

NOTES AND COMMENTS

THE OPTIMAL LEVEL OF EXPERIMENTATION

BY GIUSEPPE MOSCARINI AND LONES SMITH¹

1. INTRODUCTION

CONSIDER WALD'S CLASSICAL (1947) Bayesian formulation of sequential analysis. A given *decision maker* ($\mathcal{D.M}$) is uncertain about a payoff relevant state of the world, and before deciding, can buy multiple i.i.d. informative signals at constant marginal cost. The $\mathcal{D.M}$ should then purchase one at a time, and act when sufficiently convinced of one state.

We argue that the experimental schedule accelerates when *homo economicus* runs the laboratory in real time. We introduce two standard economic assumptions along these lines that work at cross purposes, but which jointly make sense as a theory of dynamic R&D. First, the $\mathcal{D.M}$ is assumed impatient, which encourages classical (nonsequential) behavior. Second, the cost of information is not only increasing, but strictly convex as well; this fosters spreading purchases across periods, reinforcing sequential behavior. Here, we study the variable-size experiment chosen by the $\mathcal{D.M}$ at each time, faced with such a tradeoff.

This paper has four goals. First, we show how to view the control of variance for a diffusion with uncertain mean as the continuous time extrapolation of discrete time experimentation. By adding a convex cost of raising the diffusion precision, our main modelling contribution is to formulate a simple model of experimentation with explicit and known costs, independent of beliefs or the state of the world. This contrasts with existing models with implicit (opportunity) costs, such as bandits.

Second, our main substantive finding is a robust monotonic character of the experimentation level for any increasing and strictly convex cost function. Two separate static intuitions suggest that the level peaks when the $\mathcal{D.M}$ has central, unfocused beliefs over the state. For one, he is then most uncertain about his action choice. Alternatively, the beneficial posterior belief variance is greatest for middling priors. But consider instead a dynamic perspective: To minimize the present discounted information costs, an impatient $\mathcal{D.M}$ should delay high-intensity experimentation until the end is in sight. So maximal

¹ This is a revised version of the paper circulated 1997–2000 as “Wald Revisited: The Optimal Level of Experimentation.” The current title and revision incorporates comments of the editor and three referees. We have also benefited from comments at the Warwick Dynamic Game Theory Conference, Columbia, Yale, M.I.T., Toronto, Pennsylvania, UC—Davis, Stanford, UC—Santa Cruz, Western Ontario, UC—Santa Barbara, Arizona, Arizona State, Queens, Michigan, Ohio State, Princeton, and more specifically Massimiliano Amarante, Dirk Bergemann, Lutz Busch, Prajit Dutta, Chris Harris, Paul Milgrom, Stephen Morris, Yaw Nyarko, Mike Peters, Sven Rady, Arthur Robson, Peter Sørensen, and Ennio Stacchetti. Finally, we acknowledge helpful feedback from the usenet group sci.math.research—especially, Jon Borwein of Simon Fraser University. Moscarini gratefully acknowledges support from the Yale SSFRE, and Smith from the National Science Foundation (Grant SBR-9711885) for this research.

experimentation intuitively should occur near those focused beliefs where the $\mathcal{D}\mathcal{M}$ quits experimenting and takes an action of high expected value and risk. Notice that as the value is convex in beliefs, it peaks at extreme beliefs, just like the experimentation level.

Our general finding is that the optimal experimentation level n grows in the Bellman value v prior to stopping and acting. This monotonicity of the map $v \mapsto n$ is critical to our analysis, and admits a concrete economic intuition. There are two decisions at each instant: stop or experiment, and then at what level n . The second choice equates the marginal costs and benefits of information: $c'(n) = MB(n)$. In our diffusion setting, the total continuation experimentation benefit TB derives from belief variance and is linear in the level: $MB(n) = MB$ and $TB = nMB = nc'(n)$. So the $\mathcal{D}\mathcal{M}$ acts like a neoclassical competitive firm, producing information at an increasing marginal cost and selling it to himself at the fixed price $c'(n)$. Since postponing the final decision entails a discounting cost, optimal stopping demands that the $\mathcal{D}\mathcal{M}$ equate his producer surplus from experimentation $nc'(n) - c(n)$ to the delay cost rv (given the interest rate r). Intuitively, the $\mathcal{D}\mathcal{M}$ closes down his information firm (i.e. acts) when he cannot generate profits (producer surplus) to justify his capital rental (his deferred action). Since this surplus rises in quantity with convex costs, greater experimentation is needed to generate the higher surplus for a higher value v .

The third goal of the paper is to flesh out the testable implications of the model: dynamics and comparative statics. Beliefs being a martingale, the experimentation outlays are a submartingale (positive drift) iff the flow cost of optimal experimentation is a convex function of beliefs. We show how this occurs for a large class of convex cost functions, including the geometric ones, $c(n) = n^k$. We also explore how the level and costs of experimentation respond to standard parametric shifts. For instance, a more impatient $\mathcal{D}\mathcal{M}$ here often experiments at a higher level, provided he still experiments.

The fourth goal of the paper concerns the economic content: We interpret a special case of our analysis as a pure theory of dynamic R&D. For this, we assume ordered states, with final static payoffs strictly increasing as weight is shifted toward the more favorable states. Here, we show that the value function shares the static payoff monotonicity. Consistent with the high uncertainty and skewness of patent returns (Scherer (1965)), optimal R&D outlays rise and fall as contrasting evidence accumulates, before stopping endogenously. Also, the research level also rises as beliefs shift toward better states, providing support for a commonly observed phenomenon: Research on a potential uncertain investment jumps up following a key discovery, or rises with confidence in the outcome.²

In the R&D setting, the assumption of strictly convex information costs within a period is quite realistic. For one, plausibly not all researchers are equally talented in producing information. More intensive information search then must draw on the efforts of less capable researchers. Secondly, since contemporaneously produced knowledge is based on the same current stock, identical or similar discoveries are not rare: Even if research laboratories are created at constant cost, different labs may well expend resources duplicating results. Likewise concurrent Bayesian information tends to be correlated, as it grows increasingly hard to produce i.i.d. signals. Then note that a constant marginal cost for correlated information intuitively corresponds to an increasing marginal cost of independent information.

² A rare empirical study of project-level R&D expenditures in the pharmaceutical industry (DiMasi, Grabowski, and Vernon (1995)) reveals a pattern consistent with this prediction: the FDA protocol for pre-approval clinical testing dictates three sequential phases, of growing size. Of course a major concern is side-effects, well captured by the 'bandit' model where experimentation costs depend on the state. Our model better emphasizes the time-to-market cost, in terms of potentially saved lives.

Our model may also shed light on the purchase of consulting services. Legal, financial, and medical services may be bought solely to make a more informed one-shot decision—whether to file a lawsuit, what portfolio allocation to adopt, which surgical procedure to undergo. Typically, a consultant's hourly fee is known *ex ante*, and rises with his reputation and prestige, as presumably does the quality of the purchased information.

The experimentation literature has by and large focused on 'bandit' models (i.e. broadly defined), where stage payoffs are also random signals, so that information gathering incurs an endogenous opportunity cost. We explore an experimentation problem, inspired by the statistical literature on sequential hypotheses testing, that is different in two key respects: first, information is explicitly costly, since the state-independent information costs are known and so uninformative; second, there is eventual stopping, so that delay cost drives all the results. We characterize the optimal sample size given convex costs and discounting. For this framework, Cressie and Morgan (1993) prove that *some* variable sample size is optimal, while pure sequentiality is best with superadditive costs and no discounting.

We apply continuous time mathematical methods that have recently afforded sharp characterization in 'bandit' models. Such work implicitly links the value function to the experimentation level—namely, the sacrifice of current payoffs for information. Bolton and Harris (1999) (henceforth BH) is the bandit model that is formally closest to our new framework. In BH, strategic bandit players choose the frequency of playing a risky and observable Brownian motion arm (high or low drift) rather than a safe arm with certain drift. By their Lemma 5 and Theorem 8, both the planner's and each player's optimal rate of playing the risky arm is a nondecreasing step function of the value. But playing the risky arm only constitutes experimentation for low beliefs in the high drift, where the risky arm is myopically dominated. In fact, as this belief rises from 0 to 1, the reverse monotonicity finding holds in BH: The value rises, whereas the payoff sacrifice from taking the risky arm falls, and eventually becomes negative. Next, Keller and Rady (1999) study a monopoly pricing model with a randomly shifting demand curve. Moving either direction away from the belief supporting the confounding quantity where the demand curves cross (their Figure 5), experimentation rises and then falls, whereas the convex value monotonically rises.

The Research and Development literature has generally modeled R&D as a stochastic resource allocation exercise: The value of the final prize is independent of experimental evidence, so that the optimal time pattern of R&D outlays is predetermined by features of the model. For instance, in Kamien and Schwartz (1971), a possibly uncertain distance from project completion is covered via an ad hoc knowledge technology; in Grossman and Shapiro (1986), this technology is stochastic, and discounting leads one to postpone costly R&D, as in our framework. In the Poisson R&D model of patent races (Reinganum (1981)) and growth (Aghion and Howitt (1992)), the discovery process is memoryless; when the Poisson completion rate is uncertain (Malueg and Tsutsui (1997)), beliefs and expenditures are driven only in one direction by the nonoccurrence of events. Our approach is closest to Roberts and Weitzman (1981): They too cast R&D as optimally learning about the value of a project and study an optimal stopping problem, but with an exogenous experimentation level and continuation cost.

For simplicity and expositional ease, the paper focuses on the two-action, two-state model, where we prove value-level monotonicity, establish the cost submartingale, and do sensitivity analysis. Section 5 proves value-level monotonicity in the multiple state learning model; it also extends the R&D model to this more general context, proving that experimentation is increasing in beliefs. While many proofs are appendicized, we adopt

a deliberately informal handling of the continuous time control technicalities; we refer interested readers to our 1998 working paper (MS98) for proofs.

2. THE MODEL

Terminal Payoffs. Until Section 5, we consider a two-action, two-state world. The $\mathcal{D}\mathcal{M}$ must choose between actions $a = A, B$, paying π_a^θ in states of the world $\theta = L, H$. To avoid trivialities, we assume that no action is weakly dominant, with $\pi_B^H > \pi_A^H$ and $\pi_B^L < \pi_A^L$, and WLOG that action B does strictly better in state $H : \pi_B^H > \pi_B^L$. The $\mathcal{D}\mathcal{M}$ does not know the state, but is assumed Bayesian and risk neutral, or that payoffs are in utils. If the $\mathcal{D}\mathcal{M}$'s prior belief on H is p , then his expected payoff for action a is affine in p , say $\pi_a(p) \equiv p\pi_a^H + (1-p)\pi_a^L$. The $\mathcal{D}\mathcal{M}$ is indifferent between A and B at some belief \hat{p} .

The special case with one safe action captures a stylized R&D model. Action B means 'building' a costly new prototype, and action A 'abandoning' it. If the $\mathcal{D}\mathcal{M}$ builds, it might not work: payoffs are $h = \pi_B(1) > 0$ or $\ell = \pi_B(0) < 0$ in states H, L . The $\mathcal{D}\mathcal{M}$ earns zero regardless if he abandons: $\pi_A^\theta = 0$ for $\theta = L, H$. So $\pi(p) \equiv \max\langle 0, hp + \ell(1-p) \rangle$ is increasing in p , strictly so for all beliefs $p > \hat{p} \equiv -\ell/(h-\ell)$, where the $\mathcal{D}\mathcal{M}$ invests.

More generally, with payoff discounting, if the current expected stopping payoffs are ever negative, then the $\mathcal{D}\mathcal{M}$ can still ensure himself a zero payoff by never deciding. In other words, the general payoffs for two states and two actions are $\max\langle \pi_A(p), \pi_B(p), 0 \rangle$. Technicalities here with a zero-flat in the static payoffs are fully captured by the R&D model, so that we simply assume that $\pi(p) \equiv \max\langle \pi_A(p), \pi_B(p) \rangle$ is nonnegative on $[0, 1]$.

Information. Before choosing an action, the $\mathcal{D}\mathcal{M}$ can purchase informative signals of the state θ . While variable intensity experimentation is easily understood in discrete time, it yields an intractable dynamic programming problem for the simplest signal structures.

Our observation process is a controlled diffusion $\langle \bar{x}_t \rangle$. Like BH, we parameterize the state by the drift of $\langle \bar{x}_t \rangle$, since this is simplest and corresponds to many applications. We model experimentation as the pure control of variance of $\langle \bar{x}_t \rangle$. For a fixed control, $\langle \bar{x}_t \rangle$ is a Brownian motion with constant uncertain drift chosen by Nature, μ^θ in state θ , where $\mu^H = -\mu^L = \mu > 0$. The $\mathcal{D}\mathcal{M}$ controls its flow variance σ^2/n_t .³ So $\langle \bar{x}_t \rangle$ solves the stochastic differential equation (SDE)

$$(1) \quad d\bar{x}_t^\theta = \mu^\theta dt + \frac{\sigma}{\sqrt{n_t}} dW_t$$

in state θ , where $W_t \sim N(0, t)$ is a Wiener process, with increments $W_t - W_s$ independent of θ (for $s < t$), and $n_t > 0$ is the experimentation level or intensity. As an extrapolation of discrete time sampling, this makes sense: Thinking of \bar{x}_t as a running sample mean of observations (sufficient for the unknown mean μ^θ), n_t is the number of draws at time t . Then doubling n_t halves the 'variance' $(\sigma^2/n_t)dt$ of $d\bar{x}_t^\theta$, yielding a time- t experiment that is doubly informative. The control n_t depends on the observation and intensity history $\langle \bar{x}_s, 0 \leq s \leq t \rangle \cup \langle n_s, 0 \leq s < t \rangle$, and so is a feedback and not an open loop (time-0) control.

Information Costs. Experimentation intensity level n incurs a flow cost $c(n) \geq 0$. We assume throughout that the cost function $c(n)$ is twice differentiable on $(0, \infty)$, as well as increasing ($c' > 0$) and strictly convex ($c'' > 0$) on $[0, \infty)$. It then has a right derivative $c'(0+)$, given $c(0) \geq 0$. One might venture that running a lab incurs a daily rent, for any experimentation level. In this case, there is a positive flow experimentation cost $c(0) > 0$.

³ In BH, the mixing over bandit arms induced a simultaneous linear control over mean and variance of the payoffs, and thereby a pure control of belief variance, as we find below.

Twice differentiability is merely a simplifying technical simplification. While *strict* cost convexity is a real assumption, weak convexity alone is WLOG in continuous time. For one can always rapidly ‘chatter’ between two experimentation levels. We do not delve into a technical proof of this assertion (due to Paul Milgrom) as it detracts from our focus.

A key function in our analysis is $g(n) \equiv nc'(n) - c(n)$. While g is strictly increasing by $g'(n) = nc''(n) > 0$, we must further assume that $\lim_{n \rightarrow \infty} g(n) > r \max_{a,\theta} \pi_a^\theta$. Clearly, it suffices that g be unbounded—in other words, that $nc''(n)$ not be summable.

The Dynamic Maximization. The \mathcal{DM} is impatient, discounting payoffs at an interest rate $r > 0$, and maximizing the expected final reward less costs incurred. Assuming constant unit costs and no time preference, Chernoff (1972) solves the problem of estimating the bivariate drift of a Brownian motion. This is a pure optimal stopping exercise, where the \mathcal{DM} quits at a stopping time T , choosing action A for posterior $p_T = \underline{p}$, and action B if $p_T = \bar{p}$. Our \mathcal{DM} must also solve an optimal control exercise, finding the experimentation time-path $\langle n_t \rangle$.

3. THE OPTIMAL LEVEL OF EXPERIMENTATION

3.1. The Recursive Formulation and Solution

We prove in MS98 that the current posterior belief p_t is a sufficient statistic for observed history, and that a Markov control (measurable in p only) suffices. We now state the law of motion for the belief state variable. Intuitively, the observation process $\langle \bar{x}_t, n_t \rangle$ induces a diffusion belief process $\langle p_t \rangle$. Given a prior p_0 , beliefs $\langle p_t \rangle$ evolve via Bayes rule in continuous time. Since the \mathcal{DM} does not know the true drift, he cannot observe the true noise process $\langle W_t \rangle$ that drives the observation process in (1). If $\zeta = (\mu - (-\mu))/\sigma = 2\mu/\sigma$ denotes the *signal-to-noise ratio factor* of $\langle \bar{x}_t \rangle$, we may use Theorem 9.1 of Lipster and Shirayev (1977) to find, just as in BH:

$$(2) \quad p_t = p_0 + \int_0^t p_s(1 - p_s)\zeta\sqrt{n_s}d\bar{W}_s$$

where $d\bar{W}_s \equiv (\sqrt{n_s}/\sigma)(d\bar{x}_s - [p_s\mu + (1 - p_s)(-\mu)]ds)$ defines the observation-adapted Wiener innovation process; $\langle \bar{x}_t \rangle$ is the ex ante unconditional observation process, or ex post observed path. Beliefs $\langle p_t \rangle$ are an unconditional martingale, with traps at 0 and 1.

Substituting the ex post observed history $\langle \bar{x}_t, n_t \rangle$ into these formulae reveals how beliefs are computed. The \mathcal{DM} updates beliefs upward in favor of $\mu > 0$ ($d\bar{W}_t > 0$) iff the observation process rises faster than he expects, i.e. iff $d\bar{x}_t > [p_t\mu + (1 - p_t)(-\mu)]dt$.

The \mathcal{DM} maximizes his expected discounted return less experimentation costs. Let $V(p_0)$ be the *supremum value* with respect to the stopping time T and experimentation schedule $\langle n_t \rangle$. We use the filter (2) at time $t = T$ to express V as

$$(3) \quad V(p_0) = \sup_{T, \langle n_t \rangle} \mathbb{E} \left[\int_0^T -c(n_t)e^{-rt} dt + e^{-rT} \pi \left(p_0 + \int_0^T p_t(1 - p_t)\zeta\sqrt{n_t}d\bar{W}_t \right) \middle| p_0 \right].$$

Standard for optimal learning, V is convex. That stopping occurs when V coincides with the piecewise linear payoff function π logically mandates threshold stopping rules:

LEMMA 1: *Always $V(p) \geq \pi(p)$, with $V(p) > \pi(p)$ iff $p \in (\underline{p}, \bar{p})$ for cut-offs $0 \leq \underline{p} \leq \bar{p} \leq 1$.*

The \mathcal{D}, \mathcal{M} selects action A for $p \leq \underline{p}, B$ for $p \geq \bar{p}$, and experiments at level $n(p)$ for $p \in \mathcal{D} = (\underline{p}, \bar{p})$, the open experimentation domain \mathcal{D} . Optimization (3) becomes an optimal control exercise for the schedule $n(p)$ and an optimal stopping problem for boundaries \underline{p}, \bar{p} .

We now develop equations for the recursive value v . By Lemma 2, v exists and coincides with the supremum value, or $v = V$. It can thereafter unambiguously be referred to as the value. Since beliefs $\langle p_t \rangle$ are a martingale obeying (2), for any given experimentation region \mathcal{D} , the Hamilton-Jacobi-Bellman (HJB) equation for the control problem is

$$(4) \quad rv(p) = \sup_{n>0} \{-c(n) + n\mathcal{E}(p)v''(p)\}$$

where $\mathcal{E}(p) \equiv p^2(1-p)^2\zeta^2/2$ measures ‘belief elasticity,’ plus the value matching condition:

$$(5) \quad v(\underline{p}) = \underline{p}\pi_A^H + (1-\underline{p})\pi_A^L, \quad v(\bar{p}) = \bar{p}\pi_B^H + (1-\bar{p})\pi_B^L.$$

The so-called ‘generalized Stefan’ problem for the optimal stopping problem given the control policy $n(p)$ is $rv(p) = -c(n(p)) + n(p)\mathcal{E}(p)v''(p)$, plus the value matching (5) and smooth pasting (6) conditions, that the value v is tangent to the static payoff function π at \underline{p} and \bar{p} :

$$(6) \quad v'(\underline{p}) = \pi_A^H - \pi_A^L, \quad v'(\bar{p}) = \pi_B^H - \pi_B^L.$$

While the above functional problem cannot be solved in closed form, we can still fully characterize its solution $\{v(p), \underline{p}, \bar{p}\}$ without knowing the closed form.

LEMMA 2 (Value and Policy Existence/Uniqueness/Verification):

- (a) *There exists a unique solution (p, \bar{p}, v) of (4)–(6), having strictly interior thresholds $0 < \underline{p} < \bar{p} < 1$, and a strictly convex value $v \in \mathcal{C}^2$ in \mathcal{D} , so that we have $v'' > 0$.*
- (b) *The solution v coincides with the supremum value V of (3).*
- (c) *Let producer surplus $g(n) = nc'(n) - c(n)$ have inverse $f \equiv g^{-1}$. Set $n(p) = f(rv(p))$. The solution $\{n(\cdot), \underline{p}, \bar{p}\}$ of (4)–(6) is the unique optimal control and stopping policy.*
- (d) *The level $n(p)$ is continuous and differentiable in \mathcal{D} ; it is \mathcal{C}^1 when $c(n)$ is \mathcal{C}^2 .*

PROOF SKETCH: For parts (a) and (b), fix a domain \mathcal{D} . The FOC for (4) is $c'(n) = \mathcal{E}(p)v''(p)$; the SOC is met because $-c(n) + n\mathcal{E}(p)v''(p)$ is strictly concave in n . The solution $n(p) < \infty$ then uniquely exists, as $rv(p) = -c(n) + n\mathcal{E}(p)v''(p) = -c(n) + nc'(n) = g(n)$ is solvable in n . Indeed, g is continuous and strictly increasing; also, $g(0) = -c(0) \leq 0 \leq r\pi(p) \leq rv(p)$ for $p \in \mathcal{D}$, since π is assumed nonnegative, and $g(n) > r \max(\pi(0), \pi(1)) \geq rv(p)$ for large n , by assumption. Finally, $rv(p) = g(n(p))$ implicitly defines $n(p) = f(rv(p))$, for the strictly increasing inverse function $f \equiv g^{-1}$.

We have just transformed (4)–(6) into the equivalent two-point free boundary value problem—namely, $v''(p) = c'(f(rv(p)))/\mathcal{E}(p)$, plus (5) and (6). MS98 proves that a unique solution exists, so that (4)–(6) is uniquely soluble, and that it has thresholds $0 < \underline{p} < \bar{p} < 1$ (part (a) here). Applying the standard value verification theorems for optimal control (Theorem 11.2 in Oksendal (1995)) and optimal stopping (Theorem 3.15 in Shirayev (1978)), MS98 shows this solution is the supremum value (part (b) here).

For (c), if either $\pi(p) > 0$ always or $c(0) > 0$, then the optimal control policy $n(p) = f(rv(p))$ is bounded away from zero. The filter SDE (2) is then nondegenerate and the

stopping time is finite a.s. The result then follows from the verification and uniqueness for stochastic control Theorems 11.1 and 11.2 in Oksendal (1995). For optimal stopping with the a.s. finite horizon T , we apply Theorem 3.3 in Shirayev (1978). See MS98 for details.

When $\pi(\underline{p}) = c(0) = 0$ obtains, and thus $n(\underline{p}) = f(0) = 0$, then provided control policy $n(p) = f(rv(p))$ vanishes fast enough at \underline{p} , the stopping time is infinite with positive chance (by Feller’s test for explosions; see MS98). Intuitively, the cost of a little experimentation is so small that it is worth keeping the project running forever, as beliefs may rise to \bar{p} . By standard verification theorems (see MS98), stopping near \underline{p} is ε -optimal. In fact, stopping only at \bar{p} is optimal. In the event that $p_t \rightarrow \underline{p}$, never stopping until \bar{p} yields a zero expected discounted payoff at \underline{p} , given the infinite hitting time. Since $\pi(\bar{p}) = 0$, the payoff to the limit of the ε -optimal policies—namely, only stop at \bar{p} or \underline{p} —equals the limit of the payoffs of these policies, and is therefore optimal.

For part (d), since c'' exists, so does $g'(n) = nc''(n)$. So $f = g^{-1}$ is differentiable, as is $n(p) = f(rv(p))$. If c'' is continuous, then so is g' , and thus f' and n' too. *Q.E.D.*

Notice how the marginal benefit of experimentation $MB = \mathcal{E}(p)v''(p)$ is constant in n . It factors into the ex ante variance of posterior beliefs induced by experimentation, i.e. the belief elasticity, and a payoff conversion factor, the convexity of the value function $v''(p)$. So strict cost convexity drives the unique solution. In fact, the above analysis intuitively shows that any convex, possibly piecewise linear, cost function with large enough surplus would yield a well-defined but discontinuous experimentation policy. With sufficiently bounded surplus (e.g., a linear cost function), achieving perfect information over a vanishing time interval $[0, \Delta]$ is preferred with discounting.

3.2. The Optimal Level of Experimentation

One of our main findings follows from Lemma 2-(c) and (d) and $c'' > 0$.

PROPOSITION 1 (Experimentation is Monotone in the Value): *The experimentation level $n = f(rv)$ is strictly increasing in $v(p)$ for $p \in \mathcal{D}$, and weakly exceeds $f(0) \geq 0$, where $f(0) = 0$ iff $c(0) = 0$.*

Illustrating Proposition 1, any geometric convex cost function $c(n) = n^\gamma (\gamma > 1)$ yields a surplus $g(n) = (\gamma - 1)n^\gamma$. So the optimal level $n = f(rv) = [rv/(\gamma - 1)]^{1/\gamma}$ is increasing and concave in the return. In particular, $n(p) = \sqrt{rv(p)}$ with quadratic costs $c(n) = n^2$.

Since v shares the shape of the static payoff π by convexity, value matching, and smooth pasting, the experimentation level $n = f(rv)$ also shares that shape. In particular, we have Proposition 2.

PROPOSITION 2 (Experimentation versus Beliefs): *If static payoffs $\pi(p)$ are increasing (resp., decreasing, U-shaped) in p in region \mathcal{D} , so is the experimentation level $n(p)$.*

For some context, $\pi(p)$ increases in p in the R&D model, and thus the research level rises as we approach confidence in the ‘build’ decision (see Figure 1, left panel).

Write $v(p) = \pi(p) + I(p)$, where $I(p)$ is the *expected present value of information*. Think of $I(p)$ as the option value of waiting and choosing an action after optimally experimenting. This option to change one’s plan is least valuable when one is most sure of an action to take. As noted in the introduction, the option value of information is hill-shaped. But the return to experimentation depends also on the terminal payoff $\pi(p)$.

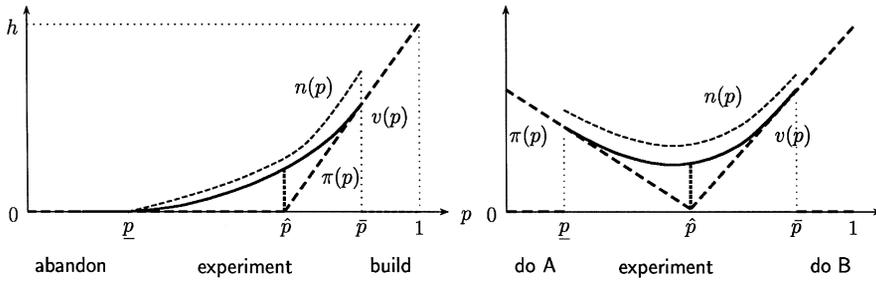


FIGURE 1.— Value Function and Experimentation Demand. Overusing the vertical axis, we depict both the static payoff function π (thick dashed line) and dynamic value function v (solid line), strictly convex in the experimentation domain $\mathcal{D} = (p, \bar{p})$, and the intensity level n (thin dashed line). The demand is increasing in v . The R&D model is illustrated on the left, and a more general decision model (no null action) on the right. The option value of experimentation—the vertical distance between the static and dynamic values—is maximized at the π kink \hat{p} .

PROPOSITION 3: *The information value $I(p)$ is single-peaked, maximized at the kink \hat{p} in π .*

PROOF: By value matching and smooth pasting (5)–(6), $v - \pi_A$ rises on $(p, 1]$, and $v - \pi_B$ falls on $[0, \bar{p})$. So $v - \pi$ is rising until π_A and π_B cross, and later falling. *Q.E.D.*

We note that the minima of the static payoff function π and value function v need not coincide. Consider the R&D example: An arbitrarily small shift in payoffs can rotate $\pi_A(p)$ clockwise, leaving \hat{p} as the strict global minimum of π ; however, the unique minimum (by strict convexity) of the value function only moves slightly away from \underline{p} .

4. TESTABLE IMPLICATIONS

4.1. Drift in the Level and Flow Costs of Experimentation

Since the belief process $\langle p_t \rangle$ is a martingale diffusion, and v is twice differentiable, the value process $\langle v(p_t) \rangle$ is a strict submartingale (drifts up) in \mathcal{D} , because $v'' > 0$. Likewise, MS98 shows that the expected remaining time until stopping $\tau(p)$ is a hill-shaped function of beliefs in \mathcal{D} , so that $\tau''(p) < 0$: Thus, $\langle \tau(p_t) \rangle$ is a strict supermartingale (drifts down). By the same reasoning, experimentation levels $\langle n(p_t) \rangle$ are a submartingale if $n(p)$ is convex. But $n = f(rv)$ may well be concave in rv , as when $c(n) = n^2$. Since v is strictly convex in $p \in \mathcal{D}$ by Proposition 2, the convexity of $n(p) = f(rv(p))$ is then unclear.

PROPOSITION 4 (Experimentation Time Series Properties):

- (a) *If the surplus $g(n) = nc'(n) - c(n)$ is concave, or $nc''(n)$ is nonincreasing in n , then $n(p)$ is strictly convex in $p \in \mathcal{D}$, and thus the experimentation level $\langle n(p_t) \rangle$ is a submartingale.*
- (b) *If $nc''(n)/c'(n)$ is nonincreasing in n , then the cost process $\langle c(n(p_t)) \rangle$ is a submartingale.*

PROOF: For part (a), the producer surplus g is weakly concave and increasing iff its inverse f is weakly convex and increasing. By Theorem 5.1 of Rockafellar (1970), the com-

position $f(rv)$ of a convex and increasing function f with a strictly convex function rv is strictly convex. Finally, simply remark that g is concave iff $g'(n) = nc''(n)$ is nonincreasing.

For (b), it suffices that $c(f(rv))$ be convex in the return $w = rv$. Put $\kappa(w) \equiv c(f(w)) = c(n)$, where $n \equiv f(w)$. Then $f'(w) = 1/g'(f(w)) = 1/nc''(n)$ implies $\kappa'(w) = c'(f(w))f'(w) = c'(n)/nc''(n)$. So κ is convex iff $\kappa'(w)$ is nondecreasing iff $nc''(n)/c'(n)$ is nonincreasing. *Q.E.D.*

Proposition 4 only offers sufficient conditions. For instance, the level $n(p)$ may be strictly convex even if surplus is slightly convex, since the value function is strictly convex.

Observe also that the conditions for costs to drift up are weaker than those for levels to drift up, since $c''(n) > 0$. Geometric cost functions fail the test in part (a) but meet the test in part (b): Any convex geometric cost function $c(n) \equiv n^\gamma$ with $\gamma > 1$ yields a convex and not concave producer surplus, since $[nc''(n)]' = \gamma(\gamma - 1)^2 n^{\gamma-2} > 0$. But by analogy to CRRA utility functions, this class of functions lies at the knife-edge for part (b), since $nc''(n)/c'(n)$ is constant. One knife-edge cost function for part (a) with linear surplus, and $g'(n) = nc''(n)$ constant, is $c(n) = 1 + n \log n$ for $n > 1$, and $c(n) = n$ for $n \leq 1$.

4.2. Sensitivity Analysis

We now explore how changes in payoffs, costs, or the interest rate affect the level of experimentation. Our main result Proposition 5 is intuited, and proved in the Appendix.

A method of tangents is developed in Appendix A and illustrated in Figure 2 (left panel) for the R&D model. When payoff ℓ rises, so does the value v , and hence the intensity level n . As is economically intuitive, the threshold \underline{p} and \bar{p} must fall: When the reward is higher, one is indifferent about adopting or quitting when slightly less optimistic, so that \bar{p} or \underline{p} both shift in the same direction—but up, if h had fallen.

Since the sole reason to pay for information is uncertainty over the state of the world, a natural thought experiment stems from raising the payoff risk (Figure 2, right panel). In other words, let h rise and ℓ fall so as to maintain a constant expected payoff $\pi_B(p)$ at the current belief p . Surely, this ought to raise the value of the dynamic problem, since the \mathcal{DM} should prefer a riskier final payoff distribution, for the static payoff frontier in the current stopping set goes up. Then by employing the same intensity level decision and stopping rule, the \mathcal{DM} 's expected payoff increases. By re-optimizing, he does no worse. Hence the return rv and thus the experimentation level n both rise.

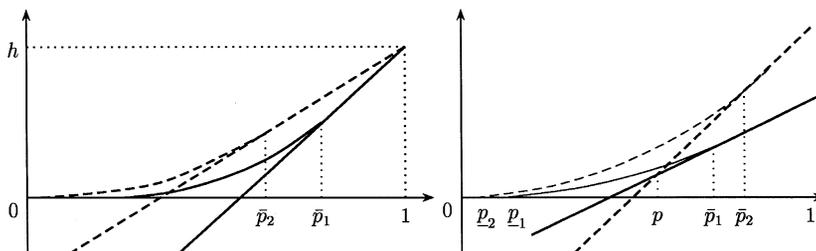


FIGURE 2.— Payoff Shifts in the R&D Model. At left: When the (unplotted) bad build payoff ℓ rises, so does the value function, and intensity levels (inside the now left-shifted interval (\underline{p}, \bar{p})). At right: When payoffs grow riskier—the static payoff frontier rotates through current belief p to the dashed line—the value (and hence intensity) rises, and thresholds shift out.

A more informative observation process raises the value and hence the experimentation level. Less convex and weakly lower costs also raise the \mathcal{DM} 's value, and similarly the intensity level. Here, there is also an indirect effect that the surplus function shifts down, and so a higher intensity n is needed to generate the same surplus. Finally, with a higher interest rate r , the \mathcal{DM} is more eager to stop and act, and therefore enjoys a lower expected payoff: Thresholds shift in. Less obviously, the value often falls proportionately less than the interest rate r rises, so that the return rv , and thus the intensity level n , rises.

PROPOSITION 5:

- (a) *Payoff Levels: The value $v(p)$ and experimentation level $n(p)$ shift up for $p \in [\underline{p}, \bar{p}]$ when any payoff π_a^0 rises. Thresholds \underline{p}, \bar{p} rise if π_A^H or π_B^H rises, and fall if π_A^L or π_B^L rises.*
- (b) *Payoff Risk: If payoffs grow riskier (the expected payoff $\pi_a(p)$ remains constant for an action a at belief p , but the payoff spread $|\pi_a^H - \pi_a^L|$ increases), then the value $v(p)$ rises, the thresholds shift out (\underline{p} falls and \bar{p} rises), and the experimentation level $n(p)$ rises.*
- (c) *Cost Convexity: Assume that the cost function grows more convex and initially weakly higher and steeper—namely, $c(n)$ is replaced by $\hat{c}(n)$, satisfying all our assumptions on the cost function, and $\hat{c}(0) \geq c(0)$, $\hat{c}'(0) \geq c'(0)$, and $\hat{c}''(n) \geq c''(n)$ for all $n \geq 0$. Then thresholds shift in, and the value $v(p)$ and hence intensity level $n(p)$ uniformly fall in \mathcal{D} .*
- (d) *Information Quality: As the signal-to-noise ratio factor ζ rises, the value $v(p)$ and the experimentation level $n(p)$ shift up, while the thresholds shift out.*
- (e) *Impatience: As the interest rate r rises, the value $v(p)$ falls, and thresholds shift out. Also, the optimal intensity level $n(p)$ rises strictly near one or both thresholds in \mathcal{D} . In the R&D model, $n(p)$ declines for all $p < p'$, and rises for all $p > p'$, for some $p' \in (\underline{p}, \bar{p})$.*

This final impatience result differs from the standard conclusion of Bayesian learning. More impatient decision makers typically ‘experiment’ less with more myopic actions. Here, greater impatience raises the \mathcal{DM} 's delay cost, and sometimes leads him to acceleration of his experimentation schedule. Proposition 6 asserts one nonstandard result: As r blows up, the \mathcal{DM} experiments at an exploding rate—albeit over a vanishing belief interval. Normally, there is never-ending experimentation, but here one eventually stops and acts.

Our analysis in the paper so far remains valid if the final payoff is an *annuity*—i.e. where $\pi(p_T)$ is an eternal flow payoff rather than a one-shot lump-sum, as we have assumed.⁴ What happens in this case is that the final decision is formally very much like the safe uninformative arm in a bandit model: It provides a constant flow current payoff, and is therefore exercised when equal to the current value. So with a higher interest rate r , the annuity is worth less, and the intensity level falls everywhere. Indeed, the Bellman equation for maximizing $\mathbb{E}[\int_0^T -c(n_t)e^{-rt} dt + e^{-rT} \pi(p_T)/r]$ is still (4). But in terms of the return $w = rv$, it becomes $w(p) = \max_{n \geq 0} \{-c(n) + n\mathcal{E}(p)w''(p)/r\}$. Hence, $n(p) = f(w(p))$, and w solves $w''(p) = rc'(f(w(p)))/\mathcal{E}(p)$. The value matching and smooth pasting conditions are still $w(\underline{p}) = \pi(\underline{p}), w(\bar{p}) = \pi(\bar{p}), w'(\underline{p}) = \pi'(\underline{p}), w'(\bar{p}) = \pi'(\bar{p})$. A higher interest rate r is then formally equivalent to a lower signal-to-noise ratio factor ζ ; thus, this diminishes the return $w(p)$, and $n(p) = f(w(p))$, by the logic of Proposition 5.

Finally, we return to the classical benchmark case of Wald. As discounting and cost convexity vanish, we recover Wald's sequential optimal policy in our model. As noted,

⁴We are grateful to Sven Rady for suggesting this payoff formulation, and explaining its implications.

these assumptions cut in opposite ways. Absent payoff discounting, the $\mathcal{D}\mathcal{M}$ sees no hurry to stack experiments, and reverts to a purely sequential mode (vanishing intensity levels) barring any fixed flow cost $c(0) > 0$. Without strict cost convexity, the $\mathcal{D}\mathcal{M}$ faces no parallel experimentation penalty, and converges upon a massive experiment at time-0.

PROPOSITION 6: *Assume an everywhere positive final payoff $\pi(p) > 0$.*

(a) *Vanishing Impatience: For fixed $c(n)$, the intensity level $n(p)$ explodes (where > 0) as $r \uparrow \infty$, and decreases to $f(0) \geq 0$ as $r \downarrow 0$. Thus, $n(p) \downarrow 0$ as $r \downarrow 0$ iff $c(0) = 0$.*

(b) *Vanishing Convexity: Fix $r > 0$. Consider a cost function sequence $c_1(n), c_2(n), \dots$. Let λ_k, Λ_k be the liminf and limsup of $c'_k(n)$, respectively. Then $n_k(p)$ uniformly explodes (where positive) if $\lim_{k \rightarrow \infty} \Lambda_k = 0$, and uniformly vanishes if $\lim_{k \rightarrow \infty} \lambda_k = \infty$.*

PROOF: To see the impatience limits, consider that since $v(p) \leq \max\{\pi_A^H, \pi_B^H, \pi_A^L, \pi_B^L\}$, the return $rv(p)$ vanishes as $r \rightarrow 0$. Because $g(0) = -c_0$, so that $f(-c_0) = 0$, the optimal intensity level must satisfy $n(p) \geq f(0) > 0$ for all $p \in \mathcal{D}$, and so $n(p)$ tends down to $f(0)$.

Likewise, since $v(p) \geq \pi(p) > 0$ for $p \in \mathcal{D}$, $rv(p)$ and $n(p)$ explode as $r \rightarrow \infty$. *Q.E.D.*

See MS98 for the quick proof of the convexity limit. For a simple example, consider $c(n) = n^k$ for $k > 1$; the producer surplus is $g(n) = nc'(n) - c(n) = (k - 1)n^k$. Its inverse $f(n) = [n/(k - 1)]^{1/k}$ blows up as $k \downarrow 1$, as we converge upon Wald's case of linear costs.

5. MANY ACTIONS AND STATES OF NATURE

We now allow for finitely many actions and states of nature. We maintain our assumption that each state θ_i corresponds to a drift μ_i of an observation process like (1). We generalize Proposition 1 (experimentation is increasing in the value), deducing the same functional form $n = f(rv)$. Then, extending the R&D model with a monotone static payoff function, Proposition 2 (monotonicity in beliefs) is generalized.

5.1. Experimentation is Monotone in the Value

Assume $K \geq 2$ states, ordered by their drifts $\mu_1 < \dots < \mu_K$. Let $\vec{p}_t = (p_{1t}, \dots, p_{Kt})$ be the vector of time- t posterior beliefs on $\vec{\mu} \equiv (\mu_1, \dots, \mu_K)$.

PROPOSITION 7: *Independently of payoffs and actions, for all beliefs in the region \mathcal{D} of positive experimentation:*

- (a) *The optimal intensity level is strictly increasing in the value: $n = f(rv)$.*
- (b) *If $g(n) = nc'(n) - c(n)$ is concave, then the intensity level $\langle n(p_t) \rangle$ is a submartingale.*
- (c) *If $nc''(n)/c'(n)$ is nonincreasing in n , then experimentation costs are a submartingale.*

PROOF: Reflecting the importance of signal-to-noise ratios, define the i th normalized drift $\bar{\mu}_i \equiv \mu_i/\sigma_i$, and the expected such drift $\bar{\mu}^{\vec{p}} \equiv \sum_{j=1}^K p_j \bar{\mu}_j$. Theorem 9.1 in Liptser and Shirayev (1977) implies that posterior chance p_{it} of θ_i solves $dp_{it} = p_{it}\sqrt{n_t}(\bar{\mu}_i - \bar{\mu}_{\vec{p}_t})d\bar{W}_t^{\vec{p}_t, n_t}$, and the Wiener process $\langle \bar{W}_t^{\vec{p}_t, n_t} \rangle$ obeys $d\bar{W}_t^{\vec{p}_t, n_t} \equiv (\sqrt{n_t}/\sigma)(d\bar{x}_t - \bar{\mu}^{\vec{p}_t} dt)$, common for all i .

The HJB equation of this higher dimensional problem of control and stopping reads:

$$rv(\vec{p}) = \sup_{n>0} \left[-c(n) + \frac{n}{2} \sum_{i=1}^K \sum_{j=1}^K (\bar{\mu}_i - \bar{\mu}^{\bar{p}^i})(\bar{\mu}_j - \bar{\mu}^{\bar{p}^j}) \frac{\partial^2 v}{\partial p_i \partial p_j}(\vec{p}) \right].$$

Observe how this reduces to (2) for $K = 2$. Just as in the proof of Lemma 2, the level n enters as a linear factor of the quadratic form in \vec{p} on the right-hand side, and hence once more yields a FOC $c'(n) = MB$. Substituting this into the HJB equation again yields $rv = -c(n) + nc'(n)$, or our earlier two state case control equation $n(\vec{p}) = f(rv(\vec{p}))$ —for any number of actions. Static payoffs $\pi(\vec{p})$ only affect the optimal stopping problem.

Finally, for parts (b) and (c), recall that the value process is a submartingale in any Bayesian learning model. So $c(n) = c(f(rv))$ is again a submartingale if (but not only if) the map $c(f)$ is convex, exactly as before. Q.E.D.

For an intuition, note that the continuation value is multiplicatively separable in the variance of posterior beliefs ($2n\mathcal{E}(p)$ in our binary model) and the Hessian of the value function. The variance of posterior beliefs, in turn, is inversely proportional to the variance of the sample mean, and so is linear in n . Hence, the marginal benefit of experimentation is still constant in the level, and the familiar trade-off between cost of waiting and information producer surplus extends to this more general case, and the equation $n = f(rv)$ with it.

A solution to the control/stopping problem via PDE methods is not always feasible, for two reasons. First, the necessity of smooth pasting conditions in many dimensions (i.e. $\nabla v(\vec{p}) = \nabla \pi(\vec{p})$ defining the stopping manifolds) is still an open problem. Second, the control PDE problem plus value matching (still necessary, by Krylov (1980, p. 279–285)) does not afford an easy characterization of the solution. However, our control finding circumvents these difficulties. Replacing $n(\vec{p}) \equiv f(rv(\vec{p}))$ into the earlier belief filter leaves a pure stopping problem of a process that depends on an unknown function v . Solving this stopping problem in many dimensions is a standard exercise. The desired supremum value is the least superharmonic majorant of the static payoff frontier π .

5.2. Experimentation is Monotone in Beliefs

The relationship between values and beliefs is model-specific. For the multi-dimensional analogue of the R&D model, the static payoff frontier must be suitably monotone in beliefs, in a way that can pass through to the value function.

With more states than actions, amalgamate states until each action is optimal in just one state. With more actions than states, our analysis is complicated by insurance actions, which we avoid. Suppose there exist K states and K actions, $K \geq 2$. Assume that action a_j is best in state θ_j , and the best payoff for each action $\pi_{a_j}^{\theta_j}$ is increasing in the action $j = 1, 2, \dots, K$. Next, consider two belief vectors \vec{p}, \vec{q} : Say that \vec{q} strictly dominates \vec{p} in the *Monotone Ratio* (MR),⁵ or $\vec{q} \succ \vec{p}$, if $q_i p_j \geq q_j p_i$ whenever $i \geq j$, with at least one strict inequality. Since beliefs lie in the unit simplex Δ^{K-1} , this implies that for some k' , we have $q_i \leq p_i$ for $i = 1, 2, \dots, k'$, and $q_i \geq p_i$ for $i = k' + 1, \dots, K$.

Crucially, MR is preserved under Bayesian updating. For every $\vec{\lambda} \in \mathbf{R}_+^K$ and $\vec{p} \in \Delta^{K-1}$, define $B_i(\vec{p}, \vec{\lambda}) \equiv p_i \lambda_i / \sum_{j=1}^K p_j \lambda_j$, with corresponding vector $\vec{B}(\vec{p}, \vec{\lambda})$. This posterior belief vector updates the prior \vec{p} after a realization of a signal with likelihood λ_i in state i .

⁵ While clearly conceptually related to the MLRP, we do not use this label, since MLRP applies to signal distributions, while MR is an ordering on beliefs.

LEMMA 3: *Bayes-updating preserves MR: $\vec{q} \succ \vec{p}$ iff $\vec{B}(\vec{q}, \vec{\lambda}) \succ \vec{B}(\vec{p}, \vec{\lambda})$ for all $\vec{\lambda} \in \mathbf{R}_+^K$.*

PROOF: Since $B_i(\vec{q}, \vec{\lambda})/B_i(\vec{p}, \vec{\lambda}) = \beta q_i/p_{i2}$ for $\beta \geq 0$ is independent of i , q_i/p_i is weakly increasing in i if and only if $B_i(\vec{q}, \vec{\lambda})/B_i(\vec{p}, \vec{\lambda})$ is, for all $\vec{\lambda} \in \mathbf{R}_+^K$, i.e. for all $\beta \geq 0$. *Q.E.D.*

The main R&D result asserts a dominance ordering on states, with higher better for expected terminal payoffs, and values. Consequently, in the experimentation domain \mathcal{D} , the optimal experimentation level is likewise ordered. Extending Proposition 2, we have the following result:

PROPOSITION 8: *Assume that the static payoff frontier π is MR-increasing: $\pi(\vec{p}) > \pi(\vec{q})$ if $\vec{p} \succ \vec{q}$. Then the value function v and the optimal intensity level n (when positive) are also MR-increasing in \vec{p} : if $\vec{p} \succ \vec{q}$, then $v(\vec{p}) > v(\vec{q})$, and $n(\vec{p}) > n(\vec{q})$ for $\vec{p}, \vec{q} \in \mathcal{D}$.*

PROOF: We use a policy improvement argument. Let $n(\cdot)$ be the optimal control policy. The optimal recursively controlled posterior belief starting from the prior \vec{q} is described by $q_{it} = q_i + \int_0^t n(\vec{q}_s) q_{is} (\bar{\mu}_i - \bar{\mu}^{\vec{q}_s}) d\bar{W}_s^{\vec{q}_s, n(\vec{q}_s)}$. Consider the following policy starting from beliefs \vec{p} : Follow the control policy $\eta_t \equiv n(\vec{q}_t)$ that is optimal for beliefs \vec{q}_t , namely, $p_{it} = p_i + \int_0^t \eta_s p_{is} (\bar{\mu}_i - \bar{\mu}^{\vec{p}_s}) d\bar{W}_s^{\vec{p}_s, \eta_s}$; quit whenever \vec{q}_t hits the stopping set $\{\vec{q} \in \Delta^{K-1} : v(\vec{q}) = \pi(\vec{q})\}$. For every observed signal path $\langle \bar{x}_s \rangle$ and associated control process $\langle \eta_s \rangle$ on $[0, t]$, we have $p_{it} = \Pr(\mu = \mu_i | \vec{p}, \langle \bar{x}_s, \eta_s, 0 \leq s \leq t \rangle)$ and $q_{it} = \Pr(\mu = \mu_i | \vec{q}, \langle \bar{x}_s, \eta_s, 0 \leq s \leq t \rangle)$. Therefore by Lemma 3, if $\vec{p} \succ \vec{q}$ then $\vec{p}_t \succ \vec{q}_t$ for any signal path \bar{x}' . By assumption this implies $\pi(p_T) \geq \pi(q_T)$ for every stopping time T dictated by the $\langle \vec{q}_t \rangle$ process. So starting from a prior \vec{p} , for each possible signal path, the \mathcal{DM} pays the same experimentation costs and stops at the same time as with the control-stopping policy that is optimal from \vec{q} . Since he obtains a higher terminal reward, the \mathcal{DM} can do no worse from \vec{p} than from \vec{q} a.s., and thus in expectation. By re-optimizing, he can do strictly better. So, $v(\vec{p}) > v(\vec{q})$. *Q.E.D.*

Consider a three-state, three-action example, with payoffs in the three states (3, 4, 5), (2, 5, 8), and (1, 3, 9). Any MR upward shift in beliefs raises the static expected payoff to the corresponding optimal action; hence, the assumptions of Proposition 8 are met, and the value v , and thus experimentation level, are both MR-monotone in beliefs too.

6. CONCLUSION

This paper has introduced and explored a continuous time model of sequential experimentation with explicit information purchases. The driving features are impatience and an increasing and strictly convex cost function of within-period experimentation. These two assumptions have yielded some robust predictions: Experimentation intensity grows with a project's expected payoff. Among falsifiable implications, we have established an upward secular drift in the experimentation level for not too convex cost functions.

Our costly control of variance for a diffusion is also a tractable decision model of information purchases or R&D. By making simple monotonicity assumptions natural in an R&D setting, our main result implies that the level is increasing in the posterior belief. We believe that this can be applied in many settings, such as strategic patent races, equilibrium R&D models, or principal-agent experimentation models.

Dept. of Economics, Yale University, P.O. Box 208268, New Haven, CT 06520-8268, U.S.A.; gm76@pantheon.yale.edu; http://www.econ.yale.edu/~mosca/mosca.html

and

Dept. of Economics, University of Michigan, 611 Tappan, St., Ann Arbor, MI 48109, U.S.A.; econ-theorist@earthling.net; http://www.umich.edu/~lones

Manuscript received April, 1998; final revision received September, 2000.

APPENDIX: PROOF OF PROPOSITION 5

PROOF OF PROPOSITION 5(a)—Increasing Payoff Levels: Write $V(p_0) = \max_{T,n} V(p_0|n, T, \pi_A, \pi_B)$, where the maximand $V(p_0|\cdot)$ —seen on the right-hand side of (3)—is differentiable in π_a^0 by inspection: $V_{\pi_B^0}(p_0|\cdot) = \mathbb{E}[e^{-rT} \bar{p} | p_T = \bar{p}, p_0, n] \Pr(p_T = \bar{p} | p_0, n) > 0$. Here the event $p_T = \bar{p}$ that the $\mathcal{D.M}$ eventually chooses action B occurs with chance $(p_0 - \underline{p})/(\bar{p} - \underline{p}) > 0$, for any $p_0 > \underline{p}$. The other payoff parametric shifts are similarly positive: $V_{\pi_a^0}(p_0|\cdot) > 0$ for all $p_0 > \underline{p}$ for a fixed policy n, T . If the $\mathcal{D.M}$ re-optimizes after a change in one of these parameters, the supremum value V cannot fall.

Consider boundary behavior at lower threshold belief \underline{p} (\bar{p} being similar), associating \underline{p}_i and payoff parameters $\pi_1 > \pi_0$. The value function $V(p|\pi)$ is continuously differentiable in p , by smooth pasting (6), and partially differentiable in π . Hence, $V(\underline{p}_0|\pi_0) - V(\underline{p}_0|\pi_1) \approx V_{\pi}(\underline{p}_0|\pi_0)(\pi_0 - \pi_1)$ is the first-order Taylor expansion, as $V_p(\underline{p}_0|\pi_0) = 0$. So $\underline{p}_0 > \underline{p}_1$. Q.E.D.

PROOF OF PROPOSITION 5(b)—Increasing Riskiness: Rotate the π_B payoff line counterclockwise through (current belief, expected payoff), as π_B^L falls and π_B^H rises. Then the value function cuts into the new π frontier on the right. By Claim 2-b in MS98, \underline{p} must fall to restore a smoothly-pasted tangency of $v_p(p)$ and $\pi(p)$ on the right side. Also, as Claim 3 in MS98 asserts, as \underline{p} falls, the value $v_p(p)$ rises, and gets steeper. Q.E.D.

CLAIM 1 (A Key Implication): *Parameterize models by Γ_1, Γ_2 , where $\Gamma_i = \{c_i(\cdot), \zeta_i, r_i\}$. Let $\{v_i, \underline{p}_i, \bar{p}_i\}$ be the corresponding solutions, with $\Psi_i(p, v_i(p)) \equiv c_i'(f_i(r_i v_i(p))) / \varphi_i(p)$. If*

$$(7) \quad v_2 \geq v_1 \Rightarrow \Psi_2(p, v_2) > \Psi_1(p, v_1)$$

independently of p , then $\underline{p}_1 < \underline{p}_2, \bar{p}_1 > \bar{p}_2$, and $v_1(p) > v_2(p)$ for all $p \in [\underline{p}_2, \bar{p}_2]$.

PROOF: Put $\Delta v \equiv v_2 - v_1$, and similarly $\Delta v', \Delta v''$, all continuous maps $(\underline{p}_1, \bar{p}_1) \cap (\underline{p}_2, \bar{p}_2) \rightarrow \mathbf{R}$.

Step 1—Ordering Thresholds: We assume $\underline{p}_1 \geq \underline{p}_2$, and obtain a contradiction. By a symmetric argument $\bar{p}_1 > \bar{p}_2$. First, we show that $\underline{p}_1 \geq \underline{p}_2$ implies $\Delta v(\underline{p}_1) \geq 0, \Delta v'(\underline{p}_1) \geq 0, \Delta v''(\underline{p}_1) > 0$, so that Δv is strictly positive, increasing, and convex at $\underline{p}_1 + \varepsilon$, for some small $\varepsilon > 0$. The first two weak inequalities follow at once from value matching and smooth pasting of each v_i at \underline{p}_i , along with strict convexity of v_2 .

Next, $\Delta v''(\underline{p}_1) = \Psi_2(\underline{p}_1, v_2(\underline{p}_1)) - \Psi_1(\underline{p}_1, v_1(\underline{p}_1)) > 0$ by (7), as v_i solves the Γ_i problem, and $\Delta v(\underline{p}_1) \geq 0$. We claim that Δv attains a global maximum at some $\bar{p} \in (\underline{p}_1, \min(\bar{p}_1, \bar{p}_2))$. To see why, consider the upper thresholds. If $\bar{p}_1 < \bar{p}_2$, then $\Delta v'(\bar{p}_1) < 0$ by smooth pasting of v_1 at \bar{p}_1 and strict convexity of v_2 . Thus, the function Δv , strictly increasing at $\underline{p}_1 + \varepsilon$ for some $\varepsilon > 0$, is strictly decreasing at $\bar{p}_1 < \bar{p}_2$: An interior global maximum \bar{p} then exists. If instead $\bar{p}_1 \geq \bar{p}_2$, then $\Delta v(\bar{p}_2) \leq 0$ by value matching of each v_i at \bar{p}_i ; the function Δv , strictly positive and increasing at $\underline{p}_1 + \varepsilon$ is nonpositive at $\bar{p}_2 \leq \bar{p}_1$ —and so has an interior global maximum \bar{p} . In either case, since $\Delta v(\underline{p}_1 + \varepsilon) > 0$, we deduce the maximum $\Delta v(\bar{p}) > 0$. But then $\Delta v''(\bar{p}) = \Psi_2(\bar{p}, v_2(\bar{p})) - \Psi_1(\bar{p}, v_1(\bar{p})) > 0$ because $\Delta v(\bar{p}) > 0$, as just deduced. This violates the second order condition $\Delta v''(\bar{p}) \leq 0$ for a maximum of Δv at \bar{p} .

Step 2—Ordering Value Functions: Value matching, plus $\underline{p}_1 < \underline{p}_2$ and $\bar{p}_1 > \bar{p}_2$ just proven, jointly imply $\Delta v(\underline{p}_2) < 0$ and $\Delta v(\bar{p}_2) < 0$. The claim then obtains near the thresholds. Assume $\Delta v(p) \geq 0$ at some $p \in (\underline{p}_1, \bar{p}_2)$. Then Δv must attain a global nonnegative maximum for some $\bar{p} \in (\underline{p}_2, \bar{p}_2)$. So $\Delta v(\bar{p}) \geq 0 \geq \Delta v''(\bar{p})$ —contradiction. Q.E.D.

PROOF OF PROPOSITION 5 (c)—Cost Convexity: Define $\xi(w) \equiv c'(f(w))$, $\Delta c(n) \equiv \hat{c}(n) - c(n)$, and similarly $\Delta c'$, $\Delta c''$, Δg , Δf , $\Delta \xi$. Convexity of Δc (since $\hat{c}'' - c'' \geq 0$) and $\Delta c(0) = 0$ imply $\Delta c'(n) > \Delta c(n)/n$ for all $n > 0$. Equivalently, $\Delta g(n) > 0$, and thus $\Delta f(w) < 0$ for all $w > 0$, and $\Delta f(0) \leq 0$ by continuity. Since $\xi'(w) = c''(f(w))f'(w) = 1/f(w)$, we have $\Delta \xi'(w) > 0$ for all $w > 0$ and $\Delta \xi'(0) \geq 0$. If we prove that $\Delta \xi(0) \geq 0$, then $\Delta \xi(w) > 0$ for all $w > 0$ and then (7) holds, yielding: $\hat{v} \geq v \Rightarrow \hat{\Psi}(p, \hat{v}) = \hat{\xi}(r\hat{v})/\mathcal{E}(p) \geq \xi(r\hat{v})/\mathcal{E}(p) \geq \xi(rv)/\mathcal{E}(p) = \Psi(p, v)$ independently of p , and thus $v > \hat{v}$.

Focus on $\Delta \xi(0) \geq 0$. A formal proof is long, but in (n, c) -space, draw the cost function $c(n)$ and the ray through the origin tangent to it. The unique tangency point v is where $vc'(v) = c(v)$, i.e. $g(v) = 0$ or $v = f(0)$; the ray has the same slope as the function at the tangency point, namely $c'(f(0)) = \xi(0)$. Then draw the cost function $\hat{c}(n)$ and the tangent ray through the origin, with slope $\hat{\xi}(0)$: The ray tangent to $\hat{c}(n)$ is steeper than that to $c(n)$, i.e. $\Delta \xi(0) > 0$. Finally, $n > \hat{n}$, as $v > \hat{v}$ and $f > \hat{f}$ from above. Q.E.D.

PROOF OF PROPOSITION 5 (d)—Information Quality: If $\xi_1 > \xi_2$, then $\mathcal{E}_1(\cdot) > \mathcal{E}_2(\cdot)$ and therefore the premise of Claim 1 holds: $\Psi_i(\cdot, v_i(\cdot)) = \xi(rv_i(\cdot))/\mathcal{E}_i(\cdot)$, so that $\xi(\cdot)$ increasing and $1/\mathcal{E}_1(\cdot) < 1/\mathcal{E}_2(\cdot)$ imply (7). Hence the value v falls and triggers shift in. Since f is unchanged, $n_i(p) = f(rv_i(p))$ falls uniformly with v_i . Q.E.D.

CLAIM 2 (Proof of Proposition 5(e): Impatience): *Given interest rates $r_2 > r_1 > 0$:*

- (i) *the r_2 -thresholds are shifted in, and the r_2 -value v is lower at all points in its domain;*
- (ii) *there exists a possibly empty interval $[\bar{q}, \underline{q}]$, strictly contained in $[\underline{p}_2, \bar{p}_2] \subset [\underline{p}_1, \bar{p}_1]$, with $n_2(p) \leq n_1(p)$ for all $p \in [\bar{q}, \underline{q}]$, and $n_2(p) > n_1(p)$ otherwise.*

PROOF OF (i): Since $r_2 > r_1$ with all else equal, ξ increasing implies (7) and Claim 1. Q.E.D.

PROOF OF (ii): Let $w_i(p) = r_i v_i(p)$ for $p \in (\underline{p}_1, \bar{p}_1)$, $i = 1, 2$. Define functions $\Delta w(p)$, $\Delta w'$, $\Delta w''$, Δv , $\Delta v'$, and $\Delta v''$ all with domain $[\underline{p}_2, \bar{p}_2]$. First, $\underline{p}_1 < \underline{p}_2$ and $\bar{p}_1 > \bar{p}_2$ from (a), smooth pasting, and strict convexity of v_1 imply $\Delta v'(\underline{p}_2) < 0 < \Delta v'(\bar{p}_2)$. By continuity, $\Delta v'$ is strictly increasing in some subset $I \subseteq [\underline{p}_2, \bar{p}_2]$. Hence, $\Delta v''(\bar{p}) = [\xi(w_2(\bar{p})) - \xi(w_1(\bar{p}))]/\mathcal{E}(\bar{p}) > 0$, and so $w_2(\bar{p}) > w_1(\bar{p})$ and $n_2(\bar{p}) > n_1(\bar{p})$. We show that the complement set, where the folk result $n_2 \leq n_1$ obtains, is a possibly empty interval.

By definition of v_i the return w_i is strictly convex and solves $\mathcal{E}(p)w_i''(p) = r_i \xi(w_i(p))$ such that $w_i(\underline{p}_i) = r_i \pi(\underline{p}_i)$, $w_i'(\underline{p}_i) = r_i(\pi_A^H - \pi_A^L)$, $w_i(\bar{p}_i) = r_i \pi(\bar{p}_i)$, $w_i'(\bar{p}_i) = r_i(\pi_B^H - \pi_B^L)$. Thus $w_2'(\underline{p}_2) = r_2(\pi_A^H - \pi_A^L) \leq r_1(\pi_A^H - \pi_A^L) = w_1'(\underline{p}_1) < w_1'(\underline{p}_2)$, where the first equality is smooth pasting, the weak inequality follows from $r_2 > r_1$ and $\pi_A^H \leq \pi_A^L$, and the strict inequality from $\underline{p}_2 > \underline{p}_1$ and $w_1' > 0$. By a symmetric argument and $\pi_B^H > \pi_B^L$, $w_2'(\bar{p}_2) > w_1'(\bar{p}_1)$. Therefore the smooth function Δw is strictly decreasing at \underline{p}_2 and increasing at \bar{p}_2 . Since we have shown that Δw is strictly positive in a nonempty set $I \subseteq [\underline{p}_2, \bar{p}_2]$, it suffices to show that Δw cannot have a local nonnegative maximum in $(\underline{p}_2, \bar{p}_2)$. By contradiction, suppose that $\Delta w(\bar{p}) \geq 0 = \Delta w'(\bar{p}) \geq \Delta w''(\bar{p})$ for some $\bar{p} \in (\underline{p}_2, \bar{p}_2)$. Then $\bar{p} \in (\underline{p}_2, \bar{p}_2) \subset [0, 1]$ implies $\mathcal{E}(\bar{p}) > 0$; furthermore $r_2 > r_1$, $\xi(\cdot)$ is increasing and $\Delta w(\bar{p}) \geq 0$, so that the familiar contradiction follows: $\Delta w''(\bar{p}) = [\mathcal{E}(\bar{p})]^{-1}[r_2 \xi(w_2(\bar{p})) - r_1 \xi(w_1(\bar{p}))] > 0 \geq \Delta w''(\bar{p})$.

Finally, we specialize to the R&D payoff specification. Here, $v_2(\underline{p}_2) = 0$, and then $\Delta w(\underline{p}_2) = w_2(\underline{p}_2) - w_1(\underline{p}_2) = -w_1(\underline{p}_2) < 0$ by $\underline{p}_2 > \underline{p}_1$. Therefore Δw is initially strictly negative and declining; since it must become strictly positive at some point below \bar{p}_2 , and it cannot have a local nonnegative maximum, it cannot change sign twice; so Δn changes sign exactly once, going from negative to positive as we raise p . Hence, there is an interior cutoff $p' \in (\underline{p}_2, \bar{p}_2)$ such that $n(p)$ declines for all beliefs $p \leq p'$, and rises for all $p > p'$. Q.E.D.

REFERENCES

AGHION, P., AND P. HOWITT (1992): "A Model of Growth through Creative Destruction," *Econometrica*, 60, 323–351.

- BOLTON, P., AND C. HARRIS (1999): "Strategic Experimentation," *Econometrica*, 67, 349–374.
- CHERNOFF, H. (1972): *Sequential Analysis and Optimal Design*, Vol. 8 of Regional Conference Series in Applied Mathematics. Philadelphia: Society for Industrial and Applied Mathematics.
- CRESSIE, N., AND P. MORGAN (1993): "The VPRT: A Sequential Testing Procedure Dominating the SPRT," *Econometric Theory*, 9, 431–450.
- DI MASI, J., H. G. GRABOWSKI, AND J. VERNON (1995): "R&D Costs, Innovative Output and Firm Size in the Pharmaceutical Industry," *International Journal of the Economics of Business*, 2, 201–219.
- GROSSMAN, G., AND C. SHAPIRO (1986): "Optimal Dynamic R&D Programs," *Rand Journal of Economics*, 17, 581–596.
- KAMIEN, M., AND N. SCHWARTZ (1971): "Expenditure Patterns for Risky R&D Projects," *Journal of Applied Probability*, 8, 60–73.
- KELLER, G., AND S. RADY (1999): "Optimal Experimentation in a Changing Environment," *Review of Economic Studies*, 66, 475–507.
- KRYLOV, N. V. (1980): *Controlled Diffusion Processes*. New York: Springer-Verlag.
- LIPTSER, R., AND A. N. SHIRYAYEV (1977): *Statistics of Random Processes: General Theory*, Vol. I. New York: Springer-Verlag.
- MALUEG, D., AND S. TSUTSUI (1997): "Dynamic R&D Competition with Learning," *Rand Journal of Economics*, 28, 751–772.
- MOSCARINI, G., AND L. SMITH (1998): "Wald Revisited: The Optimal Level of Experimentation," Yale University, Cowles Foundation D.P. 1176 [ms98].
- OKSENDAL, B. (1995): *Stochastic Differential Equations: An Introduction with Applications*. New York: Springer Verlag.
- REINGANUM, J. (1981): "Dynamic Games of Innovation," *Journal of Economic Theory*, 25, 21–41.
- ROBERTS, K., AND M. WEITZMAN (1981): "Funding Criteria for Research, Development, and Exploration Projects," *Econometrica*, 49, 1261–1288.
- ROCKAFELLAR, T. (1970): *Convex Analysis*. Princeton: Princeton University Press.
- SCHERER, F. M. (1965): "Firm Size, Market Structure, Opportunity, and the Output of Patented Inventions," *American Economic Review*, 55, 1097–1125.
- SHIRYAYEV, A. N. (1978): *Optimal Stopping Rules*. New York: Springer-Verlag.
- WALD, A. (1947): "Foundations of a General Theory of Sequential Decision Functions," *Econometrica*, 15, 279–313.