Available online at www.sciencedirect.com
ScienceDirect
Journal of Economic Theory 180 (2019) 383–437
JOURNAL OF Economic Theory
www.elsevier.com/locate/jet

Optimal learning before choice ☆

T. Tony Ke [a,*], J. Miguel Villas-Boas [b]

[a] *MIT, United States of America*
[b] *University of California, Berkeley, United States of America*

Received 1 July 2017; final version received 21 January 2019; accepted 26 January 2019
Available online 5 February 2019

Abstract

A Bayesian decision maker is choosing among two alternatives with uncertain payoffs and an outside option with known payoff. Before deciding which alternative to adopt, the decision maker can purchase sequentially multiple informative signals on each of the two alternatives. To maximize the expected payoff, the decision maker solves the problem of optimal dynamic allocation of learning efforts as well as optimal stopping of the learning process. We show that the decision maker considers an alternative for learning or adoption if and only if the expected payoff of the alternative is above a threshold. Given both alternatives in the decision maker's consideration set, we find that if the outside option is weak and the decision maker's beliefs about both alternatives are relatively low, it is optimal for the decision maker to learn information from the alternative that has a lower expected payoff and less uncertainty, given all other characteristics of the two alternatives being the same. If the decision maker subsequently receives enough positive informative signals, the decision maker will switch to learning the better alternative; otherwise the decision maker will rule out this alternative from consideration and adopt the currently more preferred alternative. We find that this strategy works because it minimizes the decision maker's learning efforts. We also characterize the optimal learning policy when the outside option is relatively high, and discuss several extensions.

JEL classification: D83; L15; M31

☆ We thank Guy Barles, Jerker Denrell, Brett Green, Craig Evans, Steven Evans, Drew Fudenberg, Nicolae Gârleanu, Yuichiro Kamada, Ioannis Karatzas, Tim Laux, Giuseppe Moscarini, Aniko Öry, Jiwoong Shin, Philipp Strack, Wenpin Tang, Johan Walden, and Jidong Zhou for helpful comments and suggestions. We thank the review team for their insightful and constructive comments, which have made the paper much improved. All remaining errors are our own.

* Corresponding author.
E-mail addresses: kete@mit.edu (T.T. Ke), villas@haas.berkeley.edu (J.M. Villas-Boas).

https://doi.org/10.1016/j.jet.2019.01.005
0022-0531/

Keywords: Information; Bayesian learning; Search theory; Dynamic allocation; Optimal stopping; Consideration set

1. Introduction

In many circumstances, agents have the opportunity to learn sequentially about multiple different alternatives before making a choice. Specifically, consider a decision maker who is deciding among several alternatives with uncertain payoffs and an outside option with a known payoff. Before deciding which one to adopt, he can gather some information on each alternative. After learning more about an alternative, the decision maker gains more precise knowledge about its payoff, with initial learning being more informative than later learning. It is costly to gather and process information. Therefore, at some point, the decision maker will decide to stop learning and make a choice on which alternative to adopt or to take the outside option.

For example, consider a consumer in the market for a car. The consumer could first learn information about a certain model A, then choose to get information about model B, then go back to get more information on model A again, and so on, until the consumer decides to either purchase model A, model B, or some other model, or not to purchase a car for now. With the recent development of information technologies, it becomes more and more important to understand this information gathering behavior of consumers.

Information gathering is not unique to the consumers' purchase process. In fact, many other important economic activities involve similar costly gradual information acquisition: companies allocating resources to R&D, individuals looking for jobs, firms recruiting job candidates, politicians evaluating better public policies, manufacturers considering alternative suppliers, etc.

We study a decision maker's optimal information gathering problem under a Bayesian framework. We consider a setting in which the payoff of each alternative follows a two-point distribution, being either high or low. The decision maker has a prior belief on the probability that an alternative is of high payoff. Each time he learns some information about an alternative, he incurs a cost and gets a noisy signal on its payoff. This signal does not reveal the payoff of this alternative completely; rather, it updates the decision maker's belief of the distribution of the payoff. The decision maker could buy another signal on the same alternative if he would like to learn more about it or, alternatively, he could buy signals on other alternatives. In this setting the precision of the decision maker's belief is a function of his belief. Therefore, we only need to account for one state variable per alternative during the learning process—the probability that each alternative is of high payoff. Despite its simplicity, this setup captures the ideas that the decision maker gains more precise information during the learning process, and that the marginal information per learning effort decreases with the cumulative information gathered so far on an alternative. This two-point distribution assumption also implies that the agent becomes more certain about the payoff of an alternative when the payoff is either relatively high or low. The decision maker solves the problem of optimal dynamic allocation of learning efforts as well as optimal stopping of the learning process, so as to maximize the expected payoff.

We consider a two-alternative continuous-time model with infinite time horizon, which enables a time-stationary characterization of the optimal learning strategy when the outside option takes relatively large or small values.[1] The result is a partition of the belief space into regions,

[1] See, for example, Bolton and Harris (1999), Moscarini and Smith (2001), for the use of the same learning process in continuous time. See also Daley and Green (2012).

within which it is optimal to learn each alternative, adopt each alternative, or take the outside option.

We find that if the belief on an alternative is below a threshold, the decision maker will never consider that alternative—he will neither learn nor adopt that alternative. This then provides an endogenous formation of a consideration set, given the decision maker's belief on the payoff distribution, his learning costs, the value of the outside option, and the noisiness of the signals received. If two alternatives have the beliefs above that threshold, then the decision maker chooses to stop learning and adopt one alternative if the expected payoff of that alternative is sufficiently higher than that of the other alternative. That is, when the expected payoffs from the alternatives are not too different, the decision maker will continue to learn.

We also investigate the effect of the value of the outside option on optimal learning behavior. We find that when the value of the outside option is relatively high, the decision maker always chooses to learn about the alternative with the higher expected payoff first. More interestingly, we find that this is not necessarily the case when the value of the outside option is sufficiently low. In this case it may be optimal to first learn about the alternative that has a lower expected payoff, all else being equal among alternatives, with the purpose of possibly ruling it out early. If the decision maker receives sufficiently poor signals on this alternative, he will stop learning and immediately adopt the other alternative. We find this strategy works because it saves learning costs.

It is interesting to understand the intuition about why a decision maker's optimal learning strategy depends on the value of the outside option. When the outside option has a relatively high value, it will be relevant for the decision maker's ultimate choice at the end of the learning process. In this case, by learning the alternative with a higher expected payoff, the decision maker keeps both the outside option and the other alternative as reservation options. Therefore, the decision maker prefers to learn the alternative with higher expected payoff. On the other hand, when the outside option has a relatively low value, it is no longer as relevant for the decision maker's ultimate choice. In this case, the decision maker is basically deciding between the two uncertain alternatives, and it may be better to learn about the alternative with a lower expected payoff first to potentially rule it out early, so as to make a clear distinction between the two uncertain alternatives with minimum learning effort.

We also compute the decision maker's likelihood to adopt each alternative given his current belief, when he follows the optimal learning strategy. We find that the probability of choosing either one of the alternatives can fall as there are more alternatives available. This is because having more alternatives makes it harder for the decision maker to make a choice, and thus leads to more learning, which can ultimately result in no adoption of any of the alternatives. Specifically, information is ex ante neutral. It is possible that with more learning, the decision maker gets positive information on either alternative, in which case he will, at best, adopt one of the alternatives. On the other hand, it is also possible that with more learning, the decision maker gets negative information, in which case he may choose to take the outside option. We also compute the decision maker's expected probability of being correct ex post given his current belief, when he follows the optimal learning strategy. Finally, we consider the optimal learning problem when alternatives have heterogeneous learning costs or payoff distributions, when the number of alternatives is greater than two, and when there is time discounting instead of learning costs.

The problem considered here is related to the multi-armed bandit problem (e.g., Rothschild 1974a; Gittins 1979; Whittle 1980; Bergemann and Välimäki 1996; Felli and Harris 1996; Bolton and Harris 1999), where a decision maker learns about different options by trying them, one in each period, while earning some stochastic rewards along the way. That problem has an elegant result that the optimal policy is to choose the arm with the highest Gittins index, which for each arm only depends on what the decision maker knows about that arm until then. However, the problem considered here is different from the bandit problem in one major aspect. In our setting, a decision maker optimally decides when to stop learning and make an adoption. Therefore, the decision horizon is endogenous, and optimally determined by the decision maker. In contrast, multi-arm bandit problems generally presume an exogenously given decision horizon, which could be either finite or infinite. In fact, our problem belongs to a general class of the *stoppable bandits problem*, first introduced by Glazebrook (1979), which generalizes simple bandit processes to allow for two arms per bandit. By adding a second "stopping arm" to each bandit, extra payoff can be generated when the stopping arm is pulled. In general, index policies are not optimal for this type of problems (Gittins et al. 2011, Chapter 4). Glazebrook (1979) shows that an index policy can be optimal under certain conditions. However, as shown below, Glazebrook's sufficient conditions are not satisfied in our setting, and the index policy is sub-optimal. We will compare our optimal policy with that of Glazebrook (1979) below. Forand (2015) considers a revised multi-armed bandit problem where the decision maker must incur maintenance costs to keep an alternative available for choice and any positive signal is fully revealing. Forand finds that it may be optimal to first try the worst alternative to rule it out, and then continue learning the other alternative. In contrast, we consider a decision maker that receives gradual signals on an alternative, either positive or negative. If he receives enough negative signals on the worst alternative, the decision maker chooses the other alternative immediately, without learning more about it.

The literature on search theory is also related to the results presented here. Although the problem considered here is central to choices in a market environment, it is quite under-researched when all of its dimensions are included. For the simpler case in which all learnable information about an alternative can be learned in one search occasion, there is a large literature on optimal search and some of its market implications (e.g., McCall 1970; Diamond 1971; Weitzman 1979; Doval 2014). This literature, however, does not consider the possibility of gradual revelation of information throughout the search process. There is also some literature on gradual learning when a single alternative is considered or information is gathered to uncover a single uncertain value (e.g., Roberts and Weitzman 1981; Moscarini and Smith 2001; Branco et al. 2012; Fudenberg et al. 2015),[2] and the choice there is between adopting the alternative or not. In the face of more than one uncertain alternative (as is the case considered in this paper) the problem becomes more complicated. This is because opting for one alternative in a choice set means giving up potential high payoffs from other alternatives about which the decision maker has yet to learn more information. This paper can then be seen as combining these two literatures, with gradual search for information on multiple alternatives. Kuksov and Villas-Boas (2010) and Doval (2014) consider a search problem where everything that can be learned about one alternative is learned in one search occasion, but the decision maker can choose an alternative without search. This feature is also present here by including an outside option and allowing for adoption of an alternative without gaining full information about it. Doval, in particular, finds that in this case the decision maker may be indifferent to search an alternative that does not have the highest reservation price, which can be seen as related to the case in this paper of choosing to learn an alternative other than the best one under some conditions. Different from Doval, this paper considers gradual learning

[2] The case with a single alternative can be traced back to the discrete costly sequential sampling in Wald (1945). The continuous time treatment of the single alternative case was also presented in Dvoretzky et al. (1953), Mikhalevich (1958), and Shiryaev (1967).

of each alternative, and we find that it could be strictly optimal to learn an alternative that has both lower expected payoff and lower uncertainty.[3]

Ke et al. (2016) considers a stationary problem of gradual search for information with multiple products, where the precision of beliefs does not increase with search, and where earlier learning is not more informative than later learning. One important result in this paper is that it can be optimal to learn about an inferior alternative, which is never optimal in Ke et al. (2016). Another important difference is that the possible payoffs are bounded here, while they are unbounded in Ke et al. (2016).[4] The problem considered here also relates to the literature studying search while learning the payoff distribution (e.g., Rothschild 1974b; Adam 2001), but there what can be learned about each alternative is learned in one search occasion, so there is no gradual learning on each alternative.

The remainder of the paper is organized as follows. Section 2 presents the optimal learning problem. Section 3 solves the problem and presents the optimal learning strategy, and Section 4 looks at the implications of the results for adoption likelihood and probability of being correct. Section 5 presents several extensions of the main model, including asymmetric alternatives, the effect of the number of alternatives, and time discounting. Section 6 compares our optimal learning strategy with Weitzman (1979) and Glazebrook (1979), and presents concluding remarks. All technical proofs are presented in the Appendix.

2. Optimal learning problem

A decision maker (DM) is uncertain about the payoffs of two alternatives.[5] It is assumed that the payoff of alternative i, π_i, is either "high" as $\overline{\pi}_i$ or "low" as $\underline{\pi}_i$, with $\overline{\pi}_i > \underline{\pi}_i$ for $i = 1, 2$. The payoff π_i is independent across alternatives. Besides the two alternatives, there is an outside option with known deterministic payoff as π_0. The DM is assumed to be Bayesian and risk-neutral, or that the payoffs are in utils. To avoid trivialities, let us assume that $\pi_0 < \overline{\pi}_i$ for all i, so that no alternative i will be dominated by the outside option in all circumstances. The payoff π_0 could be lower than $\underline{\pi}_i$ for some alternative i, in which case, the outside option would be dominated by alternative i and thus be irrelevant. Before deciding which alternative to adopt, the decision maker can gather information sequentially on the two alternatives. Specifically, we consider a continuous-time setup. At any time point t, when the DM spends extra time dt in learning some information on alternative i, he pays cost $c_i dt$ and gets an informative signal on π_i. This signal does not reveal π_i completely; instead, it updates the DM's belief of the distribution of π_i according to Bayes' rule. The DM could buy another signal on the same alternative if he would like to learn more about it or, alternatively, he could buy signals on the other alternative. Signals are assumed to be i.i.d. conditional on the payoffs, for each alternative and across both alternatives. It is assumed that the flow learning cost c_i, once paid, is sunk.[6]

[3] In a concurrent paper, Che and Mierendorff (2016) consider which type of information to collect in a Poisson-type model, when the decision maker has to choose between two alternatives, with one and only one alternative having a high payoff. See also Hébert and Woodford (2017) for a rational inattention formulation.

[4] This paper also provides proof of existence and uniqueness of the optimal solution, and a more complete characterization of the stochastic process, which is not presented in Ke et al. (2016). With the possibility of a single observation process one can also consider the case of choosing one among several alternatives whose payoffs depend on the same state variable (see, for example, Décamps et al. 2006).

[5] The case of more than two alternatives is considered in Section 5.2.

[6] Another possible dimension, not explored here, is that the DM can also choose the intensity of effort in the learning process at each moment in time, with greater effort leading to a more precise signal (as in Moscarini and Smith 2001).

Consider a dynamic decision process with infinite time horizon. Let us define the DM's posterior belief of π_i being high at time $t \in [0, \infty)$ as $x_i(t)$.

$$x_i(t) = \Pr\left(\pi_i = \overline{\pi}_i | \mathscr{F}_t\right), \tag{1}$$

where $\{\mathscr{F}_t\}_{t=0}^{\infty}$ is a filtration that represents all the observed signals by time t. The set $\mathscr{F}_t$ depends on the DM's choice of alternatives for learning up to time t. Then we have that $x_i(0)$ is the DM's prior belief at time zero before obtaining any signal. For the sake of simplicity, we will just call $x_i(t)$ the *belief* of alternative i at time t below, and in the cases without confusion, we will drop the argument t and write it as x_i. The DM's conditional expected payoff from adopting alternative i at time t is then,

$$\mathrm{E}\left[\pi_i | \mathscr{F}_t\right] = x_i(t)\overline{\pi}_i + (1 - x_i(t))\, \underline{\pi}_i = \Delta\pi_i \cdot x_i(t) + \underline{\pi}_i, \tag{2}$$

where $\Delta\pi_i \equiv \overline{\pi}_i - \underline{\pi}_i > 0$. Below, we will use $\Delta\pi_i$ and $\underline{\pi}_i$ to replace $\overline{\pi}_i$ and $\underline{\pi}_i$ as our primary parameters. When we consider symmetric alternatives below, we will drop the subscript i and use $\Delta\pi$ and $\underline{\pi}$, respectively.

At any time t, the decision maker makes a choice from five actions: to learn alternative 1, to learn alternative 2, to adopt alternative 1, to adopt alternative 2, or to take the outside option. The decision process terminates when the DM adopts either one alternative or the outside option. He makes the choice of which action to take based on the information acquired so far, $\mathscr{F}_t$. Since π_i follows a two-point distribution, intuitively, all information about π_i can be summarized by one variable—the posterior belief x_i. Formally, it is straightforward to prove that $\mathbf{x}(t) \equiv (x_1(t), x_2(t))$ is a sufficient statistic for the observed history $\mathscr{F}_t$. Therefore, the DM's learning problem is essentially for each time t, to find a mapping from the belief space where $\mathbf{x}(t)$ lives in to the set of five actions, so as to maximize the expected payoff.

We can actually classify the five actions into two categories: to continue learning, and to stop learning and make an adoption decision. From this perspective, the decision maker's optimal learning problem can be equivalently viewed as dynamically deciding which alternative to learn at each time, and when to stop learning and make an adoption decision. In other words, the DM needs to solve the problem of *optimal allocation* (which alternative to learn) and *optimal stopping* (when to stop learning) at the same time, so as to maximize his expected payoff.

To formalize the DM's optimal learning problem, let us introduce the allocation policy as $I_t(\mathbf{x}(t)) \in \{1, 2\}$ that specifies which alternative to learn at time t, based on the DM's current posterior belief $\mathbf{x}(t)$. Formally speaking, $\mathscr{I} \equiv \{I_t(\mathbf{x}(t))\}_{t=0}^{\infty}$ is a progressively measurable process adapted with respect to the natural filtration of $\{\mathbf{x}(t)\}_{t=0}^{\infty}$ (see, e.g., Karatzas 1984, p. 174), where, as noted above, at each time t the action taken only depends on the DM's current posterior belief, $\mathbf{x}(t)$, a sufficient statistic for all the observed history. Let us also introduce τ as the stopping time adapted with respect to the natural filtration of $\{\mathbf{x}(t)\}_{t=0}^{\infty}$. Given any allocation policy $\mathscr{I}$ and stopping time τ, let us define $T_i(t)$ as the cumulated time that alternative i has been engaged in learning until time $t \le \tau$, or formally,

$$T_i(t) \equiv \mu\left(0 \le z \le t;\, I_z(\mathbf{x}(z)) = i\right), \tag{3}$$

where $\mu(\cdot)$ is the Lebesgue measure on $\mathbb{R}$. The observation process for the i-th alternative is assumed to follow the stochastic differential equation (SDE),

$$ds_i(t) = \pi_i dT_i(t) + \sigma_i dW_i\left(T_i(t)\right), \tag{4}$$

where $W_i(t)$ is a standard Brownian motion, and $\sigma_i^2 > 0$ specifies the noise level for signals of alternative i (see, for example, Roberts and Weitzman 1981; Karatzas 1984; Bolton and Harris

1999; Weeds 2002, for a similar formulation). As an analog to the discrete-time counterpart, the continuous-time signal $ds_i(t)$ consists of two parts: the "true value" $\pi_i dT_i(t)$, and a white noise term $\sigma_i dW_i(T_i(t))$. Under this framework, if the DM keeps learning information from one alternative continuously, his observation will be a Brownian motion with the drift as the (unknown) payoff.

The decision maker's expected payoff given his allocation policy $\mathscr{I}$ and stopping time τ can be written as the following expression:

$$J(\mathbf{x};\mathscr{I},\tau)=\mathrm{E}\left[\max\left\{\Delta\pi_1\cdot x_1(\tau)+\underline{\pi}_1,\Delta\pi_2\cdot x_2(\tau)+\underline{\pi}_2,\pi_0\right\}-\sum_{i=1}^{2}c_iT_i(\tau)\middle|\mathbf{x}(0)=\mathbf{x}\right]. \tag{5}$$

The allocation policy $\mathscr{I}$ not only enters into the cost term in (5) directly via $T_i(\tau)$, but also influences the belief updating process, and thus $\mathbf{x}(\tau)$. Starting from any belief $\mathbf{x}$, it is obvious that the DM's expected payoff is no greater than $\max_i \overline{\pi}_i$, so the supremum of the objective function in (5) exists. Therefore, we can write the DM's optimal learning problem as the following.

$$V(\mathbf{x})=\sup_{\mathscr{I},\tau} J(\mathbf{x};\mathscr{I},\tau), \tag{6}$$

where $V(\mathbf{x})$ is the so-called *value function*.

Now let us formulate the DM's optimal learning problem in (6) as a dynamic decision problem. Given the signal generation process in (1) and (4), by applying Theorem 9.1 from Liptser and Shiryaev (2001), we can write down the following SDE for the belief updating process,

$$\begin{aligned}dx_i(t)&=\frac{\Delta\pi_i}{\sigma_i^2}x_i(t)\left[1-x_i(t)\right]\left\{ds_i(t)-\mathrm{E}\left[\pi_i|\mathscr{F}_t\right]dT_i(t)\right\}\\&=\frac{\Delta\pi_i}{\sigma_i^2}x_i(t)\left[1-x_i(t)\right]\left\{\left[\pi_i-\underline{\pi}_i\left(1-x_i(t)\right)-\overline{\pi}_i x_i(t)\right]dT_i(t)+\sigma_i dW_i(T_i(t))\right\}\end{aligned} \tag{7}$$

for $i=1,2$, and given some initial beliefs at $t=0$. Equation (7) can be understood intuitively. The first multiplicative term on the right hand side, $\Delta\pi_i/\sigma_i^2$, is the signal-to-noise ratio for alternative i. The second term, $x_i(t)\left[1-x_i(t)\right]$, illustrates that the belief updates are more significant at $x_i(t)=50\%$ than at the extremes. The third term just states that the DM will update his belief upward if and only if the observation process rises faster than his current expectation. We can also see that $dx_i(t)$ does not depend on t explicitly due to time stationarity, and therefore we can drop the argument of t in the above equation. Note also that $\mathrm{E}[dx_i(t)|\mathscr{F}_t]=0$, and $\mathrm{E}[dx_i(t)^2|\mathscr{F}_t]=(\Delta\pi_i)^2/\sigma_i^2\cdot x_i(t)^2[1-x_i(t)]^2dt$. By applying Theorem 6.1 in Karatzas (1984), we have that there exists a weak, unique-in-distribution solution to the SDE in (7) for any policy function.

Applying the principle of optimality for dynamic programming, we can obtain the following *Hamilton–Jacobi–Bellman* (*HJB*) equation that is associated with the optimal learning problem in equation (6).

$$\max\left\{\max_{i=1,2}\left\{\frac{(\Delta\pi_i)^2}{2\sigma_i^2}x_i^2(1-x_i)^2\widehat{V}_{x_ix_i}(\mathbf{x})-c_i\right\},\right.\\ \left.\max\left\{\Delta\pi_1x_1+\underline{\pi}_1,\Delta\pi_2x_2+\underline{\pi}_2,\pi_0\right\}-\widehat{V}(\mathbf{x})\right\}=0, \tag{8}$$

where $\widehat{V}_{x_i x_i}(\mathbf{x})$ represents the second-order partial derivative of $\widehat{V}(\mathbf{x})$ with respect to x_i. Let us provide some intuition on equation (8). The first term in the max operator says that the DM's maximum expected payoff should be no less than the expected payoff from continuing learning one alternative. The second term of the max operator says that the DM's maximum expected payoff should be no less than the expected payoff from adopting the current best alternative. Finally, the outside max operator says that at any time, the DM will choose to either stop learning or continue learning. Note that from equation (8), all information about each alternative that is relevant for the solution is determined by $c_i\sigma_i^2$. We keep the two different variables in the presentation because of the different conceptual implications of the learning cost c_i and the variance of the noise of the signal σ_i^2.

It can be shown that the value function, as it is continuous, is a *viscosity solution* of the partial differential equation (8) (see Crandall et al. 1992 for an extensive description of this type of solutions; see, e.g., Corollary V.3.1 in Fleming and Soner 2010). In the following lemma, we prove that the viscosity solution of the partial differential equation (8) exists and is unique. Then, this implies that the viscosity solution of equation (8) is the value function. We will construct, under some parameter conditions, an allocation policy $\mathscr{I}^*$ and a stopping time policy τ^* that generates the solution $\widehat{V}(\mathbf{x})$ to the partial differential equation (8), with $\widehat{V}(\mathbf{x})$ bounded, continuous, and with continuous first derivatives. Both the allocation and stopping time policies can be represented by a set of boundaries in the belief space, where the DM chooses different actions across the boundaries. Under these policies $\widehat{V}(\mathbf{x}) = J(\mathbf{x}; \mathscr{I}^*, \tau^*)$, $\forall \mathbf{x}$, which means the optimality of $\mathscr{I}^*$ and τ^* and we obtain the value function $V(\mathbf{x}) = \widehat{V}(\mathbf{x})$, $\forall \mathbf{x}$. To simplify the presentation, in the following sections, we will use the notation $V(\mathbf{x})$ to represent both $V(\mathbf{x})$ and $\widehat{V}(\mathbf{x})$, as we concentrate on the case of optimality of the allocation and stopping time policies. In order to show that the viscosity solution to the *HJB* equation (8) exists and is unique, we show how the *HJB* equation (8) can be adjusted to satisfy *degenerate ellipticity*, and then apply the results in Crandall et al. (1992). We state next the existence and uniqueness result, and the proof is presented in the Appendix.

Lemma 1. *There exists a unique viscosity solution $\widehat{V}(\mathbf{x})$ to the HJB equation (8).*

In the next section, we present our main results on two symmetric alternatives. We will study the case with asymmetric alternatives, and with more than two alternatives in Section 5.

3. Optimal learning strategy

We consider two symmetric alternatives in this section with $\Delta\pi_i = \Delta\pi$, $\underline{\pi}_i = 0$, $\sigma_i = \sigma$, and $c_i = c$, where it is without loss of generality to set $\underline{\pi}_i = 0$ with symmetric alternatives. Define $x_0 \equiv \pi_0/(\Delta\pi)$, which measures the relative value of the outside option. To solve the DM's optimal learning problem, we will first propose a solution, and then verify that it satisfies equation (8). Then, by Lemma 1, it must be the only solution.

Single alternative To solve the problem with two alternatives, it is illustrative to first work out the simpler case with only one alternative. In this case, the decision maker is solving an optimal stopping problem—when to stop learning information and make an adoption decision. This problem is known as the classic Wald problem (Wald 1945), and in similar settings, has been studied by Roberts and Weitzman (1981) and Branco et al. (2012).[7] The optimal learning strategy is that there exist $0 \le \underline{x} < \overline{x} \le 1$ such that the DM continues learning when $\underline{x} \le x \le \overline{x}$, and he stops learning to take the outside option when $x < \underline{x}$ or to adopt the alternative when $x > \overline{x}$. In the Appendix, we show that $\underline{x}$ and $\overline{x}$ can be solved by

$$\left(\frac{1}{\overline{x}} + \frac{1}{1-\overline{x}}\right) - \left(\frac{1}{\underline{x}} + \frac{1}{1-\underline{x}}\right) = -\frac{(\Delta\pi)^2(\Delta\pi - 2\pi_0)}{2\sigma^2 c}, \tag{9}$$

$$\Phi\left(\overline{x}\right) - \Phi\left(\underline{x}\right) = -\frac{(\Delta\pi)^3}{2\sigma^2 c}, \tag{10}$$

where,

$$\Phi(x) \equiv 2\ln\left(\frac{1-x}{x}\right) + \frac{1}{x} - \frac{1}{1-x}, \tag{11}$$

which is a strictly decreasing function in $(0, 1)$. Therefore learning is preferred only when the DM's belief is in the middle range, i.e., when the DM is uncertain between the outside option and the alternative.

Next, we will proceed to analyze the optimal learning problem with two symmetric alternatives. We find that the solution structure critically depends on the value of the outside option. We will start with the case with a sufficiently low outside option. This is the case that has the surprising result that it may be better to learn about the worse alternative under some conditions. Then, we will consider the case with a sufficiently high outside option, and finally, we will present some results on the more complex case with an intermediate outside option.

Two alternatives with low-value outside option In this subsection, we consider the case where the value of the outside option is sufficiently low such that there is essentially no outside option, i.e., $x_0 \le 0$. We will propose a solution and verify that it satisfies the *HJB* equation (8).

First, we note that the two alternatives are symmetric in the belief space of x_1-x_2, so without loss of generality, we only need to consider the case with $x_1 \ge x_2$. Within the subspace of $x_1 \ge x_2$, we will further consider two cases. When $x_1 + x_2 \ge 1$, we propose that there exists a smooth function $\overline{X}(\cdot)$ such that the DM learns alternative 1 when $x_2 \le x_1 \le \overline{X}(x_2)$, and adopts alternative 1 when $x_1 > \overline{X}(x_2)$; on the other hand, when $x_1 + x_2 < 1$, we propose that there exists a smooth function $\underline{X}(\cdot)$ such that the DM learns alternative 2 when $x_1 \ge x_2 \ge \underline{X}(x_1)$, and adopts alternative 1 when $x_2 < \underline{X}(x_1)$. Intuitively, we call $x_1 = \overline{X}(x_2)$ for $x_1 + x_2 \ge 1$ and $x_2 = \underline{X}(x_1)$ for $x_1 + x_2 < 1$ the *adoption boundary* for alternative 1, and we have $0 \le \overline{X}(x), \underline{X}(x) \le 1$ for all x. Let us analyze the two cases one by one below.

In the first case, with $x_1 + x_2 \ge 1$, when $x_2 \le x_1 \le \overline{X}(x_2)$, as proposed, it is optimal for the DM to learn alternative 1, and therefore, the HJB equation (8) implies the following partial differential equation (PDE),

$$\frac{(\Delta\pi)^2}{2\sigma^2} x_1^2(1-x_1)^2 V_{x_1 x_1}(x_1, x_2) - c = 0. \tag{12}$$

Thanks to its simple parabolic form, (12) has the following general solution,

$$V(x_1, x_2) = \frac{2\sigma^2 c}{(\Delta\pi)^2}(1 - 2x_1)\ln\left(\frac{1-x_1}{x_1}\right)$$

[7] As noted above, see also Dvoretzky et al. (1953), Mikhalevich (1958), and Shiryaev (1967).

$$+ A_1(x_2)x_1 + A_2(x_2), \text{ if } x_1 + x_2 \geq 1 \text{ and } x_2 \leq x_1 \leq \overline{X}(x_2), \tag{13}$$

where $A_1(\cdot)$ and $A_2(\cdot)$ are two undetermined functions.[8] On the other hand, when $x_1 > \overline{X}(x_2)$, as proposed, it is optimal for the DM to adopt alternative 1, and therefore we have that,

$$V(x_1, x_2) = \Delta\pi \cdot x_1, \text{ if } x_1 + x_2 \geq 1 \text{ and } x_1 > \overline{X}(x_2). \tag{14}$$

In the second case, with $x_1 + x_2 < 1$, following a similar analysis as above, we have that,

$$V(x_1, x_2) = \begin{cases} \dfrac{2\sigma^2 c}{(\Delta\pi)^2}(1 - 2x_2)\ln\left(\dfrac{1 - x_2}{x_2}\right) & \\ +B_1(x_1)x_2 + B_2(x_1), & x_1 + x_2 < 1 \text{ and } x_1 \geq x_2 \geq \underline{X}(x_1) \\ \Delta\pi \cdot x_1 & x_1 + x_2 < 1 \text{ and } x_2 < \underline{X}(x_1) \end{cases}, \tag{15}$$

where $B_1(\cdot)$ and $B_2(\cdot)$ are two undetermined functions.

Thus far we have proposed a solution of $V(x_1, x_2)$ for each region in the belief space of x_1-x_2. Next, we consider all the boundary conditions connecting all the regions. It remains an open question to prove the necessity of the smoothness conditions for a general two-dimensional mixed optimal stopping and optimal control problem.[9] In this case, we use the smoothness conditions to construct a solution, and the uniqueness result of Lemma 1 implies that the value matching and smooth pasting conditions hold for our problem, since our solution is a viscosity solution of (8) and satisfies those conditions by construction. We follow Chernoff (1968) and Chapter 9.1 of Peskir and Shiryaev (2006) to provide in the Appendix an intuitive argument on the smooth pasting conditions based on a Taylor expansion.

First, we require $V(x_1, x_2)$, $V_{x_1}(x_1, x_2)$ and $V_{x_2}(x_1, x_2)$ to be continuous at the adoption boundary of alternative 1. Let us consider the case with $x_1 + x_2 \geq 1$ first. By substituting (13) and (14) into these three conditions, we find that one of them is redundant (see also Mandelbaum et al. 1990, p. 1016, for a similar result), and the three conditions are equivalent to the following two equations,

$$A_1(x) = \frac{4\sigma^2 c}{(\Delta\pi)^2}\ln\left(\frac{1 - \overline{X}(x)}{\overline{X}(x)}\right) + \frac{2\sigma^2 c}{(\Delta\pi)^2}\frac{1 - 2\overline{X}(x)}{\overline{X}(x)\left(1 - \overline{X}(x)\right)} + \Delta\pi, \tag{16}$$

$$A_2(x) = -\frac{2\sigma^2 c}{(\Delta\pi)^2}\ln\left(\frac{1 - \overline{X}(x)}{\overline{X}(x)}\right) - \frac{2\sigma^2 c}{(\Delta\pi)^2}\frac{1 - 2\overline{X}(x)}{1 - \overline{X}(x)}. \tag{17}$$

The redundancy of one condition is not a mere coincidence. Instead, it is due to the learning problem structure that the DM can learn only one alternative at a time. Consequently, only one

[8] When $x_1 = 1$, the solution to the HJB equation is not well defined, but the optimal learning problem is trivial. Under $x_1 = 1$, the payoff of alternative 1 is known to be high, therefore, the DM's optimal strategy is to adopt alternative 1 immediately, and $V(x_1, x_2) = \overline{\pi}$. Similarly we can solve the case where $x_2 = 1$. For the sake of convenience, we only consider the case that $x_1 < 1$ and $x_2 < 1$ below. Later, we will show that $V(x_1, x_2)$ is continuous at $x_i = 1$.

[9] See Strulovici and Szydlowski (2015) for recent progress on smoothness conditions for one-dimensional mixed optimal stopping and optimal control problems. Peskir and Shiryaev (2006), p. 144, note that for optimal stopping problems if the boundary is Lipschitz-continuous, then the smooth pasting conditions will hold at the optimum. Caffarelli (1977), Theorem 2, shows that if the payoff for stopping is continuously differentiable, which occurs in this case, then the boundary is Lipschitz-continuous. Caffarelli (1977), Theorem 3, also shows that if the payoff of stopping is continuously differentiable, then the boundary is continuously differentiable. In our case, that means that $\overline{X}(x_2)$ is continuously differentiable.

Brownian motion can move at a time and we have a parabolic PDE in (12).[10] In the Appendix, we show that, due to the parabolic nature of the PDE, we will always end up with one redundant condition.

By substituting the expressions of $A_1(x)$ and $A_2(x)$ in (16) and (17) into (13), we have that,

$$\begin{aligned} V(x_1,x_2) = & \frac{2\sigma^2 c}{(\Delta\pi)^2}(1-2x_1)\left[\ln\left(\frac{1-x_1}{x_1}\right) - \ln\left(\frac{1-\overline{X}(x_2)}{\overline{X}(x_2)}\right)\right] \\ & - \frac{2\sigma^2 c}{(\Delta\pi)^2}\frac{1-2\overline{X}(x_2)}{1-\overline{X}(x_2)}\frac{\overline{X}(x_2)-x_1}{\overline{X}(x_2)} \\ & + \Delta\pi\cdot x_1, \text{ if } x_1+x_2\geq 1 \text{ and } x_2\leq x_1\leq \overline{X}(x_2). \end{aligned} \tag{18}$$

Now let us consider the other case, with $x_1+x_2<1$. Similarly, by utilizing the value match and smooth pasting conditions at the adoption boundary of $x_2=\underline{X}(x_1)$, we can solve $B_1(x)$ and $B_2(x)$, and rewrite $V(x_1,x_2)$ in (15) as the following:

$$V(x_1,x_2) = \begin{cases} \frac{2\sigma^2 c}{(\Delta\pi)^2}(1-2x_2)\left[\ln\left(\frac{1-x_2}{x_2}\right) - \ln\left(\frac{1-\underline{X}(x_1)}{\underline{X}(x_1)}\right)\right] \\ -\frac{2\sigma^2 c}{(\Delta\pi)^2}\frac{1-2\underline{X}(x_1)}{1-\underline{X}(x_1)}\frac{\underline{X}(x_1)-x_2}{\underline{X}(x_1)} + \Delta\pi\cdot x_1, \\ \quad \text{if } x_1+x_2<1 \text{ and } x_1\geq x_2\geq \underline{X}(x_1) \\ \Delta\pi\cdot x_1 \quad \text{if } x_1+x_2<1 \text{ and } x_2<\underline{X}(x_1) \end{cases} \tag{19}$$

Second, we require $V(x_1,x_2)$, $V_{x_1}(x_1,x_2)$ and $V_{x_2}(x_1,x_2)$ to be continuous at the boundary $x_1+x_2=1$. Similarly, we find that one of the three boundary conditions are redundant. Based on the expressions of $V(x_1,x_2)$ in (18) and (19), the continuity of $V(x_1,x_2)$ (value matching condition) implies that,

$$\underline{X}(x) = 1-\overline{X}(1-x). \tag{20}$$

This implies that the adoption boundary of alternative 1 is symmetric with respect to the line of $x_1+x_2=1$. In the following, we will substitute the expression of $\underline{X}(\cdot)$ in (20) into (19), so that we will work with $\overline{X}(\cdot)$ only. Based on the expressions of $V(x_1,x_2)$ in (18) and (19), the continuity of $V_{x_1}(x_1,x_2)$ (smooth pasting condition) implies that $\overline{X}(x)$ satisfies the following ordinary differential equation (ODE),

$$\overline{X}'(x) = \frac{\Phi\left(\overline{X}(x)\right)+\Phi(x)}{\Phi'\left(\overline{X}(x)\right)\left(\overline{X}(x)+x-1\right)}, \text{ for } x\leq\frac{1}{2}, \tag{21}$$

[10] We could generalize our learning problem to allow the DM at any time, to allocate a fraction of his effort to learn alternative 1 and the complementary fraction of the effort to learn alternative 2. If we impose a capacity constraint on the total number of "signals" per unit of time from the two alternatives, we can show that the optimal learning strategy will be exactly the same as that in our main model, i.e., the DM will *optimally* allocate his effort to one alternative at a time. That is to say, sequential learning is optimal. This is because we have a continuous-time learning framework. On the other hand, if we do not impose this capacity constraint on the number of total signals per time, depending on the model setting, we may end up with the case where it is optimal for the DM to learn multiple alternatives simultaneously. This would correspond to an elliptic PDE instead of the parabolic PDE that occurs here.

where $\Phi(\cdot)$ has been defined in (11). The ODE above only applies to the region with $x \le 1/2$, because in writing down (21) in terms of $\overline{X}(x_2)$, we consider the boundary of $x_1 + x_2 = 1$ under the condition that $x_1 \ge x_2$.

Third, we require $V(x_1, x_2)$, $V_{x_1}(x_1, x_2)$ and $V_{x_2}(x_1, x_2)$ to be continuous at the boundary $x_1 = x_2$. Due to the symmetry of $V(x_1, x_2)$ in x_1-x_2 space, it is not difficult to recognize that the only necessary boundary condition is the continuity of $V_{x_1}(x_1, x_2)$. Due to the symmetry with respect to the line of $x_1 + x_2 = 1$ implied by (20), we only need to consider the case with $x_1 + x_2 \ge 1$. Based on (18), continuity of $V_{x_1}(x_1, x_2)$ (smooth pasting condition) implies the following ODE,

$$\overline{X}'(x) = -\frac{\Phi\left(\overline{X}(x)\right) - \Phi(x) + (\Delta\pi)^3/(2\sigma^2 c)}{\Phi'\left(\overline{X}(x)\right)\left(\overline{X}(x) - x\right)}, \text{ for } x \ge \frac{1}{2}, \tag{22}$$

where the ODE only applies to the region with $x \ge 1/2$, because in writing down (22) in terms of $\overline{X}(x_2)$, we consider the boundary of $x_1 = x_2$ under the condition that $x_1 + x_2 \ge 1$.

Lastly, we require $V(x_1, x_2)$, $V_{x_1}(x_1, x_2)$ and $V_{x_2}(x_1, x_2)$ to be continuous at the boundary point of $x_1 = x_2 = 1/2$. In fact, recall that, we have proposed the expressions of $V(x_1, x_2)$ for the four regions divided by the two lines of $x_1 = x_2$ and $x_1 + x_2 = 1$ separately. We then paste the four regions together by continuity conditions at the boundary lines of $x_1 = x_2$ and $x_1 + x_2 = 1$. The point $x_1 = x_2 = 1/2$ is the intersection of the two boundary lines $x_1 = x_2$ and $x_1 + x_2 = 1$. Only at this point, the region of $\{x_1 \ge x_2 \text{ and } x_1 + x_2 \ge 1\}$ touches the region of $\{x_1 \le x_2 \text{ and } x_1 + x_2 \le 1\}$ (and similarly, the region of $\{x_1 \ge x_2 \text{ and } x_1 + x_2 \le 1\}$ touches the region of $\{x_1 \le x_2 \text{ and } x_1 + x_2 \ge 1\}$). Basically, at the point of $x_1 = x_2 = 1/2$, we require $V(x_1, x_2)$ to be smooth across the two regions of $\{x_1 \ge x_2 \text{ and } x_1 + x_2 \ge 1\}$ and $\{x_1 \le x_2 \text{ and } x_1 + x_2 \le 1\}$. Due to symmetry, it is straightforward to show that all but one condition is redundant, and the continuity of $V_{x_1}(x_1, x_2)$ at the boundary point of $x_1 = x_2 = 1/2$ implies that,

$$\overline{X}\left(\frac{1}{2}\right) = \Phi^{-1}\left(-\frac{(\Delta\pi)^3}{4\sigma^2 c}\right). \tag{23}$$

To summarize, $\overline{X}(x)$ can be determined by combining (21), (22) and (23) as two boundary value problems. We are unable to solve the ODEs analytically, but we prove the existence and uniqueness of the solution, as well as some useful properties by the following lemma.

Lemma 2. *There exists a unique continuous function $\overline{X}(x)$ in $(x^*, 1)$, with $x^* < 1/2$, that solves the two boundary value problems in (21), (22) and (23), where x^* is defined as the solution to $\overline{X}(x^{*+}) = 1 - x^*$. Let $\overline{X}(x^*) = \overline{X}(x^{*+})$. $\overline{X}(x)$ is smooth, increasing, and $\overline{X}(x) \ge x$ with $\overline{X}(1^-) = 1$.*

Recall that $\overline{X}(x)$ was defined for $x_1 + x_2 \ge 1$ which translates to $x \in (x^*, 1)$. The adoption boundary across the regions $x_1 + x_2 \ge 1$ and $x_1 + x_2 < 1$ is defined in (26) below.

Thus far we have considered all the boundary conditions which, together with the expressions of $V(x_1, x_2)$ given by (14), (18) and (19) entirely characterize the solution to (8). Lastly, to verify this is indeed the solution, we need to guarantee that the DM prefers to learn alternative 1 when $x_1 + x_2 \ge 1$ and $x_2 \le x_1 \le \overline{X}(x_2)$, and he prefers to learn alternative 2 when $x_1 + x_2 < 1$ and $x_1 \ge x_2 \ge \underline{X}(x_1)$. By (8) and (12), we require that,

$$x_2^2(1 - x_2)^2 V_{x_2 x_2}(x_1, x_2) \le \frac{2\sigma^2 c}{(\Delta\pi)^2}, \text{ for } \forall x_1, x_2, \text{ s.t. } x_1 + x_2 \ge 1 \text{ and } x_2 \le x_1 \le \overline{X}(x_2), \tag{24}$$

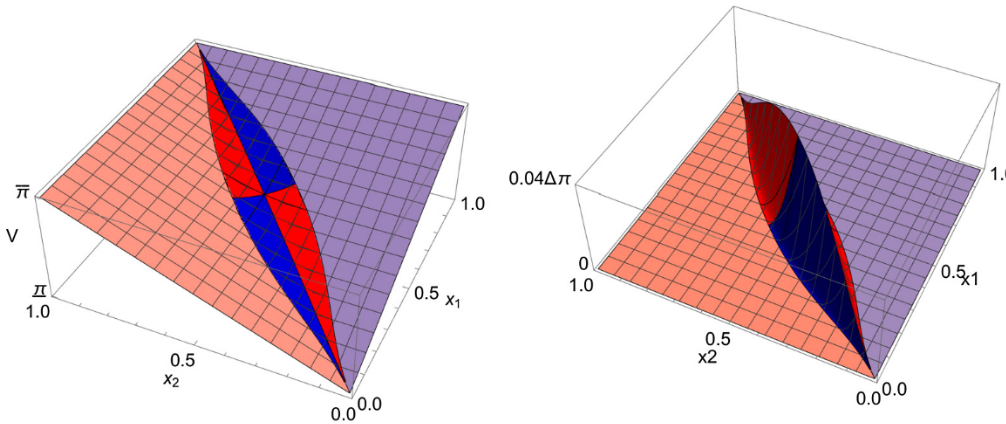

Fig. 1. Maximum expected payoff (left panel), payoff from learning (right panel), given a DM's current belief, with relatively low value of outside option, $x_0 \leq 0$. $c\sigma^2/\Delta\pi = 0.1$. Both panels are color coded in the following way: blue represents the area where it is optimal to learn alternative 1, red for learning alternative 2, light blue for adopting alternative 1, and light red for adopting alternative 2. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

$$x_1^2(1-x_1)^2 V_{x_1x_1}(x_1,x_2) \leq \frac{2\sigma^2 c}{(\Delta\pi)^2}, \text{ for } \forall x_1, x_2, \text{ s.t. } x_1+x_2<1 \text{ and } x_1 \geq x_2 \geq \underline{X}(x_1). \tag{25}$$

By (20), it is easy to show that the condition in (25) is equivalent to that in (24). Therefore, we only need to verify that (24) is valid. We prove this in the Appendix. We provide more intuition on equation (24) below, when we present the DM's optimal learning strategy. Finally, to summarize everything so far for this subsection on the value function, Theorem A1 in the Appendix presents the value function for this case when the outside option is relatively low, i.e., $x_0 \leq 0$ (the optimal learning strategy is characterized in Theorem 1 below). The existence and uniqueness of the solution has been established in Lemma 1 and, based on the way we construct the solution, we have already verified that it satisfies (8) except for lines, of measure zero, where the second derivatives may not exist. As the second derivative exists almost everywhere, and we have smooth pasting, we can apply Itô's Lemma, and verify that the function obtained is indeed the value function (along the lines of Theorem 5.7 in Touzi 2010).[11]

Fig. 1 shows the value function $V(x_1, x_2)$, as well as the payoff from learning $L(x_1, x_2)$ under some parameter setting. The payoff from learning is defined as $L(x_1, x_2) \equiv V(x_1, x_2) - \Delta\pi \max\{x_1, x_2, x_0\}$, which is basically the difference between a DM's value function and his expected payoff in the case that learning is not allowed. As the figure illustrates $V(x_1, x_2)$ is increasing in x_1 and x_2 and convex, and $L(x_1, x_2)$ is always non-negative. $L(x_1, x_2)$ peaks at $x_1 = x_2 = 1/2$, at which point the DM is most uncertain between adopting alternative i and taking the outside option. This is the point where the DM benefits most from learning.

Based on the value function given by Theorem A1, we can construct the optimal learning strategy that achieves the value function, and the corresponding belief updating process under the optimal strategy.

[11] Note also that the second derivative is continuous at the points where the DM is learning and switches between which alternative to learn about.

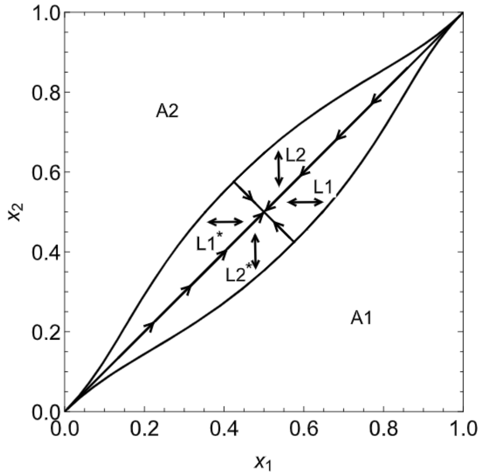

Fig. 2. Optimal learning strategy with relatively low value of the outside option, $x_0 \leq 0$, and $c\sigma^2/\Delta\pi = 0.1$. It is optimal for the decision maker to adopt alternative i in region Ai, and to learn alternative i in regions Li and Li*, for $i = 1, 2$.

To simplify notation, we can unify the adoption boundary for alternative i in two regions of $x_1 + x_1 \geq 1$ and $x_1 + x_2 < 1$ by defining

$$\overline{Y}(x) \equiv \begin{cases} \overline{X}(x), & x \in [x^*, 1] \\ 1 - \overline{X}^{-1}(1-x), & x \in [0, x^*] \end{cases}, \tag{26}$$

and the following theorem formalizes the description of the optimal policy.

Theorem 1. *Consider the decision maker's optimal learning problem in (6) with two symmetric alternatives, and suppose that the value of the outside option is relatively low, i.e.,* $x_0 \leq 0$. *Then there is a function* $\overline{Y}(x)$ *with* $\overline{Y}'(x) > 0$, $\overline{Y}(x) > x$ *for all* $x \in (0, 1)$, $\overline{Y}(0) = 0$, *and* $\overline{Y}(1) = 1$, *such that: (1) the DM adopts alternative i if* $x_i \geq \overline{Y}(x_{3-i})$ *and continues learning if that condition does not hold for any alternative; (2) When learning, the DM learns about the alternative with the higher belief if* $x_1 + x_2 \geq 1$, *and learns about the alternative with the lower belief if* $x_1 + x_2 < 1$.

Fig. 2 shows the optimal learning strategy under some parameter setting. The belief space is divided into six regions by black lines. A1 and A2 stand for the regions, where it is optimal to adopt alternative 1 and 2, respectively. Both L1 and L1* stand for the regions where it is optimal to learn alternative 1; similarly, both L2 and L2* stand for the regions where it is optimal to learn alternative 2. As noted above, it is optimal for the DM to learn the worse alternative when $x_1 + x_2 \leq 1$, as shown in the starred regions (L1* and L2*). If this alternative turns out to be really poor, the DM will adopt the other alternative immediately; otherwise, if this alternative turns out to be not as poor as thought earlier, the DM will switch to learn the other alternative.

It is interesting to understand why sometimes it is optimal for a DM to prioritize learning on the worse alternative. Let us consider a DM with belief in L2* in Fig. 2. One may suspect that while alternative 2 has a lower expected payoff than alternative 1, the DM may be more uncertain about alternative 2 and thus prefer to learn alternative 2 first. This speculation turns out to be incorrect. In fact, given (x_1, x_2) in L2*, we have $x_1 > x_2$ and $x_1 + x_2 \leq 1$, which implies that

$$\mathrm{E}[\pi_1] = \Delta\pi x_1 > \Delta\pi x_2 = \mathrm{E}[\pi_2], \tag{27}$$

$$\mathrm{Var}[\pi_1] = (\Delta\pi)^2 x_1(1-x_1) \geq (\Delta\pi)^2 x_2(1-x_2) = \mathrm{Var}[\pi_2]. \tag{28}$$

Therefore, the DM not only expects a higher payoff from alternative 1 but also is more uncertain about the payoff of alternative 1.

It turns out that the reason for the DM to learn a worse and more certain alternative first is to save expected learning costs. Here is the intuition. Given the DM's belief (x_1, x_2) in region L2*, if he learns alternative 2, he will either end the learning process by adopting alternative 1 if he accumulates enough bad news on alternative 2, or continue to learn alternative 1 if he accumulates enough good news on alternative 2. As the DM aims to save learning costs (the outside option is dominated, by assumption), he would like to end the learning process as soon as possible, so he cares about the probability of getting bad news on alternative 2, which is equal to $1 - x_2$. On the other hand, if the DM instead chooses to learn alternative 1, he will either end the learning process by adopting alternative 1 if he accumulates enough good news on alternative 1, or continue to learn alternative 2 if he accumulates enough bad news on alternative 1. With the aim of saving learning costs, the DM would like to end the learning process as soon as possible, so he cares about the probability of getting good news on alternative 1, which is equal to x_1. To minimize future learning costs, the DM chooses to learn alternative 2 if $1 - x_2 \geq x_1$, i.e., $x_1 + x_2 \leq 1$, which is exactly the upper-right boundary condition for L2*.[12] To summarize, in the starred regions, the DM finds it optimal to learn the worse alternative first with the possibility of ruling it out early so as to choose the other alternative immediately.

We now consider the stochastic dynamics. We use region L1 as an example for illustration, and other regions can be understood similarly. Suppose that the process starts in region L1, with $x_1 > x_2 > 1/2$. In this region, the optimal alternative to learn on is alternative 1, and therefore, the stochastic process moves horizontally. If the process moves sufficiently to the right, region A1 is hit and the DM stops the learning process and chooses alternative 1. On the other hand, if the process moves sufficiently to the left, it hits the line $x_1 = x_2$ and the DM stops gathering information on alternative 1 and starts gathering information on alternative 2. There are two cases to consider.

First, if the DM gets positive news on alternative 2, we are now in region L2, with the stochastic process moving vertically, and if the process reaches region A2 then the DM stops the learning process and chooses alternative 2.

Second, the DM receives bad news about alternative 2. To get an intuition of what happens in this situation, let us consider a discrete approximation of the problem. Starting in region L1 close to $x_1 = x_2$, if the DM learns bad news about alternative 1, the process moves horizontally to the left of $x_1 = x_2$. Then the DM checks alternative 2, and if receiving bad news on alternative 2, the process moves vertically below the line of $x_1 = x_2$ and the DM checks on alternative 1 again at a level of x_2 that is lower than the original level. Therefore, when the DM keeps getting bad news on the two alternatives, he will move down the line of $x_1 = x_2$. In other words, $\min\{x_1, x_2\}$ falls over time with positive probability, and it never increases, as noted by the arrows in Fig. 2. To see this more formally in our continuous-time diffusion process, note that $\min\{x_1, x_2\}$ can be written as $\min\{x_1, x_2\} = \frac{1}{2}(x_1 + x_2 - |x_1 - x_2|)$. The existence of the absolute-value term $|x_1 - x_2|$ requires the use of the Tanaka's formula (e.g., Karatzas and Shreve 1991, p. 205) to obtain the local time at $x_1 = x_2$, which then leads to the result that $\min\{x_1, x_2\}$ falls over time,

[12] We have also analyzed a two-period discrete time learning problem, and find that the intuition of saving learning costs here becomes the exact condition there. The analysis is available upon request.

and it never increases. In the literature this is known as a case of switching Brownian motions (e.g., Mandelbaum et al. 1990). The Appendix presents further details.

Now consider the other region in L1 with $x_1 > 1/2 > x_2$. In this region the process moves horizontally, and if it moves to the right, it hits region A1 where the process is stopped and alternative 1 is chosen; and if it moves to the left, it hits the line of $x_1 + x_2 = 1$ and reaches region L2*, where the DM switches to learning about the worse alternative, alternative 2. In that case, the process starts moving vertically, and if there are enough bad news about alternative 2, the process hits region A1, and alternative 1 is chosen. If there are good news on alternative 2, the process moves back to the line $x_1 + x_2 = 1$, and alternative 1 gets learnt again. By the same local time argument as above, each time the process hits $x_1 + x_2 = 1$, which occurs infinitely many times given the Brownian motion processes, the process moves up along the line of $x_1 + x_2 = 1$ towards the point $x_1 = x_2 = 1/2$, as noted by the arrows in Fig. 2.

Note that in all regions, if it is not optimal for the DM to stop the learning process and choose one of the alternatives, the process keeps going toward $x_1 = x_2 = 1/2$. That is, if the DM receives signals that do not allow for a sufficient separation of the beliefs regarding the two alternatives, both beliefs converge to the maximum uncertainty. This implies that the longer the DM is searching for information, the more uncertain the DM becomes due to a selection effect—those who become certain enough stop the learning process and make an adoption. Once the process is at $x_1 = x_2 = 1/2$, the DM is indifferent about which alternative to learn information on, but he will continue learning, and therefore $x_1 = x_2 = 1/2$ is not an absorbing state.

Lastly, we can compute the comparative statics of the optimal learning strategy with respect to the learning costs, noise of the signals, and value of the outside option. The results are intuitive.

Proposition 1. *As $c\sigma^2$ increases, the adoption boundary $\overline{Y}(x)$ decreases, and the value function $V(x_1, x_2)$ decreases.*

Two alternatives with high-value outside option In this subsection, we consider the case of two symmetric alternatives with the value of the outside option sufficiently high. Similarly, we will propose a solution, and then verify that it satisfies the *HJB* equation (8). After that, we will come back and determine the exact threshold for the outside option to be considered sufficiently high. Similarly, due to symmetry, we only need to consider the case with $x_1 \geq x_2$, where there are further two cases to consider.

In the first case, when $x_2 \leq \underline{x}$, we propose that alternative 2 will not be considered for either learning or adoption, and one goes back to the case with a single uncertain alternative—alternative 1. Correspondingly, we have

$$V(x_1, x_2) = U(x_1), \text{ if } x_2 \leq \underline{x}, \tag{29}$$

where $U(x)$, given by equation (vi) in the Appendix, is the value function for the optimal learning problem in the single-alternative case. In the second case, when $x_1 \geq x_2 > \underline{x}$, we propose that there exists a smooth function $\overline{X}(\cdot)$ such that the DM learns alternative 1 when $x_2 \leq x_1 \leq \overline{X}(x_2)$, and adopts alternative 1 when $x_1 > \overline{X}(x_2)$. We have used $\overline{X}(x_2)$ in the case of a low-value outside option above to denote the adoption boundary for alternative 1. Here we have slightly abused the notation of $\overline{X}(x_2)$, which also denotes the adoption boundary for alternative 1 in the case of a high-value outside option. By applying a similar analysis as above, we can determine $V(x_1, x_2)$ as the following by solving the partial differential equation (12) subject to the value matching and smooth pasting conditions at the adoption boundary of $x_1 = \overline{X}(x_2)$.

$$V(x_1,x_2) = \frac{2\sigma^2 c}{(\Delta\pi)^2}(1-2x_1)\left[\ln\left(\frac{1-x_1}{x_1}\right) - \ln\left(\frac{1-\overline{X}(x_2)}{\overline{X}(x_2)}\right)\right] - \frac{2\sigma^2 c}{(\Delta\pi)^2}\frac{1-2\overline{X}(x_2)}{1-\overline{X}(x_2)}\frac{\overline{X}(x_2)-x_1}{\overline{X}(x_2)} + \Delta\pi\cdot x_1 \text{ if } x_2 \le x_1 \le \overline{X}(x_2), \tag{30}$$

where $\overline{X}(x)$ is determined by the following boundary value problem of an ODE:

$$\overline{X}'(x) = -\frac{\Phi\left(\overline{X}(x)\right) - \Phi(x) + (\Delta\pi)^3/(2\sigma^2 c)}{\Phi'\left(\overline{X}(x)\right)\left(\overline{X}(x)-x\right)}, \tag{31}$$

$$\overline{X}(\underline{x}) = \overline{x}, \tag{32}$$

where $\underline{x}$ and $\overline{x}$ are defined by (9) and (10), and $\Phi(\cdot)$ is defined in (11). Equations (31) and (32) result from the value matching and smooth pasting conditions at $x_1 = x_2$ and $x_2 = \underline{x}$ respectively. We are unable to solve the ODE analytically, but we can similarly prove its existence and uniqueness, and some useful properties of the solution in the following lemma. (The proof is similar to the one of Lemma 2 and thus is omitted.)

Lemma 3. *There exists a unique continuous function $\overline{X}(x)$ in $[\underline{x},1]$ that solves the boundary value problem of an ODE in (31) and (32). $\overline{X}(x)$ is smooth, increasing, and $\overline{X}(x) \ge x$ with $\overline{X}(1^-) = 1$.*

Thus far we have considered all the boundary conditions which, together with the expressions of $V(x_1,x_2)$ given by (29) and (30), entirely characterize the solution to (8). Lastly, to verify this is indeed the solution, we need to guarantee that the DM prefers to learn alternative 1 for $\underline{x} \le x_2 \le x_1 \le \overline{X}(x_2)$. By (8) and (12), we require that

$$x_2^2(1-x_2)^2 V_{x_2x_2}(x_1,x_2) \le \frac{2\sigma^2 c}{(\Delta\pi)^2}, \text{ for } \forall x_1,x_2, \text{ s.t. } \underline{x} \le x_2 \le x_1 \le \overline{X}(x_2). \tag{33}$$

Without solving $\overline{X}(x)$, it is difficult to translate the inequality above into expressions in terms of model primitives. Nevertheless, we are able to prove the following necessary and sufficient condition for (33).

Lemma 4. *The DM prefers learning alternative 1 for $\underline{x} \le x_2 \le x_1 \le \overline{X}(x_2)$, if and only if*

$$\overline{X}'(x_2) \le \frac{\overline{X}(x_2)\left(1-\overline{X}(x_2)\right)}{x_2(1-x_2)}, \textit{ for } \forall x_2 \in [\underline{x},1). \tag{34}$$

There exists a $x_0^* < 1/2$, such that this condition holds if and only if $x_0 \equiv \pi_0/(\Delta\pi) \ge x_0^*$.

The variable x_0 measures the relative value of the outside option compared with the uncertain alternatives. This lemma implies that the solution that we just proposed exists only when the outside option is sufficiently high.

Finally, to summarize everything so far for this subsection, we have Theorem A2, presented in the Appendix, which fully characterizes the value function for the case of a high outside option (the optimal learning strategy is characterized in Theorem 2 below). The existence and uniqueness of the solution has been established in Lemma 1 and, based on the way we construct the solution, we have already verified that it satisfies (8). In order to define the adoption boundary

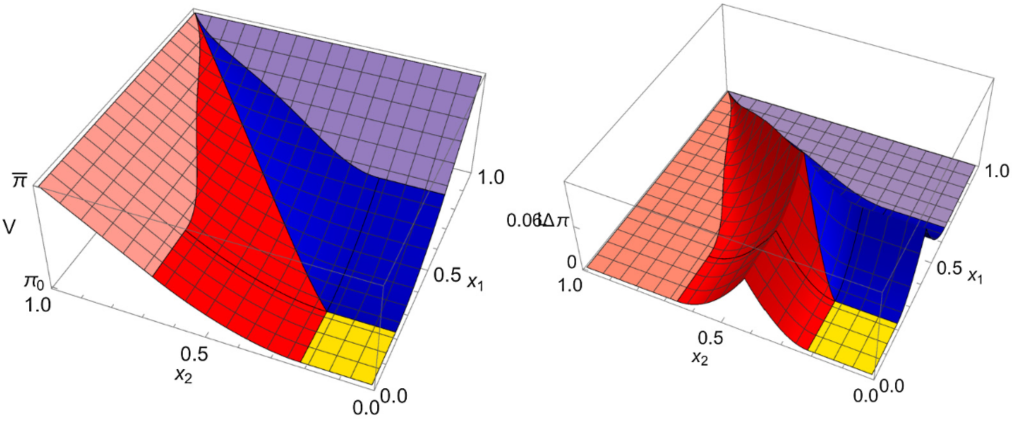

Fig. 3. Maximum expected payoff (left panel), payoff from learning (right panel), given a DM's current belief, with relatively high value of outside option, $x_0 = 0.5$. $c\sigma^2/\Delta\pi = 0.1$. Both panels are color coded in the following way: blue represents the area where it is optimal to learn alternative 1, red for learning alternative 2, light blue for adopting alternative 1, light red for adopting alternative 2, and yellow for adopting the outside option.

and write down the value function in a uniform way, we expand the support of $\overline{X}(\cdot)$ to [0, 1], by defining $\overline{X}(x) \equiv \overline{x}$ for $x \in [0, \underline{x})$ and $\overline{X}(1) \equiv 1$. It is easy to verify that $\overline{X}(x)$ is still smooth, increasing, and $\overline{X}(x) \geq x$ for $x \in [0, 1]$.

Because $x_0^* < 1/2$ and $x_0 = \pi_0/\Delta\pi$, we have that a sufficient condition for the function defined by (xix) in Theorem A2 to be the value function is that $\pi_0 \geq \overline{\pi}/2$.

Fig. 3 shows the value function $V(x_1, x_2)$, as well as the payoff from learning $L(x_1, x_2)$ under some parameter setting. This illustrates how $L(x_1, x_2)$ peaks at $x_1 = x_2 = x_0$, at which point the DM is most uncertain between adopting alternative i and taking the outside option.

Based on the value function given by Theorem A2, we can construct the optimal learning strategy that achieves the value function. Formally, the following theorem summarizes the DM's optimal learning strategy.

Theorem 2. *When the value of the outside option is relatively high, i.e., $x_0 \geq x_0^*$, a decision maker considers alternative i for learning or adoption if and only if his belief of i is sufficiently high, i.e., $x_i \geq \underline{x}$. Given one alternative i under his consideration, the DM keeps learning it if $\underline{x} \leq x_i \leq \overline{x}$, adopts it if $x_i > \overline{x}$, and takes the outside option if $x_i < \underline{x}$. Given two alternatives under his consideration, the DM always learns information from the alternative with higher belief.*[13] *He stops learning and adopts it if and only if his belief of this alternative is above a threshold. This threshold (i.e., the adoption boundary) increases with the DM's belief of the other alternative, and gets closer to the line $x_1 = x_2$ as x_2 increases, with the threshold converging to x_2 as $x_2 \to 1$.*

Fig. 4 shows the optimal learning strategy under some parameter setting, which can be understood intuitively. There are five regions. When the DM's belief of alternative i is quite high, it is optimal to adopt the alternative immediately without learning (regions A1 and A2 in the figure).

[13] In the case of equal beliefs, the DM is indifferent about which alternative to search, and we assume that the DM learns alternative 1.

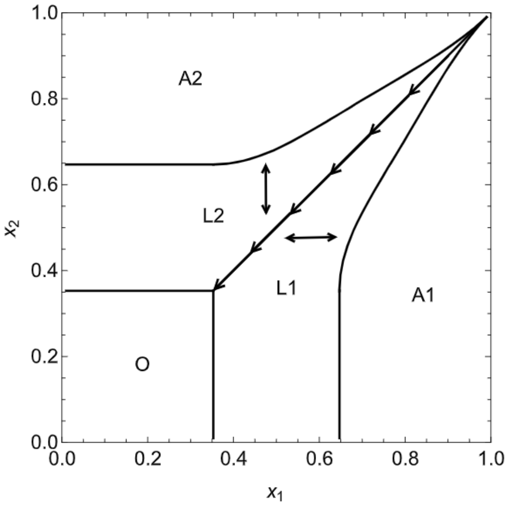

Fig. 4. Optimal learning strategy with relatively high value of the outside option, $x_0 = 0.5$, and $c\sigma^2/\Delta\pi = 0.1$. It is optimal for the decision maker to adopt alternative i in region Ai, and to learn alternative i in region Li, for $i = 1, 2$.

When the DM's beliefs of both alternatives are quite low, it is optimal to take the outside option (region O in the figure). Roughly speaking, learning becomes preferred when the DM's beliefs of the two alternatives are similar or take medium values (regions L1 and L2 in the figure).

Note that when $x_i < \underline{x}$, the DM only focuses on the comparison between the outside option and the other alternative $j \neq i$. He decides only among learning j, adopting j, and taking the outside option ($j \neq i$), and he never considers alternative i for learning or adoption. This means the DM optimally uses a simple *cutoff strategy* when forming a *consideration set* of alternatives for learning or adoption. Given both alternatives in his consideration set, i.e., $x_1, x_2 \geq \underline{x}$, the DM always prioritizes learning about the alternative with higher belief, and adopts it when the belief is sufficiently high. As shown above, this learning strategy critically depends on the assumption that the value of the outside option is sufficiently high.

Similarly, we can construct the belief updating process under the optimal learning strategy, which completes the analysis to show that the value function $V(x_1, x_2)$ given by Theorem A2 is attainable. We can show that during the learning process, $\min\{x_1, x_2\}$ falls over time with positive probability and that it never increases. This means that starting from a prior belief in the regions that it is optimal to learn, there is a positive probability of the outside option being chosen.

That is, if the DM starts in either L1 or L2, with some probability the process will hit the line $x_1 = x_2$. Then, every time that line is hit, the process goes down that line (as $\min\{x_1, x_2\}$ falls over time), and with some positive probability can reach region O, where the DM takes the outside option. This is noted by the arrows in Fig. 4

Lastly, we can compute the comparative statics of the optimal learning strategy with respect to the learning costs, noise of the signals, and value of the outside option. The results are intuitive.

Proposition 2. *As $c\sigma^2$ increases, $\underline{x}$ increases, $\overline{X}(x)$ decreases, and $V(x_1, x_2)$ decreases. As $\overline{\pi}$ increases, $\underline{x}$ decreases, $\overline{X}(x)$ decreases, and $V(x_1, x_2)$ increases. As π_0 increases, $\underline{x}$ increases, $\overline{X}(x)$ increases, and $V(x_1, x_2)$ increases.*

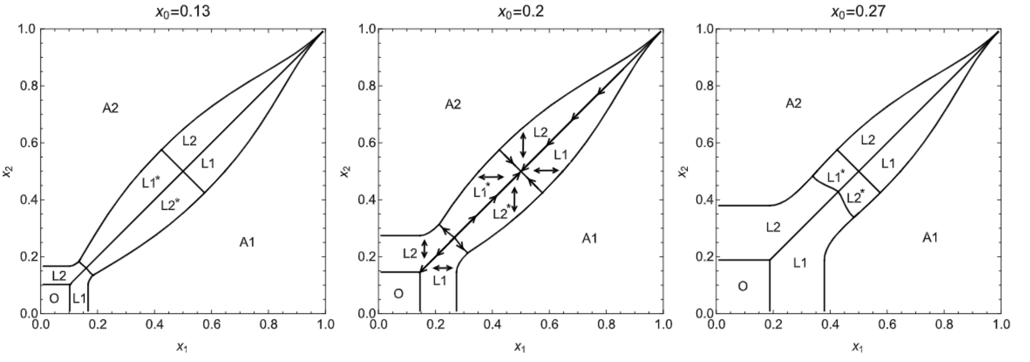

Fig. 5. Optimal learning strategy with medium values of the outside option, $x_0 = 0.13, 0.2, 0.27$ from left to the right, and $c\sigma^2/\Delta\pi = 0.1$. It is optimal for the decision maker to adopt alternative i in region Ai, and to learn alternative i in regions Li and Li*, for $i = 1, 2$. We mark the directions for belief trajectories only for the case with $x = 0.2$, while the other two cases with $x = 0.13$ and $x = 0.27$ follow a similar pattern.

Two alternatives with medium-value outside option So far, we have completely solved the optimal learning problem for the cases where $x_0 \leq 0$ and $x_0 \geq x_0^*$. For the intermediate case with $0 < x_0 < x_0^*$, we are not able to characterize the value function analytically. However, we can show that the optimal policy does not involve always learning about the better alternative. That is, in this intermediate case there has to be a region on the belief space where the DM chooses to learn about the worse alternative.

Computing the optimal policy numerically, Fig. 5 illustrates the optimal policy of the learning problem under some parameter setting where the outside option takes a medium value. The optimal learning strategy becomes more complex.[14] For both high and low beliefs, it is optimal for the DM to learn the better alternative first; however, for the medium beliefs, it is optimal for the DM to learn the worse alternative first. For low beliefs, it is optimal for the DM to learn about the superior alternative because the beliefs about the worse alternative could fall below the consideration set threshold, $\underline{x}$.

By comparing the optimal learning strategies under high, medium, and low value of the outside options, we can evaluate numerically that the higher the outside option, the smaller the starred regions. That is to say, a good outside option makes the better alternative more attractive for learning.[15] The reason is that with a good outside option, if the DM learns the better alternative and ends up with bad news on it, he can use the outside option instead of the worse alternative as the backup option. In contrast, if the DM chooses to learn the worse alternative, the outside option may not be useful as a backup option, since it can be dominated by the better alternative.

4. Adoption likelihood and probability of being correct

In this section, we continue the analysis on the case with two symmetric alternatives. Under the two cases of a high-value and low-value outside option, we derive a decision maker's adoption

---

[14] The features of this more complex learning strategy can also be obtained analytically in a two-period model.

[15] Quah and Strulovici (2013) consider the possibility of counterintuitive comparative statics with respect to discounting in dynamic problems, and present simple conditions such that those counterintuitive results stop holding. In this case, such a condition could be the outside option having a sufficiently high value.

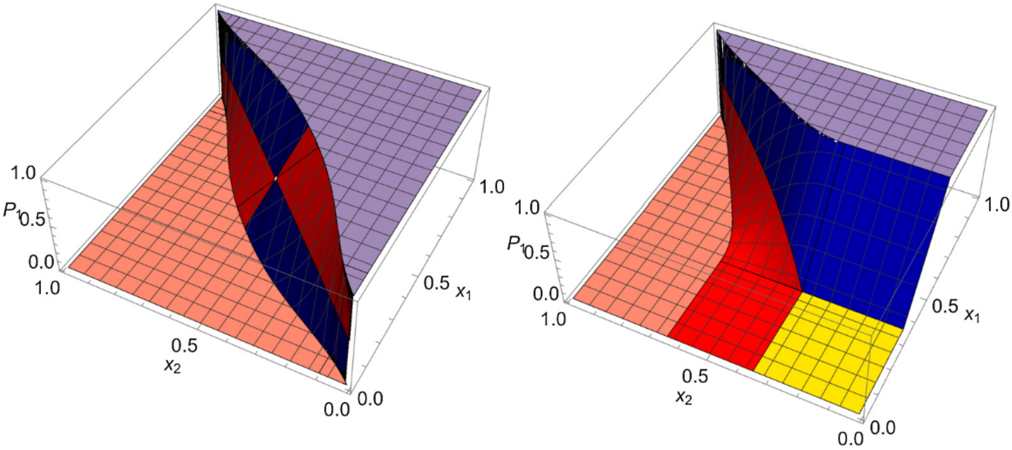

Fig. 6. Adoption likelihood of alternative 1, given a DM's current belief, under $c\sigma^2/\Delta\pi = 0.1$, and $x_0 \leq 0$ for the left panel and $x_0 = 0.5$ for the right panel. The figures are color coded as in Fig. 3.

likelihood of each alternative as well as expected probability of being ex-post correct, based on his optimal learning strategy.

It is interesting to understand the adoption likelihood in many applications. For example, in product markets, at the aggregated level, the adoption likelihood corresponds to the demand function of multiple products under consumer learning. Given a DM's current posterior beliefs as x_1 and x_2, we denote his adoption likelihood of alternative i as $P_i(x_1, x_2)$. By symmetry, we have $P_2(x_1, x_2) = P_1(x_2, x_1)$, therefore, we only need to focus on $P_1(x_1, x_2)$ below. We calculate $P_1(x_1, x_2)$ by invoking the *Optional Stopping Theorem* for martingales, and solving an ordinary differential equation. From this one can obtain the adoption likelihood for either low or high outside option which is presented in Theorem A3 in the Appendix.

Fig. 6 presents $P_1(x_1, x_2)$ under two sets of parameter settings with $x_0 \geq x_0^*$ and $x_0 \leq 0$. The figure illustrates the intuitive result that the DM is more likely to adopt an alternative if his belief of the alternative is higher, or his belief of the other alternative is lower.

We are also interested in the adoption likelihood of (either) one alternative, which is defined as $P(x_1, x_2) \equiv P_1(x_1, x_2) + P_2(x_1, x_2)$. It is interesting to note that $P(x_1, x_2)$ does not always increase with x_1 or x_2. This means that a higher belief of one alternative may lead to a lower adoption likelihood of the two alternatives combined. To understand the intuition, let us consider a special case. Given the DM's beliefs of the two alternatives as x_1 and x_2, if x_1 is high enough such that it exceeds the adoption boundary $\overline{X}(x_2)$, the DM will adopt alternative 1 immediately. In this case, the adoption likelihood is one. Now consider an exogenous increase of x_2 such that x_1 is now below the adoption boundary $\overline{X}(x_2)$. In this case, the DM will optimally learn more information before making an adoption decision. After obtaining more information, it is possible that the DM likes the alternatives more, in which case, he will adopt at most one of them; it is also possible that the DM gets some negative signals and decides to take the outside option. In general, the adoption likelihood will be lower than one after the increase of x_2. Fig. 7, which presents $P(x_1, x_2)$ under some parameter setting with $x_0 \geq x_0^*$, illustrates this point.

The introduction of a new alternative can be equivalently viewed as increasing its belief from below $\underline{x}$ to some level above $\underline{x}$. By the same argument, we can show that more alternatives available for learning and adoption may decrease the adoption likelihood. Applying this result to

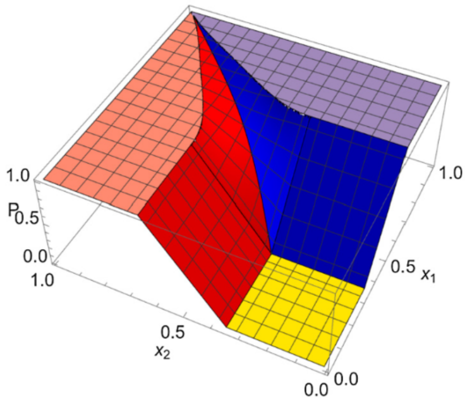

Fig. 7. Adoption likelihood of either one alternative, given a DM's current belief, under $c\sigma^2/\Delta\pi = 0.1$ and $x_0 = 0.5$. The figure is color coded as in Fig. 3.

product markets, can be seen as providing a rational explanation to *consumer choice overload*,[16] under the circumstance that a consumer engages in costly information gathering before making a choice. More options to choose from may potentially lead a consumer to exert a greater effort to distinguish the best from the rest, resulting possibly in a lower probability of choosing anything.

It is also interesting to investigate a decision maker's expected probability of being correct. The DM decides whether to learn more information and which alternative to adopt based on his current imperfect information. Therefore, it is possible that he will make mistakes in an ex post point of view, when all the uncertainties about both alternatives have been resolved. Theorem A4 in the Appendix characterizes a decision maker's expected probability of being correct ex post, $Q(x_1, x_2)$, given his current beliefs of the two alternatives (x_1, x_2), in both cases with $x_0 \leq 0$ and $x_0 \geq x_0^*$.

Fig. 8 presents $Q(x_1, x_2)$ and $\Delta Q(x_1, x_2)$ under some parameter setting, where $\Delta Q(x_1, x_2) \equiv Q(x_1, x_2) - Q_0(x_1, x_2)$ represents the improvement in probability of being correct ex post due to learning, and where $Q_0(x_1, x_2)$ is the probability of being correct ex post when learning is not allowed before adoption. Fig. 8 illustrates that it is easier to make mistakes ex post if two alternatives have similar intermediate beliefs ex ante. This is also the case where optimal learning before adoption reduces the probability of ex-post mistakes the most, similar to Fig. 3.

5. Extensions

In this section we study several extensions, including asymmetric alternatives, the effect of the number of alternatives, and time discounting.

5.1. Asymmetric alternatives

We consider the case that the outside option has a sufficiently high value. We allow the two alternatives to differ in their payoff distributions, noise levels of signals, as well as learning

[16] For lab and field experiments on choice overload, refer to a review by Scheibehenne et al. (2010). Different from our setting, Fudenberg and Strzalecki (2015) consider a dynamic logit model with choice aversion where a consumer may prefer to have a smaller choice set ex ante.

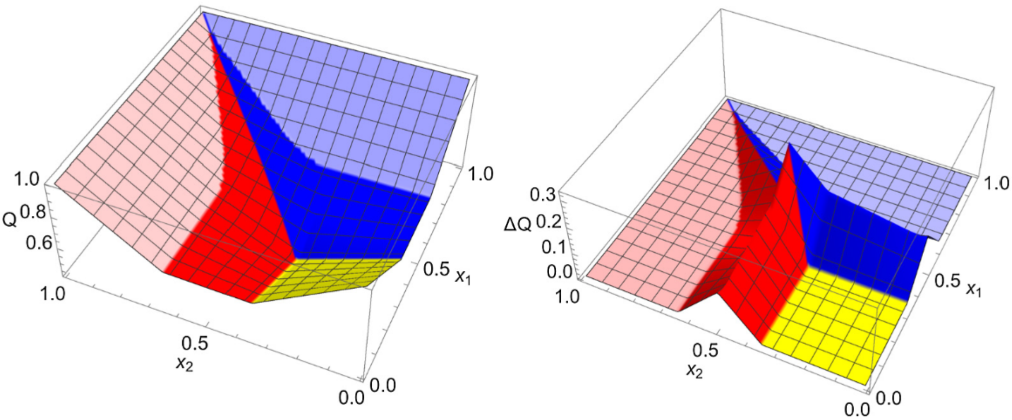

Fig. 8. Probability of being correct ex post (left panel) and the improvement in probability of being correct due to learning (right panel), under $x_0 = 0.5$ and $c\sigma^2/\Delta\pi = 0.1$.

costs. We aim to solve for the decision maker's optimal learning strategy or, equivalently, we want to understand when to switch learning between the two alternatives, and when to stop learning and make an adoption decision. We should expect that for the asymmetric alternatives case the solution structure will be similar to that of the symmetric alternatives case. We must have the adoption boundaries for the two alternatives, $X_1(x_2)$ and $X_2(x_1)$, different from each other. Moreover, the boundary separating "Learn 1" and "Learn 2" will no longer be a 45-degree line, and thus needs to be determined at the same time. This is a more difficult problem from a technical point of view.

Let us first look at the case that one alternative is "better" than the other, in terms of a higher information-to-noise ratio or a lower learning cost. Without loss of generality, let us assume that alternative 1 is better, i.e., we consider the case that $\Delta\pi_1 > \Delta\pi_2$, or $c_1 < c_2$.[17] Intuitively, in this case, we expect that alternative 1 will be preferred for learning, and thus the boundary separating "Learn 1" and "Learn 2" should be above the 45-degree line. We further assume that there is a point (x_1^*, x_2^*) such that the boundary separating "Learn 1" and "Learn 2" intersects with the adoption boundary of alternative 2 under this belief. Based on this assumption, we can characterize the solution to the *HJB* equation (8) analytically, and we present the solution in the Appendix. Fig. 9 illustrates the optimal learning strategy under some parameter settings. The left panel shows the case of different learning costs between the two alternatives, $c_1 < c_2$, with everything else being the same. The right panel shows the case with different payoff ranges between the two alternatives, $\Delta\pi_1 > \Delta\pi_2$, with everything else the same.

Consistent with the symmetric alternatives case, when $x_i < \underline{x}_i$, alternative i is not considered for learning or adoption. When $x_1 \geq \underline{x}_1$ and $x_2 \geq \underline{x}_2$, both alternatives are in the consideration set, and the decision maker's optimal learning strategy is now slightly more complicated.

Interestingly, when the expected value of both alternatives is high enough, i.e., when $x_2 \geq x_2^*$, the DM only learns alternative 1. The DM keeps on learning alternative 1 until either x_1 exceeds the adoption boundary of alternative 1, $\overline{X}_1(x_2)$, in which case he adopts alternative 1, or x_1 drops below the adoption boundary of alternative 2, $\underline{X}_1(x_2)$, in which case he adopts alternative 2. In

[17] Mathematically, $\sigma_1^2 < \sigma_2^2$ will be exactly the same as $c_1 < c_2$, because in our model only $\sigma_i^2 c_i$ is identified, as noted above. In this subsection, as the alternatives are not symmetric, we consider the possibility of $\underline{\pi}_i \neq 0$.

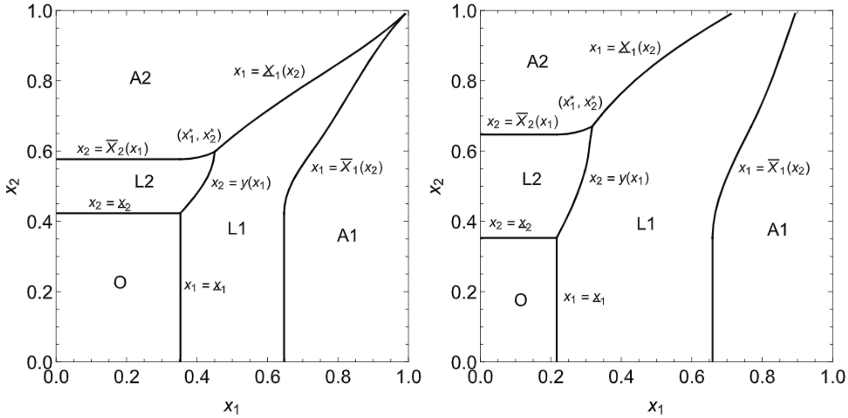

Fig. 9. Optimal learning strategy with asymmetric alternatives. For the left panel: $\Delta\pi_1 = \Delta\pi_2$, $(\pi_0 - \underline{\pi}_1)/\Delta\pi_1 = (\pi_0 - \underline{\pi}_2)/\Delta\pi_2 = 0.5$, and $c_1\sigma_1^2/\Delta\pi_1 = 0.1 < c_2\sigma_2^2/\Delta\pi_2 = 0.2$. For the right panel: $\Delta\pi_1 = 1.2\Delta\pi_2$, $(\pi_0 - \underline{\pi}_1)/\Delta\pi_2 = (\pi_0 - \underline{\pi}_2)/\Delta\pi_2 = 0.5$, and $c_1\sigma_1^2/\Delta\pi_2 = c_2\sigma_2^2/\Delta\pi_2 = 0.1$. The figures are color coded as in Fig. 3.

this case, as the expected value of both alternatives is high enough, the DM searches only the alternative for which either the search costs or the noise of the signal is lower, or for which the benefit of having the high outcome is higher. This can be seen as intuitive: If the expected value of both alternatives is high enough, the question is mostly one of which alternative to end up choosing, and therefore the DM learns information on the alternative for which there is a greater information to search cost ratio. Note then, that in comparison with the case of the previous sections where the noisiness of signals, search costs, and possible payoffs were symmetric across alternatives, we have now a situation in which if the expected value of both alternatives is sufficiently high the DM makes a choice between the two alternatives with probability one, while in the previous sections the DM could end up choosing the outside option with some positive probability.

When the expected value on the worse alternative is not so high, i.e., when $x_2^* > x_2 \geq \underline{x}_2$, and the outside option has a high value, we are back in the situation of the previous sections where the DM chooses to learn the alternative with relatively higher expected value. In the case of previous section, this was just learning about the alternative with the higher expected value. In the case here, with asymmetric alternatives, this means learning about alternative 1, the alternative with an advantage in search for information, if the expected value of alternative 1 is not much lower than the expected value of alternative 2. The results are then that the DM learns alternative 1 if and only if

$$\left(\overline{X}_1(x_2) - x_1\right)\left[\frac{2\sigma_1^2 c_1}{(\Delta\pi_1)^2}\left(\Phi\left(\overline{X}_1(x_2)\right) - \Phi(x_1)\right) + \Delta\pi_1\right] \leq \left(\overline{X}_2(x_1) - x_2\right)\left[\frac{2\sigma_2^2 c_2}{(\Delta\pi_2)^2}\left(\Phi\left(\overline{X}_2(x_1)\right) - \Phi(x_2)\right) + \Delta\pi_2\right], \quad (35)$$

where $\overline{X}_1(x_2)$ and $\overline{X}_2(x_1)$ are the adoption boundaries for alternatives 1 and 2, respectively, for the region that $x_1 \geq \underline{x}_1$ and $x_2^* > x_2 \geq \underline{x}_2$. While it is not intuitive to understand the condition above, it does show that a simple index policy is not optimal.

Lastly, we want to point out that the boundary separating "Learn 1" and "Learn 2", denoted as $x_2 = y(x_1)$ in the figure, cannot be obtained with value matching and smooth pasting conditions alone. This is intuitive, because even with the boundary exogenously given, we find that both value matching and smooth pasting conditions are necessary to pin down the solution. Now, to maximize the expected payoff by choosing the boundary optimally is equivalent to imposing an extra set of conditions, so called *super contact* conditions (Dumas 1991; Strulovici and Szydlowski 2015), which guarantee that $V_{x_1x_1}(x_1, x_2)$, $V_{x_2x_2}(x_1, x_2)$ and $V_{x_1x_2}(x_1, x_2)$ are continuous across the boundary. We find that two of the three super contact conditions are redundant, and we only need to guarantee $V_{x_1x_2}(x_1, x_2)$ to be continuous across the boundary that separates "Learn 1" and "Learn 2". Similarly, we can show that the redundancy is due to the learning problem structure that the DM can only learn one alternative at a time and the resulting equation (12) is a parabolic PDE.

On the right panel of Fig. 9, note also that because $\Delta\pi_1 > \Delta\pi_2$, the DM is more likely to learn and adopt alternative 1 than alternative 2, and the adoption thresholds do not converge anymore when the beliefs converge to $(1, 1)$.

So far, we have characterized the DM's optimal learning strategy when one alternative is "better" than the other in terms of a higher information-to-noise ratio and a lower learning cost. More generally, there are also cases where one alternative has a higher information-to-noise ratio, but at a higher learning cost. This case is complicated with discrete "jumps" of the DM's optimal learning strategy with respect to the parameters (e.g., learning costs). We briefly discuss this case in the Appendix. We have also considered asymmetric alternatives when the outside option is low. In that case, it is difficult to characterize the optimal solution analytically, but with numerical analysis, we can find examples that with small heterogeneity between the two alternatives, we still have the starred regions in the optimal learning strategy where it is optimal to learn an alternative with lower belief first.

5.2. More than two alternatives

In this subsection, we consider the optimal learning problem with more than two alternatives. The optimal learning problem and the associated *HJB* equation can be formulated similarly. Although it is difficult to solve the general problem analytically with more than two alternatives (except for the infinite number of alternatives case presented below), we can see numerically how the optimal policy is affected. Consider three symmetric alternatives without an outside option. By symmetry, we only need to obtain the optimal learning strategy under the belief $x_1 \geq x_2 \geq x_3$. Fig. 10 illustrates the optimal learning strategy. In this case it is optimal to learn the second best alternative (alternative 2) under some belief only when x_3 is relatively low. Comparing the three panels in Fig. 10 with Fig. 4, 5, and 2 respectively, we find the impact of the third alternative on the optimal learning strategy to be similar to that of the outside option.

Moreover, we could not find cases where it was optimal to learn the third best alternative (alternative 3). It is interesting that the DM only considers learning from the top two alternatives, and sometimes does not learn about the top alternative, in order to possibly rule out the second best alternative, to then decide to adopt the top alternative.

We have also numerically solved the optimal learning problem with three alternatives and an outside option, and in that case the cutoff strategy is still valid—a decision maker considers an alternative for learning or adoption if and only if his belief of this alternative is sufficiently high.

Lastly, we consider the optimal learning problem of infinite symmetric alternatives with the same prior belief $\hat{x}$. By the standard analysis, we can show that the optimal learning strategy

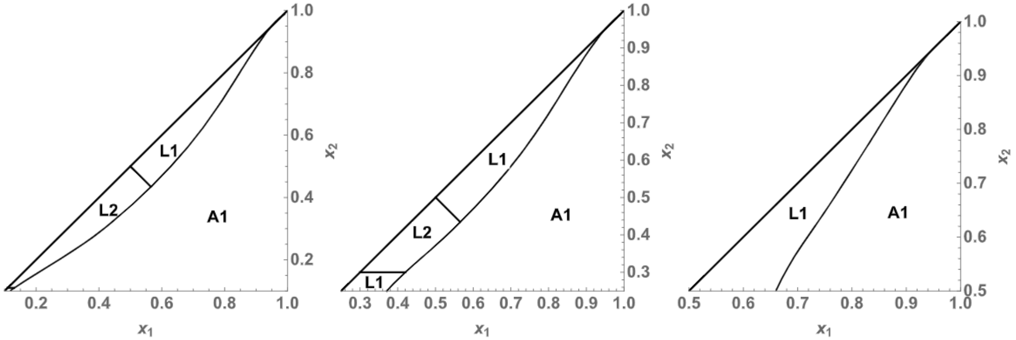

Fig. 10. Optimal learning strategy with three alternatives and no outside option. $c\sigma^2/\Delta\pi = 0.1$. $x_3 = 0.1, 0.25, 0.5$ from left to the right.

in this setting follows a simple *reservation value* policy. A DM continues to learn information on alternative i until either x_i exceeds x_R, in which case he adopts alternative i, or x_i drops below $\hat{x}$, in which case he moves to learn the next alternative. For $\hat{x} \leq x \leq x_R$, the value function $U(x)$ is given by equation (v) in the Appendix. We can determine x_R by the following set of boundary conditions: $U'(\hat{x}) = 0$, $U(x_R) = \Delta\pi \cdot x_R$, and $U'(x_R) = \Delta\pi$. By combining these three conditions, we have x_R as:

$$x_R = \Phi^{-1}\left(\Phi(\hat{x}) - \frac{(\Delta\pi)^3}{2\sigma^2 c}\right). \tag{36}$$

We also get that x_R is greater than $\overline{X}(\hat{x})$, illustrating that with more alternatives the DM is more demanding on what an alternative has to deliver before the DM settles on it.

That is, when there is the possibility of learning about other alternatives, the DM becomes more demanding on the threshold on the difference between the top two alternatives that would make the DM stop the learning process and adopt one alternative. This can be seen as intuitive as the DM has now more options to get a higher expected payoff. This also illustrates how the expected value of the alternatives beyond the second best affect the decision of the DM of when to adopt an alternative. As the DM could get enough bad news on the first and second best alternative, the third best alternative matters, and so forth.

5.3. Time discounting

In this subsection, we consider a similar problem setting with no flow cost for learning but with time discounting. Similar to (8), we can write down the following *HJB* equation.

$$\max\left\{\max_{1\leq i\leq n}\left\{\frac{(\Delta\pi_i)^2}{2\sigma_i^2}x_i^2(1-x_i)^2 V_{x_i x_i}(\mathbf{x}) - rV(\mathbf{x})\right\}, g(\mathbf{x}) - V(\mathbf{x})\right\} = 0. \tag{37}$$

Along the same lines as above, we can show that there exists a unique solution to this problem.

For the case of a single alternative, the optimal learning strategy is that there exist $0 \leq \underset{\sim}{x} < \tilde{x} \leq 1$ such that the DM continues learning when $\underset{\sim}{x} \leq x \leq \tilde{x}$, and he stops learning to take the outside option when $x < \underset{\sim}{x}$ or to adopt the alternative when $x > \tilde{x}$.

For the case of two symmetric alternatives, similarly to above, we propose a solution, and then verify that it satisfies the *HJB* equation (37). Due to symmetry, we only need to consider the case with $x_1 \geq x_2$. There are two cases to consider. First, if $x_2 \leq \underset{\sim}{x}$, we propose that alternative 2 will not be considered for either learning or adoption, and one goes back to the case with a single uncertain alternative—alternative 1. Second, if $x_1 \geq x_2 > \underset{\sim}{x}$, we propose that there exists a smooth function $\widetilde{X}(\cdot)$ such that the DM learns alternative 1 when $x_2 \leq x_1 \leq \widetilde{X}(x_2)$, and adopts alternative 1 when $x_1 > \widetilde{X}(x_2)$. By applying a similar analysis as above, we can determine $V(x_1, x_2)$ by solving the partial differential equation

$$\frac{(\Delta\pi)^2}{2\sigma^2} x_1^2(1-x_1)^2 V_{x_1x_1}(x_1, x_2) - rV(x_1, x_2) = 0 \tag{38}$$

subject to the value matching and smooth pasting conditions at the adoption boundary of $x_1 = \widetilde{X}(x_2)$. This yields

$$V(x_1, x_2) = \frac{\alpha+1}{2\alpha}\pi Z(x_2)^{\frac{\alpha-1}{2}} x_1^{\frac{\alpha+1}{2}}(1-x_1)^{-\frac{\alpha-1}{2}} + \frac{\alpha-1}{2\alpha}\pi Z(x_2)^{-\frac{\alpha+1}{2}} x_1^{-\frac{\alpha-1}{2}}(1-x_1)^{\frac{\alpha+1}{2}}$$
$$\text{if } \underset{\sim}{x} \leq x_2 \leq x_1 \leq \widetilde{X}(x_2), \tag{39}$$

where $\alpha \equiv \sqrt{1 + \frac{8r\sigma^2}{(\Delta\pi)^2}} > 1$, $\widetilde{X}(x) = (1 + Z(x))^{-1}$, and $Z(x)$ is determined by the following boundary value problem:

$$\left(\alpha^2 - 1\right) x(1-x)\frac{Z'(x)}{Z(x)} = \alpha^2 + \alpha(1-2x)\frac{Z(x)^\alpha + \left(\frac{1-x}{x}\right)^\alpha}{Z(x)^\alpha - \left(\frac{1-x}{x}\right)^\alpha} + (1-2x)$$
$$+ \alpha\frac{Z(x)^\alpha + \left(\frac{1-x}{x}\right)^\alpha}{Z(x)^\alpha - \left(\frac{1-x}{x}\right)^\alpha}, \tag{40}$$
$$Z(\underset{\sim}{x}) = \frac{1-\tilde{x}}{\tilde{x}}.$$

Lastly, to verify this is indeed the solution, we need to guarantee that the DM prefers to learn alternative 1 for $\underset{\sim}{x} \leq x_2 \leq x_1 \leq \widetilde{X}(x_2)$. By (37), we need to have that

$$\frac{(\Delta\pi)^2}{2\sigma^2} x_2^2(1-x_2)^2 V_{x_2x_2}(x_1, x_2) \leq rV(x_1, x_2) \text{ for } \forall x_1, x_2, \text{ s.t. } \underset{\sim}{x} \leq x_2 \leq x_1 \leq \tilde{X}(x_2), \tag{41}$$

which holds for all x_0. This was the condition that failed for the case with no discounting, and that leads to the optimal strategy having the DM learning on the worse alternative when x_0 was sufficiently low, and x_1 and x_2 are not too high. In this case of discounting, this condition is satisfied for all x_0, x_1, and x_2, and we then have that the DM when choosing to learn about an alternative, always chooses to learn about the alternative that has a higher expected value.

The intuition is that it does not help to minimize the learning costs by learning an inferior alternative. Specifically, with time discounting, the learning cost depends on the current expected payoff. By learning the inferior alternative first, if the DM receives negative signals, the learning costs will fall, which will give an incentive to continue learning. On the other hand, if he receives

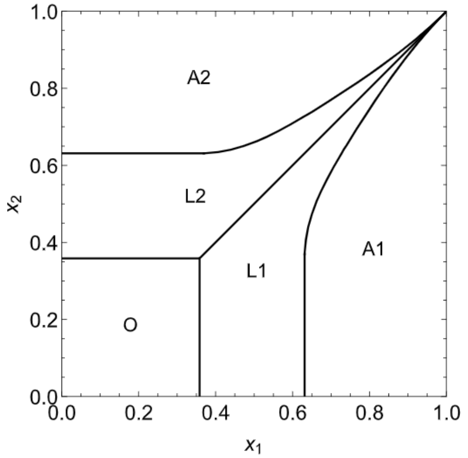

Fig. 11. Optimal learning strategy with time discounting under the parameter setting that $\pi_0/\Delta\pi = 0.5$, $\sigma/\Delta\pi = 1$, and $r = 0.2$.

positive signals, he will continue learning. However, his learning cost is also high because he just received positive signals. Therefore, the DM will not save costs by not learning the best alternative first.

From this analysis we can obtain the value function which is presented in Theorem A6 in the Appendix. Following an almost identical analysis as in Section 3, we can construct the optimal allocation and stopping rule and the corresponding belief updating process to attain the value function. Formally, we have the following theorem that characterizes the optimal learning strategy.

Theorem 3. *Consider the DM's optimal learning problem defined in (6) but with time discounting instead of a constant flow learning cost. Given one alternative i under his consideration, the DM keeps learning it if* $\underset{\sim}{x} \le x_i \le \tilde{x}$*, adopts it if* $x_i > \tilde{x}$*, and takes the outside option if* $x_i < \underset{\sim}{x}$*. Given two alternatives under his consideration, the DM always learns information from the alternative with higher belief. He stops learning and adopts it if and only if his belief of this alternative is above a threshold, which increases with his belief of the other alternative.*

Fig. 11 presents a DM's optimal learning strategy under some parameter setting, illustrating that with discounting it is always optimal to learn the best alternative first.

One could also have a model with both cost of information gathering and discounting. Such a model would replicate some of the results above of the case with no discounting and with positive cost of information gathering, but with more complicated characterization of the value functions and of the optimal policy. In such a model, the result above that the DM would prefer to learn first about the worse alternative under some conditions would continue to be a possibility.

6. Discussion and conclusion

This paper examines a canonical problem of optimal information gathering. We allow gradual learning on multiple alternatives before one's choice decision. The decision maker solves the problem of optimal dynamic allocation of learning efforts as well as optimal stopping of the learning process, so as to maximize the expected payoff. We show that if the outside option is low enough, a decision maker may choose to gather information on the worse alternative, that has lower expected payoff and less uncertainty, in order to rule out that alternative, and that the optimal policy is not a single-index policy.

It is interesting to compare the optimal learning rule in our setting with that under "complete learning" in Weitzman (1979), where everything learnable about an alternative is learned with the first signal. For complete learning, one can show that the optimal rule is to learn about the alternative with the highest belief if everything else is the same. Our result highlights that with the possibility of gradual learning, a decision maker's optimal learning strategy can be quite different and complex—sometimes it is not optimal to learn the best alternative first. It is true that under complete learning, it can be optimal for the DM to first learn an alternative that has low expected payoff but high variance, in the hopes of capturing its high payoff. Strictly speaking, in that possibility, the alternative with low expected payoff is "seemingly" worse, but is, in fact, different from the other alternatives in terms of different payoff distributions. In contrast, we find that a DM may optimally learn a worse alternative first, given everything else between the two alternatives the same in this binary environment. Note also that this worse alternative is also the alternative with lower variance, so that the DM may optimally learn an alternative with lower expected value and lower variance. Also, our underlying mechanism works in a different way. With complete learning, after learning an alternative with low expected payoff and high variance, the DM will either adopt that alternative if he gets a very positive signal, or learn other alternative if he gets a negative signal. In contrast, in our setting with gradual learning, when learning a worse alternative, the DM will either potentially switch to learning the better alternative if he gets enough positive signals on the worse alternative, or adopt the better alternative immediately if he gets enough negative signals on the worse alternative.

It is also worthwhile to revisit the index policy proposed by Glazebrook (1979) as the optimal policy for stoppable bandits problem under some conditions. Under the index policy, a DM chooses alternative i to act on (either learn or adopt) if it has the greatest index, which, like the Gittins index, only depends on the state of alternative i. Given that alternative i is chosen, the DM chooses to learn or to adopt it based on some strategy that only depends on the state of alternative i. Here, we have shown that the DM chooses to stop learning and adopt alternative i if and only if $x_i \geq \overline{Y}(x_j)$, which depends on x_j as well. Therefore, Glazebrook's index policy is suboptimal for our problem.

The model that we present can be applicable to consumer search for information across different products prior to purchase, with the consumer only being able to learn about each product at a time. In terms of the result above, a consumer may potentially try to get information on a seemingly worse product in order to rule it out and decide to purchase the seemingly best product.

Given this environment of consumer search for information, it would also be interesting to investigate how a monopolist carrying both products should behave in terms of information to provide, and prices to charge. It would also be interesting to investigate what would happen in an oligopoly competition setting with each firm carrying one product.

Another interesting application is the case of a policy maker who has to make a decision between different policies and can research each policy separately before deciding which policy to choose. The policy maker can decide to research a worse policy to potentially rule it out, and opt for the seemingly better option if the policy maker receives bad information about the worse policy.

In some situations the technology of search for information is such that when learning about a category, a decision-maker learns about the fit with several alternatives at the same time. For example, by learning about the automobile market in general, for example, by reading an automobile magazine, a consumer may learn about his relative preferences between two car models. It would be interesting to study such a situation in future research.

Appendix A

Proof of Lemma 1: existence and uniqueness of the solution. For this proof we consider the results for existence and uniqueness of a solution of a nonlinear second-order partial differential equation (Crandall et al. 1992, which we will denote as CIL[18]). Note also that given that the value function is a viscosity solution, and we know that the value function exists given that it is bounded, existence is straightforward. Consider now the question of uniqueness. Consider a nonlinear second-order PDE of the form

$$H(\mathbf{x}, u, Du, D^2u) = 0, \tag{i}$$

where $Du \in \mathbb{R}^n$ is the differential of u (at $\mathbf{x}$), $D^2u \in \mathcal{S}_n(\mathbb{R})$ is the Hessian matrix of u (at $\mathbf{x}$), and $H : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \times \mathcal{S}_n(\mathbb{R}) \to \mathbb{R}$, with $\mathcal{S}_n(\mathbb{R})$ the set of symmetric $n \times n$ matrices. The function H is said to satisfy *degenerate ellipticity* if and only if $H(\mathbf{x}, u, p, X) \leq H(\mathbf{x}, u, p, Y)$ for all $\mathbf{x}, u, p, X$ and Y such that $X - Y$ is positive semi-definite or $X = Y$. Another property of the function H that we will need is that it is nondecreasing in u.

To see that the HJB equation (8) satisfies *degenerate ellipticity* and that it is nondecreasing in u, note that in our case $n = 2$, and it is equivalent to consider

$$\max\left\{\max_{i=1,2}\left\{-\widehat{a}_i(\mathbf{x})\overline{V}_{x_ix_i}(\mathbf{x}) - c_i\right\}, g(\mathbf{x}) + \overline{V}(\mathbf{x})\right\} = 0. \tag{ii}$$

where $\widehat{a}_i(\mathbf{x}) \equiv \frac{(\Delta\pi_i)^2}{2\sigma_i^2}x_i^2(1 - x_i)^2$, $g(\mathbf{x}) \equiv \max\left\{\Delta\pi_1 x_1 + \underline{\pi}_1, \Delta\pi_2 x_2 + \underline{\pi}_2, \pi_0\right\}$ and $\overline{V}(\mathbf{x}) \equiv -\widehat{V}(\mathbf{x})$, for all $\mathbf{x}$.

Note then that the left hand side of (ii) is nondecreasing in $\overline{V}$ and nonincreasing in $\overline{V}_{x_ix_i}$ for all i, which means that (ii) satisfies degenerate ellipticity.

A continuous function u is said to be a *viscosity solution* to equation (i) if for all twice continuously differentiable functions $\phi : \mathbb{R}^n \to \mathbb{R}$: (1) if $u - \phi$ attains a local maximum at $x_0 \in \mathbb{R}^n$ then $H(x_0, \phi(x_0), D\phi(x_0), D^2\phi(x_0)) \leq 0$ (*viscosity subsolution*); and (2) if $u - \phi$ attains a local minimum at $x_0 \in \mathbb{R}^n$, then $H(x_0, \phi(x_0), D\phi(x_0), D^2\phi(x_0)) \geq 0$ (*viscosity supersolution*). Note that this definition uses the fact that equation (i) satisfies degenerate ellipticity. For later use let us also define that u is a *strict viscosity subsolution* if for all twice continuously differentiable functions ϕ such that $u - \phi$ attains a local maximum at $x_0 \in \mathbb{R}^n$, then $H(x_0, \phi(x_0), D\phi(x_0), D^2\phi(x_0)) \leq -\eta < 0$.

We apply the modification to Theorem 3.3 described in Section 5.C in CIL,[19] a comparison result, which implies uniqueness of the viscosity solution. We re-state here Theorem 3.3 in CIL:

Theorem *(Theorem 3.3 in CIL). Let Ω be a bounded open space of $\mathbb{R}^N$, $\partial\Omega$ be the boundary of Ω, $\overline{\Omega} \equiv \Omega \cup \partial\Omega$, and $H : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \times \mathcal{S}_n(\mathbb{R}) \to \mathbb{R}$ be a continuous function satisfying*

[18] See also Øksendal and Reikvam (1998) for the analysis of viscosity solutions in optimal stopping problems.

[19] See also Barles (1997), p. 38.

(C1) There exists a $\gamma > 0$ *such that* $\gamma(r-s) \leq H(\mathbf{x}, r, p, X) - H(\mathbf{x}, s, p, X)$ *for* $r \geq s$, $(\mathbf{x}, p, X) \in \overline{\Omega} \times \mathbb{R}^n \times \mathcal{S}_n(\mathbb{R})$,
(C2) H *satisfies degenerate ellipticity,*
and
(C3) H *satisfies the following structure condition for any* $\alpha > 0$. *There is a continuous function* ω: $[0, \infty) \to [0, \infty)$, *satisfying* $\omega(0^+) = 0$ *such that*

$$H(\mathbf{y}, r, \alpha(\mathbf{x}-\mathbf{y}), Y) - H(\mathbf{x}, r, \alpha(\mathbf{x}-\mathbf{y}), X) \leq \omega(\alpha|\mathbf{x}-\mathbf{y}|^2 + |\mathbf{x}-\mathbf{y}|) \tag{iii}$$

when $\mathbf{x}, \mathbf{y} \in (0,1)^2, r \in \mathbb{R}, X, Y \in \mathcal{S}_n(\mathbb{R})$ *and*

$$-3\alpha \begin{pmatrix} I & 0 \\ 0 & 1 \end{pmatrix} \leq \begin{pmatrix} X & 0 \\ 0 & -Y \end{pmatrix} \leq 3\alpha \begin{pmatrix} I & I \\ -I & I \end{pmatrix}. \tag{iv}$$

Let u *be an upper semicontinuous function and a subsolution of* $H = 0$ *in* Ω, v *be a lower semicontinuous function and a supersolution of* $H = 0$, *and* $u \leq v$ *on* $\partial\Omega$. *Then* $u \leq v$ *in* $\overline{\Omega}$.

Note that if we know that the solution is continuous and is unique on the boundary of Ω, we can apply this theorem to two any possible solutions u and v to show that $u \leq v$ and $v \leq u$, to obtain $u = v$. We state this condition of uniqueness at the boundary of Ω as the following condition:
(C4) In the boundary of $(0,1)^2$ the viscosity solution is unique.

Condition (C1) is not satisfied in our case, as we only have that H is nondecreasing in u, so we have to modify this condition as discussed in Section 5.C in CIL.[20] We discuss below how to make this modification in our case.

Condition (C2) is satisfied as shown above.

To check condition (C3) note that, as discussed in CIL, p. 20, the max operation of functions satisfying the structure condition with a common ω, satisfies the structure condition. The part $g(\mathbf{x}) + \overline{V}(\mathbf{x})$ satisfies directly the structure condition. So, the only remaining part is to show that $-\frac{(\Delta\pi_i)^2}{2\sigma_i^2} x_i^2(1-x_i)^2 \overline{V}_{x_i x_i}(\mathbf{x}) - c_i$ satisfies the structure condition. In that case the structure condition (iii) becomes

$$\frac{(\Delta\pi_i)^2}{2\sigma_i^2}[x_i^2(1-x_i)^2 X_{ii} - y_i^2(1-y_i)^2 Y_{ii}] \leq \omega(\alpha|\mathbf{x}-\mathbf{y}|^2 + |\mathbf{x}-\mathbf{y}|).$$

Note that starting from the left hand side of the equation above we can obtain

$$\frac{(\Delta\pi_i)^2}{2\sigma_i^2}[x_i^2(1-x_i)^2 X_{ii} - y_i^2(1-y_i)^2 Y_{ii}]$$
$$= \frac{(\Delta\pi_i)^2}{2\sigma_i^2} \text{trace}\left(\begin{bmatrix} x_i^2(1-x_i)^2 & x_i y_i (1-x_i)(1-y_i) \\ x_i y_i (1-x_i)(1-y_i) & y_i^2(1-y_i)^2 \end{bmatrix} \begin{bmatrix} X_{ii} & 0 \\ 0 & Y_{ii} \end{bmatrix} \right)$$
$$\leq 3\alpha \frac{(\Delta\pi_i)^2}{2\sigma_i^2} \text{trace}\left(\begin{bmatrix} x_i^2(1-x_i)^2 & x_i y_i (1-x_i)(1-y_i) \\ x_i y_i (1-x_i)(1-y_i) & y_i^2(1-y_i)^2 \end{bmatrix} \begin{bmatrix} I & -I \\ -I & I \end{bmatrix} \right)$$

[20] Note that Condition (C1) is satisfied if we add discounting to the model. A possible argument is then also to add discounting to the model, and make the discount rate go to zero.

$$= 3\frac{(\Delta\pi_i)^2}{2\sigma_i^2}\alpha[(x_i(1-x_i) - y_i(1-y_i)]^2$$

$$\leq 3\frac{(\Delta\pi_i)^2}{2\sigma_i^2}(x_i - y_i)^2,$$

where the first inequality comes from equation (iv), and the second inequality is obtained by noting that $[x_i(1-x_i) - y_i(1-y_i)]^2 = (x_i - y_i)^2(1 - x_i - y_i)^2$ and $|1 - x_i - y_i| < 1$ for all $(x_i, y_i) \in (0, 1)^2$. Then, with $\omega(r) = 3r \max_i \frac{(\Delta\pi_i)^2}{2\sigma_i^2}$, condition (iii) is satisfied.

To check condition (C4), note that for $n = 1$, at the boundary we have the value function equal to $\max\{\pi_0, \underline{\pi}\}$ if $x = 0$, and equal to $\max\{\pi_0, \overline{\pi}\}$ if $x = 1$. Then, applying the result we have a unique viscosity solution, and therefore the value function, for the case of $n = 1$. Suppose now that we have a unique viscosity solution, which is also the value function, for the case in which $n = k \geq 1$. Then, the value function at the boundary is known for the case of $n = k + 1$. To see this, let $V^k(\mathbf{x}^k)$ be the value function when $n = k$ and the beliefs for those k alternatives are $\mathbf{x}^k$. Consider now the value function at the boundary when $n = k + 1$. We have that if $x_{k+1} = 0$, $V^{k+1}(\mathbf{x}^{k+1}) = \max\{V^k(\mathbf{x}^k), \underline{\pi}_{k+1}\}$ and, if $x_{k+1} = 1$, $V^{k+1}(\mathbf{x}^{k+1}) = \max\{V^k(\mathbf{x}^k), \overline{\pi}_{k+1}\}$.

Finally, we have to adjust condition (C1) in Theorem 3.3 in CIL to the case in which there is no strict monotonicity (following Section 5.C in CIL). The uniqueness result is obtained through a comparison of a subsolution and supersolution of (i) in $(0, 1)^2$. Suppose that V is a subsolution of (i) and that W is a supersolution of (i). Set $K = -\max\{\overline{\pi}_1, \overline{\pi}_2, \pi_0\} - 1$ which we know to be strictly below the lower bound of $\overline{V}(\mathbf{x})$. Then $V_\mu(\mathbf{x}) = (1 - \mu)V(\mathbf{x}) + \mu K$ is a strict subsolution of (i) in our case for $\mu \in (0, 1]$, as (i) is convex in (u, Du, D^2u) in our case, and K, a constant, is a strict subsolution to (i). This implies that $H(\mathbf{x}, V_\mu(\mathbf{x}), DV_\mu(\mathbf{x}), D^2V_\mu(\mathbf{x})) \leq -\eta$ for some $\eta > 0$.

We can then follow the proof of Theorem 3.3 in CIL along the same lines. The proof is by contradiction. Suppose that there is a $\mathbf{z} \in (0, 1)^2$ such that $V(\mathbf{z}) > W(\mathbf{z})$. Then, there is a $\mu \in (0, 1)$ such that $V_\mu(\mathbf{z}) > W(\mathbf{z})$. Defining $(\widehat{\mathbf{x}}, \widehat{\mathbf{y}}) = \arg\max_{\mathbf{x},\mathbf{y}} V_\mu(\mathbf{x}) - W(\mathbf{Y}) - \alpha/2|\mathbf{x} - \mathbf{y}|^2$ for $\alpha > 0$, we have that $V_\mu(\widehat{\mathbf{x}}) > W(\widehat{\mathbf{y}})$ as there is a $\mathbf{z} \in (0, 1)^2$ such that $V_\mu(\mathbf{z}) > W(\mathbf{z})$. Then, there is a $\mu \in (0, 1)$ such that $V_\mu(\mathbf{z}) > W(\mathbf{z})$. Then, because V_μ is a strict subsolution and W is a supersolution, we know that there is a η such that, modifying (3.15) in CIL,

$$\begin{aligned}
0 < \eta &\leq H(\widehat{\mathbf{y}}, W(\widehat{\mathbf{y}}, \alpha(\widehat{\mathbf{x}} - \widehat{\mathbf{y}}), Y) - H(\widehat{\mathbf{x}}, V_\mu(\widehat{\mathbf{x}}, \alpha(\widehat{\mathbf{x}} - \widehat{\mathbf{y}}), X) \\
&\leq H(\widehat{\mathbf{y}}, W(\widehat{\mathbf{y}}, \alpha(\widehat{\mathbf{x}} - \widehat{\mathbf{y}}), Y) - H(\widehat{\mathbf{x}}, W(\widehat{\mathbf{y}}, \alpha(\widehat{\mathbf{x}} - \widehat{\mathbf{y}}), X) \\
&\leq \omega(\alpha|\widehat{\mathbf{x}} - \widehat{\mathbf{y}}|^2 + |\widehat{\mathbf{x}} - \widehat{\mathbf{y}}|)
\end{aligned}$$

where the third inequality come from the function H in (i) being nondecreasing in u, and the fourth inequality comes from the structure condition (iii). As in CIL (p. 18), we can then make $\alpha \to \infty$, have $\omega(\alpha|\widehat{\mathbf{x}} - \widehat{\mathbf{y}}|^2 + |\widehat{\mathbf{x}} - \widehat{\mathbf{y}}|) \to 0$, which yields the contradiction.

Analysis of the single alternative case For the single-alternative case, we denote the value function as $U(x)$, so $V(\mathbf{x})$ is used only for $n = 2$, to make the presentation clearer.

To solve for the value function $U(x)$, let us consider three cases. In the first case, when $\underline{x} \leq x \leq \overline{x}$, equation (8) is reduced to the following second-order ordinary differential equation (ODE)

$$\frac{(\Delta\pi)^2}{2\sigma^2}x^2(1-x)^2U''(x) - c = 0.$$

The general solution to this ODE is

$$U(x) = \frac{2\sigma^2 c}{(\Delta\pi)^2}(1-2x)\ln\left(\frac{1-x}{x}\right) + C_1 x + C_2, \quad \text{(v)}$$

where C_1 and C_2 are two undetermined constants. In the second case, when $x < \underline{x}$, we have $U(x) = \pi_0$. Finally, in the third case, when $x > \overline{x}$, we have $U(x) = \Delta\pi \cdot x$. To summarize the three cases, we have the following expression for the value function

$$U(x) = \begin{cases} \pi_0, & x < \underline{x} \\ \frac{2\sigma^2 c}{(\Delta\pi)^2}(1-2x)\ln\left(\frac{1-x}{x}\right) + C_1 x + C_2, & \underline{x} \le x \le \overline{x} \\ \Delta\pi \cdot x, & x > \overline{x} \end{cases}. \quad \text{(vi)}$$

To complete the computation of the optimal policy, we need a set of boundary conditions to determine $\underline{x}$, $\overline{x}$, C_1 and C_2. In fact, $U(x)$ has to be continuous and smooth at $\underline{x}$ and $\overline{x}$, which implies value matching and smooth pasting conditions. Value matching and smooth pasting are standard conditions for optimal stopping problems of diffusion processes. See, e.g., Dixit (1993), Dixit and Pindyck (1994), Chapter 9.1 of Peskir and Shiryaev (2006) and Strulovici and Szydlowski (2015) for more discussion. Particularly, in the single alternative case, the value matching and smooth pasting conditions are,

$$\begin{aligned} U(\underline{x}^+) &= U(\underline{x}^-), \\ U(\overline{x}^+) &= U(\overline{x}^-), \\ U'(\underline{x}^+) &= U'(\underline{x}^-), \\ U'(\overline{x}^+) &= U'(\overline{x}^-). \end{aligned}$$

We can determine $\underline{x}$, $\overline{x}$, C_1 and C_2 by substituting $U(x)$ into the four equations above, which leads to (9) and (10).

Smooth pasting conditions at the adoption boundary and derivation of (16) and (17) We provide an intuitive argument for the smooth-pasting condition at the adoption boundary of alternative 1. Consider a DM spending time dt in learning alternative 1 at the boundary $\left(\overline{X}(x_2), x_2\right)$. The corresponding belief update dx_1 can be either positive or negative, with equal odds. If $dx_1 \ge 0$, the DM will adopt alternative 1 immediately; otherwise if $dx_1 < 0$, the DM will stay in the market learning more information on alternative 1. Therefore, the DM's expected payoff given that he will spend dt to learn alternative 1 would be:

$$\begin{aligned} V_1\left(\overline{X}(x_2), x_2\right) &\equiv -c\,dt + \frac{1}{2}\left[\Delta\pi\left(\overline{X}(x_2) + \mathrm{E}[dx_1 | dx_1 \ge 0]\right) + \underline{x}\right] \\ &\quad + \frac{1}{2}\mathrm{E}\left[V(\overline{X}(x_2) + dx_1, x_2) | dx_1 < 0\right] \\ &= V\left(\overline{X}(x_2), x_2\right) + \frac{\Delta\pi}{2\sigma}\overline{X}(x_2)\left[1 - \overline{X}(x_2)\right]\sqrt{\frac{dt}{2\pi}}\left[\Delta\pi - V_{x_1}\left(\overline{X}(x_2), x_2\right)\right] \\ &\quad + o(\sqrt{dt}), \end{aligned}$$

where we have used that (1) $\mathrm{E}[dW | dW \ge 0] = -\mathrm{E}[dW | dW < 0] = \sqrt{\frac{dt}{2\pi}}$ for $\{W(t) | t \ge 0\}$ being a standard Brownian motion; (2) the value matching condition $V\left(\overline{X}(x_2), x_2\right) = \Delta\pi\overline{X}(x_2)$.

On the other hand, let us consider a DM who spends dt in learning alternative 2 at the boundary $\left(\overline{X}(x_2), x_2\right)$. If the resulting belief update $dx_2 \ge 0$, the DM's adoption threshold for alternative 1 increases, so he will continue to learn alternative 1; otherwise, if $dx_2 < 0$, the DM will adopt alternative 1 immediately. Therefore, the DM's expected payoff given that he will spend dt to learn alternative 2 would be:

$$V_2\left(\overline{X}(x_2), x_2\right) \equiv -cdt + \frac{1}{2}\left[\Delta\pi \overline{X}(x_2) + \underline{x}\right] + \frac{1}{2}\mathrm{E}\left[V(\overline{X}(x_2), x_2 + dx_2)|dx_2 < 0\right]$$
$$= V\left(\overline{X}(x_2), x_2\right) - \frac{\Delta\pi}{2\sigma}x_2(1-x_2)\sqrt{\frac{dt}{2\pi}}V_{x_2}\left(\overline{X}(x_2), x_2\right) + o(\sqrt{dt}).$$

The DM chooses which alternative to learn on based on expected payoff maximization. By definition, his value function should satisfy:

$$V\left(\overline{X}(x_2), x_2\right) = \max\{V_1\left(\overline{X}(x_2), x_2\right), V_2\left(\overline{X}(x_2), x_2\right)\}.$$

By substituting the expression of $V_1\left(\overline{X}(x_2), x_2\right)$ and $V_2\left(\overline{X}(x_2), x_2\right)$ into the above equation, we have

$$\max\{\overline{X}(x_2)\left[1-\overline{X}(x_2)\right]\left[\Delta\pi - V_{x_1}\left(\overline{X}(x_2), x_2\right)\right], x_2(1-x_2)V_{x_2}\left(\overline{X}(x_2), x_2\right)\} = 0.$$

Meanwhile, by taking derivative of both sides of the value matching condition $V\left(\overline{X}(x_2), x_2\right) = \Delta\pi\overline{X}(x_2)$ with respect to x_2, we have

$$\overline{X}'(x_2)\left[\Delta\pi - V_{x_1}\left(\overline{X}(x_2), x_2\right)\right] = V_{x_2}\left(\overline{X}(x_2), x_2\right).$$

Combining the above two equations, we obtain that $V_{x_1}\left(\overline{X}(x_2), x_2\right) = \Delta\pi$ and $V_{x_2}(\overline{X}(x_2), x_2) = 0$, which are the smooth pasting conditions at the adoption boundary of alternative 1. It is straightforward to derive (16) and (17) based on the value matching and smooth pasting conditions.

Redundancy of one smoothness condition As shown in (13), the general solution to the parabolic PDE in (12) can be written as the following:

$$V(x_1, x_2) = A_0(x_1) + A_1(x_2)x_1 + A_2(x_2),$$

where $A_0(x_1) = \frac{2\sigma^2 c}{(\Delta\pi)^2}(1-2x_1)\ln\left(\frac{1-x_1}{x_1}\right)$. The value matching and smooth pasting conditions at the adoption boundary of alternative 1, $x_1 = \overline{X}(x_2)$ are:

$$A_0(\overline{X}(x_2)) + A_1(x_2)\overline{X}(x_2) + A_2(x_2) = \Delta\pi\overline{X}(x_2) + \overline{\pi}, \tag{vii}$$
$$A_0'(\overline{X}(x_2)) + A_1(x_2) = \Delta\pi, \tag{viii}$$
$$A_1'(x_2)\overline{X}(x_2) + A_2'(x_2) = 0. \tag{ix}$$

We will show that given (vii) and (viii), (ix) is redundant. In fact, taking derivative of both sides of (viii) with respect to x_2, we have,

$$A_0'(\overline{X}(x_2))\overline{X}'(x_2) + A_1'(x_2)\overline{X}(x_2) + A_1(x_2)\overline{X}'(x_2) + A_2'(x_2) = \Delta\pi\overline{X}'(x_2).$$

i.e.,

$$\left[A_0'(\overline{X}(x_2)) + A_1(x_2) - \Delta\pi\right]\overline{X}'(x_2) + A_1'(x_2)\overline{X}(x_2) + A_2'(x_2) = 0.$$

By substituting (viii) to the above equation, we get (ix). □

Proof of Lemma 2: existence and uniqueness of a solution to the ODE boundary value problem in (21), (22), and (23). Denote the right hand sides of (21) and (22) as $f_1(x, \overline{X}(x))$ and $f_2(x, \overline{X}(x))$, respectively. As $\overline{X}\left(\frac{1}{2}\right) = \Phi^{-1}\left(-\frac{(\Delta\pi)^3}{4\sigma^2 c}\right) > \Phi^{-1}(0) = 1/2$, we have that there exists $\varepsilon > 0$ such that f_1 is Lipschitz continuous on $\overline{X}$ and continuous on x for all $x \in [1/2-\varepsilon, 1/2]$ and that f_2 is Lipschitz continuous on $\overline{X}$ and continuous on x for all $x \in [1/2, 1/2+\varepsilon]$. Note also that $\overline{X}(\cdot)$ is continuous, and continuously differentiable at $x = 1/2$. By applying the Picard–Lindelöf theorem, we have that the solution to the boundary value problems in (21), (22), and (23) exists uniquely in the neighborhood of $x = 1/2$.

Our objective is to show that the solution exists uniquely for all $x \in (x^*, 1)$. Let us consider the boundary value problem of (22) and (23) first, and we will prove that the solution exists uniquely in $[1/2, 1)$ and satisfies that $\overline{X}(x) > x$ for $x \in [1/2, 1)$. To prove this, we can apply the Picard–Lindelöf theorem iteratively for $[1/2, z)$ for any $z \in (1/2, 1)$. Consequently, there are two possibilities.

(1) $\overline{X}(x) > x$ for $x \in [1/2, z)$. This implies that f_2 is Lipschitz continuous on $\overline{X}$ and continuous on x for all $x \in [1/2, z)$. In this case, we will be able to cover the entire interval of $[1/2, z)$ by applying Picard–Lindelöf theorem iteratively, because there is a positive lower bound on the neighborhood size ε_0 for each iteration. Particularly, $\varepsilon_0 \geq \frac{1}{2}\max\{1/M, 1/L\}$, where M is the maximum absolute value of the slope of f_1 with respect to $x \in [1/2, z)$ and $\overline{X} \in (x, 1)$, and L is the Lipschitz constant of f_1 with respect to $\overline{X}$. For any $z \in (1/2, 1)$, and given $\overline{X}(x) > x$, we can show that both M and L are positive and finite, so ε_0 is positive and finite. This implies that by applying Picard–Lindelöf theorem iteratively, we are able to show that there exists a unique solution in $[1/2, z)$.

(2) There exists $x' \in [1/2, z)$ such that $\overline{X}(x') \leq x'$. We will show that this is impossible. First, let us define $\hat{x}$ as the infimum of all x' that satisfies that $\overline{X}(x') \leq x'$. By definition, we must have that $\overline{X}(x) > x$ for $x \in [1/2, \hat{x})$. Moreover, $\lim_{x\to\hat{x}^-} \overline{X}(x) = \hat{x}$; otherwise, the solution will exist uniquely at $\hat{x}$, and $\overline{X}(\hat{x}) > \hat{x}$. Meanwhile, by (22), we have that $\lim_{x\to\hat{x}^-} \overline{X}'(x) = +\infty$. This implies that there exists $\varepsilon > 0$, such that $\overline{X}'(x) > 1$ for $x \in [\hat{x} - \varepsilon, \hat{x})$. As a result, we have,

$$\lim_{x\to\hat{x}^-} \overline{X}(x) - \hat{x} = \overline{X}(\hat{x} - \varepsilon) - (\hat{x} - \varepsilon) + \int_{\hat{x}-\varepsilon}^{\hat{x}^-} (\overline{X}'(x) - 1)dx > 0,$$

which contradicts the fact that $\lim_{x\to\hat{x}^-} \overline{X}(x) = \hat{x}$.

So far, we have proved that there exists a unique solution $\overline{X}(x)$ to (22) and (23) and $\overline{X}(x) > x$ for all $x \in [1/2, z)$ for $\forall z \in (1/2, 1)$. We can let z go to one, and this completes the proof. Following a similar analysis, we can show that there exists a unique solution $\overline{X}(x)$ to the other boundary value problem of (21) and (23) and $\overline{X} > 1 - x \geq x$ for $x \in (x^*, 1/2]$. Notice that as x goes to x^*, $\overline{X}(x)$ goes to $1 - x^*$, and the denominator of (21) goes to zero. Moreover, the numerator of (21) also goes to zero. By L'Hôpital's rule, we can show that $\overline{X}'(x^*) = 1$.

Smoothness of $\overline{X}(x)$ is easy to verify by taking derivatives for $x \in (x^*, 1/2)$ and $x \in (1/2, 1)$, and it is also straightforward to verify that at $x = 1/2$, $\overline{X}(x)$ is smooth. Therefore, $\overline{X}(x)$ is smooth in $x \in (x^*, 1)$.

Let us prove the monotonicity of $\overline{X}(x)$. By (18) and the monotonicity of $V(x_1, x_2)$, we have that

$$\frac{\partial V(x_1, x_2)}{\partial x_2} = \frac{2\sigma^2 c}{(\Delta\pi)^2} \frac{\left(\overline{X}(x_2) - x_1\right)\overline{X}'(x_2)}{\overline{X}(x_2)^2\left(1 - \overline{X}(x_2)\right)^2} \geq 0, \text{ for } \underline{x} \leq x_2 \leq x_1 \leq \overline{X}(x_2),$$

which implies $\overline{X}'(x) \geq 0$ for $x \in [1/2, 1)$. Similarly, by (19) and the monotonicity of $V(x_1, x_2)$, we can show that $\overline{X}'(x) \geq 0$ for $x \in (x^*, 1/2]$. Now let us prove $\overline{X}(1^-) = 1$. In fact, we have proved that $\overline{X}(x) > x$ for any $x \in [1/2, 1)$ and we know that $\overline{X}(x) \leq 1$ because it is the adoption boundary. When x goes to one, we then obtain that $\overline{X}(1^-) = 1$. □

Theorem A1. *Consider the decision maker's optimal learning problem in (6) with two symmetric alternatives. When the value of the outside option is relatively low, i.e., $x_0 \leq 0$, the value function $V(x_1, x_2)$ is given by*

$$V(x_1, x_2) = \begin{cases} \dfrac{2\sigma^2 c}{(\Delta\pi)^2}(1-2x_1)\left[\ln\left(\dfrac{1-x_1}{x_1}\right) - \ln\left(\dfrac{1-\overline{X}(x_2)}{\overline{X}(x_2)}\right)\right] \\ -\dfrac{2\sigma^2 c}{(\Delta\pi)^2}\dfrac{1-2\overline{X}(x_2)}{1-\overline{X}(x_2)}\dfrac{\overline{X}(x_2)-x_1}{\overline{X}(x_2)} + \Delta\pi\cdot x_1, \quad 1-x_1 \leq x_2 \leq x_1 \leq \overline{X}(x_2) \\ \dfrac{2\sigma^2 c}{(\Delta\pi)^2}(1-2x_2)\left[\ln\left(\dfrac{1-x_2}{x_2}\right) + \ln\left(\dfrac{1-\overline{X}(1-x_1)}{\overline{X}(1-x_1)}\right)\right] \\ +\dfrac{2\sigma^2 c}{(\Delta\pi)^2}\dfrac{1-2\overline{X}(1-x_1)}{1-\overline{X}(1-x_1)}\dfrac{1-\overline{X}(1-x_1)-x_2}{\overline{X}(1-x_1)} + \Delta\pi\cdot x_1, \\ x_2 \leq 1-x_1 \leq 1-x_2 \leq \overline{X}(1-x_1) \\ \dfrac{2\sigma^2 c}{(\Delta\pi)^2}(1-2x_2)\left[\ln\left(\dfrac{1-x_2}{x_2}\right) - \ln\left(\dfrac{1-\overline{X}(x_1)}{\overline{X}(x_1)}\right)\right] \\ -\dfrac{2\sigma^2 c}{(\Delta\pi)^2}\dfrac{1-2\overline{X}(x_1)}{1-\overline{X}(x_1)}\dfrac{\overline{X}(x_1)-x_2}{\overline{X}(x_1)} + \Delta\pi\cdot x_2, \quad 1-x_2 \leq x_1 \leq x_2 \leq \overline{X}(x_1) \\ \dfrac{2\sigma^2 c}{(\Delta\pi)^2}(1-2x_1)\left[\ln\left(\dfrac{1-x_1}{x_1}\right) + \ln\left(\dfrac{1-\overline{X}(1-x_2)}{\overline{X}(1-x_2)}\right)\right] \\ +\dfrac{2\sigma^2 c}{(\Delta\pi)^2}\dfrac{1-2\overline{X}(1-x_2)}{1-\overline{X}(1-x_2)}\dfrac{1-\overline{X}(1-x_2)-x_1}{\overline{X}(1-x_2)} + \Delta\pi\cdot x_2, \\ x_1 \leq 1-x_2 \leq 1-x_1 \leq \overline{X}(1-x_2) \\ \Delta\pi\cdot\max\{x_1, x_2\}, \quad \textit{otherwise}, \end{cases} \tag{x}$$

where $\overline{X}(x)$ is determined by the two boundary value problems in (21), (22) and (23).

Proof of Theorem A1. According to the way we constructed the solution, the only remaining thing in need of a proof is the condition in (24). In fact, given $x_1 + x_2 \geq 1$ and $x_2 \leq x_1 \leq \overline{X}(x_2)$, $V(x_1, x_2)$ is given by (18). There are two cases to consider: $1/2 \leq x_2 \leq x_1 \leq \overline{X}(x_2)$ and $1/2 \leq 1 - x_2 \leq x_1 \leq \overline{X}(x_2)$. The proofs are very similar for the two cases. Below, we only provide the proof for the first case with $1/2 \leq x_2 \leq x_1 \leq \overline{X}(x_2)$, where $\overline{X}(x)$ is determined by the boundary value problem in (22) and (23).

Let us simplify the condition in equation (24) first. By taking partial derivative with respect to x_2 on both sides of (18), we have that,

$$V_{x_2}(x_1, x_2) = \frac{2c\sigma^2}{(\Delta\pi)^2} \frac{(\overline{X}(x_2) - x_1)\overline{X}'(x_2)}{\overline{X}(x_2)^2(1 - \overline{X}(x_2))^2}$$
$$= \frac{2c\sigma^2}{(\Delta\pi)^2} \frac{\overline{X}(x_2) - x_1}{\overline{X}(x_2) - x_2} \left(\Phi\left(\overline{X}(x)\right) - \Phi(x) + \frac{(\Delta\pi)^3}{2\sigma^2 c}\right),$$

where the second equality is by using (22). By taking the partial derivative with respect to x_2 and using (21) again, we have that,

$$V_{x_2x_2}(x_1, x_2) = \frac{2c\sigma^2}{(\Delta\pi)^2}\left[\frac{\overline{X}(x_2) - x_1}{\overline{X}(x_2) - x_2}\frac{1}{x_2^2(1 - x_2)^2} + \frac{(x_1 - x_2)\overline{X}(x_2)^2\left(1 - \overline{X}(x_2)\right)^2}{\left(\overline{X}(x_2) - x_2\right)^3}\left(\Phi\left(\overline{X}(x)\right) - \Phi(x) + \frac{(\Delta\pi)^3}{2\sigma^2 c}\right)^2\right].$$

Using this expression of $V_{x_2x_2}(x_1, x_2)$, the condition in (24) can be written as

$$\frac{\overline{X}(x_2) - x_1}{\overline{X}(x_2) - x_2} + \frac{(x_1 - x_2)x_2^2(1 - x_2)^2\overline{X}(x_2)^2\left(1 - \overline{X}(x_2)\right)^2}{\left(\overline{X}(x_2) - x_2\right)^3}\left(\Phi\left(\overline{X}(x)\right) - \Phi(x) + \frac{(\Delta\pi)^3}{2\sigma^2 c}\right)^2 \leq 1.$$

By multiplying $\left(\overline{X}(x_2) - x_2\right)$ and rearranging, the inequality above is equivalent to,

$$(x_1 - x_2)\left[\frac{x_2^2(1 - x_2)^2\overline{X}(x_2)^2\left(1 - \overline{X}(x_2)\right)^2}{\left(\overline{X}(x_2) - x_2\right)^2}\left(\Phi\left(\overline{X}(x)\right) - \Phi(x) + \frac{(\Delta\pi)^3}{2\sigma^2 c}\right)^2 - 1\right] \leq 0.$$

Because $x_1 \geq x_2$, this is equivalent to

$$\frac{x_2^2(1 - x_2)^2\overline{X}(x_2)^2\left(1 - \overline{X}(x_2)\right)^2}{\left(\overline{X}(x_2) - x_2\right)^2}\left(\Phi\left(\overline{X}(x)\right) - \Phi(x) + \frac{(\Delta\pi)^3}{2\sigma^2 c}\right)^2 \leq 1,$$

which is equivalent to

$$\frac{x_2(1 - x_2)\overline{X}(x_2)\left(1 - \overline{X}(x_2)\right)}{\overline{X}(x_2) - x_2}\left(\Phi\left(\overline{X}(x)\right) - \Phi(x) + \frac{(\Delta\pi)^3}{2\sigma^2 c}\right) \leq 1,$$

which in turn, is equivalent to

$$H(x_2) \equiv x_2(1 - x_2)\left(\Phi\left(\overline{X}(x_2)\right) - \Phi(x_2) + \frac{(\Delta\pi)^3}{2\sigma^2 c}\right) - \frac{\overline{X}(x_2) - x_2}{\overline{X}(x_2)\left(1 - \overline{X}(x_2)\right)} \leq 0. \quad \text{(xi)}$$

Notice that the condition in (xi) is in terms of x_2 solely. To ensure (24) for any x_1 and x_2 that satisfy $1/2 \leq x_2 \leq x_1 \leq \overline{X}(x_2)$, we only need to ensure (xi) for any $x_2 \in [1/2, 1)$.

Let us define $m = (\Delta\pi)^3/(4\sigma^2 c)$. By (23), we can calculate that $H(1/2) = m/4 - (\Phi^{-1}(-m) - 1/2)/\left[\Phi^{-1}(-m)(1 - \Phi^{-1}(-m))\right]$. Using the definition of $\Phi(\cdot)$, it is not difficult to show that $H(1/2) < 0$. Now, let us prove (xi) by contradiction. Suppose it does not hold for some $x_2 \in [1/2, 1)$. Since $H(x)$ is continuous, this implies that there exists $x_2^* \equiv \min_{1/2 \leq x_2 < 1}\{x_2 | H(x_2) = 0\}$ such that $H'(x_2^*) > 0$. By taking derivative on both sides of (xi), and using (21) as well as $H(x_2^*) = 0$, we can simplify and get the following expression:

$$H'(x_2^*) = -2\left(\Phi\left(\overline{X}(x_2^*)\right) - \Phi(x_2^*) + \frac{(\Delta\pi)^3}{2\sigma^2 c}\right)\left(x_2^* + \overline{X}(x_2^*) - 1\right) > 0.$$

However, from the proof of Lemma 2, we know that $\Phi\left(\overline{X}(x_2^*)\right) - \Phi(x_2^*) + \frac{(\Delta\pi)^3}{2\sigma^2 c} \geq 0$, and $x_2^* + \overline{Y}(x_2^*) - 1 \geq 2x_2^* - 1 \geq 0$. This is a contradiction to the inequality above, which implies that (xi) holds for any $x_2 \in [1/2, 1)$. □

Construction of the belief updating process under the optimal learning strategy Consider the case of Theorem A1 when the value of the outside option is relatively low. We have already divided the belief space of x_1-x_2 into six regions corresponding to the DM's five choices of action. Updating the beliefs of both alternatives ceases when entering the regions where it is optimal to adopt 1, or adopt 2. Therefore, we only need to construct the belief updating process for the regions where it is optimal to learn 1 or learn 2. Equation (7) can be equivalently rewritten as the following:

$$dx_i(t) = \frac{\Delta\pi}{\sigma^2} x_i(t)\left[1 - x_i(t)\right] \mathbb{1}_{L_i}(x_1(t), x_2(t)) \left\{\left[\pi_i - \overline{\pi} x_i(t)\right] dt + \sigma dW_i(t)\right\}, \qquad \text{(xii)}$$

where $\mathbb{1}_{L_i}(x_1(t), x_2(t))$ is an indicator function. It is equal to one when $(x_1(t), x_2(t)) \in L_i$, i.e., when the DM's beliefs are in the region where it is optimal to learn alternative i; it takes zero otherwise. We have $L_1 \equiv \left\{x_1, x_2 | 1 - x_1 \leq x_2 \leq x_1 \leq \overline{X}(x_2) \text{ or } x_1 \leq 1 - x_2 \leq 1 - x_1 \leq \overline{X}(1 - x_2\right\}$ and $L_2 \equiv \left\{x_1, x_2 | x_2 < 1 - x_1 < 1 - x_2 < \overline{X}(1 - x_1) \text{ or } 1 - x_2 < x_1 < x2 < \overline{X}(x_1)\right\}$.[21] Furthermore, $\{W_i(t); t \geq 0\}$ is a standard Brownian motion, and $W_1(\cdot)$ is independent of $W_2(\cdot)$.

It is interesting to see what happens to the stochastic process when we are at $x_1 = x_2$ or at $x_1 + x_2 = 1$. Let us consider the case where we start at $x_1, x_2 > 1/2$. Given the optimal learning strategies, let us investigate what happens to $\min\{x_1(t), x_2(t)\}$ conditional on the DM's belief being in L_1 or L_2 and $x_1, x_2 > 1/2$. In terms of the SDE in (xii), let us define $\widetilde{a}_i(x_i(t)) \equiv \frac{\Delta\pi}{\sigma^2} x_i(t)[1 - x_i(t)][\pi_i - \overline{\pi} x_i(t)]$, for $i = 1, 2$, and $\widetilde{b}(x_i(t)) \equiv \frac{\Delta\pi}{\sigma} x_i(t)[1 - x_i(t)]$ such that we can write equation (xii) as

$$dx_i(t) = \mathbb{1}_{L_i}(x_1(t), x_2(t))[\widetilde{a}_i(x_i(t))dt + \widetilde{b}(x_i(t))dW_i(t)] \quad (i = 1, 2).$$

The particular case where $\widetilde{a}_1 = \widetilde{a}_2$, and $\widetilde{b}$ are constants is considered in Fernholz et al. (2013), which presents pathwise unique solutions for $x_1(t)$ and $x_2(t)$.

To see the stochastic behavior of $\min\{x_1(t), x_2(t)\}$, we follow the presentation in Fernholz et al. (2013), p. 357, with the adjustments that $\widetilde{a}_1(x_1(t)) \neq \widetilde{a}_2(x_1(t))$, and $\widetilde{a}_i(x_i(t))$ for $i = 1, 2$, and $\widetilde{b}(x_i(t))$ are bounded and continuous functions of $x_i(t)$, and where $x_1(t)$ and $x_2(t)$ are unique-in-distribution solutions to (xii) for $i = 1, 2$. Let us define $Y(t) \equiv x_1(t) - x_2(t)$. Then we can obtain

$$Y(t) = Y(0) + \int_0^t \widetilde{a}_1(x_1(s))\mathbb{1}_{[Y(s)\geq 0]}ds - \int_0^t \widetilde{a}_2(x_2(s))\mathbb{1}_{[Y(s)<0]}ds + \widetilde{W}(t), \qquad \text{(xiii)}$$

[21] This sets up the model such that in the points of indifference alternative 1 is the one searched. This set-up with one alternative moving is as in Fernholz et al. (2013). Alternatively, one could have that in the points of indifference alternative 2 is the one moving, or that either alternative is searched with equal probability.

where $\widetilde{W}(t) \equiv \int_0^t \widetilde{b}(x_1(s))\mathbb{1}_{[Y(s)\geq 0]}dW_1(s) - \int_0^t \widetilde{b}(x_2(s))\mathbb{1}_{[Y(s)<0]}dW_2(s)$. We can also obtain

$$x_1(t)+x_2(t) = x_1(0)+x_2(0)+\int_0^t \widetilde{a}_1(x_1(s))\mathbb{1}_{[Y(s)\geq 0]}ds+\int_0^t \widetilde{a}_2(x_2(s))\mathbb{1}_{[Y(s)<0]}ds+\widehat{W}(t), \tag{xiv}$$

where $\widehat{W}(t) \equiv \int_0^t \widetilde{b}(x_1(s))\mathbb{1}_{[Y(s)\geq 0]}dW_1(s) + \int_0^t \widetilde{b}(x_2(s))\mathbb{1}_{[Y(s)<0]}dW_2(s)$. The processes $\widetilde{W}(t)$ and $\widehat{W}(t)$ are continuous martingales, and then $Y(t)$ is a semimartingale. Using the definitions $\widetilde{W}(t)$ and $\widehat{W}(t)$ we can also obtain

$$\widehat{W}(t) = \int_0^t \mathrm{sgn}(Y(s))d\widetilde{W}(s), \tag{xv}$$

where the signum function is defined as $\mathrm{sgn}(y) \equiv \mathbb{1}_{[0,\infty)}(y) - \mathbb{1}_{(-\infty,0)}(y)$.

Note now that

$$|Y(t)| = |Y(0)| + \int_0^t \mathrm{sgn}(Y(s))dY(s) + 2L^Y(t) \tag{xvi}$$

by Tanaka's formula (Karatzas and Shreve 1991, p. 205) applied to semimartingales, and where $L^Y(t)$ is the local time of $Y(t)$ at $Y(t) = 0$, which is defined as $L^Y(t) \equiv \lim_{\varepsilon\downarrow 0} \frac{1}{2\varepsilon}\int_0^t \mathbb{1}_{\{0\leq Y(s)<\varepsilon\}}ds$. Substituting the derivative of (xiii) and (xv) into (xvi) we can obtain

$$|Y(t)| = |Y(0)| + \int_0^t \widetilde{a}_1(x_1(s))\mathbb{1}_{[Y(s)\geq 0]}ds + \int_0^t \widetilde{a}_2(x_2(s))\mathbb{1}_{[Y(s)<0]}ds + \widehat{W}(t) + 2L^Y(t). \tag{xvii}$$

Noting that $\min\{x_1(t), x_2(t)\} = 1/2[x_1(t) + x_2(t) - |x_1(t) - x_2(t)|]$, we can then use (xiv) and (xvii) to obtain

$$\min\{x_1(t), x_2(t)\} = \min\{x_1(0), x_2(0)\} - L^Y(t),$$

which shows that $\min\{x_1(t), x_2(t)\}$, conditional on the DM being in the region where it is optimal to learn more information and $x_1, x_2 > 1/2$, falls over time under the optimal learning strategy.[22]

By the same argument we can also look at the evolution of the stochastic process in the regions L_i where $x_1 < 1/2$ and/or $x_2 < 1/2$. For example, for the case of $x_1, x_2 < 1/2$, the evolution close to $x_1(t) = x_2(t)$ has to do with the $\max\{x_1(t), x_2(t)\}$ and we can obtain that, given the optimal learning in those regions, $\max\{x_1(t), x_2(t)\} = \max\{x_1(0), x_2(0)\} + L^Y(t)$, increases over time. In the region where $x_1 > x_2$ and $x_2 < 1/2$, we can obtain that $\max\{x_2, 1 - x_1\}$ increases over time, and in the region where $x_2 > x_1$ and $x_1 < 1/2$, we can obtain that $\max\{x_1, 1 - x_2\}$ increases over time.

In those all these regions in L_1 and L_2 the process evolves towards $x_1 = x_2 = 1/2$. Once $x_1 = x_2 = 1/2$ is reached the DM is indifferent on which alternative to gather information on. Given

[22] Note that local time appears on the process $\min\{x_1(t), x_2(t)\}$ but it does not appear in the formulation of the processes $x_1(t)$ and $x_2(t)$, in (7). This is exactly as in, for example, Mandelbaum et al. (1990) and Fernholz et al. (2013).

our definition of the optimal policy, where in case of indifference the DM gathers information on alternative 1, once the process reaches $x_1 = x_2 = 1/2$ the DM only gathers information on alternative 1 going forward.

Proof of Proposition 1: comparative statics. Let us first show that $V(x_1, x_2)$ decreases with c. This is straightforward by a simple argument. Consider $c' > c''$. Under learning cost c'', the DM can always replicate the optimal strategy under c'' but now incurs less learning cost.

Next, we prove that $\overline{X}(x)$ decreases with c. In fact, let us consider any x_1 and x_2 satisfying $1 - x_1 \leq x_2 \leq x_1 \leq \overline{X}(x_2)$. By (18), we have,

$$\frac{\partial V(x_1, x_2)}{\partial c} = \frac{V(x_1, x_2) - \Delta\pi \cdot x_1}{c} + \frac{2\sigma^2 c}{(\Delta\pi)^2} \frac{\overline{X}(x_2) - x_1}{\overline{X}(x_2)^2 \left(1 - \overline{X}(x_2)\right)^2} \frac{\overline{X}(x_2)}{\partial c} \leq 0.$$

Since $V(x_1, x_2) \geq \Delta\pi \cdot x_1$, the inequality above implies that $\overline{X}(x_2)/\partial c \leq 0$ for $\forall x_2 \in [x^*, 1]$. This further implies that $\overline{Y}(x)/\partial c \leq 0$ for $\forall x \in [0, 1]$.

Lastly, notice that the HJB equation (8) can be equivalently written as,

$$\max\left\{ \max_{1 \leq i \leq n} \left\{ \frac{(\Delta\pi_i)^2}{2\sigma_i^2 c_i} x_i^2 (1 - x_i)^2 \widehat{V}_{x_i x_i}(\mathbf{x}) - 1 \right\}, g(\mathbf{x}) - \widehat{V}(\mathbf{x}) \right\} = 0,$$

so only $\sigma^2 c$ is identified in the model. Therefore, the comparative statics with respect to σ will be the same with that with c. □

Proof of Lemma 4. Let us prove the condition in (34) first. In fact, following the same derivation, we can show that the condition in (33) is equivalent to that in (xi). By using (31), (xi) can be equivalently written as (34).

Now we prove the existence of x_0^* from inequality (xi). To guarantee inequality (xi) for any $x_2 \in [\underline{x}, 1]$, we need

$$\max_{\underline{x} \leq x_2 \leq 1} H(x_2) \leq 0. \tag{xviii}$$

Note that $H(\underline{x}) = -(\overline{x} - \underline{x})/[\overline{x}(1 - \overline{x})] < 0$ and $H(x)$ is continuous, so the inequality (xviii) will be violated if and only if there exists $x_2^* \equiv \min_{\underline{x} \leq x_2 \leq 1}\{x_2 | H(x_2) = 0\}$, such that $H'(x_2^*) > 0$. By using (31) and the condition that $H(x_2^*) = 0$, we can simplify and get the following expression,

$$H'(x_2^*) = -2\left(\Phi\left(\overline{X}(x_2^*)\right) - \Phi(x_2^*) + \frac{(\Delta\pi)^3}{2\sigma^2 c}\right)\left(x_2^* + \overline{X}(x_2^*) - 1\right).$$

From Lemma 3, we know that $\Phi\left(\overline{X}(x_2^*)\right) - \Phi(x_2^*) + \frac{(\Delta\pi)^3}{2\sigma^2 c} \geq 0$. Let us define $h(x) \equiv x + \overline{X}(x) - 1$. Therefore, the equality above implies that

$$H'(x_2^*) > 0 \Leftrightarrow h(x_2^*) < 0.$$

Meanwhile, from (9) and (10), it is easy to show that

$$\underline{x} + \overline{x} \geq 1 \Leftrightarrow x_0 \geq 1/2.$$

Therefore, when $x_0 \geq 1/2$, $h(\underline{x}) = \underline{x} + \overline{x} - 1 \geq 0$. According to Lemma 3, we also know that $h(x)$ is strictly increasing. This implies that when $x_0 \geq 1/2$, $h(x) \geq 0$ for any $x \in [\underline{x}, 1]$. This in turn implies that $h(x_2^*) \geq 0$, and inequality (xviii) must hold. Therefore, if x_0^* exists, it must be less than 1/2.

Next, we will prove the existence of x_0^*. We also note that

$$\begin{aligned}
\frac{dh(x_2^*)}{d\pi_0} &= \frac{\partial \overline{X}(x)}{\partial \pi_0}\Big|_{x=x_2^*} + \left[1+\overline{X}'(x_2^*)\right]\frac{dx_2^*}{d\pi_0} \\
&= \frac{\partial \overline{X}(x)}{\partial \pi_0}\Big|_{x=x_2^*} - \frac{1+\overline{X}'(x_2^*)}{H'(x_2^*)} \times \frac{dH}{d\overline{X}}\Big|_{x=x_2^*} \times \frac{\partial \overline{X}(x)}{\partial \pi_0}\Big|_{x=x_2^*} \\
&= \left\{1 - \frac{\left[(\overline{X}(x_2^*)-x_2^*)^2 + 2x_2^*(1-x_2^*)\right]\left[\overline{X}(x_2^*)\left(1-\overline{X}(x_2^*)\right) + x_2^*(1-x_2^*)\right]}{2h(x_2^*)\left(\overline{X}(x_2^*)-x_2^*\right)\overline{X}(x_2^*)\left(1-\overline{X}(x_2^*)\right)}\right\} \\
&\quad \times \frac{\partial \overline{X}(x)}{\partial \pi_0}\Big|_{x=x_2^*}
\end{aligned}$$

As shown in Proposition 2, $\partial \overline{X}(x)/\partial \pi_0 \geq 0$, so $dh(x_2^*)/d\pi_0 \geq 0$ when $h(x_2^*) < 0$. Consider $x_0' > x_0''$. If under x_0', inequality (xviii) is violated, then there exists $x_2^* \in [\underline{x}, 1]$ such that $H'(x_2^*) > 0$, which implies that $h(x_2^*) < 0$, which in turn implies that $dh(x_2^*)/d\pi_0 \geq 0$. This implies that under $x_0'' < x_0'$, $h(x_2^*)$ will get even smaller, and inequality (xviii) will also be violated. In summary, to guarantee inequality (xviii), we must have $x_0 \geq x_0^*$ for some x_0^*.

The threshold x_0^* can be obtained as the x_0 that makes $\max_{x_2} H(x_2) = 0$, which is a function of x_0 as $\overline{X}(.)$ is a function of both $\underline{x}$ and $\overline{x}$ and both $\underline{x}$ and $\overline{x}$ are a function of x_0. □

Theorem A2. *Consider the decision maker's optimal learning problem in (6) with two symmetric alternatives. When the value of the outside option is relatively high, i.e., $x_0 \geq x_0^*$, the value function $V(x_1, x_2)$ is given by*

$$V(x_1,x_2) = \begin{cases} \frac{2\sigma^2 c}{(\Delta\pi)^2}(1-2x_1)\left[\ln\left(\frac{1-x_1}{x_1}\right) - \ln\left(\frac{1-\overline{X}(x_2)}{\overline{X}(x_2)}\right)\right] \\ -\frac{2\sigma^2 c}{(\Delta\pi)^2}\frac{1-2\overline{X}(x_2)}{1-\overline{X}(x_2)}\frac{\overline{X}(x_2)-x_1}{\overline{X}(x_2)} + \Delta\pi\cdot x_1, & \max\{x_2,\underline{x}\} \leq x_1 \leq \overline{X}(x_2) \\ \frac{2\sigma^2 c}{(\Delta\pi)^2}(1-2x_2)\left[\ln\left(\frac{1-x_2}{x_2}\right) - \ln\left(\frac{1-\overline{X}(x_1)}{\overline{X}(x_1)}\right)\right] \\ -\frac{2\sigma^2 c}{(\Delta\pi)^2}\frac{1-2\overline{X}(x_1)}{1-\overline{X}(x_1)}\frac{\overline{X}(x_1)-x_2}{\overline{X}(x_1)} + \Delta\pi\cdot x_2, & \max\{x_1,\underline{x}\} \leq x_2 \leq \overline{X}(x_1) \\ \Delta\pi\cdot\max\{x_1,x_2\}, & \textit{otherwise}, \end{cases} \tag{xix}$$

where $\overline{X}(x)$ is determined by the boundary value problem in (31) and (32) for $x \in [\underline{x}, 1]$, and $\overline{X}(x) = \overline{x}$ for $x \in [0, \underline{x}]$.

Proof of Proposition 2: comparative statics. We first show the comparative statics with respect to c. Let us initially show that $\underline{x}$ increases with c. In fact, by taking derivative with respect to c on both sides of (9) and (10) and combining the two resulting equations to cancel $\partial\overline{x}/\partial c$, we have

$$\frac{\partial \underline{x}}{\partial c} = \frac{\underline{x}^2(1-\underline{x})^2}{\overline{x}-\underline{x}} \frac{(\Delta\pi)^2 \Delta\pi \cdot \overline{x}}{2\sigma^2 c^2}.$$

To show $\partial \underline{x}/\partial c > 0$, we only need $\Delta\pi \cdot \overline{x} > 0$. In fact, let us consider any $x \in (\underline{x}, \overline{x})$. First, we know $U(x) \geq 0$, because the DM can always take the outside option and get zero. Also, by following the optimal strategy, the DM ends up either adopting the alternative, or taking the outside option, after paying the learning cost, so we have

$$U(x) = P(x)\Delta\pi \cdot \overline{x} + (1 - P(x))\, \Delta\pi \cdot \underline{x} - cT(x) \geq 0,$$

where $P(x)$ is the probability that the DM ends up adopting the alternative given his current posterior belief as x, and $T(x)$ is the expected time spent on learning. By the inequality above and $\overline{x} > \underline{x}$, we know that $\Delta\pi \cdot \overline{x} > 0$. Therefore, $\partial \underline{x}/\partial c > 0$.

The comparative statics of $\overline{X}(x)$ and $V(x_1, x_2)$ with respect to c can be similarly proved as in the proof of Proposition 1. Similarly, only $\sigma^2 c$ is identified in the model, so the comparative statics with respect to σ is the same with that of c.

Next, we prove the comparative statics with respect to $\overline{\pi}$. Let us first show that $\partial V(x_1, x_2)/\partial\overline{\pi} \leq 1$. In fact, given $\Delta\pi$ fixed, consider an increase of $\overline{\pi}$ by $\Delta\overline{\pi} > 0$. If the outside option π_0 also increased $\Delta\overline{\pi}$, then obviously, $V(x_1, x_2)$ would be shifted upward by exactly $\Delta\overline{\pi}$. Since the outside option is kept unchanged, $V(x_1, x_2)$ must increase no more than $\Delta\overline{\pi}$. This implies that $\partial V(x_1, x_2)/\partial\overline{\pi} \leq 1$. Consider any x_1 and x_2 satisfying $x_2 \leq x_1 \leq \overline{X}(x_2)$ and $x_1 \geq \underline{x}$. By (xix), we have,

$$\frac{\partial V(x_1, x_2)}{\partial\overline{\pi}} = 1 + \frac{2\sigma^2 c}{(\Delta\pi)^2} \frac{\overline{X}(x_2) - x_1}{\overline{X}(x_2)^2 \left(1 - \overline{X}(x_2)\right)^2} \frac{\overline{X}(x_2)}{\partial\overline{\pi}} \leq 1.$$

This implies that $\partial\overline{X}(x_2)/\partial\overline{\pi} \leq 0$ for $x_2 \in [0, 1]$. Because $\overline{X}(x_2) = \overline{x}$ for $x_2 \in [0, \underline{x}]$, we have $\partial\overline{x}/\partial\overline{\pi} \leq 0$. By taking derivative with respect to $\overline{\pi}$ on both sides of (10), we have

$$\frac{\partial \underline{x}}{\partial\overline{\pi}} = \frac{\underline{x}^2(1-\underline{x})^2}{\overline{x}^2(1-\overline{x})^2} \frac{\partial\overline{x}}{\partial\overline{\pi}} \leq 0.$$

Finally, we prove the comparative statics with respect to π_0. Obviously, $V(x_1, x_2)$ increases with the outside option π_0, so we have,

$$\frac{\partial V(x_1, x_2)}{\partial\pi_0} = \frac{2\sigma^2 c}{(\Delta\pi)^2} \frac{\overline{X}(x_2) - x_1}{\overline{X}(x_2)^2 \left(1 - \overline{X}(x_2)\right)^2} \frac{\overline{X}(x_2)}{\partial\pi_0} \geq 0.$$

This implies that $\partial\overline{X}(x)/\partial\pi_0 \geq 0$ for $x \in [0, 1]$, which in turn implies that $\partial\overline{x}/\partial\pi_0 \geq 0$. So we also have

$$\frac{\partial \underline{x}}{\partial\pi_0} = \frac{\underline{x}^2(1-\underline{x})^2}{\overline{x}^2(1-\overline{x})^2} \frac{\partial\overline{x}}{\partial\pi_0} \geq 0. \quad \square$$

Theorem A3. *Consider two symmetric alternatives. Under the optimal learning strategy, given his current posterior beliefs as* (x_1, x_2)*, a decision maker's probability of adopting alternative 1 is* $P_1(x_1, x_2)$*, where if* $x_0 \leq 0$*,*

$$P_1(x_1,x_2) = \begin{cases} \frac{1}{2}\frac{\overline{X}(x_2)+x_1-2x_2}{\overline{X}(x_2)-x_2}, & 1/2 \le x_2 \le x_1 \le \overline{X}(x_2) \\ 1-\frac{1}{2}\frac{\overline{X}(x_2)-x_1}{\overline{X}(x_2)+x_2-1}e^{\int_{\frac{1}{2}}^{x_2}\frac{2}{\overline{X}(\xi)+\xi-1}d\xi}, & 1/2 \le 1-x_2 \le x_1 \le \overline{X}(x_2) \\ \frac{1}{2}\frac{\overline{X}(1-x_1)+2x_1-x_2-1}{\overline{X}(1-x_1)+x_1-1}, & 1/2 \le 1-x_1 \le 1-x_2 \le \overline{X}(1-x_1) \\ 1-\frac{1}{2}\frac{\overline{X}(1-x_1)+x_2-1}{\overline{X}(1-x_1)-x_1}e^{\int_{\frac{1}{2}}^{1-x_1}\frac{2}{\overline{X}(\xi)+\xi-1}d\xi}, & 1/2 \le x_1 \le 1-x_2 \le \overline{X}(1-x_1) \\ 1-\frac{1}{2}\frac{\overline{X}(x_1)+x_2-2x_1}{\overline{X}(x_1)-x_1}, & 1/2 \le x_1 \le x_2 \le \overline{X}(x_1) \\ \frac{1}{2}\frac{\overline{X}(x_1)-x_2}{\overline{X}(x_1)+x_1-1}e^{\int_{\frac{1}{2}}^{x_1}\frac{2}{\overline{X}(\xi)+\xi-1}d\xi}, & 1/2 \le 1-x_1 \le x_2 \le \overline{X}(x_1) \\ 1-\frac{1}{2}\frac{\overline{X}(1-x_2)+2x_2-x_1-1}{\overline{X}(1-x_2)+x_2-1}, & 1/2 \le 1-x_2 \le 1-x_1 \le \overline{X}(1-x_2) \\ \frac{1}{2}\frac{\overline{X}(1-x_2)+x_1-1}{\overline{X}(1-x_2)-x_2}e^{\int_{\frac{1}{2}}^{1-x_2}\frac{2}{\overline{X}(\xi)+\xi-1}d\xi}, & 1/2 \le x_2 \le 1-x_1 \le \overline{X}(1-x_2) \\ \mathbb{1}_{x_1 \ge x_2}, & \textit{otherwise,} \end{cases} \tag{xx}$$

and if $x_0 \ge x_0^*$,

$$P_1(x_1,x_2) = \begin{cases} 0, & x_1 \le \underline{x} \text{ or } x_2 \ge \overline{X}(x_1) \\ \frac{1}{2}\frac{\overline{X}(x_1)-x_2}{\overline{X}(x_1)-x_1}\left[1-e^{-\int_{\underline{x}}^{x_1}\frac{2d\xi}{\overline{X}(\xi)-\xi}}\right], & \underline{x} < x_1 \le x_2 < \overline{X}(x_1) \\ \frac{x_1-x_2}{\overline{X}(x_2)-x_2}+\frac{1}{2}\frac{\overline{X}(x_2)-x_1}{\overline{X}(x_2)-x_2}\left[1-e^{-\int_{\underline{x}}^{x_2}\frac{2d\xi}{\overline{X}(\xi)-\xi}}\right], & \underline{x} < x_2 < x_1 < \overline{X}(x_2) \\ \frac{x_1-\underline{x}}{\overline{x}-\underline{x}}, & \underline{x} < x_1 < \overline{x} \text{ and } x_2 \le \underline{x} \\ 1, & x_1 \ge \overline{X}(x_2). \end{cases} \tag{xxi}$$

Proof of Theorem A3: adoption likelihood. Let us consider the case with $x_0 \ge x_0^*$ first. When $x_1 \ge \overline{X}(x_2)$, the DM adopts alternative 1 immediately, so $P_1(x_1,x_2)=1$. When $x_1 \le \underline{x}$ or $x_2 \ge \overline{X}(x_1)$, the DM will never adopt alternative 1, therefore $P_1(x_1,x_2)=0$. Otherwise when $\underline{x} < x_1 < \overline{X}(x_2)$ and $x_2 < \overline{X}(x_1)$, there are two cases, depending on the value of x_2.

In the first case with $x_2 \le \underline{x}$, the DM considers alternative 1 only. Given his current belief x_1, he will continue learning until his belief equals either $\underline{x}$ or $\overline{x}$. According to the Optional Stopping Theorem, we have $x_1 = P_1(x_1,x_2)\overline{x} + [1-P_1(x_1,x_2)]\underline{x}$, therefore

$$P_1(x_1,x_2) = \frac{x_1-\underline{x}}{\overline{x}-\underline{x}}, \quad \underline{x} < x_1 \le \overline{x}, x_2 \le \underline{x}.$$

In the second case, $x_2 > \underline{x}$. When $x_1 \geq x_2$, the DM keeps learning alternative 1 until his belief reaches either $\overline{X}(x_2)$ or x_2. Let us define the probability of reaching $\overline{X}(x_2)$ as $q_1(x_1, x_2)$. Then by invoking the Optional Stopping Theorem, we similarly get

$$q_1(x_1, x_2) = \frac{x_1 - x_2}{\overline{X}(x_2) - x_2}.$$

According to symmetry, the probability of reaching $\overline{X}(x_1)$, starting from (x_1, x_2) with $x_1 < x_2$, would be $q_1(x_2, x_1)$. Let us further define $P_0(x)$ as the probability of taking the outside option, given the DM's current beliefs as (x, x). Let us consider an infinitesimal learning on alternative 1 at (x, x), with belief update as dx. By conditioning on the sign of dx, we have

$$\begin{aligned}
P_0(x) &\equiv \Pr[\text{outside option}|(x, x)] \\
&= \frac{1}{2}\Pr[\text{outside option}|(x, x), dx \geq 0] + \frac{1}{2}\Pr[\text{outside option}|(x, x), dx < 0] \\
&= \frac{1}{2}[1 - q_1(x + |dx|, x)]\, P_0(x) + \frac{1}{2}[1 - q_1(x, x - |dx|)]\, P_0(x - |dx|) \\
&= P_0(x) - \frac{|dx|}{2}\left[P_0'(x) - \left(\frac{\partial q_1(x_1, x_2)}{\partial x_2} - \frac{\partial q_1(x_1, x_2)}{\partial x_1}\right)\Big|_{x_1 = x_2 = x} P_0(x)\right] + o(dx),
\end{aligned}$$

where the last equality is obtained by doing a Taylor expansion of $q_1(x + |dx|, x)$, $q_1(x, x - |dx|)$, and $P_0(x - |dx|)$. By canceling out $P_0(x)$, dividing by dx and taking limit of dx going to zero for the equation above, we have

$$\frac{P_0'(x)}{P_0(x)} = \left(\frac{\partial q_1(x_1, x_2)}{\partial x_2} - \frac{\partial q_1(x_1, x_2)}{\partial x_1}\right)\Big|_{x_1 = x_2 = x} = -\frac{2}{\overline{X}(x) - x}.$$

Combining the differential equation above with the initial condition $P_0(\underline{x}) = 1$, we can solve $P_0(x)$ as

$$P_0(x) = e^{-\int_{\underline{x}}^{x} \frac{2}{\overline{X}(\xi) - \xi} d\xi}.$$

Starting from (x_1, x_2) with $x_1 \geq x_2$, the DM learns information on alternative 1. With probability $q_1(x_1, x_2)$, he reaches the adoption boundary $\overline{X}(x_2)$, and immediately adopts alternative 1. With probability $1 - q_1(x_1, x_2)$, he reaches x_2. Then starting from (x_2, x_2), he eventually adopts alternative 1 with probability $1/2[1 - P_0(x_2)]$. Therefore, we have

$$P_1(x_1, x_2) = q_1(x_1, x_2) + [1 - q_1(x_1, x_2)]\frac{1}{2}[1 - P_0(x_2)], \quad \underline{x} < x_2 < x_1 < \overline{X}(x_2).$$

Similarly, starting from (x_1, x_2) with $x_1 < x_2$, the DM learns information on alternative 2. With probability $1 - q_1(x_2, x_1)$, he reaches x_1, upon which, he eventually adopts alternative 1 with probability $1/2[1 - P_0(x_1)]$.

$$P_1(x_1, x_2) = [1 - q_1(x_2, x_1)]\frac{1}{2}[1 - P_0(x_1)], \quad \underline{x} < x_1 < x_2 < \overline{X}(x_1).$$

Now, let us consider the other case with $x_0 \leq 0$. Obviously, in the region A1, $P_1(x_1, x_2) = 1$, and in the region A2, $P_1(x_1, x_2) = 0$.

Consider the region L1, there are two cases. In the first case, $1/2 \leq x_2 \leq x_1 \leq \overline{X}(x_2)$, the DM keeps learning alternative 1 until his belief reaches either $\overline{X}(x_2)$ or x_2. By the similar analysis above, we have that

$$P_1(x_1,x_2) = q_1(x_1,x_2) + \frac{1}{2}\left[1 - q_1(x_1,x_2)\right], \quad 1/2 \le x_2 \le x_1 \le \overline{X}(x_2).$$

In the second case, $1/2 \le 1 - x_2 \le x_1 \le \overline{X}(x_2)$, the DM keeps learning alternative 1 until his belief reaches either $\overline{X}(x_2)$ or $1 - x_2$. Given the DM's belief (x_1, x_2), the probability that he reaches $\overline{X}(x_2)$ is

$$r_1(x_1,x_2) = \frac{x_1 + x_2 - 1}{x_2 + \overline{X}(x_2) - 1}.$$

To calculate $P_1(x_1, x_2)$ in this case, we need to first consider another case with $1 - \overline{X}(1 - x_1) \le x_2 \le 1 - x_1 \le 1/2$ in the region of L2*. In this case, the DM keeps learning alternative 2 until his belief reaches either $1 - \overline{X}(1 - x_1)$ or $1 - x_1$. Given the DM's belief (x_1, x_2), the probability that he reaches $1 - \overline{X}(1 - x_1)$ is $r_1(1 - x_2, 1 - x_1)$.

Let us further define $P_D(x_2)$ as the probability of adopting alternative 1 eventually, given the DM's current beliefs as $(1 - x_2, x_2)$. Let us consider an infinitesimal learning on alternative 1 at $(1 - x_2, x_2)$, with belief update as dx. By conditioning on the sign of dx, we have

$$\begin{aligned}
P_D(x_2) &\equiv \Pr[\text{alternative } 1 | (1-x_2, x_2)] \\
&= \frac{1}{2}\Pr[\text{alternative } 1 | (1-x_2, x_2), dx \ge 0] \\
&+ \frac{1}{2}\Pr[\text{alternative } 1 | (1-x_2, x_2), dx < 0] \\
&= \frac{1}{2}\left[r_1(1 - x_2 + |dx|, x_2) + (1 - r_1(1 - x_2 + |dx|, x_2))\, P_D(x_2)\right] \\
&+ \frac{1}{2}\left[r_1(1 - x_2, x_2 + |dx|) + (1 - r_1(1 - x_2, x_2 + |dx|))\, P_D(x_2 + |dx|)\right] \\
&= P_D(x_2) + \frac{|dx|}{2}\left[P_D'(x_2) + \left(\frac{\partial r_1(x_1,x_2)}{\partial x_1} + \frac{\partial r_1(x_1,x_2)}{\partial x_2}\right)\Big|_{x_1 = 1 - x_2}(1 - P_D(x_2))\right] \\
&+ o(dx).
\end{aligned}$$

Similarly, we have the following ODE:

$$\frac{P_D'(x_2)}{1 - P_D(x_2)} = -\left(\frac{\partial r_1(x_1,x_2)}{\partial x_1} + \frac{\partial r_1(x_1,x_2)}{\partial x_2}\right)\Big|_{x_1 = 1 - x_2} = -\frac{2}{\overline{X}(x_2) + x_2 - 1}.$$

Combining the differential equation above with the initial condition $P_D(1/2) = 1/2$, we can solve $P_D(x)$ as

$$P_D(x) = 1 - \frac{1}{2} e^{\int_{\frac{1}{2}}^{x} \frac{2}{\overline{X}(\xi) + \xi - 1} d\xi}.$$

In this case, we have that,

$$P_1(x_1,x_2) = r_1(x_1,x_2) + \left[1 - r_1(x_1,x_2)\right] P_D(x_2), \quad 1/2 \le 1 - x_2 \le x_1 \le \overline{X}(x_2).$$

Similarly, we can calculate the adoption probability for regions L2, L1*, and L2*. □

Theorem A4: probability of being correct

Theorem A4. *Consider two symmetric alternatives. Under the optimal learning strategy, given his current posterior beliefs as (x_1, x_2), a decision maker's expected probability of being correct ex post is $Q(x_1, x_2)$, where if $x_0 \le 0$,*

$$
Q(x_1,x_2)=
\begin{cases}
\frac{\overline{X}(x_2)-x_1}{\overline{X}(x_2)-x_2}\left[\frac{1}{2}\left[\frac{1}{2}+\Phi^{-1}\left(-\frac{(\Delta\pi)^3}{4\sigma^2 c}\right)\right]e^{-\int_{\frac{1}{2}}^{x_2}\frac{2d\xi}{\overline{X}(\xi)-\xi}}\right. \\
\left.+\int_{\frac{1}{2}}^{x_2} e^{-\int_{\eta}^{x_2}\frac{2d\xi}{\overline{X}(\xi)-\xi}}\frac{2\overline{X}(\eta)}{\overline{X}(\eta)-\eta}d\eta\right]+\frac{x_1-x_2}{\overline{X}(x_2)-x_2}\overline{X}(x_2), & 1/2\le x_2\le x_1\le \overline{X}(x_2) \\
\frac{\overline{X}(x_2)-x_1}{x_2+\overline{X}(x_2)-1}\left[\frac{1}{2}\left[\frac{1}{2}+\Phi^{-1}\left(-\frac{(\Delta\pi)^3}{4\sigma^2 c}\right)\right]e^{\int_{\frac{1}{2}}^{x_2}\frac{2d\xi}{\overline{X}(\xi)+\xi-1}}\right. \\
\left.-\int_{\frac{1}{2}}^{x_2} e^{\int_{\eta}^{x_2}\frac{2d\xi}{\overline{X}(\xi)+\xi-1}}\frac{\overline{X}(\eta)-\eta+1}{\overline{X}(\eta)+\eta-1}d\eta\right]+\frac{x_1+x_2-1}{x_2+\overline{X}(x_2)-1}\overline{X}(x_2), & 1/2\le 1-x_2\le x_1\le \overline{X}(x_2) \\
\frac{\overline{X}(1-x_1)+x_2-1}{\overline{X}(1-x_1)+x_1-1}\left[\frac{1}{2}\left[\frac{1}{2}+\Phi^{-1}\left(-\frac{(\Delta\pi)^3}{4\sigma^2 c}\right)\right]e^{-\int_{\frac{1}{2}}^{1-x_1}\frac{2d\xi}{\overline{X}(\xi)-\xi}}\right. \\
\left.+\int_{\frac{1}{2}}^{1-x_1} e^{-\int_{\eta}^{1-x_1}\frac{2d\xi}{\overline{X}(\xi)-\xi}}\frac{2\overline{X}(\eta)}{\overline{X}(\eta)-\eta}d\eta\right]+\frac{x_2-x_1}{\overline{X}(1-x_1)+x_1-1}\overline{X}(1-x_1), & 1/2\le 1-x_1\le 1-x_2\le \overline{X}(1-x_1) \\
\frac{\overline{X}(1-x_1)+x_2-1}{\overline{X}(1-x_1)-x_1}\left[\frac{1}{2}\left[\frac{1}{2}+\Phi^{-1}\left(-\frac{(\Delta\pi)^3}{4\sigma^2 c}\right)\right]e^{\int_{\frac{1}{2}}^{1-x_1}\frac{2d\xi}{\overline{X}(\xi)+\xi-1}}\right. \\
\left.-\int_{\frac{1}{2}}^{1-x_1} e^{\int_{\eta}^{1-x_1}\frac{2d\xi}{\overline{X}(\xi)+\xi-1}}\frac{\overline{X}(\eta)-\eta+1}{\overline{X}(\eta)+\eta-1}d\eta\right]+\frac{1-x_1-x_2}{\overline{X}(1-x_1)-x_1}\overline{X}(1-x_1), & 1/2\le x_1\le 1-x_2\le \overline{X}(1-x_1) \\
\frac{\overline{X}(x_1)-x_2}{\overline{X}(x_1)-x_1}\left[\frac{1}{2}\left[\frac{1}{2}+\Phi^{-1}\left(-\frac{(\Delta\pi)^3}{4\sigma^2 c}\right)\right]e^{-\int_{\frac{1}{2}}^{x_1}\frac{2d\xi}{\overline{X}(\xi)-\xi}}\right. \\
\left.+\int_{\frac{1}{2}}^{x_1} e^{-\int_{\eta}^{x_1}\frac{2d\xi}{\overline{X}(\xi)-\xi}}\frac{2\overline{X}(\eta)}{\overline{X}(\eta)-\eta}d\eta\right]+\frac{x_2-x_1}{\overline{X}(x_1)-x_1}\overline{X}(x_1), & 1/2\le x_1\le x_2\le \overline{X}(x_1) \\
\frac{\overline{X}(x_1)-x_2}{x_1+\overline{X}(x_1)-1}\left[\frac{1}{2}\left[\frac{1}{2}+\Phi^{-1}\left(-\frac{(\Delta\pi)^3}{4\sigma^2 c}\right)\right]e^{\int_{\frac{1}{2}}^{x_1}\frac{2d\xi}{\overline{X}(\xi)+\xi-1}}\right. \\
\left.-\int_{\frac{1}{2}}^{x_1} e^{\int_{\eta}^{x_1}\frac{2d\xi}{\overline{X}(\xi)+\xi-1}}\frac{\overline{X}(\eta)-\eta+1}{\overline{X}(\eta)+\eta-1}d\eta\right]+\frac{x_2+x_1-1}{x_1+\overline{X}(x_1)-1}\overline{X}(x_1), & 1/2\le 1-x_1\le x_2\le \overline{X}(x_1) \\
\frac{\overline{X}(1-x_2)+x_1-1}{\overline{X}(1-x_2)+x_2-1}\left[\frac{1}{2}\left[\frac{1}{2}+\Phi^{-1}\left(-\frac{(\Delta\pi)^3}{4\sigma^2 c}\right)\right]e^{-\int_{\frac{1}{2}}^{1-x_2}\frac{2d\xi}{\overline{X}(\xi)-\xi}}\right. \\
\left.+\int_{\frac{1}{2}}^{1-x_2} e^{-\int_{\eta}^{1-x_2}\frac{2d\xi}{\overline{X}(\xi)-\xi}}\frac{2\overline{X}(\eta)}{\overline{X}(\eta)-\eta}d\eta\right]+\frac{x_1-x_2}{\overline{X}(1-x_2)+x_2-1}\overline{X}(1-x_2), & 1/2\le 1-x_2\le 1-x_1\le \overline{X}(1-x_2) \\
\frac{\overline{X}(1-x_2)+x_1-1}{\overline{X}(1-x_2)-x_2}\left[\frac{1}{2}\left[\frac{1}{2}+\Phi^{-1}\left(-\frac{(\Delta\pi)^3}{4\sigma^2 c}\right)\right]e^{\int_{\frac{1}{2}}^{1-x_2}\frac{2d\xi}{\overline{X}(\xi)+\xi-1}}\right. \\
\left.-\int_{\frac{1}{2}}^{1-x_2} e^{\int_{\eta}^{1-x_2}\frac{2d\xi}{\overline{X}(\xi)+\xi-1}}\frac{\overline{X}(\eta)-\eta+1}{\overline{X}(\eta)+\eta-1}d\eta\right]+\frac{1-x_2-x_1}{\overline{X}(1-x_2)-x_2}\overline{X}(1-x_2), & 1/2\le x_2\le 1-x_1\le \overline{X}(1-x_2) \\
\max\{x_1,x_2\}, & \textit{otherwise,}
\end{cases}
\tag{xxii}
$$

and if $x_0 \geq x_0^*$,

$$Q(x_1,x_2) = \begin{cases} \dfrac{\overline{x}(x_1-\underline{x})+(1-\underline{x})(\overline{x}-x_1)(1-x_2)}{\overline{x}-\underline{x}}, & x_2 \le \underline{x} < x_1 < \overline{x} \\ \dfrac{\overline{x}(x_2-\underline{x})+(1-\underline{x})(\overline{x}-x_2)(1-x_1)}{\overline{x}-\underline{x}}, & x_1 \le \underline{x} < x_2 < \overline{x} \\ \dfrac{(x_1-x_2)\overline{X}(x_2)}{\overline{X}(x_2)-x_2} + \dfrac{\overline{X}(x_2)-x_1}{\overline{X}(x_2)-x_2}\Big[(1-\underline{x})^2 e^{-\int_{\underline{x}}^{x_2}\frac{2d\xi}{\overline{X}(\xi)-\xi}} & \\ \quad + \displaystyle\int_{\underline{x}}^{x_2} e^{-\int_{\eta}^{x_2}\frac{2d\xi}{\overline{X}(\xi)-\xi}} \frac{2\overline{X}(\eta)}{\overline{X}(\eta)-\eta}d\eta\Big], & \underline{x} < x_2 < x_1 < \overline{X}(x_2) \\ \dfrac{(x_2-x_1)\overline{X}(x_1)}{\overline{X}(x_1)-x_1} + \dfrac{\overline{X}(x_1)-x_2}{\overline{X}(x_1)-x_1}\Big[(1-\underline{x})^2 e^{-\int_{\underline{x}}^{x_1}\frac{2d\xi}{\overline{X}(\xi)-\xi}} & \\ \quad + \displaystyle\int_{\underline{x}}^{x_1} e^{-\int_{\eta}^{x_1}\frac{2d\xi}{\overline{X}(\xi)-\xi}} \frac{2\overline{X}(\eta)}{\overline{X}(\eta)-\eta}d\eta\Big], & \underline{x} < x_1 < x_2 < \overline{X}(x_1) \\ x_1, & x_1 \ge \overline{X}(x_2) \\ x_2, & x_2 \ge \overline{X}(x_1) \\ (1-x_1)(1-x_2), & \textit{otherwise.} \end{cases} \tag{xxiii}$$

Proof. Let us consider the case with $x_0 \ge x_0^*$ first. By symmetry, we only need to consider the case that $x_1 \ge x_2$. There are four cases to consider.

In the first case, when $x_1 \ge \overline{X}(x_2)$, the DM adopts alternative 1 right away. In this case, he will be correct if and only if the alternative 1 turns out to be of high value. Therefore, $Q(x_1,x_2) = x_1$.

In the second case, when $x_2 \le x_1 \le \underline{x}$, the DM will adopt the outside option, which is ex-post correct if and only if both alternatives are of low value, i.e., $Q(x_1,x_2) = (1-x_1)(1-x_2)$.

In the third case, when $x_2 \le \underline{x} \le x_1 \le \overline{x}$, the DM will continue learning until his belief equals either $\underline{x}$ or $\overline{x}$. Therefore,

$$\begin{aligned} Q(x_1,x_2) &= \frac{x_1-\underline{x}}{\overline{x}-\underline{x}}\overline{x} + \left[1-\frac{x_1-\underline{x}}{\overline{x}-\underline{x}}\right](1-\underline{x})(1-x_2), \\ &= \frac{\overline{x}(x_1-\underline{x})+(1-\underline{x})(\overline{x}-x_1)(1-x_2)}{\overline{x}-\underline{x}}. \end{aligned}$$

In the last case, when $\underline{x} < x_2 < x_1 < \overline{X}(x_2)$, the DM keeps learning alternative 1 until his belief reaches either $\overline{X}(x_2)$ or x_2. Let us define $\widehat{Q}_0(x)$ as the probability of being correct, given the DM's beliefs as (x,x). Let us consider infinitesimal learning on alternative 1 at (x,x), with belief update as dx. By conditioning on the sign of dx, we have

$$\begin{aligned} \widehat{Q}_0(x) &\equiv \Pr[\text{correct}|(x,x)] \\ &= \frac{1}{2}\Pr[\text{correct}|(x,x), dx \ge 0] + \frac{1}{2}\Pr[\text{correct}|(x,x), dx < 0] \\ &= \frac{1}{2}q_1(x+|dx|,x)\overline{X}(x) + \frac{1}{2}\left[1-q_1(x+|dx|,x)\right]\widehat{Q}_0(x) \end{aligned}$$

$$+\frac{1}{2}q_1(x, x-|dx|)\overline{X}(x-|dx|)+\frac{1}{2}\left[1-q_1(x, x-|dx|)\right]\widehat{Q}_0(x-|dx|).$$

Similar to the proof of Theorem A3, by doing a Taylor expansion of $q_1(x+dx, x)$, $q_1(x, x-dx)$, $\widehat{Q}_0(x-dx)$, and $\overline{X}(x-dx)$, and simplifying, we have

$$\widehat{Q}_0'(x)=\left[\overline{X}(x)-\widehat{Q}_0(x)\right]\left(\frac{\partial q_1(x_1, x_2)}{\partial x_1}-\frac{\partial q_1(x_1, x_2)}{\partial x_2}\right)\Big|_{x_1=x_2=x}=\frac{2\left[\overline{X}(x)-\widehat{Q}_0(x)\right]}{\overline{X}(x)-x}. \tag{xxiv}$$

Combining the differential equation above with the initial condition $\widehat{Q}_0(\underline{x})=(1-\underline{x})^2$, we can solve $\widehat{Q}_0(x)$ as

$$\widehat{Q}_0(x)=(1-\underline{x})^2 e^{-\int_{\underline{x}}^{x}\frac{2d\xi}{\overline{X}(\xi)-\xi}}+\int_{\underline{x}}^{x} e^{-\int_{\eta}^{x}\frac{2d\xi}{\overline{X}(\xi)-\xi}}\frac{2\overline{X}(\eta)}{\overline{X}(\eta)-\eta}d\eta.$$

Given $\widehat{Q}_0(x)$, we have,

$$Q(x_1, x_2)=q_1(x_1, x_2)\overline{X}(x_2)+\left[1-q_1(x_1, x_2)\right]\widehat{Q}_0(x_2).$$

• Now, let us consider the other case with $x_0 \leq 0$. Due to symmetry, we only need to consider the case with $x_1 \geq x_2$. Let us first consider the case with $1/2 \leq x_2 \leq x_1 \leq \overline{X}(x_2)$. Similarly, we can define $\widehat{Q}_0(x)$ as above, and verify that it satisfies the ordinary differential equation (xxv). However, the initial condition is different with that in the case of $x_0 \geq x_0^*$. Particularly, we have that,

$$\widehat{Q}_0\left(\frac{1}{2}\right)=\frac{1}{2}\overline{X}\left(\frac{1}{2}\right)+\frac{1}{2}\frac{1}{2}=\frac{1}{2}\left[\frac{1}{2}+\Phi^{-1}\left(-\frac{(\Delta\pi)^3}{4\sigma^2 c}\right)\right].$$

We can solve $\widehat{Q}_0(x)$ as

$$\widehat{Q}_0(x)=\frac{1}{2}\left[\frac{1}{2}+\Phi^{-1}\left(-\frac{(\Delta\pi)^3}{4\sigma^2 c}\right)\right]e^{-\int_{\frac{1}{2}}^{x}\frac{2d\xi}{\overline{X}(\xi)-\xi}}+\int_{\frac{1}{2}}^{x} e^{-\int_{\eta}^{x}\frac{2d\xi}{\overline{X}(\xi)-\xi}}\frac{2\overline{X}(\eta)}{\overline{X}(\eta)-\eta}d\eta.$$

Given $\widehat{Q}_0(x)$, we have,

$$Q(x_1, x_2)=q_1(x_1, x_2)\overline{X}(x_2)+\left[1-q_1(x_1, x_2)\right]\widehat{Q}_0(x_2).$$

Now, let us consider the case with $1/2 \leq 1-x_2 \leq x_1 \leq \overline{X}(x_2)$. Let us define $\widehat{Q}_D(x_2)$ as the probability of being correct, given the DM's beliefs as $(1-x_2, x_2)$. Let us consider an infinitesimal learning on alternative 1 at $(1-x_2, x_2)$, with belief update as dx. By conditioning on the sign of dx, we have

$$\begin{aligned}
\widehat{Q}_D(x_2) &\equiv \Pr[\text{correct}|(1-x_2, x_2)] \\
&= \frac{1}{2}\Pr[\text{correct}|(1-x_2, x_2), dx \geq 0]+\frac{1}{2}\Pr[\text{correct}|(1-x_2, x_2), dx<0] \\
&= \frac{1}{2}\left[r_1(1-x_2+|dx|, x_2)\overline{X}(x_2)+(1-r_1(1-x_2+|dx|, x_2))\,\widehat{Q}_D(x_2)\right] \\
&+\frac{1}{2}\big[r_1(1-x_2, x_2+|dx|)(1-x_2-|dx|)
\end{aligned}$$

$$
\begin{aligned}
&+ (1 - r_1(1-x_2, x_2 + |dx|))\widehat{Q}_D(x_2 + |dx|)\Big] \\
&= \widehat{Q}_D(x_2) + \frac{|dx|}{2}\Big[\widehat{Q}'_D(x_2) + \frac{\partial r_1(x_1,x_2)}{\partial x_1}\Big|_{x_1=1-x_2}\overline{X}(x_2) \\
&+ \frac{\partial r_1(x_1,x_2)}{\partial x_2}\Big|_{x_1=1-x_2}(1-x_2) \\
&- \left(\frac{\partial r_1(x_1,x_2)}{\partial x_1} + \frac{\partial r_1(x_1,x_2)}{\partial x_2}\right)\Big|_{x_1=1-x_2}\widehat{Q}_D(x_2)\Big] + o(dx).
\end{aligned}
$$

We have the following ODE,

$$\frac{\widehat{Q}'_D(x)}{1+\overline{X}(x) - x - 2\widehat{Q}_D(x)} = -\frac{1}{\overline{X}(x) + x - 1}. \tag{xxv}$$

Combining the differential equation above with the initial condition that

$$\widehat{Q}_D\left(\frac{1}{2}\right) = \frac{1}{2}\left[\frac{1}{2} + \Phi^{-1}\left(-\frac{(\Delta\pi)^3}{4\sigma^2 c}\right)\right],$$

we can solve $\widehat{Q}_D(x)$ as

$$\widehat{Q}_D(x) = \frac{1}{2}\left[\frac{1}{2} + \Phi^{-1}\left(-\frac{(\Delta\pi)^3}{4\sigma^2 c}\right)\right]e^{\int_{\frac{1}{2}}^{x}\frac{2d\xi}{\overline{X}(\xi)+\xi-1}} - \int_{\frac{1}{2}}^{x} e^{\int_{\eta}^{x}\frac{2d\xi}{\overline{X}(\xi)+\xi-1}}\frac{\overline{X}(\eta)-\eta+1}{\overline{X}(\eta)+\eta-1}d\eta.$$

Given $\widehat{Q}_D(x)$, we have,

$$Q(x_1,x_2) = r_1(x_1,x_2)\overline{X}(x_2) + [1 - r_1(x_1,x_2)]\,\widehat{Q}_D(x_2). \quad \square$$

Lastly, it is straightforward to show that, if $x_0 \geq x_0^*$, we have that,

$$Q_0(x_1,x_2) = \begin{cases} x_1, & x_1 \geq \max\{x_2, x_0\} \\ x_2, & x_2 \geq \max\{x_1, x_0\} \\ (1-x_1)(1-x_2), & x_0 \geq \max\{x_1, x_2\} \end{cases},$$

and if $x_0 \leq 0$, we have that,

$$Q_0(x_1,x_2) = \max\{x_1, x_2\}.$$

Solution to the problem with two asymmetric alternatives The following theorem characterizes the optimal learning strategy in the case of two asymmetric alternatives under the assumption that alternative 1 is “better” than alternative 2, and (x_1^*, x_2^*) exists. Similar to the proof of Theorem A2, it is straightforward but tedious to show that the variational inequalities in (8) are satisfied. The details are thus omitted here.

Theorem A5. *Consider the Decision Maker’s optimal learning problem in (6) with two asymmetric alternatives, $\Delta\pi_1 > \Delta\pi_2$ (or $c_1 < c_2$). Suppose that (x_1^*, x_2^*) exits. When the value of the outside option is relatively high, the value function $V(x_1, x_2)$ is given by*

$$V(x_1,x_2)=\begin{cases}\frac{2\sigma_1^2c_1}{(\Delta\pi_1)^2}(1-2x_1)\left[\ln\left(\frac{1-x_1}{x_1}\right)-\ln\left(\frac{1-\overline{X}(x_2)}{\overline{X}(x_2)}\right)\right]\\ -\frac{2\sigma_1^2c_1}{(\Delta\pi_1)^2}\frac{1-2\overline{X}(x_2)}{1-\overline{X}(x_2)}\frac{\overline{X}(x_2)-x_1}{\overline{X}(x_2)}+\Delta\pi_1\cdot x_1+\underline{\pi}_1,\\ \quad\max\{\underline{x}_1,\underline{X}_1(x_2)\}\le x_1\le\overline{X}_1(x_2),x_2\le y(x_1)\\ \frac{2\sigma_2^2c_2}{(\Delta\pi_2)^2}(1-2x_2)\left[\ln\left(\frac{1-x_2}{x_2}\right)-\ln\left(\frac{1-\overline{X}(x_1)}{\overline{X}(x_1)}\right)\right]\\ -\frac{2\sigma_2^2c_2}{(\Delta\pi_2)^2}\frac{1-2\overline{X}(x_1)}{1-\overline{X}(x_1)}\frac{\overline{X}(x_1)-x_2}{\overline{X}(x_1)}+(\Delta\pi_2)x_2+\underline{\pi}_2,\\ \quad\underline{x}_2\le x_2\le\overline{X}_2(x_1),x_2>y(x_1)\\ \Delta\pi_1\cdot x_1+\underline{\pi}_1,\quad x_1>\overline{X}_1(x_2)\\ \Delta\pi_2\cdot x_2+\underline{\pi}_2,\quad x_2>\overline{X}_2(x_1)\ or\ x_1<\underline{X}_1(x_2)\\ \pi_0,\quad otherwise.\end{cases}$$

The adoption boundary of alternative 1, $\overline{X}_1(x_2)$ is given by

$$\overline{X}_1(x_2)=\begin{cases}\overline{\chi}_1(x_2) & x_2\ge x_2^*\\ \widetilde{\chi}_1\left(y^{-1}(x_2)\right), & x_2^*>x_2\ge\underline{x}_2\\ \overline{x}_1, & otherwise.\end{cases}$$

The adoption boundary of alternative 2, $\overline{X}_2(x_1)$ supported in $[0,x_1^*]$ is given by

$$\overline{X}_2(x_1)=\begin{cases}\overline{\chi}_2(x_1), & x_1^*\ge x_1\ge\underline{x}_1\\ \overline{x}_2, & x_1<\underline{x}_1.\end{cases}$$

The adoption boundary of alternative 2, $\underline{X}_1(x_2)$, and $\overline{\chi}_1(x_2)$, both supported in $[x_2^*,1]$, are given by the following equations:

$$F(\overline{\chi}_1(x_2))-F(\underline{X}_1(x_2))=-\frac{(\Delta\pi_1)^3}{2\sigma_1^2c_1}$$
$$\left(\frac{1}{\overline{\chi}_1(x_2)}+\frac{1}{1-\overline{\chi}_1(x_2)}\right)-\left(\frac{1}{\underline{X}_1(x_2)}+\frac{1}{1-\underline{X}_1(x_2)}\right)$$
$$=\frac{(\Delta\pi_1)^2\left[2\left(\Delta\pi_2\cdot x_2+\underline{\pi}_2-\underline{\pi}_1\right)-\Delta\pi_1\right]}{2\sigma_1^2c_1}.$$

$\overline{\chi}_2(x_1)$, $\widetilde{\chi}_1(x_1)$, and $y(x_1)$, all supported in $[\underline{x}_1,x_1^*]$, are determined by the following boundary value problem:

$$\frac{dF\left(\overline{\chi}_2(x_1)\right)}{dx_1}=-\frac{\frac{2\sigma_1^2c_1}{(\Delta\pi_1)^2}\left(F\left(\widetilde{\chi}_1(x_1)\right)-F(x_1)\right)+\Delta\pi_1}{\frac{2\sigma_2^2c_2}{(\Delta\pi_2)^2}\left(\overline{\chi}_2(x_1)-y(x_1)\right)},$$
$$\frac{dF\left(\widetilde{\chi}_1(x_1)\right)}{dx_1}=-y'(x_1)\frac{\frac{2\sigma_2^2c_2}{(\Delta\pi_2)^2}\left(F\left(\overline{\chi}_2(x_1)\right)-F(y(x_1))\right)+\Delta\pi_2}{\frac{2\sigma_1^2c_1}{(\Delta\pi_1)^2}\left(\widetilde{\chi}_1(x_1)-x_1\right)},$$
$$\left(\widetilde{\chi}_1(x_1)-x_1\right)\left[\frac{2\sigma_1^2c_1}{(\Delta\pi_1)^2}\left(F\left(\widetilde{\chi}_1(x_1)\right)-F(x_1)\right)+\Delta\pi_1\right]$$

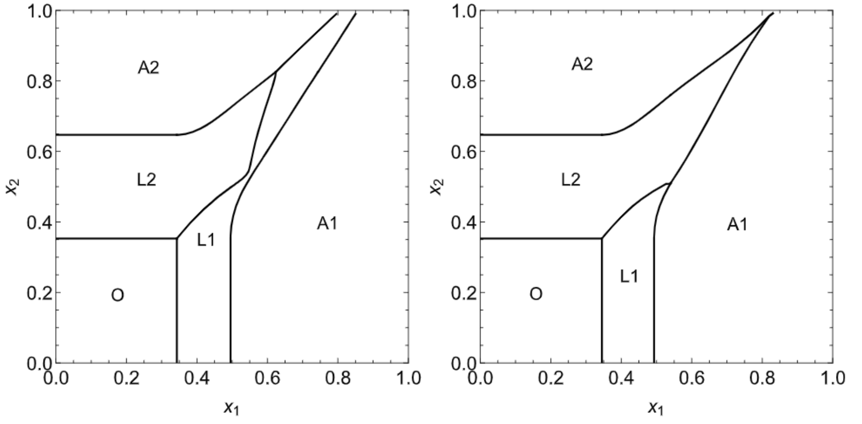

Fig. A1. Optimal learning strategy with asymmetric alternatives. We have $c_1\sigma_1^2/\Delta\pi = 0.33$ for the left panel and $c_1\sigma_1^2/\Delta\pi = 0.34$ for the right panel. For both panels: $\Delta\pi_1 = 1.2\Delta\pi_2 = 1.2\Delta\pi$, $(\pi_0 - \underline{\pi}_1)/\Delta\pi = (\pi_0 - \underline{\pi}_2)//\Delta\pi = 0.5$, and $c_2\sigma_2^2/\Delta\pi = c_2\sigma_2^2/\Delta\pi = 1$.

$$= \left(\overline{\chi}_2(x_1) - y(x_1)\right)\left[\frac{2\sigma_2^2 c_2}{(\Delta\pi_2)^2}\left(F\left(\overline{\chi}_2(x_1)\right) - F(y(x_1))\right) + \Delta\pi_2\right],$$

$$\overline{\chi}_2(\underline{x}_1) = \overline{x}_2,$$
$$\widetilde{\chi}_1(\underline{x}_1) = \tilde{x}_1,$$
$$y(\underline{x}_1) = \underline{x}_2.$$

Case of asymmetric alternatives with $\Delta\pi_1 > \Delta\pi_2$ and $c_1 > c_2$ Now, let us have a look at the more complicated case that alternative 1 may have a higher information-to-noise ratio, while at the same time a higher learning cost. Fig. A1 shows the optimal learning strategy under some parameter settings. As we can see, the solution structure remains mostly unchanged. We still have a simple cutoff policy to construct consideration set, but the boundary separating "Learn 1" and "Learn 2" can become more complicated now, and can be sensitive to the parameter setting.

Solution to the problem with time discounting

Theorem A6. *There exists a unique solution $V(x_1, x_2)$ to (37):*

$$V(x_1,x_2) = \begin{cases} \frac{\alpha+1}{2\alpha}\overline{\pi} Z(x_2)^{\frac{\alpha-1}{2}} x_1^{\frac{\alpha+1}{2}}(1-x_1)^{-\frac{\alpha-1}{2}} & \\ +\frac{\alpha-1}{2\alpha}\overline{\pi} Z(x_2)^{-\frac{\alpha+1}{2}} x_1^{-\frac{\alpha-1}{2}}(1-x_1)^{\frac{\alpha+1}{2}}, & x_2 \le x_1 \le \widetilde{X}(x_2), x_1 \ge \underset{\sim}{x} \\ \frac{\alpha+1}{2\alpha}\overline{\pi} Z(x_1)^{\frac{\alpha-1}{2}} x_2^{\frac{\alpha+1}{2}}(1-x_2)^{-\frac{\alpha-1}{2}} & \\ +\frac{\alpha-1}{2\alpha}\overline{\pi} Z(x_1)^{-\frac{\alpha+1}{2}} x_2^{-\frac{\alpha-1}{2}}(1-x_2)^{\frac{\alpha+1}{2}}, & x_1 \le x_2 \le \widetilde{X}(x_1), x_2 \ge \underset{\sim}{x} \\ \Delta\pi \cdot x_1, & x_1 > \widetilde{X}(x_2) \\ \Delta\pi \cdot x_2, & x_2 > \widetilde{X}(x_1) \\ \pi_0, & \text{otherwise,} \end{cases}$$

where $\alpha \equiv \sqrt{1+\frac{8r\sigma^2}{(\Delta\pi)^2}} > 1$, *and* $\widetilde{X}(x)$ *is given by*

$$\widetilde{X}(x) = \begin{cases} \frac{1}{Z(x)+1}, & x \geq \underset{\sim}{x} \\ \tilde{x}, & \text{otherwise}. \end{cases}$$

$Z(x)$ *is determined by the following boundary value problem:*

$$\begin{aligned} \left(\alpha^2-1\right)x(1-x)\frac{Z'(x)}{Z(x)} &= \alpha^2 + \alpha(1-2x)\frac{Z(x)^\alpha + \left(\frac{1-x}{x}\right)^\alpha}{Z(x)^\alpha - \left(\frac{1-x}{x}\right)^\alpha} + (1-2x) \\ &\quad + \alpha\frac{Z(x)^\alpha + \left(\frac{1-x}{x}\right)^\alpha}{Z(x)^\alpha - \left(\frac{1-x}{x}\right)^\alpha}, \\ Z(\underset{\sim}{x}) &= \frac{1-\tilde{x}}{\tilde{x}}. \end{aligned} \tag{xxvi}$$

$\tilde{x}$ *and* $\underset{\sim}{x}$ *are determined by the following equations:*

$$\left(\frac{\tilde{x}}{\underset{\sim}{x}}\right)^{-\frac{\alpha-1}{2}} \left(\frac{1-\tilde{x}}{1-\underset{\sim}{x}}\right)^{\frac{\alpha+1}{2}} = \frac{\Delta\pi(\alpha-1)\tilde{x}}{\pi_0(\alpha+1-2\underset{\sim}{x})}, \tag{xxvii}$$

$$\left(\frac{\tilde{x}}{\underset{\sim}{x}}\right)^{\frac{\alpha+1}{2}} \left(\frac{1-\tilde{x}}{1-\underset{\sim}{x}}\right)^{-\frac{\alpha-1}{2}} = \frac{\Delta\pi(\alpha+1)\tilde{x}}{\pi_0(\alpha-1+2\underset{\sim}{x})}. \tag{xxviii}$$

To prove Theorem A6, we can construct the solution following the similar steps in the proof of Theorem A2. We only need to show that for $\underset{\sim}{x} \leq x_2 \leq x_1 \leq \widetilde{X}(x_2)$, it is always optimal to learn alternative 1. Equivalently, we need to prove the following inequality given $\underset{\sim}{x} \leq x_2 \leq x_1 \leq \widetilde{X}(x_2)$,

$$\frac{2\sigma^2 r}{\Delta\pi^2}V(x_1,x_2) = x_1^2(1-x_1)^2 V_{x_1x_1}(x_1,x_2) \geq x_2^2(1-x_2)^2 V_{x_2x_2}(x_1,x_2),$$

where the first equation above is due to (37).

Given $\underset{\sim}{x} \leq x_2 \leq x_1 \leq \widetilde{X}(x_2)$, we have that,

$$\begin{aligned} &V(x_1,x_2) \\ &= \left[\frac{\alpha+1}{2\alpha}x_1^{\frac{\alpha+1}{2}}(1-x_1)^{-\frac{\alpha-1}{2}}Z(x_2)^{\frac{\alpha-1}{2}} + \frac{\alpha-1}{2\alpha}x_1^{-\frac{\alpha-1}{2}}(1-x_1)^{\frac{\alpha+1}{2}}Z(x_2)^{-\frac{\alpha+1}{2}}\right]\overline{\pi}. \end{aligned}$$

By taking the partial derivative with respect to x_2 on both sides of the equation above, we have that,

$$\begin{aligned} V_{x_2}(x_1,x_2) = \frac{\alpha^2-1}{4\alpha}&\left[\frac{\alpha+1}{2\alpha}x_1^{\frac{\alpha+1}{2}}(1-x_1)^{-\frac{\alpha-1}{2}}Z(x_2)^{\frac{\alpha-1}{2}} - \frac{\alpha-1}{2\alpha}x_1^{-\frac{\alpha-1}{2}}(1-x_1)^{\frac{\alpha+1}{2}}\right. \\ &\left.\times Z(x_2)^{-\frac{\alpha+1}{2}}\right]\overline{\pi}\frac{Z'(x_2)}{Z(x_2)} \end{aligned}$$

$$= \frac{\overline{\pi}}{4\alpha} \frac{1}{x_2(1-x_2)} \frac{\left(\frac{1-x_2}{x_2}\right)^{\alpha} - \frac{\alpha+1}{\alpha-1}\frac{\alpha+(1-2x_2)}{\alpha-(1-2x_2)} Z(x_2)^{\alpha}}{\left(\frac{1-x_2}{x_2}\right)^{\alpha} - Z(x_2)^{\alpha}}$$
$$\times \left[\frac{\alpha+1}{2\alpha} x_1^{\frac{\alpha+1}{2}} (1-x_1)^{-\frac{\alpha-1}{2}} Z(x_2)^{\frac{\alpha-1}{2}} - \frac{\alpha-1}{2\alpha} x_1^{-\frac{\alpha-1}{2}} (1-x_1)^{\frac{\alpha+1}{2}} Z(x_2)^{-\frac{\alpha+1}{2}} \right],$$

where the second equality is by using (xxvi). By taking the partial derivative with respect to x_2 and using (xxvi) again, we have that,

$$\frac{2\sigma^2 r}{\Delta\pi^2} V(x_1, x_2) - V_{x_2 x_2}(x_1, x_2)$$
$$= \frac{\overline{\pi}}{\alpha^2 - 1} (1-x_2) x_1^{\frac{\alpha+1}{2}} (1-x_1)^{-\frac{\alpha-1}{2}} \left[\left(\frac{1-x_2}{x_2}\right)^{\alpha} - \left(\frac{1-x_1}{x_1}\right)^{\alpha} \right]$$
$$\times \frac{Z(x_2)^{\frac{\alpha-1}{2}} \left[(\alpha-1) \left(\frac{1-x_2}{x_2}\right)^{\alpha} + (\alpha+1) Z(x_2)^{\alpha} \right]}{\left[\left(\frac{1-x_2}{x_2}\right)^{\alpha} - Z(x_2)^{\alpha} \right]^3}$$
$$\times \left[(\alpha-1)(\alpha+x_2) \left(\frac{1-x_2}{x_2}\right)^{\alpha} - (\alpha+1)(\alpha-x_2) Z(x_2)^{\alpha} \right].$$

Notice that $x_1 \geq x_2$, so we have that $\left(\frac{1-x_2}{x_2}\right)^{\alpha} - \left(\frac{1-x_1}{x_1}\right)^{\alpha} \geq 0$. Also, $\widetilde{X}(x_2) = (1 + Z(x_2))^{-1} > x_2$, so we have that $\frac{1-x_2}{x_2} > Z(x_2)$, and thus $\left(\frac{1-x_2}{x_2}\right)^{\alpha} - Z(x_2)^{\alpha} > 0$. This implies that $\frac{2\sigma^2 r}{\Delta\pi^2} V(x_1, x_2) - V_{x_2 x_2}(x_1, x_2) \geq 0$ if and only if $\left[(\alpha-1)(\alpha+x_2) \left(\frac{1-x_2}{x_2}\right)^{\alpha} - (\alpha+1)(\alpha-x_2) \times Z(x_2)^{\alpha} \right] \geq 0$, or equivalently,

$$Z(x_2)^{\alpha} \leq \left(\frac{1-x_2}{x_2}\right)^{\alpha} \frac{\alpha-1}{\alpha+1} \frac{\alpha+x_2}{\alpha-x_2}. \tag{xxix}$$

Let us prove (xxix). In fact, similar to the proof of the monotonicity of $\overline{X}(x)$ in Lemma 2, we can prove that $\widetilde{X}(x)$ increases with x. This implies that $Z(x)$ decreases with x. By (xxvi), $Z'(x_2) \leq 0$ implies that,

$$\left(\frac{1-x_2}{x_2}\right)^{\alpha} - \frac{\alpha+1}{\alpha-1} \frac{\alpha+(1-2x_2)}{\alpha-(1-2x_2)} Z(x_2)^{\alpha} \leq 0.$$

This implies that,

$$Z(x_2)^{\alpha} \leq \left(\frac{1-x_2}{x_2}\right)^{\alpha} \frac{\alpha-1}{\alpha+1} \frac{\alpha-(1-2x_2)}{\alpha+(1-2x_2)} \leq \left(\frac{1-x_2}{x_2}\right)^{\alpha} \frac{\alpha-1}{\alpha+1} \frac{\alpha+x_2}{\alpha-x_2}.$$

To summarize, we have proved that for $\underset{\sim}{x} \leq x_2 \leq x_1 \leq \widetilde{X}(x_2)$, it is always optimal to learn alternative 1. □

References

Adam, Klaus, 2001. Learning while searching for the best alternative. J. Econ. Theory 101 (1), 252–280.
Barles, Guy, 1997. Solutions de viscosité et équations elliptiques du deuxième ordre. Manuscript, Université de Tours.

Bergemann, Dirk, Välimäki, Juuso, 1996. Learning and strategic pricing. Econometrica 64795, 773.
Bolton, Patrick, Harris, Christopher, 1999. Strategic experimentation. Econometrica 67 (2), 349–374.
Branco, Fernando, Sun, Monic, Villas-Boas, J. Miguel, 2012. Optimal search for product information. Manag. Sci. 58 (11), 2037–2056.
Caffarelli, Luis A., 1977. The regularity of free boundaries in higher dimensions. Acta Math. 139, 155–184.
Che, Yeon-Koo, Mierendorff, Konrad, 2016. Optimal Sequential Decision with Limited Attention. Working paper. Columbia University.
Chernoff, Herman, 1968. Optimal stochastic control. Sankhya, Ser. A, 221–252.
Crandall, Michael G., Ishii, Hitoshi, Lions, Pierre-Louis, 1992. User's guide to viscosity solutions of second order partial differential equations. Bull., New Ser., Am. Math. Soc. 27 (1), 1–67.
Daley, Brendan, Green, Brett S., 2012. Waiting for news in the market for lemons. Econometrica 80 (4), 1433–1504.
Décamps, Jean-Paul, Mariotti, Thomas, Villeneuve, Stéphane, 2006. Irreversible investment in alternative projects. Econ. Theory 28 (2), 425–448.
Diamond, Peter A., 1971. A model of price adjustment. J. Econ. Theory 3 (2), 156–168.
Dixit, A., 1993. Art of Smooth Pasting, vol. 55. Harwood Academic, Chur, Switzerland.
Dixit, Avinash K., Pindyck, Robert S., 1994. Investment Under Uncertainty. Princeton University Press.
Doval, Laura, 2014. Whether or Not to Open Pandora's Box. Working paper. Northwestern University.
Dumas, Bernard, 1991. Super contact and related optimality conditions. J. Econ. Dyn. Control 15 (4), 675–685.
Dvoretzky, A., Kiefer, F., Wolfowitz, J., 1953. Sequential decision problems for processes with continuous time parameter. Ann. Math. Stat. 24 (2), 254–264.
Felli, Leonardo, Harris, Christopher, 1996. Learning, wage dynamics, and firm-specific human capital. J. Polit. Econ. 104 (4), 838–868.
Fernholz, E. Roberts, Ichiba, Tomoyuki, Ioannis, Karatzas, Prokaj, Vilmos, 2013. Planar diffusions with rank-based characteristics and perturbed Tanaka equations. Probab. Theory Relat. Fields 156 (1), 343–374.
Fleming, Wendell H., Soner, H.M., 2010. Controlled Markov Processes and Viscosity Solutions. Springer.
Forand, Jean G., 2015. Keeping your options open. J. Econ. Dyn. Control 53, 47–68.
Fudenberg, Drew, Strzalecki, Tomasz, 2015. Dynamic logit with choice aversion. Econometrica 83 (2), 651–691.
Fudenberg, Drew, Strack, Philipp, Strzalecki, Tomasz, 2015. Stochastic Choice and Optimal Sequential Sampling. Working paper. Harvard University.
Gittins, John C., 1979. Bandit processes and dynamic allocation indices. J. R. Stat. Soc. B 41, 148–164.
Gittins, John C., Glazebrook, Kevin D., Weber, Richard, 2011. Multi-Armed Bandit Allocation Indices, 2nd edition. Wiley Online Library.
Glazebrook, Kevin D., 1979. Stoppable families of alternative bandit processes. J. Appl. Probab., 843–854.
Hébert, Benjamin, Woodford, Michael, 2017. Rational Inattention with Sequential Information Sampling. Working paper. Stanford University and Columbia University.
Karatzas, Ioannis, 1984. Gittins indices in the dynamic allocation problem for diffusion processes. Ann. Probab., 173–192.
Karatzas, Ioannis, Shreve, Steven E., 1991. Brownian Motion and Stochastic Calculus. Springer, New York.
Ke, T. Tony, Shen, Zuo-Jun Max, Villas-Boas, J. Miguel, 2016. Search for information on multiple products. Manag. Sci. 62 (12), 3576–3603.
Kuksov, Dmitri, Villas-Boas, J. Miguel, 2010. When more alternatives lead to less choice. Mark. Sci. 29 (3), 507–524.
Liptser, Robert S., Shiryaev, Albert N., 2001. Statistics of Random Processes: General Theory, vol. 1. Springer.
Mandelbaum, Avi, Shepp, Larry A., Vanderbei, Robert J., 1990. Optimal switching between a pair of Brownian motions. Ann. Probab. 18 (3), 1010–1033.
McCall, John Joseph, 1970. Economics of information and job search. Q. J. Econ., 113–126.
Mikhalevich, A.S., 1958. A Bayes test of two hypotheses concerning the mean of a normal process. Visn. Kyiv. Univ. 1, 101–104.
Moscarini, Giuseppe, Smith, Lones, 2001. The optimal level of experimentation. Econometrica 69 (6), 1629–1644.
Øksendal, Bernt, Reikvam, Kristin, 1998. Viscosity solutions of optimal stopping problems. Stoch. Stoch. Rep. 62, 285–301.
Peskir, Goran, Shiryaev, Albert, 2006. Optimal Stopping and Free-Boundary Problems. Birkhäuser Verlag, Basel, Switzerland.
Quah, John K.-H., Strulovici, Bruno, 2013. Discounting, values, and decisions. J. Polit. Econ. 121 (5), 896–939.
Roberts, Kevin, Weitzman, Martin L., 1981. Funding criteria for research, development, and exploration projects. Econometrica 49 (5), 1261–1288.
Rothschild, Michael, 1974a. A two-armed bandit theory of market pricing. J. Econ. Theory 9 (2), 185–202.

Rothschild, Michael, 1974b. Searching for the lowest price when the distribution of prices is unknown. J. Polit. Econ. 82 (4), 689–711.
Scheibehenne, Benjamin, Greifeneder, Rainer, Todd, Peter M., 2010. Can there ever be too many options? A meta-analytic review of choice overload. J. Consum. Res. 37 (3), 409–425.
Shiryaev, A.N., 1967. Two problems of sequential analysis. Cybernetics 3, 63–69.
Strulovici, Bruno, Szydlowski, Martin, 2015. On the smoothness of value functions and the existence of optimal strategies in diffusion models. J. Econ. Theory 159, 1016–1055.
Touzi, Nizar, 2010. Optimal Stochastic Control, Stochastic Target Problems, and Backward SDE. Springer.
Wald, Abraham, 1945. Sequential tests of statistical hypotheses. Ann. Math. Stat. 16 (2), 117–186.
Weeds, Helen, 2002. Strategic delay in a real options model of R&D competition. Rev. Econ. Stud. 69 (3), 729–747.
Weitzman, Martin L., 1979. Optimal search for the best alternative. Econometrica 47 (3), 641–654.
Whittle, P., 1980. Multi-armed bandits and the Gittins index. J. R. Stat. Soc. 42 (2), 143–149.