Learning about the arrival of sales

Robin Mason*   Juuso Välimäki [†]

9 July 2010

Abstract

We propose a simple model of optimal stopping where the economic environment changes as a result of learning. A primary application of our framework is the problem of how optimally to sell an asset, when the demand for that asset is initially uncertain. In the model that we consider, the seller learns about the arrival rate of buyers to the market. As time passes without a sale, the seller becomes more pessimistic about the arrival rate. When the seller does not observe the arrival of a buyer to the market, the rate at which the seller revises her beliefs is affected by the price she sets. We show that learning then leads to a higher posted price by the seller. When the seller does observe the arrival of buyers, she sets an even higher price.

1 Introduction

Models of optimal stopping capture the trade-off between known current payoffs and the opportunity cost of future potentially superior possibilities. In many of their economic applications, the environment in which the stopping decisions are taken is assumed to be stationary. In the standard job search model, for example, a rejected offer is followed by another draw from the same distribution of offers. See Rogerson, Shimer, and Wright (2005) for a survey. The predictions of these models are at odds with empirical observations. Prices of houses decline as a function of time on the market; see e.g., Merlo and Ortalo-Magne (2004, p. 214). Reservation wages of unemployed workers decline in the duration of the unemployment spell; see e.g., Lancaster and Chesher (1983).

We propose a simple and tractable model of optimal stopping where the economic environment changes as a result of learning. Rather than using the deterministic model of optimal stopping typically used in optimal search models, we generalize the other canonical model of stopping in economic theory, where stopping occurs probabilistically (as in e.g., models of R&D).

We cast the model in the language of the classic problem of how to sell an asset (Karlin (1962)). (In the concluding section, we discuss alternative interpretations of the model.) A seller posts a take-it-or-leave-it price; potential buyers arrive according to a Poisson process and observe the posted price. They buy if and only if their valuation exceeds the posted price. The seller is uncertain about the arrival rate of buyers. The seller does not observe whether a buyer is present: she sees only whether a sale occurs or not. The seller updates her beliefs about the arrival rate in a Bayesian fashion, becoming more pessimistic about future arrivals after each period when a sale does not occur. The rate at which beliefs decline depends on the current posted price. Hence, when choosing the current price of her asset, the seller controls her immediate expected profit, conditional on a sale, as well as the beliefs about future demand if no sale occurs.

Even though our model is a standard Bayesian learning model, the belief dynamics are a little unfamiliar. When we describe the evolution of beliefs over time, we are implicitly conditioning on the event that no sale occurred in the current period. Since this event

1

is bad news about the prospect of a future sale, the seller becomes more pessimistic over time. The seller can control the downward drift of her beliefs through her choice of price.

We identify two key effects. A more pessimistic future implies a lower current value of being a seller and hence a lower current price. We call this feature the *controlled stopping effect*. By setting a lower price, the seller can cause her beliefs to fall further in the event of no sale occurring. A more pessimistic seller has a lower value than a more optimistic seller—there is a capital loss from learning. Holding fixed the probability of a sale occurring, the seller has an incentive to increase her price in order to reduce this capital loss. We call this feature the *controlled learning effect*.

Which effect dominates? Does learning cause the seller to raise or lower her price, relative to the case when no learning occurs? The benchmark is a model where the rate of contact between the seller and buyers is fixed at her current expected level. We show that in our model, the controlled learning effect dominates: the optimal posted price exceeds the price posted in the equivalent model with no learning.

We gain further insight into the effect of learning by looking at the case where the seller does observe the arrival of buyers. In this case, the updating of the seller's beliefs is independent of the price that she sets. We show that the seller optimally sets an even higher price than when she cannot observe arrivals. The key difference when arrivals are observed is that now the seller can raise her price after positive news i.e., after a buyer arrival, even when there is no sale. We therefore obtain an upper bound for the price increase resulting from controlled learning, when arrivals are not observed: it is no larger than the price increase resulting from the observability of arrivals.

A number of papers have studied the problem of a monopolist who learns about the demand it faces through its pricing experience: see Rothschild (1974), McLennan (1984), Easley and Kiefer (1988), Aghion, Bolton, Harris, and Jullien (1991) and Mirman, Samuelson, and Urbano (1993).[1] The general difficulty in analysing this type of situation is the updating that occurs when the seller observes that the object has failed to sell at a

---

[1]In these papers, the demand curve is fixed and the learning process narrows down the initial uncertainty about it. In other papers, the object of the learning is not constant, but changes over time in some random way. See, for example, Balvers and Cosimano (1990), Rustichini and Wolinsky (1995) and Keller and Rady (1999).

particular price. There are certain valuation distributions for which updating can be done analytically. Even for these cases, the amount of progress that can be made is minimal.

A notable exception is Trefler (1993), who provides a characterization of the expected value of information and the direction of experimentation in several cases of learning by a monopoly seller. There are two main differences between our work and Trefler's. First, we consider the sale of a single item; as a result, the monopolist faces a stopping problem (albeit one where stopping occurs probabilistically). In Trefler's model, the true (but unknown) economic environment remains stationary across periods, and the model is thus one of repeated sales. Secondly, in out set-up, the Bellman equations are first-order differential equations; they are therefore particularly easy to analyse. This allows us to give more characterization than Trefler of the seller's optimal policy.

Our paper is also related to the literature on search with learning. Burdett and Vishwanath (1988), Van den Berg (1990) and Smith (1999) analyse an optimal stopping problem when wage offers come from exogenous sources. Anderson and Smith (2006) analyse a matching model where each match results in a current immediate payoff and reveals new information about the productivity of the partners. In all these papers, beliefs change i.e., the environment is non-stationary; but the change in beliefs is not affected by the action chosen by the economic agent. We develop a tractable framework in which agents affect their learning through their actions.

Finally, there are other papers, from a number of different fields, that involve learning about a Poisson process. Amongst others, see Driffill and Miller (1993), Malueg and Tsutsui (1997), Keller, Rady, and Cripps (2005), Bergemann and Hege (2005), and Bonatti and Horner (2010).

The paper is organised as follows. Section 2 presents the model. In section 3, we look at a couple of benchmarks without learning, which are useful when assessing the effects of learning in the model. Section 4 characterises the optimal pricing policy when learning takes place. Section 5 looks at the case when the seller can observe the arrival of buyers to the market. Section 6 discusses alternative interpretations of the model and concludes. The appendix contains longer proofs.

2 The model

Consider a seller setting a sequence of take-it-or-leave-it prices over time for a single unit of a good of known quality. Time is discrete. Time periods are denoted by $t = 0, 1, \ldots, \infty$; each time interval is of length $dt > 0$, which we take to be arbitrarily short. It is commonly known that the seller's valuation is 0. The seller announces at the start of each period the price for that period. If a buyer announces that she is willing to buy at that price, then the seller sells and the problem stops. The seller is infinitely lived and discounts the future at rate r.

The probability that the seller encounters a buyer in a given period depends on two factors: whether a buyer is present during the period; and whether that buyer is willing to pay the seller's posted price.

Buyers arrive randomly to the market according to a Poisson process. With an arrival rate λ, in any given time interval $(t, t + dt]$, there is a probability λdt that a buyer arrives to the market; and a probability of $O((dt)^2)$ that more than one buyer arrives to the market. Since we take the continuous-time limit, the realization of two buyers arriving is ignored. There are two states. In the low state L, the Poisson arrival rate is $\lambda_L > 0$. In the high state H, the Poisson arrival rate is $\lambda_H > \lambda_L$. Let $\Delta\lambda := \lambda_H - \lambda_L$. Each buyer is present (if at all) for one period, and then disappears.

Buyers' valuations are drawn independently from a distribution $F(\cdot)$ with a density function $f(\cdot)$. Given that the buyers are short-lived, they purchase the good immediately if and only if their valuation exceeds the price. Conditional on a buyer being present, the probability of this is $1 - F(p)$. We assume that $F(p)$ has an increasing hazard rate.[2]

The seller does not observe whether a buyer is present: it observes only when a sale occurs. (We comment further on this below.) So in the seller's view, the probability of a sale in any given period, given an arrival rate $\lambda \in \{\lambda_L, \lambda_H\}$, is $\lambda dt(1 - F(p))$.

The state variable in this problem is the seller's scalar belief $\pi \in [0, 1]$ that the state of the world is high. The only event that is relevant for updating the seller's beliefs is

---

[2]We could allow the overall arrival rate of buyers to respond to the seller's price, with little change to the analysis.

when no sale has occurred in a period. In the continuous time limit $dt \to 0$,

$$d\pi(t) := \pi(t + dt) - \pi(t) = -\pi(t)(1 - \pi(t))\Delta\lambda(1 - F(p))dt \leq 0. \tag{1}$$

So the seller becomes more pessimistic about the state of the world when no sale occurs. The seller affects the updating of beliefs through the price that it sets: a high price implies a higher probability of no sale occurring, and a higher posterior.

One modelling assumption deserves further comment: the seller observes only sales and not the arrivals of buyers. (For example, the seller places an advertisement in a newspaper each period; the seller cannot tell if a buyer has seen the advertisement.) This assumption ensures two things. First, the probability of stopping, $\lambda dt(1 - F(p))$, is of order dt. Secondly, the change in beliefs $d\pi$ also is of order dt. In the continuous-time limit, as $dt \to 0$, this ensures that the dynamic programming problem yields a differential, rather than a difference equation; this makes the analysis much simpler. In Section 5, we compare this model to the one where arrivals are observed.

3   A non-learning benchmark

We start with a benchmark intertemporal model where the seller's belief π does not change over time. Since the problem remains the same in all periods, the dynamic programming problem of the seller is particularly easy to solve. This solution will serve as a natural point of comparison to the model that incorporates learning i.e., with π changing over time. We refer to this case as the *repeated* problem. (See also Keller (2007); our repeated benchmark corresponds to his "naive learner" case.)

For the repeated problem, the seller's Bellman equation is

$$V_R(\pi) = \max_p V_R(p, \pi) \tag{2}$$

where

$$V_R(p, \pi) := \left\{ \frac{\lambda(\pi)(1 - F(p))p}{r + \lambda(\pi)(1 - F(p))} \right\}.$$

The first-order condition for the repeated problem is

$$p - \frac{1 - F(p)}{f(p)} = V_R(\pi).$$ (3)

For any π, equation (3) has a unique solution, because $F(\cdot)$ has an increasing hazard rate. Denote this solution as $p_R(\pi)$.

It will be helpful for later arguments to establish a couple of facts about $p_R(\pi)$. First, denote the static price as p_S; it solves $p_S - (1 - F(p_S))/f(p_S) = 0$. Since $V_R(\cdot) \geq 0$, the following result is immediate.

Proposition 1 *The repeated price is greater than the static price: $p_R(\pi) \geq p_S$ for all $\pi \in [0, 1]$.*

The intuition is straightforward: the prospect of continuation in the repeated problem acts as an opportunity cost to a current sale; see equation (3). This opportunity cost leads the seller to set a higher price.

Next, we establish the monotonicity of the repeated price $p_R(\pi)$.

Proposition 2 *The repeated price $p_R(\pi)$ is non-decreasing in the level of the seller's belief π for all $\pi \in [0, 1]$.*

Proof. Total differentiation of equation (3) means that $p_R(\pi)$ is non-decreasing in π iff $V_R(\pi)$ is non-decreasing in π. Since the value in any program is realized only at the stopping moment, the value of any fixed policy must be increasing in the arrival rate and hence π. Hence the optimal policy starting at π yields a higher value when started at state $\pi' > \pi$ and as a consequence $V_R(\pi') > V_R(\pi)$. ∎

The proof relies on two facts: the prospect of continuation acts as an opportunity cost to a current sale; and the value of continuation increases with the belief π. Hence the repeated price increases in π.

4 The effect of learning

Consider next the model with learning: the *dynamic* problem. The crucial difference compared to the repeated problem is that now π evolves over time according to equation (1). The seller anticipates this and adjusts the optimal price to control both the current probability of a sale occurring and the future value of sales.

We show in the appendix (see Lemma 1) that $V_D(\pi)$ is convex; as a result, $V_D(\pi)$ is differentiable almost everywhere. Hence we can use a Taylor series expansion for $V_D(\pi + d\pi)$, and take the continuous-time limit $dt \to 0$, to write the Bellman equation as

$$V_D(\pi) = \max_p \left\{ V_R(p, \pi) - \frac{\Delta\lambda(1 - F(p))\pi(1 - \pi)V_D'(\pi)}{r + \lambda(\pi)(1 - F(p))} \right\}. \tag{4}$$

(See the appendix for a derivation of this equation.) Equation (4) shows that the dynamic value function is equal to a repeated value, $V_R(p_D, \pi)$ (evaluated at the optimal dynamic price, p_D), minus a term that arises due to learning. The learning term is the present value of the infinite stream of the capital losses due to learning, using the effective discount rate $r + \lambda(\pi)(1 - F(p))$.

The capital loss from learning has two parts. $\pi(1 - \pi)V_D'(\pi)$ represents the impact of a unit of learning on future profits. It is unaffected by the current choice of p and hence we ignore its effect for the moment. The term

$$-\frac{\Delta\lambda(1 - F(p))}{r + \lambda(\pi)(1 - F(p))}$$

is affected by p in two ways. The first way measures the effect of p on the effective discount rate $r + \lambda(\pi)(1 - F(p))$:

$$-\frac{\Delta\lambda\lambda(\pi)f(p)(1 - F(p))}{(r + \lambda(\pi)(1 - F(p)))^2} \leq 0,$$

We call this effect the *controlled stopping effect.* It captures the increasing pessimism of the seller. Since it has a negative sign, this effect on its own leads to a lower posted price by the seller.

7

The second way is related to the change in the seller's posterior, $\pi(1-\pi)\Delta\lambda f(p)$, resulting from a change in p, normalised by the effective discount rate $r + \lambda(\pi)(1 - F(p))$:

$$\frac{\Delta\lambda f(p)}{r + \lambda(\pi)(1 - F(p))} \geq 0.$$

We call this effect the *controlled learning effect*. This term is positive, and so on its own leads to a higher posted price by the seller.

In this model, the sum of these two effects is

$$\frac{r\Delta\lambda f(p)}{(r + \lambda(\pi)(1 - F(p)))^2} \leq 0.$$

That is, the controlled learning effect dominates the controlled stopping effect.[3]

Hence the first-order condition for the dynamic problem is

$$p - \frac{1 - F(p)}{f(p)} = V_R(p, \pi) + \frac{\Delta\lambda f(p)}{r + \lambda(\pi)(1 - F(p))}\pi(1 - \pi)V_D'(\pi). \tag{5}$$

Equation (5) can be contrasted with equation (3). Since $V_D(\pi)$ is non-decreasing (from the same argument used in the proof of Proposition 2), the price posted by the seller who learns about the state of the world is *above* the price of the non-learning seller. As a result, the probability that the learning seller makes a sale is below that of the non-learning seller.

We summarise this discussion in the following Theorem.

Theorem 1 *In the case of learning about the arrival rate of buyers, the learning price $p_D(\pi)$ is greater than the repeated price $p_R(\pi)$ for all $\pi \in [0, 1]$.*

We can compare this result to Trefler (1993)'s. By Theorem 4 of Trefler, the seller moves her price in the direction that is more informative, in the Blackwell sense, relative to the static price p_S. In our model, the expected continuation value of information,

---

[3]This will not be the case in all models. In an earlier version of this paper, we considered also the case where the seller learns about the distribution from which buyers' valuations are drawn. We showed that either effect may then dominate.

(DeGroot (1962)) is

$$I(p, \pi) := \mathrm{E}[V_D(\pi + d\pi)] - V_D(\pi) = -\lambda(\pi)(1 - F(p))\left(V(\pi) + \frac{\Delta\lambda}{\lambda(\pi)}\pi(1 - \pi)V'(\pi)\right)dt,$$

ignoring terms in $(dt)^2$. Notice that this expression conditions on no sales in the period—an event which is bad news for the seller. So, with this definition, the expected continuation value of information is negative. Hence

$$\frac{\partial I(p, \pi)}{\partial p} = \lambda(\pi)f(p)\left(V(\pi) + \frac{\Delta\lambda}{\lambda(\pi)}\pi(1 - \pi)V'(\pi)\right)dt \geq 0.$$

This means that a higher price is more informative, in the Blackwell sense. Trefler's result therefore implies that $p_D(\pi) \geq p_S$. This is indeed what we find. We know from Proposition 1 that the repeated price is greater than the static price: $p_R(\pi) \geq p_S$; Theorem 1 tells us that the dynamic price is higher still. The comparison between the repeated and dynamic price does not appear in Trefler, since he does not consider the sale of a single asset.

The next two results give some intuitive properties of the seller's optimal decision.

Property 1 (Monotonicity) *The dynamic learning price, $p_D(\pi)$, and the probability of sale, $\lambda(\pi)(1 - F(p_D(\pi)))$, are non-decreasing in the seller's belief π.*

Proof. See the appendix. ∎

Property 2 (Discounting) *The dynamic learning price $p_D(\pi)$ is decreasing in the discount rate r.*

Proof. Equation (9) implies that $p_D(\pi)$ is decreasing in r iff $rV_D(\pi)$ is increasing in r. To show the latter, differentiate the Bellman equation (4), using the Envelope Theorem:

$$\frac{\partial}{\partial r}(rV_D(\pi)) = \max_p\left\{-\lambda(\pi)(1 - F(p))\frac{\partial V_D(\pi)}{\partial r} - \pi(1 - \pi)\Delta\lambda(1 - F(p))\frac{\partial^2 V_D(\pi)}{\partial\pi\partial r}\right\}.$$

Basic arguments establish that $\partial V_D(\pi)/\partial r \leq 0$ and $\partial^2 V_D(\pi)/\partial\pi\partial r \leq 0$. The result follows. ∎

An infinitely patient seller ($r = 0$) will set his price at the upper bound of the support of the valuation distribution. In contrast, a myopic seller ($r = \infty$) will set his price at the static level, which is clearly lower. The dynamic price moves between these two extremes monotonically in r.[4]

4.1 Two-point example

We can illustrate these results in a simple case where buyers' valuations can take two values, $\overline{v} > \underline{v} > 0$, with probabilities $\beta \in (0, 1)$ and $1 - \beta$ respectively. For the problem to be interesting, assume that $\beta\overline{v} < \underline{v}$, so that in the static problem, the seller would set a price equal to $\underline{v}$. Also assume that if the seller knows that $\lambda = \lambda_H$, her optimal price is $\overline{v}$; but if the seller knows that $\lambda = \lambda_L$, her optimal price is $\underline{v}$.

By standard arguments, the seller's price strategy then takes a simple form: set the price at $\overline{v}$ when the posterior is above some cut-off; and $\underline{v}$ for posteriors below this level. Let the cut-off level of the posterior be $\pi^* \in (0, 1)$. We illustrate the seller's value function in figure 1; in the appendix, we show how to determine the seller's value function and the cut-off π^*. The value function is a strictly convex function of π in the region $\pi \geq \pi^*$, while for $\pi < \pi^*$, it is a linear function.

For a seller who is certain of the arrival rate, the optimal price is $\overline{v}$ for all $\lambda \geq \lambda_0$, and $\underline{v}$ for $\lambda < \lambda_0$, for some λ_0. (See the appendix for a derivation of λ_0.) Then define π_0 to be such that $\pi_0\lambda_H + (1 - \pi_0)\lambda_L := \lambda_0$. We show in the appendix that $\pi^* \leq \pi_0$: that is, learning causes the seller to set a higher price, as Theorem 1 predicts.

5 Observed arrivals

Now suppose that the state of the world i.e., λ is not known, but that the seller observes when buyers arrive to the market. The seller's beliefs are updated according to Bayes'

---

[4]Perhaps surprisingly, the dynamic price is not monotonic in the arrival rate parameters λ_H and λ_L, a fact confirmed in numerical examples.
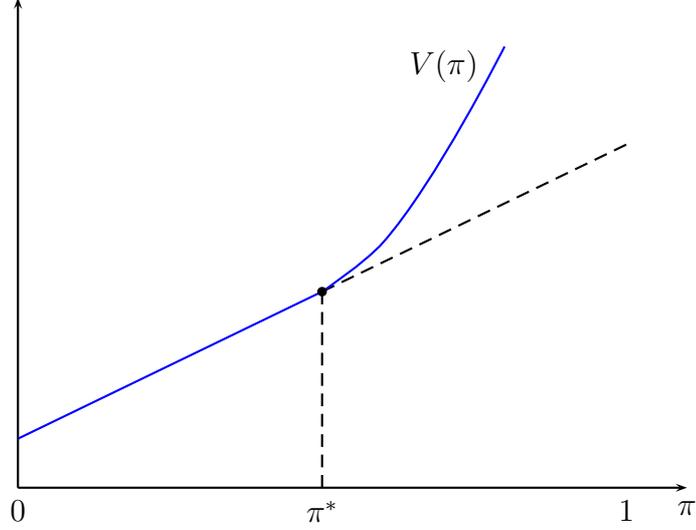
Figure 1: The value function with a 2-point valuation distribution

rule. Given an initial belief π, the posterior after an arrival is

$$\pi + d\pi_1 = \frac{\pi \lambda_H dt}{\pi \lambda_H dt + (1 - \pi)\lambda_L dt} = \frac{\lambda_H}{\lambda(\pi)}\pi. \tag{6}$$

Notice that $\pi + d\pi_1 > \pi$: the seller becomes more optimistic after an arrival. In fact,

$$d\pi_1 = \pi(1 - \pi)\frac{\Delta\lambda}{\lambda(\pi)}.$$

The posterior after no arrival is

$$\pi + d\pi_0 = \frac{\pi(1 - \lambda_H dt)}{\pi(1 - \lambda_H dt) + (1 - \pi)(1 - \lambda_L dt)} = \frac{1 - \lambda_H dt}{1 - \lambda(\pi)dt}\pi. \tag{7}$$

Since $\pi + \Delta\pi_0 < \pi$ in this case, the seller becomes more pessimistic in the event of no arrival:

$$d\pi_0 = -\pi(1 - \pi)\frac{\Delta\lambda dt}{1 - \lambda(\pi)dt}.$$

Note two things about the seller's beliefs in this case. First, since beliefs must be a martingale, the expected posterior equals the prior: $\lambda(\pi)dt(\pi + d\pi_1) + (1 - \lambda(\pi)dt)(\pi +$

11

$d\pi_0) = \pi$. Secondly, beliefs are unaffected by the price that the seller sets.

After ignoring higher order terms in dt, the seller's Bellman equation can be written as:

$$(1 + rdt)U(\pi) = \max_p \left\{ \lambda(\pi)dt \left(p(1 - F(p)) + F(p)U(\pi + d\pi_1) \right) \right\}$$
$$+ (1 - \lambda(\pi)dt)U(\pi + d\pi_0).$$

where $U(\cdot)$ is the value function when the seller observes the arrivals of buyers.

Notice that the term in the second line of this expression does not depend on the seller's price p. Hence the first-order condition (which is necessary and sufficient for an interior optimum) is

$$p - \frac{1 - F(p)}{f(p)} = U(\pi + d\pi_1). \tag{8}$$

Denote the solution to this first-order condition $p_O(\pi)$. In the next Proposition, we compare equations (5) and (8) to conclude that the seller's price is higher when it can observe the arrivals of buyers.

Proposition 3 $p_O(\pi) \geq p_D(\pi)$: *the seller's price is higher when it observes the arrival of buyers.*

Proof. First note that $U(\pi) \geq V_D(\pi)$ for all $\pi \in [0, 1]$ (with strict equality when $\pi \in \{0, 1\}$: the seller's value when it observes buyers' arrivals (i.e., has more information) is never less than when it cannot. Secondly, recall that $V_D(\pi)$ is a convex function. The first order condition for the dynamic problem with unobserved arrivals may be alternatively written as:

$$p_D - \frac{1 - F(p_D)}{f(p_D)} = V_D(\pi) + d\pi_1 V_D'(\pi),$$

where the expression for $d\pi_1$ has been substituted. Using the first two facts stated at the start of the proof, it follows that $V_D(\pi) + \Delta\pi_1 V_D'(\pi) \leq U(\pi + d\pi_1)$. This proves the
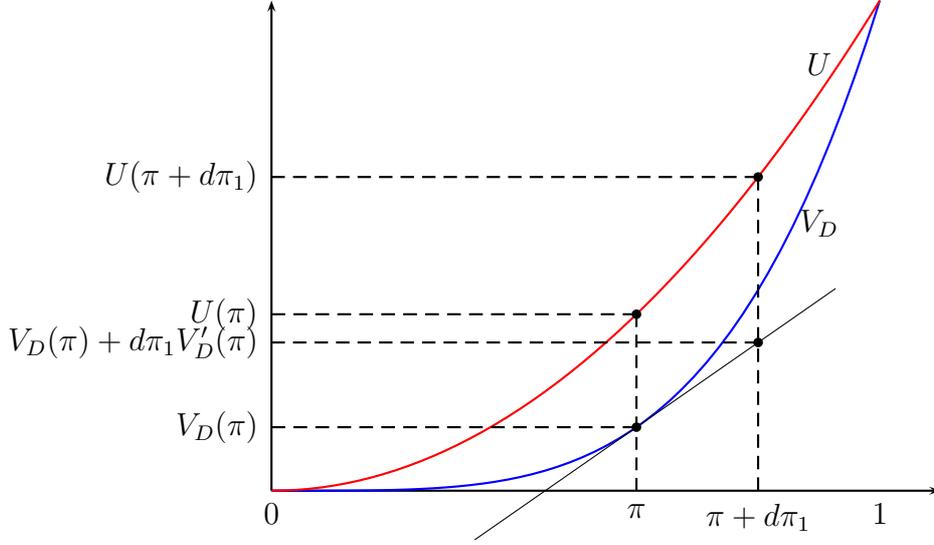
12

Figure 2: The proof of proposition 3

proposition.   ∎

The proof of the Proposition is illustrated in figure 2. The first-order condition (8) shows the intuitive fact that, when the seller observes arrivals, the price is based on the opportunity cost of a sale when an arrival occurs. (The price when no arrival occurs is irrelevant.) In contrast, when the seller does not observe arrivals, the price is based on an opportunity cost conditional on no sale occurring. This involves a capital loss from learning, $d\pi_1 V'_D(\pi)$: the marginal change in the seller's value arising from the change in the seller's beliefs. Bayes' rule tells us that the downward drift in beliefs following no sale, when arrivals are not observed, is proportional to the upward jump in beliefs when an arrival is observed. The factor relating the two is the probability of a sale $\lambda(\pi)dt(1-F(p))$ with unobserved arrivals.

This result establishes an upper bound for the price that the seller charges when arrivals are unobserved. The extent to which learning is controlled when arrivals are not observed is bounded above by the price increase resulting from full observability of arrivals.

6 Conclusion

We have written the model in terms of a seller learning about demand for her good. The model can easily be generalized to accommodate other standard economic models. For example, we can interpret the model as a job search problem when an unemployed worker sets his reservation wage while learning about employment opportunities. By introducing a flow cost for actions, it is possible to transform the model of learning about arrival rates into a model of R&D where the true success probability is initially unknown. Furthermore, there is no particularly compelling reason to concentrate only on Bayesian learning models. The change in posteriors could equally well reflect accumulation of past accumulated actions. Hence this model can be used as a basis for a non-Poisson model of R&D.[5]

The paper is written in terms of ever more pessimistic continuation beliefs. Obviously we could have taken an application where continuation beliefs drift upwards conditional on no stopping. As an example, consider the maintenance problem of a machine whose durability is uncertain. (Or the problem of finding the optimal effort level to demand from an agent or seller.) As long as the machine does not break down, beliefs about its longevity become more optimistic. Apart from the obvious change in signs, the controlled stopping and learning effects are present in such models as well.

Appendix

Lemma 1 $V_D(\pi)$ *is convex in* π.

Proof. Consider two points, π and π' and let $\pi^\alpha := \alpha\pi + (1-\alpha)\pi'$. We want to show that for all π, π' and π^α, we have $V_D(\pi^\alpha) \leq \alpha V_D(\pi) + (1-\alpha)V_D(\pi')$. Denote by $\mathbf{p}^\alpha$ the path of optimal prices starting from π^α. Denote by $W^H(\mathbf{p})$ the expected profit from an arbitrary path of prices $\mathbf{p}$ conditional on the true state being $\lambda = \lambda_H$ and similarly for

---

[5]A previous version of this paper includes such an application and is available from the authors upon request.

$W^L(\mathbf{p})$. Then we have $V_D(\pi^\alpha) = \pi^\alpha W^H(\mathbf{p}^\alpha) + (1 - \pi^\alpha)W^L(\mathbf{p}^\alpha)$. We also know that

$$V_D(\pi) \geq \pi W^H(\mathbf{p}^\alpha) + (1 - \pi)W^L(\mathbf{p}^\alpha),$$

$$V_D(\pi') \geq \pi' W^H(\mathbf{p}^\alpha) + (1 - \pi')W^L(\mathbf{p}^\alpha)$$

since $\mathbf{p}^\alpha$ is a feasible price path.

Hence

$$\alpha V_D(\pi) + (1 - \alpha)V_D(\pi') \geq \alpha(\pi W^H(\mathbf{p}^\alpha) + (1 - \pi)W^L(\mathbf{p}^\alpha))$$
$$+ (1 - \alpha)(\pi' W^H(\mathbf{p}^\alpha) + (1 - \pi')W^L(\mathbf{p}^\alpha))$$
$$= (\alpha\pi + (1 - \alpha)\pi')W^H(\mathbf{p}^\alpha) + (1 - \alpha\pi - (1 - \alpha)\pi')W^L(\mathbf{p}^\alpha)$$
$$= \pi^\alpha W^H(\mathbf{p}^\alpha) + (1 - \pi^\alpha)W^L(\mathbf{p}^\alpha) = V_D(\pi^\alpha).$$

This proves the claim.　　　　　■

The continuous-time Bellman equation

In section 4, we state the continuous-time Bellman equation (4). We now show how this Bellman equation arises as the limit of a discrete-time Bellman equation. In discrete time, the Bellman equation is

$$V_D(\pi) = \max_p \left\{ \lambda(\pi)dt(1 - F(p))p + (1 - rdt)\big(1 - \lambda(\pi)dt(1 - F(p))\big)V_D(\pi + d\pi) \right\}.$$

Noting that $V_D(\cdot)$ is differentiable, write

$$V_D(\pi + d\pi) = V_D(\pi) + d\pi V_D'(\pi) + (d\pi)^2 V_D''(\pi) + \ldots$$

Next note that $d\pi$ is of order dt. Since we take the limit $dt \to 0$, ignore terms in $(dt)^2$ or higher. The Bellman equation is then

$$V_D(\pi) = \max_p \left\{ \lambda(\pi)dt(1 - F(p))p + \left(1 - rdt - \lambda(\pi)dt(1 - F(p))\right)V_D(\pi) + d\pi V_D'(\pi)\right\}.$$

Substituting from equation (1) for $d\pi$ and re-arranging, we arrive at equation (4).

Proof of Property 1

To prove the first part, rewrite the first-order condition (5) as

$$\frac{(1 - F(p))^2}{f(p)} = \frac{rV_D(\pi)}{\lambda(\pi)} \tag{9}$$

Equation (9) implies that $p_D(\pi)$ is non-decreasing in π iff $rV_D(\pi)/\lambda(\pi)$ is non-decreasing in π. But

$$\frac{\partial}{\partial \pi}\left(\frac{rV_D(\pi)}{\lambda(\pi)}\right) = \frac{r}{\lambda(\pi)^2}\left(\pi\Delta\lambda\left(V_D'(\pi) - \frac{V_D(\pi)}{\pi}\right) + \lambda_L V'(\pi)\right);$$

by the facts that V_D is increasing and convex in π, this is non-negative and the result follows.

To prove the second part, consider the Bellman equation:

$$rV_D(\pi) = \max_p \left\{ \lambda(\pi)(1 - F(p))\left(p - V_D(\pi) - \pi(1 - \pi)\frac{\Delta\lambda}{\lambda(\pi)}V_D'(\pi)\right)\right\}$$

$$\equiv P(\pi)\Lambda(\pi)$$

where $P(\pi) \equiv \lambda(\pi)(1 - F(p_D(\pi)))$, $\quad \Lambda(\pi) \equiv p_D(\pi) - V_D(\pi) - \pi(1 - \pi)\frac{\Delta\lambda}{\lambda(\pi)}V_D'(\pi)$.

Then $P_D'(\pi) = (rV_D'(\pi) - P_D(\pi)\Lambda'(\pi))/\Lambda(\pi)$, assuming that $\Lambda \neq 0$. From the first-order condition, $\Lambda(\pi) = (1 - F(p_D(\pi)))/f(p_D(\pi))$. By assumption, $(1 - F(p))/f(p)$ is decreasing in p. From the first part of the Property, $p_D'(\pi) \geq 0$. This implies that $\Lambda'(\pi) \leq 0$. ■

The numerical example in section 4.1

We assume that if the seller knows that $\lambda = \lambda_H$, its optimal price is $\overline{v}$; but if the seller knows that $\lambda = \lambda_L$, its optimal price is $\underline{v}$. This is equivalent to

$$\frac{\beta\lambda_H\overline{v}}{r + \beta\lambda_H} \geq \frac{\lambda_H\underline{v}}{r + \lambda_H} > \frac{\lambda_L\underline{v}}{r + \lambda_L} \geq \frac{\beta\lambda_L\overline{v}}{r + \beta\lambda_L}. \tag{10}$$

For $\pi \geq \pi^*$, the seller's value function is the solution to the ODE

$$\pi(1 - \pi)\Delta\lambda V'(\pi) + \left(\frac{r}{\beta} + \lambda(\pi)\right)V(\pi) - \overline{v}\lambda(\pi) = 0.$$

The general solution to this ODE is of the form

$$V(\pi) = k_1\left(\frac{1 - \pi}{\pi}\right)^{\frac{r+\beta\lambda_H}{\beta\Delta\lambda}}\pi + \frac{\beta\lambda_L\overline{v}}{r + \beta\lambda_L} + \frac{r\beta\Delta\lambda\overline{v}}{(r + \beta\lambda_L)(r + \beta\lambda_H)}\pi$$

for $\pi \geq \pi^* > 0$. The first term represents the option value from learning to the seller. (k_1 is a constant of integration that will be determined below.) It is a convex function of π.

For $\pi < \pi^*$, the seller sets its price at $\underline{v}$; its value function is then given by the solution to a second ODE: $\pi(1 - \pi)\Delta\lambda V'(\pi) + (r + \lambda(\pi))V(\pi) - \underline{v}\lambda(\pi) = 0$. The general solution to this ODE is of the form

$$V(\pi) = k_2\left(\frac{1 - \pi}{\pi}\right)^{\frac{r+\lambda_H}{\Delta\lambda}}\pi + \frac{\lambda_L\underline{v}}{r + \lambda_L} + \frac{r\Delta\lambda\underline{v}}{(r + \lambda_L)(r + \lambda_H)}\pi.$$

The first term represents the option value from learning to the seller. It is unbounded as π approaches zero; hence the constant of integration k_2 must equal zero. In words, the seller has no option value when $\pi \leq \pi^*$. The reason is that there is no prospect of the seller's posterior increasing above π^*; hence the seller's price is constant (equal to $\underline{v}$) once the posterior falls below π^*. There is then no option for the seller to value. So, for $\pi < \pi^*$,

$$V(\pi) = \frac{\lambda_L\underline{v}}{r + \lambda_L} + \frac{r\Delta\lambda\underline{v}}{(r + \lambda_L)(r + \lambda_H)}\pi.$$

There are two remaining parameters to determine: k_1 and the optimal boundary π^*. Since π^* is chosen optimally, value matching and smooth pasting apply; these two boundary conditions are sufficient to determine the remaining parameters. In particular,

$$\pi^* = \frac{\left(-(\underline{v} - \beta\bar{b})r^2 + (\beta(\overline{v} - \underline{v})\lambda_L - (\underline{v} - \beta\overline{v})\lambda_H)r + \beta(\overline{v} - \underline{v})\lambda_L\lambda_H\right)\lambda_L}{\left(-(\underline{v} - \beta\overline{v})r^2 + \beta(\overline{v} - \underline{v})(\lambda_L + \lambda_H)r + \beta(\overline{v} - \underline{v})\lambda_L\lambda_H\right)\Delta\lambda}.$$

Let λ_0 be such that

$$\frac{\beta\lambda_0\overline{v}}{r + \beta\lambda_0} = \frac{\lambda_0\underline{v}}{r + \lambda_0}, \qquad \text{i.e.,} \quad \lambda_0 = \frac{r(\underline{v} - \beta\overline{v})}{\beta(\overline{v} - \underline{v})}.$$

For a seller who is certain of the arrival rate, the optimal price is $\overline{v}$ for all $\lambda \geq \lambda_0$, and $\underline{v}$ for $\lambda < \lambda_0$. Then define π_0 to be such that $\pi_0\lambda_H + (1 - \pi_0)\lambda_L = \lambda_0$. If $\pi^* \leq \pi_0$, then uncertainty causes the seller to set a higher price in this case with a two-point valuation distribution. Lengthy but straightforward calculation shows that this inequality holds if

$$\frac{\underline{v}}{r + \lambda_H} < \frac{\beta\overline{v}}{r + \beta\lambda_H}$$

which was assumed at the outset: see the inequalities in (10). In summary: learning about an unknown arrival rate causes the seller to set a higher price.

References

AGHION, P., P. BOLTON, C. HARRIS, AND B. JULLIEN (1991): "Optimal learning by experimentation," *Review of Economic Studies*, 58, 621–654.

ANDERSON, A., AND L. SMITH (2006): "Explosive Convexity of the Value Function in Learning Models," May 21.

BALVERS, R. J., AND T. F. COSIMANO (1990): "Actively learning about demand and the dynamics of price adjustment," *Economic Journal*, 100, 882–898.

BERGEMANN, D., AND U. HEGE (2005): "The financing of innovation: learning and stopping," *RAND Journal of Economics*, 36(4), 719–752.

BONATTI, A., AND J. HORNER (2010): "Collaborating," *American Economic Review*, Forthcoming.

BURDETT, K., AND T. VISHWANATH (1988): "Declining Reservation Wages and Learning," *Review of Economic Studies*, 55, 655–665.

DEGROOT, M. (1962): "Uncertainty, Information and Sequential Experiments," *Annals of Mathematical Statistics*, 33, 404–419.

DRIFFILL, J., AND M. MILLER (1993): "Learning and Inflation Convergence in the ERM," *The Economic Journal*, 103, 369–378.

EASLEY, D., AND N. KIEFER (1988): "Controlling a stochastic process with unknown parameters," *Econometrica*, 56, 1045–1064.

KARLIN, S. (1962): "Stochastic models and optimal policy for selling an asset," in *Studies in Applied Probability and Management Science*, ed. by K. J. Arrow, S. Karlin, and H. Scarf, pp. 148–158. Stanford University Press.

KELLER, G. (2007): "Passive Learning: a critique by example," *Economic Theory*, 33, 263–269.

KELLER, G., AND S. RADY (1999): "Optimal Experimentation in a Changing Environment," *Review of Economic Studies*, 66, 475–507.

KELLER, G., S. RADY, AND M. CRIPPS (2005): "Strategic Experimentation with Exponential Bandits," *Econometrica*, 73, 39–68.

LANCASTER, T., AND A. CHESHER (1983): "An Econometric Analysis of Reservation Wages," *Econometrica*, 51, 1661–1676.

MALUEG, D. A., AND S. O. TSUTSUI (1997): "Dynamic R&D Competition with Learning," *RAND Journal of Economics*, 28(4), 751–772.

McLennan, A. (1984): "Price dispersion and incomplete learning in the long run," *Journal of Economic Dynamics and Control*, 7, 33l–347.

Merlo, A., and F. Ortalo-Magne (2004): "Bargaining over Residential Real Estate: Evidence from England," *Journal of Urban Economics*, 56, 192–216.

Mirman, L., L. Samuelson, and A. Urbano (1993): "Monopoly Experimentation," *International Economic Review*, 34, 549–563.

Rogerson, R., R. Shimer, and R. Wright (2005): "Search-Theoretic Models of the Labor Market: A Survey," *Journal of Economic Literature*, XLIII, 959–988.

Rothschild, M. (1974): "A Two-armed Bandit Theory of Market Pricing," *Journal of Economic Theory*, 9, 185–202.

Rustichini, A., and A. Wolinsky (1995): "Learning about variable demand in the long run," *Journal of Economic Dynamics and Control*, 19, 1283–1292.

Smith, L. (1999): "Optimal job search in a changing world," *Mathematical Social Sciences*, 38, 1–9.

Trefler, D. (1993): "The Ignorant Monopolist: Optimal Learning with Endogenous Information," *International Economic Review*, 34(3), 565–581.

Van den Berg, G. (1990): "Nonstationarity in Job Search Theory," *Review of Economic Studies*, 57, 255–277.