our precision q, which is also our belief at level 1. Similarly, the first multiplier of the second term is a difference between our precision and the public belief at level $-N$. And the second multiplier is the likelihood ratio of observing N signals *Low*.

Given Theorem 10 we can solve for the optimal N and the utility at $1/2$ for any q and δ. For example when $\delta = 0.9$, $q = 0.7$ the expected utility at $1/2$ as a function of stopping level N is dipicted in Figure 3. The optimal $N^* = 4$ and $u(0.5) = 0.802$ in this case. If we do not have any prices this expected utility is 0.7. Furthermore, in the optimal pricing case we end up with probability around $p_t = 0.97$ that we choose the right action (stopping public belief), comparing it to no price case where we end up with $p_t = 0.7$.

Moreover, it is going to be shown in the next subsection that for a fixed δ the optimal N^* does not go to infinity no matter how we change q. This implies that it is impossible to learn the underlying state unless $\delta = 1$.

Furthermore, we not only increase the social welfare but also end up with a much higher public belief about one of the states. In this case, instead of stopping at $q = 0.7$ we are going to stop at $p_t = 0.9674$ conditioned on the public belief being absorbed at the upper bound. Recall, that if we do not have any price then the random walk stops after one (two) steps because a cascade occurs. Given that we have a drift towards the underlying state, making boundaries further away from $1/2$ significantly decreases probability of ending up in the wrong cascade. We have this lemma to formalize it.

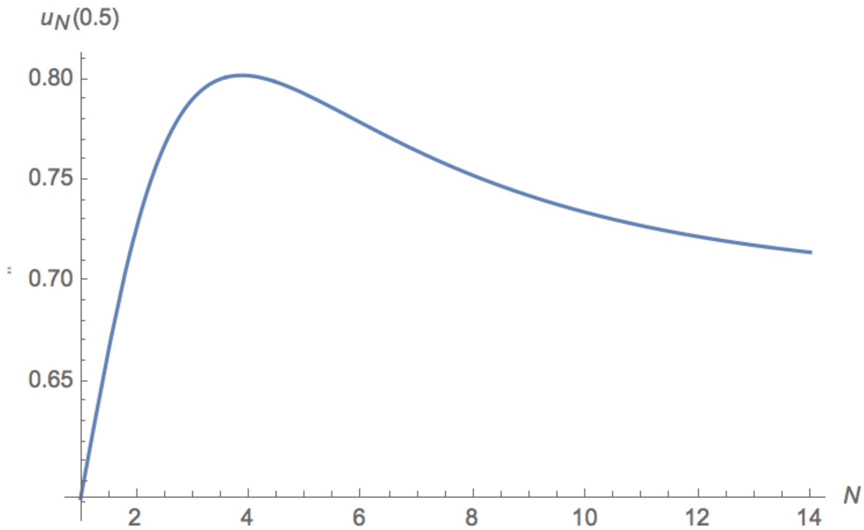


FIGURE 3. Expected utility at $p_0 = 1/2$ as a function of stoping time N.

Lemma 11. *If it is optimal to stop learning phase upon reaching level N or $-N$ then the probability of ending up in the wrong cascade is*

$$\frac{\left(\frac{1-q}{q}\right)^{2N} - \left(\frac{1-q}{q}\right)^{N}}{\left(\frac{1-q}{q}\right)^{2N} - 1}$$

Again, for $q = 0.7$, $\delta = 0.9$ the optimal N is 4 and so the probability of ending up in the wrong cascade is about 0.03, comparing to $1 - q = 0.3$ in a situation without prices. It is easy to see that this probability goes to 0 as $((1-q)/q)^N$. Thus, acquiring information for a few steps dramatically decreases the probability of the wrong cascade.

One more thing should be mentioned before we start analyzing asymptotic properties of $u(p)$, when $\delta \to 1$. As we vary stopping level N, the expected utilities at different levels increase (decrease) simultaneously. This again happens due to stationarity of the social planner's problem.

Proposition 12. *If $u(p(i))$ changes when we vary N then it increases (decreases) iff $u(p(0))$ increases (decreases).*

In the proceeding section we are going to analyze behavior of the optimal N and corresponding $u(0.5)$ when the social planner becomes patient, i.e. $\delta \to 1$.

3.1. **Patient planner.** It seems unlikely that there exists a closed form solution for the optimal N due to the complex form of u and because we are maximizing over natural numbers, which automatically prevents us from using the usual techniques. With numerical calculations this optimal N can be easily found.

However, we can understand the behavior of N for a patient social planner. Let us unfold this statement. Consider the function $u(0.5)$ from Theorem 10 when q is close to 1. If δ is bounded away from 1 then $a_2 > 1$ and as $a_1 > 0$ and all terms in nominator are bounded by $1-q$ $\forall N$, hence, N^* does not go to infinity (stays "finite"). This is a very natural result as if the precision of the private signals is high then we are more likely to get to a high public belief quicker and lose a significant utility for waiting extra rounds. If $q = 0.999$ then after observing just one *High* signal public belief becomes $p_1 = 0.999$ and we know it can never be above 1 so there is no incentive to wait for more signals, the social planner should stop learning now.

On the other hand, if the social planner is extremely patient, i.e. δ goes to 1, then denominator converges to 1 and nominator is an increasing function of N, so N^* goes to ∞. There is also a good intuition for this fact. If he is very patient his utility today matters less and also gaining some small amount, i.e. 0.01 extra, in the public belief can result in a significant increment of the total utility, even if we have to spend 10 extra periods to get it.

A natural question occurs: what happens to N^* if both q and δ go to 1? Does it go to infinity? It turns out that in this regime, N^* is of order $\ln(1-\delta)/\ln(1-q)$. We prove this by first showing that the optimal $N \in \mathbb{R}$ is of this order and then connect it to the optimal $N^* \in \mathbb{N}$.

Proposition 13. *$\exists\, \varepsilon, \gamma > 0$ such that for $q > 1-\varepsilon$ and $\delta > 1-\gamma$ the optimal N^* satisfies the following inequalities for some constants r_1, r_2*

$$r_1 \frac{\ln(1-\delta)}{\ln(1-q)} \geq N^* \geq r_2 \frac{\ln(1-\delta)}{\ln(1-q)}.$$

This result tells us that for N^* to remain constant $(1-\delta)$ must go to 0 as fast $(1-q)^c$ for some constant c. It is an interesting relationship between the precision of private signals and the patience of the social planner.

Another interesting case is when $\delta \to 1$, and q is fixed. We conjecture, that as the planner becomes more patient, absorbing public beliefs become more extreme and $u(0.5)$ gets closer to 1, maximum possible utility. We calculate the rate at which it approaches 1 using a few facts from the proof of Proposition 13.

Proposition 14. *As δ goes to* 1*, utility at belief* $1/2$ *goes to* 1 *with the following rate*

$$u_N(0.5) = 1 - O((1-\delta)\ln(1-\delta)).$$

This result tells us that as we increase δ, and consequently increase N^*, $1 - u(0.5)$ goes almost exponentially quickly to 0.

4. Continuous signal distribution

We now consider a continuous distribution of private signals rather than binary. Continuity gives us more freedom in our actions as the social planner. For example we have a finer trade-offs between the expected loss and expected gain in period t. Also, it will allow us to make the expected loss that occurs due to a non zero price as small as we want.

In the binary case the expected loss is bounded away from 0 unless $k_t = 1$. Recall, that in the binary case the actuall loss, if it occured, was $p_t - q > 0$ if $p_t > 1/2$ ($1 - p_t - q > 0$ if it is less) and increasing with the public belief, when the latter gets closer to the boundary levels. Moreover, when we had non zero prices the loss occured with probability $1-q$, making the expected loss bounded away from 0. This is an aspect of the discrete distributions because even if you have a very high public belief, there is still a probability of making the wrong action (conditioned on being in the learning period) which is bounded away from zero.

The likelihood ratio of the private signal $g_H(x)/g_L(x)$ has cumulative distribution function F_H in the $High$ state and F_L in the Low state. Again, as in the classical model we are going to consider two cases: when signals have bounded and unbounded strength. Let us start with the former one.

4.1. Bounded private signals.

Assume that the private signals are bounded in a sense that[4]

$$1 > \frac{q}{1-q} \geq \frac{g_H(x)}{g_L(x)} \geq \frac{1-q}{q} > 0\cdot$$

So the agents can not get arbitrary strong information about either state. This implies, that if $l_t > q/(1-q)$ and $k_t = 1$ then player t disregards her private signal as her total likelihood ratio is always above 1. Analogously, if $l_t < (1-q)/q$ and $k_t = 1$ then the total likelihood ratio is always below 1.

---

[4] We assume symmetry of private signals without loss of generality.

The main result in this case is similar to the one in the classical literature. Unless we have an extremely patient social planner, $\delta = 1$, it is impossible to learn the underlying state of the world. To put it differently, it is optimal to stop the LP before p_t reaches 1 or 0, as the expected loss is going to be bigger than the expected gain for high enough p_t. Notice, that it also applies to situations when there are finitely many types of private signals, that are not completely revealing.

Theorem 15. *If distributions of the private belief are bounded and* $\delta < 1$ *then there exists* $\overline{p} < 1$ *and* $\underline{p} > 0$ *such that* $\underline{p} \leq p_t \leq \overline{p}$.

To prove this theorem we first bound the expected gain in the public belief and then translate it to the expected gain in the utility. The latter one is compared to the expected loss.

This result can be explained by the choice of the form of utility function, discounted sum. It makes the social planner care more about the current generations rather than the ones that are far away in the future, even if δ is close to 1. Still, we can see from the proof that as δ increases, the public belief is able to get closer to the extreme beliefs. Which makes the society more certain about the realized state. We also conjecture that the expected utility significantly increases from the "no prices" case, as it did in the previous section.

In the next subsection we are going to investigate how prices affect the outcome in the case of unbounded signals. In particular, we are going see what happens to the asymptotic speed of learning comparing to the classical model [10].

4.2. **Unbounded private signals.** Now we would like to see whether the prices in fact increase the asymptotic speed of learning at the same time as they increase the social welfare.

To remind what the asymptotic speed of learning, recall that in the classical model with unbounded signals people eventually start choosing the correct action, 1 if the state is $High$ and 0 otherwise. Then we can see how quickly does l_t converge to the boundary value, i.e. find a function $f(t)$ such that $\lim_{t\to\infty} l_t/f(t) = 1$. This function $f(t)$ is called the *asymptotic speed of learning* or ASL.

To do this analysis in full generality, for any signal distribution, seems to be a very complex problem, so we choose a particular, well-known in the social learning literature, pair of distribution. Consider the following distributions of private signals for the $High$ and the Low

states

$$G_H(x) = x^2$$
$$G_L(x) = 1 - (1-x)^2.$$

Then corresponding distributions of the likelihood ratios are

$$F_H(y) = \mathbb{P}\left(\frac{g_H(s)}{g_L(s)} \le y \middle| \theta = High\right) = \frac{y^2}{(1+y)^2}$$
$$F_L(y) = \mathbb{P}\left(\frac{g_H(s)}{g_L(s)} \le y \middle| \theta = Low\right) = \frac{y^2+2y}{(1+y)^2},$$

for $y \in [0, \infty]$. It is easy to see that this pair of distributions actually correspond to the likelihood beliefs, as $F'_H(y)/F'_L(y) = f_H(y)/f_L(y) = y$[12].

The following theorem tells us that it is indeed profitable for the social planner to choose prices that increase the asymptotic speed of learning. For example if $p_t > 1/2$ then it is *better* to choose some constant price greater than 1 rather than to stick with no price. Furthermore, in this situation it is *not optimal* to choose the price that slows down the asymptotic speed, $k_t < 1$. And, the optimal prices are bounded.

The fact that k_t is greater than 1 and $p_t > 1/2$ implies that the ASL increases in k_t times.

Theorem 16. *For high enough p_t there exists $k > 1$ such that $k_t = k$ gives a higher utility than no price at all. Furthermore, when it is better to choose $k_t > 1$ rather than $k_t = 1$ it is not optimal to choose $k_t < 1$. Moreover, prices are bounded: $\exists\ \underline{k} > 0$ and $\overline{k} < \infty$ such that $\underline{k} < k_t^* < \overline{k}$ for the optimal k_t^*.*

Under some mild conditions on $u'(p)$ we can relax the first statement "for high enough p_t" to "for $p_t > 1/2$".

This theorem does not only provide the desirable result about the increased ASL, as formally stated in Corollary 18, but also a surprising fact: the optimal prices k_t are bounded from 0 and ∞. This implies that as the public belief goes to 1 the optimal price c_t does not go to 1 (equivalent to k_t not going to ∞) in order to extract a lot of information from the action, as it is too costly.

It also implies that we can approximate the optimal outcome with a good precision by fixing some constant price when p_t is close to the boundary values, saving the cost of updating.

To understand why this theorem is very interesting let us look at the big picture. The main problem with continuous signals is that p_t has a very complex behavior and so the analytical solution of the utility function and hence, the optimal pricing policy can not be obtained.

Given that we do not know the utility function we are still able to provide a nice description of the optimal policy: the price should be against the belief, i.e. if $p_t > 1/2$ then $k_t > 1$, some constant pricing will already give us more utility then no prices at all and this prices are bounded. As we will see in the next section, where we calculate $u(p)$ and the optimal k_t^* numerically, this is indeed a very good description of the optimal policy.

There are two main corollaries of Theorem 16. The first one is that the full learning occurs.

Corollary 17. *As t goes to infinity p_t converges to 1 if $\theta = High$ and to 0 if $\theta = Low$.*

For the second one, notice that Theorem 16 tells us that from some point on, the optimal price, k_t, is above 1. If it stays above some $k > 1$, which we will see in the next section is true, then the ASL is going to increase at least by factor k.

Corollary 18. *If the social planner chooses a constant price $k_t = k > 1$ and $\theta = High$ then the asymptotic speed of learning increases by a factor k.*

Therefore, prices indeed increase the social welfare and the ASL at the same time.

In the next section we try a different approach to solve for the utility function u and the optimal pricing policy.

5. Numerical calculations

As we saw above, all theoretical calculations are already fairly complicated and it seems impossible to obtain the analytical solution for (1). Thus, we are going to provide a numerical solution of (1), which reaffirms claims and intuitions that we had before.

Again, suppose that conditional on the underlying state the private likelihood ratio has CDF either $F_H(y)$ or $F_L(y)$ as in previous subsection and $\delta = 0.9$

$$F_H(y) = \frac{y^2}{(1+y)^2}$$
$$F_L(y) = \frac{y^2+2y}{(1+y)^2}.$$

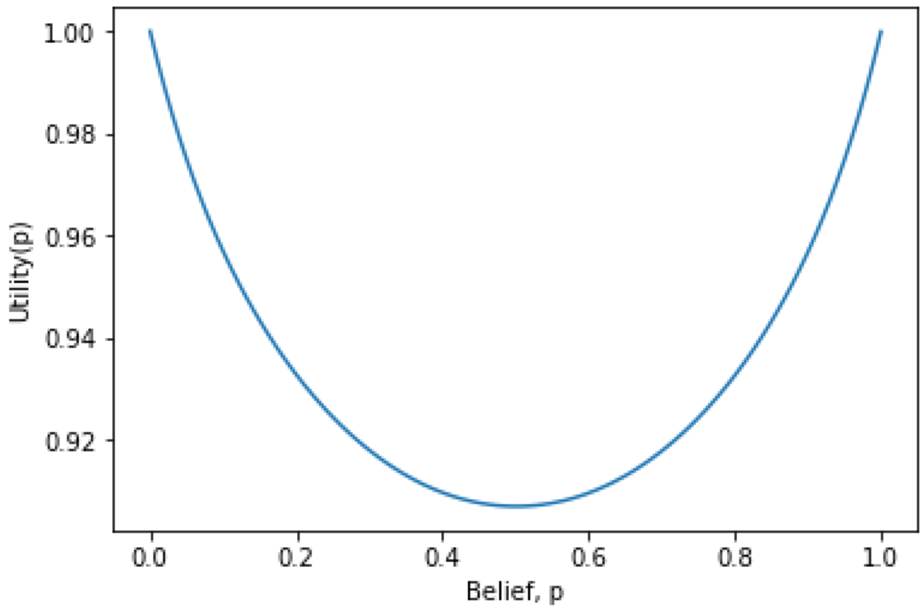


FIGURE 4. Approximation of the utility function with the optimal pricing, $u(p)$.

We are going to look for a solution of the following form

$$\widetilde{u}(p) = \sum_{i=1}^{m} t_i T(i,p),$$

where t_i are constants and $T(i,p)$ is i's Chebyshev polynomial of the first kind. To remind readers $T(0,p) = 1$, $T(1,p) = p$, $T(2,p) = 2p^2 - 1$ and $T(k,p) = 2pT(k-1,p) - T(k-2,p)$.

What we need to do is to find these constants t_i's. This is done by a collocation method [1]. The corresponding coefficients are

$$(t_i)_{i=0}^9 = (136.6, -250.2, 195.8543, -129.0, 70.4, -30.9, 10.4, -2.4, 0.3, 0).$$

The graph of $\widetilde{u}$ is presented in Figure 4.

Let us check that this is a good approximation. To do so we calculate the right hand side of (1) given $u(p)$ and compare the maximal distance between these two functions. It is equal to $8.3 \cdot 10^{-4}$ which is extremely small.

Moreover, we can now see that as public belief approaches 1 optimal k_t increases and is significantly above 1. Here are a few examples

p_t	c_t	k_t	$p_t(c_t)$
0.7	0.12	1.26	0.64
0.8	0.2	1.49	0.73
0.9	0.32	1.94	0.82
0.95	0.4	2.37	0.89
0.99	0.5	3.04	0.97
0.999	0.53	3.28	0.996

Here, $p(c_t)$ is the modified public belief.

We can also calculate the expected utility without any price in a similar way. Let us denote it by $u_{k=1}(p)$. Then for $\delta = 0.9$ the difference in expected utilities as a percentage of $1 - u_{k=1}(p)$ is around 10%. The reason we are normalizing by $1 - u_{k=1}(p)$ and not by 1 or $u_{k=1}(p)$ is that this formula better captures how the expected utility functions flattens. In other words, it becomes more convex comparing to a 45 degree cone $(u_{k=1}(p))$ that we have when there is no price.

This difference may not be as dramatic as one could think. The reason for this is that we apply a significant non zero price ($k_t \neq 1$) for high public beliefs, which results in utility gain. In order to bring back this utility growth to $p_t = 1/2$ from those high public beliefs we need to take a large number of steps. Thus, the gain is significantly discounted. We can observe bigger improvements when we take higher values of δ.

These means that the optimal prices significantly increase the asymptotic speed of learning as well as the social welfare.

6. Conclusion

In this paper we improve the main inefficiencies of the classical sequential learning model. We introduce a social planner whose objective is to maximize the social welfare of the agents by choosing the optimal prices for the new good in each period. We show that the optimal prices indeed increase the social welfare as well as have other positive effects on the public belief. In the case of bounded signals society ends up with a much higher public belief than in the classical case and for the unbounded case it significantly increases the convergence speed.

We manage to provide a complete characterization of the utility function and the optimal strategy in the binary case. Even though calculations are very complex for general distributions of private signals we provide the main properties and description of both the expected utility function and the optimal strategy for a specific continuous distribution. For example, the new product should cost more than the old one if the former one is believed to be better. Furthermore, when the public belief goes to its extreme value (1 or 0) the optimal price is bounded by a constant.

References

[1] *Collocation method*. Available at http://hdl.handle.net/1721.1/64049.

[2] Abhijit V Banerjee, *A simple model of herd behavior*, The Quarterly Journal of Economics (1992), 797–817.

[3] Manaswini Bhalla, *Dynamic pricing under social learning with strategic consumers*.

[4] Sushil Bikhchandani, David Hirshleifer, and Ivo Welch, *A theory of fads, fashion, custom, and cultural change as informational cascades*, Journal of political Economy (1992), 992–1026.

[5] Orosel G. Ottaviani M. Vesterlund L. Bose S., *Dynamic monopoly pricing and herding. the rand journal of economics*, The RAND Journal of Economics (2006), 910–928.

[6] A. Campbell, *Word-of-mouth communication and percolation in social networks*, American Economic Review (2013), 2466–98.

[7] Bimpikis K. Ozdaglar Candogan O., *Optimal pricing in networks with externalities*, Operations Research (2012), 919–939.

[8] Ifrach B. Maglaras C. Scarsini M. Crapis D., *Monopoly pricing in the presence of social learning*, Management Science (2016), 3586–3608.

[9] W. Feller, *An introduction to probability theory and its applications: Volume 1*, Wiley, 1968.

[10] Martynov V. V. Tamuz O Hann-Caruthers W., *The speed of sequential asymptotic learning*, Journal of Economic Theory **173** (2018), 383–409.

[11] Savva N. Papanastasiou Y., *Dynamic pricing in the presence of social learning and strategic consumers*, Management Science (2016), 919–939.

[12] Lones Smith and Peter Sørensen, *Pathological outcomes of observational learning*, Econometrica **68** (2000), no. 2, 371–398.

Appendix A. Binary distributions

The primary goal of this chapter is to prove Theorem 10, which gives the expected utility when the initial public belief equal to $1/2$ and the social planner chooses to stop learning upon arrival at levels N or $-N$. The proof uses some properties of the public belief, when signals are binary. For example, we show that it is a random walk in Lemma 5 and that the utility in the current period, that is in LP, is constant and equals to signal's precision, q. The latter one is showed in Lemma 9. Furthermore, we prove some general facts about the expected utility function and how it changes for different beliefs by varying stopping N.

Proof of Proposition 2. This is a standard fact as each strategy $\{k_t\}_{t=1}^{\infty}$ gives a linear function in belief p and maximum of convex functions is convex.

Moreover, if $p \in \{0, 1\}$ then if we just go with our belief, always choose 1 when $p = 1$ and always choose 0 otherwise, then our discounted utility is already going to be 1 which is the maximal possible utility we can get.

Also, the game is completely symmetric around $1/2$ and so our utility function $u(p)$ is also symmetric around $1/2$.

□

Proof of Lemma 4. If t is in the learning period then player t has to act according to her private signal: if $s_t = High$ choose 1, if $s_t = Low$ choose 0. Otherwise, she either goes against her signal and chooses a non optimal action or chooses the same action regardless of the s_t. Both situations are not allowed either by the statement of the lemma or by the assumption that agents are the expected utility maximizers. □

Proof of Lemma 5. Suppose there were n actions 1 and k actions 0 between times t_0 and t_{n+k}. Moreover, $\forall t' \in \{0, \dots n + k\}$ $t' \in LP$, i.e. players' actions non trivially depended on their private signals. By lemma 4 we know that there were n $High$ signals and k Low ones.

Let us think in terms of the likelihood ratio. As we said before, l_t is going to be multiplied by $q/(1 - q)$ n times and $(1 - q)/q$ k times. These factors do not depend on l_t, therefore it also does not matter in which order we observe these actions, the final likelihood is going to be equal to

$$l_t \cdot \left(\frac{q}{1-q}\right)^n \left(\frac{1-q}{q}\right)^k = l_t \cdot \left(\frac{q}{1-q}\right)^{n-k}.$$

Therefore, the public belief is a random walk on a fixed lattice and its position is defined by the difference in the number of times we observed signals $High$ and Low and $l_0 = 1$.

□

Proof of Lemma 8. First we show that once a we stop learning (no prices, $k_t = 1$) there is no reason to reenact LP latter. Notice that from the social planner's perspective this game is stationary. Therefore, if it is optimal to stop at time t_0 at levels N and $-N$ then at any time $t' > t_0$ there is no incentives to choose k_t different from 1 as we face the same problem as at time t_0 when it is optimal to stop.

Second, suppose we choose to stop upon arrival at either level $h > 0$ or $l < 0$. Then the expected utility at $1/2$ satisfies the following relation

$$u(0.5) = a_l + \mathbb{E}(\delta^t)u(l) + a_h + \mathbb{E}(\delta^t)u(h),$$

where a_h and a_l correspond to the utility that we get on the way to the corresponding boundary and does not depend on u. Let us sum up the first two terms and the last two and choose the biggest one. If they are equal - pick one. Let us say it is the second one, that corresponds to level h. Then due to the symmetry of u it is profitable for the social planner to choose the lower level to be $-h$ instead of l. This concludes the proof.

□

Proof of Lemma 9. Suppose at time t the public belief $p_t > q$ and we are still acquiring information/in the learning period. Depending on the private signal we get, our posterior will be equal to

$$\mu_t = \begin{cases} \dfrac{p_t q}{p_t q + (1-q)(1-p_t)} & \text{if we get the } High \text{ signal} \\ \dfrac{p_t(1-q)}{p_t(1-q) + q(1-p_t)} & \text{if we get the } Low \text{ signal.} \end{cases}$$

Moreover, the former signal happens with probability $(p_t q + (1 - p_t)(1-q))$ as with probability p_t we are in the $High$ state and with $(1-p_t)$ - in the Low state. The analogous calculation gives that the latter signal occurs with probability $p_t(1-q)+q(1-p_t)$. Also, remember when we get the Low private signal we take action 0, so our expected utility is $1 - \mu_t$. Therefore, the expected utility today is equal to

$$\frac{p_t q}{(p_t q + (1-q)(1-p_t))}(p_t q + (1-q)(1-p_t)) +$$
$$+ \left(1 - \frac{p_t(1-q)}{(p_t(1-q) + q(1-p_t))}\right)(p_t(1-q) + q(1-p_t)) = q.$$

And similarly we get the same expected utility if $p_t < 1 - q$.

□

Now we are ready to prove the main result.

Proof of Theorem 10. For notation convenience we write u_n instead of $u_H(p(n))$. We are going to start with solving (4) and then explain why $u_L(0.5)$ from (5) is the same. We know that the solution for the recurrence equation

$$u_n = (1-\delta)q + \delta(qu_{n+1} + (1-q)u_{n-1})$$

is

$$u_n = c_1\left(\frac{1}{2}\left(\frac{1}{\delta q} - \frac{\sqrt{1-4q\delta^2+4q^2\delta^2}}{\delta q}\right)\right)^n + c_2\left(\frac{1}{2}\left(\frac{1}{\delta q} + \frac{\sqrt{1-4q\delta^2+4q^2\delta^2}}{\delta q}\right)\right)^n + q,$$

for some constants c_1 and c_2. It is also straightforward to verify that these u_n's indeed satisfy our recurrence equation. Now we need to solve for c_1 and c_2 using boundary conditions $u_{-N} = (1-q)^N/((1-q)^N + q^N)$ and $u_N = q^N/((1-q)^N + q^N)$. This results in two equations

$$\begin{cases} \dfrac{(1-q)^N}{q^N + (1-q)^N} = c_1 a_1^{-N} + c_2 a_2^{-N} + q \\ \dfrac{q^N}{q^N + (1-q)^N} = c_1 a_1^{N} + c_2 a_2^{N} + q \end{cases}$$

where

$$a_i = \left(\frac{1}{2}\left(\frac{1}{\delta q} \pm \frac{\sqrt{1-4q\delta^2+4q^2\delta^2}}{\delta q}\right)\right).$$

Thus,

$$\begin{cases} c_1 = \left(\dfrac{(1-q)^N}{q^N + (1-q)^N} - q\right) a_1^N - c_2 a_2^{-N} a_1^N \\ c_2\left(a_2^N - a_1^{2N} a_2^{-N}\right) = \dfrac{q^N}{q^N + (1-q)^N} - q + \left(q - \dfrac{(1-q)^N}{q^N + (1-q)^N}\right) a_1^{2N}. \end{cases}$$

It follows that

$$\begin{cases} c_2 = \dfrac{\left(\frac{q^N}{q^N+(1-q)^N} - q\right) + \left(q - \frac{(1-q)^N}{q^N+(1-q)^N}\right) a_1^{2N}}{a_2^N - a_1^{2N} a_2^{-N}} \\ c_1 = \dfrac{-\left(\frac{q^N}{q^N+(1-q)^N} - q\right) a_1^N a_2^N - \left(q - \frac{(1-q)^N}{q^N+(1-q)^N}\right) a_2^{-N} a_1^N}{a_2^N - a_1^{2N} a_2^{-N}} \end{cases}$$

Now we can plug this into our solution

$$
\begin{aligned}
u_0 &= c_1 a_1^0 + c_2 a_2^0 + q \\
&= \frac{-\left(\frac{q^N}{q^N+(1-q)^N} - q\right) a_1^N a_2^N - \left(q - \frac{(1-q)^N}{q^N+(1-q)^N}\right) a_2^N a_1^N}{a_2^N - a_1^{2N} a_2^{-N}} + \\
&+ \frac{\left(\frac{q^N}{q^N+(1-q)^N} - q\right) + \left(q - \frac{(1-q)^N}{q^N+(1-q)^N}\right) a_1^{2N}}{a_2^N - a_1^{2N} a_2^{-N}} \\
&= \frac{(a_2^N - a_1^N)\left(\left(\frac{q^N}{q^N+(1-q)^N} - q\right) - \left(q - \frac{(1-q)^N}{q^N+(1-q)^N}\right) a_1^N a_2^N\right)}{a_2^{2N} - a_1^{2N}} \\
&= \frac{\left(\frac{q^N}{q^N+(1-q)^N} - q\right) - \left(q - \frac{(1-q)^N}{q^N+(1-q)^N}\right)\left(\frac{1-q}{q}\right)^N}{a_2^N + a_1^N},
\end{aligned}
$$

where in the third equation we multiplied everything by $a_2 \neq 0$ and in the last one we used the fact that

$$
a_1 a_2 = \frac{1}{4}\left(\frac{1}{\delta^2 q^2} - \frac{1 - 4q\delta^2 + 4q^2\delta^2}{q^2\delta^2}\right) = \frac{1-q}{q}.
$$

What is left to explain is why $u_L(0.5)$ would be the same as $u_H(0.5)$ which would save us time solving the second analogous system. Before, in state *High*, we had a random walk with a drift q towards the boundary with higher utility, level N. When we condition on state being *Low* we have the same drift but in the opposite direction, towards level $-N$. But notice that utilities of level $-N$ in state *Low* and level N in state *High* are equal to each other. The same is true for the other two boundary utilities. Moreover, in both states we have the same underlying lattice for the random walk. Therefore, problem (5) is the same problem as (4) up to renaming levels. This concludes the proof.

□

Proof of Lemma 11. This is a classic Gambler's ruin problem [9]

□

Proof of Proposition 12 . For notation simplicity we will write $u(n)$ instead of $u(p(n))$.

We know, that the social planner adopts a symmetric strategy of stopping at levels N or $-N$. Suppose that if we increase the stopping boundaries from N to $N+1$ and from $-N$ to $-N-1$ correspondingly such that the utility at level N, $u(N)$, increases. Notice that the utility at level 0, given this strategy, is equal to the utility at the boundary levels multiplied by the expected discount factor plus some utility that we collect on the way to it. The randomness comes from the hitting the boundary levels time. We divide this expression in two parts: for the level N and $-N$.

$$u(0.5) = a_{-N} + u(-N)\,\mathbb{E}_{-N}(\delta^t) + a_N + u(N)\,\mathbb{E}_N(\delta^t),$$

where a_{-N}, a_N are constants $\mathbb{E}_{\pm N}(\delta^t)$ is the expected discounted factor until we hit the corresponding boundary.

Recall that u is symmetric and so $u(-N) = u(N)$, therefore if we increase $u(N)$ then we increase $u(0.5)$ also and vice versa. This concludes the proof.

□

A.1. **Patient planner.** Now we would like to see what happens to the optimal N^* and the expected utility, when the social planner becomes more patient. In other words, how does the optimal N^* and $u(0.5)$ behave, when $\delta \to 1$. We first establish a condition for when N^* does not go to ∞ when $\delta \to 1$ in Proposition 13: precision also has to go to 1 at a certain speed in this case. Secondly, we look at our expected utility as δ increases, but q stays fixed in Proposition 14. There, we stumble upon $(1-q)/q$ factor for the third time.

Proof of Proposition 13. Recall that the expected utility when we start with prior 1/2 (at level N) and stop the random walk upon arrival at levels $2N$ or 0 is

$$u(N) = \frac{-\left(-\frac{(1-q)^N}{(1-q)^N+q^N} + q\right)\left(\frac{1-q}{q}\right)^N + \left(\frac{q^N}{q^N+(1-q)^N} - q\right)}{a_1^N + a_2^N} + q,$$

where $a_1 < a_2$ and

$$a_i = \frac{1}{2}\left(\frac{1}{\delta q} \pm \sqrt{\frac{1}{(\delta q)^2} - \frac{4}{q} + 4}\right).$$

In order to find bounds on optimal $N \in \mathbb{R}$ we are going to focus on the first term of $u(N)$ as the second one is just a constant that does not affect optimal N, and let $q = 1 - \varepsilon$, $\delta = 1 - \gamma$ where $\varepsilon, \gamma \to 0$.

Notice, that $u(N)$ is single-peaked in N, therefore, if N maximizes it over $\mathbb{R}$ and N^* maximizes over $\mathbb{N}$ then N^* is either $\lceil N \rceil$ or $\lfloor N \rfloor$. Thus, bounds on N are going to give very tight bounds on N^*. First, consider a_i:

$$\begin{aligned}
a_i &= \frac{1}{2}\frac{1 \pm \sqrt{1-4\delta^2 q+4q^2\delta^2-4q\delta+4q\delta}}{\delta q}\\
&= \frac{1}{2}\frac{1 \pm \sqrt{(2q\delta-1)^2+4q\delta(1-\delta)}}{\delta q}\\
&= \frac{1}{2}\frac{1 \pm (2q\delta-1+2q\delta\gamma c_1)+o(\gamma)}{\delta q},
\end{aligned}$$

where the thirds equality comes from Taylor series expansion of $\sqrt{a^2+x}$ around 0. Therefore,

$$\begin{aligned}
a_1 &= \frac{1}{\delta q} - 1 - O(\gamma)\\
a_2 &= 1 + O(\gamma)
\end{aligned}$$

Now we can take derivative of g.

$$\begin{aligned}
g'(N) = &\frac{\left(\left(\frac{(1-q)^N \ln(1-q)((1-q)^N+q^N)-(1-q)^N((1-q)^N\ln(1-q)+q^N\ln q)}{((1-q)^N+q^N)^2}\frac{(1-q)^N}{q^N}+m_1\left(\frac{1-q}{q}\right)^N\ln\frac{q}{1-q}\right)\right.}{(a_1^N+a_2^N)^2}+\\
&+\frac{\left.\left(\frac{q^N\ln q((1-q)^N+q^N)-q^N((1-q)^N\ln(1-q)+q^N\ln q)}{((1-q)^N+q^N)^2}\right)\right)(a_1^N+a_2^N)-(a_1^N\ln a_1+a_2^N\ln a_2)m_2}{(a_1^N+a_2^N)^2},
\end{aligned}$$

where

$$\begin{aligned}
q \geq m_1 &= \left(-\frac{(1-q)^N}{(1-q)^N+q^N}+q\right) \geq 2q-1\\
1-q \geq m_2 &= -(-\frac{(1-q)^N}{(1-q)^N+q^N}+q\left(\frac{1-q}{q}\right)^N+\left(\frac{q^N}{q^N+(1-q)^N}-q\right) \geq 0^{5}
\end{aligned}$$

In order to satisfy F.O.C. nominator should be equal to 0

$$\begin{aligned}
&\left(-\frac{(1-q)^N q^N \ln\frac{q}{1-q}}{((1-q)^N+q^N)^2}\frac{(1-q)^N}{q^N}+m_1\left(\frac{1-q}{q}\right)^N\ln\frac{q}{1-q}+\frac{q^N(1-q)^N\ln\frac{q}{1-q}}{((1-q)^N+q^N)^2}\right)(a_1^N+a_2^N)-\\
&-(a_1^N\ln a_1+a_2^N\ln a_2)m_2=0
\end{aligned}$$

[5]As N increases m_1 gets very close to q and m_2 to $1-q$.

Notice that the left-hand side is smaller than

$$\le \left(m_1\left(\frac{1-q}{q}\right)^N \ln\frac{q}{1-q}\right)(c_1 a_2^N) - (a_2^N \ln(1+O(\gamma))m_2$$
$$\le \left(m_1 c_1\left(\frac{1-q}{q}\right)^N \ln\frac{q}{1-q}\right) a_2^N - a_2^N \gamma m_2,$$

where $c_i > 1$.

In order for this to be non negative we need N to satisfy the following constraint

$$N \ge \frac{\ln\dfrac{m_2\gamma}{c_1 m_1 \ln(q/(1-q))}}{\ln\dfrac{1-q}{q}} \ge \frac{r_2 \ln\gamma}{\ln\varepsilon},$$

where r_2 goes to 1 as γ and δ go to 0. At the same time, notice that LHS is bigger than

$$\ge m_1\left(\frac{1-q}{q}\right)^N \ln\left(\frac{q}{1-q}\right) a_2^N - a_2^N \gamma c_3 m_2,$$

where $c_3 > 1$. In order for this to be non positive N should satisfy the following constraint

$$N \le \frac{\ln\dfrac{\gamma c_3 m_2}{m_1 \ln(q/(1-q))}}{\ln\dfrac{1-q}{q}}.$$

Notice, that if $(\ln c_3\gamma)$ is smaller or proportional to $\ln\left(-\ln\left((1-q)/q\right)\right)$ then N is always finite and the lower bound is satisfied when $q, \delta \to 1$, as

$$\frac{\ln\left(-\ln\frac{1-q}{q}\right)}{\ln\frac{1-q}{q}} \to 0$$

Otherwise, we get that

$$N \le \frac{\ln\gamma}{r_1 \ln\varepsilon},$$

for some constant r_1. Moreover, if N satisfies these constraints then N^* satisfies

$$\left\lfloor \frac{\ln\gamma}{r_1 \ln\varepsilon} r \right\rfloor \ge N^* \ge \left\lceil \frac{r_2 \ln\gamma}{\ln\varepsilon} \right\rceil.$$

This concludes the proof.

∎

Proof of Proposition 14. Recall that

$$u(N) = \frac{-\left(-\frac{(1-q)^N}{(1-q)^N+q^N}+q\right)\left(\frac{1-q}{q}\right)^N + \left(\frac{q^N}{q^N+(1-q)^N}-q\right)}{a_1^N+a_2^N} + q.$$

$$a_1 = \frac{1}{(1-\gamma)q} - 1 - O(\gamma)$$

$$a_2 = 1 + O(\gamma)$$

After plugging expressions for a_1 and a_2 in $u(N)$ we get

$$\begin{aligned} u(N) - q &= \frac{-\left(-\frac{(1-q)^N}{(1-q)^N+q^N}+q\right)\left(\frac{1-q}{q}\right)^N + \left(\frac{q^N}{q^N+(1-q)^N}-q\right)}{\left(\frac{1-q}{q}-O(\gamma)\right)^N + (1+O(\gamma))^N} \\ &= \frac{-\left(-\frac{(1-q)^N}{(1-q)^N+q^N}+q\right)\left(\frac{1-q}{q}\right)^N + \left(\frac{q^N}{q^N+(1-q)^N}-q\right)}{\left(\frac{1-q}{q}\right)^N - o(\gamma) + 1 + O(\gamma N)} \\ &= \frac{-\left(-\frac{(1-q)^N}{(1-q)^N+q^N}+q\right)\left(\frac{1-q}{q}\right)^N + \left(\frac{q^N}{q^N+(1-q)^N}-q\right)}{\left(\frac{1-q}{q}\right)^N + 1} - O(\gamma N). \end{aligned}$$

From the proof of Proposition 13 we know that when q is fixed N behaves as $O(\ln\gamma)$. Now, let us transform the first term into a more clear form

$$u(N) - q = \left(\frac{q^N}{q^N+(1-q)^N} - q\right) - \frac{\left(\frac{q^N-(1-q)^N}{q^N+(1-q)^N}\right)\left(\frac{1-q}{q}\right)^N}{\left(\frac{1-q}{q}\right)^N+1} - O(\gamma N)$$

As $N \to \infty$ and $q > \frac{1}{2}$, $((1-q)/q)^N \to 0$. Suppose $((1-q)/q)^N = \nu$ then

$$\frac{q^N}{q^N+(1-q)^N} \geq 1 - \nu$$

and

$$\frac{\left(\frac{q^N-(1-q)^N}{q^N+(1-q)^N}\right)\left(\frac{1-q}{q}\right)^N}{\left(\frac{1-q}{q}\right)^N+1} \leq \frac{\nu}{\nu+1}$$
$$\leq \nu.$$

From the proof of Theorem 13 $\nu = O(\gamma)$. Hence,

$$u(N) - q \geq 1 - q - 2\nu - O(\gamma N) = 1 - q - O(\gamma \ln \gamma).$$

It means that as $N \to \infty$, $u(N)$ goes to its maximal value as $C((1-q)/q)^N \cdot N \to 0$, which is quicker than $((1-q)/q)^{Nk}$, for any $k < 1$. Also, as optimal N^* increases then absorbing beliefs are further away from $1/2$.

□

Appendix B. Continuous distributions

In this section we study continuous distributions. For our purposes they differ in a few aspects from the binary case. First of all, now we have to deal with more complicated expressions for expected loss and gain. Second of all, the set of prices that we can choose and that result in different outcomes is now continuous rather than binary. This also complicates the analysis, especially if we do not have an explicit expression for the expected utility. And the last, and probably the main difference, is that the public belief is not a random walk anymore, but has an intricate behavior.

We start with bounded signals and establish an analogous result, to the one in the classical sequential model [12]: the underlying state of the world is never fully revealed if signals have bounded strength. To prove this, we look at what are the expected loss and gain, when the social planner chooses a price $k_t \neq 1$. As we do not know the exact formula for the expected utility function, we first bound the expected gain in the public belief and then translate it to the expected utility gain. The latter one uses the fact that $u(p)$ is convex and the absolute value of its derivative is less than 1.

B.1. Bounded private signals.

Suppose without loss of generality that the modified LR satisfies $q/(1-q) \geq \frac{p_t}{(1-p_t)k_t} \geq 1/2$, so the *High* state is more likely and we are still in the learning period (less than $q/(1-q)$). Given this, the total modified belief of agent t will be in the interval

$$\left[\frac{p_t}{(1-p_t)k_t}\cdot\frac{1-q}{q}; \frac{p_t}{(1-p_t)k_t}\cdot\frac{q}{1-q}\right].$$

In order to get the final modified likelihood-ratio equal to y in the interval above we need

$$\frac{g_H(x)}{g_L(x)} = y\cdot\frac{(1-p_t)k_t}{p_t}.$$

Before we calculate the expected loss and the expected gain from the price k_t we need a few more facts.

The public belief and the corresponding likelihood ratio satisfy the following relation

$$p_t = \frac{l_t}{1+l_t}.$$

Furthermore, if the total belief $\mu_t > 1/2$ and agent t takes the action 0, then the expected loss is $\mu_t - (1-\mu_t) = 2\mu_t - 1$.

The following lemma helps us understand how the change in the LR translates to the change in the public belief.

Lemma 19. *If $l_{t+1} = l_t(1+\delta)$ then*

$$p_{t+1} - p_t = \delta p_t(1-p_t) + o(\delta(1-p_t)).$$

Proof of Lemma 19. By the definition of the likelihood ratio

$$\frac{p_t}{1-p_t} = x$$
$$\frac{p_{t+1}}{1-p_{t+1}} = x(1+\delta),$$

then

$$\begin{aligned} p_{t+1} - p_t &= \frac{x+x\delta}{1+x+x\delta} - \frac{x}{1+x} \\ &= \frac{x\delta}{(1+x)(1+x+x\delta)} \\ &= \frac{\delta p_t(1-p_t)}{(1-p_t+p_t)(1-p_t+p_t(1+\delta))} \\ &= \delta p_t(1-p_t) + o(\delta(1-p_t)). \end{aligned}$$

□

Furthermore, the following relation is satisfied between the PDFs of the likelihood ratio [12]

$$f_H(x) = x f_L(x).$$

Therefore,

$$
\begin{aligned}
F_H(x) &= \int_{\frac{1-q}{q}}^{x} f_H(y)dy \\
&= \int_{\frac{1-q}{q}}^{x} y f_L(y)dy \qquad (6)\\
&\leq xF_L(x)
\end{aligned}
$$

These facts help us prove one of the main results, that if private signals are bounded then the full revelation of the underlying state is not possible, unless $\delta = 1$.

Proof of Theorem 15. Let us start with the expected loss which is equal to

$$
\begin{aligned}
&\int_a^b \left(2\frac{x\frac{p_t}{1-p_t}}{1+x\frac{p_t}{1-p_t}} - 1\right)(p_t f_H(x) + (1-p_t)f_L(x))dx \\
&\geq c_1 p_t \left(F_H\left(\frac{1-p_t}{p_t}k_t\right) - F_H\left(\frac{1-q}{q}\right)\right) + (1-p_t)\left(F_L\left(\frac{1-p_t}{p_t}k_t\right) - F_L\left(\frac{1-q}{q}\right)\right) \\
&= c_1\left(p_t F_H\left(\frac{1-p_t}{p_t}k_t\right) + (1-p_t)F_L\left(\frac{1-p_t}{p_t}k_t\right)\right),
\end{aligned}
$$

where $a = (1-q)/q$, $b = (1-p_t)k_t/p_t$ and c_1 gets arbitrary close to 1 as p_t goes to 1. Notice that we can make $(1-p_t)k_t/p_t$ as close to $(1-q)/q$ as we want, which is the lowest value for the likelihood ratio.

Now let us find the expected gain. Remember that the modified public belief is $p_t/((1-p_t)k_t)$. Hence we get the following expressions for $p_{t+1}/(1-p_{t+1})$

$$
\overline{\left(\frac{p_{t+1}}{1-p_{t+1}}\right)} = \frac{p_t}{1-p_t}\cdot\frac{1-F_H\left(\frac{(1-p_t)k_t}{p_t}\right)}{1-F_L\left(\frac{(1-p_t)k_t}{p_t}\right)}
$$

if player t buys the new product (random walk goes up) and

$$
\underline{\left(\frac{p_{t+1}}{1-p_{t+1}}\right)} = \frac{p_t}{1-p_t}\cdot\frac{F_H\left(\frac{(1-p_t)k_t}{p_t}\right)}{F_L\left(\frac{(1-p_t)k_t}{p_t}\right)}
$$

if she does not (random walk goes down). Here, underline and overline represent possible LRs that can result from the current public belief p_t and price k_t depending on the action of player t.

As utility function on $[0.5, 1]$ is convex and symmetric around 0.5 we know that expected gain is less than $u(\overline{p_{t+1}}) - u(p_t)$ (as if we go up with probability 1). Which in its turn is less than $\overline{p_{t+1}} - p_t$ as u is also above 45 degree line and at 1 is equal to 1. In order to calculate the latter we should find $\overline{(p_{t+1}/(1-p_{t+1}))}$. We are thinking about $(1-p_t)k_t/p_t$ as a point close to the left end of the domain of F_H and F_L, so in the following calculation we are using notation $\varepsilon = (1-p_t)k_t/p_t$.

$$\begin{aligned}
\overline{\left(\frac{p_{t+1}}{1-p_{t+1}}\right)} &= \frac{p_t}{1-p_t}\left(\frac{1-F_H\left(\frac{(1-p_t)k_t}{p_t}\right)}{1-F_L\left(\frac{(1-p_t)k_t}{p_t}\right)}\right) \\
&\leq \frac{p_t}{1-p_t}\big(1-c_2F_L(\varepsilon))(1+F_L(\varepsilon)+O(F_-^2(\varepsilon))\big) \\
&= \frac{p_t}{1-p_t}(1+F_L(\varepsilon)(1-c_2)+O(F_L^2(\varepsilon)))
\end{aligned}$$

where $c_2 > 0$ is a constant comes from (6). And hence,

$$\begin{aligned}
\delta(\overline{p_{t+1}} - p_t) &\leq \delta\frac{F_L(\varepsilon)(1-c_2)(1-p_t)}{p_t} + O(F_L^2(\varepsilon)) \\
&= \frac{F_L(\varepsilon)(\delta(1-p_t+p_tc_2-c_2))}{p_t} + O(F_L^2(\varepsilon)).
\end{aligned}$$

Comparing it to $F_L(\varepsilon)((1-\delta)(1-p_t+p_tc_2))c_1$ tells us that for $p_t \to 1$ coefficient $\delta(1-p_t+p_tc_2-c_2)/(p_t)$ goes to 0 whereas the other one goes to $(1-\delta)c_2 > 0$. $\square$

B.2. **Unbounded private signals.** Now we are going to consider an example of unbounded private signals. This is the only situation, when it is appropriate to talk about the asymptotic speed of learning, as in other cases the public belief does not converge to 0 or 1. The second main result is Theorem 16, which says that the optimal prices, k_t, are bounded away from 0 and ∞ and that it is optimal to choose a positive price, $k_t > 1$, when the public belief is high enough. We need to assume that p_t is high due to the lack of information about u and its derivative. As we mentioned above, if we use the results from the numerical calculation, it can be generalized for $p_t \geq 1/2$.

One of the corollaries is that we indeed have the asymptotic learning and the asymptotic speed of learning is also significantly increased, due to taxation of the good that is more likely to be better. The situation when $p_t < 1/2$ is symmetric.

Recall that the CDFs of the likelihood ratios are

$$F_H(y) = \mathbb{P}\left(\frac{g_H(s)}{g_L(s)} \le y \middle| \theta = High\right) = \frac{y^2}{(1+y)^2}$$
$$F_L(y) = \mathbb{P}\left(\frac{g_H(s)}{g_L(s)} \le y \middle| \theta = Low\right) = \frac{y^2+2y}{(1+y)^2},$$

for $y \in [0, \infty]$.

Proof of Theorem 16. Recall that if we apply non zero price k_t then we have the expected loss due to non optimal actions and the expected gain from bigger expected increase in public belief p_{t+1}. We start with the former one.

Suppose we have a public belief p_t and a price at this period is k_t. As was stated in the Section 2, agent t is going to buy the new product **iff**

$$\frac{p_t}{(1-p_t)k_t} \cdot \frac{g_H(x)}{g_L(x)} \ge 1.$$

Notice, that a player t will take a non optimal action only if her likelihood belief with price k_t is less than 1 and her likelihood belief without the price is above 1. More formally private likelihood ratio can take the following values

$$\frac{(1-p_t)k_t}{p_t} \ge \frac{g_H(x)}{g_L(x)} \ge \frac{1-p_t}{p_t}$$

If a person t gets a private likelihood belief in this interval and her total likelihood (without accounting for the price) is equal to y then the loss, due to taking a non optimal action, is equal to

$$\frac{2y}{1+y} - 1.$$

Therefore, expected loss $l(p_t, k_t)$ has the following form

$$\begin{aligned}
l(p_t, k_t) &= \int_a^b \left(\frac{2x\frac{(1-p_t)}{p_t}}{x\frac{(1-p_t)}{p_t}+1} - 1\right)(p_t f_H(x) + (1-p_t) f_L(x))\, dx \\
&\le \left(\frac{2k_t}{k_t+1} - 1\right)\int_a^b (p_t f_H(x) + (1-p_t) f_L(x))\, dx \\
&= \left(\frac{k_t-1}{k_t+1}\right)\left(p_t\left(F_H\left(\frac{(1-p_t)k_t}{p_t}\right) - F_H\left(\frac{1-p_t}{p_t}\right)\right) + \right. \\
&\left. + (1-p_t)\left(F_L\left(\frac{(1-p_t)k_t}{p_t}\right) - F_L\left(\frac{1-p_t}{p_t}\right)\right)\right),
\end{aligned}$$

where $a = \frac{(1-p_t)}{p_t}$ and $b = \frac{(1-p_t)k_t}{p_t}$. After pluging in the expressions for F_i we get

$$l(p_t, k_t) = \frac{((-1+k_t)^2 p_t^2 (1-p_t)^2)}{(k_t + p_t - k_t p_t)^2}.$$

Now let us go to the expected gain term. Suppose that in period t the public belief $p_t > 1/2$ is high enough. If there is no price, $k_t = 1$, then the public belief in the next period goes to either a or b depending on the action of player t (buying the new and the old products correspondingly). And if we apply a price k_t then the public belief in the next period goes to a' and b' correspondingly

$$a = \frac{2-p_t}{3-2p_t},\ b = \frac{p_t}{(1+2p_t)},$$
$$a' = \frac{(2k_t(1-p_t)+p_t)}{(2k_t(1-p)+1)},\ b' = \frac{k_t p_t}{(k_t+2p_t)}$$

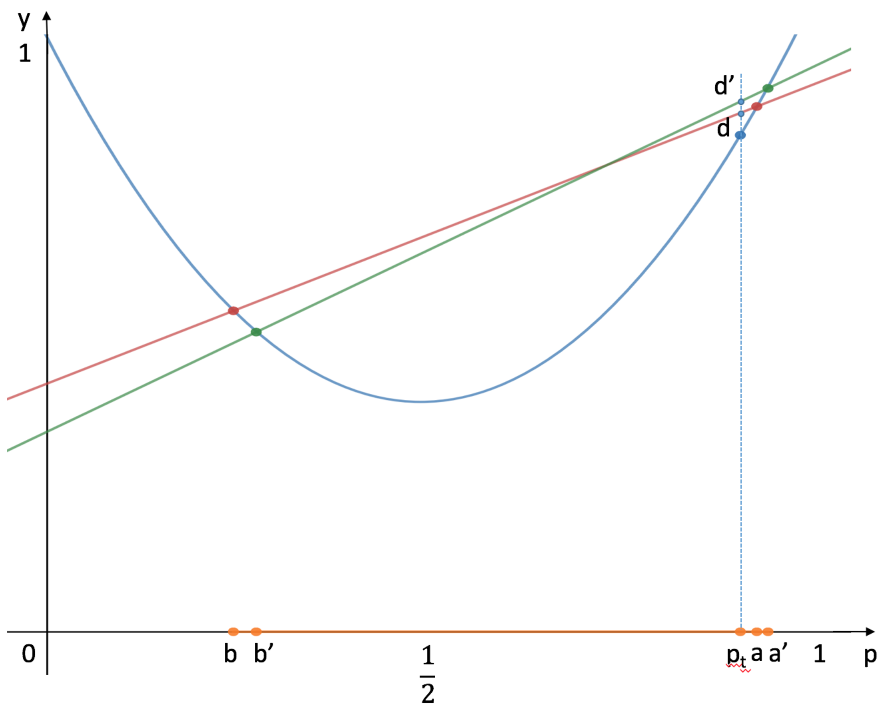
As the public belief is a martingale then in expectation it does not move. This means, that the expected gain is equal to the distance between two chords that connect $u(a), u(b)$ and $u(a'), u(b')$ at point p_t. This is depicted in Figure 5. In other words, the expected gain is equal to the distance between d and d', where d belongs to the chord $u(a), u(b)$ and d' to the chord $u(a')u(b')$ and their p-coordinate is p_t. So the distance between d' and d is alongside the y-coordinate.

Let us calculate this distance. Denote by y_a and y_b - y-coordinates of a and b. Notice, that the y-coordinate of a' and b' are equal to $y_a + a_1(a'-a)$ and $y_b - b_1(b'-b)$ where a_1, b_1 are some constants that are in $[u'(a), u'(a')]$ and $[-u'(b'), -u'(b)]$.

$$d = y_b + \frac{y_a - y_b}{a-b}(p_t - b)$$
$$d' = y_b - b_1(b'-b) + \frac{y_a - y_b + a_1(a'-a) + b_1(b'-b)}{a'-b'}(p-b').$$

Pluging in the expressions for a, b, a', b' gives us the following formula

$$\text{expected gain}(p_t, k_t) = d' - d = \frac{((2a_1(-1+k_t)(-1+p_t)^2 p_t^2)}{((3-2p_t)(k_t+p_t-k_t p_t)^2))} - \frac{(2b_1(-1+k_t)k_t(-1+p_t)^2 p_t^2)}{((1+2p_t)(k_t+p_t-k_t p_t)^2)} +$$
$$+ \frac{((-1+k_t)(-1+p_t)^2 p_t^2(-1-3k_t+2(-1+k_t)p_t)(y_a-y_b))}{(k_t+p_t-k_t p_t)^2} =$$
$$= \frac{(-1+k_t)(1-p_t)^2 p_t^2}{(k_t+p_t-k_t p_t)^2}\left(\frac{2a_1}{3-2p_t} - \frac{2b_1 k_t}{1+2p_t} - (1+3k_t-2(-1+k_t)p_t)(y_a-y_b))\right).$$



Where d and d' are the y-coordinates of point p_t on the cords $u(a)u(b)$ and $u(a')u(b')$ correspondingly.

FIGURE 5. The expected gain induced by the cost.

Denote by P the terms in the parenthesis

$$P(p_t, k_t) = \left(\frac{2a_1}{3-2p_t} - \frac{2b_1k_t}{1+2p_t} - (1+3k_t-2(-1+k_t)p_t)(y_a-y_b))\right).$$

Notice, that $y_a \geq y_b$ as a is closer to 1 than b is to 0 for $p_t \geq 1/2$

$$\frac{2-p_t}{3-2p_t} - \left(1 - \frac{p_t}{1+2p_t}\right) = \frac{2p_t-1}{(3-2p_t)(1+2p_t)} \geq 0.$$

Let us now compare the expected gain and the expected loss due to the price k_t. We do this by subtracting the former one by the latter and taking the discount factor into account

$$\frac{\delta}{1-\delta}\text{expected gain} - \text{expected loss} = \frac{(-1+k_t)(1-p_t)^2p_t^2}{((k_t+p_t-k_tp_t)^2}\cdot$$

$$\left(\frac{\delta}{1-\delta}\left(\frac{2a_1}{3-2p_t} - \frac{2b_1k_t}{1+2p_t} - (1+3k_t-2(-1+k_t)p_t)(y_a-y_b)\right) - (k_t-1)\right) =$$

$$= \left(\frac{\delta}{1-\delta}P(p_t,k_t) - k_t + 1\right) \cdot \frac{(-1+k_t)(1-p_t)^2p_t^2}{((k_t+p_t-k_tp_t)^2}$$

We can see that the sign of this expression is defined by the parenthesis and $k_t - 1$ terms.

Suppose that the expression in the parenthesis is positive for some $k_t > 1$ (and $p_t \notin \{0, 1\}$)

(7)
$$\frac{\delta}{1-\delta}\left(\frac{2a_1}{3-2p_t} - \frac{2b_1k_t}{1+2p_t} - (1+3k_t-2(-1+k_t)p_t)(y_a-y_b)\right) - (k_t-1) > 0$$

Then $d' - d$ is positive as the sign here is defined by this parenthesis and $(k_t - 1)$. If we now make k less then 1, the expression above will increase and $k_t - 1$ will become negative, which will make the expected gain also negative. Therefore, if (7) holds for some $k_t > 1$ then the optimal k_t^* is not less than 1.

Now let see whether there exists $k_t > 1$ which gives higher utility than $k_t = 1$ for high enough p_t. As we said before, it is enough to show that (7) holds for some $k_t > 1$. The main challenge here is that we do not the utility function u and how convex it is.

Still, we can say the following: $y_a - y_b < a_1(a-b) < a_1/3$ and $b_1 < a_1$. This means that for p_t close enough to 1

$$P(p_t, k_t) > \left(\frac{2a_1}{3-2p_t} - \frac{2a_1k_t}{1+2p_t} - (1+3k_t-2(-1+k_t)p_t)\frac{a_1}{3})\right).$$

For $k_t = 1$ this function is increasing in p_t and at $p_t = 1$ it is 0. The expected loss is also 0 for $k_t = 1$. But as the bounds for $y_a - y_b$ and b_1 are not tight, $P > 0$ at 1. As P and $-(k_t - 1)$ are also continuous functions of k_t there exists a threshold p such that for any $p_t > p$ the expected gain is positive for some $k_t = 1 + \varepsilon$. In other words, there exists $k_t > 1$ such that it is better to choose than $k_t = 1$ for p_t high enough. Moreover, if we use tighter constraints that we get from section 5 we can see that it holds for $p_t > 1/2$.

Furthermore, there exists neighborhoods of 0 and ∞ such that the optimal k_t^* does not lie in them for any p_t.

Let us start with 0

$$\begin{aligned} P(p_t, 0) =& \frac{2a_1}{3-2p_t} - (1+2p_t)(y_a-y_b) \\ &\geq \frac{2a_1}{3-2p_t} - \frac{(1+2p_t)a_1}{3}, \end{aligned}$$

where the inequality comes from the fact that $(y_a - y_b) \leq a_1(a-b) \leq a_1/3$. Furthermore,

$$\frac{2}{3-2p_t} > \frac{1+2p_t}{3},$$

as $(3-2p_t)(1+2p_t) \leq 4$. Thus, for k in some neighborhood of 0 $P(p_t,k)>0$ and $k-1<0$. Hence, there exists $\underline{k}>0$ such that $k^*>\underline{k}$.

Also, it is obvious that for high enough k_t, $P(p_t,k_t)$ is negative. There is another case when b' jumps to the right side of $1/2$ such that y_b' becomes higher than y_b. Then we have a bit different expression for d'

$$d' = y_b + b_1(b'-(1-b)) + \frac{y_a - y_b + a_1(a'-a) - b_1(b'-(1-b))}{a'-b'}(p-b').$$

And hence,

$$d' - d - \text{expected loss}(p_t,k_t) = \frac{(-1+p_t)^2}{(k_t+p_t-k_tp_t)^2}\left(\frac{2a_1(-1+k_t)p_t^2}{(3-2p_t)} + \right.$$
$$\left. + \frac{b1k_t(-2p_t(1+p_t)+k_t(-1+2p_t^2))}{1+2p_t} + (-1+k_t)p_t^2(-1-3k_t+2(-1+k_t)p_t)(y_a-y_b)) - \right.$$
$$\left. -(k_t-1)^2p_t^2\right).$$

Define the term in the big parenthesis by $P'(p_t,k_t)$ and notice that

$$P'(p_t,k_t) \leq \left(2a_1(k_t-1)p_t^2 + \frac{b_1k_t\cdot k_t(2p_t^2-1)}{2} - (k_t-1)^2p_t^2\right).$$

If we look at this as a function of k_t then it is a downward looking parabola, which implies that for k_t high enough P' is negative and, as a consequence, the expected gain is smaller than the expected loss. Thus, k_t does not go to ∞ as p_t increases.

Therefore there exist $\overline{k}>1$ such that $k^*<\overline{k}$. This concludes the proof.

□

Proof of Corollary 17. This is an implication of two facts. The first one is that prices are bounded, so it is impossible to stop aggregating information. And second, public belief is a martingale and hence, converges. For more details see [12].

□

Proof of Corollary 18. Suppose at time t the likelihood ratio is l_t and $\theta = High$. Let us calculate l_{t+1} when agent t buys the new product with price k

$$l_{t+1} = l_t \cdot \frac{1-F_H\left(\frac{k}{l_t}\right)}{1-F_L\left(\frac{k}{l_t}\right)} = l_t\left(2\frac{k}{l_t}+1\right) = 2k + l_t.$$

This means that when everybody start taking the same correct action instead of adding 2 to the LR we going to add $2k$. Therefore, the convergence speed increases in k times.

$\square$

Sub optimal strategy: for k signals (where we optimize k later a bit) : choose prices such that modified public belief is 1/2 and calculate conditional (on state? for some class of distributions) probability that you go up or down and compare it to expeceted utility without prices.

California Institute of Technology