LEVELING UP WITHOUT BREAKING DOWN: STRATEGIES FOR SUCCESS UNDER STRESS*

JIN LI   TAIQI LI   ZHENGLI WANG

JUNE 2, 2024

This paper explores strategies for achieving goals under stress by revisiting Herbert Simon's Berlitz model. Higher effort results in faster progress but also increases the stress level, causing a higher breakdown rate. We characterize the optimal effort trajectory by using an exchangeability technique that compares different sequences of effort choices. The comparison reflects a trade-off between speed (for faster recovery) and risk (of breakdown). One consequence of this trade-off is that the individual should engage in preemptive stress management. Optimal strategies, therefore, feature planned strategic retreats, such as cheat days in weight loss programs and rest days in marathon training. *JEL* Codes: C61, D15, J24

KEYWORDS: Skill Development, Goal Attainment, Stress Management, Planned Strategic Retreat, Stochastic Control.

I. INTRODUCTION

This paper explores strategies for achieving goals under stress. We emphasize the dual effect of effort: while effort acts as a driving force toward the goal, it can also induce negative effects such as stress, causing the individual to break down and quit. We characterize the optimal strategy by building on Herbert Simon's Berlitz model (Simon (1954)). Our analysis uses an *exchangeability* technique that compares different sequences of effort choices, i.e., rest-then-sprint vs sprint-then-rest. The comparison of effort sequences reflects a trade-off between speed (for faster recovery) and risk (of breakdown). The resulting optimal strategies feature planned strategic retreats that manage stress pre-emptively.

The Berlitz Model receives its name because Simon uses the example of an individual who learns French using the Berlitz method. As the individual practices French, it gradually becomes easier for him. But if the individual practices too much, it becomes unpleasant, causing him to reduce practice in the future. Successful learning then hinges on the balance between these two factors: the ease gained through practice and the potential aversion caused by over-practicing. The prevailing factor will determine the individual's success in reaching the goal.

The two forces in the Berlitz model are present in numerous settings, from mastering a musical instrument to reaching a weight-loss goal, building the skill for running a race, and so on. Albert Hirschman (1958), in "The Strategy of Economic Development," found the model helpful in describing certain aspects of economic development: *Somewhat like a person who decides in a fit of enthusiasm to learn a foreign language, a country that sets out on the road to development often does not realize the difficulties of the task ahead. As these difficulties appear, as it becomes clear that the price of development is a high one in terms of human suffering, social tensions, forced abandonment of traditional behavior and values, etc., "practice" may be reduced, contradictory and harmful economic policies are being adopted, and development*

*Jin Li: Faculty of Business and Economics, The University of Hong Kong. E-mail: jli1@hku.hk. Taiqi Li: Faculty of Business and Economics, The University of Hong Kong. E-mail: litaiqi@hku.hk. Zhengli Wang: Faculty of Business and Economics, The University of Hong Kong. E-mail: wzl1@hku.hk. We thank Yeon-Koo Che, Florian Englmaier, Shuaizhang Feng, Mira Frick, Guido Friebel, Jie Gong, Carl Heese, Keri Hu, John Klopfer, Michael Kosfeld, Ruby Lu, Masaki Miyashita, Wing Suen, Balas Szentes, Can Urgun, Brian Viard, Yufu Wong, Yanhui Wu, Chen Zhao, and seminar participants at Fudan University, Goethe University of Frankfurt, Jinan University, University of Hong Kong, and University of Utah for helpful comments.

will be slowed down and halted. In general, many initiatives within organizations experience similar patterns: the resources and work invested in the initiative, while helpful in pushing it forward, also fuel resistance that can stall the initiative midway.

Simon's Berlitz model, however, assumes that the individual's behavior adjusts automatically, meaning that the amount of effort exerted tomorrow is completely determined by today's effort and skill level. This assumption implies that once the initial effort of the individual is known, the entire effort trajectory is determined. Consequently, it is known from the start whether the goal can be attained or not. Simon's model, as such, informs strategies for achieving goals only through how much initial effort is required.

In this paper, we incorporate effort choices into the Berlitz model so that the forward-looking individual can voluntarily change his effort levels over time to manage future challenges. To do so, we introduce an underlying state variable, which we denote as the *stress level* for convenience. The stress level serves as a sufficient statistic that captures the the entire past choices of effort.

We assume that the stress level determines the likelihood of the individual experiencing a breakdown at each point in time. The stress level goes up if the individual's effort exceeds a threshold, and it decreases otherwise. In addition, this threshold is increasing in the individual's existing skill. These assumptions maintain the spirit of Simon's original model.

We solve the individual's optimal strategies and show that they can be put in two distinct categories, depending on the level of the goal. The first category–the *modest-goal* case–occurs when the gap between the starting skill and the targeted skill is small. In this scenario, the optimal strategies are simple. The individual either *surge*, meaning that he always puts in the maximal effort. Or, if his initial stress level is high, the individual puts in no effort until the stress level drops below a level, and then he surges.

The second category, which we label as the *ambitious-goal* case, occurs when the difference between the starting skill and the targeted skill is large. The ambitious-goal case poses a greater challenge for the individual, leading to more varied strategies. Depending on the initial stress level of the individual, there are three types of optimal strategies. First, when the stress level is low, the optimal strategy is again to surge.

Second, when the stress level is in the intermediate range, the optimal strategy takes the form of *HIIT*[1], a term we use because a key feature of the optimal strategy bears resemblance to the popular workout exercise. The HIIT strategy has three phases. In the first phase, the individual puts in maximal effort, leading to rapid skill improvement but also a rapid increase in stress levels. When stress gets too high, the second phase starts. Here, the individual drops his effort to a low (but positive) level at the beginning of the phase and then increases it. Yet the effort level remains relatively low, and the stress level drops steadily. The second phase ends when the individual's skill is sufficiently high so the final goal is close in sight. The strategy then transitions into the third phase, where the individual surges again.

The second phase of HIIT is a planned strategic retreat, which is akin to cheat days in weight management programs and rest days in marathon training regimes. Planned strategic retreat underscores the importance of pre-emptive stress management, contrasting the myopic idea that stress is managed only when it reaches a sufficiently high level. The HIIT strategy implies that when the goal is ambitious, rather than waiting for the stress level to escalate, the individual should lower the stress level at an earlier stage. The advantage of doing so is that when the goal is eventually close in sight, the individual's stress level is sufficiently low so that he can surge.

[1] HIIT stands for high-intensity interval training. It features low-intensity movements between high-intensity movements.

Finally, when the individual's initial stress level is high, the optimal strategy takes the form of *phased escalation*. This strategy again has three phases. Initially, the individual puts in no effort, enabling the stress level to drop swiftly. Once the stress level drops below a skill-dependent level, the strategy transitions into the second phase, which can be viewed as a warm-up phase. Here, the effort gradually increases, and the stress continues to decrease. Once the skill level reaches a high enough level, the third phase starts, and the effort jumps to the maximal level.

We solve for the optimal strategies by using an exchangeability argument that compares the payoff between sprint (maximal effort)-then-rest (zero effort) versus rest-then-sprint. This comparison reveals a trade-off between speed (for faster recovery) and risk (of breakdown). Sprint-then-rest has the cost of a higher risk of breakdown. However, it has a novel dynamic benefit of faster recovery (stress reduction) associated with a higher skill level.

Importantly, this benefit is higher when the individual's skill level is lower, implying that the individual is better off sprinting when his skill is lower. Because the individual is also better off sprinting when the goal is near in sight (when he has higher skills), two types of non-monotonicities arise. First, the individual's effort level can be non-monotone over time, as the HIIT strategy indicates. Second, for a fixed initial stress level, the optimal initial effort of the individual can be non-monotone at the goal level. The individual's initial action is to start small when the goal is at the intermediate level and to do a big push when both the goal levels are low or high.

Our model is related to several dynamic models that feature the negative effect of effort on future efforts. A closely related one is Urgun (2021), who assumes that the marginal cost of effort increases when he works and drops when he is idle. Different from our paper, Urgun (2021) is interested in how to allocate tasks across different agents.

Our model is also related to papers that model discrete jumps (in continuation payoffs) with Poisson process; see, for example, Keller et al. (2005), Rosenberg et al. (2007), Keller and Rady (2010), Strulovici (2010), Bonatti and Hörner (2011), Klein and Rady (2011), Murto and Välimäki (2011), Board and Meyer-ter Vehn (2013), Guo (2016), Halac et al. (2017), Che and Mierendorff (2019), Zhong (2022), Che et al. (2023) and Liu and Wong (2023). In these models, Poisson arrivals reflect discrete jumps in beliefs. In contrast, our Poisson process describes the likelihood that the individual quits.

Finally, we characterize the optimal strategies by using a technique that relies on exchanging the sequence of actions. This exchangability technique is instrumental in solving HJB equations that involve more than one state variable. The exchangeability technique has been used in Li et al. (2023) to characterize optimal experimentation strategy with multiple bandits. Unlike Li et al. (2023), the application of exchangeability argument in this paper requires adjusting the length of the time intervals.

The rest of the paper is organized as follows. We setup the model in Section 2 and carry out the main analysis in Section 3. Section 4 concludes.

II. MODEL SETUP

In this section, we first review Simon's original Berlitz model and then describe our setup.

II.A. Benchmark: Simon's Berlitz Model

Simon (1954) considers the example of an individual learning French through the Berlitz method. Let X_t be the amount of time (effort) the individual practices in period t. Simon's model has three key assumptions. First, as the individual practices it, his skill improves, and the corresponding difficulty level of the language decreases. In particular, denote the difficulty level at period t to be D_t. Simon assumes that it decreases logarithmically in effort with learning rate $a > 0$:

$$dD_t/dt = -aD_tX_t. \tag{1}$$

Second, Simon assumes that "at any level of difficulty, practice is pleasurable up to a certain point, and unpleasant beyond that point." He calls this point the individual's *satiation level* $\bar{X}(D_t)$ and assumes that the individual practices more when he finds it pleasurable (effort is below the satiation level) and that he practices less otherwise. Specifically, the individual's effort adjusts according to

$$dX_t/dt = -b(X_t - \bar{X}(D_t)). \tag{2}$$

Finally, Simon assumes that the individual's satiation level $\bar{X}(D_t)$ decreases with the problem's difficulty. In other words, the higher the individual's skill level is, the higher his satiation level is. Simon does not give an explicit expression for $\bar{X}(D_t)$ but draws it as a linear function in Figure 1 in his paper (on page 403 of Simon 1954).

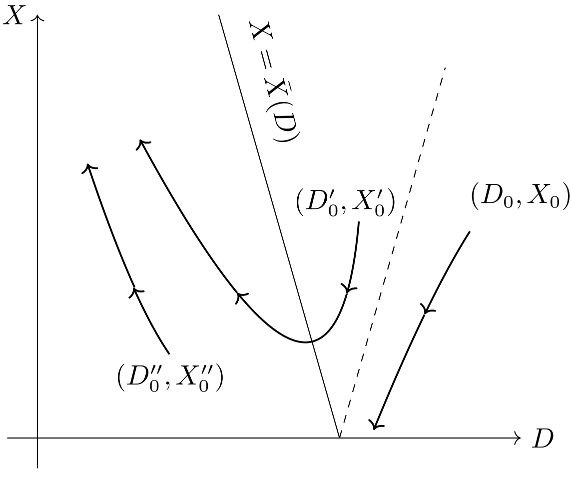


FIGURE I. Simon's Original Berlitz Model

We reproduce here Figure I of Simon (1954), which captures the key insights of the original Berlitz model. The solid decreasing line $X = \bar{X}(D)$ represents the satiation level for any given difficulty level D. On the left side of the satiation line, practice is pleasurable and will be increased in the following period; on the right side, however, practice is unpleasant and will be reduced. Simon draws three examples of effort trajectories to illustrate how the success (of skill development) depends on the battle between the ease gained through practice and the aversion caused by over-practicing. First, if the individual starts at (D_0, X_0), he over-practices all the time. His effort drops down to zero before making enough progress, failing to acquire the skill. Second, if the individual starts at (D_0'', X_0''), practice is always pleasurable. His effort increases, and he succeeds. Third, if the individual starts at (D_0', X_0'): practice is initially unpleasant for him, and he reduces his effort. However, his practice effort leads to a decrease in the difficulty level, causing the satiation level to increase. At some point, his practice level falls below the satiation level. He then practices more and more and succeeds eventually. This last example showcases how the success of skill development depends on the battle between the cost of over-practicing and the dynamic benefit (in raising the satiation level).

Despite its insights, Simon's model does not allow the individual to choose his actions so that the effort dynamic evolves automatically. As a result, once the initial condition is determined, so is the eventual success or failure. The dashed line in Figure I is the boundary for this: if the initial position is on the left side of the dashed line, then the process is guaranteed to succeed; otherwise, it is guaranteed to fail. Simon's model, therefore, only provides insight into how the initial condition affects success and not how the individual should choose the effort over time.

II.B. Berlitz Model with Effort Choice

We now incorporate effort choice into the Berlitz model. In every period t, the individual can choose to spend a certain amount of effort $X_t \in [0, x_U]$. Let D_t be the level of difficulty at period t. In accordance with Simon's framework, we assume that the difficulty decreases logarithmically in effort:

$$dD_t/dt = -aD_tX_t. \tag{3}$$

Once the difficulty level diminishes to 1, the individual succeeds and receives a reward $\mathcal{R}$.

Next, to capture the idea that an increased effort level can create effort aversion in the future, we assume that effort can affect the individual's stress level, affecting the probability of quitting. Let Y_t denote the individual's stress level in period t. The occurrence of a "quit" event follows a non-homogeneous Poisson process whose arrival rate is a non-negative, increasing, and convex function of Y_t. For simplicity, we assume the rate is given by e^{Y_t}. When a "quit" event occurs, the individual fails and receives no rewards.

To model how the stress level Y_t changes, we assume that, as in Simon's model, the individual has a satiation level denoted as $\bar{X}(D_t)$. When the individual's effort exceeds the satiation level, his stress level increases. Otherwise, his stress level decreases. Specifically,

$$dY_t/dt = b[X_t - \bar{X}(D_t)]. \tag{4}$$

In other words, instead of affecting tomorrow's effort level directly, effort affects the underlying stress level, and $b > 0$ is the rate at which the stress level changes. We follow Simon's illustration and assume that $\bar{X}(D_t) = k_1 - k_2D_t$. This captures the idea that the individual's satiation level is higher when his skill level is higher (the difficulty level is lower).

Note that our modeling choice of the stress level allows the breakdown probability in each moment to depend on the entire past effort choices. In addition, it captures the idea that the marginal negative effect of effort is lower if the individual's skill level is higher. An alternative modeling choice is to allow the marginal cost of effort to depend on the past effort choices (with more weights on recent efforts). But then, capturing the additional dependence of marginal cost of effort on the skill level becomes cumbersome.

To describe the individual's maximization problem formally, denote $\mathcal{N}_t$ as the number of "quit" event occurrences during $[0, t]$, and denote the terminal time T as

$$T = \inf\{t \mid D_t \leq 1 \text{ or } \mathcal{N}_t \geq 1\}. \tag{5}$$

The individual optimizes the following objective

$$\max_{X_t \in [0, x_U]} \mathbb{E}\left[e^{-\lambda T}\mathcal{R}\mathbf{1}_{\{D_T = 1\}}\right], \tag{6}$$

subject to equations (3) and (4). Here, $\lambda > 0$ represents the discount factor, and $\mathbf{1}_{\{\cdot\}}$ denotes the indicator function.

We impose the following parameter restrictions to make the model interesting. First, we assume $k_1 - k_2 D_t > 0$ so that the individual's stress level always decreases when no effort is exerted. Second, we assume $k_1 - k_2 < x_U$ to ensure that the stress level always increases when maximal effort is exerted. Third, we assume the parameter b exceeds a certain threshold, so that the stress level Y_t adjusts meaningfully: otherwise, the individual always exerts maximal effort in every period; see Appendix for details.

We end the section with two remarks on our model. First, similar to Simon's model, our model is intentionally simple and stylized. In particular, we assume that the individual's sole objective is to develop the skill (quickly). We do not incorporate considerations of effort costs or the utility loss associated with stress levels. While these are obviously relevant features of many applications, adding them complicates the analysis significantly. The simplicity of our model gives the advantage that the optimal strategy can be solved analytically.

Second, unlike Simon's model, where the individual fails by gradually lowering his effort, our model features sudden failure through a Poisson process of quitting. Our model, therefore, is more relevant to examples where failures take the form of breakdown.

III. ANALYSIS

In this section, we characterize the optimal strategy for skill development. We do so by using a dynamic programming approach. In Section 3.1, we reformulate the problem as a dynamic programming problem. In Section 3.2, we describe our approach to solve the problem. Section 3.3 describes the optimal strategy.

III.A. Preliminary Analysis and HJB

The characterization of the optimal strategy requires a description of the entire effort sequence over time. As in many dynamic optimization problems, we can solve this problem with dynamic programming. Notice that in every period t, the current difficulty and stress level summarizes all information that is relevant for the individual's effort choice X_t. Therefore, the difficulty level D_t and the stress level Y_t are the state variables, and we can define the following value function[2]

$$V(D_t, Y_t) \triangleq \max_{X_t \in [0, x_U]} \mathbb{E}\left[e^{-\lambda T} \mathcal{R} \mathbf{1}_{\{D_T = 1\}}\right]. \tag{7}$$

We now derive, heuristically, the Hamilton-Jacobi-Bellman (HJB) equation associated with the value function. The HJB equation is a continuous-time analog of the Bellman equation in discrete-time models. Consider a small time interval $[t, t + \Delta t]$. Starting at any difficulty level D_t and stress level Y_t, the process may terminate (a "quit" event occurs) with probability $e^{Y_t}\Delta t + o(\Delta t)$ or proceed smoothly (no "quit" event occurs) with probability $1 - e^{Y_t}\Delta t + o(\Delta t)$. Hence,

$$\begin{aligned} V(D_t, Y_t) &= \max_{X_t \in [0, x_U]} e^{-\lambda \Delta t} \left[\left(1 - e^{Y_t}\Delta t + o(\Delta t)\right) \cdot V(D_{t+\Delta t}, Y_{t+\Delta t}) \right. \\ &\qquad\qquad \left. + \left(e^{Y_t}\Delta t + o(\Delta t)\right) \cdot 0\right] + o(\Delta t) \\ &= \max_{X_t \in [0, x_U]} \left[1 - (\lambda + e^{Y_t})\Delta t\right] \cdot V(D_{t+\Delta t}, Y_{t+\Delta t}) + o(\Delta t). \end{aligned} \tag{8}$$

[2]In the appendix, we show that the value function is well defined because the action set is compact.

Applying Taylor expansion to $V(D_{t+\Delta t}, Y_{t+\Delta t})$ and sending $\Delta t \to 0$, we arrive at

$$0 = \max_{x\in[0,x_U]} \left\{ -(e^y+\lambda)V(D,y) - aDx\frac{\partial V(D,y)}{\partial D} + b[x-(k_1-k_2D)]\frac{\partial V(D,y)}{\partial y} \right\}, \quad (9)$$

where we substitute $D_t = D$, $Y_t = y$, $X_t = x$ and use the expressions of how the difficulty and the stress level change with effort in (3) and (4).

The right-hand side of equation (9) contains three terms. The first term, $-(e^y+\lambda)V(D,y)$, represents the usual discounting effect. Different from most models, there is an additional e^y. This reflects the fact that the effective discount rate takes into account that the process breaks down with rate e^y. The second term, $-aDx\frac{\partial V(D,y)}{\partial D}$, involves the changes in the difficulty level of the problem: $-aDx$ reflects how the rate of the difficulty level changes, and $\partial V/\partial D$ is the marginal value (of the difficulty level) to the individual. The third term, $b[x-(k_1-k_2D)]\frac{\partial V(D,y)}{\partial y}$, relates the changes in the stress level: $b[x-(k_1-k_2D)]$ reflects how the rate of the stress level changes, and $\partial V/\partial y$ is the marginal value of the stress level.

Equation (9) shows that the effort level x affects the right-hand side of HJB in two ways: on the one hand, an increase in x benefits the individual by reducing the difficulty level; on the other hand, an increase in x hurts him by reducing the rate the stress level decreases. The optimal choice of x, therefore, depends on the relative magnitude of these two effects.

III.B. The Exchangeability Curve and Speed-Risk Trade-off

The HJB is a partial differential equation with a maximization operator. The typical way to solve the equation is to first conjecture a solution and then verify it. Notice that the right-hand side of the HJB (equation (9)) is linear in x. A natural guess, then, is that the optimal solution is bang-bang. However, it turns out that this is not the case, which complicates the analysis.

To deal with the difficulty, we take a step back and take advantage of the sequential nature of the problem, that is, the individual chooses his effort one at a time. Our key solution technique relies on the following exchangeability argument. If under the optimal strategy, the individual, in two consecutive (short) time intervals, chooses to rest (no effort) first and sprint (maximal effort) next, his payoff should be weakly higher than an alternative effort trajectory where he first sprint and then rest (with the rest of the effort trajectory unchanged). In other words, the optimal strategy cannot be improved upon with the type of perturbation where one exchanges the sequence of effort choices.

Before applying the exchangeability argument, we simplify the analysis by isolating a region in the space of state variables where the individual chooses the maximal effort all the time. Specifically, we first identify points (D,y) where a) the optimal strategy specifies maximal effort at (D,y), and b) the optimal strategy specifies maximal effort along the trajectory following (D,y). We refer to these points as *surge points*, which describes the management practice that individuals try their utmost until the task is done.

LEMMA 1: *(Surge Region) There exists a decreasing function $g_1(\cdot)$ ("surge curve") such that (D,y) is a surge point if and only if $(D,y) \in \mathcal{M}_1 \triangleq \{(D,y); y \le g_1(D)\}$.*

Lemma 1 shows that it is optimal for the individual to surge (keep making the highest effort level x_U) when the initial stress level is low and the difficulty level is low. This result follows from the fact that the individual's quitting probability is increasing and convex in the stress level. When the initial stress level is low, the cost of working—the increase in the quitting rate—is low, so the individual is better off putting in effort. Similarly, when the difficulty level

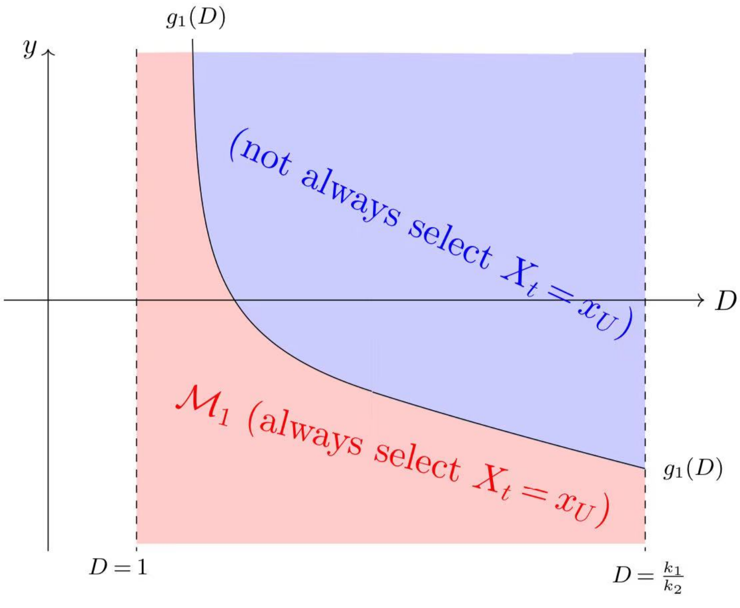


FIGURE II. Below $g_1(D)$: surge (always select $X_t = x_U$) is optimal. Above $g_1(D)$: surge is not optimal.

is low, the time it takes to reach the goal is short when the individual keeps choosing the highest effort. This implies that the cumulative increase in the stress level (and therefore the quitting rate) will be small, again implying that it is optimal for the individual to surge.

Notice that, starting with any (D, y) in this region (which we call $\mathcal{M}_1$), any path traversed by the individual will continue in this region. That is, for all initial levels $(D, y) \in \mathcal{M}_1$, if we always select $X_t = x_U$, the trajectory of the whole process stays within $\mathcal{M}_1$. Otherwise, the initial point (D, y) fails to be a surge point.

Lemma 1 shows that for $(D, y) \in \mathcal{M}_1$, the individual's optimal strategy is to always put in the highest effort. For points outside $\mathcal{M}_1$, the optimal strategy is more complicated, and we solve them by using an exchangeability argument.

Specifically, suppose the initial state variable of the individual is represented by point A in Figure III. Also assume that the individual will arrive at point B at some point. The exchangeability argument consider two ways for the individual to reach point B. The first one is represented by the blue curve, where the individual first chooses $X_t = x_U$ until the difficulty decreases by ΔD, then chooses $X_t = 0$ until the stress level decreases by Δy. We denote this curve as SR because the individual sprints first and then rests. The second one is represented by the red curve, where the individual first selects $X_t = 0$ until the stress decreases by Δy, then chooses $X_t = x_U$ until the difficulty decreases by ΔD. We denote the red curve as RS: rests first and then sprints.

We denote the individual's payoffs along the two curves as P_{SR} and P_{RS}, and the exchangeability argument compares these two payoffs. The comparison reflects a trade-off between risk (of breakdown) and speed (for recovery). For the individual, the disadvantage of choosing SR is that his stress level is higher than that of RS. This means that the cost of SR (relative to RS) is a higher risk of breaking down.

The advantage of choosing SR is that it takes less time to lower the stress level than the path RS. In other words, one benefit of sprinting first is that, by building up his skill first, the individual saves the future recovery time (of stress). In other words, the speed of recovery is faster for individuals with higher skills. As an illustration for this point, consider the following example. An individual tries to get fit by starting a training exercise. In the first several sessions of the exercise, the individual's muscles may get exceedingly sore, and it takes a long time for him to recover. As the individual builds up his skill gradually, he still gets tired after the exercise, but it takes less time for him to recover. Therefore, building up the skill has the benefit of a faster recovery.

The risk-speed trade-off determines when the individual should choose SR and when he should choose RS. The next lemma shows how the comparison depends on the difficulty and stress levels of the individual.

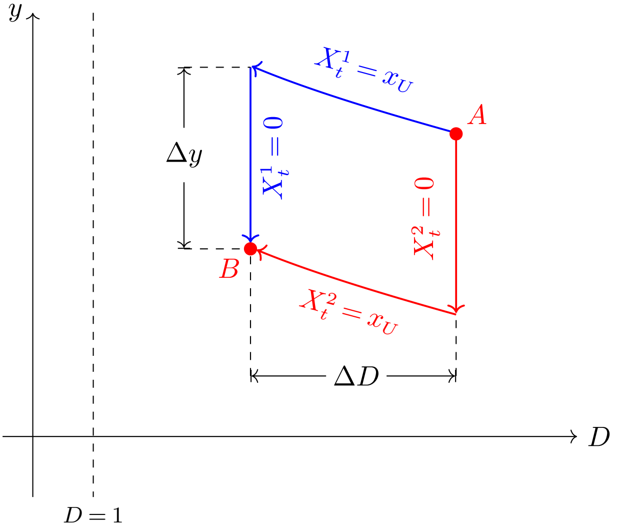


FIGURE III. The blue curve represents the trajectory when applying policy X_t^1, while the red curve represents the trajectory when applying X_t^2.

LEMMA 2: *(Exchangeability curve) When $\Delta D \to 0^+$ and $\Delta y \to 0^+$, there exists an increasing function $g_2(D)$ such that $P_{SR} = P_{RS}$ for the set of points $\{(D,y); y = g_2(D)\}$; $P_{SR} > P_{RS}$ for the set of points $\{(D,y); y < g_2(D)\}$; $P_{SR} < P_{RS}$ for the set of points $\{(D,y); y > g_2(D)\}$.*

Lemma 2 shows that the risk-speed trade-off favors sprinting (first) when the stress level is low and when the difficulty is high (so that the individual's skill level is low). To see why this is the case, note that when the stress level is low, the cost of sprinting first is small, favoring SR over RS. In addition, when the difficulty is high (so that the skill level is low), the gain from faster recovery is larger. The bigger gain arises because the recovery time is longer for lower-skilled individuals. Using the example earlier: when the individual starts out the exercise, it takes him a long time to recover, and the reduction in recovery time from becoming more fit is significant, but when he is sufficiently fit, it takes him less time to recover, and the reduction in recovery time is smaller. In other words, the gain from reducing the recovery time is bigger when the difficulty level is high (skill level is low), favoring SR over RS.

In Lemma 2, we denote $g_2(D)$ as the *exchangeability curve*, which represents the function that at the set $\{(D,y); y = g_2(D)\}$, path SR and RS are equally preferable. We now use the exchangeability curve, together with the surge curve ($g_1(D)$ derived in Lemma 1), to partition the state space and derive the associated optimal actions.

III.C. Optimal Strategies of Skill Development

We now characterize the optimal strategies for skill development. Recall that $\mathcal{M}_1$ is the region below the surge curve. Now define $\mathcal{M}_2$ as the region that is above both the surge and exchangeability curve, and $\mathcal{M}_3$ as the region in between:

$$\mathcal{M}_2 \triangleq \{(D,y); y > \max\{g_1(D), g_2(D)\}\} \tag{10}$$

$$\mathcal{M}_3 \triangleq \{(D,y); g_1(D) < y < g_2(D)\} \tag{11}$$

LEMMA 3: *It is optimal for the individual to select $X_t = 0$ whenever $(D,y) \in \mathcal{M}_2$, and to select $X_t = x_U$ whenever $(D,y) \in \mathcal{M}_3$.*

Lemma 3 states that the individual always rests in $\mathcal{M}_2$ and sprints in $\mathcal{M}_3$. To give a heuristic argument for why the individual chooses to rest in $\mathcal{M}_2$, suppose to the contrary that there exists a point $(D^*, y^*) \in \mathcal{M}_2$ such that the individual sprints. By the definition of $\mathcal{M}_2$, SR is better than RS. This means that the individual cannot rest in the next moment because he would be better off switching the order and resting today. But if the individual sprints in the next moment, he will remain above the exchangeability curve, implying that he has to continue to sprint the moment after. In other words, once the individual starts to sprint in $\mathcal{M}_2$, he will sprint all the way. But this would imply that (D^*, y^*) is in region $\mathcal{M}_1$, which is a contradiction.

A similar argument shows why the individual must sprint in $\mathcal{M}_3$. Again suppose to the contrary that the individual finds it optimal to rest in a point $(D^*, y^*) \in \mathcal{M}_3$. By the definition of $\mathcal{M}_3$, RS is better than SR. This means that the individual cannot sprint in the next moment. But if the individual rests in the next moment, he will remain below the exchangeability curve, implying that he has to continue to rest the moment after. But the individual cannot rest forever because then he will never reach the goal. This is a contradiction.

Lemma 1 and Lemma 3 describe the optimal actions for all points in the state space other than those on the exchangeability curve. To state the optimal strategies, note that because the surge curve is downward sloping and the exchangeability curve is upward sloping, let D_c be the unique difficulty level where the two curves intersect. This difficulty level divides the goals into two categories. We call a goal "modest" if the initial level of difficulty $D \leq D_c$ and "ambitious" if its initial level of difficulty $D > D_c$. We also partition $\mathcal{M}_1$ and $\mathcal{M}_2$ depending on whether the difficulty level exceeds $D \leq D_c$. We define $\mathcal{M}_{1m}$, $\mathcal{M}_{1a}$, $\mathcal{M}_{2m}$, and $\mathcal{M}_{2a}$ accordingly, where the subscript "*m*" stands for modest, and "*a*" stands for ambitious. Figure IV provides an illustration of the regions.

THEOREM 1: *(Characterization of Optimal Policy). Given the initial states (D,y), the optimal policy has the following structures.*

I (Modest Goal). Suppose $D \leq D_c$:

*(i) **[Surge]** If $(D,y) \in \mathcal{M}_{1m}$, the optimal policy is to surge (always select $X_t = x_U$).*

*(ii) **[Rest then Surge]** If $(D,y) \in \mathcal{M}_{2m}$, the optimal policy has two phases*

$$X_t = \begin{cases} 0, & t \in [0, \frac{y - g(D)}{b(k_1 - k_2 D)}) \\ x_U, & \textit{otherwise} \end{cases} \tag{12}$$

II (Ambitious Goal). Suppose $D > D_c$:

*(i) **[Surge]** If $(D,y) \in \mathcal{M}_{1a}$, the optimal policy is to surge (always select $X_t = x_U$).*

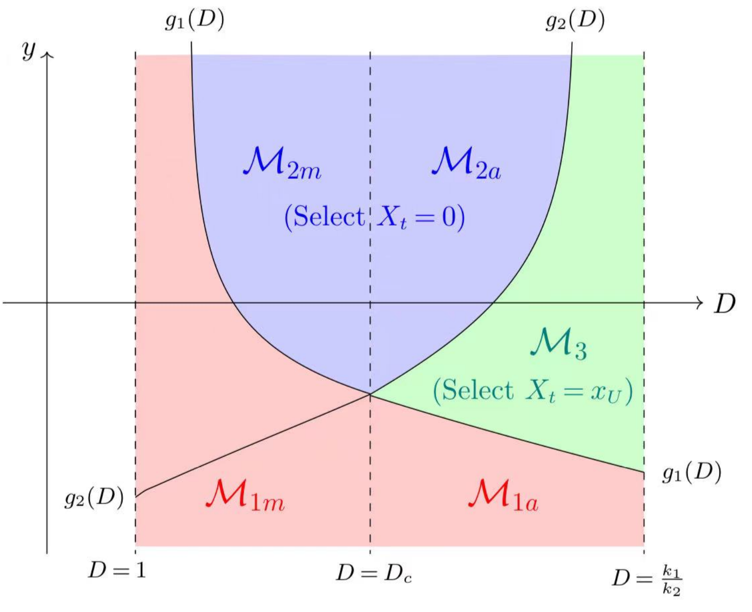


FIGURE IV. $g_1(D)$ and $g_2(D)$ intersects when $D = D_c$. It is optimal to select $X_t = 0$ in $\mathcal{M}_2$ (i.e. $\mathcal{M}_{2m} \bigcup \mathcal{M}_{2a}$) and to select $X_t = x_U$ in $\mathcal{M}_3$.

(ii) ***[Phased Escalation]*** *If $(D, y) \in \mathcal{M}_{2a}$, there exists $t_1, t_2 \geq 0$ such that the optimal policy has three phases*

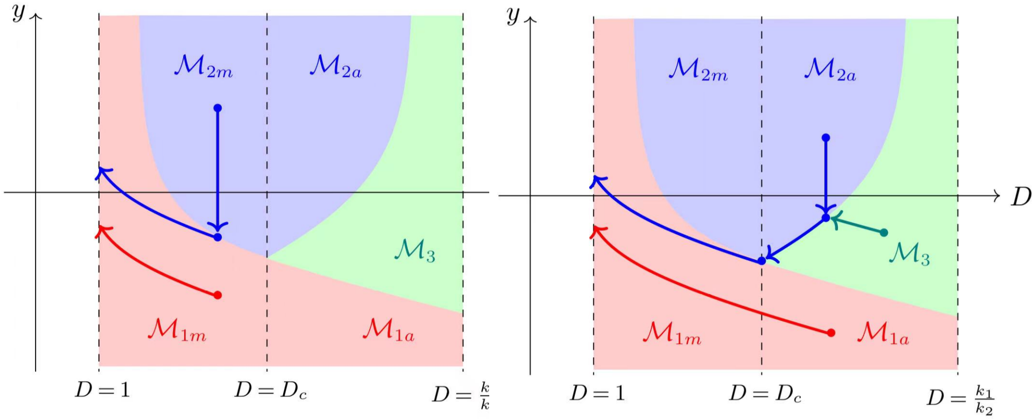
$$X_t = \begin{cases} 0 & , t \in [0, t_1) \\ \frac{bk_1}{b+1} - k_2 D_t & , t \in [t_1, t_1 + t_2) \\ x_U & , \textit{otherwise} \end{cases} \tag{13}$$

(iii) ***[HIIT]*** *If $(D, y) \in \mathcal{M}_3$, there exists $t_1', t_2' \geq 0$ such that the optimal policy has three phases*

$$X_t = \begin{cases} x_U & , t \in [0, t_1') \\ \frac{bk_1}{b+1} - k_2 D_t & , t \in [t_1', t_1' + t_2') \\ x_U & , \textit{otherwise} \end{cases} \tag{14}$$

Figure V illustrates the two different cases of Theorem 1. The left panel (Figure V(a)) describes the case when the goal is modest. In this case, the optimal policy can be of two forms, depending on whether the initial stress level y is low or high relative to the surge curve $g_1(D)$. When the initial stress level y is relatively low, the optimal policy, illustrated by the red arrow, is to surge by consistently selecting the maximal effort $x = x_U$. Doing so allows the individual to reach the goal as fast as possible without increasing the stress level unduly.

When the initial stress level y is relatively high, surge is no longer the optimal policy because doing so will result in an excessively high stress level. The optimal policy, illustrated by the blue arrow, is to rest first until the stress level falls below the surge curve. Once the individual is sufficiently rested, he then sprints.



(a) Modest Goals.

(b) Ambitious Goals.

FIGURE V. Trajectory of optimal policy with initial states (D, y) in different regions.

The right panel (Figure V(b)) describes the case when the goal is ambitious. In this case, the optimal policy can be of three forms. The first one, similar to the modest-goal case, is to surge. This happens, again, when the initial stress level is below the surge curve $g_1(D)$.

When the initial stress level is in the intermediate range (i.e. when the stress level is between the surge curve $g_1(D)$ and the exchangeability curve $g_2(D)$), the optimal policy takes the form of "HIIT". In particular, the individual starts with maximum effort. Doing so allows for fast skill development, but it also increases stress rapidly. At some point, as illustrated by the green arrow, the stress level reaches the exchangeability curve. At this point, the individual instantly reduces his effort to a low (but positive) effort level. Once the effort is reduced, the individual then adjusts his effort gradually so that the resulting trajectory stays on the exchangeability curve. Both the difficulty level and the stress level decrease steadily. The individual's effort, however, gradually increases. Finally, once the difficulty level drops to D_c, the individual transitions back to the maximal effort level.

When the initial stress level is very high, i.e. the initial stress level y exceeds the threshold $g_2(D)$, the optimal policy takes the form of Phased Escalation. The individual initially rests to reduce the stress level. When the stress level is reduced to $g_2(D)$, the individual instantly increases his effort to a moderate effort level. As in the "HIIT" case, the individual increases his effort steadily so that the resulting trajectory stays on the exchangeability curve $g_2(D)$. And once the difficulty level drops to D_c, the individual instantly increases his effort to the maximal level $x = x_U$.

III.D. Implications of the Optimal Strategies

We now summarize the key implications derived from our analysis, offering insights on how individuals should choose their effort trajectories to best develop their skills.

Planned Strategic Retreat as Pre-emptive Stress Management. Our optimal strategies highlight the importance of a planned strategic retreat, which occurs when the individual has an ambitious goal and a high initial stress level, that is, in the region that calls for HIIT strategy. In this case, the individual's effort drops following an initial sprint. The advantage of the planned strategic retreat is that it manages stress preemptively. As a result, the individual's stress level is sufficiently low so that he can surge when the goal is within reach. In other words, by managing stress levels in advance, the individual avoids a rest-and-surge scenario when the final goal is in sight.

In addition to calling for pre-emptive stress management, our strategies also prescribe how to carry it out. First, the individual's effort is not zero during the strategic retreat. That is, he continues to build his skill during the strategic retreat, albeit at a low speed. Second, the individual gradually increases his effort during the retreat. Therefore, the stress reduction rate decreases, and the skill-building rate increases. Third, the effort level is strictly bounded away from the maximal level. Consequently, the individual's effort level experiences a discontinuous increase over the strategic retreat ends.

Big Push vs Starting Small. Our model also provides guidance on how an individual should initiate his journey toward the goal. There are two distinct approaches: "Big Push" and "Starting Small". The Big-Push strategy involves initially putting in the maximum effort. It occurs when the goal level is either very high or very low. For intermediate goal levels, the individual starts small. He waits for the stress level to drop and gradually increases his effort.

This implies that under the optimal strategies, the initial effort level is U-shaped at the goal level, suggesting that the effect of the goal level can be subtle. When the goal level is very low, Big Push has the advantage of reaching the goal as soon as possible. However, when the goal distance increases to an intermediate level, Big Push would raise the stress level so high at some point that the breakdown becomes likely. The optimal strategy then calls for starting small. When the goal level is very high, starting small becomes suboptimal because doing so will take too long to reach the goal—the individual gains by front-loading some of the efforts earlier. That is, the individual carries out big push initially. But rather than sprinting all the way, he plans a strategic retreat midway to manage the stress.

Strategic Choice and Stress Factors. Our model shows that the choice of optimal strategies depends on many factors. Naturally, the goal's distance is a crucial factor. But individual characteristics also matter. For example, for the same goal distance, an individual who learns fast may sprint all the way. In contrast, an individual who learns slowly will find the goal ambitious and plan a strategic retreat midway. However, learning speed aside, our model also shows that stress matters a lot in determining the optimal strategy.

Our model highlights three types of stress-related characteristics. First, how fast does the stress level increase in effort? Second, what is the individual's baseline stress tolerance level? Finally, what is the rate at which the stress tolerance level changes with skill? The choice of the optimal strategy requires taking these three types of stress characteristics into account. Specifically, we can show that when the stress tolerance level increases faster in skill, the HIIT strategy is more likely (in the sense that the exchangeability curve moves up). The faster build-up in tolerance level then calls for both a Big Push initially and a planned strategic retreat.

IV. CONCLUSION

This paper revisits the Berlitz model by allowing the individual to adjust his effort over time. We adopt an exchangeability argument to characterize the optimal strategy. The results show that the optimal strategy depends both on the individual characteristic and the targeted goal. When the goal is modest, the individual either sprints all the way (surge) or rests first and then surges. When the goal is ambitious, the individual may still surge, but rest-then-surge is no longer optimal. Instead, the individual may carry out phased escalation, or he does HIIT: sprint initially, followed by a planned strategic retreat, and sprint again. The planned strategic retreat highlights the importance of pre-emptive stress management, which sets the stage for the later surge.

As in the original Berlitz model, we intentionally make the model simple and stylized. Many extensions are possible. For example, the goal is given in this paper, and one may study how best to set the goal given the individual characteristics. Second, there is only one goal in the model, and a natural follow-up is to consider how the individual should dynamically allocate effort (and leisure) across multiple goals. Finally, we may apply the model to an organizational setting, where the increase in the resistance from a project (which corresponds to the increase in the stress level in this model) depends on the actions of the other players. Understanding this may help organizations to implement changes more effectively.

REFERENCES

BOARD, SIMON AND MORITZ MEYER-TER VEHN (2013): "REPUTATION FOR QUALITY," *Econometrica*, 81 (6), 2381–2462. [3]

BONATTI, ALESSANDRO AND JOHANNES HÖRNER (2011): "Collaborating," *American Economic Review*, 101 (2), 632–663. [3]

CHE, YEON-KOO, KYUNGMIN KIM, AND KONRAD MIERENDORFF (2023): "Keeping the listener engaged: a dynamic model of bayesian persuasion," *Journal of Political Economy*, 131 (7), 1797–1844. [3]

CHE, YEON-KOO AND KONRAD MIERENDORFF (2019): "Optimal dynamic allocation of attention," *American Economic Review*, 109 (8), 2993–3029. [3]

GUO, YINGNI (2016): "Dynamic delegation of experimentation," *American Economic Review*, 106 (8), 1969–2008. [3]

HALAC, MARINA, NAVIN KARTIK, AND QINGMIN LIU (2017): "Contests for experimentation," *Journal of Political Economy*, 125 (5), 1523–1569. [3]

HIRSCHMAN, A.O. (1958): *The Strategy of Economic Development*, Yale paperbound, Yale University Press. [1]

KELLER, GODFREY AND SVEN RADY (2010): "Strategic experimentation with Poisson bandits," *Theoretical Economics*, 5 (2), 275–311. [3]

KELLER, GODFREY, SVEN RADY, AND MARTIN CRIPPS (2005): "Strategic experimentation with exponential bandits," *Econometrica*, 73 (1), 39–68. [3]

KLEIN, NICOLAS AND SVEN RADY (2011): "Negatively correlated bandits," *The Review of Economic Studies*, 78 (2), 693–732. [3]

LI, JIN, YE LUO, XIAOWEI ZHANG, AND HANMO WANG (2023): "Optimal Experimentation with Complementary Bandits," *Working Paper*. [3]

LIU, QINGMIN AND YU FU WONG (2023): "Strategic Exploration: Pre-emption and Prioritization," *Review of Economic Studies*, rdad084. [3]

MURTO, PAULI AND JUUSO VÄLIMÄKI (2011): "Learning and information aggregation in an exit game," *The Review of Economic Studies*, 78 (4), 1426–1461. [3]

ROSENBERG, DINAH, EILON SOLAN, AND NICOLAS VIEILLE (2007): "Social learning in one-arm bandit problems," *Econometrica*, 75 (6), 1591–1611. [3]

SIMON, HERBERT A (1954): "Some Strategic Considerations in the Construction of Social Science Models," *Mathematical thinking in the social sciences*, 2, 388–415. [1]

STRULOVICI, BRUNO (2010): "Learning While Voting: Determinants of Collective Experimentation," *Econometrica*, 78 (3), 933–971. [3]

URGUN, CAN (2021): "Restless Contracting," Tech. rep. [3]

ZHONG, WEIJIE (2022): "Optimal dynamic information acquisition," *Econometrica*, 90 (4), 1537–1582. [3]

APPENDIX A: SCALING OF a

For the original problem formulation (3) and (4), we can perform a change of scale to make $a = 1$. This change of scale, described below, allows us to only consider the effect of b, k_1, k_2, and x_U, therefore simplifying the subsequent analysis.

Recall the evolution equation (3). If we define $\tilde{X}_t = aX_t$, then we have

$$dD_t/dt = -D_t\tilde{X}_t,$$

and correspondingly, we define $\tilde{x}_U = ax_U$. Then we define $\tilde{b} = \frac{b}{a}$, $\tilde{k_1} = ak_1$, and $\tilde{k_2} = ak_2$. We now have

$$dY_t/dt = \tilde{b}(\tilde{X}_t - (\tilde{k_1} - \tilde{k_2}D_t)).$$

Therefore, in our problem formulation, we can assume $a = 1$ without loss of generality. Throughout the remainder of the Appendix, we will proceed under the assumption that $a = 1$, unless specified otherwise.

APPENDIX B: PROOF OF LEMMA 1, 2 AND 3

In this section, we establish Lemma 1, 2 and 3. To achieve this, we first introduce two useful functions $f(D)$ (defined in (15)) and $\Gamma(D)$ (defined in (22)) and show that they satisfy certain properties in a series of Lemmas. In particular, Lemma 4 and 5 shows the properties of $f(D)$ and Lemma 6 shows the properties of $\Gamma(D)$. After that, using the results of Lemma 4, 5 and 4, we prove Lemma 1, 2 and 3.

We start by defining the function

$$f(D) = x_U + bZ(D)(k_1 - k_2 D)e^{b(x_U - k_1)\frac{\ln D}{x_U} + \frac{bk_2}{x_U}D}, \tag{15}$$

where

$$Z(D) = \int_1^D -\frac{1}{s}e^{-b(x_U - k_1)\frac{\ln s}{x_U} - \frac{bk_2}{x_U}s}ds.$$

LEMMA 4: *There exists a threshold b_+ such that, if $b \leq b_+$, then $f(D) \geq 0$ for all $D \in [1, \frac{k_1}{k_2})$; $b > b_+$, then there exists $D \in [1, \frac{k_1}{k_2})$ such that $f(D) < 0$.*

Proof of Lemma 4. We can do the following transformations:

$$\begin{aligned} f(D,b) &= x_U + bZ(D)(k_1 - k_2 D)e^{b(x_U - k_1)\frac{\ln D}{x_U} + \frac{bk_2}{x_U}D} \\ &= x_U - (k_1 - k_2 D)\int_1^D \frac{b}{s}e^{\frac{b(x_U - k_1)}{x_U}(\ln D - \ln s) + \frac{bk_2}{x_U}(D-s)}ds \end{aligned}$$

When $b = 0$, obviously we have $f(D) = x_U$, i.e. $f(D) > 0$ for all $D \in [1, \frac{k_1}{k_2})$. When $b \to +\infty$, $f(D,b) \to -\infty$ for any $D \in (1, \frac{k_1}{k_2})$. Also, note that

$$\begin{aligned} \frac{\partial f(D,b)}{\partial b} = -(k_1 - k_2 D)\int_1^D \frac{1}{s}\left(\frac{b(x_U - k_1)}{x_U}(\ln D - \ln s) + \frac{bk_2}{x_U}(D-s) + 1\right) \\ e^{\frac{b(x_U - k_1)}{x_U}(\ln D - \ln s) + \frac{bk_2}{x_U}(D-s)}ds, \end{aligned} \tag{16}$$

so $\frac{\partial f(D,b)}{\partial b}$ has the sign opposite of to $b(x_U - k_1)(\ln D - \ln s) + bk_2(D-s) + x_U$. Since we require $x_U > k_1 - k_2$ and $s \in [1, D]$, we have

$$\begin{aligned} b(x_U - k_1)(\ln D - \ln s) + bk_2(D-s) &> b(k_1 - k_2 - k_1)(\ln D - \ln s) + bk_2(D-s) \\ &= bk_2(D - s - \ln D + \ln s) \end{aligned} \tag{17}$$

Since we have

$$\frac{\partial}{\partial s}\left(D - s - \ln D + \ln s\right) = \frac{1}{s} - 1 \leq 0$$

for $s \in [1, D]$, and when $s = D$, $D - s - \ln D + \ln s = 0$. Hence for $s \in [1, D]$, we always have $D - s - \ln D + \ln s \geq 0$. Plugging it in equation (17) yields

$$b(x_U - k_1)(\ln D - \ln s) + bk_2(D - s) + x_U > x_U. \tag{18}$$

Plugging the inequality (18) into equation (16) yields $\frac{\partial f(D,b)}{\partial b} < 0$, i.e. $f(D,b)$ is monotone in b. Hence there must exists a threshold b_+, such that, if $b \leq b_+$, then $f(D) \geq 0$ for all $D \in [1, \frac{k_1}{k_2})$; $b > b_+$, then there exists $D \in [1, \frac{k_1}{k_2})$ such that $f(D) < 0$. □

In the remainder of the appendix, we always assume that $b > b_+$. Such assumption is without loss of generality. This is because if the assumption does not hold, then surging is always optimal for any (D, y).

LEMMA 5: *$f(D)$ has exactly two zeros D_L and D_U with $1 < D_L < \frac{bk_1}{(b+1)k_2} < D_U < \frac{k_1}{k_2}$ such that we have $f(D) < 0$ over (D_L, D_U) and $f(D) \geq 0$ over $[1, \frac{k_1}{k_2}) \setminus (D_L, D_U)$.*

Proof of Lemma 5. By Lemma 4, $b > b_+$ indicates there exists $D \in [1, \frac{k_1}{k_2})$ such that $f(D) < 0$. Also note that

$$f(1) = x_U, \ f(\frac{k_1}{k_2}) = x_U,$$

so by continuity, there must exist at least two zeros D_L, D_U that complies $D_L < D_U$, $f'(D_L) < 0$, and $f'(D_U) > 0$. We denote the zeros as D_i, where $i = 1, 2, 3, \ldots$ and we investigate the derivatives at the zeros. Take derivative, we have

$$\begin{aligned} f'(D) = & -\frac{b}{D}(k_1 - k_2 D) - bk_2 Z(D) e^{b(x_U - k_1)\frac{\ln D}{x_U} + \frac{bk_2}{x_U} D} \\ & + \frac{b(x_U - (k_1 - k_2 D))}{D x_U} bZ(D)(k_1 - k_2 D) e^{b(x_U - k_1)\frac{\ln D}{x_U} + \frac{bk_2}{x_U} D}. \end{aligned} \tag{19}$$

Notice that, for zeros D_i, by definition we have

$$x_U + bZ(D_i)(k_1 - k_2 D_i) e^{b(x_U - k_1)\frac{\ln D_i}{x_U} + \frac{bk_2}{x_U} D_i} = 0,$$

so we can substitute $bZ(D_i)(k_1 - k_2 D_i) e^{b(x_U - k_1)\frac{\ln D_i}{x_U} + \frac{bk_2}{x_U} D_i}$ by $-x_U$ to equation (19), and obtain

$$f'(D_i) = \frac{(b+1)k_2 D_i - bk_1}{D_i(k_1 - k_2 D_i)} x_U \tag{20}$$

Let D_L be the smallest D_i that complies $f'(D_i) < 0$, we have $D_L \in [1, \frac{bk_1}{(b+1)k_2})$; let D_U be the largest D_i that complies $f'(D_i) > 0$, we have $D_U \in (\frac{bk_1}{(b+1)k_2}, \frac{k_1}{k_2})$. Then we cannot have any zeros in the interval $D \in [1, D_L)$. This is because D_L is the smallest D_i that complies $f'(D_i) < 0$, and by equation (20), $f'(D_i) \geq 0$ requires $D_i \geq \frac{bk_1}{(b+1)k_2} > D_L$. Similarly, we cannot have any zeros in the interval $D \in [D_U, \frac{k_1}{k_2})$. Recall that $f(1) = f(\frac{k_1}{k_2}) = x_U > 0$. Consequently, we have $f(D) \geq 0$ over $[1, \frac{k_1}{k_2}) \setminus (D_L, D_U)$.

Next we need to prove that there are no zeros in the interval $D \in (D_L, D_U)$. We suppose by contradiction that there exists a zero $D_p \in (D_L, D_U)$ and we discuss the following scenarios.

Scenario 1: $f'(D_p) < 0$. Note that $\lim_{D \to D_L^+} f(D) < 0$ and $\lim_{D \to D_p^-} f(D) > 0$, which implies there must exist another zero D_q over (D_L, D_p) such that $f'(D_q) > 0$. However, by equation (20), $f'(D_p) < 0$ implies that $D_p < \frac{bk_1}{(b+1)k_2}$, and $f'(D_q) > 0$ implies that $D_q > \frac{bk_1}{(b+1)k_2}$, i.e. $D_q > D_p$, which leads to contradiction.

Scenario 2: $f'(D_p) > 0$. Note that $\lim_{D \to D_p^+} f(D) > 0$ and $\lim_{D \to D_U^-} f(D) < 0$, which implies there must exist another zero D_q over (D_p, D_U) such that $f'(D_q) < 0$. However, by equation (20), $f'(D_p) > 0$ implies that $D_p > \frac{bk_1}{(b+1)k_2}$, and $f'(D_q) < 0$ implies that $D_q < \frac{bk_1}{(b+1)k_2}$, i.e. $D_q < D_p$, which leads to contradiction.

Scenario 3: $f'(D_p) = 0$. Then by equation (20) we know that $D_p = \frac{bk_1}{(b+1)k_2}$. Note that, by equation (20),

$$\lim_{D \to D_p^-} f'(D) < 0, \quad \lim_{D \to D_p^+} f'(D) > 0,$$

which indicates $f(D_p)$ is a local minimum, i.e. $\lim_{D \to D_p^-} f(D) > 0$ and $\lim_{D \to D_p^+} f(D) > 0$, so there must exist at least a zero D_j over (D_L, D_p) such that $f'(D_j) > 0$ and another zero D_q over (D_p, D_U) such that $f'(D_q) < 0$. However, by equation (20), $f'(D_j) > 0$ implies that $D_j > \frac{bk_1}{(b+1)k_2} = D_p$, and $f'(D_q) < 0$ implies that $D_q < \frac{bk_1}{(b+1)k_2} = D_p$, which leads to contradiction.

Hence there are no zeros in the interval $D \in (D_L, D_U)$, and $f(D) < 0$ over (D_L, D_U). □

Next, we define the following function

$$\Gamma(D) = \begin{cases} \infty, \text{ if } f(D) \geq 0 \\ \ln\left(\frac{-\lambda x_U}{f(D)}\right), \text{ otherwise} \end{cases} \tag{21}$$

This function has the following properties.

LEMMA 6: *The function $\Gamma(D)$ has following properties:*
(a). The equation

$$\Gamma(D) = \ln\left(-\frac{\lambda k_2 D}{k_2(b+1)D - bk_1}\right), \tag{22}$$

has a unique solution $D = D_c$ where $D_c \in (D_L, \frac{bk_1}{(b+1)k_2})$, and we have

$$\Gamma'(D_c) = -\frac{bk_2}{x_U} - \frac{b(x_U - k_1)}{D_c x_U}.$$

(b). For $D_L < D < D_c$, we always have

$$\Gamma'(D) < -\frac{bk_2}{x_U} - \frac{b(x_U - k_1)}{D x_U}.$$

Proof of Lemma 6. We first prove part (a). Note that

$$\Gamma'(D) = \frac{f(D)}{-\lambda x_U} \cdot \frac{\lambda x_U f'(D)}{(f(D))^2} = -\frac{f'(D)}{f(D)}.$$

We can transform the equation (22) as follows:

$$-x_U = k_2 D Z(D) e^{b(x_U - k_1)\frac{\ln D}{x_U} + \frac{bk_2}{x_U}D},$$

where the right-hand side is monotone in D. Note that when $D = 1$, RHS is 0 which is larger than LHS; when $D = \frac{bk_1}{(b+1)k_2}$, we have $b(k_1 - k_2 D) = k_2 D$, so we obtain that

$$k_2 D Z(D) e^{b(x_U - k_1)\frac{\ln D}{x_U} + \frac{bk_2}{x_U}D} = b(k_1 - k_2 D) Z(D) e^{b(x_U - k_1)\frac{\ln D}{x_U} + \frac{bk_2}{x_U}D} < -x_U.$$

The inequality holds because, by Lemma 5, $f(D) < 0$ when $D = \frac{bk_1}{(b+1)k_2}$. Thus, there must exist a unique $D = D_c$. Then we prove $\Gamma'(D_c) = -\frac{bk_2}{x_U} - \frac{b(x_U - k_1)}{D_c x_U}$. Or equivalently, we should prove

$$\frac{f'(D_c)}{f(D_c)} = \frac{bk_2}{x_U} + \frac{b(x_U - k_1)}{D_c x_U}. \tag{23}$$

Since

$$f'(D_c) = \frac{b(x_U - (k_1 - k_2 D_c))}{D_c}\left(1 - \frac{b(k_1 - k_2 D_c)}{k_2 D_c}\right),$$

and

$$f(D_c) = x_U\left(1 - \frac{b(k_1 - k_2 D_c)}{k_2 D_c}\right),$$

we have equation (23) holds. Hence $D = D_c$ is the unique solution.

Then we prove part (b). We consider another function

$$h(x) = -\frac{-\dfrac{b}{D}(k_1 - k_2 D) - bk_2 x + \dfrac{b(x_U - (k_1 - k_2 D))}{D x_U} b(k_1 - k_2 D)x}{x_U + b(k_1 - k_2 D)x}$$

and note that, when $x = Z(D)e^{b(x_U - k_1)\frac{\ln D}{x_U} + \frac{bk_2}{x_U}D}$, we have $h(x) = \Gamma'(D)$. Also, note that when $x = -\frac{x_U}{k_2 D}$, we have $h(x) = -\frac{bk_2}{x_U} - \frac{b(x_U - k_1)}{D x_U}$. Take derivative, we have

$$h'(x) = -\frac{b x_U (bk_1 - (b+1)k_2 D)}{D(x_U + b(k_1 - k_2 D)x)^2}.$$

Recall that $D_c < \frac{bk_1}{(b+1)k_2}$, so for any $D \in (D_L, D_c)$, $h'(x) < 0$. Since $D < D_c$, we have $Z(D)e^{b(x_U - k_1)\frac{\ln D}{x_U} + \frac{bk_2}{x_U}D} > -\frac{x_U}{k_2 D}$, and therefore we have $\Gamma'(D) < -\frac{bk_2}{x_U} - \frac{b(x_U - k_1)}{D x_U}$. $\square$

We are now ready to prove Lemma 1, 2 and 3.

B.A. Proof of Lemma 1.

Proof of Lemma 1. In this proof, we consider the following two strategies:

- Policy 1: Always select $X_t = x_U$.
- Policy 2: First select $X_t = 0$ and let Y_t decrease by Δy, then always select $X_t = x_U$.

Suppose the initial point is (D, y) and we take policy 1 as our standard. Then the strength of policy 2 is the relatively lower risk in the sprint ($X_t = x_U$) stage, but it suffers from the time disadvantage and excess risk caused by the rest ($X_t = 0$) stage.

Letting $\Delta y \to 0$ and applying Taylor expansion, the second policy has to endure excess discount and risk $\frac{\lambda + e^y}{b(k_1 - k_2 D)}\Delta y + O(\Delta y)^2$, but has the advantage in the sprint session for $-\frac{1}{x_U} Z(D) e^{b(x_U - k_1)\frac{\ln D}{x_U} + \frac{bk_2}{x_U}D + y}\Delta y + O(\Delta y)^2$. Taking the difference and plugging $f(D)$ we defined in Lemma 4 in, we have

$$\frac{(\lambda + e^y)\Delta y}{b(k_1 - k_2 D)} + \frac{Z(D)}{x_U} e^{b(x_U - k_1)\frac{\ln D}{x_U} + \frac{bk_2}{x_U}D + y}\Delta y = \frac{\Delta y}{x_U b(k_1 - k_2 D) f(D)}\left(e^y + \frac{\lambda x_U}{f(D)}\right).$$

Hence we need to discuss the sign of $f(D)$. When $f(D) \geq 0$, we always have the first policy dominates the second one. When $f(D) < 0$, there is a curve $y = \ln\left(-\frac{\lambda x_U}{f(D)}\right) = \Gamma(D)$ such that, above it, policy 2 is better than policy 1; below it, policy 1 is better than policy 2.

We construct the surge curve $g_1(D)$:

$$g_1(D) = \begin{cases} \Gamma(D) & \text{, for } 1 < D \leq D_c, \\ b(x_U - k_1)\frac{\ln D_c - \ln D}{x_U} + \frac{bk_2}{x_U}(D_c - D) + g_1(D_c) & \text{, for } D_c < D < \frac{k_1}{k_2}, \end{cases}$$

where D_c is defined as in Lemma 6. We will prove that $g_1(D)$ complies with the statement of Lemma 1. Note that, suppose the initial point is (D, y), then the expression of its trajectory can be written as

$$Y_t(D_t) = -b(x_U - k_1)\frac{\ln D_t}{x_U} - \frac{bk_2}{x_U}D_t + b(x_U - k_1)\frac{\ln D}{x_U} + \frac{bk_2}{x_U}D + y$$

Note that, we always have $g_1'(D_t) \leq Y_t'(D_t) < 0$ by Lemma 6 (b). Suppose $y \leq g_1(D)$, then along the whole trajectory we will have $Y_t \leq g_1(D_t)$, i.e. we will always have $Y_t \leq \Gamma(D_t)$, which indicates it is optimal to surge (i.e. always select $X_t = x_U$).

Meanwhile, if $y > g_1(D)$,then the trajectory will interfere the area above $\Gamma(D)$, i.e. there must exist t such that $Y_t > \Gamma(D_t)$. Note that at such (D_t, Y_t), rest then sprint is strictly better than sprint. Thus it is always suboptimal to surge. □

B.B. Proof of Lemma 2.

Proof of Lemma 2. Consider the following two policies (see Figure III):

- Policy SR: First select $X_t = x_U$ until difficulty level decreases by ΔD, and then select $X_t = 0$ until stress level decreases by Δy.
- Policy RS: First select $X_t = 0$ until stress level decreases by Δy, and then select $X_t = x_U$ until difficulty level decreases by ΔD.

Suppose the payoffs are P_{SR} and P_{RS}, respectively, and the individual's initial level is (D, y). We first compute the following time duration:

- Duration of sprint session in policy SR: $t_{SR}^h = -\frac{1}{x_U}\ln\left(\frac{D - \Delta D}{D}\right)$.
- Duration of rest session in policy SR: $t_{SR}^v = \frac{\Delta y}{b(k_1 - k_2(D - \Delta D))}$.

- Duration of sprint session in policy RS: $t^h_{RS} = -\frac{1}{x_U}\ln\left(\frac{D-\Delta D}{D}\right)$.
- Duration of rest session in policy RS: $t^v_{RS} = \frac{\Delta y}{b(k_1-k_2 D)}$.

Letting $\Delta D \to 0$ and $\Delta y \to 0$ and applying Taylor expansion, we have

$$\begin{aligned} t^v_{RS} &= \frac{\Delta y}{b(k_1-k_2D)} \\ t^v_{SR} &= \frac{\Delta y}{b(k_1-k_2D)} - \frac{k_2\Delta y}{b(k_1-k_2D)^2}\Delta D + O(\Delta D^2\Delta y) \\ t^h_{RS} &= t^h_{SR} = \frac{\Delta D}{Dx_U} + O(\Delta D)^2 \end{aligned}$$

For policy SR, we define the stress level right after the sprint session as Y'. Take policy RS as our standard. The time advantage of policy SR is

$$-\lambda(t^v_{SR} - t^v_{RS}) = \frac{\lambda k_2}{b(k_1-k_2D)^2}\Delta D\Delta y + O(\Delta D^2\Delta y),$$

while the excess risk of policy SR is

$$e^y t^h_{SR} + e^{Y'} t^v_{SR} - e^y t^v_{RS} - e^{y-\Delta y} t^h_{RS} = e^y \frac{bk_1-(b+1)k_2D}{Db(k_1-k_2D)^2}\Delta D\Delta y + o(\Delta D\Delta y).$$

We need to discuss the sign of $bk_1 - k_2(b+1)D$. If $bk_1 - k_2(b+1)D < 0$, we always have $P_{SR} > P_{RS}$. For $bk_1 - k_2(b+1)D > 0$, we can compute

$$e^y \frac{bk_1-(b+1)k_2D}{Db(k_1-k_2D)^2}\Delta D\Delta y = \frac{\lambda k_2}{b(k_1-k_2D)^2}\Delta D\Delta y,$$

which yields

$$y = \left(\frac{\lambda k_2 D}{bk_1-(b+1)k_2D}\right).$$

Thus, we have the exchangeability curve $g_2(D) = \ln\left(\frac{\lambda k_2 D}{bk_1-k_2(b+1)D}\right)$. When $y < g_2(D)$, $P_{SR} > P_{RS}$; when $y > g_2(D)$, $P_{SR} < P_{RS}$. □

B.C. Proof of Lemma 3.

Proof of Lemma 3. Lemma 3 holds directly as a consequence of Theorem 1, and we will prove Theorem 1 in the following section. □

APPENDIX C: PROOF OF THEOREM 1.

In this section, We prove Theorem 1 through a series of Lemmas. We first introduce an important function $\tilde{g}_2(D)$ (defined in (24)) and show that this function satisfies certain properties in Lemma 7. With this result, Lemma 8 derives, by first principle, the value function associated with the proposed policy. Then, Lemma 9 shows that the value function is C^1 and Lemma 10 shows it solves the HJB equation (9). These Lemmas, together, complete the proof of Theorem 1.

We first define

$$\tilde{g}_2(D) = \ln\left(\frac{\lambda k_2 D}{bk_1 - k_2(b+1)D}\right) + b(x_U - k_1)\frac{\ln D}{x_U} + \frac{bk_2}{x_U}D, \tag{24}$$

which has the following properties.

LEMMA 7: *The function* $\tilde{g}_2 : [D_c, \frac{bk_1}{(b+1)k_2}) \to [g_1(D_c) + b(x_U - k_1)\frac{\ln D_c}{x_U} + \frac{bk_2}{x_U}D_c, \infty)$ *is a bijection.*

Proof of Lemma 7.

$$\tilde{g}_2'(D) = -\frac{bk_1}{D(k_2(b+1)D - bk_1)} + \frac{b(x_U - k_1)}{Dx_U} + \frac{bk_2}{x_U} > 0. \tag{25}$$

$\tilde{g}_2(D)$ is continuously monotonically increasing as $D \in [D_c, \frac{bk_1}{(b+1)k_2})$, so it is a bijection. □

To facilitate the presentation of subsequent results, we define several additional functions. In particular, define

$$Y_c(D) = -b(x_U - k_1)\frac{\ln D}{x_U} - \frac{bk_2}{x_U}D + b(x_U - k_1)\frac{\ln D_c}{x_U} + \frac{bk_2}{x_U}D_c + g_1(D_c),$$

and note that $Y_c(D_c) = g_1(D_c)$. Moreover, define

$$\gamma(w) = \tilde{g}_2^{-1}(w),$$

for any $w \geq g_1(D_c) + b(x_U - k_1)\frac{\ln D_c}{x_U} + \frac{bk_2}{x_U}D_c$. Note that $\gamma(w)$ is well-defined by Lemma 7. Furthermore, define

$$Q_1(D) = \int_{D_c}^{D}\left(\frac{\lambda g_2(s) + e^{g_2(s)}}{b(k_1 - k_2 s)^2}k_2 - \frac{\lambda + e^{g_2(s)}}{s(k_1 - k_2 s)}\right)ds - \frac{\lambda \ln D_c}{x_U} + \frac{\lambda g_1(D_c) - \lambda}{b(k_1 - k_2 D_c)},$$

and

$$Q_2(D,y) = -\int_{g_2(D_c)+b(x_U-k_1)\frac{\ln D_c}{x_U}+\frac{bk_2}{x_U}D_c}^{y+b(x_U-k_1)\frac{\ln D}{x_U}+\frac{bk_2}{x_U}D} \frac{Z(\gamma(w))}{x_U}e^w + \frac{\lambda}{b(k_1 - k_2\gamma(w))} + \frac{\exp(w - b(x_U - k_1)\frac{\ln\gamma(w)}{x_U} - \frac{bk_2}{x_U}\gamma(w))}{b(k_1 - k_2\gamma(w))}dw.$$

LEMMA 8: *For Theorem 1's proposed policy, the corresponding value function is*

$$V(D,y) = \begin{cases} \mathcal{R}\exp\left(\frac{Z(D)}{x_U}e^{y+b(x_U-k_1)\frac{\ln D}{x_U}+\frac{bk_2}{x_U}D} - \frac{\lambda \ln D}{x_U}\right) & \text{, if } (D,y) \in \mathcal{M}_1 \\ \mathcal{R}\exp\left(-\frac{\lambda \ln D}{x_U} + \frac{\lambda g_1(D) - \lambda - \lambda y - e^y}{b(k_1 - k_2 D)}\right) & \text{, if } (D,y) \in \mathcal{M}_{2m} \\ \mathcal{R}\exp\left(Q_1(D) - \frac{\lambda y + e^y}{b(k_1 - k_2 D)}\right) & \text{, if } (D,y) \in \mathcal{M}_{2a} \\ \mathcal{R}\exp\left(Q_2(D,y) + \frac{Z(D)}{x_U}e^{y+b(x_U-k_1)\frac{\ln D}{x_U}+\frac{bk_2}{x_U}D} - \frac{\lambda \ln D}{x_U}\right) & \text{, if } (D,y) \in \mathcal{M}_3 \end{cases}$$

Proof of Lemma 8. If $(D,y) \in \mathcal{M}_1$, we have

$$V(D,y) = \mathcal{R}\exp\left(-\int_0^{\frac{\ln D}{x_U}} e^{Y_t} + \lambda\, dt\right)$$

Note that $dD = -X_t D_t dt$ and $Y_t(D_t) = -b(x_U - k_1)\frac{\ln D_t}{x_U} - \frac{bk_2}{x_U}D_t + b(x_U - k_1)\frac{\ln D}{x_U} + \frac{bk_2}{x_U}D + y$, so we apply the change of variable and obtained that

$$\begin{aligned} V(D,y) &= \mathcal{R}\exp\left(\int_1^D -\frac{1}{sx_U}e^{-b(x_U-k_1)\frac{\ln s}{x_U} - \frac{bk_2}{x_U}s + b(x_U-k_1)\frac{\ln D}{x_U} + \frac{bk_2}{x_U}D + y}\, ds - \int_0^{\frac{\ln D}{x_U}} \lambda\, dt\right) \\ &= \mathcal{R}\exp\left(\frac{Z(D)}{x_U}e^{y + b(x_U-k_1)\frac{\ln D}{x_U} + \frac{bk_2}{x_U}D} - \frac{\lambda \ln D}{x_U}\right) \end{aligned}$$

If $(D,y) \in \mathcal{M}_{2m}$, we have

$$\begin{aligned} V(D,y) &= V(D,g_1(D)) \cdot \exp\left(-\int_0^{\frac{y-g_1(D)}{b(k_1-k_2D)}} e^{y-b(k_1-k_2D)t} + \lambda\, dt\right) \\ &= \mathcal{R}\exp\left(-\frac{\lambda \ln D}{x_U} + \frac{\lambda g_1(D) - \lambda - \lambda y - e^y}{b(k_1 - k_2 D)}\right) \end{aligned}$$

If $(D,y) \in \mathcal{M}_{2a}$, we first compute the duration of Phase 2. Moving along with the exchangeability curve with initial level $(D, g_2(D))$, we have:

$$D_t = \frac{\dfrac{bk_1}{b+1}}{k_2 - (k_2 - \dfrac{bk_1}{D(b+1)})e^{\frac{bk_1}{b+1}t}}$$

and the duration of Phase 2 t_2 lasts for

$$t_2 = \frac{b+1}{bk_1}\ln\left(\frac{k_2 - \dfrac{bk_1}{D_c(b+1)}}{k_2 - \dfrac{bk_1}{D(b+1)}}\right),$$

and the resulting value function is

$$\begin{aligned} V(D,y) &= V(D_c, g_1(D_c)) \cdot \exp\left(-\int_0^{\frac{y-g_2(D)}{b(k_1-k_2D)}} e^{y-b(k_1-k_2D)t} + \lambda\, dt - \int_{D_c}^{D} \frac{e^{g_2(s)}}{sx}\, ds - \lambda t_2\right) \\ &= \mathcal{R}\exp\left(-\frac{\lambda + e^{g_2(D_c)}}{b(k_1 - k_2 D_c)} - \frac{\lambda \ln D_c}{x_U} + \frac{\lambda g_2(D) + e^{g_2(D)} - \lambda y - e^y}{b(k_1 - k_2 D)}\right. \\ &\qquad \left. - \int_{D_c}^{D} \frac{e^{g_2(s)}}{s(\dfrac{bk_1}{b+1} - k_2 s)}\, ds - \lambda t_2\right). \end{aligned}$$

Note that we have

$$t_2 = \frac{b+1}{bk_1} \ln \left(\frac{k_2 - \dfrac{bk_1}{D_c(b+1)}}{k_2 - \dfrac{bk_1}{D(b+1)}} \right) = \frac{b+1}{bk_1} \int_{D_c}^{D} \frac{\dfrac{bk_1}{b+1}}{s\left(\dfrac{bk_1}{b+1} - k_2 s\right)} ds.$$

Hence we can compute that

$$V(D,y) = \mathcal{R} \exp \left(-\frac{\lambda + e^{g_2(D_c)}}{b(k_1 - k_2 D_c)} - \frac{\lambda \ln D_c}{x_U} + \frac{\lambda g_2(D) + e^{g_2(D)} - \lambda y - e^y}{b(k_1 - k_2 D)} \right.$$
$$\left. - \int_{D_c}^{D} \frac{\lambda + e^{g_2(s)}}{s(\dfrac{bk_1}{b+1} - k_2 s)} ds \right).$$

Note that

$$\frac{d}{dD} \left(\frac{\lambda g_2(D) + e^{g_2(D)}}{b(k_1 - k_2 D)} - \int_{D_c}^{D} \frac{\lambda + e^{g_2(s)}}{s(\dfrac{bk_1}{b+1} - k_2 s)} ds \right)$$
$$= \frac{\lambda g_2(D) + e^{g_2(D)}}{b(k_1 - k_2 D)^2} k_2 + \frac{\dfrac{k_1}{k_1 - k_2 D} - (b+1)}{D(bk_1 - (b+1)k_2 D)} (\lambda + e^{g_2(D)})$$
$$= \frac{\lambda g_2(D) + e^{g_2(D)}}{b(k_1 - k_2 D)^2} k_2 - \frac{\lambda + e^{g_2(D)}}{D(k_1 - k_2 D)}$$
$$= Q_1'(D),$$

and

$$-\frac{\lambda + e^{g_2(D_c)}}{b(k_1 - k_2 D_c)} - \frac{\lambda \ln D_c}{x_U} + \frac{\lambda g_2(D_c) + e^{g_2(D_c)}}{b(k_1 - k_2 D_c)} = -\frac{\lambda \ln D_c}{x_U} + \frac{\lambda g_1(D_c) - \lambda}{b(k_1 - k_2 D_c)} = Q_1(D_c).$$

By the fundamental theorem of Calculus, we have

$$V(D,y) = \mathcal{R} \exp \left(Q_1(D) - \frac{\lambda y + e^y}{b(k_1 - k_2 D)} \right).$$

If $(D,y) \in \mathcal{M}_3$, we assume $u = y + b(x_U - k_1)\frac{\ln D}{x_U} + \frac{bk_2}{x_U} D$, then $\gamma(u)$ is the horizontal coordinate of the intersection point of $g_2(D)$ and the trajectory. The value function is

$$V(D,y) = \mathcal{R} \exp \left(-\int_1^{D_c} \frac{\lambda + e^{Y_c(s)}}{s x_U} ds - \int_{\gamma(u)}^{D} \frac{\lambda + e^{Y(s)}}{s x_U} ds - \int_{D_c}^{\gamma(u)} \frac{\lambda + e^{g_2(s)}}{s(\dfrac{bk_1}{b+1} - k_2 s)} ds \right).$$

We first consider phase 1 and phase 3. Note that

$$-\int_1^{D_c} \frac{\lambda + e^{Y_c(s)}}{s x_U} ds = \frac{Z(D_c)}{x_U} e^{g_2(D_c) + b(x_U - k_1)\frac{\ln D_c}{x_U} + \frac{bk_2}{x_U} D_c} - \frac{\lambda \ln D_c}{x_U}$$

and

$$-\int_{\gamma(u)}^{D} \frac{\lambda + e^{Y(s)}}{s x_U}\, ds = -\int_{1}^{D} \frac{\lambda + e^{Y(s)}}{s x_U}\, ds + \int_{1}^{\gamma(u)} \frac{\lambda + e^{Y(s)}}{s x_U}\, ds$$
$$= \frac{Z(D)}{x_U} e^{y + b(x_U - k_1)\frac{\ln D}{x_U} + \frac{bk_2}{x_U} D} - \frac{\lambda \ln D}{x_U}$$
$$- \frac{Z(\gamma(u))}{x_U} e^{g_2(\gamma(u)) + b(x_U - k_1)\frac{\ln \gamma(u)}{x_U} + \frac{bk_2}{x_U}\gamma(u)} + \frac{\lambda \ln \gamma(u)}{x_U},$$

so $-\int_1^{D_c} \frac{\lambda + e^{Y_c(s)}}{s x_U}\, ds - \int_{\gamma(u)}^{D} \frac{\lambda + e^{Y(s)}}{s x_U}\, ds$ can be converted to

$$\frac{Z(D)}{x_U} e^{y + b(x_U - k_1)\frac{\ln D}{x_U} + \frac{bk_2}{x_U} D} - \frac{\lambda \ln D}{x_U} + \left(\frac{Z(s)}{x_U} e^{g_2(s) + b(x_U - k_1)\frac{\ln s}{x_U} + \frac{bk_2}{x_U} s} - \frac{\lambda \ln s}{x_U} \right) \Big|_{s=\gamma(u)}^{D_c}.$$

By taking derivative, we can convert $\left(\frac{Z(s)}{x_U} e^{g_2(s) + b(x_U - k_1)\frac{\ln s}{x_U} + \frac{bk_2}{x_U} s} - \frac{\lambda \ln s}{x_U} \right) \Big|_{s=\gamma(u)}^{D_c}$ into the following integral form:

$$\int_{D_c}^{\gamma(u)} \frac{\lambda + e^{g_2(s)}}{s x_U} - \left(\frac{k_1}{bk_1 - k_2(b+1)s} + \frac{x_U - (k_1 - k2s)}{x_U} \right) \frac{bZ(s)}{s x_U} e^{g_2(s) + b(x_U - k_1)\frac{\ln s}{x_U} + \frac{bk_2}{x_U} s}\, ds$$

Hence we can use a change of variable by letting $w = g_2(s) + b(x_U - k_1)\frac{\ln s}{x_U} + \frac{bk_2}{x_U} s$, and therefore $dw = \left(\frac{bk_1}{s(bk_1 - k_2(b+1)s)} + \frac{b(x_U - k_1)}{s x_U} + \frac{bk_2}{x_U} \right) ds$, $s = \gamma(w)$, which yields

$$\frac{Z(s)}{x_U} e^{g_2(s) + b(x_U - k_1)\frac{\ln s}{x_U} + \frac{bk_2}{x_U} s} - \frac{\lambda \ln s}{x_U} \Big|_{s=\gamma(u)}^{D_c} = -\int_{\tilde{g}_2(D_c)}^{y + b(x_U - k_1)\frac{\ln D}{x_U} + \frac{bk_2}{x_U} D} \frac{Z(\gamma(w))}{x_U} e^{w}\, dw$$

Finally, for phase 2, we have

$$\int_{D_c}^{\gamma(u)} \frac{\lambda + e^{g_2(s)}}{s x_U} - \frac{\lambda + e^{g_2(s)}}{s(\frac{bk_1}{b+1} - k_2 s)}\, ds = \int_{D_c}^{\gamma(u)} (\lambda + e^{g_2(s)}) \frac{bk_1 - k_2(b+1)s - (b+1)x_U}{s x_U (bk_1 - k_2(b+1)s)} ds.$$

Since we can do the following transformation

$$\frac{bk_1 - k_2(b+1)s - (b+1)x_U}{s x_U (bk_1 - k_2(b+1)s)} b(k_1 - k_2 s) = \frac{dw}{ds},$$

we can transform $\int_{D_c}^{\gamma(u)} \frac{\lambda + e^{g_2(s)}}{s x_U} - \frac{\lambda + e^{g_2(s)}}{s(\frac{bk_1}{b+1} - k_2 s)}\, ds$ as

$$-\int_{g_2(D_c) + b(x_U - k_1)\frac{\ln D_c}{x_U} + \frac{bk_2}{x_U} D_c}^{y + b(x_U - k_1)\frac{\ln D}{x_U} + \frac{bk_2}{x_U} D} \frac{\lambda + \exp(w - b(x_U - k_1)\frac{\ln \gamma(w)}{x_U} - \frac{bk_2}{x_U}\gamma(w))}{b(k_1 - k_2\gamma(w))}\, dw.$$

Thus,

$$V(D,y) = \mathcal{R}\exp\left(-\int_1^{D_c} \frac{\lambda + e^{Y_c(s)}}{sx_U}\,ds - \int_{\gamma(u)}^{D} \frac{\lambda + e^{Y(s)}}{sx_U}\,ds - \int_{D_c}^{\gamma(u)} \frac{\lambda + e^{g_2(s)}}{s(\frac{bk_1}{b+1} - k_2 s)}\,ds\right)$$
$$= \mathcal{R}\exp\left(Q_2(D,y) + \frac{Z(D)}{x_U} e^{y + b(x_U - k_1)\frac{\ln D}{x_U} + \frac{bk_2}{x_U}D} - \frac{\lambda \ln D}{x_U}\right) \quad \square$$

LEMMA 9: *$V(D,y)$ defined in Lemma 8 is C^1.*

Proof of Lemma 9. It suffices to show that $V(D,y)$ is C^1 at the region boundaries. Also, since all four pieces of $V(D,y)$ are exponential functions, it suffices to show the exponential parts are C^1. For simplicity, we denote the exponential parts of $V(D,y)$ in $\mathcal{M}_1$, $\mathcal{M}_{2m}$, $\mathcal{M}_{2a}$, and $\mathcal{M}_3$ by V_1, V_2, V_3, and V_4 respectively.

On the boundary of $\mathcal{M}_1$ and $\mathcal{M}_{2m}$, i.e. the set of (D,y) such that $y = g_1(D) = \Gamma(D)$, note that

$$e^{g_1(D) + b(x_U - k_1)\frac{\ln D}{x_U} + \frac{bk_2}{x_U}D} = -\frac{x_U(\lambda + e^{g_1(D)})}{bZ(D)(k_1 - k_2 D)}.$$

We can compute that

$$\frac{\partial}{\partial D}(V_1 - V_2)\big|_{y=g_1(D)} = \frac{Z'(D)}{x_U} e^{g_1(D) + b(x_U - k_1)\frac{\ln D}{x_U} + \frac{bk_2}{x_U}D} - \frac{Z'(D)}{x_U} e^{g_1(D) + b(x_U - k_1)\frac{\ln D}{x_U} + \frac{bk_2}{x_U}D}$$
$$+ \frac{\lambda g_1'(D) + g_1'(D) e^{g_1(D)}}{b(k_1 - k_2 D)} - \frac{\lambda g_1'(D) + g_1'(D) e^{g_1(D)}}{b(k_1 - k_2 D)}$$
$$= 0,$$

and

$$\frac{\partial}{\partial y}(V_1 - V_2)\big|_{y=g_1(D)} = \frac{Z(D)}{x_U} e^{g_1(D) + b(x_U - k_1)\frac{\ln D}{x_U} + \frac{bk_2}{x_U}D} + \frac{\lambda + e^{g_1(D)}}{b(k_1 - k_2 D)} = 0,$$

so we have $V(D,y)$ is C^1 on this boundary.

On the boundary of $\mathcal{M}_{2m}$ and $\mathcal{M}_{2a}$, i.e. the set of (D,y) such that $D = D_c$, note that we have $g_1(D_c) = g_2(D_c)$.

$$\frac{\partial}{\partial D}(V_2 - V_3)\big|_{D=D_c} = -\frac{\lambda + e^{g_1(D_c)}}{D_c x_U} - \frac{\lambda + e^{g_1(D_c)}}{b(k_1 - k_2 D_c)}\left(\frac{b(x_U - k_1)}{D_c x_U} + \frac{bk_2}{x_U}\right) + \frac{\lambda + e^{g_2(D_c)}}{D_c(k_1 - k_2 D_c)}$$
$$= 0,$$

and

$$\frac{\partial}{\partial y}(V_2 - V_3)\big|_{D=D_c} = -\frac{\lambda + e^y}{b(k_1 - k_2 D_c)} + \frac{\lambda + e^y}{b(k_1 - k_2 D_c)} = 0,$$

so we have $V(D,y)$ is C^1 on this boundary.

On the boundary of $\mathcal{M}_{2a}$ and $\mathcal{M}_3$, i.e. the set of (D,y) such that $y = g_2(D)$, we have

$$\begin{aligned}\frac{\partial}{\partial D}(V_3 - V_4)\big|_{y=g_2(D)} &= -\frac{\lambda + e^{g_2(D)}}{D(k_1 - k_2 D)} + \frac{\lambda + e^{g_2(D)}}{b(k_1 - k_2 D)}\left(\frac{b(x_U - k_1)}{Dx_U} + \frac{bk_2}{x_U}\right) + \frac{e^{g_2(D)} + \lambda}{Dx_U} \\ &= 0\end{aligned}$$

and

$$\begin{aligned}\frac{\partial}{\partial y}(V_3 - V_4)\big|_{y=g_2(D)} &= -\frac{\lambda + e^{g_2(D)}}{b(k_1 - k_2 D)} + \frac{\lambda + e^{g_2(D)}}{b(k_1 - k_2 D)} \\ &= 0,\end{aligned}$$

so we have $V(D,y)$ is C^1 on this boundary.

On the boundary of $\mathcal{M}_1$ and $\mathcal{M}_3$, i.e. the set of (D,y) such that $y = g_1(D) = Y_c(D)$, we have

$$\frac{\partial}{\partial D}(V_1 - V_4)\big|_{y=Y_c(D)} = \left(\frac{\lambda + e^{g_1(D_c)}}{b(k_1 - k_2 D_c)} - \frac{\lambda + e^{g_1(D_c)}}{b(k_1 - k_2 D_c)}\right)\left(\frac{b(x_U - k_1)}{Dx_U} + \frac{bk_2}{x_U}\right) = 0.$$

The equation holds because $Y_c(D) + b(x_U - k_1)\frac{\ln D}{x_U} + \frac{bk_2}{x_U}D = g_1(D_c) + b(x_U - k_1)\frac{\ln D_c}{x_U} + \frac{bk_2}{x_U}D_c$. Also,

$$\frac{\partial}{\partial y}(V_1 - V_4)\big|_{y=Y_c(D)} = -\frac{\lambda + e^{g_1(D_c)}}{b(k_1 - k_2 D_c)} + \frac{\lambda + e^{g_1(D_c)}}{b(k_1 - k_2 D_c)} = 0.$$

so we have $V(D,y)$ is C^1 on this boundary, which completes the proof of $V(D,y)$ is C^1. $\square$

Now we prove that $V(D,y)$ satisfies the HJB equation.

LEMMA 10: *$V(D,y)$ solves the HJB equation* (9).

Proof of Lemma 10. Note that, $-D\frac{\partial V(D,y)}{\partial D} + b\frac{\partial V(D,y)}{\partial y}$ has the same sign with $-D\frac{\partial V_i}{\partial D} + b\frac{\partial V_i}{\partial y}$ for $i = 1,\ldots,4$ and it suffices to show $e^y + \lambda = -DX_t\frac{\partial V_i}{\partial D} + b(X_t - (k_1 - k_2 D))\frac{\partial V_i}{\partial y}$ for $i = 1,\ldots,4$.

In the first region $\mathcal{M}_1$, note that we always have $y \le g_1(D)$. Thus we have

$$-D\frac{\partial V_1}{\partial D} + b\frac{\partial V_1}{\partial y} = \frac{\lambda + e^y}{x_U} + b(k_1 - k_2 D)\frac{Z(D)}{x_U^2}e^{y + b(x_U - k_1)\frac{\ln D}{x_U} + \frac{bk_2}{x_U}D}$$

If $x_U + bZ(D)(k_1 - k_2 D)e^{b(x_U - k_1)\frac{\ln D}{x_U} + \frac{bk_2}{x_U}D} < 0$, we have:

$$\begin{aligned}-D\frac{\partial V_1}{\partial D} + b\frac{\partial V_1}{\partial y} &= \frac{\lambda}{x_U} + \frac{e^y}{x_U e^{g_1(D)}}\left(e^{g_1(D)} + b(k_1 - k_2 D)\frac{Z(D)}{x_U}e^{g_1(D) + b(x_U - k_1)\frac{\ln D}{x_U} + \frac{bk_2}{x_U}D}\right) \\ &= \frac{\lambda}{x_U}\left(1 - \frac{e^y}{e^{g_1(D)}}\right) \\ &\ge 0.\end{aligned}$$

Otherwise, we have

$$-D\frac{\partial V_1}{\partial D}+b\frac{\partial V_1}{\partial y}\geq\frac{\lambda+e^y}{x_U}-\frac{e^y}{x_U}=\frac{\lambda}{x_U}\geq 0.$$

Hence, we can compute

$$-Dx_U\frac{\partial V_1}{\partial D}+b(x_U-(k_1-k_2D))\frac{\partial V_1}{\partial y}=-Dx_U\left(\frac{e^y}{-Dx_U}-\frac{\lambda}{Dx_U}\right)=e^y+\lambda$$

In the second region $\mathcal{M}_{2m}$, suppose that

$$f_2(y)=(k_2(b+1)D-bk_1)(e^y-e^{g_1(D)})+\lambda k_2D(y-g_1(D))$$

and

$$f_2'(y)=(k_2(b+1)D-bk_1)e^y+\lambda k_2D.$$

Note that we have $y>g_1(D)$ and $k_2(b+1)D-bk_1<0$, so we can compute that

$$\begin{aligned}f_2'(y)&<(k_2(b+1)D-bk_1)e^{g_1(D)}+\lambda k_2D\\&=\frac{x_U+k_2DZ(D)e^{b(x_U-k_1)\frac{\ln D}{x_U}+\frac{bk_2}{x_U}D}}{x_U+bZ(D)(k_1-k_2D)e^{b(x_U-k_1)\frac{\ln D}{x_U}+\frac{bk_2}{x_U}D}}\lambda b(k_1-k_2D)\end{aligned}$$

Recall that D_c has the property that

$$x_U+k_2D_cZ(D_c)e^{b(x_U-k_1)\frac{\ln D_c}{x_U}+\frac{bk_2}{x_U}D_c}=0.$$

Region $\mathcal{M}_{2m}$ indicates $D\leq D_c$, so we have

$$x_U+k_2DZ(D)e^{b(x_U-k_1)\frac{\ln D}{x_U}+\frac{bk_2}{x_U}D}\geq 0.$$

Hence we have

$$\frac{x_U+k_2DZ(D)e^{b(x_U-k_1)\frac{\ln D}{x_U}+\frac{bk_2}{x_U}D}}{x_U+bZ(D)(k_1-k_2D)e^{b(x_U-k_1)\frac{\ln D}{x_U}+\frac{bk_2}{x_U}D}}\lambda b(k_1-k_2D)\leq 0,$$

and therefore,

$$f_2'(y)<0$$

for $y>g_1(D)$ and

$$f_2(y)<f_2(g_1(D))=0.$$

Hence we obtain that

$$-D\frac{\partial V_2}{\partial D}+b\frac{\partial V_2}{\partial y}<0.$$

Hence we have

$$-b(k_1 - k_2D)\frac{\partial V_2}{\partial y} = \lambda + e^y$$

In $\mathcal{M}_{2a}$, we have

$$-D\frac{\partial V_3}{\partial D} + b\frac{\partial V_3}{\partial y} = \frac{(k_2(b+1)D - bk_1)(e^y - e^{g_2(D)}) + \lambda k_2 D(y - g_2(D))}{b(k_1 - k_2D)^2}$$

and obviously $b(k_1 - k_2D)^2 > 0$. Suppose that

$$f_3(y) = (k_2(b+1)D - bk_1)(e^y - e^{g_2(D)}) + \lambda k_2 D(y - g_2(D))$$

and

$$f_3'(y) = (k_2(b+1)D - bk_1)e^y + \lambda k_2 D.$$

Note that we have $y \geq g_2(D)$ and $k_2(b+1)D - bk_1 < 0$, so we can compute that

$$f_3'(y) \leq (k_2(b+1)D - bk_1)e^{g_2(D)} + \lambda k_2 D = 0.$$

Obviously that $f_3(g_2(D)) = 0$, so we have

$$f_3(y) \leq f_3(g_2(D)) = 0$$

and therefore

$$-D\frac{\partial V_3}{\partial D} + b\frac{\partial V_3}{\partial y} \leq 0.$$

Hence we have

$$-b(k_1 - k_2D)\frac{\partial V_3}{\partial y} = \lambda + e^y$$

For the region $\mathcal{M}_3$, we introduce $w = y + \frac{b(x_U - k_1)}{x_U}\ln D + \frac{bk_2}{x_U}D$ for our convenience. In the fourth region, we have

$$\begin{aligned}
-D\frac{\partial V_4}{\partial D} + b\frac{\partial V_4}{\partial y} = &-\frac{b(k_1 - k_2D)}{x_U}\left(\frac{Z(\gamma(w))}{x_U}e^w - \frac{Z(D)}{x_U}e^w\right) \\
&- \frac{b(k_1 - k_2D)}{x_U}\frac{\lambda + \exp\left(w - b(x_U - k_1)\frac{\ln\gamma(w)}{x_U} - \frac{bk_2}{x_U}\gamma(w)\right)}{b(k_1 - k_2\gamma(w))} \\
&- D\left(\frac{Z'(D)}{x_U}e^w - \frac{\lambda}{Dx_U}\right)
\end{aligned}$$

Suppose that

$$f_4(D, w) = -D\frac{\partial V_4}{\partial D} + b\frac{\partial V_4}{\partial y}$$

Note that, although at first, w is a function of D and y, we can now view w and D are independent variables, with the constraint $D > \gamma(w)$. It is like we project the original region from the (D,y)-plane to a (D,w)-plane. This transformation will benefit us in learning the lower bound of $-D\frac{\partial V_4}{\partial D} + b\frac{\partial V_4}{\partial y}$.

Note that we have

$$f_4(\gamma(w),w) = 0,$$

so we compute

$$\begin{aligned}\frac{\partial f_4}{\partial D} = &\frac{1}{Dx_U(k_1 - k_2\gamma(w))}\Bigg(\lambda k_2 D + k_2 D\exp(w - b(x_U - k_1)\frac{\ln\gamma(w)}{x_U} - \frac{bk_2}{x_U}\gamma(w)) \\ &-b(k_1 - k_2\gamma(w))e^{w-b(x_U-k_1)\frac{\ln D}{x_U} - \frac{bk_2}{x_U}D}\Bigg) + \frac{bk_2}{x_U}\left(\frac{Z(\gamma(w))}{x_U}e^w - \frac{Z(D)}{x_U}e^w\right)\end{aligned}$$

Note that $D > \gamma(w)$, so we have $\frac{Z(\gamma(w))}{x_U}e^w - \frac{Z(D)}{x_U}e^w > 0$ and

$$-b(k_1 - k_2\gamma(w))e^{w-b(x_U-k_1)\frac{\ln D}{x_U} - \frac{bk_2}{x_U}D} > -b(k_1 - k_2\gamma(w))e^{w-b(x_U-k_1)\frac{\ln\gamma(w)}{x_U} - \frac{bk_2}{x_U}\gamma(w)}$$

and therefore,

$$\begin{aligned}&\lambda k_2 D + k_2 D\exp\left(w - b(x_U - k_1)\frac{\ln\gamma(w)}{x_U} - \frac{bk_2}{x_U}\gamma(w)\right) - b(k_1 - k_2\gamma(w))e^y \\ >&\lambda k_2\gamma(w) + (k_2\gamma(w) - b(k_1 - k_2\gamma(w)))\exp\left(w - b(x_U - k_1)\frac{\ln\gamma(w)}{x_U} - \frac{bk_2}{x_U}\gamma(w)\right) \\ =&0\end{aligned}$$

Hence we always have $\frac{\partial f_4}{\partial D} > 0$ when $D > \gamma(w)$, so we know

$$f_4(D,w) > f_4(\gamma(w),w) = 0,$$

and therefore,

$$-D\frac{\partial V_4}{\partial D} + b\frac{\partial V_4}{\partial y} > 0.$$

and

$$-Dx_U\frac{\partial V_4}{\partial D} + b(x_U - (k_1 - k_2 D))\frac{\partial V_4}{\partial y} = -Dx_U\left(\frac{Z'(D)}{x_U}e^w - \frac{\lambda}{Dx_U}\right) = e^y + \lambda \quad \square$$

Proof of Theorem 1. Theorem 1 holds as a result of Lemma 8, Lemma 9, and Lemma 10. □