Birkbeck Working Papers in Economics & Finance

ISSN 1745-8587

Department of Economics, Mathematics and Statistics

BWPEF 1505

Dynamic Project Selection

Arina Nikandrova
Birkbeck, University of London

Romans Pancs
University of Rochester

April 2015

Dynamic Project Selection

Arina Nikandrova and Romans Pancs*

April 2015

Abstract

We study a normative model of an internal capital market, used by a company to choose between its two divisions' pet projects. Each project's value is initially unknown to all but can be dynamically learned by the corresponding division. Learning can be suspended or resumed at any time and is costly. We characterize an internal capital market that maximizes the company's expected cash flow. This market has indicative bidding by the two divisions, followed by a spell of learning and then firm bidding, which occurs at an endogenous deadline or as soon as either division requests it.

Keywords: internal capital market, irreversible project selection

JEL codes: D82, D83, G320, G310.

*Nikandrova (a.nikandrova@gmail.com) is at Birkbeck; Pancs (rpancs@gmail.com) is at the University of Rochester. We thank Dmitry Orlov, Vladimir Parail, Marek Pycia, Jeanine Miklos-Thal, Juuso Toikka, and Michael Waldman for valuable input at different stages of this project. This paper would not have been written without George Georgiadis's incisive modelling suggestion, made while Pancs was on sabbatical at UCLA. For their feedback, we are also grateful to seminar audiences in Rochester, ITAM, Bielefeld, the University of Surrey, the Australian National University, the University of Sydney, and the University of Queensland.

1 Introduction

A corporate finance textbook (e.g., Webster, 2003, Chapter 12) would recommend that a company invest in a project if and only if the project's internal rate of return exceeds the cost of capital. If companies operated in this manner, their investment decisions would be independent across projects within the company, conditional on the projects' cash flows. In practice, such independence is an exception rather than the rule (Ozbas and Scharfstein, 2010).

Investment decisions can be interdependent for two reasons: projects may be mutually exclusive, or internal capital, used to finance these projects, may be scarce. We use the term "internal capital market" to describe a project-selection mechanism that deals with either situation. We are interested in the design of an optimal internal capital market.

We focus on problems in which project values are initially unknown but can be learned over time. Before deciding which project to finance, a company carries out due diligence on each project. If due diligence were unnecessary or infeasible, optimal project selection would be trivial because the company would immediately choose the project with the highest expected value, without any deliberation.

The situation in which Universal Music Group found itself in 2011 fits our model's environment exactly. Universal was considering two alternative projects: the purchase of EMI Music and the purchase of Warner Music Group.[1] Purchasing both was infeasible, if only because of antitrust concerns. Assessing the profitability of each purchase required costly due diligence by the teams of lawyers, consultants and accountants, to evaluate music catalogs, potential synergies, and antitrust risks.

EMI is based in London; Warner Music is based in New York. Universal (headquartered in Santa Monica) also happens to have two divisions: one in London and one in New York. Universal could charge the London division with carrying out due diligence

[1] See http://www.completemusicupdate.com/article/bmg-and-universal-may-co-bid-for-warner-and-emi/ and http://www.completemusicupdate.com/timeline-emisale/.

pertaining to the purchase of EMI and could charge the New York division with carrying out due diligence pertaining to the purchase of Warner Music. Our first question asks how Universal should orchestrate its divisions' due diligence to maximize its expected cash flow.

The example of Universal carries one further complication. The London division favors buying EMI because this purchase would increase the influence of the London division. The New York division similarly favors buying Warner Music. If Universal's headquarters cannot monitor each division's due diligence, then each division is likely to be strategic when deciding whether to follow the headquarters' recommendations regarding due diligence (moral hazard) and when deciding whether to report truthfully the outcomes of its due diligence (adverse selection). Hence, our second question asks how the optimal policy can be implemented in the presence of both moral hazard and adverse selection.

To address the two design questions raised above, we study Universal's problem in an auction-like environment. HQ (the headquarters) allocates an item (the requisite funds to pursue an acquisition) to one of two divisions, denoted by D_1 and D_2. The value of each division's project (the profitability of the acquisition) is either 0 or 1 and is distributed independently across the two divisions. Initially, each division has a belief about its project's value and revises this belief as it learns (carries out due diligence).

Time is continuous, and the time horizon is infinite. At each instant, each division can learn at a cost. A division's learning affects the arrival intensity of "good news," which reveals the project's value to be 1. The alternative, "no news," means that the project's value can be either 0 or 1 and causes the division to revise its value estimate downward.

HQ maximizes the expected cash flow, defined as the expected value of the winning project net of both divisions' expected cumulative costs of learning. Assuming HQ can directly control each division's learning, observe learning outcomes, and select the winning project, HQ's problem is a stochastic-control and optimal-stopping problem. This

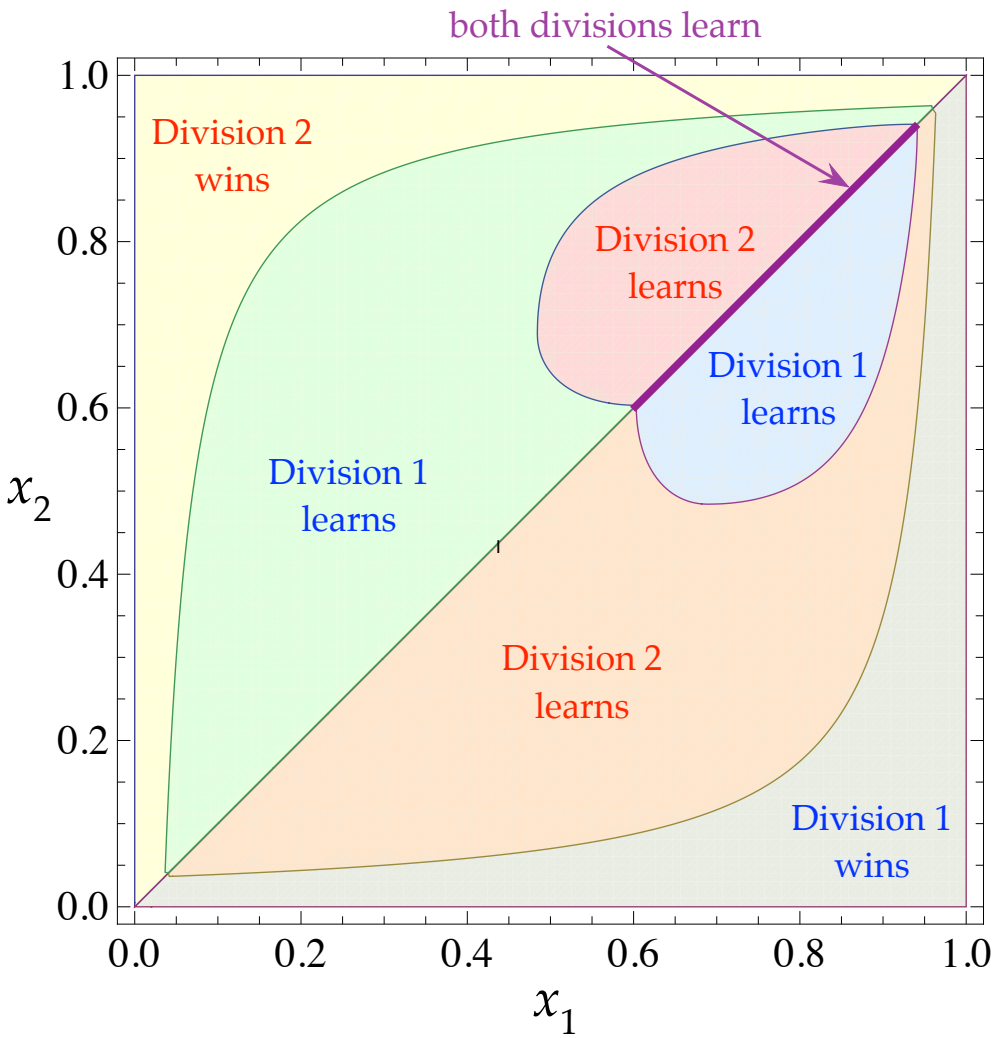


Figure 1: An optimal policy's prescription for each pair (x_1, x_2) of the project's expected values.

problem's solution—an optimal policy—is this paper's first contribution. The second contribution is the optimal policy's implementation in the presence of moral hazard and adverse selection.

Figure 1 summarizes an optimal policy when the cost of learning is sufficiently small.[2] This optimal policy is stationary and prescribes, for every pair (x_1, x_2) of the two projects' expected values, whether either division should win immediately, and if not, which division should learn. Normalizing $x_2 \geq x_1$, four prescriptions are possible:

[2]For larger costs, an optimal policy is a special case, and is characterized later in the paper.

1. Division 2 wins immediately

   D_2 wins immediately whenever x_1 and x_2 are either both close to 0 or both close to 1. In this case, there is little uncertainty about each project's value, and so learning is not worth the cost. D_2 also wins immediately whenever x_2 is substantially larger than x_1. In this case, there is little uncertainty about the fact that the value of project 2 exceeds the value of project 1, so learning is unlikely to affect the decision regarding which project to select and hence is suboptimal.

2. Division 2 learns

   D_2 learns when x_1 and x_2 are close to each other (so that which project is more valuable is highly uncertain), and when both x_1 and x_2 are far away from 0 and 1 (so that each project's value is highly uncertain). In this case, the need for information is so great that it is worthwhile to ask D_2 to learn first, and to plan on also asking D_1 to learn later if D_2 does not observe good news. Asking D_2 to learn without ever planning to ask D_1 to learn is suboptimal, as is explained below.

   D_2's learning is more informative than D_1's, which suggests that asking D_2 to learn may be optimal. Indeed, because $x_2 > x_1$, D_2 stands a higher chance of observing good news than D_1 does. At the same time, a spell of no news leads to a rapid downward revision of D_2's belief because a likely event—the arrival of good news—has failed to occur. Either way, as D_2 learns, its belief changes fast.

   Asking only D_2 to learn, without planning to ask D_1 to learn later on, amounts to committing to choose D_2's project. Such a commitment is suboptimal because D_2's belief is a martingale, and consequently, D_2's learning entails costs, but does not affect the (ex-ante) expected value of D_2's project. By contrast, if D_2's learning, with positive probability, is followed by D_1's learning, then the selected project's identity is contingent on the outcomes of learning, thereby justifying learning.

3. Both divisions learn

Both divisions learn simultaneously if (i) $x_1 = x_2$ (i.e., which project is more valuable is unclear), (ii) x_1 and x_2 are sufficiently large (i.e., learning by either division is rather informative), and (iii) x_1 and x_2 are bounded away from 0 and 1 (i.e., each project's value is highly uncertain).

4. Division 1 learns

   D_1 learns whenever the values of x_1 and x_2 are complementary to those described in scenarios 1–3. In this case, by having D_1 learn, HQ bets on having D_1 observe the good news. HQ is "insured" by D_2, which does not learn, and whose project can be selected if D_1 observes no news.

We show that the described optimal policy is implementable in the presence of moral hazard and adverse selection. That is, the prescriptions of the optimal policy coincide with equilibrium outcomes of a carefully constructed dynamic game, called optimal auction game. In this game, each division learns privately and hence must be motivated to conform with the optimal policy. The optimal auction begins with indicative bidding: each division publicly announces a real number, which does not directly affect the division's payment or the chances of getting its project selected. HQ uses the announced indicative bids to compute the firm-bidding deadline, which HQ then publicly announces. Until the deadline, each division can learn and, irrespective of whether it is learning, may be asked by HQ to pay a fee for the right to remain in competition. At any time, either division may request early firm bidding. Firm bidding takes the form of the second-price auction, which determines the divisions' payments and the winning project.

The fees and the firm-auction rules are chosen to ensure that each division pays the "externality" that its participation imposes on the other division, by analogy to a Vickrey-Clarke-Groves mechanism. Each division thus becomes a "residual claimant" to the cash flow. As a result, at equilibrium, each division's indicative bid equals its prior belief about its project's value. Furthermore, each division learns as prescribed by the optimal policy and requests firm bidding if and only if it observes good news. Each division's firm

bid is its final belief about its project's value. Thus, the highest expected-value project is selected.

Our paper contributes to two literatures: corporate finance literature on internal capital markets and economic theory literature on irreversible project selection in the presence of uncertainty, including the literature on auctions with information acquisition. The assumptions underlying our model of the internal capital market are motivated by the vision described by Stein (1997).[3] In particular, because internal capital is scarce (e.g., because of informational frictions associated with raising outside capital), not all profitable projects can be financed and so HQ must ration. At the same time, even unprofitable projects may end up being financed (e.g., because of HQ's empire-building tendencies), and so HQ invests all available internal capital. Accordingly, we assume that HQ selects exactly one project.

Existing literature on internal capital markets is predominantly positive. Among the positive models, in addition to Stein (1997), are Harris and Raviv (1996), Rajan et al. (2000), Scharfstein and Stein (2000), de Motta (2003), and Inderst and Laux (2005). The only normative dynamic model of an internal capital market that we are aware of is that of Malenko (2012). Whereas our focus is on learning about, and selection from, two given projects, Malenko (2012) studies selection from dynamically arriving projects and does not model learning.

The economic theory literature on irreversible project selection can be interpreted to model internal capital market. The real-option approach, exemplified by the work of Dixit and Pindyck (1994), assumes that the values of projects evolve exogenously. We extend their approach to situations in which these values evolve endogenously, as a result of learning. Learning is the focus of multi-armed bandit problems (Bolton and Harris, 1999; Keller et al., 2005; Klein and Rady, 2010). Bandit problems model reversible project selection because a selected arm can be unselected as new information arrives. In the

[3]For a textbook introduction to internal capital markets, see Tirole (2006, Section 10.5) and Gertner and Scharfstein (2012).

bandit problems in which no new information can ever justify unselecting an arm (and so investment irreversibility is irrelevant), there is no way to incorporate the cost of learning in a manner consistent with our model.[4]

Persico (2003), Compte and Jehiel (2007), Crémer et al. (2009), Shi (2012), and Krähmer and Strausz (2011) analyze information acquisition in auctions, but do not investigate situations in which learning can occur over an infinite horizon, information is cumulable over time, and no structure is imposed on the timing of learning. The richness of the set of admissible learning policies distinguishes our model and leads to novel insights. For example, the insight that sometimes it is optimal for the higher-expected-value division to learn would be lost in a model with a one-off chance to learn. The richness of the policy set comes at a cost; we work with a rather special, good-news, learning technology.

Athey and Segal (2007) and Bergemann and Välimäki (2010) are our inspirations for the auction implementation of the optimal policy.

The rest of the paper is structured as follows. Section 2 describes the environment. Section 3 conjectures and informally justifies an optimal policy. Section 4 verifies the conjecture by appealing to the viscosity-solution techniques. Section 5 describes the optimal policy's implementation in a dynamic auction. Section 6 concludes.

2 Model

Time is continuous and is indexed by $t \geq 0$. The time horizon is infinite.

Valuations

HQ holds an indivisible item, which it values at zero. HQ allocates this item to one of two divisions, indexed by $i \in \mathcal{N} \equiv \{1, 2\}$ and denoted by D_i. D_i's valuation $v_i \in \{0, 1\}$ is

[4]The challenge within the bandit framework is to make the cost of learning vanish once an irreversible investment has been made.

a random variable with $\Pr\{v_i = 1\} = X_i(0)$, for some prior belief $X_i(0) \in [0,1]$, $i \in \mathcal{N}$. Valuations v_1 and v_2 are statistically independent.

Learning

At any time t, each D_i can acquire information about v_i, or **learn**. Let a_i be the indicator function with values in $\{0,1\}$ such that $a_i(t) = 1$ indicates that D_i learns at time t. The cumulative cost of learning incurred by D_i from time 0 to time t is $c \int_0^t a_i(s)\,\mathrm{d}s$, for some cost parameter $c > 0$.

D_i's learning process $\{a_i(t) \mid t \geq 0\}$, denoted by a_i, controls the arrival-intensity process $\{\lambda a_i(t) v_i \mid t \geq 0\}$ of a Poisson process $\left\{N_i^{a_i}(t) \mid t \geq 0\right\}$ that has $N_i^{a_i}(0) = 0$, where $\lambda > 0$ is interpreted as the precision of learning. The event when $N_i^{a_i}(t)$ is incremented is called **good news** (about v_i). The event when $N_i^{a_i}(t)$ is not incremented is called **no news**. Because the event $N_i^{a_i}(t) > 0$ can occur only if $v_i = 1$, the good news reveals $v_i = 1$. Processes $N_1^{a_1}$ and $N_2^{a_2}$ are assumed to be independent.

Define $X_i^{a_i}(t)$, D_i's time-t **type**, or **belief**, to be the expectation of v_i conditional on the information revealed up to time t and on some learning process a_i:

$$X_i^{a_i}(t) \equiv \mathbb{E}\left[v_i \mid \left\{N_i^{a_i}(s) \mid 0 \leq s \leq t\right\}\right] = \mathbb{E}\left[v_i \mid N_i^{a_i}(t)\right].$$

The last equality in the display above obtains because $N_i^{a_i}(t)$ summarizes all the history that is relevant for learning about v_i. For any learning-process profile $a \equiv (a_1, a_2)$, the tuple $X^a(t) \equiv \left(X_1^{a_1}(t), X_2^{a_2}(t)\right)$ is a time-t **type profile**. By construction (by the Law of Iterated Expectations), the process X^a is a martingale.

The Evolution of Types

For any t and $t' > t$, D_i's type $X_i^{a_i}(t')$ is derived from $X_i^{a_i}(t)$ by application of the Bayes rule. According to the Bayes rule, $N_i^{a_i}(t') > 0$ implies $X_i^{a_i}(t') = 1$, whereas $N_i^{a_i}(t') = 0$

implies

$$\frac{X_i^{a_i}(t')}{1-X_i^{a_i}(t')} = \frac{X_i^{a_i}(t)}{1-X_i^{a_i}(t)} e^{-\lambda \int_t^{t'} a_i(s)\mathrm{d}s}. \tag{1}$$

For future reference, let us describe the stochastic type process $X_i^{a_i}$ in its differential form. Because $X_i^{a_i}$ is not conditional on v_i, it is convenient to define the good-news Poisson process $\tilde{N}_i^{a_i}$ whose arrival intensity is also unconditional, and equals $\lambda a_i(t) X_i^{a_i}(t)$. The stochastic differential equation for $X_i^{a_i}$ becomes

$$\mathrm{d}X_i^{a_i}(t) \;=\; \left(1-X_i^{a_i}(t)\right)\mathrm{d}\tilde{N}_i^{a_i}(t) - \lambda a_i(t) X_i^{a_i}(t)\left(1-X_i^{a_i}(t)\right)\mathrm{d}t \tag{2}$$

subject to $X_i^{a_i}(0) = X_i(0)$. The differential representation in (2) is a difference of two terms. The first term is the upward jump in the belief caused by the arrival of good news. The second, negative, term is the downward revision of the belief caused by the failure of the good news to arrive during learning; the magnitude of this revision follows from the Bayes formula in (1).

An Optimal Policy

The environment is stationary, and so no generality is lost by focusing on stationary policies. A (stationary) **policy** is a tuple (α, τ), where the learning policy $\alpha \equiv (\alpha_1, \alpha_2)$ maps a type profile $x \equiv (x_1, x_2)$ into learning decisions $(\alpha_1(x), \alpha_2(x))$ in $\{0,1\}^2 \setminus (0,0)$, and where τ is the stopping time that designates when the item is allocated to the highest-type division.[5] A policy (α, τ) induces the type process denoted by $\{X^{\alpha,\tau}(t) \mid t \geq 0\}$.

A policy (α, τ) is **admissible** if the learning process $\{\alpha(X(t)) \mid t \geq 0\}$, induced by the learning policy α, is predictable and integrable,[6] and if, for every $i \in \mathcal{N}$ and every

[5]No generality is lost by requiring that at least one division learn at any time until the item has been allocated; $(\alpha_1(x), \alpha_2(x)) \neq (0,0)$.

[6]A continuous-time stochastic process is **predictable** if it is measurable with respect to the σ-algebra generated by all left-continuous adapted processes. In the current setting, predictability means that the induced learning process is adapted with left-continuous paths; that is, at every t, $\alpha(X(t)) = \alpha(X(t^-))$, where $X(t^-) \equiv \lim_{s\to t, s<t} X(s)$. In other words, when armed with an admissible policy (α, τ) and approaching time t, HQ can foresee which division, if any, will learn at time t, according to α. HQ cannot

$X_i(0) \in [0,1]$, the stochastic differential equation (2) has a unique strong solution.[7]

A policy (α, τ) and an initial type profile x induce the expected **cash flow**

$$J(x, \alpha, \tau) \equiv \mathbb{E}\left[\max_{i \in \mathcal{N}} \left\{X_i^{\alpha,\tau}(\tau)\right\} - c \sum_{i \in \mathcal{N}} \int_0^{\tau} \alpha_i \left(X^{\alpha,\tau}(s)\right) \mathrm{d}s \mid X^{\alpha,\tau}(0) = x\right]. \tag{3}$$

For every initial type profile x, the **value function** ϕ is defined by

$$\phi(x) \equiv \sup_{\alpha,\tau} J(x, \alpha, \tau), \tag{4}$$

where the maximization is over all admissible policies. An **optimal policy** (α^*, τ^*) is defined to satisfy $\phi(x) = J(x, \alpha^*, \tau^*)$ for all x.

3 A Conjectured Optimal Policy and Value Function

An optimal policy and the induced value function are conjectured by appealing to HJBQVI (Hamilton-Jacobi-Bellman Quasi-Variational Inequality), a continuous-time analogue of the Bellman equation. The conjecture is subsequently verified using the viscosity approach.

Define the effective cost $\hat{c} \equiv c/\lambda$, the cost of learning per unit of precision. The optimal policy will be shown to depend on c and λ only through $\hat{c}$. Inspired by the divisions' symmetry (except for their prior beliefs), a symmetric optimal policy and a symmetric value function are conjectured. Hence, whenever the description of a policy or a value function is restricted to the case in which $x_2 \geq x_1$, the symmetric case $x_2 < x_1$ follows immediately.

foresee, however, the jump in X that may occur at t (X is not left continuous) or whether the item will be allocated at time t (even though adapted, τ need not be predictable).

[7] Protter (1990, Chapter V) discusses the conditions for existence and uniqueness of strong solutions to stochastic differential equations.

3.1 HJBQVI

Define an open convex set $\Omega \equiv (0,1)^2$, whose closure is $\bar{\Omega}$. On the boundary $\partial\Omega$ of Ω, immediate allocation is trivially optimal; the identity of the division with the highest valuation is known; $x \in \partial\Omega$ implies $\phi(x) = \max\{x_1, x_2\}$. It remains to find ϕ on Ω.

Any candidate value function, denoted by u, must satisfy HJBQVI, which can be derived as the limit of the Bellman equation for the corresponding discrete-time model as the length of each time period goes to zero. This derivation is standard and gives

$$\min_{i \in \mathcal{N}} \left\{ \hat{c} + x_i (1 - x_i) \frac{\partial u(x)}{\partial x_i} - x_i (1 - u(x)), u(x) - x_i \right\} = 0, \qquad x \in \Omega \tag{5}$$

and the **boundary condition**

$$u(x) = \max_{i \in \mathcal{N}} \{x_i\}, \qquad x \in \partial\Omega. \tag{6}$$

The quasi-variational-inequality (QVI) component of (5) is $u(x) \geq \max\{x_1, x_2\}$. QVI requires that the value provided by a candidate value function u be at least as high as the value that can be obtained from immediately allocating the item.

The Hamilton-Jacobi-Bellman (HJB) component of (5) is

$$\min_{i \in \mathcal{N}} \left\{ \hat{c} + x_i (1 - x_i) \frac{\partial u(x)}{\partial x_i} - x_i (1 - u(x)) \right\} = 0, \qquad x \in \Omega.$$

HJB requires that, whenever u suggests the suboptimality of immediate allocation (i.e., $u(x) > \max\{x_1, x_2\}$), some division, say, D_i, learns, and u satisfies the no-arbitrage condition

$$c + \underbrace{\lambda x_i (1 - x_i) \frac{\partial u(x)}{\partial x_i}}_{\text{payoff revision due to no news}} = \underbrace{\lambda x_i (1 - u(x))}_{\text{payoff revision due to good news}}, \qquad x \in \Omega. \tag{7}$$

The left-hand side of (7) is the flow cost c of learning plus $\lambda x_i (1 - x_i)$, the rate at which D_i's type declines if no news arrives, multiplied by $\partial u(x) / \partial x_i$, the change in the contin-

uation value in response to that decline. The right-hand side of (7) is the expected arrival intensity λx_i of good news times the discrete change $1 - u(x)$ in the continuation value in response to the arrival of good news. Thus, condition (7) is essentially definitional; it requires any candidate value function to promise zero expected change in the payoff from following a candidate optimal policy. In other words, any expected change in the future payoff must be incorporated into the current expected payoff.

To admit the possibility that a value function is nondifferentiable on a null set (which will be the case in our problem), we define a solution concept that ignores the points of nondifferentiability.

Definition 1. A function $u^* : \Omega \to \mathbb{R}$ is a **generalized solution** of, or **g-solves**, HJBQVI if u^* satisfies HJBQVI almost everywhere in Ω.

Our conjecture for ϕ on Ω is denoted by F. It is constructed to g-solve HJBQVI and satisfy the boundary condition (6). In principle, multiple functions may g-solve HJBQVI subject to the boundary condition. The constructed conjecture is recommended by the economic intuition for the underlying policy and, eventually, the verification argument. The qualitative features of F vary with the magnitude of $\hat{c}$, and so three cases are distinguished.

3.2 Learning Is Prohibitively Costly: $\hat{c} \geq c_2$

The Conjecture

It is natural to conjecture that if the cost of learning exceeds some threshold, then, at any type profile x, it will be suboptimal to ask any division to learn. Instead, the item will be optimally allocated immediately to the highest-type division, as is illustrated in Figure 2. The associated conjectured value function is

$$F(x) = \max\{x_1, x_2\}.$$

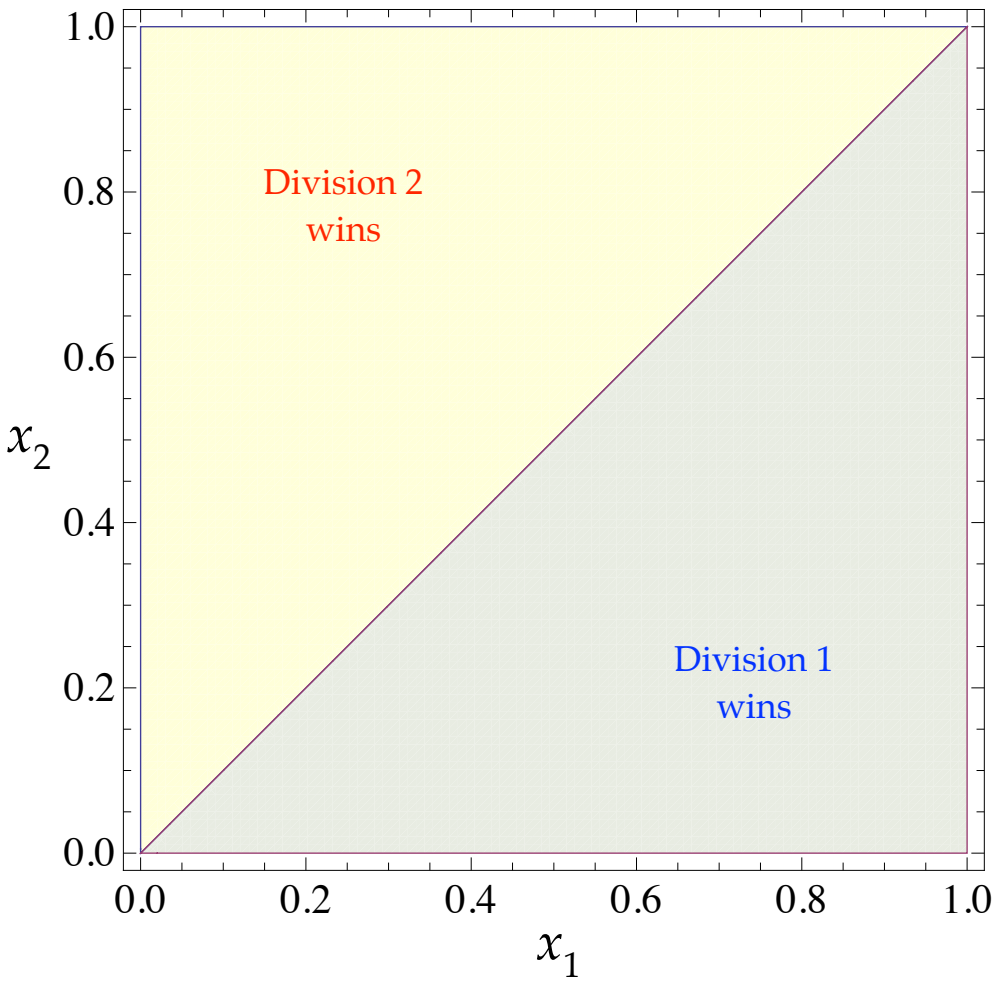


Figure 2: An optimal policy's prescription for each type profile when the learning cost is prohibitive; $\hat{c} \geq c_2$.

Define the relevant effective-cost threshold to be $c_2 = \frac{1}{2}$, or equivalently,

$$c_2 \equiv \inf \left\{ c' \geq 0 \mid \inf_{0 \leq x_1 \leq x_2 \leq 1} \left\{ c' - x_1 \left(1 - x_2\right) \right\} \geq 0 \right\}. \tag{8}$$

Threshold c_2 in (8) is motivated by the infinitesimal look-ahead rule (Ross, 1970, Section 9.6), which is a continuous-time counterpart of the one-step look-ahead rule in discrete-time stopping problems. The threshold c_2 in (8) is the smallest (effective) cost c' such that, at any type profile x with $x_2 \geq x_1$ (which is a normalization), HQ prefers allocating the item immediately to the highest-type division (D_2) to having the lowest-type division (D_1) learn for $\delta \to 0$ units of time, at cost $\lambda c' \delta$, and only then allocating the item to whomever by then has been revealed as the highest-type division. The revealed highest-type division is D_1 if and only if it observes the good news, which occurs with probability $x_1 \left(1 - e^{-\lambda\delta}\right)$. With the complementary probability, D_2 of type x_2 remains the highest-type division. The inequality associated with HQ's stated preference is

$$x_2 \geq -\lambda c' \delta + x_1 \left(1 - e^{-\lambda\delta}\right) + \left(1 - x_1 \left(1 - e^{-\lambda\delta}\right)\right) x_2, \tag{9}$$

which leads to the inequality in (8) when $\delta \to 0$.

The Conjecture G-Solves HJBQVI

Lemma 1. *Suppose that learning is prohibitively costly, or $\hat{c} \geq c_2$. The conjectured value function $F(x) = \max\{x_1, x_2\}$ g-solves HJBQVI subject to the boundary condition.*

Proof. Substituting F into HJBQVI (5) and recalling the convention $x_2 \geq x_1$ yields

$$\min\left\{\hat{c} - x_1 \left(x_1 - x_2\right), 0\right\} = 0, \qquad x \in \Omega,$$

which is implied by $\hat{c} \geq c_2$ and the definition of c_2 in (8).

The set $\{x \in \Omega \mid x_1 = x_2\}$, on which F is nondifferentiable, has measure zero. Thus,

by Definition 1, F g-solves HJBQVI.

It is immediate that F satisfies the boundary condition (6). □

3.3 Learning Is Moderately Costly: $c_1 \leq \hat{c} < c_2$

The Conjecture

If the (effective) cost $\hat{c}$ of learning is below the threshold c_2, defined in (8), asking some division to learn is optimal, which follows from the infinitesimal look-ahead perturbation formerly ruled out by (9). Further, it is natural to conjecture that any type profile x at which learning occurs has x_1 and x_2 close to each other, so that it is quite uncertain which division has a higher value, and has both x_1 and x_2 sufficiently far away from 0 and 1, so that there is considerable uncertainty about each division's value.

Which division learns depends on x and $\hat{c}$. For now, we focus on the case in which $c_1 \leq \hat{c} < c_2$, where $c_1 \approx 0.047$. Then, whenever learning occurs, the lower-type division learns. The threshold c_1 is the unique solution of[8]

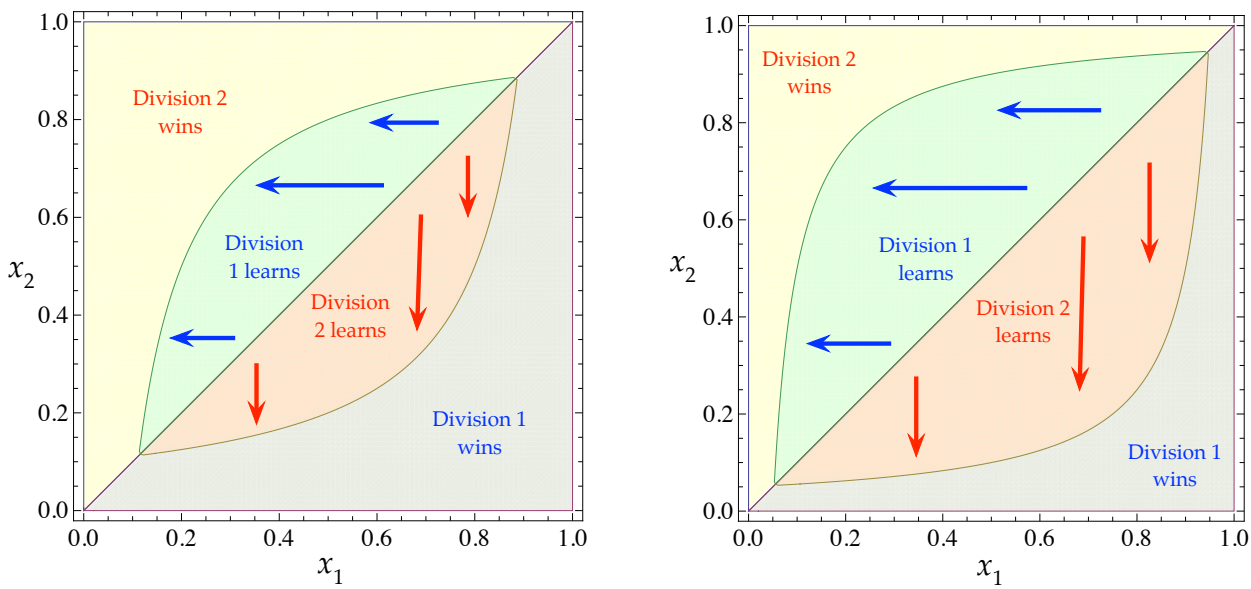
$$\log \frac{\left(1-\sqrt{c_1}\right)^2}{c_1} = \frac{2}{1-\sqrt{c_1}}. \tag{10}$$

The conjectured optimal policy is illustrated in Figure 3. To demarcate the set of type profiles at which D_1 learns—the lens-shaped region in the figure—normalize $x_2 \geq x_1$. Define a type threshold for D_1 as a function of D_2's type:

$$b\left(x_2\right) \equiv \frac{\hat{c}}{1-x_2}. \tag{11}$$

In Figure 3, D_1 learns if $x_1 > b\left(x_2\right)$; otherwise, D_2 wins immediately. When D_1 learns, it wins if and only if it observes the good news. Because b is increasing in $\hat{c}$, the lower the cost of learning, the larger the set of type profiles at which some division learns.

[8]Equation (10) admits no interpretation by inspection and emerges from a condition that ensures that asking the higher-type division to learn is suboptimal.



(a) $\hat{c} = 0.1$.

(b) $\hat{c} = 0.05$.

Figure 3: The optimal policy's prescription for each type profile when the learning cost is moderate; $c_1 \leq \hat{c} < c_2$. Within the lens-shaped region, one of the divisions learns. The arrows indicate the direction in which the type profile is revised if the division that learns observes no news. Outside the lens-shaped region, no division learns and the highest-type division wins.

In order to see the intuition for why the lower-type division, say D_1, is the one that learns, note that asking D_1 to learn amounts to betting on D_1's observation of the good news. This bet's upside is to see D_1 observe the good news and then let it win. The bet's downside is bounded by D_2's type. If D_1 observes no news for a sufficiently long period of time, HQ lets D_2 win. Because D_2 does not learn, its type does not deteriorate and so ensures against the downward revisions of D_1's type.

The rest of the subsection is concerned with demonstrating, in Lemma 3, that the conjectured value function F associated with the described conjectured policy g-solves HJBQVI. In the course of this demonstration, the origins of the boundary b in (11) and of the threshold c_1 in (17) are explained.

The Conjecture G-Solves HJBQVI

A critical element in the proof of the conjectured policy's optimality is an optimal stopping problem referred to as the **V-auxiliary problem**, defined in (11). This problem is solved by b of (11). In this problem, consistent with the conjectured optimal policy, D_1, the lower-type division, learns up to some optimally chosen deadline $t \geq 0$. At t or as soon as D_1 observes the good news, learning stops, and the division with the highest revised type wins. The value of the V-auxiliary problem is denoted by V and will be a component of the conjectured value function F. Letting $\bigstar$ abbreviate the event "D_1

observes the good news," the function V is defined

$$\begin{aligned} V(x) &\equiv \sup_{t\geq 0}\Bigg\{ \underbrace{x_1\left(1-e^{-\lambda t}\right)}_{\text{payoff from D}_1\text{'s winning}} + \underbrace{x_2\left(1-x_1\left(1-e^{-\lambda t}\right)\right)}_{\text{payoff from D}_2\text{'s winning}} \\ &\quad - \underbrace{x_1 c\int_0^t s\lambda e^{-\lambda s}\mathrm{d}s}_{\text{expected cost of learning if }\bigstar} - \underbrace{ct\left(1-x_1\left(1-e^{-\lambda t}\right)\right)}_{\text{expected cost of learning if not }\bigstar}\Bigg\} \\ &= \sup_{t\geq 0}\left\{1-(1-x_2)\left(1-x_1\left(1-e^{-\lambda t}\right)\right)-c\left((1-x_1)t+x_1\frac{1-e^{-\lambda t}}{\lambda}\right)\right\}, \end{aligned} \tag{12}$$

where the last equality is obtained by computing the integral and simplifying.

It is convenient to change variables in (12). Instead of maximizing over the deadline t, maximize over D_1's corresponding threshold type, denoted by z. The relationship between the deadline and the threshold type is found by the Bayes rule, as in (1):

$$\frac{z}{1-z}=\frac{x_1}{1-x_1}e^{-\lambda t}.$$

Substituting z from the display above into the expression for V in (12) yields

$$V(x)\equiv\sup_{z\in[x_1,0)}\left\{1-\frac{(1-x_1)(1-x_2)}{1-z}-\hat{c}(1-x_1)\left[\varphi(x_1)-\varphi(z)\right]\right\}, \tag{13}$$

where

$$\varphi(s)\equiv\frac{s}{1-s}+\ln\frac{s}{1-s},\qquad s\in(0,1).$$

In (13), the term $1-(1-x_1)(1-x_2)/(1-z)$ is the expected payoff from eventually allocating the item, either to D_1 (as soon as D_1 observes the good news) or to D_2 (if D_1 never observes the good news). The term $\hat{c}(1-x_1)\left[\varphi(x_1)-\varphi(z)\right]$ is the expected cost of learning.

Lemma 2 characterizes a solution of the V-auxiliary problem in (13) in terms of the

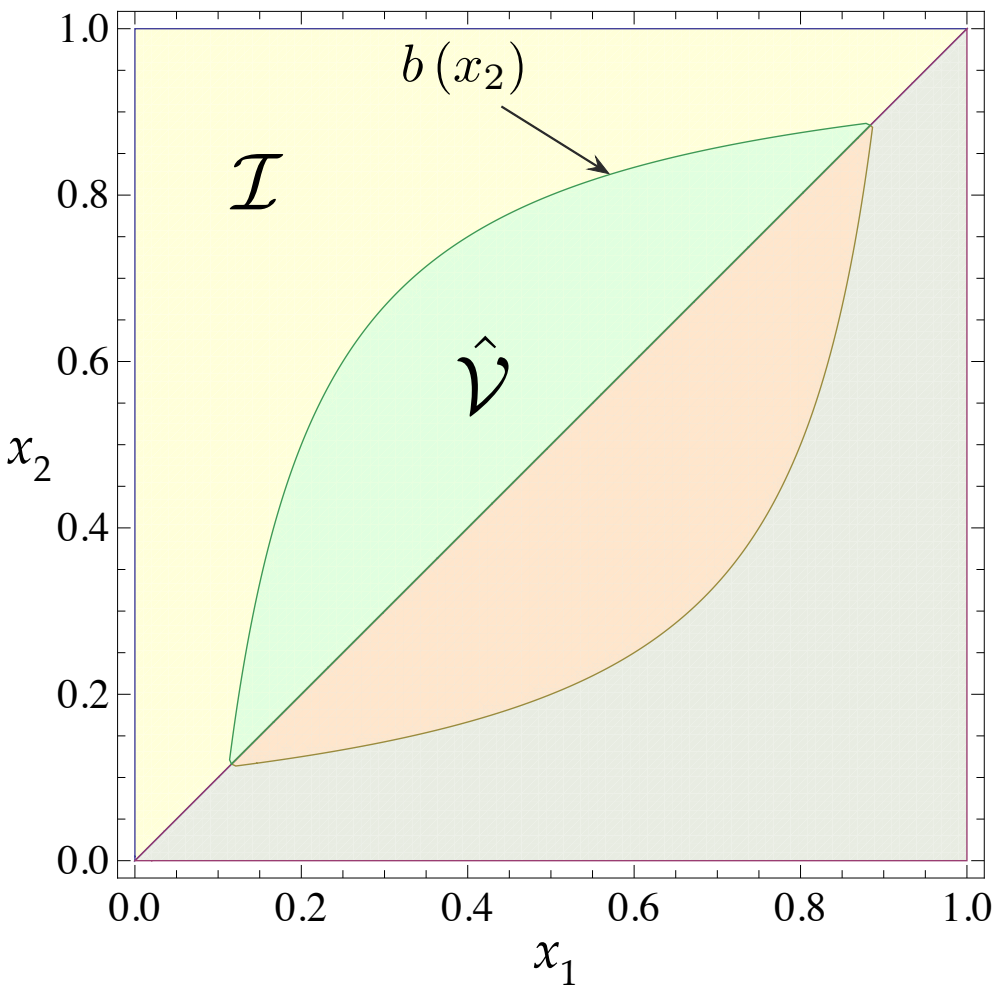


Figure 4: Each of the depicted sets $\mathcal{I}$ and $\hat{\mathcal{V}}$ depends on $\hat{c}$ and collects the type profiles at which the conjectured optimal policy makes identical prescriptions. The sets $\mathcal{I}$ and $\hat{\mathcal{V}}$ are separated by the function b.

type set

$$\hat{\mathcal{V}} \equiv \left\{ x \in [0,1]^2 \mid b(x_2) < x_1 \leq x_2 \right\}, \tag{14}$$

and the complementary set

$$\mathcal{I} \equiv \left\{ x \in [0,1]^2 \mid x_1 \leq x_2 \right\} \setminus \hat{\mathcal{V}}. \tag{15}$$

These sets are depicted in Figure 4.

Lemma 2. *Normalize $x_2 \geq x_1$. The value of the V-auxiliary problem in (13) satisfies $V(x) = x_2$*

if $x \in \mathcal{I}$, and

$$V(x) = 1 - \frac{(1-x_1)(1-x_2)}{1-b(x_2)} - \hat{c}(1-x_1)[\varphi(x_1) - \varphi(b(x_2))] \tag{16}$$

if $x \in \hat{\mathcal{V}}$. The policy that achieves V has D_2 win immediately if $x \in \mathcal{I}$ and has D_1 learn if $x \in \hat{\mathcal{V}}$.

Proof. The threshold $b(x_2)$, defined in (11), is derived using the infinitesimal look-ahead rule for optimal stopping problems (Ross, 1970, Section 9.6). According to this rule, at any type profile x, it is optimal for D_1 to learn if and only if learning for amount $\delta \to 0$ of time and then allocating the item to the division with the highest revised type delivers a weakly higher cash flow than immediately allocating the item to the higher-type division.

Assuming (and verifying shortly) that the infinitesimal look-ahead rule applies, D_1 learns if $x \in \hat{\mathcal{V}}$, and D_2 wins immediately if $x \in \mathcal{I}$. The threshold function b emerges from the indifference condition. When $x_1 = b(x_2)$, HQ is indifferent between (i) allocating the item to D_2 and (ii) having D_1 learn for $\delta \to 0$ units of time and then allocating the item to the division with the highest revised type. This indifference condition is

$$x_2 = \lim_{\delta \to 0} \left(-c\delta + b(x_2)\left(1 - e^{-\lambda\delta}\right) + \left[1 - b(x_2)\left(1 - e^{-\lambda\delta}\right)\right] x_2 \right).$$

Computing the limit and simplifying gives b as defined in (11). The value function in (16) is then derived from (13) assuming that the infinitesimal look-ahead rule applies as described.

By Theorem 9.3 of Ross (1970), to ascertain that the infinitesimal look-ahead rule indeed applies, one must verify that $\mathcal{I}$ is closed in the sense that $x \in \mathcal{I}$ implies that, for any $x_1' < x_1$, $(x_1', x_2) \in \mathcal{I}$. The set $\mathcal{I}$ is indeed closed, by inspection of (14) and (15). Hence, the infinitesimal look-ahead rule applies, and the desired result follows. □

The threshold c_1 in (17) emerges from the condition that, for any $x \in \hat{\mathcal{V}}$, rules out the profitability of an infinitesimal deviation whereby D_2 learns for an infinitesimal amount

of time before the conjectured optimal policy is resumed. This deviation can be verified to be unprofitable if and only if $\hat{c} \geq c_1$, where

$$c_1 \equiv \min\left\{c' \geq 0 \mid \min_{0 \leq x_1 \leq x_2 \leq 1} \Phi\left(x, c'\right) \geq 0\right\} \tag{17}$$

for

$$\Phi\left(x, \hat{c}\right) \equiv 1 - x_2\left(1 - x_1\right)\left(\varphi\left(x_1\right) - \varphi\left(b\left(x_2\right)\right)\right). \tag{18}$$

The right-hand side of (18) depends on $\hat{c}$ implicitly, through $b\left(\theta_2\right) \equiv \hat{c}/\left(1 - \theta_2\right)$. Here, Φ is proportional to HQ's payoff from having D_1 learn as prescribed by the conjectured policy less the payoff from asking D_2 to learn for $\delta \to 0$ and only then asking D_1 to learn as prescribed by the conjectured policy. This payoff is nonnegative for all $x \in \hat{\mathcal{V}}$ as long as $\hat{c} \geq c_1$. Lemma B.1 of Appendix B proves that c_1 defined in (17) can be equivalently and more simply characterized as the unique solution of (10).

Lemma 3. *Suppose that learning is moderately costly, or $c_1 \leq \hat{c} < c_2$. The conjectured value function $F\left(x\right) = \mathbf{1}_{\left\{x \in \hat{\mathcal{V}}\right\}} V\left(x\right) + \mathbf{1}_{\{x \in \mathcal{I}\}} x_2$ g-solves HJBQVI and satisfies the boundary condition.*

Proof. HJBQVI is verified by considering two cases: $x \in \mathcal{I}$ and $x \in \hat{\mathcal{V}}$.

1. Suppose $x \in \mathcal{I}$, or equivalently, $x_1 \leq b\left(x_2\right)$, so that $F\left(x\right) = x_2$. As in the proof of Lemma 1, the implied HJBQVI equation is

$$\min_{i \in \mathcal{N}} \left\{\hat{c} - x_1\left(1 - x_2\right), 0\right\} = 0, \qquad x \in \Omega$$

   which holds true because $x_1 \leq b\left(x_2\right) = \hat{c}/\left(1 - x_2\right)$, by $x \in \mathcal{I}$.

2. Suppose that $x \in \hat{\mathcal{V}}$, or equivalently, $x_1 > b\left(x_2\right)$, so that $F\left(x\right) = V\left(x\right)$. HJBQVI is

verified component by component of the min function:

$$\min \left\{ \begin{array}{c} \hat{c} + x_1 (1 - x_1) \frac{\partial V(x)}{\partial x_1} - x_1 (1 - V(x)), \\ \hat{c} + x_2 (1 - x_2) \frac{\partial V(x)}{\partial x_2} - x_2 (1 - V(x)), V(x) - x_2 \end{array} \right\} = 0.$$

(a) To show

$$\hat{c} + x_1 (1 - x_1) \frac{\partial V(x)}{\partial x_1} - x_1 (1 - V(x)) = 0, \tag{19}$$

differentiate V in (16) to obtain

$$\frac{\partial V(x)}{\partial x_1} = \frac{1 - x_2}{1 - b(x_2)} + \hat{c} (\varphi(x_1) - \varphi(b(x_2))) - \hat{c} (1 - x_1) \varphi'(x_1). \tag{20}$$

Substitution of the above display and of V into (19) verifies the equality.

(b) The inequality $V(x) \geq x_2$ follows by "revealed preference," that is, by construction of the V-auxiliary problem.

(c) To show

$$\hat{c} + x_2 (1 - x_2) \frac{\partial V(x)}{\partial x_2} - x_2 (1 - V(x)) \geq 0, \tag{21}$$

apply the Envelope Theorem (Milgrom and Segal, 2002) to the definition of V in (16) to obtain

$$\frac{\partial V(x)}{\partial x_2} = \frac{1 - x_1}{1 - b(x_2)}.$$

Substituting the above display and (16) into (21), dividing by $\hat{c}$, and substituting the definition of b yields an equivalent desired inequality:

$$\Phi(x, \hat{c}) \geq 0, \tag{22}$$

where Φ is defined in (18). Because Φ is increasing in $\hat{c}$, the inequality in (22) holds for all admissible x if and only if $\hat{c}$ is sufficiently large. Then, because $\hat{c} \geq c_1$, and because c_1 satisfies (17), the provisional inequality sign in (22) is

definitive.

F is non-differentiable only on a null subset of the 45-degree line in the type space. Hence, by Definition 1, F g-solves HJBQVI, as desired. It is immediate that F satisfies the boundary condition. □

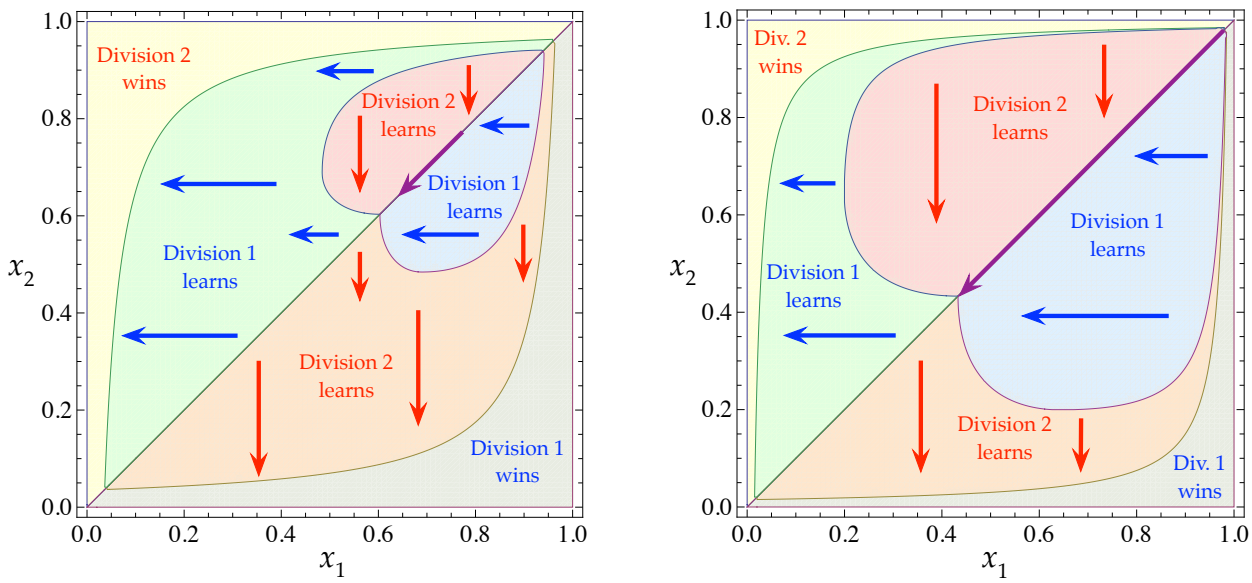
3.4 Learning Is Cheap: $\hat{c} < c_1$

The Conjecture

If the (effective) cost $\hat{c}$ of learning falls below threshold c_1, defined in (17), asking the higher-type division to learn is no longer suboptimal. Intuitively, learning drives the divisions' types apart, thereby helping HQ identify the highest-valuation division. Learning is faster for the higher-type division, say D_2, which is more likely than D_1 to observe the good news. If D_2 observes the good news, its type jumps to 1. If D_2 observes no news, its type is revised substantially because an event deemed likely (the arrival of the good news) has failed to occur. That is, in some informal sense, D_2's learning is more informative than D_1's learning.

Because each division's revised type is a martingale, the policy of asking only D_2 to learn and then selecting D_2's project regardless of the learning outcome does not affect the project's expected value but entails learning costs, and so cannot be optimal. Therefore, D_2's learning can be optimal only if it is sufficiently long to potentially flip the ranking of the divisions' revised types. Lengthy learning can be optimal only if learning is sufficiently cheap—which leads to the condition $\hat{c} < c_1$—and only if the need for information is sufficiently large. The need for information is large when the types x_1 and x_2 are close to each other (so that which project is more valuable is highly uncertain), when both x_1 and x_2 are far away from 0 and 1 (so that each project's value is highly uncertain), and when x_2 is rather large (so that D_2's learning is rather informative).

The conjectured optimal policy is illustrated in Figure 5. For the type profiles in the



(a) $\hat{c} = 0.035$.

(b) $\hat{c} = 0.015$.

Figure 5: The optimal policy's prescription for each type profile when the learning cost is small; $\hat{c} < c_1$. Within the lens-shaped region (which encompasses the heart-shaped region), at least one of the divisions learns. The arrows indicate the direction in which the type profile is revised if the division (or divisions) that learns observes no news. Outside the lens-shaped region, no division learns and the highest-type division wins.

heart-shaped region, the highest-type division learns. Elsewhere in the lens-shaped region, the lowest-type division learns. On the diagonal traversing the heart-shaped region, both divisions learn. If neither division learns, the highest-type division wins. The boundary of the lens-shaped region is demarcated by function b defined in (11), as before. The demarcation of the heart-shaped region is subtler and will be derived after additional notation has been developed.

The Conjecture G-Solves HJBQVI

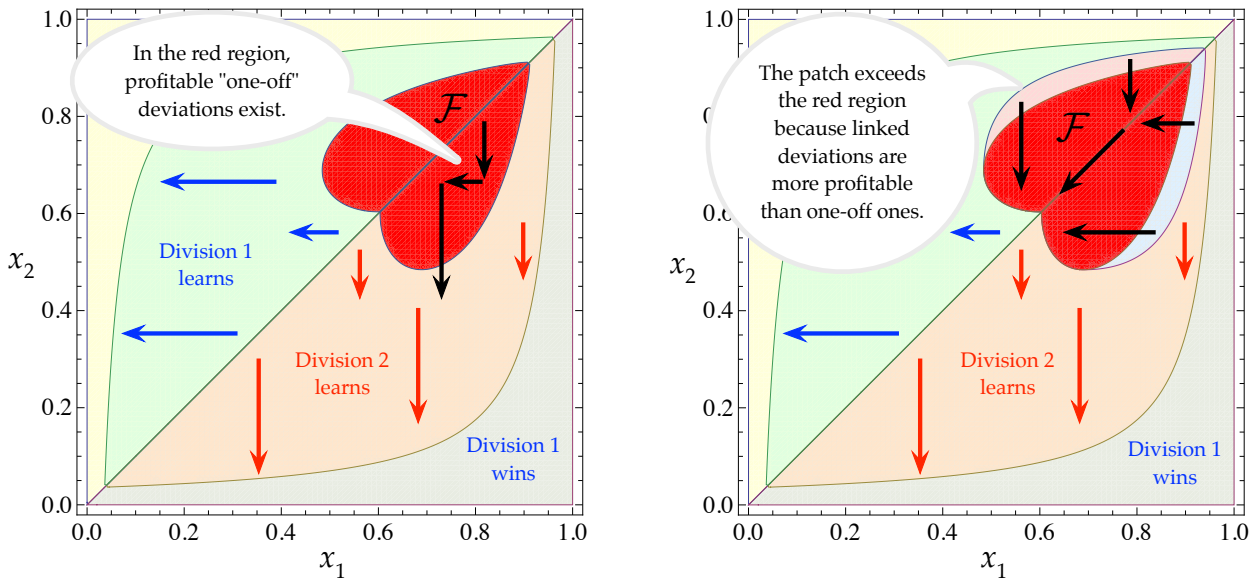
The economic intuition for D_2's occasional learning, supplied above, will be complemented by a technical intuition. The notation required to formulate the conjectured value function will be introduced along the way. The technical intuition can be informally viewed as the first step of a value-function-iteration-like procedure. Take the value function in Lemma 3 to be the initial guess for the value function for the case in which $\hat{c} < c_1$.

For every $x \in \hat{\mathcal{V}}$, Lemma 3 prescribes that D_1 learn. This prescription induces the value function that violates HJBQVI whenever $\hat{c} < c_1$. In particular, inequality (22) in Step 2c of the lemma's proof fails on the **failure set**

$$\mathcal{F} \equiv \{x \mid \Phi(x, \hat{c}) < 0\}. \tag{23}$$

Figure 6a illustrates the failure set. On and only on $\mathcal{F}$, infinitesimal learning by D_2 followed by D_1's learning is a profitable deviation from the policy in which only D_1 learns on $\hat{\mathcal{V}}$. The described identification of the failure set would be the first step in the "value-function iteration process."

Lemma 6 "patches" the failure set $\mathcal{F}$ by making D_2 learn on a certain set, called **patch**, which is a strict superset of $\mathcal{F}$, as is illustrated in Figure 6b. The infinitesimal episodes of learning by D_2 on $\mathcal{F}$ can be linked together to form longer spells of learning on $\mathcal{F}$. Each



(a) On $\mathcal{F}$, instead of asking the lower-type division to learn, HQ can achieve a higher payoff by momentarily asking the higher-type division to learn and then reverting to asking the lower-type division to learn.

(b) The heart-shaped patch (on which the highest-type division learns) exceeds $\mathcal{F}$.

Figure 6: The fallacy of the conjecture that the lower-type division learns within the lens-shaped region is exposed by the failure of HJBQVI on the (smaller) heart-shaped failure region, denoted by $\mathcal{F}$.

of these longer spells is more profitable for HQ than an infinitesimal episode (followed by D_1's learning). As a result, even though on the boundary of $\mathcal{F}$, HQ is indifferent between infinitesimal learning by D_2 and learning by D_1 all the way, HQ strictly prefers a longer episode of learning by D_2 to learning by D_1. The profitability of having D_2 learn thus extends beyond the failure set and to the patch.

What remains of $\hat{\mathcal{V}}$ once it has been patched is the new region on which we conjecture that D_1 learns. This set is denoted by

$$\mathcal{V} \equiv \hat{\mathcal{V}} \setminus (\mathcal{A} \cup \mathcal{B} \cup \mathcal{C}), \tag{24}$$

where the sets $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$ are depicted in Figure 7 and will be derived formally shortly.

To characterize the sets $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$, we consider three more auxiliary stopping problems: A-auxiliary, B-auxiliary, and C-auxiliary. In each of these three problems, the goal is to find the optimal amount of time that a division (or divisions) must learn before the learning pattern is changed.

To formulate the C-auxiliary problem, used to characterize the set $\mathcal{C}$, define the bounds

$$\underline{x} \equiv \frac{1 - \sqrt{1 - 4\hat{c}}}{2} \qquad \text{and} \qquad \bar{x} \equiv \frac{1 + \sqrt{1 - 4\hat{c}}}{2}. \tag{25}$$

The **C-auxiliary problem** is defined on the set

$$\hat{\mathcal{C}} \equiv \left\{ x \in [0,1]^2 \mid \underline{x} \leq x_1 = x_2 \leq \bar{x} \right\},$$

which is a subset of $\hat{\mathcal{V}}$. In this problem, both divisions learn until either observes the good news or until both revised types reach some optimally chosen threshold. At that threshold, the type profile enters $\mathcal{V}$, and the strategy described in Lemma 3 is followed; that is, D_1 learns. Formally, the C-auxiliary problem's value function, denoted by C, satisfies, for

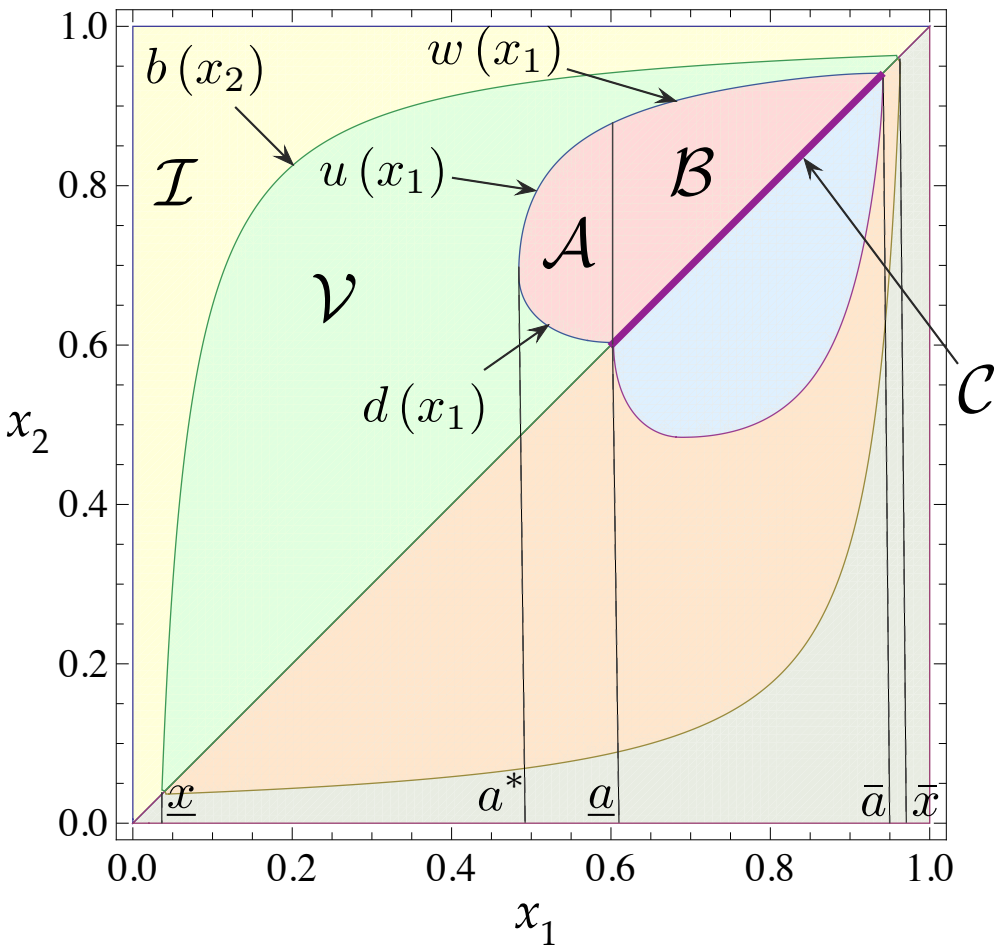


Figure 7: Each of the depicted sets $\mathcal{I}$, $\mathcal{V}$, $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$ depends on $\hat{c}$ and collects the type profiles at which the optimal policy makes identical prescriptions. The set $\mathcal{A}$ is bounded by function u above and function d below. The set $\mathcal{B}$ is bounded by function w above and the 45-degree line below. The set $\mathcal{V}$ is bounded by function b above and to the left.

any $x \in \hat{\mathcal{C}}$,

$$C(x_1) \equiv \max_{z \in [\underline{\theta}, x_1]} \left\{ \underbrace{1 - (1 - V(z,z)) \left(\frac{1 - x_1}{1 - z} \right)^2 - 2\hat{c} (1 - x_1)^2 [\sigma(x_1) - \sigma(z)]}_{\equiv M^C(z)} \right\}, \tag{26}$$

where

$$\sigma(s) \equiv \frac{2s}{1-s} + \frac{1}{2} \left(\frac{s}{1-s} \right)^2 + \ln \frac{s}{1-s}, \qquad s \in (0,1),$$

and where the maximand in (26) is denoted by $M^C(z)$, for future reference. The maximand is constructed analogously to the maximand in (13) and, to conserve space, will not be written out in detail.

Lemma 4 characterizes a solution of the C-auxiliary problem in terms of the type subset

$$\mathcal{C} \equiv \left\{ x \in \hat{\mathcal{C}} \mid x_1 \in (\underline{a}, \bar{a}) \right\}, \tag{27}$$

where

$$\underline{a} \equiv \inf \left\{ z \in (\underline{x}, \bar{x}) \mid \Phi(z, z, \hat{c}) \leq 0 \right\} \tag{28}$$

$$\bar{a} \equiv \max \left\{ z \in (\underline{a}, \bar{x}) \mid \frac{1 - V(z,z)}{(1-z)^2} - 2\hat{c}\sigma(z) = \frac{1 - V(\underline{a}, \underline{a})}{(1-\underline{a})^2} - 2\hat{c}\sigma(\underline{a}) \right\}. \tag{29}$$

The origins of $\underline{a}$ and $\bar{a}$ defined above is revealed in the lemma's proof.

Lemma 4. *Normalize $x_2 \geq x_1$. The value of the C-auxiliary problem in (26) satisfies $C(x_1) = V(x_1, x_1)$* if $x \in \hat{\mathcal{C}} \backslash \mathcal{C}$, and satisfies

$$C(x_1) = 1 - (1 - V(\underline{a}, \underline{a})) \left(\frac{1 - x_1}{1 - \underline{a}} \right)^2 - 2\hat{c} (1 - x_1)^2 [\sigma(x_1) - \sigma(\underline{a})] \tag{30}$$

if $x \in \mathcal{C}$. The policy that achieves C has D_1 learn as in Lemma 3 if $x \in \hat{\mathcal{C}} \backslash \mathcal{C}$ and has both divisions learn if $x \in \mathcal{C}$.

Proof. Recall that the maximand in (26) is denoted by $M^C(z)$. Its dependence on x_1 is

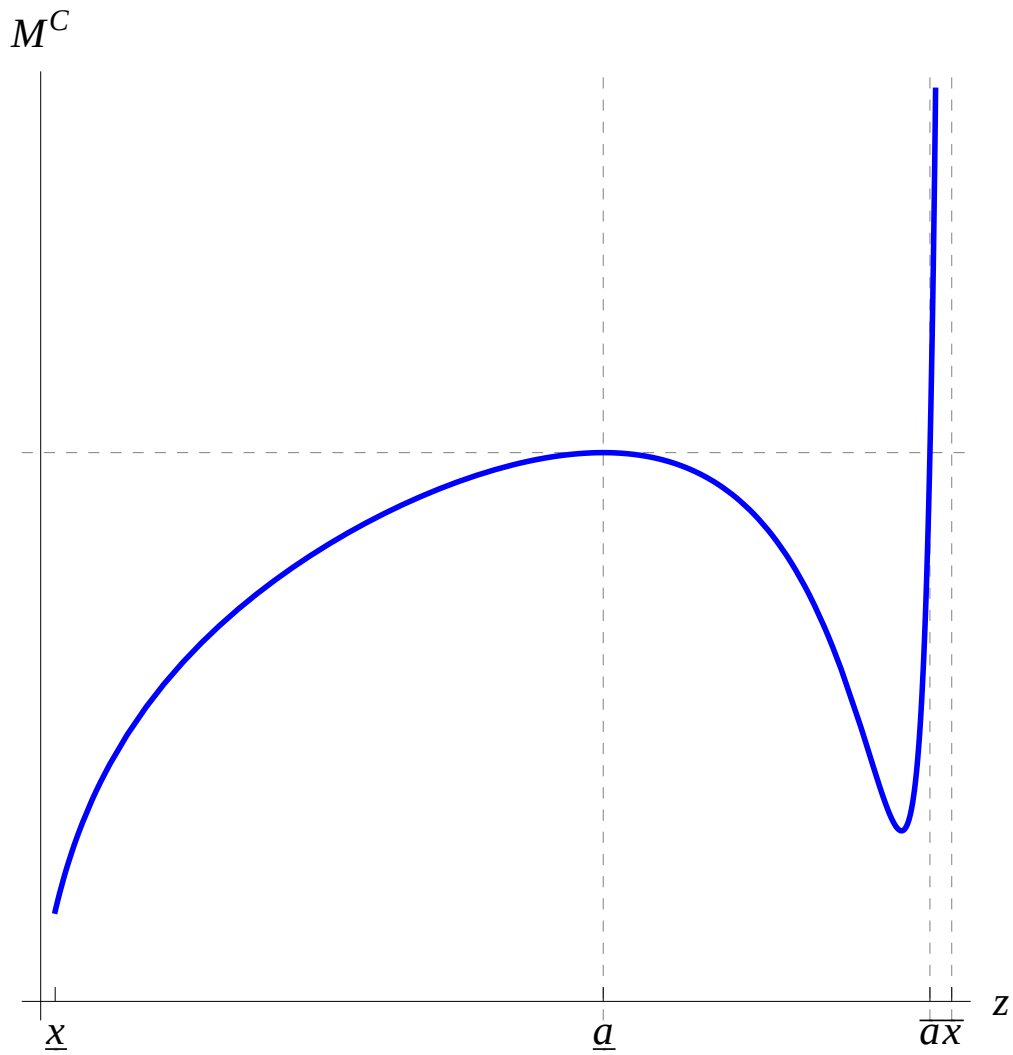


Figure 8: M^C, the maximand in the C-auxiliary problem. The sign of $\mathrm{d}M^C(z)/\mathrm{d}z$ coincides with the sign of $\Phi(z, z, \hat{c})$.

suppressed, motivated by the observation that the maximizer in (26) depends on x_1 only through the restriction $z \in [\underline{\theta}, x_1]$. To understand the solution of $\max_{z \in [\underline{\theta}, x_1]} M^C(z)$, it is instructive to study the shape of M^C on $[\underline{x}, \bar{x}]$. This shape is described by the sign of $\mathrm{d}M^C(z)/\mathrm{d}z$, which, by differentiation, can be verified to coincide with the sign of $\Phi(z, z, \hat{c})$. Figure 8 illustrates M^C.

Note that $\Phi(\underline{x}, \underline{x}, \hat{c}) = \Phi(\bar{x}, \bar{x}, \hat{c}) = 1$ (by direct substitution) and that $\Phi(z, z, \hat{c})$ is at first decreasing in z and then increasing in z (by Lemma B.2 in Appendix B) and is uniquely minimized at some $\underline{z}$, at which $\Phi(\underline{z}, \underline{z}, \hat{c}) < 0$. In order to see that $\Phi(\underline{z}, \underline{z}, \hat{c}) < 0$, note

that, by the definition of c_1 in (17) and by Lemma B.1, $\min_z \Phi(z, z, c_1) = 0$. Because Φ is strictly increasing in $\hat{c}$, $\hat{c} < c_1$ and $\min_z \Phi(z, z, c_1) = 0$ imply $\Phi(\underline{z}, \underline{z}, \hat{c}) < 0$.

To summarize, $\Phi(z, z, \hat{c})$—and hence also $\mathrm{d}M^C(z)/\mathrm{d}z$—first switches the sign from the positive to the negative and then from the negative to the positive. This positive-negative-positive sign pattern implies that M^C is wave-shaped, with local maxima at $\underline{a}$, defined in (28), and $\bar{x}$, and with local minima at $\underline{x}$ and somewhere between $\underline{a}$ and $\bar{x}$. Moreover, by Lemma B.4 of Appendix B, $\bar{x}$ is the unique global maximum, and so $M^C(\bar{x}) > M^C(\underline{a})$. Coupled with $M^C(\bar{x}) > M^C(\underline{a})$, the wave shape of M^C implies the existence of a unique $\bar{a} \in (\underline{a}, \bar{x})$ such that $M^C(\bar{a}) = M^C(\underline{a})$. Condition $M^C(\bar{a}) = M^C(\underline{a})$ can be simplified to express $\bar{a}$ equivalently (but still implicitly) by (29).

The derived properties of M^C have the following implications for the program in $C(x_1) \equiv \max_{z \in [\underline{x}, x_1]} M^C(z)$. When $x_1 \notin (\underline{a}, \bar{a})$, M^C is maximized on $[\underline{\theta}, x_1]$ at x_1, and so $C(x_1) = V(x_1, x_1)$; D_1 learns. When $x_1 \in (\underline{a}, \bar{a})$, M^C is uniquely maximized on $[\underline{\theta}, x_1]$ at $\underline{a}$, and so $C(x_1) > V(x_1, x_1)$; both divisions learn until either division observes the good news or until both types fall to $\underline{a}$, whereupon D_1 learns. □

To characterize the set $\mathcal{A}$ in Figure 7, we formulate the **A-auxiliary problem** on the set

$$\{x \in (a^*, \underline{a}] \times [0, 1] \mid x_2 \geq x_1\}.$$

In this problem, the choice is when to stop asking D_2 to learn and instead adopt the strategy described in Lemma 3, which begins with D_1 learning. The value of this stopping problem is, for any $x \in \hat{\mathcal{V}}$,

$$A(x) \equiv \max_{z \in [x_1, x_2]} \left\{ \underbrace{1 - (1 - V(x_1, z)) \frac{1 - x_2}{1 - z} - \hat{c}(1 - x_2)[\varphi(x_2) - \varphi(z)]}_{\equiv M^A(z; x_1)} \right\}, \tag{31}$$

where the maximand is denoted by $M^A(z)$ and is constructed using the arguments anal-

ogous to those used in the construction of V and C above.

Lemma 5 characterizes a solution to the A-auxiliary problem in terms of the type subset

$$\mathcal{A} \equiv \{x \in (a^*, \underline{a}] \times [0,1] \mid x_2 \in (d(x_1), u(x_1))\}, \tag{32}$$

where

$$a^* \equiv \inf\{x_1 \in [0,1] \mid \exists x_2 \in [x_1, 1] \text{ s.t. } \Phi(x, \hat{c}) = 0\}, \tag{33}$$

$$d(x_1) \equiv \inf\{z \in [x_1, 1] \mid \Phi(x_1, z, \hat{c}) \leq 0\}, \qquad x_1 \in (a^*, \bar{a}), \tag{34}$$

and, letting b^{-1} denote the inverse function of b,

$$u(x_1) \equiv \sup\left\{ z \in \left[x_1, b^{-1}(x_1)\right] \mid \begin{array}{c} \frac{1-V(x_1,d(x_1))}{1-d(x_1)} - \hat{c}\varphi(d(x_1)) \\ = \frac{1-V(x_1,z)}{1-z} - \hat{c}\varphi(z) \end{array} \right\}, \qquad x_1 \in (a^*, \bar{a}). \tag{35}$$

The origin of the objects entering the definition of $\mathcal{A}$ is revealed in the lemma's proof.

Lemma 5. *Normalize $x_2 \geq x_1$. The value of the A-auxiliary problem in (31) satisfies $A(x) = V(x)$ if $x \in (a^*, \underline{a}] \times [0,1] \setminus \mathcal{A}$, and satisfies*

$$A(x) = 1 - [1 - V(x_1, d(x_1))] \frac{1 - x_2}{1 - d(x_1)} - \hat{c}(1 - x_2)[\varphi(x_2) - \varphi(d(x_1))] \tag{36}$$

if $x \in \mathcal{A}$. The policy that achieves A has D_1 learn as in Lemma 3 if $x \in (a^*, \underline{a}] \times [0,1] \setminus \mathcal{A}$ and has D_2 learn if $x \in \mathcal{A}$.

Proof. Recall that the maximand in (31) is denoted by $M^A(z; x_1)$. By differentiating, the sign of $\mathrm{d}M^A(z; x_1)/\mathrm{d}z$ coincides with the sign of $\Phi(x_1, z, \hat{c})$. Hence, the argument analogous to the argument in the proof of Lemma 4 applies, except now, in $\Phi(x_1, z, \hat{c})$, only one argument, z, varies.

In particular, Lemma B.3 of Appendix B establishes that M^A is a wave-shaped function

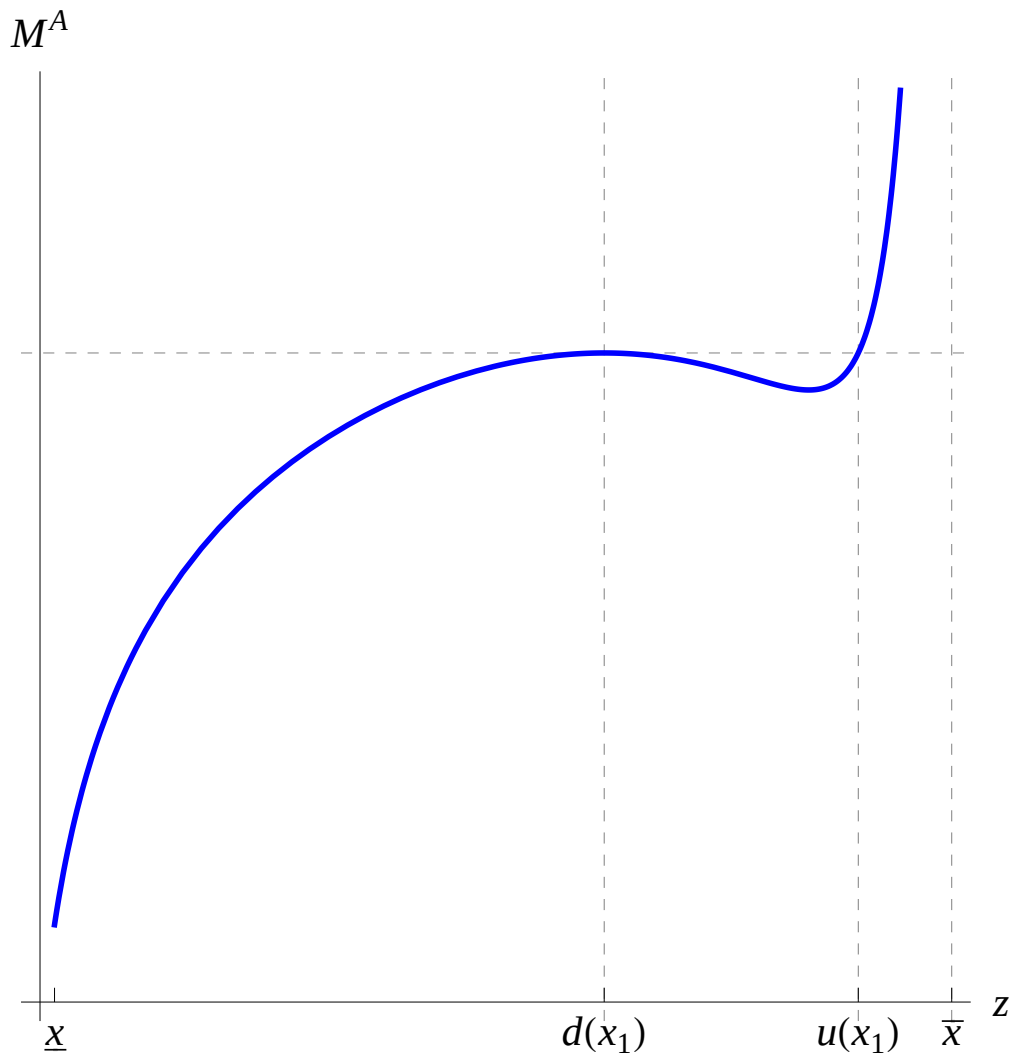


Figure 9: M^A, the maximand in the A-auxiliary problem. The sign of $\mathrm{d}M^A(z; x_1)/\mathrm{d}z$ coincides with the sign of $\Phi(x_1, z, \hat{c})$.

of z, as depicted in Figure 9. By inspection of M^A, for a given x_1, a local maximum of $M^A(z; x_1)$ is given by (34).

Coupled with $M^A\left(b^{-1}(x_1); x_1\right) > M^A(d(x_1); x_1)$ (by Lemma B.5 of Appendix B), the wave shape of M^A implies the existence of a unique $u(x_1) \in \left(m(x_1), b^{-1}(x_1)\right)$ such that $M^A(u(x_1); x_1) = M^A(d(x_1); x_1)$. Condition $M^A(u(x_1); x_1) = M^A(d(x_1); x_1)$ can be simplified to express $u(x_1)$ equivalently (but still implicitly) as in (35).[9]

The derived properties of M^A have the following implications for the program in $A(x) \equiv$

[9] As defined in (35), u satisfies $A(x_1, u(x_1)) = V(x_1, u(x_1))$.

$\max_{z\in[x_1,x_2]} M^A(z;x_1)$. When $x_2 \notin (d(x_1), u(x_1))$, $M^A(\cdot;x_1)$ is maximized on $[x_1,x_2]$ at x_2, and so $A(x) = V(x)$. When $x_2 \in (d(x_1), u(x_1))$, $M^A(\cdot;x_1)$ is uniquely maximized on $[x_1,x_2]$ at $d(x_1)$, and so $A(x) > V(x)$. □

To characterize $\mathcal{B}$, the remaining set in Figure 7, we formulate a degenerate auxiliary problem, the **B-auxiliary problem**. The problem is defined on the set

$$\mathcal{B} \equiv \{x \in (\underline{a}, \bar{a}) \times [0,1] \mid x_2 \in (x_1, w(x_1))\}, \tag{37}$$

where $\underline{a}$ is defined in (28), $\bar{a}$ is defined in (29), and

$$w(x_1) \equiv \inf\left\{z \in \left[x_1, b^{-1}(x_1)\right] \mid V(x_1,z) = B(x_1,z)\right\}, \qquad x_1 \in (\underline{a}, \bar{a}). \tag{38}$$

The problem consists in having D_2 learn until it observes the good news or until its revised type drops down to D_1's type x_1 (the "choice" of this threshold x_1 is the degenerate decision), whereupon both divisions learn as prescribed by the C-auxiliary problem. This policy's value is

$$B(x) \equiv 1 - (1 - C(x_1))\frac{1-x_2}{1-x_1} - \hat{c}(1-x_2)[\varphi(x_2) - \varphi(x_1)]. \tag{39}$$

The function B is constructed using the arguments analogous to those used in the construction of V, C, and A above.

We can now assemble the pieces.

Lemma 6. *Suppose that learning is cheap, or $\hat{c} < c_1$. The conjectured value function*

$$F(x) = \mathbf{1}_{\{x\in\mathcal{A}\}}A(x) + \mathbf{1}_{\{x\in\mathcal{B}\}}B(x) + \mathbf{1}_{\{x\in\mathcal{C}\}}C(x) + \mathbf{1}_{\{x\in\mathcal{V}\}}V(x) + \mathbf{1}_{\{x\in\mathcal{I}\}}x_2$$

g-solves HJBQVI and satisfies the boundary condition.

Proof. HJBQVI is verified by considering four cases: $x \in \mathcal{I}$, $x \in \mathcal{V}$, $x \in \mathcal{C}$, and $x \in \mathcal{A} \cup \mathcal{B}$.

1. Suppose that $x \in \mathcal{I}$, so that $F(x) = x_2$. HJBQVI holds by Step 1 in the proof of Lemma 3.

2. Suppose that $x \in \mathcal{V}$, so that $F(x) = V(x)$. HJBQVI follows from the argument in Step 2 in the proof of Lemma 3 with one modification. Part 2c of that step requires the inequality $\Phi(x, \hat{c}) \geq 0$. To show this inequality here, note that $\Phi(x, \hat{c}) < 0 \iff x \in \mathcal{F}$, by the definition of the failure set $\mathcal{F}$ in (23). Because $\mathcal{F} \subset \mathcal{A} \cup \mathcal{B} \cup \mathcal{C}$ (by Lemmas B.5 and B.6) and $\mathcal{V} \cap (\mathcal{A} \cup \mathcal{B} \cup \mathcal{C}) = \varnothing$ (by the definition of $\mathcal{V}$ in (24)), $x \in \mathcal{V}$ implies $x \notin \mathcal{F}$. As a result, $\Phi(x, \hat{c}) \geq 0$ for all $x \in \mathcal{V}$, as desired.

3. Suppose that $x \in \mathcal{C}$, so that $F(x) = C(x_2)$ and $\partial F(x) / \partial x_i = (1/2) \partial C(x_i) / \partial x_i$.[10] HJBQVI in (5) will be verified component by component of the min function:

$$\min \left\{ \hat{c} + \frac{x_2 (1 - x_2)}{2} \frac{\partial C(x_2)}{\partial x_2} - x_2 (1 - C(x_2)), C(x_2) - x_2 \right\} = 0.$$

   (a) The inequality $C(x_2) \geq x_2$ follows by "revealed preference." By construction of the C-auxiliary problem, $C(x_2, x_2) \geq V(x_2, x_2)$. By construction of the V-auxiliary problem, $V(x_2, x_2) \geq x_2$. Therefore, $C(x_2) \geq x_2$.

   (b) To show

$$\hat{c} + \frac{x_2 (1 - x_2)}{2} \frac{\partial C(x_2)}{\partial x_2} - x_2 (1 - C(x_2)) = 0, \tag{40}$$

   differentiate C in (30) to obtain

$$\frac{\partial C(x_2)}{\partial x_2} = \frac{2}{1 - x_2} \left(1 - C(x_2) - \frac{\hat{c}}{x_2} \right). \tag{41}$$

   Substitution of the above display and of C into (40) verifies the equality.

4. Suppose $x \in \mathcal{A} \cup \mathcal{B}$. If $x \in \mathcal{A}$, then $F(x) = A(x)$. If $x \in \mathcal{B}$, then $F(x) = B(x)$.

[10] Note the notational convention according to which $\partial F(x) / \partial x_i \neq \partial C(x_i) / \partial x_i$. The rationale is that $\partial F(x) / \partial x_i$ is the change in the conjectured value function in response to a marginal change in a single type, whereas $\partial C(x_i) / \partial x_i$ is the change in the conjectured value function in response to identical marginal changes in both types.

HJBQVI is verified component by component of the min function:

$$\min\left\{\begin{array}{c}\hat{c}+x_1\left(1-x_1\right)\frac{\partial F(x)}{\partial x_1}-x_1\left(1-F\left(x\right)\right),\\ \hat{c}+x_2\left(1-x_2\right)\frac{\partial F(x)}{\partial x_2}-x_2\left(1-F\left(x\right)\right),F\left(x\right)-x_2\end{array}\right\}=0.$$

(a) To verify

$$\hat{c}+x_2\left(1-x_2\right)\frac{\partial F\left(x\right)}{\partial x_2}-x_2\left(1-F\left(x\right)\right)=0, \tag{42}$$

replace F and $\partial F/\partial x_2$ in the display above by A and $\partial A/\partial x_2$ if $x\in\mathcal{A}$, and by B and $\partial B/\partial x_2$ if $x\in\mathcal{B}$. The function $\partial A/\partial x_2$ is obtained by differentiating (36):

$$\frac{\partial A\left(x\right)}{\partial x_2}=\frac{1-V\left(x_1,d\left(x_1\right)\right)}{1-d\left(x_1\right)}+\hat{c}\left(\varphi\left(x_2\right)-\varphi\left(d\left(x_1\right)\right)\right)-\frac{\hat{c}}{x_2\left(1-x_2\right)}.$$

The function $\partial B/\partial x_2$ is obtained by differentiating (39):

$$\frac{\partial B\left(x\right)}{\partial x_2}=\frac{1-C\left(x_1\right)}{1-x_1}+\hat{c}\left(\varphi\left(x_2\right)-\varphi\left(x_1\right)\right)-\frac{\hat{c}}{x_2\left(1-x_2\right)}.$$

(b) To show $F\left(x\right)\geq x_2$, it suffices to show $F\left(x\right)\geq V\left(x\right)$ and, in order to conclude that $V\left(x\right)\geq x_2$ by construction of the V-auxiliary problem, to show $x\in\mathcal{A}\cup\mathcal{B}\implies x\in\hat{\mathcal{V}}$. Hence, it will be shown that (i) $A\left(x\right)\geq V\left(x\right)$ and $B\left(x\right)\geq V\left(x\right)$, and (ii) $\mathcal{A}\subseteq\hat{\mathcal{V}}$ and $\mathcal{B}\subseteq\hat{\mathcal{V}}$.

i. When $x\in\mathcal{A}$, $A\left(x\right)\geq V\left(x\right)$ follows "by revealed preference," by construction of the A-auxiliary problem.

When $x\in\mathcal{B}$, showing $B\left(x\right)\geq V\left(x\right)$ is mildly more involved, but is also essentially a revealed-preference argument, applied twice. In particular, extend the A-auxiliary problem (originally defined on $\mathcal{A}$) onto the set $\mathcal{B}$. Using the same arguments as in Lemma 5, this problem's solution can be verified to imply that, whenever $x\in\mathcal{B}$, D_2 learns until either he observes the good news or until his belief drops down to x_1, at which point the

prescription of the V-auxiliary problem is followed, delivering the continuation value $V(x_1, x_1)$. So, by revealed preference, $A(x) \geq V(x)$.

On $\mathcal{B}$, the B-auxiliary problem has the same threshold as the extended A-auxiliary problem, but once this threshold has been reached, delivers a higher continuation value, $C(x_1)$, with $C(x_1) \geq V(x_1, x_1)$ by "revealed preference" in the C-auxiliary problem. So, $B(x) \geq A(x)$, which, combined with $A(x) \geq V(x)$, gives $B(x) \geq V(x)$, as desired.

ii. The inclusion $\mathcal{A} \subseteq \hat{\mathcal{V}}$ follows because Lemma B.5 shows that $x_1 \in (a^*, \underline{a})$ implies $u(x_1) < b^{-1}(x_1)$. That is, $u(x_1)$, the upper boundary of set $\mathcal{A}$, on which the A-auxiliary problem is defined, lies strictly below $b^{-1}(x_1)$, the upper boundary of set $\hat{\mathcal{V}}$.

The inclusion $\mathcal{B} \subseteq \hat{\mathcal{V}}$ follows because Lemma B.6 shows that $x_1 \in (\underline{a}, \bar{a})$ implies $w(x_1) < b^{-1}(x_1)$. That is, $w(x_1)$, the upper boundary of set $\mathcal{B}$, on which the B-auxiliary problem is defined, lies strictly below $b^{-1}(x_1)$, the upper boundary of set $\hat{\mathcal{V}}$.

(c) It will be shown that

$$\hat{c} + x_1(1 - x_1)\frac{\partial F(x)}{\partial x_1} - x_1(1 - F(x)) \geq 0. \tag{43}$$

If $x \in \mathcal{A}$, apply the Envelope Theorem to A in (36) to obtain

$$\frac{\partial A(x)}{\partial x_1} = \frac{\partial V(x_1, d(x_1))}{\partial x_1}\frac{1 - x_2}{1 - d(x_1)}.$$

Substituting the above display and the definition of A in (36) into the tentative

inequality (43) and dividing by $\hat{c}$ yields the equivalent tentative inequality

$$1 \geq x_1 (1 - x_2) (\varphi (x_2) - \varphi (d (x_1))) + \frac{x_1 (1 - x_2)}{(1 - d (x_1)) \hat{c}} \left[1 - V (x_1, d (x_1)) - (1 - x_1) \frac{\partial V (x_1, d (x_1))}{\partial x_1} \right],$$

which further simplifies by substituting V and $\partial V / \partial x_1$ using (19) and by dividing both sides by $1 - x_2$:

$$\frac{1}{1 - x_2} \geq x_1 (\varphi (x_2) - \varphi (d (x_1))) + \frac{1}{1 - d (x_1)}. \tag{44}$$

If $x \in \mathcal{B}$, differentiate B in (39) and use the expression for $\partial C (x_1) / \partial x_1$ to obtain

$$\frac{\partial B (x_1, x_2)}{\partial x_1} = \frac{1 - x_2}{(1 - x_1)^2} \left(1 - C (x_1) - \frac{\hat{c}}{x_1} \right).$$

Substituting the above display and the definition of B in (39) into the tentative inequality (43) leads, once again, to the tentative inequality (44).

If $x_2 = d (x_1)$, the inequality (44) holds trivially, as equality. To show that the inequality holds also for $x_2 > d (x_1)$, it suffices to show that its left-hand side increases in x_2 faster than its right-hand side. Indeed, the left-hand side's derivative, which is $1 / (1 - x_2)^2$, is larger than the right-hand side's derivative, which is

$$x_1 \varphi' (x_2) = \frac{x_1}{x_2 (1 - x_2)^2},$$

by $x_1 < x_2$. Thus, the tentative inequality (44) is definitive, as desired.

It is immediate that F satisfies the boundary condition.

□

3.5 Synthesis

To summarize, the **conjectured value function** is

$$F(x) = \begin{cases} x_2 & \text{if } c_2 \leq \hat{c} \\ \mathbf{1}_{\{x \in \mathcal{V}\}} V(x) + \mathbf{1}_{\{x \in \mathcal{I}\}} x_2 & \text{if } c_1 \leq \hat{c} < c_2 \\ \mathbf{1}_{\{x \in \mathcal{A}\}} A(x) + \mathbf{1}_{\{x \in \mathcal{B}\}} B(x) & \\ \quad + \mathbf{1}_{\{x \in \mathcal{C}\}} C(x) + \mathbf{1}_{\{x \in \mathcal{V}\}} V(x) + \mathbf{1}_{\{x \in \mathcal{I}\}} x_2 & \text{if } \hat{c} < c_1, \end{cases} \tag{45}$$

where A, B, C, and V are defined, respectively, in (36), (39), (30), and (16).

The underlying **conjectured optimal policy** can be read off Figure 2 when $c_2 \leq \hat{c}$, Figure 3 when $c_1 \leq \hat{c} < c_2$, and Figure 5 when $\hat{c} < c_1$.

4 Verification of the Conjectured Value Function

Theorem 1. *The conjectured value function F in (45) is the sought value function; that is, $\phi = F$. By implication, the conjectured optimal policy is indeed optimal.*

The theorem's proof builds on several intermediate results and is distilled toward the end of the section.

4.1 The Big Picture

Continuous-time modelling enables us to compute explicitly HQ's value function and derive the implied optimal policy. In addition, the statements of the results are not contaminated by provisos stemming from the indivisibility of information increments, which would be inherent in any discrete-time formulation. The cost of the continuous-time approach is the unfamiliar mathematics required to rigorously justify the results. In particular, the concept of viscosity solution emerges; it has no counterpart in discrete time. A

number of technical conditions are also invoked. Therefore, before delving into the formal analysis, we informally outline the main ideas and relate continuous-time optimization to its familiar discrete-time counterpart.

Consider the dynamic programming principle (DPP). DPP says that HQ's value today equals HQ's expected continuation value at an arbitrary future stopping time plus the flow payoffs enjoyed until that time. These intervening flow payoffs and the eventual continuation value depend on the intervening controls, which are chosen to maximize HQ's value today. This statement of DPP is the same in both continuous and discrete time and relies on the same backward-induction argument. The argument requires either a finite horizon or discounting. We impose neither. However, without loss of generality, we could have imposed a finite horizon. Indeed, if learning is protracted, eventually, beliefs become so precise that further learning cannot possibly justify its costs. So DPP holds in our setting.

One can show that DPP holds if and only if the DPP equation holds for the smallest possible positive stopping time. In a discrete model, this smallest time is one period, and the corresponding DPP equation is known as the Bellman equation. In a continuous model, the smallest time is operationalized by taking limits, thereby obtaining a differential counterpart of the Bellman equation, known as the HJBQVI equation. Taking limits is delicate, however, because limits may not exist, even under natural "technical" conditions. That is, the value function need not be differentiable; it may have "kinks." Kinks are not an issue in discrete time, because the Bellman equation is not a differential equation.

In continuous time, kinks cannot be neglected. A candidate value function that solves HJBQVI at all points of differentiability but with its kinks unrestricted may fail to be the sought value function. That is, extraneous generalized solutions of HJBQVI are possible. Enter viscosity solution, which uses HJBQVI to discipline the candidate value function not only at the points of differentiability, but also at the kinks.

In our maximization problem, viscosity solution restricts upward kinks, but not downward kinks.[11] It is somewhat of a folk wisdom that, in maximization problems, value functions abhor upward kinks. For instance, that, in a static problem, under natural conditions, the value function admits no upward kinks can be seen graphically by taking the upper envelope of the reward functions for various values of the control while varying the state. Clausen and Strub (2012) show how upward kinks can be ruled out in a dynamic problem (not related to ours).[12] HJBQVI's viscosity solution strengthens generalized solution by imposing the requisite optimality restrictions on the kinks. A characterization theorem obtains; the sought value function, and only this function, is the viscosity solution of HJBQVI.

The conjectured value function F in (45) has no upward kinks at all, as will be shown. (The downward kinks are where regions $\mathcal{A}$ and $\mathcal{B}$ border $\mathcal{V}$ in Figure 7 and on the 45-degree line except within the heart-shaped region in Figures 4 and 7.) So the viscosity conditions on the kinks hold automatically. All that remains to verify is that HJBQVI holds at all points of differentiability, which we have already done. Thus, the conjectured value function is the viscosity solution of HJBQVI and hence is the sought value function.

To summarize, the conditions that viscosity solution imposes on kinks are substantive, not technical. The economic content of these conditions is not peculiar to the continuous-time model. In the discrete model, the same content is already embedded into the Bellman equation, which is comprehensive because it is not restricted to the value function's points of differentiability. Just as, in discrete time, the value function is the unique solution of the Bellman equation, in continuous time, the value function is the unique viscosity solution of HJBQVI. This viscosity characterization of the value function is the essence of the verification procedure, justified and performed in the remainder of this section.

[11] The kinks interpretation is inherent in Theorem 5.6 of Bardi and Capuzzo-Dolcetta (1997, p. 80).
[12] Their Figure 1(b) is an illustration.

4.2 The Dynamic Programming Principle

Lemma 7. *For any initial type profile $x \in [0,1]^2$ and any finite stopping time τ, the value function ϕ, defined in (4), satisfies the recursive relationship that is the* ***dynamic programming principle (DPP)****:*

$$\begin{aligned}\phi(x) &= \sup_{\alpha,\tau'} \ \mathbb{E}\left[\mathbf{1}_{\{\tau'<\tau\}} \max_{i\in\mathcal{N}} \left\{X_i^{\alpha,\tau'}(\tau')\right\} + \mathbf{1}_{\{\tau'\geq\tau\}}\phi\left(X^{\alpha,\tau'}(\tau)\right)\right. \\ &\qquad \left. -c\sum_{i\in\mathcal{N}}\int_0^{\tau'\wedge\tau} \alpha_i\left(X^{\alpha,\tau'}(s)\right)ds \mid X^{\alpha,\tau'}(0) = x\right]. \end{aligned} \tag{46}$$

The critical step in DPP's proof is the ability to treat HQ's problem as if it were finite-horizon.

Lemma 8. *For any initial type profile x, there exists a finite horizon T_x such that the value of HQ's problem in (4) is unaffected by restricting attention to the policies that allocate the item no later than at time T_x.*

Proof. If the initial type profile violates $x \ll (1,1)$, then it is apparent that allocating the item immediately to division i with $x_i = 1$ is uniquely optimal. In this case, $T_x = 0$. For the remainder of the proof, assume that $x \ll (1,1)$.

Denote the set of type profiles near the western and southern boundaries of the type set by

$$S \equiv \left\{x \in [0,1]^2 \mid \min\{x_1, x_2\} \leq \underline{x}\right\},$$

where $\underline{x}$ is defined in (25). For an arbitrary type profile $x \ll (1,1)$, define the finite bound

$$T_x \equiv \mathbf{1}_{\{x\notin S\}}\frac{1}{\lambda}\ln\left[\frac{x_1}{1-x_1}\frac{x_2}{1-x_2}\left(\frac{1-\underline{x}}{\underline{x}}\right)^2\right], \tag{47}$$

where $\mathbf{1}_{\{\cdot\}}$ is the indicator function. It will be shown that T_x is the finite horizon mentioned in the lemma's statement.

We shall refer to two technologies. The **regular learning technology** is as specified

in the model (Section 2). The **superior learning technology** is obtained by modifying the regular learning technology so that, for the same flow cost c, HQ learns about both projects simultaneously. Learning using the regular learning technology is abbreviated as **r-learning**. Learning using the superior learning technology is abbreviated as **s-learning**.

The remainder of the proof has two steps. Step 1 argues that $x \in S$ implies the optimality of immediate allocation. Step 2 shows that, for any initial type profile $x \ll (1,1)$, within time T_x, r-learning leads either to the revised type profile in S or to the arrival of good news.

Step 1

The **infinitesimal look-ahead rule** (**ILAR**) is defined to prescribe immediate allocation if and only if HQ's cash flow is higher after immediate allocation than after s-learning for $\delta \to 0$ units of time followed by the allocation to the division with the highest type. Formally, ILAR prescribes immediate allocation if and only if[13]

$$\max\{x_1, x_2\} \geq -\lambda\hat{c}\delta + \max\{x_1, x_2\} + (1 - \max\{x_1, x_2\}) \min\{x_1, x_2\} \left(1 - e^{-\lambda\delta}\right).$$

Taking the limit $\delta \to 0$ gives

$$\hat{c} \geq (1 - \max\{x_1, x_2\}) \min\{x_1, x_2\}. \tag{48}$$

ILAR prescribes s-learning if and only if the inequality in (48) fails.

At all $x \in S$, the ILAR inequality (48) holds by the construction of $\underline{x}$ in (25).

The set S is **closed** in the sense that $x \in S \implies x' \in S$ for any type profile x' that is an s-learning revision of x conditional on no news. Any such revision satisfies $x' \leq x$, which implies the desired inclusion.

[13]Under the normalization $x_2 \geq x_1$, the displayed inequality is identical to (9). If no good news arrives about the lowest-type division while s-learning, ILAR allocates the item to the other division, regardless of whether good news about that other division arrives. So s-learning is equivalent to r-learning about the lowest-type division, as far as ILAR is concerned.

Suppose that HQ can use only the superior learning technology. Because S is closed and because the ILAR inequality (48) holds on S, we conclude that, whenever $x \in S$, immediate allocation is in fact optimal (Ross, 1970, Section 9.6). Downgrading the learning technology only strengthens the case for immediate allocation. That is, if HQ can use only the regular learning technology, $x \in S$ implies that immediate allocation is optimal.[14] Thus, $T_x = 0$ in (47) is an appropriate finite horizon in the lemma's statement.

Step 2

It remains to confirm that T_x in (47) is an appropriate horizon for the lemma's statement in the case in which $x \notin S$ and $x \ll (1,1)$. So assume that $x \notin S$ and $x \ll (1,1)$.

The longest r-learning can take to drive the revised belief into the set S is the time necessary to drive x_1 (almost) down to $\underline{x}$ and then to drive x_2 down to $\underline{x}$. These times are implied by the Bayes rule in (1) and are, correspondingly,

$$\frac{1}{\lambda} \ln \left(\frac{x_1}{1 - x_1} \frac{1 - \underline{x}}{\underline{x}} \right) \qquad \text{and} \qquad \frac{1}{\lambda} \ln \left(\frac{x_2}{1 - x_2} \frac{1 - \underline{x}}{\underline{x}} \right).$$

These times add up to T_x defined in (47), thereby establishing the desired bound. □

We can now prove DPP.

Proof of Lemma 7 (Sketch). The lemma's conclusion follows from the DPP in Proposition 3.1 of Pham (1998). A handful of inconsequential differences between Pham's setup and ours are worth noting.

Pham's focus on a finite-horizon problem is not restrictive for us because, by Lemma 8, without loss of generality, attention can be restricted to policies that allocate the item in finite time.

Pham assumes that the intensity of the Poisson jump process is independent of the

[14] In the proof, the superior technology helps sidestep the question as to which division HQ should ask to learn. A rigorous treatment of this question under the regular technology would require the use of the dynamic programming techniques, which rely on the present lemma, thereby leading to a circular argument.

state. By contrast, in our problem, the jump intensity, $\lambda a X$, depends on the state process, X. Nevertheless, stochastic integration with respect to this more general jump process is well-defined—which is all that matters for Pham's argument.[15]

Pham considers a discounted problem of which the non-discounted problem, such as ours, is a special case. □

4.3 Continuity of the Value Function

The viscosity characterization of the sought value function is simpler to state and easier to prove if it is known a priori that the value function is continuous.

Lemma 9. *The value function ϕ, defined in (4), is continuous on* $[0,1]^2$.

Proof. Let x and x' be two arbitrary type profiles.

1. Suppose that $x' \geq x$. It will be shown that $\phi(x') \geq \phi(x)$.

   Define $\Delta \equiv x' - x$. Fix an arbitrarily small $\varepsilon > 0$.

   Let $(\tilde{\alpha}, \tilde{\tau})$ denote an ε-optimal policy, which implicitly depends on ε:[16]

   $$J(x, \tilde{\alpha}, \tilde{\tau}) > \phi(x) - \varepsilon, \qquad \text{for all } x \in \Omega.$$

   Define the translated policy $\left(\tilde{\alpha}_{|\Delta}, \tilde{\tau}_{\Delta}\right)$, which prescribes at any type profile $y \in [\Delta_1, 1] \times [\Delta_2, 1]$ what $(\tilde{\alpha}, \tilde{\tau})$ prescribes at the translated type profile $y - \Delta$.

   The inequality

   $$\phi(x') \geq J\left(x', \tilde{\alpha}_{|\Delta}, \tilde{\tau}_{\Delta}\right)$$

   follows from the definition of the value function ϕ.

   The inequality

   $$J\left(x', \tilde{\alpha}_{|\Delta}, \tilde{\tau}_{\Delta}\right) \geq J(x, \tilde{\alpha}, \tilde{\tau})$$

[15] Process X is a finite variation process. Hence, the stochastic integral with respect to X is well-defined, as a path-by-path Riemann-Stieltjes integral (see, e.g., Protter, 1990, Chapter I.6).

[16] We are careful not to predicate our results on the existence of an optimal policy. Hence, we introduce the ε-optimal policy. Existence is a result, not an assumption.

obtains because (i) policy $\left(\tilde{\alpha}_{|\Delta}, \tilde{\tau}_{\Delta}\right)$ initiated at x' is weakly more likely to lead to good news than policy $(\tilde{\alpha}, \tilde{\tau})$ initiated at x, by $x' \geq x$, and (ii) if neither policy leads to good news, then the eventual allocation gives a weakly higher payoff under policy $\left(\tilde{\alpha}_{|\Delta}, \tilde{\tau}_{\Delta}\right)$ initiated at x' than under policy $(\tilde{\alpha}, \tilde{\tau})$ initiated at x, again, by $x' \geq x$.

Chaining together the three inequalities displayed above gives $\phi(x') > \phi(x) - \varepsilon$. For the inequality to hold for all $\varepsilon > 0$, it must be that $\phi(x') \geq \phi(x)$, as desired.

2. Suppose $x' \geq x$ and $x \ll (1,1)$. It will be shown that

$$\phi(x') \leq \phi(x) \frac{1-x_1'}{1-x_1} \frac{1-x_2'}{1-x_2} + \left(1 - \frac{1-x_1'}{1-x_1} \frac{1-x_2'}{1-x_2}\right). \tag{49}$$

Start at the type profile x'. Assume that HQ is prepared to learn about v_i until the revised type drops from x_i' to x_i. The required duration of learning is implied by the Bayes rule in (1), and the probability that, during this time, no good news about v_i arrives is

$$\frac{1-x_i'}{1-x_i}.$$

One can now interpret (49). The right-hand side of inequality (49) would be HQ's value if (counterfactually) HQ were temporarily granted the possibility of learning about both v_1 and v_2 at no cost until either good news is observed or x' is revised down to x. The inequality sign follows from the fact that HQ's is a Bayesian decision problem, and so (free) information has a nonnegative value.

3. Fix a type profile x. It will be shown that if a sequence $(x'(n))$ converges to x and satisfies $x'(n) \geq x$ for each n, then $\phi(x'(n)) \to \phi(x)$.

If $x \ll (1,1)$ fails, then $\phi(x) = 1$, and so, for any $x' \geq x$, $\phi(x) = \phi(x')$, by 1.

Suppose $x \ll 1$ holds. Combining the inequalities in 1 and 2 gives, for any $x' \geq x$,

$$0 \leq \phi\left(x'\right) - \phi\left(x\right) \leq \left(1 - \frac{1 - x_1'}{1 - x_1}\frac{1 - x_2'}{1 - x_2}\right)\left(1 - \phi\left(x\right)\right).$$

The inequalities displayed above imply

$$x'\left(n\right) \to x \implies \phi\left(x'\left(n\right)\right) \to \phi\left(x\right)$$

as long as the sequence $\left(x'\left(n\right)\right)$ has $x'\left(n\right) \geq x$ for all n.

4. Take any two type profiles x' and x, not necessarily ordered. Let $x' \vee x$ and $x' \wedge x$ denote, respectively, the pairwise maximum and pairwise minimum of x' and x. By 1,

$$\begin{aligned} \phi\left(x' \wedge x\right) &\leq \phi\left(x\right) \leq \phi\left(x' \vee x\right) \\ \phi\left(x' \wedge x\right) &\leq \phi\left(x'\right) \leq \phi\left(x' \vee x\right). \end{aligned}$$

If $x' \to x$, then $x' \vee x \to x' \wedge x$. So, By 3, $x' \to x$ implies $\phi\left(x' \vee x\right) \to \phi\left(x' \wedge x\right)$, which, combined with the display above, gives

$$\phi\left(x'\right) \to \phi\left(x\right).$$

In other words, ϕ is continuous on $[0,1]^2$.

□

4.4 Viscosity Characterization of the Value Function

The viscosity characterization of the value function has two components: the relevant HJBQVI equation and the notion of the viscosity solution for that equation. We introduce each in turn and then present the characterization.

The HJBQVI Revisited

It is convenient to transform (5), the HJBQVI equation with which we have been working so far, into the equivalent form that makes apparent the applicability of the existing techniques developed to deal with HJBQVI equations induced by convex Hamiltonians (to be clarified shortly). To do so, note that, because the right-hand side of (5) is zero, one can multiply each argument of the min operator by an arbitrary positive function of x without affecting the set of the generalized solutions. In particular, multiplying the HJB component of (5) by $1/x_i$ and taking u outside the min operator gives the transformed **HJBQVI**:

$$u(x) + H(x, \nabla u(x)) = 0, \qquad x \in \Omega, \tag{50}$$

where

$$\nabla u(x) \equiv \left(\frac{\partial u(x)}{\partial x_1}, \frac{\partial u(x)}{\partial x_2} \right),$$

and, for any $p \equiv (p_1, p_2) \in \mathbb{R}^2$,

$$H(x, p) \equiv \min_{i \in \{1,2\}} \left\{ (1 - x_i) p_i - 1 + \frac{\hat{c}}{x_i}, -x_i \right\}. \tag{51}$$

The function H is called **Hamiltonian**.[17] The associated **boundary condition** (6), unchanged, is reproduced here for convenience:

$$u(x) = \max_{i \in \mathcal{N}} \{x_i\}, \qquad x \in \partial\Omega. \tag{6}$$

All future references to HJBQVI are to the transformed equation (50).

[17] As a matter of terminology, H in (51) is not a Hamiltonian for the HJBQVI in (5), but is the Hamiltonian for the transformed HJBQVI in (50).

The Concept of Viscosity Solution

Viscosity solution refines a generalized solution (Definition 1) by imposing restrictions on the function even at the points at which no derivative exists. Superdifferential and subdifferential exist more generally, and viscosity solution disciplines them. Formally, define $D^{+}u(x)$ and $D^{-}u(x)$ to be, respectively, the (local) **superdifferential** and the (local) **subdifferential** of some function u at x:

$$\begin{aligned} D^{+}u(x) &\equiv \left\{ p \in \mathbb{R}^2 \mid \limsup_{y \to x} \frac{u(y) - u(x) - p \cdot (y - x)}{|y - x|} \leq 0 \right\} \\ D^{-}u(x) &\equiv \left\{ p \in \mathbb{R}^2 \mid \liminf_{y \to x} \frac{u(y) - u(x) - p \cdot (y - x)}{|y - x|} \leq 0 \right\}. \end{aligned}$$

Any vector p in $D^{-}u(x)$ is called a **subgradient** of u at x and has the property that, for a sufficiently small neighborhood $S \subset \Omega$ of x, at point $(x, u(x))$, vector $(p, -1)$ supports the epigraph of u restricted to S.[18] Graphically, the epigraph of u restricted to S is the "downward kink" of u at x.[19] Analogously, any vector p in $D^{+}u(x)$ is a **supergradient** of u at x and has the property that vector $(p, -1)$ supports an "upward kink" of u at x.

Definition 2. A continuous function $u^{*} : \Omega \to \mathbb{R}$ is a **viscosity solution** of (50) if, for every $x \in \Omega$,

$$p \in D^{+}u^{*}(x) \implies u^{*}(x) + H(x, p) \leq 0 \tag{52}$$

$$p \in D^{-}u^{*}(x) \implies u^{*}(x) + H(x, p) \geq 0. \tag{53}$$

Henceforth, when referring to a viscosity solution without explicitly specifying the equation, we refer to HJBQVI subject to the boundary condition.

[18] An **epigraph** is the set of points lying on or above the graph of a function.

[19] In contrast to another common definition of a (non-local) subgradient, here, we do not presume that u is convex. Therefore, the restriction of u to a sufficiently small neighborhood S is necessary to isolate the downward kink of u at x, thereby requiring a local definition of subgradient.

The Viscosity Characterization of the Value Function

Proposition 1. *A function is the sought value function ϕ, defined in problem (4), if and only if it is the (unique) continuous viscosity solution of HJBQVI subject to the boundary condition.*

The necessity and sufficiency claims in Proposition 1 can be split into two lemmas. The necessity part follows.

Lemma 10. *The value function ϕ, defined in (4), is a continuous viscosity solution of HJBQVI subject to the boundary condition.*

Proof (Sketch). The lemma's conclusion follows by combining the arguments in the proofs of Theorems 9.2 and 9.8 of Oksendal and Sulem (2005). Their theorems show that, in both the optimal-stopping problem (their Theorem 9.2) and the combined stochastic-control and impulse-control problem (their Theorem 9.8), the sought value function is a viscosity solution of the relevant HJBQVI. Our problem has both optimal stopping and stochastic control. The theorems of Oksendal and Sulem (2005) apply because, in our problem, the sought value function is continuous (Lemma 9) and DPP holds (Lemma 7).

The working assumption of Oksendal and Sulem (2005) is that the jump process has a state-independent intensity. This assumption is unnecessarily strong; their arguments admit a state-dependent jump intensity and so apply in our setting. In particular, with a state-dependent jump intensity, stochastic integration with respect to a jump processes remains valid (see, e.g., Protter, 1990, Chapter I.6) and the conclusion of Dynkin's lemma still holds (see, e.g., Hanson, 2007, Theorem 7.1). □

The sufficiency part of Proposition 1 follows.

Lemma 11. *HJBQVI subject to the boundary condition has a unique continuous viscosity solution.*

Proof. The lemma's conclusion follows from Remark 3.2 of Bardi and Capuzzo-Dolcetta (1997, p. 53) once the Hamiltonian H has been shown to satisfy condition H_1 of their Theorem 3.1 (Bardi and Capuzzo-Dolcetta, 1997, p. 51).

Condition H_1 is ascertained by verifying the conditions in Remark 3.4 of Bardi and Capuzzo-Dolcetta (1997, p. 54), which applies because H is of the form in equation (0.1) of Bardi and Capuzzo-Dolcetta (1997, p. 25).[20] Indeed, to cast our problem in the notation of that condition, define $Z \equiv \{1,2,3,4\}$ and rewrite H:

$$H(x,p) \equiv \min_{z \in Z} \{-f(x,z) \cdot p - l(x,z)\},$$

with the convention $f(x,1) = f(x,3) = (0,0)$, $f(x,2) = (-(1-x_1),0)$, and $f(x,4) = (0,-(1-x_2))$, and $l(x,1) = -x_1$, $l(x,3) = -x_2$, $l(x,2) = 1 - \hat{c}/x_1$, and $l(x,4) = 1 - \hat{c}/x_2$.

Note that the function f is continuous (in particular, because Z is a discrete set); the set Z is finite; because linear, the function $x \longmapsto f(x,z)$ is Lipschitz continuous uniformly with respect to $z \in Z$; $x \longmapsto l(x,z)$ is (uniformly) continuous in $\bar{\Omega}$ (because continuously differentiable), with a modulus of continuity independent of $z \in Z$ (because Z is a finite set). Hence, the conditions in Remark 3.4 of Bardi and Capuzzo-Dolcetta (1997, p. 54) are satisfied, so Remark 3.2 of Bardi and Capuzzo-Dolcetta (1997, p. 53) applies, and hence the lemma's conclusion follows. □

Lemmas 10 and 11 combine to yield the proof of Proposition 1.

Proof of Proposition 1. That the value function is a continuous viscosity solution follows from Lemma 10. That any continuous viscosity solution is a value function follows from the uniqueness of the continuous viscosity solution (Lemma 11) and from the fact that the value function is a continuous viscosity solution (Lemma 10). □

[20] The Hamiltonian in equation (0.1) of Bardi and Capuzzo-Dolcetta (1997, p. 25) is a sup, whereas H in (51) is an inf (indeed, a min). This discrepancy is inessential for verifying H_1 because $|\sup\{A\} - \sup\{B\}| = |\inf\{-A\} - \inf\{-B\}|$, and because each property required of each function in Remark 3.4 of Bardi and Capuzzo-Dolcetta (1997, p. 54) holds if and only if the corresponding property holds for the negative of that function.

4.5 Verification

Here, we finally prove Theorem 1, which verifies the conjectured value function. By Proposition 1, it suffices to verify that this function is a viscosity solution. To do so, we exploit the special structure of HJBQVI afforded by the fact that we are dealing with a maximization problem. This special structure is captured by the concavity of the Hamiltonian: $p \longmapsto H(x, p)$ is (weakly) concave for any $x \in \Omega$, as can be ascertained by inspection of (51). The implication is, roughly, that viscosity solution imposes no restrictions on downward kinks of the function (Bardi and Capuzzo-Dolcetta, 1997, Theorem 5.6), so only the upward kinks need to be checked.

It turns out that the conjectured solution has no upward "kinks," so there are no kinks to check. To formalize the lack of upward kinks, we define semiconvexity. A function u is **semiconvex** if there exists a constant $\kappa \geq 0$ such that the function $x \longmapsto u(x) + \kappa |x|^2$ is convex.

Lemma 12. *The conjectured value function F, defined in (45), is locally Lipschitz continuous and seminconvex.*

Proof. In Section 3, F has been constructed as an upper envelope of the surfaces derived from the value functions for various auxiliary problems. Each constituent surface is semiconvex because it is in $\mathcal{C}^2$. Indeed, any function in $\mathcal{C}^2$ has a continuously differentiable—and hence Lipschitz, and hence locally Lipschitz—gradient. Therefore, by the observation in the last sentence of Bardi and Capuzzo-Dolcetta (1997, p. 65), each constituent surface is semiconvex.

Because F is an upper envelope of the semiconvex surfaces, the conclusion of Bardi and Capuzzo-Dolcetta (1997, Exercise 4.6, p.75) applies and F itself is semiconvex.

Because F is semiconvex, it is locally Lipschitz continuous (Bardi and Capuzzo-Dolcetta, 1997, Proposition 4.6, p. 66). □

Lemma 13. *The conjectured value function F, defined in (45), is a continuous viscosity solution of HJBQVI subject to the boundary condition.*

Proof. By Lemmas 1, 3, and 6, F is a generalized solution of HJBQVI subject to the boundary condition.

Because the Hamiltonian $p \longmapsto H(x, p)$ is weakly concave for any $x \in \Omega$, and because F is a semiconvex locally-Lipschitz (Lemma 12) and a generalized solution of HJBQVI, Corollary 5.2(i) of Bardi and Capuzzo-Dolcetta (1997, p. 78) implies that F is a viscosity solution. □

The verification of the conjectured value function can now be completed.

Proof of Theorem 1. By Lemma 13, F is a viscosity solution. By Proposition 1, the viscosity solution is the value function. Hence, $\phi = F$, as desired. □

5 An Optimal-Policy Implementation in the Presence of Moral Hazard and Adverse Selection

In the preceding sections, we derived an optimal policy by assuming that HQ can both directly control divisions' learning and observe the news, if any. What if HQ can do neither? That is, can the optimal policy still be implemented in the presence of moral hazard and adverse selection? This section shows that it can be.

The proposed implementation is accomplished in a dynamic auction.[21] This auction begins with indicative bidding, followed by information acquisition and then firm bidding. The firm bidding occurs by the deadline that HQ specifies on the basis of the indicative bids. In addition, either division can call for early firm bidding, before the deadline.

[21] The auction's construction is inspired by VCG mechanism's dynamic extensions by Bergemann and Välimäki (2010), and Athey and Segal (2007).

5.1 The Optimal-Auction Game

Building on the environment described in Section 2, we define the **optimal auction game**, which features both moral hazard and adverse selection. The game has two players, D_1 and D_2, who are bidders. HQ is the auctioneer.

Timing and Actions

At time $t = 0$, each D_i privately observes his initial type $X_i(0) \in [0,1]$ and holds a commonly known belief about the full-support probability distribution from which $X_{-i}(0)$ is drawn.[22] Having observed $X_i(0)$, D_i publicly submits an **indicative bid** $b_i \in [0,1]$.

The submitted indicative-bid pair $b \equiv (b_1, b_2)$ determines the deadline by which HQ solicits firm bids. To construct this deadline, let $\bar{X}_{|b}$ denote the type process $X \equiv \{(X_1(t), X_2(t)) \mid t \geq 0\}$ that is conditional on $(X_1(0), X_2(0)) = b$ and on each division learning according to the optimal learning policy α^* without ever observing the good news. Given the optimal stopping time τ^*, define

$$\bar{\tau}_{|b} \equiv \mathbb{E}\left[\tau^* \mid X = \bar{X}_{|b}\right]$$

to be the (deterministic) time necessary for the item to be optimally allocated conditional on $X = \bar{X}_{|b}$. Time $\bar{\tau}_{|b}$ is the deadline for firm bidding.[23] Until time $\bar{\tau}_{|b}$, each division may learn.

In particular, at time $t \in \left(0, \bar{\tau}_{|b}\right)$, each D_i

- privately takes a learning action $a_i(t) \in \{0,1\}$ and, if $a_i(t) = 1$, privately observes the learning outcome (the good news or no news),[24]
- makes a flow payment $p_i(t) = c\alpha^*_{-i}\left(\bar{X}_{|b}(t)\right)$ to HQ,

[22] We do not specify this probability distribution because it is not needed in the analysis.

[23] The stopping time $\bar{\tau}_{|b}$ is finite by Lemma 8.

[24] Because no one but the division itself observes whether it learns and what it learns, the model features both moral hazard and adverse selection.

- either publicly requests firm bidding by choosing action $r_i(t) = 1$ or refrains from doing so by choosing action $r_i(t) = 0$.

Let T denote the minimum of $\bar{\tau}_{|b}$ and the first time at which any division publicly requests firm bidding:

$$T \equiv \bar{\tau}_{|b} \wedge \inf\{t \geq 0 \mid r_1(t) + r_2(t) \geq 1\}. \tag{54}$$

At time T, each D_i submits its **firm bid** $B_i \in [0, 1]$. The item is allocated and time-T payments are administered according to the rules of the second-price auction. (The ties are resolved, say, uniformly at random.)

No action can be taken by any player at any time $t > T$; the game ends at T.

Payoffs

Let $\mathbf{1}_{\{\mathrm{D}_i \text{ wins}\}}$ denote the indicator function that equals 1 if and only if D_i wins the second price auction at the firm-bidding stage. In this notation, D_i's payoff is

$$\mathbf{1}_{\{\mathrm{D}_i \text{ wins}\}} v_i - c \int_0^T a_i(s)\,\mathrm{d}s - \int_0^T p_i(s)\,\mathrm{d}s - B_{-i},$$

where T, defined in (54), denotes the time at which the item is allocated.

Consistent with the definition of HQ's objective function in the expected cash-flow maximization problem (3), the cash flow is

$$\sum_{i \in \mathcal{N}} \left(\mathbf{1}_{\{\mathrm{D}_i \text{ wins}\}} v_i - c \int_0^T a_i(s)\,\mathrm{d}s \right).$$

The payments $(p_i, B_i)_{i \in \mathcal{N}}$ do not affect the cash flow. If interpreted as monetary, these payments cancel out because the divisions' monetary loss is HQ's gain. Alternatively, these payments could be the divisions' productive efforts, which are costly to the divisions but valuable to HQ, and so would cancel out, too.

5.2 Implementation of the Optimal Policy

We show that the optimal auction game has an equilibrium whose outcome is the optimal policy.

Stationary Strategies

We restrict attention to equilibria in stationary strategies. Only stationary strategies will be defined, to conserve notation.[25] Let $\tilde{X}^i(t) \equiv \left(\tilde{X}_1^i(t), \tilde{X}_2^i(t)\right)$ denote D_i's time-t belief, where $\tilde{X}_j^i(t)$ is the expected value of v_j conditional on information available to D_i at time t. Thus, we distinguish between $\tilde{X}^i$, D_i's belief process about (v_1, v_2), and X, the actual type process that obtains conditional on the initial type profile $(X_1(0), X_2(0))$ and the subsequent learning actions of bidders. The stationarity of D_i's strategy is defined with respect to each division's belief process $\tilde{X}^i$.

D_i's (stationary) **strategy** is a tuple $\left(\tilde{b}_i, \tilde{a}_i, \tilde{r}_i, \tilde{B}_i\right)$, where

- indicative bidding strategy $\tilde{b}_i$ associates an indicative bid $\tilde{b}_i(x) \in [0,1]$ with each initial belief $\tilde{X}^i(0) = x$,
- learning strategy $\tilde{a}_i$ associates a decision $\tilde{a}_i(x) \in \{0,1\}$ of whether to learn with each belief $\tilde{X}^i(t) = x$,
- stopping strategy $\tilde{r}_i$ associates a decision $\tilde{r}_i(x) \in \{0,1\}$ of whether to call for firm bidding with each belief $\tilde{X}^i(t) = x$,
- firm bidding strategy $\tilde{B}_i$ associates a firm bid $\tilde{B}_i(x) \in [0,1]$ with each terminal belief $\tilde{X}^i(T) = x$.

[25] If D_{-i} uses a stationary strategy, D_i has a stationary best response. Thus, the focus on stationary strategies is an equilibrium refinement, not a restriction on admissible strategies.

Equilibrium Defined

The divisions' belief processes $\left(\tilde{X}^i\right)_{i\in\mathcal{N}}$ and strategies $\left(\tilde{b}_i, \tilde{a}_i, \tilde{r}_i, \tilde{B}_i\right)_{i\in\mathcal{N}}$ constitute a (Bayes-Nash) **equilibrium** if the beliefs conform with the Bayes rule given the strategies, and the strategies are best responses given these beliefs:

- (the Bayes rule) each D_i uses the Bayes rule and the knowledge of equilibrium to derive its belief process $\tilde{X}^i$ from the observations of D_{-i}'s indicative bid, D_i's initial type and own news, and D_{-i}'s decisions of whether to call for firm bidding;

- (best responses) for any belief $x \in [0,1]^2$ of D_i, its indicative bid $\tilde{b}_i(x)$, its decisions $\tilde{a}_i(x)$ and $\tilde{r}_i(x)$, and its firm bid $\tilde{B}_i(x)$ maximize the expected payoff in the continuation of the optimal auction game.

The Implementation Theorem

Theorem 2. *The optimal auction implements the optimal policy, which solves (4). That is, the optimal auction game has an equilibrium at which:*

1. *(Indicative Bidding) Each division bids its initial belief about its project's value. That is, $\tilde{b}_i(x) = b_i^*(x) \equiv x_i$ for all $x \in [0,1]^2$.*

2. *(Learning) Each division learns as prescribed by the optimal policy. Each division requests firm bidding when and only when it observes the good news. That is, $\tilde{a}_i(x) = \alpha_i^*(x)$ and $\tilde{r}_i(x) = r_i^*(x) \equiv \mathbf{1}_{\{x_i=1\}}$ for all $x \in [0,1]^2$.*

3. *(Firm Bidding) At the deadline T or as soon as either division requests firm bidding (whichever occurs first), each division bids its terminal belief about its project's value. That is, $\tilde{B}_i(x) = B_i^*(x) \equiv x_i$ for all $x \in [0,1]^2$.*

Moreover, each D_i's expected payoff conditional on the initial type profile is nonnegative.

Proof. Note that, on the path of the equilibrium described in the theorem's statement, each D_i's belief, $\tilde{X}^i$, coincides with the actual type process X, influenced by both divisions' learning outcomes: $\tilde{X}^1 = \tilde{X}^2 = X$.

The remainder of the proof proceeds in two steps. Step 1 shows that $\left(b_i^*, \alpha_i^*, r_i^*, B_i^*\right)_{i \in \mathcal{N}}$ is an equilibrium strategy profile by showing that each division is a "residual claimant" to the cash flow. Step 2 shows that each division's expected equilibrium payoff is non-negative.

<u>Step 1</u>

Throughout this step, we assume that D_{-i} adheres to its posited equilibrium strategy, whereas D_i contemplates deviations, which will be shown to be unprofitable.

Because the firm-bidding stage is a second-price auction, at any firm-bidding time T, for any belief $\tilde{X}^i(T) = x$, bidding $\tilde{B}_i(x) = B_i^*(x) \equiv x_i$ is a weakly dominant strategy for D_i. It remains to verify that deviating to some alternative strategy $\left(\tilde{b}_i, \tilde{\alpha}_i, \tilde{r}_i\right)$ prior to the firm-bidding stage cannot improve on $\left(b_i^*, \alpha_i^*, r_i^*\right)$.

From the knowledge of his own indicative-bidding strategy $\tilde{b}_i$ and D_{-i}'s strategy, D_i can infer D_{-i}'s belief process $\tilde{X}^{-i}$, D_{-i}'s learning process $\left\{\alpha_{-i}^*\left(\tilde{X}^{-i}(t)\right) \mid t \geq 0\right\}$, and hence also D_{-i}'s actual type process X_{-i}. Consequently, following any deviation $\tilde{b}_i$, D_i's belief process coincides with the actual type process, $\tilde{X}^i = X$, whereas D_{-i}'s belief process may deviate from X.

To express D_i's payoff as a function of its strategy, let $\tilde{\alpha} \equiv \left\{\left(\tilde{\alpha}_i(X(t)), \alpha_{-i}^*\left(\tilde{X}^{-i}(t)\right)\right) \mid t \geq 0\right\}$ denote the learning process induced by D_i's strategy $\left(\tilde{b}_i, \tilde{\alpha}_i, \tilde{r}_i\right)$. Let $X^{\tilde{\alpha}}$ denote the type process associated with $\tilde{\alpha}$.

Let $\tilde{\tau}$ denote the firm-bidding stopping time induced by D_i's strategy $\left(\tilde{b}_i, \tilde{\alpha}_i, \tilde{r}_i\right)$. D_i's choice of $\tilde{r}_i$ affects $\tilde{\tau}$ through the term $\inf\{t \geq 0 \mid r_1(t) + r_2(t) \geq 1\}$ in (54). D_i's choice of $\tilde{b}_i$ affects $\tilde{\tau}$ both directly and indirectly. Directly, $\tilde{b}_i$ affects $\tilde{\tau}$ by affecting HQ's choice of the firm-bidding deadline, $\bar{\tau}_{|b}$. Indirectly, $\tilde{b}_i$ affects D_{-i}'s belief $\tilde{X}^{-i}$ and thus also the learning process $\left\{\alpha_{-i}^*\left(\tilde{X}^{-i}(s)\right) \mid t \geq 0\right\}$, which, in turn, affects term $\inf\{t \geq 0 \mid r_1(t) + r_2(t) \geq 1\}$

in (54) through D_{-i}'s strategy r^*_{-i}.

D_i's continuation payoff at an (actual) type profile x, before firm bidding, is

$$\mathbb{E}\left[\max_{j\in\mathcal{N}}\left\{X_j^{\tilde{\alpha}_j}(\tilde{\tau})\right\} - c\sum_{j\in\mathcal{N}}\int_t^{\tilde{\tau}}\tilde{\alpha}_j\left(X^{\tilde{\alpha}}(s)\right)\mathrm{d}s \mid x\right] - \mathbb{E}\left[X_{-i}^{\tilde{\alpha}_{-i}}(\tilde{\tau}) \mid x\right]. \tag{55}$$

The first term in (55) is the cash flow. The second term will be shown to equal x_{-i}.

To see that the second term in (55) is x_{-i}, note that the process $\left\{X_{-i}^{\tilde{\alpha}_{-i}}(t) \mid t \geq 0\right\}$ is a bounded, and therefore uniformly integrable, martingale, and $\tilde{\tau}$ is a stopping time—both with respect to the filtration generated by $\left\{X^{\tilde{\alpha}}(t) \mid t \geq 0\right\}$. Hence, by Doob's optional sampling theorem, $\mathbb{E}\left[X_{-i}^{\tilde{\alpha}_{-i}}(\tilde{\tau}) \mid x\right] = x_{-i}$, a constant that is independent of D_i's choices.

Thus, D_i chooses its strategy $\left(\tilde{b}_i, \tilde{\alpha}_i, \tilde{r}_i\right)$ to maximize the expected cash flow, the first term in (55). Consequently, D_i 's problem is equivalent to solving HQ's problem (4) for an optimal policy (α, τ), but with additional restrictions on the class of admissible policies. Relative to HQ's policy choice in problem (4), D_i's choice of (α, τ) is restricted because D_i can affect D_{-i}'s strategy $(\tilde{\alpha}_{-i}, \tilde{r}_{-i})$ only indirectly, through $\tilde{b}_i$. Nevertheless, D_i can access HQ's (first-best) optimal policy by adopting its prescribed strategy $\left(b_i^*, \alpha_i^*, r_i^*\right)$, thereby inducing not only $\tilde{\alpha}_i = \alpha_i^*$, but also $\tilde{\alpha}_{-i} = \alpha_{-i}^*$ and $\tilde{\tau} = \tau^*$.

Thus, D_i's best response to D_{-i}'s posited strategy is D_i's posited strategy, as desired.

Step 2

To show that each D_i's expected equilibrium payoff is non-negative, a stronger result will be proved. This stronger result is that, on the equilibrium path, even if D_i could exit the optimal auction at any time t after having settled all past payments $\{p_i(s) \mid s \leq t\}$ and foregoing the opportunity to win the item, D_i would not benefit from this exit. (The theorem's statement corresponds to the special case of $t = 0$.) This stronger result follows from D_i's payoff decomposition in (55), which, following the application of Doob's

optional sampling theorem, yields the equilibrium continuation payoff:

$$\mathbb{E}\left[\max_{j\in\mathcal{N}}\left\{X_j^{\alpha^*}(\tau^*)\right\} - c\sum_{j\in\mathcal{N}}\int_t^{\tau^*}\alpha_j^*\left(X^{\alpha^*}(s)\right)\mathrm{d}s \mid x\right] - x_{-i}. \tag{56}$$

The first term in (56) is the maximized cash flow. The maximized cash flow is at least x_{-i} because allocating the item immediately to D_{-i} is always possible. Hence, (56) is nonnegative, as desired. □

In the optimal auction, the payments make each division a residual claimant to HQ's cash flow. At the final, firm-bidding, stage, the second-price-auction payments make the winner pay the "externality" that he imposes on the loser by depriving the loser of winning. The payments at the preceding, learning, stage take care of the externality that a division's presence imposes on the other division when it makes the other division engage in costly learning, instead of letting it win immediately.[26]

6 Conclusions

The paper solves for a cash-flow-maximizing policy for a company that faces an irreversible project-selection problem with information acquisition. This policy is implemented in a dynamic auction that features indicative bidding, a spell of learning, and a second-price auction. The use of an internal auction to make a decision within a company is plausible. Companies have successfully used market-like mechanisms to predict sales (Hewlett-Packard), manage manufacturing capacity (Intel), generate business ideas (General Electric), and select marketing campaigns (Starwood).[27] An internal auction would be impractical, however, if a company were unable to commit to subsequently not undoing any payments received from or made to its divisions in the course of the auction.

[26]The flow payments at the learning stage can be replaced by a side payment paid by the winner (in addition to the second-price auction payment), without affecting the divisions' incentives and the expected cash-flow.

[27]See Wikipedia entry "Prediction Market" for further examples and references.

One would expect the requisite commitment to be available to successful companies with a developed reputation for committing to cash-flow-maximizing decisions.

A Appendix: Omitted Proofs

All proofs are in the papers' main body.

References

Athey, Susan and Ilya Segal, "An Efficient Dynamic Mechanism," *Stanford University Working Paper*, 2007. 8, 54

Bardi, Martino and Italo Capuzzo-Dolcetta, *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations*, Birkhauser, 1997. 42, 51, 52, 53, 54

Bergemann, Dirk and Juuso Välimäki, "The Dynamic Pivot Mechanism," *Econometrica*, 2010, *78* (2), 771–789. 8, 54

Bolton, Patrick and Christopher Harris, "Strategic Experimentation," *Econometrica*, 1999, *67* (2), 349–374. 7

Clausen, Andrew and Carlo Strub, "Envelope Theorems for Non-Smooth and Non-Concave Optimization," 2012. 42

Compte, Olivier and Philippe Jehiel, "Auctions and information acquisition: sealed bid or dynamic formats?," *The RAND Journal of Economics*, 2007, *38* (2), 355–372. 8

Crémer, Jacques, Yossi Spiegel, and Charles Z. Zheng, "Auctions with Costly Information Acquisition," *Economic Theory*, 2009, *38* (1), 41–72. 8

de Motta, Adolfo, "Managerial Incentives and Internal Capital Markets," *The Journal of Finance*, 2003, *58* (3), 1193–1220. 7

Dixit, Avinash and Robert Pindyck, *Investment Under Uncertainty*, Princeton University Press, 1994. 7

Gertner, Robert and David Scharfstein, *The Handbook of Organizational Economics*, Princeton University Press, 2012. 7

Hanson, Floyd B., *Applied Stochastic Processes and Control for Jump-Diffusions: Modeling, Analysis, and Computation (Advances in Design and Control)*, Society for Industrial and Applied Mathematics, 2007. 51

Harris, Milton and Artur Raviv, "The Capital Budgeting Process: Incentives and Information," *The Journal of Finance*, 1996, *51* (4), 1139–1174. 7

Inderst, Roman and Christian Laux, "Incentives in Internal Capital Markets: Capital Constraints, Competition, and Investment Opportunities," *The RAND Journal of Economics*, 2005, *36* (1), 215–228. 7

Keller, Godfrey, Sven Rady, and Martin Cripps, "Strategic Experimentation with Exponential Bandits," *Econometrica*, 2005, *73* (1), 39–68. 7

Klein, Nicolas A. and Sven Rady, "Negatively Correlated Bandits," 2010. 7

Krähmer, Daniel and Roland Strausz, "Optimal Procurement Contracts with Pre-Project Planning," *Review of Economic Studies*, 2011, *78* (3), 1015–1041. 8

Malenko, Andrey, "Optimal Design of Internal Capital Markets," 2012. 7

Milgrom, Paul R. and Ilya Segal, "Envelope Theorems for Arbitrary Choice Sets," *Econometrica*, 2002, *70* (2), 583–601. 23

Oksendal, Bernt and Agnes Sulem, *Applied Stochastic Control of Jump Diffusions*, Springer, 2005. 51

Ozbas, Oguzhan and David S. Scharfstein, "Evidence on the Dark Side of Internal Capital Markets," *The Review of Financial Studies*, 2010, *23* (2), 581–599. 2

Persico, Nicola, "Information Acquisition in Auctions," *Econometrica*, 2003, *68* (1), 135–148. 8

Pham, Huyen, "Optimal Stopping of Controlled Jump Diffusion Processes: A Viscosity Solution Approach," *Journal of Mathematical Systems, Estimation, and Control*, 1998, *8* (1), 1–27. 45

Protter, Philip, *Stochastic Integration and Differential Equations: A New Approach*, Springer-Verlag, 1990. 11, 46, 51

Rajan, Raghuram, Henri Servaes, and Luigi Zingales, "The Cost of Diversity: The Diversification Discount and Inefficient Investment," *The Journal of Finance*, 2000, *55* (1), 35–80. 7

Ross, Sheldon M., *Applied Probability Models with Optimization Applications*, Dover Publications, Inc., 1970. 15, 21, 45

Scharfstein, David S. and Jeremy C. Stein, "The Dark Side of Internal Capital Markets: Divisional Rent-Seeking and Inefficient Investment," *The Journal of Finance*, 2000, *55* (6), 2537–2563. 7

Shi, Xianwen, "Optimal Auctions with Information Acquisition," *Games and Economic Behavior*, 2012, *74* (2), 666–686. 8

Stein, Jeremy C., "Internal Capital Markets and the Competition for Corporate Resources," *The Journal of Finance*, 1997, *52* (1), 111–133. 7

Tirole, Jean, *The Theory of Corporate Finance*, Princeton University Press, 2006. 7

Webster, Thomas J., *Managerial Economics: Theory and Practice*, Academic Press, 2003. 2

B Supplementary Appendix: Auxiliary Technical Lemmas

Lemma B.1. *The cost c_1 defined in (17) is the unique solution of (10). Moreover, the inner minimum in (17) is attained at $\theta_1 = \theta_2$:*

$$\min_{0 \le \theta_1 \le \theta_2 \le 1} \Phi(\theta_1, \theta_2, c_1) = \min_{0 \le x \le 1} \Phi(x, x, c_1) = 0. \tag{B.1}$$

Proof. The proof proceeds in steps. □

1. Claim: A unique $x \in (0, c_2)$ solves (10).

   Proof: The conclusion follows by observing that the left-hand side of (10) is strictly decreasing in c_1 (verified by differentiating) and maps $(0, c_2)$ onto $\mathbb{R}_+$, and that the right-hand side is strictly increasing in c_1 and maps $(0, c_2)$ onto $(2, 4)$.

2. Claim: At the unique x that solves (10), (B.1) holds.

   Proof: Differentiating gives:

$$\frac{\mathrm{d}\Phi(\theta_1, \theta_2, x)}{\mathrm{d}\theta_1} = -\theta_2 \left(1 + \frac{1}{\theta_1} + \frac{x}{1 - x - \theta_2} + \log \frac{x(1 - \theta_1)}{\theta_1 (1 - x - \theta_2)}\right). \tag{B.2}$$

   Because

$$\frac{\mathrm{d}^2\Phi(\theta_1, \theta_2, x)}{\mathrm{d}\theta_1^2} = \frac{\theta_2}{\theta_1^2 - \theta_1^3} > 0,$$

   $\mathrm{d}\Phi(\theta_1, \theta_2, x) / \mathrm{d}\theta_1 \le 0$ for all $\theta_1 \le \theta_2$ if and only if $\mathrm{d}\Phi(\theta_1, \theta_2, x) / \mathrm{d}\theta_1|_{\theta_1 = \theta_2} \le 0$, or using (B.2),

$$1 + \frac{1}{\theta_2} + \frac{x}{1 - x - \theta_2} + \log \frac{x(1 - \theta_2)}{\theta_2 (1 - x - \theta_2)} \ge 0, \qquad \theta_2 \in [\underline{\theta}, \bar{\theta}]. \tag{B.3}$$

   To establish the inequality in (B.3), note that the derivative of its left-hand side is

$$\frac{\left(x - (1 - \theta_2)^2\right)\left(1 - x - \theta_2^2\right)}{(1 - \theta_2)\, \theta_2^2 (1 - x - \theta_2)^2},$$

where $1 - x - \theta_2^2 \geq 1 - \theta_2 > 0$ and $1 - x - \theta_2 \geq (1 - \theta_2)^2 \geq 0$, both by $\theta_2 (1 - \theta_2) \geq x$, or equivalently, by $\theta_2 \in [\underline{\theta}, \bar{\theta}]$. Hence, the sign of the derivative is the sign of $x - (1 - \theta_2)^2$, which changes from negative to positive, crossing zero at $\theta_2 = 1 - \sqrt{x}$, as θ_2 increases from $\underline{\theta}$ to $\bar{\theta}$. Hence, the left-hand side of (B.3) is minimized at $\theta_2 = 1 - \sqrt{x}$ and attains the value

$$\frac{2}{1 - \sqrt{x}} - \log \frac{\left(1 - \sqrt{x}\right)^2}{x} = 0,$$

where the equality follows from (10).

3. Claim: The unique x that solves (10) satisfies $x = c_1$, where c_1 is defined in (17).

   Proof: Differentiating,

$$\frac{\mathrm{d}\Phi(\theta_1, \theta_2, \hat{c})}{\mathrm{d}\hat{c}} = \frac{(1 - \theta_1)(1 - \theta_2)^2 \theta_2}{\hat{c}(1 - \hat{c} - \theta_2)^2} > 0.$$

   Hence, $\min_{0 \leq \theta_1 \leq \theta_2 \leq 1} \Phi(\theta_1, \theta_2, x) = 0$ implies $\min_{0 \leq \theta_1 \leq \theta_2 \leq 1} \Phi(\theta_1, \theta_2, \hat{c}) < 0$ for any $\hat{c} < x$. Therefore, $x = c_1$.

Lemma B.2. Suppose that $\hat{c} < c_1$. Define $f : [\underline{\theta}, \bar{\theta}] \to \mathbb{R}$ by $f(x) = \Phi(x, x, \hat{c})$. Then, the function f is at first positive, then intersects zero at a point, then is negative, then intersects zero at a point, and then is positive again.

Proof. The proof proceeds in steps.

1. Claim: $1 - \hat{c} - x^2 > 0$ for any $x \in [\underline{\theta}, \bar{\theta}]$.

   Proof:

$$1 - \hat{c} - x^2 \geq 1 - \hat{c} - \bar{\theta}^2 = \frac{1 - \sqrt{1 - 4\hat{c}}}{2} > 0.$$

2. Claim: $(x - x^*)\left(\hat{c} - (1 - x)^2\right) > 0$ for any $x \in [\underline{\theta}, x^*) \cup (x^*, \bar{\theta}]$, where $x^* = 1 - \sqrt{\hat{c}}$ is such that $x^* \in (\underline{\theta}, \bar{\theta})$.

Proof: The inequality follows by inspection; it remains to verify that $x^* \in \left(\underline{\theta}, \bar{\theta}\right)$. Indeed,

$$\begin{aligned} x^* - \underline{\theta} &= \frac{1 + \sqrt{1 - 4\hat{c}} - 2\sqrt{\hat{c}}}{2} > 0 \\ \bar{\theta} - x^* &= \frac{2\sqrt{\hat{c}} + \sqrt{1 - 4\hat{c}} - 1}{2} > 0. \end{aligned}$$

3. Claim: If $x \in \left[\underline{\theta}, \bar{\theta}\right]$ with $f(x) < 0$ exists, then f is at first positive, then intersects zero at a point, then is negative, then intersects zero at a point, and then is again positive.

   Proof: Because $f\left(\underline{\theta}\right) = f\left(\bar{\theta}\right) = 1$, $f(x) < 0 \implies x \in \left(\underline{\theta}, \bar{\theta}\right)$. Then, by

$$f(x) = (1 - x)\, g(x), \qquad \text{where } g(x) \equiv \frac{1 - \hat{c} - x^2}{1 - \hat{c} - x} + x \log \frac{\hat{c}(1 - x)}{x(1 - \hat{c} - x)},$$

   $f(x) = 0 \iff g(x) = 0$, which justifies the focus on g. Differentiating,

$$g'(x) = \frac{\hat{c}\left(1 - \hat{c} - x^2\right)}{(1 - x)(1 - \hat{c} - x)^2} + \log \frac{\hat{c}(1 - x)}{x(1 - \hat{c} - x)}.$$

   If $g(w) = 0$ for some $w \in \left[\underline{\theta}, \bar{\theta}\right]$, then

$$g'(w) = \frac{\left(\hat{c} - (1 - w)^2\right)\left(1 - \hat{c} - w^2\right)}{(1 - w)\, w\, (1 - \hat{c} - w)^2}.$$

   By Step 1, the sign of $g'(w)$ is the sign of $\hat{c} - (1 - w)^2$, which, by Step 2, changes the sign from negative to positive at $x^* \in \left(\underline{\theta}, \bar{\theta}\right)$. Hence, if x with $g(x) < 0$ exists, then g intersects zero twice: once from above and to the left of x^*, and once from below and to the right of x^*.

4. Claim: For some $x \in \left(\underline{\theta}, \bar{\theta}\right)$, $f(x) < 0$.

Proof: Note that, at $x \in \left[\underline{\theta}, \bar{\theta}\right]$,

$$\frac{\mathrm{d}g(x)}{\mathrm{d}\hat{c}} = \frac{(1-x)^2 x}{\hat{c}(1-\hat{c}-x)^2} > 0.$$

Then, $g(x^*)|_{\hat{c}=c_1} = 0$ implies $g(x^*) < 0$ for $\hat{c} < c_1$. Combining this observation with Step 3 implies the lemma's conclusion.

□

Lemma B.3. Suppose that $\hat{c} < c_2$. Define $f : \left[\underline{\theta}, \bar{\theta}\right] \to \mathbb{R}$ by $f(x) = \Phi(\theta_1, x, \hat{c})$. Then, f is negative on an interval, possibly an empty one.

Proof. The proof proceeds in steps.

1. Claim: Define $\kappa(x) \equiv (1-x)\left[(1-\theta_1)x^2 + 2\hat{c} - (1-x)\right] - \hat{c}^2$. Then, $(x - x^*)\kappa(x) > 0$ for any $x \in [\underline{\theta}, x^*) \cup \left(x^*, \bar{\theta}\right]$, for some $x^* \in \left(\underline{\theta}, \bar{\theta}\right)$.
   Proof: First, it will be shown that $\kappa\left(\bar{\theta}\right) > 0$. Indeed,

$$\kappa\left(\bar{\theta}\right) \geq \kappa\left(\bar{\theta}\right)\big|_{\theta_1=\bar{\theta}} = \frac{1}{2}\left(-1 + \sqrt{1-4\hat{c}} + 2\left(2 - \sqrt{1-4\hat{c}}\right)\hat{c}\right) > 0.$$

   To show the second inequality in the above display, it will be shown that $b(\hat{c}) \equiv \kappa\left(\bar{\theta}\right)\big|_{\theta_1=\bar{\theta}}$ is a strictly quasi-concave function of $\hat{c}$ on $(0, c_2)$. Indeed, differentiating implies that the sign of $b'(\hat{c})$ coincides with the sign of

$$\sqrt{1-4\hat{c}} + 3\hat{c} - 1,$$

   which is a strictly concave function in $\hat{c}$ on $(0, c_2)$ (which is verified by differentiating twice), evaluates to 0 at $\hat{c} = 0$, and evaluates to $-1/4$ at $\hat{c} = c_2$. Thus, on $(0, c_2)$, the above display is at first positive and then negative. Hence, $b'(\hat{c})$ on $(0, c_2)$ is at first positive and then negative, and so b is strictly quasi-concave. Finally, because $b(0) = b(c_2) = 0$, $b(\hat{c}) > 0$ on $(0, c_2)$, and so the desired inequality follows.

It will be shown that $\kappa(\underline{\theta}) < 0$. Indeed,

$$\kappa(\underline{\theta}) \leq \kappa(\underline{\theta})|_{\theta_1=\underline{\theta}} = \left(2+\sqrt{1-4\hat{c}}\right)\hat{c} - \frac{1}{2}\left(1+\sqrt{1-4\hat{c}}\right) < 0.$$

To show the second inequality in the above display, it will be shown that $\xi(\hat{c}) \equiv \kappa(\underline{\theta})|_{\theta_1=\underline{\theta}}$ is strictly increasing in $\hat{c}$ on $(0, c_2)$. Indeed, differentiating ξ gives

$$\xi'(\hat{c}) = \frac{2\left(1-3\hat{c}+\sqrt{1-4\hat{c}}\right)}{\sqrt{1-4\hat{c}}} > 0,$$

where the inequality follows from $\hat{c} \in (0, c_2)$. Furthermore, $\xi(c_2) = 0$, and so $\xi(\hat{c}) < 0$ for all $\hat{c} \in (0, c_2)$.

By differentiation, $\kappa''(x) = -(1-\theta_1)6x - 2\theta_1 < 0$, and so κ is a strictly concave function. Hence, by $\kappa(\underline{\theta}) < 0$ and $\kappa(\bar{\theta})$, on the interval $(\underline{\theta}, \bar{\theta})$, κ crosses zero once, from below, at a point, which is denoted by $x^* \in (\underline{\theta}, \bar{\theta})$.

2. <u>Claim</u>: If $x \in [\underline{\theta}, \bar{\theta}]$ with $f(x) < 0$ exists, then f is negative on an interval.

   <u>Proof</u>: If $f(s) = 0$ for some $s \in [\underline{\theta}, \bar{\theta}]$, then

$$f'(s) = \frac{(1-s)\left[(1-\theta_1)s^2 + 2\hat{c} - (1-s)\right] - \hat{c}^2}{s(1-\hat{c}-s)^2}.$$

   The sign of $f'(s)$ is the sign of the numerator, $\kappa(s)$, which, by Step 1, changes the sign from negative to positive at $x^* \in (\underline{\theta}, \bar{\theta})$.

   Hence, if $f(x) < 0$ for some $x \in [\underline{\theta}, \bar{\theta}]$ and $f(x') > 0$ for some $x' \in [x, \bar{\theta}]$, then it must be that f intersects zero somewhere in (x, x'), at which point f' is positive. Because the slope of f' must remain positive at any larger intersection point, there can be no larger intersection point. Because x with $f(x) < 0$ can be chosen to be arbitrarily small, the argument implies that the set $\{x \in [\underline{\theta}, \bar{\theta}] \mid f(x) < 0\}$ is convex.

$\square$

Lemma B.4. *Suppose that*

$$M^C(x) \equiv 1 - [1 - V(x,x)] \left(\frac{1-\theta_1}{1-x}\right)^2 - 2\hat{c}(1-\theta_1)^2 [\sigma(\theta_1) - \sigma(x)].$$

Then, $\arg\max_{x \in [\underline{\theta}, \bar{\theta}]} M^C(x) = \{\bar{\theta}\}$.

Proof. By differentiation,

$$\begin{aligned} \frac{\mathrm{d}M^C(x)}{\mathrm{d}x} &= \left(\frac{1-\theta_1}{1-x}\right)^2 \left[\frac{\mathrm{d}V(x,x)}{\mathrm{d}x} + 2\frac{V(x,x)-1}{1-x} + 2\hat{c}\frac{1}{x(1-x)}\right] \\ &= \hat{c}\frac{(1-\theta_1)^2}{x(1-x)^3}\Phi(x,x,\hat{c}), \end{aligned}$$

where the second equality follows by substitution. Thus the sign of $\mathrm{d}M^C(x)/\mathrm{d}x$ coincides with the sign of $\Phi(x,x,\hat{c})$. Then, by Lemma B.2, $\mathrm{d}M^C(x)/\mathrm{d}x$ first switches the sign from positive to negative and then switches the sign from negative to positive. These sign switches imply that M^C is wave-shaped, with local maxima at $\underline{a} \in (\underline{\theta}, \bar{\theta})$, defined in (28), and at $\bar{\theta}$. It remains to verify that $M^C(\bar{\theta}) > M^C(\underline{a})$.

Note that by (16), $V(\bar{\theta}, \bar{\theta}) = \bar{\theta}$. Substituting $V(\bar{\theta}, \bar{\theta}) = \bar{\theta}$ into the expression for $M^C(\bar{\theta})$ gives

$$M^C(\bar{\theta}) = 1 - (1-\theta_1)^2 \left(\frac{1}{1-\bar{\theta}} + 2\hat{c}\left[\sigma(\theta_1) - \sigma(\bar{\theta})\right]\right).$$

Similarly, substituting the expression for $V(\underline{a}, \underline{a})$ from (16) into the expression for $M^C(\underline{a})$ and using $\Phi(\underline{a}, \underline{a}, \hat{c}) = 0$ implied by (28) gives

$$M^C(\underline{a}) = 1 - (1-\theta_1)^2 \left(\frac{1}{1-b(\underline{a})} + \frac{\hat{c}}{\underline{a}(1-\underline{a})^2} + 2\hat{c}\left[\sigma(\theta_1) - \sigma(\underline{a})\right]\right).$$

Thus,

$$\begin{aligned}
M^C\left(\bar{\theta}\right) - M^C\left(\underline{a}\right) &= \left(1-\theta_1\right)^2\left(\frac{1}{1-b\left(\underline{a}\right)} + \frac{\hat{c}}{\underline{a}\left(1-\underline{a}\right)^2} + 2\hat{c}\left[\sigma\left(\bar{\theta}\right) - \sigma\left(\underline{a}\right)\right] - \frac{1}{1-\bar{\theta}}\right) \\
&= \left(1-\theta_1\right)^2\left(\begin{array}{c} \frac{1}{1-b(\underline{a})} + \hat{c}\left(2\sigma\left(\bar{\theta}\right) - \frac{1}{\bar{\theta}\left(1-\bar{\theta}\right)^2}\right) \\ -\hat{c}\left(2\sigma\left(\underline{a}\right) - \frac{1}{\underline{a}(1-\underline{a})^2}\right) - \frac{\bar{\theta}\left(1-\bar{\theta}\right)-\hat{c}}{\bar{\theta}\left(1-\bar{\theta}\right)^2} \end{array}\right) \\
&= \left(1-\theta_1\right)^2\left(\frac{1}{1-b\left(\underline{a}\right)} + \hat{c}\left(2\sigma\left(\bar{\theta}\right) - \frac{1}{\bar{\theta}\left(1-\bar{\theta}\right)^2}\right) - \hat{c}\left(2\sigma\left(\underline{a}\right) - \frac{1}{\underline{a}\left(1-\underline{a}\right)^2}\right)\right),
\end{aligned}$$

where the last equality follows from the definition of $\bar{\theta}$, which implies $\bar{\theta}\left(1-\bar{\theta}\right) - \hat{c} = 0$. Note that, for $x \in (0,1)$,

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(2\sigma\left(x\right) - \frac{1}{x\left(1-x\right)^2}\right) = \frac{1}{x^2\left(1-x\right)^2} > 0.$$

Hence, $\bar{\theta} > \underline{a}$ implies

$$2\sigma\left(\bar{\theta}\right) - \frac{1}{\bar{\theta}\left(1-\bar{\theta}\right)^2} > 2\sigma\left(\underline{a}\right) - \frac{1}{\underline{a}\left(1-\underline{a}\right)^2}.$$

Combined with $1 - b\left(\underline{a}\right) > 0$, the above display implies $M^C\left(\bar{\theta}\right) - M^C\left(\underline{a}\right) > 0$. That is, the function $M^C : \left[\underline{\theta}, \bar{\theta}\right] \to \mathbb{R}$ is uniquely maximized at $\bar{\theta}$. □

Lemma B.5. *Suppose that*

$$M^A\left(x;\theta_1\right) \equiv 1 - \left(1 - V\left(\theta_1, x\right)\right)\frac{1-\theta_2}{1-x} - \hat{c}\left(1-\theta_2\right)\left[\varphi\left(\theta_2\right) - \varphi\left(x\right)\right].$$

Then, $\arg\max_{x\in[\theta_1,\theta^-(\theta_1)]} M^A\left(x;\theta_1\right) = \left\{\theta^-\left(\theta_1\right)\right\}$.

Proof. To show $M^A(\theta^-(\theta_1);\theta_1) > M^A(d(\theta_1);\theta_1)$, write

$$\begin{aligned}\frac{1-M^A(\theta^-(\theta_1);\theta_1)}{1-\theta_2} &= \frac{1-V(\theta_1,\theta^-(\theta_1))}{1-\theta^-(\theta_1)} + \hat{c}\left[\varphi(\theta_2) - \varphi\left(\theta^-(\theta_1)\right)\right] \\ &= 1 + \hat{c}\left[\varphi(\theta_2) - \varphi\left(\theta^-(\theta_1)\right)\right],\end{aligned}$$

where the second equality is by $V(\theta_1,\theta^-(\theta_1)) = \theta^-(\theta_1)$.

Evaluating V in (16) at $(\theta_1, d(\theta_1))$ and using $\Phi(\theta_1, d(\theta_1), \hat{c}) = 0$ (by the definition of d in (34)), one can write

$$\frac{1-V(\theta_1,d(\theta_1))}{1-d(\theta_1)} = \frac{1-\theta_1}{1-b(d(\theta_1))} + \frac{\hat{c}}{d(\theta_1)(1-d(\theta_1))}.$$

Then,

$$\begin{aligned}\frac{1-M^A(d(\theta_1);\theta_1)}{1-\theta_2} &= \frac{1-V(\theta_1,d(\theta_1))}{1-d(\theta_1)} + \hat{c}\left[\varphi(\theta_2) - \varphi(d(\theta_1))\right] \\ &= \frac{1-\theta_1}{1-b(d(\theta_1))} + \frac{\hat{c}}{d(\theta_1)(1-d(\theta_1))} + \hat{c}\left[\varphi(\theta_2) - \varphi(d(\theta_1))\right].\end{aligned}$$

Subtracting gives

$$\begin{aligned}&\frac{M^A(\theta^-(\theta_1);\theta_1) - M^A(d(\theta_1);\theta_1)}{1-\theta_2} \\ &\quad = \hat{c}\left(\varphi\left(\theta^-(\theta_1)\right) - \frac{1}{\theta^-(\theta_1)(1-\theta^-(\theta_1))} - \left(\varphi(d(\theta_1)) - \frac{1}{d(\theta_1)(1-d(\theta_1))}\right)\right) \\ &\qquad + \frac{1-\theta_1}{1-b(d(\theta_1))} + \frac{\hat{c}}{\theta^-(\theta_1)(1-\theta^-(\theta_1))} - 1.\end{aligned}$$

The first line of the right-hand side of the above display is positive by $\theta^-(\theta_1) > d(\theta_1)$ and by

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\varphi(x) - \frac{1}{x(1-x)}\right) = \frac{1}{x^2(1-x)} > 0.$$

The second line, using the definition of b, can be rearranged to give

$$\frac{1-\theta_1}{1-b\left(d\left(\theta_1\right)\right)}+\frac{\theta_1}{\theta^-\left(\theta_1\right)}-1=\frac{b\left(d\left(\theta_1\right)\right)\left(\theta^-\left(\theta_1\right)-\theta_1\right)+\theta_1\left(1-\theta^-\left(\theta_1\right)\right)}{\theta^-\left(\theta_1\right)\left(1-b\left(d\left(\theta_1\right)\right)\right)}>0,$$

where the inequality follows from $\theta_1<\theta^-\left(\theta_1\right)<1$. □

Lemma B.6. *For $\theta_1 \in (\underline{a}, \bar{a})$,*

$$V\left(\theta_1, \theta^-\left(\theta_1\right)\right)>B\left(\theta_1, \theta^-\left(\theta_1\right)\right).$$

Proof. From the definition of B in (39),

$$B\left(\theta_1, \theta^-\left(\theta_1\right)\right)=1-\left(1-C\left(\theta_1\right)\right)\frac{1-\theta^-\left(\theta_1\right)}{1-\theta_1}-\hat{c}\left(1-\theta^-\left(\theta_1\right)\right)\left[\varphi\left(\theta^-\left(\theta_1\right)\right)-\varphi\left(\theta_1\right)\right].$$

Then,

$$V\left(\theta_1, \theta^-\left(\theta_1\right)\right)-B\left(\theta_1, \theta^-\left(\theta_1\right)\right)=\left(1-\theta^-\left(\theta_1\right)\right)\left(\frac{1-C\left(\theta_1\right)}{1-\theta_1}+\hat{c}\left[\varphi\left(\theta^-\left(\theta_1\right)\right)-\varphi\left(\theta_1\right)\right]-1\right).$$

From the definition of $C\left(\theta_1\right)$ in (26),

$$\begin{aligned}\frac{1-C\left(\theta_1\right)}{1-\theta_1} &= \left(1-\theta_1\right)\left(\frac{1-V(\underline{a}, \underline{a})}{(1-\underline{a})^2}+2\hat{c}\left[\sigma\left(\theta_1\right)-\sigma(\underline{a})\right]\right) \\ &= \left(1-\theta_1\right)\left(\frac{1}{1-b(\underline{a})}+\frac{\hat{c}}{\underline{a}(1-\underline{a})^2}+2\hat{c}\left[\sigma\left(\theta_1\right)-\sigma(\underline{a})\right]\right).\end{aligned}$$

Thus,

$$\begin{aligned}\frac{V\left(\theta_1,\theta^-\left(\theta_1\right)\right)-B\left(\theta_1,\theta^-\left(\theta_1\right)\right)}{1-\theta^-\left(\theta_1\right)} &= \left(1-\theta_1\right)\left(\frac{1}{1-b\left(\underline{a}\right)}+\frac{\hat{c}}{\underline{a}\left(1-\underline{a}\right)^2}+2\hat{c}\left[\sigma\left(\theta_1\right)-\sigma\left(\underline{a}\right)\right]\right)\\ &\quad+\hat{c}\left[\varphi\left(\theta^-\left(\theta_1\right)\right)-\varphi\left(\theta_1\right)\right]-1\\ &= \left(\frac{\left(1-\theta_1\right)}{1-b\left(\underline{a}\right)}+\frac{\hat{c}}{\theta^-\left(\theta_1\right)\left(1-\theta^-\left(\theta_1\right)\right)}-1\right.\\ &\quad+\left(1-\theta_1\right)\hat{c}\left[2\sigma\left(\theta_1\right)-\frac{1}{\theta_1\left(1-\theta_1\right)^2}-\left(2\sigma\left(\underline{a}\right)-\frac{1}{\underline{a}\left(1-\underline{a}\right)^2}\right)\right]\\ &\quad\left.+\hat{c}\left[\varphi\left(\theta^-\left(\theta_1\right)\right)-\frac{1}{\theta^-\left(\theta_1\right)\left(1-\theta^-\left(\theta_1\right)\right)}-\left(\varphi\left(\theta_1\right)-\frac{1}{\theta_1\left(1-\theta_1\right)}\right)\right]\right).\end{aligned}$$

Note that, using the definition of b,

$$\begin{aligned}\frac{\left(1-\theta_1\right)}{1-b\left(\underline{a}\right)}+\frac{\hat{c}}{\theta^-\left(\theta_1\right)\left(1-\theta^-\left(\theta_1\right)\right)}-1 &= \frac{\left(1-\theta_1\right)}{1-b\left(\underline{a}\right)}+\frac{\theta_1}{\theta^-\left(\theta_1\right)}-1\\ &= \frac{\theta_1\left[1-\theta^-\left(\theta_1\right)\right]+b\left(\underline{a}\right)\left[\theta^-\left(\theta_1\right)-\theta_1\right]}{\theta^-\left(\theta_1\right)\left(1-b\left(\underline{a}\right)\right)}>0,\end{aligned}$$

where the inequality follows from $\theta_1<\theta^-\left(\theta_1\right)<1$. Moreover,

$$2\sigma\left(\theta_1\right)-\frac{1}{\theta_1\left(1-\theta_1\right)^2}-\left(2\sigma\left(\underline{a}\right)-\frac{1}{\underline{a}\left(1-\underline{a}\right)^2}\right)>0$$

by $\theta_1>\underline{a}$ and by

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(2\sigma\left(x\right)-\frac{1}{x\left(1-x\right)^2}\right)=\frac{1}{x^2\left(1-x\right)^2}>0$$

for $x\in(0,1)$. Finally,

$$\varphi\left(\theta^-\left(\theta_1\right)\right)-\frac{1}{\theta^-\left(\theta_1\right)\left(1-\theta^-\left(\theta_1\right)\right)}-\left(\varphi\left(\theta_1\right)-\frac{1}{\theta_1\left(1-\theta_1\right)}\right)>0$$

by $\theta^-\left(\theta_1\right)>\theta_1$ and by

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\varphi\left(x\right)-\frac{1}{x\left(1-x\right)}\right)=\frac{1}{x^2\left(1-x\right)}>0$$

for $x \in (0,1)$. Thus, $V(\theta_1, \theta^-(\theta_1)) - B(\theta_1, \theta^-(\theta_1)) > 0$ as required. $\square$