Gradual Learning from Incremental Actions

Tuomas Laiho, Pauli Murto, Julia Salmi*

January 11, 2022

Abstract

We introduce a collective experimentation problem where small agents choose the timing of an irreversible action under uncertainty and when public feedback from their actions arrives gradually over time. We solve the decentralized equilibrium where agents maximize their own payoffs and the socially optimal policy, which internalizes the social value of information. The latter entails an informational tradeoff where acting today speeds up learning but postponing capitalizes on the option value of waiting. Extending our analysis to include payoff externalities, we show that a better learning technology can reduce learning and total welfare in equilibrium.

JEL classification: C61, C73, D82, D83

Keywords: social learning, experimentation, optimal stopping, Brownian motion, dynamic markets, innovation adoption

*Laiho: Ministry of Finance, Finland (tuomas.s.laiho@gmail.com), Murto: Aalto University School of Business (pauli.murto@aalto.fi), Salmi: University of Copenhagen (julia.salmi@econ.ku.dk). We thank Aislinn Bohren, Martin Cripps, Rahul Deb, Francesc Dilme, Nils Christian Framstad, Dino Gerardi, Yingni Guo, Daniel Hauser, Jan Knoepfle, Matti Liski, Erik Madsen, Helene Mass, Konrad Mierendorff, Francesco Nava, Mariann Ollar, Marco Ottaviani, Sven Rady, Peter Norman Sørensen, Roland Strausz, Juuso Toikka, Christian Traeger, Juuso Välimäki, and various seminar and conference audiences for their helpful comments. This project has received financial support from ERC grant no. 683031 through PI Bård Harstad (Laiho), Academy of Finland (Murto), and from Emil Aaltonen Foundation, OP Group Research Foundation, Yrjö Jahnsson Foundation, and ERC grant no. 682417 through PI Vasiliki Skreta (Salmi).

1 Introduction

Many actions have long-run consequences that can be observed only gradually over time. A firm considering expanding its production capacity has to address uncertainty about future demand; a consumer buying a durable good today must evaluate her needs for it in the future; policy makers have to contemplate irreversible impacts of policies that may be hard to predict. If such uncertainties resolve exogenously over time, the decision maker's response is to postpone her actions in order to learn more. But when the actions themselves affect the learning process, there is a tradeoff between waiting for more information and acting now to speed up learning for later decisions.

This paper introduces a novel learning problem where an action taken today has a long-run impact on the flow of information. A continuum of small agents chooses when to stop – for example, when to adopt an innovation or enter a new market. An unknown binary state determines if stopping is profitable for the agents. Crucially, learning is *gradual*: upon stopping, each agent initiates a persistent flow of information over time. This is in contrast to the standard experimentation models where an action generates an *instantaneous* one-time signal.

Our main question is how the path of stopping decisions – the expansion path – is determined on one hand when agents optimize individually and on the other hand when a planner coordinates the actions. The paper's contribution is twofold. First, we derive new economic insights that are caused by gradual learning. Second, we develop a novel methodological approach with suitable solution techniques. The tractability of gradual learning enables extending the analysis to study the interaction between informational and payoff externalities. We show a surprising result that a better learning technology tends to decrease the agents' welfare and lead to slower learning in the equilibrium if there are positive payoff externalities.

Gradual learning creates a new tradeoff for the socially optimal expansions: *the information generation effect* calls for aggressive expansion in order to improve information for future decisions and *the option value effect* calls for cautious expansion in order to have better information for the current decisions. A social planner balances these two effects, but individual agents internalize only the latter effect and thus the decentralized equilibrium suffers from informational free-riding.

We approach experimentation under gradual learning by modeling the path of individual actions as a stock process, which controls the speed of learning. Each agent who has stopped produces a flow of i.i.d. signals conditional on the true state. In continuous time, the aggregate signal then follows a Brownian motion with an unknown drift, determined by the true state, and a signal-to-noise ratio proportional to the stock of agents who have stopped. Each stopping decision thus affects information generation gradually over time.

The techniques to solve the decentralized equilibrium and the socially optimal policy turn out to be quite different. The common challenge is that the problems are two-dimensional as both the belief that the state is high and the stock affect the future. Furthermore, the stock and the belief processes are interlinked as the stock determines the flow of new information. We show that the decentralized equilibrium can be solved by analyzing "shortsighted" agents who optimize their stopping decisions against the assumption that no agent stops in the future. The shortsighted approach works because information arrives smoothly over time under gradual learning.

Unlike the decentralized equilibrium, the socially optimal policy takes into account the social value of faster learning. The equivalence with shortsighted optimization breaks because the value of information depends on the expected future actions. Because of the information generation effect, socially optimal policy favors earlier and more aggressive expansions than what happens in the decentralized equilibrium. The difference between the two is especially pronounced when the learning technology is good and learning could potentially be fast. Compared to the no-learning benchmark, gradual learning tends to increase the socially optimal stock for low beliefs and to decrease it for high beliefs due to the informational tradeoff between information generation and the option value of waiting.

Thanks to the technical tractability, the model with gradual learning has a potential to work as a workhorse model that can be further extended to analyze other phenomena. We provide one important extension in this paper by including the possibility that agents' payoffs depend on the other agents' stopping behavior. This way we can analyze the joint effect of informational and payoff externalities and show that agents can be strictly better off under a worse learning technology if there are positive payoff externalities.

In the last part of the paper, we demonstrate how our modeling approach can be used to analyze different applications. We first discuss how we can incorporate mechanism design techniques to implement the social optimum, or any other policy, as a decentralized equilibrium. We solve the pricing problem of a monopolist selling durable goods and we analyze a model of capital investment in a competitive industry. A common takeaway from these applications is the potential benefit of market power to consumers: a large player can internalize the informational externality that is harmful in competitive equilibrium. We also discuss an extension where agents are heterogeneous in how much information they generate. A thorough analysis of these topics is provided as supplementary material.

1.1 Related literature

Using the framework of our paper, the previous literature on learning can be organized based on whether the information generation effect or the option value effect is present in the model. The current paper is the first to analyze the dual effect of endogenous learning.

The information generation effect is present in papers analyzing classic single agent bandit problems and experimental consumption (Gittins and Jones 1974, Rothschild 1974, Prescott 1972 and Grossman, Kihlstrom and Mirman 1977). Introducing multiple agents to these models adds an informational externality that dampens the information generation effect. Bolton and Harris (1999), Keller, Rady and Cripps (2005) and Keller and Rady (2010) analyze such models under different assumptions on the learning technology. Applications include Bergemann and Välimäki (1997, 2000) and Bonatti (2011) who analyze dynamic pricing. No option value effect exists in these papers because actions are reversible and hence learning always increases the level of optimal quantities relative to the no-learning benchmark.

When actions are irreversible but information arrives exogenously rather than endogenously, only the option value effect is present. Seminal papers in this literature include McDonald and Siegel (1986), Pindyck (1988), and Dixit (1989) and the ensuing literature on real options is summarized in Dixit and Pindyck (1994).

A few papers investigate social learning with irreversible actions, which bears similarities with informational free-riding in our decentralized solution. Frick and Ishii (2020) analyze the adoption of new technologies using a Poisson process with instantaneous feedback to model learning. The adoption rate of innovations is lower than without learning in their model because of the option value effect. 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 causing socially inefficient delays (Chamley and Gale 1994, Murto and Välimäki 2011).

Introducing a large player can overturn the effect of social learning and irreversibility on optimal quantities because a large player internalizes the information generation effect. Che and Hörner (2017) study how a social planner, who designs a recommendation system for consumers, can mitigate informational free-riding. Laiho and Salmi (2021) analyze monopoly pricing in a similar setup. Both in Che and Hörner (2017) and in Laiho and Salmi (2021), the presence of a social planner or a monopolist induces learning to increase quantities. The crucial difference from 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. More generally, there is no informational tradeoff under instantaneous learning.

Our assumption that learning is gradual implies that past actions matter for the current information flow. Two contemporaneous papers share this feature with us, although their models and key tradeoffs are otherwise different from ours. Liski and Salanié (2020) analyze a single-agent problem where a decision-maker controls the accumulation of a stock that triggers a one-time catastrophe at an unknown threshold level. The novel feature in their model is a random delay between the crossing of the threshold and the onset of the catastrophe. Martimort and Guillouet (2020) analyze a model with similar features focusing on a time-inconsistency problem under their assumptions. In these papers learning is about an unknown tipping point, whereas in our paper it is about a fixed unknown state.

2 Model

2.1 Actions and payoffs

A unit mass of small agents choose when, if ever, to take an irreversible action (to stop). We index individual agents by their type θ and assume that θ is distributed according to a continuously differentiable distribution function F with a full support on $\Theta := [\underline{\theta}, \overline{\theta}]$. Time t is continuous and goes to infinity.

An agent's stopping payoff, $v_\omega(\theta)$, depends on the state of the world $\omega \in \{H, L\}$ such that the payoff is higher in the high state of the world for all types: $v_H(\theta) \geq 0 > v_L(\theta)$.[1] Payoffs are continuously differentiable with bounded derivatives and increasing in type: for each $\theta \in \Theta$, $v'_\omega(\theta) \geq 0$ for both $\omega \in \{H, L\}$ and $v'_\omega(\theta) > 0$ for at least one $\omega = H$ or $\omega = L$. The realized payoff for an agent of type θ, who stops at time t, is $e^{-rt}v_\omega(\theta)$ where r is the common discount rate. An agent's outside option is zero and we normalize $v_H(\underline{\theta}) = 0$ so that $\underline{\theta}$ is the lowest type who would ever want to stop. The model is equivalent to a setting where agents receive a flow of state dependent payoffs $\pi_\omega(\theta) = rv_\omega(\theta)$ at every instant after stopping.

Agents are risk-neutral and maximize their expected discounted stopping payoffs. The agents do not know the state of the world ω but learn about it over time as we will describe next.

[1]The analysis easily extends to the case where $v_L(\theta) > 0$ for some types. The only change is that all types, who get a positive stopping payoff in both states of the world, stop immediately.

2.2 Learning

The key idea of *gradual* learning is that every agent who has stopped generates a flow of conditionally independent public signals. Therefore, we consider endogenous learning from the *stock* of stopped agents: let q_t denote the stock (measure) of agents who have stopped by time t.

Specifically, the public learns about the state by observing a Brownian diffusion

$$dy_t = q_t \mu_\omega dt + \sigma \sqrt{q_t} dw_t, \tag{1}$$

where we normalize $\mu_H = 1/2$ and $\mu_L = -1/2$, $\sigma > 0$ is the standard deviation of the process, and w_t is a standard Wiener process. Signal process (1) is the limit of a model where q_t is composed of discrete units that produce conditionally independent noisy signals over time and where the total informativeness is normalized to stay constant. The signals can be for example interpreted as realized individual payoffs (see Appendix A).[2]

We denote by x_t the public posterior belief $x_t = Pr(\omega = H|\mathcal{F}_t)$, where $\mathcal{F}_t$ is the natural filtration generated by the signal process (1). The unconditional law of motion for the public belief follows from Bayes' rule:

$$dx_t = \frac{\sqrt{q_t}}{\sigma} x_t(1 - x_t) d\tilde{w}_t, \tag{2}$$

where $\tilde{w}_t$ is a standard Wiener process. In equation (2), the term $\frac{\sqrt{q_t}}{\sigma}$ is the signal-to-noise ratio of the process (1) and determines how fast the belief converges to the truth. Hence, the higher the stock of stopped agents, the more informative the public signals. In Appendix F, we extend the model to allow for a more general relationship between the stock q_t and the signal-to-noise ratio.

2.3 Solution concepts

We use the term *policy* for a description of how the stock q_t evolves over time. A policy $Q = \{q_t\}_{t \geq 0}$ is an increasing stochastic process adapted to $\mathcal{F}_t$. Notice that

[2] See Bergemann and Välimäki (1997, 2000), Bolton and Harris (1999), Moscarini and Smith (2001), and Bonatti (2011) for other applications and further discussion. The difference to these papers is that they do not consider learning from the stock but from the flow of new actions.

the signal process itself depends on the evolution of q_t, so that in effect we are defining policy Q jointly with signal process Y.

Individual agents take the policy Q as given when they choose their stopping strategies. A strategy for an agent of type θ is a stopping time $\tau(\theta)$ adapted to $\mathcal{F}_t$. The payoff to type θ adopting $\tau(\theta)$ under Q is

$$\mathbb{E}\left[e^{-r\tau(\theta)}v_\omega(\theta)\middle|Q\right], \tag{3}$$

where the vertical line notation means that the expectation is for some fixed process Q.

We say that a stopping profile $\mathcal{T} = \{\tau(\theta)\}_{\theta\in\Theta}$ is *consistent with* Q if

$$Pr\left[\int_{\underline{\theta}}^{\overline{\theta}} \mathbf{1}(\tau(\theta) \le t)dF(\theta) = q_t\middle|Q\right] = 1$$

for all t. In other words, $\mathcal{T}$ is consistent with Q if the measure of agents that it commands to stop always matches the policy.

It is convenient to define solution concepts directly in terms of a policy rather than in terms of a stopping profile. We consider two solution concepts:

Definition 1. *A policy Q^E is a decentralized equilibrium if there exists a profile $\mathcal{T}^E$ such that i) it is consistent with Q^E and ii) $\tau^E(\theta)$ maximizes* (3) *for each θ when $Q = Q^E$.*

Definition 2. *A policy Q^* is socially optimal if there exists a profile $\mathcal{T}^*$ such that i) it is consistent with Q^* and ii)*

$$\mathbb{E}\left[\int_{\underline{\theta}}^{\overline{\theta}} e^{-r\tau^*(\theta)}v_\omega(\theta)dF(\theta)\middle|Q^*\right] \ge \mathbb{E}\left[\int_{\underline{\theta}}^{\overline{\theta}} e^{-r\tau(\theta)}v_\omega(\theta)dF(\theta)\middle|Q\right],$$

for any policy Q and profile $\mathcal{T} = \{\tau(\theta)\}_{\theta\in\Theta}$ consistent with Q.

3 Analysis

Our objective is to analyze how gradual learning affects stopping decisions. First, we discuss some common properties that hold regardless of whether stopping times are individually or socially optimal and present the no-learning benchmark. Then, we solve both the (unique) decentralized equilibrium and the socially optimal policy. Lastly, we compare the decentralized equilibrium and the socially optimal solution to the no-learning benchmark and provide comparative statics results on the effects of learning.

3.1 Higher types stop first

In principle, one can implement a policy Q by many different stopping profiles. However, because the stopping payoffs are increasing in θ, higher type agents want to stop whenever a lower type agent wants to stop, which leads to monotone stopping profiles:

Lemma 1. *If $\mathcal{T} = \{\tau(\theta)\}_{\theta\in\Theta}$ maximizes* (3) *for each θ for given process Q, then*

$$\Pr\left[\tau(\theta) \leq \tau(\theta') \middle| \mathcal{F}_t; Q\right] = 1$$

whenever $\theta > \theta'$.

Monotone stopping times are also socially optimal:

Lemma 2. *Any stopping profile $\mathcal{T} = \{\tau(\theta)\}_{\theta\in\Theta}$ consistent with Q satisfies:*

$$\mathbb{E}\left[\int_{\underline{\theta}}^{\overline{\theta}} e^{-r\tau(\theta)} v_\omega(\theta)\, dF(\theta) \middle| \mathcal{F}_t; Q\right] \leq \mathbb{E}\left[\int_{\underline{\theta}}^{\overline{\theta}} e^{-r\tau^{mon}(\theta)} v_\omega(\theta)\, dF(\theta) \middle| \mathcal{F}_t; Q\right],$$

where $\tau^{mon}(\theta) := \inf\{t : q_t \geq 1 - F(\theta)\}$.

We prove both Lemma 1 and Lemma 2 in Appendix A.

As it is without loss of generality to restrict attention to monotone stopping profiles, there is a one-to-one mapping between the stock q_t and the largest type θ_t who has not stopped: $q_t = 1 - F(\theta_t)$. It is useful to let the stock be a function of the current highest type, $q(\theta) := 1 - F(\theta)$, which has an inverse (current highest type): $\theta(q) := \{\theta : 1 - F(\theta) = q\}$. With a slight notational abuse, we use $v_\omega(q)$ to denote the stopping payoff of type $\theta(q)$.

3.2 Boundary policies

This subsection discusses the dynamics in our model. It turns out that both solutions can be characterized as *boundary policies*:

Definition 3. *A policy Q is a boundary policy if there exists a continuous function $\tilde{q} : [0,1] \to [0,1]$ such that $q_t = \tilde{q}(\max_{s\in[0,t]} x_s)$ where $\tilde{q}$ is strictly increasing for all x such that $\tilde{q}(x) > 0$.*

A boundary policy is Markovian: agents' stopping decisions depend only on the stock and the belief. Because stopping is irreversible, the stock at time t is determined by the highest belief reached up to t. A boundary policy hence divides the stock-belief state space into two regions: in *the expansion region*, more agents stop until the stock equals $\tilde{q}(x)$ and in *the waiting region*, everyone waits.

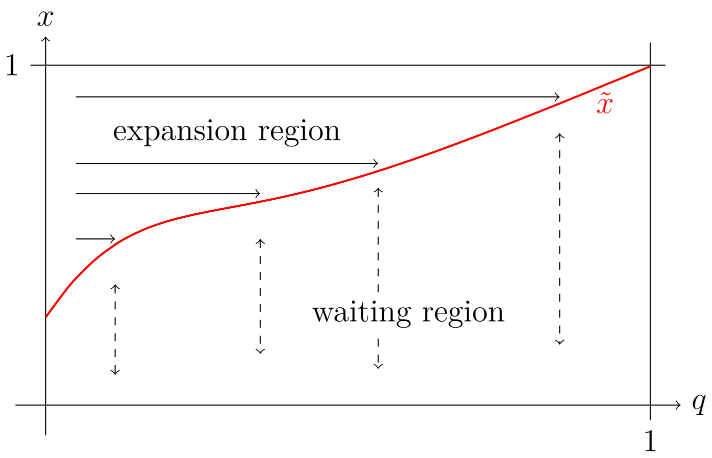

Figure 1: Dynamics in the waiting and expansion regions of the state space.

A boundary policy is fully characterized by the inverse of $\tilde{q}$, a *policy function* $\tilde{x} : [0,1] \to [0,1]$, which maps the stock to the cutoff belief. It turns out that it is easier to use policy functions to characterize our solutions than functions $\tilde{q}$. Figure 1 illustrates a boundary policy and the implied dynamics in the state space. Above the boundary, the stock increases (horizontal movement in the figure) and below it, the stock stays constant and only the belief moves (vertical movement). As soon as the belief hits the boundary from below, the quantity is pushed towards right along the boundary. The expansions in the stock are immediate (depicted by solid arrows in the figure), whereas the belief fluctuates according to the diffusion process (2) (dashed arrows). Apart from the possible initial jump, the stock process stays below the boundary and is continuous almost surely.

It is useful to note that since a boundary policy is Markovian in the stock-belief state space, we can express an individual agent's best-response to such a policy as an optimally chosen stopping region in the state space. We utilize this in establishing the existence and uniqueness of a decentralized equilibrium.

3.3 No-learning benchmark

We start our analysis with the benchmark case without learning, which allows us to disentangle how learning affects the decentralized equilibrium and the socially optimal solution.

When there is no learning but the common belief stays constant, the agents' stopping problem is static. An agent stops if and only if his type is so high that the expected payoff is positive: $xv_H(\theta) + (1-x)v_L(\theta) \geq 0$. Hence, the no-learning policy is characterized by the following cutoff:

$$x^{stat}(q) = \frac{-v_L(q)}{v_H(q) - v_L(q)},$$

where $v_\omega(q) \coloneqq v_\omega(\theta(q))$.

Individual optimization and socially optimal policies coincide when there is no learning.

3.4 Decentralized equilibrium

We next characterize the decentralized equilibrium defined in Definition 1. An optimal stopping time for an individual agent trades off the cost of waiting with the option value of waiting. Because the belief process changes endogenously as the stock of stopped agents increases, waiting not only brings more information but also faster learning. Despite this, we show that we can solve *equilibrium* stopping times by first solving a sequence of stopping problems where each agent finds the optimal time to stop when the stock is fixed. That is, we fix $q_t = \hat{q}$ for all t and find the optimal stopping time for type $\theta(\hat{q})$ (pinned down by Lemma 1), assuming that q_t is constant and equal to $\hat{q}$. This problem is a one-dimensional stopping problem and can be solved using standard techniques in the literature (Dixit and Pindyck (1994), see Appendix B.1 for details). We show that the equilibrium in the original problem corresponds to this "shortsighted" problem in which agents do not take future stopping decisions by other agents into account.[3]

The intuition for the equivalence between the shortsighted problem and the original problem is the following: because later expansions in the stock happen only when it is optimal for lower type agents to stop, all higher type agents strictly prefer stopping always when the stock expands (Lemma 1). Hence, higher type agents want to stop even before the expansion, which means that future expansions do not change the optimal stopping times. The equivalence with shortsighted optimization is hence an equilibrium property and may be violated against other (non-equilibrium) stock processes. In Appendix B.1, we formalize this argument to get the following result:

Proposition 1. *There is a unique decentralized equilibrium, which is characterized by an increasing policy function x^E:*

$$x^E(q) := \frac{-\beta(q)v_L(q)}{(\beta(q)-1)\, v_H(q) - \beta(q)v_L(q)},$$

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

According to Proposition 1, an agent of type θ waits until the belief reaches the cutoff $x^E(q(\theta))$. The decentralized equilibrium is thus a boundary policy: the policy function x^E defines a boundary so that whenever the belief is about to cross the boundary, more agents stop.

Notice that the cutoff $x^E(q)$ is increasing in the signal precision (decreasing in σ), which means that a better learning technology decreases the stock of agents who are willing to stop at any given belief. The no-learning benchmark is a special case of the decentralized equilibrium as we take $\sigma \to \infty$, which directly gives:

[3]Our method to solve the decentralized equilibrium is inspired by an industry investments paper by Leahy (1993) who shows that under exogenous uncertainty the competitive equilibrium behavior coincides with that of 'myopic' investors who ignore the effect the future investments have on the price.

Corollary 1. *For all beliefs in* $(x^{stat}(0), 1)$*, the stock of stopped agents is strictly smaller in the decentralized equilibrium than in the no-learning benchmark.*

The intuition behind Corollary 1 is that agents want to free-ride on the information provided by other agents. Proposition 1 implies that in the decentralized equilibrium only *past* stopping decisions affect individual agents' behavior because from an individual agent's perspective past actions determine the speed of learning. In the next section, we analyze the socially optimal policy which takes into account the informational externality between agents. The solution then takes into account how both *past and future* stopping decisions affect learning.

3.5 Social optimum

Next, we consider the problem in Definition 2 where a benevolent social planner seeks to maximize agents' expected joint payoff. This problem is identical to a problem of a single decision maker who controls a path of incremental expansions.

From Lemma 2, we know that the skimming property holds for the social optimum and hence the problem is reduced to finding the policy Q that maximizes the expected social welfare. We denote the planner's payoff in state (x, q) as

$$U(Q; x, q) = \mathbb{E}\Big[\int_q^1 e^{-r\tau(s)}(x_{\tau(s)}v_H(s) + (1 - x_{\tau(s)})v_L(s))ds\Big|x, q; Q\Big]. \tag{4}$$

The planner's problem is then to find $\sup_Q U(Q; x, q)$. By applying Itô's lemma and using the properties of the Brownian motion, we have the following Hamilton-Jacobi-Bellman (HJB) equation for the planner's problem:

$$rV(x, q) = \max_{q^* \geq q}\left(r\int_q^{q^*}(xv_H(s) + (1 - x)v_L(s))ds + \frac{1}{2}V_{xx}(x, q^*)\frac{x^2(1 - x)^2}{\sigma^2}q^*\right). \tag{5}$$

We solve the planner's problem by showing that the HJB equation is solved by a boundary policy that cuts the state space into an expansion region and a waiting region. A verification argument then shows that our candidate solution also maximizes the original objective (4).

The optimal policy could in principle consist of several waiting and expansion regions. We proceed by guessing that there is only one expansion and only one waiting region and then later verify this guess (in Appendix C). Let $x^* : [0,1] \to [0,1]$ denote our candidate for the socially optimal policy, which we derive next. Function x^* splits the state space in two so that for a given q the planner waits for beliefs $x < x^*(q)$ and expands for beliefs $x \geq x^*(q)$. Since the planner internalizes the value of information for further decisions, we should intuitively expect the socially optimal expansion region to be larger than in the case of decentralized equilibrium, i.e. $x^*(q) < x^E(q)$. We shall verify that this property indeed holds.

We start by solving the value function that solves the HJB equation (5). In the waiting region below x^*, the value consists of the value of potential future actions and can be solved as:

$$V(x,q) = B(q)\Phi(x,q), \tag{6}$$

where

$$\Phi(x,q) := x^{\beta(q)}(1-x)^{1-\beta(q)} \text{ and } \beta(q) = \frac{1}{2}\left(1 + \sqrt{1 + \frac{8r\sigma^2}{q}}\right) \text{ as in Proposition 1.}$$

The next step is to find functions B and x^* that maximize the right-side of the HJB equation. To characterize these, we apply value matching and smooth pasting conditions, which are necessary for the optimality of policy x^*. Because the planner controls the intensity of experimentation, the conditions apply to a marginal increase of the stock q. The value matching condition is thus $V_q(x^*(q),q) = -x^*(q)v_H(q) - (1-x^*(q))v_L(q)$ and the smooth pasting condition is $V_{qx}(x^*(q),q) = -v_H(q) + v_L(q)$. Notice that the HJB equation consists of only future, not past, stopping payoffs and therefore the value matching condition equals the marginal value of increasing the stock with the lost stopping payoff.

Using Equation (6), we can write the value matching and smooth pasting conditions as

$$x^*(q)v_H(q) + (1-x^*(q))v_L(q) + B_q(q)\Phi(x^*(q),q) + B(q)\Phi_q(x^*(q),q) = 0, \tag{7}$$

$$v_H(q) - v_L(q) + B_q(q)\Phi_x(x^*(q),q) + B(q)\Phi_{qx}(x^*(q),q) = 0. \tag{8}$$

Our candidate policy x^* must balance the direct payoff effect, 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 latter terms of each equation as the derivatives of the value function.

We show in Appendix C that the system (7) - (8) can be transformed into a non-linear differential equation that defines our candidate policy x^*:

$$x^{*\prime}(q) = g(x^*(q), q), \tag{9}$$

where

$$\begin{aligned} g(x,q) =& x(1-x)\Big[x\Big(\beta'(q)(\beta(q)-1)v_H'(q) - ((\beta(q)-1)\beta''(q) - 2(\beta'(q))^2)v_H(q)\Big) \\ &+ (1-x)\Big(\beta'(q)\beta(q)v_L'(q) - (\beta(q)\beta''(q) - 2(\beta'(q))^2)v_L(q)\Big)\Big] / \\ &\Big[\Big(x(\beta(q)-1)^2 v_H(q) + (1-x)(\beta(q))^2 v_L(q)\Big)\beta'(q)\Big]. \end{aligned}$$

The appropriate initial condition for the differential equation is $x^*(1) = 1$ because the solution must equal the no-learning benchmark when the belief equals one.

The denominator of function g is zero at $(1,1)$ and hence a potential singularity problem arises. However, we show in Appendix C that the initial value problem has a unique solution below the decentralized solution i.e. a solution satisfying $x^*(q) \leq x^E(q)$ for all $q \in [0,1]$ (proof of Lemma 5 in Appendix C.2).[4] We then verify that together with the value function in (6) this candidate policy x^* solves the HJB equation, and we further verify that it also maximizes the original objective (4). In the process, we show that the policy function x^* is continuous and strictly increasing in q and hence satisfies the requirements for a boundary policy.

Proposition 2. *The socially optimal policy is a boundary policy characterized by the unique solution to the initial value problem* $x^{*\prime}(q) = g(x^*(q), q)$ *and* $x^*(1) = 1$ *such that the solution is always below the decentralized solution:* $x^*(q) \leq x^E(q)$ *for all* $q \in [0,1]$.

Proposition 2 confirms that we can solve the potentially complicated history-dependent problem with a simple boundary policy. However, unlike the decentralized equilibrium, we cannot solve the planner's problem in closed form because

[4] To see that the uniqueness can only hold in a restricted domain, note that $g(1,q) = 0$ and hence the initial value problem has a trivial solution $x(q) = 1$ for all $q \leq 1$.

the planner is truly forward-looking. For the socially optimal policy, *both past and future* actions are relevant. The past generates information that is useful in evaluating the right decision today, whereas future decisions can be based on information generated by today's action. A socially optimal policy balances the resulting tradeoff between the efficient use of information (option value effect) and the efficient production of information (information generation effect).

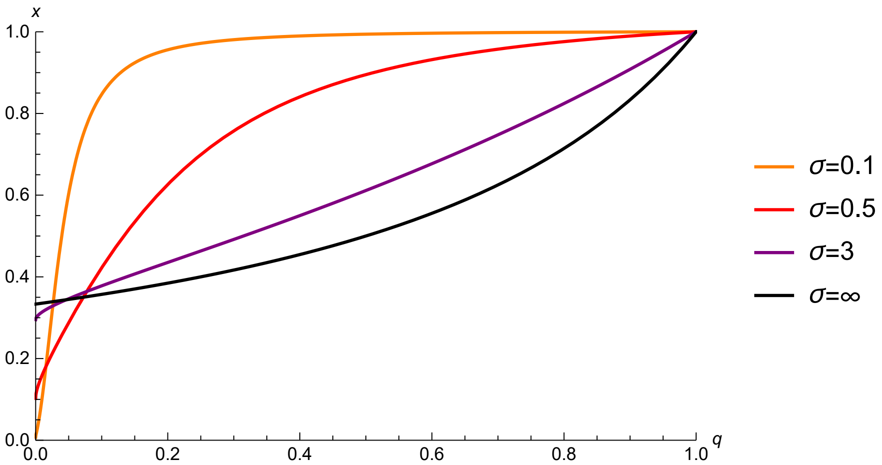

Figure 2: Socially optimal policy $x^*(q)$ for different σ when $v_H(q) = 1 - q$, $v_L(q) = -1/2$, and $r = 0.1$.

Figure 2 provides a numerical example of the effects of the signal precision. The smaller the noise parameter σ is, the more precise the signals are. Better learning technology decreases the cutoff belief $x^*(q)$ when the stock is small and increases it when the stock is high. This arises because improved learning amplifies both information generation and option value effects. The former dominates in the beginning, when the existing stock is low and there are many uncommitted agents who benefit from more information. Conversely, the option value effect dominates later when there are few such agents. Notice that the policies with learning (finite σ) are first below and later above the policy without learning ($\sigma = \infty$). Hence, gradual learning may either increase or decrease expansions as the informational tradeoff suggests. The following proposition generalizes this observation (see Appendix C.3 for the proof).

Proposition 3. *There exists* $\underline{x} \in (x^{stat}(0), 1)$ *and* $\overline{x} \in [\underline{x}, 1)$ *such that the optimal stock is strictly larger than the no-learning benchmark for all beliefs in* $(x^*(0), \underline{x})$

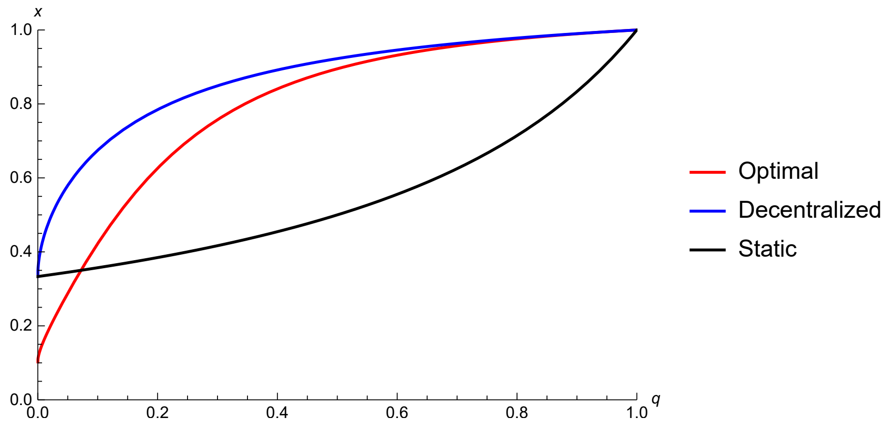

Figure 3: Different policies when $v_H(q) = 1 - q$, $v_L(q) = -1/2$, $\sigma = 0.5$, and $r = 0.1$.

and strictly lower for all beliefs in $(\overline{x}, 1)$.

Figure 3 illustrates the relationship between the solutions. Compared to the no-learning benchmark, gradual learning first increases and then decreases optimal expansions over time. The decentralized policy requires a higher belief for further expansions than the other policies.

Finally, it is illuminating to look at what happens to the actual speed of learning when the learning technology improves. To do that, let $q^*_\sigma(x)$ and $q^E_\sigma(x)$ denote the socially optimal and the decentralized stocks for signal precision σ.

Proposition 4. *(a) The socially optimal signal-to-noise ratio* $\sqrt{q^*_\sigma(x)}/\sigma \to \infty$ *as* $\sigma \to 0$ *for all* $x \in (0, 1)$. *(b) The decentralized signal-to-noise ratio* $\sqrt{q^E_\sigma(x)}/\sigma \to a(x)$ *as* $\sigma \to 0$ *where* $a(x) = 0$ *for all* $x \leq x^{stat}(0)$ *and* $a(x) \in (0, \infty)$ *for all* $x \in (x^{stat}(0), 1)$.

Learning gets arbitrarily fast in the socially optimal solution when the learning technology improves, whereas learning remains slow in the decentralized equilibrium. The latter is caused by informational free-riding: no-one wants to be the first one to stop if information arrives fast. In Appendix C.4, we prove Proposition 4 and derive the functional form for $a(x)$.

4 Payoff externalities

4.1 Model with payoff externalities

In our baseline model, the stopping payoff $v_\omega(\theta)$ depends only on state ω so that all externalities between the agents are purely informational. We now discuss an extension where the stopping payoff depends on the stopping decisions of the other agents. To do this in a meaningful way, we use the alternative interpretations of stopping payoffs where agents start receiving a flow payoff over time: as soon as an agent of type θ stops, he gets access to a payoff flow $\pi_\omega(\theta, q_t)$, where q_t is the mass of agents that have stopped by time t. The flow payoff for agent θ changes based on how the stock evolves. We assume that function $\pi_\omega(\theta, q)$ is increasing in θ and is continuously differentiable with bounded derivatives in both θ and q. As before, we normalize payoffs so that all types $\theta \in \left[\underline{\theta}, \overline{\theta}\right]$ want to stop if they know that the state is high: $\pi_H(\underline{\theta}, q) \geq 0$ for $0 \leq q \leq 1$.

Working within the class of Markovian policies, we can now write the stopping payoff of type θ in state (x, q) as:

$$v_\theta(x, q) = \mathbb{E}\Big[\int_t^\infty e^{-r(s-t)}\Big(x_s\pi_H(\theta, q_s) + (1 - x_s)\pi_L(\theta, q_s)\Big)ds\Big|x_t = x, q_t = q\Big].$$

It is important to recognize that, in contrast to our baseline model, this stopping payoff depends on the whole policy Q through expectations on future q_s. Notice also that if $\pi_\omega(\theta, q) \equiv \pi_\omega(\theta)$, i.e. if there are no payoff externalities, then we get back to our baseline model by denoting $v_\omega(\theta) := \frac{1}{r}\pi_\omega(\theta)$.

4.2 Decentralized equilibrium with payoff externalities

Let us consider the decentralized equilibrium with payoff externalities. In the current context, policy Q^E is a decentralized equilibrium if there exists a stopping profile $\mathcal{T}^E$ such that i) it is consistent with Q^E and ii) $\tau^E(\theta)$ maximizes $\mathbb{E}\left[e^{-r\tau(\theta)}v_\theta\left(x_{\tau(\theta)}, q_{\tau(\theta)}\right)|Q\right]$ for each θ when $Q = Q^E$. Analogously with the case without payoff externalities, let us define a candidate policy function for the decentralized equilibrium:

$$x^{E}\left(q\right):=\frac{-\beta\left(q\right)\pi_{L}\left(\theta\left(q\right),q\right)}{\left(\beta\left(q\right)-1\right)\pi_{H}\left(\theta\left(q\right),q\right)-\beta\left(q\right)\pi_{L}\left(\theta\left(q\right),q\right)}. \quad (10)$$

We prove in Appendix B.1 that the candidate policy characterizes the equilibrium:

Proposition 5. *If $x^{E}\left(q\right)$ defined in (10) is strictly increasing for all $q\in[0,1]$, then it characterizes the unique decentralized equilibrium in the model with payoff externalities.*

This result looks very similar to Proposition 1, but it contains a subtle additional insight. In the absence of payoff externalities, the players who ignore future changes in q_t merely forgo the future improvements in the learning process. The current result says that the players can also ignore the direct payoff effects of future stopping decisions, even if those effects remain payoff relevant after a player herself has stopped.

The main caveat is the requirement that $x^{E}\left(q\right)$ should be strictly increasing. In the absence of payoff externalities, this is always true. It is also true if payoff externalities are negative (directly follows from (10)):

Corollary 2. *There exists a unique decentralized equilibrium whenever $\pi_{\omega}\left(\theta,q\right)$ is decreasing in q for both $\omega\in\{H,L\}$.*

If payoff externalities are strongly positive, then $x^{E}\left(q\right)$ can be decreasing for some q.[5] In such a case, we should expect different types to bunch together their stopping decisions. This gives rise to a coordination problem among agents and, as is typical in coordination games, leads to multiplicity of equilibria.

In contrast to the decentralized equilibrium, the concept of social optimum is essentially unchanged by the introduction of the payoff externalities. This is because the social planner can internalize the payoff externality in her objective.

[5]More precisely, this happens when the effect from the positive payoff externality outweighs the joint effect from the decreasing type and the option value of waiting: $\beta(1-\beta)(\pi_{L,2}\pi_H-\pi_{H,2}\pi_L)>\beta(1-\beta)(\pi_{H,1}\pi_L-\pi_{L,1}\pi_H)+\pi_H\pi_L\beta'$, where $\pi_{\omega,1}$ and $\pi_{\omega,2}$ denote the derivatives with respect to the first and the second argument. Hence, a sufficient, but not necessary, condition for the monotonicity is that $\pi_{\omega,1}(\theta(q),q)\theta'(q)\geq\pi_{\omega,2}(\theta(q),q)$ for all q and ω.

4.3 Better learning technology can harm welfare

We now show that improvements in the learning technology may backfire: under positive payoff externalities, an improvement in the learning technology (lower σ) slows down learning and reduces total welfare. This shows that the interaction of informational and payoff externalities may lead to quite unexpected results.

For this subsection, consider the case where all agents are homogeneous: $\pi_\omega(\theta, q) \equiv \pi_\omega(q)$ is independent of θ.[6] Furthermore, let π_ω be strictly increasing in q and be such that all agents are willing to stop before the belief hits $1 : \pi_H(1) > 0$.

Let x^E_σ denote the policy function (10) under learning technology σ. When the positive externality is relatively mild, x^E_σ is strictly increasing and Proposition 5 holds. We then have the following welfare result that holds independent of the further properties of the solution (proof in Appendix B.2):

Proposition 6. *Let σ and $\sigma' < \sigma$ be such that both x^E_σ and $x^E_{\sigma'}$ are monotone. Then, the equilibrium with homogeneous agents is unique under both σ and σ' and the total expected welfare is strictly larger under σ than under σ' for all initial beliefs $x_0 \in (x^E_\sigma(0), x^E_\sigma(1)]$.*

The key step to prove Proposition 6 is to show that learning is paradoxically slower under the better learning technology. The intuition is that an agent does not want to stop when learning is equally fast because it means that the stock is lower under the better learning technology. A lower stock implies a lower stopping payoff due to positive payoff externalities. As a result, agents are worse off because they face both a slower increase in the stock and slower arrival of information.

Finally, let us comment what happens if the monotonicity of x^E breaks down. This happens if the positive payoff externality is strong compared to learning. Then, multiplicity of equilibria naturally arises because the positive externality creates a coordination problem between the agents. Even in that case, a result similar to Proposition 6 is likely to hold, for instance, for the comparison between the most efficient equilibria.

[6] More precisely, we take the limit of the model with heterogeneous agents in order to directly apply all results from previous chapters.

5 Applications

5.1 Mechanism design

We have solved the social planner's solution and the decentralized equilibrium separately. Next, we discuss how the planner's solution can be implemented in a decentralized manner by means of an appropriate incentive scheme. Appendix D shows in detail how a designer can implement policies – including the socially optimal policy – with transfers. We provide a short discussion here.

One way to implement a boundary policy is to use an *ex-ante* transfer rule, $P_0 : [\underline{\theta}, \overline{\theta}] \to \mathbb{R}$, where each type θ pays the transfer $P_0(\theta)$ and is assigned a given stopping rule at time 0. The *ex-ante* transfer rule is pinned down (up to a constant) by the envelope theorem:

$$\begin{aligned} P_0(\theta) &= \mathbb{E}\Big[e^{-r\tau(\theta)}(x_{\tau(\theta)}v_H(\theta) + (1 - x_{\tau(\theta)})v_L(\theta))\Big] \\ &\quad -\mathbb{E}\left[\int_{\underline{\theta}}^{\theta} e^{-r\tau(s)}\Big(x_{\tau(s)}v'_H(s) + (1 - x_{\tau(s)})v'_L(s))\Big)ds\right]. \end{aligned}$$

We further show when and how it is possible to implement the same policy by a posted price that postpones the transaction and type revelation to the moment when the agent stops. Implementation is always possible with a *dynamic posted price* rule that is a function of the current belief and hence continuously responds to news about the state. It turns out that alternative *simple posted prices*, which depend only on the stock, can often be used too. The analysis and the exact condition are in Appendix D. From the applied perspective, simple posted prices are of special interest because they are much easier to use in practice. For example, a seller of a new durable good could set the price based on the cumulative past sales instead of reviews or other feedback from past buyers.

5.2 Durable goods monopoly

We here illustrate the use of the mechanism design tools in the context of monopoly pricing of durable goods under uncertainty about the product quality. Gradual learning naturally arises in durable good markets: after a buyer purchases the product, he starts using it and observes how well it functions over time. The formal analysis are in Appendix D.5 but we present a few highlights here.

The first thing to notice is that the competitive market solution (price=marginal cost) corresponds with the decentralized equilibrium of Section 3.4. Hence, it suffers from informational free-riding and insufficient experimentation. We can utilize the planner's solution in Section 3.5 to solve the monopolist's problem, with the change that the monopolist maximizes the virtual surplus instead of the social welfare (see Sections D.4 and D.5).

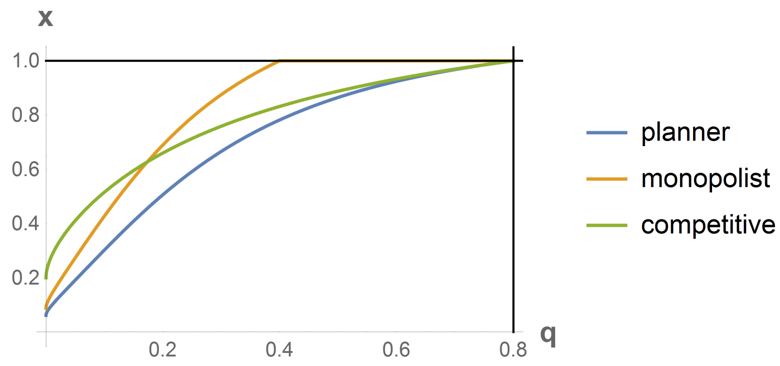

Figure 4: Different solutions. See Appendix D.5 for details.

We show that monopoly power can increase consumer welfare because the monopolist partly internalizes the informational externality. In the numerical example of Figure 4, the monopolist's and the competitive policies are everywhere above the planner's policy: both markets require inefficiently high belief for new consumers to purchase the product. The monopoly policy is below the competitive policy when the quantity is small. This is because the information generation effect encourages the monopolist to sell in the beginning and because early sales do not generate large information rents to other buyers. Later on, the monopolist reduces sales as the option value effect gets stronger and because later sales increase information rents to higher type buyers.

5.3 Capital investments in a competitive industry

Consider competitive entry to a market with unknown demand where firms make irreversible investments in productive capital. A larger installed capital mean a larger capacity and hence more sales and faster learning in the market. Furthermore, firms impose a negative payoff externality on each other because the larger the production capacity, the fiercer the competition and lower the prices. We explicitly model and analyze the case in Appendix E. The model is essentially the same as in Leahy (1993), except in our model information generation is endogenous, which drives a wedge between competitive equilibrium and social optimum.

Similar to durable goods, we find that the effect of market power on consumer welfare is ambiguous. If there are no barriers of entry, no firm has an incentive to start production just in order to find out the demand. This problem is especially pronounced when the learning technology is good and learning could potentially be fast. Figure 5 illustrates how signal precision σ affects the difference in consumer surplus between the social optimum and the competitive equilibrium as a function of the initial belief. We show in Appendix E that the designer can implement the social optimum with investment subsidies that decrease over time.

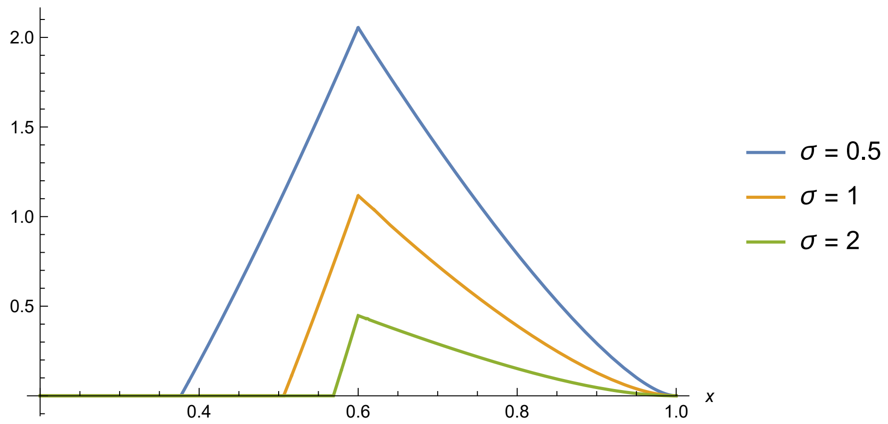

Figure 5: The difference in consumer welfare between the social optimum and the competitive equilibrium. See Appendix E for details.

5.4 Type-dependent informativeness: experts versus fanatics

Our baseline model assumes that all agents are equally informative: one unit of the stock q produces the same (marginal) amount of information. However, this might not necessarily be true in many applications of our model. For example, first buyers might be fans of the product whose experience matters less for the general population. Or conversely, the first units might be acquired by experts who are able to deduce the true value of the product much more quickly than average users. In Appendix F, we show how to extend the model to type-dependent informativeness.

The comparison between the experts and the fanatics cases provides an additional perspective of the informational tradeoff between information generation and option value of waiting. Figure 6 illustrates how heterogeneous informativeness affects the socially optimal policy. The information generation effect is more pronounced for low q in the experts environment, while it is more pronounced for higher q in the fanatics environment. We see this in Figure 6 as the solution in the experts environment is first below and then quickly rises above the fanatics solution. The experts environment favors relatively early expansions because the high types produce more information than in the fanatics environment.

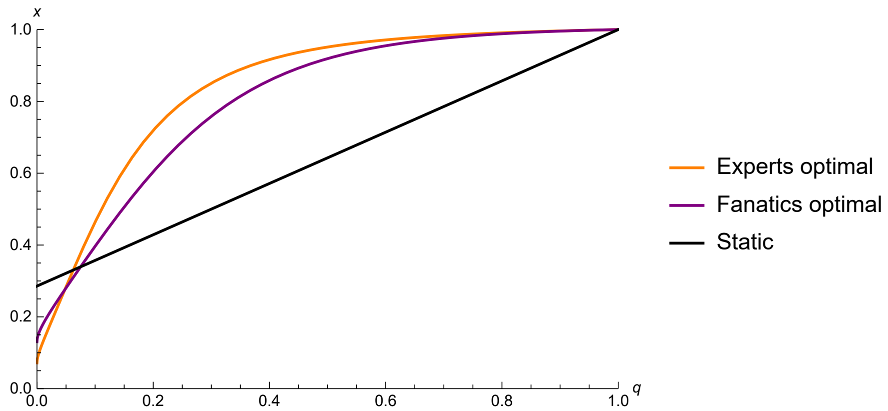

Figure 6: Socially optimal policies for experts and fanatics. See Appendix F for details.

5.5 Other applications

The main contribution of this paper is to develop a methodology for analyzing implications of irreversibility and endogenous gradual learning. We have discussed a few main applications, but gradual learning is present in many other contexts that we do not explicitly model in this paper or in its supplementary material.

We conclude by pointing out some potential avenues for further applications.

Our methodology can be applied to analyze adoption patterns of new technologies in a much broader sense than industry investment. Think of, for example, new medical devices and treatments. While they go through rigorous testing before approved for use, there are numerous examples of products which have been withdrawn due to concerns for the well-being of patients.[7] In many cases, the true effectiveness of new products is only learned gradually from experience.[8] Similarly, R&D or development aid projects tie resources for many years and information about the outcomes arrives gradually over different stages of the project.

Many actions beyond adopting new technology have uncertain and irreversible consequences. One example is the emissions of pollution which affect the environment. Once created, it is hard to reduce a stock of emissions. A prominent example is the regulation of anthropogenic greenhouse gases whose effects on the climate and society remain uncertain despite considerable research effort. Arguably, we can learn the true effects only gradually from experience. Our mechanism design approach is a useful starting point for designing environmental regulation that takes into account both informational and environmental externalities.[9]

There are many sources of irreversibility in public policy making. Policies themselves may be hard to change because of political uncertainty, potential legal consequences, or legislative lags. But even if changed, a policy may have already affected many people. As a concrete example, consider educational policy, such as the maximal size of a class room, which has an irreversible impact on children. Later labor market outcomes and other information we collect from each cohort helps to evaluate the educational policy at the time when they went to school, independent of whether the norms have changed since then. The trade-off that our framework addresses is how to balance the value of information generation

[7]See for example DePuy hip replacement recall or for example the FDA list of recalled medical devices for 2018.

[8]Some products have both instantaneous and gradual features. Consider e.g. vaccines: some side effects may get revealed immediately after taking the shot but the effectiveness of the vaccine can be learned only gradually over time as vaccinated people get exposed to the disease.

[9]Liski and Salanié (2020) analyze the optimal policy in a different model that focuses on the possibility of a one-time catastrophe that is triggered once the emission stock exceeds an unknown threshold.

and the risk of irreversible negative consequences.

6 Concluding remarks

We conclude by summarizing the most important reasons to analyze the gradual arrival of endogenous information. The first reason is that the approach enables the analysis of various real-life situations where the long-run consequences of a decision determine its profitability. The gradual arrival of information dramatically shapes the incentives of experimentation. Gradual learning creates a novel informational tradeoff on the social level between information generation and the option value of waiting.

The second motive to model the gradual arrival of information is technical. As demonstrated in this paper, the decentralized equilibrium can be solved in a closed form under gradual learning. The solution method extends to richer environments, such as to models with payoff externalities. We further show how mechanism design techniques can be utilized to conduct policy analysis in our environment.

An important takeaway from the paper is that the signal precision has subtle implications for learning. We show that even if signals get arbitrarily precise, learning remains slow in the equilibrium. This contrasts with the socially optimal solution, in which the true state is learned arbitrarily fast as the learning technology improves. As a result, under endogenous learning the signal precision may have unexpected implications for welfare and policies. The equilibrium welfare loss is particularly severe if the learning technology is good. We show that a better learning technology may even slow down learning if there are positive payoff externalities between the agents.

References

Bergemann, Dirk and Juuso Välimäki, "Market diffusion with two-sided learning," *RAND Journal of Economics*, 1997, *28* (4), 773–795.

_ **and** _ , "Experimentation in markets," *Review of Economic Studies*, 2000, *67* (2), 213–234.

Board, Simon, "Selling options," *Journal of Economic Theory*, 2007, *136* (1), 324–340.

_ **and Andrzej Skrzypacz**, "Revenue management with forward-looking buyers," *Journal of Political Economy*, 2016, *124* (4), 1046–1087.

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.

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, pp. 871–925.

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*, 2020.

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.

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.

Kruse, Thomas and Philipp Strack, "Optimal stopping with private information," *Journal of Economic Theory*, 2015, *159*, 702–727.

Laiho, Tuomas and Julia Salmi, "Coasian Dynamics and Endogenous Learning," *Working paper*, 2021.

Leahy, John V, "Investment in competitive equilibrium: The optimality of myopic behavior," *The Quarterly Journal of Economics*, 1993, *108* (4), 1105–1133.

Liski, Matti and Francois Salanié, "Catastrophes, delays, and learning," *Toulouse School of Economics Working Papers No 1148*, 2020.

Martimort, David and Louise Guillouet, "Precaution, information and time-inconsistency: on the value of the precautionary principle," *CEPR Discussion Paper Series DP15266*, 2020.

McDonald, Robert and Daniel Siegel, "The value of waiting to invest," *The quarterly journal of economics*, 1986, *101* (4), 707–727.

Milgrom, Paul and Ilya Segal, "Envelope theorems for arbitrary choice sets," *Econometrica*, 2002, *70* (2), 583–601.

Moscarini, Giuseppe and Lones Smith, "The Optimal Level of Experimentation," *Econometrica*, 2001, *69* (6), 1629–1644.

Murto, Pauli and Juuso Välimäki, "Learning and Information Aggregation in an Exit Game," *The Review of Economic Studies*, 2011.

Pindyck, Robert S., "Irreversible Investment, Capacity Choice, and the Value of the Firm," *The American Economic Review*, 1988, *78* (5), 969–985.

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.

Appendix – For Online Publication

Appendices A, B, and C contain proofs for the main text. Appendices D, E, F, and G are supplementary material with extensions and applications.

A Additional material for Sections 2 and 3.1

Learning process as the continuous limit

Consider a discrete model where the number of agents is n and where the period length is dt. Let the signal process be such that in each period, each agent who has stopped generates a normally distributed conditionally iid. signal:

$$y_t^i \sim N\left(\frac{\mu_\omega dt}{n}, \frac{\sigma^2 dt}{n}\right).$$

This normalization keeps the informativeness of the aggregate signal constant while letting the number of small agents to grow as in Bergemann and Välimäki (1997).

When the number of agents who has stopped is $k \leq n$, this implies the following aggregate signal:

$$\sum_{i=1}^{k} y_t^i \sim N\left(\mu_\omega dt \frac{k}{n}, \sigma^2 dt \frac{k}{n}\right).$$

Let $q = k/n$ denote the fraction of agents who have stopped. Now, the signal process (1) follows once we take the limit when $n \to \infty$ (and $k \to \infty$ so that k/n stays fixed) and $dt \to 0$.

Notice that the limiting distribution for the aggregate signal depends only on the mean and the variance of y_t^i (the central limit theorem). Hence, the signal process (1) is also the limiting process for the case where y_t^i is not normally distributed, including the case where agents communicate through binary signals.

Furthermore, we can rewrite the model so that the individual signals represent realized payoffs in a model where agents start receiving a stochastic flow payoff after stopping: $\pi_t(\theta) = \pi_\omega(\theta) + \epsilon_t(\theta)$ where $\epsilon_t(\theta) \sim N\left(0, \sigma^2(\pi_H(\theta) - \pi_L(\theta))^2\right)$.

The noise term is scaled so that every increment in q is equally informative. This assumption is not necessary: as Appendix F points out, both the analysis and the qualitative results remain unchanged with heterogeneous informativeness if the stopping profile is monotone. When we set $\pi_\omega(\theta) = rv_\omega(\theta)$, the expected stopping payoff is $x_t v_H(\theta) + (1 - x_t)v_L(\theta)$ just like in the main text. Since there are no further actions after stopping, it does not matter how fast the agents learn privately after they have stopped: the parameter σ can be interpreted to capture both the noise in the private learning and the noise in communication.

Proof of Lemma 1

Proof. Let policy Q be fixed. Type θ wants to stop at time t if

$$x_t v_H(\theta) + (1 - x_t)v_L(\theta) \geq \mathbb{E}[e^{-r(\tau-t)}(x_\tau v_H(\theta) + (1 - x_\tau)v_L(\theta))|\mathcal{F}_t; Q],$$

for all stopping rules τ. Or equivalently,

$$v_L(\theta)(1 - x_t - \mathbb{E}[e^{-r(\tau-t)}(1 - x_\tau)|\mathcal{F}_t; Q]) + v_H(\theta)(x_t - \mathbb{E}[e^{-r(\tau-t)}x_\tau|\mathcal{F}_t; Q]) \geq 0.$$

The left-hand side is increasing in θ because expressions $(1-x_t-\mathbb{E}[e^{-r(\tau-t)}(1-x_\tau)])$ and $(x_t-\mathbb{E}[e^{-r(\tau-t)}x_\tau])$ are positive (follows from that x_τ is a martingale and $e^{-r(\tau-t)} < 1$) and v_ω is increasing. Therefore, if type θ wants to stop, type $\theta' > \theta$ wants to stop too. □

Proof of Lemma 2

Proof. $\mathcal{T}$ and $\mathcal{T}^{mon}$ are both consistent with Q. We show that monotone stopping ordering maximizes *ex post* welfare for all realized paths of (X, Q). The claim follows once we show that for all types $\theta, \theta' \in [\underline{\theta}, \overline{\theta}]$ such that $\theta > \theta'$ and for all realized stopping times $t, t' \in \mathbb{R}_+$ such that $t \leq t'$,

$$e^{-rt}v_\omega(\theta) + e^{-rt'}v_\omega(\theta') \geq e^{-rt'}v_\omega(\theta) + e^{-rt}v_\omega(\theta').$$

The above condition is equivalent with $(e^{-rt} - e^{-rt'})(v_\omega(\theta) - v_\omega(\theta')) \geq 0$, which necessarily holds as $t \leq t'$ and $v_\omega(\theta) \geq v_\omega(\theta')$ by assumption if $\theta > \theta'$. □

B Decentralized equilibrium: Sections 3.4 and 4

B.1 Proof of Proposition 1 and Proposition 5

Proposition 1 is a special case of Proposition 5 that allows the possibility of payoff externalities. Therefore, we do not provide a separate proof for Proposition 1: without consulting Section 4, one can use the proof of Proposition 5 for Proposition 1 by plugging in $\pi_\omega(\theta(q), q) = rv_\omega(q)$.

We present the proof of Proposition 5 in the next subsection. Before that, we briefly comment the main idea of the proof and its connection with the special case without payoff externalities (Proposition 1). First, we verify that the candidate is indeed a decentralized equilibrium. This part of the proof would remain largely identical even without payoff externalities. Second, we show the uniqueness of the equilibrium. The second part is mostly needed for the case with payoff externalities. Without payoff externalities, the monotonicity basically rules out any other equilibria except the candidate policy. For both parts of the proof, we use two useful lemmas that we present first as the preliminary step of the proof.

B.1.1 Preliminaries

We will utilize the following lemma that states the optimal stopping problem of an individual agent as a minimization problem of the *opportunity cost of delay*:

Lemma 3. *Fix arbitrary policy Q. Then type θ should not stop at state point (x, q) if there exists a stopping time τ such that*

$$\mathbb{E}\left[\int_0^\tau e^{-rt}\left(x_t\pi_H\left(\theta, q_t\right) + \left(1 - x_t\right)\pi_L\left(\theta, q_t\right)\right)dt \,|(x, q; Q)\right] < 0. \tag{11}$$

If no such stopping time exists, then it is optimal for type θ to stop immediately at state point (x, q).

Proof. Given policy Q, the optimal stopping time τ_θ^* of type θ must satisfy:

$$\tau_\theta^* \in \arg\sup_\tau \mathbb{E}\left[\int_\tau^\infty e^{-r(t)}\left(x_t\pi_H\left(\theta, q_t\right) + \left(1 - x_t\right)\pi_L\left(\theta, q_t\right)\right)dt \,|(x, q; Q)\right], \tag{12}$$

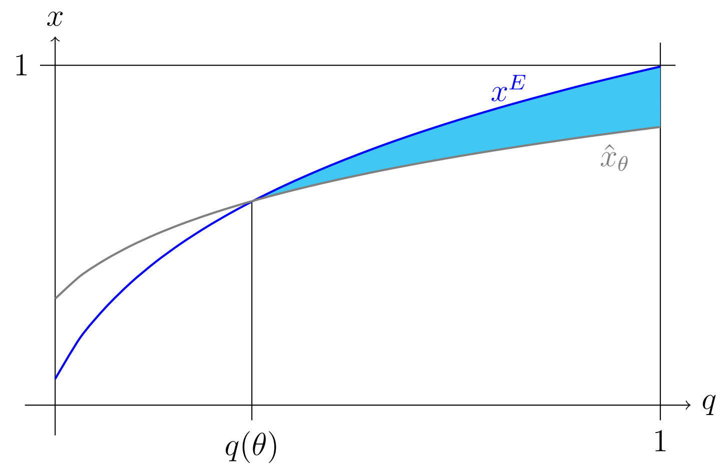

Figure 7: Optimal stopping for type θ.

where maximization is over all stopping times τ adapted to $\mathcal{F}_t$ under policy Q with initial value $(x_0, q_0) = (x, q)$. Rewriting the expectation as

$$\mathbb{E}\left[\int_0^\infty e^{-rt}\left(x_t\pi_H\left(\theta, q_t\right) + \left(1 - x_t\right)\pi_L\left(\theta, q_t\right)\right)dt \right. \\ \left. - \int_0^\tau e^{-rt}\left(x_t\pi_H\left(\theta, q_t\right) + \left(1 - x_t\right)\pi_L\left(\theta, q_t\right)\right)dt \,|(x, q; Q)\right],$$

we note that (12) is equivalent to

$$\tau_\theta^* \in \arg\inf_\tau \mathbb{E}\left[\int_0^\tau e^{-rt}\left(x_t\pi_H\left(\theta, q_t\right) + \left(1 - x_t\right)\pi_L\left(\theta, q_t\right)\right)dt \,|(x, q; Q)\right], \tag{13}$$

and so $\tau_\theta^* > 0$ only if (11) holds. □

Throughout the proof, we will also utilize the solution to the auxiliary optimal stopping problem, where the stock is assumed to be fixed at $q_t \equiv q$ forever:

Lemma 4. *Take the policy Q, where the stock is fixed at $q_t \equiv q$ forever. Then, it is optimal for θ to stop if and only if $x_t \geq \widehat{x}_\theta(q)$, where*

$$\widehat{x}_\theta(q) = \frac{-\beta(q)\,\pi_L(\theta, q)}{(\beta(q) - 1)\,\pi_H(\theta, q) - \beta(q)\,\pi_L(\theta, q)}.$$

Proof. The derivation of this result is perfectly analogous with the corresponding problem in e.g. Dixit and Pindyck (1994), where the stochastic variable follows a geometric Brownian motion instead of (2). □

Figure 7 shows $x^E(q)$ and the optimal stopping threshold $\widehat{x}_\theta(q)$ for type θ under the assumption of fixed $q_t \equiv q$. By definition, $\widehat{x}_\theta(q)$ coincides with the policy function $x^E(q)$ at $q = q(\theta)$, and $\widehat{x}_\theta(q) > (<)\, x^E(q)$ for $q < (>)\, q(\theta)$.

B.1.2 Existence and characterization

Suppose that $x^E(q)$ defined in (10) is strictly increasing for $0 \leq q \leq 1$ and denote by Q^E the corresponding policy. We will show in this section that the optimal stopping strategy τ^*_θ for type θ against Q^E is to stop whenever the current state (x_t, q_t) is in the blue shaded region in Figure 7. Since the optimal realized stopping time under this strategy for θ coincides with the moment where q_t hits $q(\theta)$ for the first time, the stopping profile $\tau^* := \{\tau^*_\theta\}_{\theta=\underline{\theta}}^{\overline{\theta}}$ is consistent with Q^E. It then follows that Q^E is a decentralized equilibrium.

We will proceed through two main steps. In step 1, we will show that whenever $q_t < q(\theta)$, it is strictly optimal for θ to wait. In step 2, we will show that if $q_t > q(\theta)$, then it is optimal for θ to stop at the boundary point $x_t = x^E(q_t)$. We show that this will further imply that it is already optimal to stop at all points $x_t \in \left[\widehat{x}_\theta(q), x^E(q)\right]$, i.e. in the blue shaded region of Figure 7.

Step 1: Optimal to wait when $q < q(\theta)$

Assume that the current state is (x, q), where $x \leq x^E(q)$, $q < q(\theta)$ and let $\tau^*_\theta := \inf\left\{t : (x_t, q_t) = \left(x^E(q(\theta)), q(\theta)\right)\right\}$ denote the first hitting time of point $x^E(q(\theta)), q(\theta)$. We will show below that this stopping time delivers a strictly negative opportunity cost of delay:

$$F^*(x,q) := \mathbb{E}\left[\int_0^{\tau^*_\theta} e^{-rt}\left(x_t \pi_H(\theta, q) + (1 - x_t)\,\pi_L(\theta, q)\right) dt\, |(x, q; Q)\right] < 0,$$

and hence by Lemma 3 it is strictly non-optimal to stop at (x, q).

We will first establish some properties of the function $F^*(x, q)$ defined above. Since stopping at state point $(x^E(q(\theta)), q(\theta))$ is immediate, we must have

$$F^*\left(x^E(q(\theta)), q(\theta)\right) = 0. \tag{14}$$

Moreover, $F^*(x, q)$ cannot be lower than the present value of obtaining the lowest possible flow payoff forever nor higher than obtaining the highest possible flow payoff forever, and therefore discounting guarantees that $F^*(x,q)$ is bounded from below and above.

The general solution to the HJB-equation for $F^*(x,q)$ that must hold whenever there is no stopping takes the form:

$$F^*(x,q) = \frac{x\pi_H(\theta,q) + (1-x)\pi_L(\theta,q)}{r} + A(q)\cdot\Phi(x,q) + B(q)\cdot\widetilde{\Phi}(x,q), \quad (15)$$

where $A(q)$ and $B(q)$ are constants and $\Phi(x,q) := x^{\beta(q)}(1-x)^{1-\beta(q)}$, $\widetilde{\Phi}(x,q) := (1-x)^{\beta(q)}x^{1-\beta(q)}$, and $\beta(q) := \frac{1}{2}\left(1+\sqrt{1+\frac{8r\sigma^2}{q}}\right)$. Noting that $\lim_{x\downarrow 0}\widetilde{\Phi}(x,q) = \infty$, the requirement that $F^*(x,q)$ is bounded implies that $B(q)=0$ and so

$$F^*(x,q) = \frac{x\pi_H(\theta,q) + (1-x)\pi_L(\theta,q)}{r} + A(q)\cdot\Phi(x,q). \quad (16)$$

Finally, we can write the derivative of $F^*(x,q)$ w.r.t. q along the boundary $x^E(q)$ as:

$$\begin{aligned}\frac{d}{dq}F^*\left(x^E(q),q\right) &= \frac{\partial}{\partial x}F^*\left(x^E(q),q\right)\frac{d}{dq}x^E(q) + \frac{\partial}{\partial q}F^*\left(x^E(q),q\right)\\ &= \left[\frac{\pi_H(\theta,q)-\pi_L(\theta,q)}{r} + A(q)\cdot\Phi_x\left(x^E(q),q\right)\right]\frac{d}{dq}x^E(q), \quad (17)\end{aligned}$$

where we have utilized the fact that q_t increases at the boundary $x^E(q)$, and hence the first-order effect of q on $F^*\left(x^E(q),q\right)$ vanishes there: $\frac{\partial}{\partial q}F^*\left(x^E(q),q\right)=0$.

We will first show that for all $q<q(\theta)$, $F^*(x,q)<0$ at the boundary $x = x^E(q)$. For contradiction, suppose that $F^*\left(x^E(q),q\right)\geq 0$ for some $q<q(\theta)$. By (16), this means that

$$A(q) \geq -\frac{x^E(q)\pi_H(\theta,q) + \left(1-x^E(q)\right)\pi_L(\theta,q)}{r}\frac{1}{\Phi\left(x^E(q),q\right)},$$

and so by (17)

$$\begin{aligned}\frac{d}{dq}F^*\left(x^E(q),q\right) \geq \frac{d}{dq}x^E(q)\cdot&\left[\frac{\pi_H(\theta,q)-\pi_L(\theta,q)}{r}\right.\\ &\left.-\frac{x^E(q)\pi_H(\theta,q)+\left(1-x^E(q)\right)\pi_L(\theta,q)}{r}\frac{\Phi_x\left(x^E(q),q\right)}{\Phi\left(x^E(q),q\right)}\right]. \quad (18)\end{aligned}$$

Notice that $\frac{d}{dq}x^E(q)>0$ and that

$$\begin{aligned}\frac{\Phi_x\left(x^E(q),q\right)}{\Phi\left(x^E(q),q\right)} &= \frac{\beta(q)-x^E(q)}{x^E(q)\left(1-x^E(q)\right)} \text{ and}\\ x^E(q) &< \frac{-\beta(q)\pi_L(\theta,q)}{(\beta(q)-1)\pi_H(\theta,q)-\beta(q)\pi_L(\theta,q)} \text{ for } q>q(\theta).\end{aligned}$$

It then follows by straightforward algebra that the right hand side of (18) is strictly positive. We can conclude that $\frac{d}{dq}F^*\left(x^E(q), q\right) > 0$, and since this holds for an arbitrary boundary point $x^E(q)$ such that $q < q(\theta)$, it further implies that $F^*\left(x^E(q(\theta)), q(\theta)\right) > 0$. This is a contradiction with (14).

We conclude that $F^*\left(x^E(q), q\right) < 0$ for all $q < q(\theta)$, which by (16) implies

$$A(q) < -\frac{x^E(q)\,\pi_H(\theta, q) + \left(1 - x^E(q)\right)\pi_L(\theta, q)}{r}\frac{1}{\Phi\left(x^E(q), q\right)}.$$

Applying (16) again for an arbitrary $x \leq x^E(q)$, $q < q(\theta)$, we have

$$\begin{aligned} F^*(x, q) &= \frac{x\pi_H(\theta, q) + (1 - x)\pi_L(\theta, q)}{r} + A(q) \cdot \Phi(x, q) \\ &< \frac{x\pi_H(\theta, q) + (1 - x)\pi_L(\theta, q)}{r} \\ &\quad - \frac{\Phi(x, q)}{\Phi\left(x^E(q), q\right)} \frac{x^E(q)\,\pi_H(\theta, q) + \left(1 - x^E(q)\right)\pi_L(\theta, q)}{r}. \end{aligned}$$

This last expression is concave in x, has a root at $x = x^E(q)$, and its derivative at $x = x^E(q)$ is the term in brackets in (18), which we already found strictly positive. It follows that $F^*(x, q) < 0$ for all $q < q(\theta)$, $x \leq x^E(q)$, and therefore it cannot be optimal for θ to stop at any $q < q(\theta)$.

Step 2: Optimal to stop at boundary for $q > q(\theta)$

We will next show that when $q > q(\theta)$, it is optimal to stop at boundary $x = x^E(q)$. Let

$$F(x, q) := \inf_{\tau} \mathbb{E}\left[\int_0^{\tau} e^{-rt}\left(x_t\pi_H(\theta, q) + (1 - x_t)\pi_L(\theta, q)\right)dt\,|(x, q; Q)\right].$$

For contradiction, suppose that there is some $q > q(\theta)$ such that $F\left(x^E(q), q\right) < 0$ so that by Lemma 3 it is not optimal to stop at that boundary point. As before, the form of the value function in the continuaton region must be

$$F(x, q) = \frac{x\pi_H(\theta, q) + (1 - x)\pi_L(\theta, q)}{r} + A(q) \cdot \Phi(x, q) + B(q) \cdot \widetilde{\Phi}(x, q),$$

where $A(q)$ and $B(q)$ are some constants. We will first show that our assumption $F\left(x^E(q), q\right) < 0$ implies $F_x\left(x^E(q), q\right) < 0$. We will do that separately for potential cases $B(q) \geq 0$ and $B(q) < 0$.

Suppose that $B(q) \geq 0$. Then $F\left(x^E(q), q\right) < 0$ implies that

$$A(q)\,\Phi\left(x^E(q), q\right) < -\frac{x^E(q)\,\pi_H(\theta, q) + \left(1 - x^E(q)\right)\pi_L(\theta, q)}{r}. \tag{19}$$

But then,

$$\begin{aligned}
F_x\left(x^E(q), q\right) &= \frac{\pi_H(\theta, q) - \pi_L(\theta, q)}{r} + \frac{\beta(q) - x^E(q)}{x^E(q)\left(1 - x^E(q)\right)} A(q)\,\Phi\left(x^E(q), q\right) \\
&+ \frac{1 - x^E(q) - \beta(q)}{x^E(q)\left(1 - x^E(q)\right)} B(q)\,\Phi\left(x^E(q), q\right) \\
&\leq \frac{\pi_H(\theta, q) - \pi_L(\theta, q)}{r} + \frac{\beta(q) - x^E(q)}{x^E(q)\left(1 - x^E(q)\right)} A(q)\,\Phi\left(x^E(q), q\right) \\
&< \frac{\pi_H(\theta, q) - \pi_L(\theta, q)}{r} \\
&- \frac{\beta(q) - x^E(q)}{x^E(q)\left(1 - x^E(q)\right)} \frac{x^E(q)\,\pi_H(\theta, q) + \left(1 - x^E(q)\right)\pi_L(\theta, q)}{r}.
\end{aligned} \tag{20}$$

where the first inequality is implied by $B(q) \geq 0$ and $\beta(q) > 1$, and the second inequality is implied by (19). Noting that the expression$(\beta(q) - 1)\, x\pi_H(\theta, q) + \beta(q)(1 - x)\,\pi_L(\theta, q)$ is increasing in x and has a root at $x = \widehat{x}_\theta(q) < x^E(q)$, it follows that

$$(\beta(q) - 1)\, x^E(q)\,\pi_H(\theta, q) + \beta(q)\left(1 - x^E(q)\right)\pi_L(\theta, q) > 0,$$

which is equivalent to

$$\frac{\pi_H(\theta, q) - \pi_L(\theta, q)}{r} < \frac{\left(\beta(q) - x^E(q)\right)\left(x^E(q)\,\pi_H(\theta, q) + \left(1 - x^E(q)\right)\pi_L(\theta, q)\right)}{r x^E(q)\left(1 - x^E(q)\right)}$$

and so it follows from (20) that $F_x\left(x^E(q), q\right) < 0$.

Suppose next that $B(q) < 0$. This is only possible if there is some $x' < x^E(q)$ where $F(x', q) \geq 0$ (otherwise it would be optimal to continue for all values $x < x^E(q)$ and $B(q) < 0$ would imply $\lim_{x \downarrow 0} F(x, q) = B(q) \lim_{x \downarrow 0} \widetilde{\Phi}(x, q) = -\infty$ contradicting the boundary condition at $x = 0$). If $A(q) \geq 0$, then $F(x, q)$ is increasing in x, but then $F(x', q) \geq 0$ contradicts our running assumption $F\left(x^E(q), q\right) < 0$. The only remaining possibility is that both $B(q)$ and $A(q)$ are strictly negative. In that case $F(x, q)$ is concave, and so $F(x', q) = 0$ and $F\left(x^E(q), q\right) < 0$ together imply $F_x\left(x^E(q), q\right) < 0$.

We have now shown that $F\left(x^E(q),q\right)<0$ implies $F_x\left(x^E(q),q\right)<0$. But then the derivative of $F\left(x^E(q),q\right)$ along the boundary is also negative:

$$\frac{d}{dq}F\left(x^E(q),q\right)=F_x\left(x^E(q),q\right)\frac{d}{dq}x^E(q)<0.$$

Iterating the argument along the boundary implies that $F\left(x^E(1),1\right)<0$. But since $q_t=1$ is an absorbing value for q_t, we can treat $q_t\equiv 1$ as fixed. Lemmas 3 and 4 imply that $F(x,1)=0$ for $x>\widehat{x}_\theta(1)$. Since $x^E(1)>\widehat{x}_\theta(1)$, we have a contradiction. It follows that $F\left(x^E(q),q\right)>0$ for all $q>q(\theta)$, and hence it is strictly optimal to stop at all boundary points for $q>q(\theta)$.

Finally, we note that θ will then optimally stop at all points within the shaded region in Figure 1. Suppose that $q_t>q(\theta)$, $x_t<x^E(q)$, and type θ has not yet stopped. Then, as we have just showed, θ will optimally stop at the latest at boundary $x^E(q_t)$, and therefore we may treat q_t as fixed until optimal stopping of type θ. It then follows from Lemma 4 that θ should stop at all state values $x_t\in\left[\widehat{x}_\theta(q),x^E(q)\right]$.

B.1.3 Uniqueness

We show here that no policy other than the boundary policy defined by $x^E(q)$ can be a decentralized equilibrium. We will do that in two steps. In the first step, we will show that the region $x>x^E(q)$ of the state space must be an expansion region in any equilibrium, and in the second step we will show that there is no equilibrium where some player wants to stop for $x<x^E(q)$.

As a preliminary step, we point out that Lemma 1 direcly extends to the case with payoff externalities. Hence, any equilibrium policy Q must be such that θ optimally stops when the stock has value $q_t=q(\theta)$.

Step 1: expansion when $x>x^E(q)$

Suppose that at time t the state is given by $(x_t,q_t)=(x,q)$, where $x>x^E(q)=\widehat{x}_{\theta(q)}(q)$. If the stock were to be fixed at $q_t\equiv q$ forever, then by Lemma 4 it would be optimal for type $\theta(q)$ to stop immediately. Combining this with Lemma 3 implies that for fixed $q_t \equiv q$, we have

$$\mathbb{E}\left[\int_0^\tau e^{-rt}\left(x_t\pi_H\left(\theta,q\right)+\left(1-x_t\right)\pi_L\left(\theta,q\right)\right)dt\left|(x,q;q_t\equiv q)\right.\right]\geq 0 \tag{21}$$

for all stopping times τ (given initial value (x,q) and $q_t \equiv q$ fixed). But since the skimming property holds in equilibrium, type $\theta\left(q\right)$ should indeed treat the stock fixed at $q_t \equiv q$ for all $t \leq \tau$, and since the expression (21) does not depend on the evolution of q_t for $t > \tau$, we can conclude that in any equibrium type $\theta\left(q\right)$ should stop whenever $x_t > x^E\left(q\right)$. Since this conclusion was for arbitrary q, it follows that the whole region $x > x^E\left(q\right)$ must be an expansion region in equilibrium.

Step 2: no stopping when $x < x^E\left(q\right)$

We will now show that it cannot be optimal for any player to stop for $x < x^E\left(q\right)$ in equilibrium. We will prove this through iterated deletion of dominated stopping strategies. Define

$$\overline{\pi}_\omega\left(\theta;q',q''\right):=\max_{q\in[q',q'']}\pi_\omega\left(\theta,q\right),$$

and define a mapping $\widetilde{x}_\theta\left(q\right):\left[q\left(\theta\right),1\right]\to[0,1]$ as

$$\widetilde{x}_\theta\left(q\right):=\sup\Big\{x:\ \text{there exists a stopping time } \tau \text{ such that}$$
$$\mathbb{E}\left(\int_0^\tau e^{-rt}\left(x\overline{\pi}_H\left(\theta;q\left(\theta\right),q\right)+\left(1-x\right)\overline{\pi}_L\left(\theta;q\left(\theta\right),q\right)\right)dt\left|x,q_t\equiv q(\theta)\right.\right)<0\Big\}.$$

The interpretation is that $\widetilde{x}_\theta\left(q\right)$ would be the optimal stopping threshold of type θ if belief x_t follows process (2) with q_t fixed at $q\left(\theta\right)$, but with the flow payoff given in each state by $\max_{s\in[q(\theta),q]}\pi_\omega\left(\theta,s\right)$. It can be solved in closed form:

$$\widetilde{x}_\theta\left(q\right) \quad = \quad \frac{-\beta\left(q\left(\theta\right)\right)\overline{\pi}_L\left(\theta;q\left(\theta\right),q\right)}{\left(\beta\left(q\left(\theta\right)\right)-1\right)\overline{\pi}_H\left(\theta;q\left(\theta\right),q\right)-\beta\left(q\left(\theta\right)\right)\overline{\pi}_L\left(\theta;q\left(\theta\right),q\right)}. \tag{22}$$

We see directly from (22) that $\widetilde{x}_\theta\left(q\right)$ is Lipschitz continuous in q and θ, decreasing in q, strictly decreasing in θ, and $\widetilde{x}_\theta\left(q\left(\theta\right)\right)=x^E\left(q\left(\theta\right)\right)$ with $\widetilde{x}_{\theta(1)}\left(1\right)=1$.

We next define iteratively a sequence $\{x_n\}_{n=0}^\infty$ of functions $x_n:[0,1]\to[0,1]$. We first set $x_0\left(q\right):=\widetilde{x}_{\theta(q)}\left(1\right)$. The assumption that q_t is fixed at $q\left(\theta\right)$ sets the learning rate at a lower bound for what it can be in the future and the assumption that the payoff is given by $\max_{s\in[q(\theta),1]}\pi_\omega\left(\theta,s\right)$ in state ω sets the flow payoff at an upper bound for what it can be in the future. Since an increase in the learning rate or a decrease in the payoff flow can only make stopping less profitable, stopping at $(x, q(\theta))$ with $x < x_0(q(\theta))$ is a dominated strategy for type θ. Since the skimming property holds in any equilibrium, this implies that no player stops below curve $x_0(q)$ in equilibrium.

For $n \geq 1$, assume that $x_{n-1}(q)$ is a Lipschitz continuous and strictly increasing function with $x_{n-1}(q) \leq x^E(q)$ for all q and $x_{n-1}(1) = 1$ (note that these properties hold for $x_0(q)$ defined above). Define

$$x_n(q) := \widetilde{x}_{\theta(q)}(q'), \text{ where } q' = \min\{s : \widetilde{x}_{\theta(q)}(s) = x_{n-1}(s)\}. \tag{23}$$

The interpretation is that $x_n(q)$ would be the optimal stopping threshold for type $\theta(q)$ under the assumption that q_t is fixed at $q(\theta)$ and the flow payoff is given by $\max_{s \in [q(\theta), q']} \pi_\omega(\theta, s)$ in state ω. Here $q' \in (q(\theta), 1)$ is chosen in such a way that if no player ever stops below $x_{n-1}(q)$, then q' is an upper bound for q_t as long as $x_t \leq x_n(q)$. Hence $x_n(q(\theta))$ is a lower bound for the optimal stopping threshold for type θ under the assumption that no player stops below $x_{n-1}(q)$. It follows from (23) that if $x_{n-1}(q) < x^E(q)$, then $x_{n-1}(q) < x_n(q) \leq x^E(q)$, and if $x_{n-1}(q) = x^E(q)$, then $x_n(q) = x^E(q)$. Further, $x_n(q)$ inherits from $x_{n-1}(q)$ the properties of being Lipschitz continuous and strictly increasing function with $x_n(1) = 1$.

We have now shown by iterated deletion of dominated strategies that no player wants to stop at any (x, q) such that $x < x_n(q)$ for some n. The final step is to show that $x_n(q) \to x^E(q)$ as $n \to \infty$. By the construction, $\{x_n\}$ is a family of equicontinuous and strictly increasing functions. Therefore, the sequence $\{x_n\}_{n=0}^{\infty}$ converges uniformly to some function $\overline{x}(q)$, which is increasing and continuous. It also satisfies $\overline{x}(q) \leq x^E(q)$ for all q and $\overline{x}(1) = 1$.

It remains to show that $\overline{x}(q) = x^E(q)$. For contradiction, assume $\overline{x}(q) < x^E(q)$ for some $q < 1$. Being increasing, $\overline{x}(q)$ is differentiable almost everywhere. Since $x^E(q)$ is strictly increasing, we can then pick q in such a way that $\overline{x}'(q) > 0$ and $\overline{x}(q) < x^E(q)$. By (23), the difference $q' - q$ is bounded away from zero for any $x_{n-1}(q) < \overline{x}(q)$. Since $\overline{x}'(q) > 0$, the difference $x_n(q) - x_{n-1}(q)$ is then also bounded away from zero by (23). But this contradicts the assumption that $x_n(q)$ converges to $\overline{x}(q)$ from below as $n \to \infty$. We can conclude that $\overline{x}(q) = x^E(q)$

and so $x_n(q) \to x^E(q)$.

To better understand the previous step, let $x^E(q) - \overline{x}(q) > \epsilon$. Let L be the Lipschitz constant for $\widetilde{x}_{\theta(q)}$ and let L' be the Lipschitz constant for x_{n-1}. Then, $q' - q \geq \frac{x^E(q) - x_{n-1}(q)}{L+L'}$. Furthermore, we may assume that $x_{n-1}(q') - x_{n-1}(q) \geq x^E(q') - x^E(q)$. If the condition fails, one can simply use q' as q in the proof because then $x^E(q') - \overline{x}(q') > \epsilon$. Then, $x_{n-1}(q') - x_{n-1}(q) \geq k(q' - q)$ for a strictly positive k (the derivative of x^E is bounded away from 0). Putting all these together, yields $x_n(q) - x_{n-1}(q) = x_{n-1}(q') - x_{n-1}(q) \geq k\frac{x^E(q) - x_{n-1}(q)}{L+L'} > \frac{k}{L+L'}\epsilon$. Since this holds for all n, we cannot have $x_n(q) - x_{n-1}(q) \to 0$ as $n \to \infty$, contradicting that $x_n(q)$ converges to $\overline{x}(q)$.

Steps 1 and 2 together show that the only candidate for equilibrium is the boundary policy defined by $x^E(q)$.

B.2 Other proofs for Section 4

Proof of Proposition 6

Define $\tilde{q}(q)$ to be such that the informativeness of stock $\tilde{q}(q)$ under learning technology σ' is the same as the the informativeness of stock q under learning technology σ: $\tilde{q}(q) := \frac{\sigma'}{\sigma}q$. Similarly, let $\tilde{x} : [0,1] \to [0,1]$ be $\tilde{x}(\tilde{q}(q)) := x^E_\sigma(q)$. The interpretation of $\tilde{x}$ is that if $x^E_{\sigma'}(q) > \tilde{x}(q)$, then the agents require a higher belief to be willing to stop under σ' than under σ, given the same arrival rate of information.

Notice that the information flow under policy $\tilde{x}$ and technology σ' is exactly the same as the information flow under policy x^E_σ and technology σ. Next, use (10) to see that $x^E_{\sigma'}(\tilde{q}(q)) > x^E_\sigma(q)$ for all q: for all beliefs, the stock is lower under σ' than under σ. Now, because $\tilde{x}(\tilde{q}(q)) = x^E_\sigma(q)$, we also have $x^E_{\sigma'}(\tilde{q}(q)) > \tilde{x}(\tilde{q}(q))$, which implies that learning is slower under policy $x^E_{\sigma'}$ and technology σ' than under policy x^E_σ and technology σ for all $x \leq x^E_\sigma(1)$.

This means that agents who stop before the stock reaches $\tilde{q}(1)$ are worse off because they face both a slower increasing stock and slower arrival of information. As agents are homogeneous, this further implies that all agents are worse off because they are indifferent between stopping and waiting, and therefore the claim follows. □

C Socially optimal policy

We use the derivatives of $\Phi(x,q)$ in many proofs of this section:

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

Deriving the differential equation

We first show that the value matching and smooth pasting conditions, (7) and (8), imply the differential equation in (9). Solving (7) and (8) for $B_q(q)$ and $B(q)$ yields

$$B_q(q) = A^1(x^*(q),q)\,x^*(q) + A^2(x^*(q),q)\,, \tag{24}$$

$$B(q) = U^1(x^*(q),q)\,x^*(q) + U^2(x^*(q),q)\,, \tag{25}$$

where

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

Differentiating (25) with respect to q and using the chain rule gives

$$\begin{aligned}
B_q(q) = &\left[U_x^1(x^*(q),q)\,x^{*\prime}(q) + U_q^1(x^*(q),q)\right]x^*(q) + U^1(x^*(q),q)\,x^{*\prime}(q)\\
&+ U_x^2(x^*(q),q)\,x^{*\prime}(q) + U_q^2(x^*(q),q)
\end{aligned} \tag{26}$$

Equating (24) and (26), solving for $x^{*\prime}(q)$, and simplifying yields the expression (9) in the text.

Any solution that satisfies the differential equation (9) must be continuous.

C.1 Proof of Proposition 2

The proof contains three parts. In part 1, we show that the initial value problem (9) has a unique solution $x^*(q)$ with the property $x^*(q) < x^E(q)$ for all $q < 1$. We also show that $x^*(q)$ is continuous and strictly increasing and hence defines a boundary policy. In part 2, we show that our candidate policy $x^*(q)$ satisfies the HJB equation (5). In part 3, we verify that the solution to the HJB equation solves the original problem.

Part 1: solution to the initial value problem (9)

We first establish some key properties of function g in (9) (all proofs of the lemmas are in Appendix C.2):

Lemma 5. *For all (x, q) such that $q < 1$ and $x \leq x^E(q)$, function $g(x, q)$ in* (9) *is strictly positive and strictly increasing in x and it is Lipschitz continuous for all $q \in [0, q_1]$ if $q_1 < 1$ and for all $x \leq x^E(q)$. Furthermore, $g(x^E(q), q) > x^{E\prime}(q)$ for $q < 1$ and $\lim_{q\to 1} g(x^E(q), q) = x^{E\prime}(1)$.*

The singularity at (1,1) prevents us from directly applying the Picard-Lindelöf theorem to show the existence and uniqueness of a solution to the initial value problem (9). Instead, we note that the requirements for the Picard-Lindelöf theorem are satisfied for all initial conditions $x(q_1) = x_1$ where $x(q_1) \leq x^E(q_1)$ and $q_1 < 1$, and hence each such initial value problem defines a unique solution. Since g is increasing in x, these solutions diverge when approaching $(1, 1)$ and hence at most one path can approach $(1, 1)$ from below the decentralized policy. The fact that $\lim_{q\to 1} g(x^E(q), q) = x^{E\prime}(1)$ implies that there is a path that approaches $(1, 1)$ from the same direction as the decentralized policy $x^E(q)$ and the fact that $g(x^E(q), q) > x^{E\prime}(q)$ for $q < 1$ implies that such a path must be strictly below the decentralized solution for all $q < 1$. It follows that the initial value problem has a unique solution below the decentralized solution.

We have now shown that the initial value problem (9) has a unique solution x^* such that $x^*(q) \leq x^E(q)$ for all $x \leq q$. This solution $x^*(q)$ is continuous and strictly increasing in q, and it is our candidate policy.

Part 2: our candidate x^* solves the HJB equation

Fix $x^*(q)$ to be the candidate policy defined in the proposition and let $q^*(x)$ be its inverse with the convention $q^*(x) = 0$ for $x \leq x^*(0)$. Its associated value function is

$$V(x,q) = \begin{cases} \int_q^{q^*(x)} (xv_H(s) + (1-x)v_L(s))\,ds + V(x, q^*(x)), \text{ for } q < q^*(q) \\ B(q)\Phi(x,q), \text{ for } q \geq q^*(x), \end{cases} \tag{27}$$

where $B(q)$ is given by (25). By construction, for $q \geq q^*(x)$, $V(x,q)$ satisfies

$$rV(x,q) = \frac{1}{2}V_{xx}(x,q)\frac{x^2(1-x)^2}{\sigma^2}q \tag{28}$$

and at the boundary $q = q^*(x)$, the value matching and smooth pasting conditions (7) and (8) hold:

$$V_q(x, q^*(x)) + xv_H(q^*(x)) + (1-x)v_L(q^*(x)) = 0, \tag{29}$$

$$V_{qx}(x, q^*(x)) + v_H(q^*(x)) - v_L(q^*(x)) = 0. \tag{30}$$

Differentiating (28) with respect to q, we have

$$rV_q(x,q) = \frac{1}{2}V_{xx}(x,q)\frac{x^2(1-x)^2}{\sigma^2} + \frac{1}{2}V_{xxq}(x,q)\frac{x^2(1-x)^2}{\sigma^2}q, \tag{31}$$

which allows us to re-write (29) as:

$$r[xv_H(q^*(x)) + (1-x)v_L(q^*(x))] + \frac{1}{2}V_{xx}(x,q^*(x))\frac{x^2(1-x)^2}{\sigma^2} + \frac{1}{2}V_{xxq}(x,q^*(x))\frac{x^2(1-x)^2}{\sigma^2}q = 0. \tag{32}$$

We next state three lemmas that concern the partials of the value function below, above, and at the boundary, respectively. Their proofs are in a separate section, Appendix C.2.

Lemma 6. *For $q \geq q^*(x)$, we have $V_q(x,q) + xv_H(q) + (1-x)v_L(q) \leq 0$.*

Lemma 7. *For $q < q^*(x)$, we have $V_{xx}(x,q) = V_{xx}(x, q^*(x))$, $V_{qq}(x,q) = V_{qq}(x, q^*(x))$, and $V_{xxq}(x,q) = 0$.*

Lemma 8. *For $q = q^*(x)$, we have $V_{xxq}(x,q) < 0$.*

We are now ready to show that our candidate policy satisfies the HJB-equation (5), which we re-write here using notation q' instead of q^* for the maximizer (this is to avoid confusion with boundary $q^*(x)$):

$$rV(x,q) = \max_{q'>q} \left(r \int_q^{q'} (xv_H(s) + (1-x)v_L(s))\, ds + \frac{1}{2} V_{xx}(x,q') \frac{x^2(1-x)^2}{\sigma^2} q' \right). \tag{33}$$

The right-hand side of this equation is a continuous function in q' and its derivative with respect to q' is

$$r(xv_H(q') + (1-x)v_L(q')) + \frac{1}{2} V_{xx}(x,q') \frac{x^2(1-x)^2}{\sigma^2} + \frac{1}{2} V_{xxq}(x,q') \frac{x^2(1-x)^2}{\sigma^2} q'. \tag{34}$$

Let us inspect the sign of this for different values of q'. For $q' \geq q^*(x)$, we can use (31) to write (34) as

$$r(xv_H(q') + (1-x)v_L(q')) + rV_q(x,q),$$

which is negative by lemma 6. It follows that whenever $q \geq q^*(x)$, the right-hand side of (33) is maximized by choosing $q' = q$, i.e. keeping q fixed.

For $q' < q^*(x)$, we can use lemma 7 to write (34) as

$$r(xv_H(q') + (1-x)v_L(q')) + \frac{1}{2} V_{xx}(x, q^*(x)) \frac{x^2(1-x)^2}{\sigma^2},$$

which is decreasing in q'. Moreover, combining (32) and Lemma 8 we can conclude that it is positive in the limit $q' \to q^*(x)$, and hence it is positive for all $q' < q^*(x)$. Since the right-hand side of (33) is continuous, and its derivative is positive (negative) for $q' < q^*(x)$ ($q' \geq q^*(x)$), it is maximized at $q' = q^*(x)$ if $q < q^*(x)$.

We have now shown that for any $x \in (0,1)$, the right-hand side of the HJB equation is maximized by choosing $q' = \max\{q, q^*(x)\}$. Furthermore, since $V(x,q)$ satisfies (28) for $q \geq q^*(x)$, the left- and right-hand sides of (33) coincide with this choice of q'. Hence, we have shown that $V(x,q)$ defined in (27) satisfies the HJB-equation.

Part 3: verification

The verification of the solution follows from the standard arguments in the literature (see e.g. Fleming and Soner (2006)). Let V^* solve the HJB equation (5) and let $q^*(x,q) = \max\{q, q^*(x)\}$ be the corresponding q^*. Then, let $T \geq t$ be the time at which the candidate value function is evaluated. From generalized Itô's formula we have[10]

$$
\begin{aligned}
e^{-rT}V^*(x_T,q_T) &= e^{-rt}V^*(x_t,q_t) - \int_t^T e^{-rs} rV^*(x_s,q_s)ds + \int_t^T e^{-rs}V_x^*(x_s,q_s)dx_s \\
&+ \int_t^T e^{-rs}V_q^*(x_s,q_s)dq_s + \frac{1}{2}\int_t^T e^{-rs}V_{xx}^*(x_s,q_s)d[x]_s + \frac{1}{2}\int_t^T e^{-rs}V_{qq}^*(x_s,q_s)d[q]_s \\
&+ \int_t^T e^{-rs}V_{qs}^*(x_s,q_s)d[q,x]_s
\end{aligned}
$$

where $d[x]_t$ and $d[q]_t$ are the quadratic variations of x and q and $d[x,y]_t$ is their quadratic covariation. The process Q_t has bounded variation and hence $d[q]_t = d[x,y]_t = 0$. Notice also that $dx_t = x_t(1-x_t)\sigma^{-1}\sqrt{q_t}dw_t$ and $d[x]_t = x_t^2(1-x_t)^2\sigma^{-2}q_t dt$. We can further simplify the equation by noting that $V_q^* dq = -(xv_H(q) + (1-x)v_L(q))dq$. The HJB equation gives an upper bound for $\frac{q_s}{\sigma^2}x_s^2(1-x_s)^2V_{xx}^*(x_s,q_s) - rV^*(x_s,q_s) \leq \int_{q_s}^{q^*(x_s,q_s)}(xv_H(q)+(1-x)v_L(q))dq$, which equals zero for almost all s. Combining gives:

$$
\begin{aligned}
e^{-rT}V^*(x_T,q_T) &\leq e^{-rt}V^*(x_t,q_t) - \int_t^T e^{-rs}(x_s v_H(q_s) + (1-x_s)v_L(q_s)))dq_s \\
&+ \int_t^T e^{-rs}V_x{}^*(x_s,q_s)\frac{\sqrt{q_s}}{\sigma}x_s(1-x_s)dw_s.
\end{aligned}
$$

Taking conditional expectations , multiplying by $-e^{rt}$ and simplifying then gives

$$
V^*(x_t,q_t) \geq \mathbb{E}\left[\int_t^T e^{-r(t-s)}(x_s\pi_H(q_s) + (1-x_s)\pi_L(q_s))ds + e^{-r(T-t)}V^*(x_T,q_T)|\mathcal{F}_t\right].
$$

The candidate value function is bounded and therefore clearly satisfies the transversality condition: $\lim_{T\to\infty}\mathbb{E}[e^{-r(T-t)}V^*(x_T,q_T)] = 0$. Hence, taking the limit $T \to \infty$ gives that $V^*(x,q) \geq \max_Q U(Q;x,q)$.

[10] To see that $V \in C^2$ check V_x at the boundary. The continuity of V_{xx} and V_{qq} follows from Lemma 7 and the continuity of V_q and V_{qx} are implied by the value matching and smooth pasting conditions.

The last step is to use the fact that Q, induced by policy x^*, achieves the pointwise maximum of the HJB-equation and thus the inequalities above become equalities: $V^*(x,q) = \max_Q U(Q;x,q)$. Our solution solves the original problem.

C.2 Proof of Lemmas 5, 6, 7, and 8

Proof of Lemma 5. Taking the derivative of $g(x,q)$ with respect x gives:

$$\begin{aligned} g_x(x,q) = &-\Big[\beta''(q)\Big(x^2(1-2x)(\beta(q)-1)^3 v_H(q)^2 - 2(1-x)x\beta(q)(\beta(q)-1) \\ &\times v_H(q)v_L(q)((1-2x)\beta(q)-x) + (1-x)^2(1-2x)\beta(q)^3 v_L(q)^2\Big) \\ &+ \beta'(q)\Big(2x^2(2x-1)(\beta(q)-1)^2 v_H(q)^2\beta'(q) + (1-x)^2\beta(q)^2 v_L(q) \\ &\times \Big(2(1-2x)v_L(q)\beta'(q) - 2x(\beta(q)-1)v_H'(q) - (1-2x)\beta(q)v_L'(q)\Big) \\ &+ xv_H(q)\Big(4(1-x)v_L(q)\beta'(q)\left((1-2x)\beta(q)^2 + 2x\beta(q) + x\right) \\ &- x(\beta(q)-1)^2\Big((1-2x)(\beta(q)-1)v_H'(q) + 2(1-x)\beta(q)v_L'(q)\Big)\Big)\Big)\Big]/ \\ &\Big[(x(\beta(q)-1)^2 v_H(q) + (1-x)(\beta(q))^2 v_L(q))^2\beta'(q)\Big]. \end{aligned}$$

Both $g(x,q)$ and $g_x(x,q)$ are bounded if their denominators are bounded away from zero. We show that this is true if $q < 1$ and $x \leq x^E(q)$ by showing that it hold at $x = x^E$:

$$x^E(q)(\beta(q)-1)^2 v_H(q) + (1-x^E(q))(\beta(q))^2 v_L(q) < 0, \tag{35}$$

for all $q \in [0,1)$. Notice that the left-side is increasing in x and hence (35) implies the same inequality for all lower x. The condition (35) is equivalent with

$$\frac{\beta(q)(\beta(q)v_L(q) - (\beta(q)-1)v_H(q))}{\beta(q)^2 v_L(q) - (\beta(q)-1)^2 v_H(q)} > 1 \iff (\beta(q)-1)v_H(q) > 0,$$

which holds as $\beta(q) > 1$ and $v_H(q) > 0$. We can conclude that g and g_x are bounded and continuous in both x and q for all (x,q) such that $q < 1$ and $x \leq x^E(q)$ and hence g is Lipschitz when we gap q away from 1. The denominator of $g(x,q)$ is strictly positive.

To see that $g(x,q) > 0$, it is now enough to show that the numerator of (9) is strictly positive. First notice that the second term inside the brackets is always positive but the first term can be negative.[11] The first term is scaled by x, while the second therm is scaled by $(1-x)$. Therefore, if the numerator is positive at a belief above the boundary, it must be positive for the belief at the boundary as well. Since the decentralized belief, $x^E(q)$, is always above the fully optimal boundary, we can use it to show that the numerator is positive.

Plugging in $x^E(q)$ to the numerator of (9) and dividing by $x(1-x)$ gives:

$$\frac{\beta(q)v_L(q)\left(\beta'(q)\left(\beta(q)-1\right)v_H'(q)-\left(\left(\beta(q)-1\right)\beta''(q)-2\left(\beta'(q)\right)^2\right)v_H(q)\right)}{\beta(q)v_L(q)+(1-\beta(q))v_H(q)}$$
$$+\frac{\left(1-\beta(q)\right)v_H(q)\left(\beta'(q)\beta(q)v_L'(q)-\left(\beta(q)\beta''(q)-2\left(\beta'(q)\right)^2\right)v_L\left(q\right)\right)}{\beta\left(q\right)v_L\left(q\right)+\left(1-\beta\left(q\right)\right)v_H\left(q\right)}.$$

Since the denominator is negative ($v_L<0$ and $\beta>1$), this is proportional to

$$[v_H(q)v_L'(q)-v_H'(q)v_L(q)]\beta'(q)\beta(q)(\beta(q)-1)-2v_H(q)v_L(q)(\beta'(q))^2,$$

which is always positive because $v_H(q)>0$ and $v_L(q),v_H'(q),v_L'(q)<0$. Hence, $q(x,q)>0$ for all $q\in[0,1)$ and $x\le x^E(q)$.

Similar direct calculations show that $g_x>0$ for all (x,q) such that $q<1$ and $x\le x^E(q)$.

Next, insert $x^E(q)$ to (9):

$$\begin{aligned}g(x^E(q),q)=&\frac{\frac{-\beta(1-\beta)v_Lv_H}{(\beta v_L+(1-\beta)v_H)^2}}{\frac{\beta'\beta(1-\beta)v_Lv_H}{\beta v_L+(1-\beta)v_H}}\Bigg(\frac{\beta'\beta(1-\beta)(v_Lv_H'-v_L'v_H)}{\beta v_L+(1-\beta)v_H}\\&+\frac{\beta v_L\gamma_H(-2\beta'^2+(\beta-1)\beta'')}{\beta v_L+(1-\beta)v_H}+\frac{(\beta-1)v_Lv_H(-2\beta'^2+\beta\beta'')}{\beta v_L+(1-\beta)v_H}\Bigg)\\=&\frac{v_H\left(2v_L\beta'-(\beta-1)\beta\gamma_L'\right)+(\beta-1)\beta v_Lv_H'}{((\beta-1)v_H-\beta v_L)^2}.\end{aligned}$$

The derivative of the decentralized policy x^E is

$$x^{E\prime}(q)=\frac{v_H\left(v_L\beta'-(\beta-1)\beta v_L'\right)+(\beta-1)\beta v_Lv_H'}{((\beta-1)v_H-\beta v_L)^2}.$$

By subtracting $x^{E\prime}(q)$ from $g(x^E(q),q)$, we get

$$g(x^E(q),q)-x^{E\prime}(q)=\frac{\beta'(q)v_L(q)v_H(q)}{(\beta(q)v_L(q)+(1-\beta(q))v_H(q))^2}.$$

This expression is strictly positive for $q<1$ and goes to zero as q goes to 1 (since $v_H(q)\to 0$). □

[11] This follows from $v_L(q)<0, v_\omega'(q)<0, \beta'(q)<0, \beta(q)>1$ and that $\beta(q)\beta''(q)>2(\beta'(q))^2$.

Proof of Lemma 6. If the claim is not true, there must be some x and $q > q^*(x)$ such that

$$V_q(x, q) + xv_H(q) + (1 - x)v_L(q) > 0. \tag{36}$$

We show that this leads to a contradiction by showing that (36) implies $V_{qx}(x, q) + v_H(q) - v_L(q) > 0$, which further implies that (36) holds also for all beliefs in $[x, x^*(q)]$, including $V_q(x^*(q), q) + x^*(q)v_H(q) + (1 - x^*(q))v_L(q) > 0$, which contradicts the value matching condition (29).

It remains to show that (36) implies $V_{qx}(x, q) + v_H(q) - v_L(q) > 0$. First notice that $V_q(x, q) = B_q(q)\Phi(x, q) + B(q)\Phi_q(x, q)$, which then together with (36) implies

$$B_q > -\frac{\Phi_q}{\Phi}B - \frac{xv_H + (1 - x)v_L}{\Phi}$$

where we have left out all dependencies to simplify notation. We now get the following lower bound:

$$\begin{aligned} V_{qx} + v_H - v_L &= B_q\Phi_x + B\Phi_{qx} + v_H - v_L \\ &> -\frac{\Phi_q\Phi_x}{\Phi}B - \frac{\Phi_x}{\Phi}(xv_H + (1 - x)v_L) + B\Phi_{qx} + v_H - v_L \\ &= \Phi^{-1}[B(\Phi_{qx}\Phi - \Phi_q\Phi_x) + \Phi(v_H - v_L) - \Phi_x(xv_H + (1 - x)v_L)]. \end{aligned} \tag{37}$$

The first term can be simplified as

$$\begin{aligned} &\Phi^{-1}B(\Phi_{qx}\Phi - \Phi_q\Phi_x) = \frac{B\Phi\beta'}{x(1-x)} = \frac{\Phi\beta'}{x(1-x)}\frac{\Phi^*_x(x^*v_H + (1 - x^*)v_L) - \Phi^*(v_H - v_L)}{\Phi^*_{qx}\Phi^* - \Phi^*_q\Phi^*_x} \\ &= \frac{x^*(1-x^*)}{x(1-x)}\frac{\Phi}{\Phi^*\Phi^*}[\Phi^*_x(x^*v_H + (1 - x^*)v_L) - \Phi^*(v_H - v_L)], \end{aligned}$$

where the notation Φ^* refers to $\Phi(x^*(q), q)$.

Now, (37) becomes

$$\begin{aligned} &\frac{x^*(1-x^*)}{x(1-x)}\frac{\Phi}{\Phi^*\Phi^*}[\Phi^*_x(x^*v_H + (1 - x^*)v_L) - \Phi^*(v_H - v_L)] \\ &- \frac{1}{\Phi}[\Phi_x(xv_H + (1 - x)v_L) - \Phi(v_H - v_L)] \\ &= \frac{1}{x(1-x)}\Big(\frac{\Phi}{\Phi^*}((\beta - 1)x^*v_H + \beta(1 - x^*)v_L) - ((\beta - 1)xv_H + \beta(1 - x)v_L)\Big), \end{aligned} \tag{38}$$

where we have used the following for both terms inside the brackets:

$$\Phi(v_H - v_L) - \Phi_x(xv_H + (1-x)v_L) = \Phi(v_H - v_L) - \Phi\frac{\beta - x}{x(1-x)}(xv_H + (1-x)v_L)$$
$$= \frac{-\Phi}{x(1-x)}((\beta-1)xv_H + \beta(1-x)v_L).$$

To conclude that (38) is larger than 0, notice first that $(\beta-1)xv_H + \beta(1-x)v_L < 0$ whenever $x < x^E(q)$ and that it is increasing in x. Then observe that $\Phi/\Phi^* \in (0,1)$ and hence $(\beta-1)xv_H + \beta(1-x)v_L < (\Phi/\Phi^*)((\beta-1)x^*v_H + \beta(1-x^*)v_L)$.

We conclude that $V_q + xv_H + (1-x)v_L > 0$ implies $V_{qx} + v_H - v_L > 0$ and the proof is complete. □

Proof of Lemma 7. Fixing some (x,q) such that $q < q^*(x)$, differentiating (27) twice with respect to x, and simplifying gives:

$$\begin{aligned} V_{xx}(x,q) &= V_{xx}(x,q^*(x)) \\ &+2(q^*)'(x)(V_{qx}(x,q^*(x)) + v_H(q^*(x)) - v_L(q^*(x))) \\ &+(q^*)''(x)(V_q(x,q^*(x)) + xv_H'(q^*(x)) + (1-x)v_L'(q^*(x))) \\ &+\left((q^*)'(x)\right)^2(V_{qq}(x,q^*(x)) + xv_H'(q^*(x)) + (1-x)v_L'(q^*(x))). \end{aligned} \tag{39}$$

The second term on the right-hand side vanishes by the value-matching condition (29) and the third term vanishes by the smooth-pasting condition (30). Let us look at the last term. First, since (29) holds along the boundary $(x, q^*(x))$, we can totally differentiate it with respect to x to get:

$$\begin{aligned} 0 &= V_{qx}(x,q^*(x)) + V_{qq}(x,q^*(x))(q^*)'(x) + v_H(q^*(x)) - v_L(q^*(x)) \\ &+[xv_H'(q^*(x)) + (1-x)v_L'(q^*(x))](q^*)'(x). \end{aligned}$$

Applying (30), several terms disappear and this reduces to

$$V_{qq}(x,q^*(x)) + xv_H'(q^*(x)) + (1-x)v_L'(q^*(x)) = 0.$$

The last term in (39) vanishes as well, and it follows that $V_{xx}(x,q) = V_{xx}(x,q^*(x))$. Since this holds for any $q < q^*(x)$, it immediately implies that $V_{xxq}(x,q) = 0$. □

Proof of Lemma 8. This is by direct computation. Recall that the value function for $q \geq q^*(x)$ is $V(x,q) = B(q)\Phi(x,q)$ and hence

$$V_{xxq} = B_q(q)\Phi_{xx}(x,q) + B(q)\Phi_{xxq}(x,q).$$

Plugging in the expressions for $B_q(q)$ and $B(q)$ from (24) and (25), multiplying by r, simplifying, and evaluating at $q = q^*(x)$ gives:

$$\begin{aligned} rV_{xxq}\left(x, q^*\left(x\right)\right) &= \frac{\left(\beta\left(q^*\left(x\right)\right) - 1\right)^2 + x\left(1-x\right)}{x^2\left(1-x\right)^2} v_L\left(q^*\left(x\right)\right) \\ &\quad - \frac{\beta\left(q^*\left(x\right)\right)\left(\beta\left(q^*\left(x\right)\right) - 1\right)}{x^2\left(1-x\right)^2}\left(v_H\left(q^*\left(x\right)\right) - v_L\left(q^*\left(x\right)\right)\right). \end{aligned}$$

Noting that $\beta\left(q^*\left(x\right)\right) > 1$, $v_L\left(q^*\left(x\right)\right) < 0$, and $v_H\left(q^*\left(x\right)\right) - v_L\left(q^*\left(x\right)\right) > 0$, it follows that $V_{xxq}\left(x, q^*\left(x\right)\right) < 0$.

□

C.3 Proof of Proposition 3

Proof. First, recall that $x^*(0) < x^E(0) = x^{stat}(0)$ by the proof of Proposition 2. Using this together with the continuity of the policy functions we find that there exists $\underline{q} > 0$ such that $x^{stat}(q) > x^*(q)$ for all $q < \underline{q}$. As the policy functions are strictly increasing and continuous, the stocks $q^*(x)$ and $q^{stat}(x)$ are pinned down as the inverse of the policy functions for all $x \geq x^*(0)$ and $x \geq x^{stat}(0)$ respectively. In addition, $q^*(x) = 0$ for all $x \leq x^*(0)$ and $q^{stat}(x) = 0$ for all $x \leq x^{stat}(0)$, and hence q^* and q^{stat} are continuous.

Let $\underline{x} := x^{stat}(\underline{q}) > x^{stat}(0)$ where the inequality follows from x^{stat} being strictly increasing. Then, $q^{stat}(x) < q^*(x)$ for all $x \in [x^{stat}(0), \underline{x})$ by that q^* and q^{stat} are the inverse functions of x^* and x^{stat}. Furthermore, $q^{stat}(x) = 0 < q^*(x)$ for all $x \in [x^*(0), x^{stat}(0)]$, which completes the proof.

Next, we show the other direction by showing that $x^*(1) = x^{stat}(1) = 1$ and $x^*_q(1) < x^{stat}_q(1)$. The first part is immediate. For the second part, use Lemma 5 and the uniqueness of the solution to get $x^*_q(1) = x^E_q(1)$. Now it is enough to show that the derivative of the equilibrium is smaller than of the static solution:

$${x^E}_q(1) - {x^{stat}}_q(1) = \frac{(\beta-1)\beta v_L v'_H}{(\beta v_L)^2} - \frac{v_L v'_H}{(v_L)^2} = -\frac{v_L v'_H}{\beta v_L^2} < 0.$$

The static and optimal solutions meet at $q = 1$ but the optimal solution reaches the point above the static solutions. Hence, by continuity there must exist $\overline{q} < 1$ such that $x^*(q) > x^{stat}(q)$ for all $q \in (\overline{q}, 1)$, which then further implies the existence of $\overline{x} < 1$ by the same argument as used above for $\underline{x}$. □

C.4 Proof of Proposition 4

Proof. Part (a): We show the result by contradiction. By using the solution from Proposition 2 and the value function derived in its proof, we show that $q^* = 0$ cannot maximize the HJB equation (5) in the limit as $\sigma \to 0$ unless $\sqrt{q^*_\sigma(x)}/\sigma \to \infty$. If $q^*(x)$ goes to any other value than 0, the claim immediately follows.

By taking the first order condition from (5), we get

$$xv_H(q^*) + (1-x)v_L(q^*) + \frac{1}{2}\frac{x^2(1-x)^2}{\sigma^2}(V_{xx}(x, q^*) + V_{xxq}(x, q^*)q^*).$$

The first order condition is necessarily strictly positive at $q^* = 0$ in the limit as $\sigma \to 0$ once we show that $V_{xx}(x, q) > 0$ and $V_{xxq}(x, q)$ is finite.

Recall that the value function is $V(x, q) = B(q)\Phi(x, q)$ and its derivatives are then $V_{xx} = B(q)\Phi_{xx}$ and $V_{xxq} = B_q(q)\Phi_{xx} + B(q)\Phi_{xxq}$. By plugging in the values of Φ_{xx}, we get

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

We know that $B > 0$ for all $q < 1$ in the optimal solution and that $\Phi > 0$ for all $x \in (0, 1)$. Then, $V_{xx} > 0$ whenever $\beta > 1$ which is true whenever the signal-to-noise ration is finite.

We can write V_{xxq} as

$$V_{xxq} = \frac{(\Phi_x\Phi_{xxq} - \Phi_{qx}\Phi_{xx})(xv_H + (1-x)v_L)}{\Phi\Phi_{qx} - \Phi_q\Phi_x} + \frac{(\Phi_q\Phi_{xx} - \Phi\Phi_{xxq})(v_H - v_L)}{\Phi\Phi_{qx} - \Phi_q\Phi_x}.$$

The first term equals $\frac{(\beta-x)^2+x(1-x)}{x^2(1-x)^2}(xv_H + (1-x)v_L)$ and the second term equals $-\frac{\beta+(\beta-1)ln\left(\frac{x}{1-x}\right)}{x(1-x)}(v_H - v_L)$. Both are finite for all $x \in (0, 1)$.

Hence, we conclude that for the first order condition to be satisfied, we must have $\sqrt{q^*_\sigma(x)}/\sigma \to \infty$ as $\sigma \to 0$.

Part (b): We fix the belief to be $x \in (0,1)$. By rearranging the solution in Proposition 1, we get

$$\beta(q) = \frac{xv_H(q)}{xv_H(q) + (1-x)v_L(q)}.$$

We take the limit $\lim_{\sigma \to 0} \beta(q^E_\sigma(x)) = \frac{xv_H(0)}{xv_H(0)+(1-x)v_L(0)}$, which is strictly larger than 1 for all $x > x^{stat}(q)$ and hence further implying that $\lim_{\sigma \to 0} \sqrt{q^E_\sigma(x)}/\sigma < \infty$. More precisely, we get the limit of the signal-to-noise ratio as $a(x)$ satisfying $\frac{xv_H(0)}{xv_H(0)+(1-x)v_L(0)} = \frac{1}{2}\left(1 + \sqrt{1 + 8ra(x)^{-2}}\right)$. □

Supplementary material: applications and extensions

Next, we present the formal analysis of the topics discussed in Section 5. We cover the models and the results separately for each topic in Sections D, E, and F. All longer proofs are then located in separate Section G.

D Mechanism design

In this section, we bridge the gap between individual optimization and the socially optimal policy by analyzing how a designer can implement policies – including the socially optimal policy – with transfers. We present our analysis in steps starting with static implementation before moving on to dynamic implementation. We conclude by analyzing revenue maximizing mechanisms.[12]

D.1 Static implementation by a direct mechanism

As a building block towards dynamic implementation, we first consider static direct implementation. This means that all transactions between the designer

[12]This section is related to the literature on dynamic mechanism design, especially in the context of optimal stopping. Board (2007), Kruse and Strack (2015) and Board and Skrzypacz (2016) analyze implementation of stopping rules for agents, whose private types evolve exogenously over time. In contrast, we analyze implementation in a game with externalities but where private types are fixed over time.

and the agents take place at time 0: the designer announces a policy Q to be implemented and offers a menu $\{\tau(\theta), P_0(\theta)\}_{\theta\in\Theta}$, where an agent who reports θ pays an upfront transfer $P_0(\theta)$ and gets as an allocation an obligation to stop at time $\tau(\theta) = \inf\{t : q_t = 1 - F(\theta)\}$.

We now outline the steps to pin down transfers that implement a given boundary policy Q and the associated monotone stopping profile as a Bayes-Nash equilibrium.[13] Let $W_0(\theta)$ be the time-0 value of an agent θ under policy Q. Requiring truthful reporting to be optimal, we can use the envelope theorem of Milgrom and Segal (2002) to write out an agent's value (and setting $W_0(\underline{\theta}) = 0$):

$$W_0(\theta) = \int_{\underline{\theta}}^{\theta} W_0'(s)ds = \mathbb{E}\left[\int_{\underline{\theta}}^{\theta} e^{-r\tau(s)}\Big(x_{\tau(s)}v_H'(s) + (1 - x_{\tau(s)})v_L'(s))\Big)ds\right],$$

where $\tau(\theta) = \inf(t : q_t = 1 - F(\theta))$ and the expectation is taken over belief process X induced by policy Q. An *ex-ante* transfer rule, $P_0 : [\underline{\theta}, \overline{\theta}] \to \mathbb{R}$, where each type θ is assigned a given stopping rule at time 0, is hence pinned down as

$$\begin{aligned} P_0(\theta) &= \mathbb{E}\Big[e^{-r\tau(\theta)}(x_{\tau(\theta)}v_H(\theta) + (1 - x_{\tau(\theta)})v_L(\theta))\Big] \\ &\quad -\mathbb{E}\left[\int_{\underline{\theta}}^{\theta} e^{-r\tau(s)}\Big(x_{\tau(s)}v_H'(s) + (1 - x_{\tau(s)})v_L'(s))\Big)ds\right], \end{aligned} \tag{40}$$

an expression that is easy to compute numerically for a fixed Q.

To finish the argument, we have to check that incentive compatibility holds globally. This can be done by noting that we associate policy Q with a monotone stopping profile where higher type agents always stop before lower type agents. We show formally in Appendix G.1 that this implies an increasing differences property that guarantees global incentive compatibility.

Lemma 9. *Given a boundary policy Q, a direct mechanism $\{\tau(\theta), P_0(\theta)\}_{\theta\in\Theta}$, where an agent reporting type θ pays transfer $P_0(\theta)$ defined in (40) and commits to stopping at time $\tau(\theta) = \inf\{t : q_t = 1 - F(\theta)\}$, satisfies incentive compatibility and participation constraints for every type.*

[13] This ex-ante implementation applies even if the policy Q is not a boundary policy.

D.2 Dynamic implementation by posted prices

In the previous section, we reduced the dynamic problem into a static one by assuming that a report made at time 0 involves a commitment to follow a pre-specified stopping rule. However, committing to such behavior is not likely to be feasible in practice. We next investigate how to implement the same policy by a posted price that postpones the transaction and type revelation to the moment when the agent stops.

As a preliminary step, consider a restricted dynamic implementation where the agents are only allowed to choose amongst the stopping times $\{\tau(\theta)\}_{\theta\in\Theta}$ where the state hits boundary points $(\widetilde{x}\,(q)\,,q)$ for some q. Denote by $P\,(\widetilde{x}\,(q)\,,q)$ the payment requested at the moment of stopping from an agent who chooses to stop at the boundary point $(\widetilde{x}\,(q)\,,q)$. To guarantee that agent $\theta(q)$ wants to stop exactly at $(\widetilde{x}\,(q)\,,q)$ rather than some other boundary point $(\widetilde{x}\,(q')\,,q')$, we can design the payment $P\,(\widetilde{x}\,(q)\,,q)$ so that from an ex-ante perspective it replicates the corresponding static payment $P_0\,(\theta\,(q))$ for type $\theta\,(q)$:

$$
\begin{aligned}
&P\,(\widetilde{x}\,(q)\,,q) = \widetilde{x}(q)(v_H(\theta(q)) + (1-\widetilde{x}(q))v_L(\theta(q))) \\
&\quad - \mathbb{E}\Big[\int_{\underline{\theta}}^{\theta(q)} e^{-r(\tau(s)-\tau(\theta(q)))}\Big(\widetilde{x}_{\tau(s)}v_H'(s) + (1-\widetilde{x}_{\tau(s)})v_L'(s)\Big)ds\Big|x(q),q\Big]. \qquad (41)
\end{aligned}
$$

Since this satisfies

$$
\mathbb{E}\Big[e^{-r\tau(\theta)}P\,(\widetilde{x}(q),q)\,\Big|x_0,q_0\Big] = P_0(\theta(q)),
$$

an agent who intends to stop at $(\widetilde{x}\,(q)\,,q)$ is ex-ante indifferent between paying $P_0\,(\theta(q))$ at time zero and paying $P\,(\widetilde{x}\,(q)\,,q)$ at the moment of stopping. As long as the agents are not allowed to stop outside of the stopping boundary, incentive compatibility continues to hold in this restricted dynamic implementation.

It remains to be shown that the agents do not have an incentive to stop below the boundary when allowed to do so.[14] We consider here two natural transfer schemes that coincide with $P(\widetilde{x}(q),q)$ at the boundary but differ below it.

[14] A trivial but non-practical way to do this is to fix the transfer payment to be arbitrarily high whenever $x<\widetilde{x}\,(q)$ making the cost of stopping prohibitive below the boundary.

By a *dynamic posted price*, we refer to a transfer that is a function of the current belief and hence continuously responds to news about the state. The dynamic posted price, $P^D(x)$, is fully pinned down by (41) as the two transfer schemes must coincide at boundary points:

$$P^D(x) := \begin{cases} P(x, \widetilde{q}(x)) & \text{for } x \geq \widetilde{x}(0) \\ P(\widetilde{x}(0), 0) & \text{for } x < \widetilde{x}(0) \end{cases} \tag{42}$$

where $\widetilde{q}(x) : [\widetilde{x}(0), 1] \longrightarrow [0, q_1]$ is the inverse of $\widetilde{x}(\cdot)$. We allow for the possibility that the designer wants to implement a restricted maximal stock, i.e. $q_1 < 1$.

If agent θ stops at state (x, q), his stopping payoff is

$$u_\theta^D(x) = x v_H(\theta) + (1 - x) v_L(\theta) - P^D(x).$$

The term $P^D(x)$ makes the optimal stopping problem more complicated than the corresponding problem in the context of the decentralized equilibrium. Yet, without even explicitly solving the individual agents' stopping problems we will prove below that this pricing rule is immune to deviations to stopping in state (x, q) that is not on the intended boundary and hence it dynamically implements the intended policy.

When using the dynamic posted price rule (42), the designer needs to continuously observe the news process and carry out detailed Bayesian calculation to adjust the transfer. To make the job of the designer easier and the commitment assumption more palatable we consider as an alternative *simple posted prices* that depend only on the stock q. For example, a seller of a new durable good could set the price based on the cumulative past sales instead of reviews or other feedback from past buyers. As with the dynamic posted price, we set the simple posted price $P^S(q)$ to coincide with (41) at boundary points:

$$P^S(q) := P(\widetilde{x}(q), q) \text{ for } q \in [0, q_1]. \tag{43}$$

Now agents face an optimal stopping problem where the stopping payoff depends on q as well as x:

$$u_\theta^S(x, q) = x v_H(\theta) + (1 - x) v_L(\theta) - P^S(q).$$

An important property of this scheme is that the stopping payoff changes abruptly at the boundary. Whether the transfer is increasing or decreasing turns out to be critical for providing sufficient stopping incentive at the boundary without inviting deviations to stop below it. If the transfer is increasing, even a slight postponement would make stopping more expensive for the agent at the boundary, providing an additional incentive to stop at the boundary relative to the states below it. However, if the transfer is decreasing, stopping at the boundary is less attractive as a delay would be rewarded with a reduction in transfer payment. As a result, if the decreasing simple posted price scheme is such that an agent is willing to stop at the boundary, he wants to stop even before reaching it.

We summarize our findings in the proposition below:

Proposition 7. *Let Q denote a boundary policy with a strictly increasing policy function $\widetilde{x}$ and $\widetilde{x}(q_1) = 1$. Then:*

- *Q can be implemented by a dynamic posted price $P^D(x)$ in (42).*
- *Q can be implemented by a simple posted price $P^S(q)$ in (43) if and only if $\left(P^S\right)'(q) \geq 0$ for all $q \in [0, q_1]$.*

The proof of Proposition 7 is in Appendix G.2. The key issue is to rule out deviations to stop too early below the intended stopping boundary. We show that we can always rule such deviations out for the dynamic posted price, but in the case of the simple posted price they are ruled out if and only if $P^S(q)$ is everywhere increasing in q. As a part of the next subsection, we give an example where $P^S(q)$ is not everywhere increasing.

D.3 Socially optimal transfers

Now we are ready to connect decentralized optimization and the socially optimal policy and demonstrate the properties of socially optimal transfers. The left panel of Figure 8 depicts the socially optimal boundary policy $\widetilde{x}$ for stopping payoffs $v_H(q) = 1 - q$, $v_L(q) = -\eta - q$, where the three cases correspond to different values for parameter η which increases the cost of stopping in the low state. The right panel shows the corresponding socially optimal transfer function $P(\widetilde{x}(q), q)$ as a function of q. We see that $P(\widetilde{x}(q), q)$ is always negative: the designer pays each agent so that they internalize the information generation effect. The optimal transfer goes smoothly to zero when q approaches 1 because the information generation effect disappears and the decentralized and the optimal policies coincide even without transfers. These properties hold generally for the socially optimal transfer.

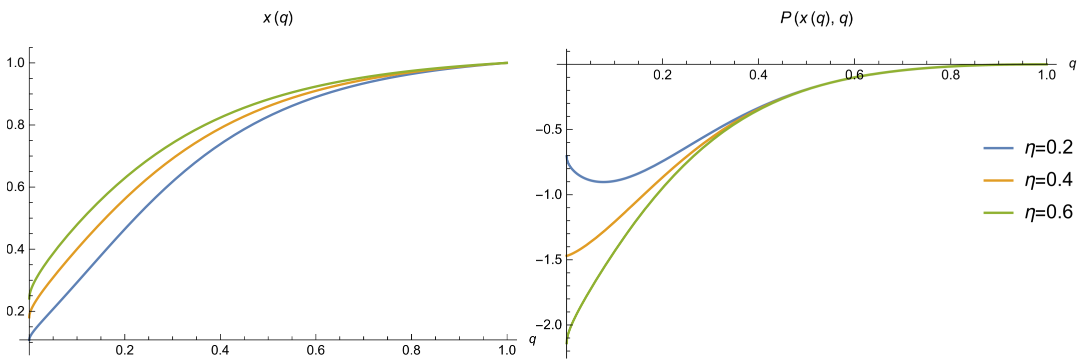

Figure 8: The left panel shows the socially optimal policy for different values of η. The right panel shows the corresponding transfers. $v_H(q) = 1 - q, v_L(q) = -\eta - q, r = 0.1, \sigma = 1$.

We know from Proposition 7 that we can always use a dynamic posted price to implement the policy, but a simple posted price works only if $P^S(q) := P(\widetilde{x}(q), q)$ is monotone. In this example, the optimal policy can only be implemented with simple posted prices when the risk parameter is large enough. That is, transfers are non-monotone when $\eta = 0.2$, but motone when $\eta = 0.4$ or $\eta = 0.6$. The transfer rule becomes non-monotone When η is small because the highest types are almost willing to stop even without transfers whereas lower types need larger transfers as they get a large negative payoff if $\omega = L$ and need to be compensated for the option value of waiting.

D.4 Revenue maximizing designer

In this section, we show how the techniques that we developed extend to the case of a revenue maximizing designer. We then use the results in Appendix D.5 to solve the problem of a durable good monopolist. Throughout we assume that the type distribution F is twice continuously differentiable and has monotone hazard rate.

Consider a designer whose objective is to maximize the expected sum of transfers, $\mathbb{E}\left[\int_0^\infty e^{-rt} P_t dq_t\right]$, where P_t is the transfer that the agent pays if he stops at time t. Using the techniques in the previous section, the incentive compatible posted price is pinned down by (41). To back out the incentive compatible revenue, we change the order of integration to get a virtual surplus representation for the designer's payoff (see Appendix G.3 for the proof):[15]

Proposition 8. *Incentive compatibility implies that the designer's expected revenue is:*

$$\mathbb{E}\Big[\int_0^\infty e^{-rt} P_t dq_t\Big] = \mathbb{E}\Big[\int_{\underline{\theta}}^{\overline{\theta}} e^{-r\tau(\theta)} \phi_\omega(\theta) d\theta\Big], \tag{44}$$

where $\phi(\theta)$ is the virtual stopping payoff: $\phi_\omega(\theta) := v_\omega(\theta) - v'_\omega(\theta)\frac{1-F(\theta)}{f(\theta)}$.

Now, we can solve the revenue maximizing designer's problem by using Proposition 2 if we replace the stopping payoffs with virtual stopping payoffs and use a different initial value: $x^*(q_1) = 1$ where q_1 solves $\phi_H(\theta(q_1)) = 1$. Here the monotone hazard rate condition is important as it guarantees that the virtual valuation is increasing in θ and hence the problem satisfies all our assumptions in Section 2. Because we know that the planner's solution is a boundary policy, Proposition 7 guarantees that posted prices are without loss of generality and hence the revenue maximizing mechanism can be implemented.

D.5 Durable goods monopoly

We apply our results for durable goods monopoly when there is uncertainty about the product quality. Gradual learning naturally arises in durable good markets: after a buyer purchases the product, he starts using it and observes how well it functions for a long time. This is in contrast to experience goods, such as movies,

[15] The approach extends a result in our earlier paper Laiho and Salmi (2021) and shares similar features with Board (2007) who analyzes the optimal sale of options.

where consumption is one-shot and hence instantaneous learning is a natural way of formulating such markets. The existing literature on experimentation has focused solely on experience goods. One aim of this subsection is to point out that the informational trafeoff between the information generation and option value effects is likely to be present in the markets for goods that are used for a long time, such as new cars, home appliances, and investment goods. In addition, we present novel welfare results that are likely to hold independent of whether learning is instantaneous or gradual.

The durable good monopolist's problem under commitment is a special case of a revenue maximizing designer's problem of Appendix D.4. We use the following formulation. Neither the monopolist nor the buyers know the true quality of the product, $\omega \in \{H, L\}$, but they observe a public signal process (1), generated by past sales with the natural interpretation that the signal present's experienced payoffs together with some noise in communication. Each buyer wants to purchase one unit and exits after purchase. Similar to the general model a buyer's utility from consumption depends on his private type, $\theta \in [\underline{\theta}, \overline{\theta}]$, and the common quality: $\mathbb{E}[u(\theta, \omega)] = \mathbb{E}[\mathbf{1}_{\omega=H} \cdot \theta] = x_t\theta$, where x_t is the current belief that the quality is high. In addition, we assume that the type distribution satisfies the conditions in Appendix D.4. The monopolist faces marginal cost of production $c > 0$ and commits to a pricing scheme, P_t.

We can use Proposition 8 to write the monopolist's objective by using the expected virtual valuation net of the cost of production. Then, posted prices from Equation (41) can be used to implement the desired policy. The monopolist's problem becomes a special case of the model in Section 2 where $v_H(\theta) = \theta - (1 - F(\theta))/f(\theta) - c$ and $v_L(\theta) = -c$. This allows us to use Proposition 2 to characterize the profit maximizing policy:

Corollary 3. *The monopolist's policy is to sell when the belief is above $x^M(q)$ and to wait when it is below. The policy x^M is characterized by $x^M(\overline{q}^M) = 1$ and $x^{M\prime}(q) = g(q, x^M(q))$, where $\overline{q}^M$ solves $\theta(\overline{q}^M) - (1 - F(\theta(\overline{q}^M)))/f(\theta(\overline{q}^M)) = c$ and g is given in* (9)*.*

We contrast the monopoly solution with the planner's socially optimal solution and the competitive market equilibrium. The planner's solution can be found by applying Proposition 2 for the case where $v_H(q) = \theta(q) - c$ and $v_L(q) = -c$:

Corollary 4. *The social planner's policy* x^P *is characterized by* $x^P(\overline{q}^P) = 1$ *and* $x^{P\prime}(q) = g(q, x^P(q))$, *where* $\overline{q}^P$ *solves* $\theta(\overline{q}^P) = c$ *and* g *is given in* (9).

Suppose next that there are no barriers of entry to the market so that the price equals the marginal cost: $P_t = c$. An individual buyer's purchasing problem then coincides with the decentralized equilibrium in Section 3.4, with $v_H(q) = \theta(q) - c$ and $v_L(q) = -c$, and we have the following corollary to Proposition 1:

Corollary 5. *The competitive market policy is*

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

The planner's solution and the competitive equilibrium are the socially optimal and the decentralized solutions of the same problem, whereas the monopoly solution uses different stopping payoffs. This difference leads to different inefficiencies in monopoly and competitive markets.

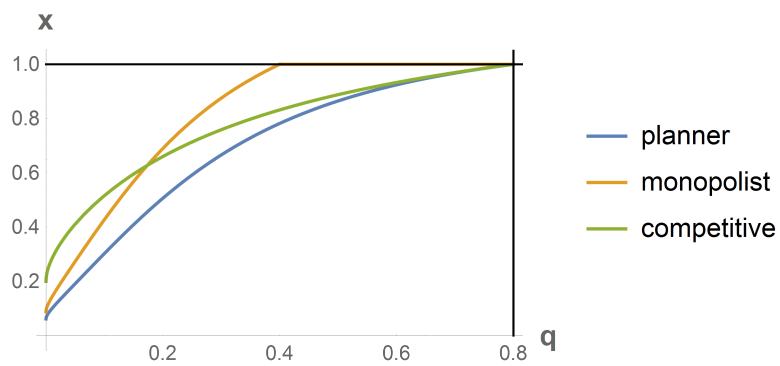

Figure 9: Different solutions for uniform $(0, 1)$ types, $c = 0.2$, $r = 0.1$, and $\sigma = 0.5$.

In the numerical example of Figure 9, the monopolist's and the competitive policies are everywhere above the planner's policy: both markets require inefficiently high belief for new consumers to purchase the product. The monopoly policy is first below and then above the competitive policy. This is because the information generation effect encourages the monopolist to sell in the beginning and because early sales do not generate large information rents to other buyers.

Later on, the monopolist reduces sales as the option value effect gets stronger and because later sales impose information rents to higher type buyers. The competitive market ignores the information generation effect but is otherwise efficient and therefore competitive sales are larger for high beliefs. This comparison generalizes to all type distributions and parameter values (proof in Appendix G.4):

Proposition 9. *There exist cutoffs $x_a \in (0,1)$ and $x_b \in (0,1)$ such that the monopoly quantity is larger if the initial belief is below x_a and the competitive market quantity is larger if the initial belief is above x_b.*

Notice that because the monopolist sells to lower type buyers only if the belief increases, the implied prices can be non-monotone. Endogenous learning favors introductory offers, and even pricing below the marginal cost because incentives to generate information are the strongest in the beginning. However, prices are still inefficiently high. The monopolist's incentives to generate information are weaker than the planner's because the monopolist cannot capture all the value from the buyers. This creates a distortion at the top of the type distribution, which is not present when the quality is known. In other words, a higher initial belief is needed for the monopolist to be willing to launch the product: $x^M(0) > x^P(0)$.

As a final remark, notice that a regulator can implement the socially efficient consumption in both monopoly and competitive markets by using appropriate subsidies but the subsidy schemes differ qualitatively. To encourage the monopolist to sell more, the regulator should use *back-loaded* subsidy that increases over sales: $s(x,q) = x(1 - F(\theta(q)))/f(\theta(q))$. A back-loaded subsidy scheme incentivizes the monopolist to sell the socially optimal amount because she internalizes the benefits of information generation. If the market is competitive, however, the subsidy must be *front-loaded* because competition eliminates dynamic incentives (see Figure 8 in Appendix D.3).

E Capital investments in a competitive industry

In this section, we consider an important application with negative payoff externalities: entry to a market with unknown demand. In markets for new products and services, firms investing in productive capital face uncertainty about market demand. Learning the true long-term demand takes time because of noise in consumer behavior and because the initial capacity may be too low to capture all important market segments. In the language of our model, the total capital installed is the stock variable that facilitates market experimentation and generates revenue.

We demonstrate that 1) the effect of market power on consumer welfare is ambiguous, 2) the quality of the learning technology drives the degree of distortion in competitive equilibrium, 3) social optimum can be implemented with investment subsidies that decrease over time.

E.1 Model of investments in a new market

Consider an investment problem where firms indexed by $\theta \in \left[\underline{\theta}, \overline{\theta}\right]$ choose when to make an irreversible investment to a market with unknown demand. Let the market use capital as the only input and assume firm θ can invest in a unit of capital at lump sum cost $c(\theta)$, where $c'(\theta) \leq 0$. Equivalently, cost of investment amounts to a perpetual stream of rental cost flow $rc(\theta)$. We assume that all capital is in production so that we can normalize the production level to equal the current capital stock q_t.[16] Let the inverse demand be

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

where x_t is the belief that the state of the world is high and D_ω is the inverse demand in state $\omega \in \{H, L\}$. We assume $D'_\omega(q) \leq 0$ for all $q \geq 0$.

Let the belief follow the diffusion process in (2). The interpretation of the learning process is now that only new purchases generate information: consumers

[16]This arises, for example, when the market is competitive and the marginal cost of using capacity is zero.

make repeat purchases and stay in the market forever. The rate of purchases is pinned down by investment decisions in the past.

Now, an investing firm imposes a negative payoff externality on other firms as an investment pushes the price down.[17] The payoff flow of type θ at time t can be written as: $\pi_\omega(\theta, q) = D_\omega(q) - rc(\theta)$.

E.2 Competitive industry equilibrium, monopoly solution, and social optimum

The decentralized equilibrium, which in this context we call the competitive industry equilibrium, can be easily computed using Proposition 5 (note that $\pi_\omega(\theta, q)$ is increasing in θ and decreasing in q):

Corollary 6. *The policy function for the decentralized equilibrium, x^D, is*

$$x^D(q) = \frac{-\beta(q)\left(D_L(q) - rc(\theta(q))\right)}{(\beta(q) - 1)\, D_H(q) - \beta(q)\, D_L(q) + rc(\theta(q))}.$$

We will contrast the competitive equilibrium to two different "centrally optimized" solutions, the monopoly solution and the social planner's solution. Since in these cases a central planner internalizes both the informational and payoff externalities, they can be solved using the techniques we developed in Section 3.5 for the social optimum. To see this, note that the total industry profit flow is

$$\pi_\omega^M(q) = qD_\omega(q) - r\int_{\theta(q)}^{1} c(s)\, dF(s).$$

Similarly, the total surplus flow can be written as

$$\pi_\omega^*(q) = \int_0^q D_\omega(s)\, ds - r\int_{\theta(q)}^{1} c(s)\, dF(s).$$

We find the monopoly and the social planner's solutions by following the solution in Section 3.5.

Corollary 7. *The policy functions for the social optimum, x^*, and for the monopolist, x^M, are characterized by Proposition 2 when using the following:*

[17] Notice that there is no payoff externality on the social level when we take into account consumer surplus. However, to solve for the competitive outcome, we need to take into account the effect between the firms.

- *For x^*, set $v_\omega(q) = r^{-1}D_\omega(q) - c(\theta(q))$ and use the initial value $x^*(\overline{q}^*) = 1$ where $\overline{q}^*$ solves $r^{-1}D_H(\overline{q}^*) = c(\theta(\overline{q}^*))$.*

- *For x^M, set $v_\omega(q) = r^{-1}(D_\omega(q) + D'_\omega(q)q = c(\theta(q))$ and use the initial value $x^M(\overline{q}^M) = 1$ where $\overline{q}^M$ solves $r^{-1}(D_H(\overline{q}^M) + D'_H(\overline{q}^M)\overline{q}^M) = c(\theta(\overline{q}^M))$.*

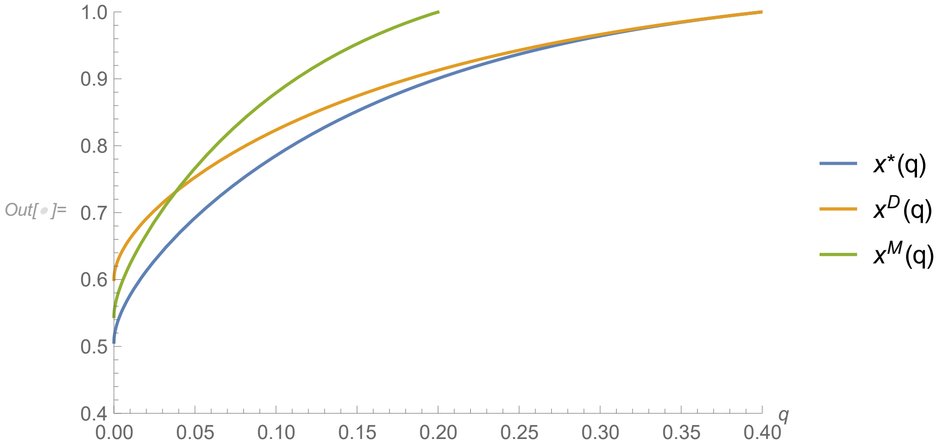

Figure 10: Socially optimal, competitive, and monopoly solutions.

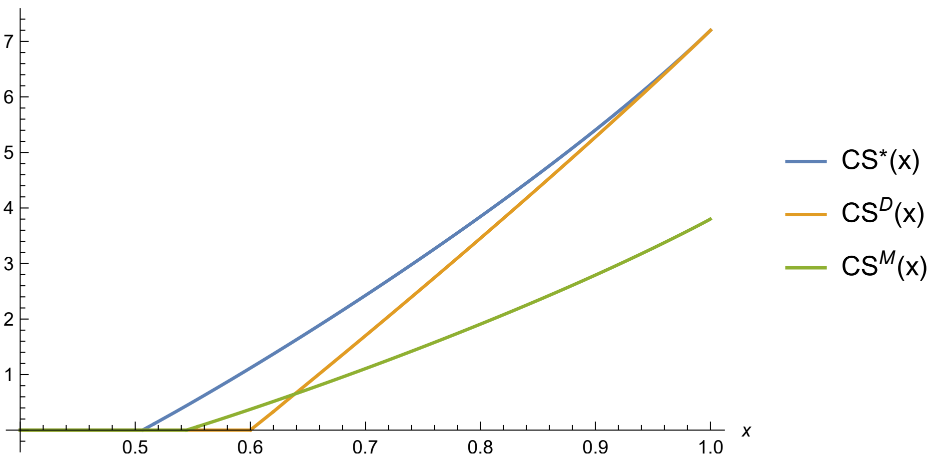

Figure 11: Consumer surplus as the function of the initial belief.

To illustrate these solutions, consider a parametric example: $D_\omega(q) = d_\omega - q$ and $c(\theta) = c$ for all $\theta \in [\underline{\theta}, \overline{\theta}]$ where d_ω and c are positive constants.[18] Figure

[18]Note that this formulation assumes that $\pi_\omega(\theta, q) = d_\omega - q - rc$ is independent of θ. We made it just to simplify computations. It rules out information rents for the firms and hence results in a competitive equilibrium where all profits are competed away.

10 shows the socially optimal, competitive equilibrium, and monopoly solutions for the parameter values $d_H = 2$, $d_L = 1$, $c = 16$, $r = 0.1$, and $\sigma = 1$. Figure 11 shows the resulting consumer surplus computed as a function of initial belief in the three cases. The solutions differ largely because of the same reasons as already discussed in the durable goods application in Appendix D.5: if there is a single monopolist, she never installs a large amount of capital because she wants to keep the production low and the price high; the competitive industry equilibrium suffers from insufficient investments when there is a lot of uncertainty because the potential entry of other firms eliminates all gains from information generation.

Notice that if learning were *exogenous*, the decentralized, or competitive, equilibrium corresponds to the social optimum, as established by Leahy (1993). With endogenous learning, this correspondence is no longer true because of the informational externality (see the difference between Proposition 1 and Proposition 2).

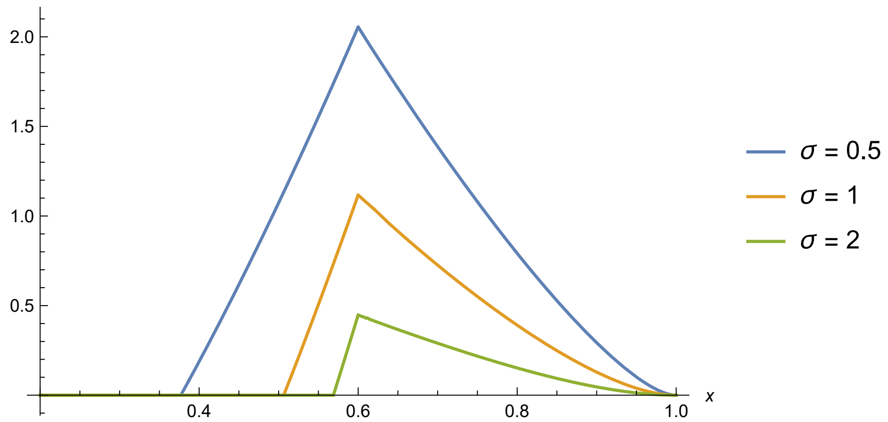

Figure 12: The difference in consumer welfare between the social optimum and the competitive equilibrium.

The problem with insufficient experimentation in the industry equilibrium is especially pronounced when the learning technology is good and learning could potentially be fast. One way to see this is to consider what happens to the lowest initial belief such that some investments take place under different learning technologies. When the learning technology improves, the threshold for the socially optimal initial belief decreases while it remains unchanged in the industry equilibrium. Even if the initial belief is high enough for some firms to enter, informational free-riding limits how much the actual speed of learning increases as the learning technology improves (see Proposition 4). Figure 12 depicts the difference in consumer welfare between the socially optimal solution and the industry equilibrium for different signal precisions: the better the learning technology, the larger the wedge between the two solutions. The welfare loss in the equilibrium is always the largest for the initial belief that makes the first firm indifferent whether to enter or not. For that belief, the social gain of information generation is large, while the individual gain from an investment is zero.

E.3 Taxes and subsidies

Let us next investigate the use of different transfer policies. Translating the terminology used above to the current context, a simple posted price policy corresponds to a lump sum investment tax/subsidy. Figure 13 shows the posted price policy computed for the socially optimal policy, following the steps in Section D.4. The parameter values are the same as in the previous subsection. Since the simple posted price policy is increasing, it implements the socially optimal investment plan by Proposition 7. The transfer is negative throughout, i.e. it is an investment subsidy. Intuitively, it is optimal to subsidise investment heavily in the beginning to compensate for the strong information generation effect, while the subsidy is reduced as the information generation gets gradually less important.

F Type-dependent informativeness: experts versus fanatics

Our baseline model assumes that all agents are equally informative: one unit of the stock q produces the same (marginal) amount of information. However, this might not necessarily be true in many applications of our model. For example, first buyers might be fans of the product whose experience matters less for the general population. Or conversely, the first units might be acquired by experts who are able to deduce the true value of the product much more quickly than

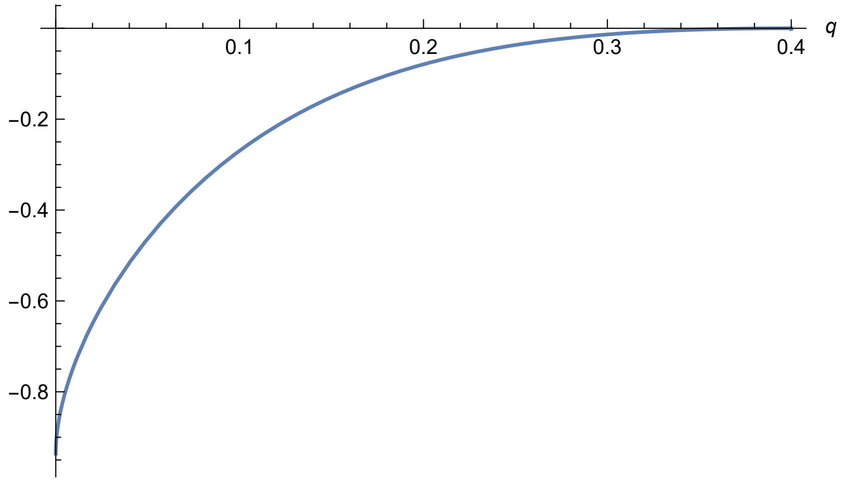

Figure 13: Transfers $P(q)$ for the social optimum.

average users. In this section, we show that as long as we can ensure that agents' stopping decision are monotone so that higher types stop first, the analysis in the baseline model is still valid.

Let the marginal informativeness of agent θ be $i(\theta) > 0$ so that the total "information stock" at time t is $z_t := \int_{\theta \in S_t} i(\theta) dF(\theta)$ where S_t is the set of agents who have stopped by time t. The evolution of the news process Y_t is then given by $dy_t = z_t \mu_\omega dt + \sigma \sqrt{z_t} dw_t$. The two especially interesting cases are when the high types are more informative, $i'(\theta) > 0$ ("experts"), and when the low types are more informative, $i'(\theta) < 0$ ("fanatics"). Throughout we assume that function i is continuously differentiable.

Notice first that the stopping profile in the decentralized equilibrium must be monotone in type because individual agents ignore the effect their stopping has on information, and thus Lemma 1 holds as in the baseline model. Because of this, we have a one-to-one relationship between the stock and the information stock: $z_t = h(q_t)$ where h is an "informativeness" function that satisfies $h'(q) = i(\theta(q))$. The analysis of the decentralized equilibrium then stays essentially the same as before.[19]

The socially optimal stopping profile is monotone in the experts environment

[19] We only need to adjust β function (plug in $h(q)$ instead of q). With this change Proposition 1 still characterizes decentralized behavior.

as both the marginal informativeness and the stopping payoff are increasing in the type. It is also monotone in the fanatics environments when marginal informativeness is not changing too extremely. In these cases, we can solve the socially optimal solution with a change of variables. The relevant problem is otherwise identical to the problem solved in Section 3.5 but the stock process Q is replaced with the information stock process Z and the stopping payoffs are scaled so that they take into account how many agents need to stop to increase the information stock by one unit: $\hat{v}_\omega(z) := v_\omega(h^{-1}(z))h^{-1\prime}(z)$. We show in Appendix G.5 that solving this problem solves the planner's original problem when infomativeness of stopping is type-dependent:[20]

Proposition 10. *Suppose either condition (i) or condition (ii) holds:*

(i) Marginal informativeness is increasing in type (experts).

(ii) Marginal informativeness is decreasing in type (fanatics) but informativeness does not change too extremely: $\frac{-i'(\theta)}{i(\theta)} \leq \frac{-v_L'(\theta)}{v_L(\theta)}$ *holds for all* θ.

Then the socially optimal policy is characterized by $x^*(z(q))$ *where* x^* *is as defined in Proposition 2.*

When high types are extremely fanatical, we may run into trouble: because low types are more informative, the social planner may want to use them first for experimentation. If we restrict to monotone allocations over types, this problem disappears. Then, the optimal policy function may be non-increasing because the scaled payoffs are non-monotone. Essentially the social planner may want to bunch some agents to stop together because lower types may have a larger social value of stopping.

Figure 14 illustrates how heterogeneous informativeness affects the socially optimal policy. As before, the comparison between optimal quantities depends on the informational tradeoff between option value and information generation effects. The information generation effect is more pronounced for low q in the expert

[20] The key steps are to show that the optimal stopping profile is monotone in type and that the policy function x^* is monotone in z.

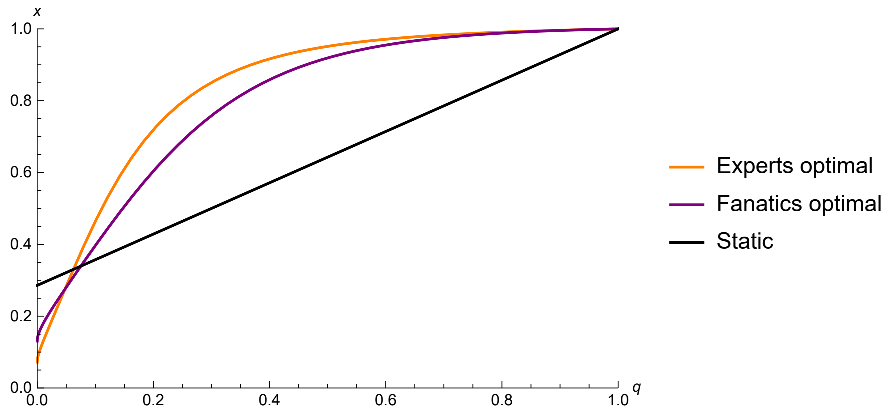

Figure 14: Socially optimal policies for experts (concave $h(q) = 3/2q - 1/2q^2$) and fanatics (convex $h(q) = 1/2q + 1/2q^2$) when $v_H(q) = 1 - q$, $v_L(q) = -0.4 - q$, $\sigma = 0.5$, and $r = 0.1$.

environment, while it is more pronounced for higher q in the fanatics environment. We see this in Figure 14 as the solution in the expert environment is first below and then quickly rises above the fanatics solution. The expert environment favors relatively early expansions because the high types produce more information than in the fanatics environment.

G Proofs for Appendices D, E, and F

G.1 Static implementation: proof of Lemma 9

Proof. We first prove the increasing differences property that is crucial for global incentive compatibility. Denote by $U\left(\theta, \widetilde{\theta}\right)$ the expected payoff at time $t = 0$ for type θ that reports $\widetilde{\theta}$:

$$U\left(\theta, \widetilde{\theta}\right) = \mathbb{E}\left[e^{-r\tau\left(\widetilde{\theta}\right)}\left(x_{\tau\left(\widetilde{\theta}\right)} v_H(\theta) + (1 - x_{\tau\left(\widetilde{\theta}\right)}) v_L(\theta)\right)\right] - P_0\left(\widetilde{\theta}\right).$$

Its partial derivative with respect to θ is

$$U_1\left(\theta, \widetilde{\theta}\right) = \mathbb{E}\left[e^{-r\tau\left(\widetilde{\theta}\right)}\left(x_{\tau\left(\widetilde{\theta}\right)} v_H'(\theta) + (1 - x_{\tau\left(\widetilde{\theta}\right)}) v_L'(\theta)\right)\right]. \tag{45}$$

The property that we want to prove is that $U_1\left(\theta, \widetilde{\theta}\right)$ is increasing in $\widetilde{\theta}$. To do that, note that applying the law of iterated expectations, we can write $U_1\left(\theta, \widetilde{\theta}\right)$

in terms of the initial belief x_0 as

$$\begin{aligned} U_1\left(\theta,\widetilde{\theta}\right) &= x_0\mathbb{E}\left(e^{-r\tau\left(\widetilde{\theta}\right)}v'_H\left(\theta\right)|\omega=H\right)+\left(1-x_0\right)\mathbb{E}\left(e^{-r\tau\left(\widetilde{\theta}\right)}v'_L\left(\theta\right)|\omega=L\right) \\ &= x_0v'_H\left(\theta\right)\mathbb{E}\left(e^{-r\tau\left(\widetilde{\theta}\right)}|\omega=H\right)+\left(1-x_0\right)v'_L\left(\theta\right)\mathbb{E}\left(e^{-r\tau\left(\widetilde{\theta}\right)}|\omega=L\right). \end{aligned}$$

Since $v'_H\left(\theta\right)\geq 0$ and $v'_L\left(\theta\right)\geq 0$ (with at least one of the inequalities strict), both terms in the above expression are positive. Reported type $\widetilde{\theta}$ enters the expression only through the discounting terms $\mathbb{E}\left(e^{-r\tau\left(\widetilde{\theta}\right)}|\omega\right)$. Since $\theta''>\theta'$ implies that $\tau\left(\theta''\right):=\inf\left\{t:q_t=1-F\left(\theta''\right)\right\}<\tau\left(\theta'\right):=\inf\left\{t:q_t=1-F\left(\theta'\right)\right\}$ with probability 1, it follows that $\mathbb{E}\left(e^{-r\tau\left(\widetilde{\theta}\right)}|\omega\right)$ is strictly increasing in $\widetilde{\theta}$ irrespective of state ω, and hence $U_1\left(\theta,\widetilde{\theta}\right)$ is strictly increasing in $\widetilde{\theta}$ as well.

We now utilize the above property to show that incentive compatibility holds for all types θ. Note first that we used the envlope theorem to set the transfer payment in (40) so that the value of θ satisfies:

$$\begin{aligned} W_0(\theta) &= \mathbb{E}\Big[e^{-r\tau(\theta)}(x_{\tau(\theta)}v_H(\theta)+(1-x_{\tau(\theta)})v_L(\theta))\Big]-P_0(\theta) \\ &= \mathbb{E}\left[\int_{\underline{\theta}}^{\theta}e^{-r\tau(s)}\Big(x_{\tau(s)}v'_H(s)+(1-x_{\tau(s)})v'_L(s))\Big)ds\right]. \end{aligned}$$

Therefore, for arbitrary θ' and θ'', we have

$$W_0(\theta'')-W_0(\theta')=\mathbb{E}\left[\int_{\theta'}^{\theta''}e^{-r\tau(s)}\Big(x_{\tau(s)}v'_H(s)+(1-x_{\tau(s)})v'_L(s))\Big)ds\right]. \quad (46)$$

We can now write the payoff of type θ who reports $\widetilde{\theta}$ as:

$$\begin{aligned} U\left(\theta,\widetilde{\theta}\right) &= U\left(\widetilde{\theta},\widetilde{\theta}\right)+\int_{\widetilde{\theta}}^{\theta}U_1\left(s,\widetilde{\theta}\right)ds=W_0\left(\widetilde{\theta}\right)+\int_{\widetilde{\theta}}^{\theta}U_1\left(s,\widetilde{\theta}\right)ds \\ &\leq W_0\left(\widetilde{\theta}\right)+\int_{\widetilde{\theta}}^{\theta}U_1\left(s,s\right)ds= \\ &= W_0\left(\widetilde{\theta}\right)+\mathbb{E}\left[\int_{\widetilde{\theta}}^{\theta}e^{-r\tau(s)}\Big(x_{\tau(s)}v'_H(s)+(1-x_{\tau(s)})v'_L(s))\Big)ds\right] \\ &= W_0\left(\theta\right), \end{aligned}$$

where the inequality uses the property that $U_1\left(\theta,\widetilde{\theta}\right)$ is increasing in $\widetilde{\theta}$, the second last equality uses (45), and the last equality uses (46). This shows that it is optimal for an arbitrary type θ to report truthfully.

To see that participation constraint is satisfied it sufficies to note that by reporting $\underline{\theta}$ any type gets a weakly positive payoff.

□

G.2 Dynamic implementation: proof of Proposition 7

Proof. Fix a boundary policy Q with a strictly increasing boundary $\widetilde{x}(q)$ and $\widetilde{x}(q_1) = 1$. We analyze the optimal stopping problem of arbitrary type θ. We will show that under $P^D(x)$ it is optimal to stop at the first hitting time of the point $(\widetilde{x}(q(\theta)), q(\theta))$ and the same conclusion holds under $P^S(q)$ if $\left(P^S\right)'(q) \geq 0$ for all $0 \leq q \leq q_1$. Finally, we show that if $\left(P^S\right)'(q) < 0$ for some q, then type $\theta(q)$ has a profitable deviation to stopping at some (x', q') with $x' < \widetilde{x}(q')$, $q' < q$.

Step 1: Optimal stopping when stopping below boundary prohibited

As a preliminary step we confirm that under both $P^D(x)$ and $P^S(q)$ it is optimal to stop at $(\widetilde{x}(q(\theta)), q(\theta))$ if stopping below the boundary is prohibited. This restricted stopping problem is still a Markovian stopping problem where the optimal solution is a first-hitting time of some of the boundary points. Under both $P^D(x)$ and $P^S(x)$, stopping at boundary point $(\widetilde{x}(q), q)$ entails transfer payment $P(\widetilde{x}(q), q)$ given in (41), which is designed in such a way that ex-ante expected payoff is equivalent to reporting $\theta(q)$ in the static direct mechanism. According to Lemma 9 it is optimal for θ to report truthfully in the static direct mechanism and so it follows that stopping at $(\widetilde{x}(q(\theta)), q(\theta))$ is optimal in the restricted stopping problem.

Step 2: Optimal stopping under dynamic posted price $P^D(x)$

We will utilize the property that the stopping value $u_\theta^D(x)$ is indepenent of q to show that even if stopping below the boundary is allowed, an agent will optimally stop at a first-hitting time of *some* boundary point. For contradiction, assume that type θ stops with a strictly positive probability at the first-hitting time of some point below the boundary and let (x', q') be the "left-most" such point. Formally, denoting by $F_\theta(x, q)$ the optimal value function of type θ, let (x', q'), $x' < \widetilde{x}(q')$, be a state point such that $F_\theta(x', q') = u_\theta^D(x')$ and $F_\theta(x, q) > u_\theta^D(x)$ for all points with $q < q'$.[21] We can write $F_\theta(x, q) = B_\theta(q)\,\Phi(x, q)$ for some function $B_\theta(q)$ so

[21] Note that we must have $F_\theta(\widetilde{x}(q), q) > u_\theta^D(\widetilde{x}(q))$ also at all boundary points with $q \leq q'$ because all those boundary points are visited before point (x', q') and hence otherwise stopping could not take place with positive probability at (x', q') .

that

$$\begin{aligned}\frac{\partial}{\partial q}F_\theta(x,q) &= B_\theta'(q)\,\Phi(x,q)+B_\theta(q)\,\Phi_q(x,q)\\ &= \Phi(x,q)\left[B_\theta'(q)+B_\theta(q)\,\beta'(q)\ln\left(\frac{x}{1-x}\right)\right],\end{aligned}$$

where we have used that $\Phi_q(x,q)=\beta'(q)\ln\left(\frac{x}{1-x}\right)\Phi(x,q)$.

Similar to the proof of Proposition 5, we note that the partial of a value function w.r.t. q must be zero at the boundary $x=\widetilde{x}(q)$, i.e. for $q<q'$ we have $\frac{\partial}{\partial q}F_\theta(x,q)_{|x=\widetilde{x}(q)}=0$ and so

$$B_\theta'(q)+B_\theta(q)\,\beta'(q)\ln\left(\frac{\widetilde{x}(q)}{1-\widetilde{x}(q)}\right)=0.$$

But since $\beta'(q)<0$ and $\widetilde{x}(q)>x$, this implies that

$$B_\theta'(q)+B_\theta(q)\,\beta'(q)\ln\left(\frac{x}{1-x}\right)>0,$$

and so we have $\frac{\partial}{\partial q}F_\theta(x,q)>0$. This is a contradiction with our assumption that $F_\theta(x',q')=u_\theta^D(x')$ and $F_\theta(x,q)>u_\theta^D(x)$ for all points with $q<q'$.

We can conclude that the optimal stopping time must be the first hitting time of some boundary point. This means that the solution must be the same as if stopping below the bounday is prohibited. By step 1 above, it is then optimal for θ to stop at the first-hitting time of point $(\widetilde{x}(q(\theta)),q(\theta))$.

Step 3: Optimal stopping under simple posted price $P^S(q)$

Since the stopping value $u_\theta^S(x,q)$ is now a function of q as well as x, the argument in the second step above does not hold and we will use a more direct approach. We will first directly show that if $P^S(q)$ is increasing everywhere, it is optimal for type θ to stop at the first hitting time of point $(\widetilde{x}(q(\theta)),q(\theta))$ that we denote by

$$\tau_\theta^*:=\inf\{t:(x_t,q_t)=(\widetilde{x}(q(\theta)),q(\theta))\}.$$

After that we will show that if $\left(P^S\right)'(q)<0$ for some q, then type $\theta(q)$ has a profitable deviation to stop at some (x',q') with $q'<q(\theta)$, $x'<\widetilde{x}(q')$.

Let us first investigate the implication of the condition $\left(P^S\right)'(q)\geq 0$. We can write

$$P^S(q)=x(q)\,v_H(\theta(q))+(1-x(q))\,v_L(\theta(q))-S(q),$$

where $S(q)$ is the information rent obtained by type $\theta(q)$ evaluated at the moment of stopping:

$$S(q) := \mathbb{E}\left[\int_{\underline{\theta}}^{\theta(q)} e^{-r\tau(s)} \left(x_{\tau(s)} v_H'(s) + \left(1 - x_{\tau(s)}\right) v_L'(s)\right) ds \,|\widetilde{x}(q), q\right].$$

We next differentiate $P^S(q)$ with respect to q. Note first that for any $s < \theta(q)$ we can write

$$\mathbb{E}\left[e^{-r\tau(s)} \,|\widetilde{x}(q), q\right] = A_s(q)\, \Phi(\widetilde{x}(q), q) \tag{47}$$

for some function $A_s(q)$. Since q_t increases at the boundary, the partial derivative of (47) w.r.t. q must be zero there, i.e.: $\frac{\partial}{\partial q}\mathbb{E}\left[e^{-r\tau(s)} \,|\widetilde{x}(q), q\right] = 0$. Therefore, the change of (47) when moving the initial point along the boundary is

$$\begin{aligned}
\frac{d}{dq}\mathbb{E}\left[e^{-r\tau(s)} \,|\widetilde{x}(q), q\right] &= \frac{\partial}{\partial x}\mathbb{E}\left[e^{-r\tau(s)} \,|x, q\right]_{x=\widetilde{x}(q)} \widetilde{x}'(q) \\
&= A_s(q)\, \Phi_x(\widetilde{x}(q), q)\, \widetilde{x}'(q) \\
&= \frac{\Phi_x(\widetilde{x}(q), q)}{\Phi(\widetilde{x}(q), q)} \widetilde{x}'(q)\, \mathbb{E}\left[e^{-r\tau(s)} \,|\widetilde{x}(q), q\right]. \qquad (48)
\end{aligned}$$

Using this, we get:

$$\begin{aligned}
S'(q) &= \Big(\widetilde{x}(q)\, v_H'(\theta(q)) + (1 - \widetilde{x}(q))\, v_L'(\theta(q))\Big)\theta'(q) \\
&+ \int_{\underline{\theta}}^{\theta(q)} \frac{\Phi_x(\widetilde{x}(q), q)}{\Phi(\widetilde{x}(q), q)} \widetilde{x}'(q)\, \mathbb{E}\left[e^{-r\tau(s)} \left(\widetilde{x}_{\tau(s)} v_H'(s) + \left(1 - \widetilde{x}_{\tau(s)}\right) v_L'(s)\right) |\widetilde{x}(q), q\right] ds \\
&= \Big(\widetilde{x}(q)\, v_H'(\theta(q)) + (1 - \widetilde{x}(q))\, v_L'(\theta(q))\Big)\theta'(q) + \frac{\Phi_x(\widetilde{x}(q), q)}{\Phi(\widetilde{x}(q), q)} S(q)\, \widetilde{x}'(q), \qquad (49)
\end{aligned}$$

where the first term is from differentiation with respect to the integral bound and the second term is from differentiation with respect to the initial point $\widetilde{x}(q)$ using (48). The derivative of $P^S(q)$ with respect to q is then

$$\begin{aligned}
\left(P^S\right)'(q) =& \widetilde{x}'(q)\,(v_H(\theta(q)) - v_L(\theta(q))) \\
&+ \left[\widetilde{x}(q)\, v_H'(\theta(q)) + (1 - \widetilde{x}(q))\, v_L'(\theta(q))\right] \theta'(q) - S'(q) \\
=& \left[v_H(\theta(q)) - v_L(\theta(q)) - \frac{\Phi_x(\widetilde{x}(q), q)}{\Phi(\widetilde{x}(q), q)} S(q)\right] \widetilde{x}'(q).
\end{aligned}$$

Hence, $\left(P^S\right)'(q) \geq 0$ is equivalent to

$$v_H(\theta(q)) - v_L(\theta(q)) \geq \frac{\Phi_x(\widetilde{x}(q), q)}{\Phi(\widetilde{x}(q), q)} S(q), \tag{50}$$

a condition that we will utilize below.

We now show that it cannot be optimal to stop below the boundary $\widetilde{x}(q)$ for $q < q(\theta)$. Fix (x,q) with $q < q(\theta)$ and $x < \widetilde{x}(q)$. We show that stopping at (x,q) is dominated by waiting until x hits $\widetilde{x}(q)$. Denote the value under that stopping rule by

$$\underline{F}_{\theta}(x,q) := \mathbb{E}\left[e^{-r\tau(\widetilde{x}(q))}u_{\theta}^{S}(x,q)\,|x,q\right],$$

where $\tau(\widetilde{x}(q)) = \inf\{t : x_t = \widetilde{x}(q)\}$. Writing $\underline{F}_{\theta}(x,q) := \underline{A}_{\theta}(q)\,\Phi(x,q)$ and solving $\underline{A}_{\theta}(q)$ from the boundary condition $\underline{F}_{\theta}(\widetilde{x}(q),q) = u_{\theta}^{S}(\widetilde{x}(q),q)$ gives us

$$\underline{F}_{\theta}(x,q) = \frac{\Phi(x,q)}{\Phi(\widetilde{x}(q),q)}u_{\theta}^{S}(\widetilde{x}(q),q).$$

We aim to show that below the boundary, i.e. for $x < \widetilde{x}(q)$, we have $\underline{F}_{\theta}(x,q) > u_{\theta}^{S}(x,q)$. Let us differentiate these functions with respect to x, and evaluate the derivative at the boundary:

$$\frac{\partial}{\partial x}\left[u_{\theta}(x,q)\right]_{x=\widetilde{x}(q)} = v_H(\theta) - v_L(\theta)$$

and

$$\begin{aligned}
&\frac{\partial}{\partial x}\left[\underline{F}_{\theta}(x,q)\right]_{x=\widetilde{x}(q)} = \frac{\Phi_x(\widetilde{x}(q),q)}{\Phi(\widetilde{x}(q),q)}u_{\theta}(\widetilde{x}(q),q)\\
&= \frac{\Phi_x(\widetilde{x}(q),q)}{\Phi(\widetilde{x}(q),q)}\left[\widetilde{x}(q)\left(v_H(\theta) - v_H(\theta(q))\right) + \left(1-\widetilde{x}(q)\right)\left(v_L(\theta) - v_L(\theta(q))\right) + S(q)\right]\\
&\leq v_H(\theta)\frac{\beta(q) - \widetilde{x}(q)}{1-\widetilde{x}(q)} + v_L(\theta)\frac{\beta(q) - \widetilde{x}(q)}{\widetilde{x}(q)} - v_H(\theta(q))\frac{\beta(q) - 1}{1-\widetilde{x}(q)} - v_L(\theta(q))\frac{\beta(q)}{\widetilde{x}(q)}\\
&= v_H(\theta) - v_L(\theta) + \left(\frac{\beta(q) - 1}{1-\widetilde{x}(q)}\right)\left(v_H(\theta) - v_H(\theta(q))\right) + \frac{\beta(q)}{\widetilde{x}(q)}\left(v_L(\theta) - v_L(\theta(q))\right)\\
&< v_H(\theta) - v_L(\theta),
\end{aligned}$$

where the first inequality utilizes the fact that $(P^s)'(q) \geq 0$ is equivalent to (50) and the second inequality utilizes the fact that $q < q(\theta)$ implies that $v_H(\theta) - v_H(\theta(q)) \leq 0$ and $v_L(\theta) - v_L(\theta(q)) \leq 0$ with at least one of the inequalities strict It now follows that

$$\frac{\partial}{\partial x}\left[\underline{F}_{\theta}(x,q)\right]_{x=\widetilde{x}(q)} < \frac{\partial}{\partial x}\left[u_{\theta}^{S}(x,q)\right]_{x=\widetilde{x}(q)},$$

and since $\underline{F}_{\theta}(x,q)$ is strictly convex in x and positive for $x < \widetilde{x}(q)$ while $u_{\theta}^{S}(x,q)$ is linear in x, we have

$$\underline{F}_{\theta}(x,q) > u_{\theta}^{S}(x,q) \text{ for } x < \widetilde{x}(q),$$

and so stopping at (x,q) is strictly dominated by waiting until x hits $\widetilde{x}(q)$. Since (x,q) was arbitrarily chosen, it can never be optimal to stop strictly below the boundary for $q<q(\theta)$. But we know from Step 1 of the proof that stopping at point $(\widetilde{x}(q(\theta)),q(\theta))$ dominates stopping at other boundary points. This implies that it can never be optimal for θ to stop ealier than at τ_θ^*.

The above argument ruled out stopping below the boundary only for $q<q(\theta)$. It remains to show that θ cannot benefit from delaying stopping beyond time τ_θ^* to the hitting time of some (x,q) with $q>q(\theta)$, $x<\widetilde{x}(q)$. For that it sufficies to show that it is optimal to stop at all boundary points for $q\geq q(\theta)$, i.e. at all $(\widetilde{x}(q),q)$ for $q\geq q(\theta)$. The proof follows similar reasoning as the proof of Proposition 5.

Suppose, to the contrary, that there is some $(\widetilde{x}(q),q)$ such that $q\geq q(\theta)$ where it is not optimal to stop. At that point we therefore have $\widetilde{F}_\theta(\widetilde{x}(q),q)>u_\theta^S(\widetilde{x}(q),q)$, where $\widetilde{F}_\theta(x,q)$ is the value function under the optimal stopping rule (whatever that may be). We will show next that this implies that along the boundary, the rate of change in $\widetilde{F}_\theta(\widetilde{x}(q),q)$ is higher than in $u_\theta^S(\widetilde{x}(q),q)$:

$$\frac{d}{dq}\widetilde{F}_\theta(\widetilde{x}(q),q)>\frac{d}{dq}u_\theta^S(\widetilde{x}(q),q). \tag{51}$$

We prove this separately in two possible cases. First, suppose that it is optimal to stop for some (x',q), where $x'<\widetilde{x}(q)$. In that case $\widetilde{F}_\theta(x',q)=u_\theta^S(x',q)$. Since $\widetilde{F}_\theta(x,q)$ must be strictly convex in x whenever it is optimal to wait (i.e. when $\widetilde{F}_\theta(x,q)>u_\theta^S(x,q)$), whereas $u_\theta^S(x,q)$ is linear in x with slope $v_H(\theta)-v_L(\theta)$, we must have

$$\frac{\partial}{\partial x}\left[\widetilde{F}_\theta(x,q)\right]_{x=\widetilde{x}(q)}>v_H(\theta)-v_L(\theta). \tag{52}$$

Let us now compare the rates of change in $\widetilde{F}_\theta(\widetilde{x}(q),q)$ and $u_\theta^S(\widetilde{x}(q),q)$ along the boundary. We have

$$\begin{aligned}\frac{d}{dq}\widetilde{F}_\theta(\widetilde{x}(q),q) &= \frac{\partial}{\partial x}\left[\widetilde{F}_\theta(x,q)\right]_{x=\widetilde{x}(q)}\widetilde{x}'(q)+\frac{\partial}{\partial q}\widetilde{F}_\theta(\widetilde{x}(q),q)\\ &= \frac{\partial}{\partial x}\left[\widetilde{F}_\theta(\widetilde{x},q)\right]_{x=\widetilde{x}(q)}\widetilde{x}'(q),\end{aligned}$$

where we have again utilized the fact that the partial of a waiting value w.r.t q must vanish at the boundary where q is increased, i.e. $\frac{\partial}{\partial q}\widetilde{F}_\theta(\widetilde{x}(q),q)=0$.

The stopping value at (x, q) is

$$\begin{aligned} u_\theta^S(x,q) &= xv_H(\theta) + (1-x)\, v_L(\theta) - P^S(q) \\ &= xv_H(\theta) + (1-x)\, v_L(\theta) \\ &\quad -x(q)\, v_H(\theta(q)) - (1-x(q))\, v_L(\theta(q)) + S(q) \end{aligned}$$

and so at the boundary $x = \widetilde{x}(q)$ the stopping value is

$$\begin{aligned} u_\theta^S(\widetilde{x}(q), q) &= \widetilde{x}(q)\,(v_H(\theta) - v_H(\theta(q))) \\ &\quad + (1-\widetilde{x}(q))\,(v_L(\theta) - v_L(\theta(q))) + S(q) \end{aligned}$$

and we can compute its rate of change along the boundary as:

$$\begin{aligned} &\frac{d}{dq} u_\theta^S(\widetilde{x}(q), q) = \widetilde{x}'(q)\,(v_H(\theta) - v_H(\theta(q)) - v_L(\theta) + v_L(\theta(q))) \\ &- [\widetilde{x}(q)\, v_H'(\theta(q)) + (1-\widetilde{x}(q))\, v_L'(\theta(q))]\, \theta'(q) + S'(q) \\ &= \widetilde{x}'(q) \left[(v_H(\theta) - v_L(\theta)) - (v_H(\theta(q)) - v_L(\theta(q))) + \frac{\Phi_x(\widetilde{x}(q), q)}{\Phi(\widetilde{x}(q), q)} S(q) \right], \quad (53) \end{aligned}$$

where the latter equality uses (49). Using (52) and (50), we then see that (51) holds.

We move to the second case. Suppose that it is optimal to wait for all (x, q), where $x < \widetilde{x}(q)$, in which case $\widetilde{F}_\theta(x,q) > u_\theta^S(x,q)$ for all $x < \widetilde{x}(q)$ and $\widetilde{F}_\theta(0,q) = 0$. In that case, function $\widetilde{F}_\theta(x,q)$ must take the form

$$\widetilde{F}_\theta(x,q) = \widehat{A}_\theta(q)\, \Phi(x,q)$$

for some function $\widehat{A}_\theta(q)$. Our assumption $\widetilde{F}_\theta(\widetilde{x}(q), q) > u_\theta^S(\widetilde{x}(q), q)$ is equivalent to

$$\widehat{A}_\theta(q)\, \Phi(\widetilde{x}(q), q) > \widetilde{x}(q)\,(v_H(\theta) - v_H(\theta(q))) + (1-\widetilde{x}(q))\,(v_L(\theta) - v_L(\theta(q))) + S(q)$$

or

$$\widehat{A}_\theta(q) > \frac{\widetilde{x}(q)\,(v_H(\theta) - v_H(\theta(q))) + (1-\widetilde{x}(q))\,(v_L(\theta) - v_L(\theta(q))) + S(q)}{\Phi(\widetilde{x}(q), q)}.$$

This implies

$$\frac{d}{dq}\widetilde{F}_\theta\left(\widetilde{x}\left(q\right),q\right)>\widetilde{x}'\left(q\right)\frac{\Phi_x\left(\widetilde{x}\left(q\right),q\right)}{\Phi\left(\widetilde{x}\left(q\right),q\right)}$$
$$\times\left[\widetilde{x}\left(q\right)\left(v_H\left(\theta\right)-v_H\left(\theta\left(q\right)\right)\right)+\left(1-\widetilde{x}\left(q\right)\right)\left(v_L\left(\theta\right)-v_L\left(\theta\left(q\right)\right)\right)+S\left(q\right)\right]$$
$$=\widetilde{x}'\left(q\right)\left[\frac{\beta\left(q\right)-\widetilde{x}\left(q\right)}{\left(1-\widetilde{x}\left(q\right)\right)}\left(v_H\left(\theta\right)-v_H\left(\theta\left(q\right)\right)\right)\right.$$
$$\left.+\frac{\beta\left(q\right)-\widetilde{x}\left(q\right)}{\widetilde{x}\left(q\right)}\left(v_L\left(\theta\right)-v_L\left(\theta\left(q\right)\right)\right)+\frac{\Phi_x\left(\widetilde{x}\left(q\right),q\right)}{\Phi\left(\widetilde{x}\left(q\right),q\right)}S\left(q\right)\right].$$

Noting that $q>q\left(\theta\right)$ means that $v_H\left(\theta\right)-v_H\left(\theta\left(q\right)\right)\geq 0$ and $v_L\left(\theta\right)-v_L\left(\theta\left(q\right)\right)\geq 0$, and comparing the expression above to (53), we note again that (51) holds.

Having proved that $\widetilde{F}_\theta\left(\widetilde{x}\left(q\right),q\right)>u_\theta\left(\widetilde{x}\left(q\right),q\right)$ implies (51), we note that this would imply that $\widetilde{F}_\theta\left(\widetilde{x}\left(q'\right),q'\right)>u_\theta\left(\widetilde{x}\left(q'\right),q'\right)$ along the boundary $\left(\widetilde{x}\left(q'\right),q'\right)$ for all $q\leq q'\leq q_1$, and so $\widetilde{F}_\theta\left(\widetilde{x}\left(q_1\right),q_1\right)>u_\theta\left(\widetilde{x}\left(q_1\right),q_1\right)$. This is a contradiction since we know that it must be optimal to stop at state point $\left(\widetilde{x}\left(q_1\right),q_1\right)$. We conclude that it is optimal to stop at all $\left(\widetilde{x}\left(q\right),q\right)$, $q\left(\theta\right)\leq q\leq q_1$. It now follows that the optimal stopping time for θ is τ_θ^*, i.e. the first hitting time of $\left(\widetilde{x}\left(q\left(\theta\right)\right),q\left(\theta\right)\right)$.

As a final point we note that our conclusion hinges critically on the assumption that $\left(P^S\right)'\left(q\right)\geq 0$ for all q. If, in contrast, $\left(P^S\right)'\left(q\right)<0$ for some $q\in\left(0,q_1\right)$, then there is a profitable deviation for type $\theta=\theta\left(q\right)$ to stop ealier than at time τ_θ^*. To see this, note that $\left(P^S\right)'\left(q\right)<0$ implies that for $\theta=\theta\left(q\right)$, we have

$$v_H\left(\theta\left(q\right)\right)-v_L\left(\theta\left(q\right)\right)<\frac{\Phi_x\left(\widetilde{x}\left(q\right),q\right)}{\Phi\left(\widetilde{x}\left(q\right),q\right)}S\left(q\right). \tag{54}$$

Consider then the value of type $\theta\left(q\right)$ who plans to stop at $\left(\widetilde{x}\left(q\right),q\right)$:

$$F_{\theta(q)}\left(x,q\right)=\widetilde{A}_{\theta(q)}\left(q\right)\Phi\left(x;q\right).$$

Since

$$F_{\theta(q)}\left(\widetilde{x}\left(q\right),q\right)=u^S_{\theta(q)}\left(\widetilde{x}\left(q\right),q\right)$$
$$=v_H\left(\theta\left(q\right)\right)-v_L\left(\theta\left(q\right)\right)-\left(v_H\left(\theta\left(q\right)\right)-v_L\left(\theta\left(q\right)\right)-S\left(q\right)\right)=S\left(q\right),$$

we have

$$\widetilde{A}_{\theta(q)}\left(q\right)=\frac{S\left(q\right)}{\Phi\left(\widetilde{x}\left(q\right),q\right)}$$

and so

$$\frac{\partial}{\partial x}\left[F_{\theta(q)}\left(x,q\right)\right]_{x=\widetilde{x}(q)}=\frac{\Phi_{x}\left(\widetilde{x}\left(q\right),q\right)}{\Phi\left(\widetilde{x}\left(q\right),q\right)}S\left(q\right).$$

Noting that

$$\frac{\partial}{\partial x}\left[u^{S}_{\theta(q)}\left(x,q\right)\right]_{x=\widetilde{x}(q)}=v_{H}\left(\theta\left(q\right)\right)-v_{L}\left(\theta\left(q\right)\right),$$

equation (54) implies

$$\frac{\partial}{\partial x}\left[u^{S}_{\theta(q)}\left(x,q\right)\right]_{x=\widetilde{x}(q)}<\frac{\partial}{\partial x}\left[F_{\theta(q)}\left(x;q\right)\right]_{x=\widetilde{x}(q)}$$

and since $u^{S}_{\theta(q)}\left(\widetilde{x}\left(q\right),q\right)=F_{\theta(q)}\left(\widetilde{x}\left(q\right),q\right)$, it follows that for some $x<\widetilde{x}\left(q\right)$ we have

$$u^{S}_{\theta(q)}\left(x,q\right)>F_{\theta(q)}\left(x,q\right).$$

By continuity of F and u, this implies that there is some $q'<q$, $x'<\widetilde{x}\left(q'\right)$, such that $u^{S}_{\theta(q)}\left(x',q'\right)>F_{\theta(q)}\left(x',q'\right)$, and hence type $\theta\left(q\right)$ has a strictly beneficial deviation to stopping at that state, and that state is reached before τ^{*}_{θ} with a strictly positive probability.

We have now shown that $\widetilde{x}\left(q\right)$ can be implemented by the simple posted price $P^{s}\left(q\right)$ if and only if $\left(P^{s}\right)'\left(q\right)\geq 0$ for all $0\leq q\leq q_{1}$ and the proof is complete.

□

G.3 Revenue maximizing designer: proof of Proposition 8

Proof. Here, we show how to derive the virtual valuation representation for the designer's value. Suppose transfers follow an arbitrary policy, P_{τ}, adapted to $\mathcal{F}_{t}$. We denote the realization of P_{τ} at time t with p_{t}. The designer's expected revenue can be written as

$$\begin{aligned}
&\int_{\underline{\theta}}^{\overline{\theta}}\mathbb{E}\Big[e^{-r\tau(\theta)}\Big(x_{\tau(\theta)}v_{H}(\theta)+(1-x_{\tau(\theta)})v_{L}(\theta)-W(\theta,x)\Big)\Big]f(\theta)d\theta\\
&=\int_{\underline{\theta}}^{\overline{\theta}}\Big(\mathbb{E}\Big[e^{-r\tau(\theta)}\Big(x_{\tau(\theta)}v_{H}(\theta)+(1-x_{\tau(\theta)})v_{L}(\theta)\Big)\Big]f(\theta)d\theta\\
&-\int_{\underline{\theta}}^{\overline{\theta}}\mathbb{E}\Big[\int_{\underline{\theta}}^{\theta(q)}e^{-r\tau(s)}\Big(x_{\tau(s)}v'_{H}(s)+(1-x_{\tau(s)})v'_{L}(s)\Big)ds\Big]f(\theta)d\theta,
\end{aligned}$$

where we have used the envelope theorem for the agent's value (see Section 4 in the main text).

We can use Fubini's theorem to change the order of integration in the second term:

$$
\begin{aligned}
&\int_{\underline{\theta}}^{\overline{\theta}} \mathbb{E}\Big[\int_{\underline{\theta}}^{\theta(q)} e^{-r\tau(s)}\Big(x_{\tau(s)}v'_H(s) + (1-x_{\tau(s)})v'_L(s)\Big)ds\Big]f(\theta)d\theta \\
=&\mathbb{E}\Big[\int_{\underline{\theta}}^{\overline{\theta}}\int_{s}^{\overline{\theta}} e^{-r\tau(s)}\Big(x_{\tau(s)}v'_H(s) + (1-x_{\tau(s)})v'_L(s)\Big)f(\theta)d\theta ds\Big] \\
=&\mathbb{E}\Big[\int_{\underline{\theta}}^{\overline{\theta}} e^{-r\tau(s)}\Big(x_{\tau(s)}v'_H(s) + (1-x_{\tau(s)})v'_L(s)\Big)(1-F(s))ds\Big].
\end{aligned}
$$

The rest is simply to plug the above expression back into the designer's payoff and to write the integral over quantities rather than types (where we use $1 - F(\theta(q)) = q$). The profit maximizing designer's objective becomes

$$
\mathbb{E}\Big[\int_0^1 e^{-r\tau(q)}(x_{\tau(q)}\phi_H(q) + (1-x_{\tau(q)})\phi_L(q))dq\Big],
$$

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

$$
\phi_\omega(q) := v_\omega(\theta(q)) - v'_\omega(\theta(q))\frac{1-F(\theta(q))}{f(\theta(q))}.
$$

□

G.4 Durable goods monopoly: proof of Proposition 9

Proof. The existence of $x_b < 1$ follows from the continuity of the policy functions and from the complete information quantity being larger in the competitive market: $q^C(1) > q^M(1)$, where $q^C(1)$ solves $\theta(q^C(1)) = c$ and $q^M(1)$ solves $\theta(q^M(1)) - (1 - F(\theta(q^M(1))))/f(\theta(q^M(1))) = c$.

The existence of $x_a > 0$ follows from the same argument that was used to show that the socially optimal policy is below the decentralized policy (details omitted): 1) To see that $x^M(0) \neq x^C(0)$, observe that the smooth pasting and value matching conditions for both the decentralized and the planner's policies cannot hold simultaneously when we approach $q \to 0$ along x^M. The reason why the same proof works for the monopolist's policy as for the planner's policy is that the monopolist's flow payoff is the same as the social planner's when $q = 0$. 2) To rule out $x^M(0) > x^C(0)$, notice that the monopolist gets strictly positive profits by selling to some small q whenever the initial belief is above $x^C(0) = x^{stat}(0)$. 3) Now, it is enough to use the continuity and monotonicity of the policy functions the same way as in the proof of Proposition 3 in the main text to conclude that there exists $x_a > 0$ such that the monopolists sells more for all beliefs below x_a. □

G.5 Type-dependent informativeness: proof of Proposition 10

Proof. **Part (ii): fanatics.** Suppose $-i'(\theta)/i(\theta) \in [0, -v_L'(\theta)/v_L(\theta)]$.

First, we argue that the socially optimal stopping profile is monotone. We can use the same argument as in Lemma 2 but we need to normalize the informativeness of different agents. We show that for all agents $\theta, \theta' \in [\underline{\theta}, \overline{\theta}]$ such that $\theta > \theta'$ and for all realized stopping times $t, t' \in \mathbb{R}_+$ such that $t \leq t'$,

$$e^{-rt}\frac{v_\omega(\theta)}{i(\theta)} + e^{-rt'}\frac{v_\omega(\theta')}{i(\theta')} \geq e^{-rt'}\frac{v_\omega(\theta)}{i(\theta)} + e^{-rt}\frac{v_\omega(\theta')}{i(\theta')}.$$

When the above condition holds, the planner always wants to implement any information stock process Z so that higher types stop first. The condition is equivalent to $(e^{-rt} - e^{-rt'})(v_\omega(\theta)/i(\theta) - v_\omega(\theta')/i(\theta')) \geq 0$. To see that this is satisfied, notice that $-i'(\theta)/i(\theta) \in [0, -v_L'(\theta)/v_L(\theta)]$ implies that $v_\omega(\theta)/i(\theta)$ is increasing in θ for both $\omega \in \{H, L\}$.

Now, we can use montonicity and define the planner's problem as finding the information stock policy Z that maximixes

$$\mathbb{E}\Big[\int_z^1 e^{-r\tau(s)}(x\hat{v}_H(s) + (1-x)\hat{v}_L(s))ds\Big|x, z; Z\Big], \tag{55}$$

where $\hat{v}_\omega(z) = v_\omega(h^{-1'}(z)))$. Notice that the problem is equivalent to (4).

All assumptions in Section 2 hold when $-i'(\theta)/i(\theta) \in [0, -v_L'(\theta)/v_L(\theta)]$ because then $\hat{v}_\omega(z)$ is decreasing in z for both $\omega \in \{H, L\}$, and hence the claim in the proposition immediately follows for this case.

Part (i): experts. Suppose $i'(\theta) \geq 0$.

The monotonicity of the stopping profile follows directly from Lemma 2 in the experts environment because the planner can always decide not to use the additional information from the higher type θ in the proof of Lemma 2. Therefore, we can use (55) as the planner's objective.

In the experts environment, $\hat{v}_H(z)$ need not be decreasing and therefore we need to verify that the policy function x^* we get from Proposition 2 is increasing in z and therefore defines a boundary policy. All other parts of the proof of Proposition 2 remain unchanged even when the stopping payoffs are not monotone.

To show the monotonicity of x^*, we redo the part of the proof of Proposition 2 that shows that $g(x, z) > 0$.

$$\begin{aligned}
g(x,z) =& x(1-x)\Big[x\Big(\beta'(z)(\beta(z)-1)\hat{v}'_H(z) - ((\beta(z)-1)\beta''(z) - 2(\beta'(z))^2)\hat{v}_H(z)\Big) \\
&+ (1-x)\Big(\beta'(z)\beta(z)\hat{v}'_L(z) - (\beta(z)\beta''(z) - 2(\beta'(z))^2)\hat{v}_L(z)\Big)\Big]/ \\
&\Big[\Big(x(\beta(z)-1)^2\hat{v}_H(z) + (1-x)(\beta(z))^2\hat{v}_L(z)\Big)\beta'(z)\Big] \\
=& x(1-x)\Big[x\Big(\beta'(z)(\beta(z)-1)(v'_H(z) + v_H(z)h^{-1\prime\prime}(z)(h^{-1\prime}(z))^{-1}) \\
&- ((\beta(z)-1)\beta''(z) - 2(\beta'(z))^2)v_H(z)\Big) \\
&+ (1-x)\Big(\beta'(z)\beta(z)(v'_L(z) + v_L(z)h^{-1\prime\prime}(z)(h^{-1\prime}(z))^{-1}) \\
&- (\beta(z)\beta''(z) - 2(\beta'(z))^2)v_L(z)\Big)\Big]/ \\
&\Big[\Big(x(\beta(z)-1)^2 v_H(z) + (1-x)(\beta(z))^2 v_L(z)\Big)\beta'(z)\Big],
\end{aligned}$$

where we use $v_\omega(z)$ for $v_\omega(h^{-1\prime}(z)))$ and use that $\hat{v}'_\omega(z) = v'_\omega(z)h^{-1\prime}(z) + v_\omega(z)h^{-1\prime\prime}(z)$.

The expression is otherwise equivalent to $g(x, q)$ in (9) but with an additional term in the numerator:

$$x(1-x)\beta'(z)h^{-1\prime\prime}(z)(h^{-1\prime}(z))^{-1})\Big[x(z)(\beta(z)-1)v_H(z) + (1-x)\beta(z)v_L(z)\Big]. \tag{56}$$

First, notice that $\beta' < 0$, $h^{-1\prime} < 0$, and $h^{-1\prime\prime} < 0$ where the last part follows from the assumption that we are in the experts environment. The term inside the brackets is negative for all $x < x^E(z)$ and hence (56) is weakly positive at $x^*(z)$. Then, $g(x, z) > 0$ follows from the proof of Proposition 2 for the original model because the additional term only makes it larger.

□