COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
AT YALE UNIVERSITY

Box 2125, Yale Station
New Haven, Connecticut 06520

COWLES FOUNDATION DISCUSSION PAPER NO. 1176

WALD REVISITED: THE OPTIMAL LEVEL OF EXPERIMENTATION

Giuseppe Moscarini and Lones Smith

May 1998

Wald Revisited: The Optimal Level of Experimentation[*†‡]

Giuseppe Moscarini[§]
Economics Department
Yale University

Lones Smith[¶]
Economics Department
M.I.T.

April 22, 1998
(first version: July, 1997)

Abstract

The paper revisits Wald's (1947) sequential experimentation paradigm, now assuming that an *impatient* decision maker can run variable-size experiments each period at some *increasing and strictly convex cost* before finally choosing an irreversible action. We translate this natural discrete time experimentation story into a tractable control of variance for a continuous time diffusion. Here we robustly characterize the optimal experimentation level: It is rising in the confidence about the project outcome, and for not very convex cost functions, the random process of experimentation levels has a positive drift over time. We also explore several parametric shifts unique to our framework. Among them, we discover what is arguably an 'anti-folk' result: Where the experimentation level is positive, it is often higher for a more impatient decision maker.

This paper more generally suggests that a long-sought economic paradigm that delivers a sensible law of demand for information is our dynamic one — namely, allowing the decision maker an eternal repurchase (resample) option.

*This paper was initially entitled "Wald Revisited: A Theory of Optimal R&D". We have 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, Queens, and more specifically Massimiliano Amarante, Dirk Bergemann, Lutz Busch, John Conley, Prajit Dutta, Drew Fudenberg, Peter Hammond, Chris Harris, Angelo Melino, Paul Milgrom, Stephen Morris, Klaus Nehring, Yaw Nyarko, Mike Peters, Sven Rady, Arthur Robson, John Rust, Chris Sims, Peter Sørensen, Ennio Stacchetti, and Steve Tadelis. 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 SSFRF, and Smith from the National Science Foundation (grant SBR-9711885) for this research.

†Abraham Wald left Vienna in 1938 and went to the USA on a Cowles Commission Fellowship. Five years later, while working in the Statistical Research Group in New York, he produced in one afternoon the celebrated Sequential Probability Ratio Test, which is the inspiration for this paper.

‡JEL: C11, C12, C44, C61, D81, D83. Keywords: learning, experimentation, sequential analysis, R&D.

§Email: `gm76@pantheon.yale.edu`. Snail mail: P.O. Box 208268, New Haven CT 06520-8268

¶Email: `econ-theorist@earthling.net`. Snail mail: 50 Memorial Drive, Cambridge MA 02142-1347.

1. INTRODUCTION

This paper revisits a classic decision theory contribution around its semicentenial — Wald's (1947a) *sequential probability ratio test.*[1] More than any other work, this has defined the paradigm of optimal sequential experimentation, either statistical or Bayesian. For it tackled afresh the simplest of decision problems — for instance, choosing among two actions (perhaps accepting hypotheses), each optimal in one of two states of the world.

In his Bayesian formulation, Wald posits that the *decision maker* ($\mathcal{DM}$), ever uncertain of the state, can buy multiple i.i.d. informative signals at constant marginal cost. Wald shows that the $\mathcal{DM}$ should purchase sequentially, and act when sufficiently convinced of one state. Yet we venture that the experimental schedule should accelerate when *homo economicus* runs the laboratory in real time. We offer here a compelling, economically-motivated twist along these lines that also creates a pure theory of dynamic R&D. In our spin on Wald's tale, the $\mathcal{DM}$ is assumed *impatient,* but may elect a variable-size experiment each period only at an *increasing, strictly convex cost* of information. We argue that this richer experimentation story provides some intriguing new testable implications, and that it also points economists to a long-sought well-behaved theory of information demand.

This paper has three goals, attacked in sequence after an essential literature review. First, we motivate our impatient $\mathcal{DM}$ and convex cost spin on Wald's discrete time tale, and then introduce and solve a tractable continuous time equivalent model of experimentation — which we argue is a control of variance for a diffusion with uncertain mean.

Second, we investigate the robust character of the experimentation level. One might suppose that it peaks when the $\mathcal{DM}$ has middling, elastic beliefs over the state. In fact, we find the opposite: Given a convex cost function, experimentation grows in the expected payoff, which is greatest at extreme beliefs. For example, R&D levels are least when one is most discouraged, peaking just prior to project approval. We argue that the $\mathcal{DM}$ acts like a neoclassical competitive firm producing information at an increasing marginal cost, with an increasing "producer surplus." Optimal stopping demands that this surplus equal the cost of delaying the final decision, i.e. the portion of current value killed by discounting. *Ipso facto,* the research level that generates this surplus rises in the value. We also show that the level is convex in beliefs for not too convex costs, and so drifts up rise over time.

Finally, we explore how experimentation level responds to some natural parametric shifts. Among our findings, we discover that contrary to established folk wisdom, a more impatient $\mathcal{DM}$ often experiments at a higher level provided that he still experiments.

[1]The original source is a 1943 war-classified mimeo by Wald. Wald (1945a) and (1945b) were the first published versions of the program. In his (1947) preamble, he credits Milton Friedman and Allen Wallis for proposing sequential analysis. Wallis details this history in his (1980) retrospective paper.

2. THREE RELATED LITERATURE THREADS

2.1 Statistical Testing

The (1933) Neyman-Pearson theorem says that the best critical test of two competing hypotheses given a fixed sample has the form: accept H_0 (vs. H_1) iff the sample likelihood ratio exceeds a threshold. Wald and Wolfowitz (1948) proved that the sequential probability ratio test (SPRT) is optimal — minimizing the expected number of observations for a given power, even given either hypothesis. Arrow, Blackwell, and Girshick (1949) further developed this idea, and was among the earliest of contributions to dynamic programming.

The SPRT has been explored and generalized by several authors (see Chernoff (1972)). But in a half century, surprisingly *no one has characterized* the optimal sample size or its dynamics with discounting, where a unit sample size is no longer optimal — our main contribution. Cressie and Morgan (1993), a recent sampling from the frontier, prove that a variable size probability ratio test is unconditionally optimal among sequential design procedures, for instance, while a SPRT is best with superadditive costs and no discounting.

2.2 Optimal Experimentation

Optimal experimentation was born out of sequential analysis. Among the canonical Bayesian learning models, those with a discrete action space are essentially stopping rule problems, or more generally, bandit models; continuous action models are richer, and not so easily pigeon-holed. Our model is rare as it grafts together these two contexts: variable experimentation with an eventual binary stopping decision. Crucially, because our $\mathcal{DM}$ makes a pure information purchase until stopping, it differs from all other experimentation models where myopic expected stage payoffs are themselves informative random signals; here, payoffs are known and negative (the information cost), until the final decision stage.

This paper investigates the level and drift of experimentation, as opposed to the long-run question, 'Is learning complete?'[2] The literature has also studied three main short-run questions. For finite action models like bandits, there has been work on the *sequential ordering* of experiments, or stochastic scheduling. Second, the *direction* of experimentation has been explored in continuous action models like monopoly pricing.[3] Finally, the *secular diminishing trend* in experimentation has been remarked in papers on search with learning (eg. rising reservation prices). This comes closest to our thrust, and yet runs exactly counter to the rising experimentation levels that we find. For with price search, the information purchase is blended with, and thus colored by, immediate payoff concerns.

[2]See, for instance, Easley and Kiefer (1988) or Kihlstrom, Mirman, and Postlewaite (1984).

[3]McLennan (1984) and Trefler (1993) are good examples. Keller and Rady (1997), who posit a randomly shifting demand curve, assume continuous time learning; it is therefore also a technically relevant paper.

2.3 Research and Development

We interpret a special case of our analysis as a pure theory of dynamic R&D. While this is obviously a well-studied subject, other work differs in a key dimension. During 1971–82, Kamien and Schwartz modelled R&D as (very roughly) covering a possibly uncertain distance in 'progress space', given cost and effort functions. Dutta (1997) has recently assumed a related budget constrained model. It may help to imagine extant work as capturing 'D', and ours 'R'. For in our setting, the actual state of the world ("Is this feasible?") always hangs in the balance, while these papers posit a goal that is eventually attainable, if desired. We cast research as optimal learning rather than resource allocation; this is a theory of science (the search for truth) and not engineering (its implementation). An eventual comprehensive theory of dynamic R&D will no doubt embed both phases.

3. MOTIVATING THE CONTINUOUS TIME MODEL

3.1 Bayesian Sequential Analysis

A. The Final Static Decision Problem. In a standard problem, the $\mathcal{DM}$ eventually must choose between two actions: $a = A, B$ pays π_a^θ in the state of the world $\theta = L, H$. Action A is better in state L, and B in state H, and the $\mathcal{DM}$ does not know the state.

Assume that the $\mathcal{DM}$ is Bayesian and risk neutral (simply treat utilities as payoffs). If the $\mathcal{DM}$'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 $\mathcal{DM}$ is indifferent between A and B at some belief $\hat{p}$. We also admit the costless option of never deciding, and thus need a default null action yielding constant zero payoff. His static value is then $\pi(p) \equiv \max\langle \pi_A(p), \pi_B(p), 0\rangle$. Given the null action, we can assume WLOG (without loss of generality) that action B optimally pays a positive payoff $\pi_B^H > 0$. If state L is dominated, then $\pi(p)$ is always increasing in p.

The special case of one risky and one safe action conveniently captures a stylized R&D problem: action B means 'building' a costly new prototype, and action A 'abandoning' it. If the $\mathcal{DM}$ builds, it might or might not work: payoffs are $h = \pi_B(1) > 0$ or $\ell = \pi_B(0) < 0$ in states $\theta = H$ and $\theta = L$. The $\mathcal{DM}$ earns zero regardless if he abandons: $\pi_A(p) \equiv 0$. So $\pi(p) \equiv \max\langle 0, hp + \ell(1-p)\rangle$, and the $\mathcal{DM}$ invests iff $p > \hat{p} \equiv -\ell/(h-\ell)$.

B. Information Acquisition. Before choosing an action, let the $\mathcal{DM}$ initially acquire informative signals of θ at a fixed unit cost. Assume he maximizes the expected final reward less costs incurred. As Wald and Wolfowitz proved that sequential purchases are optimal, this is a pure optimal stopping exercise: The $\mathcal{DM}$ quits at the stopping time T and chooses action A (or B, or the null action, i.e. quits) with posterior $p \leq \underline{p}$ (or $p \geq \bar{p}$, or $p \in [\underline{p}_0, \bar{p}_0]$).

3.2 Discrete Time Experimentation with Impatience and Convex Costs

Our goal is to extend Wald's setting along two dimensions: payoff discounting and cost convexity of experimentation. These assumptions are of particular interest for economists, less so for statisticians. This may explain why this stone has been left unturned.

First assume an impatient $\mathcal{DM}$, who maximizes the expected present discounted value of wealth. Faced with the time cost of delay, the $\mathcal{DM}$ does not necessarily wish to proceed purely sequentially, but may opt to stack his information purchases. After seeing his signal outcomes in any period, he either purchases more, or stops and chooses an action. In period k, the $\mathcal{DM}$ may buy N_k *i.i.d.* signals $\tilde{X}_1, \ldots, \tilde{X}_{N_k}$ at cost $C(N_k)$, with $C(0) \geq 0$. One might venture that running a lab incurs a daily rent, independent of the experimentation level. In this case, there is a positive fixed flow experimentation cost $C(0) > 0$. But, the analogy with Wald's setting is obviously closest with no fixed costs $C(0) = 0$.

Next assume strictly convex information costs within a period; this fosters more equal purchases across periods, reinforcing Wald's sequential conclusion, absent discounting. Such an assumption makes economic sense on two grounds. First, plausibly not all researchers are equally talented in producing information. More intensive information search then must draw on the efforts of less capable researchers. Second, as with non-Bayesian inventive activity, since contemporaneously-produced knowledge is based on the same current stock, identical or similar discoveries are not rare.[4] 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.

3.3 Developing the Continuous Time Experimentation Model

Although variable intensity experimentation is easily understood and formulated in discrete time using dynamic programming, the solution is intractable even for the very simplest signal structures. Perhaps this explains why no one has seriously pursued it. We now describe a tractable continuous time learning paradigm that captures the quintessence of the discrete time story. Below we sketch and motivate it and focus on its economic substance. Its recursive solution is found in section 4.2, and a formal justification in Appendix B. In a work in progress, we argue that the model and its solution is the limit of a rich class of discrete time models with a vanishing time interval between experiments.

[4] A recurring theme in science is that great minds simultaneously think alike (eg. Newton and Leibniz' codevelopment of calculus). Indeed, those seeing further often stand on the shoulders of the same giants.

A. The Signal Process. In discrete time, choosing the number of i.i.d. signals to buy, each with distinct state-dependent means, is an apt description of variable intensity experimentation. For instance, if $\tilde{X}$ has mean $\pm\mu$ in in states H, L, then experiments offer a noisy but informative glimpse of the signal mean. In fact, the very goal of experimentation is to infer this mean, for in so doing, the $\mathcal{DM}$ learns the state. Since the average signal $\bar{X} = \sum X_i/N$ is sufficient for the unobserved mean, greater experimentation only serves to decrease the variance of $\bar{X}$: Doubling the sample size precisely halves the variance.

Motivated by this general observation, we model continuous time experimentation as the *control of variance* of a diffusion observation process $\langle \bar{x}_t \rangle$. For a fixed control, $\langle \bar{x}_t \rangle$ is a Brownian motion with constant uncertain drift and known variance. Nature chooses its drift, μ^θ in state θ, where $\mu^H = -\mu^L = \mu > 0$, while the $\mathcal{DM}$ controls its flow variance σ^2/n_t, with the intensity n_t. The observation diffusion process thus solves the stochastic differential equation (SDE)

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

in state θ. Here, we think of n_t as the flow of information purchases, and call it the experimentation *level* or *intensity*. Doubling n_t halves the "variance" of $d\bar{x}_t^\theta$, yielding a doubly informative time-t experiment. As usual, $dW_t \sim N(0, dt)$ is the Wiener increment, while the control n_t depends on the observation and intensity history $\langle \bar{x}_s, 0 \le s \le t \rangle \cup \langle n_s, 0 \le s < t \rangle$. It is a feedback and not open loop control, not decided at time-0.[5]

See §17.6 in Liptser and Shiryayev (1978), §17.5 in Chernoff (1972), or §4.2 in Shiryayev (1978) for the pure problem of estimating the bivariate drift of a Brownian motion. Their motivation is its link to the heat equation. Neither source discusses control of variance. To be sure, without time preference, there is no pressing reason to consider such an exercise.

REMARK. Our choice of process $\langle \bar{x}_t^\theta \rangle$ yields a motivational pure control of variance: But a realization $\bar{x}_t = \int_0^t d\bar{x}_s$ is an unweighted running integral of *sample means,* and so is not a sufficient statistic for $\langle (\bar{x}_s, n_s), 0 \le s \le t \rangle$ (w.r.t. the drift μ^θ). Consider instead the running *sample totals* observation process $\langle x_t^\theta \rangle$, obeying $dx_t^\theta = n_t \mu^\theta dt + (\sigma\sqrt{n_t}) dW_t$. In that case, $\int_0^t n_s ds$ can be thought of as the running *sample size,* so that $(x_t, \int_0^t n_s ds) \in \mathbb{R}^2$ is a simple sufficient statistic for the mean. This process also yields a clearly well-defined level-0 experimentation, as will our belief filter (2) in §4.2. We see also that a higher intensity level n_t essentially accelerates time — advancing the schedule that the $\mathcal{DM}$ observes future samples. The two processes $\langle \bar{x}_t^\theta \rangle$ and $\langle x_t^\theta \rangle$ are mere conceptual devices, and a choice between them is not critical: Both are sufficient for the drift μ^θ (and hence the state θ), and crucially yield the same belief filter (2), which is what we work with.

[5]This is why the signal must be defined recursively via a SDE rather than as an exogenous Ito process.

B. The Cost of Flow Experimentation. Intensity level n incurs a flow cost $c(n)$.

($\bigstar$): *The cost function $c(n)$ is finite, increasing, strictly convex (and thus continuous) on $[0, \infty)$, differentiable on $(0, \infty)$,*[6] *with nonnegative fixed costs $c(0) \geq 0$. Marginal costs are 'Lipschitz-down' on $(0, \infty)$: $c'(n_2) - c'(n_1) \geq \lambda(n_2 - n_1)$, for some $\lambda > 0$, and any $n_2 > n_1 > 0$.*

Strict cost convexity means $c(\gamma n_1 + (1-\gamma)n_2) > \gamma c(n_1) + (1-\gamma)c(n_2)$ for $n_1 \neq n_2$ and $0 < \gamma < 1$. It precludes undesirable bang-bang control solutions. Both the Lipschitz-down property — which is true if $c''(>0)$ exists — and differentiability play purely technical roles, ensuring that our observation process (1), and solution is well-defined.

REMARK. Weak (though not strict) cost convexity remarkably obtains *WLOG* in continuous time.[7] For the $\mathcal{DM}$ can achieve any cost function arbitrarily close to the lower convex hull $\text{vex}(c) = \sup\{c_0 \leq c | c_0 \text{ is convex}\}$. Indeed, for any $\varepsilon > 0$, any average cost at least $\gamma c(n_1) + (1-\gamma)c(n_2) - \varepsilon$ is achieved in any time interval $[t_0, t_1]$ by sufficiently rapidly chattering between n_1 and n_2 with weights $(\gamma, 1-\gamma)$ at small payoff loss. Then, let $\varepsilon \to 0$.

C. The Objective Function. At each time t, the impatient $\mathcal{DM}$, facing an interest rate $r > 0$, chooses whether to stop and earn the final expected payoff $\pi(p_t)$, or to continue at a chosen intensity level $n_t \geq 0$. Admissibility demands that the random stopping time T and the level n_t each be functions of the observed history. In the *Optimal Control and Stopping* ($\mathcal{OCS}$) problem, the $\mathcal{DM}$ maximizes his expected discounted return less incurred experimentation costs: $E[\int_0^T -c(n_t)e^{-rt}dt + e^{-rT}\pi(p_T)|p_0]$. We write the *optimized value*, or supremum w.r.t. T and $\langle n_t \rangle$, as $V(p_0)$, since Appendix B.1 proves that the current posterior belief p_0 on state H is a sufficient statistic for observed history to that moment.

4. THE OPTIMAL LEVEL OF EXPERIMENTATION

4.1 Competing Static and Dynamic Intuitions for the Value of Information

Let the convex function Π describe the expected payoff of a one-shot Bayesian program. At the prior belief p, a signal yielding the random posterior belief $\tilde{q}$ is worth $\mathcal{I}(\tilde{q}|p) = E[\Pi(\tilde{q}) - \Pi(p)]$. This admits a motivational visual depiction. Beliefs being a martingale, we have $p = E(\tilde{q})$; therefore, we may tack on any multiple of $[E\tilde{q} - p] = 0$. As Π is convex, it is differentiable for a.e. p. Put $d = \Pi'(p)$ when defined, and otherwise choose any subdifferential $d \in \partial\Pi(p)$, i.e. a slope between the left and right derivatives. Then $\mathcal{I}(\tilde{q}|p) = E[\Pi(\tilde{q}) - \Pi(p) - d(\tilde{q} - p)]$ is the weighted area between Π and any supporting tangent line, with weights given by the density over $\tilde{q}$ (see Figure 1).

[6] Thus, the cost function is $\mathcal{C}^1$ (continuously differentiable); by convexity, it has a right derivative $c'(0+)$.

[7] We thank Paul Milgrom for this nice insight. We do not wish to delve into a technical proof of this assertion, as it would detract from our focus. We intend the point and explanation to be intuitive.

true noise process $\langle W_t \rangle$ that drives the signal process in (1). The precise measure-theoretic statement, in terms of signal filtrations, appears in Appendix B.1. As a driftless diffusion, beliefs $\langle p_t \rangle$ are an unconditional martingale, with least variance near the extremes 0 and 1.

Substituting the ex post observed history $\langle \bar{x}_t, n_t \rangle$ into these formulae reveals how beliefs are computed. Intuitively, 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 $d\bar{x}_t > [p_t\mu + (1-p_t)(-\mu)]dt$.

REMARK. Using (1), we find that $d\bar{W}_t^H = 2(\sqrt{n_t}/\sigma)\mu(1-p_t)dt + dW_t$ and $d\bar{W}_t^L = -2(\sqrt{n_t}/\sigma)\mu p_t dt + dW_t$. Plugging either *Ito process* (i.e. Ito stochastic integral) $\langle \bar{W}_t^\theta \rangle$ into the belief filter (2) yields the SDE solved by the conditional belief processes $\langle p_t^H \rangle$ or $\langle p_t^L \rangle$. Not surprisingly, beliefs have a positive drift in state H, and a negative drift in state L.

B. The Value Function. The supremum value of the $\mathcal{OCS}$ problem is

$$V(p_0) = \sup_{T,\langle n_t \rangle} E\left[\int_0^T -c(n_t)\, e^{-rt} dt + e^{-rT}\pi\left(p_0 + \int_0^T p_t(1-p_t)\zeta\sqrt{n_t} d\bar{W}_t\right)\Big| p_0\right] \tag{4}$$

Standard for optimal learning, this value function is convex. For intuitively, a signal spreading the belief $p_0 = \gamma_1 p_1 + \gamma_2 p_2$ to p_i with chance γ_i ($\gamma_1 + \gamma_2 = 1$) cannot possibly hurt the $\mathcal{DM}$, as he can ignore it: If he optimizes at each p_i, he then gets $\gamma_1 V(p_1) + \gamma_2 V(p_2) \geq V(p_0)$. The next lemma, proved in Appendix B.1, mandates threshold stopping rules.

Lemma 1 (Convexity) *Consider any cost function $c(n) > 0$. Then the supremum value V is convex in p. Also, $V(p) = \pi(p)$ for $p \leq \underline{p}$ and $p \geq \bar{p}$, for some cut-offs $0 \leq \underline{p} \leq \bar{p} \leq 1$. If the null action is ever exercised, then $V(p) = \pi(p)$ in $[\underline{p}_0, \bar{p}_0]$, where $\underline{p} < \underline{p}_0 < \bar{p}_0 < \bar{p}$.*

C. Optimality Conditions. The $\mathcal{DM}$ selects action A for $p \leq \underline{p}$, action B for $p \geq \bar{p}$, and absent the null action, he experiments at level $n(p)$ for $p \in (\underline{p}, \bar{p})$, an open set. If $\pi < 0$ ever, then the null action may be chosen in a subinterval $[\underline{p}_0, \bar{p}_0] \subset (\underline{p}, \bar{p})$, where $\underline{p} < \underline{p}_0 < \bar{p}_0 < \bar{p}$. We then partition the $\mathcal{OCS}$ problem into an Optimal Control ($\mathcal{OC}$) exercise for the schedule $n(p)$ (Appendix B.1.c proves this Markovian form) and an Optimal Stopping ($\mathcal{OS}$) problem for the boundaries $\underline{p}, \bar{p}$, and perhaps $\underline{p}_0, \bar{p}_0$. The *experimentation domain* is then $\mathcal{E} = (\underline{p}, \bar{p})$, or $\mathcal{E} = (\underline{p}, \underline{p}_0) \cup (\bar{p}_0, \bar{p})$ if the null action is viable. We now develop recursive equations for what we call the *value* v; we then prove in Proposition 1 that v exists, and coincides with the supremum value, or $v = V$. Hence, we solve for the optimal dynamic policy using ordinary differential equations methods (ODE) in Proposition 2.

Since beliefs $\langle p_t \rangle$ are a martingale obeying (2), for any given experimentation region $\mathcal{E}$, the Hamilton-Jacobi-Bellman ($\mathcal{HJB}$) equations associated to the $\mathcal{OC}$ problem are

$$\begin{aligned} rv(p) &= \sup_{n \geq 0}\{-c(n) + 0 \cdot v'(p) + (1/2)np^2(1-p)^2\zeta^2 v''(p)\} \\ \Rightarrow rv(p) &= \sup_{n \geq 0}\{-c(n) + n\Sigma(p)v''(p)\} \end{aligned} \tag{5}$$

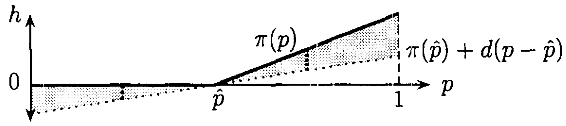

Figure 1: **Static Value of Information.** With a continuous signal support, the shaded area between a value function and any supporting line, *appropriately weighted,* is the value of information. With discrete signals, we instead have a weighting of vertical line segments in this shaded region (eg. the thick dashed lines). In either case, for a purely static value function π, the information value is maximized in the middle at $\hat{p}$, and generally is quasiconcave in p.

For the specific case of the static payoff frontier $\Pi = \pi$, this visual information value is maximal at $\hat{p}$, when the $\mathcal{DM}$ is indifferent between the two actions. Intuitively, he values information most when he is most uncertain about his *action* choice. Alternatively, the $\mathcal{DM}$ likes to spread his posterior beliefs; the impact of new information on the variance of posterior beliefs is increasing in $p(1-p)$, and peaks at $p = 1/2$, where the $\mathcal{DM}$ is most uncertain as to which *state* is true. Either static logic suggests that information value and thus its demand (extrapolating to margins) are quasiconcave 'hill-shaped' functions of p.

The above logic fails for our impatient $\mathcal{DM}$ in a dynamic setting. For the $\mathcal{DM}$ has an incentive minimize the present discounted cost of information, and therefore wishes to delay his costly high-intensity experimentation until just prior to stopping. As this occurs only at extreme beliefs, information demand ought to be greatest for extreme beliefs. Of course, this intuition cannot establish the shape of the experimentation schedule, let alone its intertemporal trend. For that, we must consider the recursively-formulated problem.

4.2 The Recursive Formulation and Solution

A. The Belief Filter and Bayes Problem. Intuitively (and formally, as we later show), 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 according to Bayes rule in continuous time. If $\zeta = 2\mu/\sigma$ denotes the *signal-to-noise ratio factor* of $\langle \bar{x}_t \rangle$, then Theorem 9.1 of Liptser and Shiryayev (1977) (LS77) asserts:

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

where the Wiener increment $d\bar{W}_s = p_s d\bar{W}_s^H + (1-p_s)d\bar{W}_s^L$, and for $\theta = L, H$:

$$d\bar{W}_s^\theta \equiv \frac{\sqrt{n_s}}{\sigma}[d\bar{x}_s^\theta - [p_s\mu + (1-p_s)(-\mu)]ds] \tag{3}$$

Alternatively, LS77's Theorem 9.1 implies that $\langle \bar{W}_t \rangle$ is a Wiener process from the $\mathcal{DM}$'s unconditional perspective,[8] as he does not know the true drift, and so cannot observe the

[8]Namely: As of time s, given observed history $\langle \bar{x}_{t'}, 0 \le t' \le s \rangle \cup \langle n_{t'}, 0 \le t' < s \rangle$, or more simply, just $(x_t, \int_0^t n_s ds)$ the increment $\bar{W}_t - \bar{W}_s$ is Gaussian $N[0, t-s]$ and independent of $\bar{W}_s$, for all $t > s$.

where $\Sigma(p) \equiv p^2(1-p)^2\zeta^2/2$ measures "belief elasticity", plus the *value matching* condition:

$$v(\underline{p}) = \underline{p}\pi_A^H + (1-\underline{p})\pi_A^L \qquad v(\bar{p}) = \bar{p}\pi_B^H + (1-\bar{p})\pi_B^L \quad \text{and possibly} \quad v(\underline{p}_0) = v(\bar{p}_0) = 0 \quad (6)$$

For any given control policy $n(p)$, the Stefan problem ($\mathcal{ST}$) associated to the $\mathcal{OS}$ problem is: $rv(p) = -c(n(p)) + n(p)\Sigma(p)v''(p)$, plus (6), and the free boundary *smooth pasting* (7) conditions, that the value v be tangent to the static payoff function π at $\underline{p}$ and $\bar{p}$:

$$v'(\underline{p}) = \pi_A^H - \pi_A^L \qquad v'(\bar{p}) = \pi_B^H - \pi_B^L \qquad \text{and possibly} \qquad v'(\underline{p}_0) = v'(\bar{p}_0) = 0 \quad (7)$$

While the functional problems $\mathcal{ST}$ and $\mathcal{HJB}$ cannot be solved in closed form, we can create a unified equivalent (second order nonautonomous) ODE problem ($\mathcal{EU}$) below with free boundaries, and fully characterize its solution $\{v(p), \bar{p}, \bar{p}\}$ without the closed form.

Proposition 1 (Value Existence / Uniqueness / Verification) *Assume* (★).
(*a*) *There exists a unique solution* $(\underline{p}, \bar{p}, v)$ *or* $(\underline{p}, \bar{p}, \underline{p}_0, \bar{p}_0, v)$ *of* $\mathcal{HJB}$+$\mathcal{ST}$, *namely* (5)-(7), *with value* $v \in \mathcal{C}^2$ *strictly convex* ($v'' > 0$) *in* $\mathcal{E}$, *and interior thresholds* $0 < \underline{p} < \bar{p} < 1$.
(*b*) *The solution* v *coincides with the maximized value* V *of* $\mathcal{OCS}$: *objective* (4).
(*c*) *The value* v *is jointly increasing/decreasing/U-shaped in* p *with the static payoff* π.

Proof Sketch: The proof is largely appendicized. Here, we provide some easier and more intuitive insights. For a fixed domain $\mathcal{E}$, the FOC for $\mathcal{HJB}$ (5) is $c'(n) = \Sigma(p)v''(p)$; the SOC is met because $c(n)$ is strictly convex, and thus $-c(n) + n\Sigma(p)v''(p)$ is strictly concave in n. The solution $n(p) < \infty$ then uniquely exists if $rv(p) = -c(n) + n\Sigma(p)v''(p) = -c(n) + nc'(n) = g(n)$ is soluble in n. Now, g is clearly continuous as c and c' are; it is strictly increasing too by Claim 1 in §A.1. Since $c'(0+) < \infty$ if $c(n)$ is everywhere finite and convex, then $g(n) \downarrow -c(0)$ as $n \downarrow 0$, and we just set $g(0) = -c(0) \leq 0$. Thus, $g(0) < rv(p)$ for $p \in \mathcal{E}$, because then $v(p) > 0$. Also, $g(n) > r\max\langle\pi(0), \pi(1)\rangle \geq rv(p)$ for large n, because $g(n)$ is unbounded above by Claim 1 in §A.1. Given $rv(p) = g(n(p))$, we have $n(p) = f(rv(p))$ for the strictly increasing inverse $f \equiv g^{-1}$.

We have established that $\mathcal{HJB}$+$\mathcal{ST}$ is equivalent to the two-point boundary value problem $\mathcal{EU}$: $v'' = c'(f(rv))/\Sigma$, cum (6) and (7). Theorem 1 in §A.2 proves that a unique solution to $\mathcal{EU}$ exists, and thereby $\mathcal{HJB}$+$\mathcal{ST}$ is uniquely soluble (part (*a*) here), while this solution is the supremum value of $\mathcal{OCS}$ (part (*b*) here) by Theorem 3 in §B.3.

For (*c*), v shares the shape of π by convexity, value matching, and smooth pasting. □

REMARK. Strict cost convexity rules out one undesirable and not implausible outcome: suddenly exploding the experimentation level over a vanishing time interval $[0, \Delta]$, $\Delta \to 0$. With a linear (i.e. not strictly convex) cost function, such a policy that quickly achieves arbitrarily perfect information is preferred with discounting, absent any cost premium.

Proposition 2 (Policy Existence / Uniqueness) *Assume* ($\bigstar$).
(*a*) *If payoffs* $\pi(p) > 0$ *for all* $p \in [0,1]$, *then the solution* $\{f(rv(\cdot)), \underline{p}, \bar{p}\}$ *from* $\mathcal{HJB}+\mathcal{ST}$ *is the unique optimal policy for* $\mathcal{OCS}$.
(*b*) *The* marginal cost $\xi(w) \equiv c'(f(w))$ of the optimal intensity level $n = f(w)$ *is increasing, strictly concave, and differentiable in the return* $w = rv > 0$ — *even if* c' *is not differentiable.*
(*c*) *If* $\pi(p) = 0$ *at some* $p \in [0,1]$, *then the policy* $\{f(rv(\cdot)), \underline{p}, \bar{p}\}$ *is uniquely optimal for* $\mathcal{OCS}$ *if either* $c(0) > 0$, *or if* $c(0) = 0$ *but there exists* $\eta > 0$ *with* $\lim_{w \downarrow 0} w^{1-\eta}\xi'(w)/\xi(w) < \infty$.
(*d*) *The level* $n(p)$ *is continuous in* $\mathcal{E}$; $n'(\cdot)$ *exists if* $c''(\cdot)$ *does;* $n(\cdot)$ *is* $\mathcal{C}^1$ *when* $c(\cdot)$ *is* $\mathcal{C}^2$.

Proof Sketch: Part (*a*) is proven in Theorems 4–5 in §B.3, and (*c*) in Theorem 6 in §B.4. Part (*b*) is established in the appendicized Claims 2–4. For (*d*), assume that $c''(n) > 0$ exists, and hence so does $g'(n) = nc''(n) > 0$. Since f is differentiable if $g = f^{-1}$ is, so is $n(p) = f(rv(p))$. If c'' is continuous, then so is g', and thus f' and n' too, as claimed. □

Any geometric convex cost function $c(n) = n^k$ ($k > 1$) violates the elasticity condition in Proposition 2-*c* — met by exponential functions like $c(n) = e^{\alpha n} - 1$ (some $\alpha > 0$).

REMARK. We note after Theorem 6 of §B.4 that if $\pi = 0$ somewhere and $c'(0) = c(0) = 0$, so that the final proviso in Proposition 2(*c*) fails, the supremum $V(p)$ of $\mathcal{OCS}$ is not attained; therefore, no optimal policy exists. This pathological case, an undesirable by-product of the continuous time approximation, arises since the $\mathcal{DM}$ may never choose A (or the null action) in finite time given the negligible cost of running very small experiments.

4.3 The Optimal Experimentation Level: The $\mathcal{DM}$ as an Information Firm

Even though the FOC $c'(n) = \Sigma(p)v''(p)$ and the policy equation $n(p) = f(rv(p))$ are jointly insoluble in closed form, the monotonicity of f allows us to conclude that:

Lemma 2 (Monotonicity) *Given* ($\bigstar$), *the optimal intensity level* $n(p)$ *is strictly increasing in the value* $v(p)$ *for* $p \in \mathcal{E}$. *It weakly exceeds* $f(0) \geq 0$, *with* $f(0) = 0$ *iff* $c(0) = 0$.

For instance, quadratic costs $c(n) = n^2$ yields surplus $g(n) = n^2$, and thus $f(n) = \sqrt{n}$. The experimentation level is then an increasing concave function of the return $\sqrt{rv(p)}$.

Here is a concrete economic intuition for the monotonicity of $v \mapsto n$. We formally argue that the bang per research dollar is greater with a higher value, and therefore the level n rises with v. There are two decisions at each moment in time: experiment or stop ($\mathcal{OS}$), and if to experiment, at what level ($\mathcal{OC}$). Focus first on the level choice. Optimality demands that the marginal cost $c'(n)$ of information equal its marginal benefit. Since belief precision is linear in the experimentation level n, the marginal benefit of experimentation $MB = \Sigma(p)v''(p)$ is constant, and the total benefit is then linear: nMB. So at an optimum,

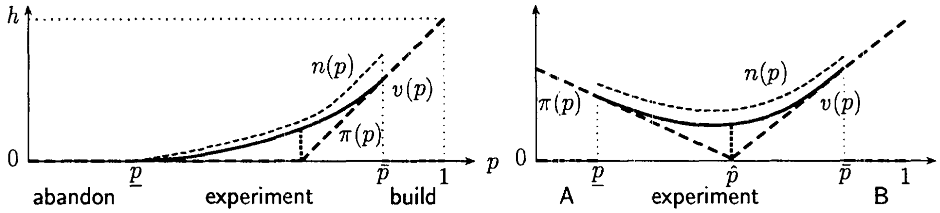

Figure 2: **Value Function and Experimentation Demand.** Overusing the vertical axis, we depict both the static payoff function π (thick dashed line) and dynamic value function v (solid line), strictly convex in the experimentation domain $\mathcal{E} = (\underline{p}, \bar{p})$, and the intensity level 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 $\hat{p}$.

$g(n) \equiv nc'(n) - c(n) = nMB - c(n)$ equals total (flow) benefits less total (flow) costs of experimentation, or the (flow) *producer surplus* of information. Imagine the $\mathcal{DM}$ as a competitive firm, facing an increasing marginal cost curve, "selling" himself information at the constant "price" $\Sigma(p)v''(p)$. The surplus $g(n)$ rises in the optimal quantity if c is strictly convex. Next for the optimal stopping decision, the cost of delay must equal the surplus from optimally experimenting, or $rv(p) = g(n(p))$. This surplus rises in n with convex costs. So as the value v and thus delay cost rv rises, an experimenting $\mathcal{DM}$ must choose a higher level n to generate the requisite higher surplus. Simply put, the $\mathcal{DM}$ only closes down his informational firm (and acts) when he cannot generate sufficient net profits (producer surplus) to justify the time cost of his capital rental (his deferred action).

Proposition 3 (The Optimal Level of Experimentation) *Assume the static payoff frontier $\pi(p)$ is increasing (resp. decreasing, U-shaped) in p. Given (★), the optimal level of experimentation $n(p)$ is increasing (resp. decreasing, U-shaped) in p in the domain $\mathcal{E}$.*

For immediate context, consider our R&D spin. Here, $\pi(p)$ strictly increases in p, and thus the research level rises as we approach confidence in the 'build' decision (see Figure 2, left panel). This provides optimizing foundations for a commonly-observed phenomenon: A potential uncertain investment has new life/money breathed into it by a key discovery or finding; research expenditures stochastically grow over time, only later on to (i) shrink because of discouragement (eg. cold fusion), or (ii) continue growing, as the project is likely headed for development (recently, spinal cord research or a new shuttle rocket engine design). A similar pattern also emerges in the how clinical tests of new drugs proceed: first, small tests, and sometimes later, larger tests, and then quite expensive field trials.[9]

[9]A rare empirical study of project-level R&D expenditures in the pharmaceutical industry (DiMasi, Grabowski, and Vernon (1995)) reveals a pattern strikingly similar to our theoretical prediction. Time-intensive pre-FDA clinical testing usually occurs over three sequential conditional phases, of growing size.

4.4 The Option Value of Experimentation

Write $v(p) = \pi(p) + I(p)$, where $I(p)$ is the *expected present value of information* (or experimentation). Since learning costs time and money, the value v necessarily inherits the general shape of the static payoffs π. For simplicity, assume $\pi(p) > 0$ always, obviating a null action. We next show that $I(p)$ peaks at the kink $\hat{p}$ in π when the myopically-best action switches: Experimentation is most valuable when the $\mathcal{DM}$ is most uncertain as of which action to take, as no action is dominant. 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 worth the least when one is most sure of an action to take. So the suggestion in section 4.1 of a hill-shaped value of information is an apt description of this option value; however, any implication that it alone determines the experimentation level was false, for the flow return to experimentation also includes the terminal payoff $\pi(p)$.

Proposition 4 *Assume (★). If the null action is never exercised, then the option value of experimentation $I(p)$ is single-peaked in p, maximized at $\hat{p}$ where π_A and π_B cross. If the null action is exercised, then $I(p)$ has two peaks: one at $\pi_A = 0$ and one at $\pi_B = 0$.*

Proof of first case: By value matching and smooth pasting (6)–(7), $v - \pi_A$ rises on $(\underline{p}, 1]$, and $v - \pi_B$ falls on $[0, \bar{p})$. So $v - \pi$ is rising until π_A and π_B cross, and later falling. □

REMARK. Since information value owes to $v'' > 0$, some have suggested that the sign of v''' should be relevant for the experimentation derivative. It is instructive to see that it is not. Note that since $v''(p) = \xi(rv(p))/\Sigma(p)$, and ξ' exists by Proposition 2, v''' exists, and equals

$$v'''(p) = \frac{rv'(p)}{\Sigma(p)f(rv(p))} + \frac{2(2p-1)\xi(rv(p))}{\Sigma(p)^{3/2}}$$

Consider the R&D model, where $n' > 0$ always. Clearly, if payoffs are such that $\underline{p} < 1/2 < \bar{p}$, then $v''' > 0$ on $(1/2, \bar{p})$, while $v''' < 0$ just above $\underline{p}$, since $v'(\underline{p}) = 0$ by smooth pasting.

4.5 Expected Remaining Time and Experimentation Costs

We now analyze the behavior of the only two costs of experimentation in our model: time and money. Assume a stopping time $T < \infty$ a.s., which is true under the assumptions of Proposition 2-a or c. Then by §15.3 of Karlin and Taylor (1981) (KT81), since $\langle p_t \rangle$ has zero drift and variance $2\Sigma(p)n(p)$, the expected remaining time $\tau(p) \equiv E[T|p_0 = p]$ until stopping obeys the boundary conditions $\tau(\underline{p}) = \tau(\bar{p}) = 0$, as well as the ODE: $-1 = 0 + \Sigma(p)n(p)\tau''(p)$. Hence, $\tau''(p) < 0$, and consequently, $\tau(p)$ is hill-shaped.

We next write $v(p) = R(p) - \kappa(p)$, or the expected present value of final rewards $R(p) \equiv E[e^{-rT}\pi(p_T)|p_0 = p]$ less that of experimentation costs $\kappa(p) \equiv E[\int_0^T e^{-rt}c(n(p_t))dt|p_0 = p]$.

Since $\kappa(p)$ eventually falls near the extremes given the vanishing expected time horizon, it is clearly nonmonotonic with $\kappa(\bar{p}) = \kappa(\underline{p}) = 0$. By KT81, it also satisfies the ODE $\Sigma(p)n(p)\kappa''(p) = r\kappa(p) - c(n(p))$. When $n(p)$ and thus $c(n(p))$ is everywhere positive and U-shaped, $\kappa(p)$ is smaller and thus concave near $\underline{p}$ and $\bar{p}$. If $\kappa(p) < c(n(p))/r$ always then κ is everywhere concave, and is single-peaked. But if $\kappa(p)$ ever exceeds $c(n(p))/r$ — as seems quite plausible for middling p when $n(p)$ and thus $c(n(p))$ is U-shaped — then it must become convex, and any crossing is an inflection point; since $\kappa \in \mathcal{C}^1$, it then strangely must then be 'molar-tooth' shaped: two local maxima, bracketing two inflection points at $p_1 < p_2$, with $\kappa(p) > c(n(p))/r$ on $(p_1, p_2) \subset (\underline{p}, \bar{p})$. Which scenario arises is an open question. Quite plausibly, κ is molar-tooth shaped only for low enough r (so that experimentation lasts a long time), and disparate payoffs (extremely U-shaped costs).

4.6 Experimentation Drift

Since the belief process $\langle p_t \rangle$ is a martingale diffusion, and $v \in \mathcal{C}^2$, values $\langle v(p_t) \rangle$ are an Ito process by Ito's Lemma, and also a strict submartingale (drifts up) in $\mathcal{E}$, since $v'' > 0$ in $\mathcal{E}$. Likewise, $\langle \tau(p_t) \rangle$ is a strict supermartingale Ito process (drifts down), as $\tau'' < 0$ in $\mathcal{E}$. Finally, $\langle \kappa(p_t) \rangle$ is everywhere a strict submartingale Ito process if κ is hill-shaped, and otherwise, it must switch to a strict supermartingale inside (p_1, p_2).

By the same reasoning, the experimentation level process $\langle n(p_t) \rangle$ is a submartingale if $n(p)$ is convex. But n may well be concave in rv, as when $c(n) = n^2$. Since v is strictly convex in $p \in \mathcal{E}$ by Proposition 1, the convexity of $n(p) = f(rv(p))$ is then unclear But if the producer surplus g is weakly concave and increasing in n, then its inverse f is weakly convex and increasing in n. By Theorem 5.1 of Rockafellar (1970), the composition $f(rv)$ of a convex and increasing function f with a strictly convex function rv is strictly convex.

A more refined statement is possible when c''' exists. Indeed, differentiate the Bellman equation, that surplus equals the delay cost, $g(n(p)) = rv(p)$, to get $g'(n)n'(p) = rv'(p)$. Differentiating once more, and applying the optimal control FOC $v''(p) = c'(n(p))/\Sigma(p)$, yields a simple nonautonomous second order differential equation in the level n alone:[10]

$$[nc''(n)]n'' + [nc''(n)]'(n')^2 \equiv [g'(n)n'(p)]' = rv''(p) = rc'(n)/\Sigma(p) \tag{8}$$

Since $c' > 0$, (8) implies $n'' > 0$ at least when $[nc''(n)]' \leq 0$ — i.e., when the information producer surplus is itself concave in the experimentation level, as already asserted.

Proposition 5 *Assume (★). If the producer surplus $g(n) = nc'(n) - c(n)$ is concave, then the intensity level $n(p)$ is strictly convex in beliefs $p \in \mathcal{E}$, and thus $\langle n(p_t) \rangle$ is a submartingale.*

[10]This more simply yields the nonautonomous second order differential equation in the value v alone: $c'(f(rv(p))) = \Sigma(p)v''(p)$. Equation (8) does not necessarily hold even though v'' exists since the composition $g(n(p))$ may be twice differentiable even if we cannot write its first derivative as $g'(n(p))n'(p)$.

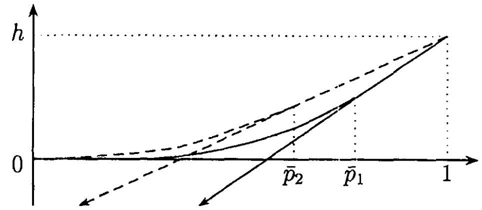

Figure 3: **A Payoff Shift in the R&D Model.** When the (unplotted) bad build payoff ℓ rises, so does the value function, and intensity levels (inside the now left-shifted interval $(\underline{p}, \bar{p})$).

For instance, any convex geometric cost function $c(n) \equiv n^k$ $(k > 1)$ yields a convex producer surplus $[nc''(n)]' = k(k-1)^2 n^{k-2} > 0$. The knife-edge case $c(n) = 1 + n \log n$ for $n > 1$ ($c(n) = n$ for $n \leq 1$) yields $nc''(n) = 1$ constant, and so linear surplus. Note that a concave surplus function $g(n)$ is *sufficient* that intensity levels $\langle n_t \rangle$ be a submartingale. In light of (8), for large enough v'', we may have still have $n'' > 0$ even if $g(n)$ is slightly convex.

REMARK. If $[nc''(n)]'$ vanishes in (8), then n is locally convex. In particular, for the R&D model, if $\lim_{n \downarrow 0} nc'''(n) = 0$ and $c''(0) < \infty$, then one can show by l'Hôpital's rule that $n''(p) > 0$ in a neighborhood $[\underline{p}, \underline{p} + \varepsilon)$ for some $\varepsilon > 0$; therefore, the experimentation level $\langle n(p_t) \rangle$ is locally a submartingale. In particular, if $c(n) \equiv n^k$, then $\lim_{n \downarrow 0} c''(n) = \lim_{n \downarrow 0} nc'''(n) = \lim_{n \downarrow 0} O(n^{k-2}) = 0$ when $k \geq 2$. So for barely profitable R&D projects and not too convex costs, the experimentation level is at least initially expected to rise.

5. SENSITIVITY ANALYSIS

5.1 Parametric Shifts

We now explore how the experimentation schedule moves with changes in the payoffs, information cost, or interest rate. Appendix C develops a method of tangents, which we brutally illustrate in Figure 3 for the R&D model. When ℓ rises, so does v, and hence n. But the tangent line $ph + (1-p)\ell$ tilts upward too, and thus the thresholds $\underline{p}$ and $\bar{p}$ must fall, since (rather loosely) the value function is less curved. This is intuitive: 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 most natural thought experiment stems from raising the payoff risk. In other words, assume that h rises, and ℓ falls so as to maintain a constant expected payoff $\pi_B(p)$ from that action at the current belief p. Intuitively, 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 level decision and stopping rule, the $\mathcal{DM}$'s current value intuitively increases. By re-optimizing, he does no worse. Hence the return rv and thus the experimentation level n both rise.

The $\mathcal{DM}$ also benefits from less convex information costs, and from more powerful signals (higher ζ). Since the value is greater, so is the intensity level. Finally, with a higher interest rate r, the $\mathcal{DM}$ enjoys a higher expected payoff (since discounting is a key cost), and is more eager to stop and act: Thresholds both shift in. Less obviously, the value often falls proportionately less than the interest rate rises, so that the return rv, and thus the intensity level n, rises.

To avoid a case-by-case analysis that adds little to the general story, the portmanteau result that follows (proved in Appendix C.1) ignores complications due to the null action.

Proposition 6 *Assume* ($\bigstar$), *and payoffs* π_a^θ *so high that the null action is never taken.*
(*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^θ *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" — namely,* $c(n)$ *is replaced by* $\hat{c}(n)$, *where* $c(0) = \hat{c}(0)$, $\hat{c}(n) - c(n)$ *convex, and corresponding slope of* $\xi = c'(f)$ *(the marginal costs at the optimum) at 0 is higher:* $\hat{\xi}'(0+) \geq \xi'(0+)$. *Then thresholds shift in, and the value* $v(p)$ *and intensity level* $n(p)$ *uniformly fall in* $\mathcal{E}$. *With no fixed costs* $\hat{c}(0) = c(0) = 0$, *this is true if* $\hat{c}(n) - c(n)$ *is convex and* $\hat{c}'(0) \geq c'(0) \geq 0$.
(*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* $\mathcal{E}$. *In the R&D model,* $n(p)$ *declines for all* $p < p'$, *and rises for all* $p > p'$, *for some* $p' \in (\underline{p}, \bar{p})$.

Surprisingly, the impatience result runs counter to the folk wisdom of Bayesian learning. More impatient decision makers typically 'experiment' less (eg. in the canonical settings of §2.B), with more myopic actions. In our model, greater impatience raises the $\mathcal{DM}$'s delay cost, and somewhere induces him to accelerate his experimentation. Further, as Proposition 7 will assert, as r blows up, the $\mathcal{DM}$ experiments at an exploding rate — albeit over a vanishing belief interval. The 'folk intuition' depends on never-ending experimentation, whereas experimentation has a finite purpose here, that one eventually stops and acts. Information accrual here is a means to an end; it is not a payoff-generating lifestyle.

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. 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.[11] Indeed, the Bellman equation for the maximization $E[\int_0^T -c(n_t)e^{-rt}dt + e^{-rt}\pi(p_T)/r]$ is still (5). But in terms of the *return* $w = rv$, it becomes $w(p) = \max_{n\geq 0}\{-c(n) + n\Sigma(p)w''(p)/r\}$. Hence, $n(p) = f(w(p))$, and w solves $w''(p) = rc'(f(w(p)))/\Sigma(p)$. 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 6.

5.2 The Return to Wald's World

We finally turn full circle, recalling our two paired twists on Wald's sequential paradigm: impatience and strict cost convexity. As noted, these assumptions cut in opposite ways. Absent payoff discounting, the $\mathcal{DM}$ sees no hurry to stack experiments, and reverts to a purely sequential mode (vanishing intensity levels) barring any fixed flow experimentation cost $c(0) > 0$. Without strict cost convexity, the $\mathcal{DM}$ faces no parallel experimentation penalty, and he becomes a classical statistician, running a massive experiment at time-0.

Proposition 7 *Assume the final payoff $\pi(p) > 0$ everywhere.*
(*a*) **Vanishing Impatience:** *For fixed $c(n)$ obeying ($\bigstar$), 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 $c_1(n), c_2(n), \ldots$ obeying ($\bigstar$). Let λ_k, Λ_k be the lower and upper Lipschitz constants of the marginal cost $c'_k : \lambda_k(n_2 - n_1) \leq c'_k(n_2) - c'_k(n_1) \leq \Lambda_k(n_2 - n_1)$ for all $n_2 > n_1 \geq 0$. Then $n_k(p)$ uniformly explodes (where positive) if $\lim_{k\to\infty} \Lambda_k = 0$, and uniformly vanishes if $\lim_{k\to\infty} \lambda_k = \infty$.*

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 $r \to 0$. 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 $p \in \mathcal{E}$, and so $n(p)$ tends down to $f(0)$.

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

Proof of the convexity limit is appendicized, but for a powerful 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.

[11] We are very grateful to Sven Rady for discovering this key difference.

6. CONCLUSION

Summary. This paper has explored the basic two action, two state decision problem under uncertainty. We have modified Wald's model by assuming an impatient $\mathcal{DM}$, and particularized it to the case of an increasing and strictly convex cost function of within-period experimentation. These two plausible economic assumptions have jointly afforded a reasonably complete Bayesian characterization of information demand: Research levels grow with the optimism over the project. For not too convex cost functions, they drift upwards over time. This yields some simple falsifiable implications in an R&D context.

Our conclusions have not come in the standard discrete time setting, but in a new continuous time control of variance for a diffusion. We consider this framework a key contribution of the paper, as it is a tractable new off-the-shelf decision model of information purchases or R&D. We believe that it can be applied in many settings, such as strategic patent races, general equilibrium R&D models, or principal-agent experimentation models.

Robustness. While we have studied a model with just two actions and two states, our main intensity-value monotonicity result $n = f(rv)$, and the associated information producer surplus intuition extends to finitely many actions and states. It also obtains in a normal learning model with state space $\Theta = \mathbb{R}$; however, the resulting problem is not stationary in the posterior mean alone, as total research outlays so far increase one's belief precision. Such a $\mathcal{DM}$ eventually becomes convinced of anything he learns, and quits.

An Extension. It is surprising that first order conditions nailed down the optimal experimentation level in Proposition 1. Apart from normal learning models, such regularity properties on the demand for information are very rare. For instance, Radner and Stiglitz (1984) have pointed out that with a one-shot information purchase, a 'non-concavity' in the value of information may emerge — its marginal value is initially zero. In a work in progress, we argue that when information can be repeatedly purchased, and its value is endogenous, a well-behaved dynamic demand theory emerges. The resulting value of information is concave in the quantity (and linear in our continuous time limit, underscored on page 10), as economists prefer. Information is quite unlike other goods, being only economically well-behaved *if the decision maker can exercise an eternal option of repurchase.* This paper might be seen as an application of this general principle to a somewhat historically important Bayesian decision problem.

Finally, we also believe that this same framework will afford a dynamic extension of Blackwell's (1953) theorem on the value of information with a much finer ordering. For large purchases of very weak and cheap signals, all the $\mathcal{DM}$ cares about is the signal-to-noise ratio. A careful formulation and development of this idea awaits our future work.

APPENDICES

A. PRIMARY MATHEMATICAL RESULTS

A.1 Properties Related to the Cost Function

Claim 1 *The surplus $g(n)$ is continuous, increasing, and unbounded given ($\bigstar$). Since $g(0) \leq 0$ ($=0$ iff $c(0)=0$), its inverse $f \equiv g^{-1}$ is continuous and increasing, with $f(0) \geq 0$.*

Proof:
$$\begin{aligned} g(n+\varepsilon) - g(n) &= [c(n) - c(n+\varepsilon)] + (n+\varepsilon)c'(n+\varepsilon) - nc'(n) \\ &> -\varepsilon c'(n+\varepsilon) + (n+\varepsilon)c'(n+\varepsilon) - nc'(n) \\ &> nc'(n+\varepsilon) - nc'(n) \geq n\lambda\varepsilon \qquad \forall \varepsilon > 0 \qquad \square \end{aligned}$$

Claim 2 *Given ($\bigstar$), the marginal cost of experimentation $\xi(w) \equiv c'(f(w))$ is increasing, strictly concave, and* Lipschitz *(i.e.* locally, *not globally so) in the return $w = rv$, for $w > 0$.*

Proof: Clearly, $\xi(0) \geq 0$ and $\xi > 0$ on $(0, \infty)$. Since $g(n) = nc'(n) - c(n)$ is continuous and strictly increasing by Claim 1, its inverse f exists and is continuous. Let $c^*(n^*) \equiv \sup_n n^* n - c(n)$ be the (Legendre-Fenchel) conjugate dual of the cost function. By Theorems 12.2 and 26.3 in Rockafellar (1970) (R70), c^* is convex since c is convex, and strictly convex as its dual $c^{**} = c$ is smooth. His Theorem 23.5-*d* then yields $c^*(n^*) + c(n) = n^* n$ for any subgradient n^* of c at n (the Fenchel-Young equality). Taking $n^* = c'(n)$, we have $c^*(c'(n)) + c(n) = nc'(n)$, or $c^*(c'(n)) = g(n)$. Finally, at $n = f(w)$, we get $c^*(\xi(w)) = c^*(c'(f(w))) = g(f(w)) = w$. The inverse relationship $\xi = (c^*)^{-1}$ obtains, and since c^* is increasing and convex, ξ is increasing and concave. Finally, ξ is locally Lipschitz away from $w = 0$, on the relative interior of its domain, by Theorem 10.4 of R70. $\square$

Observe that even though the marginal cost function $c'(n)$ need not be concave in n, the function $\xi(w) \equiv c'(f(w))$ is surprisingly concave in w. When $c''(>0)$ exists, this is easily seen: For then $f'(w) = [g'(f(w))]^{-1} = [f(w)c''(f(w)]^{-1} > 0$, and consequently, $\xi'(w) = c''(f(w))f'(w) = 1/f(w) > 0$ and $\xi''(w) = -f'(w)/f^2(w) < 0$.

Claim 3 *Given ($\bigstar$), both the inverse surplus function f and $\sqrt{f}$ are Lipschitz on $(0, \infty)$.*

Proof: If false, the Lipschitz inequality for f fails for $w_1 > w_0 > 0$ close. By the relations $w \equiv g(f(w)) = g(y) = yc'(y) - c(y)$, plus the Lipschitz-down constant λ for $c'(n)$,

$$\begin{aligned} \frac{g(f(w_1)) - g(f(w_0))}{f(w_1) - f(w_0)} &= c'(f(w_0)) - \frac{c(f(w_1)) - c(f(w_0))}{f(w_1) - f(w_0)} + f(w_1)\frac{c'(f(w_1)) - c'(f(w_0))}{f(w_1) - f(w_0)} \qquad (9) \\ &> c'(f(w_0)) - \frac{c(f(w_1)) - c(f(w_0))}{f(w_1) - f(w_0)} + \lambda f(w_1) \uparrow \lambda f(w_0) > 0 \text{ as } w_1 \downarrow w_0 \end{aligned}$$

where the limit exists as c' exists and f is continuous, and is approached from below by concavity of $\xi \equiv c'(f)$ (see Claim 2). The Lipschitz constant $1/\lambda f(w_0) \in (0, \infty)$ suffices.

For $\sqrt{f}$, the LHS of (9) is divided by $\sqrt{f(w_1)}-\sqrt{f(w_0)}$. Expressing this on the RHS as $[f(w_1)-f(w_0)]/[\sqrt{f(w_1)}+\sqrt{f(w_0)}]$, this yields the Lipschitz constant $1/2\lambda(f(w_0))^{3/2}$. □

Claim 4 *Given (★), the function $\xi(w)$ is differentiable for all $w>0$, with $\xi'(w)=1/f(w)$.*

Proof: Since ξ is concave on $(0,\infty)$, the right derivative $\xi'(w+)$ exists by Theorem 24.1 of R70. Let

$$D(y) \equiv \frac{\xi(y)-\xi(w)}{y-w} - \frac{1}{f(w)}.$$

Clearly, $\lim_{y\downarrow w} D(y) = \xi'(w+) - 1/f(w)$ exists. To see why $\xi'(w+) = 1/f(w)$, observe that using $g(n) = nc'(n) - c(n)$ and $g(f(w)) = w$ (as well as algebraic simplification), we have the identity

$$D(y) = \frac{1}{f(w)}\frac{f(y)-f(w)}{y-w}\left[\frac{c(f(y))-c(f(w))}{f(y)-f(w)} - c'(f(y))\right]$$

Given f continuous, and c differentiable, the [..] term vanishes as $y-w\downarrow 0$. We only need $1/f$ bounded ($f>0$ by Claim 1) and f Lipschitz (Claim 3). Similarly, $\xi'(w-)=1/f(w)$. □

Claim 5 *Given (★), $\Psi(p,v) \equiv \xi(rv)/\Sigma(p)$ is continuous and Lipschitz on $(0,1)\times(0,\infty)$.*

Proof: By Claim 4, ξ' exists, so that Ψ is partially differentiable in v, with $\partial\Psi/\partial v = r\xi'(rv)/\Sigma(p) = r/[f(rv)\Sigma(p)] \in (0,\infty)$ for $(p,v) \in (0,1)\times(0,\infty)$. Similarly, $\partial\Psi/\partial p = 4\xi(rv)(2p-1)/(\zeta^2 p^3(1-p)^3) \in (-\infty,\infty)$ for $(p,v) \in (0,1)\times(0,\infty)$. □

A.2 Two-Point Boundary Value Problems

a. An ODE Problem with Value Matching and Smooth Pasting. Let $\mathcal{EU}$ denote the ff. existence / uniqueness problem: For any strictly convex cost function $c \in \mathcal{C}^1$ with $c(0), c'(0) \geq 0$, there exist unique $\underline{p}, \bar{p}$ with $0 < \underline{p} < \bar{p} < 1$ and a $\mathcal{C}^2$ function v with $v''(p) = \Psi(p,v) \equiv c'(f(rv(p)))/\Sigma(p) \equiv \xi(rv(p))/\Sigma(p)$, such that (6)–(7) hold (no null action). With value matching and smooth pasting, this is not a standard two-point boundary value ODE, amenable to known free boundary results. It needs an ad hoc proof.

Observe that $\bar{p} > \hat{p}$ in $\mathcal{EU}$, where π is kinked at $\hat{p}$, since $v'' > 0$ (given $c' > 0$) and $0<\underline{p}<\bar{p}<1$. Towards solving $\mathcal{EU}$, we first consider a second order one point boundary value problem $\mathcal{IV}(\bar{p})$ of the Cauchy type: For any fixed $\bar{p}\in(\hat{p},1)$, solve $v_{\bar{p}}''(p) = \Psi(p, v_{\bar{p}}(p))$ given $v_{\bar{p}}(\bar{p}) = \bar{p}\pi_B^H + (1-\bar{p})\pi_B^L > 0$ and $v_{\bar{p}}'(\bar{p}) = \pi_B^H - \pi_B^L > 0$. Clearly, any such $v_{\bar{p}}$ is $\mathcal{C}^2$.

We now start our attack on the $\mathcal{EU}$ problem, by first assuming the R&D payoffs.

Claim 6 *(a) For all $\bar{p} \in (\hat{p},1)$, $\mathcal{IV}(\bar{p})$ has a unique solution $v_{\bar{p}}$, that can be continued on $(0,1)\times(0,\infty)$, with $v_{\bar{p}}$ and $v_{\bar{p}}'(p)$ uniformly continuous in $\bar{p}$.*

(b) For any $\bar{p} \in (\hat{p},1)$, $v_{\bar{p}}$ is positive, strictly convex, and Lipschitz either (i) on $(0,\bar{p})$, or (ii) on $(p_0,\bar{p}]$, for $p_0>0$, with $v_{\bar{p}}(p_0)=\lim_{p\downarrow p_0} v_{\bar{p}}(p)=0$.

Proof of (*a*): All claims follow given Ψ continuous and Lipschitz on $(0,1) \times (0,\infty)$ (see Claim 5), after reducing the second order ODE to a two-dimensional system of first-order ODEs. See Theorems 1.1–2 and 2.1, and Observation 2 in §1.6 of Elsgolts (1970).

Proof of (*b*): Either $v_{\bar{p}} > 0$, and $v_{\bar{p}}$ is Lipschitz on $(0,1)$ as it continues left, or $v_{\bar{p}}$ Lipschitz (valid on an open set) fails at some supremum $p_0 \in (0,\bar{p})$. If so, by Theorem 1-4.1 of Brock and Malliaris (1989), the domain of $v_{\bar{p}}$ can be compactified via $v_{\bar{p}}(p_0) = \lim_{p \downarrow p_0} v_{\bar{p}}(p) = 0$ — because Ψ is Lipschitz when $v_{\bar{p}}(p) > 0$, and Ψ is continuous and bounded in $(0,1) \times [0, \bar{p}\pi_B^H + (1-\bar{p})\pi_B^L]$. Strict convexity follows from $v_{\bar{p}}'' = \Psi(p, v(p)) > 0$ given $v_{\bar{p}}(p) > 0$.

Claim 7 (Limit Solution Behavior) (*a*) *For any $\bar{p} > \hat{p}$ that is close enough to $\hat{p}$, there exists $p_0 > 0$ such that $v_{\bar{p}}(p_0) = \lim_{p \downarrow p_0} v_{\bar{p}}(p) = 0$, with $v_{\bar{p}} > 0$ and $v_{\bar{p}}' > 0$ on $(p_0, \bar{p}]$.*
(*b*) *There is a unique $\bar{p} \in (\hat{p}, 1)$, such that for all $\bar{q} \in (\bar{p}, 1)$, a unique, strictly positive solution $v_{\bar{q}}$ exists on $(0, \bar{q}]$. Also, $\lim_{p \downarrow 0} v_{\bar{q}}'(p) = -\infty$ for all $\bar{q} > \bar{p}$.*

Proof of (*a*): If $v_{\bar{p}} > 0$ on $(0, \bar{p}]$ for all $\bar{p}$ near $\hat{p}$, then for p just below $\bar{p}$, by a Taylor expansion

$$v_{\bar{p}}(p) \approx v_{\bar{p}}(\bar{p}) - v_{\bar{p}}'(\bar{p}-)(\bar{p} - p) = v_{\bar{p}}(\bar{p}) - (\pi_B^H - \pi_B^L)(\bar{p} - p)$$

Since $v_{\bar{p}}(\bar{p}) = \bar{p}\pi_B^H + (1-\bar{p})\pi_B^L \downarrow \hat{p}\pi_B^H + (1-\hat{p})\pi_B^L = 0$ as $\bar{p} \downarrow \hat{p}$, continuity, the limit in Claim 6-*b*, case (*ii*), and $\pi_B^H - \pi_B^L > 0$ force $v_{\bar{p}}(p) < 0$ for $\bar{p}$ near $\hat{p}$ — a contradiction.

Next let p_0 be the supremum (and largest, by continuity) point where $v_{\bar{p}}(p_0) = 0$. We claim that $v_{\bar{p}}'(p_0) \geq 0$. If not, then $v_{\bar{p}}(p_0 + \varepsilon) < 0$ for small $\varepsilon > 0$. Since $v_{\bar{p}}(\bar{p}) = \pi_B^H - \pi_B^L > 0$ by construction, $v_{\bar{p}}$ crosses the p axis at some $q_0 \in (p_0, \bar{p})$, with non-negative slope. But then $v_{\bar{p}}(q_0) = 0 \leq v_{\bar{p}}'(q_0)$, and $q_0 > p_0$ has all the properties of p_0 — contrary to p_0 maximal. Finally, since $v''(p) = \Psi(p, v) > 0$, the slope $v_{\bar{p}}' > 0$ on $(p_0, \bar{p}]$, as claimed. □

Proof of (*b*): Consider any $p > \max(0, p_0)$, i.e. wherever the solution $v_{\bar{q}}(p)$ uniquely exists.

$$v_{\bar{q}}'(p) = \pi_B^H - \pi_B^L - \int_p^{\bar{q}} \Psi(z, v_{\bar{q}}(z))dz = \pi_B^H - \pi_B^L - \frac{\sigma^2}{2\mu^2}\int_p^{\bar{q}} \frac{\xi(r v_{\bar{q}}(z))}{z^2(1-z)^2}dz$$

Because $v_{\bar{q}}(z) > \varepsilon > 0$ for all $z \in [p, \bar{q}] \subset (0,1)$, this last integral strictly exceeds $\xi(r\varepsilon)\int_p^{\bar{q}} z^{-2}(1-z)^{-2}dz$, and thus blows up either as $\bar{q} \uparrow 1$, or as $p \downarrow 0$ for any $\bar{q} > \bar{p}$. So $\lim_{\bar{q} \uparrow 1} v_{\bar{q}}'(p) = -\infty$ for fixed p, and $\lim_{\bar{p} \downarrow 0} v_{\bar{q}}'(p) = -\infty$ for $\bar{q} > \bar{p}$. Since $v_{\bar{q}}'(\bar{q}) = \pi_B^H - \pi_B^L > 0$, $v_{\bar{q}}(p) > 0$ is strictly convex and U-shaped in $(0, \bar{q}]$ for $\bar{q}$ close enough to 1.

Claim 8 (Monotonicity Properties) *The solution $v_{\bar{p}}$ rises and flattens as $\bar{p}$ rises: For all threshold pairs $\bar{p}_1 < \bar{p}_2$ in $(\hat{p}, 1)$, we have (a) $v_{\bar{p}_1} < v_{\bar{p}_2}$, and (b) $v_{\bar{p}_1}' > v_{\bar{p}_2}'$.*

Proof: Suppose (*a*) fails at some $p' < \bar{p}_1$ — say, the largest p' where $v_{\bar{p}_1}, v_{\bar{p}_2}$ cross: $\Delta v(p') \equiv v_{\bar{p}_1}(p') - v_{\bar{p}_2}(p') = 0$. Because $\Delta v < 0$ on $(p', 1)$, this requires $\Delta v'(p') \leq 0$. Since $\bar{p}_2 > \bar{p}_1$, strict convexity of $v_{\bar{p}_2}$ in $(\bar{p}_1, \bar{p}_2)$ yields $v_{\bar{p}_1}'(\bar{p}_1) = \pi_B^H - \pi_B^L = v_{\bar{p}_2}'(\bar{p}_2) > v_{\bar{p}_2}'(\bar{p}_1)$, and so

$\Delta v'(\bar{p}_1) > 0$. By continuity of $\Delta v'$, $\Delta v'(p'') = 0$ for some $p'' \in (p', \bar{p}_1)$, i.e. $v'_{\bar{p}_1}(p'') = v'_{\bar{p}_2}(p'')$, or

$$v'_{\bar{p}_1}(\bar{p}_1) - \int_{p''}^{\bar{p}_1} \Psi(z, v_{\bar{p}_1}(z))dz = v'_{\bar{p}_2}(\bar{p}_2) - \int_{p''}^{\bar{p}_2} \Psi(z, v_{\bar{p}_2}(z))dz \tag{10}$$

Cancelling $v'_{\bar{p}_1}(\bar{p}_1) = \pi_B^H - \pi_B^L = v'_{\bar{p}_2}(\bar{p}_2)$ from both sides of (10), the integrals on either side of (10) are equal. But this is impossible, because $\bar{p}_1 < \bar{p}_2$, and $\Psi(z, v_{\bar{p}_1}(z)) < \Psi(z, v_{\bar{p}_2}(z))$ since $v_{\bar{p}_1}(z) < v_{\bar{p}_2}(z)$ given $\Delta v(z) < 0$ for $z \geq p'' > p'$.

Finally, part (b) follows from $v'_{\bar{p}}(p) = (\pi_B^H - \pi_B^L) - \int_p^{\bar{p}} \Psi(z, rv_{\bar{p}}(z))dz$. For the domain of integration rises with $\bar{p}$, and thus so does the integral, given $\Psi \geq 0$. And we've just shown that $v_{\bar{p}}(z)$, and thus $\Psi(z, v_{\bar{p}}(z))$ rises uniformly for z in the domain of the solution. □

Theorem 1 ($\mathcal{EU}$ Existence/Uniqueness) *Given (★), a unique $\bar{p} < 1$ exists s.t. $\mathcal{IV}(\bar{p})$ has a unique solution $v_{\bar{p}}$ for a unique $\underline{p} > 0$ obeying the boundary conditions* (6), (7) *of $\mathcal{EU}$.*

Proof: Choose $\bar{q} > \bar{p}$ close enough to 1 that $v_{\bar{q}}(p)$ uniquely exists and is strictly positive on $(0, \bar{q}]$. As $\bar{q}$ decreases, $v_{\bar{q}}(p)$ uniformly shifts down by Claim 8. But as long as $v_{\bar{q}} > 0$ over $(0, \bar{q}]$, it is U-shaped by Claim 7-b, with $v_{\bar{q}}(0+) = -\infty < 0 < \pi_B^H - \pi_B^L = v_{\bar{q}}(\bar{q})$. Since it is also strictly convex by Claim 6-b, with exactly one global minimum in $(0, \bar{q})$, it can first cross the horizontal axis only in the way required by $\mathcal{EU}$. The assertion then fails only if $v_{\bar{q}}$ always stays strictly positive for all $\bar{q}$ — which we ruled out in Claim 7-a.

Finally, the above argument only made essential use of the rising $\pi_B(p)$ curve. A falling $\pi_A(p)$ line could easily substitute for the horizontal axis, which was thus WLOG. □

b. An ODE Problem with Value Matching Alone.

Claim 9 *Fix $\underline{p}' < \widehat{p} < \bar{p}'$. The two-point boundary value problem $\widetilde{v}''(p) = \Psi(p, r\widetilde{v}(p))$ s.t. $\widetilde{v}(\underline{p}') = \pi(\underline{p}')$ and $\widetilde{v}(\bar{p}') = \pi(\bar{p}')$ has a solution $\widetilde{v}$.*

Proof: Consider the Cauchy problem $\widetilde{v}''_s(p) = \Psi(p, r\widetilde{v}_s(p))$ s.t. $\widetilde{v}_s(\bar{p}') = \pi(\bar{p}') > 0$, $\widetilde{v}'_s(\bar{p}') = s$, for some (slope) $s \in \mathbb{R}$. By the properties of Ψ in Claim 5, a unique solution $\widetilde{v}_s$ to this problem exists locally for each $s \in \mathbb{R}$, is uniformly continuous in s, and can be extended left as long as it remains positive: $\widetilde{v}_s(p) = \pi(\bar{p}') + s(p - \bar{p}') + \int_{\bar{p}'}^p \int_{\bar{p}'}^x \Psi(y, \widetilde{v}_s(y))dxdy$. Next,

$$\widetilde{v}_{s_2}(p) - \widetilde{v}_{s_1}(p) = (s_2 - s_1)(p - \bar{p}') + \int_{\bar{p}'}^p \int_{\bar{p}'}^x [\Psi(y, \widetilde{v}_{s_2}(y)) - \Psi(y, \widetilde{v}_{s_1}(y))]dxdy < 0 \tag{11}$$

for every $s_2 > s_1$, and all p, *where both solutions $\widetilde{v}_{s_1}, \widetilde{v}_{s_2}$ exist.* This obtains because $0 = \widetilde{v}_{s_2}(\bar{p}') - \widetilde{v}_{s_1}(\bar{p}') < s_2 - s_1 = \widetilde{v}'_{s_2}(\bar{p}') - \widetilde{v}'_{s_1}(\bar{p}')$. Inequality then holds by continuity for p near $\bar{p}'$; and so $\Psi(p, \widetilde{v}_{s_2}(p)) < \Psi(p, \widetilde{v}_{s_1}(p))$ by monotonicity of Ψ; continuing leftward, $\widetilde{v}_{s_2} < \widetilde{v}_{s_1}$. Hence, $\lim_{s \to \pm\infty} \widetilde{v}_s(p)$ exists for all $p \leq \bar{p}'$ provided $\widetilde{v}_s(p) \geq 0$. In fact, $\lim_{s \to -\infty} \widetilde{v}_s(p) = \infty$ because $\widetilde{v}_{s_2}(p) - \widetilde{v}_s(p) < (s_2 - s)(p - \bar{p}')$ from (11), provided $\widetilde{v}_s, \widetilde{v}_{s_2} \geq 0$. Likewise, $\widetilde{v}_s(p) = 0$ for large enough s. Therefore by continuity and monotonicity of $\widetilde{v}_s(\cdot)$ in s, for all $p < \widehat{p}$ there exists s with $\widetilde{v}_s(p) = \pi(p) \in [0, \infty)$, and in particular for $p = \underline{p}'$. □

B. DERIVATION OF OPTIMAL CONTROL FOUNDATIONS

This appendix rigorously formulates the control/stopping problem and the sequence of proofs that establish the existence and uniqueness properties of its solution. We first summarize our plan of attack, and highlight the absence of any logical circularity. We ignore the null action throughout, as its additional difficulties (with a horizontal tangency because $\pi = 0$ somewhere) are fully captured by the R&D case: a vanishing intensity level.

§**B.1.** We progressively restrict the control process to simpler domains. First, we posit the SDE met by the observation process $\langle \bar{x}_t \rangle$, given any control functional $\mathfrak{n}(\cdot)$ of past observations. Next, we derive the Bayes filter SDE describing the evolution of $\langle p_t \rangle$. Then we show how to formulate the $\mathcal{OCS}$ problem with current beliefs p_t as state variable: Since the best existing theory does not do this for joint optimal control and stopping, we first find weak conditions to write the value function as $V = V(p)$, and then prove Lemma 1 (V convex in p); this yields a simple belief exit set $[0, \underline{p}] \cup [\bar{p}, 1]$, and a Markov stopping time T. Finally, we show that we may restrict WLOG to Markov control policies $n(p)$.

§**B.2.** We establish key properties of the belief diffusion to prove that any candidate optimal policy of beliefs alone (as §B.1 allows) a.s. stops experimenting in finite time.

§**B.3.** Assuming $c(0) > 0$ or $\pi > 0$ everywhere (or both), we verify that the candidate optimal Markov control $n(p) = f(rv(p))$ from $\mathcal{EU}$ is admissible in the sense of §B.1, justifying uniqueness presumed so far. We then check that the unique solution to $\mathcal{EU}$ (and thus to $\mathcal{HJB}$+$\mathcal{ST}$) also solves $\mathcal{OCS}$ ($V = v$ works, with $\underline{p}, \bar{p}$), and also is its only solution.

§**B.4.** We specialize §B.3 to the R&D model with $c(0) = 0$ and $\pi = 0$ somewhere.

B.1 Developing the Markov Control Model

a. Drift and Noise Notation. Let $\{\Omega, \mathcal{F}, \mathcal{P}\}$ be the underlying probability space. Let $h_t(\omega)$ or $h(t, \omega)$ denote the same continuous time *stochastic process* $h = \langle h_t \rangle_{t \geq 0}$, namely a collection of $\mathcal{F}$-measurable real-valued functions (random variables) $h : \mathbb{R}_+ \times \Omega \to \mathbb{R}$. We consider only processes with continuous sample paths, i.e. where $h(\cdot, \omega)$ is continuous for every $\omega \in \Omega$. We denote by $\mathcal{C}^*_{[0,\infty)}$ the space of continuous functions on any time subset $[0, t] \subseteq [0, \infty)$; by $h^t(\omega) = \langle (s, h(s, \omega)) \rangle_{s \leq t}$ the (graph of the) process restricted to $[0, t]$, whose (superscripted) elements are sample paths or trajectories in $\mathcal{C}^*_{[0,\infty)}$; by h^{t-} the restriction of h^t to $[0, t)$; $\langle \mathcal{F}^h_t \rangle$ is the filtration generated by h. Finally, $W = \langle W_t, \mathcal{F}^W_t \rangle$ is the standard Wiener process or *standard Brownian Motion* (BM) with zero drift and unit variance, adapted only to $\langle \mathcal{F}^W_t \rangle$, and (canonically) with continuous sample paths.

Let $p_0 \in (0, 1)$. Think of Nature as initially independently drawing a drift $m \in \{-\mu, \mu\}$, with chances $p_0, 1 - p_0$, and a BM noise path realization $W^\infty \in \mathcal{C}_{[0,\infty)}$. To render our problem amenable to existing results in nonlinear filtering theory, we must define an underlying unobserved *process* that the $\mathcal{DM}$ wishes to infer (not simply the stationary drift), and an observable process that provides noisy information on the former (the controlled signal). This construction works even though the unobservable process is time-invariant — e.g. Oksendal (1995) (hereafter, O95), Example 6.11.

Put $\Omega = (0,2]$, and endow it with the Borel σ-algebra $\mathcal{F} = \mathcal{B}_{(0,2]}$, and partition $\Omega \equiv \Omega_L \cup \Omega_H \equiv (0,1] \cup (1,2]$. Let $W : \mathbb{R}_+ \times (0,1] \to \mathbb{R}$ be a standard BM on $\{(0,1], \mathcal{B}_{(0,1]}, \mathcal{P}_W\}$, where $\mathcal{P}_W$ and W exist by the Kolmogorov extension theorem. Then W extends to a standard BM on $\{(1,2], \mathcal{B}_{(1,2]}, \mathcal{P}_W\}$ by translation. Next, define probability measures $\mathcal{Q}(\omega) = p_0 \mathbb{I}_{\omega\in\Omega_H} + (1-p_0)\mathbb{I}_{\omega\in\Omega_L}$ and $\mathcal{P}(\omega) = \mathcal{Q}(\omega)\mathcal{P}_W(\omega)$ on $\{\Omega, \mathcal{F}\}$. Finally choose $\mu > 0$ and define the time-invariant process $m : \mathbb{R}_+ \times \Omega \to \{-\mu, \mu\}$ such that $m_t(\omega) \equiv m(\omega) = \mu$ for all $\omega \in \Omega_H$ and $m(\omega) = -\mu$ otherwise. Note the equivalence: payoff-relevant state $\theta \in \{H, L\} \leftrightarrow$ drift $m \in \{\mu, -\mu\} \leftrightarrow$ partition element $\Omega_\theta \in \{\Omega_H, \Omega_L\}$.

The *non-anticipatory control functional* of time and sample paths $\mathfrak{n} : \mathbb{R}_+ \times \mathcal{C}^*_{[0,\infty)} \to \mathbb{R}_{++}$ satisfies (i) $\mathfrak{n}(t, h^\infty) = \mathfrak{n}(t, h^t)$ for all t,[12] and (ii) $\mathfrak{n}(t, h^t(\omega))$ is also $\mathcal{F}_t$-adapted for each $\mathcal{F}_t$-adapted process h. Fix $\sigma > 0$. The controlled signal process $\langle \bar{x}_t(\omega)\rangle$ is a diffusion with $\bar{x}_0(\omega) = 0$, solving:

$$d\bar{x}_t(\omega) = m(\omega)dt + \frac{\sigma}{\sqrt{\mathfrak{n}(t, \bar{x}^t(\omega))}} dW_t(\omega). \tag{12}$$

b. The Continuous-time Non-Linear Bayes Rule Filter. Given state-contingent payoffs and vNM preferences, the $\mathcal{DM}$'s expected payoff from stopping depends on the controlled signal diffusion only via the posterior probability of $m = \mu$ (state H), given time and past observations of the signal and own control $p_t = \mathcal{P}(m = \mu \mid p_0, t, \bar{x}^t, n^{t-})$.

Our model, with the constant state process m and signal process obeying (12), meets all conditions of Theorem 9.1 in LS77 provided that there exists a unique strong solution to the SDE (given a prespecified Wiener process $\langle \bar{W}_t\rangle$), established in §B.3. Their result, a special case of the so-called 'fundamental theorem of non-linear filtering', asserts that: (i) the process $p_t = \mathcal{P}(\Omega_H \mid \mathcal{F}^{\bar{x}}_t, n^{t-})$ obeys $p_t(\omega) = p_0$, and solves the SDE:[13]

$$\begin{aligned} dp_t(\omega) &= p_t(\omega)(1-p_t(\omega))\frac{\mu-(-\mu)}{\sigma}\sqrt{\mathfrak{n}(t,\bar{x}^t(\omega))}d\bar{W}_t(\omega) \\ \bar{W}_t(\omega) &\equiv (1/\sigma)\textstyle\int_0^t \sqrt{\mathfrak{n}(s,\bar{x}^s(\omega))}\{d\bar{x}^s(\omega) - [p_s(\omega)\mu + (1-p_s(\omega))(-\mu)]ds\} \end{aligned} \tag{13}$$

and that (ii) $\bar{W} = \langle \bar{W}_t, \mathcal{F}^{\bar{x}}_t\rangle$ is a Wiener process, but *not* a standard BM. That is, it is adapted not only to its own filtration, but also to the smaller one $\langle \mathcal{F}^{\bar{x}}_t\rangle$.

c. Beliefs as a Sufficient Statistic.

[12] We shall choose $h = \bar{x}$, and soon after $h = p$. Of course, at time t, we need the joint history of $\bar{x}^t$ and n^{t-} to compute the current optimal control n_t. However, given the functional $\mathfrak{n}$ and $\bar{x}^t$, the control history n^{t-} is implicitly recursively embedded; therefore, we shall omit it altogether for notational simplicity.

[13] More recently, Bolton and Harris (1993) provide alternative insights into this formula.

• THE CONTROL AS A FUNCTION JUST OF BELIEF HISTORY. Since we want to work with beliefs rather than signals as a state variable, we establish:

Claim 10 *There is a bijection between information sets $\langle p^t, n^{t-}\rangle$ and $\langle \bar{x}^t, n^{t-}\rangle$. Hence, p^t is sufficient for $\bar{x}^t$, and we may WLOG replace $\mathfrak{n}(t, \bar{x}^t(\omega))$ by $\mathfrak{n}(t, p^t(\omega))$.*

Proof: We first prove that $\langle \bar{x}^t, n^{t-}\rangle \mapsto \langle p^t, n^{t-}\rangle$ is an injection. Fix the control functional $\mathfrak{n}$, and states $\omega, \omega' \in \Omega$ whose resulting signal paths differ: $\bar{x}_s(\omega) \neq \bar{x}_s(\omega')$ for some $s \in [0, t]$. By continuity, this holds on a time set of positive measure. Then we either have (i) $\mathfrak{n}(t, \bar{x}^t(\omega)) = n^{t-}(\omega) = n^{t-}(\omega') = \mathfrak{n}(t, \bar{x}^t(\omega'))$, and given the filter (13), belief paths differ, $p^t(\omega) \neq p^t(\omega')$, or (ii) $n^{t-}(\omega) \neq n^{t-}(\omega')$. Either way, $\{p^t(\omega), n^{t-}(\omega)\} \neq \{p^t(\omega'), n^{t-}(\omega')\}$.

That $\{p^t, n^{t-}\} \mapsto \{\bar{x}^t, n^{t-}\}$ is an injection follows from strong existence and uniqueness of the diffusion $p_t(\omega)$, which imply that every signal process path $\bar{x}^t(\omega)$ and associated control process $\mathfrak{n}(t, \bar{x}^t(\omega))$ yield a unique belief process $p_t(\omega)$ via the filter. □

• THE VALUE AS A FUNCTION JUST OF CURRENT BELIEFS. Given the absence of suitable sufficiency theorems for optimal control joint with optimal stopping — Krylov (1980) being the best source here — we proceed indirectly, via the value function.

Claim 11 *The supremum (4) can be written as $V(p)$, a function of the current belief p.*

Proof: The random variable $T : \Omega \mapsto [0, \infty]$ is a *Markov time* relative to $\langle \mathcal{F}_t^W\rangle$ if $\{\omega : T(\omega) \leq t\} \in \mathcal{F}_t^W$ for all $t \geq 0$. The $\mathcal{DM}$ must optimally stop the one-dimensional controlled process $\langle p_t(\omega)\rangle$ solving (13) for an adapted control process $n_t(\omega)$ and a Markov time $T(\omega)$, w.r.t. the measure $\mathcal{P}(\omega)$ on Ω. The general results in Krylov (1980) assume a continuous terminal reward function π (like us; his more powerful verification theorems require $\pi \in \mathcal{C}^2$). His Theorem 3.1.9 provides a recursive equation for $V(t, p_t|\bar{T})$, namely the supremum value at time t of the length-$\bar{T}$ horizon problem. His Theorem 3.1.10 states that this upper bound is achieved via a feedback control functional $\mathfrak{n}(t, p^t)$. Next, by Theorem 6.4.4, since $\lim_{\bar{T}\to\infty} E_\omega[e^{-r(\bar{T}-t)}\pi(p_{\bar{T}-t})] = 0$ for all $t > 0$ and adapted control processes n_t, the value of the infinite horizon problem is $V(t, p_t) := \lim_{\bar{T}\to\infty} V(t, p_t|\bar{T})$. Finally, by his Remark 6.4.13, $V(t, p_t) = V(p_t)$ for all t, since our $\Sigma(\cdot)$, $c(\cdot)$, $\pi(\cdot)$, and r are stationary. □

• BELIEF THRESHOLDS AND VALUE FUNCTION CONVEXITY: PROOF OF LEMMA 1. We can now prove that the optimal stopping decision is described by a stopping set in current belief space $[0, 1]$. By Claim 11, the supremum value can be written as $V(p)$. By using a policy that is ε-optimal for the $V(p_0)$ problem (i.e. it achieves a payoff $\geq V(p_0) - \varepsilon$) for any $p \in (p_1, p_2)$, the $\mathcal{DM}$ can ensure himself an expected payoff that is affine in p. For such a strategy yields a constant payoff in either state $\theta = H, L$, and adjusting p merely weights these payoffs — since the Wiener probability law $\mathcal{P}_W$ is independent of the ω partition element $\{\Omega_\theta\}$, while $m = \pm\mu$ in states H, L. Thus, V everywhere weakly exceeds some supporting ε-tangent line at p. As $\varepsilon > 0$ and p are arbitrary, V is convex.

Next, define stopping sets S_A, S_B, and S_0 for A, B, and the null action. Since point beliefs are stationary for (13), $V(0) = \pi(0)$ and $V(1) = \pi(1)$. As stopping is always an option, we have $V \geq \pi$; equality $V(p) = \pi(p)$ holds iff $p \in S_A \cup S_H \cup S_0$. If ever $V(p) = \pi$ so that $p \in S_A$, then convexity, $V \geq \pi$, and $0 \in S_B$, together force $V = \pi$ on $[0, p]$. Thus, $S_A = [0, \underline{p}]$, and similarly $S_B = [\bar{p}, 1]$ for some $0 \leq \underline{p} \leq \bar{p} \leq 1$. Similarly, since $\pi_A^L > 0$ and $\pi_B^H > 0$, the null action is never exercised at $p=0$ and $p=1$, so that $S_0 = [\underline{p}_0, \bar{p}_0]$. □

• THE CONTROL AS A FUNCTION JUST OF CURRENT BELIEFS. We have now proven that the stopping rule is characterized by an exit set in the space $[0,1]$ of current beliefs. We may therefore simplify the supremum operator in (4) as $\sup_{T,\mathrm{n}} \equiv \sup_{\{\underline{p},\bar{p}\}}[\sup_{\mathrm{n}}]$. For each $\{\underline{p}, \bar{p}\}$, resolving the inner supremum is crucially a pure control exercise. Under very weak conditions guaranteed by our simple two-point boundaries and boundedly finite maximand, for any optimal $\mathcal{F}_t^{\bar{x}}$-adapted control, there exists a Markov control that performs as well, by Theorem 11.3 in O95. Hence, only the *current* belief p_t matters for control. Hereafter, we restrict to stationary Markov controls, of the type $\mathrm{n}(t, p^t(\omega)) = n(p_t(\omega))$, where $n : [0,1] \to \mathbb{R}_+$ is a $\mathcal{B}_{[0,1]}$-measurable function. Thus, we consider beliefs as the solution to the SDE $dp_t = \sqrt{2\Sigma(p_t)n(p_t)}d\bar{W}_t$ (given p_0), on any continuation set $\varnothing \neq (\underline{p}, \bar{p}) \subset [0,1]$.

Definition *An* admissible *stationary Markov control policy $n \in \mathcal{M}$ is a strictly positive $\mathcal{B}_{[0,1]}$-measurable function $n : [0,1] \to \mathbb{R}_{++}$ yielding a unique strong solution to* (12)–(13).

d. The Properly Formulated Optimization Problem. Replacing (4) given (2), the $\mathcal{OCS}$ problem when WLOG restricting to stationary Markov controls is therefore now:

$$V(p_0) = \sup_{n(\cdot)\in\mathcal{M};\underline{p},\bar{p}\in[0,1]} \int_\Omega \left[\int_0^{T(\omega|\underline{p},\bar{p})} -c(n(p_t(\omega)))e^{-rt}dt + e^{-rT(\omega|\underline{p},\bar{p})}\pi\left(p_{T(\omega|\underline{p},\bar{p})}\right) \right] d\mathcal{P}(\omega) \quad (14)$$

$$\begin{aligned} \text{s.t.} \quad & T(\omega \mid \underline{p}, \bar{p}) = \inf\{t \geq 0 : p_t(\omega) \leq \underline{p} \text{ or } p_t(\omega) \geq \bar{p}\} \\ & p_t(\omega) = p_0 + \textstyle\int_0^t \sqrt{2n(p_s(\omega))\Sigma(p_s(\omega))}d\bar{W}_t(\omega) \qquad (15) \\ & \bar{W}_t(\omega) \equiv (1/\sigma)\textstyle\int_0^t \sqrt{n(p_s(\omega))}\{d\bar{x}_s(\omega) - [p_s(\omega)\mu + (1 - p_s(\omega)(-\mu))]ds\} \end{aligned}$$

e. Brief Aside: Making Sense of the Bayesian Model of §4.2*A*. In the text, we formulated this result without measure theory for illustrative purposes. Yet our summary was not without basis. We defined a conditional signal process $\langle\bar{x}_t^\theta\rangle$, which we can now more generally write $d\bar{x}_t^\theta(\omega) = \mathbb{I}_{\omega\in\Omega_\theta}[\mu^\theta(\omega)dt + (\sigma/n_t)dW_t(\omega)]$, for $\theta = L, H$; the belief driving force $d\bar{W}_t$, from the unconditional point of view of the $\mathcal{DM}$, was really the mixture of two maps, $\bar{W}_t(\omega) = p_t(\omega)\bar{W}_t^H(\omega) + (1 - p_t(\omega))\bar{W}_t^L(\omega)$, where (3) is equivalent to this definition if premultiplied by the indicator function $\mathbb{I}_\omega$ of Ω_θ (suppressed in the text, along with ω).

B.2 Experimentation Ends Almost Surely in Finite Time

We now assume that a stationary optimal policy $\langle n(\cdot), \underline{p}, \bar{p}\rangle$ exists, with $n(\cdot)$ admissible. We then show that experimentation a.s. ends in finite time, as Theorem 5 will need. Note that positivity, demanded by admissibility, is not a restriction on optimality. For if $n(p')=0$ at $p' \in (\underline{p}, \bar{p}) \subseteq [0,1]$, then $dp|_{p'} = n(p')\Sigma(p')d\bar{W} = 0$ since $\Sigma < \infty$, contrary to $p' \in \mathcal{E}$.

a. Background Terminology, Notation, and Theory. Here, we hew closely to KT81, §15.6–7, and Karatzas and Shreve (1991) (KS91), §5.5.C (adapting both for clarity, and to avoid overuse of letters). Below, p_0, p, y, and z are arbitrary points in $(\underline{p}, \bar{p})$. The *hitting time* of $z \in [\underline{p}, \bar{p}]$ from p_0 is the random variable $T_{p_0 z} \equiv \inf\{t \geq 0 : p_t = z \mid p_0\}$. Next, the diffusion $\langle p_t\rangle$ is *regular* if, for any pair of interior points $y, z \in (\underline{p}, \bar{p})$, y can be reached in *finite time with positive probability* (hereafter, FTPP) from z, or $\Pr(T_{zy} < \infty) > 0$. This is a critical notion, analogous to the communicating property for Markov processes.

We now define several important integrals of $1/n(p)$, needed both here and in §B.3,B.4. Though they may be unbounded, they are well-defined, because the control $n(\cdot) > 0$ is $\mathcal{B}_{[0,1]}$-measurable. Given a general belief diffusion process $dp = \beta(p)dt + \alpha(p)dW$ with range $(\underline{p}, \bar{p}) \subseteq (0,1)$, the *scale function* of KT81, KS91 is $S(y) \equiv \int_{p_0}^{y} \exp(-2\int_{p_0}^{x} \beta(z)dz/\alpha^2(z))dx$. By KT81's two convenient abuses of notation, $S[y,p] \equiv S(p) - S(y)$ denotes the *scale measure,* and $S(\underline{p}, p] \equiv \lim_{y\downarrow \underline{p}} S[y,p]$ its left-side limit. The boundary $\underline{p}$ is *attracting* if $S(\underline{p}, p] < \infty$, independently of $p \in (\underline{p}, \bar{p})$. A similar definition applies at $\bar{p}$. By an equivalent formulation in KT81's Lemma 6.1, an attracting boundary is strictly speaking non-repelling: $\langle p_t\rangle$ approaches the boundary (say $\underline{p}$) arbitrarily close with positive chance starting from any interior point p_0, before any larger interior $b > p_0$ is hit.

As an attracting boundary might not be reached in FTPP, we need a harsher concept: Similar to the scale function, define the *speed function* $M(y) \equiv \int_{p_0}^{y}[S_\theta'(z)\alpha^2(z)]^{-1}dz$, the *speed measure* $M[y,p]$, and $M(\underline{p},p] \equiv \lim_{y\downarrow\underline{p}} M[y,p]$. Next, let $J[y,p] \equiv \int_y^p S[y,z]M'(z)dz$ and $K[y,p] \equiv \int_y^p S[z,p]M'(z)dz$, with right limits $J(\underline{p},p]$ and $K(\underline{p},p]$. The boundary $\underline{p}$ is *attainable* if $J(\underline{p},p] < \infty$ for any $p \in (\underline{p}, \bar{p})$. Intuitively, by Lemma 6.2 in KT81 this means $\Pr(T_{p\underline{p}} < \infty) > 0$. An attainable boundary is thus not just approached but hit in FTPP.

b. Attracting Boundaries. We explore the unconditional driftless ($\beta(p)=0$) belief process, with diffusion term $\alpha(p) = \sqrt{2n(p)\Sigma(p)}$ in $(\underline{p}, \bar{p})$. Then $S[y,p] = p - y$ as $n > 0$.

Claim 12 *Assume $\pi > 0$ in $[0,1]$, or $c(0) > 0$. For an optimal control $n(p)$, the belief process $\langle p_t = p_0 + \int_0^t \sqrt{2n(p_s)\Sigma(p_s)}d\bar{W}_s\rangle$ is regular, and any triggers $0 \leq \underline{p} \leq \bar{p} \leq 1$ attracting.*

Proof: That boundaries are attracting follows from $S(\underline{p}, p] < \infty$ and $S[p, \bar{p}) < \infty$, both essentially stated above. Regularity is proven by contradiction. Suppose that there exists a pair $y, z \in (\underline{p}, \bar{p})$ with $\Pr(T_{yz} < \infty) = 0$. Obviously $y \neq z$; WLOG order $\underline{p} < y < z < \bar{p}$.

Proof Step 1: Regularity on Some Interval. Given sample path continuity and $y < z < \bar{p}$, all diffusion paths from y to $\bar{p}$ first hit z, and so $\Pr(T_{yz} \leq T_{y\bar{p}}) = 1$. Hence, $\Pr(T_{y\bar{p}} < \infty) \leq \Pr(T_{yz} < \infty) = 1 - 1 = 0$. But then $\Pr(T_{y\underline{p}} < \infty) > 0$, for otherwise starting at belief y, the $\mathcal{DM}$ strictly suboptimally experiments at cost without (a.s.) any prospect of positive discounted returns. Once more sample path continuity yields $\Pr(T_{y''y'} < \infty) \geq \Pr(T_{y''\underline{p}} < \infty) > 0$ for any $y' < y''$ with $[y', y''] \subseteq [\underline{p}, y]$. To wit, the diffusion process $\langle p_t \rangle$ traverses *down* across any such $[y', y'']$ to y' in FTPP. For regularity on $[y', y'']$, $\langle p_t \rangle$ also must transit *up* in FTPP. Pick any $y' = \underline{y}$ in $(\underline{p}, y)$, and assume to the contrary that no $y'' > \underline{y}$ in $(\underline{p}, y)$ is hit in FTPP from $\underline{y}$. Then the belief process drifts down but not up in FTPP, which violates the martingale property.[14]

Proof Step 2: The Maximal Regularity Interval. By Step 1, $\langle p_t \rangle$ is a regular process on a non-empty interval $[\underline{y}, y''] \subset [\underline{p}, y)$. We next claim that the largest connected regularity set with lower bound $\underline{y}$ is right-open, of the form $[\underline{y}, \bar{y})$. To see why, we prove that if $\langle p_t \rangle$ is regular on $[\underline{y}, y'']$, then it is also regular on $[y'', y'' + \eta]$ for small $\eta > 0$. Assume not. Specifically, suppose $\Pr(T_{y''y''+\eta} < \infty) = 0$ for all $\eta > 0$. Because $\Pr(T_{\underline{y}z} < \infty) = 0$ with $\underline{y} < y$, and $\Pr(T_{\underline{y}y''} < \infty) > 0$, we have $y'' \leq z < \bar{p}$, and so $y'' + \eta < \bar{p}$ for small enough $\eta > 0$. As in Step 1, this contradicts the martingale property of $\langle p_t \rangle$.

Proof Step 3. Contradiction to Optimality. Since $\bar{y}$ is unattainable in FTPP starting from any $p \in [\underline{y}, \bar{y})$, so is $\bar{p} > \bar{y}$. Then the $\mathcal{DM}$'s only reason to experiment is reaching the lower boundary $\underline{p}$ in FTPP. We now argue that this yields a negative payoff.

Since $S(\underline{y}, \bar{y}] = \bar{y} - \underline{y} < \infty$, the boundary $\bar{y}$ is attracting. Steps 1–2 proved that $\bar{y}$ is an unattainable boundary of a regular process on $[\underline{y}, \bar{y})$; thus, $J[\underline{y}, \bar{y}] = \infty$ by KT81's Lemma 6.2. Because also $J[\underline{y}, \bar{y}] + K(\underline{y}, \bar{y}] = S(\underline{y}, \bar{y}]M(\underline{y}, \bar{y}]$ by KT81's Lemma 6.3-*v*, $0 < S(\underline{y}, \bar{y}] < \infty$ forces $M(\underline{y}, \bar{y}] = \infty$. By KT81's Lemma 6.3-*iii*, this implies $K(\underline{y}, \bar{y}] = \infty$. Thus, $\bar{y}$ is a *natural* (*Feller*) boundary and so their Theorem 7.2-*ii* conveniently says $\lim_{p \uparrow \bar{y}} \lim_{q \uparrow \bar{y}} E[e^{-rT_{qp}}] = 0$. Both limits exist as the expectation is monotone in p and q, and lies in $[0, 1]$. Respecting the order of limits, this means: For each $\varepsilon > 0$, there exist p_ε and q_ε, with $\underline{y} < p_\varepsilon < q_\varepsilon < \bar{y}$, such that $E[e^{-rT_{q_\varepsilon p_\varepsilon}}] < \varepsilon$. Since $\underline{p} < p_\varepsilon < q_\varepsilon$, the expected gross discounted returns starting from q_ε are $R(q_\varepsilon) = E[e^{-rT_{q_\varepsilon \underline{p}}}]\pi(\underline{p}) \leq E[e^{-rT_{q_\varepsilon p_\varepsilon}}]\pi(\underline{p}) \leq \varepsilon\pi(\underline{p})$.

Since the $\mathcal{DM}$ stops only at the lower threshold $\underline{p}$, the expected total discounted experimentation costs $\kappa(q_\varepsilon)$ are positive for all $q_\varepsilon > \underline{p}$, increasing as q_ε rises to $\bar{y}$ as $\varepsilon \to 0$,

[14] Here's a proof. First, for any $t > 0$, if $T \geq 0$ is a stopping time, then so is $t \wedge T \equiv \min\langle t, T \rangle$. Since $\Pr(T_{y\underline{p}} < \infty) > 0$, we have $t \wedge T_{y\underline{p}} = T_{y\underline{p}}$ with positive chance for large enough $t < \infty$. Unless $t \wedge T_{y\underline{p}} \wedge T_{\underline{y}y''} = T_{\underline{y}y''}$ with positive probability for some $y'' > \underline{y}$ (and so y'' is hit in FTPP from $\underline{y}$), $\Pr(t \wedge T_{\underline{y}\underline{p}} \wedge T_{\underline{y}y''} = T_{\underline{y}\underline{p}}) = \Pr(t \wedge T_{\underline{y}\underline{p}} = T_{\underline{y}\underline{p}}) > 0$ for all $y'' > \underline{y}$. Then $E[p_{t \wedge T_{\underline{y}\underline{p}} \wedge T_{\underline{y}y''}} | \underline{y}] \leq \underline{y} + \Pr(t \wedge T_{\underline{y}\underline{p}} = T_{\underline{y}\underline{p}})[\underline{p} - \underline{y}] < \underline{y}$. But Wald's Optional Stopping Theorem guarantees equality because $\langle p_t \rangle$ is a martingale.

and hence boundedly positive. Thus, the value $v(q_\varepsilon) = R(q_\varepsilon) - \kappa(q_\varepsilon) \leq \varepsilon\pi(\underline{p}) - \kappa(q_\varepsilon) < \pi(q_\varepsilon)$ for small enough $\varepsilon > 0$, because either $c(0) > 0$ and thus $\kappa(q_\varepsilon) > 0$, or $\pi(q_\varepsilon) > 0$, or both. This contradicts the assumption $q_\varepsilon \in (\underline{y}, \bar{y}) \subset (\underline{p}, \bar{p})$. □

c. Finite Hitting Times. In summary, regularity only fails if the intensity level "nearly vanishes" around some interior belief $\bar{y} \in (\underline{p}, \bar{p})$. This implies that the belief process could approach $\bar{y}$, WLOG from below, only at a vanishing speed; and by the martingale property, the speed also vanishes moving downward, towards the only attainable threshold $\underline{p}$. But then all *discounted* payoffs starting at belief $p_0 \in (\underline{p}, \bar{y})$ come from hitting $\underline{p}$, which vanish as p_0 nears $\bar{y}$. The $\mathcal{DM}$ then wishes to stop immediately, contradicting $p_0 \in (\underline{p}, \bar{p})$.

Theorem 2 *Assume $\pi > 0$ in $[0,1]$, or $c(0) > 0$, or both. If a solution $\{n(\cdot), \underline{p}, \bar{p}\}$ to $\mathcal{OCS}$ exists, then the expected time to hit $\underline{p}$ or $\bar{p}$ is finite. The hitting time is thus a.s. finite.*

Proof Step 1: Both boundaries $\underline{p}, \bar{p}$ are attainable from any interior belief: Indeed, optimality at $p_0 \in (\underline{p}, \bar{p})$ precludes $\Pr(T_{p_0\underline{p}} < \infty) = \Pr(T_{p_0\bar{p}} < \infty) = 0$. If either is zero, say $\underline{p}$ is unattainable in FTPP, then by the logic of Step 3 above, $v(p_0) < \pi(p_0)$ for p_0 close to $\underline{p}$.

Proof Step 2: With attainable boundaries, $J(\underline{p}, p], J[p, \bar{p}) < \infty$ by KT81's Lemma 6.2. But by Proposition 5.5.32-i in KS91,[15] these two inequalities are iff for our result *assuming* (i) *nondegeneracy* (ND'), or a positive diffusion term $n\Sigma > 0$ in $(\underline{p}, \bar{p})$, and (ii) *local integrability* (LI'), or $\forall p \in (\underline{p}, \bar{p})\ \exists \varepsilon > 0$ with $M[p_0 - \varepsilon, p_0 + \varepsilon] = \int_{p-\varepsilon}^{p+\varepsilon} dy/[n(y)\Sigma(y)] < \infty$. Clearly, ($ND'$) obtains because $n(p) > 0$ by optimality, and $\Sigma(p) > 0$ inside $(0,1)$.

For a contradiction, assume (LI') fails for some $p_0 \in (\underline{p}, \bar{p})$. Then $M[p_0 - \varepsilon, p_0 + \varepsilon] = \infty$ for all $\varepsilon > 0$ — say, small enough that $\underline{p} < p_0 - \varepsilon < p_0 + \varepsilon < \bar{p}$. Next, all points in $[p_0 - \varepsilon, p_0 + \varepsilon]$ are attainable in FTPP from p_0, as the diffusion is regular in its superinterval $(\underline{p}, \bar{p})$. Thus, $J[p_0 - \varepsilon, p_0 + \varepsilon] < \infty$ by attainability, while $M[p_0 - \varepsilon, p_0 + \varepsilon] = \infty$. By KT81's classification (end of §15.6-b), both $p_0 \pm \varepsilon$ are then *exit boundaries* of $[p_0 - \varepsilon, p_0 + \varepsilon]$. The contradiction proof from Claim 12 (step 3) with KT81's Theorem 7.2-ii still applies. □

B.3 Verification Theorem: Existence and Uniqueness of a Solution.

a. Admissibility Conditions. We check the conditions (see the definition in §B.1-c) for our candidate Markov control $n(p) = f(rv(p))$ to be admissible. Until §B.4, we assume the non-R&D and non- null action case with $\pi \gg 0$, to avoid some technical issues.

- Positivity. This has already been addressed at the outset of §A.2.

- Measurability. The composition of a Borel measurable map with any measurable map X_t preserves $\mathcal{F}_t^X$ measurability (O95, Lemma 2.1); also, being constant in time, $n(\cdot)$

[15]In our notation, their result says: *The stopping time T is finite a.s. iff (i) both boundaries are attainable* $[J(\underline{p}, p] < \infty$ *and* $J[p, \bar{p}) < \infty$ (same as their v function)]*; in this case, more strongly $E(T) < \infty$.*

is measurable in the time Borel sets. So for any $\mathcal{B}_{[0,1]}$-measurable function $n(\cdot)$, the control process $n(p_t(\omega))$ is $\mathcal{B}_{\mathbb{R}_+} \otimes \mathcal{F}$-measurable, and $\mathcal{F}_t^W$- and $\mathcal{F}_t^{\bar{x}}$-adapted, like $p_t(\omega)$. Since the Markov control function $f(rv(p))$ is continuous, it is therefore $\mathcal{B}_{[0,1]}$-measurable.

• STRONG EXISTENCE AND UNIQUENESS (SEU). Given our candidate optimal policy $\{f(rv, \underline{p}, \bar{p}\}$, the belief process has a unique strong solution. Indeed, replace the functional $\mathrm{n}(t, \bar{x}^t)$ with our candidate Markov control $f(rv(p_t))$ in (12), and substitute into (15) for $\bar{W}_t(\omega)$, and then $p_t(\omega)$. Suppressing ω,

$$dp_t = p_t(1-p_t)(2\mu/\sigma)\left([f(rv(p_t))/\sigma]\,[m - \mu(2p_t - 1)]dt + \sqrt{f(rv(p_t))}dW_t\right) \quad (16a)$$

$$= \begin{cases} p_t(1-p_t)^2\zeta^2 f(rv(p_t))dt + p_t(1-p_t)\zeta\sqrt{f(rv(p_t))}dW_t \\ -p_t^2(1-p_t)\zeta^2 f(rv(p_t))dt + p_t(1-p_t)\zeta\sqrt{f(rv(p_t))}dW_t \end{cases} \quad (16b)$$

where $\zeta = 2\mu/\sigma$. As the probability space and Wiener process W are primitives, we need only verify that there exists a unique *strong* (KS91, §5.2.1) solution to each corresponding SDE $dp_t = \beta_\theta(p, m)dt + \alpha(p)dW$ on Ω_θ, $\theta = L, H$, where α and β_θ are implicitly defined by (16b). If so, the process $\langle p_t(\omega)\rangle$ solving (16a) is a diffusion, being a mixture (weights $p_0, 1-p_0$) of the conditional diffusions (16b), and so inherits the defining properties (KS91, §5.1.1). An adapted control process $\langle f(rv(p_t(\omega)))\rangle$ is induced, and a uniquely defined observation process $\bar{x}_t$, after substitution into (12).

Because our belief SDEs are defined on a bounded open interval $(\underline{p}, \bar{p})$, all requirements below need only hold up to any hitting time T, at which point the process is absorbed. Existence of a *weak* solution on $(\underline{p}, \bar{p})$ — i.e. up to any hitting time — for $\theta = L, H$ follows by Skorohod's Theorem (KS91, Theorem 5.4.22) from $p_0 \in (0,1)$, and the continuity and boundedness of α and β_θ on $(\underline{p}, \bar{p})$. Just as with ODEs, an SDE is uniquely soluble given a Lipschitz condition, here joint on the drift and variance terms, β_θ and α. This requirement is met because f and $\sqrt{f}$ Lipschitz (by Claim 3) and v differentiable imply that $f(rv(p))$ and $\sqrt{f(rv(p))}$ are Lipschitz in $p \in (\underline{p}, \bar{p})$. (A proof mimicks that of the composition chain rule in calculus.) By Proposition 5.2.13 in KS91 (with $h(x) = x$ in their equation (2.25)), this implies *strong uniqueness* and then *pathwise* uniqueness (by KS91, Remark 5.3.3). Skipping around their book, along with weak existence, this yields SEU (KS91, p.310).

b. Value Function Characterization and Verification of the $\mathcal{OCS}$ Policy.

Theorem 3 *The supremum value V of $\mathcal{OCS}$ (14)–(15) equals v, and $\{p : V(p) > \pi(p)\} = (\underline{p}, \bar{p})$, where $\{v, \underline{p}, \bar{p}\}$ solves $\mathcal{EU}$ (5)–(7).*

PROOF STEP 1: CONTROL. The supremum value $V(p)$ of $\mathcal{OCS}$ clearly exists — and by Lemma 1, it uniquely defines a continuation region $\mathcal{E} = (\underline{p}^*, \bar{p}^*) \subseteq (0,1)$, by $[0, \underline{p}^*] \cup [\bar{p}^*, 1] \equiv$

$\{p \in [0,1] : V(p) = \pi(p)\}$. For any admissible Markov control $n(p)$ and resulting belief process $\langle p_t(\omega)\rangle$, define $V(p_0)$ as in (14), except using stopping time $T(\omega|\underline{p}^*, \bar{p}^*)$. Since π is continuous, this is a pure control problem with given boundaries $\underline{p}^*, \bar{p}^*$.

For a sufficiency verification theorem, we check three conditions. First, there exists a solution $\tilde{v}(p)$ of the $\mathcal{HJB}$ problem with value matching only at $\underline{p}^*, \bar{p}^*$. By Claim 9, a solution $\tilde{v}(p)$ to the ODE $\tilde{v}''(p) = c'(f(r\tilde{v}(p)))/\Sigma(p)$ and value matching (6) exists for any thresholds in $[0,1]$, such as $\underline{p}^*, \bar{p}^*$, on the boundary of the set $V(p) = \pi(p)$. Second, $\tilde{v} \in \mathcal{C}^2$ in $[\underline{p}, \bar{p}]$ because Σ is continuous in $[\underline{p}, \bar{p}]$, and $\xi = c'(f)$ is continuous on $[0, \infty)$ by Claim 2, so $\tilde{v}''(p)$ is continuous. Third, the family of functions $\{\tilde{v}(p_t)\}_{t \leq T}$ is uniformly integrable for all Markov controls $n(\cdot)$, any stopping time realization T, and any process $\langle p_t\rangle$ starting in $(\underline{p}, \bar{p})$, simply because $|v(p_t)|$ is boundedly finite (see O95, Appendix C.3). Therefore, by Theorem 11.2 in O95: $V(p) = \tilde{v}(p) \in \mathcal{C}^2$ in $[\underline{p}, \bar{p}]$, and there exists an optimal (WLOG Markov, as proved in §B.1.c) control $n(p) = f(r\tilde{v}(p)) = f(rV(p))$ for $\mathcal{OCS}$.

Proof Step 2: Stopping. Next, replace $n(p)$ in (13) with our candidate $f(rV(p))$. Given this controlled diffusion $p_t(\omega)$, consider the resulting pure optimal stopping problem:

$$V(p_0) = \sup_{T \geq 0} E_\omega \left[\int_0^T -c(f(rV(p)))e^{-rs}ds + e^{-rT}\pi(p_T(\omega))\right]$$

with $p_0(\omega) = p_0$ — where the value of this stopping problem is our supremum $V(p)$, because $f(rV(p))$ is an optimal control for $\mathcal{OCS}$, from Step 1. By Theorem 3.15 in Shiryayev (1978) (S78), extended to a continuous flow cost $c(f(rV(\cdot)))$ and final payoff $\pi(\cdot)$, and geometric discounting by his Remark 3.8.3, for all $p \in \mathcal{E}$, V solves the generalized Stefan problem $rV(p) = -c(f(rV(p))) + f(rV(p))\Sigma(p)V''(p)$ s.t. value matching at $\partial\mathcal{E}$. Hence, $V'' > 0$ in $\mathcal{E}$, and $V = \pi$ in $[0,1] \setminus \bar{\mathcal{E}}$ (where $V'' = 0$). Therefore V is continuous and convex in $[0,1]$, and by Theorem 24.1 in R70 it has right and left derivatives in $[0,1]$, and so at $\underline{p}^*, \bar{p}^*$. Also, as shown in §B.2.a, both boundaries are attracting for any positive control. The necessity of smooth pasting then follows from Theorem 3.16 in S78, i.e. V solves $\mathcal{ST}$, i.e. $rv(p) = -c(n(p)) + n(p)\Sigma(p)v''(p)$, plus (6)–(7), for the given control $n(p) = f(rV(p))$.

Summing up, the triple $\{V, \underline{p}^*, \bar{p}^*\}$ solves $\mathcal{HJB}$ (Step 1) and $\mathcal{ST}$ (Step 2). As sketched after Proposition 1, the Bellman equation (5) has a unique maximizer $f(rv)$, and so is equivalent to $v'' = c'(f(rv(p)))/\Sigma(p)$. Thus, $\mathcal{HJB}$+$\mathcal{ST}$ are jointly equivalent to $\mathcal{EU}$, and a unique solution $\{v(p), \underline{p}, \bar{p}\}$ exists; therefore $v(p) = V(p)$, $\underline{p}^* = \underline{p}$, $\bar{p}^* = \bar{p}$, as asserted. □

Theorem 4 *Assume $\pi > 0$ in $[0,1]$, or $c(0) > 0$, or both. The Markov control policy $n(p) = f(rv(p))$ and thresholds $\underline{p}, \bar{p}$ from $\mathcal{EU}$ are optimal for the $\mathcal{OCS}$ problem.*

Proof: Since $v(p) = V(p)$ and $f(rV(p))$ is an optimal control (Step 1 of Theorem 3), the *$\mathcal{EU}$ control* $f(rv(p))$ is not only admissible (§B.3.a) but also optimal for $\mathcal{OCS}$. By Theorem 3.3 in S78, the stopping time $T = T_{p_0\underline{p}} \wedge T_{p_0\bar{p}}$ — and then the whole $\mathcal{EU}$ policy — is optimal for $\mathcal{OCS}$ if $T < \infty$ a.s, which we now establish. First, $f(rv(p)) > 0$ for all $p \in (\underline{p}, \bar{p})$ implies that conditions (ND') and (LI') in KS91 5.5 hold for the resulting controlled belief SDE (see the proof of Theorem 2), and $S(\underline{p}, p] < \infty$, $S[p, \bar{p}) < \infty$. Therefore by Proposition 5.5.32-(i) in KS91, $T < \infty$ a.s. iff $J(\underline{p}, p] = \int_{\underline{p}}^{p}(z - \underline{p})dz/[f(rv(z))\Sigma(z)] < \infty$ and $J[p, \bar{p}) = \int_{p}^{\bar{p}}(\bar{p} - z)dz/[f(rv(z))\Sigma(z)] < \infty$. Since $(\underline{p}, \bar{p}) \subset (0, 1)$, then $\Sigma(z) > 0$ and these inequalities follow from $f(rv(\cdot)) \gg 0$ in $(\underline{p}, \bar{p})$ — true given $c(0) > 0$ or $\pi > 0$ always. □

c. Verification: Uniqueness of the Optimal Policy for $\mathcal{OCS}$.

Theorem 5 *Assume $\pi > 0$ in $[0, 1]$, or $c(0) > 0$, or both. Any solution to the $\mathcal{OCS}$ problem (14)–(15) solves $\mathcal{EU}$ (5)–(7). Since $\{f(rv(\cdot)), \underline{p}, \bar{p}\}$ solves $\mathcal{OCS}$ by Theorem 4, the unique solution of $\mathcal{EU}$ is the unique optimal policy for $\mathcal{OCS}$.*

Proof: Because $\pi(p)$ and $V(p)$ are uniquely defined, so is the boundary of the region where $V(p) = \pi(p)$, i.e. $\{\underline{p}^*, \bar{p}^*\} \equiv \{\underline{p}, \bar{p}\}$. Consider the uniqueness of the optimal control $f(rv(p))$. We have proven above that $V(p) = v(p) \in \mathcal{C}^2$ in $(\underline{p}^*, \bar{p}^*) = (\underline{p}, \bar{p})$, and in Theorem 2 that the stopping time is finite a.s. for *any* Markov optimal policy $n(p)$ — not just for $f(rv(p))$. So by Theorem 11.1 in O95 any optimal $n(p)$ must solve $\mathcal{HJB}$ with value $V(p)$, i.e. $rV(p) = -c(n(p)) + n(p)\Sigma(p)V''(p) = \max_n[-c(n) + n\Sigma(p)V''(p)]$ for $p \in (\underline{p}, \bar{p})$. Thus $c'(n(p)) = \Sigma(p)V''(p) = c'(f(rV(p)))$. □

B.4 The R&D Model / Null Action Case ($\pi = 0$ Somewhere)

When $\pi(p) = 0$ somewhere and $c(0) = 0$, then boundary problems arise: For $n(\underline{p}) = f(rv(\underline{p})) = 0$ as in Figure 2 (left), and $\underline{p}$ may be unattainable in FTPP: So Theorem 2 fails.

Theorem 6 *Assume payoffs $\pi(p) = 0$ for some $p \in [0, 1]$. Then the $\mathcal{EU}$ policy is uniquely optimal if either $\rho \equiv \lim_{w \downarrow 0} w^{1-\eta}\xi'(w)/\xi(w) < \infty$ for some $\eta > 0$, or $c(0) > 0$, or both.*

Proof: Clearly, $f(rv(\cdot)) > 0$ in $(\underline{p}, \bar{p})$, while $f(rv(z)) \gg 0$ near $\bar{p}$ since $\pi(\bar{p}) > 0$ WLOG. As in the proof of Theorem 4, the result obtains iff $J(\underline{p}, p] < \infty$, i.e. $\int_{\underline{p}}^{p}(z - \underline{p})/f(rv(z))\, dz < \infty$. If $c(0) > 0$ then $f(rv(\cdot)) \gg 0$ in $(\underline{p}, \bar{p})$ and the integral is finite. If $c(0) = 0$ and $\pi = 0$ somewhere, so that $f(rv(\underline{p})) = 0$, we show this is still true. Indeed, near $\underline{p}$, we have

$$\frac{z - \underline{p}}{f(rv(z))} \simeq \frac{1}{f(rv(z))}\frac{2v(z)}{(z - \underline{p})v''(z)} = \xi'(rv(z))\frac{2rv(z)\Sigma(z)}{(z - \underline{p})r\xi(rv(z))} \tag{17}$$

The approximate equality follows from a Taylor expansion $v(z) \simeq v''(z)(z - \underline{p})^2/2$ (valid because $v \in \mathcal{C}^2$) with $v(\underline{p}) = v'(\underline{p}) = 0$, the second equality from $v''(z) = \xi(rv(z))/\Sigma(z)$ and from $\xi'(w) = 1/f(w)$, as established in §A.1. Notice that $2\Sigma(z)/r \to 2\Sigma(\underline{p})/r \equiv \underline{\Sigma} \in (0, \infty)$

as $r > 0$ and $\underline{p} \in (0,1)$. Observe that:

$$\lim_{z \downarrow \underline{p}} \frac{z-\underline{p}}{f(rv(z))}(z-\underline{p})^{1-\eta} = \Sigma \lim_{z \downarrow \underline{p}} \frac{\xi'(rv(z))rv(z)}{\xi(rv(z))(z-\underline{p})^{\eta}} = \rho\Sigma \lim_{z \downarrow \underline{p}} \left[\frac{rv(z)}{z-\underline{p}}\right]^{\eta} = \rho\Sigma \lim_{z \downarrow \underline{p}} (rv'(z))^{\eta} = 0$$

where the first equality follows from (17), the second from the premise about ξ, and the third from a Taylor expansion $0 = rv(\underline{p}) \simeq rv(z) + rv'(z)(\underline{p} - z)$ which is valid because $v \in \mathcal{C}^2$. It follows that the integrand $(z-\underline{p})/f(rv(z))$ is bounded above by $(z-\underline{p})^{\eta-1}$, whose integral on $[\underline{p}, p]$ is bounded for any $\eta > 0$, which establishes the claim. □

REMARK. If $c(0) = c'(0) = 0$, then in the proof of Claim 14, $\xi(rv(z))/rv(z) \simeq \xi'(rv(z)) = 1/f(rv(z))$ for $rv(z)$ small. Thus the integrand in the proof of Claim 14 is of order $(z-\underline{p})^{-1}$, whence $J(\underline{p}, p] = \infty$. Hence, $\Pr(T_{p_0\underline{p}} \wedge T_{p_0\bar{p}} < \infty) < 1$, and no optimal stopping time exists. More precisely, the $\mathcal{EU}$ control $n(p) = f(rv(p))$ with the new thresholds $\{\underline{p} + \varepsilon, \bar{p}\}$ is an ε-optimal policy. See Theorem 3.3 in S78.

C. OMITTED PROOFS: SENSITIVITY ANALYSIS

C.1 Comparative Statics: Proof of Proposition 6

We first consider payoff shifts, which just affect the boundary conditions of $\mathcal{EU}$. We then employ a different methodology for shifts in r, ζ, and $c(\cdot)$ that skew the ODE itself.

• PROOF OF PROPOSITION 6(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 RHS of (4) — is differentiable in π_a^θ by inspection: $V(p_0|\cdot)_{\pi_B^H} = E[e^{-rT}\bar{p} \mid p_T = \bar{p}, p_0, n]Pr(p_T = \bar{p} \mid p_0, n) > 0$. Here the event $p_T = \bar{p}$ that the $\mathcal{DM}$ 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: $V_{\pi_a^\theta}(p_0|\cdot) > 0$ for all $p_0 > \underline{p}$ for a fixed policy n, T. If the $\mathcal{DM}$ 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 (7), 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 $\underline{p}_0 > \underline{p}_1$. □

• PROOF OF PROPOSITION 6(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 a left-right reflection of Claim 7-b, $\underline{p}$ must fall to restore a smoothly-pasted tangency of $v_{\underline{p}}(p)$ and $\pi(p)$ on the right side. Also, as Claim 8 asserts, as $\underline{p}$ falls, the value $v_{\underline{p}}(p)$ rises, and gets steeper. □

Claim 13 (A Key Implication) *Parametrize models by* Γ_1, Γ_2, *where* $\Gamma_i = \{c_i(\cdot), \zeta_i, r_i\}$. *Let* $\{v_i, \underline{p}_i, \bar{p}_i\}$ *denote the corresponding solutions, with* $\Psi_i(p, v_i(p)) \equiv c_i'(f_i(r_i v_i(p)))/\Sigma_i(p)$. *If*

$$v_2 \geq v_1 \Rightarrow \Psi_2(p, v_2) > \Psi_1(p, v_1) \tag{18}$$

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 $(\underline{p}_1, \bar{p}_1) \cap (\underline{p}_2, \bar{p}_2) \to \mathbb{R}$.

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 (18), 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 $\tilde{p} \in (\underline{p}_1, \min\langle \bar{p}_1, \bar{p}_2 \rangle)$. 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 $\tilde{p}$ 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 $\underline{p}_i$; 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 $\tilde{p}$. In either case, since $\Delta v(\underline{p}_1 + \varepsilon) > 0$, we deduce the maximum $\Delta v(\tilde{p}) > 0$. But then $\Delta v''(\tilde{p}) = \Psi_2(\tilde{p}, v_2(\tilde{p})) - \Psi_1(\tilde{p}, v_1(\tilde{p})) > 0$ because $\Delta v(\tilde{p}) > 0$, as just deduced. This violates the second order condition $\Delta v''(\tilde{p}) \leq 0$ for a maximum of Δv at $\tilde{p}$.

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. Suppose that $\Delta v(p) \geq 0$ at some $p \in (\underline{p}_2, \bar{p}_2)$. Then Δv must attain a global non-negative maximum for one such $\tilde{p} \in (\underline{p}_2, \bar{p}_2)$. This requires $\Delta v(\tilde{p}) \geq 0 \geq \Delta v''(\tilde{p})$ — yielding the same contradiction as before. □

• PROOF OF PROPOSITION 6 (*c*): COST CONVEXITY. Define $\Delta c(n) \equiv c_2(n) - c_1(n)$. Convexity of Δc and $\Delta c(0) = 0$ imply $\Delta c'(n) > \Delta c(n)/n$ for all $n > 0$. Equivalently, we have $g_2(n) > g_1(n)$, and thus $f_2(w) < f_1(w)$ for all $w > 0$, so that $\xi_2'(w) = 1/f_2(w) > 1/f_1(w) = \xi_1'(w)$. Since $\xi_2'(0) \geq \xi_1'(0)$ by assumption, $\xi_2(w) > \xi_1(w)$ for all $w > 0$. [Or, if $c_i(0) = 0$, then $f_i(0) = 0$ for $i = 1, 2$, so that $\Delta c'(0+) \geq 0$ implies $\xi_2(0) = c_2'(f_2(0)) \geq c_1'(f_1(0)) = \xi_1(0)$.] Thus, (18) holds: $\Psi_i(\cdot, v_i(\cdot)) = \xi_i(rv_i(\cdot))/\Sigma(\cdot)$, so that $\xi_i(\cdot)$ increasing and $\xi_2(\cdot) > \xi_1(\cdot)$. This proves value and threshold monotonicity. Since also $f_2(w) < f_1(w)$, and each f_i is increasing, we have $n_2(p) = f_2(rv_2(p)) < f_2(rv_1(p)) < f_1(rv_1(p)) = n_1(p)$. □

• PROOF OF PROPOSITION 6 (*d*): INFORMATION QUALITY. If $\zeta_1 > \zeta_2$, then $\Sigma_1(\cdot) > \Sigma_2(\cdot)$ and therefore the premise of Claim 13 holds: $\Psi_i(\cdot, v_i(\cdot)) = \xi(rv_i(\cdot))/\Sigma_i(\cdot)$, so that

$\xi(\cdot)$ increasing and $1/\Sigma_1(\cdot) < 1/\Sigma_2(\cdot)$ imply (18). 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. □

Claim 14 (Proof of Proposition 6(*e*): Impatience) *Given interest rates $r_2 > r_1 > 0$.*
(a) The r_2-thresholds are shifted in, and the r_2-value v lower at all points in its domain;
(b) 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 (a): Since $r_2 > r_1$ with all else equal, ξ increasing implies (18) and Claim 13. □

Proof of (b): Let $w_i(p) = r_i v_i(p)$ for $p \in (\underline{p}_i, \bar{p}_i)$, $i = 1, 2$. 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, $\Delta v''(p) = [\xi(w_2(p)) - \xi(w_1(p))]/\Sigma(p) > 0$, and so $w_2(p) > w_1(p)$ and $n_2(p) > n_1(p)$. 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 $\Sigma(p) w_i''(p) = r_i \xi(w_i(p))$ s.t. $w_i(\underline{p}_i) = r_i \pi(\underline{p}_i)$, $w_i'(\underline{p}_i) = r_i(\pi_A^H - \pi_A^L)$, $w_i(\bar{p}_i) = r_i \pi(\bar{p}_i)$, $w_i'(\bar{p}_i) = r_i(\pi_B^H - \pi_B^L)$. 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 non-empty set $I \subseteq [\underline{p}_2, \bar{p}_2]$, it suffices to show that Δw cannot have a local non-negative maximum in $(\underline{p}_2, \bar{p}_2)$. By contradiction, suppose that $\Delta w(\tilde{p}) \geq 0 = \Delta w'(\tilde{p}) \geq \Delta w''(\tilde{p})$ for some $\tilde{p} \in (\underline{p}_2, \bar{p}_2)$. Then $\tilde{p} \in (\underline{p}_2, \bar{p}_2) \subset [0, 1]$ implies $\Sigma(\tilde{p}) > 0$; furthermore $r_2 > r_1$, $\xi(\cdot)$ is increasing and $\Delta w(\tilde{p}) \geq 0$, so that the familiar contradiction follows: $\Delta w''(\tilde{p}) = [\Sigma(\tilde{p})]^{-1}[r_2 \xi(w_2(\tilde{p})) - r_1 \xi(w_1(\tilde{p}))] > 0 \geq \Delta w''(\tilde{p})$.

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 non negative 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 $p > p'$. □

C.2 Vanishing / Exploding Convexity: Completion of Proof of Proposition 7

Assume $n_2 > n_1$, with $\lambda_k \to \infty$. By Claim 1, $g_k(n_2) - g_k(n_1) > n_1 \lambda_k (n_2 - n_1)$. Then $g_k(n) > g_k(0) + \int_0^n dg_k(n') = 0 + \int_0^n \lambda_k n' \, dn' = \lambda_k n^2/2 \to \infty$ as $k \to \infty$. Similarly, $g_k(n) \to 0$ as

$k \to \infty$ when $\lambda_k \to 0$. Hence, the inverse f_k of g_k explodes away from 0. Since $v \geq \pi \gg 0$, $n_k(p) = f(rv(p)) \geq f(r\pi(p))$ explodes. The proof for $\Lambda_k \to 0$ is symmetric. $\square$

References

ARROW, K., D. BLACKWELL, and M. GIRSHICK (1949): "Bayes and Minimax Solutions of Sequential Design Problems," *Econometrica*, 17, 213–244.

BLACKWELL, D. (1953): "Equivalent Comparison of Experiments," *Annals of Mathematics and Statistics*, 24, 265–272.

BOLTON, P., and C. HARRIS (1993): "Strategic Experimentation," Nuffield College mimeo.

BROCK, W. A., and A. G. MALLIARIS (1989): *Differential Equations, Stability and Chaos in Dynamic Economics.* North Holland, New York.

CHERNOFF, H. (1972): *Sequential Analysis and Optimal Design*, vol. 8 of *Regional Conference Series in Applied Mathematics.* Society for Industrial and Applied Mathematics, Philadelphia.

CRESSIE, N., and P. MORGAN (1993): "The VPRT: A Sequential Testing Procedure Dominating the SPRT," *Econometric Theory*, 9, 431–450.

DIMASI, 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.

DUTTA, P. K. (1997): "Optimal Management of an R&D Budget," *Journal of Economic Dynamics and Control*, 21, 575–602.

EASLEY, D., and N. KIEFER (1988): "Controlling a Stochastic Process with Unknown Parameters," *Econometrica*, 56, 1045–1064.

ELSGOLTS, L. (1970): *Differential Equations and Calculus of Variations.* Mir Publishers, Moscow.

KAMIEN, M., and N. SCHWARTZ (1971): "Expenditure Patterns for Risky R&D Projects," *Journal of Applied Probability*, 8, 60–73.

KARATZAS, I., and S. E. SHREVE (1991): *Brownian Motion and Stochastic Calculus.* Springer-Verlag, New York [KS91].

KARLIN, S., and H. M. TAYLOR (1981): *A Second Course in Stochastic Processes.* Academic Press, San Diego [KT81].

KELLER, G., and S. RADY (1997): "Optimal Experimentation in a Changing Environment," Stanford GSB Research Paper No. 1443.

KIHLSTROM, R., L. MIRMAN, and A. POSTLEWAITE (1984): "Experimental Consumption and the 'Rothschild Effect'," in *Bayesian Models in Economic Theory*, ed. by M. Boyer, and R. Kihlstrom, pp. 279–302. Elsevier Science Publishers, New York.

KRYLOV, N. V. (1980): *Controlled Diffusion Processes*. Springer-Verlag, New York.

LIPTSER, R., and A. N. SHIRYAYEV (1977): *Statistics of Random Processes: General Theory*, vol. I. Springer-Verlag, New York [LS77].

——— (1978): *Statistics of Random Processes: Applications*, vol. II. Springer-Verlag, New York.

MCLENNAN, A. (1984): "Price Dispersion and Incomplete Learning in the Long Run," *Journal of Economic Dynamics and Control*, 7, 331–347.

NEYMAN, J., and E. S. PEARSON (1933): "On the Problem of the Most Efficient Tests of Statistical Hypotheses," *Philosophical Transactions of the Royal Society*, 231, 140–185.

OKSENDAL, B. (1995): *Stochastic Differential Equations: An Introduction with Applications*. Springer Verlag, New York [O95].

RADNER, R., and J. STIGLITZ (1984): "A Nonconcavity in the Value of Information," in *Bayesian Models in Economic Theory*, ed. by M. Boyer, and R. Kihlstrom, pp. 33–52. Elsevier Science Publishers, New York.

ROCKAFELLAR, T. (1970): *Convex Analysis*. Princeton University Press, Princeton [R70].

SHIRYAYEV, A. N. (1978): *Optimal Stopping Rules*. Springer-Verlag, New York [S78].

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

WALD, A. (1945a): "Sequential Method of Sampling for Deciding Between Two Courses of Action," *Journal of the American Statistical Association*, 40.

——— (1945b): "Sequential Tests of Statistical Hypotheses," *Annals of Mathematics and Statistics*, 16.

——— (1947a): *Sequential Analysis*. Wiley, New York.

WALD, A., and J. WOLFOWITZ (1948): "Optimum Character of the Sequential Probability Ratio Test," *Annals of Mathematics and Statistics*, 19, 326–329.

WALLIS, W. A. (1980): "The Statistical Research Group, 1942–45," *The American Statistician*, 75.