Managing Customer Churn via Service Mode Control

Yash Kanoria, Ilan Lobel, Jiaqi Lu

January 4, 2020

Abstract

We consider the problem of a service firm interacting over time with one of its customers. The firm has two service modes available, which differ in their expected reward rates as well as their volatilities (risk). The firm’s objective is to maximize the rewards generated over the customer’s lifetime. Meanwhile, the customer is unsophisticated and might abandon the system if unsatisfied with recent rewards. We show that the firm’s optimal policy is either myopic or a sandwich policy. A sandwich policy is one where the firm utilizes the myopically optimal service mode when the customer is either very happy or very unhappy but that utilizes the service mode with inferior reward rate when the customer happiness is in a specific interval near the satisfaction threshold. Specifically, the firm should be risk averse when the customer is marginally satisfied and risk seeking when the customer is marginally unsatisfied. Our results extend to a setting where the firm can mix between the two service modes, and to one where the rewards accrue according to a geometric (instead of a regular) Brownian motion with drift. We show numerically that our results are robust to adding a small switching cost, and to different specifications of the hazard rate of customer abandonment as a function of happiness. We also demonstrate numerically that the customer lifetime value under the optimal policy is large relative to that under the myopic policy.

1 Introduction

Online firms spend significant sums attracting customers via digital advertising, but they often find that customers do not stay engaged for very long. This is particularly problematic for service firms that hope their newfound customers will sign up for a service subscription, or at least engage them repeatedly over time. Keeping customers from quitting (reducing churn) is therefore seen by many online firms and platforms as key to their long-term viability and success.

Why do customers quit online platforms or services they have previously signed up for? In many instances it is because they are dissatisfied with the recent service they were provided. There

*Columbia Business School, Columbia University — ykanoria@gsb.columbia.edu
†Stern School of Business, New York University — ilobel@stern.nyu.edu
‡Columbia Business School, Columbia University — jlu20@gsb.columbia.edu
§We thank Rene Caldentey and Iosif Pinelis for pointing out relevant research and techniques.
are several strategies companies can adopt to try to better retain customers, including providing discounts and improving customer service. In this paper, we study the impact of the company’s *service strategy* on customer retention. That is, when facing customers who are prone to leaving the system, should companies play it safe or take risks?

As a focal example, let us consider the case of an investment manager, which could either be a human financial advisor or a robot-advisor platform. Suppose the investment allocation chosen by the investment manager has been performing poorly in the last few months. Given this poor performance in the recent past, the client is a flight risk. How should the investment manager respond? Should the manager shift the portfolio toward higher-risk, higher-return assets like stocks or toward safer ones such as bonds? Our paper aims to address questions of this form, and to understand the business value of this kind of service mode control.

**Our formulation (an informal preview).** Our model contains two players, a service firm and a customer. At each point in time, the firm chooses a service mode: Risky or Safe. In our Base model formulation, the Safe mode generates rewards at a constant rate, while the Risky mode produces Brownian rewards\(^1\) with a drift that can be either higher or lower than the Safe mode. We assume the rewards accrue to both the firm and the customer according to some fixed proportion, so that both parties care about the rewards. The firm is patient and rational and wants to maximize the long-term rewards produced by its interaction with the customer. Meanwhile, the customer is assumed to be subject to recency bias. More precisely, the customer’s happiness follows a goodwill model that weighs recent experiences more heavily, which we model as an exponentially weighted moving average of recent experiences (as in Nerlove and Arrow [19]). The customer has a hazard rate of leaving that is a function of his happiness (throughout the text, we use male pronouns to refer to the customer). The hazard rate is assumed to be a step function where the customer quits with positive hazard rate if the happiness is below a given satisfaction threshold and with zero hazard rate if the happiness is at or above the threshold (we relax this functional form in our numerics in Section 8 and find that our main results are robust). The firm’s problem is to determine a policy that maximizes the rewards generated before the customer exits the system.

We assume the firm can observe how happy the customer is with the service. This assumption is valid in the investment manager example, since market performance is public, and is a reasonable approximation in settings where the firm can estimate customer happiness with some accuracy. In Section 2, we discuss methods in the literature for estimating customer happiness and churn probability.

**Our results.** For the Base model, we find (Theorem 1) that the optimal policy is of one of two kinds: it is either the *myopic policy*, which always chooses the superior mode (superior in terms of instantaneous reward rate) regardless of customer happiness, or a *sandwich policy* that chooses the inferior mode in an intermediate happiness interval and the superior mode elsewhere. The

\(^1\)We also consider a geometric Brownian motion reward formulation in Section 6.
emergence of the sandwich policy as the optimal one in a large part of the parameter space is interesting. In particular, when the Safe service mode is myopically superior, the optimal policy is always of sandwich type, where the firm utilizes the Risky mode when the customer happiness is slightly below the customer satisfaction threshold. This is in contrast with the customary wisdom that a high volatility and low return option is dominated by a low volatility and high return option. (We discuss this case further under “Technical contributions” below.) When the Risky mode is superior, the optimal policy is either myopic (Risky always) or of sandwich type, where the firm utilizes the Safe mode when the customer happiness is at or slightly above the customer satisfaction threshold.

Another equivalent but nevertheless useful way to describe the optimal policy is as follows. The firm should be risk averse if the customer is currently not a flight risk but may soon become one (if the next experiences are negative), and be risk seeking if the customer is currently a flight risk but will no longer be one if the next experiences are positive (see Figure 1 in Section 4). Whereas the former feature of the optimal policy aligns with conventional wisdom, the latter feature is novel. The reason for being risk seeking is that the Risky mode increases the chance that the customer will soon have a few positive experiences and hence become satisfied and no longer a flight risk.

Next, we establish several comparative statics results, which may be summarized as follows: the deviation from the myopic policy is more substantial in cases where the alternative service mode has a similar instantaneous reward rate, and in cases where the difference in riskiness is larger.

Numerical investigations reveal a very substantial customer lifetime value (CLV) improvement compared with a myopic policy that doesn’t internalize the service mode’s impact on the customer’s probabilistic churn, in a large part of the parameter space. For example, we see CLV increases of 100% or more in many cases (see Figure 5 in Section 4.3).

Model variants and robustness checks. We also consider several variants and generalizations of the Base model. Our main findings remain intact in each case; the structure of the optimal policy remains similar, and the CLV increase relative to the myopic policy remains large in much of the parameter space. First, in the “Investor model” presented in Section 5, the firm is allowed to mix between the Safe and Risky modes at each point in time, and the Risky mode provides a superior reward rate. This larger action set is important to capture the investment manager example discussed above. We prove that the optimal policy, if not myopic (i.e., pure Risky everywhere), is again a sandwich policy with the middle interval containing a mixed policy rather than using only the Safe mode. In Section 6, we consider a variant of our model where rewards accrue according to a geometric Brownian motion instead of a Brownian motion, corresponding to typical models of financial asset returns (Harrison and Pliska [12], Broadie and Glasserman [7], Richardson [26], among others). We formally establish optimality conditions for this setting and then show via numerical solutions that an optimal policy is again either the myopic one or a sandwich policy. In Section 7, we examine the effect of a fixed positive cost of switching from one service mode to the other. We find for small switching costs that the firm’s optimal policy still preserves the sandwich
structure, with switching thresholds replaced by buffer intervals where the policy retains the current service mode. Also, as the switching cost increases, the optimal policy transforms smoothly. We characterize the firm’s optimal policy under different switching costs and find that a large CLV increase over the myopic policy persists. The strong performance of the optimal sandwich policy with buffers further indicates that the sandwich policy is robust to errors in estimating customer happiness (see the discussion at the end of Section 7). Lastly, we test robustness of our results for different (abandonment) hazard rate functions in Section 8 and numerically show that our main results continue to hold; in particular, a myopic policy or a sandwich policy is again optimal for all hazard rate functions tested.

**Contribution to customer relationship management (managerial insights).** Our paper makes the following key contributions towards enabling service firms such as investment managers to better manage their relationships with customers.

We introduce the novel idea of managing customer churn via controlling the riskiness of the service mode deployed. We build and solve (despite technical challenges, discussed below) a natural model which captures the customer’s recency bias and propensity to depart as a function of the recent experiences, as well as the firm’s different service options.

We find a remarkably large increase in the customer lifetime value (CLV) from using the optimal policy relative to the myopic policy, in a large part of the parameter space. Furthermore, the optimal policy has a simple and practical structure, and the policy structure as well as the size of the increase in CLV are robust to a number of different modifications to our Base model, which address the various practical concerns in the case of an investment manager (as well as in other applications where the firm is able to estimate customer happiness reasonably well). Meanwhile, the control lever we propose appears to be unexplored in practice, i.e., current practice in effect relies on the myopic policy (deploy the service mode with highest expected rewards) and ignores riskiness of the service mode deployed. Thus, our findings strongly suggest that service firms such as investment managers can unlock a large increase in CLV by correctly choosing between service modes with different risk levels based on the customer’s current happiness. Crucially, unlocking this value requires no additional expense to the firm. The absence of additional expense is in sharp contrast to the popular levers for managing customer churn such as discounts, and improved quality or effort level, which can be very expensive.

We provide clean structural insights on when and how the firm should deviate from using the myopically optimal service mode. The optimal policy is very simple and hence practically appealing: it is either myopic or a sandwich policy which uses the inferior mode for intermediate happiness values. The firm should be risk averse when the customer is currently not a flight risk but may soon become one if the next experiences are negative, and be risk seeking when the customer is currently a flight risk but might no longer be one if the next experiences are positive.

---

2To develop an understanding of industry practice in managing customer churn, we spoke with a variety of practitioners, including data scientists at Pandora and Instagram.

4
Technical contributions. Our model poses the firm’s problem as a stochastic control problem, with two levels of stochasticity. The first type of stochasticity is in the reward process, which we model using a Brownian motion. In addition there is a second layer of uncertainty to capture the customer’s probabilistic churn, and both sources of stochasticity are impacted by the firm’s service mode. This kind of additional uncertainty (captured by the hazard rate of churn in our model) is new to the stochastic control literature, to the best of our knowledge. This feature makes our problem difficult to analyze, as we discuss below. The payoff from solving this technically challenging formulation is in the clean structure of the optimal policy derived, leading to clear managerial insights.3

Obtaining a closed-form solution in our setting was a challenging task, one that required subtle modeling that balanced algebraic hardness with stochastic control tractability. To reduce the algebraic hardness in analytically obtaining the optimal policy, we choose to model the Safe mode as having zero volatility and the hazard rate of churn to be a step function of the happiness state. However, these assumptions invalidate the usual sufficient conditions for the processes to be well-defined semimartingales and for the HJB equation to have a twice-continuously differentiable solution, as it turns out that in our model the customer happiness may spend positive time at a single happiness value at the boundary where the policy switches between the Safe and Risky modes. We built upon a very recent advance in stochastic processes (Salins and Spiliopoulos [28]) to be able to modify and apply classic stochastic control methodologies to our setting.

Next, to work around the algebra when solving the HJB equation (specifically while verifying optimality of the derived policy by showing nonpositivity in Condition 4 of Proposition 2; see Section 4.2), we crucially drew upon L’Hopital-type rules for monotonicity (Pinelis [22]) and Chernoff-type bounds for the error function (Chang [8]) to establish properties of functions involving the error function (see Appendix B.1 for the various results). In general, our analysis showcases a way to deal with algebraic hardness in the verification step when solving stochastic control applications, via the levers provided by Pinelis [22]. This method can potentially be applied to other stochastic control applications where the verification step is challenging.

What we do not study. In this work, we developed a model to study how customer churn can be managed via a company’s service strategy. Managing churn is an important business problem, and there are many interesting aspects of it that we do not study in this paper. We do not study service effort level control, a topic explored in Aflaki and Popescu [2]. We also do not focus on discounts as a technique for managing churn, a problem that would lead to similar incentive misalignments between the firm and the customer as an effort model, since providing a discount is costly for the firm but beneficial to the customer. We do not incorporate these features in our model so that we can isolate and determine the value of service mode control. In our model, the incentives of the players are aligned, except that the firm is a more careful long-term thinker than the customer.

3We expect the optimal policy to be messier and challenging to interpret under alternate modelling approaches where time, customer happiness and/or rewards are discretized, though the analysis may be simpler.
2 Related Literature

The most closely related paper to ours is Aflaki and Popescu [2]. That paper studies a relationship between a service firm and a customer who might abandon its service. The service firm aims to maximize the value it obtains from the relationship, and its control is the service effort level at each point in time. Increasing the service level is costly to the firm but increases the customer’s likelihood of staying in the system. As in our method, they use a goodwill model to model happiness (they also study a habituation model we do not consider). They find that an optimal effort policy leads to stationary effort and happiness levels. The two key differences between our model and theirs is that our model is stochastic (happiness is a deterministic function of effort in Aflaki and Popescu) and that our control is the firm’s service mode rather than its effort level. That is, the firm cannot simply increase its effort when the customer is unhappy, but it can choose to increase/decrease the risk level of its service.

Goodwill models were first introduced in a seminal paper by Nerlove and Arrow [19]. That paper aimed to study the dynamic impact of advertising on demand. Nerlove and Arrow’s goodwill model had two important properties: it captured the fact that advertising has persistent effect on demand but that, at the same time, the effect of advertising disappears exponentially fast. We make the same assumption regarding how consumer happiness responds to past quality of service. Goodwill models have been validated empirically (Zeithaml [35]) and have the support of celebrated behavioral economics experiments that show that individuals overweight recent experiences (Kahneman et al. [15]). Goodwill models (and their variants) have been used in operations and marketing to model how customers respond to past fill rates (Gaur and Park [11], Adelman and Mersereau [1], Liu and van Ryzin [16]), how customers recall service experiences (Ho et. al. [13], Das Gupta et al. [10]), and how customers learn about prices (Ovchinnikov and Milner [21]).

In terms of methodology, we build on techniques from the stochastic control literature. We assume that consumer happiness is a continuous-time stochastic process that is affected by the firm’s choice of service mode at each point in time. This is a departure from the discrete time modeling of Aflaki and Popescu [2]. Moving from discrete time to continuous time increases the technical complexity of our paper, but it improves the paper in that it allows us to obtain crisper theorems. To solve our stochastic control model, we use a “smooth pasting” technique that matches the value function and its derivatives at happiness values where the optimal service mode changes. To prove the optimality of the policy we obtain, we use the verification technique that is typical of the stochastic control literature (Borkar [6], Mirică [17], Radner and Shepp [23], Touzi [33], and Ata et al. [4]). We also build on ideas and results from recent papers on stochastic calculus, such Strulovici and Szyslowski [31]’s sufficient conditions for a value function to be twice continuously differentiable (which our models do not quite satisfy because the Safe service mode has zero volatility), and Salins and Spiliopoulos [28]’s characterization of Markov processes with spatial delays.

A concurrent work by Johari and Schmit [14] studies the problem of learning about a customer
who may abandon. A customer leaves the first time the firm does not meet her expectations, and the firm tries to balance the risk of customer departure with rewards earned. The focus of the paper, including the model and results, are quite different from the present paper.

Our paper assumes the firm has full information about the customer. This requires the firm to have a fair understanding about the customer's churn behavior as a function of the state of the relationship between the customer and the firm (which we call it the customer happiness state in our paper). There are a few different methods in the marketing literature for dealing with this challenge. Netzer et al. [20] introduced an approach to model latent customer relationship dynamics using hidden Markov models (HMM). An HMM can capture the transitions among the latent relationship states as well as the probabilistic customer behavior that depends on the current customer state. This model can be calibrated using real data. For example, Netzer et al. used an HMM to estimate university-alumni relations and the alumni's behavior conditional on this relation. In another example, Ascarza et al. [3] analyze customer activities data from a daily deal website using HMMs. They identify different states of the customer relationship on which customer's propensity to churn or engage depends. From the observed activity data, they are able to estimate what state the customer relationship is in. Similarly, Montoya et al. [18] utilize an HMM and prescription activities data from a pharmaceutical firm to estimate the underlying prescription state physicians are in. These estimates are then used to optimize sales efforts by solving a partially observable Markov decision process. These methodologies provide guidance on how a firm can estimate model parameters and customer happiness in our setting. In Section 9, we enhance the connection with the marketing literature by arguing that our main insights extend to a non-contractual setting, where the customer visits with happiness-dependent interarrival times.

It is worthwhile to contrast our results with Radner and Shepp [23], a paper that uses somewhat similar mathematics to study optimal dividend control given bankruptcy risk. Radner and Shepp find that firms should use increasingly risky policies (with increasing reward rate) the more cash they have on hand. In contrast, we find that optimal policies are oftentimes sandwich polices where risk-taking is a non-monotonic function of the state. A key driver of this difference is that unsatisfied customers in our model quit according to a hazard rate while, in Radner and Shepp, a firm without cash goes bankrupt immediately. Based on our findings, we believe that if bankruptcy were probabilistic in Radner and Shepp (customers could survive for some time with negative assets) rather than deterministic, the optimal control policy would be non-monotonic.

3 The Base Model

We study a continuous-time model of one firm repeatedly interacting with one customer. The firm has two alternative service modes that determine the drift and volatility of a reward process. The first is a Risky mode $R$, identified by drift and volatility parameters $(\mu_R, \sigma_R)$, and the second one is a Safe mode $S$, identified by $(\mu_S, \sigma_S)$, where $\sigma_S = 0$. At any given time $t \geq 0$, the firm chooses
service mode $u_t \in \{R, S\}$ which determines the drift $\mu_{u_t}$ and volatility $\sigma_{u_t}$ of the reward process. The firm is able to switch between two service modes over time. The reward $Y_t$ accrues according to the following stochastic differential equation:

$$dY_t = \mu_{u_t} dt + \sigma_{u_t} dB_t,$$

where $B_t$ is a standard Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$ with $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ and $Y_0 = 0$. We are interested in cases where the expected rewards are positive. Therefore, we assume $\mu_R > 0$ and $\mu_S > 0$. We call the service mode with higher drift the (myopically) superior mode. Namely, the Risky mode is the superior mode if $\mu_R > \mu_S$ and the Safe mode is the superior mode if $\mu_R < \mu_S$. Likewise, we call the service mode with lower drift the (myopically) inferior mode.

We think of the rewards as accruing to both the customer and the firm, in the sense that both players seek higher rewards. The firm is rational and aims to maximize the total reward generated from its interaction with the customer over time. We give a precise description of the firm’s objective in Eq. (8) below. The customer, meanwhile, is assumed to be unsophisticated in the sense that he adaptively learns from past experience but is subject to recency bias (as in a goodwill model, à la Nerlove and Arrow [19] and Aflaki and Popescu [2]). We describe the customer at each point in time via $H_t$, a one-dimensional state that we term happiness. To model recency bias, we model customer happiness as an exponentially weighted moving average of recent rewards. That is, the customer happiness follows a stochastic differential equation:

$$dH_t = dY_t - H_t dt,$$

where $H_0 = x$ is the initial customer happiness. This implies that the customer happiness at time $t$ is equal to

$$H_t = x \cdot e^{-t} + \int_0^t e^{-(t-s)} dY_s.$$  

Thus, customer happiness would follow an Ornstein–Uhlenbeck (O-U) process if the firm were to select the same service mode over the entire time horizon (for the Safe mode, the O-U process is degenerate, since $\sigma_S = 0$). The value $H_t$ can be interpreted as the customer’s unsophisticated prediction at time $t$ of the reward rate he expects to get in the future, and Eq. (2) captures that the customer iteratively modifies this happiness by the difference between the realized reward $dY_t$ and the predicted reward $H_t dt$. There is no private information in our model, so the firm is able to observe the history of customer happiness and rewards. Note that $\mathcal{F}_t$ contains all information of the reward process $Y_t$ and the happiness process $H_t$ up to time $t$, since $(u_t)_{t \geq 0}$ and $(B_t)_{t \geq 0}$ are both adapted to $\mathbb{F}$.

The firm cares about customer happiness because the customer is likely to quit the system if he is unhappy. We model customer abandonment via a hazard rate function $Q(\cdot)$ which acts on the

---

4If $\mu_R = \mu_S$, both service modes classify as superior. Section 4 provides discussion for this boundary case.
current happiness. In the Base model, we assume there is a satisfaction threshold \( q \) (or happiness threshold) such that the customer does not quit the system when her happiness is above or equal to it. We say the customer is satisfied when her happiness is above or equal to \( q \), borderline satisfied when her happiness is exactly equal to \( q \), and unsatisfied when her happiness is below \( q \). Specifically, we assume the hazard rate \( Q(\cdot) \) takes the functional form of an indicator function; that is

\[
Q(H_t) \equiv 1\{H_t < q\}. \tag{4}
\]

The literature on managing customer churn has considered different hazard rate functions, such as logit functions (Rust et. al. [27]) and exponential functions (Berger and Nasr [5]). The step function in Eq. (4) can be thought of as the limit of logit functions and is chosen for the sake of tractability. In Section 8, we numerically demonstrate robustness of our results to different hazard rate functions.

We denote the customer lifetime by \( T \). The hazard rate assumption implies that the customer survival probability \( S_t \) at time \( t \) is equal to

\[
S_t \equiv P(T > t \mid \mathcal{F}_t) = e^{-\int_0^t Q(H_s)ds}. \tag{5}
\]

More formally, we let \( z \) be a uniform random variable over \([0,1]\) independent of \( \mathcal{F} \) and let the lifetime \( T \) be defined as

\[
T \equiv \inf \left\{ t \geq 0 : e^{-\int_0^t Q(H_s)ds} = z \right\}. \tag{6}
\]

Here, \( \int_0^t Q(H_s)ds \) is a Lebesgue integral that measures the amount of time up to \( t \) that the customer happiness process has spent in the unsatisfied zone. In order to ensure that the customer lifetime is finite\(^6\), we assume that the threshold \( q \) is greater than the drift of the Safe mode, that is, \( q > \mu_S \).

We now formally describe the firm’s problem. An admissible policy defines a mapping from \( \mathcal{F}_t \) to the action space \( \{R, S\} \) for all \( t \leq T \). Formally, a policy \( \pi \) is admissible if the firm’s action process \( (u_t)_{t \geq 0} \) (by following this policy) is adapted to the filtration \( \mathcal{F} \), takes value in \( \{R, S\} \), and is such that the stochastic process \( (H_t)_{t \geq 0} \) is an \( \mathcal{F} \)-adapted semimartingale specified uniquely in law.\(^7\) Below, we show a large class of admissible policies.

We denote the space of admissible policies by \( \Pi \). For a given starting happiness \( x \) and a given admissible policy \( \pi \in \Pi \), the happiness process \((H_t^x,\pi)_{t \geq 0}\) is uniquely specified in law by Eqs.

\[^5\]Formally, from here on we consider the product probability space \((\Omega, \mathcal{F}, P) \otimes ([0,1], \mathcal{B}, U)\), where \( \mathcal{B} \) is the \( \sigma \)-algebra of Borel sets of \([0,1]\), \( U \) denotes the uniform probability measure on \([0,1]\), and \( z \in [0,1] \) is the “state of the world” in the second component probability space.

\[^6\]Proposition 1 in Section 4 formalizes that the expected lifetime is finite under a big class of admissible policies. Also in Theorem 1 (in Section 4) we establish that there exists an optimal policy in that big admissible policy class. This also means that no other admissible policy can produce a larger customer lifetime value. Therefore the expected lifetime value as well as the expected lifetime must be finite for any admissible policy.

\[^7\]Note that the happiness process \((H_t)_{t \geq 0}\) uniquely determines the reward process \((Y_t)_{t \geq 0}\) and survival probability \((S_t)_{t \geq 0}\), as per \( Y_t = H_t - H_0 + \int_0^t H_s ds \) and \( S_t = e^{-\int_0^t Q(H_s)ds} \).
(2) and (1). Similarly, for a given initial happiness $x$ and a given admissible policy $\pi \in \Pi$, the abandonment time $T^{x,\pi}$ is a properly defined random variable, as specified by Eq. (6). When clear from context, we suppress the superscript and denote these quantities simply by $Y_t$, $H_t$, $S_t$, and $T$. The firm’s objective is to maximize the customer lifetime value (CLV) it earns from interacting with the customer. For a given starting happiness value $x$ and policy $\pi$, the CLV is equal to

$$V(x, \pi) = \mathbb{E} \left[ \int_0^\infty 1\{t < T\}dY_t \bigg| H_0 = x \right]$$

$$= \mathbb{E} \left[ \int_0^\infty \mu \pi(H_t) 1\{t < T\}dt + \int_0^\infty \sigma \pi(H_t) 1\{t < T\}dB_t \bigg| H_0 = x \right]$$

(7)

where we used (1). The firm wants to find a policy that maximizes the CLV. The optimal CLV given a starting happiness value $x$ is given by

$$V^*(x) = \sup_{\pi \in \Pi} V(x, \pi).$$

(8)

This completes our definition of the firm’s problem. (Note that in our definition of admissible policies, the policy does not have access to $z$, and hence does not know whether or not the customer has already left. This assumption clearly has no impact on the CLV since rewards stop accruing after $T$, and is made for technical convenience.)

We say a policy is stationary Markov if it is a time-invariant mapping from the current happiness state to the service mode $\pi : \mathbb{R} \rightarrow \{R, S\}$. Given that the platform is facing a stationary Markov control problem, one would expect there to be a stationary Markov optimal control policy, and indeed in Section 3 we will formally establish that this is the case. Note that under a stationary Markov policy $\pi$, we can write $u_t = \pi(H_t)$, and the CLV in Eq. (7) can be restated as

$$V(x, \pi) = \mathbb{E} \left[ \int_0^\infty \mu \pi(H_t)e^{-\int_0^t \lambda(H_s)ds}dt \bigg| H_0 = x \right].$$

(9)

Observe from Eq. (9) that the expectation is with respect to the probability space $(\Omega, F, F, P)$ only, and the customer lifetime $T$ does not play any role here. In Section 9, we will use Eq. (9) to informally argue that our model also captures an alternate “non-contractual” setting where the customer never leaves, but rather interacts with the firm less frequently if they are unhappy.

**Formal specification of our stochastic processes.** Our setting presents some atypical technical features as a result of our assumption that the volatility under the Safe mode $\sigma_S$ is zero.

---

8The expectation in Eq. (7) is with respect to the product probability space $(\Omega, F, F, P) \otimes ([0,1], B, U)$.

9The second integral in Eq. (7) has zero expectation since the integrand is bounded. By the Dominated Convergence Theorem and then the tower property, we can write $V(x, \pi)$ as

$$V(x, \pi) = \lim_{t \to \infty} \mathbb{E} \left[ \int_0^t \mu \pi(H_s) 1\{s < T\}ds \bigg| H_0 = x \right] = \lim_{t \to \infty} \mathbb{E} \left[ \mathbb{E} \left[ \int_0^t \mu \pi(H_s) 1\{s < T\}ds \bigg| F_t \right] \bigg| H_0 = x \right] = \lim_{t \to \infty} \mathbb{E} \left[ \int_0^t \mu \pi(H_s)P(1\{s < T\} | F_s)ds \bigg| H_0 = x \right] = \lim_{t \to \infty} \mathbb{E} \left[ \int_0^t \mu \pi(H_s)e^{-\int_0^t \lambda(H_s)ds}ds \bigg| H_0 = x \right].$$

The last step follows from Eq. (5).
However, we are able to formally specify the stochastic processes for reward and happiness resulting from a broad set of control policies by drawing on the work of Salins and Spiliopoulos [28] on Markov processes with spatial delay. In Lemma 1 below we establish admissibility for this set of control policies, which we call interval policies. First we give the definition of interval policies.

**Definition 1** (Interval policy). A policy $\pi$ is an interval policy if it is stationary Markov (that is, the corresponding action process is a function of current happiness $u_t = \pi(H_t)$) such that it divides the happiness real line into a countable number of alternating Risky and Safe intervals, and such that the Safe mode is adopted at each boundary point between intervals.

Recall the definition of admissible policies earlier in this section. The firm’s action process under an interval policy as defined in Definition 1 is clearly adapted to the filtration $\mathcal{F}$ and takes value in $\{R, S\}$. If we can show the corresponding happiness process is an $\mathcal{F}$-adapted semimartingale specified uniquely in law, we can conclude an interval policy is admissible. The next lemma formally establishes this result, and further deduces that we can apply the Itô-Tanaka formula to the happiness process $(H_t)_{t \geq 0}$ under an interval policy.

**Lemma 1.** If $\pi$ is an interval policy, then for any starting happiness $x$, the process $H_t^{x, \pi}$ defined by Eq. (2) is an $\mathcal{F}$-adapted semimartingale specified uniquely in law. Hence interval policies are admissible. Let $(u_t)_{t \geq 0}$ be the corresponding action process. Assume $f$ is a function that is continuously differentiable on $\mathbb{R}$ and twice continuously differentiable on $\mathbb{R} \setminus \mathcal{E}$ for some countable set $\mathcal{E}$. Then $f(H_t)$ is also an $\mathcal{F}$-adapted semimartingale, with

$$f(H_t) = f(x) + \int_0^t (\mu_{u_s} - H_s) f'(H_s)ds + \int_0^t \sigma_{u_s} f'(H_s)dB_s + \frac{1}{2} \int_0^t \sigma_{u_s}^2 f''(H_s)1_{\{H_s \notin \mathcal{E}\}} ds.$$  (10)

The proof of Lemma 1 is in Appendix A. In fact, the Itô-Tanaka formula Eq. (10) holds for the happiness process under any admissible policy, since by definition of admissibility $H_t$ is a semimartingale.

**Discussion of model assumptions.** We briefly discuss specific assumptions in our model. Our model assumes that customer happiness is an exponentially weighted moving average of recent rewards with the window size parameter $w = 1$; that is, rewards from $\Delta t$ time ago are given a weight $(1/w) \exp(-\Delta t/w)$ for $w = 1$ in determining happiness. If instead a customer uses an exponentially weighted moving average with arbitrary window size $w \in (0, \infty)$, then we can rescale time to ensure that Eq. (2) still holds, so our assumption $w = 1$ is without loss of generality. However, having fixed the scaling of time thus, we notice that Eq. (4) represents a specific assumption that the size of the step in the hazard rate function is 1. This assumption enables us to obtain analytical results for our stochastic control problem, but as we demonstrate via numerics in Section 8, our structural insights hold for other step size values as well, and in fact also for other hazard rate functions.

---

10We define $f''(H_s)1_{\{H_s \notin \mathcal{E}\}}$ to be zero if $H_t \in \mathcal{E}$.
4 Analysis of the Base Model

In this section, we solve the firm’s optimization problem posed in Eq. (8) and describe the structure of an optimal policy. We find that there is an optimal policy that achieves the optimal CLV that is a simple interval policy. In particular, the optimal policy will be an interval policy that consists of at most three intervals.

The next proposition establishes that under interval policies, the expected customer lifetime and the expected customer lifetime value (CLV) are both finite.\textsuperscript{11}

**Proposition 1.** For any $x \in \mathbb{R}$ and any interval policy $\pi$ with a finite number of intervals, both the customer lifetime and the CLV are finite, i.e., $E[T^{x,\pi}] < \infty$ and $V(x, \pi) < \infty$.

To prove Proposition 1, we just need to show that the happiness process eventually will spend enough time in the unsatisfied zone (below $q$), where the customer churns at hazard rate 1. This is true since the happiness process $H_t$ drifts toward $\mu_S < q$ deterministically under the Safe mode, and moves stochastically over the entire real line under the Risky mode. The full proof is in Appendix C.

The rest of this section is organized as follows. In Section 4.1, we present the structure of an optimal policy. In Section 4.2, we give an overview of our proof of the optimality of the policy. In Section 4.3, we state some formal comparative statics results and use numerics to understand both the structure of the optimal policy (including comparative statics) as well as the CLV increase from the optimal policy relative to the myopic policy. We find that the benefit can be very large.

4.1 Structure of an Optimal Policy

In our main theorem of this section, we prove that there exists an optimal policy where the firm should always use the myopically superior service mode except (possibly) for some intermediate customer happiness levels. In fact, we show that regardless of which service mode is the superior one, the firm should choose the Safe service mode if the customer happiness level is in some region (possibly empty) $[q, \theta_b]$ just above the happiness threshold $q$, and choose the Risky service mode if the customer happiness level is in some nonempty region $(\theta_G, q)$ just below $q$. In particular, the values $\theta_b$ and $\theta_G$ are defined in Lemma 2 and Lemma 3 below. We name the happiness region $[q, \theta_b]$ where the firm should always use the Safe mode the risk-averse region. Likewise, we name $(\theta_G, q)$ where the firm should always use the Risky mode the risk-seeking region.

**Lemma 2.** If $\mu_R \leq \mu_S$, let $\theta_b = \infty$. If $\mu_R > \mu_S$, let $\Theta$ be the set of values of $\theta$ that satisfy

$$\frac{\mu_R \sqrt{\pi} (\mu_S - \theta)}{\sigma_R} e^{\frac{(\theta - \mu_R)^2}{\sigma_R^2}} \left(1 - \text{erf} \left( \frac{\theta - \mu_R}{\sigma_R} \right) \right) + \mu_S = 0.$$

\textsuperscript{11}In fact, if \textit{any} stationary Markov policy (not necessarily an interval policy) is admissible, the associated CLV is finite. This follows from our main result (Theorem 1 below) which says that an interval policy is optimal, since we do not restrict admissible policies to be interval policies.
Then, the set \( \Theta \cap (\mu_S, \infty) \) contains a single element, which we label \( \theta_b \).

From Lemma 2, note that the risk-averse region \([q, \theta_b]\) is bounded only if the Risky mode is superior \( \mu_R > \mu_S \). Moreover, in a subset of cases where it is bounded, the risk-averse region is in fact empty; this occurs if \( \theta_b < q \).

**Lemma 3.** If \( \mu_R \geq \mu_S \), let \( \theta_G = -\infty \). If \( \mu_R < \mu_S \), let \( \Theta_G^{\text{small}} \) be the set of values of \( \theta \) that satisfy

\[
e^{-\frac{(q-\mu_R)^2}{\sigma_R^2}} \sqrt{\pi} (\theta - \mu_R)(q - \mu_R) \left( \text{erf} \left( \frac{q - \mu_R}{\sigma_R} \right) - \text{erf} \left( \frac{\theta - \mu_R}{\sigma_R} \right) \right)
+ e^{-\frac{(q-\mu_R)^2}{\sigma_R^2}} \frac{q - \mu_R}{\sigma_R} - \frac{\theta - \mu_R}{\sigma_R} = 0,
\]

and let \( \Theta_G^{\text{big}} \) be the set of values of \( \theta \) that satisfy

\[
2e^{-\frac{(q-\mu_R)^2}{\sigma_R^2}} \left( 1 + \frac{(q - \mu_R)(\theta - \mu_R)}{\sigma_R^2} \right)
+ e^{-\frac{(q-\mu_R)^2}{\sigma_R^2}} \sqrt{\pi} \left( \frac{\mu_S + 2\mu_R - 3\theta_R}{\sigma_R} + \frac{2(\theta - \mu_R)(\mu_S - \theta)}{\sigma_R^2} \right) \left( \text{erf} \left( \frac{q - \mu_R}{\sigma_R} \right) - \text{erf} \left( \frac{\theta - \mu_R}{\sigma_R} \right) \right)
+ \left( \frac{\mu_S + 2\mu_R - 3\theta_R}{\sigma_R} + \frac{2(\theta - \mu_R)(\mu_S - \theta)}{\sigma_R^2} \right) \frac{\mu_S}{q - \mu_R} \frac{\sigma_R^2 + (\mu_S - \mu_R)(\theta - \mu_R)}{(\mu_S - \mu_R)(q - \mu_R)(q - \mu_S)} = 0.
\]

Then, the set \( \left( \Theta_G^{\text{big}} \cap (-\infty, \mu_S) \right) \cup \left( \Theta_G^{\text{small}} \cap [\mu_S, q] \right) \) contains a single element, which we label \( \theta_G \).

Observe from Lemma 3 that the risk-seeking region \((\theta_G, q)\) is always nonempty. It is bounded only when the Safe mode is the superior mode \( \mu_S > \mu_R \). In fact, for any values of \( \mu_S < q \) and \( \mu_R \neq \mu_S \), exactly one of the two regions \([q, \theta_b]\) and \((\theta_G, q)\) is bounded, while the other is unbounded. The bounded region corresponds to the intermediate happiness level where the firm should use the inferior mode. Everywhere else (including the other, unbounded region) the firm should use the myopically superior service mode. In the special case of \( \mu_R = \mu_S \), both the risk-averse region and the risk-seeking region are unbounded, and they partition the entire happiness real line. We postpone a description of how we arrive at the above \( \theta_b \) and \( \theta_G \) until just before Subsection 4.2.

We are now ready to formally state our first theorem.

**Theorem 1.** Fix \( \mu_R, \mu_S, \sigma_R \) and satisfaction threshold \( q \) (recall Eq. (4)) such that \( \mu_S < q \). Consider the firm’s problem as presented in Eq. (8). Let \( \theta_b \) be as defined in Lemma 2 and let \( \theta_G \) be as defined in Lemma 3. Then there is an optimal policy where the firm chooses the Safe mode on \([q, \theta_b]\), the Risky mode on \((\theta_G, q)\), and the superior of the two modes elsewhere.\(^\text{12}\) Also, the customer’s expected lifetime and the CLV are finite under the optimal policy (hence the CLV is finite under any policy).

Consider the optimal policy described in Theorem 1. In this policy, the firm chooses the Safe mode on \([q, \theta_b]\), the Risky mode on \((\theta_G, q)\), and the superior mode elsewhere (see Figure 1 for a

\(^{12}\text{Note that if } \mu_R = \mu_S \text{ then } \theta_G = -\infty \text{ and } \theta_b = +\infty, \text{ so the specified policy uses Risky on } (-\infty, q) \text{ and Safe on } [q, \infty). \text{ In } \mu_R \neq \mu_S \text{ then there is a unique superior mode. Hence the policy is uniquely specified in all cases.} \)
Figure 1: Stylized representation of the firm’s optimal policy as per Theorem 1.

Now let us take a closer look at this policy under different cases of $\mu_R$ and $\mu_S$. When the Risky service mode is the superior one ($\mu_R > \mu_S$), the optimal policy uses the Risky mode for all happiness values below $q$. Moreover, depending on the value of $\theta_b$ (see Lemma 2), the risk-averse region $[q, \theta_b]$ might degenerate to an empty set. This happens if $\theta_b < q$, in which case the optimal policy becomes a myopic one where the firm always uses the Risky service mode (see Figure 2a). On the other hand when $\theta_b \geq q$, the risk-averse region $[q, \theta_b]$ is nonempty and bounded, and the optimal policy is a sandwich policy, i.e., an interval policy with three exactly intervals, where the firm uses the Safe mode for happiness values in $[q, \theta_b]$ just above the happiness threshold (see Figure 2b), and the Risky mode elsewhere. In the other case where the Safe mode is the superior one ($\mu_R < \mu_S$), the optimal policy uses the Safe mode for all happiness values (weakly) above $q$. Meanwhile, the risk-seeking region $(\theta_G, q)$ is always nonempty and bounded (see Lemma 3). Therefore in this case the firm’s optimal policy is always a sandwich policy (see Figure 3), where the Safe mode is used everywhere except for happiness levels in $(\theta_G, q)$ just below the happiness threshold. When customer happiness value is in this region, the firm should switch to the Risky service mode.

Figure 2: Optimal policy when $\mu_R > \mu_S$.

Figure 3: Optimal policy when $\mu_R < \mu_S$. 
In the special case of \( \mu_R = \mu_S \), we have \( \theta_b = \infty \) and \( \theta_G = -\infty \). In this case both the Risky mode and the Safe mode are superior, and the optimal policy becomes a particular myopic one, where the firm uses the Risky mode for happiness values on \((-\infty, q)\) and the Safe mode on \([q, \infty)\).

**Trade off between immediate payoffs and customer lifetime.** Observe that except for the situation depicted in Figure 2a, in all other cases, the firm’s optimal policy exhibits a risk-averse region just above the happiness threshold. In this region, the firm should use the Safe mode even if it generates lower immediate rewards. If the optimal policy is sacrificing rewards in the short run, it must be that it confers some long-term benefits. The intuition is that when the happiness level is close to the unsatisfied zone, the Safe mode prolongs customer lifetime by delaying entry into the unsatisfied zone, compared with the Risky mode. When the volatility \( \sigma_R \) is high, the mean first passage time into the unsatisfied zone from above will be longer under the Safe mode than under the Risky mode. Therefore, using the Safe mode in the risk-averse region just above \( q \) serves to delay the inevitable — entry into the unsatisfied zone — when the volatility of the Risky mode is high.\(^{13}\) A preference for low volatility in relatively low states is familiar from other settings, for example Radner and Shepp [23].

Perhaps more unexpected is the existence of a risk-seeking region. In particular, surprisingly, no matter by how much the Risky mode is inferior to the Safe mode, there is some happiness region below the happiness threshold \( q \) in which the firm should use the Risky mode. This challenges the customary idea that a high volatility and low return combination should be dominated by a low volatility and high return one. In fact, it turns out that the Risky mode outperforms the Safe mode in that region via its ability to push the customer out of the unsatisfied zone (with positive probability) when his happiness is below but close to the threshold \( q \), thereby extending his lifetime. We highlight that the risk-seeking region is always non-empty, regardless of model primitives.

An immediate consequence of the sandwich policy featuring both a risk-averse region just above the happiness threshold and a risk-seeking region just below it is that, surprisingly, once it enters the unsatisfied zone, the customer happiness will never again reemerge above the happiness threshold \( q \) (see Figure 4). Recall that by assumption, \( \mu_S < q \), so that switching to the Safe mode when the happiness process hits \( q \) from below leads to negative drift \( dH_t/dt = \mu_S - q \), and so the happiness process immediately drops back into the unsatisfied zone. This might seem counterintuitive. After all, we want the customer to stay satisfied with the firm, to prolong his lifetime. However, the optimal (sandwich) policy seems to advise us to keep his happiness low and prevent him from being satisfied. How can such a policy maximize customer lifetime value? Remember that the myopic policy always produces the maximum payoff rate. Therefore we can deduce that the reason for a non-myopic policy to be optimal must be that the customer lives longer under this policy. In other words, the customer spends more time in the satisfied zone under this policy. Counterintuitively, the optimal sandwich policy indeed increases the customer’s time spent in the satisfied zone even

\(^{13}\)In Theorem 2 later, we show that \( \theta_b \) is increasing in \( \sigma_R \), holding other primitives fixed.
after he enters the unsatisfied zone. This occurs because it causes the customer happiness to spend a *positive measure of time exactly at the borderline satisfied level of* $q$.

![Sample path of the happiness process under the optimal sandwich policy](image)

Figure 4: Sample path of the happiness process under the optimal sandwich policy: $\mu_S = 8$, $\mu_R = 9$, $\sigma_R = 10$, $q = 10$. Here, the risk-averse region where the optimal sandwich policy uses the Safe mode is $[10, 22.1]$.

**Comparison to a reflected Ornstein-Uhlenbeck process.** As in Section 3, let us consider the happiness process as an infinite horizon stochastic process. That is, let us ignore the customer abandonment at time $T^{x,\pi}$ and continue to track the evolution of $H^{x,\pi}_t$ under policy $\pi$. Assume that the optimal policy is a sandwich policy and consider a starting happiness level $x \leq q$, so that the happiness level never exceeds $q$ for any time $t$. At first sight, the happiness process appears to be the well-understood reflected O-U process (Reed et al. [24]). However, this is not the case. The reflected O-U process and the reflected Brownian motion both spend a measure zero of time at the reflecting boundary. In our process, the time spent at the reflecting boundary has a positive measure with probability one. While this process is less well known, such a *delayed reflected* process was first introduced by Skorokhod [30] and was recently shown to be a semimartingale by Salins and Spiliopoulos [28]. The measure of time spent by a delayed reflected process at a reflection boundary is proportional to the local time of the process at the boundary, and the constant of proportionality is the inverse of the drift at the boundary. In our setting, the drift is $-(q - \mu_S)$ under the Safe mode at the threshold happiness $q$. The standard reflected O-U process corresponds to a negative infinity drift at the reflecting boundary.

**The switching thresholds of the sandwich policy.** We now turn our attention to the values of $\theta_b$ and $\theta_G$, as defined in Lemmas 2 and 3. We provide here a very informal argument to explain
how these values arise. We present the key ideas used to build a formal proof in the next subsection, with the details of the argument deferred to Appendix C.

Suppose $\mu_R > \mu_S$. The value of $\theta_b$ can be established informally by comparing the value functions of two policies. Intuitively, the optimal policy should choose the Risky mode for a sufficiently high levels of happiness, given that the Risky mode is myopically superior. We want to find the happiness level $\theta_b > q$ below which the firm should switch to the Safe mode. Let $x$ be some starting happiness level and let $\Delta x$ be some value such that $x - \Delta x \geq q$. Consider two stationary policies, $\pi_S$ and $\pi_R$, that both choose the Risky mode above $x$ and make identical decisions for values of happiness below $x - \Delta x$ (for our argument, it does not matter what decisions the policies make below $x - \Delta x$; we are only concerned with the expected reward accumulated in the period it takes for the happiness to drop from $x$ to $x - \Delta x$). Within the interval $[x - \Delta x, x]$, the two policies choose different modes, with $\pi_S$ choosing the Safe mode and $\pi_R$ choosing the Risky mode.

We now compare the marginal benefits of these two policies, which we define according to

\[
m(x, \pi) = \lim_{\Delta x \to 0} \frac{V(x, \pi) - V(x - \Delta x, \pi)}{\Delta x}.
\]

The quantity $m(x, \pi)$ establishes how much differential value the policy $\pi$ generates from happiness being $x$ rather than $x - \Delta x$, which is equal to the reward (per unit $\Delta x$) accumulated by the policy while the happiness falls from $x$ to $x - \Delta x$. This is fairly easy to compute: the numerator is simply the product of the expected reward rate and the expected first passage time of an O-U process from $x$ to $x - \Delta x$. Note that the happiness reaches $x - \Delta x$ in finite time with probability 1, and the customer does not depart before this occurs. If there exists a level of happiness $x \geq q$ such that $m(x, \pi_S) \geq m(x, \pi_R)$, then at happiness $x$, the firm weakly prefers the Safe action, assuming it uses the Risky action above it. In fact, $\theta_b$ can be determined by finding $x$ such that the marginal benefits of these two policies are equal; that is

\[
m(\theta_b, \pi_S) = m(\theta_b, \pi_R).
\]

Lemma 2 encodes this condition in terms of model primitives and tells us that it uniquely specifies $\theta_b$.

Similarly, to derive the value of $\theta_G$, the conditions in Eqs. (11) and (12) come from guessing a sandwich policy that uses the Risky mode in $(\theta_G, q)$ and the Safe mode elsewhere and equating the marginal value of happiness (i.e., the slope of the value function) just to the left and right of $\theta_G$. As evidenced by the complexity of the definition of $\theta_G$, matching marginal values in the unsatisfied zone is a harder task here than in the satisfied zone. The extra difficulty arises from having to account for the customer abandonment risk while computing first passage times.

### 4.2 Key Ideas Behind the Proof of Theorem 1

Theorem 1 is a statement about the optimal control policy for our stochastic control problem. Establishing the optimal control is essentially equivalent to determining the value function $V^*$.
(see Eq. (8)). A classical technique for determining a continuous-time, continuous-space value function $V^*$ such as ours is called the verification technique. The verification technique involves first obtaining a candidate value function and subsequently proving its optimality. There are many papers in the literature that use the verification technique for solving stochastic control problems, including Borkar (1989), Mirică (1992), Radner and Shepp (1996), and Touzi (2002). We cannot immediately use a method from the literature for two reasons. First, our problem involves stochastic abandonment rather than deterministic discounting. Second, as will become clear, the value function $V^*$ in our problem does not satisfy the standard smoothness condition, which is that $V^*$ should be twice continuously differentiable everywhere. We modify the standard approach from the literature in order to create a methodology that works for our problem. In addition, the verification step turns out to be extremely challenging, and we make innovative use of L'Hopital style rules for monotonicity [22] along with a tour de force analysis to accomplish it.

To start with, the HJB equation of the Base model (formally derived in the proof of Proposition 2, see Appendix C) yields the following ordinary differential equation (ODE):

$$\max_{i=S,R} \left\{ -Q(x)V(x) + (\mu_i - x)V'(x) + \frac{1}{2}\sigma_i^2 V''(x) + \mu_i \right\} = 0 \quad (13)$$

for all $x \in \mathbb{R}$ where $V''$ exists.

We cannot deduce that Eq. (13) has a twice continuously differentiable solution, since $Q(x)$ is discontinuous and the second order term can be zero. In fact, the value function of this problem is never twice continuously differentiable at $q$. We shall conclude, by the end of this section, that $V^*$ is continuously differentiable everywhere on $\mathbb{R}$, and twice continuously differentiable everywhere on $\mathbb{R} \setminus \mathcal{E}$, with $\mathcal{E}$ containing the switching boundaries and $q$.

The following proposition provides conditions for optimality.

**Proposition 2.** Suppose a function $\bar{V} : \mathbb{R} \to \mathbb{R}$ satisfies

1. the function $\bar{V}$ is non-negative;
2. the function $\bar{V}$ is continuously differentiable on $\mathbb{R}$ and twice continuously differentiable on $\mathbb{R} \setminus \mathcal{E}$ for some countable set $\mathcal{E}$;
3. the function $\bar{V}'$ is bounded;
4. for any $x \in \mathbb{R}$ for $i = S$ and for any $x \in \mathbb{R} \setminus \mathcal{E}$ for $i = R$, the following inequality holds$^{14}$

$$-Q(x)\bar{V}(x) + (\mu_i - x)\bar{V}'(x) + \frac{1}{2}\sigma_i^2 \bar{V}''(x) + \mu_i \leq 0; \quad (14)$$

5. for some interval policy $\bar{\pi}$ (see Definition 1) such that $\bar{\pi}(y) = S$ for all $y \in \mathcal{E}$, the process $\bar{V}(H_{t,x}^{\pi,\bar{\pi}})$ is an $\mathcal{F}$-adapted semimartingale, and for all $x \in \mathbb{R}$ it holds that

$$-Q(x)\bar{V}(x) + (\mu_{\bar{\pi}(x)} - x)\bar{V}'(x) + \frac{1}{2}\sigma_{\bar{\pi}(x)}^2 \bar{V}''(x) + \mu_{\bar{\pi}(x)} = 0. \quad (15)$$

$^{14}$For any $x$ including $x \in \mathcal{E}$, we define $\frac{1}{2}\sigma_S^2 \bar{V}''(x) \geq 0$ consistent with $\sigma_S = 0$. 

18
6. For all \( x \in \mathbb{R} \) and the policy \( \bar{\pi} \) in Condition 5,

\[
\lim_{t \to \infty} \mathbb{E} \left[ (1 + |H_t^{x, \bar{\pi}}|) e^{-\int_0^t Q(H_s^{x, \bar{\pi}}) ds} \right] = 0.
\]

Then, the function \( \bar{V} \) is the value function \( V^* \), and \( \bar{\pi} \) is an optimal policy.

This proposition tells us that if an interval policy \( \bar{\pi} \) together with a function \( \bar{V} \) satisfies the stated conditions, then \( \bar{\pi} \) is optimal among the class of all admissible policies \( \Pi \) (not just among interval policies), and \( \bar{V} = V^* \).

We now provide a short summary of the proof of Proposition 2; the proof is in Appendix C. Conditions 1–4 imply that the function \( \bar{V} \) is an upper bound of the optimal value function \( V^* \). This is established by constructing a stochastic process

\[
X_t^{x, \bar{\pi}} = \bar{V}(H_t^{x, \bar{\pi}}) e^{-\int_0^t Q(H_s^{x, \bar{\pi}}) ds} + \int_0^t e^{-\int_0^s Q(H_u^{x, \bar{\pi}}) du} dY_s^{x, \bar{\pi}},
\]

and showing \( \bar{V}(x) \geq \limsup_{t \to \infty} \mathbb{E} X_t^{x, \bar{\pi}} \geq V(x, \pi) \) for any admissible policy \( \pi \in \Pi \). Note that Condition 3 serves to ensure that the stochastic part of \( X_t^{x, \bar{\pi}} \) defined in (17) is a martingale and hence has zero expectation. Conditions 5 and 6 further imply that the bound is tight under the interval policy \( \bar{\pi} \), i.e., \( \bar{V} \) is the optimal value function and the optimal policy is \( \bar{\pi} \). In particular, we use Condition 5 to show that \( \bar{V}(x) = \limsup_{t \to \infty} \mathbb{E} X_t^{x, \bar{\pi}} \). We then use Condition 6 together with Condition 3 to deduce \( \lim_{t \to \infty} \bar{V}(H_t^{x, \bar{\pi}}) e^{-\int_0^t Q(H_s^{x, \bar{\pi}}) ds} = 0 \) and hence \( \limsup_{t \to \infty} \mathbb{E} X_t^{x, \bar{\pi}} = V(x, \bar{\pi}) \).

With Proposition 2 in place, our task is to find \( \bar{V} \) and \( \bar{\pi} \) satisfying conditions 1–6 there. The following lemma shows that Condition 6 is satisfied by any interval policy with a finite number of intervals. Note that the sandwich policy is an interval policy with three intervals.

**Lemma 4.** For any starting happiness level \( x \in \mathbb{R} \) and any interval policy with a finite number of intervals, we have

\[
\lim_{t \to \infty} \mathbb{E} \left[ (1 + |H_t^{x, \bar{\pi}}|) e^{-\int_0^t Q(H_s^{x, \bar{\pi}}) ds} \right] = 0.
\]

This lemma implies that the policy described in Theorem 1 satisfies Condition 6 in Proposition 2.

Next we find a candidate optimal value function by solving the HJB equation (13). Observe that Conditions 4–5 in Proposition 2 imply \( \bar{V} \) and \( \bar{\pi} \) solve the HJB equation (13). We obtain a candidate value function by the following informal reasoning. First of all, we conjecture that the superior mode should be used myopically for sufficiently large and sufficiently small happiness values. This is a straightforward conjecture since the only possible reason to use the inferior mode is to prolong customer lifetime, but this benefit (if there is such benefit) is minimal when the customer’s happiness value is far away from the happiness threshold \( q \). In particular, when the happiness state is far below, there is almost no chance to escape the unsatisfied zone before \( T \) is hit. On the other hand, when the happiness state is far above, it descends slower under the superior mode. Thus inspired, we start by solving Eq. (15) for a control policy that always uses the superior mode.
mode, while asking for continuity of $V'$ at $q$ and boundedness of $V'$ everywhere. Next, we check at which $x$ values the superior mode achieves the maximum in the LHS of the HJB equation (13). We then update our policy by retaining the superior mode at such happiness locations and replacing it with the inferior mode at the complementary locations where the inferior mode achieves the maximum. For the updated policy, we compute the updated $V$ by solving Eq. (15) again, while satisfying the boundary conditions — continuity of $V'$ as well as boundedness of $V'$ everywhere.

Having computed the updated $V$, we again update the policy, and so on. This policy iteration process leads to a monotone improving sequence of $V$ and associated policies, and it can be carried on until convergence. In our case, we find a final candidate value function and candidate policy via at most one update from the initial myopic policy. The candidate value function that we denote by $W^*$ is defined in the following proposition, where we also verify that it satisfies not only the HJB equation (13), but all conditions in Proposition 2.

**Proposition 3.** Let $W(x, C_1, C_2, C_3, C_4, C_5)$ be defined as

$$W(x, C_1, C_2, C_3, C_4, C_5) = \begin{cases} 
V_1(x, C_1) & \text{if } x \leq \theta_G; \\
V_2(x, C_2, C_3) & \text{if } \theta_G < x < q; \\
V_3(x, C_4) & \text{if } q \leq x \leq \theta_b; \\
V_4(x, C_5) & \text{if } x > \max\{q, \theta_b\},
\end{cases}$$

where $q$ is the happiness threshold (see Eq. (4)), $\theta_b$ is as defined in Lemma 2, $\theta_G$ is as defined in Lemma 3, and

$$V_1(x, C_1) = \frac{C_1}{x - \mu_S} + \mu_S;$$
$$V_2(x, C_2, C_3) = C_2 e^{\frac{(x-\mu_R)^2}{\sigma_R^2}} + C_3 \text{erf}\left(\frac{x-\mu_R}{\sigma_R}\right) e^{\frac{(x-\mu_R)^2}{\sigma_R^2}} + \mu_R;$$
$$V_3(x, C_4) = C_4 + \mu_S \log(x - \mu_S).$$
$$V_4(x, C_5) = C_5 + \int_{0}^{x} \frac{\mu_R \sqrt{\pi}}{\sigma_R} e^{-\frac{(x-\mu_R)^2}{\sigma_R^2}} (1 - \text{erf}\left(\frac{z-\mu_R}{\sigma_R}\right)) dz.$$

Then there exist $C_1^*, C_2^*, C_3^*, C_4^*, C_5^*$ (defined explicitly in the proof of this proposition) such that $W^*(\cdot) := W(\cdot, C_1^*, C_2^*, C_3^*, C_4^*, C_5^*)$ together with the policy specified in Theorem 1 satisfies all conditions in Proposition 2.

The proof of this proposition is in Appendix C. Note that verifying the conditions in Proposition 2 is a challenging task, the hardest component being to show non-positivity Eq. (14) of parametric functions involving four free parameters ($\mu_S, \mu_R, \sigma_R, q$) plus a variable $x$ representing the happiness level. We achieve this by decomposing the function of interest into different pieces and showing non-positivity for each piece, utilizing Chernoff-type bounds on the error function [8] and L'Hôpital-type rules for monotonicity [22] to facilitate our analysis. Appendix B.1 establishes supporting lemmas on the error function for us to complete the verification step.
4.3 Comparative Statics and Numerics

In the previous section, we constructed an optimal policy that is either the myopic policy (superior mode everywhere) or a sandwich policy (superior mode everywhere except for some intermediate happiness range). In this section, we explore how the optimal policy and the optimal CLV depend on model primitives. There are two results we highlight in this section: (1) the improvement in CLV under the optimal policy compared to the CLV under the myopic policy can be very large; (2) an increase in volatility of the Risky mode substantially increases the CLV in some cases.

We first look at some numerical examples regarding CLVs under the optimal policy versus under the myopic policy. Theorem 1 tells us that the myopic policy is not always optimal. Figure 5 below shows how much improvement in CLV the optimal policy provides over the myopic policy for $\mu_R = 5$, $q = 6$, initial happiness $q$ (as an example), and for different $\mu_S$ and $\sigma_R$. The figure shows that the magnitude of improvement can be very large, especially when the Safe mode is the superior mode. Also, as the Risky mode’s volatility increases, the improvement in CLV gets larger.

![Figure 5: The ratio of CLV under the optimal policy to CLV under the myopic policy (on a logarithmic scale) versus $\mu_S$, for $\mu_R = 5$, $q = 6$, and for initial happiness $q$.](image)

It is interesting to investigate how the optimal policy varies with problem primitives. Figure 6 below shows that both the risk-seeking region described by value $\theta_G$, and the risk-averse region by value $\theta_b$ vary monotonically in $\mu_R$, holding $\mu_S$, $\sigma_R$ and $q$ fixed. In the plot, the horizontal coordinate is the Risky mode’s drift $\mu_R$ (a problem primitive), and the vertical coordinate is the customer’s current happiness value (the “state” of the customer). For each $\mu_R$ (a fixed location on the horizontal axis), Theorem 1 gives us the optimal policy. The shading in the figure is used to represent when the firm should use the Risky mode under the optimal policy. Therefore, the boundaries between the shaded and unshaded areas are the switching thresholds between the two service modes. The top left unshaded block with a right-pointing tail corresponds to the risk-averse
region in the optimal policy. The bottom right shaded block with a left pointing tail corresponds to the risk-seeking region in the optimal policy. The plot shows that when $\mu_R$ increases, the size of the risk-averse region decreases and the size of the risk-seeking region increases. We further show the impact of $\sigma_R$ on the size of the sandwich in Figure 7. It is not surprising that as $\sigma_R$ increases, both the risk-averse region and the risk-seeking region get larger. This is because the firm wants to avoid volatility in the risk-averse region and seek it in the risk-seeking region.

Figure 6: The optimal sandwich policies for different values of $\mu_R$. Fix $\mu_S = 8$, $\sigma_R = 10$ and $q = 10$. The horizontal axis corresponds to the value of $\mu_R$, and the vertical line marks the happiness value. The two curves are the switching boundaries between the Risky mode and the Safe mode.

We formalize our monotonicity results in the next theorem. Again, to prove these results, we draw upon the L’Hospital-type rules for monotonicity [22] and establish supporting lemmas on the error function in Appendix B.1.

**Theorem 2.** Denote by $V^*$ the optimal value function under the Base model parameters. Consider $\theta_b$ and $\theta_G$ defined in Lemmas 2 and 3 respectively. Then if the Risky mode is superior, i.e., $\mu_R > \mu_S$,

1. the value $\theta_b$ is strictly increasing in $\mu_S$, and for any $x \in \mathbb{R}$, $V^*$ is increasing in $\mu_S$;
2. the value $\theta_b$ is strictly decreasing in $\mu_R$, and for any $x \in \mathbb{R}$, $V^*$ is increasing in $\mu_R$;
3. the value $\theta_b$ is strictly increasing in $\sigma_R$ and, for any $x > q$, we have that $\frac{\partial V^*}{\partial x}$ is weakly decreasing in $\sigma_R$.

On the other hand, when the Safe mode is superior, i.e., $\mu_S > \mu_R$,

4. the value $\theta_G$ is decreasing in $\sigma_R$.

Parts 1 and 2 of Theorem 2 show that a higher mean reward rate makes the corresponding service mode more attractive (so the risk-averse region, where the Safe mode is used, is larger if
Figure 7: The optimal switching curves for different model primitives. Fix $\mu_S = 8$, $q = 10$. The horizontal axis corresponds to the value of $\mu_R$, and the vertical line marks the happiness value.

$\mu_S$ is larger, and smaller if $\mu_R$ is larger) and improves the overall CLV. Part 3 and 4 show that the greater the volatility of the Risky mode, the larger the size of both the risk-averse region and the risk-seeking region in the optimal sandwich policy. In part 3 we also show that the marginal value of customer happiness is lower when the Risky mode is more volatile. This is due to the fact that the mean first passage time to the unsatisfied zone is smaller when the Risky mode is more volatile.

So far we characterized the optimal policy — either the myopic policy or a sandwich policy that contains an intermediate region when the firm utilizes the inferior mode — for the Base model, under assumptions that 1) the customer’s hazard rate of churn is a step function of step size one; 2) the reward accumulates as a brownian motion; 3) the firm only utilizes exactly one of the two service modes at any time; 4) there is no switching cost for switching between service modes. Under these assumptions we were able to obtain a full characterization of the optimal policy and its dependence on problem primitives. However, returning to the investment manager’s problem, we notice that some of the above assumptions are not satisfied. For example, 1) the hazard rate of churn is expected to be some monotone decreasing function of happiness, but not our simple step function; 2) the reward from investing in a financial asset should accumulate as a geometric brownian motion (GBM); 3) the manager can, in fact, mix his investment between the Risky asset and the Safe asset; 4) there should be a switching cost every time the manager adjusts his portfolio.

In the next sections, we incorporate these features (one at a time), and we show that our main findings are robust to each of them. In particular, in the following section, we prove formally that when the manager is allowed to mix between two service modes, the optimal policy structure is still qualitatively the same. Moreover, we explicitly characterize the firm’s optimal mixed portfolio as a function of the customer’s happiness state.
5 The Investor Model

In the previous sections, we studied a Base model where the action space at any time is binary. In this section, we consider a variant of the Base model where the firm is allowed to mix between the two service modes. That is, at each point in time $t \geq 0$, the firm chooses the proportion of the Risky mode $p_t \in [0, 1]$ in its service, so that the reward is generated according to

$$dY_t = ((1 - p_t)\mu_S + p_t\mu_R)dt + p_t\sigma_R dB_t.$$

We restrict attention to $\mu_R > \mu_S > 0$. This is motivated by the real world application of an investment manager, where he can mix between assets and where the riskier asset usually provides better rewards. We also require $\mu_S < q$ as in the Base model to ensure finite lifetime.

Note that over time, the investor may dynamically adjust the fraction of assets invested in the stock market. The investment manager, who is paid proportionally to returns, would prefer the higher-return option — to be fully invested in the stock market at all times — but is aware that a period of poor returns could cause the customer to leave.

Analogous to the Base model, a policy $\lambda$ is admissible if the firm’s action process $(p_t)_{t \geq 0}$ (by following this policy) is adapted to the filtration $\mathbb{F}$, takes value in $[0, 1]$, and is such that the corresponding happiness processes is an $\mathbb{F}$-adapted semimartingale specified uniquely in law. We denote the set of admissible policies by $\Lambda$.

The optimal value function under the new policy space $\Lambda$ is given by

$$V I(x) = \sup_{\lambda \in \Lambda} \mathbb{E} \left[ \int_0^\infty (\mu_S + p_t(\mu_R - \mu_S))1\{t < T\}dt + \int_0^\infty p_t\sigma_R1\{t < T\}dB_t \bigg| H_0 = x \right].$$

(19)

We call this model under policy space $\Lambda$ the Investor model.

As in the Base model, we expect that there is a stationary Markov optimal policy $\lambda : \mathbb{R} \to [0, 1]$. Interval policies, suitably generalized (see Definition 2 in Appendix A), are a subclass of stationary Markov policies that are admissible as before (see the last paragraph in Appendix A).

The main result in this section shows that, similar to the optimal policy of the Base model, there exists an optimal policy in the Investor model that is either the myopic policy (one that always uses the Risky mode) or a sandwich policy. The structure of a sandwich policy in the Investor model is slightly different than in the Base model: the firm still uses the Risky mode for low and high levels of happiness, but instead of using purely the Safe mode for an intermediate happiness interval, the firm mixes the Risky mode with the Safe mode in $[q, \theta_I]$, where $\theta_I$ is defined in the lemma below.

We call the interval $[q, \theta_I]$ the risk-averse region.

Lemma 5. Let $\Theta_I$ be the set of values of $\theta$ that satisfy

$$\frac{\mu_R\sqrt{\pi}}{\sigma_R} e^{-\frac{(\theta - \mu_R)^2}{\sigma_R^2}} \left( 1 - \text{erf} \left( \frac{\theta - \mu_R}{\sigma_R} \right) \right) (2 \theta - \mu_R - \mu_S) - (\mu_R + \mu_S) = 0.$$

Then, the set $\Theta_I \cap (\frac{\mu_S + \mu_R}{2}, \infty)$ contains a single element, which we label $\theta_I$. 

24
The idea behind this definition is that at happiness $\theta_I$, the marginal benefit from happiness is the same whether the investor uses purely the Risky mode or includes a tiny proportion of the Safe mode (similar to the idea behind $\theta_b$ in the Base model explained at the end of Section 4.1).

Moreover, inside the risk-averse region $[q, \theta_I]$, the proportion $\lambda^*(x)$ of the Risky mode which the firm employs at happiness level $x$ is specified as follows:

**Lemma 6.** Let $X(\cdot)$ be the following function:

$$X(y) = e^{-2a \left( \log(y+1) + \frac{1}{y+1} \right)} K_0 - (2a)^{1+2a} ye^{-2a \left( \log(y+1) + \frac{1}{y+1} \right)} \Gamma \left( -2a, -\frac{2a}{y+1} \right),$$

where

$$K_0 = \theta_I e^{2a \left( \log(g)+\frac{1}{g} \right) b (-2a)^{1+2a} \Gamma \left( -2a, -\frac{2a}{g} \right)},$$

$$a = \frac{\sigma_R^2}{(\mu_R - \mu_S)^2}, \quad b = \mu_S, \quad g = \frac{2\theta_I}{2\theta_I - \mu_S - \mu_R},$$

and $\Gamma(s, z)$ is the upper incomplete gamma function (see Chaudry and Zubair [9]). Then, the inverse function of function $X(\cdot)$ is properly defined on $(\mu_S, \theta_I]$ and is represented by $G(\cdot) = X^{-1}(\cdot)$. Then, on $(\mu_S, \theta_I]$, the function $G(\cdot)$ is positive, strictly decreasing, differentiable, and satisfies the following ODE:

$$(G(x) + 1)^2 + 2a(x - b)G(x)G'(x) - 2abG'(x) = 0.$$

Finally, define the function

$$\lambda^*(x) \triangleq \frac{(\mu_S - \mu_R)(G(x) + 1)}{\sigma_R^2 G'(x)}.$$  \hspace{1cm} (20)

Then, $\lambda^*$ is strictly increasing on $[q, \theta_I]$ with $\lambda^*(q) > 0$ and $\lambda^*(\theta_I) = 1$.

The function $G(x)$ in Lemma 6 captures the marginal benefit of happiness at $x$, when the firm uses the conjectured optimal proportion $\lambda^*(\cdot)$ of the Risky mode. The function $X(\cdot)$ is the inverse function of $G(\cdot)$.

We are now ready to state our main theorem of this section.

**Theorem 3.** Suppose $\mu_S < \mu_R$ and $\mu_S < q$. Consider the firm’s problem as presented in Eq. (19). Let $\theta_I$ be as defined in Lemma 5. If $\theta_I \leq q$, then the myopic (pure Risky-everywhere) policy is optimal. If $\theta_I > q$, then a sandwich policy is optimal, where the proportion of the Risky mode is $\lambda^*(x)$ as defined in Lemma 6 for happiness levels $x \in [q, \theta_I]$, and $\lambda^*(x) = 1$ for $x \notin [q, \theta_I]$.

The proof of Theorem 3 follows the same methodology as the proof of Theorem 1. The HJB equation for the firm’s objective Eq. (19) is

$$\max_{p \in [0,1]} \left\{ -Q(x)V(x) + (\mu_S + p(\mu_R - \mu_S) - x)V'(x) + \frac{1}{2} p^2 \sigma_R^2 V''(x) + \mu_S + p(\mu_R - \mu_S) \right\} = 0. \hspace{1cm} (21)$$

Therefore, the optimality conditions for the Investor model are as in Proposition 2, except that Conditions 4-5 are replaced by
4’ the following inequality\textsuperscript{15} is true for any $x \in \mathbb{R}$ when $p = 0$ and for any $x \in \mathbb{R} \setminus \mathcal{E}$ when $p \in (0, 1]$:

$$-Q(x)\tilde{V}(x) + (\mu_S + p(\mu_R - \mu_S) - x)\tilde{V}'(x) + \frac{1}{2}p^2\sigma_R^2\tilde{V}''(x) + \mu_S + p(\mu_R - \mu_S) \leq 0; \quad (22)$$

and

5’ for all $x \in \mathbb{R}$ and some interval policy (see Definition 2 in Appendix A) $\bar{\lambda} \in \Lambda$ such that $\bar{\lambda}(y) = 0$ for all $y \in \mathcal{E}$, the process $(\tilde{V}(H_t^{x, \bar{\lambda}}))_{t \geq 0}$ is an $\mathbb{F}$-adapted semimartingale, and

$$-Q(x)\tilde{V}(x) + (\mu_S + \bar{\lambda}(x)(\mu_R - \mu_S) - x)\tilde{V}'(x) + \frac{1}{2}\bar{\lambda}(x)^2\sigma_R^2\tilde{V}''(x) + \mu_S + \bar{\lambda}(x)(\mu_R - \mu_S) = 0. \quad (23)$$

We obtain a candidate value function $W^I$ (see Appendix D for the explicit expression) by solving the HJB equation (21). Then we verify that the candidate value function $W^I$ together with the specified policy satisfies the optimality conditions 1-3, 4’-5’ and 6. The theorem is proved in Appendix D.

5.1 Comparative Statics and Numerics for the Investor Model

We now provide monotonicity results regarding the size of the risk-averse region and the firm’s behavior inside that region.

**Theorem 4.** Let $\theta_I$ be as defined in Lemma 5 and $\lambda^*(\cdot)$ be as defined in Lemma 6. Then, the following properties hold:

1. The threshold $\theta_I$ increases in $\mu_S$ and $\sigma_R$.

2. Assume the optimal policy for a given set of parameters (see Theorem 3) is a sandwich policy. Then, the proportion of the Risky mode $\lambda^*(\cdot)$ in the risk-averse region $[q, \theta_I]$ is strictly increasing with happiness, and the firm strictly mixes at the satisfaction threshold $q$ but not at $\theta_I$; that is, we have $\lambda^*(q) > 0$ and $\lambda^*(\theta_I) = 1$.

The first part of this theorem simply states that the size of the risk-averse region in the Investor model possesses the same monotonicities with regard to $\mu_S$ and $\sigma_R$ as in the Base model. The second part of Theorem 4 states that when the customer happiness level is in the unsatisfied zone, the firm is more risk averse closer to the happiness threshold $q$. That is, the firm prefers a lower risk profile closer to $q$ even though this generates a lower current reward rate. Note that $\lambda^*(q-) = 1$, and $\lambda^*(q+)$ has a value such that $(V^I)'$ is continuous at $q$ despite the step in $Q(\cdot)$.

An interesting implication of part 2 of Theorem 4 is that the optimal policy never uses the Safe service mode alone. It always mixes the Risky mode with the Safe mode when it is risk averse.

\textsuperscript{15}When $p = 0$, we define $\frac{1}{2}p^2\sigma_R^2\tilde{V}''(x)$ to be zero for any $x$ including $x \in \mathcal{E}$.
The intuition is that the Risky mode has a higher drift \((\mu_R > \mu_S)\), and as specified in Eq. (18), the variance is only quadratic in \(\lambda(\cdot)\) while the drift is the \(\lambda(\cdot)\)-weighted convex combination of \(\mu_R\) and \(\mu_S\), so it is always beneficial to include at least a small proportion of Risky.

We give a few numerical examples of the optimal policy in the following graphs (Figures 8–10). Notice that the model parameters considered here are the same as those in the Base Model. In comparison, the size of the risk-averse regions here are larger than in the Base model.

![Graphs showing proportion of Risky mode for different model parameters](image)

Figure 8: Optimal policies for the Investor model: \(\mu_R = 9, \sigma_R = 10, q = 10\).

![Graphs showing proportion of Risky mode for different model parameters](image)

Figure 9: Optimal policies for the Investor model: \(\mu_S = 5, \sigma_R = 10, q = 10\).

![Graphs showing proportion of Risky mode for different model parameters](image)

Figure 10: Optimal policies for the Investor model: \(\mu_S = 5, \mu_R = 9, q = 10\).

We omit to provide the gap between the optimal policy and the myopic policy (pure Risky everywhere). Observe that the Investor model expands the policy space, hence increasing optimal CLV. That is, the gap for the Investor model exceeds that under the Base model (Figure 5).
In the next section, we discuss the robustness of our optimal policy structure to the reward process being a GBM.

6 Geometric Brownian Motion

Motivated by the investment manager application, in this section we consider a variant of our model where rewards accrue according to a geometric Brownian motion. Specifically, we assume that under the firm’s choice of service mode $u_t \in \{R, S\}$, instead of Eq. (1) in the Base model, the total reward $\tilde{Y}_t$ evolves according to a Geometric Brownian Motion (GBM)

$$d\tilde{Y}_t = \mu_{u_t} \tilde{Y}_t dt + \sigma_{u_t} \tilde{Y}_t dB_t$$

with $\tilde{Y}_0 = 1$. This formulation is motivated by an investment problem where a risk-free asset generates $\mu_S$ returns as interest and a risky asset generates $\mu_R$ returns in capital gains and dividends, with the dividends being immediately reinvested in the portfolio. (Here $\mu_{u_t}$ may be termed the “percentage drift” as per standard terminology in a GBM setting, but we simply call it the rate of return.) We are interested in cases where $\mu_R > \mu_S > 0$, since typically, the risky asset generates higher return on average and both assets give positive returns. (However, we do not impose the restriction $\mu_R > \mu_S$ for our analytical development.) We assume customer happiness $\tilde{H}_t$ follows a stochastic differential equation:

$$d\tilde{H}_t = \frac{d\tilde{Y}_t}{\tilde{Y}_t} - \tilde{H}_t dt,$$

where $\tilde{H}_0 = x$ is the initial customer happiness. Comparing Eqs. (24) and (25) with Eqs. (1) and (2), we see that under the same action process $u_t$, customer happiness $\tilde{H}_t$ follows the same dynamics as $H_t$ does in the Base model. We assume that the hazard rate of customer churn is still a step function but is positive even for $\tilde{H}_t \geq q$. Specifically,

$$\hat{Q}(\tilde{H}_t) \triangleq Q_1 1\{H_t < q\} + Q_2 1\{\tilde{H}_t \geq q\}$$

with $Q_1 > Q_2 > 0$. Denote by $\tilde{T}$ the customer lifetime:

$$\tilde{T} \triangleq \inf\left\{ t \geq 0 : e^{-\int_0^t \hat{Q}(\tilde{H}_s) ds} = w \right\},$$

where $w$ is a uniform random variable over $[0, 1]$ independent of filtration $\mathcal{F}$.

We require the next condition on $Q_1$ and $Q_2$ to ensure that the expected CLV (which will be defined next) is finite$^{17}$.

**Condition 1.** $Q_1 > Q_2 > \max\{\mu_S, \mu_R\}$.

$^{16}$The solution to Eq. (24) is $\tilde{Y}_t = \exp\left(\int_0^t (\mu_{u_s} - \sigma^2_{u_s}/2) ds + \int_0^t \sigma_{u_s} dB_s\right)$.

$^{17}$Suppose the hazard rate is $Q$ for all happiness states and the firm always uses the Risky service mode, then it is not hard to show that the expected CLV is $\frac{Q}{Q-R}$ if $Q > R$, and $\infty$ if $Q \leq R$. 

28
Now we state the firm’s problem. Denote by $\tilde{\Pi}$ the space of admissible policies that satisfy the usual conditions as in the Base model, which is that under the policy, the corresponding action process $u_t$ should be adapted to filtration $\mathcal{F}$, takes value in $\{S, R\}$, and the corresponding $\tilde{H}_t$ is an $\mathcal{F}$-adapted semimartingale uniquely specified in law. For a given starting happiness $x$ and admissible policy $\pi$, let $\tilde{Y}^{x,\pi}$ denote the reward gained up to time $t$ and $\tilde{T}^{x,\pi}$ be the customer lifetime. Then the CLV is equal to

$$\tilde{V}(x, \pi) = \mathbb{E} \left[ 1 + \int_0^\infty \mathbb{1}_{\{t < \tilde{T}^{x,\pi}\}} d\tilde{Y}^{x,\pi}_t \mid \tilde{H}_0 = x \right].$$

(28)

The firm’s objective is to maximize the CLV it earns from interacting with the customer. The optimal CLV given a starting happiness $x$ is given by

$$\tilde{V}^*(x) = \sup_{\pi \in \tilde{\Pi}} \tilde{V}(x, \pi).$$

(29)

### 6.1 The Optimality Conditions

As in the Base model, we first present the optimality conditions for a function $\tilde{W}: \mathbb{R} \to \mathbb{R}$ to be the optimal value function.

**Proposition 4.** Suppose a function $\tilde{W}: \mathbb{R} \to \mathbb{R}$ satisfies

1. the function value $\tilde{W}(x) > 1$ for any $x \in \mathbb{R}$;
2. the function $\tilde{W}$ is continuously differentiable everywhere on $\mathbb{R}$ and twice continuously differentiable everywhere on $\mathbb{R} \setminus \mathcal{E}$ for some countable set $\mathcal{E}$;
3. the function $\tilde{W}$ is bounded;
4. the function $\tilde{W}'$ is bounded;
5. for any $x \in \mathbb{R}$ for $i = S$ and for any $x \in \mathbb{R} \setminus \mathcal{E}$ for $i = R$ the following holds$^{18}$:

$$\tilde{Q}(x) + (\mu_i - \tilde{Q}(x))\tilde{W}(x) + (\mu_i + \sigma_i^2 - x)\tilde{W}'(x) + \frac{1}{2}\sigma_i^2\tilde{W}''(x) \leq 0;$$

6. for some interval policy $\tilde{\pi}$ (see Definition 1) such that $\tilde{\pi}(y) = S$ for all $y \in \mathcal{E}$, the process $\tilde{W}(\tilde{H}_t)$ is an $\mathcal{F}$-adapted semimartingale, and for all $x \in \mathbb{R}$ it holds that

$$\tilde{Q}(x) + (\mu_{\tilde{\pi}(x)} - \tilde{Q}(x))\tilde{W}(x) + \left(\mu_{\tilde{\pi}(x)} + \sigma_{\tilde{\pi}(x)}^2 - x\right)\tilde{W}'(x) + \frac{1}{2}\sigma_{\tilde{\pi}(x)}^2\tilde{W}''(x) = 0.$$  

(30)

Then, the function $\tilde{W}$ is the optimal value function $\tilde{V}^*$, and $\tilde{\pi}$ is an optimal policy.

---

$^{18}$For any $x \in \mathcal{E}$, we define the term $\frac{1}{2}\sigma_S^2\tilde{W}''(x)$ to be zero consistent with $\sigma_S = 0$.  

29
Note that Conditions 5 and 6 together establish the HJB equation for this new setting:

$$\max_{i \in \{S, R\}} \left\{ \tilde{Q}(x) + (\mu_i - \tilde{Q}(x)) \tilde{V}(x) + (\mu_i + \sigma_i^2 - x) \tilde{V}'(x) + \frac{1}{2} \sigma_i^2 \tilde{V}''(x) \right\} = 0$$  \hspace{1cm} (31)

for all $x \in \mathbb{R}$ where $\tilde{V}''$ exists.

Recall from the analysis of the Base model (Section 4.2), we would like to solve the HJB equation to get a candidate value function such that all optimality conditions are satisfied. A general solution to the HJB equation (31) (for a single service mode) is of the following form:

$$\tilde{W}(x, C_1, C_2) = \frac{\tilde{Q}(x)}{\tilde{Q}(x) - \mu_R} + C_1 H \left( \mu_R - \tilde{Q}(x), \frac{x - \mu_R - \sigma_R^2}{\sigma_R} \right)$$

$$+ C_2 M \left( \frac{\tilde{Q}(x) - \mu_R}{2}, \frac{1}{2}, \frac{(x - \mu_R - \sigma_R^2)^2}{\sigma_R^2} \right)$$  \hspace{1cm} (32)

under the Risky service mode, and

$$\tilde{W}(x, C_3, S) = \frac{\tilde{Q}(x)}{\tilde{Q}(x) - \mu_S} + C_3 (\mu_S - x)^{\mu_S - \tilde{Q}(x)}$$

under the Safe service mode, where $C_1$, $C_2$, $C_3$ are free parameters, and in Eq. (32) $H(\cdot, \cdot)$ is a Hermite polynomial function, and $M(\cdot, \cdot, \cdot)$ is the Kummer confluent hypergeometric function.\(^{19}\) We then numerically find the values of $C_1$, $C_2$ and $C_3$ such that the optimality conditions in Proposition 4 are satisfied. In the next section, we provide our numerical findings, which shows that our main structural results for the Base model also hold in the GBM setting.

6.2 Numerics

In this section, we provide numerical solutions to this modified model with GBM reward. In particular, we are interested in the regime where $q > \mu_S$, which implies that the customer is eventually not satisfied if the firm uses the Safe mode all the time. We also let $\mu_R > \mu_S$, so that the Risky asset accumulates higher rewards on average. We solve for the optimal policy for random instances (the model primitives are randomly generated, see details in Appendix E.). We make careful use of a reflecting boundary $q - B$ for some large positive $B$ to ensure numerical stability while ensuring that the effect on CLV is very small. Next we present our findings.

**Optimal policy is either myopic or sandwich.** As in the Base model (see Theorem 1 for the $\mu_R > \mu_S$ case), the optimal policy is either a myopic one or a sandwich policy. In particular, among all the randomly generated instances, the (numerically solved) optimal solutions are either a myopic policy that always uses the Risky mode everywhere, or a sandwich policy that uses the Risky mode for all happiness states except for some intermediate happiness range $[q, \tilde{\theta}_b]$ (for some numerically specified $\tilde{\theta}_b$). It is worth noting that in the GBM setting, the optimal sandwich policy

\(^{19}\)The two functions $H(\lambda, x)$ and $M(\frac{-1}{2}, \frac{1}{2}, x^2)$ are the two linearly independent solutions to the Hermite Differential Equation $y''(x) - 2xy'(x) + 2\lambda y(x) = 0$. 

30
once again provides substantial CLV increase over the myopic policy. For example, consider the model primitives $\mu_S = 0.12$, $\mu_R = 0.14$, $\sigma_R = 0.3$, $q = 0.13$, $Q_1 = 1.5$ and $Q_2 = 0.5$. (Here one may think of 1 time unit in the model as being a period of about 2 years.) This means the Safe asset’s rate of return is 12% (with continuous compounding), and the Risky asset’s expected rate of return is 14%, with volatility 30%. Also the customer is not satisfied with a rate of return below 13%, and his hazard rate of churn increases from 0.5 to 1.5 if he estimates (as quantified by his happiness) that the rate of return is below 13%. Under this set of model primitives, the optimal sandwich policy (see the dotted vertical line in Figure 11) is to use the Safe mode for happiness values on $[0.13, 0.147]$, and use the Risky mode elsewhere. The CLV increase from using the optimal sandwich policy relative to the myopic policy is 7.0%. Though we do not permit mixed strategies in this section, we briefly observe that the CLV increase relative to the myopic policy will be even larger if mixed strategies are permitted, since the myopic strategy remains unaffected.

Figure 11: The optimal sandwich policies for different values of $\mu_R$, under $\mu_S = 0.12$, $\sigma_R = 0.3$, $q = 0.13$, $Q_1 = 1.5$ and $Q_2 = 0.5$. The horizontal axis corresponds to the value of $\mu_R$, and the vertical axis is the happiness value. The two curves are the switching boundaries between the Risky mode and the Safe mode.

Optimal switching threshold exhibits similar monotonicity as in the Base model. The optimal switching threshold $\tilde{\theta}_b$ shows similar monotonicity in model primitives as in the Base model (see Theorem 2). That is, $\tilde{\theta}_b$ decreases as we increase $\mu_R$, and increases as we increase $\sigma_R$. Figure 11 plots the prescribed optimal sandwich policy for various model primitives. In particular, we fix $\mu_S = 0.12$, $\sigma_R = 0.3$, $q = 0.13$, $Q_1 = 1.5$, $Q_2 = 0.5$, and vary $\mu_R$ in $[0.12, 0.20]$. The horizontal axis is $\mu_R$, and the vertical axis is the happiness value. The two curves on the plot are the two switching thresholds, separating the happiness regions where the firm should use different service modes. The white region represents the risk-averse region where the firm should choose the
Safe service mode. Notice that $\hat{\theta}_b$ here has the same monotonicity (i.e., decreasing in $\mu_R$) as the switching threshold in the right half (corresponding to $\mu_R > \mu_S$) of Figure 6 in the Base model.

In the following section, we discuss the impact of adding switching costs to the optimal policy structure.

7 Switching Costs

Consider the Base model in the case $\mu_R > \mu_S$, and recall that the firm’s optimal policy is either a myopic policy or a sandwich policy. In the cases where a sandwich policy is optimal, the optimal policy prescribes to use the Safe mode at and just above the happiness threshold $q$, and to use the Risky mode below the happiness threshold. As a result, during a customer’s lifetime, with positive probability the firm will switch between the Risky and Safe modes at $q$ infinitely many times (see Figure 4). This is a concern in practical settings where each transition from one mode to the other incurs a cost. Given any fixed positive cost for each transition, the optimal policy we constructed will incur infinite cost (with positive probability and in expectation), which is obviously undesirable and suboptimal.

This section will provide a brief discussion and numerical investigation of a setting with fixed positive switching cost, and find the way in which the structure of the optimal policy gets modified to deal with positive switching costs. In the interest of space, we will omit formal technical details in this section. Notably, we will find that our main insights are preserved for small positive switching costs. At the end of this section, we will briefly discuss implications for an alternate setting (with no switching costs) where the firm is unable to exactly measure the happiness of a customer, namely, that our findings can nevertheless help the firm increase the customer CLV.

Model with switching costs. Assume that each transition from one mode to the other incurs a fixed cost, denoted by $K$, and keep all other assumptions in the Base model unchanged. In this setting, the decision of which service mode to adopt should not only depend on the customer’s current happiness level, but should also depend on the firm’s current mode of service. In other words, the setting is stationary Markov with respect to a two-variable state which includes both the happiness and the current mode of service, and so, without loss of optimality, we can restrict attention to stationary Markov policies with respect to this state. Let $\pi$ be such a stationary Markov policy, which maps from $\mathbb{R} \times \{S, R\}$ to $\{S, R\}$, i.e., $\pi(x, i)$ prescribes which service mode to use when the customer’s happiness level is $x$ and the firm’s current service mode is $i \in \{S, R\}$. Denote by $u_t$ the firm’s service mode prescription at time $t$. Given an initial happiness $H_0 = x$ and initial service mode $u_{0-} = i$, under policy $\pi$, the firm’s service mode prescription at each time $t \geq 0$ is $u_t = \pi(H_t, u_{t-})$. (As before, $H_t$ and $u_t$ are defined for all $t \geq 0$ independent of the customer lifetime $T$.) We restrict attention to policies such that, with probability 1, the resulting $u_t$ process is right-continuous with left limits (cadlag).

The firm’s objective is to find the policy that maximizes the difference between the expected
reward earned during the customer’s lifetime and the total switching cost incurred. Accordingly, the optimal expected CLV for a starting happiness state \( x \) and starting service mode \( i \in \{S, R\} \) is

\[
V^*_i(x) \triangleq \sup_{\pi \in \Pi} \mathbb{E} \left[ \int_0^\infty \mathbbm{1}_{\{t < T\}} dY_t - K \sum_{k=1}^{\infty} \mathbbm{1}_{\{\tau_k < T\}} \right | H_0 = x, u_{0-} = i ,
\]

where \( \tau_k \) denotes the time point of the \( k \)th switch in service mode in the \( u_t \) sample path, and the customer lifetime \( T \) is defined as per Eq. (6) as before. (Since \( u_t \) is cadlag under admissible policy \( \pi \), the set of time points where \( u_t \) switches is countable.)

**HJB equations.** The optimal CLV under switching cost should satisfy the following HJB equations (derived informally in Appendix F):

\[
0 = \max \left\{ -Q(x)V^*_R(x) + (\mu_R - x)V^*_R(x) + \frac{\sigma_R^2}{2} V^{*''}_R(x) + \mu_R, V^*_S(x) - V^*_R(x) - K \right\} \quad \forall x \in \mathbb{R} \text{ where } V^{*''}_R(x) \text{ exists}; \tag{33}
\]

\[
0 = \max \left\{ -Q(x)V^*_S(x) + (\mu_S - x)V^*_S(x) + \mu_S, V^*_R(x) - V^*_S(x) - K \right\} \quad \forall x \in \mathbb{R}. \tag{34}
\]

We now summarize how we obtain these equations: We have two equations in this setting, each corresponding to one of the two possible service modes (the Risky mode or the Safe mode) being used currently. In particular, Eq. (33) comes from comparing the continuation value of staying with the Risky mode with the value of switching to the Safe mode while incurring a switching cost \( K \). Similarly, Eq. (34) comes from comparing the continuation value of staying with the Safe mode with the value of switching to the Risky mode while incurring a switching cost \( K \).

**Findings.** We focus on the case \( \mu_R > \mu_S \), and numerically solve the HJB equations (33) and (34) to find the firm’s optimal policy under switching cost \( K \). In short, we find that adding a small switching cost results in an optimal policy which is very similar to the sandwich policy we find to be optimal in the Base model.

Recall our result for the Base model for \( \mu_R > \mu_S \), namely, that in cases where a sandwich policy is optimal, \( \theta_b \) and \( q \) are the thresholds separating the happiness values where the firm should use the Safe service mode from those where the firm should use the Risky service mode (Figure 2b). With a small positive switching cost \( K \), our numerics reveal that each switching threshold is replaced by a buffer interval. Specifically, above and below a buffer the firm should prefer opposite service modes (as in the Base model, regardless of which service mode is currently in use), whereas inside a buffer interval the firm should not switch service modes. (Intuitively, the CLV benefit of having buffers in place of sharp switching thresholds for \( K > 0 \) is to reduce the number of switches between service modes.)

Figure 12 illustrates the optimal policy as a function of switching cost \( K \). In the plot, the horizontal coordinate is the switching cost, and the vertical coordinate is the customer’s current happiness value. All model primitives besides \( K \) are held fixed as per \( \mu_S = 8, \mu_R = 9, \sigma_R = 10 \), and \( q = 10 \). The shaded areas in Figure 12 represents the buffers of the optimal policy. Observe
that the buffers grow when we increase the switching costs. The special case $K = 0$ corresponds to the Base model, and leads to an interval optimal policy as before with sharp thresholds $q$ and $\theta_b = 22.10$. Consider $K = 0.05$. The optimal policy has buffer zones $(q,10.37) = (10, 10.37)$ and $(\theta_b,28.22) = (22.10, 28.22)$, corresponding to the intersection between the $K = 0.05$ line and the shaded areas in Figure 12. Outside the buffer zones, the policy is identical to that in the Base model, i.e., it uses the Risky mode for happiness values above 28.22 and below $q$, and it uses the Safe mode for happiness values in $(10.37,22.10)$. To illustrate the role of the buffer zones: if the starting service mode is Risky and the starting happiness is $q$ (or anywhere below 10.37), then the policy prescribes to stay with Risky until the happiness exceeds 10.37. If the latter occurs, the policy immediately switches to Safe and stays with Safe until the happiness either exceeds 28.22 or drops below $q = 10$, and so on. Despite the positive switching cost, the optimal policy produces a substantially larger expected CLV than the myopic policy. Specifically, the payoff is nearly 79% larger under the optimal policy $V^*_R(q) = 1.79$, where $V(q, \text{Myopic})$ is the CLV of using the Risky mode always.

Figure 12: Firm’s optimal policy as a function of switching cost $K$. We fix $\mu_S = 8, \mu_R = 9, \sigma_R = 10$, and $q = 10$. The shaded areas represent the buffers where the optimal policy retains the current service mode. The white areas represents the regions where the policy employs a specific service mode, even if this would incur a switching cost. For each $K$, the buffers of the optimal policy are given by the intervals where the (horizontal coordinate = $K$) vertical line intersects the shaded areas, e.g, the dotted line in the left subfigure gives the optimal policy for $K = 0.05$.

In the cases where the myopic (Risky always) policy is optimal under the Base model (with $K = 0$), it is clear that the Risky always policy remains optimal even for $K > 0$, since the policy does not incur any switching cost.\footnote{In fact, we also find in numerical solutions that as the switching cost $K$ increases to above $\mu_R - \mu_S$, there emerges a threshold $\theta_0 < q$, increasing with $K$, such that for happiness states under this threshold, the firm should not switch to the Risky mode, if currently using the Safe mode. Here $\mu_R - \mu_S$ is the CLV difference between the Risky-always policy and Safe-always policy in the limit of initial happiness $x \to -\infty$, since $Q(x') = 1 \forall x' < q$.}\footnote{There is technical caveat here: if the starting service mode is Safe, then if $K > 0$ there is a nontrivial decision}
Implications for a setting where the firm cannot perfectly measure happiness. So far we have assumed that the firm is able to perfectly estimate the customer’s current happiness state. A reader might be concerned about the robustness of our findings to estimation errors (or delays). We now argue that the results we have obtained under switching costs suggest that small to medium-sized errors in estimating customer happiness would not significantly impair the CLV benefits of the optimal policy (relative to the myopic policy). In each case where the optimal policy for the case of switching costs is one where switching is postponed by a buffer (see Figure 12), by definition this policy produces higher CLV than the myopic policy. We can hence conclude this policy also produces higher CLV than the myopic policy if there were no switching costs (for \( K = 0 \), the CLV under the former policy is even larger whereas under the CLV under the latter policy stays the same). To give a quantitative example, for \( K = 0.70 \) the lower buffer interval of the optimal policy is \((10, 12.03)\) as per Figure 12. Thus the policy decisions are loosely similar to those of the optimal (sandwich) policy under the Base model acting on estimated customer happiness when there are estimation errors of size similar to the size of the lower buffer interval \( \sim 2 \) (the upper buffer interval plays only a small role since the customer happiness only rarely rises to that level). The CLV increase from using the optimal (buffer) policy relative to the myopic policy is 30.0% for \( K = 0.70 \), and so the CLV increase from using the same policy in the absence of switching costs is even larger. This gives us confidence that our proposed policies still substantially increase the CLV in the face of small to medium-sized errors in estimating customer happiness. Along similar lines, interpreting the effect of the buffer intervals as delays in switching, one can argue that (small) delays in estimating customer happiness are unlikely to erode the benefits of using our proposed policies.

In the next section, we also test robustness of the sandwich policy to the shape of the hazard rate function.

8 Other Hazard Rate Functions

We now provide numerical evidence for our results’ robustness to alternative specifications of the hazard rate function (recall the original step function in Eq. (4)) in the case \( \mu_R > \mu_S > 0 \). The most important takeaway of this section is that, for a variety of hazard rate functions and different model primitives, the optimal policy is still either myopic or a sandwich policy, and moreover, the switching thresholds in the optimal sandwich policy are fairly robust to the shape of hazard rate functions.

First, we consider a variety of different hazard rate specifications in the unsatisfied zone, while keeping the original assumption of zero hazard rate in the satisfied zone. Let \( q \) be the happiness threshold separating the unsatisfied zone and the satisfied zone. Consider the following four types regarding whether to switch to Risky and under what conditions. This case has limited practical relevance, so we avoid discussing it in the interest of space.
of hazard rate functions:

1. constant $k$: $Q(x) = k 1\{x < q\}$;
2. $n$th power: $Q(x) = (q - x)^n 1\{x < q\}$;
3. exponential: $Q(x) = (e^{q-x} - 1) 1\{x < q\}$;
4. logit: $Q(x) = \left(\frac{e^{q-x}}{1 + e^{q-x}} - \frac{1}{2}\right) 1\{x < q\}$.

The value function and optimal policy associated with each hazard rate function can be established by solving the HJB equation (13) and checking that the optimality conditions in Proposition 2 are still satisfied. Under all these different choices of the hazard rate function in the unsatisfied zone (including several different constants, and powers) and different model primitives with $\mu_R > \mu_S$, we
find that (similar to Theorem 1) the optimal policy is either myopic or a sandwich policy. Figure 13 presents the firm’s optimal policy and the associated CLV for some of the numerical instances.

One interesting observation from Figure 13 is that the size of the risk-averse region depends on how fast the hazard rate of leaving increases as the customer happiness level descends into the unsatisfied zone. In the Base model, the firm switches from Safe to Risky as soon the customer happiness crosses from the satisfied zone into the unsatisfied zone. This is not always the case with arbitrary hazard rate functions. When the hazard rate grows relatively fast as customer happiness goes down, the lower switching threshold remains at \( q \). However, in the plotted cases where the hazard rate does not grow swiftly when customer happiness crosses into the unsatisfied zone, the switching point to Risky is strictly below \( q \). In our numerics, this occurs for the cases of hazard rate functions growing as the fourth and eighth powers (see Figures 13(f) and 13(g)), where the lower boundary of the risk-averse region is strictly below \( q \).

In Figure 14, we also show that the gap in CLVs between under the optimal policy and the myopic policy remains large for different hazard rate functions. In particular, the second curve from the bottom corresponds to the original step function hazard rate in Eq. (4). The other three curves correspond to the hazard rate function being linear, exponential, and logit in the unsatisfied zone (and zero in the satisfied zone) as introduced earlier in this section.

Other than the choice of hazard rate functions listed above, we also numerically examined cases where the hazard rate is strictly positive everywhere (including in the satisfied zone), again in the Base model and restricting \( \mu_R > \mu_S \), and find our main results remain intact. In particular, we considered hazard rate functions of the form \( Q(x) = 1 \{ x < q \} + \epsilon 1 \{ x \geq q \} \) for \( \epsilon \in (0, 1) \), and various model primitives such that \( \mu_R > \mu_S \), and found that the optimal policy is still either myopic or a sandwich policy, where (in the optimal sandwich policy) the upper switching threshold decreases smoothly as we increase the value of \( \epsilon \). The CLV increase from using the optimal sandwich policy relative to the myopic policy is still large for small values of \( \epsilon \). We omit more details of this robustness check for brevity.

9 Discussion

We first summarize our findings and then argue that our findings should be relevant in an alternate “non-contractual” setting as well, where unhappy customers don’t leave, but use the firm less often.

Summary of Findings. In this paper, we introduced and studied a behavioral model to incorporate customer abandonment when her happiness depends on recent experiences, and we asked how the firm should choose between a Risky and a Safe service mode. Although volatility — in our setting, uncertainty in the customer’s response — is often considered undesirable, we proved that there always exists a risk-seeking region where the firm prefers volatility, even in cases where the corresponding average reward rate is lower than under the Safe service mode. We showed that the optimal policy has a risk-seeking region just to the left of the happiness threshold separating the
Figure 14: The ratio of CLV under the optimal policy to CLV under the myopic policy versus $\mu_S$ under different hazard rate functions (see Section 8), for $\mu_R = 5$, $\sigma_R = 2$, $q = 6$, and initial happiness $q$.

unsatisfied and satisfied zones and a “symmetric” risk-averse region to the right of this threshold (Figure 1).

Specifically, the risk-seeking and risk-averse regions both lie at intermediate customer happiness levels, where the next few experiences can significantly alter the customer’s risk of leaving. When the customer happiness is currently in the unsatisfied zone but close to the satisfied zone, the firm should be risk-seeking and use the Risky service mode. On the other hand, when the customer happiness is currently in the satisfied zone but close to the unsatisfied zone, the firm should be risk-averse and use the Safe service mode (unless its reward rate is very low). However, when customer happiness level is far from the happiness threshold, the firm should use whichever service mode that generates a higher reward rate.

We showed (see Figure 5) that the optimal service policy can produce a large multiple of the CLV achieved by the myopic policy (which sticks with the service mode that has a higher reward rate). We extended our results to the case where the firm can mix between the service modes. We also considered geometric Brownian motion reward process instead of arithmetic Brownian motion reward, and showed robustness of our results. Moreover, we examined the impact of having switching cost on the optimal policy structure, and showed a “buffered” sandwich policy is optimal for switching costs small enough (in particular, when the switching cost is no larger than $\mu_R - \mu_S$). We also showed the robustness of our results to alternate specifications of the hazard rate function governing customer abandonment. As a result, despite certain stylized features of our model, we are optimistic that our insights can provide significant value to a variety of service firms that grapple with customer churn and have flexibility in their choice of service mode.

**Connection with a non-contractual setting.** Now we introduce an alternate setting where the customer interacts with the firm at discrete epochs, with time between successive interactions
governed by the customer’s happiness. This is of interest because, while the model in the paper mirrors the so-called “contractual” setting studied in the marketing literature (where customers either continue to subscribe or leave entirely), the alternate model we describe here captures the complementary “non-contractual” setting (where a customer’s frequency of interaction with the firm drops if they are unhappy). We observe a close connection between the platform’s problem in the alternate setting and that in the model studied in the paper, which suggests that our main insights should extend to non-contractual settings as well.

Suppose the customer interacts with a firm at discrete epochs, indexed as \( j = 1, 2, \ldots \). We label the rewards generated from the \( j \)th interaction \( \hat{Y}_j \). Based on the firm’s service mode during the \( j \)th interaction, the reward is either a deterministic value \( \mu_S \delta \), or a random value independently generated from a Normal\((\mu_R \delta, \sigma_R^2 \delta)\) distribution (fix some \( \delta \) small). Compare this with the original definition of reward \( Y_t \) (see Eq. (1)). The new reward \( \hat{Y}_j \) approximates the original infinitesimal reward rate \( dY_t \), where \( \delta \) approximates \( dt \). The customer’s happiness level during the time interval between the \( j \)-th and the \((j + 1)\)-th interaction, \( \hat{H}_j \), is an exponentially weighted moving average of his past rewards: \( \hat{H}_j = (1 - \delta) \hat{H}_{j-1} + \hat{Y}_j \).

**Remark 1.** Assume the customer’s initial happiness state is \( \hat{H}_0 = x \). Then, after the \( j \)th interaction, the happiness \( \hat{H}_j \) is given by

\[
\hat{H}_j = x(1 - \delta)^j + \sum_{k=1}^{j} (1 - \delta)^{j-k} \hat{Y}_k
\]

(35)

Compare \( \hat{H}_j \) in Eq. (35) with the original definition of happiness \( H_t \) in Eq. (3). Note how \( \hat{H}_j \) relates to \( \hat{Y}_j \) in the same way that \( H_t \) relates to \( dY_t \).

Next we define the inter-arrival times of the customer and the platform objective in the alternate setting. Here we assume that the customer does not churn, but his inter-arrival time between the \( j \)-th and the \((j + 1)\)-th visit is exponentially distributed with mean \( e^{\delta \hat{Q}(\hat{H}_j)} - 1 \). Assume \( \hat{Q}(x) > 0 \) for all \( x \in \mathbb{R} \), in particular, we will consider

\[
\hat{Q}(x) \triangleq \left\{ \begin{array}{ll}
1 & \text{if } x < q, \\
\epsilon & \text{if } x \geq q.
\end{array} \right.
\]

(36)

Suppose the firm wants to maximize the present value of all future payoffs with exponential discount factor 1. This is again a stationary Markov setting, so the firm can restrict attention to stationary Markov policies \( \pi : \mathbb{R} \to \{R, S\} \) that map from the current happiness to the service mode employed. Then the firm’s objective is

\[
\max_{\pi} \hat{V}(x, \pi), \quad \text{where } \hat{V}(x, \pi) \triangleq \mathbb{E} \left[ \sum_{j=1}^{\infty} e^{-\hat{T}_j} \hat{Y}_j \bigg| \hat{H}_0 = x \right],
\]

where \( \hat{T}_j \) is the time of the \( j \)th interaction. We are now ready to make our key observation.
Remark 2. In the alternate model, the customer lifetime value is given by

\[ \hat{V}(x, \hat{\pi}) = \mathbb{E} \left[ \sum_{j=1}^{\infty} e^{-\sum_{k=1}^{j} \delta \hat{Q}(\hat{H}_{k-1})} \mu_{\hat{\pi}(\hat{H}_j)} \delta \bigg| \hat{H}_0 = x \right]. \]  

(37)

Proof. Let \( \hat{T}_0 \triangleq 0 \). We know that the interarrival time is exponentially distributed as per

\[ \hat{T}_j - \hat{T}_{j-1} \sim \text{Exp} \left( \frac{1}{e^{\delta \hat{Q}(\hat{H}_j - 1)}} \right) \quad \forall j = 1, 2, \ldots \]

Also, since the interarrival times are independently drawn given the happiness values so far, we have:

\[ \mathbb{E} \left[ e^{-\hat{T}_j} \mid \hat{H}_0, \hat{H}_1, \ldots, \hat{H}_{j-1} \right] = \mathbb{E} \left[ \prod_{k=1}^{j} e^{-(\hat{T}_k - \hat{T}_{k-1})} \mid \hat{H}_0, \hat{H}_1, \ldots, \hat{H}_{j-1} \right] \]

\[ = \prod_{k=1}^{j} \mathbb{E} \left[ e^{-(\hat{T}_k - \hat{T}_{k-1})} \mid \hat{H}_{k-1} \right] \]

\[ = \prod_{k=1}^{j} e^{-\delta \hat{Q}(\hat{H}_{k-1})} \]

\[ = e^{-\sum_{k=1}^{j} \delta \hat{Q}(\hat{H}_{k-1})}. \]

The customer’s CLV hence becomes

\[ \mathbb{E} \left[ \sum_{j=1}^{\infty} e^{-\hat{T}_j} \hat{Y}_j \bigg| \hat{H}_0 = x \right] = \mathbb{E} \left[ \sum_{j=1}^{\infty} e^{-\hat{T}_j} \hat{Y}_j \bigg| \hat{H}_0, \hat{Y}_j, \forall j = 1, 2, \ldots \bigg| \hat{H}_0 = x \right] \]

\[ = \mathbb{E} \left[ \sum_{j=1}^{\infty} \hat{Y}_j \mathbb{E} \left[ e^{-\hat{T}_j} \mid \hat{H}_0, \hat{Y}_j, \forall j = 1, 2, \ldots \big| \hat{H}_0 = x \right] \right] \]

\[ = \mathbb{E} \left[ \sum_{j=1}^{\infty} e^{-\sum_{k=1}^{j} \delta \hat{Q}(\hat{H}_{k-1})} \hat{Y}_j \bigg| \hat{H}_0 = x \right] \]

\[ = \mathbb{E} \left[ \sum_{j=1}^{\infty} e^{-\sum_{k=1}^{j} \delta \hat{Q}(\hat{H}_{k-1})} \mu_{\hat{\pi}(\hat{H}_j)} \delta \bigg| \hat{H}_0 = x \right]. \]

Observe the above CLV function and compare it with the CLV in the original model Eq. (9), restated here for convenience:

\[ V(x, \pi) = \mathbb{E} \left[ \int_{0}^{\infty} \mu_{\pi(H_t)} e^{-\int_{0}^{t} \hat{Q}(H_s) ds} dt \bigg| H_0 = x \right]. \]  

(38)

The new CLV function in Eq. (37) can be viewed as a discrete time version of the original continuous time problem in Eq. (38), if we consider the limit \( \epsilon \to 0^+ \) for hazard rate function \( \hat{Q}(x) \) given by Eq. (36). Therefore, we conclude that the insights produced by our “contractual” setting extend to more settings, such as the discrete-time “non-contractual” one presented here.
References


Online Appendix

Our appendix is subdivided into 5 sections. In Appendix A, we describe a class of admissible polices called interval policies. The optimal policies we find are members of this class. In Appendix B, we introduce and prove several supporting lemmas. Appendices C–F contain the proofs and technical material associated with Sections 4–7, respectively.

A Interval policies: A class of admissible control policies

In this section, we formally specify the stochastic processes for reward and happiness resulting from a broad set of “interval” control policies, by drawing upon the work of Salins and Spiliopoulos [28] on Markov processes with spatial delay. As a consequence, all policies in this class are admissible and in fact, the optimal policies we find, the myopic policy and the sandwich policy, are among the simplest members of this class, consisting of one and three intervals respectively.

We first establish a technical lemma, which will be used in the proof of Lemma 1, as well as later in the proofs of Proposition 2 and Proposition 4.

Lemma A.1. Fix any admissible policy \( \pi \in \Pi \) and starting happiness level \( x \in \mathbb{R} \). Let \( H_t \) be the corresponding happiness process as defined in Eq. (3) and let \( u_t \) be the corresponding action process. For \( y \in \mathbb{R} \), denote by \( L^H(t,y) \) the symmetric local time of \( H_t \) at \( y \). Then for any \( t \geq 0 \) and any \( y \in \mathbb{R} \), it holds that

\[
L^H(t,y) < \infty \text{ almost surely.}
\]

Let \( E \) be any countable set in \( \mathbb{R} \). Then for any \( t \geq 0 \) it holds that

\[
\int_0^t \mathbb{1}\{H_s \in E \ & u_s \neq S\} ds = 0 \text{ almost surely.}
\]

Proof. By the definition of local time (see Definition A.1 in Salins and Spiliopoulos [28]), we have

\[
L^H(t,y) = |H_t - y| - (x - y) - \int_0^t \text{sign}(H_s - y)(\mu_{u_s} - H_s) ds - \int_0^t \text{sign}(H_s - y) \sigma_{u_s} dB_s,
\]

which implies

\[
E L^H(t,y) \leq E|H_t| + |y| + |x| + |y| + \int_0^t (\mu_R + E|H_s|) ds.
\]

If we take expectations on both sides of Eq. (3), one can easily see that \( E H_t \leq x + \mu_R t \). Therefore the above inequality implies that \( E L^H(t,y) < \infty \), and hence \( L^H(t,y) < \infty \) almost surely. To complete the proof it remains to show the occupation time \( \int_0^t \mathbb{1}\{H_s \in E \ & u_s \neq S\} ds = 0 \), or equivalently, \( \int_0^t \mathbb{1}\{H_s \in E\}(dH_s)^2 = 0 \) almost surely. This is indeed true, since if otherwise, \( L^H(t,y) = \infty \) with positive probability by definition of local time (see Definition A.1 in Salins and Spiliopoulos [28]), contradicting with \( L^H(t,y) < \infty \) almost surely. \( \square \)
Lemma A.1 is stated in terms of definitions in the Base model (Section 3). However, this requirement is not restrictive. The proof above works for processes (the happiness process and action process) as defined in the Investor model (Section 5) and in the GBM setting (Section 6) as well, under any admissible policy defined in these alternative settings. Therefore, the results of Lemma A.1 — $L^H(t, y) < \infty$ and $\int_0^t \mathbf{1}\{H_s \in E \& u_s \neq S\}ds = 0$ almost surely — extend to the Investor model and to the GBM setting.

Next we prove Lemma 1.

**Proof of Lemma 1.** As mentioned in Section 3, we expect that for our stochastic control problem it is sufficient to consider the class of stationary Markov policies since the happiness and reward processes evolve based only on the current happiness and the service mode chosen, see Eqs. (1) and (2). A stationary Markov policy $\pi$ is a mapping from happiness to service mode, i.e., $\pi : \mathbb{R} \rightarrow \{R, S\}$. Here, we will focus on a large subclass of stationary Markov policies: this class of *interval* policies includes each stationary Markov policy $\pi$ that is piecewise constant with a countable number of pieces, and such that the Safe mode is adopted at each boundary between a Safe piece and a Risky piece (in other words, each of the Risky pieces is an open set). We will ignore customer abandonment in this section, and specify the evolution of the happiness process for all time in $[0, \infty)$. We would have been able to use Salins and Spiliopoulos [28] directly to specify the happiness process for any control policy in our class except for one wrinkle — Salins and Spiliopoulos [28] assumes that the volatility of the stochastic process is bounded below by a positive constant everywhere. So the strategy we adopt is to use Salins and Spiliopoulos [28] to specify the happiness process during the time intervals when it is in the closure of a Risky piece, and to combine this specification with the obvious, deterministic trajectory converging exponentially at rate 1 to $\mu_S$ when the happiness is in the interior of a Safe piece.

Our specification of the happiness process is inductive on the number of pieces. The simplest case is where there is just one piece: If the policy is to use the Risky mode everywhere, the happiness process is simply an O-U process with parameters $\mu_R$ and $\sigma_R$. If the policy is to use the Safe mode everywhere, the happiness process is deterministic and converges exponentially at rate 1 to $\mu_S$.

Now let us consider the case where the policy has a Risky piece and a Safe piece. One possible case (among four possibilities, we will consider the other three cases below) is that the policy uses the Risky mode on $(-\infty, \theta)$ and the Safe mode on $[\theta, \infty)$ for some $\theta > \mu_S$. Now if the happiness starts at $x > \theta$, there is an initial deterministic transient

$$H_t = \mu_S + (x - \mu_S)e^{-t}, \quad \text{for all } t \in \left[0, \log\left(\frac{x - \mu_S}{\theta - \mu_S}\right)\right].$$

(39)

where the happiness decays exponentially at rate 1 from $x$ to $\theta$ under the Safe mode, arriving at $\theta$ at time $t_0 = \log\left(\frac{x - \mu_S}{\theta - \mu_S}\right)$. After this transient, the happiness process resembles a reflected O-U process with parameters $\mu_R$ and $\sigma_R$ that lives in $(\infty, \theta)$ and is reflected downwards at $\theta$, but with a crucial difference. The reflected O-U process spends a measure zero of time at the reflecting boundary. In our process, the time spent at the reflecting boundary has positive measure with probability one, conditioned on the customer happiness being $q$ at some time. While this process is less well-known, such a *delayed reflected* process was introduced by Skorokhod [30], and shown to be
a semimartingale by Salins and Spiliopoulos [28]. The measure of time spent by a delayed reflected process at a reflection boundary is proportional to the local time of the process at the boundary, and the constant of proportionality (termed the delay parameter in Salins and Spiliopoulos [28]) is the inverse of the drift at the boundary. In our setting the drift is negative with magnitude $\theta - \mu_S$ under the Safe mode at the boundary happiness $\theta$, and hence the delay parameter is $1/(\theta - \mu_S)$.

Having established this connection, we rely on Theorem 3.4 from Salins and Spiliopoulos [28] to conclude that the distribution of the happiness process for $[t_0, \infty)$ can be defined as the law of the unique solution (guaranteed to exist) of the SDE and local time pair

$$dH_t = (\mu_R - H_t)1\{H_t < \theta\}dt + \sigma_R1\{H_t < \theta\}dB_t - L^H(dt, \theta),$$

$$H_{t_0} = \theta,$$

$$\frac{1}{\theta - \mu_S}L^H(t, \theta) = \int_{t_0}^t 1\{H_s = \theta\}ds,$$

where $L^H(t, \theta)$ is the symmetric local time of the happiness process at $\theta$. And being a (weak) solution to an SDE, the happiness process is guaranteed to be a semimartingale. Notice that we also obtain a unique specification of (the distribution of) time spent at the boundary: this will be helpful when the boundary is exactly at $q$, so that time spent at the boundary is identical to time spent in the satisfied zone. For example, integrals such as $\int_0^t Q(H_s)ds$ will be meaningful even when the boundary is at $q$; this integral will simply be the time spent away from the boundary inside the unsatisfied zone (since the hazard rate there is 1, whereas it is 0 in the satisfied zone). As a result, we can, in fact, conclude that $(H_t)_{t \in [0, \bar{t}]}$ and $(S_t)_{t \in [0, \bar{t}]}$ are semimartingales adapted to $(\mathcal{F}_t)_{t \geq 0}$ for any $\bar{t} < \infty$, since the survival probability process $S_t = e^{-\int_0^t Q(H_s)ds}$ is uniquely specified in terms of the happiness process.

Now we discuss the other three cases for one Risky and one Safe piece along similar lines. In the interest of space, we focus on the happiness process $H_t$ and skip $Y_t$ and $S_t$, though they are immediate to specify.

- If the policy uses the Risky mode on $(-\infty, \theta)$ and the Safe mode on $[\theta, \infty)$ for some $\theta \leq \mu_S$: If the happiness starts at $x \geq \theta$, then it decays exponentially at rate 1 towards $\mu_S$ as per Eq. (39) for all $t \in [0, \infty)$. If the happiness starts at $x < \theta$ it follows an O-U process with parameters $\mu_R$ and $\sigma_R$ until the (finite w.p. 1) time at which it hits $\theta$, after which it remains in the Safe piece forever and decays exponentially at rate 1 towards $\mu_S$.

- If the policy uses the Safe mode on $(-\infty, \theta]$ and the Risky mode on $(\theta, \infty)$ for some $\theta \leq \mu_S$: this is analogous to the case discussed at length above. The trajectory is deterministic upwards inside the Safe piece, whereas it is a delayed reflected O-U process inside the Risky piece with upward reflection with delay parameter $1/(\mu_S - \theta)$ at the boundary $\theta$.

- If the policy uses the Safe mode on $(-\infty, \theta]$ and the Risky mode on $(\theta, \infty)$ for some $\theta > \mu_S$: this is analogous to the first bullet with a deterministic trajectory for all time inside the Safe piece, and an O-U process in the Risky piece such that after the first instant when it hits the boundary $\theta$, the happiness thereafter remains within the Safe piece forever.
At this point, it is straightforward to see how the construction extends inductively to any countable number of pieces with the Risky pieces being open.

Since $H_t$ is a semimartingale, we can apply the Itô-Tanaka formula (see Theorem A.3 in Salins & Spiliopolous [28]) for any function $f$ that is continuously differentiable on $\mathbb{R}$ and twice continuously differentiable on $\mathbb{R} \setminus \mathcal{E}$ for some countable set $\mathcal{E}$:

$$f(H_t) = f(H_0) + \int_0^t f'(H_s) dH_s + \frac{1}{2} \int_0^t f''(H_s) \mathbb{1}_{\{H_s \notin \mathcal{E}\}} (dH_s)^2 + \frac{1}{2} \sum_{y \in \mathcal{E}} (f'_r(y) - f'_l(y)) L^H(t, y),$$

where $f'_r$ and $f'_l$ are the right and left derivatives of $f$, and $L^H(t, y)$ is the symmetric local time of $H_t$ at value $y$. Since $f$ is continuously differentiable everywhere on $\mathbb{R}$, and since local time $L^H(t, y) < \infty$ almost surely by Lemma A.1, we can conclude that $\sum_{y \in \mathcal{E}} (f'_r(y) - f'_l(y)) L^H(t, y) = 0$ almost surely. Therefore, from the above equation and by definition of $dH_t$ in Eq. (2), almost surely,

$$f(H_t) = f(H_0) + \int_0^t f'(H_s) (\mu_{u_s} - H_s) ds + \int_0^t f'(H_s) \sigma_{u_s} dB_s + \frac{1}{2} \int_0^t f''(H_s) \mathbb{1}_{\{H_s \notin \mathcal{E}\}} \sigma_{u_s}^2 ds.$$

This completes the proof of Lemma 1.

Note that the definition of interval policies in Definition 1 as well as the happiness process construction in the proof of Lemma 1 apply to both the Base model and the GBM setting (Section 6). This demonstrates that interval policies are admissible in both the Base model and the GBM setting. In fact, interval policies for the Investor model (see Definition 2 below) are also admissible. Here, the policy may choose an arbitrary blend of the service modes at different points in the “Risky” pieces as long as the fraction of the Risky mode $\lambda(x)$ is Lipschitz continuous in the happiness level $x$ within each piece, and $\lambda(x)$ is uniformly bounded below everywhere on the union of all “Risky” pieces (so that Salins and Spiliopoulos [28] still applies on the closure of each “Risky” piece). It will turn out (see Theorem 3) that the optimal policy for the Investor model is either myopic (pure Risky everywhere) or has a modified sandwich structure with two “Risky” pieces: one piece below $q$ where the firm uses purely the Risky policy, and the other piece above $q$ where the firm uses a blend between Risky and Safe for an interval and pure Risky at all happiness values to the right of the interval. The proportion of Risky to the right of $q$ is continuous (and increasing) in happiness and bounded below so Salins and Spiliopoulos [28] applies in that piece (as well as the other piece).

**Definition 2** (Interval policy in Investor model). In the Investor model (see Section 5), a policy $\lambda$ is an interval policy if:

- it is stationary Markov, that is, the corresponding action process is given by a mapping from current happiness to the proportion of the Risky mode, which we denote by $p_t = \lambda(H_t)$.

- there is a partition of the happiness real line into a countable number of intervals, such that $\lambda(\cdot)$ is Lipschitz continuous within each interval, and that there exists some $c > 0$ such that $\lambda(x) \in \{0\} \cup [c, 1]$ for all $x \in \mathbb{R}$.
B Supporting Lemmas

We present a sequence of lemmas in this appendix that are used to prove results throughout the paper. We start with the definition of a reflected O-U process for reader’s reference. This is from Reed et al. [24] Definition 3.1. WLOG, we set $\gamma = 1$.

**Definition 3 (Reflected O-U Process).** Let $B = (B_t : t \geq 0)$ be a standard Brownian motion, and let $\sigma > 0$, and $\theta \in \mathbb{R}$. We say that the process $Z$ is a $(\sigma, \theta)$ reflected O-U process starting from $x \geq 0$, if the following four conditions are satisfied.

1. $Z_t = x + \theta t - \int_0^t Z_s ds + \sigma B_t + L_t$ for $t \geq 0$,
2. $Z_t \geq 0$ for $t \geq 0$,
3. $L$ is non-decreasing with $L_0^- = 0$,
4. $\int_0^\infty 1\{Z_t > 0\} dL_t = 0$.

Then we present the following results on the stochastic ordering of reflected O-U processes.

**Lemma B.1.** Let $X_t$ be a $(\sigma_0, \theta_0)$ reflected O-U process on $[0, \infty)$ starting from $x > 0$. Also define the unreflected process $Y_t$ starting from $y \leq x$,

$$Y_t = y + \theta_0 t - \int_0^t Y_s ds + \sigma_0 B_t.$$  \hspace{1cm} (40)

Then $X_t \geq_{st} Y_t$ and $X_t \geq_{st} Y_t^+$ for all $t \geq 0$, where $\geq_{st}$ is the usual stochastic order and $Y_t^+$ is the positive part of $Y_t$.

**Proof.** If $X_t \geq_{st} Y_t$ then $X_t \geq_{st} Y_t^+$ follows easily since $X_t \geq 0$. Let $B = (B_t : t \geq 0)$ be a standard Brownian motion and let

$$X_t = x + \theta_0 t - \int_0^t X_s ds + \sigma_0 B_t + L_t, \hspace{1cm} (41)$$

$$Y_t = y + \theta_0 t - \int_0^t Y_s ds + \sigma_0 B_t, \hspace{1cm} (42)$$

where $L_t$ is the local time processes as defined in Definition 3. It suffices to show $X_t \geq Y_t$ for all $t \geq 0$.

Suppose there exists $t > 0$ such that $X_t < Y_t$. Then by continuous paths there exists $0 \leq s < t$ such that $X_s = Y_s$, and $0 \leq X_u < Y_u$ for all $s < u < t$.

By definition, we have

$$X_t - Y_t = X_s + \theta_0(t - s) - \int_s^t X_u du + \sigma_0(B_t - B_u) + (L_t - L_s) - Y_s - \theta_0(t - s) + \int_s^t Y_u du - \sigma_0(B_t - B_u)$$

$$= (X_s - Y_s) + \int_s^t (Y_u - X_u) du + (L_t - L_s) > L_t - L_s.$$
Observe the last inequality follows from $X_s = Y_s$ and $Y_u > X_u \geq 0$ for all $s < u < t$. Also the last term is nonnegative since $L_t$ is nondecreasing. This is a contradiction with $X_t < Y_t$. \hfill $\square$

Lemmas B.2 and B.3 are key to the characterization of the optimal Risky proportion $\lambda^*$ (see Lemma 6).

**Lemma B.2.** Let $X(y)$ be as defined in Lemma 6. Then $X(y)$ is strictly decreasing on $\left[\frac{\mu_S+\mu_R}{2\theta_I-\mu_S-\mu_R}, \infty\right)$. Moreover, $X(\frac{\mu_S+\mu_R}{2\theta_I-\mu_S-\mu_R}) = \theta_I$ and $\lim_{y \to \infty} X(y) = \mu_S$.

**Proof of Lemma B.2.** First we restate $X(y)$ here:

$$X(y) = e^{-2a\left(\frac{\log(y+1)+\frac{1}{y+1}}{y+1}\right)}K_0 - (-2a)^{1+2a}be^{-2a\left(\frac{\log(y+1)+\frac{1}{y+1}}{y+1}\right)}\Gamma\left(-2a, -\frac{2a}{y+1}\right). \quad (43)$$

Then one can verify that $X\left(\frac{\mu_S+\mu_R}{2\theta_I-\mu_S-\mu_R}\right) = \theta_I$ and $\lim_{y \to \infty} X(y) = b$. Recall by definition of $\theta_I$ in Lemma 5 that $\theta_I > \frac{\mu_S+\mu_R}{2} > \mu_S$. Next we want to show that $X'(y) < 0$ on $\left[\frac{\mu_S+\mu_R}{2\theta_I-\mu_S-\mu_R}, \infty\right)$. Compute this derivative:

$$X'(y) = 2aye^{-\frac{2a}{y+1}(y+1)^{-2(a+1)}F(y)}$$

where

$$F(y) = \frac{be^{-\frac{2a}{y+1}(y+1)^{2a+1}}}{y} + 2^{a+1}(-a)^{2a+1}b\Gamma\left(-2a, -\frac{2a}{y+1}\right) - K_0.$$ 

Observe that for $y \in \left[\frac{\mu_S+\mu_R}{2\theta_I-\mu_S-\mu_R}, \infty\right)$, we have $2aye^{-\frac{2a}{y+1}(y+1)^{-2(a+1)}} > 0$. Therefore we want to show $F(y) < 0$ on $\left[\frac{\mu_S+\mu_R}{2\theta_I-\mu_S-\mu_R}, \infty\right)$. In fact, this is true since

$$F'(y) = -\frac{be^{-\frac{2a}{y+1}(1+y)^{2a}}}{y^2} < 0$$

for $y \geq \frac{\mu_S+\mu_R}{2\theta_I-\mu_S-\mu_R} > 0$ and

$$F(\frac{\mu_S+\mu_R}{2\theta_I-\mu_S-\mu_R}) = \frac{e^{2a}g^{2a}(bg - g\theta_I + \theta_I)}{g-1} < 0,$$

where the last step follows from $g = \frac{2\theta_I}{2\theta_I-\mu_S-\mu_R} > 1$ and $b = \mu_S < \frac{\mu_S+\mu_R}{2}$. We have thus completed the proof. \hfill $\square$

**Lemma B.3.** The function $\lambda^*(x)$ as defined in Lemma 6 is strictly increasing on $[q, \theta_I]$, and $\lambda^*(q) > 0$, $\lambda^*(\theta_I) = 1$.

**Proof of Lemma B.3.** Let $p^*(\cdot)$ be defined as Eq. (20), but on the domain $(\mu_S, \theta_I]$. Then on $(\mu_S, \theta_I]$, we have

$$p^*(x) = \frac{(\mu_S - \mu_R)(G(x) + 1)}{\sigma_R^2 G'(x)}.$$
Also from Lemma 6, $G(x)$ satisfies

$$(G(x) + 1)^2 + 2a(x - b)G(x)G'(x) - 2abG'(x) = 0.$$ 

on this interval. Therefore we have

$$G'(x) = \frac{(G(x) + 1)^2}{2a(b - (x - b)G(x))}$$

on $(\mu_S, \theta_I]$, and thus we can rewrite $p^*(x)$ as

$$p^*(x) = \frac{2a(\mu_S - \mu_R)(b - (x - b)G(x))}{\sigma_R^2(G(x) + 1)}.$$  \hspace{1cm} (44)

Since $\mu_S < \mu_R$, to show $p^*(x)$ is strictly increasing on $(\mu_S, \theta_I]$ is equivalent to showing $\frac{b - (x - b)G(x)}{G(x) + 1}$ is strictly decreasing on $(\mu_S, \theta_I]$. Also $G(\cdot)$ is the inverse function of $X(\cdot)$, and by Lemma B.2, $X$ is strictly decreasing on $[g - 1, \infty)$, therefore it is equivalent if we can show $\frac{b - y(X(y) - b)}{y + 1}$ is strictly increasing on $[g - 1, \infty)$.

Denote $L(y) \triangleq \frac{b - y(X(y) - b)}{y + 1}$. Compute its derivative:

$$L'(y) = 2ae^{-\frac{2a}{y+1}}(y + 1)^{-2}a^{-3} (2a)^2a(-a)^2b(2ay^2 - y - 1) \Gamma \left( -2a, -\frac{2a}{y + 1} \right)$$

$$- 2abye^{-\frac{2a}{y+1}}(y + 1)^{2a+1} + K_0 (2ay^2 - y - 1).$$

We want to show that $L'(y) > 0$ on $[g-1, \infty)$. Consider $y$ sufficiently large such that $2ay^2 - y - 1 > 0$. We want to show

$$f(y) = 2^{2a+1}a(-a)^{2a}b \Gamma \left( -2a, -\frac{2a}{y + 1} \right) - \frac{2abye^{-\frac{2a}{y+1}}(y + 1)^{2a+1}}{2ay^2 - y - 1} + K_0 > 0.$$ 

One can check that $f'(y) = \frac{2ab(y+2)e^{-\frac{2a}{y+1}}(y+1)^{2a+1}}{(-2ay^2+y+1)^2} > 0$. Also

$$f(g - 1) = 2^{2a+1}a(-a)^{2a}b \Gamma \left( -2a, -\frac{2a}{g} \right) - \frac{2ab(g - 1)e^{\frac{2a}{g}}g^{2a+1}}{2a(g - 1)^2 - g} + K_0$$

$$= -\frac{2e^{\frac{2a}{g}}g^{2a+1}(2a(g - 1)^2(\mu_S - \mu_R) + g(\mu_R + \mu_S))}{2(g - 1)(2a(g - 1)^2 - g)}.$$

If we can show $g$ is such that $2a(g - 1)^2 - g > 0$ (so that $-2ay^2 + y + 1 > 0$ for all $y \geq g - 1$) and $2a(g - 1)^2(\mu_S - \mu_R) + g(\mu_R + \mu_S) \leq 0$, then we are done. In fact, we can check that the second inequality implies the first, therefore we only need to show the second one is true. Now let $\alpha \triangleq \frac{1}{2} \sqrt{1 + \beta^2}$, $\beta \triangleq \frac{\mu_S + \mu_R}{2\mu_R}$, $\theta_I \triangleq \frac{\theta_I - \mu_R}{\sigma_R}$, and recall the definition of $g \triangleq \frac{2\theta_I}{2\sigma_I - \mu_S - \mu_R}$. After solving the desired inequality on $g$, we can get an equivalent desired inequality on $\hat{\theta}_I$:

$$\hat{\theta}_I \leq \frac{\alpha(\beta - 2) + \sqrt{\alpha^2\beta^2 + 2\beta(1 - \beta)}}{2(1 - \beta)}.$$ 

7
By Lemma 5, \( \hat{\theta}_I \) is the unique root of \( F(\theta) = (x + \alpha)e^{\theta^2}\text{erfc}(\theta) - \frac{\beta}{\sqrt{\pi}} \) on \((-\alpha, \infty)\). Then apply Lemma B.10, we get that the desired inequality of \( \hat{\theta}_I \) is true. Therefore we have proved that \( p^*(x) \) is strictly increasing on \((\mu_S, \theta_I]\).

Finally, consider \( \lim_{x \to \mu_S^+} p^*(x) \) and \( p^*(\theta_I) \). Since \( G(\cdot) \) is the inverse function of \( X(\cdot) \), then by Lemma B.2 we know that \( \lim_{x \to \mu_S^+} G(x) = \infty \), and \( G(\theta_I) = \frac{\mu_S + \mu_R}{2\theta_I - \mu_S - \mu_R} \). Combine with Eq. (44), we get \( \lambda^*(q) = p^*(q) > 0 \) and \( \lambda^*(\theta_I) = p^*(\theta_I) = 1 \).

### B.1 Properties of the error function

Lemmas B.4 through B.20 are results related to the error function \( \text{erf}(\cdot) \) or its complement \( \text{erfc}(\cdot) \).

**Lemma B.4.** \( e^{x^2}\text{erfc}(x)(1 + 2x^2) - \frac{2x}{\sqrt{\pi}} \geq 0 \) for any \( x \in \mathbb{R} \).

**Proof.** The Chernoff-type lower bound for \( \text{erfc}(x) \) by Chang et al. [8] gives:

\[
\text{erfc}(x) \geq \sqrt{\frac{2e^\sqrt{\beta - 1}}{\beta}} e^{-\beta x^2} \text{ for } \beta > 1.
\]

Choose \( \beta(x) = 1 + \frac{1}{2x^2} \). Then,

\[
e^{x^2}\text{erfc}(x)(1 + 2x^2) - \frac{2x}{\sqrt{\pi}} \geq \frac{1}{\sqrt{\pi}} \left( \sqrt{\frac{2e^\sqrt{\beta(x) - 1}}{\beta(x)}} e^{x^2(\beta^{-1}(x) - 1)} (1 + 2x^2) - 2x \right) = 0\]

\[\blacksquare\]

**Lemma B.5.** The following limits are true:

1. \( e^{x^2}\text{erfc}(x) \to 0 \) as \( x \to \infty \);
2. \( xe^{x^2}\text{erfc}(x) \to \frac{1}{\sqrt{\pi}} \) as \( x \to \infty \).

**Proof.** Consider the asymptotic expansion in [25]:

\[
\text{erfc}(x) = \frac{e^{-x^2}}{x\sqrt{\pi}} (1 + o\left(\frac{1}{x^2}\right)). \tag{45}
\]

Then put (45) into LHS, the limit results follow. \[\blacksquare\]

**Lemma B.6.** \( r(x) = (x + \alpha)e^{\theta^2}\text{erfc}(x) \) is first increasing then decreasing on \((-\alpha, \infty)\) if \( \alpha > 0 \); and increasing on \((-\alpha, \infty)\) if \( \alpha \geq 0 \).

**Proof.** \( r(-\alpha) = 0 \). Also from Lemma B.5, \( r(\infty) = \frac{1}{\sqrt{\pi}} \) follows easily. Let

\[
f(x) = \text{erfc}(x), \quad g(x) = \frac{e^{-x^2}}{x + \alpha}
\]

on \((-\alpha, \infty)\). Then

\[
r(x) = \frac{f(x)}{g(x)}. \tag{46}
\]
Consider the “derivative ratio”

\[ \rho(x) \triangleq \frac{f'(x)}{g'(x)} = \frac{2(x + \alpha)^2}{\sqrt{\pi}(1 + 2\alpha x + 2x^2)}. \]  

Then

\[ \rho'(x) = \frac{4(\alpha + x)(1 - \alpha^2 - \alpha x)}{\sqrt{\pi}(2x^2 + 2\alpha x + 1)^2} \]  

Case 1: \( \alpha > 0 \).
\( \rho'(x) > 0 \) on \((-\alpha, -\alpha + \frac{1}{\alpha})\), and \( \rho'(x) < 0 \) on \((-\alpha + \frac{1}{\alpha}, \infty)\). Hence \( \rho(x) \) first increases then decreases on \((-\alpha, \infty)\). Also both \( f \) and \( g \) vanish at \( \infty \). Use “L’Hospital-type rules for monotonicity” ([22] Proposition 4.3), \( r(x) \) first increases then decreases on \((-\alpha, \infty)\).

Case 2: \( \alpha \leq 0 \).
\( \rho'(x) > 0 \) on \((-\alpha, \infty)\). Hence \( \rho(x) \) increases on \((-\alpha, \infty)\). Use “L’Hospital-type rules for monotonicity” (Pinelis [22] Proposition 4.1), \( r(x) \) also increases on \((-\alpha, \infty)\).

The next three corollaries follows easily from Lemma B.6.

**Corollary B.7.** \( r(x) = (x + \alpha)e^{x^2} \text{erfc}(x) - \frac{\beta}{\sqrt{\pi}} \) has a unique root (denote by \( x_0 \)) on \((-\alpha, \infty)\) for any \( 0 < \beta < 1 \). Moreover, \( r(x) < 0 \) on \((-\alpha, x_0)\), \( r(x) > 0 \) on \((x_0, \infty)\).

**Corollary B.8.** \( r(x) = -xe^{x^2} \text{erfc}(x) + \frac{1}{\sqrt{\pi}} > 0 \) for any \( x \in \mathbb{R} \).

**Corollary B.9.** \( e^{x^2} \text{erfc}(x) \) is monotonically decreasing in \( x \).

**Proof.** Take the derivative of \( e^{x^2} \text{erfc}(x) \), we get \( 2xe^{x^2} \text{erfc}(x) - \frac{2}{\sqrt{\pi}} \). Then apply Corollary B.8. \( \square \)

**Lemma B.10.** Let \( r(x) \) and \( x_0 \) be as defined in Corollary B.7. Moreover suppose \( \alpha > 0 \) and \( 0 < \beta < 1 \). Then we have the following inequality:

\[ x_0 \leq \frac{\alpha(\beta - 2) + \sqrt{\alpha^2 \beta^2 + 2\beta(1 - \beta)}}{2(1 - \beta)}. \]

**Proof.** Denote \( x_1 \triangleq \frac{\alpha(\beta - 2) + \sqrt{\alpha^2 \beta^2 + 2\beta(1 - \beta)}}{2(1 - \beta)}. \) Since \( 0 < \beta < 1 \), we have \( x_1 > \frac{\alpha(\beta - 2) + \alpha \beta}{2(1 - \beta)} = -\alpha \). Also by Corollary B.7, \( \{ x > -\alpha : r(x) \geq 0 \} = [x_0, \infty) \). Hence to show \( x_1 \geq x_0 \), we only need to show \( r(x_1) \geq 0 \).

Denote

\[ f(\beta) \triangleq -\left( \alpha \beta + \sqrt{\alpha^2 \beta^2 + 2\beta(1 - \beta)} \right) \cdot e^{-\left( \alpha(\beta - 2) + \sqrt{\alpha^2 \beta^2 + 2\beta(1 - \beta)} \right)^2/4(\beta - 1)^2} \]

\[ + \text{erfc} \left( \frac{\alpha(\beta - 2) + \sqrt{\alpha^2 \beta^2 + 2\beta(1 - \beta)}}{2(1 - \beta)} \right). \]

Since \( x_1 > -\alpha \), showing \( r(x_1) \geq 0 \) is equivalent as showing \( f(\beta) \geq 0 \) for any \( \beta \in (0, 1) \). By letting \( \hat{\beta}(\beta) \triangleq \frac{\alpha(\beta - 2) + \sqrt{\alpha^2 \beta^2 + 2\beta(1 - \beta)}}{2(\beta - 1)} \), one can verify that \( \lim_{\beta \to 1^-} \hat{\beta}(\beta) = \alpha - \frac{1}{\alpha} \). Then substitute
\( \beta(\beta) \) into \( f(\beta) \) and take the limit, we get \( g(\alpha) \triangleq \lim_{\beta \to 1^-} f(\beta) = \text{erfc}(\frac{1}{2\alpha} - \alpha) - \frac{2ae^{-\left(\alpha - \frac{1}{a}\right)^2}}{\sqrt{\pi}}. \) In fact, \( g(\alpha) > 0 \) for all \( \alpha > 0 \) since \( \lim_{\alpha \to 0^+} g(\alpha) = 0 \) and \( g'(\alpha) = \frac{4a^2e^{-\left(\alpha - \frac{1}{a}\right)^2}}{\sqrt{\pi}} > 0. \) Therefore we have proved \( \lim_{\beta \to 1^-} f(\beta) > 0, \) and if we can also show \( f'(\beta) \leq 0 \) then we have \( f(\beta) > 0 \) for any \( \beta \in (0, 1). \) In fact,

\[
 f'(\beta) = -\frac{(\alpha + \sqrt{\alpha^2 + 2\beta(1 - \beta)})}{\beta \sqrt{\pi}} e^{-\frac{(\alpha(\beta - 2) + \sqrt{\alpha^2 + 2\beta(1 - \beta)})^2}{4(\beta - 1)^2}} < 0.
\]

\( \square \)

**Lemma B.11.** Consider \( q > a > 0. \) Then we have the following inequality:

\[
 (a^2 - 2)\left(1 + e^q \sqrt{\pi}q \left(\text{erf}(q) - \text{erf} \left(\alpha - \frac{1}{a}\right)\right)\right) - 2ae^{-q^2 - \left(\frac{1}{a}\right)^2} q < 0.
\]

**Proof.** Since \( e^q > 0, \) we can divide both sides by \( e^q \) and show the following inequality instead:

\[
 F(q) \triangleq (a^2 - 2)\left(\frac{e^{-q^2}}{q} + \sqrt{\pi} \left(\text{erf}(q) - \text{erf} \left(\alpha - \frac{1}{a}\right)\right)\right) - 2ae^{-\left(\frac{1}{a}\right)^2} q < 0.
\]

In fact, if \( a^2 - 2 \leq 0, \) then since \( \text{erf}(\cdot) \) is an increasing function and that \( q > a > a - \frac{1}{a}, \) \( F(q) < 0 \) is obviously true. Now consider the case where \( a^2 - 2 > 0. \) Compute the derivative of \( F(\cdot):\)

\[
 F'(q) = -\frac{(a^2 - 2) e^{-q^2}}{q^2} < 0 \quad \text{for any} \quad q \in \mathbb{R},
\]

which implies that for \( q > a, \)

\[
 F(q) < F(a) < F(a - \frac{1}{a}) = -\frac{a^3 e^{-\left(\frac{1}{a} - a\right)^2}}{a^2 - 1} < 0.
\]

Therefore we have completed the proof. \( \square \)

**Lemma B.12.** The function \( F(x) = \frac{1}{\sqrt{\pi}} + xe^x \text{erf}(x) \) satisfies \( F(x) > 0 \) for all \( x \in \mathbb{R}. \)

**Proof.** To show \( F(x) > 0 \) for all \( x \in \mathbb{R}, \) it is equivalent to show \( f(x) \triangleq F(x)e^{-x^2} > 0 \) for all \( x \in \mathbb{R}. \) In fact, \( f'(x) = \text{erf}(x). \) Therefore \( f(x) \) is decreasing on the interval \((-\infty, 0),\) increasing on the interval \((0, \infty),\) and the minimum value is obtained at \( x = 0, \) with \( f(0) = \frac{1}{\sqrt{\pi}} > 0. \) Thus we have finished the proof. \( \square \)

**Lemma B.13.** Consider functions

\[
 F_1(x, a) = 2 - 2ax + 2x^2 + (-a + 3x - 2ax^2 + 2x^3) e^x \sqrt{\pi} \text{erf}(x),
\]

\[
 F_2(x, a) = 2 - 2ax + 2x^2 - (-a + 3x - 2ax^2 + 2x^3) e^x \sqrt{\pi} \text{erfc}(x),
\]

\[
 F_3(x, a) = 2 - 2ax + 2x^2 + (-a + 3x - 2ax^2 + 2x^3) e^x \sqrt{\pi}(1 + \text{erf}(x)).
\]

Then for all x and a such that \( a > 0, \) \( x \in (a - \frac{1}{a}, a), \) it is true that \( F_1(x, a) > 0, F_2(x, a) > 0, \) and \( F_3(x, a) > 0. \)
Proof. If \( a > 0 \) and \( x \in (a - \frac{1}{a}, a) \), then it must be that \( a \in \left( \max\{x, 0\}, \frac{x}{2} + \sqrt{1 + \frac{x^2}{4}} \right) \). Therefore it is even sufficient if we can prove \( F_1(x, a) > 0 \), \( F_2(x, a) \), and \( F_3(x, a) \) for \( x \) and \( a \) such that 
\[
a \in \left( x, \frac{x}{2} + \sqrt{1 + \frac{x^2}{4}} \right).\]

Consider the partial derivatives with regard to \( a \):
\[
\frac{\partial F_1}{\partial a}(x, a) = -\sqrt{\pi}e^{x^2} (2x^2 + 1) \text{erf}(x) - 2x,
\]
\[
\frac{\partial F_2}{\partial a}(x, a) = \sqrt{\pi}e^{x^2} (2x^2 + 1) \text{erfc}(x) - 2x,
\]
\[
\frac{\partial F_3}{\partial a}(x, a) = -\sqrt{\pi}e^{x^2} (2x^2 + 1) \text{erfc}(-x) - 2x.
\]

Observe that all of them are independent of \( a \). Therefore \( F_1(x, \cdot), F_2(x, \cdot) \) and \( F_3(x, \cdot) \) are all linear in \( a \). Moreover, by Lemma B.4 we can get \( \frac{\partial F_2}{\partial a}(x, a) \geq 0 \) and \( \frac{\partial F_1}{\partial a}(x, a) \leq 0 \). To prove the result, it then suffices to show \( F_2(x, x) > 0 \), \( F_3(x, \frac{x}{2} + \sqrt{1 + \frac{x^2}{4}}) > 0 \), \( F_1(x, x) > 0 \) as well as \( F_1(x, \frac{x}{2} + \sqrt{1 + \frac{x^2}{4}}) > 0 \) for all \( x \in \mathbb{R} \). We have
\[
F_1(x, x) = 2\sqrt{\pi}e^{x^2} \text{erf}(x) + 2 > 0
\]
by Lemma B.12, and
\[
F_2(x, x) = 2 - 2\sqrt{\pi}e^{x^2} \text{erfc}(x) > 0
\]
by Corollary B.8. Also consider
\[
f_1(x) \triangleq F_1(x, \frac{x}{2} + \sqrt{1 + \frac{x^2}{4}}) = -\frac{1}{2}\sqrt{\pi}e^{x^2} \left( -2x^3 + 2\sqrt{x^2 + 4x^2} + \sqrt{x^2 + 4} - 5x \right) \text{erf}(x)
+ x^2 - \sqrt{x^2 + 4}x + 2,
\]
and
\[
f_3(x) \triangleq F_3(x, \frac{x}{2} + \sqrt{1 + \frac{x^2}{4}}) = -\frac{1}{2}\sqrt{\pi}e^{x^2} \left( -2x^3 + 2\sqrt{x^2 + 4x^2} + \sqrt{x^2 + 4} - 5x \right) \text{erfc}(-x)
+ x^2 - \sqrt{x^2 + 4}x + 2.
\]

It remains to show \( f_1(x) > 0 \) and \( f_3(x) > 0 \) for any \( x \in \mathbb{R} \). In fact, one can check that \(-2x^3 + 2\sqrt{x^2 + 4x^2} + \sqrt{x^2 + 4} - 5x \) has a single root at \( \frac{1}{\sqrt{2}} \), and that \(-2x^3 + 2\sqrt{x^2 + 4x^2} + \sqrt{x^2 + 4} - 5x > 0 \) for \( x < \frac{1}{\sqrt{3}} \) and \(-2x^3 + 2\sqrt{x^2 + 4x^2} + \sqrt{x^2 + 4} - 5x < 0 \) for \( x > \frac{1}{\sqrt{2}} \). Therefore we can consider the cases separately.

First check that \( f_1(\frac{1}{\sqrt{2}}) = 1 > 0 \) and \( f_3(\frac{1}{\sqrt{2}}) \approx 2 > 0 \). Then consider \( x \neq \frac{1}{\sqrt{2}} \). Let \( g_1(x) \triangleq \frac{e^{-x^2}f_1(x)}{-2x^3 + 2\sqrt{x^2 + 4x^2} + \sqrt{x^2 + 4} - 5x} \) and let \( g_3(x) \triangleq \frac{e^{-x^2}f_3(x)}{-2x^3 + 2\sqrt{x^2 + 4x^2} + \sqrt{x^2 + 4} - 5x} \). Compute their derivatives:
\[
g_1'(x) = g_3'(x) = \frac{2e^{-x^2} \left( \sqrt{x^2 + 4} - x \right)}{\sqrt{x^2 + 4} \left( -2x^3 + 2\sqrt{x^2 + 4x^2} + \sqrt{x^2 + 4} - 5x \right)^2} > 0.
\]
Then \( g_1(x) > \lim_{x \to -\infty} g_1(x) = \frac{\sqrt{\pi}}{2} > 0 \) for \( x < \frac{1}{\sqrt{2}} \), \( g_1(x) < \lim_{x \to \infty} g_1(x) = -\frac{\sqrt{\pi}}{2} < 0 \) for \( x > \frac{1}{\sqrt{2}} \), \( g_3(x) > \lim_{x \to -\infty} g_3(x) = 0 \), and \( g_3(x) < \lim_{x \to \infty} g_3(x) = -\sqrt{\pi} \). This means \( f_1(x) > 0 \) and \( f_3(x) > 0 \) everywhere. We have thus completed the proof.

**Lemma B.14.** Consider a function

\[
F(x) = \frac{C}{\sqrt{\pi}} + xe^{x^2}(1 + \text{C erf}(x)),
\]

where \( C \in [-1, 1] \) is some constant. Then \( F(x) \) is non-decreasing in \( x \).

**Proof.** If \( C = 0 \), then the result trivially holds. If \( C = 1 \), then the result is obvious, since \( 1 + \text{erf}(x) \) is positive and increasing. Compute the derivative of \( F \), we get:

\[
F'(x) = e^{x^2}(2x^2 + 1)(1 + \text{erf}(x)) + \frac{2Cx}{\sqrt{\pi}}.
\]

Then if \( C = -1 \), we can apply Lemma B.4 to obtain \( F'(x) \geq 0 \), and therefore \( F \) is non-decreasing. Now it only remains to consider the case where \( C \in (-1, 0) \) and where \( C \in (0, 1) \). Let \( f(x) \triangleq F'(x)e^{-x^2} \) and compute its second order derivative:

\[
f''(x) = 4C\text{erf}(x) + 4 > 0.
\]

Therefore \( f'(x) \) is monotonically increasing, and \( \lim_{x \to -\infty} f'(x) = -\infty \), \( \lim_{x \to \infty} f'(x) = \infty \) if \( C \in (-1, 1) \). Thus \( f(x) \) is first decreasing then increasing on \( \mathbb{R} \) if \( C \in (-1, 1) \). We want to show that \( f(x) \geq 0 \) for all \( x \in \mathbb{R} \), and where \( C \in (-1, 1) \setminus \{0\} \). Consider the minimum of \( f(x) \), which is obtained at some level \( x_0 \in \mathbb{R} \). In other words, at \( x_0 \) we have

\[
f'(x_0) = 4 \left( Cx_0\text{erf}(x_0) + \frac{Ce^{-x_0^2}}{\sqrt{\pi}} + x_0 \right) = 0.
\]

(49)

Also one can check that \( f'(0) = \frac{4C}{\sqrt{\pi}} \), so that \( f'(0) > 0 \) if \( C \in (0, 1) \), and \( f'(0) < 0 \) if \( C \in (-1, 0) \). Recall that \( f'(x) \) is monotonically increasing, hence \( x_0 < 0 \) if \( C \in (0, 1) \) and \( x_0 > 0 \) if \( C \in (-1, 0) \). In both cases \( Cx_0 \neq 0 \), and thus we can divide both sides of Eq. (49) by \( Cx_0 \) and get

\[
\text{erf}(x_0) = \frac{Ce^{-x_0^2} + x_0}{-Cx_0}.
\]

Then we can compute

\[
\min_{x \in \mathbb{R}} f(x) = f(x_0) = (2Cx_0^2 + C)\text{erf}(x_0) + \frac{2Ce^{-x_0^2}x_0}{\sqrt{\pi}} + 2x_0^2 + 1
\]

\[
= \frac{2Ce^{-x_0^2}(x_0^2 - x_0 + 1)}{-x_0}.
\]

Observe that \( x_0^2 - x_0 + 1 > 0 \), and that \( \frac{C}{-x_0} > 0 \) from the above argument. Therefore we have \( f(x) > 0 \) everywhere for \( C \in (-1, 1) \setminus \{0\} \), and the proof is complete.
Lemma B.15. Consider a function
\[ F(z) = 2(a - z) \left( a - 3\theta + 2a\theta^2 - 2\theta^3 \right) - 2e^{z^2-\theta^2} \left( 1 - a\theta + \theta^2 \right) (1 - 2az + 2z^2) + e^{z^2} \sqrt{\pi} \left( a - 3\theta + 2a\theta^2 - 2\theta^3 \right) (1 - 2az + 2z^2) (\text{erf}(\theta) - \text{erf}(z)), \]
where \( a > 0 \) and \( \theta \in (a - \frac{1}{a}, a) \) are constants. Then \( F(z) \leq -2 + 2a^2 - 2a\theta \) for all \( z \in (\theta, a) \).

Proof. Observe that \( F(z) + 2 - 2a^2 + 2a\theta \) is quadratic in \( a \). In fact, as we will show next, it is convex in \( a \). Consider the second order partial derivative of \( G(a, \theta, z) \triangleq F(z) + 2 - 2a^2 + 2a\theta \) with respect to \( a \):
\[ \frac{\partial^2 G(a, \theta, z)}{\partial a^2} = 4\sqrt{\pi} (2\theta^2 + 1) e^{z^2} z (\text{erf}(z) - \text{erf}(\theta)) + 8\theta \left( \theta - z e^{z^2-\theta^2} \right). \]
We will show next that \( g(\theta) \triangleq \frac{\partial^2 G(a, \theta, z)}{\partial a^2} \geq 0 \) for \( \theta \leq z \). Compute \( g''(\theta) \):
\[ g''(\theta) = 16\sqrt{\pi} e^{z^2} z (\text{erf}(z) - \text{erf}(\theta)) + 16. \]
When \( z \geq 0 \), it is obvious that \( g''(\theta) > 0 \) for all \( \theta \leq z \). When \( z < 0 \), since \(-\text{erf}(\cdot) < 1\), we have
\[ g''(\theta) > 16\sqrt{\pi} \left( \frac{1}{\sqrt{\pi}} - ze^{z^2} \text{erfc}(z) \right) \geq 0 \]
by Corollary B.8, which then implies that for \( \theta \leq z \),
\[ g'(\theta) = 16\theta \left( \sqrt{\pi} e^{z^2} z (\text{erf}(z) - \text{erf}(\theta)) + 1 \right) - 16ze^{z^2-\theta^2} \]
\[ \leq g'(z) = 0. \]
Therefore, for \( \theta \leq q \),
\[ g(\theta) = 4\sqrt{\pi} (2\theta^2 + 1) e^{z^2} z (\text{erf}(z) - \text{erf}(\theta)) + 8\theta \left( \theta - z e^{z^2-\theta^2} \right) \]
\[ \geq g(z) = 0, \]
and \( G \) is convex in \( a \). Recall that we want to show \( G(a, \theta, z) \leq 0 \) for all \( a \in \left[ \max\{0, z\}, \frac{\theta}{2} + \sqrt{1 + \frac{\theta^2}{4}} \right] \), \( z \in \left( \theta, \frac{\theta}{2} + \sqrt{1 + \frac{\theta^2}{4}} \right) \), and \( \theta \in \mathbb{R} \). Since \( G \) is convex in \( a \) and \( z \leq \max\{0, z\} \), it is sufficient if we can show \( G(z, \theta, z) \leq 0 \) and \( G\left( \frac{\theta}{2} + \sqrt{1 + \frac{\theta^2}{4}}, \theta, z \right) \leq 0 \) for all \( \theta < z \).
First consider \( f_1(\theta, z) \triangleq G(z, \theta, z) \). Starting from its third order partial derivative with respect to \( \theta \):
\[ \frac{\partial^3 f_1(\theta, z)}{\partial \theta^3} = 4e^{z^2} \left( 3\sqrt{\pi} (\text{erf}(z) - \text{erf}(\theta)) + 2e^{-\theta^2} \right). \]
If \( z \geq 0 \), then obviously \( \frac{\partial^3 f_1(\theta, z)}{\partial \theta^3} > 0 \) for all \( \theta < z \), and \( \frac{\partial^2 f_1(\theta, z)}{\partial \theta^2} < \frac{\partial^2 f_1(\theta, z)}{\partial \theta^2} = -12 < 0 \). On the other hand, if \( z < 0 \), then for all \( \theta < z < 0 \),
\[ \frac{\partial^3 f_1(\theta, z)}{\partial \theta^3} = -8e^{z^2} \theta^2 (2\theta z + 3) < 0, \]
which implies that \( \frac{\partial^3 f_1(\theta, z)}{\partial \theta^3} \) is strictly decreasing in \( \theta \). Compute \( \frac{\partial^3 f_1(z, z)}{\partial \theta^3} = 8z < 0 \), and \( \lim_{\theta \to -\infty} \frac{\partial^3 f_1(\theta, z)}{\partial \theta^3} \) 

\[ 12\sqrt{\pi}e^{\frac{1}{2}} \text{erfc}(-z) > 0, \]

then there is a unique root \( \theta_0 \) of \( \frac{\partial^3 f_1(z, z)}{\partial \theta^3} \) on \((-\infty, z)\), which is also the maximizer of \( \frac{\partial^3 f_1(z, z)}{\partial \theta^3} \) on \((-\infty, z)\). Therefore, \( \theta_0 \) satisfies \( \frac{\partial^3 f_1(\theta_0, z)}{\partial \theta^3} = 0 \), i.e., \( \text{erf}(\theta_0) - \text{erf}(z) = \frac{2e^{-\theta_0^2 z}}{3\sqrt{\pi}} \), and for \( \theta < z < 0 \),

\[
\frac{\partial^2 f_1(\theta, z)}{\partial \theta^2} = 4e^{2z} \left( \sqrt{\pi}(z - 3\theta)(\text{erf}(\theta) - \text{erf}(z)) - 3e^{-\theta^2} \right) \\
\leq 4e^{2z} \left( \sqrt{\pi}(z - 3\theta_0)(\text{erf}(\theta_0) - \text{erf}(z)) - 3e^{-\theta_0^2} \right) \\
= 4e^{2z - \theta_0^2} \left( \frac{2z(z - 3\theta_0)}{3} - 3 \right) \\
< 4e^{2z - \theta_0^2} \left( \frac{-4z^2}{3} - 3 \right) < 0.
\]

We have just shown that \( \frac{\partial^2 f_1(\theta, z)}{\partial \theta^2} < 0 \) for all \( \theta < z \) and \( z \in \mathbb{R} \). This implies that for \( \theta < z \),

\[
\frac{\partial f_1(\theta, z)}{\partial \theta} = \sqrt{\pi}e^{2z} \left( 60\theta^2 - 4\theta z + 3 \right) (\text{erf}(z) - \text{erf}(\theta)) + 2e^{2z - \theta^2} (2z - 3\theta) + 2z \\
> \frac{\partial f_1(z, z)}{\partial \theta} = 0,
\]

and thus

\[
f_1(\theta, z) = \sqrt{\pi}e^{2z} \left( -2\theta^3 - 3\theta + 2\theta^2 z + z \right) (\text{erf}(\theta) - \text{erf}(z)) - 2e^{2z - \theta^2} (\theta^2 - \theta z + 1) \\
- 2z^2 + 2\theta z + z < f_1(z, z) = 0.
\]

We have shown \( f_1(\theta, z) = G(z, \theta, z) < 0 \) for all \( \theta < z \). It only remains to show \( G(\theta + \sqrt{1 + \frac{\theta^2}{a}}, \theta, z) \leq 0 \) for \( \theta < z \). This is equivalent to showing \( f_2(a, z) \triangleq G(a, a - \frac{1}{a}, z) a^3 e^{-z^2} \leq 0 \) for all \( z \in (a - \frac{1}{a}, a] \) and \( a > 0 \). Consider the third order derivative of \( f_2 \) with respect to \( z \):

\[
\frac{\partial^3 f_2(a, z)}{\partial z^3} = 8(a^2 - 2) e^{-z^2} (az + 1).
\]

Observe that since \( z > a - \frac{1}{a} \) and \( a > 0 \), we have \( az + 1 > a^2 > 0 \), and thus \( \frac{\partial^3 f_2(a, z)}{\partial z^3} \leq 0 \) if \( a \in (0, \sqrt{2}] \), \( \frac{\partial^3 f_2(a, z)}{\partial z^3} > 0 \) if \( a > \sqrt{2} \). We write the expression for \( \frac{\partial^2 f_2(a, z)}{\partial z^2} \):

\[
\frac{\partial^2 f_2(a, z)}{\partial z^2} = -4\sqrt{\pi} \left( a^2 - 2 \right) \left( \text{erf} \left( a - \frac{1}{a} \right) - \text{erf}(z) \right) - 4 \left( a^2 - 2 \right) a e^{-z^2} - 8e^{-a^4 - \frac{1}{a^2}} a.
\]

When \( a \in (0, \sqrt{2}] \), \( \frac{\partial^2 f_2(a, z)}{\partial z^2} \leq \frac{\partial^2 f_2(a, a - \frac{1}{a})}{\partial z^2} = -4a^3 e^{-\left(a - \frac{1}{a}\right)^2} < 0 \) for all \( z > a - \frac{1}{a} \). When \( a > \sqrt{2} \),

\[
\frac{\partial^2 f_2(a, z)}{\partial z^2} < \hat{f}_2(a) \triangleq \lim_{z \to \infty} \frac{\partial^2 f_2(a, z)}{\partial z^2} = 4\sqrt{\pi} (a^2 - 2) \text{erfc}(a - \frac{1}{a}) - 8ae^{-\left(a - \frac{1}{a}\right)^2} < 0,
\]

where the last negativity result follows from

\[
\hat{f}_2(a) = 8a \left( e^{-\frac{a^4 - 1}{a^2}} a + \sqrt{\pi} \left( \text{erf} \left( a - \frac{1}{a} \right) + 1 \right) \right) > 0 \quad \text{for} \quad a > \sqrt{2},
\]

which implies that for any \( a > \sqrt{2} \),

\[
\hat{f}_2(a) \to \hat{f}_2(a) = 0.
\]
Then since \( \frac{\partial^2 f_2(a,z)}{\partial a^2} < 0 \) for all \( a > 0 \) and \( z \in (a - \frac{1}{a}, a] \), we have
\[
\frac{\partial f_2(a,z)}{\partial z} = -2\sqrt{\pi} (a^2 - 2) (a - 2z) \left( \text{erf} \left( \frac{1}{a} - a \right) + \text{erf}(z) \right) + 4 (a^2 - 2) e^{-z^2} \\
+ 4ae^{2-\frac{a^4+1}{a^2}} (a - 2z)
\]
and thus
\[
f_2(a,z) = \sqrt{\pi} (a^2 - 2) (2z(z-a)+1) \left( \text{erf} \left( \frac{1}{a} - a \right) + \text{erf}(z) \right) \\
- 2 \left( a^2 - 2 \right) e^{-z^2} (a - z) + 2ae^{2-\frac{a^4+1}{a^2}} (2z(a-z) - 1)
\]
\[
< f_2(a,a - \frac{1}{a}) = 0.
\]
\[\square\]

**Lemma B.16.** Consider a function
\[
F(\theta) = e^{q^2} q - e^{\theta^2} q e^{\theta^2+q^2} \theta \sqrt{\pi} (\text{erfc}(\theta) - \text{erfc}(q)),
\]
where \( q > 0 \) is a constant. Then \( F(\theta) \geq 0 \) for \( \theta < q \).

*Proof.* Since \( \text{erfc}(\cdot) \) is a monotonically decreasing function, if \( \theta < q \), then \( \text{erfc}(\theta) - \text{erfc}(q) \geq 0 \). Therefore if \( \theta \leq 0 \), it is clear that \( F(\theta) \geq 0 \). Now it only remains to show \( F(\theta) \geq 0 \) for \( \theta \in (0, q) \). Since both \( q \) and \( \theta \) are strictly positive, we can equivalently show \( \frac{F(\theta)}{e^{q^2+\theta^2} q \theta} \) is non-negative. In fact, we have
\[
\frac{F(\theta)}{e^{q^2+\theta^2} q \theta} = 1 - \sqrt{\pi} e^{\theta^2} \theta \frac{\text{erfc}(\theta)}{e^{q^2} q} - 1 - \sqrt{\pi} e^{\theta^2} q \frac{\text{erfc}(q)}{e^{q^2} q}.
\]
Therefore we only need to show that \( f(\theta) \equiv \frac{1-\sqrt{\pi} e^{\theta^2} \theta \text{erfc}(\theta)}{e^{q^2} q} \) is monotonically non-increasing in \( \theta \) for \( \theta \in (0, q) \). It is clear that the denominator is positive and increasing in \( \theta \) for \( \theta \in (0, q) \). Also the numerator is decreasing in \( \theta \) by an application of lemma B.6. Therefore \( f(\theta) \) is decreasing in \( \theta \) for \( \theta \in (0, q) \), and we have proved the lemma. \[\square\]

**Lemma B.17.** Consider a function
\[
F(\theta) = 2e^{q^2-\theta^2} q \theta - (1 + 2\theta^2) \left( 1 + e^{q^2} q \sqrt{\pi} (\text{erfc}(\theta) - \text{erfc}(q)) \right),
\]
where \( q > 0 \) is a constant. Then \( F(\theta) < 0 \) for \( \theta < q \).

*Proof.* Since \( q > \theta \) and \( \text{erf}(\cdot) \) is an increasing function, we have \( \text{erf}(q) - \text{erf}(\theta) > 0 \) and hence
\[
F''(\theta) = -4 - 4e^{q^2} q \sqrt{\pi} (\text{erfc}(q) - \text{erfc}(\theta)) < 0.
\]
This implies that
\[
F'(\theta) = 4e^{q^2-\theta^2} q - 4\theta - 4e^{q^2} q \theta \sqrt{\pi} (\text{erfc}(q) - \text{erfc}(\theta))
\]
\[
> F'(q) = 0
\]
on \((\infty, q)\). Therefore \( F(\theta) < F(q) = 0 \) for \( \theta < q \). \[\square\]
Therefore, there exist scalars $x$ which implies

Proof. Since $q > \theta$ and $\text{erf}(\cdot)$ is an increasing function, we have $\text{erf}(q) - \text{erf}(\theta) > 0$ and hence it can be easily observed that $F(\theta) < 0$ for $\theta \leq 0$. Now consider $\theta \in (0, q)$, then $q^2 - \theta^2 > 0$ and we have

$F''(\theta) = -2e^{q^2 - \theta^2}q(3 + 2q^2 - 2\theta^2) < 0,$

which implies

$F'(\theta) = 1 + 2q^2 - 2e^{q^2 - \theta^2}q\theta + 2e^{q^2}q(q + q^2)\sqrt{\pi} (\text{erf}(q) - \text{erf}(\theta))$

$> F'(q) = 1 > 0.$

Therefore, $F(\theta) < F(q) = 0$ for $\theta \in (0, q)$, and hence the proof is done.

**Lemma B.19.** Consider two real-valued functions $f_L(x), f_H(x)$ that are both differentiable with regard to $x$ on $(-\infty, x_0)$ for some $x_0 \in \mathbb{R}$. Suppose that $-\infty < \lim_{x \to -\infty} f_L(x) < \lim_{x \to -\infty} f_H(x) < \infty$ and that $f'_L(x) \leq f'_H(x)$ for all $x$ on $(-\infty, x_0)$. Then $f_L(x) < f_H(x)$ for all $x$ on $(-\infty, x_0)$.

Proof. Denote $y_L \triangleq \lim_{x \to -\infty} f_L(x)$ and $y_H \triangleq \lim_{x \to -\infty} f_H(x).$ Also define $\epsilon \triangleq \frac{y_H - y_L}{3}$. Then $\epsilon > 0$ and there exist scalars $x_L < x_0$ and $x_H < x_0$ such that $|f_L(x) - y_L| < \epsilon$ for all $x \leq x_L$ and $|f_H(x) - y_H| < \epsilon$ for all $x \leq x_H$. This implies that $f_L(x) < f_H(x)$ for all $x \leq \min\{x_L, x_H\}$. Hence we only need to show $f_L(x) < f_H(x)$ for $\min\{x_L, x_H\} < x < x_0$. This is true since $f_L(\min\{x_L, x_H\}) < f_H(\min\{x_L, x_H\})$ and $f'_L(x) \leq f'_H(x)$ for all $x < x_0$.

The next Lemma is used to prove monotonicity results of $\theta_b$ in the optimal policy and the corresponding value function $V^*$ in the Base model.

**Lemma B.20.** Consider a function

$F(y) = \log \left( ye^{y^2} \text{erfc}(-y) + \frac{1}{\sqrt{\pi}} \right).$

Then $F'(y)$ is increasing in $y$.

Proof. By Corollary B.8, $F(y)$ is well-defined for all $y \in \mathbb{R}$. Let $r(y) \triangleq F'(y)$. Want to show $f(y)$ is increasing in $y$. First we write the expression for $r(y)$:

$r(y) = \frac{f(y)}{g(y)},$

where

$f(y) \triangleq \sqrt{\pi} (2y^2 + 1) \text{erfc}(-y) + 2e^{-y^2}y$

and

$g(y) \triangleq \sqrt{\pi} ye^{y^2} + e^{-y^2}.$
Consider the “derivative ratio”:
\[ r_1(y) \triangleq \frac{f'(y)}{g'(y)} = \frac{4 \left( \sqrt{\pi}y \text{erfc}(-y) + e^{-y^2} \right)}{\sqrt{\pi} \text{erfc}(-y)} \]
and the “second-order derivative ratio”:
\[ r_2(y) \triangleq \frac{f''(y)}{g''(y)} = 2 \sqrt{\pi} \text{erfc}(-y) e^{y^2} \]
which is increasing in \( y \) by Corollary B.9. One can also check by using the asymptotic expansion of \( \text{erfc}(x) = \frac{e^{-x^2}}{x \sqrt{\pi}} (1 + o\left(\frac{1}{x^2}\right)) \) that as \( y \to -\infty, f(y), f'(y), g(y) \) and \( g'(y) \) all go to 0. Therefore we can apply the “L’Hospital-type rules for monotonicity” (Pinelis [22] Proposition 4.1) on \( r_1(y) \) and \( r \) to get that \( r(y) \) is increasing in \( y \), i.e., \( F'(y) \) is increasing in \( y \).

C Appendix to the Base Model

In this appendix, we give the proofs for the propositions, lemmas and the theorems of Section 4. These includes Propositions 1-3, Lemmas 2-4, and Theorems 1 and 2.

We first prove Proposition 1, which establishes that the expected customer lifetime and CLV are finite for any interval policy with a finite number of intervals.

**Proof of Proposition 1**. The time customer happiness spends in the unsatisfied zone prior to the customer’s departure is exponentially distributed and has a finite expectation, given that the hazard rate of customer departure is 1 in the unsatisfied zone and zero in the satisfied zone. We want to show the time customer happiness spends in the satisfied zone also has finite expectation. Our approach will be to show that the long-run average ratio of time in satisfied zone to time in unsatisfied zone is finite (or zero).

**Case (i): Policy \( \pi \) uses the Safe mode somewhere in \([\mu_S, q]\)**. First, consider the case where there the policy \( \pi \) uses the Safe mode at some happiness value in \([\mu_S, q]\), i.e., the set \( \{x \in [\mu_S, q] : \pi(x) = S\} \) is nonempty. Define \( l \triangleq \max\{x \in [\mu_S, q - \epsilon] : \pi(x) = 0\} \) where \( \epsilon \in (0, (q - \mu_S)/2) \) is chosen to be small enough that \( \{x \in [\mu_S, q - \epsilon] : \pi(x) = 0\} \) is nonempty. If the starting happiness \( x \leq l \), then it is easy to see that the happiness level never rises above \( l \), hence the customer is always unsatisfied and we are done. So suppose \( x > l \). Let \( \tau \) be the first time at which the happiness reaches \( l \). We will show that \( E\tau < \infty \). This will complete our proof for this case, since the happiness will never rise above \( l \) thereafter.

It is easy to bound the total time spent in the interior of Safe intervals of policy \( \pi \) prior to \( \tau \), since the happiness process decays exponentially towards \( \mu_S \) at such happiness values, and these happiness values exceed \( q - \epsilon > \mu_S + \epsilon \) by definition of \( l \). It remains to bound the total time spent traversing Risky intervals of \( \pi \) above \( l \), including the time spent on delayed reflections at the upper boundary of such intervals (at the lower boundary of a Risky interval above \( l \), the process enters a Safe interval and never again returns to the Risky interval). But this bound is also easy, since the happiness process inside each Risky interval is merely a \((\sigma_R, \mu_R)\) O-U process, possibly
with a delayed reflecting upper boundary (if the interval has an upper boundary) \( b \) where the drift is \( -(b - \mu_S) \). Any standard (unreflected) O-U process has first passage times with finite expectation (see Thomas [32]). Also the fraction of time a delayed reflected O-U process spends at its reflecting boundary is bounded away from one. Hence using the fact that the unreflected O-U process dominates the reflected O-U process as per Lemma B.1, we conclude that the time for our reflected O-U process to reach the lower end of the Risky interval (or \( l \), whichever is larger) has finite expectation. Hence, we have shown \( \mathbb{E} \tau < \infty \).

**Case (ii): Policy \( \pi \) uses the Risky mode everywhere in \([\mu_S, q]\).** In this case, policy \( \pi \) has a Risky interval \((a, b)\) that contains \([\mu_S, q]\). By the argument above, the time taken to enter this Risky interval has finite expectation. Having entered this interval, the happiness process remains within the closure of this interval, and is simply a \((\sigma_R, \mu_R)\) delayed reflected O-U process with drift \( -(b - \mu_S) < 0 \) at the upper boundary and drift \( \mu_S - a > 0 \) at the lower boundary. This process has a steady state distribution [34] with positive measure everywhere in \((a, b)\) and atoms at \( a \) and \( b \). Since \( a < \mu_S \Rightarrow a < q \), it follows that the happiness process spends a positive fraction of its time in the unsatisfied zone in the long run. This completes our proof. \( \square \)

Next we prove Lemma 2 and Lemma 3, which define the boundaries of the sandwich interval.

**Proof of Lemma 2.** Define

\[
\alpha := \frac{\mu_R - \mu_S}{\sigma_R}, \quad \beta := \frac{\mu_S}{\mu_R}, \quad z = \frac{\theta - \mu_R}{\sigma_R}.
\]

Then we have \( \alpha > 0 \) and \( 0 < \beta < 1 \). Also define

\[
r(z) = (z + \alpha)e^{z^2} \text{erfc}(z) - \frac{\beta}{\sqrt{\pi}}.
\]

Then our goal is to show that \( r(z) \) has a unique root on \((-\alpha, \infty)\). This is true by Corollary B.7. \( \square \)

**Proof of Lemma 3.** Let \( a \Delta:= \frac{\mu_S - \mu_R}{\sigma_R}, \quad b = \mu_S - \mu_R, \quad \hat{q} \triangleq \frac{q - \mu_R}{\sigma_R}, \) and \( \hat{\theta} \triangleq \frac{\theta - \mu_R}{\sigma_R} \). Define also

\[
F_{\text{small}}(\hat{\theta}) \triangleq 2 \left( e^{\hat{\theta}^2 - \hat{\theta}^2 \hat{q}} - \hat{\theta} + \hat{\theta}^2 \sqrt{\pi} \hat{q} \hat{\theta} \left( \text{erf}(\hat{\theta}) - \text{erf}(\hat{q}) \right) \right) - \frac{\mu_S}{b(\hat{q} - a)}
\]

and

\[
F_{\text{big}}(\hat{\theta}) \triangleq \left( a - 3\hat{\theta} + 2a\hat{\theta}^2 - 2\hat{\theta}^3 \right) + 2e^{\hat{\theta}^2 - \hat{\theta}^2 \hat{q}} \left( 1 - a\hat{\theta} + \hat{\theta}^2 \right)
\]

\[
+ e^{\hat{\theta}^2} \sqrt{\pi} \hat{q} \left( a - 3\hat{\theta} + 2a\hat{\theta}^2 - 2\hat{\theta}^3 \right) \left( \text{erf}(\hat{q}) - \text{erf}(\hat{\theta}) \right) - \frac{\mu_S \left( 1 - a^2 + a\hat{\theta} \right)}{b(\hat{q} - a)}.
\]

Observe that \( F_{\text{small}}(\hat{\theta}) \) corresponds to the LHS of Eq. (11), multiplied by a factor of 2. Also \( F_{\text{big}}(\hat{\theta}) \) corresponds to the LHS of Eq. (12), multiplied by a factor \( \frac{q - \mu_R}{\sigma_R} \) (\( q - \mu_R > 0 \) since \( q > \mu_S > \mu_R \)). Then proving this lemma is equivalent to showing that exactly one of the following two cases is true: (1) \( F_{\text{small}}(\hat{\theta}) \) has a unique root on \([a, \hat{q}]\) and \( F_{\text{big}}(\hat{\theta}) \) has no root on \((a - \frac{1}{a}, a)\); (2) \( F_{\text{small}}(\hat{\theta}) \) has no root on \([a, \hat{q}]\) and \( F_{\text{big}}(\hat{\theta}) \) has a unique root on \((a - \frac{1}{a}, a)\).
First, observe that

$$F_{\text{small}}(\hat{q}) = -\frac{\mu S}{b(\hat{q} - a)} < 0.$$ 

and that

$$F_{\text{big}}(a - \frac{1}{a}) = \frac{(a^2 - 2) \left(1 + e^{\hat{q}^2} \sqrt{\pi a} \left(\text{erf}(a - \frac{1}{a})\right) - 2ae^{\hat{q}^2}(a - \frac{1}{a})\right)}{-a^3} > 0$$

by Lemma B.11. Since \( a > 0 \), for \( \hat{\theta} > a - \frac{1}{a} \), we have \( 1 - a^2 + a\hat{\theta} > 0 \). Also, since \( F_{\text{big}}(\cdot) \) is continuous at \( a - \frac{1}{a} \), we have \( \lim_{x \to (a - 1/a)^+} \frac{F_{\text{big}}(\hat{\theta})}{1-a^2+a\hat{\theta}} > 0 \) as well. Consider \( F_{\text{small}}(a) \) and \( F_{\text{big}}(a) \). In fact, one can check that

$$F_{\text{small}}(a) = F_{\text{big}}(a) = 2 \left(e^{\hat{q}^2-a^2} - a + e^{\hat{q}^2} \sqrt{\pi a} \left(\text{erf}(a) - \text{erf}(\hat{q})\right)\right) - \frac{\mu S}{b(\hat{q} - a)}.$$ 

Therefore, if we can show that both \( F_{\text{small}}(\hat{\theta}) \) and \( \hat{F}_{\text{big}}(\hat{\theta}) = \frac{F_{\text{big}}(\hat{\theta})}{1-a^2+a\hat{\theta}} \) are strictly decreasing functions of \( \hat{\theta} \) on \((a - \frac{1}{a}, q)\), then we are done. In fact, one can check that \( F_{\text{small}}''(\hat{\theta}) = 2\hat{q}e^{\hat{q}^2-\hat{\theta}^2} > 0 \). Thus, \( F_{\text{small}}'(\hat{\theta}) < F_{\text{small}}(\hat{q}) = -1 < 0 \) for \( \hat{\theta} \in (a - \frac{1}{a}, q) \), and hence \( F_{\text{small}}(\hat{\theta}) \) is strictly decreasing for \( \hat{\theta} \in (a - \frac{1}{a}, q) \).

It remains to be shown that \( \hat{F}_{\text{big}}(\hat{\theta}) \) is strictly decreasing in \( \hat{\theta} \in (a - \frac{1}{a}, q) \). The derivative of \( \hat{F}_{\text{big}} \) is:

$$\hat{F}_{\text{big}}'(\hat{\theta}) = -\frac{e^{-\hat{\theta}^2}(3 - 2a^2 + 2a\hat{\theta})}{(1-a^2+a\hat{\theta})^2} \cdot f(\hat{\theta}, a),$$

where

$$f(\hat{\theta}, a) = e^{\hat{\theta}^2} \left(1 - 2a\hat{\theta} + 2\hat{\theta}^2\right) \left(1 + e^{\hat{\theta}^2} \sqrt{\pi \hat{q}} \left(\text{erfc}(\hat{\theta}) - \text{erfc}(\hat{q})\right)\right) + 2e^{\hat{\theta}^2} \hat{q}(a - \hat{\theta}).$$

We want to show that \( \hat{F}_{\text{big}}'(\hat{\theta}) \leq 0 \) for \( \hat{\theta} \in (a - \frac{1}{a}, q) \). Observe that \( 3 - 2a^2 + 2a\hat{\theta} > 0 \) since \( \hat{\theta} > a - \frac{1}{a} \) and \( a > 0 \). Therefore, we only need to show \( f(\hat{\theta}, a) > 0 \) for \( \hat{\theta} \in (a - \frac{1}{a}, q) \). In fact,

$$\frac{\partial f}{\partial x_2}(\hat{\theta}, a) = 2e^{\hat{\theta}^2} \hat{q} - 2e^{\hat{\theta}^2} \hat{\theta} - 2e^{\hat{\theta}^2+\hat{\theta}^2} \hat{q}\hat{\theta} \sqrt{\pi} \left(\text{erfc}(\hat{\theta}) - \text{erfc}(\hat{q})\right) \geq 0$$

for \( \hat{\theta} < \hat{q} \) by an application of Lemma B.16. Thus, since \( a > 0 \), if we can show \( f(\hat{\theta}, 0) > 0 \) for \( \hat{\theta} \in (a - \frac{1}{a}, q) \), then we are done. Applying Lemma B.17 we get

$$f(\hat{\theta}, 0) = -2e^{\hat{\theta}^2} \hat{q}\hat{\theta} + e^{\hat{\theta}^2} \left(1 + 2\hat{\theta}^2\right) \left(1 + e^{\hat{\theta}^2} \sqrt{\pi} \left(\text{erfc}(\hat{\theta}) - \text{erfc}(\hat{q})\right)\right) > 0.$$ 

\( \square \)

Now we prove Proposition 2, which gives the optimality conditions.
Proof of Proposition 2. To show that a function $\tilde{V}$ as described in Proposition 2 is the optimal value function $V^*$, we will first show that it is an upper bound for $V^*$, and then show that the bound is tight.

To show that $\tilde{V}$ is an upper bound for $V^*$, it suffices to show $\tilde{V}(x) \geq V(x, \pi)$ for all $x \in \mathbb{R}$ and for any admissible policy $\pi \in \Pi$. Now fix any $x \in \mathbb{R}$ and any $\pi \in \Pi$. Define a process $X_t, t \geq 0$ by

$$X_t = \tilde{V}(H_t)e^{-\int_0^t Q(H_s)ds} + \int_0^t e^{-\int_0^s Q(H_z)dz}dY_s,$$

where $H_t$ is the happiness process under policy $\pi$ with $H_0 = x$, and $Y_t$ the corresponding cumulative reward (conditional on no quitting) up to time $t$. Next we will rewrite $X_t$ in integral form.

Since $\tilde{V}$ is continuously differentiable everywhere and twice continuously differentiable almost everywhere except for a countable set $E$ (Condition 3 in the proposition) and since $H_t$ is a semimartingale under the admissible policy, we can apply the Itô-Tanaka formula to obtain that $\tilde{V}(H_t)$ is also a semimartingale (see Eq. (10), denote $u_t$ the corresponding action process under policy $\pi$):

$$\tilde{V}(H_t) = \tilde{V}(x) + \int_0^t \tilde{V}'(H_s)(\mu_{u_s} - H_s)ds + \int_0^t \tilde{V}'(H_s)\sigma_{u_s}dB_s + \frac{1}{2} \int_0^t 1\{H_s \notin E\}\tilde{V}''(H_s)\sigma_{u_s}^2 ds.$$

Since both $\tilde{V}(H_t)$ and $e^{-\int_0^t Q(H_s)ds}$ are semimartingales (the latter follows from the fact that $H_t$ is a semimartingale), we can then apply the multidimensional Itô formula on semimartingales to the function $g(\tilde{V}(H_t), e^{-\int_0^t Q(H_s)ds}) = \tilde{V}(H_t)e^{-\int_0^t Q(H_s)ds}$ and rewrite $X_t$ as

$$X_t = \tilde{V}(x) + \int_0^t e^{-\int_0^s Q(H_z)dz} \tilde{V}'(H_s)(\mu_{u_s} - H_s)ds + \int_0^t e^{-\int_0^s Q(H_z)dz} \tilde{V}'(H_s)\sigma_{u_s}dB_s + \int_0^t e^{-\int_0^s Q(H_z)dz} \mu_{u_s} ds + \int_0^t e^{-\int_0^s Q(H_z)dz} \sigma_{u_s} d(B_s).$$

(51)

Since $\tilde{V}'$ is bounded (Condition 3 in the proposition) and $\sigma_{u_s} \in \{0, \sigma_R\}$, the two stochastic integral terms in the above equation have bounded integrands hence have zero expectations. Now take expectations on both sides of Eq. (51) and write 1 as $1\{H_s \notin E\} + 1\{H_s \in E \& u_s = S\} + 1\{H_s \in E \& u_s = R\}$, we get:

$$\mathbb{E}X_t$$

$$= \tilde{V}(x) + \mathbb{E} \int_0^t e^{-\int_0^s Q(H_z)dz} \left( -Q(H_s)\tilde{V}(H_s) + \tilde{V}'(H_s)(\mu_{u_s} - H_s) + \frac{1}{2} \sigma_{u_s}^2 \tilde{V}''(H_s) + \mu_{u_s} \right) 1\{H_s \notin E\} ds$$

$$+ \left( -Q(H_s)\tilde{V}(H_s) + \tilde{V}'(H_s)(\mu_R - H_s) + \mu_R \right) 1\{H_s \in E \& u_s = S\} ds$$

$$+ \left( -Q(H_s)\tilde{V}(H_s) + \tilde{V}'(H_s)(\mu_R - H_s) + \mu_R \right) 1\{H_s \in E \& u_s = R\} ds$$

$$\leq \tilde{V}(x) + \mathbb{E} \int_0^t e^{-\int_0^s Q(H_z)dz} \left( -Q(H_s)\tilde{V}(H_s) + \tilde{V}'(H_s)(\mu_R - H_s) + \mu_R \right) 1\{H_s \in E \& u_s = R\} ds$$

$$= \tilde{V}(x).$$

(52)
The inequality results from a direct application of Condition 4. The last step follows from the fact that \(-Q(H_s)\bar{V}(H_s) + \bar{V}'(H_s)(\mu_R - H_s) + \mu_R\) is bounded for \(H_s \in \mathcal{E}\), and \(\int_0^1 1\{H_s \in \mathcal{E} \& u_s = R\}ds = 0\) almost surely by Lemma A.1. Since Eq. (52) holds for any \(t \geq 0\), in particular it also holds in the limit:

\[
\limsup_{t \to \infty} \mathbb{E}X_t \leq \bar{V}(x).
\]  

(53)

We will now show \(\limsup_{t \to \infty} \mathbb{E}X_t \geq V(x, \hat{\pi})\). Observe that since \(\mu_R > 0\) and \(\mu_S > 0\), the integral \(\int_0^t e^{-\int_0^s Q(H_u)du} \mu_{u_s} ds\) is pathwise monotone increasing in \(t\) and hence converges pathwise to \(\int_0^\infty e^{-\int_0^s Q(H_u)du} \mu_{u_s} ds\) as \(t \to \infty\). It follows from Eq. (9) that

\[
V(x, \hat{\pi}) = \mathbb{E}\left[\int_0^\infty e^{-\int_0^s Q(H_u)du} \mu_{u_s} ds\right] = \lim_{t \to \infty} \mathbb{E}\left[\int_0^t e^{-\int_0^s Q(H_u)du} dY_s\right].
\]

(54)

Take expectations on both sides of Eq. (50) and let \(t \to \infty\), we get

\[
\limsup_{t \to \infty} \mathbb{E}X_t = \limsup_{t \to \infty} \mathbb{E}\left[\bar{V}(H_t)e^{-\int_0^t Q(H_s)ds}\right] + V(x, \hat{\pi}) \geq V(x, \bar{\pi}),
\]

(55)

where we used Eq. (54) and that \(\bar{V}\) in non-negative (Condition 1 in the proposition). Combining Eqs. (53) and (55), we obtain the desired result \(\bar{V}(x) \geq V(x, \bar{\pi})\) for \(\forall x \in \mathbb{R}\) and any admissible policy \(\bar{\pi} \in \Pi\). That is, \(\bar{V}\) is an upperbound of the optimal value function \(V^*\).

Now it remains to show this upper bound is tight, i.e., the inequalities in Eqs. (53) and (55) are binding under policy \(\bar{\pi}\) (\(\bar{\pi}\) defined in Condition 5). By Condition 5 of the proposition, the inequality in Eq. (52) is binding under policy \(\bar{\pi}\) (where \(u_s = \pi(H_t)\)). Hence its limit, Eq. (53) is also binding under policy \(\bar{\pi}\). On the other hand, by Condition 3 and 6 of the proposition, (Condition 3 implies that \(\bar{V}(x) \leq K(1 + |x|)\) for some \(K \leq \infty\) we have

\[
\lim_{t \to \infty} \mathbb{E}\left[\bar{V}(H_t)e^{-\int_0^t Q(H_s)ds}\right] + V(x, \bar{\pi}) \leq \lim_{t \to \infty} \mathbb{E}\left[K(1 + |H_t|)e^{-\int_0^t Q(H_s)ds}\right] + V(x, \bar{\pi})
\]

\[
= V(x, \bar{\pi}).
\]

Combining the above inequality with Eq. (55), we have that the inequality in Eq. (55) is binding. Therefore we have proved that \(\bar{V}\) is the value function \(V^*\) and that \(\bar{\pi}\) is an optimal policy.

\[
\square
\]

Now we prove Lemma 4, which establishes that Condition 6 is satisfied by a subset of admissible policies including the optimal policy.

**Proof of Lemma 4.** We want to show \(\lim_{t \to \infty} \mathbb{E}\left[ (1 + |H_t|)e^{-\int_0^t Q(H_s)ds} \right] = 0\) for the described policies. This is equivalent to showing \(\lim_{t \to \infty} \mathbb{E}\left[ e^{-\int_0^t Q(H_s)ds} \right] = 0\) and \(\lim_{t \to \infty} \mathbb{E}\left[ |H_t|e^{-\int_0^t Q(H_s)ds} \right] = 0\). By the Cauchy-Schwarz inequality, we have \(\lim_{t \to \infty} \mathbb{E}\left[ |H_t|e^{-\int_0^t Q(H_s)ds} \right] \leq \lim_{t \to \infty} \sqrt{\mathbb{E}[H_t^2] \mathbb{E}\left[ e^{-2\int_0^t Q(H_s)ds} \right]}\).

Hence it suffices to show

\[
\lim_{t \to \infty} \mathbb{E}\left[ e^{-\int_0^t Q(H_s)ds} \right] = 0,
\]

(56)
\[
\limsup_{t \to \infty} \mathbb{E}[H_t^2] < \infty, \quad (57)
\]

and
\[
\limsup_{t \to \infty} \mathbb{E}\left[ e^{-2 \int_0^t Q(H_s)ds} \right] = 0. \quad (58)
\]

First consider Eq. (56) and Eq. (58). Since (Eq. (5)) \( P(T > t | \mathcal{F}_t) = e^{- \int_0^t Q(H_s)ds} \), we have LHS of Eq. (56) equivalent to \( \lim_{t \to \infty} P(T > t) = P(T > \infty) \), and LHS of Eq. 58 is
\[
\limsup_{t \to \infty} \mathbb{E}\left[ e^{-2 \int_0^t Q(H_s)ds} \right] = \limsup_{t \to \infty} \mathbb{E}\left[ (P(T > t))^2 \right] = \limsup_{t \to \infty} P(T > t)^2 = P(T > \infty)^2,
\]
both of which are zero since by Proposition 1 the customer lifetime under interval policies with finite intervals is finite in expectation.

It only remains to show Eq. (57).

Fix any starting happiness value \( x \in \mathbb{R} \) and any interval policy with finite intervals. Denote by \( u_t \) the corresponding action process. By definition of \( H_t \) in Eq. (3), we have
\[
H_t = x e^{-t} + \int_0^t e^{-(t-s)} \mu_{u_s} ds + \int_0^t e^{-(t-s)} \sigma_{u_s} dB_s,
\]
which gives
\[
H_t^2 = x^2 e^{-2t} + \int_0^t e^{-(t-s)} \mu_{u_s} ds + \int_0^t e^{-(t-v)} \mu_{u_v} dvds + \int_0^t e^{-(t-s)} \sigma_{u_s}^2 ds + 2xe^{-t} \int_0^t e^{-(t-s)} \mu_{u_s} ds + 2xe^{-t} \int_0^t e^{-(t-s)} \sigma_{u_s} dB_s + 2 \int_0^t e^{-(t-s)} \mu_{u_s} \int_0^t e^{-(t-s)} \sigma_{u_s} dB_s ds.
\]
Take expectation to \( H_t^2 \), we can remove both stochastic integral terms on the second line above since the their integrands are both bounded. Note that \( \mu_{u_s} \leq \mu \triangleq \max\{\mu_S, \mu_R\} \) and \( \sigma_{u_s} \leq \sigma_R \) for any \( t \geq 0 \). Therefore we have
\[
\mathbb{E}[H_t^2] \leq x^2 e^{-2t} + \int_0^t e^{-(t-s)} \mu \int_0^t e^{-(t-v)} \mu dvds + \int_0^t e^{-(t-s)} \sigma_R^2 ds + 2xe^{-t} \int_0^t e^{-(t-s)} \mu ds = x^2 e^{-2t} + \mu^2(e^{-2t} + 1 - 2e^{-t}) + \frac{\sigma_R^2 (1 - e^{-2t})}{2} + 2xe^{-t}(e^{-t} - e^{-2t}).
\]
Let \( t \to \infty \) in the above equation, we get \( \limsup_{t \to \infty} \mathbb{E}[H_t^2] \leq \mu^2 + \frac{\sigma_R^2}{2} < \infty \), and hence Eq. (57) is true. Therefore we have proved Lemma 4.

Now we prove Proposition 3, which establishes the optimal value function.

**Proof of Proposition 3.** The policy associated with the \( W \) function belongs to the policy types in Lemma 4. Therefore by Lemma 4, Condition 6 of Proposition 3 is satisfied. We only need to show the other conditions are true. We consider the two cases \( \mu_R \geq \mu_S \) and \( \mu_R < \mu_S \) separately.
First consider the case $\mu_R \geq \mu_S$. By the definition of $\theta_G$ in Lemma 3, we know that when $\mu_R \geq \mu_S$, $\theta_G = -\infty$. Therefore the function $W$ as defined in Proposition 3 reduces to:

$$W(x, C_2, C_3, C_4, C_5) = \begin{cases} V_2(x, C_2, C_3) & \text{if } x < q; \\ V_3(x, C_4) & \text{if } q \leq x \leq \theta_b; \\ V_4(x, C_5) & \text{if } x > \max\{q, \theta_b\}, \end{cases}$$

(59)

where $\theta_b$ is as defined in Lemma 2 and

$$V_2(x, C_2, C_3) = C_2 e^{\frac{(x-\mu_R)^2}{\sigma_R^2}} + C_3 \text{erf} \left( \frac{x-\mu_R}{\sigma_R} \right) e^{\frac{(x-\mu_R)^2}{\sigma_R^2}} + \mu_R;$$

(60)

$$V_3(x, C_4) = C_4 + \mu_S \log(x - \mu_S);$$

(61)

$$V_4(x, C_5) = C_5 + \int_0^x \frac{\mu_R \sqrt{\pi}}{\sigma_R} e^{\frac{(x-y)^2}{\sigma_R^2}} (1 - \text{erf} \left( \frac{y-\mu_R}{\sigma_R} \right)) dy. \tag{62}$$

Our objective here is to find $C^*_2, C^*_3, C^*_4, C^*_5$ such that $W^*(\cdot) \triangleq W(\cdot, C^*_2, C^*_3, C^*_4, C^*_5)$ together with the associated policy satisfy Conditions 1-6. First we restate Eq. (13) here:

$$\max_{i=S,R} \left\{ -Q(x)V(x) + (\mu_i - x)V'(x) + \frac{1}{2} \sigma_i^2 \mathbb{1}\{x \notin \mathcal{E}\} V''(x) + \mu_i \right\} = 0$$

for all $x$, where $\mathcal{E}$ is the set of values at which $V''$ doesn’t exist.

For easier writing, we introduce the following notations:

$$B_i(x, V) \triangleq -Q(x)V(x) + (\mu_i - x)V'(x) + \frac{1}{2} \sigma_i^2 \mathbb{1}\{x \notin \mathcal{E}\} V''(x) + \mu_i, \quad \text{for } i = R, S; \tag{63}$$

$$\delta(x, V) \triangleq B_R(x, V) - B_S(x, V) = (\mu_R - \mu_S)V'(x) + \frac{1}{2} \sigma_R^2 \mathbb{1}\{x \notin \mathcal{E}\} V''(x) + \mu_R - \mu_S. \tag{64}$$

One can check that for any $C_2, C_4, C_5$ in $\mathbb{R}$, $V_2(x, C_2, C_2)$ satisfies $B_R(x, V_2(x, C_2, C_2)) = 0$ for $x \in (-\infty, q)$, $V_3(x, C_4)$ satisfies $B_S(x, V_3(x, C_4)) = 0$ for $x \in [q, \max\{q, \theta_b\}]$, and $V_4(x, C_5)$ satisfies $B_R(x, V_4(x, C_5)) = 0$ for $x \in (\max\{q, \theta_b\}, \infty)$. This constitutes “half” of Eq. (13). Also $W(\cdot, C_2, C_2, C_4, C_5)$ is almost everywhere twice continuously differentiable except possibly at $q$ and $\theta_b$. Hence to complete the proof of Proposition 3, we need to show that there exist $C^*_2, C^*_4$ and $C^*_5$ in $\mathbb{R}$ such that the following six statements are true:

**Statement 1.** $W^*(\cdot)$ is non-negative;

**Statement 2.** $W^*(\cdot)$ is continuously differentiable at $q$ and $\theta_b$;

**Statement 3.** $W^*(\cdot)$ has bounded first derivative;

**Statement 4.** $\delta(x, V_2(x, C^*_2, C^*_2)) \geq 0$ for any $x \in (-\infty, q)$,

**Statement 5.** $\delta(x, V_3(x, C^*_4)) \leq 0$ for any $x \in (q, \max\{q, \theta_b\})$,

**Statement 6.** $B_S(x, V_4(x, C^*_5)) \leq 0$ for any $x \in [\max\{q, \theta_b\}, \infty)$.
Observe that Statements 4–6 will complete the HJB equation (13).

Before we give the values of \( C_2^*, C_4^* \) and \( C_5^* \), first observe that the first derivatives of \( V_3(\cdot, C_4) \) and \( V_4(\cdot, C_5) \) does not depend on \( C_4 \) and \( C_5 \). We therefore denote them \( V_3'(\cdot) \) and \( V_4'(\cdot) \), respectively. Now we give the values of \( C_2^*, C_4^* \) and \( C_5^* \) here, and will then verify that Statements 1-6 are true. Define

\[
C_2^* = \begin{cases} 
\frac{2}{\sigma_R} \left( \frac{1}{\sqrt{\pi}} + e^{\frac{(q-\mu_R)^2}{2\sigma_R^2}} \right) & \text{if } \theta_b < q \\
\frac{2}{\sigma_R} \left( \frac{1}{\sqrt{\pi}} + e^{\frac{(q-\mu_R)^2}{2\sigma_R^2}} \right) & \text{if } \theta_b \geq q 
\end{cases}
\]

(65)

\[
C_4^* = C_2^* e^{\frac{(q-\mu_R)^2}{2\sigma_R^2}} \left( 1 + \text{erf} \left( \frac{q - \mu_R}{\sigma_R} \right) \right) + \mu_R - \mu_S \log(q - \mu_S),
\]

(66)

and

\[
C_5^* = \begin{cases} 
V_2(q, C_2^*, C_4^*) - \frac{\mu_R}{\sigma_R} e^{\frac{(x-\mu_R)^2}{2\sigma_R^2}} \int_0^\infty \frac{\mu_{RV}}{\sigma_R} e^{\frac{(z-\mu_R)^2}{2\sigma_R^2}} \left( 1 - \text{erf} \left( \frac{z - \mu_R}{\sigma_R} \right) \right) dz & \text{if } \theta_b < q \\
V_3(\theta_b, C_4^*) - \frac{\mu_b}{\sigma_R} e^{\frac{(x-\mu_R)^2}{2\sigma_R^2}} \int_0^{\theta_b} \frac{\mu_{RV}}{\sigma_R} e^{\frac{(z-\mu_R)^2}{2\sigma_R^2}} \left( 1 - \text{erf} \left( \frac{z - \mu_R}{\sigma_R} \right) \right) dz & \text{if } \theta_b \geq q 
\end{cases}
\]

(67)

We define \( V_2^*(\cdot) \equiv V_2(\cdot, C_2^*, C_4^*) \), \( V_3^*(\cdot) \equiv V_3(\cdot, C_4^*) \), \( V_4^*(\cdot) \equiv V_4(\cdot, C_5^*) \) and \( W^*(\cdot) \equiv W(\cdot, C_2^*, C_4^*, C_5^*) \). Next we will show Statements 1-6 are true.

First consider Statement 2. Since \( V_2^*(\cdot), V_3^*(\cdot) \) and \( V_4^*(\cdot) \) are all continuously differentiable everywhere, to show \( W^*(\cdot) \) is continuously differentiable at \( q \) and \( \theta_b \), it suffices to show

\[
V_2^*(q) = V_4^*(q), \quad V_2^*(q) = V_4^*(q)
\]

(68)

if \( \theta_b < q \). On the other hand, if \( \theta_b \geq q \), it suffices to show

\[
V_2^*(q) = V_3^*(q), \quad V_2^*(q) = V_3^*(q), \quad V_3^*(\theta_b) = V_4^*(\theta_b), \quad V_3^*(\theta_b) = V_4^*(\theta_b).
\]

(69)

It is easy to verify that Eqs. (68) and (69) are true by the construction of \( C_2^*, C_4^* \) and \( C_5^* \) in Eqs. (65) – (67).

Now consider Statement 3. Since we have just proved that \( W^*(\cdot) \) is continuous in \( \mathbb{R} \), to show that \( W^*(\cdot) \) is bounded, it suffices to show

\[
\lim_{x \to -\infty} W^*(x) < \infty \quad \text{and} \quad \lim_{x \to \infty} W^*(x) < \infty.
\]

This is equivalent to showing \( \lim_{x \to -\infty} V_2^*(x) < \infty \) and \( \lim_{x \to \infty} V_4^*(x) < \infty \). By definition of \( V_2^* \) (see Eqs. (60) and (65)) and \( V_4^* \) (see Eqs. (62) and (67)), we have

\[
\lim_{x \to -\infty} V_2^*(x) = \lim_{x \to -\infty} \frac{2C_2^*}{\sigma_R} \left( \frac{1}{\sqrt{\pi}} + e^{\frac{(x-\mu_R)^2}{2\sigma_R^2}} \left( 1 + \text{erf} \left( \frac{x - \mu_R}{\sigma_R} \right) \right) \right)
\]

\[
= \lim_{z \to \infty} \frac{2C_2^*}{\sigma_R} \left( \frac{1}{\sqrt{\pi}} - ze^{z^2 \text{erfc}(z)} \right)
\]

24
and

\[
\lim_{x \to \infty} V_1'(x) = \lim_{x \to \infty} \frac{\mu_R \sqrt{\pi}}{\sigma_R} e^{-\frac{(x-\mu_R)^2}{\sigma^2_R}} \left( 1 - \text{erf} \left( \frac{x-\mu_R}{\sigma_R} \right) \right) \\
= \lim_{z \to \infty} \frac{\mu_R \sqrt{\pi}}{\sigma_R} e^{z^2} \text{erfc}(z).
\]

From Lemma B.5, the limits above are 0.

Now consider Statement 4. For \( x \leq q \), substitute \( V_2(x, C_2^*, C_2^*) \) (see Eqs. (60) and (65)) into Eq. (64), we get

\[
\delta(x, V_2^*) = (\mu_R - \mu_S) V_2'(x) + \frac{1}{2} \sigma_R^2 V_2''(x) + \mu_R - \mu_S.
\]

Consider the expressions for \( V_2'(x) \) and \( V_2''(x) \):

\[
V_2'(x) = \frac{2C_2^*}{\sigma_R} \left( \frac{1}{\sqrt{\pi}} + e^{-\frac{(x-\mu_R)^2}{\sigma^2_R}} \frac{x-\mu_R}{\sigma_R} \text{erfc} \left( \frac{x-\mu_R}{\sigma_R} \right) \right);
\]

\[
V_2''(x) = \frac{2C_2^*}{\sigma^2_R} \left( 2 \frac{(x-\mu_R)}{\sigma_R} + e^{-\frac{(x-\mu_R)^2}{\sigma^2_R}} \left( 1 + \frac{2(x-\mu_R)^2}{\sigma^2_R} \right) \text{erfc} \left( \frac{x-\mu_R}{\sigma_R} \right) \right).
\]

By Lemma B.4, Corollary B.8, the terms in both parenthesis above are non-negative. Therefore, if we can show \( C_2^* > 0 \), then both \( V_2'(x) \) and \( V_2''(x) \) are non-negative, which implies (since \( \mu_R \geq \mu_S \)) \( \delta(x, V_2^*) \geq 0 \) for \( x \leq q \). The next claim establishes this result.

**Claim 1.** The value \( C_2^* \) as defined in Eq. (65) is strictly positive.

**Proof.** By Eq. (65), we have

\[
C_2^* = \begin{cases} 
\frac{V_2'(q)}{\frac{1}{\sqrt{\pi}} + e^{-\frac{(q-\mu_R)^2}{\sigma^2_R}} \frac{q-\mu_R}{\sigma_R} \left( 1 + \text{erf} \left( \frac{q-\mu_R}{\sigma_R} \right) \right)} & \text{if } \theta_b < q; \\
\frac{V_2'(q)}{\frac{1}{\sqrt{\pi}} + e^{-\frac{(q-\mu_R)^2}{\sigma^2_R}} \frac{q-\mu_R}{\sigma_R} \left( 1 + \text{erf} \left( \frac{q-\mu_R}{\sigma_R} \right) \right)} & \text{if } \theta_b \geq q,
\end{cases}
\]
Since $q > \mu_s$, the numerators in both cases are positive. In fact the denominator, which is the same for both cases, is also positive by an application of Corollary B.8. Hence $C_2^2 > 0$.

Therefore we have proved Statement 4 is true.

Now we only need to show Statements 1, 5 and 6. We first consider Statement 6. Note that if $\mu_R = \mu_S$, then $\theta_b = \infty$ by Lemma 2 and hence Statement 6 is trivially satisfied. Suppose $\mu_R > \mu_S$. By Eq. (63), for any $x > \max\{q, \theta_b\}$, we have

$$B_S(x, V_4^\ast) = (\mu_S - x)V_4^\ast(x) + \mu_S.$$  

Also by Eq. (62), we have

$$V_4^\ast(x) = \frac{\mu_R \sqrt{\pi}}{\sigma_R} e^{-\frac{(x-\mu_R)^2}{\sigma_R^2}} \text{erfc}\left(\frac{x - \mu_R}{\sigma_R}\right).$$  

Combining the above two equations, we get

$$B_S(x, V_4^\ast) = \frac{\mu_R \sqrt{\pi}}{\sigma_R} (\mu_S - x) e^{-\frac{(x-\mu_R)^2}{\sigma_R^2}} \text{erfc}\left(\frac{x - \mu_R}{\sigma_R}\right) + \mu_S.$$  

Observe that the RHS of the above equation is the same as the LHS of the equation in Lemma 2. Therefore $B_S(\theta_b, V_4^\ast) = 0$. Using Corollary B.7, we get $B_S(x, V_4^\ast) < 0$ on $\theta_b, \infty$, i.e., Statement 6 is true.

Now it remains to show Statement 1 and 5. Consider Statement 5. By Eqs. (64) and (61), for any $x \in (q, \max\{q, \theta_b\})$ we have

$$\delta(x, V_3^\ast) = (\mu_R - \mu_S) \cdot \frac{\mu_S}{x - \mu_S} + \frac{\sigma_R^2}{2} \cdot \frac{-\mu_S}{(x - \mu_S)^2} + \mu_R - \mu_S$$

$$= 2(\mu_R - \mu_S)x^2 - 2\mu_S(\mu_R - \mu_S)x - \mu_S\sigma_R^2 = 2(\mu_R - x)^2.$$  

If $\mu_R = \mu_S$, then $\delta(x, V_3^\ast) \leq 0$ easily follows. Now consider $\mu_R > \mu_S$. Solving $\delta(x, V_3^\ast) \leq 0$, we get

$$\frac{\mu_S}{2} - \sqrt{\frac{\mu_S^2}{4} + \frac{\mu_S\sigma_R^2}{2(\mu_R - \mu_S)}} \leq x \leq \frac{\mu_S}{2} + \sqrt{\frac{\mu_S^2}{4} + \frac{\mu_S\sigma_R^2}{2(\mu_R - \mu_S)}}.$$  

(70)
Since $\mu_S \leq q$, we have \( \frac{\mu_S}{2} - \sqrt{\frac{\mu_S^2}{4} + \frac{\mu_S \sigma_R^2}{2(\mu_R - \mu_S)}} \leq \mu_S < q \). Hence if we can show
\[
\theta_b \leq \frac{\mu_S}{2} + \sqrt{\frac{\mu_S^2}{4} + \frac{\mu_S \sigma_R^2}{2(\mu_R - \mu_S)}},
\]
then this means $\delta(x, V^*_3) \leq 0$ for any $x \in (q, \max\{q, \theta_b\})$, i.e., Statement 5 is true. Let $\alpha = \frac{\mu_R - \mu_S}{\sigma_R}, \beta = \frac{\mu_S}{\mu_R}$, and $\hat{\theta}_b = \frac{\theta_b - \mu_R}{\sigma_R}$. We can rewrite Eq. (71) as
\[
\hat{\theta}_b \leq \frac{\alpha (\beta - 2) + \sqrt{\alpha^2 \beta^2 + 2\beta (1 - \beta)}}{2(1 - \beta)}.
\]
Moreover, by Lemma 2, we have $\hat{\theta}_b > -\alpha$ and $\hat{\theta}_b$ satisfies $(\hat{\theta}_b + \alpha) e^{\hat{\theta}_b} \text{erfc}(\hat{\theta}_b) - \frac{\beta}{\sqrt{\pi}} = 0$. Using Lemma B.10 we deduce Eq. (72). Therefore Statement 5 is proved. It only remains to show, under the $\mu_R \geq \mu_S$ scheme, Statement 1 is true.

By construction (see Eq. (60)), $V_2(x, C_2^*, C_2^*) > 0$ if $C_2^* > 0$. By Claim 1, we have $C_2^* > 0$, which implies that $W^*(x) > 0$ for all $x \leq q$. In particular, $W^*(q) > 0$. Also by definition in Eq. (61), $V_3(x, C_1^*)$ increases in $x$ on $[q, \infty)$. Therefore by continuity of $W^*$ at $q$, we have $W^*(x) > 0$ for all $x \leq \max\{q, \theta_b\}$. Similarly $V_4(x, C_5^*)$ increases in $x$ as well (see Eq. (62)), hence by continuity of $W^*$ at $\theta_b$, we have $W^*(x) > 0$ for all $x \in \mathbb{R}$. Thus we have proved Statement 1, and hence the proposition for under the $\mu_R \geq \mu_S$ scheme.

Now it remains to prove the proposition for the other scheme, $\mu_R < \mu_S$.

By the definition of $\theta_b$ in Lemma 2, we know that when $\mu_R < \mu_S$, $\theta_b = \infty$. Therefore in this case, the function $W$ as defined in Proposition 3 reduces to:
\[
W(x, C_1, C_2, C_3, C_5) = \begin{cases} 
V_1(x, C_1) & \text{if } x \leq \theta_G \\
V_2(x, C_2, C_3) & \text{if } \theta_G < x < q; \\
V_4(x, C_5) & \text{if } x \geq q,
\end{cases}
\]
where $V_2(x, C_2, C_3), V_4(x, C_5)$ are as defined in Eqs. (61) and (62), and $V_1(x, C_1)$ defined as:
\[
V_1(x, C_1) = \frac{C_1}{x - \mu_S} + \mu_S.
\]
We need to show that there exists $C_1^G, C_2^G, C_3^G, C_5^G$ such that $W^G(\cdot) \triangleq W(\cdot, C_1^G, C_2^G, C_3^G, C_5^G)$ satisfy Conditions 1-6 of Proposition 2. Recall the definition of $B_i(x, V)$ in Eq. (63):
\[
B_i(x, V) \triangleq -Q(h)V(x) + (\mu_i - x)V'(x) + \frac{1}{2} \sigma_t^2 \mathbf{1}\{x \notin \mathcal{E}\} V''(x) + \mu_i, \quad \text{for } i = R, S,
\]
where $\mathcal{E}$ is the set of values at which $V$ is not twice continuously differentiable. $B_i(x, \hat{V})$ is the reduced-form of HJB equation (13) under service mode $i = R, S$. One can check that for any $C_1, C_2, C_3, C_5$, functions $V_1(x, C_1)$ and $V_4(x, C_5)$ satisfy ODE $B_S(x, V) = 0$ and function $V_2(x, C_2, C_3)$ satisfy ODE $B_R(x, \hat{B}) = 0$, on their domains. Also $V_1(\cdot, C_1), V_2(\cdot, C_2, C_3), V_4(\cdot, C_5)$ are continuously differentiable on their domains. Hence to complete the proof, we only need to show there exist $C_1^G, C_2^G, C_3^G, C_5^G$ such that the following statements are true:

**Statement 7.** $W^G(\cdot)$ is non-negative;
**Statement 8.** $W^G(\cdot)$ is continuously differentiable at $q$ and $\theta_G$;

**Statement 9.** $W^G(\cdot)$ has bounded first derivative;

**Statement 10.** $B_R(x, V_1(x, C_i^G)) \leq 0$ for any $x < \theta_G$;

**Statement 11.** $B_S(x, V_2(x, C_i^G, C_i^G)) \leq 0$ for any $x \in (\theta_G, q)$;

**Statement 12.** $B_R(x, V_4(x, C_5)) \leq 0$ for any $x > q$.

First we specify $C_1^G, C_2^G, C_3^G$ and $C_5^G$. To shorten the expressions a bit, we introduce the following notations:

$$a \triangleq \frac{\mu_S - \mu_R}{\sigma_R}, \quad b \triangleq \mu_S - \mu_R, \quad \hat{\theta}_G \triangleq \frac{\theta_G - \mu_R}{\sigma_R}.$$  

We can then define

$$C_1^G = \begin{cases} 
0 & \text{if } \theta_G \geq \mu_S; \\
-\frac{b(a-\hat{\theta}_G)}{\sigma_R} & \text{if } \theta_G < \mu_S, 
\end{cases} \quad (74)$$

$$C_2^G = \begin{cases} 
\theta_G \sqrt{2\pi} e^{-\hat{\theta}_G^2} \text{erf}(\hat{\theta}_G) & \text{if } \theta_G \geq \mu_S; \\
\theta_G \sqrt{2\pi} e^{-\hat{\theta}_G^2} \text{erf}(\hat{\theta}_G) & \text{if } \theta_G < \mu_S, 
\end{cases} \quad (75)$$

$$C_3^G = \begin{cases} 
-\frac{b\sqrt{\pi}}{2\sigma_R} & \text{if } \theta_G \geq \mu_S; \\
\frac{b\sqrt{\pi}}{2\sigma_R} & \text{if } \theta_G < \mu_S, 
\end{cases} \quad (76)$$

and

$$C_5^G = V_2(q, C_2^G, C_3^G) - \mu_S \log(q - \mu_S). \quad (77)$$

We also define $V_1^G(\cdot) \triangleq V_1(\cdot, C_i^G)$, $V_2^G(\cdot) \triangleq V_2(\cdot, C_i^G, C_i^G)$ and $V_4^G(\cdot) \triangleq V_4(\cdot, C_5^G)$.

Next we show Statement 7-12 are true.

Consider first Statement 8. By the construction of $C_5^G$ in Eq. (77), we know that $V_2^G(q) = V_4^G(q)$. Therefore, to complete the proof of Statement 8 we only need to show

$$V_1^G(\theta_G) = V_2^G(\theta_G), \quad V_1^G(\theta_G) = V_2^G(\theta_G), \quad V_2^G(q) = V_4^G(q).$$

The derivatives of $V_1^G, V_2^G$ and $V_4^G$ are

$$V_1^G'(x) = -\frac{C_1^G}{(x - \mu_S)^2},$$

$$V_2^G'(x) = 2\frac{\sigma_R}{\sqrt{\pi}} e^{-\frac{(x-\mu_R)^2}{\sigma_R^2}} \left( \text{erf}\left(\frac{x-\mu_R}{\sigma_R}\right) \right).$$
and

\[ V_4^{G'}(x) = \frac{\mu_S}{x - \mu_S} \]

By the construction of \(C_1^G, C_2^G\) and \(C_3^G\), one can check that \(V_1^G(\theta_G) = V_2^G(\theta_G)\) and \(V_1^{G'}(\theta_G) = V_2^{G'}(\theta_G)\) hold. Now it only remains to prove \(V_2^{G'}(q) = V_4^{G'}(q)\). This is equivalent to showing

\[ \frac{2 \left( \frac{C_2^G}{\sqrt{\pi}} + \hat{q} e^{\hat{\theta}^2} \left( C_2^G + C_3^G \text{erf}(\hat{q}) \right) \right)}{\sigma_R} = \frac{\mu_S}{q - \mu_S}, \]

where \(\hat{q} \triangleq \frac{q - \mu_R}{\sigma_R}\). After substituting \(C_2^G\) and \(C_3^G\) in the expression, we get an equivalent equality that remains to be verified:

\[ \frac{\mu_S}{2a(q - \mu_S)} = -e^{\hat{q}^2} \sqrt{\pi} \hat{q} \hat{\theta}_G \left( \text{erf}(\hat{q}) - \text{erf}(\hat{\theta}_G) \right) + \hat{q} e^{\hat{\theta}^2 - \hat{\theta}_G^2} - \hat{\theta}_G, \quad \text{if } \theta_G \geq \mu_S, \]

or

\[ \frac{\mu_S \left( 1 + a \hat{\theta}_G - a^2 \right)}{a(q - \mu_S)} = 2\hat{q} e^{\hat{\theta}^2 - \hat{\theta}_G^2} \left( 1 + \hat{\theta}_G^2 - a \hat{\theta}_G \right) + \left( a - 3\hat{\theta}_G + 2a \hat{\theta}_G^2 - 2\hat{\theta}_G^3 \right) \]

\[ + \hat{q} e^{\hat{\theta}^2} \sqrt{\pi} \left( a - 3\hat{\theta}_G + 2a \hat{\theta}_G^2 - 2\hat{\theta}_G^3 \right) \left( \text{erf}(\hat{q}) - \text{erf}(\hat{\theta}_G) \right), \quad \text{if } \theta_G < \mu_S. \]

By the definition of \(\theta_G\) in Lemma 3, the above equalities are true. Therefore we have proved Statement 8.

Next consider Statement 9, i.e., want to show \(W^{G'}\) is bounded everywhere. By Statement 8, \(W^{G'}\) exists and is continuous everywhere. Therefore to show \(W^{G'}\) is bounded everywhere, it suffices to show \(\lim_{x \to \infty} W^{G'}(x)\) and \(\lim_{x \to -\infty} W^{G'}(x)\) are bounded. In fact,

\[ \lim_{x \to \infty} W^{G'}(x) = \lim_{x \to \infty} \frac{\mu_S}{x - \mu_S} = 0 \]

and

\[ \lim_{x \to -\infty} W^{G'}(x) = \lim_{x \to -\infty} -\frac{C_1^G}{(x - \mu_S)^2} = 0, \]

since \(C_1^G\) is a constant. Hence, we have proved Statement 9.

Now consider Statement 10. We want to show \(B_R(x, V_1^G) \leq 0\) for \(x < \theta_G\). In fact, for any \(x < \theta_G < q\), we have \(Q(x) = 1\) and hence

\[ B_R(x, V_1^G) = -V_1^G(x) + (\mu_R - x)V_1^{G'}(x) + \frac{1}{2}\sigma_R^2 V_1^{G''}(x) + \mu_R \]

\[ = \begin{cases} -b & \text{if } \theta_G \geq \mu_S; \\ \sigma_R^2 bf \left( \frac{x - \mu_R}{\sigma_R} \right) & \text{if } \theta_G < \mu_S, \end{cases} \]

where

\[ f(z) = (a - \hat{\theta}_G)^3(1 - a^2 + az) - (a - z)^3(1 - a^2 + a\hat{\theta}_G). \]
Since we know that $1 - a^2 + a\tilde{\theta}_G > 0$ and $x < \theta_G < \mu_S$, to show $B_R(x, V^G_1) \leq 0$ we only need to show $f(x - \mu_R) \leq 0$ for $x < \theta_G$, or equivalently, to show $f(z) \leq 0$ for $z < \hat{\theta}_G$, and where $\hat{\theta}_G \in (a - \frac{1}{a}, a)$.

Observe that

$$f'(z) = a(a - \hat{\theta}_G)^3 + 3(1 - a^2 + a\tilde{\theta}_G)(a - z)^2.$$  

Since $\hat{\theta}_G < a$, we have $(a - \hat{\theta}_G)^3 > 0$. Also $a > 0$ and $\tilde{\theta}_G > a - \frac{1}{a}$ imply $1 - a^2 + a\tilde{\theta}_G > 0$. Therefore $f'(z) > 0$ for $a < \hat{\theta}_G$. Now to show $f(z) \leq 0$ for $z < \hat{\theta}_G$, we only need to show $f(\hat{\theta}_G) \leq 0$. One can easily verify that $f(\hat{\theta}_G) = 0$. Hence, we have proved Statement 10.

Next, consider Statement 11. We want to show $B_S(x, V^G_2) \leq 0$ for $x \in (\theta_G, q)$. In fact, for any $x \in (\theta_G, q)$, we have $Q(x) = 1$ and hence

$$B_S(x, V^G_2) = -V^G_2(x) + (\mu_S - x)V^G_2(x) + \mu_S.$$  

If we can show $W^G'(x) \geq 0$ for all $x \in \mathbb{R}$, then the proof of Statement 11 is easy. Also Statement 7 would be true since $\lim_{x \to \infty} W^G(x) = \mu_S > 0$. Now we show $W^G'(x) \geq 0$ for all $x \in \mathbb{R}$ is true.

From Lemma 3, we know that $\hat{\theta}_G > a - \frac{1}{a}$ (recall that $a = \frac{\mu_S - \mu_R}{\sigma_R}$), i.e., we have $1 - a^2 + a\tilde{\theta}_G > 0$. Then one can check Eqs. (73) and (74) to see that $V^G_1 \geq 0$ on $(-\infty, \theta_G)$. Also that $V^G_4 \geq 0$ on $[q, \infty)$. Therefore now it only remains to show $V^G_2 \geq 0$ on $(\theta_G, q)$, i.e., $\frac{C^G}{\sqrt{\pi}} + ze^{z^2}(C^G_2 + C^G_3 \text{erf}(z)) \geq 0$ for $z \in (\hat{\theta}_G, q)$. Recall the definition of $C^G_2$ in Eq. (75) and apply Lemma B.12 and B.13, we can get $C^G_2 > 0$. Therefore to show $V^G_2(x)$ is non-negative for $x \in (\theta_G, q)$, it is equivalent to show $\frac{C}{\sqrt{\pi}} + ze^{z^2}(1 + \text{Cerf}(z)) \geq 0$ for $z \in (\hat{\theta}_G, q)$, where $C \triangleq \frac{C^G_2}{C^G_3}$. Computing

$$C^G_3 + C^G_2 = \begin{cases} 
\frac{1}{2(1 - a^2 + a\tilde{\theta}_G)} \left( b \left( e^{-\theta_G^2} - \sqrt{\pi} \tilde{\theta}_G \text{erf}(\theta_G) \right) 
+ \theta_G e^{\theta_G^2} \left( 2 - 2a\tilde{\theta}_G + 2\theta_G^2 + e^{\theta_G^2} \sqrt{\pi} (a - 3\tilde{\theta}_G + 2a\tilde{\theta}_G^2 - 2\tilde{\theta}_G^3) \right) \right) & \text{if } \theta_G \geq \mu_S; \\
\frac{1}{2(1 - a^2 + a\tilde{\theta}_G)} \left( b \left( e^{-\theta_G^2} - \sqrt{\pi} \tilde{\theta}_G \text{erf}(\theta_G) \right) 
\theta_G e^{\theta_G^2} \left( -2 + 2a\tilde{\theta}_G - 2\theta_G^2 + e^{\theta_G^2} \sqrt{\pi} (a - 3\tilde{\theta}_G + 2a\tilde{\theta}_G^2 - 2\tilde{\theta}_G^3) \right) \right) & \text{if } \theta_G < \mu_S,
\end{cases}$$

and

$$C^G_3 - C^G_2 = \begin{cases} 
\frac{1}{2(1 - a^2 + a\tilde{\theta}_G)} \left( b \left( e^{-\theta_G^2} - \sqrt{\pi} \tilde{\theta}_G \text{erf}(\theta_G) \right) 
+ \theta_G e^{\theta_G^2} \left( -2 + 2a\tilde{\theta}_G - 2\theta_G^2 + e^{\theta_G^2} \sqrt{\pi} (a - 3\tilde{\theta}_G + 2a\tilde{\theta}_G^2 - 2\tilde{\theta}_G^3) \right) \right) & \text{if } \theta_G \geq \mu_S; \\
\frac{1}{2(1 - a^2 + a\tilde{\theta}_G)} \left( b \left( e^{-\theta_G^2} - \sqrt{\pi} \tilde{\theta}_G \text{erf}(\theta_G) \right) 
\theta_G e^{\theta_G^2} \left( 2 - 2a\tilde{\theta}_G + 2\theta_G^2 + e^{\theta_G^2} \sqrt{\pi} (a - 3\tilde{\theta}_G + 2a\tilde{\theta}_G^2 - 2\tilde{\theta}_G^3) \right) \right) & \text{if } \theta_G < \mu_S.
\end{cases}$$

By Corollary B.8, if $\theta_G \geq \mu_S$, then $C^G_3 + C^G_2 > 0$ and $C^G_3 - C^G_2 < 0$, hence $C = \frac{C^G_3}{C^G_2} \in (-1, 1)$. This is also true if $\theta_G < \mu_S$, since $1 - a^2 + a\tilde{\theta}_G > 0$ and Lemma B.13 together imply $C^G_3 + C^G_2 > 0$ and $C^G_3 - C^G_2 < 0$.

Now since $C \in (-1, 1)$ holds, we can apply Lemma B.14 to get $\frac{C}{\sqrt{\pi}} + ze^{z^2}(1 + \text{Cerf}(z))$ is increasing in $z$. Therefore to show $\frac{C}{\sqrt{\pi}} + ze^{z^2}(1 + \text{Cerf}(z)) \geq 0$ for $z \in (\hat{\theta}_G, q)$, we only need to show $\frac{C}{\sqrt{\pi}} + \tilde{\theta}_Ge^{\theta_G^2}(1 + \text{Cerf}(\theta_G)) \geq 0$. Since $C = \frac{C^G_3}{C^G_2}$ is explicitly characterized by $a, b$ and $\tilde{\theta}_G$, we denote

$$F(\tilde{\theta}_G) \triangleq \frac{C}{\sqrt{\pi}} + \tilde{\theta}_Ge^{\theta_G^2}(1 + \text{Cerf}(\theta_G))$$

$$\begin{cases} 
0 & \text{if } \hat{\theta}_G \geq a; \\
\frac{e^{\theta_G^2}(a - \tilde{\theta}_G)}{2 - 2a\tilde{\theta}_G + 2\theta_G^2 + e^{\theta_G^2} \sqrt{\pi} (a - 3\tilde{\theta}_G + 2a\tilde{\theta}_G^2 + 2\tilde{\theta}_G^3) \text{erf}(\theta_G)} & \text{if } \hat{\theta}_G \in (a - \frac{1}{a}, a),
\end{cases}$$

30
To complete the proof of \( W^G(x) \geq 0 \), it remains to show \( F(\hat{\theta}_G) \geq 0 \). If \( \hat{\theta}_G \geq a \) this is trivial. Hence we only need to consider the case where \( \hat{\theta}_G \in (a - \frac{1}{a}, a) \). In fact, by Lemma B.13, the denominator is non-negative. Furthermore, the numerator is non-negative since \( \hat{\theta}_G < a \). Therefore, we have proved \( W^G(x) \geq 0 \), and hence Statement 7 is true.

Since \( \lim_{x \to -\infty} W^G(x) = \mu_S \) and that \( W^G(x) \geq 0 \) for all \( x \in \mathbb{R} \), we have that if \( \theta_G \geq \mu_S \), then for all \( x \in (\theta_G, q) \), we have \( \mu_S - x < 0 \), \( \mu_S - V_2^G(x) \leq 0 \), \( V_2^G(x) \geq 0 \), and hence \( B_S(x, V_2^G) \leq 0 \) holds. Now in order to complete the proof of Statement 11, we only need to consider the case where \( \theta_G < \mu_S \). By the same logic above, for all \( x \in (\mu_S, q) \), we have the desired inequality \( B_S(x, V_2^G) \leq 0 \). Therefore, it only remains to show the inequality for \( x \in (\theta_G, \mu_S) \).

We first determine the expression of \( B_S(x, V_2^G) \) for \( x \in (\theta_G, \mu_S) \) and \( \theta_G < \mu_S \):

\[
B_S(x, V_2^G) = \frac{b}{2 \left( 1 - a^2 + a\hat{\theta}_G \right)} \cdot \left( 2 - 2a^2 + 2a\hat{\theta}_G + F \left( \frac{x - \mu_R}{\sigma_R} \right) \right),
\]

where

\[
F(z) = 2(a - z) \left( a - 3\hat{\theta}_G + 2a\hat{\theta}_G - 2\hat{\theta}_G^2 \right) - 2e^{z^2} - \hat{\theta}_G \left( 1 - a\hat{\theta}_G + \hat{\theta}_G^2 \right) (1 - 2az + 2z^2) + e^{z^2} \sqrt{\pi} \left( a - 3\hat{\theta}_G + 2a\hat{\theta}_G - 2\hat{\theta}_G^2 \right) (1 - 2az + 2z^2) \left( \text{erf}(\hat{\theta}_G) - \text{erf}(z) \right).
\]

Since \( \theta_G < \mu_S \), then \( \hat{\theta}_G \in (a - \frac{1}{a}, a) \) where \( a > 0 \), and hence \( 1 - a^2 + a\hat{\theta}_G > 0 \). Also we know \( b = \mu_S - \mu_R > 0 \). Therefore to complete the proof of Statement 11, we only need to show \( F(z) \leq -2 + 2a^2 - 2a\hat{\theta}_G \) for all \( z \in (\hat{\theta}_G, a) \). Applying Lemma B.15, we obtain that this statement is true.

Now we only need to prove Statement 12 to finish the proof of Proposition 3 under the scheme of \( \mu_R < \mu_S \). Statement 12 says \( B_R(x, V_4^G) \leq 0 \) for any \( x > q \). In fact, for any \( x > q \), we have \( Q(x) = 0 \) and hence

\[
B_R(x, V_4^G) = (\mu_R - x)V_4^G(x) + \frac{1}{2}\sigma_R^2 V_4''(x) + \mu_R
= -\frac{b}{(\mu_S - x)^2} \left( x(x - \mu_S) + \frac{\mu_S \sigma_R^2}{2b} \right).
\]

Since \( x > q > \mu_S > 0 \) and \( b > 0 \), it is clear that \( B_R(x, V_4^G) \leq 0 \). Therefore we have completed the proof of Proposition 3.

\[ \square \]

The proof of Theorem 1 is now immediate.

**Proof of Theorem 1.** The main part of the theorem (the structure of the optimal policy) is an immediate consequence of Proposition 3.

The upper bound on the value function (see Proposition 2 Conditions 1-4) is a bound on the CLV under any admissible policy. A bound on expected customer lifetime under any admissible policy also follows using Wald’s identity, since the expected reward rate is at least \( \mu_S > 0 \) at every instant while the customer is alive.

\[ \square \]
**Proof of Theorem 2.** We first prove the monotonicity results when the Risky mode is superior ($\mu_R > \mu_S$). Consider the parts regarding $\theta_b$. Define the function

$$F(\mu_S, \mu_R, \sigma_R, x) \triangleq \frac{(x - \mu_S)}{\sigma_R} e^{\frac{(x-\mu_R)^2}{\sigma_R^2}} \left(1 - \text{erf}\left( \frac{x - \mu_R}{\sigma_R \sqrt{\pi}} \right) \right) - \frac{\mu_S}{\mu_R \sqrt{\pi}}.$$  

This function is the same one in the left-hand side of Lemma 2, with $\theta$ replaced by $x$ and divided by $-\mu_R \sqrt{\pi}$. By Lemma 2, $F(\mu_S, \mu_R, \sigma_R, \theta_b) = 0$. Moreover, by evaluating the function $r(\cdot)$ from Corollary B.7 at $(x - \mu_R)/\sigma_R$ with parameters $\alpha = (\mu_R - \mu_S)/\sigma_R$ and $\beta = \mu_S/\mu_R$, we obtain that $F(\mu_S, \mu_R, \sigma_R, x) < 0$ for all $x \in (\mu_S, \theta_b)$, and $F(\mu_S, \mu_R, \sigma_R, x) > 0$ for all $x \in (\theta_b, \infty)$. Hence to prove $\theta_b$ is strictly decreasing in $\mu_R$, it suffices to show that $\frac{\partial F}{\partial \mu_R}(\mu_S, \mu_R, \sigma_R, x) > 0$ for all $x > \mu_S$. Similarly, to show $\theta_b$ is strictly increasing in $\mu_S$ and $\sigma_R$, it suffices to show $\frac{\partial F}{\partial \mu_S}(\mu_S, \mu_R, \sigma_R, x) < 0$ for all $x > \mu_S$ and $\frac{\partial F}{\partial \sigma_R}(\mu_S, \mu_R, \sigma_R, x) < 0$ for all $x > \mu_S$. We now write the partial derivatives explicitly:

$$\frac{\partial F}{\partial \mu_R}(\mu_S, \mu_R, \sigma_R, x) = \frac{2(x - \mu_S)}{\sigma_R} \left[ -\frac{x - \mu_R}{\sigma_R} e^{\frac{(x-\mu_R)^2}{\sigma_R^2}} \text{erf}\left( \frac{x - \mu_R}{\sigma_R \sqrt{\pi}} \right) + \frac{1}{\sqrt{\pi}} \right] + \frac{\mu_S}{\mu_R^2 \sqrt{\pi}};$$

$$\frac{\partial F}{\partial \mu_S}(\mu_S, \mu_R, \sigma_R, x) = -\frac{1}{\sigma_R} e^{\frac{(x-\mu_R)^2}{\sigma_R^2}} \text{erf}\left( \frac{x - \mu_R}{\sigma_R \sqrt{\pi}} \right) - \frac{1}{\sqrt{\pi} \mu_R};$$

$$\frac{\partial F}{\partial \sigma_R}(\mu_S, \mu_R, \sigma_R, x) = -\frac{x - \mu_S}{\sigma_R} \left[ e^{\frac{(x-\mu_R)^2}{\sigma_R^2}} \text{erf}\left( \frac{x - \mu_R}{\sigma_R \sqrt{\pi}} \right) \left(1 + \frac{2(x - \mu_R)^2}{\sigma_R^2} \right) \right].$$

Fix any $x > \mu_S$. Then consider a change of variable $\hat{x} \triangleq \frac{x - \mu_R}{\sigma_R}$. Applying Corollary B.8, we get $\frac{\partial F}{\partial \mu_R}(\mu_S, \mu_R, \sigma_R, x) > 0$. The derivative $\frac{\partial F}{\partial \mu_S}(\mu_S, \mu_R, \sigma_R, x) < 0$ follows from the fact that erfc($\cdot$) is non-negative. Finally, applying Lemma B.4, we obtain $\frac{\partial F}{\partial \sigma_R}(\mu_S, \mu_R, \sigma_R, x) < 0$. This completes the proof of monotonicity with regard to $\theta_b$.

Next, we show monotonicity results of the value function and its first order derivative. We already know the functional form of the value function. Hence, we can check the partial derivatives with regard to model parameters directly.

Let us consider the monotonicity of $V^*$ with regard to $\mu_S$. Consider two possible valid values for the Safe mode drift $\mu_S^L < \mu_S^H$. Let $\theta_b^L$ and $\theta_b^H$ be the cutoffs defined in Lemma 2 for $\mu_S^L$ and $\mu_S^H$, respectively. Also, denote the corresponding value functions by $V_L$ and $V_H$. We have just proved that $\theta_b$ is strictly increasing in $\mu_S$, hence $\theta_b^L < \theta_b^H$. There are three cases based on the location of $q$: (1) $\theta_b^L < \theta_b^H < q$, (2) $\theta_b^L < q \leq \theta_b^H$, (3) $q \leq \theta_b^L < \theta_b^H$. In the first case, the myopic policy is optimal for both $\mu_S^L$ and $\mu_S^H$. Therefore, the value functions $V_L$ and $V_H$ are identical. In the second case, $V_L$ is the value function for the myopic policy regardless of the value of the Safe mode drift. Hence, $V_L \leq V_H$ by the optimality of $V_H$ under Safe mode drift $\mu_S^L$. We now consider the last case, $q \leq \theta_b^L < \theta_b^H$. It is easy to verify that for any $x \geq q$, $\frac{\mu_S^L}{x-\mu_S^L} < \frac{\mu_S^H}{x-\mu_S^H}$, i.e., $V_L(x) < V_H(x)$ on $[q, \theta_b^L]$. In particular, $V_L(q) < V_H(q)$. As a result, $V_L(x) < V_H(x)$ (see Eq. (60)) for $x < q$. By continuity, $V_L(q) < V_H(q)$. Therefore $V_L(x) < V_H(x)$ (see Eq. (61)) for $q \leq x \leq \theta_b^L$ since $V_L'(x) < V_H'(x)$ for $q \leq x \leq \theta_b^L$ and $V_L(q) < V_H(q)$. From Lemma 2 and Corollary B.7, we know
that $\frac{\mu_H}{x-\mu_R^L} \geq \frac{\mu_R\sqrt{\pi}}{\sigma_R} e^{\frac{(x-\mu_R)^2}{\sigma_R^2}} \text{erfc}\left(\frac{x-\mu_R}{\sigma_R}\right)$ for $q \leq x \leq \theta_b^H$. Hence $V'_L(x) \leq V'_H(x)$ (see Eqs. (61) and (62)) for $\theta_b^L < x \leq \theta_b^H$, which implies $V_L(x) < V_H(x)$ for $\theta_b^L < x \leq \theta_b^H$. Finally, for $x > \theta_b^H$, we have $V'_L(x) = V'_H(x)$, therefore $V_L(x) < V_H(x)$ since $V_L(\theta_b^H) < V_H(\theta_b^H)$.

We have now proved that the value function and is increasing in $\mu_S$. Next, we will prove the monotonicity of the value function in $\mu_R$ using the same method. Consider two valid values for the Risky mode drift $\mu_R^L < \mu_R^H$. Let $\theta_b^L$ and $\theta_b^H$ be as defined in Lemma 2 for $\mu_R^L$ and $\mu_R^H$, respectively, and denote $V_L$ and $V_H$ as the corresponding value functions. We have proved that $\theta_b$ is strictly decreasing in $\mu_R$, hence $\theta_b^H < \theta_b^L$. Again, there are three cases based on the location of $q$: (1) $\theta_b^L < \theta_b^H < q$, (2) $\theta_b^H < q < \theta_b^L$, (3) $q < \theta_b^H < \theta_b^L$.

Let us first consider case (1). In both the high Risky reward and low Risky reward scenarios, the myopic policy is optimal. Applying Corollary B.9 to $V'_L(x)$ (see Eq. (62)) for $x \geq q$, we obtain that $V'_L(x) < V'_H(x)$ for $x \geq q$. Next we will show that $V'_L(x) < V'_H(x)$ for $x < q$ as well, so that since $\lim_{x \to -\infty} V_L(x) = \mu_R^L < \mu_R^H = \lim_{x \to -\infty} V_H(x)$, we can apply Lemma B.19 and get $V_L(x) < V_H(x)$ for all $x \in \mathbb{R}$.

The explicit expression for $V'_L(x)$ and $V'_H(x)$ on $(-\infty, q)$ is given by

$$V'_L(x) = \frac{\sqrt{\pi} \mu_R^L e^{\frac{x^2}{2}} \text{erfc}(\frac{\hat{q}_L}{y_L}) \left(e^{\frac{x^2}{2}} y_L \text{erfc}(-y_L) + \frac{1}{\sqrt{\pi}}\right)}{\sigma_R \left(e^{\frac{x^2}{2}} \hat{q}_L \text{erfc}(-\hat{q}_L) + \frac{1}{\sqrt{\pi}}\right)}$$

and

$$V'_H(x) = \frac{\sqrt{\pi} \mu_R^H e^{\frac{x^2}{2}} \text{erfc}(\frac{\hat{q}_H}{y_H}) \left(e^{\frac{x^2}{2}} y_H \text{erfc}(-y_H) + \frac{1}{\sqrt{\pi}}\right)}{\sigma_R \left(e^{\frac{x^2}{2}} \hat{q}_H \text{erfc}(-\hat{q}_H) + \frac{1}{\sqrt{\pi}}\right)},$$

where $\hat{q}_L \triangleq \frac{x-\mu_R^L}{\sigma_R}$, $y_L \triangleq \frac{x-\mu_R^L}{\sigma_R}$, $\hat{q}_H \triangleq \frac{x-\mu_R^H}{\sigma_R}$, and $y_H \triangleq \frac{x-\mu_R^H}{\sigma_R}$. Observe that $\hat{q}_L > \hat{q}_H$ and $y_L > y_H$. Since $\mu_R^L < \mu_R^H$ and $e^{\frac{x^2}{2}} \text{erfc}(\hat{q}_L) < e^{\frac{x^2}{2}} \text{erfc}(\hat{q}_H)$ by Corollary B.9, to show $V'_L(x) < V'_H(x)$ for $x < q$, it is sufficient if we can show

$$\frac{e^{\frac{x^2}{2}} y_L \text{erfc}(-y_L) + \frac{1}{\sqrt{\pi}}}{e^{\frac{x^2}{2}} \hat{q}_L \text{erfc}(-\hat{q}_L) + \frac{1}{\sqrt{\pi}}} < \frac{e^{\frac{x^2}{2}} y_H \text{erfc}(-y_H) + \frac{1}{\sqrt{\pi}}}{e^{\frac{x^2}{2}} \hat{q}_H \text{erfc}(-\hat{q}_H) + \frac{1}{\sqrt{\pi}}},$$

or equivalently,

$$\log \left(e^{\frac{x^2}{2}} y_L \text{erfc}(-y_L) + \frac{1}{\sqrt{\pi}}\right) - \log \left(e^{\frac{x^2}{2}} \hat{q}_L \text{erfc}(-\hat{q}_L) + \frac{1}{\sqrt{\pi}}\right)$$

is increasing in $\mu_R^L$. Define $G(y_L) \triangleq \log \left(e^{\frac{x^2}{2}} y_L \text{erfc}(-y_L) + \frac{1}{\sqrt{\pi}}\right)$. By the chain rule, the derivative of $\log \left(e^{\frac{x^2}{2}} y_L \text{erfc}(-y_L) + \frac{1}{\sqrt{\pi}}\right)$ with respect to $\mu_R^L$ is $-\frac{G'(y_L)}{\sigma_R}$, which is decreasing in $y_L$ by Lemma B.20. Therefore, since $y_L < \hat{q}_L$, the derivative of $\log \left(e^{\frac{x^2}{2}} y_L \text{erfc}(-y_L) + \frac{1}{\sqrt{\pi}}\right)$ with respect to $\mu_R^L$ is non-negative.

Now consider case (2). Here, the myopic policy is optimal in the low Risky reward scenario but suboptimal in the high Risky reward scenario. In case (1), we showed that the value function corresponding to the myopic policy in the low Risky reward scenario, $V_L$, is lower than the value
function corresponding to the myopic policy in the high Risky reward scenario. By optimality, the latter must be lower than the optimal value function in the high Risky reward scenario, $V_H$. Hence, $V_L \leq V_H$.

Now consider case (3), where for both low Risky reward scenario and high Risky reward scenarios, sandwich policies are optimal. By Corollary B.9, we have $V'_L(x) < V'_H(x)$ for $x > \theta_b^L$. We also know that $V'_L(x) = V'_H(x)$ for $q \leq x \leq \theta_b^H$. We only need to consider $x < q$ and $\theta_b^H < x \leq \theta_b^L$. For the latter, we can refer to the definition of $\theta_b^L$ in Lemma 2 and Corollary B.7 to get

\[
\frac{\mu_S}{x - \mu_S} \leq \frac{\mu_H}{\sigma_R} \frac{(x - \mu_R)^2}{\sigma_R^2} \text{erfc} \left( \frac{x - \mu_R}{\sigma_R} \right) \text{erfc} \left( \frac{x - \mu_R}{\sigma_R} \right)
\]

for $x \geq \theta_b^H$, which implies $V'_L(x) \leq V'_H(x)$ for $x \in [\theta_b^H, \theta_b^L]$ (see Eqs. (61) and (62)). For the former, we will now show $V'_L(x) < V'_H(x)$ for $x < q$. The expressions for $V'_L(x)$ is given by

\[
V'_L(x) = \frac{\mu_S}{(q - \mu_S)} \left( e^{y_L^2} y_L \text{erfc}(-y_L) + \frac{1}{\sqrt{\pi}} \right)
\]

where $y_L \triangleq \frac{x - \mu_L}{\sigma_L}$ and $y_H \triangleq \frac{x - \mu_H}{\sigma_H}$. Therefore to show $V'_L(x) < V'_H(x)$ for $x < q$, it is equivalent to show $\log \left( e^{y_L^2} y_L \text{erfc}(-y_L) + \frac{1}{\sqrt{\pi}} \right) - \log \left( e^{y_H^2} y_H \text{erfc}(-y_H) + \frac{1}{\sqrt{\pi}} \right)$ is increasing in $\mu_R$. By the same argument in case (1), we know that this is true.

Finally since $\lim_{x \to -\infty} V_L(x) = \mu_R^L < \mu_R^H = \lim_{x \to -\infty} V_H(x)$, we can apply Corollary B.9 to get $V_L(x) < V_H(x)$ for all $x \in \mathbb{R}$. This completes the proof of monotonicity with regard to $\mu_R$.

Lastly we want to show monotonicity results of the first order derivative of the value function for $x \geq q$. Consider two valid Risky volatility values $\sigma_R^L < \sigma_R^H$. Let $\theta_b^L$ and $\theta_b^H$ be as be defined in Lemma 2 for $\sigma_R^L$ and $\sigma_R^H$, respectively, and denote $V_L$ and $V_H$ as the corresponding value functions. We have proved that $\theta_b$ is strictly increasing in $\sigma_R$, hence $\theta_b^L < \theta_b^H$ Again, there are three cases based on the location of $q$: (1) $\theta_b^L < \theta_b^H < q$, (2) $\theta_b^L < q \leq \theta_b^H$, (3) $q \leq \theta_b^L < \theta_b^H$.

In case (1), for $x \geq q$, we have

\[
f(\sigma_R^L, x) \triangleq V'_L(x) = \frac{\mu_R \sqrt{\pi}}{\sigma_R^L} \frac{(x - \mu_R)^2}{\sigma_R^L} \text{erfc} \left( \frac{x - \mu_R}{\sigma_R^L} \right)
\]

Take its derivative with respect to $\sigma_R$, we get

\[
\frac{\partial f(\sigma_R^L, x)}{\partial \sigma_R^L} = -\frac{\mu_R \hat{f}(\sigma_R^L)}{(\sigma_R^L)^2},
\]

where

\[
\hat{f}(\sigma_R^L) \triangleq e^{(x - \mu_R)^2} \left( 1 + \frac{2(x - \mu_R)^2}{(\sigma_R^L)^2} \right) \text{erfc} \left( \frac{x - \mu_R}{\sigma_R^L} \right) - \frac{2(x - \mu_R)}{\sigma_R^L}.
\]

By Lemma B.4, $\hat{f}(\sigma_R^L) \geq 0$, which implies that $\frac{\partial f(\sigma_R^L, x)}{\partial \sigma_R^L} \leq 0$, i.e., $V'_L(x) \geq V'_H(x)$ for $x \geq q$.

Similarly, in case (2), we have $V'_L(x) \geq V'_H(x)$ for $x \geq \theta_b^H$. Hence, we only need to show $V'_L(x) \geq V'_H(x)$ for $q \leq x < \theta_b^H$. In fact, by the definition of $\theta_b^L$, we get $\frac{\mu_S}{x - \mu_S} \leq \frac{\mu_H}{\sigma_R} \frac{(x - \mu_R)^2}{\sigma_R^2} \text{erfc} \left( \frac{x - \mu_R}{\sigma_R} \right)$ for $x \geq \theta_b^L$, which implies $V'_L(x) \leq V'_H(x)$ for $x \in [q, \theta_b^H]$ (see Eqs. (61) and (62)).
Lemma 3. Following the notations:

\[ V(x) \]

where

\[ f(\theta) \]

We introduced the notation

\[ \theta \]

In fact, by the definition of \( \theta^L \), we get

\[ \frac{\mu_S}{x-\mu_S} \leq \frac{\mu_R \sqrt{\pi}}{\sigma_R} e^{-\frac{(x-\mu_R)^2}{2\sigma_R^2}} \text{erfc}\left(\frac{x-\mu_R}{\sigma_R}\right) \text{ for } x \geq \theta^L, \]

which implies \( V'_L(x) \leq V'_L(x) \) for \( x \in [\theta^L, \theta^H] \) (see Eqs. 61 and 62).

Next we prove the monotonicity result when the Safe mode is superior \( (\mu_S > \mu_R) \).

We want to show that \( \theta_G \) is monotonically decreasing in \( \sigma_R \). Recall the definition of \( \theta_G \) in Lemma 3. Following the notations:

\[ a \triangleq \frac{\mu_S - \mu_R}{\sigma_R}, \quad b \triangleq \mu_S - \mu_R, \quad \hat{\theta}_G \triangleq \frac{\theta_G - \mu_R}{\sigma_R}. \]

We already know that \( \theta_G \) is the unique root of either of two monotone strictly-decreasing functions on \( (a - \frac{1}{a}, q) \). Therefore if we can show the two function values are decreasing in \( \sigma_R \), then we are done. Consider a fixed value of \( \theta \) such that \( \theta \in (a - \frac{1}{a}, q) \).

We first put down the values of the two functions, both evaluated at \( \theta \):

\[
\begin{align*}
\text{small}(\sigma_R) &= \frac{\mu_S \sigma_R}{(\mu_S - \mu_R)(q - \mu_S)} + 2\left( e^{\frac{(q - \mu_R)^2 - (q - \mu_S)^2}{2\sigma_R^2}} \frac{q - \mu_R - \theta - \mu_R}{\sigma_R} + e^{\frac{(q - \mu_R)^2}{\sigma_R^2}} \frac{(q - \mu_R)(\theta - \mu_R)}{\sigma_R^2} \sqrt{\pi} \left( \text{erf}\left(\frac{q - \mu_R}{\sigma_R}\right) \right) \right), \\
\text{big}(\sigma_R) &= -\frac{\mu_S \sigma_R}{(\mu_S - \mu_R)(q - \mu_S)} + \left( \frac{\mu_S - \mu_R}{\sigma_R} - 3 \frac{\theta - \mu_R}{\sigma_R} + 2 \frac{(\mu_S - \theta)(\theta - \mu_R)^2}{\sigma_R^3} \right) + 2e^{\frac{(q - \mu_R)^2 - (q - \mu_S)^2}{2\sigma_R^2}} \frac{q - \mu_R}{\sigma_R} \cdot \left( 1 - \frac{(\theta - \mu_R)(\mu_S - \theta)}{\sigma_R} \right) + e^{\frac{(q - \mu_R)^2}{\sigma_R^2}} \frac{q - \mu_R}{\sigma_R} \sqrt{\pi} \left( \frac{\mu_S - \mu_R}{\sigma_R} \right) - 3 \frac{\theta - \mu_R}{\sigma_R} + 2 \frac{(\mu_S - \theta)(\theta - \mu_R)^2}{\sigma_R^3} \left( \text{erf}\left(\frac{q - \mu_R}{\sigma_R}\right) \right) - \text{erf}\left(\frac{\theta - \mu_R}{\sigma_R}\right)) \right).
\end{align*}
\]

We introduced the notation \( f_{\text{small}}(\sigma_R) \) and \( f_{\text{big}}(\sigma_R) \) so that we can compute the derivatives with regard to \( \sigma_R \). We want to show that \( f'_{\text{small}}(\sigma_R) \) and \( f'_{\text{big}}(\sigma_R) \) are non-positive.

Consider first \( f'_{\text{small}}(\sigma_R) \):

\[
f'_{\text{small}}(\sigma_R) = \frac{1}{\sigma_R} \cdot g_1\left(\frac{\theta - \mu_R}{\sigma_R}, \frac{q - \mu_R}{\sigma_R}, \frac{\mu_S - \mu_R}{\sigma_R}, \frac{q - \mu_R}{\sigma_R}\right),
\]

where

\[
g_1(\hat{\theta}, \hat{q}, a) = \frac{\mu_S}{(\mu_S - \mu_R)(a - \hat{q})} - 2\hat{q}(1 + 2\hat{q}^2)e^{\hat{q}^2 - \hat{q}^2} + 2\hat{q}(1 + 2\hat{q}^2) \\
+ 4\hat{q}\hat{\theta}(1 + \hat{q}^2)\sqrt{\pi}e^{\hat{q}^2} \left( \text{erf}(\hat{q}) - \text{erf}(\hat{\theta}) \right).
\]
To show $f'_{\text{small}}(\sigma_R) \leq 0$, it is equivalent to show $g_1(\hat{\theta}, \hat{q}, a) \leq 0$ for $\hat{q} > a > 0$, $\hat{\theta} \in (a - \frac{1}{a}, \hat{q})$. Since $\mu_S > \mu_R > 0$ and $\hat{q} > a > 0$, we have $\frac{\mu_S}{\mu_S - \mu_R} > 1$ and $\frac{1}{a-\hat{q}} < \frac{1}{a}$, and hence $\frac{\mu_S}{(\mu_S - \mu_R)(a-\hat{q})} < \frac{1}{a}$.

Thus, we have

$$g_1(\hat{\theta}, \hat{q}, a) < -\frac{1}{\hat{q}} - 2\hat{q}(1 + 2\hat{q}^2)e^{\hat{q}^2 - \hat{\theta}^2} + 2\hat{\theta}(1 + 2\hat{q}^2) + 4\hat{q}\hat{\theta}(1 + \hat{q}^2)\sqrt{\pi}e^{\hat{\theta}^2} \left( \text{erf}(\hat{q}) - \text{erf}(\hat{\theta}) \right).$$

Observe that the RHS of the above inequality only depends on $\hat{\theta}$ and $\hat{q}$, hence we denote it by $\tilde{g}_1(\hat{\theta}, \hat{q})$. Since $g_1(\hat{\theta}, \hat{q}, a) < \tilde{g}_1(\hat{\theta}, \hat{q})$, to show $g_1(\hat{\theta}, \hat{q}, a) \leq 0$ it is sufficient to show $\tilde{g}(\hat{\theta}, \hat{q}) \leq 0$. If $\hat{\theta} \leq 0$, then this is true, since $\hat{q} > 0$ and $\text{erf}(\cdot)$ is an increasing function. On the other hand, if $\hat{\theta} > 0$, since $\hat{\theta} < \hat{q}$, it must be that $\hat{\theta} \in (0, \hat{q})$. In this case, we have

$$\frac{\partial^2 g_1}{\partial \hat{\theta}^2}(\hat{\theta}, \hat{q}) = -4e^{\hat{q}^2 - \hat{\theta}^2}\hat{q}\left(3 + 2\hat{q}^2 - 2\hat{\theta}^2\right) < 0,$$

and hence

$$\frac{\partial \tilde{g}_1}{\partial \hat{\theta}}(\hat{\theta}, \hat{q}) = 2 + 4\hat{q}^2 - 4e^{\hat{q}^2 - \hat{\theta}^2}\hat{q}\hat{\theta} + 4e^{\hat{\theta}^2}\hat{q}(1 + \hat{q}^2)\sqrt{\pi} \left( \text{erf}(\hat{q}) - \text{erf}(\hat{\theta}) \right) > \frac{\partial \tilde{g}_1}{\partial \hat{\theta}}(\hat{\theta}, \hat{q}) = 2 > 0.$$

Therefore

$$\tilde{g}_1(\hat{\theta}, \hat{q}) < \tilde{g}_1(\hat{q}, \hat{q}) = -\frac{1}{\hat{q}} < 0,$$

which is what we want. We have thus proved $f'_{\text{small}}(\sigma_R) \leq 0$.

Now it only remains to show $f'_\text{big}(\sigma_R) \leq 0$. Since $f'_\text{big}(\sigma_R)$ is only relevant to the value of $\theta_G$ in the “big” sandwich case, i.e., when $\theta_G < \mu_S$, it suffices to consider $\theta \in \left(\mu_S - \frac{\sigma^2}{\mu_S - \mu_R}, \mu_S\right)$ in the rest of this proof.

Compute $f'_\text{big}(\sigma_R)$, we get

$$f'_\text{big}(\sigma_R) = a \frac{a^2 - a \hat{\theta} + 1}{a - \hat{q}} \cdot h_1 \left( \frac{\theta - \mu_R}{\sigma_R}, \frac{q - \mu_R}{\sigma_R}, \frac{\mu_S - \mu_R}{\sigma_R} \right),$$

where

$$h_1(\hat{\theta}, \hat{q}, a) = \frac{\mu_S}{\mu_S - \mu_R} \cdot \frac{a^2 - a \hat{\theta} + 1}{a - \hat{q}} - a + 3\hat{\theta} - 6a\hat{\theta}^2 + 6\hat{\theta}^3 - 2\hat{q}^2 \left( a - 3\hat{\theta} + 2a\hat{\theta}^2 - 2\hat{\theta}^3 \right) + e^{\hat{q}^2 - \hat{\theta}^2}\hat{q} \left( 6a\hat{\theta} - 6\hat{\theta}^2 - (1 + 2\hat{q}^2) \left( 2 - 2a\hat{\theta} + 2\hat{\theta}^2 \right) \right) + e^{\hat{\theta}^2}\hat{q}\sqrt{\pi} \left( -a + 3\hat{\theta} - 6a\hat{\theta}^2 + 6\hat{\theta}^3 - (1 + 2\hat{q}^2) \left( a - 3\hat{\theta} + 2a\hat{\theta}^2 - 2\hat{\theta}^3 \right) \right) \left( \text{erf}(\hat{q}) - \text{erf}(\hat{\theta}) \right).$$

To show $f'_\text{big}(\sigma_R) \leq 0$, it is equivalent to show $h_1(\hat{\theta}, \hat{q}, a) \leq 0$ for $\hat{q} > a > 0$, $\hat{\theta} \in (a - \frac{1}{a}, a)$. Observe that $a^2 - a \hat{\theta} + 1 > 0$ since $\hat{\theta} < a$ and $a > 0$. Also $a - \hat{q} < 0$ and $\frac{\mu_S}{\mu_S - \mu_R} > 1$. Hence we have $\frac{\mu_S}{\mu_S - \mu_R} \cdot \frac{a^2 - a \hat{\theta} + 1}{a - \hat{q}} < a^2 - a \hat{\theta} + 1$, and

$$h_1(\hat{\theta}, \hat{q}, a) < a^2 - a \hat{\theta} + 1 - a + 3\hat{\theta} - 6a\hat{\theta}^2 + 6\hat{\theta}^3 - 2\hat{q}^2 \left( a - 3\hat{\theta} + 2a\hat{\theta}^2 - 2\hat{\theta}^3 \right) + e^{\hat{q}^2 - \hat{\theta}^2}\hat{q} \left( 6a\hat{\theta} - 6\hat{\theta}^2 - (1 + 2\hat{q}^2) \left( 2 - 2a\hat{\theta} + 2\hat{\theta}^2 \right) \right) + e^{\hat{\theta}^2}\hat{q}\sqrt{\pi} \left( -a + 3\hat{\theta} - 6a\hat{\theta}^2 + 6\hat{\theta}^3 - (1 + 2\hat{q}^2) \left( a - 3\hat{\theta} + 2a\hat{\theta}^2 - 2\hat{\theta}^3 \right) \right) \left( \text{erf}(\hat{q}) - \text{erf}(\hat{\theta}) \right).$$
Denote the RHS of the above by $\tilde{h}_1(a)$. If we can show $\tilde{h}_1(a) \leq 0$ then we are done. Next we will show $\tilde{h}_1(\cdot)$ is monotone decreasing on $(0, \hat{q})$. Since $\hat{q} > a > 0$ and $\hat{q} > a > \hat{\theta}$, we have

$$\tilde{h}_1''(a) = -\frac{2(1 + \hat{q}(\hat{q} - \hat{\theta}))}{(q - a)^3} < 0,$$

which implies

$$\tilde{h}_1'(a) = -1 - 6\hat{\theta}^2 - 2\hat{q}^2(1 + 2\hat{\theta}^2) + \frac{a^2 - 2a\hat{q} + \hat{\theta}\hat{q} - 1}{(a - \hat{q})^2} + e^{\hat{q}^2 - \hat{\theta}^2} \hat{q} \left(6\hat{\theta} + 2\hat{\theta}(1 + 2\hat{q}^2)\right)$$

$$- e^{\hat{q}^2} \hat{q} \left((1 + 2\hat{q}^2)(1 + 2\hat{\theta}^2) + 1 + 6\hat{\theta}^2\right) \sqrt{\pi} \left(\text{erf}(\hat{q}) - \text{erf}(\hat{\theta})\right)$$

$$< \tilde{h}_1'(0)$$

$$= -1 - 6\hat{\theta}^2 - 2\hat{q}(1 + 2\hat{\theta}^2) + \frac{\hat{\theta}}{\hat{q}} - \frac{1}{\hat{q}^2} + e^{\hat{q}^2 - \hat{\theta}^2} \hat{q} \left(6\hat{\theta} + 2\hat{\theta}(1 + 2\hat{q}^2)\right)$$

$$- e^{\hat{q}^2} \hat{q} \left((1 + 2\hat{q}^2)(1 + 2\hat{\theta}^2) + 1 + 6\hat{\theta}^2\right) \sqrt{\pi} \left(\text{erf}(\hat{q}) - \text{erf}(\hat{\theta})\right)$$

$$:= \hat{h}_1(\hat{\theta}).$$

We will now show $\hat{h}_1(\hat{\theta}) < 0$ for $\hat{\theta} < a < \hat{q}$. First one can verify that $\hat{h}_1(\hat{q}) = -\frac{1}{\hat{q}^2} < 0$. Therefore we will only show $\hat{h}_1'(\hat{\theta}) \geq 0$ for $\hat{\theta} < \hat{q}$. Since $\hat{\theta} < \hat{q}$ and that $\text{erf}(\cdot)$ is a monotone increasing function, if $\hat{\theta} \leq 0$, then

$$\hat{h}_1''(\hat{\theta}) = -12 - 8\hat{q}^2 + 8e^{\hat{q}^2 - \hat{\theta}^2} \hat{q} - 8e^{\hat{q}^2} \hat{q} \left(2 + \hat{q}^2\right) \sqrt{\pi} \left(\text{erf}(\hat{q}) - \text{erf}(\hat{\theta})\right) < 0.$$

On the other hand, if $\hat{\theta} > 0$, then for $\hat{\theta} \in (0, \hat{q})$, we have

$$\hat{h}_1'''(\hat{\theta}) = 8e^{\hat{q}^2 - \hat{\theta}^2} \left(5 + 2\hat{q}^2 - 2\hat{\theta}^2\right) > 0,$$

which implies that

$$\hat{h}_1''(\hat{\theta}) < \hat{h}_1''(\hat{q}) = -12 < 0.$$

Hence $\hat{h}_1'(\hat{\theta})$ is monotone decreasing for $\hat{\theta} < \hat{q}$, and we have

$$\hat{h}_1'(\hat{\theta}) = \frac{1}{\hat{q}} + e^{\hat{q}^2 - \hat{\theta}^2} \left(12\hat{q} + 8\hat{q}^3\right) - 12\hat{\theta} - 8\hat{q}^2\hat{\theta} - 8e^{\hat{q}^2} \hat{q} \left(2 + \hat{q}^2\right) \hat{\theta} \sqrt{\pi} \left(\text{erf}(\hat{q}) - \text{erf}(\hat{\theta})\right)$$

$$> \hat{h}_1'(\hat{q}) = \frac{1}{\hat{q}} > 0,$$

which is what we want in order to have $\hat{h}_1(\hat{\theta}) < 0$ for $\hat{\theta} < a < \hat{q}$, and thus $\tilde{h}_1'(a) < 0$. Recall that we want to show $\tilde{h}_1(a) \leq 0$ so that $f'_\text{big}(\sigma_R) \leq 0$. Since we now know that $\tilde{h}_1'(a) < 0$ for $a \in (\hat{\theta}, \hat{q})$, we have

$$\tilde{h}_1(a) < \tilde{h}_1(\hat{\theta})$$

$$= 2\hat{\theta}(1 + 2\hat{q}^2) - \frac{1}{\hat{q} - \hat{\theta}} - 2e^{\hat{q}^2 - \hat{\theta}^2} \hat{q}(1 + 2\hat{q}^2) + 4e^{\hat{q}^2} \hat{q}(1 + \hat{q}^2) \hat{\theta} \sqrt{\pi} \left(\text{erf}(\hat{q}) - \text{erf}(\hat{\theta})\right)$$

$$< 2\hat{\theta}(1 + 2\hat{q}^2) - 2e^{\hat{q}^2 - \hat{\theta}^2} \hat{q}(1 + 2\hat{q}^2) + 4e^{\hat{q}^2} \hat{q}(1 + \hat{q}^2) \hat{\theta} \sqrt{\pi} \left(\text{erf}(\hat{q}) - \text{erf}(\hat{\theta})\right)$$

$$\leq 0,$$

where the last step follows from Lemma B.18.

This completes the proof. □

37
D Appendix to Section 5 (Investor Model)

In this appendix, we give the proofs of results of Section 5. These include Lemmas 5 and 6, and Theorems 3 and 4.

Proof of Lemma 5. Define

\[ \alpha \triangleq \frac{\mu_R - \mu_S}{\sigma_R}, \quad \beta \triangleq \frac{\mu_S}{\mu_R} \quad \text{and} \quad z \triangleq \frac{\theta - \mu R}{\sigma_R}. \]

We have \( \alpha > 0 \) and \( 0 < \beta < 1 \). Define also

\[ r(z) \triangleq (z + \frac{\alpha}{2})e^{z^2} \text{erfc}(z) - \frac{1 + \beta}{2\sqrt{\pi}}. \]

The lemma is equivalent to showing that \( r(z) \) has a unique root on \( (-\frac{\alpha}{2}, \infty) \). Applying Corollary B.7, we get that \( r(z) \) has a unique root on \( (-\frac{\alpha}{2}, \infty) \).

Proof of Lemma 6. By Lemma B.2, the function \( X(\cdot) : [g - 1, \infty) \rightarrow (\mu_S, \theta_I) \) as specified in Lemma 6 is strictly decreasing and differentiable, and \( X(g - 1) = \theta_I \), and \( \lim_{y \rightarrow \infty} X(y) = \mu_S \). Therefore, its inverse function \( G(\cdot) : (\mu_S, \theta_I) \rightarrow [g - 1, \infty) \) is well-defined, strictly decreasing and differentiable, and \( G'(x) = \frac{1}{X(G(x))} \) by the inverse function theorem. Let \( G(x) = y \), \( x = X(y) \), and \( G'(x) = \frac{1}{X(y)} \), one can easily check that \( (y + 1)^2 + 2a(x(y) - b)X(y) - 2ab = 0 \). Therefore \( G(\cdot) \) satisfies the ODE \( (G(x) + 1)^2 + 2a(x-b)G(x)G'(x) - 2abG'(x) = 0 \) on \( [\mu_S, \infty) \). Since \( g - 1 > 0 \), we have that \( G(\cdot) \) is strictly positive on \( (\mu_S, \theta_I] \). Finally by Lemma B.3, the function \( \lambda^*(\cdot) \) is strictly increasing on \([q, \theta_I] \), with \( \lambda^*(q) > 0 \) and \( \lambda^*(\theta_I) = 1 \).

Proof of Theorem 3. The proof technique is similar to one we used prove Theorem 1 (see Section 4.2). Recall that in Section 4.2, we first obtain a candidate value function \( W(\cdot) \) by solving the HJB equation (13), then we prove its optimality by showing that it satisfies a set of optimality conditions (Conditions 1-6 in Proposition 2). Similarly, for the Investor model, we will first provide a candidate value function \( W^I(\cdot) \) by solving the HJB equation (21), and then prove its optimality by showing that \( W^I(\cdot) \) together with the policy stated in Theorem 3 satisfies Conditions 1-6, with Condition 4'-5' replacing Condition 4-5.

Definition of \( W^I(\cdot) \). Define

\[ W^I(x) = \begin{cases} V_1(x,C_1^I) & \text{if } x < q; \\ V_2^I(x,C_2^I) & \text{if } q \leq x \leq \theta_I; \\ V_3(x,C_3^I) & \text{if } x > \max\{q, \theta_I\}; \end{cases} \]

(78)

for some uniquely specified \( C_1^I, C_2^I \) and \( C_3^I \), where \( V_1 \) and \( V_3 \) are as defined in Eqs. (60) and (62), restated here:

\[ V_1(x,C_1) = C_1 e^{\frac{(x-\mu R)^2}{\sigma_R}} \left( 1 + \text{erf}\left( \frac{x-\mu R}{\sigma_R} \right) \right) + \mu R; \]

\[ V_3(x,C_3) = C_3 + \int_0^x \frac{\mu R\sqrt{\pi}}{\sigma_R} e^{\frac{(x-\mu R)^2}{\sigma_R}} (1 - \text{erf}(\frac{z-\mu R}{\sigma_R}))dz. \]
$V_2^I$ is defined as follows:

$$V_2^I(x, C_2) = C_2 + \int_q^x G(z)dz,$$  \hspace{1cm} (79)

where $G$ is as defined in Lemma 6. Observe that $V_2^I(\cdot, C_2)$ and $V_3^I(\cdot, C_3)$ are independent of $C_2$ and $C_3$, hence we will use $V_2^I(\cdot)$ and $V_3^I(\cdot)$ to denote the two functions, respectively.

We now provide the explicit expressions of $C_1^I, C_2^I$ and $C_3^I$:

$$C_1^I = \begin{cases} \frac{\mu_R \sqrt{\pi}}{\sigma_R} e^{\frac{(q-\mu_R)^2}{\sigma_R^2}} (1-\text{erf}(\frac{q-\mu_R}{\sigma_R})) & \text{if } \theta_I < q; \\ \frac{2}{\sqrt{\pi}} + \frac{q-\mu_R}{\sigma_R} (1+\text{erf}(\frac{q-\mu_R}{\sigma_R})) & \text{if } \theta_I \geq q, \end{cases}$$

and

$$C_2^I = C_1^I e^{\frac{(q-\mu_R)^2}{\sigma_R^2}} \left(1 + \text{erf}(\frac{q-\mu_R}{\sigma_R})\right) + \mu_R,$$  \hspace{1cm} (80)

and

$$C_3^I = \begin{cases} V_1(q, C_1^I) - \int_0^q \frac{\mu_R \sqrt{\pi}}{\sigma_R} e^{\frac{(z-\mu_R)^2}{\sigma_R^2}} (1 - \text{erf}(\frac{z-\mu_R}{\sigma_R}))dz & \text{if } \theta_I < q; \\ V_2^I(\theta_I, C_2^I) - \int_0^{\theta_I} \frac{\mu_R \sqrt{\pi}}{\sigma_R} e^{\frac{(z-\mu_R)^2}{\sigma_R^2}} (1 - \text{erf}(\frac{z-\mu_R}{\sigma_R}))dz & \text{if } \theta_I \geq q. \end{cases}$$

To reduce the burden of notation, we define $V_1^I(\cdot) \triangleq V_1(\cdot, C_1^I)$, $V_2^I(\cdot) \triangleq V_2^I(\cdot, C_2^I)$ and $V_3^I(\cdot) \triangleq V_3^I(\cdot, C_3^I)$.

**Conditions 1, 2, 3, 4, 5 and 6.** Now we have an explicitly defined candidate value function $W^I(\cdot)$ and an explicit stationary Markov policy $\lambda^*(\cdot)$ (see Lemma 6). Next we want to show that $W^I(\cdot)$ and $\lambda^*(\cdot)$ together satisfy Conditions 1-3 (see Proposition 2), 4, 5 and 6 (see (22) and (23)) and 6 (see Proposition 2).

We start with Condition 1. We want to show that $W^I(\cdot) \geq 0$. By construction of $V_1$ (see Eq. (60)), if $C_1^I > 0$ then $V_1^I(\cdot) > 0$. Indeed this is true since $G(q) > 0$ (see Lemma 6). Then, in particular $V_1^I(q) > 0$ and $W^I(\cdot) > 0$, since $W^I(\cdot)$ is continuous everywhere (see below) including at $q$ and increases on $[q, \infty)$, see Eq. (78).

Condition 2 requires $W^I(\cdot)$ to be continuously differentiable everywhere and twice continuously differentiable almost everywhere. By construction, $W^I(\cdot)$ is continuously differentiable and twice continuously differentiable everywhere except possibly at $q$ and $\theta_I$ (note that by Lemma 6, $G$ is differentiable on $[q, \theta_I]$). Hence, we only need to show that $W^I(\cdot)$ is differentiable at $q$ and $\theta_I$. Equivalently, we want to show

$$V_1^I(q) = \begin{cases} V_3^I(q) & \text{if } \theta_I < q; \\ V_2^I(q) & \text{if } \theta_I \geq q, \end{cases}$$

and

$$V_1^I(\theta_I) = \begin{cases} V_3^I(q) & \text{if } \theta_I < q; \\ V_2^I(q) & \text{if } \theta_I \geq q, \end{cases}$$

where $G$ is as defined in Lemma 6. Observe that $V_2^I(-, C_2)$ and $V_3^I(-, C_3)$ are independent of $C_2$ and $C_3$, hence we will use $V_2^I(\cdot)$ and $V_3^I(\cdot)$ to denote the two functions, respectively.
and if \( \theta_I \geq q \),

\[
V_2^I(\theta_I) = V_3^I(\theta_I), \tag{85}
\]

\[
V_2''(\theta_I) = V_3''(\theta_I). \tag{86}
\]

Eq. (83) is implied by the definitions of \( C_2^I \) and \( C_3^I \) (see Eqs. (81) and (82)). Eq. (84) is implied by the definition of \( C_1^I \) (see Eq. (80)). Eq. (85) is implied by the definition of \( C_3^I \) (see Eq. (82)). Eq. (86) is implied by the fact that \( G^{-1}(V_3''(\theta_I)) = \theta_I \) (see Lemma 6).

Condition 3 requires that \( W''(.) \) be bounded. Since we have just proved that \( W''(.) \) is continuous in \( \mathbb{R} \), to show that \( W''(.) \) is bounded, it suffices to show \( \lim_{x \to -\infty} W''(x) < \infty \) and \( \lim_{x \to \infty} W''(x) < \infty \). This is equivalent to showing \( \lim_{x \to -\infty} V_1''(x) < \infty \) and \( \lim_{x \to \infty} V_3''(x) < \infty \). By the definitions of \( V_1^I \) (see Eqs. (60) and (80)) and \( V_3^I \) (see Eqs. (62) and (82)), we have

\[
\lim_{x \to -\infty} V_1''(x) = \lim_{x \to -\infty} \frac{2C_1^I}{\sigma_R} \left( \frac{1}{\sqrt{\pi}} + e^{-\frac{(x-\mu_R)^2}{\sigma_R^2}} x - \mu_R \left( 1 + \text{erf} \left( \frac{x - \mu_R}{\sigma_R} \right) \right) \right)
\]

\[
= \lim_{x \to -\infty} \frac{2C_1}{\sigma_R} \left( \frac{1}{\sqrt{\pi}} - ze^z \text{erfc}(z) \right)
\]

and

\[
\lim_{x \to \infty} V_3''(x) = \lim_{x \to \infty} \frac{\mu_R \sqrt{\pi}}{\sigma_R} e^{-\frac{(x-\mu_R)^2}{\sigma_R^2}} \left( 1 - \text{erf} \left( \frac{x - \mu_R}{\sigma_R} \right) \right)
\]

\[
= \lim_{x \to \infty} \frac{\mu_R \sqrt{\pi}}{\sigma_R} e^z \text{erfc}(z),
\]

for \( z = (x - \mu_R)/\sigma_R \). From Lemma B.5, the limits above are zero.

Conditions 5\( I \) and 4\( I \) respectively require \( W^I(.) \) and \( \lambda^*(.) \) to satisfy

\[
-Q(x)W^I(x) + (\mu_S(1 - \lambda^*(x)) + \mu_R \lambda^*(x) - x)W''(x) + \frac{1}{2}(\sigma_R \lambda^*(x))^2 W''(x)
\]

\[
+ \mu_S(1 - \lambda^*(x)) + \mu_R \lambda^*(x) = 0 \tag{87}
\]

and

\[
-Q(x)W^I(x) + (\mu_S(1 - p) + \mu_R p - x)W''(x) + \frac{1}{2}(\sigma_R p)^2 W''(x)
\]

\[
+ \mu_S(1 - p) + \mu_R p \leq 0 \quad \text{for all } p \in [0, 1] \tag{88}
\]

for all \( x \in \mathbb{R} \), except possibly at \( q \) and \( \theta_I \). Eq. (87) is true by the construction of \( W^I(.) \) and \( \lambda^*(.) \). Specifically, \( V_1^I(x) \) satisfies

\[
-V_1^I(x) + (\mu_R - x) V_1''(x) + \frac{1}{2} \sigma_R^2 V_1''(x) + \mu_R = 0 \quad \text{for all } x < q;
\]

\( V_3^I(x) \) satisfies

\[
(\mu_R - x) V_3''(x) + \frac{1}{2} \sigma_R^2 V_3''(x) + \mu_R = 0 \quad \text{for all } x \geq \theta_I;\]
and \( G(x) \) satisfies (see Lemma 6)
\[
(G(x) + 1)^2 + 2a(x-b)G(x)G'(x) - 2abG'(x) = 0,
\]
which is equivalent to (since \( G'(\cdot) \) is nonzero)
\[
(\mu_S(1 - \lambda^*(x)) + \mu_R\lambda^*(x) - x)G(x) + \frac{1}{2}(\sigma_R\lambda^*(x))^2G'(x)
+ \mu_S(1 - \lambda^*(x)) + \mu_R\lambda^*(x) = 0 \text{ for all } x \in [q, \theta_I].
\]

Next we will show Eq. (88) to complete Condition 4'. Denote the LHS of Eq. (88) by \( B(p, x, W^I) \). Then, we need to show \( B(p, x, V^I_1) \leq 0 \) for \( x < q \), \( B(p, x, V^I_2) \leq 0 \) for \( x \in [q, \theta_I] \), and \( B(p, x, V^I_3) \leq 0 \) for \( x > \max\{q, \theta_I\} \); each for all \( p \in [0, 1] \).

Let us first start with \( x < q \). We already know that \( B(1, x, V^I_1) = 0 \). Therefore, to show \( B(p, x, V^I_1) \leq 0 \) is equivalent to showing \( B(1, x, V^I_1) - B(p, x, V^I_1) \geq 0 \). Rearranging the terms, we can get
\[
B(1, x, V^I_1) - B(p, x, V^I_1) = (1 - p) \left( (\mu_R - \mu_S)V^I_1(x) + \frac{(1 + p)\sigma_R^2 V^I_1''(x) + \mu_R - \mu_S}{2} \right),
\]
In fact, by Lemma B.4, Corollary B.8 and the fact that \( C^I_1 > 0 \), we have
\[
V^I_1''(x) = 2C^I_1 \frac{1}{\sigma_R} \left( \frac{1}{\sqrt{\pi}} + \frac{e^{(x-\mu_R)^2/\sigma_R}}{\sigma_R \sqrt{\pi}} \text{erfc} \left( - \frac{x-\mu_R}{\sigma_R} \right) \right) \geq 0
\]
and
\[
V^I_1'''(x) = 2C^I_1 \frac{2(x-\mu_R)}{\sigma_R^2} + \frac{e^{(x-\mu_R)^2/\sigma_R}}{\sigma_R \sqrt{\pi}} \left( 1 + 2 \frac{(x-\mu_R)^2}{\sigma_R^2} \right) \text{erfc} \left( - \frac{x-\mu_R}{\sigma_R} \right) \geq 0.
\]

Therefore since \( p \in [0, 1] \) and \( \mu_R > \mu_S \), following Eq. (89), we obtain the desired inequality \( B(1, x, V^I_1) - B(p, x, V^I_1) \geq 0 \), and we have proved Eq. (88) for \( x < q \).

Now consider the case \( x > \max\{q, \theta_I\} \). We know that \( B(1, x, V^I_2) = 0 \). Therefore, to show \( B(p, x, V^I_2) \leq 0 \) is equivalent to proving \( B(1, x, V^I_2) - B(p, x, V^I_2) \geq 0 \). Rearranging the terms, we get
\[
B(1, x, V^I_2) - B(p, x, V^I_2) = (1 - p) \left( (\mu_R - \mu_S)V^I_2(x) + \frac{(1 + p)\sigma_R^2 V^I_2''(x) + \mu_R - \mu_S}{2} \right) = (1 - p)\Lambda(p),
\]
where
\[
\Lambda(p) \triangleq \frac{\mu_R \sqrt{\pi}}{\sigma_R} e^{(x-\mu_R)^2/\sigma_R} \text{erfc} \left( \frac{x-\mu_R}{\sigma_R} \right) \left( -p\mu_R - \mu_S + (p + 1)x - p\mu_R - \mu_S \right).
\]

By an application of Corollary B.8, we obtain
\[
\Lambda'(p) = \frac{\mu_R \sqrt{\pi}}{\sigma_R} e^{(x-\mu_R)^2/\sigma_R} \text{erfc} \left( \frac{x-\mu_R}{\sigma_R} \right) (x - \mu_R) - \mu_R \leq 0.
\]
Therefore, since $p \in [0, 1]$, following Eq. (90) we have

$$B(1, x, V^I_c) - B(p, x, V^I_c) = (1 - p) \Lambda(p)$$

$$\geq (1 - p) \Lambda(1)$$

$$= (1 - p) \left( \frac{\mu R \sqrt{\pi}}{\sigma_R} e^{\frac{(x - \mu R)^2}{\sigma_R^2}} \text{erf} \left( \frac{x - \mu R}{\sigma_R} \right) (2x - \mu_R - \mu_S) - \mu_R - \mu_S \right)$$

$$\geq 0,$$

where the last step follows from the definition of $\theta_I$ (see Lemma 5) and Corollary B.7.

Now we are only left with $x \in [q, \theta_I]$. Similar to the argument above, we want to show

$$B(\lambda^*(x), x, V^I_c) - B(p, x, V^I_c) \geq 0.$$  

Rearranging the terms, we get

$$B(\lambda^*(x), x, V^I_c) - B(p, x, V^I_c)$$

$$= (\lambda^*(x) - p) \left( (\mu_R - \mu_S) G(x) + \frac{\lambda^*(x) + p}{2} \sigma_R^2 G'(x) + \mu_R - \mu_S \right)$$

$$= -\frac{(\lambda^*(x) - p)^2 \sigma_R^2}{2} G'(x) \geq 0.$$

The second step follows from the definition of $\lambda^*(x)$ (see Lemma 6). The last step follows from the fact that $G(\cdot)$ is nonincreasing (see Lemma 6). Finally, it remains to be shown that Condition 6 holds. Condition 6 holds straightforwardly by an application\(^{22}\) of Lemma 4.

\(\square\)

**Proof of Theorem 4.** The monotonicity of $\lambda^*(x)$ on $[q, \theta_I]$ and the boundary values follow from Lemma 6. It remains to show the monotonicity of $\theta_I$. Let

$$F(\mu_S, \mu_R, \sigma_R, x) \triangleq \frac{2x - \mu S - \mu R}{2 \sigma_R} e^{\frac{(x - \mu R)^2}{\sigma_R^2}} \left( 1 - \text{erf} \left( \frac{x - \mu R}{\sigma_R} \right) \right) - \mu S + \mu R \frac{2 \sigma_R^2}{\sqrt{\pi}}$$

$$= (y + \frac{a}{2}) e^{\sigma^2} (1 - \text{erf}(y)) - \frac{b + 1}{2 \sqrt{\pi}},$$

where

$$y \triangleq \frac{x - \mu R}{\sigma_R}, \quad a \triangleq \frac{\mu_R - \mu_S}{\sigma_R}, \quad b \triangleq \frac{\mu_S}{\mu_R}.$$

By Lemma 5, $\theta_I$ is the only root of $F(\mu_S, \mu_R, \sigma_R, \cdot)$ on $(\frac{\mu S + \mu R}{2}, \infty)$. From Lemma B.6, we obtain that $F(\mu_S, \mu_R, \sigma_R, \cdot) < 0$ for all $x \in (\frac{\mu S + \mu R}{2}, \theta_I)$, and $F(\mu_S, \mu_R, \sigma_R, \cdot) > 0$ for all $x \in (\theta_I, \infty)$. Hence to prove $\theta_I$ is increasing in $\mu_S$ and $\sigma_R$, it suffices to show $F(\cdot, \mu_R, \sigma_R, x)$ is decreasing and $F(\mu_S, \mu_R, \cdot, x)$ is decreasing. Equivalently, we want to show $\frac{\partial F}{\partial \mu_S}(\mu_S, \mu_R, \sigma_R, x) < 0$, and $\frac{\partial F}{\partial \sigma_R}(\mu_S, \mu_R, \sigma_R, x) < 0$

---

\(^{22}\)Lemma 4 applies here as long as the customer lifetime is finite in expectation. Recall the proof of Proposition 1 for finite lifetime. Consider the happiness process under the policy stated in Theorem 3. If the initial happiness level is below $q$, then we are done since happiness process spends positive measure of time below $q$. On the other hand, if the initial happiness level is above $q$, then the first passage time to $q$ must have finite expectation since both $\mu$ and $\sigma$ terms are Lipschitz and bounded below, and hence we are done.\]
for any $\mu_R > \mu_S > 0$ and $x > \frac{\mu_S + \mu_R}{2}$. Suppose now we fix such $\mu_R$, $\mu_S$, $\sigma_R$ and $x$. Then we have $y > \frac{a}{2}$. Compute the partial derivatives:

\[
\begin{align*}
\frac{\partial F}{\partial \mu_S} (\mu_S, \mu_R, \sigma_R, x) &= - \frac{1}{2\sigma_R} e^{y^2} \text{erfc}(y) - \frac{1}{2\mu_R} \frac{1}{\sqrt{\pi}}; \\
\frac{\partial F}{\partial \sigma_R} (\mu_S, \mu_R, \sigma_R, x) &= - \frac{1}{2\sigma_R} \left[ e^{y^2} \text{erfc}(y)(2y^2 + 1) - \frac{2y}{\sqrt{\pi}} \right] (a + 2y).
\end{align*}
\]

The first result needed, $\frac{\partial F}{\partial \mu_S} (\mu_S, \mu_R, \sigma_R, x) < 0$, follows from the fact that $\text{erfc}(\cdot) > 0$ everywhere. The second, $\frac{\partial F}{\partial \sigma_R} (\mu_S, \mu_R, \sigma_R, x) \leq 0$, follows from the fact that $x > \frac{\mu_S + \mu_R}{2}$ (so that $a + 2y > 0$) and using a Chernoff-type bound of the error function (Chang et al. [8]), $\text{erfc}(y) \geq \frac{2y}{\sqrt{\pi(2y^2+1)}} e^{-y^2}$. This completes the proof. 

E Appendix to Section 6 (Geometric Brownian Motion)

In this appendix, we give the proof of Proposition 4, explain how we numerically solve for the optimal value function, and finish by presenting our numerical findings.

Before we proceed, we note that the customer’s survival probability $\tilde{S}_t$ at time $t$ is given by

\[
\tilde{S}_t \triangleq P\left(\tilde{T} > t \mid \tilde{F}_t\right) = e^{-\int_0^t \tilde{Q}(\tilde{H}_s)ds}.
\]

(91)

Proof of Proposition 4. In order to prove Proposition 4, we first establish the following lemma:

Lemma E.1. For any admissible policy $\pi \in \Pi$, any starting happiness $x \in \mathbb{R}$ and any $t > 0$, the following holds:

\[
\mathbb{E} \left[ \int_0^t \left\{ \tilde{Y}_{s}^{x,\pi} \right\}^2 ds \right] < \infty,
\]

where $\tilde{Y}_s$ is as defined in Eq. (24). In other words, the process $\tilde{Y}_{s}^{x,\pi} \in L^2[0,t]$, and the stochastic integral $\int_0^t \tilde{Y}_{s}^{x,\pi} dB_s$ is a martingale for any $t > 0$, which then implies that

\[
\mathbb{E} \left[ \int_0^t \tilde{Y}_{s}^{x,\pi} dB_s \right] = 0.
\]

Proof. Fix an admissible policy $\pi \in \Pi$, a starting happiness $x \in \mathbb{R}$, and a time $t > 0$. Denote $u_t$ the corresponding action process. Since the solution to Eq. (24) is

\[
\tilde{Y}_t = \exp \left( \int_0^t (\mu_{u_s} - \sigma_{u_s}^2/2) ds + \int_0^t \sigma_{u_s} dB_s \right),
\]

43
we have
\[
\mathbb{E} \left[ \int_0^t \{ \tilde{Y}^x, \pi \}^2 \, ds \right] = \int_0^t \mathbb{E} \left[ \{ \tilde{Y}^x, \pi \}^2 \right] \, ds
\]
\[
= \int_0^t \mathbb{E} \left[ e^{\int_0^t (2\mu u_z - \sigma^2 u_z) \, dz + \int_0^t 2\sigma u_z dB_s} \right] \, ds
\]
\[
= \int_0^t \mathbb{E} \left[ e^{\int_0^t (2\mu u_z + \sigma^2 u_z) \, dz} \right] \, ds
\]
\[
\leq \int_0^t e^{(2\mu_R + \sigma_R^2)s} \, ds
\]
\[
= \frac{e^{(2\mu_R + \sigma_R^2)t} - 1}{2\mu_R + \sigma_R^2} < \infty,
\]
where the inequality results from the fact that \( \mu_S < \mu_R \) and \( \sigma_S < \sigma_R \). Hence we have shown that \( \tilde{Y}^{x, \pi} \in \mathcal{L}^2[0, t] \), and that by the Martingale property of stochastic integrals, the process \( \int_0^t \tilde{Y}^{x, \pi} \, dB_s \) is a martingale for any \( t > 0 \).

Now we are ready to prove Proposition 4.

**Proof of Proposition 4.** As in the proof of Proposition 2, we will show that a function \( \tilde{W} \) as described in Proposition 4 is an upper bound for \( \tilde{V}^* \), and that the gap \( \tilde{W} - \tilde{V}^* \) is zero.

To show that a function \( \tilde{W} \) as described in Proposition 4 is an upper bound for \( \tilde{V}^* \), it suffices to show \( \tilde{W}(x) \geq \tilde{V}(x, \pi) \) for \( \forall x \in \mathbb{R} \) and for any admissible policy \( \pi \in \tilde{\Pi} \). Now fix any \( x \in \mathbb{R} \) and any \( \pi \in \tilde{\Pi} \). Let \( u_t \) denote the action process under policy \( \pi \). Define a process \( X_t, t \geq 0 \) by
\[
X_t = \tilde{W}(\tilde{H}_t)\tilde{Y}_t\tilde{S}_t + \int_0^t \tilde{Y}_s\tilde{Q}(\tilde{H}_s)\tilde{S}_s \, ds,
\]
where \( \tilde{H}_t \) is the happiness process under policy \( \pi \) and initial happiness \( x \) (see Eq. (25)), \( \tilde{Y}_t \) the corresponding cumulative reward (conditional on no quitting) up to time \( t \) (see Eq. (24)), and \( \tilde{S}_t \) is the corresponding customer survival probability at time \( t \) (see Eq. (91)). Since \( \pi \) is admissible, the corresponding process \( \tilde{H}_t \) is a semimartingale (and hence \( \tilde{Y}_t \) and \( \tilde{S}_t \) are also semimartingales). Next we follow a similar process as in Section 6 to expand \( X_t \) in integral form.

Since \( \tilde{W} \) is continuously differentiable everywhere on \( \mathbb{R} \) and twice continuously differentiable everywhere on \( \mathbb{R} \setminus \mathcal{E} \) for some countable set \( \mathcal{E} \) (Condition 2 in Proposition 4), we can apply the Itô-Tanaka formula to conclude that \( \tilde{W}(\tilde{H}_t) \) is also a semimartingale:
\[
\tilde{W}(\tilde{H}_t) = \tilde{W}(x) + \int_0^t \tilde{W}'(\tilde{H}_s)(\mu u_x - \tilde{H}_s) \, ds + \int_0^t \tilde{W}''(\tilde{H}_s)\sigma u_x \, dB_s
\]
\[
+ \frac{1}{2} \int_0^t \{ \tilde{H}_s \notin \mathcal{E} \} \tilde{W}'''(\tilde{H}_s)\sigma^2 u_x \, ds + \frac{1}{2} \sum_{y \in \mathcal{E}} (\tilde{W}'_r(y) - \tilde{W}'_l(y)) L^\tilde{H}(t, y),
\]
where \( \tilde{W}'_r \) and \( \tilde{W}'_l \) are the right and left derivatives of \( \tilde{W} \), and \( L^\tilde{H}(t, y) \) is the symmetric local time of \( \tilde{H}_t \) at \( y \). In fact, we can still apply the results of Lemma A.1 to the GBM setting, since the
The inequality in Eq. (95) results from applying Condition 5 of Proposition 4 and the fact that the last term involving the local time in Eq. 93 is zero almost surely. Since \( \tilde{W}(\tilde{H}_t), \tilde{Y}_t \) and \( \tilde{S}_t \) are all semimartingales, we can then apply the multi-dimensional Itô’s formula on semimartingales to \( g(\tilde{W}(\tilde{H}_t), \tilde{Y}_t, \tilde{S}_t) = \tilde{W}(\tilde{H}_t)\tilde{Y}_t\tilde{S}_t \) and rewrite \( X_t \) as:

\[
X_t = \tilde{W}(x) + \int_0^t \tilde{Y}_s\tilde{S}_s\tilde{W}'(\tilde{H}_s) \left( \mu_{u_s} - \tilde{H}_s \right) ds + \int_0^t \tilde{Y}_s\tilde{S}_s\tilde{W}'(\tilde{H}_s)\sigma_{u_s} dB_s \\
- \int_0^t \tilde{Y}_s\tilde{S}_s\tilde{W}(\tilde{H}_s)\tilde{Q}(\tilde{H}_s) ds + \frac{1}{2} \int_0^t \tilde{Y}_s\tilde{S}_s\tilde{W}''(\tilde{H}_s)\mathbb{1}\{\tilde{H}_s \notin \mathcal{E}\} \sigma_{u_s}^2 ds \\
+ \int_0^t \tilde{Y}_s\tilde{S}_s\tilde{W}(\tilde{H}_s)\mu_{u_s} ds + \int_0^t \tilde{Y}_s\tilde{S}_s\tilde{W}(\tilde{H}_s)\sigma_{u_s} dB_s \\
+ \int_0^t \tilde{Y}_s\tilde{S}_s\tilde{W}'(\tilde{H}_s)\sigma_{u_s}^2 ds + \int_0^t \tilde{Y}_s\tilde{S}_s\tilde{Q}(\tilde{H}_s) ds. \tag{94}
\]

Note that by the Martingale property of stochastic integrals, the two stochastic integrals above have zero expectations if \( \tilde{Y}_s\tilde{S}_s\tilde{W}'(\tilde{H}_s)\sigma_{u_s} \) and \( \tilde{Y}_s\tilde{S}_s\tilde{W}(\tilde{H}_s)\sigma_{u_s} \) are in the \( L^2[0, t] \) space. This is indeed true since \( \tilde{Y}_s^{x\pi} \in L^2[0, t] \) (see Lemma E.1), \( \tilde{S}_s \in [0, 1] \), \( \tilde{W} \) is bounded (see Condition 3 in Proposition 4), \( \tilde{W}' \) is bounded (see Condition 4 in the proposition), and \( \sigma_{u_s} \in \{0, \sigma_R\} \). Hence we can take expectation on both sides of Eq. (94) and remove the two stochastic integrals, while replacing 1 with \( 1\{\tilde{H}_s \notin \mathcal{E}\} + 1\{\tilde{H}_s \in \mathcal{E} \& u_s = S\} + 1\{\tilde{H}_s \in \mathcal{E} \& u_s = R\} \), and get

\[
\mathbb{E}X_t = \tilde{W}(x) + \mathbb{E}\int_0^t \tilde{Y}_s\tilde{S}_s \left[ \left( \tilde{Q}(\tilde{H}_s) + \left( \mu_{u_s} - \tilde{Q}(\tilde{H}_s) \right) \tilde{W}(\tilde{H}_s) + \left( \mu_{u_s} + \sigma_{u_s}^2 - \tilde{H}_s \right)\tilde{W}'(\tilde{H}_s) \right) \right. \\
+ \frac{1}{2} \sigma_{u_s}^2 \tilde{W}''(\tilde{H}_s) \mathbb{1}\{\tilde{H}_s \notin \mathcal{E}\} \\
+ \left( \tilde{Q}(\tilde{H}_s) + \left( \mu_{u_s} - \tilde{Q}(\tilde{H}_s) \right) \tilde{W}(\tilde{H}_s) + \left( \mu_{u_s} + \sigma_{u_s}^2 - \tilde{H}_s \right)\tilde{W}'(\tilde{H}_s) \right) \mathbb{1}\{\tilde{H}_s \in \mathcal{E} \& u_s = S\} \\
+ \left( \tilde{Q}(\tilde{H}_s) + \left( \mu_{u_s} - \tilde{Q}(\tilde{H}_s) \right) \tilde{W}(\tilde{H}_s) + \left( \mu_{u_s} + \sigma_{u_s}^2 - \tilde{H}_s \right)\tilde{W}'(\tilde{H}_s) \right) \mathbb{1}\{\tilde{H}_s \in \mathcal{E} \& u_s = R\} \right] ds \\
\leq \tilde{W}(x) + \mathbb{E}\int_0^t \tilde{Y}_s\tilde{S}_s \left[ \tilde{Q}(\tilde{H}_s) + \left( \mu_{u_s} - \tilde{Q}(\tilde{H}_s) \right) \tilde{W}(\tilde{H}_s) + \left( \mu_{u_s} + \sigma_{u_s}^2 - \tilde{H}_s \right)\tilde{W}'(\tilde{H}_s) \right. \\
+ \left( \tilde{Q}(\tilde{H}_s) + \left( \mu_{u_s} - \tilde{Q}(\tilde{H}_s) \right) \tilde{W}(\tilde{H}_s) + \left( \mu_{u_s} + \sigma_{u_s}^2 - \tilde{H}_s \right)\tilde{W}'(\tilde{H}_s) \right) \mathbb{1}\{\tilde{H}_s \in \mathcal{E} \& u_s = R\} ds \\
= \tilde{W}(x). \tag{95}
\]

The inequality in Eq. (95) results from applying Condition 5 of Proposition 4 and the fact that
\( \hat{Y}_s \hat{S}_s > 0 \). The last step in Eq. (95) follows from applying the Cauchy-Schwarz inequality:

\[
\left| \mathbb{E} \int_0^t \hat{Y}_s \hat{S}_s \left( \hat{Q}(\hat{H}_s) + \left( \mu_{u_s} - \hat{Q}(\hat{H}_s) \right) \hat{W}(\hat{H}_s) + (\mu_{u_s} + \sigma_{u_s}^2 - \hat{H}_s)\hat{W}'(\hat{H}_s) \right) \mathbb{1} \{ \hat{H}_s \in E \text{ & } u_s = R \} ds \right|
\leq \sqrt{ \left( \mathbb{E} \int_0^t \hat{Y}_s^2 \hat{S}_s^2 ds \right) \cdot \mathbb{E} \int_0^t \left( \hat{Q}(\hat{H}_s) + \left( \mu_{u_s} - \hat{Q}(\hat{H}_s) \right) \hat{W}(\hat{H}_s) + (\mu_{u_s} + \sigma_{u_s}^2 - \hat{H}_s)\hat{W}'(\hat{H}_s) \right)^2 \mathbb{1} \{ \hat{H}_s \in E \text{ & } u_s = R \} ds }
\]

\[= 0. \]

In the last step above, the first squared root term is bounded since \( \hat{Y}_s \in L^2[0, t] \) by Lemma E.1 and \( \hat{S}_s \in [0, 1] \). The second squared root term above is zero, since \( \int_0^t \mathbb{1} \{ \hat{H}_s \in E \text{ & } u_s = R \} ds = 0 \) almost surely by Lemma A.1, and \( \hat{Q}(\hat{H}_s) + \left( \mu_{u_s} - \hat{Q}(\hat{H}_s) \right) \hat{W}(\hat{H}_s) + (\mu_{u_s} + \sigma_{u_s}^2 - \hat{H}_s)\hat{W}'(\hat{H}_s) \) is bounded for \( \hat{H}_s \) in a countable set \( E \). Since Inequality (95) holds for any \( t \geq 0 \), it also holds in the limit:

\[
\lim_{t \to \infty} \mathbb{E} X_t \leq \hat{W}(x). \quad (96)
\]

We will now show \( \lim_{t \to \infty} \mathbb{E} X_t \geq \hat{V}(x, \pi) \) to complete the proof of \( \hat{W}(x) \geq \hat{V}(x, \pi) \). Observe that since \( \hat{Y}_t > 0, \hat{S}_t > 0 \) for any \( t > 0 \) and \( \mu_R > 0, \mu_S > 0 \), the integral \( \int_0^t \hat{Y}_s \hat{S}_s \mu_{u_s} ds \) is pathwise monotone increasing in \( t \) and hence converges pathwise to \( \int_0^\infty \hat{Y}_s \hat{S}_s \mu_{u_s} ds \) as \( t \to \infty \). Therefore

\[
\lim_{t \to \infty} \mathbb{E} \left[ \int_0^t \hat{Y}_s \hat{S}_s \mu_{u_s} ds \right] = \mathbb{E} \left[ \int_0^\infty \hat{Y}_s \hat{S}_s \mu_{u_s} ds \right].
\]

Define \( \hat{V}_t(x, \pi) = \hat{Y}_0 + \mathbb{E} \left[ \int_0^t \hat{S}_s d\hat{Y}_s \right] \). It follows that

\[
\hat{V}_t(x, \pi) = \hat{Y}_0 + \mathbb{E} \left[ \int_0^t \hat{S}_s d\hat{Y}_s \right]
\]

\[= 1 + \mathbb{E} \left[ \int_0^t \hat{Y}_s \hat{S}_s \mu_{u_s} ds + \int_0^t \hat{Y}_s \hat{S}_s \sigma_{u_s} dB_s \right]
\]

\[= 1 + \mathbb{E} \int_0^\infty \hat{Y}_s \hat{S}_s \mu_{u_s} ds \]

is monotone increasing in \( t \). Note that the stochastic integral has zero expectation since \( \hat{Y}_s \hat{S}_s \sigma_{u_s} \in L^2[0, t] \), the same reasoning as before. Then by the Monotone Convergence Theorem, we have

\[
\hat{V}(x, \pi) = \hat{V}_\infty(x, \pi)
\]

\[= \lim_{t \to \infty} \hat{V}_t(x, \pi)
\]

\[= 1 + \lim_{t \to \infty} \mathbb{E} \int_0^t \hat{Y}_s \hat{S}_s \mu_{u_s} ds
\]

\[= 1 + \mathbb{E} \int_0^\infty \hat{Y}_s \hat{S}_s \mu_{u_s} ds \] .

(97)

where the first equation follows from Eq. (28) and replacing \( \mathbb{1} \{ t < T \} \) with \( \hat{S}_t \) by the tower property of conditional expectation and the applying the definition of \( \hat{S}_t \) in Eq. (91). With Eq. 97, we are ready to take expectations on both sides of Eq. (92) and let \( t \to \infty \). Apply Itô's formula on \( \hat{Y}_t \hat{S}_t \),
The inequality results from the fact that $\tilde{W}(\cdot) > 1$ on $\mathbb{R}$ (see Condition 1 of Proposition 4) and that $\tilde{Y}_t, \tilde{S}_t > 0$. Combining Eqs. (96) and (98), we obtain the desired result $\tilde{W}(x) \geq \tilde{V}(x, \pi)$ for $\forall x \in \mathbb{R}$ and any admissible policy $\pi$. Hence $\tilde{W}$ as described in Proposition 4 is an upper bound for the optimal CLV $\tilde{V}^*$.

Now it remains to show that the gap $\tilde{W} - \tilde{V}^*$ is zero when the policy is chosen to be $\tilde{\pi}$, as described in Proposition 4. Observe in the above analysis that it suffices to show inequalities (95) and (98) are tight. Inequality (95) (see Eqs. (24) and (91)), this is true if we can show $\limsup_{t \to \infty} \mathbb{E} \left[ \tilde{Y}_t \tilde{S}_t \right] = 0$. From definitions of $\tilde{Y}_t$ and $\tilde{S}_t$ (see Eqs. (24) and (91)), we know that (since $\tilde{\pi}$ is an interval policy by Condition 6 of Proposition 4, we use $\tilde{\pi}(\tilde{H})$ to denote the policy’s choice of service mode at time $t$)

$$
\mathbb{E} \left[ \tilde{Y}_t \tilde{S}_t \right] = \mathbb{E} \left[ e^{\int_0^t \left( \mu_{\tilde{H}_z} - \frac{\sigma^2_{\tilde{H}_z}}{2} - \tilde{Q}(\tilde{H}_z) \right) dz + \int_0^t \sigma_{\tilde{H}_z} dW_z} \right] = \mathbb{E} \left[ e^{\int_0^t \left( \mu_{\tilde{H}_z} - \tilde{Q}(\tilde{H}_z) \right) dz} \right].
$$

Since $Q_1 > Q_2 > \max\{\mu_S, \mu_R\}$ (see Condition 1), the exponent $\int_0^t \left( \mu_{\tilde{H}_z} - \tilde{Q}(\tilde{H}_z) \right) dz$ is bounded above by $(\max\{\mu_S, \mu_R\} - Q_2)t$. Hence

$$
0 \leq \limsup_{t \to \infty} \mathbb{E} \left[ e^{\int_0^t \left( \mu_{\tilde{H}_z} - \tilde{Q}(\tilde{H}_z) \right) dz} \right] \leq \limsup_{t \to \infty} \mathbb{E} \left[ e^{\int_0^t \left( \max\{\mu_S, \mu_R\} - Q_2 \right) dz} \right] = 0.
$$

Thus we have proved the desired result.
How we numerically solve for the optimal policy and value function. We need to numerically find the values of $C_1$, $C_2$ and $C_3$ of the following functions (that solves Eq. (30) for the Risky service mode and for the Safe mode, respectively,)

$$W(x, C_1, C_2, R) = \frac{\dot{Q}(x)}{\dot{Q}(x) - \mu_R} + C_1 H \left( \mu_R - \dot{Q}(x), \frac{x - \mu_R - \sigma^2_R}{\sigma_R} \right) + C_2 M \left( \frac{\dot{Q}(x) - \mu_R}{2}, \frac{1}{2}, \frac{(x - \mu_R - \sigma^2_R)^2}{\sigma^2_R} \right)$$

and

$$W(x, C_3, S) = \frac{\dot{Q}(x)}{\dot{Q}(x) - \mu_S} + C_3 (\mu_S - x)^{\mu_S - Q(x)}$$

in different happiness regions where the firm chooses different service modes, such that the optimality conditions in Proposition 4 are satisfied.

In the above expressions, $H(\cdot, \cdot)$ is a Hermite Polynomial, and $M(\cdot, \cdot, \cdot)$ is the Kummer confluent hypergeometric function: the two functions $H(\lambda, x)$ and $M(\frac{1}{2}, \frac{1}{2}, x^2)$ are the two linearly independent solutions to the Hermite Differential Equation $y''(x) - 2xy'(x) + 2\lambda y(x) = 0$. One challenge of calculating this $W$ is that, as the happiness value $x$ decreases, both $H()$ and $M()$ grow exponentially. Since both functions cannot be evaluated to their exact values, the error in calculation $W$ (which is the difference of two very large numbers) can get large for negative $x$ with large magnitude. To control for this error, we place a reflecting boundary at $q - B$ in the unsatisfied zone for some large $B > 0$ and let the happiness process only evolve on $[q - B, \infty)$. Recall that the happiness process is an Ornstein-Uhlenbeck (O-U) process in the unsatisfied zone, if the firm always utilizes the Risky service mode there. With a reflecting boundary $q - B < q$, it becomes a reflected O-U process. To preserve our insights, we want to choose $B$ large enough that the customer churns before hitting the reflecting boundary $q - B$ with probability close to 1.

We use the following method to choose the reflecting boundary $q - B$. Consider a reflected O-U process $\tilde{X}_t$ on $(-\infty, q]$ with infinitesimal drift $\mu_R - \tilde{X}_t$, infinitesimal volatility $\sigma_R$, initial value $\tilde{X}_0 = q$, and reflecting boundary at the happiness threshold $q$. Note that this is an approximation of the happiness process $\tilde{H}_t$ under the sandwich policy with the risk-averse region right above the unsatisfied zone, which we conjecture to be optimal, if not the myopic policy. Under this policy, the happiness process becomes a delayed reflected O-U process on $(-\infty, q]$ once it hits the unsatisfied zone. Notice the difference between process $\tilde{X}_t$ and the happiness process just stated — $\tilde{X}_t$ has instantaneous reflection at $q$ while $\tilde{H}_t$ has delayed reflection at $q$. Nevertheless, this means that the stationary probability of $\tilde{X}_\infty < q - B$ is an overestimation of the stationary probability of $\tilde{H}_\infty < q - B$, and we can bound the latter by bounding $\Pr\{\tilde{X}_\infty < q - B\}$. From Ward and Glynn’s paper [34], we know the stationary distribution of $\tilde{X}_t < q - B$ is $\Pr \left[ N \left( \mu_R, \frac{\sigma^2_R}{2} \right) < q - B \right] N \left( \mu_R, \frac{\sigma^2_R}{2} \right) \leq q$. Hence we would like to choose $B$ such that this probability is small.
Numerical findings. Next we give details of how random instances are generated to compute the optimal policies. In each iteration, the parameters are randomly generated in the following sequence:

1. randomly generate $\sigma_R \sim \text{Uniform}[0, 2, 1]$;
2. randomly generate $\mu_R \sim \text{Uniform}[0, 0.8]$;
3. randomly generate $\mu_S \sim \text{Uniform}[0, \mu_R]$;
4. randomly generate $q = \mu_S + \text{Uniform}[0, 0.8]$;
5. assign $B = q - \mu_R + 5\sigma_R$;
6. randomly generate $Q_1 \sim \text{Uniform}[\mu_R, \mu_R + 5]$;
7. randomly generate $Q_2 \sim \text{Uniform}[\mu_R, Q_1]$.

Note that the range of the specifications of $\mu_S$, $\mu_R$ and $\sigma_R$ are chosen to have a similar magnitude with the GBM drift and volatility estimations from the financial market (for example, see Schneider et al. [29]). Also note that the choice of $B$ is made to ensure that the second argument $\frac{x - \mu_R - \sigma_R^2}{\sigma_R}$ inside the Hermite Polynomials is always bounded below on $[q - B, \infty)$, in order to ensure numerical stability. We numerically solve 1000 random instances generated by the above procedure, by solving for the free parameters $C_1$ and $C_2$ in Eq. (32) and verify that all the conditions in Proposition 4 are satisfied. In each instance, the stationary probability of a reflected O-U process $\tilde{H}_t < q - B$ is calculated. In fact, they are all less than $10^{-8}$. Therefore we can say that with very high probability, the customer churns before hitting the boundary $q - B$, and that placing a boundary at $q - B$ will very likely not affect the firm’s optimal policy. In fact, we also check this by perturbing the choice of $B$ within $[q - \mu_R + 4\sigma_R, q - \mu_R + 5\sigma_R]$ and showing that both the value function and the sandwich structure are extremely insensitive to the choice of $B$. Remarkably, in all the randomly generated instances, the optimal policy is either a myopic policy (Risky mode everywhere) or a sandwich policy (Safe mode only in an interval just above $q$).

F Appendix to Section 7: HJB Equation with Switching Costs

In this appendix, we go over the heuristic steps to obtain the HJB equation.

Consider a starting happiness value at $H_0 = x$ and the firm’s current service mode being the Risky mode. We answer the following question: should the firm stick with the Risky mode for a very short time $t$ and then continue optimally, or should the firm immediately switch to the Safe mode but incur a switching cost of $K > 0$? In the first option, the total reward collected is

$$\int_0^t 1\{T > s\}(\mu_R ds + \sigma_R dB_s) + 1\{T > t\}V^*_R(H_t),$$
where \( V_i^*(x) \) is the continuation value when happiness level starts at \( x \) and the firm’s starting service mode is \( i \in \{S, R\} \). Take expectation of the above expression and apply Itō’s formula\(^{23} \) on \( e^{-\int_0^t Q(H_s)ds}V^*_R(H_t) \), we get

\[
E\int_0^t \mathbb{1}\{T > s\}(\mu_R ds + \sigma_R dB_s) + E[\mathbb{1}\{T > t\}V^*_R(H_t)] = E\int_0^t e^{-\int_0^s Q(H_v)dv} \mu_R ds + E\int_0^t e^{-\int_0^s Q(H_v)dv}V^*_R(H_t)
\]

\[
= E\int_0^t e^{-\int_0^s Q(H_v)dv} \mu_R ds + E\left[V^*_R(H_0) + \int_0^t e^{-\int_0^s Q(H_v)dv}(\mu_R - H_s)V^*_R(H_s)ds + \int_0^t e^{-\int_0^s Q(H_v)dv} \frac{\sigma^2}{2} V^**_R(H_s)ds - \int_0^t Q(H_s) e^{-\int_0^s Q(H_v)dv} V^*_R(H_s)ds\right].
\]

On the other hand, if the firm immediately switches to the Safe mode by incurring a cost of \( K \), the total reward (minus cost) collected is \( V^*_S(H_0) - K \). Since the firm wants to maximize total reward, we must have

\[
V^*_R(H_0) = \max \left\{ E\int_0^t e^{-\int_0^s Q(H_v)dv} \mu_R ds + E\left[V^*_R(H_0) + \int_0^t e^{-\int_0^s Q(H_v)dv}(\mu_R - H_s)V^*_R(H_s)ds + \int_0^t e^{-\int_0^s Q(H_v)dv} \frac{\sigma^2}{2} V^**_R(H_s)ds - \int_0^t Q(H_s) e^{-\int_0^s Q(H_v)dv} V^*_R(H_s)ds\right], V^*_S(H_0) - K \right\}.
\]

Consider the limit as \( t \to 0 \), the above equation reduces to

\[
0 = \max \left\{ -Q(x)V^*_R(x) + (\mu_R - x)V^*_{R'}(x) + \frac{\sigma^2}{2} V^**_R(x) + \mu_R, V^*_S(x) - V^*_R(x) - K \right\}.
\]

Similarly, if the firm’s current service mode is the Safe mode and the customer’s happiness value is \( H_0 = x \), we have

\[
0 = \max \left\{ -Q(x)V^*_S(x) + (\mu_S - x)V^*_{S'}(x) + \mu_S, V^*_R(x) - V^*_S(x) - K \right\}.
\]

\(^{23}\)Fix \( i \in \{S, R\} \). Since \( H_t \) is a semimartingale, by the Itō-Tanaka formula, if \( V_i^*(\cdot) \) is sufficiently smooth, then \( V_i^*(H_t) \) is also a semimartingale. Also \( e^{-\int_0^t Q(H_v)dv} \) is a semimartingale. Therefore by the multidimensional Itō formula, \( e^{-\int_0^t Q(H_v)dv}V^*_R(H_t) \) is also a semimartingale.