

# Forward-looking Experimentation of Correlated Alternatives\*

Yu Fu Wong<sup>†</sup>

January 22, 2025

This paper studies how a forward-looking decision maker experiments on unknown alternatives of correlated utilities. The utilities are modeled by a Brownian motion such that similar alternatives yield similar utilities. Experimentation trades off between the continuation value of exploration and the opportunity cost of exploitation. The optimal strategy is to continuously explore unknown alternatives, and then exploit the best known alternative when the one being explored is found to be sufficiently worse than the best one. The decision maker explores unknown alternatives more quickly as they prove to be worse than the best known one. Applied to firm experimentation, my model predicts a conditional version of Gibrat's law and a linear relation between firm size and profitability.

Keywords: Experimentation, Brownian motion, Firm dynamics, Gibrat's law

JEL classifications: D83, D30, D21

---

\*First draft: April 28, 2017. This paper is based on my master's thesis, "Spatial Experimentation," submitted to the Toulouse School of Economics. I am indebted to Thomas Mariotti for his supervision. I am grateful to Arjada Bardhi, Jess Benhabib, Alberto Bisin, Adrien Blanchet, Mike Borns, Steven Callander, Yeon-Koo Che, Ricardo Fernholz, Robert Fernholz, Xavier Gabaix, Matthieu Gomez, Duarte Gonçalves, Johannes Hörner, Boyan Jovanovic, Ioannis Karatzas, Navin Kartik, Elliot Lipnowski, Qingmin Liu, Laurent Miclo, Marcel Nutz, Bernard Salanié, and Bruno Strulovici for insightful comments. I thank Martsella Davitaya for research assistance.

<sup>†</sup>University of Pittsburgh. Contact: [yufuwong0@gmail.com](mailto:yufuwong0@gmail.com).

# 1 Introduction

This paper studies the experimentation problem of a long-lived decision maker who chooses alternatives of unknown correlated utility. The decision maker learns the utility of her chosen alternative, which informs her future choices. To capture the correlation, I model utility as a Brownian motion over a continuum of alternatives such that similar alternatives yield similar utilities. Prior work on correlated utility, including [Jovanovic and Rob \(1990\)](#) and [Callander \(2011\)](#), studies *short-lived* agents who can exploit a better alternative based on past explorations. My innovation is to study a *long-lived* decision maker who can also continue to explore.

The optimal experimentation strategy describes how quickly the decision maker continuously explores unknown alternatives and when she stops to exploit the best known one (Theorem 1). The speed of exploration depends on the drawdown, i.e., the utility difference between the best known alternative and the one being explored, and the decision maker exploits the best alternative when the drawdown exceeds a threshold.

The innovation of my analysis is the application of time change to the domain of alternatives. A time change indexes a stochastic process by an increasing sequence of stopping times. The experimentation literature on finitely many risky arms, such as [Karatzas \(1984\)](#), [Moscarini and Smith \(2001\)](#), and [Keller, Rady, and Cripps \(2005\)](#), performs time change in the domain of information that accumulates at a controlled rate. The most up-to-date information is a sufficient statistic because it encompasses past information. In my model, the decision maker controls the speed of exploring new alternatives. The alternative being explored is not a sufficient statistic because previously explored alternatives could be better. I show that the optimal strategy also takes the best known alternative into account.

The tradeoff between the continuation value of exploration and the flow opportunity cost of exploitation implies faster exploration for more negative drawdowns (Proposition 1). On the one hand, faster exploration acquires information more quickly and realizes the continuation value sooner. When the drawdown becomes more negative, the impatient decision maker finds it less pressing to realize the decreased continuation value. On the other hand, faster exploration shortens the du-

ration of exploration as the flow opportunity cost of exploitation accrues. When the drawdown is more negative, the increased flow opportunity cost prompts the decision maker to explore in shorter time. I show that the opportunity cost incentive always dominates the continuation value incentive. This is because the option to exploit partially insures the decision maker from the drawdown, and so the continuation value is less sensitive to the drawdown than the flow opportunity cost.

In the time series, i.e., how experimentation evolves over time, the speed's monotonicity with respect to drawdown implies that the speed of exploration increases over time under conditions on model parameters (Corollary 1).

I study the comparative statics of the optimal strategy (Proposition 2). Intuitively, the value of exploration increases when the decision maker is more patient, the cost of exploration is lower, or the utility has higher on average. To realize the increased value, the decision maker explores alternatives more quickly for any given drawdown and tolerates more negative drawdowns, i.e., the drawdown threshold increases. The comparative statics also hold when the utility is more volatile because of the option value of exploitation.

My model applies to a firm that experiments on its firm size to maximize profits. I interpret the log number of employees as a proxy for firm size and log profit as a proxy for profitability. Under the Brownian specification, firm-size elasticity of profitability is modeled to be independent and identically distributed over all firm sizes.

In the cross section, i.e., how experimentation differs across Brownian realizations, I derive two results that correspond to predictions on firm dynamics. The first result is that, as the decision maker explores more alternatives, the speed of exploration converges asymptotically in distribution (Proposition 3). In terms of firm dynamics, this translates to a conditional version of Gibrat's law: among large firms, the percentage change in firm size *conditional on growth* follows essentially the same distribution that is invariant to absolute firm size. The percentage change depends on absolute size only through the extensive margin, i.e., firms that stop growing, but not the intensive margin, i.e., firms that keep growing. By comparison, the original, unconditional version of Gibrat's law states that the distribution is invariant at both the extensive margin and intensive margin.

The second result is that the exploited alternative and its utility are asymptotically

linear (Proposition 4). In context of firm experimentation, this result means that long-run profitability is proportional to long-run firm size among large firms. The result highlights an endogenous selection effect: firms *choose* to operate at large sizes because they experience increasing profits along the growth path. Therefore, the regression of firm size on profits may overestimate average profit elasticity.

## Related literature

My paper incorporates a long-lived decision maker into the experimentation problem where the utilities of unknown alternatives are modeled by a Brownian motion. [Jovanovic and Rob \(1990\)](#) analyze such a problem with overlapping generations of short-lived agents, each of which can learn the Brownian realization of one alternative before choosing another for utility. [Callander \(2011\)](#) examines myopic agents who reveal the utility of an alternative to their choice. [Garfagnini and Strulovici \(2016\)](#) extend [Callander \(2011\)](#) to agents who live for two periods.<sup>1</sup> In these papers, the agents acquire information at most once before exploitation and therefore do not incorporate the information value of continued exploration.<sup>2</sup> By contrast, the long-lived decision maker in my model can continue to explore. As a result, past exploration becomes instrumental to future exploration, and the information value feeds back to the decision maker's experimentation strategy.

[Urgun and Yariv \(2025\)](#) considers a related search model in which an agent searches continuously among alternatives of Brownian utility. The objective is to maximize the expected utility of the best alternative subject to a search cost. As in my model, their optimal strategy is to search continuously until the drawdown reaches a threshold. The optimal speed of search is constant in order to minimize the search cost per alternative. By contrast, the objective in my model is to maximize the expected discounted flow utility from exploration and exploitation subject to a

---

<sup>1</sup>[Garfagnini and Strulovici \(2016\)](#) show that the agents may alternate between exploration and exploitation. This is because the current agent's remaining lifespan is non-stationary in their overlapping generations model.

<sup>2</sup>In [Garfagnini and Strulovici \(2016\)](#), the agents can in principle continue to explore in the second period. However, it is weakly suboptimal to do so because, under their driftless Brownian specification, the expected utility is maximized at some known alternative. Therefore, the experiment in the first period depends on the marginal distribution of utilities only, but not the joint distribution. Consequently, their optimal strategy can be derived from the Gittins index as in [Weitzman \(1979\)](#), as if the alternatives were independent.

learning cost. The interaction between the continuation value of exploration and the flow opportunity cost of exploration generates the dynamics of exploration.

The time change enables me to derive the time-series and cross-sectional results from the literature on Brownian motions that are stopped at a drawdown threshold. [Taylor \(1975\)](#) and [Lehoczky \(1977\)](#) characterize the joint distribution of the running maximum and the stopping time, for a Brownian motion whose drift and volatility may depend on its value. However, their framework does not accommodate my optimal strategy because both the drift and volatility in my case depend on the drawdown. To leverage their results, I apply a time change to the Brownian utility, transforming it from the time domain to the domain of alternatives. Because the Brownian utility over alternatives has constant drift and volatility, their results can be applied to this process, as the threshold strategy remains invariant under the time change.

The remainder of the paper is organized as follows. Section 2 introduces the experimentation problem. Section 3 derives the optimal experimentation strategy. Section 4 studies the properties of optimal experimentation. Section 5 discusses key modeling assumptions.

## 2 Experimentation problem

I study the experimentation problem of a forward-looking decision maker who continuously explores new alternatives of unknown and correlated utility subject to a learning cost.

I model utility as a Brownian motion with drift over the domain of alternatives to capture their correlation. Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a probability space that can support a standard Brownian motion. The utility process follows

$$U(\omega, x) := \mu x + \sigma B(\omega, x)$$

for realization  $\omega \in \Omega$  and alternative  $x \in [0, \infty)$ , where  $\mu \in (-\infty, \infty)$  is the drift,  $\sigma > 0$  the volatility, and  $B$  a standard Brownian motion.

An experimentation strategy specifies the choice of alternatives based on past ex-

periments. In realization  $\omega \in \Omega$ , the decision maker chooses alternative  $x_t(\omega)$  at time  $t \in [0, \infty)$ . As in [Callander \(2011\)](#) and [Garfagnini and Strulovici \(2016\)](#), the decision maker learns the realized utility  $U(\omega, x)$  upon her choice via learning-by-doing. Therefore, her information at time  $t$  consists of the history of calendar time, the chosen alternatives, and the corresponding utilities, i.e.,  $\mathcal{F}_t := \Sigma\left(\{(\eta, x_\eta(\omega), U(\omega, x_\eta))\}_{\eta \leq t}\right)$ , where  $\Sigma(\cdot)$  denotes the generated sigma-algebra. Let  $\mathbb{F}$  denote the augmented filtration and  $r > 0$  the decision maker's discount rate. I define the strategy space as follows.

**Definition 1** *An experimentation strategy,  $\{x_t(\omega) : \omega \in \Omega, t \in [0, \infty)\}$  or simply  $\{x_t\}$ , is a stochastic process that satisfies:*

- *Initial condition: For each  $\omega \in \Omega$ ,  $x_0(\omega) = 0$ .*
- *Continuous exploration condition:  
For each  $\omega \in \Omega$ , the frontier  $X_t := \sup_{\eta \in [0, t]} x_\eta(\omega)$  is absolutely continuous.*
- *Measurability condition:  
 $\{x_t\}$  is predictable with respect to the augmented filtration  $\mathbb{F}$ .*
- *Growth condition:  
There exists  $k_1 \in (0, r)$  and  $k_2 > 0$  such that  $x_t(\omega) \leq k_2 \exp(k_1 t)$  for all  $\omega \in \Omega$  and  $t \in [0, \infty)$ .*

An important and immediate consequence of the initial condition and the continuous exploration condition is that the (closure of the) set of explored alternatives at time  $t$  is the interval  $[0, X_t]$ .

When the decision maker explores new alternatives, she incurs a flow learning cost  $c(s_t)$ , where  $s_t := \frac{d}{dt} X_t$  is the speed of exploration. The cost function  $c : [0, \infty) \rightarrow [0, \infty)$  is twice continuously differentiable and strictly convex. It is assumed to satisfy  $c(0) = 0$ ,  $c'(\infty) > \mu^+/r$ , and the Inada condition:  $\lim_{s \rightarrow \infty} c'(s)s - c(s) = \infty$ . Note that the cost is increasing in the speed of exploration, and is zero when the speed is zero, e.g., when the decision maker exploits a previously chosen alternative. Thus, the flow learning cost can be interpreted as the utility loss incurred by the decision maker when she first learns to extract utility from a given alternative.

The forward-looking decision maker chooses an experimentation strategy to maximize the expected discounted Brownian utility from the chosen alternatives net of the learning cost of exploration. Formally:

$$\max_{\{x_t\}} \mathbb{E} \left[ \int_0^\infty r e^{-rt} (U(x_t) - c(s_t)) dt \right]. \quad (1)$$

I denote the value attained by strategy  $\{x_t\}$  as  $V\{x_t\}$ , and call  $\{x_t\}$  an optimal strategy if it solves the maximization problem (1).

**Remark 1 (Key assumptions)** I comment on the key assumptions in the model. First, the utility process with stationary independent increments allows the recursive formulation of the optimal strategy. In particular, the Brownian specification with constant drift and volatility lends itself to the well-developed mathematics literature. Second, on the strategy space, the continuous exploration condition permits jumps to known alternatives and forbids jumps to unknown alternatives to the right of the frontier. I argue in Section 5.1 that this condition is without loss of optimality for convex learning costs. Third, the growth condition rules out strategies with explosive explorations that attain infinite utility at infinite cost.

### 3 Optimal experimentation strategy

I derive the optimal experimentation strategy in three steps. First, I show that the objective function is bounded. Second, I construct a candidate strategy based on several conjectures. Third, I prove the optimality of this strategy through verification. The primary departure from the standard argument is that the strategy  $\{x_t\}$  cannot be characterized by a stochastic differential equation due to the possibility of exploitation.

#### 3.1 Bounded objective function

Unlike myopic or short-lived agents, the long-lived decision maker might attain infinite utility by exploring an arbitrarily large set of alternatives for an arbitrarily good alternative. Because such a strategy would also incur an infinite learning cost,

the objective function would be ill-defined as infinity minus infinity.<sup>3</sup>

The growth condition guarantees a bounded objective function by limiting exploration in finite time. The decision maker can choose from an exponentially growing but bounded set of alternatives whose running maximum increases at most exponentially in expectation. The restriction on  $k_1$  in the condition guarantees that the running maximum grows at a slower rate than the exponential discounting, leaving the objective bounded from above.

**Lemma 1 (Bounded objective function)** *Over the set of strategies that satisfy the growth condition with fixed constants  $k_1, k_2$ , the objective function (1) is bounded from above.*

## 3.2 Verification argument

I construct a candidate strategy based on some conjectures, and then verify the strategy's optimality. The key observation is that, among the many explored alternatives, the utility of the best known alternative and that of the frontier (i.e., rightmost alternative) are state variables, due to the Markovian property of the Brownian utility.

I first claim that exploitation is a lookback option that gives the running maximum,  $M_t := \max_{x \in [0, X_t]} U(x)$ , when executed. A lookback option is an option with a backward-looking and history-dependent payoff.<sup>4</sup> Because exploitation does not generate information or advance the frontier for future experiments, once the decision maker starts to exploit she will continue to do so by recalling the best known alternative.

Together with the running maximum, the frontier utility,  $W_t := U_{X_t}$ , characterizes the continuation value, i.e.,  $V_t = V(W_t, M_t)$ . This is because all unknown alternatives are to the right of  $X_t$ . Conditional on available information, the utility distribution of

---

<sup>3</sup>The utility is unbounded only for  $\mu \geq 0$ . In such a case, an example strategy (which violates the growth condition) is

$$x_t = \begin{cases} 2 \exp(2rn) (t - (n - 1)) & \text{for } t \in [n - 1, n - 1/2) \\ \arg \max\{U(x) : x \in [0, \exp(2rn)]\} & \text{for } t \in [n - 1/2, n). \end{cases}$$

<sup>4</sup>See [Conze and Viswanathan \(1991\)](#) for a discussion of lookback options in finance.



unknown alternatives depends only on  $W_t$  due to the Markov property of Brownian motion.

Next, I claim that the option value is a function of the drawdown. The option value is the difference between the continuation value and the value of exploitation, i.e.,  $v_t := V_t - M_t$ , and the drawdown is the difference between the frontier utility and the running maximum, i.e.,  $y_t := W_t - M_t$ . Note that the drawdown is weakly negative by definition. Suppose that both the frontier utility and the running maximum increase by the same constant. The utility distribution of unknown alternatives then shifts up by the same amount due to the stationary independent increments of Brownian utility. Therefore, both the value of exploration and the value of exploitation increase by the same amount, leaving the option value unchanged. The option value function can thus be written as  $v_t = v(y_t)$ .

I derive the laws of the frontier utility, running maximum, and drawdown by applying a time change to the utility process.

**Lemma 2 (Time change)** *The frontier utility satisfies  $W_t := \mu X_t + Z_t$  where the process  $Z_t$  is a continuous square-integrable  $\mathcal{F}$ -martingale with quadratic variation  $\sigma^2 \int_0^t s_q d\eta$ .*

The drift and volatility of  $W_t$  are both linear in the speed of exploration because the time change controls how quickly the frontier advances and experiences the variations in the domain of alternatives. The Dambis–Dubins–Schwarz theorem further implies that  $dZ_t = \sigma \sqrt{s_t} dB'_t$  and therefore  $dW_t = \mu s_t dt + \sigma \sqrt{s_t} dB'_t$ , where  $B'_t$  is a standard Brownian motion in a possibly enlarged probability space.<sup>5</sup> As the time change is continuous by the continuous exploration condition, the running maximum of explored alternatives equals that of the frontier utility, i.e.,  $M_t = \max_{\eta \in [0, t]} W_\eta$ . When the drawdown is strictly negative, the running maximum does not increase and therefore the drawdown shares the same law as the frontier utility, i.e.,  $dM_t = 0$  and  $dy_t = dW_t = \mu s_t dt + dZ_t$  for  $y_t < 0$ .

The decision maker executes the option to exploit when the drawdown exceeds a threshold, i.e., when the opportunity cost of exploitation exceeds the continuation

---

<sup>5</sup>I do not use this corollary of Lemma 2 because the optimal strategy it identifies may not be adapted to the original probability space; i.e., the strategy may be a weak solution to the decision maker's problem.

value of exploration. When the frontier utility decreases, the utility distribution of unknown alternatives shifts down due to the stationary independent increments. The continuation value therefore decreases. At the same time, the opportunity cost increases. Therefore, there exists a threshold  $\underline{y} < 0$  such that it is optimal to exploit if and only if  $y \leq \underline{y}$ . In addition, the stationary independent increments further implies that the threshold does not depend on the frontier utility or the running maximum. I denote by  $\tau$  the exploitation time, i.e.,

$$\tau = \inf\{t \geq 0 : y_t \leq \underline{y}\}. \quad (2)$$

I posit that, during exploration, the option value satisfies a Hamilton–Jacobi–Bellman equation (HJB) that captures the opportunity cost of exploitation and the flow value of exploration. To maximize the option value, the speed of exploration must maximize the Hamiltonian, i.e.,

$$v = \max_s y - c(s) + \frac{1}{r}\mu v' s + \frac{1}{r}\frac{\sigma^2}{2}v'' s \quad \forall y \in (\underline{y}, 0). \quad (3)$$

The option value consists of three components. First, the drawdown,  $y < 0$ , captures the opportunity cost of exploring at the frontier instead of exploiting the best known alternative. Second, the learning cost,  $c(s)$ , increases with the speed of exploration. Third, the flow value of exploration characterizes how the option value function changes with the drawdown. Because  $dy_t = \mu s_t dt + dZ_t$  for  $y_t < 0$ , the variation in drawdown translates to the option value via the drift term  $\frac{1}{r}\mu v' s$  and the Itô term  $\frac{1}{r}\frac{\sigma^2}{2}v'' s$ . These two terms are controlled by the speed of exploration by Lemma 2.

The optimal speed of exploration maximizes the flow value of exploration subject to the learning cost. As the speed controls how quickly the frontier advances in the domain of alternatives over time, the impatient decision maker can reap the value of exploration earlier in time by exploring at a higher speed subject to the convex learning cost. The optimal speed therefore solves the first-order condition (FOC),

$$c'(s) = \frac{1}{r}\mu v' + \frac{1}{r}\frac{\sigma^2}{2}v'', \quad (4)$$

which equates the marginal learning cost to the marginal value of exploration.

In addition to Equation (3), the option value function should satisfy three standard boundary conditions. First, the value-matching condition defines the drawdown threshold by equalizing the opportunity cost and the value of exploration. The option value function therefore vanishes at the threshold, i.e.,

$$v(\underline{y}) = 0. \tag{5}$$

Second, the smooth-pasting condition characterizes the optimal execution of the option to exploit. For  $y < \underline{y}$ , the decision maker exploits immediately and so  $v(y) = v'(y) = 0$ . If the marginal option value were positive at the threshold, i.e.,  $v'(\underline{y}) > 0$ , the kink would imply a positive infinite flow value of exploration and so the decision maker would explore more, and vice versa. Thus, optimal execution implies that

$$v'(\underline{y}) = 0. \tag{6}$$

Third, the super-contact condition determines the marginal value of frontier utility when the decision maker marks a new maximum. Because the running maximum increases by the same amount as the frontier utility, so does the continuation value, i.e.,

$$v'(0) = 1. \tag{7}$$

By constructing a candidate strategy that satisfies these conditions, I characterize the value function and the unique optimal strategy by a verification argument.

**Theorem 1 (Optimal experimentation)**

- *The continuation value satisfies  $V_t = V(W_t, M_t) = M_t + v(y_t)$ . The option value function  $v$  is the unique solution to Equation (3) subject to boundary conditions (5)–(7), with extension  $v(y) = 0$  for  $y < \underline{y}$ . Moreover,  $v$  is increasing and convex.*

- The optimal experimentation strategy  $\{x_t^*\}$  is unique and is given by

$$x_t^* = \begin{cases} \int_0^t s(y_\eta) d\eta & \text{for } t < \tau \\ \arg \max_{x \leq X_\tau} U(x) & \text{for } t \geq \tau, \end{cases}$$

where Equation (2) defines exploitation time  $\tau$ , and Equation (4) defines speed of exploration  $s$ .

## 4 Properties of optimal experimentation

I analyze how optimal experimentation evolves over time, depends on model parameters, and differs in the cross section.

### 4.1 Time-series properties

By analyzing two opposing incentives, I show that the more negative the drawdown, the higher speed of exploration. The speed controls the drawdown, which in turn determines the dynamics of exploration.

**Proposition 1** *The speed of exploration  $s$  is higher for the more negative drawdown  $y$ .*

Define  $g(s) := sc'(s) - c(s)$ . It is continuously differentiable and strictly increasing, and so is its inverse  $g^{-1}$  which exists. Plugging Equation (3) into Equation (4), I write the speed of exploration as a function of  $v - y$ , i.e.,

$$s = g^{-1}(v - y). \tag{8}$$

The function  $g^{-1}$  maps the sum of the option value and the opportunity cost (recall that  $y < 0$ ) to the optimal speed, subject to the convex learning cost. Note that neither  $v'$  or  $v''$  enters Equation (8). This is because the drift and volatility terms in Equation (3) are both linear in the rate of time change,  $s$ , and therefore can be substituted by Equation (4) simultaneously.

The intertemporal tradeoff between the flow opportunity cost of exploitation and the continuation value of exploration determines the optimal speed of exploration. On the one hand, the decision maker incurs the flow opportunity cost,  $|y|$ , when she explores new alternatives instead of exploiting the best known one. To avert the flow cost, she would like to explore quickly over a short period of time, and more so for more negative drawdowns when the opportunity cost is higher. On the other hand, the impatient decision maker would like to realize the continuation value,  $v$ , sooner by acquiring more information. For less negative drawdowns, the value is higher and therefore prompts faster exploration.

The speed of exploration is higher for more negative drawdowns because the opportunity cost is more sensitive to the drawdown than the continuation value, i.e.,  $v' < 1$ . When the drawdown experiences a negative shock at frontier  $X_t$ , the utilities of all alternatives to the right of  $X_t$  decrease while those to the left, including the best alternative, remain unchanged. As a result, the flow opportunity cost increase by the size of the shock.

However, the continuation value decreases by a smaller amount because the option to exploit partially insures against the shock. Given optimal execution of the option to exploit, the envelope theorem implies<sup>6</sup>  $v'(y_t) = \int_t^\infty r e^{-r(\eta-t)} \mathbb{P}_t [x_\eta \geq X_t] d\eta \in [0, 1]$ . In words, the shock at the current frontier is relevant to the continuation value only when the chosen alternative in the future,  $x_\eta$ , lies to the right of  $X_t$ . If the frontier utility reaches the drawdown threshold before marking a new maximum, the decision maker will exploit the best alternative, which lies to the left of  $X_t$ , and therefore insulates herself from the negative shock. Only when  $y_t = 0$ , alternative  $X_t$  is the best and so all future alternatives chosen during exploration and exploitation will lie to its right. The negative shock at  $X_t$  factors in the decision maker's flow utility permanently, implying that  $v'(0) = 1$ .

I provide a sufficient condition for the speed of exploration to increase on average in the time series.

**Corollary 1** *If  $\mu \leq 0$  and  $sc''(s)$  is decreasing over  $[s(0), s(y)]$ , then the speed of exploration  $s(y_t)$  is a submartingale over  $[0, \tau)$ .*

---

<sup>6</sup>One can also show  $v' < 1$  from the super-contact condition (7) and the strict convexity of  $v$ .

By Proposition 1, the negative drift in drawdown contributes to a higher speed of exploration over time. The volatility also contributes to it via the Itô term because  $v$  is convex. The condition on  $c$  guarantees that  $g^{-1}$  in Equation (8) is convex as well.<sup>7</sup> Note that the change in the running maximum does not factor in the calculation because  $s'(0) = 0$  by the smooth pasting condition (6).

## 4.2 Comparative statics

I derive the comparative statics of the option value function, the speed of exploration, and the drawdown threshold with respect to model parameters. I write that  $c$  increases if  $c'(0)$  increases and  $c''$  increases pointwise, and that  $v$  and respectively  $s$  increase if  $v$  and  $s$  increase pointwise. Moreover, a set of parameters is said to be more *favorable* if either the drift  $\mu$  is higher, the volatility  $\sigma^2$  is higher, the discount rate  $r$  is lower, or the cost function  $c$  is lower.

**Proposition 2 (Comparative statics)** *Option value  $v$ , speed of exploration  $s$ , and drawdown threshold  $|y|$  are increasing in the drift  $\mu$  and volatility  $\sigma^2$  of the utility process, and are decreasing in the discount rate  $r$  and the learning cost function  $c$ .*

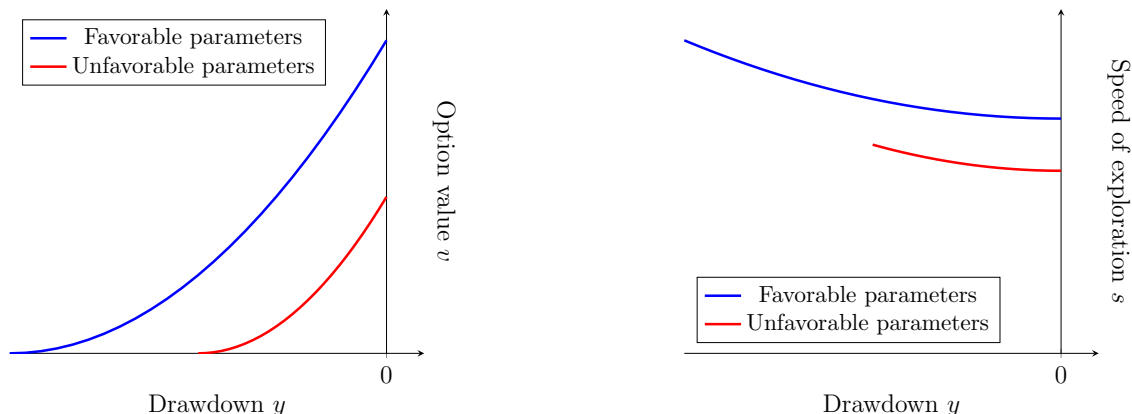


Figure 1: An illustration of the option value functions and the speed of exploration functions for two sets of parameters.

Intuitively, the decision maker derives higher value from exploration when it yields higher utility, the decision maker is more patient, or learning is less costly (Figure 1).

<sup>7</sup>The function  $sc''(s)$  is positive and attains 0 at  $s = 0$ , and therefore it is not decreasing monotonically.

As a result, she also tolerates a more negative drawdown threshold. Moreover, she is motivated to explore at a higher speed at any given drawdown in order to realize the increased value more quickly. The comparative statics holds when the utility process is more volatile as well, because the option to exploit insures against negative shocks.

### 4.3 Cross-sectional properties

In the application of firm experimentation, a firm experiments on its size,  $x \in [0, \infty)$ , to maximize expected discounted profitability,  $U(x)$ . I use the log number of employees as a proxy for firm size and log profit for profitability. The Brownian specification means that, over all firm sizes, the firm-size elasticity of profitability is independent and identically distributed. I interpret the learning cost,  $c$ , as the loss in profits when a firm first operating at a given firm size fails to allocate human resources to maximize profit. Starting at the minimal firm size, the optimal strategy is to grow continuously until the percentage drawdown,  $y$ , exceeds a threshold and then to scale down to exploit the most profitable size in the long run. The percentage growth rate,  $s$ , depends on the percentage drawdown.

Stochastic utility generates heterogeneous experimentation behaviors in the cross section under the optimal strategy. In the context of firm experimentation, I obtain two cross-sectional predictions on firm dynamics: the conditional version of Gibrat's law and the linear relation between long-run firm size and profitability.

First, I present the asymptotic convergence of the speed of exploration, which corresponds to the conditional version of Gibrat's law. With a slight abuse of notation, I index the drawdown and speed of exploration by the alternative rather than calendar time, i.e.,  $y_x := U(x) - \max_{[0,x]} U$  and  $s_x := s(y_x)$ , by undoing the time change  $t \mapsto X_t$ . This is possible because whether the drawdown has exceeded the threshold is invariant to the continuous time change, i.e.,  $y_s < \underline{y} \quad \forall s \leq t \Leftrightarrow y_x < \underline{y} \quad \forall x \leq X_t$ .

**Proposition 3 (Asymptotic convergence)** *As  $x \rightarrow \infty$ , the speed of exploration conditional on exploration,  $s_x | X_\tau > x$ , converges in distribution.*

The proposition concerns the few firms that grow beyond a large size even when most other firms have already stopped exploration. The growth rate of these remaining

few firms follows a stable distribution regardless of absolute firm size. In other words, the asymptotic convergence implies the conditional version of Gibrat’s law: among large firms, the percentage growth in firm size *conditioned on growth* is essentially invariant to absolute firm size.

The conditional version of Gibrat’s law can explain an empirical deviation from the original law, which states that the percentage growth in firm size is (unconditionally) invariant to absolute firm size. [Daunfeldt and Elert \(2013\)](#) reject the original, unconditional version of Gibrat’s law on the aggregate level for Swedish firms during 1998–2004 because small firms on average grow more quickly than large firms. By contrast, the conditional version of Gibrat’s law suggests that its deviation from the original law derives from the extensive margin: large and old firms are more likely to have stopped exploration and growth. Moreover, the conditional law weakens the original law parsimoniously in that the intensive margin—the percentage growth in firm size conditional on growth—is invariant.

The proof of Proposition 3 also implies that the decision maker quickly stops exploring.

**Corollary 2 (Short exploration)** *There exist  $K, \alpha > 0$  such that  $\mathbb{P}[\tau > t] \leq Ke^{-\alpha t}$  for all  $t \in [0, \infty)$ .*

Corollary 2 states that the tail probability is subexponential; i.e., exploration stops at least exponentially quickly.<sup>8</sup> Most firms stop growing and maintain a fixed size soon after establishment.

My second cross-sectional result is that the exploited alternative is linear in its utility. This result corresponds to the linear relation between long-run firm size and profitability. To present a stark contrast, I restrict attention to a driftless utility process, i.e.,  $\mu = 0$ , which means that the average elasticity of profit is zero. Denote the exploited alternative by  $x_\infty := \lim_{t \rightarrow \infty} x_t = x_\tau$  and its utility by  $M_\infty := \lim_{t \rightarrow \infty} M_t = M_\tau = U_{x_\tau}$ .

**Proposition 4 (Linear relation)** *For  $\mu = 0$ , the pair  $(x_\infty, M_\infty)$  satisfies the linear*

---

<sup>8</sup>Therefore, the decision maker will eventually stop exploring, as [Callander \(2011\)](#) and [Garfagnini and Strulovici \(2016\)](#) show in their models with short-lived agents.



relation

$$\mathbb{E}[M_\infty | x_\infty > x] = \frac{\pi^2 \sigma^2}{8 \underline{y}} x + o(x) \quad \text{as } x \rightarrow \infty.$$

Among large established firms, the average firm-size elasticity of profit is  $\pi^2 \sigma^2 / 8 \underline{y}$ , which is positive.<sup>9</sup>

Proposition 4 highlights an endogenous selection effect: large firms *choose* to operate at large sizes because they experience increasing profits along the growth path. As a result, the naïve regression of profitability on firm size would estimate a positive average elasticity despite the zero mean in the profit process, because the estimation overlooks the selection bias. The proposition implies that the bias persists even for very large firms.

## 5 Discussions on modeling assumptions

In this section, I discuss how my analysis depends on three modeling assumptions: the continuous exploration condition, Inada condition, and learning cost.

### 5.1 Continuous exploration condition

The continuous exploration condition in Definition 1 is without loss of optimality under convex learning costs. The idea is that, in continuous time, fast continuous exploration approximates discontinuous jumps, and small but frequent jumps can approximate continuous explorations.

Suppose that the decision maker can explore discontinuously subject to a discrete learning cost  $C(D_t)$ , where  $D_t$  is the distance of the chosen alternative from previous explorations and  $C : [0, \infty) \rightarrow [0, \infty)$  is the discrete cost function. I assume that  $C(0) = 0$  and that  $C$  is convex as in [Garfagnini and Strulovici \(2016\)](#). This formulation is consistent with the flow learning cost of continuous exploration because the

---

<sup>9</sup>The proof of Proposition 4 establishes that  $M_\infty$ ,  $x_\infty$ , and  $X_\tau$  all exhibit exponential tails. Moreover, the proof shows that all three variables satisfy pairwise linear relations, and thus implies selection biases in the pairwise regressions in the long-run profitability, long-run firm size, and historic maximum firm size in the context of firm experimentation.

speed of exploration measures the infinitesimal distance from previous explorations.

I claim that the discrete cost function is effectively linear. Because  $C$  is convex, it suffices to show subadditivity, i.e.,  $C(D) \leq C(D') + C(D - D')$  for  $D' < D$ . Instead of exploring an alternative at distance  $D$  in one step, the decision maker can first explore the middle point at distance  $D'$  and then explore the chosen alternative, now at distance  $D - D'$ . Over an arbitrarily short time lapse, the two-step exploration reveals additional information at the middle point while accumulating zero flow utility. The cost  $C(D') + C(D - D')$  therefore bounds the effective learning cost from below.

The linearity of the discrete cost function implies the optimality of continuous exploration. For the case  $c'(\infty) \leq C'$ , the statement follows because the decision maker can explore continuously at an arbitrarily high speed instead of discontinuously. Such continuous exploration incurs strictly lower learning cost because of the Inada condition. Now, let us consider the case  $c'(\infty) > C'$ . Instead of exploring continuously at speed  $s > \bar{s} := (c')^{-1}(C')$ , the decision maker can mix or “chatter” between exploring continuously at speed  $\bar{s}$  and discontinuously in arbitrarily small steps.<sup>10</sup> For the same distance per unit time, the mixed exploration incurs a flow cost of  $c(\bar{s}) + (s - \bar{s})C'$ , which is lower than  $c(s)$ . Because the utility process is continuous, the mixed exploration provides essentially the same information as continuous exploration when the step size vanishes. The effective flow learning cost function

$$c_E(s) := \begin{cases} c(s) & \text{for } s \leq \bar{s} \\ c(\bar{s}) + (s - \bar{s})C' & \text{for } s > \bar{s} \end{cases}$$

falls into the first case described above and so continuous exploration is weakly optimal, but not necessarily strictly so because the effective cost function fails the Inada condition.

**Remark 2** The above argument fails if  $C$  is not convex. When the decision maker explores discontinuously, the (closure of the) set of explored alternatives may no longer be an interval. In addition to the utility at the rightmost alternative, the decision maker also needs to take the utility at each boundary point into account. The number of state variables grows with each discontinuous exploration, leaving the

---

<sup>10</sup>Moscarini and Smith (2001) make a related convexification argument and attribute it to Paul Milgrom.

recursive analysis intractable.

## 5.2 Inada condition

The Inada condition guarantees an interior solution to Equation (4). If the learning cost function  $c$  fails the condition, it may be optimal to explore at infinite speed. Such instantaneous exploration realizes the continuation value immediately and therefore averts the flow opportunity cost. The decision maker does so over an interval of alternatives until the drawdown reaches the threshold or becomes less negative whereupon the optimal speed becomes finite again.<sup>11</sup> Despite the more involved technicalities, such instantaneous explorations yield the same economic insights.

## 5.3 Learning cost

My analysis extends to the experimentation problem with an adjustment friction in place of the learning cost. I model the adjustment friction by restricting the chosen alternative to be continuous, with a given Lipschitz constant; in other words, the speed of change is uniformly bounded. The key difference is that the decision maker can no longer exploit the best known alternative instantaneously, but can only take time to backtrack from the frontier.

I denote the highest value from backtracking from an given alternative by the exploitation value. The exploitation value–frontier utility pair can be shown to be a Markov process over the domain of alternatives.

Similar to the case with a learning cost, the optimal strategy is to explore until the generalized drawdown—i.e., the difference between the exploitation value and the frontier utility—exceeds a threshold, and then backtrack to exploit the alternative that attains the exploitation value. While the maximum utility increases monotonically during exploration, the exploitation value decreases at a rate proportional to the generalized drawdown, reflecting the lower flow utility when backtracking from the frontier. As a result, the exploited alternative may differ from the best one.

---

<sup>11</sup>Technically, the value is not attained; instead, it is approached by continuous exploration at an arbitrarily high speed.

## References

- Borodin, Andrei N and Paavo Salminen (2012). *Handbook of Brownian Motion: Facts and Formulae*. Birkhäuser.
- Callander, Steven (2011). “Searching and learning by trial and error”. In: *American Economic Review* 101.6, pp. 2277–2308.
- Conze, Antoine and Vish Viswanathan (1991). “Path dependent options: The case of lookback options”. In: *Journal of Finance* 46.5, pp. 1893–1907.
- Daunfeldt, Sven-Olov and Niklas Elert (2013). “When is Gibrat’s law a law?” In: *Small Business Economics* 41.1, pp. 133–147.
- Garfagnini, Umberto and Bruno Strulovici (2016). “Social experimentation with interdependent and expanding technologies”. In: *Review of Economic Studies* 83.4, pp. 1579–1613.
- Jovanovic, Boyan and Rafael Rob (1990). “Long waves and short waves: Growth through intensive and extensive search”. In: *Econometrica*, pp. 1391–1409.
- Karatzas, Ioannis (1984). “Gittins indices in the dynamic allocation problem for diffusion processes”. In: *Annals of Probability* 12.1, pp. 173–192.
- Keller, Godfrey, Sven Rady, and Martin Cripps (2005). “Strategic experimentation with exponential bandits”. In: *Econometrica* 73.1, pp. 39–68.
- Lehoczky, John P (1977). “Formulas for stopped diffusion processes with stopping times based on the maximum”. In: *Annals of Probability*, pp. 601–607.
- Moscarini, Giuseppe and Lones Smith (2001). “The optimal level of experimentation”. In: *Econometrica* 69.6, pp. 1629–1644.
- Peskir, Goran (2006). “On reflecting Brownian motion with drift”. In: *Proceedings of the ISCIE International Symposium on Stochastic Systems Theory and its Applications*. Vol. 2006. The ISCIE Symposium on Stochastic Systems Theory and Its Applications, pp. 1–5.
- Revuz, Daniel and Marc Yor (2013). *Continuous Martingales and Brownian Motion*. Vol. 293. Springer Science & Business Media.
- Strauss, Walter A (2007). *Partial Differential Equations: An Introduction*. John Wiley & Sons.
- Taylor, Howard M (1975). “A stopped Brownian motion formula”. In: *Annals of Probability*, pp. 234–246.

- Urgun, Can and Leeat Yariv (2025). “Contiguous Search: Exploration and Ambition on Uncharted Terrain”. In: *Journal of Political Economy* 133.2, pp. 000–000.
- Weitzman, Martin L (1979). “Optimal search for the best alternative”. In: *Econometrica*, pp. 641–654.

# A Proofs

## A.1 Proof of Lemma 1

**Proof.** By the growth condition, the expectation of the running maximum of  $U$  over  $[0, k_2 \exp(k_1 t)]$  is bounded from above, i.e.,

$$\mathbb{E} \left[ \sup_{[0, k_2 \exp(k_1 t)]} U \right] \leq \mu^+ k_2 e^{k_1 t} + \sqrt{\frac{2}{\pi} \sigma^2 k_2} \exp\left(\frac{k_1}{2} t\right),$$

where  $\mu^+ k_2 e^{k_1 t}$  is the maximum cumulative drift and  $\sqrt{2\sigma^2 k_2 / \pi} \exp(k_1 t / 2)$  is the expected running maximum of the driftless Brownian motion  $\sigma B$ . Invoking the Fubini theorem, I bound the objective function from above by

$$\int_0^\infty e^{-rt} \left( \mu^+ k_2 e^{k_1 t} + \sqrt{\frac{2}{\pi} \sigma^2 k_2} \exp\left(\frac{k_1}{2} t\right) \right) dt = \mu^+ k_2 (r - k_1)^{-1} + \sqrt{\frac{2}{\pi} \sigma^2 k_2} \left( r - \frac{k_1}{2} \right)^{-1} < \infty.$$

■

## A.2 Proof of Lemma 2

This lemma adapts Proposition 1.1.5 in Chapter V of [Revuz and Yor \(2013\)](#) to my setting where  $\mathcal{F}$  is defined over  $t$  instead of  $x$ . The time change may remind the reader of the celebrated Dambis–Dubins–Schwarz theorem (see, for example, Theorem 1.1.6 in Chapter V of [Revuz and Yor \(2013\)](#)). However, my model specifies a standard Brownian motion  $B_x$  (and utility process  $U_x$ ) as a primitive and then constructs its time-change  $W_t := U_{X_t}$ . By comparison, the theorem specifies a continuous martingale and then shows the existence of a Brownian motion in a possibly enlarged probability space such that the martingale is a time-change of that Brownian motion. In particular, my lemma does not require such enlargement.

**Proof.** Define  $Z := W - \mu X$ . The process  $Z$  is continuous by composition and

is adapted to  $\mathcal{F}$  by the definition of  $\mathcal{F}$ . For  $\eta \geq t$ , its expectation satisfies

$$\mathbb{E}[Z_\eta | \mathcal{F}_t] = \mathbb{E}[Z_\eta | \mathcal{F}_{X_t}^B] = Z_t$$

where  $\mathcal{F}^B$  is the natural filtration of  $B$ . The first equality follows because  $x$  is predictable by the measurability condition. The second is due to the optional sampling theorem because  $U_x - \mu x = \sigma B_x$  is a uniformly integrable  $\mathcal{F}^B$ -martingale on  $[0, k_2 e^{k_1 \eta}]$  by the growth condition. Therefore,  $Z$  is an  $\mathcal{F}$ -martingale.

The same argument shows that  $Z^2 - \sigma^2 X$  is a  $\mathcal{F}$ -martingale, because  $(U_x - \mu x)^2 - \sigma^2 x$  is a uniformly integrable  $\mathcal{F}^B$ -martingale on  $[0, k_2 e^{k_1 \eta}]$  for all  $\eta \geq 0$ . Therefore,  $Z$  is squared-integrable and its quadratic variation is  $\sigma^2 X$  by definition. ■

### A.3 Proof of Theorem 1

Following the conjectures in the main text, I construct a candidate strategy (Proposition 5 and Corollary 3) and then verify its optimality and uniqueness.

**Proposition 5** *There exist constants  $\underline{y} < 0$ ,  $\bar{v} > 0$ , and a convex function  $v : [\underline{y}, 0] \rightarrow [0, \infty)$  with  $v(0) = \bar{v}$  such that  $v$  satisfies HJB (3) and boundary conditions (5), (6), and (7). For such  $v$ , the unique maximizer  $s^*(y)$  is given by FOC (4).*

**Proof.** I prove the existence to the free boundary problem by the shooting method. First, I transform the HJB to an autonomous ODE system by a change of variables. Second, I prove the existence to an initial value problem that omits the value matching condition and the smooth pasting condition (Lemma 3). Third, I derive the monotonicity (Lemma 4) and asymptotic properties (Lemma 5 and Lemma 6) of such solutions. Fourth, I show the two omitted conditions hold for one such solution, which therefore solves the free boundary problem (Lemma 7).

Consider the change of variables  $\phi(y) := v(y) - y$ . By substituting the speed of exploration (8) into HJB (3), I obtain

$$\phi'' + \frac{\mu}{\sigma^2/2} \phi' = \frac{r}{\sigma^2/2} f(\phi) - \frac{\mu}{\sigma^2/2} \quad (9)$$

on  $y \leq 0$ , where  $f := c' \circ g^{-1} : [0, \infty) \rightarrow [c'(0), c'(\infty))$  is increasing. The boundary

conditions translate to  $\phi'(0) = 0$ ,  $\phi(\underline{y}) = -\underline{y}$ , and  $\phi'(\underline{y}) = -1$ .

Define  $\underline{v}$  to be 0 for  $\mu \leq 0$  and  $f^{-1}(\mu/r)$  for  $\mu > 0$ . For  $v_0 > \underline{v}$  so that  $\phi''(0) > 0$ , consider the initial value problem of ODE (9) with initial conditions  $\phi(0) = v_0$  and  $\phi'(0) = 0$ . Such  $v_0$  exists because  $c'(\infty) > \mu^+/r$ .

**Lemma 3** *There exists a unique solution  $\phi$  to the initial value problem. Moreover,  $\phi'' > 0$ .*

**Proof.** By the Picard–Lindelöf theorem, there exists a unique solution to the initial value problem on  $\phi \geq 0$ , where  $f(\phi)$  is well-defined.

To show  $\phi'' > 0$  by contradiction, I suppose  $\tilde{y} := \sup\{y \in (-\infty, 0) : \phi''(y) = 0\} > -\infty$ . For  $[\tilde{y}, 0]$ , I have  $\phi \geq 0$  because  $\phi'' > 0$  by continuity. The supremum is attained also by continuity. Moreover, I have  $\phi'(y) = 0 + \int_0^y \phi'' < 0$  and thus  $\phi(y) = \phi(0) + \int_0^y \phi' > \phi(0)$ . Differentiating ODE (9) at  $\tilde{y}$ , I obtain

$$\phi''' = \frac{r}{\sigma^2/2} f'(\phi) \phi' - \frac{\mu}{\sigma^2/2} \phi'' = \frac{r}{\sigma^2/2} f'(\phi) \phi' < 0.$$

Therefore,  $\phi''(\tilde{y} + \epsilon) < 0$  for sufficiently small  $\epsilon > 0$ , which is a contradiction.

Because  $\phi'' > 0$  implies  $\phi > 0$ , the solution exists and is unique on  $(-\infty, 0]$ . ■

**Lemma 4**  *$\phi(y; v_0)$  is continuously differentiable with respect to  $v_0$ , and its  $v_0$ -derivative is positive and strictly decreasing in  $y$ .*

**Proof.** Because  $f \in \mathcal{C}^1$ , the derivative of  $\phi$  with respect to  $v_0$  is the solution to the variational equation

$$z'' + \frac{\mu}{\sigma^2/2} z' = \frac{r}{\sigma^2/2} f'(\phi) z$$

with initial conditions  $z(0) = 1$  and  $z'(0) = 0$ . It then suffices to show that  $z' < 0$  on  $(-\infty, 0)$ .

To obtain a contradiction, I suppose  $\tilde{y} := \sup\{y < 0 : z'(y) = 0\} > -\infty$ . The supremum is attained because  $z''(0) = \frac{r}{\sigma^2/2} f'(\phi(0)) > 0$ . I have  $z'(y) < 0$  for  $y \in (\tilde{y}, 0)$  and  $z(y) > 0$  for  $y \in [\tilde{y}, 0]$ . At  $\tilde{y}$ , I have  $z'' = (2r/\sigma^2) f(\phi) z - (2\mu/\sigma^2) z' > 0$  as  $f(\phi) > 0$ . Therefore,  $z'(\tilde{y} + \epsilon) > z'(\tilde{y}) = 0$  for sufficiently small  $\epsilon > 0$ , which is a contradiction. ■



**Lemma 5** For  $v_0$  sufficiently close to  $\underline{v}$ , the solution  $\phi(\cdot; v_0)$  intersects with the 45-degree line,  $-y$ .

**Proof.** For the case of  $\mu \leq 0$ , I have  $\underline{v} = 0$ . The initial conditions satisfy  $\phi(0; 0) = 0$  and  $\phi'(0; 0) = 0 > -1$ . Moreover,  $\phi''(0; 0) = (2r/\sigma^2)f(0) - (2\mu/\sigma^2) \geq 0$  and so  $\phi$  exists and is unique in a neighborhood of 0. I have  $\phi(y; \underline{v}) < -y$  for some  $y < 0$  close enough to zero.

For the case of  $\mu > 0$ , I have  $\underline{v} := f^{-1}(\mu/r)$ . The solution is the constant function  $\phi(\cdot; \underline{v}) = \underline{v}$ . It satisfies  $\phi(y; \underline{v}) = \underline{v} < -y$  for some  $y < -\underline{v}$ .

In either case, I have  $\phi(y; \underline{v}) < -y$  for some  $y < 0$ . By the continuity of  $\phi$  with respect to  $v_0$ , I have  $\phi(y; \underline{v} + \epsilon) < -y$  for sufficiently small  $\epsilon > 0$ . Because  $\phi(0; \underline{v} + \epsilon) \geq \epsilon > 0$ ,  $\phi(\cdot; \underline{v} + \epsilon)$  intersects with  $-y$  by the intermediate value theorem.

■

**Lemma 6** For sufficiently large  $v_0$ , the solution  $\phi(\cdot; v_0)$  does not intersect with the 45-degree line,  $-y$ .

**Proof.** For the case of  $\mu \leq 0$ , define

$$L := \inf \left\{ l \in (0, \infty) : \frac{r}{\sigma^2/2} f(l) - \frac{\mu}{\sigma^2/2} > l^{-1} \right\}.$$

The infimum is finite and strictly positive because the LHS of the inequality is increasing, and the RHS is decreasing and spans  $(0, \infty)$ . For  $v_0 > L$ , I have

$$\phi'' = \frac{r}{\sigma^2/2} f(\phi) - \frac{\mu}{\sigma^2/2} - \frac{\mu}{\sigma^2/2} \phi' > \frac{r}{\sigma^2/2} f(v_0) - \frac{\mu}{\sigma^2/2} > v_0^{-1}.$$

Therefore,  $\phi'(-v_0) = \int_0^{-v_0} \phi'' < \int_0^{-v_0} v_0^{-1} = -1$ . On  $[-v_0, 0)$ , I have  $\phi > v_0 > -y$  since  $\phi'' > 0$  and  $\phi'(0) = 0$ ; on  $(-\infty, -v_0)$ , I also have  $\phi > -y$  because  $\phi'(-v_0) < -1$  and  $\phi'' > 0$ . The lemma therefore obtains.

For the case of  $\mu > 0$ , define

$$L := \inf \left\{ l \in (0, \infty) : \frac{r}{\sigma^2/2} f(l) > l^{-1} \right\}.$$

As in the previous case, the infimum is finite and strictly positive. For  $v_0 > L$ , I show  $\phi'(-v_0) < -1$  by contradiction. Suppose otherwise. I have

$$\phi'' = \frac{r}{\sigma^2/2}f(\phi) - \frac{\mu}{\sigma^2/2}(\phi' + 1) > \frac{r}{\sigma^2/2}f(v_0) > v_0^{-1}.$$

Therefore,  $\phi'(-v_0) = \int_0^{-v_0} \phi'' < \int_0^{-v_0} v_0^{-1} = -1$  which is a contradiction. On  $[-v_0, 0]$ , I have  $\phi > v_0 \geq -y$ ; on  $(-\infty, -v_0)$ , I have  $\phi > -y$  because  $\phi'(-v_0) < -1$ . The lemma therefore obtains. ■

**Lemma 7** *There exists  $\bar{v} > \underline{v}$  and  $\underline{y} < 0$  such that  $\phi(y; \bar{v}) \geq -y$  on  $[\underline{y}, 0]$  with equality at  $\underline{y}$  and  $\phi'(\underline{y}; \bar{v}) = -1$ .*

**Proof.** Let  $\bar{v}$  be the supremum of initial values  $v_0$ 's solutions of which intersect with the 45-degree line. It exists by Lemma 5 and Lemma 6. Let  $\bar{\phi} := \phi(\cdot; \bar{v})$  denote the corresponding solution. Let  $\{v_n : n \in \mathbb{N}\}$  denote a maximizing sequence. For each  $n$ , let  $\phi_n := \phi(\cdot; v_n)$  denote the solution for  $v_n$ , and  $y_n := \max\{y : \phi_n(y) = -y\} < 0$  denote the first intersection. The sequence  $\{y_n\}$  is uniformly bounded by  $L$  by the proof of Lemma 6 and therefore admits a converging subsequence. With a slight abuse of notation, let  $\{v_n, y_n\}$  denote the converging subsequence. Define  $\underline{y} := \inf_n y_n$ .

For all  $n$ , I have  $\phi'_n(y_n) \geq -1$  because  $\phi_n$  intersects  $-y$  from above.  $\{\phi'_n\}$  is uniformly bounded from below by  $-1$  on  $[\underline{y}, 0]$ , because  $\phi'_n$  is increasing in the initial value  $v_n$  and decreasing in  $y$  by Lemma 4.

The first derivative  $\{\phi'_n\}$  is uniformly bounded by  $[-1, 0]$  and so  $\{\phi_n\}$  is uniformly Lipschitz. Moreover,  $\{\phi_n\}$  is uniformly bounded by  $L + |\underline{y}|$  on  $[\underline{y}, 0]$  because  $\phi'_n \geq -1$ . The Arzelà–Ascoli theorem therefore implies that the sequence admits a uniformly converging subsequence. The uniqueness of the initial value problem further implies that the limit is  $\bar{\phi}$ . Moreover, the uniform convergence implies  $\bar{\phi}(\underline{y}) = \lim_n \phi_n(y_n) = \lim_n -y_n = -\underline{y}$ ; i.e.,  $\bar{\phi}$  intersects with the 45-degree line. It also implies that  $\bar{\phi}(y) > -y$  on  $(\underline{y}, 0]$

I show that  $\{\phi''_n\}$  is uniformly bounded on  $[\underline{y}, 0]$ . For the case of  $\mu \leq 0$ , I have

$$\phi''_n = \frac{r}{\sigma^2/2}f(\phi_n) - \frac{\mu}{\sigma^2/2} - \frac{\mu}{\sigma^2/2}\phi'_n < \frac{r}{\sigma^2/2}f(-\underline{y}) - \frac{\mu}{\sigma^2/2}.$$

because  $\phi'_n < 0$ . For the case of  $\mu > 0$ , I have

$$\phi''_n = \frac{r}{\sigma^2/2}f(\phi_n) - \frac{\mu}{\sigma^2/2}(\phi'_n + 1) \leq \frac{r}{\sigma^2/2}f(-\underline{y})$$

because  $\phi'_n$  is bounded from below by  $-1$ .

The uniformly bounded  $\{\phi''_n\}$  implies that  $\{\phi'_n\}$  is uniformly Lipschitz. Because the first derivative is uniformly bounded between  $[-1, 0]$ , the Arzelà–Ascoli theorem implies that  $\{\phi'_n\}$  admits a uniformly converging subsequence. By the uniqueness of the initial value problem, the limit is  $\bar{\phi}'$ . Moreover, the uniform convergence implies  $\bar{\phi}'(\underline{y}) = \lim_n \phi'_n(\underline{y}_n) \geq -1$ .

It remains to show  $\bar{\phi}'(\underline{y}) \leq -1$ . Suppose  $\bar{\phi}'(\underline{y}) > -1$ . I have  $\bar{\phi}(\underline{y}-\delta) < -(\underline{y}-\delta)$  for sufficiently small  $\delta > 0$ . By the continuity of the initial value problem,  $\phi(\underline{y}-\delta; \bar{v}+\epsilon) < -(\underline{y}-\delta)$  so  $\phi(\cdot; \bar{v}+\epsilon)$  intersects the 45-degree line, which contradicts the fact that  $\bar{v}$  is the supremum. ■

Because  $\phi(\cdot; \bar{v})$  solves ODE (9) and the corresponding boundary conditions, I obtain the desired function  $v$  by reverting the change of variables via  $v := \phi - y$ . ■

Proposition 5 allows me to extend the construction from  $[\underline{y}, 0]$  to  $(-\infty, 0]$ .

**Corollary 3** *Any function  $v$  defined in Proposition 5, with the extension of value 0 for  $y \in (-\infty, \underline{y})$ , solves HJB equation*

$$\max \left\{ \max_s -rv + ry - rc(s) + \mu v's + \frac{\sigma^2}{2}v''s, 0 \right\} = 0 \quad \text{for } y \in (-\infty, 0]$$

subject to the super-contact condition  $v'(0) = 1$ .

**Remark 3** When  $\mu = 0$ , the solution to the free boundary problem can also be characterized by

$$\int_{v(0)}^{v(y)+y} \left( 2 \int_{v(0)}^a \frac{2r}{\sigma^2} f^{-1}(b) db \right)^{-\frac{1}{2}} da = -y \quad \forall y \in (\underline{y}, 0).$$

This integral equation characterization is possible because the first-order derivative vanishes in HJB (9).

**Proof of Part 1 of Theorem 1.** Define the candidate continuation value

function as  $\bar{V}(W, M) := M + v(W - M)$  for  $W \leq M$ . Note that the function is convex. Recall that  $Z := W - \mu X$ .

For any strategy  $\{x_t\}$ , Itô's lemma gives

$$\begin{aligned}
& e^{-rt} \bar{V}(W_t, M_t) \\
&= \bar{V}(W_0, M_0) + \int_0^t e^{-r\eta} \left( -r\bar{V} d\eta + \mu\bar{V}_W dX_\eta + \frac{1}{2}\bar{V}_{WW} d\langle Z \rangle_\eta + \bar{V}_W dZ_\eta + \bar{V}_M dM_\eta \right) \\
&= \bar{V}(W_0, M_0) + \int_0^t e^{-r\eta} \left( \left( -r(M_\eta + v) + \mu v' s + \frac{\sigma^2}{2} v'' s \right) d\eta + v' dZ_\eta + (1 - v') dM_\eta \right) \\
&= \bar{V}(W_0, M_0) + \int_0^t e^{-r\eta} \left( \left( -r(M_\eta + v) + \mu v' s + \frac{\sigma^2}{2} v'' s \right) d\eta + v' dZ_\eta \right) \tag{10}
\end{aligned}$$

for any  $t \in \mathbb{R}_+$ . In the first equality, the Itô terms for  $M$  are zero because  $M$  is an increasing process. The quadratic variation of  $Z$  in the second equality follows from Lemma 2. In the third equality, the  $dM_\eta$  term vanishes because both its components are zero:

$$\int_0^t e^{-r\eta} (1 - v') dM_\eta = \int_0^t e^{-r\eta} (1 - v') \mathbf{1}_{W_\eta = M_\eta} dM_\eta + \int_0^t e^{-r\eta} (1 - v') \mathbf{1}_{W_\eta < M_\eta} dM_\eta.$$

The first component on the right-hand side is identically zero by the super contact condition. The second component is also zero according to the Skorokhod equation; i.e., the maximum of a continuous process does not increase when the process is away from its maximum.

Taking expectation (over  $\mathcal{F}$ ) of Equation (10), I obtain

$$\mathbb{E}_0 [e^{-rt} \bar{V}(W_t, M_t)] = \bar{V}(W_0, M_0) + \mathbb{E}_0 \left[ \int_0^t e^{-r\eta} \left( -r(M_\eta + \bar{v}) + \frac{1}{2} s \bar{v}'' \right) d\eta \right].$$

The expectation of the  $dZ_\eta$  term vanishes because it is a martingale.

Including the flow utility and learning costs over  $[0, t]$  and taking expectations, I

obtain

$$\begin{aligned} & \mathbb{E}_0 \left[ e^{-rt} \bar{V}(W_t, M_t) + \int_0^t r e^{-r\eta} (U(x_\eta) - c(s_\eta)) d\eta \right] \\ = & \bar{V}(W_0, M_0) + \mathbb{E}_0 \left[ \int_0^t e^{-r\eta} \left( -rv + r(U(x_\eta) - M_\eta) - rc(s_\eta) + \mu v' s_\eta + \frac{\sigma^2}{2} v'' s_\eta \right) d\eta \right]. \end{aligned}$$

For  $s_\eta > 0$ , the chosen alternative is at the frontier  $x_\eta = X_\eta$  and so  $U(x_\eta) = W_\eta$ . In this case, Corollary 3 implies that the integrand is weakly negative. For  $s_\eta = 0$ , I have  $U(x_\eta) \leq M_\eta$ . The integrand is also weakly negative since  $v$  is weakly positive.

By construction,  $\{x_t^*\}$  attains the maximum integrand (at zero) over all  $s$  because it satisfies the HJB equation in Corollary 3. For this strategy, the speed is bounded and so the learning cost is uniformly integrable. As  $t \rightarrow \infty$ , the terminal value,  $e^{-rt} \bar{V}(W_t, M_t)$ , vanishes and so the value of  $\{x_t^*\}$  is given by  $V\{x_t^*\} = \bar{V}(W_0, M_0)$ .

Consider an arbitrary strategy  $\{x_t\}$ . The growth condition implies that  $e^{-r\eta} U(x_\eta)$  is uniformly integrable. Fatou's lemma gives

$$\begin{aligned} & \mathbb{E}_0 \left[ \int_0^\infty r e^{-r\eta} (U(x_\eta) - c(s_\eta)) d\eta \right] \\ \leq & \bar{V}(W_0, M_0) + \limsup_{t \rightarrow \infty} \mathbb{E}_0 \left[ \int_0^t e^{-r\eta} \left( -rv + r(U(x_\eta) - M_\eta) - rc(s_\eta) + \mu v' s_\eta + \frac{\sigma^2}{2} v'' s_\eta \right) d\eta \right] \\ & (11) \\ \leq & \bar{V}(W_0, M_0) \end{aligned}$$

where the second inequality follows from Corollary 3. Therefore,  $V = \bar{V}$  and  $\{x_t^*\}$  is an optimal strategy. ■

**Proof of Part 2 of Theorem 1.** For any optimal strategy, the integrand in Equation (11) is zero  $t$ -almost everywhere with probability one. Within the threshold  $W_t - M_t \in (\underline{y}, 0]$ , it is uniquely maximized by  $s_t = s^*(W_t - M_t)$ . Beyond the threshold  $W_t - M_t < \underline{y}$ , the integrand is maximized if and only if  $s_t = 0$  and  $U(x_t) = M_t$ . Therefore, the process  $X_t = \int_0^t s_\eta d\eta$  is unique. Within the threshold,  $x_t = X_t$  and is therefore unique. For the unique  $X$ , the threshold is reached only when  $t \geq \tau$ . Then,  $x_t \in \arg \max_{[0, X_t]} U = \arg \max_{[0, X_\tau]} U$  is unique almost surely. ■

## A.4 Proof of Proposition 2

The proof of Proposition 2 relies on two single-crossing lemmata about the solution to initial value problem (9) with respect to model parameters (Lemma 9 and Lemma 10). They imply the monotonicity of  $\bar{v}$  and  $\underline{y}$  (Corollary 4 and Corollary 5). Given the monotonic endpoints, the single-crossing lemmata strengthen the monotonicity results to the functions  $v$  (Lemma 11) and  $s$  point-wise (Lemma 12).

Consider two sets of parameters labeled by  $i = 1, 2$  that are identical except either  $r_1 > r_2$ ,  $c_1 > c_2$ ,  $\mu_1 < \mu_2$ , or  $\sigma_1^2 < \sigma_2^2$ . Recall that  $f := c' \circ g^{-1}$ .

**Lemma 8** *If  $c_1 > c_2$ , then  $f_1 > f_2$  point-wise.*

**Proof.** See the proof of Proposition 5 (c) in [Moscarini and Smith \(2001\)](#). ■

For  $i = 1, 2$ , let  $\phi_i$  denote the solution to ODE (9) subject to initial condition  $\phi_i(0) = v_i > \underline{v}_i$  and  $\phi_i'(0) = 0$  from  $y = 0$  until  $\phi_i' \geq -1$ , for parameters  $i$ .

**Lemma 9 (Single crossing, starting from 0)**

- *If  $v_1 \geq v_2$ , then  $\phi_1 > \phi_2$  and  $\phi_1' < \phi_2'$  for all  $y < 0$ .*
- *If  $v_1 < v_2$ , then there is at most one  $\bar{y} < 0$  such that  $\phi_1(\bar{y}) = \phi_2(\bar{y})$ . Moreover,  $\phi_1 > \phi_2$  and  $\phi_1' < \phi_2'$  for all  $y < \bar{y}$ .*

**Proof.** I first prove the lemma for either  $r_1 > r_2$  or  $c_1 > c_2$ . By Lemma 8,  $r_1 f_1 > r_2 f_2$  point-wise.

For the case of  $v_1 \geq v_2$ , I prove  $\phi_1 > \phi_2$  on  $(-\infty, 0)$  by showing  $\phi_1' < \phi_2'$  there. Suppose on the contrary  $\tilde{y} := \sup\{y < 0 : \phi_1'(y) = \phi_2'(y)\} > -\infty$ . It is attained by continuity of  $\phi'$ 's and  $\phi_1' < \phi_2'$  in the neighborhood of 0, because  $\phi_1''(0) = (2r_1/\sigma^2)f_1(v_1(0)) - 2\mu/\sigma^2 > (2r_2/\sigma^2)f_2(v_2(0)) - 2\mu/\sigma^2 = \phi_2''(0)$ . Moreover,  $\phi_1 > \phi_2$  on  $[\tilde{y}, 0)$  because  $\phi_1' < \phi_2'$  there. At  $\tilde{y}$ , I have

$$\phi_1'' = \frac{r_1}{\sigma^2/2}f_1(\phi_1) - \frac{\mu}{\sigma^2/2}(1 + \phi_1') > \frac{r_2}{\sigma^2/2}f_2(\phi_2) - \frac{\mu}{\sigma^2/2}(1 + \phi_2') = \phi_2''.$$

Therefore,  $\phi_1'(\tilde{y} + \epsilon) > \phi_2'(\tilde{y} + \epsilon)$  which is a contradiction.

For the case of  $v_1 < v_2$ , I show that  $\phi_1$  and  $\phi_2$  crosses at most once. Let  $\bar{y} := \max\{y < 0 : \phi_1(y) = \phi_2(y)\}$  denote the first crossing. Because  $\phi_1(0) < \phi_2(0)$ , I have  $\phi_1'(\bar{y}) \leq \phi_2'(\bar{y})$ . Whenever  $\phi_1'(\bar{y}) = \phi_2'(\bar{y})$ , I have  $\phi_1''(\bar{y}) = (2r_1/\sigma^2)f_1(\phi_1) - (2\mu/\sigma^2)(1 + \phi_1') > (2r_2/\sigma^2)f_1(\phi_2) - (2\mu/\sigma^2)(1 + \phi_2') = \phi_2''(\bar{y})$ , and so  $\phi_1'(\bar{y} - \epsilon) < \phi_2'(\bar{y} - \epsilon)$  for sufficiently small  $\epsilon > 0$ .<sup>12</sup> I continue to show  $\phi_1 > \phi_2$  for all  $y < \bar{y}$  by showing  $\phi_1' < \phi_2'$  there. Suppose on the contrary  $\tilde{y} := \sup\{y < \bar{y} : \phi_1'(y) = \phi_2'(y)\} > -\infty$ . It is attained by continuity of  $\phi'$ 's. In addition, I have either  $\phi_1' < \phi_2'$ , or  $\phi_1' = \phi_2'$  and  $\phi_1'' > \phi_2''$ . Moreover,  $\phi_1 > \phi_2$  on  $[\tilde{y}, \bar{y})$  because  $\phi_1' < \phi_2'$  there. At  $\tilde{y}$ , I have

$$\phi_1'' = \frac{r_1}{\sigma^2/2}f_1(\phi_1) - \frac{\mu}{\sigma^2/2}(1 + \phi_1') > \frac{r_2}{\sigma^2/2}f_2(\phi_2) - \frac{\mu}{\sigma^2/2}(1 + \phi_2') = \phi_2''.$$

Therefore,  $\phi_1'(\tilde{y} + \epsilon) > \phi_2'(\tilde{y} + \epsilon)$  which is a contradiction.

The proofs for single crossing with respect to  $\mu$  and  $\sigma^2$  are analogous—the only difference lies on the derivation of the ordering between  $\phi''$ 's. For  $\mu$ , I obtain the ordering by noting  $\phi' > -1$  and thus  $1 + \phi' > 0$  for all interior  $y$ . For  $\sigma^2$ , I note  $\phi'' > 0$  and so  $rf(\phi) - \mu(1 + \phi') > 0$ . ■

**Corollary 4** *Let  $\bar{v}_i$  denote the option value at 0 for parameters  $i$ . Then  $\bar{v}_1 < \bar{v}_2$ .*

**Proof.** Suppose otherwise. I have  $\phi_1 > \phi_2$  by Lemma 9. Because  $\phi_2$  touches the 45-degree line, the function  $\phi_1$  will never intersect the line, which is a contradiction. ■

For  $i = 1, 2$ , let  $\phi_i$  denote the solution to ODE (9) for parameters  $i$  from  $y = y_i$  until  $\phi_i' \leq 0$ , subject to initial condition  $\phi_i(y_i) = -y_i$  and  $\phi_i'(y_i) = -1$  where  $y_i < 0$ . The solution exists at least for  $y_i = \underline{y}_i$  by construction in Proposition 5.

**Lemma 10 (Single crossing, starting from  $y_i$ )**

- If  $|y_1| \geq |y_2|$ , then  $\phi_1 > \phi_2$  and  $\phi_1' > \phi_2'$  for  $y \in (y_2, 0]$ .
- If  $|y_1| < |y_2|$ , then there exists at most one  $\bar{y} \in (y_1, 0]$  such that  $\phi_1(\bar{y}) = \phi_2(\bar{y})$ .  
Moreover,  $\phi_1 > \phi_2$  and  $\phi_1' > \phi_2'$  for  $y \in (\bar{y}, 0)$ .

---

<sup>12</sup>If  $\bar{y}$  is the left endpoint of one of the  $\phi$ 's, then the lemma already obtains.

**Proof.** The proof is analogous to Lemma 9. The only difference is the starting point  $y$ 's instead of 0. For  $|y_1| > |y_2|$ , I have  $\phi_1(y_2) > y_2 = \phi_2(y_2)$  and  $\phi_1'(y_2) > -1 = \phi_2'(y_2)$  by the convexity of  $\phi_1$ . The reverse goes for  $|y_1| < |y_2|$ . ■

**Corollary 5** *Let  $\underline{y}_i$  denote the drawdown threshold for parameter  $i$ . Then  $|\underline{y}_1| < |\underline{y}_2|$ .*

**Proof.** Suppose otherwise. I have  $0 = \phi_1'(0) > \phi_2'(0) = 0$ , a contradiction. ■

**Lemma 11** *Let  $v_i$  denote the option value for parameters  $i$ . Then  $v_1 < v_2$  point-wise.*

**Proof.** It suffices to show that  $\phi_1 < \phi_2$  point-wise.

Suppose otherwise. There exists  $y < 0$  such that  $\phi_1(y) = \phi_2(y)$  by the intermediate value theorem. By Lemma 9, I have  $\phi_1 > \phi_2$  and so  $\phi_1$  does not intersect the 45-degree line on  $[\underline{y}_2, 0]$ . Moreover,  $\phi_1'(\underline{y}_2) < \phi_2'(\underline{y}_1) = -1$  and so the convex  $\phi_1$  does not intersect the 45-degree line on  $(-\infty, \underline{y}_2)$ . Therefore,  $\phi_1$  does not intersect with the 45-degree line, which is a contradiction. ■

**Lemma 12** *Let  $s_i$  denote the speed of exploration for parameters  $i$ . Then  $s_1 < s_2$  point-wise.*

**Proof.** By Equation (8), the monotonicity of  $s$  with respect to  $\mu$ ,  $\sigma^2$ , and  $r$  follows directly from that of  $\phi$ . With respect to  $c$ , Lemma 8 implies that  $g^{-1}$  and therefore  $s = g^{-1}(\phi)$  are both decreasing point-wise. ■

## A.5 Proof of Corollary 1

Define  $h := g^{-1}$ . It is increasing because  $g$  is increasing. Moreover, elementary calculus shows that  $h'$  is decreasing whenever  $sc''(s)$  is increasing.

Applying Itô's lemma on Equation (8) gives

$$\begin{aligned} ds &= h'(v - y)(v' - 1)dy + \frac{1}{2} \left( h''(v - y)(v' - 1)^2 + h'(v' - y)v'' \right) d\langle y \rangle \\ &= \left( h'(v - y)(v' - 1)\mu + \frac{\sigma^2}{2} (h''(v - y)(v' - 1)^2 + h'(v - y)v'') \right) sdt \\ &\quad - h'(v(0) - 0)(v'(0) - 1)dM + h'(v - y)(v' - 1)dZ \end{aligned}$$



where the second equality follows from Lemma 2 and the fact that  $\mathbf{1}_{y < 0} dM = 0$  by the Skorokhod equation. The three  $dt$  terms are positive because (i)  $h$  is increasing,  $v' \leq 1$ , and  $\mu < 0$ ; (ii)  $h'$  is decreasing because  $sc''(s)$  is increasing by the hypothesis; (iii)  $v$  is convex. The  $dM$  term vanishes because  $v'(0) = 1$  by the smooth pasting condition (6). Therefore, the corollary obtains because  $Z$  is a martingale by Lemma 2.

## A.6 Proof of Proposition 3

**Lemma 13** *Let  $z$  denote a Brownian motion with drift  $\mu$  and volatility  $\sigma^2$  reflected at  $z = 0$ . Then the pairs of processes  $(U - M, M)$  and  $(z, 2L^z(0))$  share the same law, where  $L^z(0)$  is the local time of  $z$  at 0.*

**Proof.** See [Peskir \(2006\)](#). ■

**Proof of Proposition 3.** Lemma 13 implies that the measure  $\mathbb{P}[y_x \in \cdot, X_\tau > x]$  is absolutely continuous and its density  $p(t, \cdot)$  satisfies the Kolmogorov forward equation

$$\partial_x p = \frac{\sigma^2}{2} \partial_y^2 p - \mu \partial_y p \quad \forall t > 0, y \in (\underline{y}, 0)$$

subject to boundary conditions  $p(x, \underline{y}) = 0$  and  $\partial_y p(x, 0) = (2\mu/\sigma^2)p(x, 0)$  and initial condition  $p(0, y) = \delta_0(y)$ .

This diffusion-drift problem with absorption/reflection boundary conditions can be solved by Fourier series. See Chapter 4 of [Strauss \(2007\)](#). The solution converges asymptotically to the leading eigenfunction, denoted by  $q(y) \geq 0$ , exponentially at the rate of the leading eigenvalue, denoted by  $\lambda > 0$ .

I note that  $\lambda = \pi^2 \sigma^2 / 8 |\underline{y}|^2$  for  $\mu = 0$ . ■

## A.7 Proof of Corollary 2

The proof of Proposition 3 implies that the distribution of  $X_\tau$  decays to zero exponentially at rate  $\lambda$ . Because  $X_\tau = \int_0^\tau s_t dt$ , I have  $X_\tau/s(\underline{y}) \leq \tau \leq X_\tau/s(0)$ . Therefore, the distribution of  $\tau$  also decays to zero exponentially at rate  $\lambda$ .

## A.8 Proof of Proposition 4

Similar to Proposition 3, I index the utility process  $U$  by the alternatives instead of time by undoing the time change  $X_t$ . Because the threshold strategy is invariant to the continuous time change, the utility process is stopped at the alternative  $x = X_\tau$ , where  $y_x = \underline{y}$ .

I prove Proposition 4 by establishing the linear relationship between  $X_\tau$  and  $M_\infty$ , and then translating the relationship from  $X_\tau$  to  $x_\infty$ . I first approximate the tail distribution of  $x_\infty$  (Lemma 14). I continue to derive the expectation of  $x_\infty$  conditional on  $M_\infty$ , using that of  $X_\tau$  conditional on  $M_\infty$  (Lemma 15). I use this indirect method because  $x_\infty$  is not adapted to the natural filtration  $\mathcal{F}^B$  but  $X_\tau$  is. Finally, I translate the conditional expectation to  $M_\infty$  conditional on  $x_\infty$  by the marginal distribution of  $x_\infty$  approximated in Lemma 14 and the Markov inequality .

**Lemma 14** *Suppose  $\mu = 0$ . Then  $x_\infty$  has an exponential tail with exponent  $\lambda$ .*

**Proof.** Proposition 3 implies that  $X_\tau$  has an exponential tail, and so there exist  $K_1, K_2 > 0$  such that  $K_1 e^{-\lambda x} \leq \mathbb{P}[X_\tau > x] \leq K_2 e^{-\lambda x}$ . Because  $x_\infty < X_\tau$ , the tail probability of  $x_\infty$  is bounded from above by  $\mathbb{P}[x_\infty > x] \leq K_2 e^{-\lambda x}$ .

The difference  $z := X_\tau - x_\infty$  is independent of  $x_\infty$  due to the strong Markov property. Recall that Brownian motion near an extrema follows a Bessel-3 process starting at zero. Therefore,  $z$  follows the distribution of hitting time at  $|\underline{y}|/\sigma^2$  of that process. Its density follows an exponential distribution asymptotically.<sup>13</sup> Because the stopping alternative is distributed as the hitting time at  $|\underline{y}|/\sigma^2$  of a Bessel-1 process, i.e., reflected Brownian motion,  $z$  has a thinner exponential tail than  $X_\tau$ . Because  $x_\infty$  and  $z$  are independent and sum to  $X_\tau$ , the exploited alternative  $x_\infty$  must have a tail at least as thick as  $X_\tau$ 's. ■

**Lemma 15** *Suppose  $\mu = 0$ . Then  $x_\infty$  and  $M_\infty$  satisfy*

$$\mathbb{E}[x_\infty | M_\infty > m] = \frac{2|\underline{y}|}{3\sigma^2} m + \frac{2\underline{y}^2}{3\sigma^2} \quad \forall m \geq 0.$$

**Proof.** Because  $y_{x_\infty} = 0$  by definition,  $X_\tau - x_\infty$  is independent of  $(x_\infty, M_\infty)$  due to

<sup>13</sup>See Formula 2.0.2, Chapter 4, Part II of [Borodin and Salminen \(2012\)](#) for the exact distribution.

the strong Markov property of Brownian motion. Therefore, it suffices to compute  $\mathbb{E}[X_\tau - x_\infty]$  and  $\mathbb{E}[X_\tau | M_\infty > m]$ .

First, I compute  $\mathbb{E}[X_\tau - x_\infty]$ . To simplify notation, I will compute it under the normalization  $\sigma^2 = 1$ . For standard Brownian motion starting at  $y \in (z, 0)$ , let  $H$  denote the exit time from  $(z, 0)$ . The expectation of  $e^{-\beta H}$  conditional on exiting at  $z$  is<sup>14</sup>  $\mathbb{E}_y[e^{-\beta H} | B_H = z] = z \sinh(x\sqrt{2\beta}) / x \sinh(z\sqrt{2\beta})$ . Because  $H | B_H = z$  has finite expectation, the dominated convergence theorem applies and the conditional expectation of  $H$  is given by  $\mathbb{E}_y[H | B_H = z] = -\partial_\beta \mathbb{E}_y[e^{-\beta H} | B_H = z] = \frac{1}{3}(z^2 - y^2)$ .

Due to the quadratic variation of Brownian motion, the escape time converges to  $X_\tau - x_\infty$  almost surely as  $y \rightarrow 0$ . By the dominated convergence theorem, the conditional expectation converges as well, i.e.,  $\mathbb{E}[X_\tau - x_\infty] = \lim_{y \rightarrow 0} \mathbb{E}_y[H | B_H = z] = \frac{1}{3}z^2$ .

Second, I compute  $\mathbb{E}[X_\tau | M_\infty > m]$ . For a standard Brownian motion and  $\beta > 0$ , the conditional Laplace transform satisfies

$$\mathbb{E}[e^{-\beta X_\tau} | L_{X_\tau}(0) = l] = \frac{z\sqrt{2\beta}}{\sinh(z\sqrt{2\beta})} \exp\left(-\frac{1}{2z}\left(z\sqrt{2\beta} \coth(z\sqrt{2\beta}) - 1\right)l\right)$$

where  $L_{X_\tau}(0)$  is the local time at zero.<sup>15</sup> Because  $X_\tau$  has finite expectation by Proposition 3, the dominated convergence theorem implies that the conditional expectation of  $X_\tau$  is  $\mathbb{E}[X_\tau | L_{X_\tau}(0) = l] = -\partial_\beta \mathbb{E}[e^{-\beta X_\tau} | L_{X_\tau}(0) = l] \Big|_{\beta \rightarrow 0} = zl/3 + z^2/3$ . Lemma 13 implies that the running maximum is twice the local time at zero, thus  $\mathbb{E}[X_\tau | M_\infty = m] = \mathbb{E}[X_\tau | L_{X_\tau}(0) = 2m] = 2zm/3 + z^2/3$ .

[Lehoczky \(1977\)](#) shows that  $M_\infty$  is exponentially distributed with mean  $|y|$ . Using his result, I obtain the tail conditional expectation of  $X_\tau$  by integrating the conditional expectation

$$\mathbb{E}[X_\tau | M_\infty > m] = \frac{\int_m^\infty \mathbb{E}[X_\tau | M_\infty = \eta] \mathbb{P}[M_\infty \in d\eta]}{\int_m^\infty \mathbb{P}[M_\infty \in d\eta]} = \frac{2}{3}zm + z^2.$$

Therefore, the tail conditional expectation of  $x_\infty$  is  $\mathbb{E}[x_\infty | M_\infty > m] = \mathbb{E}[X_\tau | M_\infty > m]$ —

<sup>14</sup>See Formula 3.0.4(b) and Formula 3.0.5(b), of Chapter 1, Part II in [Borodin and Salminen \(2012\)](#).

<sup>15</sup>See Formula 2.3.4 together with Formula 2.3.2, of Chapter 3, Part II in [Borodin and Salminen \(2012\)](#).

$$\mathbb{E}[X_\tau - x_\infty] = 2zm/3 + 2z^2/3.$$

For the utility process  $U$  with  $\sigma > 0$ , the Brownian scaling property implies that the conditional expectation for volatility  $\sigma^2$  and threshold  $\sigma z$  is

$$\mathbb{E}[x_\infty | M_\infty > \sigma m']|_{\sigma^2, \sigma z} = \mathbb{E}[x_\infty | M_\infty > m']|_{1, z} = \frac{2}{3}zm' + \frac{2}{3}z^2.$$

I obtain the desired equality by taking  $z = |y|/\sigma$  and  $m' = m/\sigma$ . ■

**Proof of Proposition 4.** Note that the coefficient of  $x$  is in fact  $\pi^2\sigma^2/8|y| = |y|\lambda$ . See the proof of Proposition 3.

I first show  $\mathbb{E}[M_\infty | x_\infty > x] \leq |y|\lambda x + o(x)$ . The Markov inequality applied to Lemma 15 yields

$$x\mathbb{P}[x_\infty > x | M_\infty > m] \leq \mathbb{E}[x_\infty | M_\infty > m] = \frac{2|y|}{3\sigma^2}m + \frac{2y^2}{3\sigma^2} \quad \forall x \geq 0.$$

Substituting the tail probability for  $M_\infty$  from [Lehoczky \(1977\)](#), I obtain

$$\mathbb{P}[M_\infty > m | x_\infty > x] \leq \left( \frac{2|y|}{3\sigma^2} \frac{m}{x} + \frac{2y^2}{3\sigma^2} \frac{1}{x} \right) e^{-\frac{m}{|y|}} \mathbb{P}[x_\infty > x]^{-1} =: P(m).$$

The integral over  $[0, \infty)$  gives

$$\int_0^\infty \mathbb{P}[M_\infty > m | x_\infty > x] dm \leq \int_0^\infty \min\{P(m), 1\} dm.$$

As both sides are integrable, the boundary terms from the integration by parts vanish

$$-\int_0^\infty m d\mathbb{P}[M_\infty > m | x_\infty > x] \leq -\int_{m^*}^\infty m dP(m)$$

where  $m^*$  is defined by

$$m^* := \sup\{r : P(r) \geq 1\} = |y|\lambda x + o(x).$$

It is well-defined for sufficiently large  $x$ . The approximation is due to Lemma 14. The

conditional expectation is therefore bounded by

$$\begin{aligned}
\mathbb{E}[M_\infty | x_\infty > x] &= - \int_0^\infty m d\mathbb{P}[M_\infty > m | x_\infty > x] \\
&\leq - \int_{m^*}^\infty m dP(m) \\
&= \mathbb{P}[x_\infty > x]^{-1} \int_{m^*}^\infty m \left( \frac{2}{3} \frac{1}{\sigma^2} \frac{m}{x} + \frac{2}{3} \frac{|y|}{\sigma^2} \frac{1}{x} \right) e^{-\frac{m}{|y|}} dm \\
&= \mathbb{P}[x_\infty > x]^{-1} e^{-\frac{m^*}{|y|}} \left( \frac{2}{3} \frac{|y|}{\sigma^2} \frac{(m^*)^2}{x} + \frac{1}{x} o((m^*)^2) \right) \\
&= P(m^*) (m^* + o(m^*)) \\
&= |y| \lambda x + o(x).
\end{aligned}$$

I continue to show  $\mathbb{E}[M_\infty | x_\infty > x] \geq |y| \lambda x + o(x)$ . For a standard Brownian motion stopped at drawdown  $z > 0$ , the Laplace transform of  $X_\tau$  conditional on  $M_\infty$  satisfies<sup>16</sup>

$$\mathbb{E}[e^{-\beta X_\tau} | M_\infty = m] = \frac{z\sqrt{2\beta}}{\sinh(z\sqrt{2\beta})} \exp\left(-\frac{1}{4z} \left(z\sqrt{2\beta} \coth(z\sqrt{2\beta}) - 1\right) m\right).$$

For the utility process with volatility  $\sigma^2$  stopped at  $|y|$ , I take  $\beta = x^{-2}$  to obtain

$$\mathbb{E}\left[e^{-X_\tau/x^2} | M_\infty = m\right] = 1 - \frac{1}{6} \frac{|y|}{\sigma} \frac{m}{x^2} + o\left(\frac{1}{x} + \frac{m}{x^2}\right).$$

The independence between  $(x_\infty, M_\infty)$  and  $X_\tau - x_\infty$  implies

$$\mathbb{E}[e^{-x_\infty/x^2} | M_\infty] = \mathbb{E}[e^{-X_\tau/x^2} | M_\infty] / \mathbb{E}\left[e^{-(X_\tau - x_\infty)/x^2}\right] = 1 - \frac{1}{6} |y| \frac{m}{x^2} + o\left(\frac{1}{x} + \frac{m}{x^2}\right).$$

The Markov inequality applied to  $e^{-x_\infty/x^2}$  conditional on  $M_\infty$  reads  $k\mathbb{P}\left[e^{-x_\infty/x^2} > k | M_\infty\right] \leq$

---

<sup>16</sup>See Formula 2.3.4 and Formula 2.3.2, of Chapter 3, Part II in [Borodin and Salminen \(2012\)](#).

$\mathbb{E} \left[ e^{-x_\infty/x^2} | M_\infty \right]$ . The tail probability is then bounded by

$$\mathbb{P} [x_\infty > x | M_\infty \in dm] \geq 1 - e^{1/x} \mathbb{E} \left[ e^{-x_\infty/x^2} | M_\infty \in dm \right] = \frac{1}{6} |\underline{y}| \frac{m}{x^2} + o \left( \frac{1}{x} + \frac{m}{x^2} \right) dm.$$

Given the distribution of  $M_\infty$  in [Lehoczky \(1977\)](#), the density of  $M_\infty$  conditional on  $x_\infty > x$  is bounded by

$$\mathbb{P} [M_\infty \in dm | x_\infty > x] \geq \mathbb{P} [x_\infty > x]^{-1} \frac{1}{|\underline{y}|} e^{-\frac{m}{|\underline{y}|}} \left( 1 - e^{1/x} \mathbb{E} \left[ e^{-x_\infty/x^2} | M_\infty \in dm \right] \right) dm =: q(m) dm.$$

Define  $Q(m) := \int_m^\infty q(m) dm$  and  $\tilde{m} := Q^{-1}(1) = |\underline{y}| \lambda x + o(x)$ . The integration by parts yields

$$\mathbb{E} [M_\infty | x_\infty > x] \geq \int_{\tilde{m}}^\infty m q(m) dm = Q(\tilde{m}) (\tilde{m} + o(\tilde{m})) = |\underline{y}| \lambda x + o(x).$$

■