

# Endogenous Learning from Incremental Actions

Job Market Paper

[Click here for the latest version](#)

Julia Salmi\*, Tuomas Laiho†, Pauli Murto‡

January 31, 2019

## Abstract

We study an experimentation problem where actions today have a long-lasting impact on information generation in the future. Actions are irreversible and generate information gradually over time. We solve for the optimal path of actions when the decision maker does not know the payoff-relevant state of the world. Because current choices have persistent effects, the problem has two state variables: a summary of past actions and the current belief on the state of the world. There is a novel informational trade-off as acting today speeds up information generation but postponing actions results in more informed choices. Our two leading examples cover the monopoly pricing of durable goods with social learning and capacity expansion in a market with uncertain optimal size. We show that since the monopolist can internalize future benefits from learning, the monopolist's optimal solution may result in a higher social surplus than the competitive market in both examples.

JEL classification: C61, D41, D42, D83, L1, O3

Keywords: Endogenous learning, experimentation, Brownian motion, dynamic monopoly, innovation adoption

## 1 Introduction

Many economic actions have irreversible effects but must be taken under uncertainty. A firm considering expanding its production capacity has to address uncertainty about

---

\*Aalto University School of Business (julia.salmi@aalto.fi, <https://sites.google.com/view/juliasalmi>).

†University of Oslo (tuomas.laiho@econ.uio.no).

‡Aalto University School of Business (pauli.murto@aalto.fi). We thank Nils Christian Framstad, Daniel Hauser, Matti Liski, Christian Traeger, Juuso Välimäki, and the seminar audiences at the Econometric Society European Winter Meeting and Helsinki GSE for their helpful comments. Salmi acknowledges financial support from Emil Aaltonen Foundation, OP Group Research Foundation, and Yrjö Jahansson Foundation.

future demand; a consumer buying a durable good today must evaluate her needs for it far in the future; emissions of pollution have irreversible impacts which may be hard to predict. When such uncertainties resolve exogenously over time, the decision maker should postpone her actions in order to learn more. But when the actions themselves affect the learning process, there is a trade-off between waiting for more information and acting now to speed up learning for later decisions.

This paper introduces a novel learning problem where actions today affect information generation in the future. Each action generates information gradually over time. Incentives to experiment under gradual learning differ from those with instantaneous feedback as each action affects the learning process far in the future. Many important economic applications such as durable good markets and environmental policy feature delays in learning as the consequences of past actions are revealed gradually over time.

We present a tractable model of experimentation where a decision maker optimizes the timing of her actions. The actions have immediate payoff consequences and future informational consequences. Our analytical innovation is to model the path of actions as a (weakly) increasing process. The process can be interpreted as the stock of installed capacity in an industry with unknown optimal size, the cumulated past sales of a durable good of unknown quality or the cumulated amount of environmental pollution with unknown long-run damages. In each of these applications, it is natural to imagine that the speed of learning depends on the current stock variable. As a result, we are led to consider a model with a two-dimensional state variable that includes the current level of action and the current belief on the underlying state of the world.

We use a continuous time Normal learning model as our informational set-up. An underlying binary state of the world together with the current action determines a signal process that we take to be a Brownian motion with an unknown drift. Since the process of current actions is assumed to be increasing, any decision to increase the action results in long-run effects on the signal process.

The optimal expansion rule can be fully characterized by a stationary policy, which gives the optimal quantity as a function of the current belief. The decision maker then expands until reaching the optimal quantity if the current quantity is below the optimal level. If the current quantity is above the optimal level, the decision maker does not expand but waits for a potential increase in the belief. The current quantity acts as a lower bound restricting the decision maker from reacting to bad signal realizations by decreasing the quantity. Learning from past actions shapes the optimal policy through two opposite channels: an expansion today speeds up information generation in the future, but postponing enables the current expansion decision to be made with more information. The former is *the information generation effect*, which increases the optimal quantity, and the latter is *the option value effect*, which decreases the optimal quantity.

The feature that learning from each action is gradual rather than instantaneous is important for the comparison between information generation and option value. If all information from one unit of quantity becomes instantaneously available, there is

no incentive to delay decisions. Gradual learning creates the dynamics in our model. The optimal expansion policy in the no-learning benchmark is to implement the static optimum instantly in the beginning. When we compare the optimal quantities under gradual learning and without learning, learning tends to increase the quantity for low beliefs and decrease it for high beliefs. This stems from the fact that the information generation effect dominates when the optimal quantity is low and the option value effect dominates for large quantities. Over time, learning tends to increase the optimal quantity in the beginning, when the need to speed up learning is largest, and to decrease it later as the option value effect starts to dominate. As the quantity increases, there are both fewer decisions left and more information available and hence less need for faster information generation. Furthermore, better learning technology strengthens both effects: experimentation in the beginning is more valuable and there is less need to experiment later as the speed of learning increases more rapidly with quantity.

We contrast the fully optimal solution with the case when the decision maker is *shortsighted*: she ignores how her action affects the learning process. The distinction between the full and the shortsighted solutions is important for the information generation versus option value comparison. The information generation effect is not present in the shortsighted solution and thus it incorporates only the option value effect. While quantity in the full solution can be either higher or lower than the static optimum, the shortsighted quantity lies always below both the static and fully optimal quantities.

The shortsighted solution of our model gives the equilibrium outcome when small individuals decide the quantity expansions. By expanding, each individual produces information valuable to others. However, they do not take into account that informational externality in their decision making. To see why the shortsighted solution is suboptimal note that the decision maker in our model is doing two things at the same time: producing and using a public good. Increasing the quantity today is equivalent to producing a public good as it allows later decisions to be made with better information. Exercising the option to wait – not increasing the quantity – is a way to use this public good as waiting allows to take the action with better information. In the full solution, these two sides of the problem weld together efficiently whereas the shortsighted solution will always exhibit too much free-riding as the value of information to later decision makers is not internalized.

We apply our general model to three intrinsically important setups: the pricing of a new durable good, the adoption of a new production technology and optimal environmental taxation. An additional contribution of the paper is to show how dynamic multiple agent problems can be modelled as a single-player quantity expansion problem by applying familiar mechanism design techniques in our stochastic and dynamic environment.

In the investment application, we show that endogenous learning breaks the link between decentralized and optimal investments. When uncertainty is exogenous, Leahy (1993) shows the equivalence between the competitive equilibrium, the social planner's solution, and the 'myopic' solution, which ignores the effect of individual firms on

prices. The last case corresponds to the shortsighted solution in our model. However, while the equivalence between the competitive and shortsighted solutions remains, the connection to the planner's solution is no longer valid and the shortsighted solution in our model is not efficient as the investment levels are too low due to the *endogenous* nature of learning. Therefore, small individual firms adopt new technologies too slowly and some innovations worth trying remain unadopted. This is because small investors focus on the option value of learning and do not internalize the information generation effect, which always distorts the market outcome from the first best.

The trade-off between the option value and the information generation effect is also present in durable good markets when there is social learning about initially unknown product quality. With endogenous learning about a durable good, a competitive market leads to similar inefficiencies as small investors do in the investment application. Competition forces prices to equal the marginal cost and the firms do not benefit from speeding up information production. As for consumers, waiting enables optimization with more information, and hence the competitive production level equals the shortsighted optimum. In contrast, a monopolist takes into account both the option value and the information generation effects.

We show that in a durable good market, the monopoly solution is more efficient than the competitive equilibrium if the initial belief is low. If the quality is likely to be high, standard monopoly distortions exceed the gains from the better internalization of the information effect. Furthermore, the monopolist tends to produce more in the beginning but less later on, when compared with the competitive market *at the same belief level*. This is important because while the true state will be eventually learned almost surely, how fast the belief convergences crucially depends on the rate of sales. A fast start triggers faster learning.

In the environment application, we solve for a policy maker's problem when she sets taxes in order to control emissions when damages from the emissions are uncertain. The emissions decisions are made by small individual producers whose intertemporal incentives are affected by the tax. Our model predicts that increased evidence of large damages should have an even larger effect on taxation as both the expected damages are higher and there is less uncertainty about them. Optimal taxation makes individuals to internalize two externalities: the informational one and the standard negative common pool externality. If there are no taxes, the latter feature implies individual producers emit too much pollution. This is because while individual producers internalize the benefits from emissions, they only internalize infinitesimally small part of the expected damages. Both the optimal solution and static solution are always better in terms of welfare than the no tax equilibrium.

## 1.1 Related literature

Using the framework of our paper, the previous literature on learning can be divided whether the option value effect or the information generation effect is present in the

model. To the best of our knowledge, the current paper is the first to analyze the dual effect of endogenous learning. The dual effect arises because the decision maker benefits from learning in the future and because learning from each action is gradual.

Models with exogenous uncertainty have only the option value effect. The decision of taking an irreversible action under exogenous uncertainty is identical to a problem of when to exercise a real option, which again bears close similarity to the problem of when to exercise an American option (see e.g. Dixit and Pindyck (1994)). Because the option can be exercised later, the threshold for exercising it is strictly above the static optimum. The main trade-off under exogenous uncertainty is between the cost of waiting and the option value of staying uncommitted. This trade-off is extensively analyzed in the investment literature.<sup>1</sup> Lucas and Prescott (1971) show that the competitive investment equilibrium is efficient under exogenous uncertainty. Leahy (1993) further shows that the competitive equilibrium behavior coincides with that of ‘myopic’ investors who ignore the effect the future investments have on the price. That result is behind our notion of the shortsighted solution.

The other extreme are the models of endogenous learning and reversible actions, which is the setup in the classic bandit problems (Gittins and Jones (1974), Rothschild (1974)) and in the early papers in experimental consumption (Prescott (1972) and Grossman et al. (1977)). The trade-off in these models is between experimentation and exploitation, i.e. between producing new information and using current information. The possibility of learning creates the information generation effect and hence increases the optimal quantity. The quantities chosen are always above the optimal static quantities. There is no option value effect since the decision maker is free to optimize her actions in each period and switch after bad news.

There are many papers studying setups where individuals learn from the consequences of each others’ actions. The literature on this kind of social learning finds that learning has an ambiguous effect on quantities. When actions are reversible, learning encourages taking the action. Whether experimentation is above or below the efficient level depends on the underlying strategic interaction. In the multi-agent Brownian bandit problem by Bolton and Harris (1999), the strategic dimension is limited to the learning externality between different experimenting agents. The equilibrium exhibits more experimentation than the static solution but less than the first best. Interestingly, the presence of other agents creates incentives to affect their future experimentation and may increase the level of experimentation above of what it would be in isolation (‘encouragement’ effect). In papers using a Poisson learning process, Keller, Rady and Cripps (2005) and Keller and Rady (2010) find roughly corresponding results but they depend on the details of the learning environment. Bergemann and Välimäki (2000) analyze the incumbent-entrant oligopoly behavior with uncertainty about the entrant’s quality and find that the equilibrium experimentation tends to be inefficiently high. This stems from oligopoly competition which favors large differences in quality

---

<sup>1</sup>See Dixit (1989), Pindyck (1991), Caballero and Pindyck (1996), Baldursson and Karatzas (1996), and Kogan (2001).

compared with the first best. Bergemann and Välimäki (2002) extend the analysis to markets with heterogeneous buyers, enabling vertical differentiation. In a similar environment, Bonatti (2011) studies monopoly menu pricing. There the information effect encourages the monopolist to serve a larger share of the market under uncertainty than when the quality is known to be high.

There are a few papers looking at social learning with irreversible actions. Frick and Ishii (2016) analyze informational free-riding in the adoption of new technologies. While they use the Poisson process to model learning, the model bears similarities with our shortsighted solution. The adoption rate is lower than without learning because individuals do not internalize the information effect of social learning but do take into account the option value it creates. An early paper by Rob (1991) makes a similar observation when analyzing sequential entry into a market of unknown size. Similarly, in the models of optimal timing under observational learning, the option value creates an incentive to wait and free-ride on information from others, causing socially inefficient delays (Chamley and Gale (1994), Murto and Välimäki (2011)).

The implications of social learning can be overturned once a large player is introduced. This is because there is an information generation effect at the social level. Che and Hörner (2017) study how a social planner, who designs a recommendation system for the consumers, can mitigate informational free-riding. Laiho and Salmi (2018) analyze monopoly pricing in a similar setup. Both in Che and Hörner (2017) and in Laiho and Salmi (2018), the presence of a social planner or a monopolist changes the situation compared to social learning with small agents so that learning increases quantities. This arises because a large player can internalize the informational externality. The crucial difference to the present paper is that these papers model instantaneous learning from each consumption decision. The planner and the monopolist do not face the option value effect since they get more information only by attracting new consumers. Therefore, learning decreases the initial price in Laiho and Salmi (2018) and the planner over-recommends new products in Che and Hörner (2017). In the present paper, gradual learning creates the option value effect also at the social level.

One aim of the current paper is to understand the driving forces behind the various effects of learning in the existing literature. Decomposing the effect of learning into the option value and information generation effects leads to a taxonomy of the existing literature. Papers where decision makers face only the option value, such as exogenous learning models and social learning models where individuals make irreversible one-time decisions, find that learning decreases the quantity. In models that have only the information generation effect, the opposite is true and quantities are larger than without learning. This is the case when actions are reversible or when a collective makes decisions under social learning from the flow of new actions. In our model, both effects are present and we show that information generation effect dominates in the beginning while the option value effect dominates towards the end.

The rest of the paper is organized as follows: Section 2 presents the general model, which is analyzed in Section 3; Sections 4 and 5 discuss the implications of the general

level results in several applications.

## 2 Model

For the general analysis, we model a simple quantity expansion problem with one decision maker. In Sections 4 and 5, we show that this general model is a reduced form of many important economic problems with multiple agents.

A decision maker chooses quantity  $q_t \in [0, \infty)$  over continuous time  $t$ , which goes from zero to infinity. Quantity expansions are irreversible so that  $q_t \geq q_{t'}$  for all  $t > t'$ . The decision maker discounts the future with rate  $r > 0$ .

There is uncertainty about the state of the world which is either high or low,  $\omega \in \{L, H\}$ . Let  $x_t$  denote the belief at time  $t$  that  $\omega = H$ . Given the processes for the quantity,  $Q_t$ , and the belief,  $X_t$ , the decision maker's payoff is

$$V(Q_t; x, q) = \mathbb{E}_x \left[ \int_0^\infty e^{-rt} (X_t \pi_H(Q_t) + (1 - X_t) \pi_L(Q_t)) dt \right], \quad (1)$$

where  $\pi_\omega$  is a differentiable concave function for state  $\omega$ . We denote its derivative by  $\gamma_\omega(q) := \pi'_\omega(q)$ . The derivatives satisfy  $\gamma_H(q) > \gamma_L(q)$  and  $\gamma'_\omega(q) \leq 0$  for all  $q$ , and that  $\gamma_H(0) > 0$  and  $\gamma_\omega(\infty) < 0$ . The assumptions guarantee the optimality of an interior solution and that the optimal quantity is higher in the high state than in the low state.

News about the state arrive via a Brownian diffusion process:

$$dY_t = q_t \mu_\omega dt + \sigma \sqrt{q_t} dW_t, \quad (2)$$

where  $\mu_H = 1/2$ ,  $\mu_L = -1/2$  and  $W_t$  is the standard Wiener process. This process is the continuous time limit of a discrete time process where each unit produces a conditionally independent normally distributed signal with mean  $y_{t-1} + \mu_\omega dt$  and variance  $\sigma^2 dt$ .

Let the current belief be  $x_t$ . Then, the unconditional belief evolves according to

$$dX_t = \frac{\sqrt{q_t}}{\sigma} x_t (1 - x_t) dW_t, \quad (3)$$

where the capital  $X_t$  denotes the stochastic process for the belief and the lower case  $x_t$  the current realization of this stochastic process. We refer the reader to Bolton and Harris (1999) for more details.

The signal to noise ratio of the process is the quantity,  $\sqrt{q_t}/\sigma$ . The quantity scales the variance of the belief process and determines how fast the belief converges to the true state. The law of motion implies that over time the belief has drift zero and variance equal to  $(q_t/\sigma^2)(x_t(1 - x_t))^2 dt$ . In other words, the speed of learning is *endogenous* in the model and controlled by the quantity  $q_t$ .

We are often interested in how the underlying informational environment affects optimal quantities, and we denote the inverse of the volatility by  $\lambda := \sigma^{-1}$ , which we refer to as *learning technology*. A higher  $\lambda$  implies better learning technology as there is less noise and thus signals are more informative.

### 3 Optimal quantity path

The decision maker chooses a quantity expansion rule  $Q_t$ , which is a random process measurable and adapted to the natural filtration generated by past news and quantities. The expansion rule maps the realized histories to present quantities.

As the learning process is stationary, the current belief and the current quantity contain all relevant information about the history, and we refer to them as the state. We can restrict to expansion rules which depend on the current state only and satisfy irreversibility:  $Q_t \geq Q_{t'}$  for all  $t > t'$ .

When using an optimal expansion rule, the decision maker's value (1) is

$$V(x, q) := \max_{Q_t} V(Q_t; x, q). \quad (4)$$

The problem is to find  $Q_t^*$  such that  $V(x, q) = V(Q_t^*; x, q)$ .

By applying Itô's lemma and using the properties of the Brownian motion, we get the following Hamilton-Jacobi-Bellman (HJB) equation for the decision maker's problem:

$$rV(x, q) = \max_{q^* \geq q} \left( x\pi_H(q^*) + (1-x)\pi_L(q^*) + \frac{1}{2}V_{xx}(x, q^*) \frac{x^2(1-x)^2}{\sigma^2} q^* \right). \quad (5)$$

We show the existence of a solution to the HJB equation by formulating an equivalent free boundary problem and showing that it has a unique solution. Then a standard verification argument shows that the solution maximizes (1).

We can recast this problem as finding a boundary in the  $(q, x)$  state space below which the decision maker increases the quantity and above which she waits for better news. This follows simply from the optimality of the quantity  $q^*$ : for each belief there is a quantity such that given the continuation value (or the derivative  $V_{xx}$ ) the decision maker is indifferent between increasing the quantity and waiting. One can think that the decision maker owns a stock of options (which yield both profit and information in the future) and the boundary determines how many options she would like to exercise at each point of the state space.

The boundary is fully characterized by a policy function  $x : \mathbb{R}_+ \rightarrow [0, 1]$ , which gives the cutoff belief for each quantity. Let  $x^*(q)$  denote the optimal policy. If the current quantity is  $q$ , the decision maker expands for beliefs above  $x^*(q)$ . The corresponding boundary,  $(q, x^*(q))$ , then splits the state space into two regions: a waiting region in which irreversibility constraint binds and the quantity stays constant, and an expansion region in which the decision maker increases the quantity so that the state reaches the boundary. Figure 1 illustrates the optimal policy in the quantity-belief state space. Above the boundary  $(q, x^*(q))$ , the decision maker increases the quantity (horizontal movement in the figure) and below it, she waits and only the belief moves (vertical movement) until the boundary is reached and the decision maker starts expanding along it. As soon as the belief hits the boundary, the quantity is 'reflected' upwards



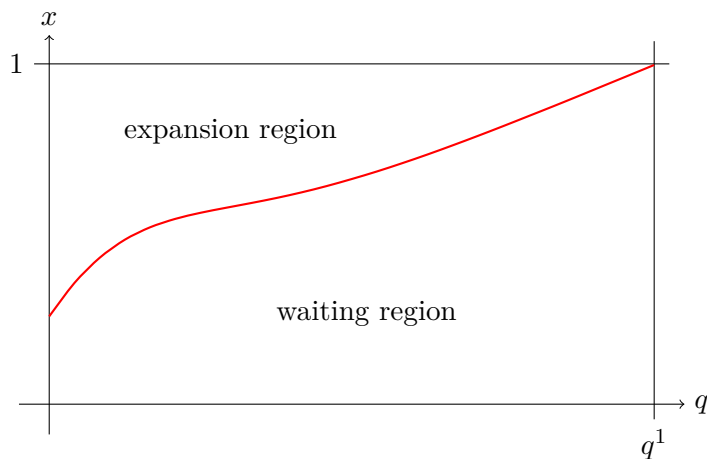


Figure 1: Waiting and expansion regions in the state space.

i.e. at the boundary the decision maker reacts to good news by increasing the quantity immediately.

To find the boundary, we can apply the usual value matching and smooth pasting conditions familiar in the literature. In our model, these conditions are set for the derivative of the value function. For the existence of an interior solution for  $q^*$  in (5), the marginal effect of increasing  $q$  must be zero at the optimal level. Therefore we get  $V_q(q, x) = 0$  as our value matching condition. Similarly, the smooth pasting condition is now set for the marginal effect so that  $V_{qx}(q, x) = 0$ . Smooth pasting means that the value of increasing the quantity matches smoothly to the value of waiting at the boundary.<sup>2</sup> We can write them as follows:

$$\begin{aligned} V_q(x, q) &= 0, \\ V_{qx}(x, q) &= 0. \end{aligned}$$

To apply these, we first note that the HJB-equation (5) can be solved explicitly in the continuation region, and the solution takes the form:<sup>3</sup>

$$V(x, q) = \frac{x\pi_H(q) + (1-x)\pi_L(q)}{r} + B(q)\Phi(x, q),$$

---

<sup>2</sup>Smooth pasting follows from standard arguments that if two functions meet at an upward kink, randomization between two points in different sides of the kink produces higher outcome than the kink point itself. In our model, waiting enables this kind of randomization between different levels of the belief, and hence contradicts indifference. See e.g. Dixit and Pindyck (1994) for a demonstration that the gain from randomization necessarily outperforms the cost of waiting for Brownian motion.

<sup>3</sup>We have discarded the other root of the proposed solution as we must have that the value converges to the static solution as  $x \rightarrow 0$  and  $x \rightarrow 1$ .

where

$$\beta(q) := \frac{1}{2} \left( 1 + \sqrt{1 + \frac{8r\sigma^2}{q}} \right),$$

$$\Phi(x, q) := x^{\beta(q)}(1-x)^{1-\beta(q)}.$$

We can see here that the value function consists of a discounted stream of the current flow (the first term) and the value of potential future actions (the second term). Given this form for the value function, we can write the value matching and smooth pasting conditions as:

$$\frac{x^*(q)\gamma_H(q) + (1-x^*(q))\gamma_L(q)}{r} + B_q(q)\Phi(x^*(q), q) + B(q)\Phi_q(x^*(q), q) = 0, \quad (6)$$

$$\frac{\gamma_H(q) - \gamma_L(q)}{r} + B_q(q)\Phi_x(x^*(q), q) + B(q)\Phi_{qx}(x^*(q), q) = 0. \quad (7)$$

These two equations are going to determine the optimal boundary  $x^*(q)$  and the value of the unknown function  $B$  for all  $q$ . The optimal policy must balance the direct effect shown in the first term in both equations against both the option value of waiting and the value of information generation. The last two show up in the derivatives of the value function, i.e. in the latter terms of each equation.

The difficulty in finding the optimal boundary is that it must be solved together with the function  $B$ . Since (6) and (7) tie together  $B(q)$  and the derivative  $B_q(q)$ , it is inevitable that the solution must involve solving a differential equation. The one piece still missing is the appropriate initial condition for such a differential equation. This we get by considering the optimal quantity at the end of the learning process when the state of the world is known to be high. When  $x = 1$ , the problem of finding the optimal quantity reduces to a static optimization problem. Solving it yields:

$$q^1 = \inf\{q \in \mathbb{R}_+ : \gamma_H(q) = 0\}.$$

The appropriate initial value for the optimal boundary is therefore  $x^*(q^1) = 1$ .

We show in the Appendix (A) that the system (6) - (7) can be transformed into a non-linear differential equation for the optimal boundary  $x^*(q)$ . Appendix (A) also verifies that the solution to this differential equation maximizes the original objective (1). The proposition below summarizes the optimal policy:

**Proposition 1.** *The optimal policy  $x^*(q)$  is characterized by the differential equation  $(x^*)'(q) = g(x^*(q), q)$  and the initial condition  $x^*(q^1) = 1$ , where*

$$g(x, q) = x(1-x) \left[ x(\beta'(q)(\beta(q) - 1)\gamma'_H(q) - ((\beta(q) - 1)\beta''(q) - 2(\beta'(q))^2)\gamma_H(q)) \right. \\ \left. + (1-x)(\beta'(q)\beta(q)\gamma'_L(q) - (\beta(q)\beta''(q) - 2(\beta'(q))^2)\gamma_L(q)) \right] / \\ \left[ (x(\beta(q) - 1)^2\gamma_H(q) + (1-x)(\beta(q))^2\gamma_L(q))\beta'(q) \right]. \quad (8)$$

We cannot solve the boundary  $x^*(q)$  in closed form because it is governed by a complicated nonlinear ordinary differential equation. The next lemma establishes an important property of the boundary.

**Lemma 1.** *The optimal policy  $x^*(q)$  is increasing in  $q$ .*

Lemma 1 verifies the intuition that the decision maker chooses a larger quantity when the belief is high. Importantly,  $x^*(q)$  defines an increasing quantity path  $Q_t$  in terms of a reflecting barrier in the  $(q, x)$  state space: every time  $X_t$  hits  $x^*(q)$ , output is increased so as to prevent  $X_t$  from crossing  $x^*(q)$ . If decision maker receives bad news (i.e. low signal) after increasing the quantity she will wait for a positive amount of time before increasing quantity again. Thus with a positive probability the decision maker waits before increasing the quantity again. A realized quantity path then moves along the boundary when there is a large enough increase in the belief and departs from when the belief decreases.

### 3.1 Shortsighted solution

Next, we analyze a slightly different solution to the problem, which we call the *short-sighted solution*. It captures the idea of a decision maker who takes correctly into account the learning generated by the existing quantity but fails to account for how her actions provide future benefit by changing the learning process. The shortsighted solution serves two roles in the analysis: first, it gives a tractable lower bound for the optimal quantity path, providing a useful tool in proofs; second, the shortsighted solution captures a decentralized equilibrium, where small agents fail to internalize the informational externality that they provide to others.

Instead of considering a single decision maker who chooses the quantity over time, we define the problem here as a sequence of infinitesimally small decision makers indexed by  $q \in (0, q^1)$ , where each decision maker chooses the optimal time to implement her expected individual contribution,  $dq$ , to the total welfare. Suppose that the current output is fixed at  $q$  and the belief process follows (3). The decision maker indexed by  $q$  decides when to increase quantity from  $q$  to  $q + dq$  in return of an perpetual surplus flow of  $\gamma_\omega(q) dq$ . Since state  $\omega$  is unknown, the expected present value of this increment at the current belief  $x_t$  is

$$dq \int_t^\infty e^{-r(s-t)} [X_t \gamma_H(q) + (1 - X_t) \gamma_L(q)] ds = \frac{dq}{r} [X_t \gamma_H(q) + (1 - X_t) \gamma_L(q)].$$

The problem is hence to find the optimal stopping time  $\tau$ , adapted to the natural filtration of the past news and quantities, for each  $q$  to solve

$$\sup_\tau \mathbb{E} \left( e^{-r\tau} \frac{1}{r} [X_t \gamma_H(q) + (1 - X_t) \gamma_L(q)] \right) \quad (9)$$

subject to (3).

Before solving (9), let us discuss its connection to the original control problem (4). The solution to (9) finds the optimal time to implement the marginal contribution of

each potential increment in output to the total surplus. In effect, the solution is solving a sequence of optimal stopping problems when to exercise an option. By solving such an optimization problem for every  $q$  and tying those solutions together appears to solve the optimization problem presented (4) in a decentralized manner. This would indeed be the case if the learning process for  $X_t$  was exogenous (i.e., if the evolution of  $X_t$  was independent on  $q$ ).<sup>4</sup> However, it is important to note that in our model an increment  $dq$  has an additional social benefit on top of the direct contribution  $\gamma_\omega(q) dq$ : increasing  $q$  to  $q + dq$  also changes the signal to noise ratio in the learning process from  $\frac{\sqrt{q}}{\sigma}$  to  $\frac{\sqrt{q+dq}}{\sigma}$ . The failure of the problem (9) to account for this additional benefit breaks the connection between the original control problem and the shortsighted solution.

We now proceed to solve (9) for each  $q$ . By standard arguments, the solution must be a threshold  $x^S(q)$  such that it is optimal to stop as soon as the belief process  $X_t$  hits  $x^S(q)$  from below. Denote by  $V^S(x; q)$  the value of the problem for  $x_t = x$ . By standard dynamic programming steps, this value function must satisfy the following differential equation for all  $x < x^S(q)$ :

$$\frac{1}{2}V_{xx}^S(x; q)\frac{x^2(1-x)^2}{\sigma^2}q - rV^S(x; q) = 0. \quad (10)$$

We also know that for  $x = 0$ , the value must be zero and so we have a boundary condition  $V^S(0; q) = 0$ . With this condition, the general solution to (10) is

$$V^S(x; q) = B^S(q)\Phi(x, q),$$

where,

$$\begin{aligned} \Phi(x, q) &= x^{\beta(q)}(1-x)^{1-\beta(q)}, \\ \beta(q) &: = \frac{1}{2} \left( 1 + \sqrt{1 + \frac{8r\sigma^2}{q}} \right) > 1, \end{aligned}$$

and where  $B^S(q)$  is an arbitrary function. We can solve it together with the optimal stopping threshold  $x^S(q)$  from the value matching and smooth pasting conditions:

$$\begin{aligned} V^S(x^S(q); q) &= \frac{x^S(q)\gamma_H(q) + (1-x^S(q))\gamma_L(q)}{r}, \\ V_x^S(x^S(q); q) &= \frac{\gamma_H(q) - \gamma_L(q)}{r}. \end{aligned}$$

The value matching and smooth pasting conditions are different from the optimal solution. This follows from the fact that the decision maker(s) is now choosing a stopping time for each individual increment of  $q$ . Thus, the value matching condition is equivalent to the decision maker being indifferent between waiting and exercising her option. The smooth pasting condition is equivalent to saying the decision maker cannot gain by randomizing her decision: the value of waiting smoothly to the value of exercising the option.<sup>5</sup>

---

<sup>4</sup>The corresponding connection between socially optimal capacity expansion policy and individually optimal investment timing policies for small firms has been discussed in, e.g., Lucas and Prescott (1971), Leahy (1993), Baldursson and Karatzas (1996), and Dixit and Pindyck (1994).

<sup>5</sup>As previously, these conditions follow from the standard arguments in the literature.

By substituting in  $V^S$  and  $V_x^S$  and solving the pair of equations above, we get a closed-form solution for the stopping threshold  $x^S(q)$ ; see Proposition 2 below. The shortsighted dynamics are then determined by a similar reflecting boundary as in the optimal solution: the quantity is held constant below the boundary and the process never crosses it as the quantity increases to balance a possible increase in the belief, forcing the process to climb up the boundary instead. However, the form of the boundary differs from the optimal boundary since the trade-offs are different in the two solutions.

**Proposition 2.** *The shortsighted policy is given by threshold*

$$x^S(q) = \frac{\beta(q)\gamma_L(q)}{\beta(q)\gamma_L(q) + (1 - \beta(q))\gamma_H(q)}.$$

*The threshold  $x^S(q)$  is continuous, positive, and strictly increasing for  $q \in (0, q^1]$ , and  $x^S(q^1) = 1$ .*

### 3.2 Comparison of different solutions

In this subsection, we compare the optimal solution with the static and shortsighted solutions. Without learning the decision maker implements the static optimum, which solves the first order condition:

$$x\gamma_H(q) + (1 - x)\gamma_L(q) = 0.$$

Let the static policy be denoted by  $x^{stat} = -\gamma_L(q)/(\gamma_H(q) - \gamma_L(q))$ .

We can use the HJB-equation (5) to get an intuitive understanding of the trade-offs the decision maker is facing. We evaluate the value matching condition (6) with the help of the HJB-equation and get

$$x\gamma_H(q) + (1 - x)\gamma_L(q) + \frac{1}{2}V_{xx}(x, q)\frac{x^2(1 - x)^2}{\sigma^2} + \frac{1}{2}V_{xxq}(x, q)\frac{x^2(1 - x)^2}{\sigma^2}q = 0,$$

where  $x = x^*(q)$  at the boundary. The first part is the direct effect on the flow. In an environment without learning, the decision maker implements the static optimum:  $x\gamma_H(q) + (1 - x)\gamma_L(q) = 0$ . Learning adds the information generation and the option value effects, which are reflected by the two latter terms in the equation.

The term proportional to  $V_{xx}(x, q)$  is always positive (see Appendix B). It captures the gain of the increased volatility in belief, that is the information generation effect. The last term, proportional to  $V_{xxq}(x, q)q$ , is always negative (see Appendix B). This is driven by the option value. To understand why the option value is captured by the third cross derivative of the value function,  $V_{xxq}(x, q)$ , notice that learning is more valuable if the current quantity is low and there are still many decisions to be made. An increase in  $q$  means that the stock of uncommitted actions,  $q^1 - q$ , is decreased, and hence an option to make the use of the increased volatility is lost.

How optimal quantities compare with the static optimum is determined by the relative effect of the last two terms. The quantity is above (below) the static level when the following has a positive (negative) sign:

$$V_{xx}(x, q) + V_{xxq}(x, q)q.$$

When the expression is negative, i.e. when the option value dominates,  $x\gamma_H(q) + (1-x)\gamma_L(q)$  must be positive and hence the quantity must be below the static optimum. The opposite is true when information generation dominates.

Figure 2 illustrates the optimal policy and how it differs from the static and short-sighted solutions. The option value dominates when the static solution is larger than the dynamic one; the information effect dominates in the opposite case. The option value term  $V_{xxq}(x, q)q$  is zero when  $q = 0$ . Therefore, the effect of learning is positive in the beginning. The option value dominates close to  $x = 1$  as the information generation effect disappears faster than the option value effect when the belief approaches 1. We state and prove these results formally in the end of this subsection.

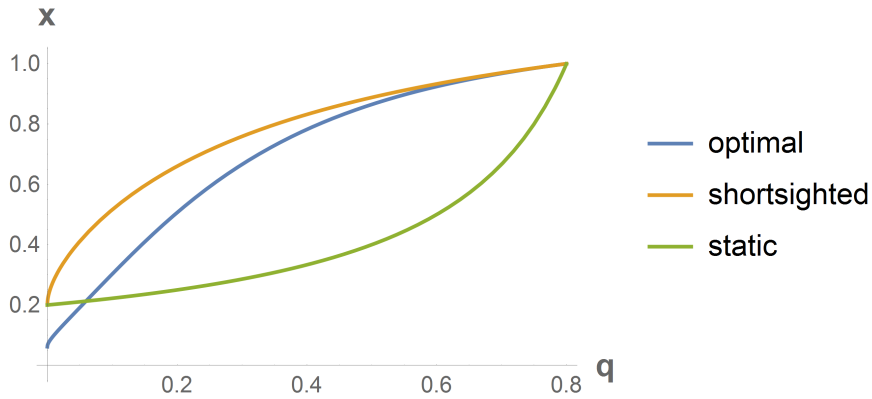


Figure 2: Numerical example of the boundaries.

A shortsighted decision maker takes into account only the option value while the optimal solution has both and the static solution has neither of the two effects. In the shortsighted solution, quantities are everywhere lower than in the static solution and in the optimal dynamic solution (see Appendix B for the proof):

**Lemma 2.** *The shortsighted boundary is everywhere above the optimal and static boundaries:  $x^S(q) > x^{stat}(q)$  for all  $q \in (0, q^1)$  and  $x^S(q) > x^*(q)$  for all  $q \in [0, q^1)$ .*

The difference between the optimal and shortsighted solution smoothly disappears close to  $x = 1$  (see Appendix B):

**Lemma 3.** *The full solution and the shortsighted solution coincide and have the same derivative at  $(q^1, 1)$ :  $x^*(q^1) = x^S(q^1) = 1$  and  $(x^*)'(q^1) = (x^S)'(q^1)$ .*

The intuition behind Lemma 3 is the same as for why the static solution gives larger quantities for high beliefs: information generation effect goes to zero very fast

when  $q$  approaches  $q^1$ . When  $q$  is large and hence the potential increase small, there is little value of speeding up learning.

To show that the optimal solution is above the static optimum near  $q^1$  we next argue that the optimal solution reaches  $x = 1$  above the static solution. We do this by showing that the derivative of the static solution is strictly larger than that of the optimal solution. By Lemma 3, it is enough to compare the derivatives of the shortsighted and static solutions at  $q^1$ . Because the derivatives of the shortsighted and optimal solutions coincide, the result can be directly applied also to the comparison between the optimal and static solutions. The difference between the derivatives of the shortsighted and static policies evaluated at  $q^1$  is

$$\begin{aligned} x^{S'}(q^1) - x^{stat'}(q^1) &= \frac{(\beta - 1)\beta\gamma_L\gamma'_H}{(\beta\gamma_L)^2} - \frac{\gamma_L\gamma'_H}{(\gamma_L)^2} \\ &= -\frac{\gamma_L\gamma'_H}{\beta\gamma_L^2} < 0. \end{aligned}$$

We can conclude that the static optimum has larger derivative at  $q^1$  than the optimal solution, which further implies that it must approach 1 below the shortsighted boundary. We can now combine the discussion of this subsection into the following proposition (the detailed proof is in Appendix B):

**Proposition 3.** *There exist  $\underline{x}$  and  $\bar{x}$  such that  $0 < \underline{x} \leq \bar{x} < 1$  and that the optimal quantity is larger than the static optimum for all beliefs below  $\underline{x}$  and smaller for all beliefs above  $\bar{x}$ .*

## 4 Durable goods monopoly

We apply the results from the general model to solve the problem of a durable good monopolist when there is common uncertainty about the quality of the product the monopolist is selling. Both the monopolist and the buyers learn from past purchases so the problem involves *social learning*. Each buyer wants to purchase only one unit, which makes the selling decision irreversible. The monopolist, therefore, faces a trade-off between postponing sales until she can charge higher prices and selling now in order to speed up learning, which may lead to higher prices in the future.

A buyer's utility from consumption depends on his private type,  $\theta \in [\underline{\theta}, \bar{\theta}]$ , and the common quality,  $\omega \in \{L, H\}$ :

$$\mathbb{E}(\mathbf{1}_{\omega=H} \cdot \theta) = x_t\theta,$$

where  $x_t$  is the current belief that the quality is high. There is a mass of buyers (normalized to one), whose types follow a twice continuously differentiable density function  $f$ , which satisfies the monotone hazard rate condition:  $\frac{1-F(\theta)}{f(\theta)}$  is decreasing in  $\theta$ . Each buyer wants to buy one unit of the good and exits after purchasing.

All players observe the amount of sales and the public news process. Let the news follow the Brownian diffusion process specified in (2), and hence the implied process for the belief is governed by (3).

The monopolist sets a pricing strategy, which is allowed to depend on the realized news and past sales, and the buyers choose whether to purchase or wait. Buyers who purchase, exit the market and other buyers stay in the market. We solve the optimal pricing scheme with full commitment for the monopolist.<sup>6</sup>

#### 4.1 Envelope theorem representation

In this subsection we show that the buyers' incentive compatibility implies a virtual surplus representation for the monopolist's payoff function, a result familiar from static mechanism design. Our treatment is based on our earlier paper Laiho and Salmi (2018) and shares similar features with Board (2007) who uses the Milgrom-Segal (2002) envelope theorem to analyze optimal sales of options.

Let the monopolist follow an arbitrary pricing policy,  $P_\tau$ , adapted to the natural filtration  $\mathcal{H}_t$  of the public history  $h_t$ , which consists of the path of past sales, prices set, and news arrived. We denote the realization of  $P_\tau$  at time  $t$  with  $p_t$ . Buyers are infinitesimally small and hence an individual buyer has no effect on the aggregates, including the learning process. Therefore, an individual buyer faces an optimal stopping problem under exogenous learning. Let  $\mathcal{T}$  denote the set of  $\mathcal{H}_t$ -adapted stopping rules. We can write the buyer's payoff as a function of the stopping rule,  $\tau \in \mathcal{T}$ :

$$U(\tau; \theta) = \mathbb{E} [e^{-r\tau} (\theta X_\tau - P_\tau)]$$

subject to (3).

The buyer then uses a stopping rule to maximize  $U$ :

$$W(\theta) := \sup_{\tau} U(\tau; \theta).$$

Let  $\tau(\theta)$  be such that  $\sup_{\tau} U(\tau; \theta) = U(\tau(\theta); \theta)$ . A collection of such stopping rules  $(\tau(\theta))_{\theta=\theta}$  is an incentive compatible allocation.

First we show that incentive compatible allocations are monotone:

**Lemma 4.** *Given a public history  $h_t$ , let  $\tau(\theta) = t$ . Then,  $\tau(\theta') \leq t$  for all  $\theta' > \theta$ .*

*Proof.* The incentive compatibility of the stopping rule  $\tau(\theta)$  implies

$$\theta x_t - p_t \geq \mathbb{E} [e^{-r(\tau-t)} (\theta X_\tau - P_\tau) | h_t],$$

for all stopping rules  $\tau$ . Or equivalently,

$$\theta \left( x_t - \mathbb{E} [e^{-r(\tau-t)} X_\tau | h_t] \right) \geq p_t - \mathbb{E} [e^{-r(\tau-t)} P_\tau | h_t].$$

---

<sup>6</sup>Possible sources of commitment power include money-back guarantees, leasing, capacity constraints, and the need to affect buyers' expectations in nonstationary equilibria.



The right-hand side is independent of  $\theta$ . The left-hand side is increasing in  $\theta$  if expression  $\left(x_t - \mathbb{E}\left[e^{-r(\tau-t)}x_\tau|h_t\right]\right)$  is positive, which is always true since  $x_\tau$  is a posterior belief and hence a martingale, and discounting then makes the expression strictly positive for all  $\tau$  which do not stop at  $t$ . Therefore, incentive compatibility holds as a strict inequality for all  $\theta' > \theta$ .  $\square$

Under strictly increasing differences, global incentive compatibility is equivalent with local incentive compatibility together with monotonicity. The local incentive compatibility takes the envelope theorem form:<sup>7</sup>

$$W(\theta) = \int_{\underline{\theta}}^{\theta} W_{\theta}(s)ds = \int_{\underline{\theta}}^{\theta} U_{\theta}(\tau(s); s)ds = \int_{\underline{\theta}}^{\theta} \mathbb{E}\left[e^{-r\tau(s)}X_{\tau(s)}\right] ds,$$

where we have already plugged in  $W(\underline{\theta}) = 0$ .

We write the monopolist's problem in terms of choosing an incentive compatible stopping rule for each buyer. In the spirit of Myerson (1981), we use the envelope theorem form of the buyers' incentive compatibility to get a virtual surplus representation for the monopolist's payoff:

**Lemma 5.** *Given an incentive compatible stopping rule  $(\tau(\theta))$ , the monopolist's payoff is*

$$\mathbb{E}\left[\int_{\underline{\theta}}^{\bar{\theta}} e^{-r\tau(\theta)}\left(X_{\tau(\theta)}\theta - X_{\tau(\theta)}\frac{1-F(\theta)}{f(\theta)} - c\right)f(\theta)d\theta\right].$$

Proof: repeated use of Fubini's theorem (see Appendix C).

By monotonicity, there is a one-to-one mapping between the quantity sold and the highest type still in the market,  $\theta(q) : 1 - F(\theta(q)) = q$ . The payoff can thus be equivalently written over quantities:

$$\mathbb{E}\left[\int_0^{\bar{q}} e^{-r\tau(q)}\left(X_{\tau(q)}\phi(q) - c\right)dq\right], \tag{11}$$

where  $\tau(q)$  is the stopping rule of the  $q$  highest type buyer and  $\phi(q)$  is his virtual valuation:

$$\phi(q) = \theta(q) - \frac{1 - F(\theta(q))}{f(\theta(q))}.$$

The monotone hazard rate condition guarantees that the virtual valuation is decreasing in  $q$ .

## 4.2 Optimal dynamic pricing

The monopolist's problem is to allocate incentive compatible stopping rules  $\tau(\theta)$  to each type  $\theta$  to maximize her payoff. The integral cannot be maximized pointwise with

<sup>7</sup>See Milgrom and Segal (2002).

endogenous learning because the allocation affects the news process. Therefore, even though we can apply static mechanism design techniques to reduce the monopolist's problem into maximizing virtual surplus subject to monotonicity, the optimization has to be dynamic. Next, we argue that the monopolist's problem is a special case of the model in Section 2.

In the case of durable goods, the monopolist receives a one-time payoff of sales, but we can equivalently write the payoff as a flow over time. Normalized to be a flow payoff, the monopolist's payoff (11) equals

$$r\mathbb{E}\left[\int_0^\infty e^{-rt}\left(\int_0^{Q_t}(X_t\phi(q)-c)dq\right)dt\right].$$

To see that this is a special case of our general model, set

$$\begin{aligned}\gamma_H(q) &= r(\phi(q) - c), \\ \gamma_L(q) &= -rc.\end{aligned}$$

We get the following as a direct corollary of Proposition 1 and Lemma 1:

**Corollary 1.** *The monopolist's policy is to sell for beliefs above  $x^M(q)$  and to wait for beliefs below. The boundary  $(x^M(q), q)$  is monotone and is characterized by*

$$\begin{aligned}x^M(q^1) &= 1, \\ x^{M'}(q) &= f(q, x^M(q)),\end{aligned}$$

where  $f$  is the nonlinear function of  $x$  and  $q$  specified in Proposition 1.

The monopolist chooses quantities following the general logic described in the previous section. She is first willing to sell even if the virtual surplus is negative. In the beginning, this means that she is willing to price below the marginal cost. The monopolist is willing to subsidize consumption in order to generate learning. Later on, the monopolist reduces sales as the option value effect starts to dominate. Then the monopolist is less willing to sell to low type buyers because of the information rents they impose on higher type buyers and because there is less need to further speed up learning. Notice that since the monopolist sells to lower type buyers only if the belief increases, the implied prices can be non-monotone.

The option value effect creates delays in sales. The monopolist wants to spread sales over time because postponing enables sales to be made with more information. Compared with markets with exogenous learning, endogenous learning favors a faster start and steeper decline in quantity of sales if news is not good enough. The predicted price paths also differ in that initial offers are suggested to feature the markets with endogenous learning.

Figure 3 shows the price (blue) as a function of the optimal quantity the monopolist wants to set. The figure also plots the valuation of the marginal buyer (orange) and

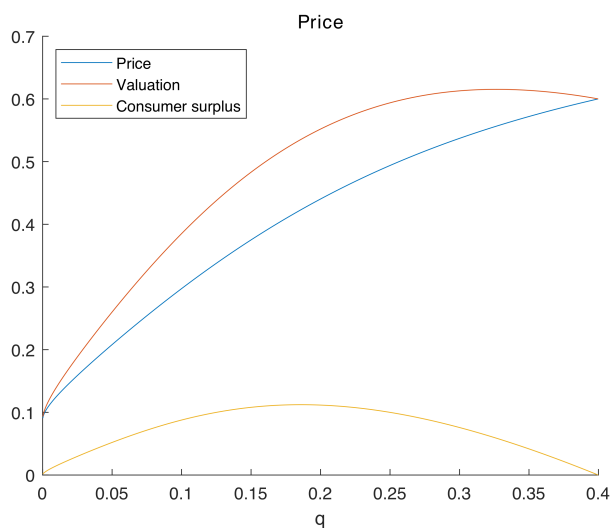


Figure 3: Numerical example of the price, the marginal buyer’s valuation and the consumer surplus as a function of the optimal quantity sold.

the consumer surplus (yellow) given by the value of buyers,  $W(\theta)$ . The price is the difference between these two (details how to calculate it are in Appendix C). It is important to note here that as quantity changes so do *both the marginal type and the belief* because the figure depicts the optimal price given the optimal quantity. As we see, the valuation of the buyers is first increasing in the quantity as the belief has to be higher for the monopolist to find it optimal to sell larger quantities. However, since the type of the marginal buyer is decreasing over quantity at some point the valuation starts to decrease, which means the valuation is non-monotonic. In contrast, the price in the figure is monotonic because the monopolist has to take the incentives to delay purchases into account in her pricing. An upward sloping price incentivizes the buyers to purchase today rather than later and is the only way the monopolist can implement the optimal quantity path. Finally, the consumer surplus starts from zero and ends up in zero because the value of waiting is zero at both very low and very high beliefs. In the former case, the monopolist is very unlikely to sell more units and in the latter case the monopolist has already sold to almost all buyers she is willing to sell to.

### 4.3 Competitive market

We contrast the monopoly solution with the competitive market. Suppose there are no barriers of entry to the market. Then firms will enter the market as long as the price exceeds the marginal cost. When the belief is high, this means that there will be entry to the market until the price equals marginal cost. When the belief is low, the price must still equal the marginal cost because no firm is willing to produce below the marginal cost as there are no gains from learning. Therefore, in a competitive market with free entry, the price equals the marginal cost independent of learning.

From the perspective of an individual buyer, the information process is exogenous and therefore he ignores the information effect in his decision making. The buyers still have an option to wait. Waiting enables purchasing with more information but comes with the cost of delayed consumption. Aggregate consumption consists of a sequence of stopping problems, similar to those in presented in the short-sighted solution (Subsection 3.1). Using Proposition 2 in the durable goods model leads to the following proposition:

**Corollary 2.** *The competitive market quantity equals*

$$x^S(q) = \frac{\beta(q)c}{(\beta(q) - 1)\theta(q) + c},$$

where  $\theta(q) : 1 - F(\theta(q)) = q$ .

Notice that  $x^S(q)$  is guaranteed to be always between 0 and 1 because  $\theta(q) - c > 0$  and  $\beta(q) > 1$  for all  $q < q^1$ .

The competitive market clearly exhibits the option value effect: it is always above the static optimum,  $x^{stat}(q) = c/\theta(q)$ . The smaller  $\beta(q)$  is, the further away the competitive solution is from the static solution. This happens when the learning technology is good, i.e.  $\lambda$  is large. This should not be interpreted that a better learning technology is harmful for the overall efficiency. Even though there is a larger incentive for informational free-riding, a better learning technology also speeds up learning and hence the belief converges to the true value faster, which naturally improves efficiency of the solution. Nevertheless, we can conclude that the inefficiency in the competitive market increases as the learning technology improves.

#### 4.4 Welfare effects

We can use the general model and its optimal solution to understand the socially efficient consumption path. A benevolent social planner faces again the trade-off between information generation and the option value of waiting. Her problem is otherwise equivalent with the monopolist's except that the planner maximizes the expected surplus (e.g. uses first degree price discrimination), not the virtual surplus. The planner's objective is thus

$$r\mathbb{E}\left[\int_0^\infty e^{-rt} \left(\int_0^{Q_t} (X_t\theta(q) - c)dq\right) dt\right],$$

where  $\theta(q) : 1 - F(\theta(q)) = q$ .

The planner's policy, captured by  $x^P(q)$ , is such that the planner always sells at lower beliefs than the monopolist. The monopolist takes into account how price expectations affect today's willingness to pay and hence commits to lower quantities in the future. Without endogenous learning, the monopoly allocation is fully efficient for the highest type buyer. When taking into account the informational externality

consumption creates, this is no longer true when learning is endogenous. Because the value of generating information is lower under the monopoly solution, a higher belief is needed for the monopolist to be willing to launch the product: there is a distortion at the top as well as at the bottom of the type distribution. Figure 4 illustrates the planner's, the monopolist's, and the competitive market quantities as functions of the belief.

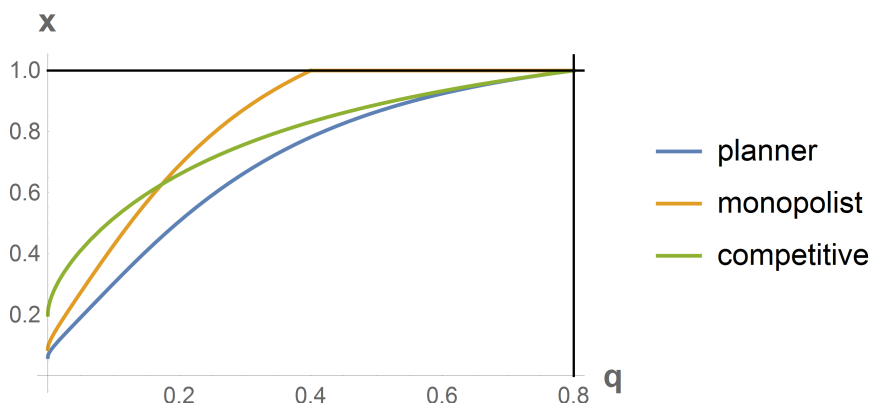


Figure 4: Different solutions for uniform  $(0, 1)$  types,  $c = 0.2$ ,  $r = 0.1$ , and  $\sigma = 0.5$ .

The monopolist creates the usual monopoly distortion by limiting the amounts sold in order to charge higher prices from higher type buyers. The competitive market, on the other hand, does not internalize the informational externality. Thus, whether the competitive market produces more consumer surplus than the monopolist greatly depends on the initial belief. With a low enough initial belief, the product is not launched in the competitive market at all, while a monopolist is willing to suffer a short-term loss in order to enable profitable sales in the future. When the product is likely to be of high quality, both the information generation and the option value effects disappear. Then, the competitive market approaches the planner's solution but the monopoly solution has the usual monopoly distortion in quantity.

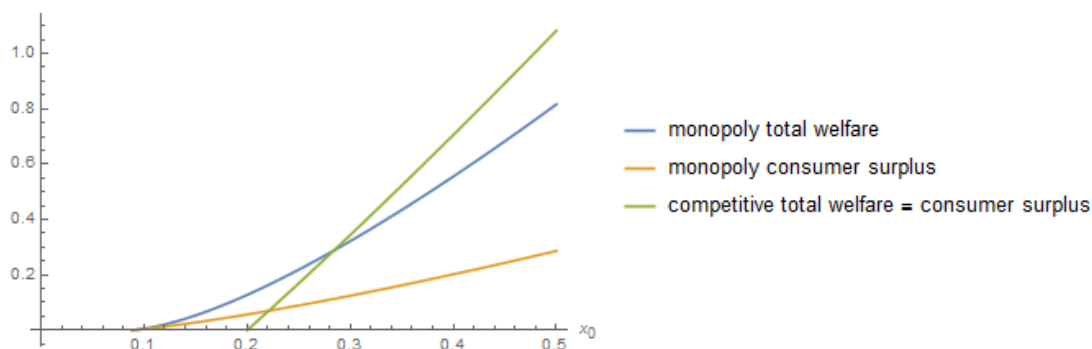


Figure 5: Expected total and consumer surplus as a function of the initial belief. Parameter values: type distribution is uniform on  $[0, 1]$ ,  $c = 0.2$ ,  $r = 0.1$ , and  $\sigma = 0.5$ .

In the terms of the total and consumer welfare, the monopoly solution is better

than the competitive market for low initial beliefs. Figure 5 presents the expected total and consumer surplus as functions of the initial belief, when using the same parameter values as in Figure 4. In the Figure, there is a cutoff for the initial belief such that below it the monopolist provides more total welfare than the competitive market below the cutoff and less above it. Similarly, there is a lower cutoff determining whether the monopolist or the competitive market produces more consumer surplus. For a general type distribution we obtain a slightly weaker result with potentially multiple cutoffs:

**Proposition 4.** *There exist cutoffs  $x_a \in (0, 1)$  and  $x_b \in (0, 1)$  such that the monopoly is more efficient if the initial belief is below  $x_a$  and the competitive market is more efficient if the initial belief is above  $x_b$ .*

*Proof.* The existence of  $x_b < 1$  directly follows from that the shortsighted and the optimal solution are continuous and meet at  $x = 1$  (Lemma 3) and that the monopoly boundary,  $x^M$ , is everywhere above the planner's boundary,  $x^P$ .

The existence of  $x_a > 0$  can be seen by comparing the value matching conditions at  $q = 0$ . For the competitive market, we use the shortsighted condition and plug in  $q = 0$ :

$$x\theta(0) - c = 0.$$

For the monopolist, the equivalent condition is obtained by using the full solution and the monopolist's objective:

$$x\phi(0) - c + \frac{1}{2r} V_{xx}(x, q) \frac{x^2(1-x)^2}{\sigma^2} = 0.$$

Since  $\phi(0) = \theta(0)$  and  $V_{xx}(x, q) > 0$ , we have  $x^M(0) < x^C(0)$ . The monopolist launches the product with lower beliefs. Since we also have  $x^M(0) > x^P(0)$ , this improves total efficiency.  $\square$

Notice that we can use an almost identical proof to show the existence of cutoffs determining which solution produces more consumer surplus. It is evident that the monopoly solution produces more consumer surplus when the competitive market does not launch the product or launches it very slowly.

As a final point, we analyze sales over time. The monopolist tends to produce more in the beginning but less later on, when compared with the competitive market *at the same belief level*. The belief process is, however, endogenously determined and hence it is not the same for the two solutions. Conditional on the true quality being high, the process reaches higher beliefs faster with larger quantities, which makes the comparison over time more favorable for the monopoly solution. Figure 6 illustrates this effect.

In Figure 6, the figure at the top shows how the belief changes over time in the monopoly and competitive solutions for a realized path of the Wiener process. We have drawn the sample path conditional on the true quality being high. As we can

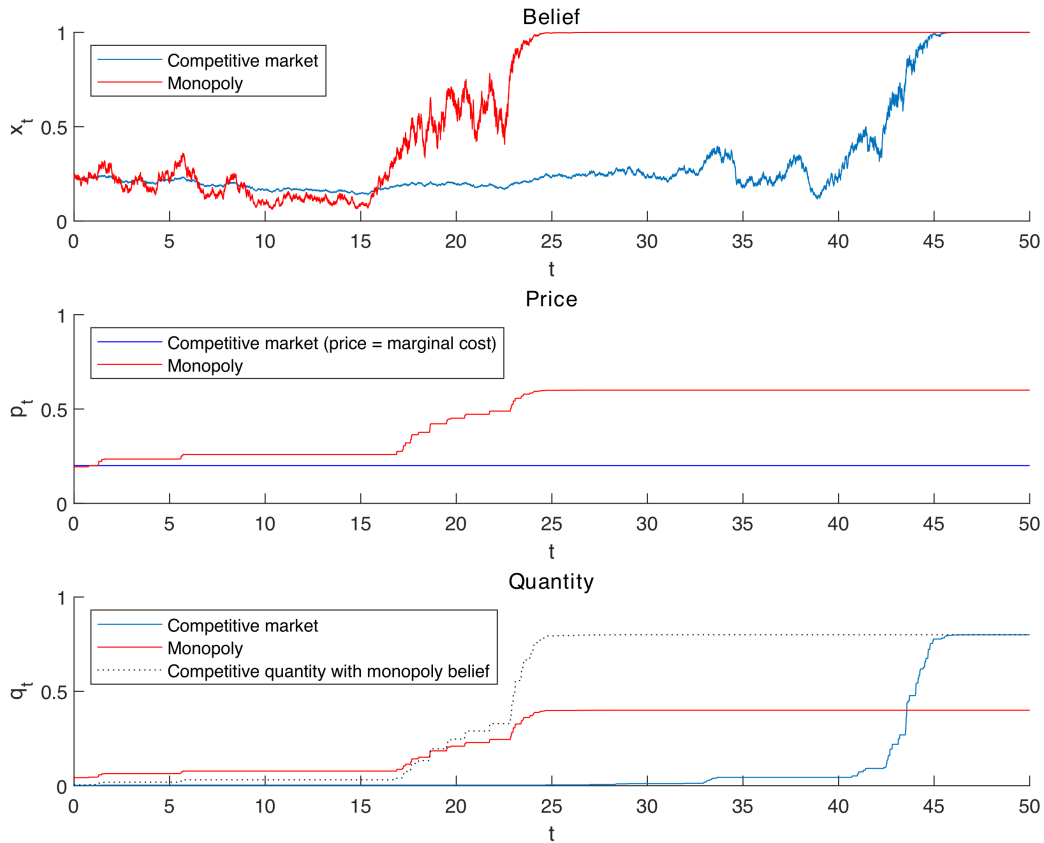


Figure 6: Sample paths conditional on the true quality being high for the competitive and monopoly solutions. The type distribution is uniform on  $[0, 1]$ ,  $c = 0.2$ ,  $r = 0.1$ , and  $\sigma = 0.5$ . The top figure shows how the belief evolves over time, the middle figure shows how the price evolves over time and the bottom figure shows how quantity evolves over time. The dotted line shows the competitive quantity when the belief follows the law of motion from the monopoly solution (the red  $x_t$  path). All paths are based on the same realized path of a Wiener process.

see, the monopoly belief changes more rapidly over time because it entails selling larger quantities earlier. The middle figure the price at each instant of time. For the competitive market the price equals the marginal cost whereas the monopolist sets a price that depends on the current belief and thus changes over time. The figure below gives the corresponding quantity paths: a red path for the monopolist, a blue path for the competitive market, and a dotted grey path which tells what the competitive quantities would be if the belief was taken from the monopoly solution. The point at which the red and the dotted paths intersect is the time when the belief in the monopoly solution is such that the monopoly and the competitive quantities are equal. The true competitive path reaches the monopoly path later as the belief changes more slowly in the competitive market.

## 5 Other applications

### 5.1 Incremental investments

In this subsection, we show that the general model can be used to analyze the problem of optimal investments to production capacity when demand is uncertain. The difference to the durable goods model is that there is an irreversible cost to build capacity and consumers make repeat purchases. With exogenous uncertainty, Leahy (1993) shows the equivalence between the competitive equilibrium, the social planner's solution, and the case where firms do not take into account how future investments affect the price. The last case is the shortsighted solution in our model, referred to as 'the myopic problem' in Leahy (1993). We show that while the equivalence between the competitive and shortsighted solutions remains, the equivalence with the planner's problem breaks down when learning is endogenous.

#### 5.1.1 Model

Consider a market using capital as the only input. There is a constant marginal cost  $c$  to increase capital, and we assume that all capital is in production. We normalize and let the production level equal the current capital stock  $q$ .<sup>8</sup> Let the inverse demand be

$$p_t = x_t D_H(q_t) + (1 - x_t) D_L(q_t) \leq 0,$$

where  $x_t$  is the belief that the state of the world is high and  $D_\omega$  is the inverse demand in state  $\omega \in \{H, L\}$ .

The market learns about the state of the world from consuming goods produced by the invested capital: the rate of learning is scaled by the current production capacity  $q_t$ . More specifically, let the belief follow the diffusion process in (3). The interpretation of the learning process is now that only new purchases generate information: consumers make *repeat* purchases and return to the market at each period. Furthermore, the rate of purchases is pinned down by investment decisions in the past.

#### 5.1.2 Optimal investments

We start our analysis with the competitive equilibrium. The market consists of a sequence of infinitesimally small firms indexed by  $q \in (0, q^1)$ . A firm chooses the optimal time to invest  $\tau$ , which is a stopping rule adapted to the natural filtration of the information process and past quantities, to solve

$$\sup_{\tau} \mathbb{E} \left[ e^{-r\tau} \left( \int_{\tau}^{\infty} e^{-r(t-\tau)} (X_t D_H(Q_t) + (1 - X_t) D_L(Q_t)) dt - c \right) \right]$$

subject to (3).

---

<sup>8</sup>This is equivalent to the marginal cost of using capacity being zero and the consumer market being competitive.



Leahy (1993) shows that the firms can neglect the effect of future investments to the price. This holds for our model as well because individual firms do not internalize the learning externality and optimize as if learning is exogenous. This together with the martingale property of the belief greatly simplifies the firm's problem, which we can write more simply as:

$$\sup_{\tau} \mathbb{E} \left[ e^{-r\tau} \left( \frac{1}{r} (X_{\tau} D_H(Q_{\tau}) + (1 - X_{\tau}) D_L(Q_{\tau})) - c \right) \right].$$

We already know from Subsection 3.1 that this kind of sequence of stopping problems is equivalent with the shortsighted solution of the problem in section 3.

As for the social planner, the problem is to choose an investment rule  $Q_t$  to maximize the social surplus<sup>9</sup>:

$$\mathbb{E}_x \left[ \int_0^{\infty} \int_0^{Q_t} e^{-rt} (X_t D_H(q) + (1 - X_t) D_L(q) - rc) dq dt \right].$$

Now, this is a special case of the model in Section 2 when we specify  $\gamma_{\omega}(q) = D_{\omega}(q) - rc$ . Therefore, we can readily conclude that the competitive market, which we have argued to resemble the shortsighted solution to the same problem, produces too little at every belief level.

A monopoly problem is similar to the planner's problem except that the objective function is different. The monopolist maximizes her own profits:  $(x_t D_H(q_t) + (1 - x_t) D_L(q_t) - rc) q_t$ . For  $q_t = 0$ , the marginal payoffs of the monopolist and the planner coincide but the incentives to generate information do not as the planner is willing to expand more later on and hence gains more from speeding up learning. Nevertheless, the net effect of endogenous learning for total welfare in the monopoly solution is positive for low enough beliefs as shown by Proposition 3. We therefore conclude that for low beliefs, the monopolist invests more whereas the competitive market invests less than they would without learning.

We have modeled the investor's problem so that there is uncertainty about future demand for capacity. We could equally well analyze uncertainty with regards to production technology, such as operation costs or other aspects of the productivity of capital. We could also incorporate learning from firms using the same technology but operating in a different market. With that interpretation, an individual firm's problem would be similar to the buyers' problem in the durable goods market with marginal cost pricing. As discussed in this subsection and formally shown in Leahy (1993), the effect investments have on prices can be neglected in the equilibrium analysis. Hence, the equilibrium behavior with social learning about production technology or about demand can both be analyzed by using the shortsighted solution.

---

<sup>9</sup>We can write the cost of investment as a flow cost because  $X_t$  is a martingale.

## 5.2 Environmental policy

In environmental economics, many actions have irreversible impacts. Once made, it is hard to to reduce the stock of emissions. As an application of our general model, in this section we solve for an optimal carbon tax when damages from CO2 emissions are uncertain.

### 5.2.1 Model

Let  $z \in [0, \bar{z}]$  be the stock of fossil fuel (polluting resource) in the ground. The fuel is owned by a continuum of producers and each producer chooses when to extract and sell the fuel they own. Let  $\theta \in [\underline{\theta}, \bar{\theta}]$  with  $\underline{\theta} > 0$  denote a producer's (private) marginal payoff from extracting a unit of the resource and selling it in the market. The marginal payoffs are distributed according to  $F$ , which satisfies the same assumptions as in the durable goods application.<sup>10</sup>

Let  $q$  denote the stock of emissions caused by the use of the fossil fuel. For simplicity, assume that one unit of fossil fuel causes one unit of emissions:  $q = \bar{z} - z$ .<sup>11</sup> Emissions cause damages, which depend on the state of the world and the stock of emissions. In the low state of the world marginal damages are  $D_L(q)$  and in the high state they are  $D_H(q)$  where  $q$  is the stock of emissions. Both  $D_\omega(q)$  are twice differentiable and weakly convex functions such that  $D_H(q) < D_L(q)$  for all  $q$ . News about damages arrive according to equation (2) so the belief about the state of the world is governed by equation (3).

A social planner can influence the producers decisions to emit pollution by setting a tax  $P_\tau$ , which is allowed to depend on the realized news and past emissions. Effectively, the planner is deciding the stock of emissions, e.g. setting a 'carbon budget' for the economy.

### 5.2.2 Producers' problem

A producer decides when to extract the resource he owns. Because each individual producer is infinitesimally small and thus disregards aggregate effects, we can model their problem as choosing a stopping rule for each  $z \in [0, \bar{z}]$ . As in Section 4 given the planner's tax policy  $P_\tau$ , the stopping problem of an individual producer is as if learning is exogenous. Thus, we can formulate the problem as choosing a stopping rule  $\tau \in \mathcal{T}$  for  $\theta \in [\underline{\theta}, \bar{\theta}]$  such that it maximizes

$$\mathbb{E}[e^{-r\tau}(\theta - P_\tau)].$$

The results from Section 4 follow identically for the problem here: producers' decisions are monotone in  $\theta$  and thus there is one-to-one mapping from the stock of

<sup>10</sup> $F$  has twice continuously differentiable density  $f$  and satisfies the monotone hazard rate condition.

<sup>11</sup>We assume  $\bar{z}$  to be large enough to have an interior optimum.

emissions to types (Lemma 4). We have  $z = F(\theta)\bar{z}$  or  $q = (1 - F(\theta))\bar{z}$ . Furthermore, we can represent the value of a producer (or the value of having a marginal payoff of  $\theta$ ),  $W(\theta)$  with the help of the envelope theorem. The producers' problem is otherwise identical with the buyers' problem in the durable goods model except that the producers do not care about the state of the world directly. It affects their value only through taxes the planner sets.

### 5.2.3 Planner's problem

The social planner sets the tax on emissions to maximize the total welfare of the economy. Payoffs to the producers represent the benefit of emissions and damages the cost. We can write the planner's problem over the stock of emissions  $q \in [0, \bar{z}]$  as

$$r\mathbb{E}\left[\int_0^q e^{-r\tau(q)}\left(S(q) - X_{\tau(q)}D_H(q) - (1 - X_{\tau(q)})D_L(q)\right)dq\right], \quad (12)$$

where we have normalized the payoff to a flow and  $S(q)$  is the benefit from emissions defined as

$$S(q) = \theta(q).$$

Since  $\theta(q)$  is determined by  $q = (1 - F(\theta))\bar{z}$ , the shape of the distribution  $F$  determines also the shape of the benefit function  $S(q)$ . An immediate consequence of the monotonicity from the producers decisions is that the benefits are decreasing in quantity. The realized tax  $p(x, q) = \theta(q) - W(x, q)$  is set so that the marginal extracting producer at state  $(x, q)$  is indifferent between extracting and not extracting. The producer's value of waiting,  $W(x, q)$ , depends on the belief through expected future taxes.

### 5.2.4 Optimal taxation of emissions

The planner's problem is a special case of our general model. In this application, the decision maker's payoffs are

$$\begin{aligned} \gamma_H(q) &= r(S(q) - D_H(q)), \\ \gamma_L(q) &= r(S(q) - D_L(q)). \end{aligned}$$

Given that this application is just a special case of our general model, the optimal quantity is then given by  $x^*(q)$  in Proposition 1. The discussion regarding the information generation and the option value effects is relevant here as well: for low beliefs the planner might want to emit simply to speed up learning and to do so she sets a less strict emissions policy than in the static solution. In contrast, for high beliefs she wants to set a much stricter policy than in the static solution to take into account the option value of delaying emissions.

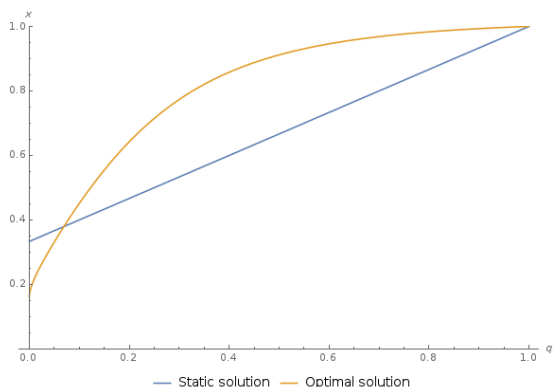


Figure 7: Optimal emissions

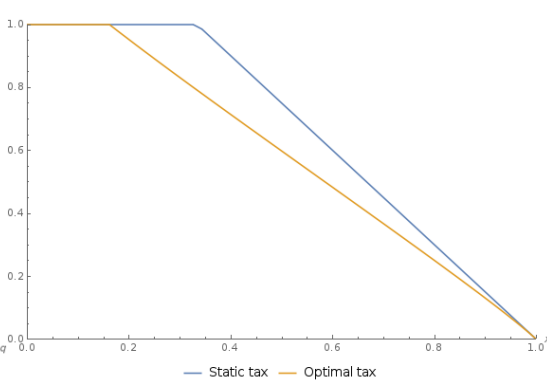


Figure 8: Optimal taxes

Figure 7 depicts the optimal emissions policy (orange) and the corresponding static solution (blue) in a numerical example.<sup>12</sup> As we see, the optimal emissions are first larger than in the static solution (the belief boundary is lower) but for higher beliefs it quickly becomes optimal to emit less than in the static solution. This is because after enough information is being produced, the irreversibility of emissions means that it is optimal to limit emissions until the state is known with more certainty. The static policy in this example is simply a linear function of the emissions because we have assumed constant marginal damages.<sup>13</sup>

Figure 8 depicts the tax the planner must set to implement the optimal policy (orange) and the corresponding static tax (blue) in the same numerical example as Figure 7. Both are given as functions of the belief given the (optimal) behavior of the planner. Details how to calculate the tax can be found in the Appendix (section C). Except for very low beliefs, when there are no emissions, we see that the planner sets a strictly *decreasing* path of taxes over the belief, which is natural as the optimal emissions are increasing over the belief. Furthermore, as soon as there is a positive amount of emissions, the optimal and static taxes diverge as can be seen in the figure. It is clear that the optimal tax must be below the static tax for low enough beliefs – emissions are larger – but it is less clear why this should be the case for higher beliefs when the static emissions are larger. The reason for the latter is that the optimal tax is constrained by the incentive compatibility of the producers and they accrue rents from their private type and the option to emit later. Thus, to ensure incentive compatibility the planner has to set taxes lower than the marginal producer’s valuation, which in this example means that the optimal tax stays (weakly) below the static tax for all beliefs.

What happens if there is no tax on emissions? As we explain in Section 3.1, the short-sighted solution entails informational free-riding: small sequential players do not

<sup>12</sup>Parameters:  $\sigma = 1$ ,  $r = 0.05$ , constant marginal damages with  $D_H(q) = 0$ ,  $D_L = -1.5$  and uniform distribution of producers’ values between  $[0, 1]$ .

<sup>13</sup>Constant marginal damages from emissions to welfare are fairly common in the literature on climate change and CO2 emissions see for example Golosov et al. (2014) or Traeger (2019).

internalize the gains from faster information production. In a game of a public bad, such as in this section, there is another crucial externality: individual players do not internalize the damages to other players. Thus, if there is no tax on emissions, then the equilibrium to the game is such that we have  $q = \bar{z}$  immediately. This is because while the owners of the resource internalize the benefits from emissions, they only internalize an infinitesimally small part of the expected damages. Both the optimal solution and static solution are always better in terms of welfare than the no tax equilibrium.

We have here modelled the problem of emissions. Equivalently, we could have modelled the problem as a providing public good by cutting emissions from some ‘business as usual’ trajectory. The lessons remain the same in this setup: while it is optimal to let emissions rise in the beginning, irreversibility means that sooner or later it is optimal for the government to implement a stricter environmental policy than in the static solution.

### 5.3 Additional applications

The trade-off between the option value and the information generation is present in many contexts we do not explicitly model here. Prominent other applications include R&D and development aid projects, education and medical treatments.

Consider a development aid agency facing a problem whether it should start a new aid project or stick to the well-trying ones. A new project requires long-lasting commitment and the potential gains arrive gradually over time. The results of this paper suggest that a strictly milder criterion should be used for launching the new project than for the established ones. But then it is reasonable to require strictly larger achievements and more evidence of the impact before scaling it up. The decision over the size of a R&D project is similar, suggesting fast starts and conservative extensions.

There are many sources of irreversibility in public policy making. The policies themselves may be hard to change because of political uncertainty, potential legal consequences, or preparation time needed for a legislative proposal. Even if changed, the policy has already affected many people and those effects may be irreversible. Think of, for example, policy norms limiting the size of a class room. Each cohort is affected by the norm at the time when they go to school. Education targets to long-term effects and hence the complete policy evaluation need to be based on both short-term and long-term outcomes. Each cohort affected by changed norms keeps producing more information over time, independent of whether the policy is still active. A pure market outcome may not produce enough experimentation because individual schools and parents do not internalize the information generation benefits to other schools. One interpretation of our results is that it can be a good idea to first mandate the implementation of a new norm and then to suspend it until there is information available from the first cohorts.

Many medical treatments and devices such as artificial joint replacements require a long time before their full effects can be assessed. While they go through rigorous

testing before approved for use, the market is full examples where operations or products have been withdrawn due to concerns for the well-being of the patients.<sup>14</sup> Because this is a problem of gradual learning and undoing such operations is either impossible or very costly our results have direct implications how to phase such products to the market: a large enough initial launch and then wait for evidence of success.

## 6 Conclusions

This paper analyzes an optimal control problem with endogenous learning about the profitability of quantity expansions. We extend the literature to cover situations where each action produces information gradually over time. Compared to earlier literature, gradual learning changes the incentives of the decision maker as she needs to balance the increased rate of learning generated by early activity with the risk of later discovering that the irreversible decision was suboptimal. Endogenous learning calls for fast expansions when the current quantity is low but caution when the quantity is already high.

One of the implications of learning from past actions is that it breaks the link between the optimal control problem and the series of optimal stopping problems, which has been established in the previous literature. When learning is endogenous, the competitive equilibrium is no longer efficient. This is because the sequence of small firms does not take into account the information generation effect of production. This is true both in durable goods markets and in markets with costly capacity expansions and repeat purchases. In a competitive environment, decision makers do not internalize the potential benefit of information generation, and thus a less competitive market may be better prepared to balance informational trade-offs. This suggests that copyrights and patents may have positive welfare effects even *after* the innovation has been made. In a market with a lot of initial uncertainty, exclusive rights incentivize the launching of new products.

---

<sup>14</sup>See for example DePuy hip replacement recall or the FDA list of recalled medical devices for 2018. All links retrieved on November 27, 2018.

## Appendix

### A Optimal quantity path

#### Proof of Proposition 1

Recall that the value matching and smooth pasting conditions are

$$\begin{aligned} \frac{x\gamma_H(q) + (1-x)\gamma_L(q)}{r} + B_q(q)\Phi(x, q) + B(q)\Phi_q(x, q) &= 0, \\ \frac{\gamma_H(q) - \gamma_L(q)}{r} + B_q(q)\Phi_x(x, q) + B(q)\Phi_{qx}(x, q) &= 0, \end{aligned}$$

for  $x = x^*(q)$  at the boundary. For notational convenience we do not substitute  $x^*(q)$  in.

Solving the equations above for  $B_q(q)$  and  $B(q)$  yields

$$B_q(q) = A^1(x, q)x + A^2(x, q), \quad (13)$$

$$B(q) = U^1(x, q)x + U^2(x, q), \quad (14)$$

where

$$\begin{aligned} A^1(x, q) &: = \frac{-\Phi_{qx}(x, q) \left( \frac{\gamma_H(q) - \gamma_L(q)}{r} \right)}{\Phi(x, q)\Phi_{qx}(x, q) - \Phi_q(x, q)\Phi_x(x, q)}, \\ A^2(x, q) &: = \frac{\Phi_{qx}(x, q) \left( \frac{-\gamma_L}{r} \right) + \Phi_q(x, q) \left( \frac{\gamma_H(q) - \gamma_L}{r} \right)}{\Phi(x, q)\Phi_{qx}(x, q) - \Phi_q(x, q)\Phi_x(x, q)}, \\ U^1(x, q) &: = \frac{\Phi_x(x, q) \left( \frac{\gamma_H(q) - \gamma_L}{r} \right)}{\Phi(x, q)\Phi_{qx}(x, q) - \Phi_q(x, q)\Phi_x(x, q)}, \\ U^2(x, q) &: = \frac{-\Phi_x(x, q) \left( \frac{-\gamma_L}{r} \right) - \Phi(x, q) \left( \frac{\gamma_H(q) - \gamma_L}{r} \right)}{\Phi(x, q)\Phi_{qx}(x, q) - \Phi_q(x, q)\Phi_x(x, q)}. \end{aligned}$$

The value matching and smooth pasting conditions do not simplify as in the short-sighted case because we have both the unknown function  $B$  and its derivative in the equations. Differentiating (14) with respect to  $q$  and using the chain rule gives

$$\begin{aligned} B_q(q) &= \left[ U_1^1(x, q) \frac{\partial x}{\partial q} + U_2^1(x, q) \right] x \\ &\quad + U^1(x, q) \frac{\partial x}{\partial q} + U_1^2(x, q) \frac{\partial x}{\partial q} + U_2^2(x, q) \\ &=: U^3(x, q) \frac{\partial x}{\partial q} + U^4(x, q), \end{aligned} \quad (15)$$

where subscripts denote derivatives with respect to first and second argument.

Equating (13) and (15) then gives:

$$A^1(x, q)x + A^2(x, q) = U^3(x, q) \frac{\partial x}{\partial q} + U^4(x, q).$$

Or

$$x^{*'}(q) = \frac{A^1(x, q)x + A^2(x, q) - U^4(x, q)}{U^3(x, q)}.$$

This then yields the expression in the text by plugging in  $A^1, A^2, U^3$  and  $U^4$ . The existence of a solution to the differential equation determined by (8) and the initial value  $x^*(q^1) = 1$  follows from Peano existence theorem.

**Uniqueness of the boundary.** The solution to the differential equation given by (8) and the initial value  $x(q^1) = 1$  is unique from *below*. To see this, note first that derivative of  $g(x, q)$  with regards to  $x$  equals

$$\begin{aligned} g_x(x, q) = & - \left[ \beta''(q) \left( x^2(1-2x)(\beta(q)-1)^3 \gamma_H(q)^2 - 2(1-x)x\beta(q)(\beta(q)-1) \times \right. \right. \\ & \gamma_H(q)\gamma_L(q)((1-2x)\beta(q)-x) + (1-x)^2(1-2x)\beta(q)^3 \gamma_L(q)^2 \left. \right) \\ & + \beta'(q) \left( 2x^2(2x-1)(\beta(q)-1)^2 \gamma_H(q)^2 \beta'(q) + (1-x)^2 \beta(q)^2 \gamma_L(q) \left( 2(1-2x) \right. \right. \\ & \gamma_L(q)\beta'(q) - 2x(\beta(q)-1)\gamma'_H(q) - (1-2x)\beta(q)\gamma'_L(q) \left. \right) \\ & + x\gamma_H(q) \left( 4(1-x)\gamma_L(q)\beta'(q) \left( (1-2x)\beta(q)^2 + 2x\beta(q) + x \right) \right. \\ & \left. \left. - x(\beta(q)-1)^2 \left( (1-2x)(\beta(q)-1)\gamma'_H(q) + 2(1-x)\beta(q)\gamma'_L(q) \right) \right) \right] / \\ & \left[ (x(\beta(q)-1)^2 \gamma_H(q) + (1-x)(\beta(q))^2 \gamma_L(q))^2 \beta'(q) \right]. \end{aligned}$$

The derivative is finite (denominator is non-zero) as long as  $x < x^d(q)$ , where  $x^d(q) = (\beta(q)^2 \gamma_L(q)) / (\beta(q)^2 \gamma_L(q) - (\beta(q)-1)^2 \gamma_H(q))$ , which is true for the optimal boundary  $x^*(q)$  as long as  $q < q^1$  (see Lemma 1). Thus  $g(x^*(q), q)$  is Lipschitz continuous for all  $q < q^1$ .

From Lemma 3,  $\lim_{q \rightarrow q^1} (x^*(q) - x^S(q)) = 0$  and  $\lim_{q \rightarrow q^1} (x^*(q) - x^S(q)) / (q - q^1) = 0$ . Any differential equation satisfying  $x_q(q) = g(x, q)$  and having an initial value  $x(q^1 - \varepsilon) = x^S(q^1 - \varepsilon)$  has a unique solution (by the Picard-Lindelöf theorem), where  $x^S$  is the shortsighted solution. Therefore, the sequence of the unique solutions of initial value problems  $x_q(q) = g(x, q)$  and  $x(q^1 - \varepsilon) = x^S(q^1 - \varepsilon)$  converges uniformly to  $x^*(q)$  as  $\varepsilon \rightarrow 0$ . Hence the limit is unique.

**Verification.** We still need to verify that the optimal boundary together with the solution to the HJB equation solve the original problem. The verification of the solution follows from the standard arguments in the literature (see e.g. Fleming and Soner (2006)).

First note that value matching and smooth pasting are necessary conditions for a strategy to maximize the HJB equation pointwise. A pointwise maximum exists by continuity of the HJB equation and compactness of the strategy space ( $x^*(q) \in [0, 1]$ ). Value matching and smooth pasting together imply the differential equation for the optimal boundary which has a unique solution as we have shown. Thus, the optimal boundary given above is the unique pointwise maximizer of the HJB equation.



Let  $h(x, q) = B(q)\Phi(x, q)$  denote our solution to the HJB equation and let  $T \geq t$  be an arbitrary stopping time at which the game ends (we will take the limit as  $T$  goes to infinity). From Itô's formula we have

$$\begin{aligned} e^{-rT}h(x_T, q_T) &= e^{-rt}h(x_t, q_t) + \int_t^T e^{-rs} \left[ -rh(X_s, Q_s) \right. \\ &\quad \left. + \frac{1}{2} \frac{X_s^2(1-X_s)^2}{\sigma^2} Q_s h_{xx}(X_s, Q_s) \right] ds \\ &\quad + \int_t^T e^{-rs} \frac{X_s(1-X_s)\sqrt{Q_s}}{\sigma} h_{xx}(X_s, Q_s) dW_s \\ &\leq e^{-rt}h(x_t, q_t) - \int_t^T e^{-rs} (X_s \pi_H(Q_s) + (1-X_s) \pi_L(Q_s)) ds \\ &\quad + \int_t^T e^{-rs} \frac{X_s(1-X_s)\sqrt{Q_s}}{\sigma} h_{xx}(X_s, Q_s) dW_s. \end{aligned}$$

The inequality follows from the HJB equation for any policy that induces admissible  $Q_t$  process. Since we let  $T$  denote the stopping time for the game ending we have  $h(x_T, q_T) = 0$ . Taking conditional expectations, multiplying with  $e^{rt}$  and simplifying then gives

$$h(x_t, q_t) \geq \mathbb{E} \left[ \int_t^T e^{-r(t-s)} (X_s \pi_H(Q_s) + (1-X_s) \pi_L(Q_s)) ds \middle| \mathcal{F}_t \right].$$

The value to the DM is bounded and therefore clearly satisfies the following transversality condition:

$$\lim_{T \rightarrow \infty} \mathbb{E}[e^{-rT}h(x_T, q_T)] = 0.$$

This implies that the limit of the right-hand side is well defined as  $T \rightarrow \infty$  so we can take the limit. Therefore, we have that  $h(x_t, q_t) \geq \max_{Q_t} V(Q_t; x, q)$  always.

The last step is to use the fact that  $Q_t$  induced by  $x^*(q)$  achieves the pointwise maximum of the HJB-equation and thus the inequalities above become equalities. Thus,  $h(x_t, q_t) = \max_{Q_t} V(Q_t; x, q)$ . This verifies that our solution indeed solves the original problem.

The verification of the shortsighted solution follows from nearly identical steps. The main difference is that we only need that it is optimal to stop at  $x = x^S(q)$ , which follows from the necessity of the value matching and smooth pasting conditions.

## Lemma 6

**Lemma 6.** *For all  $q < q^1$  on the boundaries  $(q, x^*(q))$  and  $(q, x^S(q))$ ,  $\gamma_H(q) > 0$  and  $\gamma_L(q) < 0$ .*

*Proof.* Assume on the contrary that  $\gamma_H(q) \leq 0$ . But  $\gamma_H(q) > \gamma_L(q)$  and hence  $\gamma_L(q) < 0$ , which then implies that the flow payoff would be increased by decreasing  $q$  for

$x < 1$ . Hence the DM would strictly prefer decreasing it and thus  $q$  cannot be on the boundary. The proof for the case  $\gamma_L \geq 0$  is the same but relies on the DM strictly preferring to increase the quantity when both  $\gamma_H(q)$  and  $\gamma_L(q)$  are positive.  $\square$

## Proof of Lemma 1

*Proof.* Recall first the equation (8):

$$\begin{aligned} g(x, q) = & x(1-x) \left[ x(\beta'(q)(\beta(q)-1)\gamma'_H(q) - ((\beta(q)-1)\beta''(q) - 2(\beta'(q))^2)\gamma_H(q)) \right. \\ & \left. + (1-x)(\beta'(q)\beta(q)\gamma'_L(q) - (\beta(q)\beta''(q) - 2(\beta'(q))^2)\gamma_L(q)) \right] / \\ & \left[ (x(\beta(q)-1)^2\gamma_H(q) + (1-x)(\beta(q))^2\gamma_L(q))\beta'(q) \right]. \end{aligned} \quad (16)$$

And the shortsighted solution:

$$x^S(q) = \frac{\beta(q)\gamma_L(q)}{\beta(q)\gamma_L(q) + (1-\beta(q))\gamma_H(q)}.$$

We have that  $\gamma_H(q) > 0$  and  $\gamma_L(q) < 0$  for all boundary  $q < q^1$  and that  $\beta(q) > 1$  (see Lemma 6 in the appendix for a proof). Because  $\beta'(q) < 0$  for all  $q \geq 0$ , the denominator is then strictly positive for all  $x \in (0, x^d(q))$ , where

$$x^d(q) = \frac{\beta(q)^2\gamma_L(q)}{\beta(q)^2\gamma_L(q) - (\beta(q)-1)^2\gamma_H(q)}.$$

To see that the optimal policy  $x^*(q)$  is below  $x^d(q)$ , we make use of the shortsighted cutoff  $x^S(q)$ . The shortsighted solution is always above the optimal solution. Hence, it is enough to show the inequality for beliefs below the shortsighted policy. The condition  $x^d(q) > x^S(q)$  equals

$$\frac{(\beta(q))^2\gamma_L(q)}{(\beta(q))^2\gamma_L(q) - (\beta(q)-1)^2\gamma_H(q)} > \frac{\beta(q)\gamma_L(q)}{\beta(q)\gamma_L(q) + (\beta(q)-1)\gamma_H(q)},$$

which can be further rewritten as

$$\frac{\beta(q)(\beta(q)\gamma_L(q) + (1-\beta(q))\gamma_H(q))}{(\beta(q))^2\gamma_L(q) - (\beta(q)-1)^2\gamma_H(q)} > 1 \iff (1-\beta(q))\gamma_H(q) < 0.$$

Recall that  $\beta(q) > 1$  for all  $q \in \mathbb{R}_+$ , and hence the condition is satisfied whenever  $\gamma_H(q) > 0$ , which is always the case for  $q < q^1$ . We conclude that the denominator is positive everywhere on the full solution boundary.

Consider the numerator next. First notice that the second term inside the brackets is always positive but the first term can be negative.<sup>15</sup> The first term is scaled by  $x$ , while the second term is scaled by  $(1-x)$ . Therefore, if the numerator is positive at a

<sup>15</sup>This follows from  $\gamma'_\omega(q) < 0$ ,  $\beta'(q) < 0$ ,  $\beta(q) > 1$  and that  $\beta(q)\beta''(q) > 2(\beta'(q))^2$  together with that  $\gamma_L(q) < 0$  at the boundary.

belief above the boundary, it must be positive for the belief at the boundary as well. Since the the shortsighted belief,  $x^S(q)$ , is always above the fully optimal boundary, we can use it to show that the numerator is positive.

Plugging in  $x^S(q)$ , to the expression inside the brackets gives:

$$\frac{\beta(q) \gamma_L(q) \left( \beta'(q) (\beta(q) - 1) \gamma'_H(q) - \left( (\beta(q) - 1) \beta''(q) - 2 (\beta'(q))^2 \right) \gamma_H(q) \right)}{\beta(q) \gamma_L(q) + (1 - \beta(q)) \gamma_H(q)} + \frac{(1 - \beta(q)) \gamma_H(q) \left( \beta'(q) \beta(q) \gamma'_L(q) - \left( \beta(q) \beta''(q) - 2 (\beta'(q))^2 \right) \gamma_L(q) \right)}{\beta(q) \gamma_L(q) + (1 - \beta(q)) \gamma_H(q)}.$$

Since the denominator is negative ( $\gamma_L < 0$  at the boundary and  $\beta > 1$ ), this expression is proportional to

$$-\beta(q) \gamma_L(q) \left( \beta'(q) (\beta(q) - 1) \gamma'_H(q) - \left( (\beta(q) - 1) \beta''(q) - 2 (\beta'(q))^2 \right) \gamma_H(q) \right) + (\beta(q) - 1) \gamma_H(q) \left( \beta'(q) \beta(q) \gamma'_L(q) - \left( \beta(q) \beta''(q) - 2 (\beta'(q))^2 \right) \gamma_L(q) \right).$$

We simplify to get

$$[\gamma_H(q) \gamma'_L(q) - \gamma'_H(q) \gamma_L(q)] \beta'(q) \beta(q) (\beta(q) - 1) - 2 \gamma_H(q) \gamma_L(q) (\beta'(q))^2,$$

which is always positive because  $\gamma_H(q) > 0$  and  $\gamma_L(q), \gamma'_H(q), \gamma'_L(q) < 0$ . We have now shown that both the numerator and denominator are positive at the boundary and thus the lemma follows.  $\square$

## B Comparison of different solutions

### Properties of $V_{xx}$ and $V_{xxq}$

We need the derivatives of  $\Phi$  to prove the following lemma:

$$\begin{aligned} \Phi &= \left( \frac{x}{1-x} \right)^{\beta(q)} (1-x), \\ \Phi_q &= \Phi \beta'(q) \ln \left( \frac{x}{1-x} \right), \\ \Phi_x &= \Phi \frac{(\beta(q) - x)}{x(1-x)}, \\ \Phi_{qx} &= \Phi \beta'(q) x^{-1} (1-x)^{-1} \left[ 1 + (\beta(q) - x) \ln \left( \frac{x}{1-x} \right) \right], \\ \Phi_{xx} &= \Phi \beta(q) \frac{(\beta(q) - 1)}{x^2(1-x)^2}, \\ \Phi_{xxq} &= \Phi \frac{\beta'(q)}{x^2(1-x^2)} \left[ \beta(q) + (\beta(q) - 1) (1 + \beta(q) \ln \left( \frac{x}{1-x} \right)) \right]. \end{aligned}$$

**Lemma 7.** For all  $q \in [0, q^1]$  and  $x = x^*(q)$ ,  $V_{xx}(x, q) > 0$  and  $V_{xxq}(x, q) < 0$ .

*Proof.* Recall that the value function is

$$V(x, q) = \frac{x\pi_H(q) + (1-x)\pi_L(q)}{r} + B(q)\Phi(x, q),$$

The derivatives of the value function are then

$$\begin{aligned} V_{xx} &= B(q)\Phi_{xx}, \\ V_{xxq} &= B_q(q)\Phi_{xx} + B(q)\Phi_{xxq}. \end{aligned}$$

By plugging in the values of  $\Phi$  and its derivatives, the second derivative of the value function becomes

$$V_{xx} = B(q)\beta(q)\Phi \frac{(\beta(q) - x)}{x^2(1-x)^2},$$

which we can further rewrite as

$$V_{xx} = \frac{(\beta(q) - 1)\beta(q)[(\beta(q) - 1)x\gamma_H + \beta(q)(1-x)\gamma_L]}{r(1-x)^2x^2\beta'(q)}.$$

We know that  $B(q) > 0$  for all  $q < q^1$  because the decision maker would get the expected value of  $\frac{x\pi_H(q) + (1-x)\pi_L(q)}{r}$  by never expanding again. It is clear that the optimal solution does strictly better than that. Then  $\beta(q) > 1$  and  $\Phi > 0$  guarantee that  $V_{xx} > 0$ .

We can write  $V_{xxq}$  as

$$\begin{aligned} rV_{xxq} &= \left( \frac{-\Phi_{qx}(\gamma_H - \gamma_L)}{\Phi\Phi_{qx} - \Phi_q\Phi_x} x + \frac{-\Phi_{qx}\gamma_L + \Phi_q(\gamma_H - \gamma_L)}{\Phi\Phi_{qx} - \Phi_q\Phi_x} \right) \Phi_{xx} \\ &\quad + \left( \frac{\Phi_x(\gamma_H - \gamma_L)}{\Phi\Phi_{qx} - \Phi_q\Phi_x} x + \frac{\Phi_x\gamma_L - \Phi(\gamma_H - \gamma_L)}{\Phi\Phi_{qx} - \Phi_q\Phi_x} \right) \Phi_{xxq}. \end{aligned}$$

which we can further rewrite as

$$rV_{xxq} = \frac{(\Phi_x\Phi_{xxq}x - \Phi\Phi_{xxq} - \Phi_{qx}\Phi_{xx}x + \Phi_q\Phi_{xx})(\gamma_H - \gamma_L) + (\Phi_x\Phi_{xxq} - \Phi_{qx}\Phi_{xx})\gamma_L}{\Phi\Phi_{qx} - \Phi_q\Phi_x}.$$

We prove that  $V_{xxq} < 0$  by showing that the coefficient in the front of the negative term  $\gamma_L$  is negative and the coefficient in the front of the positive term  $(\gamma_H - \gamma_L)$  is positive and that the denominator is negative. We start with the former coefficient. After plugging in the expressions for  $\Phi$  and its derivatives, it becomes:

$$\Phi_x\Phi_{xxq} - \Phi_{qx}\Phi_{xx} = \frac{\Phi^2\beta'}{x^3(1-x)^3}((\beta - x)^2 + x(1-x)).$$

Recall that  $\beta' < 0$  and therefore the expression is negative.

Let us consider the former coefficient next:

$$\begin{aligned}
& \Phi_x \Phi_{xxq}x - \Phi \Phi_{xxq} - \Phi_{qx} \Phi_{xx}x + \Phi_q \Phi_{xx} \\
&= \Phi \frac{(\beta(q) - x)}{x(1-x)} \Phi \frac{\beta'(q)}{x^2(1-x^2)} \left[ \beta(q) + (\beta(q) - 1)(1 + \beta(q) \ln \left( \frac{x}{1-x} \right)) \right] x \\
&\quad - \Phi \Phi \frac{\beta'(q)}{x^2(1-x^2)} \left[ \beta(q) + (\beta(q) - 1)(1 + \beta(q) \ln \left( \frac{x}{1-x} \right)) \right] \\
&\quad - \Phi \beta'(q) x^{-1} (1-x)^{-1} \left[ 1 + (\beta(q) - x) \ln \left( \frac{x}{1-x} \right) \right] \Phi \beta(q) \frac{(\beta(q) - 1)}{x^2(1-x)^2} x \\
&\quad + \Phi \beta'(q) \ln \left( \frac{x}{1-x} \right) \Phi \beta(q) \frac{(\beta(q) - 1)}{x^2(1-x)^2} \\
&= -\Phi^2 \beta'(q) \beta(q) \frac{(\beta(q) - 1)}{x^2(1-x)^3} > 0.
\end{aligned}$$

The denominator is:

$$\begin{aligned}
& \Phi \Phi_{qx} - \Phi_q \Phi_x \\
&= \Phi^2 \beta'(q) x^{-1} (1-x)^{-1} \left[ 1 + (\beta(q) - x) \ln \left( \frac{x}{1-x} \right) \right] - \Phi^2 \beta'(q) \ln \left( \frac{x}{1-x} \right) \frac{(\beta(q) - x)}{x(1-x)} \\
&= \Phi^2 \beta'(q) x^{-1} (1-x)^{-1} < 0.
\end{aligned}$$

Combining the three parts leads to  $V_{xxq} < 0$  everywhere on the boundary.  $\square$

## Proof of Lemma 2

*Proof.* Let's start by proving the first part:  $x^S(q) > x^{stat}(q)$  for all  $q \in (0, q^1)$ .

When we drop the dependencies on  $q$ , the difference  $x^S - x^{stat}$  is

$$\frac{\beta \gamma_L}{\beta \gamma_L - (\beta - 1) \gamma_H} + \frac{\gamma_L}{\gamma_H - \gamma_L} = \frac{\gamma_L \gamma_H}{(\beta \gamma_L - (\beta - 1) \gamma_H)(\gamma_H - \gamma_L)} \geq 0.$$

Since the numerator and the denominator are both negative, this is strictly positive for all  $q \in (0, q^1)$  and equals 0 at  $q = 0$  and  $q = q^1$ .

Let's move on to the second part:  $x^S(q) > x^*(q)$  for all  $q \in [0, q^1)$ . First note that we have  $x^S(q) > x^*(q)$  at  $q = 0$ . From above we already have that  $x^S(0) = x^{stat}(0)$ . Then using the HJB equation (5) to evaluate the value matching condition (6) at  $q = 0$  we have

$$x^*(0) \gamma_H(0) + (1 - x^*(0)) \gamma_L(0) + \frac{1}{2} V_{xx}(x^*(0), 0) \frac{(x^*(0))^2 (1 - x^*(0))^2}{\sigma^2} = 0$$

Since we have  $V_{xx} > 0$  (Lemma 7) we must have that  $x^*(0) > x^S(0) = x^{stat}(0)$ . We then only need to show that  $x^S(q)$  and  $x^*(q)$  never cross.

Assume to the contrary that  $x^S(q) = x^*(q) = x$  for some  $q \in [0, q^1)$  (we already know that the policies are continuous and hence it is enough to show a contradiction

with this claim). The value matching and smooth pasting conditions hold at the boundary and we get

$$\begin{aligned} B^S \Phi - \frac{x\gamma_H + (1-x)\gamma_L}{r} &= 0, \\ B^S \Phi_x - \frac{\gamma_H - \gamma_L}{r} &= 0, \\ \frac{x\gamma_H + (1-x)\gamma_L}{r} + B_q \Phi + B\Phi_q &= 0, \\ \frac{\gamma_H - \gamma_L}{r} + B_q \Phi_x + B\Phi_{qx} &= 0, \end{aligned}$$

where we have again dropped all dependencies. This implies

$$\begin{aligned} -B^S \Phi + B_q \Phi + B\Phi_q &= 0, \\ -B^S \Phi_x + B_q \Phi_x + B\Phi_{qx} &= 0, \end{aligned}$$

and further that

$$\Phi \Phi_{qx} = \Phi_x \Phi_q.$$

When we plug in  $\Phi$  and its derivatives, the difference  $\Phi \Phi_{qx} - \Phi_x \Phi_q$  becomes

$$\beta'(q) \left( \frac{x}{1-x} \right)^{2\beta(q)-1} < 0.$$

This is strictly negative for all  $q < q^1$  as  $\beta'(q) < 0$  and approaches 0 as  $q \rightarrow q^1$  (and hence  $x \rightarrow 1$ ). This is a contradiction and thus the claim follows.  $\square$

### Proof of Lemma 3

The first part of Lemma 3 is pure house keeping since  $x(q^1) = 1$  is defined to be the complete information solution. For the second part about equal derivatives, let's first prove a stronger version:

**Lemma 8.** *For all  $q \in [0, q^1]$ ,*

$$g(x^S(q), q) - x^{S'}(q) = \frac{\beta'(q)\gamma_L(q)\gamma_H(q)}{(\beta(q)\gamma_L(q) + (1-\beta(q))\gamma_H(q))^2}.$$

*Proof.* The claim follows after straightforward calculations. Insert  $x^S(q)$  to  $x$  in (8):

$$\begin{aligned} g(x^S(q), q) &= \frac{-\beta(1-\beta)\gamma_L\gamma_H}{(\beta\gamma_L + (1-\beta)\gamma_H)^2} \left( \frac{\beta'\beta(1-\beta)(\gamma_L\gamma'_H - \gamma'_L\gamma_H)}{\beta\gamma_L + (1-\beta)\gamma_H} \right. \\ &\quad \left. + \frac{\beta\gamma_L\gamma_H(-2\beta'^2 + (\beta-1)\beta'')}{\beta\gamma_L + (1-\beta)\gamma_H} + \frac{(\beta-1)\gamma_L\gamma_H(-2\beta'^2 + \beta\beta'')}{\beta\gamma_L + (1-\beta)\gamma_H} \right), \end{aligned}$$

where we have dropped dependencies of  $q$ . After combining terms, this becomes

$$g(x^S(q), q) = \frac{\gamma_H (2\gamma_L \beta' - (\beta - 1)\beta\gamma'_L) + (\beta - 1)\beta\gamma_L\gamma'_H}{((\beta - 1)\gamma_H - \beta\gamma_L)^2}.$$

The derivative of the shortsighted policy  $x^S$  is

$$x^{S'}(q) = \frac{\gamma_H (\gamma_L \beta' - (\beta - 1)\beta\gamma'_L) + (\beta - 1)\beta\gamma_L\gamma'_H}{((\beta - 1)\gamma_H - \beta\gamma_L)^2}.$$

The claim follows by subtracting  $x^{S'}(q)$  from  $g(x^S(q), q)$ .  $\square$

The second part of Lemma 3 then follows by taking the limit  $\gamma_H(q) \rightarrow 0$  as  $q \rightarrow q^1$ .

### Proof of Proposition 3

The main text discusses the main parts of the proof. Here, we provide the proof with full detail. For that, we split the proposition into two and prove them separately.

**Lemma 9.** *There exists  $\underline{x} > 0$  such that the fully optimal quantity is larger than the static optimum for all beliefs below  $\underline{x}$ .*

*Proof.* We can use the value matching condition:

$$0 = x\gamma_H(q) + (1 - x)\gamma_L(q) + \frac{1}{2}V_{xx}(x, q)\frac{x^2(1 - x)^2}{\sigma^2} + \frac{1}{2}V_{xxq}(x, q)\frac{x^2(1 - x)^2}{\sigma^2}q.$$

When we evaluate it at  $(x(0), 0)$ , it equals

$$0 = x\gamma_H(0) + (1 - x)\gamma_L(0) + \frac{1}{2}V_{xx}(x, 0)\frac{x^2(1 - x)^2}{\sigma^2}.$$

$V_{xx}(x, q)$  is positive everywhere on the boundary (see Lemma 7). This then implies that the first term must be negative and hence the quantity must be above the static optimum.

The claim follows by observing that both the optimal and the static boundaries are continuous.  $\square$

**Lemma 10.** *There exists  $\bar{x} < 1$  such that the fully optimal quantity is smaller than the static optimum for all beliefs above  $\bar{x}$ .*

*Proof.* We proof this by showing that  $x^*(q^1) = x^{stat}(q^1)$  and  $x_q^*(q^1) < x_q^{stat}(q^1)$ . The first part is immediate.

For the second part, use Lemma 3 to get  $x^*(q^1) = x^S(q^1)$ .

$$\begin{aligned} x^S_q(q^1) - x^{stat}_q(q^1) &= \frac{(\beta - 1)\beta\gamma_L\gamma'_H}{(\beta\gamma_L)^2} - \frac{\gamma_L\gamma'_H}{(\gamma_L)^2} \\ &= -\frac{\gamma_L\gamma'_H}{\beta\gamma_L^2} < 0. \end{aligned}$$

Hence, the optimal and static boundaries meet at  $q^1$  but the static boundary reaches the point from below the optimal boundary.

□

## C Durable goods application

### Proof of Lemma 5

*Proof.* The monopolist's payoff is

$$\begin{aligned}\Pi(\tau) &= \int_{\underline{\theta}}^{\bar{\theta}} \mathbb{E} \left[ e^{-r\tau(\theta)} (\theta X_{\tau}(\theta) - c) - W(\theta) \right] f(\theta) d\theta \\ &= \int_{\underline{\theta}}^{\bar{\theta}} \left( \mathbb{E} \left[ e^{-r\tau(\theta)} (X_{\tau(\theta)}\theta - c) - \int_{\underline{\theta}}^{\theta} \mathbb{E} \left[ e^{-r\tau(s)} X_{\tau(s)} \right] ds \right) f(\theta) d\theta.\end{aligned}$$

The requirements for Fubini's theorem are satisfied and we can change the order of integration in the second term:

$$\begin{aligned}& \int_{\underline{\theta}}^{\bar{\theta}} \mathbb{E} \left[ \int_{\underline{\theta}}^{\theta} \mathbb{E} \left[ e^{-r\tau(s)} X_{\tau(s)} \right] ds \right] f(\theta) d\theta \\ &= \mathbb{E} \left[ \int_{\underline{\theta}}^{\bar{\theta}} \int_{\underline{\theta}}^{\theta} \mathbb{E} \left[ e^{-r\tau(s)} X_{\tau(s)} \right] f(\theta) ds d\theta \right] \\ &= \mathbb{E} \left[ \int_{\underline{\theta}}^{\bar{\theta}} \int_s^{\bar{\theta}} \mathbb{E} \left[ e^{-r\tau(s)} X_{\tau(s)} \right] f(\theta) d\theta ds \right] \\ &= \mathbb{E} \left[ \int_{\underline{\theta}}^{\bar{\theta}} \mathbb{E} \left[ e^{-r\tau(s)} X_{\tau(s)} \right] (1 - F(s)) ds \right] \\ &= \mathbb{E} \left[ \int_{\underline{\theta}}^{\bar{\theta}} e^{-r\tau(s)} X_{\tau(s)} (1 - F(s)) ds \right].\end{aligned}$$

After plugging this into the monopolist's payoff function and rearranging, we obtain

$$\Pi(\tau) = \mathbb{E} \left[ \int_{\underline{\theta}}^{\bar{\theta}} e^{-r\tau(\theta)} \left( X_{\tau(\theta)}\theta - X_{\tau(\theta)} \frac{1 - F(\theta)}{f(\theta)} - c \right) f(\theta) d\theta \right].$$

□

### Calculating prices

We can invert the monopolist's (or planner's) policy  $x^*(q)$  from quantities to prices. The policy is implementable with prices that keep the current marginal buyer of type  $\theta$  indifferent between buying and waiting. Given quantity  $q$ , the marginal buyer is determined by  $q = 1 - F(\theta)$  or  $\theta = F^{-1}(1 - q)$ . Then note that the policy can be expressed in terms of types as a function  $\hat{\theta}(x)$ , which tells the marginal type which



buys at belief  $x$ . Let the inverse of this (monotone) function be  $\hat{x}(\theta) = x^*(q(\theta))$ . We can then write the price as

$$P(x) = x\hat{\theta}(x) - W(\hat{\theta}(x)),$$

where

$$W(\theta) = \int_{\underline{\theta}}^{\theta} \mathbb{E} \left[ e^{-r\tau(s)} \hat{x}(s) \right] ds.$$

We next derive an expression for  $W(\theta)$ . For types that never buy,  $\theta < F^{-1}(1 - q^1)$ , we have  $\tau(\theta) = \infty$ . Thus assume  $F^{-1}(1 - q^1) \leq \theta \leq \bar{\theta}$ . Then start by defining

$$\psi(x, \theta, \theta') := \mathbb{E} \left( e^{-r\tau(\theta')} | x_0 = x, \theta_0 = \theta \right).$$

First fix  $\theta'$ . Then by the same arguments as for the HJB-equation we get the following partial differential equation

$$r\psi(x, \theta, \theta') = \frac{1}{2}q(\theta) \frac{x^2(1-x)^2}{\sigma^2} \psi_{xx}(x, \theta, \theta')$$

It is easy to show that the following function is a solution to this PDE:

$$\psi(x, \theta, \theta') = A(\theta, \theta') \Phi(x, q(\theta)),$$

where  $A(\theta)$  is an arbitrary function to be determined from the boundary conditions. At  $\theta' = \theta$  (the type is the marginal buyer) we must have that discounting equals one

$$\psi(\hat{x}(\theta'), \theta', \theta') = 1.$$

Or

$$A(\theta') = \frac{1}{\Phi(\hat{x}(\theta'), q(\theta'))}.$$

We also must have that

$$\frac{\partial \psi(x, \theta, \theta')}{\partial \theta} = 0 \text{ for } x = \hat{x}(\theta).$$

This condition follows because the quantity adjusts immediately along the boundary and so does the current marginal type  $\theta$ . Therefore, the change in the current marginal type cannot affect expected stopping time for type  $\theta'$ .

Plugging in  $\psi(x, \theta, \theta')$  then gives

$$A'(\theta) \Phi(\hat{x}(\theta), q(\theta)) + A(\theta) \Phi_q(\hat{x}(\theta)) \frac{dq(\theta)}{d\theta} = 0.$$

Noting that  $\Phi_q(\hat{x}(\theta)) = \beta'(q) \ln(x/(1-x)) \Phi(x, q)$  and  $dq(\theta)/d\theta = -f(\theta)$  we get a differential equation for  $A(\theta)$ :

$$A'(\theta) = A(\theta) \beta'(q(\theta)) \ln \left( \frac{\hat{x}(\theta)}{1 - \hat{x}(\theta)} \right).$$

This is a linear ordinary differential equation which can be solved with the help of an integrating factor. Using  $A(\theta') = \Phi(x, q(\theta'))^{-1}$  as an initial condition, the solution is

$$A(\theta) = \frac{1}{\Phi(\hat{x}(\theta'), q(\theta'))} \exp \int_{\theta'}^{\theta} \beta'(q(h)) \ln \left( \frac{\hat{x}(h)}{1 - \hat{x}(h)} \right) dh.$$

This gives then

$$\mathbb{E}[e^{-r\tau(\theta')} | x_0 = \hat{x}(\theta), \theta_0 = \theta] = \frac{\Phi(\hat{x}(\theta), q(\theta'))}{\Phi(\hat{x}(\theta'), q(\theta'))} \exp \int_{\theta'}^{\theta} \beta'(q(h)) \ln \left( \frac{\hat{x}(h)}{1 - \hat{x}(h)} \right) dh.$$

The value for a buyer of type  $\theta$  is then

$$\begin{aligned} W(\theta) &= \int_{\underline{\theta}}^{\theta} \mathbb{E}[e^{-r\tau(s)} \hat{x}(s)] ds \\ &= \int_{\underline{\theta}^1}^{\theta} \hat{x}(s) \frac{\Phi(\hat{x}(\theta), q(\theta'))}{\Phi(\hat{x}(s), q(s))} \exp \int_s^{\theta} \beta'(q(h)) \ln \left( \frac{\hat{x}(h)}{1 - \hat{x}(h)} \right) dh, \end{aligned}$$

where  $\theta^1 = F^{-1}(1 - q^1)$ . This then allows us to calculate the price from  $P(x) = x\hat{\theta} - W(\hat{\theta})$  for  $x \geq \hat{x}(\bar{\theta})$ .

### Calculating the environmental tax

The tax in the environmental application follows from an otherwise identical calculation as above but is of course affected by the belief only indirectly through the hitting time  $\tau(\theta)$ . That is, we can calculate it from  $P(x) = \hat{\theta}(x) - W(\hat{\theta}(x))$ , where  $W(\theta) = \int_{\underline{\theta}}^{\theta} \mathbb{E}[e^{-r\tau(s)}] ds$ .

## References

- Baldursson, Fridrik M. and Ioannis Karatzas**, “Irreversible investment and industry equilibrium,” *Finance and Stochastics*, Dec 1996, *1* (1), 69–89.
- Bergemann, Dirk and Juuso Välimäki**, “Experimentation in markets,” *Review of Economic Studies*, 2000, *67* (2), 213–234.
- and –, “Entry and Vertical Differentiation,” *Journal of Economic Theory*, 2002, *106* (1), 91–125.
- Board, Simon**, “Selling options,” *Journal of Economic Theory*, 2007, *136* (1), 324 – 340.
- , “Durable-Goods Monopoly with Varying Demand,” *The Review of Economic Studies*, 2008, *75* (2), 391–413.
- Bolton, Patrick and Christopher Harris**, “Strategic Experimentation,” *Econometrica*, 1999, *67* (2), 349–374.
- Bonatti, Alessandro**, “Menu pricing and learning,” *American Economic Journal: Microeconomics*, 2011, *3* (3), 124–163.
- Caballero, Ricardo J and Robert S Pindyck**, “Uncertainty, Investment, and Industry Evolution,” *International Economic Review*, August 1996, *37* (3), 641–662.
- Chamley, Christophe and Douglas Gale**, “Information Revelation and Strategic Delay in a Model of Investment,” *Econometrica*, 1994, *62* (5), 1065–1085.
- Che, Yeon-Koo and Johannes Hörner**, “Recommender Systems as Mechanisms for Social Learning\*,” *The Quarterly Journal of Economics*, 2017, p. qjx044.
- Daley, Brendan and Brett Green**, “Waiting for News in the Market for Lemons,” *Econometrica*, 2012, *80* (4), 1433–1504.
- Dixit, Avinash**, “Entry and Exit Decisions under Uncertainty,” *Journal of Political Economy*, 1989, *97* (3), 620–638.
- Dixit, Avinash K. and Robert S. Pindyck**, *Investment under Uncertainty*, Princeton: Princeton University Press, 1994.
- Fleming, Wendell H and Halil Mete Soner**, *Controlled Markov processes and viscosity solutions*, Vol. 25, Springer Science & Business Media, 2006.
- Frick, Mira and Yuhta Ishii**, “Innovation Adoption by Forward-Looking Social Learners,” *Working Paper*, 2016.
- Fuchs, William and Andrzej Skrzypacz**, “Bargaining with arrival of new traders,” *American Economic Review*, 2010, *100* (3), 802–36.

- Gittins, J. and D. Jones**, “A Dynamic Allocation Index for the Sequential Allocation of Experiments,” *Progress in Statistics*, ed. by J. Gani, 1974, pp. 241–266.
- Golosov, Mikhail, John Hassler, Per Krusell, and Aleh Tsyvinski**, “Optimal taxes on fossil fuel in general equilibrium,” *Econometrica*, 2014, 82 (1), 41–88.
- Grossman, Sanford J., Richard E. Kihlstrom, and Leonard J. Mirman**, “A Bayesian Approach to the Production of Information and Learning by Doing,” *The Review of Economic Studies*, 1977, 44 (3), 533–547.
- Keller, Godfrey and Sven Rady**, “Strategic experimentation with Poisson bandits,” *Theoretical Economics*, 2010, 5 (2), 275–311.
- , – , and **Martin Cripps**, “Strategic Experimentation with Exponential Bandits,” *Econometrica*, 2005, 73 (1), 39–68.
- Kogan, Leonid**, “An equilibrium model of irreversible investment,” *Journal of Financial Economics*, November 2001, 62 (2), 201–245.
- Laiho, Tuomas and Julia Salmi**, “Social Learning and Monopoly Pricing with Forward Looking Buyers,” *Working paper*, 2018.
- Leahy, John V.**, “Investment in competitive equilibrium: The optimality of myopic behavior,” *The Quarterly Journal of Economics*, 1993, 108 (4), 1105–1133.
- Lucas, Robert E. and Edward C. Prescott**, “Investment Under Uncertainty,” *Econometrica*, 1971, 39 (5), 659–681.
- Milgrom, Paul and Ilya Segal**, “Envelope theorems for arbitrary choice sets,” *Econometrica*, 2002, 70 (2), 583–601.
- Murto, Pauli and Juuso Välimäki**, “Learning and Information Aggregation in an Exit Game,” *The Review of Economic Studies*, 2011.
- Myerson, Roger**, “Optimal auction design,” *Mathematics of operations research*, 1981, 6 (1), 58–73.
- Pindyck, Robert S.**, “Irreversibility, Uncertainty, and Investment,” *Journal of Economic Literature*, 1991, 29 (3), 1110–1148.
- Prescott, Edward C.**, “The Multi-Period Control Problem Under Uncertainty,” *Econometrica*, 1972, 40 (6), 1043–1058.
- Rob, Rafael**, “Learning and Capacity Expansion under Demand Uncertainty,” *The Review of Economic Studies*, 1991, 58 (4), 655–675.
- Rothschild, Michael**, “A two-armed bandit theory of market pricing,” *Journal of Economic Theory*, 1974, 9 (2), 185–202.

**Traeger, Christian P.**, “ACE – Analytic Climate Economy (with Temperature and Uncertainty),” *Available at SSRN: <https://ssrn.com/abstract=3307622> or <http://dx.doi.org/10.2139/ssrn.3307622>*, 2019.