Consumption-investment decisions with endogenous reference point and drawdown constraint

Zongxia Liang^{a*} Xiaodong Luo^{a†} Fengyi Yuan^{a‡}

^{a}*Department of Mathematical Sciences, Tsinghua University, China*

November 23, 2022

Abstract

We propose a consumption-investment decision model where past consumption peak h plays a crucial role. There are two important consumption levels: the lowest constrained level and a reference level, at which the risk aversion in terms of consumption rate is changed. We solve this stochastic control problem and derive the value function, optimal consumption plan, and optimal investment strategy in semi-explicit forms. We find five important thresholds of wealth, all as functions of h, and most of them are nonlinear functions. As can be seen from numerical results and theoretical analysis, this intuitive and simple model has significant economic implications, and there are at least three important predictions: the marginal propensity to consume out of wealth is generally decreasing but can be increasing for intermediate wealth levels, and it jumps inversely proportional to the risk aversion at the reference point; the implied relative risk aversion is roughly a smile in wealth; the welfare of the poor is more vulnerable to wealth shocks than the wealthy. Moreover, locally changing the risk aversion influences the optimal strategies globally, revealing some risk allocation behaviors.

2010 **Mathematics Subject Classification:** 91B08, 91B42, 91G10, 93E20.

Keywords: Consumer behavior; Past consumption peak; Drawdown constraint; Endogenous reference point; Stochastic control.

1   Introduction

It is intuitive that the historical peak of past consumption has great impact on individual consumption decision. For example, to consume below a certain ratio of the historical peak brings an impulsion to "reclaim the past glory". If the consumption is forced to further decline to a level that is far below the historical peak, it becomes unbearable and people will try to increase their wealth (by financing, borrowing or selling illiquid assets) at any cost to satisfy consumption at a certain (but low) ratio of past peak. Therefore it is not surprising that there has been literature studying so-called drawdown constraint (Dybvig (1995) and Arun (2012)).

---

*Email: `liangzongxia@mail.tsinghua.edu.cn`

†Corresponding author, email: `luoxd21@mails.tsinghua.edu.cn`

‡Email: `yfy19@mails.tsinghua.edu.cn`

We develop a theoretical model based on dynamic portfolio choice theory to incorporate aforementioned psychological insights and study the effects of them on consumption and risky investment decisions. In our model, the process of *standard of living* represented by running maximum of consumption, $h_t = h_0 \vee \sup_{0 \le s \le t} \{c_s\}$, plays a central role. Here $c \triangleq \{c_s, s \ge 0\}$ is the process of consumption rate, and h_0 is the *inherited* running maximum level, which is usually determined by exogenous factors, say, family fortunes. To be specific, the consumption is *constrained to be no less than* λh_t, and the preference on consumption is assumed to be of the form:

$$U(c, h) = \begin{cases} \frac{1}{\beta_1} \left[1 - e^{-\beta_1(c - \alpha h)}\right], & \lambda h \le c < \alpha h, \\ \frac{1}{\beta_2} \left[1 - e^{-\beta_2(c - \alpha h)}\right], & \alpha h \le c, \end{cases}$$

where λ and α with $0 < \lambda < \alpha < 1$ are exogenous constants, representing two important thresholds of consumption level, and β_1 and β_2 with $\beta_1, \beta_2 > 0$ are absolute risk aversions. We allow $\beta_1 > \beta_2$, $\beta_1 < \beta_2$ or $\beta_1 = \beta_2$.

In our model, the utility is produced by the difference between the agent's instantaneous consumption and a reference point αh. Besides, we also consider a change of risk aversions on different sides of the reference point. This reflects psychological effects when consumption rises above or falls below the reference point. In economical literature, changes of the individual risk aversions are discussed and empirical evidences have been found. On the one hand, based on both naturally occurring data and lab data, it is acknowledged that people become more risk averse when experiencing crisis or fear, see Cohn et al. (2015), Guiso et al. (2018) and references therein. In this paper, choosing $\beta_1 > \beta_2$, we model such crisis as the decline of consumption level. In this case, people are more risk averse in crisis ($c < \alpha h$). On the other hand, as an important part of the Nobel-wining prospect theory, loss aversion is an effect that people become risk seeking when the pay-offs fall below some certain reference point. When incorporating such extreme gambling behavior into consumption decision, somewhat extreme optimal decisions are derived: people never consume between 0 (or lowest constrained level) and the reference point, see Van Bilsen et al. (2020) and Li et al. (2022) for examples. One possible reason is that consumption falling below certain reference level is not generally treated as loss, but rather as bad luck or temporary crisis. To incorporate gambling effect into consumption decision, we can conveniently choose $\beta_1 < \beta_2$ in our model. In this case, people are willing to take more risk when their consumption is in danger ($c < \alpha h$). Our main interests are investigating consumption and portfolio behaviors under the aforementioned preference change at the reference point. The reference point itself, however, can be further generalized from the particular choice αh. We just list main results of this generalization in Appendix D because solution techniques we use are still applicable.

It turns out that our simple risk-aversion-changing preference leads to consumption and investment decisions with significant economic implications.

We find five important thresholds of wealth, all depending on h and denoted respectively by $W_{\text{bkrp}}(h), W_{\text{low}}(h), W_{\text{ref}}(h), W_{\text{peak}}(h)$ and $W_{\text{updt}}(h)$, which are crucial to describe the derived consumption and investment decisions. For those with $x \ge W_{\text{updt}}(h)$, i.e., the very wealthy ones, the best choice of consumption is to update the consumption peak in order to move to or maintain on the line $x = W_{\text{updt}}(h)$. If $x < W_{\text{bkrp}}(h)$, the wealth of the agent is insufficient for keeping a consumption rate $c \ge \lambda h$ considering his standard of living in the past and leads to bankruptcy, hence is not allowed in our discussion. The above results indicate that only those $x \in [W_{\text{bkrp}}(h), W_{\text{updt}}(h)]$ are of special interests, and they are said to be in *the effective region*. For the wealthiest and poorest people in effective region, the optimal consumption rate is clear: the wealthiest ($W_{\text{peak}}(h) < x \le W_{\text{updt}}(h)$, named as satisfactory region) revisit the historical peak
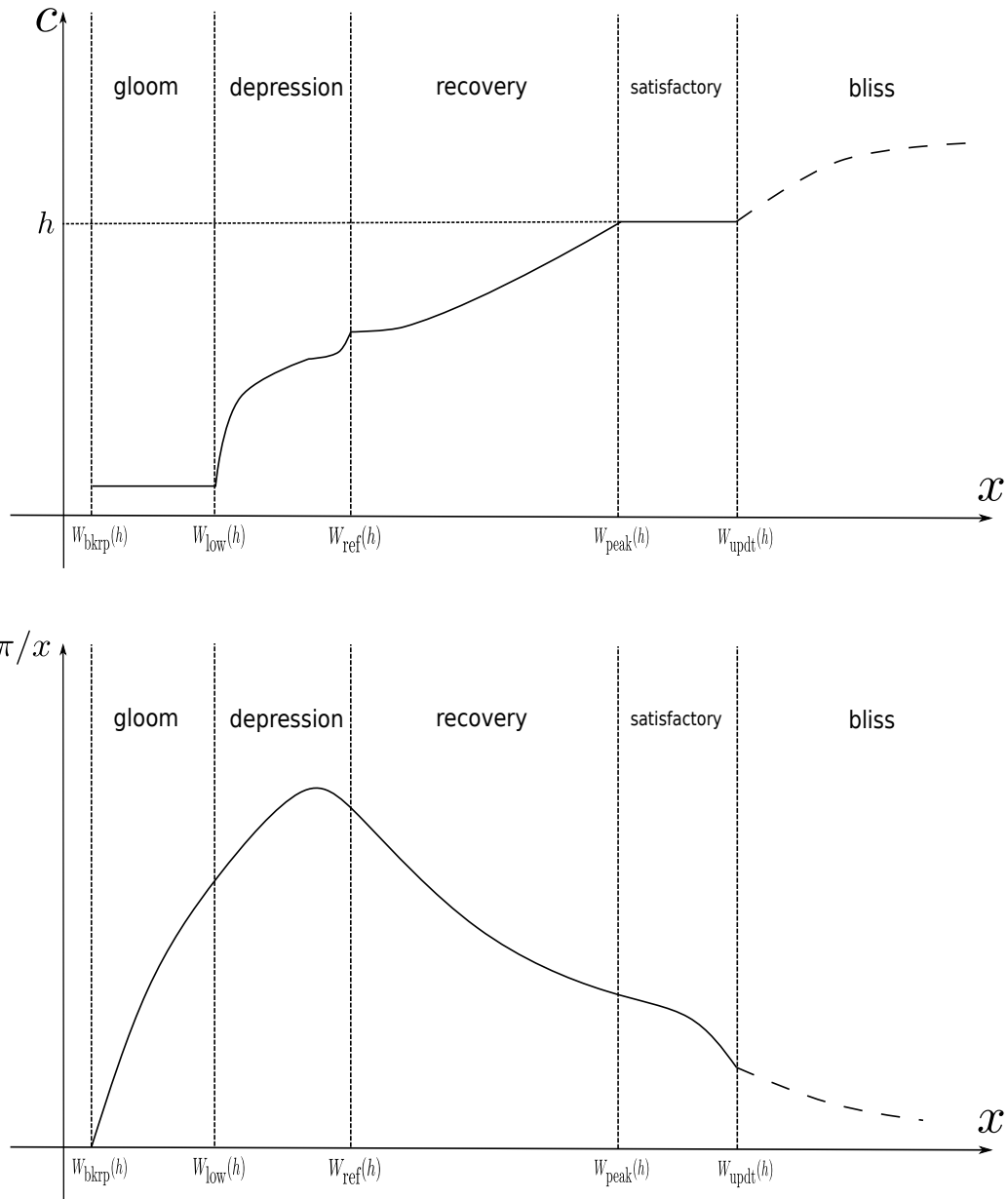
Figure 1: Optimal consumption (above) and optimal risky investment proportion (below) when fixing an h.

$c = h$, and the poorest ($W_{\mathrm{bkrp}}(h) \leq x \leq W_{\mathrm{low}}(h)$, named as gloom region) choose to consume at the lowest level $c = \lambda h$. As for the region of intermediate wealth $W_{\mathrm{low}}(h) < x \leq W_{\mathrm{peak}}(h)$, by the reference threshold $W_{\mathrm{ref}}(h)$, we divide it into two sub-regions: depression and recovery. We document two important phenomena. *First*, the (instantaneous) marginal propensity to consume (MPC) out of wealth $\frac{\partial c^*}{\partial x}$ is decreasing for people in lower part of depression-recovery region (most or all part of depression, sometimes plus part of recovery), which is consistent with empirical studies (e.g. Arrondel et al. (2015)), and we predict that it is increasing for wealthier people, i.e., those in the rest of depression-recovery region. However, if the wealth level is $x = W_{\mathrm{updt}}(h)$, called bliss curve by us, the MPC out of wealth is then decreasing again. *Second*, the MPC out of wealth jumps by a fixed proportion β_1/β_2 at the threshold $W_{\mathrm{ref}}(h)$, at which the risk aversion is changed. This fact reveals one of salient features of our model and indicates that, generally, when getting wealthier, people have lower or higher MPC based on its particular type of risk attitude (gambling type or stop-loss type). See Figure 1 for a graphic illustration of optimal consumption decision.

We also obtain the optimal investment strategy π^*, representing the *amount of wealth* invested in risky assets, and optimal investment proportion π^*/x is obtained as a by-product. Recall that in classical Merton's problem, the optimal investment proportion is $\frac{\mu - r}{\sigma^2 \gamma}$ if relative risk aversion is γ. Therefore, the inverse of π^*/x can be treated equivalently as relative risk aversion, which we call implied relative risk aversion (see also Jeon and Park (2020)). Keeping this in mind, we find decreasing relative risk aversion (DRRA) and increasing relative risk aversion (IRRA) are both possible even for a single agent and the implied relative risk aversion is roughly a U-shaped curve (smile) in variable x with trough around $W_{\mathrm{ref}}(h)$. Equivalently, risky investment proportion is a hump with peak around $W_{\mathrm{ref}}(h)$. We further predict that this effect is more pronounced for poorer people. In economic literature, there has been long standing debate on how relative risk aversion varies in wealth distribution, and evidence for both DRRA and IRRA are found (Siegel and Hoban (1982) and Bellante and Green (2004)). We provide an explanation for this: the RRA can decrease in wealth because there is impulsion to get back to a higher consumption above αh if x is not so large; the RRA can also increase in x if x has been enough for maintaining a satisfactory standard of living. As for the portfolio choice itself, although it is widely admitted that for *macro* data, the wealthier people tends to invest more proportion of their wealth in risky assets, there is no solid agreement on the same question in *micro* aspect. That is to say, what happens to risky investment if the wealth increases for *a given* household? Our model predicts that it is very likely that people proportionally reduce risky investment if their wealth grows, which is consistent with empirical studies (Brunnermeier and Nagel (2008) and Paya and Wang (2016)) or other possible alternative models (Wachter and Yogo (2010)). Our model also allows opposite result, provided that the wealth is not enough, which provides explanations for co-existence of both phenomena in some literature such as Brunnermeier and Nagel (2008). See Figure 1 for a graphic illustration of optimal risky assets allocation.

The literature most closely related to the present paper is Deng et al. (2022). In the aspect of reference point, we adopt the setting of Deng et al. (2022) and rely on their solving techniques as well as other celebrated tools such as dynamic programming, dual transformation and region-wise solving method. Our choice of the reference point is for simplicity. It turns out that the set of solving techniques we have used, inspired by Deng et al. (2022), is widely applicable to other forms of reference point, see Appendix D. In addition to the reference point itself, our main interests include investigating consumption and portfolio behaviors under the preference change at the reference point. There are several distinctive features of our model, serving as complement to the one proposed and solved in Deng et al. (2022). First of all, the preference change leads to an upward or downward turn of optimal consumption at the threshold $W_{\mathrm{ref}}(h)$, instead of a relatively

smooth curve in Deng et al. (2022). Moreover, we emphasize that the influences of preference change on optimal consumption choice are global. For example, adjusting the risk aversion from β_2 to $\beta_1(< \beta_2)$ in the region $x < W_{\mathrm{ref}}(h)$ even changes the consumption behavior in the region $x > W_{\mathrm{ref}}(h)$, revealing a kind of strategical risk allocation among different states of economics. In terms of risky investment proportion, the preference change results in significant increase or decrease with a wide range of wealth level. The preference change together with drawdown constraint leads to a totally different optimal investment proportion curve comparing with Deng et al. (2022). Another advantage of the present model is that considering preference change highlights the importance of the wealth threshold $W_{\mathrm{ref}}(h)$ that has been more or less neglected before. For example, the risky investment proportion attains its maximum around $W_{\mathrm{ref}}(h)$, and the value function has very different sensitivity to wealth shock on different sides of the curve $x = W_{\mathrm{ref}}(h)$. The detailed differences between our paper and Deng et al. (2022), as well as other implications of our model, will become clear in Section 5.

To sum up, the model studied in this paper has considered and combined three aspects of the economical and psychological background of the consumption and investment problem: (1) the running maximum of past consumption is taken as the habit formation process; (2) a drawdown constraint is imposed on consumption; (3) risk aversion is changing in the model and a reference point is added. We have applied the solution method inspired by Deng et al. (2022) as well as other celebrated tools. We have used dynamic programming principle to derive the HJB equation of the problem, solved the optimal consumption and portfolio strategy thanks to the duality method and numerically analyzed the solution and sensitivity of thresholds to certain parameters. From this simple and intuitive model, we have found several interesting economic implications such as MPC jump and RRA smile.

The rest of the paper is organized as follows: Section 2 is devoted to mathematically formulating the optimal consumption and investment problem focused in this paper. We deduce the HJB equation and obtain the feedback form of solution in dual form in Section 3. In Section 4, the verification theorem is established and the optimal strategy is obtained in primal form. Numerical analysis with fixed parameters are in Section 5, while in Section 6 we present sensitivity analysis. Section 7 gives a brief conclusion. Technical proofs and some generalizations are presented in Appendices.

Related literature. Our model is based on dynamic consumption-investment decision model, whose classical form dates back to Merton (1969). A wide range of literature extends this problem by taking habit formation into account, see Pollak (1970), Detemple and Zapatero (1992) and Chapman (1998) for instance. The habit formation is usually modelled through habit formation process and habit formation preference. Habit formation process is a process whose value at time t is determined by the consumption process up to time t. Habit formation preference suggests that the decision maker's utility is depending on both the consumption process and the habit formation process, which sketches how consumption habit affects current consumption behavior.

One may just set the habit formation process as the average of the past consumption process. A more reasonable and flexible choice of the habit formation process is the so-called linear habit formation, i.e., the weighted average of the past consumption where more weight is placed on the recent consumption. Such form of habit formation process has dominated the research in habit formation setting since early literature such as Ryder and Heal (1973). Recently, taking the running maximum process of past consumption as the habit formation process has opened another stream of research in the study of habit formation. The running maximum process is non-decreasing and only updates if the consumption level exceeds the historical running maximum, which brings

about mathematical challenges because of the connection to singular control theory. We follow the running maximum habit formation model recently studied by Guasoni et al. (2020) and Deng et al. (2022). However, our model takes into account more factors that may influence the decision maker's consumption and portfolio choice, including the constraint on consumption and the change of risk aversion.

How utility depends on the consumption c and habit h is another topic in habit formation. The habit formation preference usually refers to the case that the utility is a function of $c - h$, which suggests that the habit has an inverse impact on the utility. A wide range of literature, Chapman (1998) for instance, adopts the above habit formation preference setting. A more flexible model, in Deng et al. (2022), requires the utility to depend on $c - \alpha h$ where α may vary in $[0, 1]$. Guasoni et al. (2020) uses a different approach by setting the utility to be a function of $\frac{c}{h^\alpha}$ where $\alpha \in (0, 1)$. All the above settings insure the inverse impact of the habit. How consumption is constrained, on the other hand, is also important when studying habit-related model. Usually, the constraint imposes a lower bound on consumption. In the extreme case, the consumption is prohibited from falling below the habit (Muraviev (2011) and Yu (2015) for linear habit, Dybvig (1995) and Jeon and Park (2020) for running maximum), which is termed as addictive habit formation. In other literature it is required instead that the consumption is at least a fraction of the habit, such as Arun (2012) and Angoshtari et al. (2022). In this paper, we also impose such drawdown constraint, in which we require $c \geq \lambda h$, where $\lambda \in (0, \alpha)$. We assume $\lambda < \alpha$ to ensure the validity of the reference point αh where risk aversion changes. For $\lambda = 0$, our model reduces to a model with no drawdown constraint; for $\lambda = \alpha$, our model reduces to a model with no risk aversion change. It is worth mentioning that both Arun (2012) and Angoshtari et al. (2022) obtain a threshold of wealth/habit ratio below which the agent chooses to consume the lowest. We derive similar phenomenon, but with a rather complicated threshold curve ($x = W_{\text{low}}(h)$) of wealth-habit pair, instead of a simple ray.

The paradigm of our setting is utility with reference points, which has been widely studied in different economical problems, such as Jin and Zhou (2008) and He and Yang (2019). Another common model under this paradigm is the S-shaped utility developed by D. Kahneman and A. Tversky. Li et al. (2022) has studied such S-shaped utility in the context of running maximum habit formation. However, such non-concave utility results in an extreme optimal strategy where the decision maker never consumes between 0 and the reference point. Instead, we analyze a utility with risk aversion change but in a concave form. As a result, we obtain a more reasonable optimal strategy where the optimal consumption varies from the lower bound λh to the running maximum h in continuous values. Another related work is Van Bilsen et al. (2020), which permits the agent to be risk averse in loss domain (setting $\gamma_L > 1$ therein). This setting is consistent with ours, but they model risk aversion change in terms of relative risk aversions.

2   Model Formulation

The financial market consists of one risk-free asset and one risky asset in our model. The risk-free asset $\{S_t^0, t \geq 0\}$ satisfies the dynamic

$$dS_t^0 = S_t^0 r dt,$$

where $r > 0$ is the constant interest rate. The risky asset $\{S_t^1, t \geq 0\}$ satisfies

$$dS_t^1 = S_t^1 \left[\mu dt + \sigma dB_t \right],$$

where $\mu \geq r$ is the expected return and $\sigma > 0$ is the volatility. The Brownian motion B is defined on the filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$ satisfying the usual conditions.

The decision maker determines his dynamic spending rate $c \triangleq \{c_t, t \geq 0\}$ and the dynamic amount of his wealth invested in the risky asset denoted by $\pi \triangleq \{\pi_t, t \geq 0\}$. Let x_0 be the investor's initial wealth, then his wealth process $\{X_t, t \geq 0\}$ follows the following stochastic differential equation (abbr. SDE):

$$\begin{cases} dX_t = rX_t dt + \pi_t(\mu - r)dt + \pi_t \sigma dB_t - c_t dt, \\ X_0 = x_0. \end{cases} \tag{2.1}$$

Given the consumption process $\{c_t, t \geq 0\}$, the running maximum process of past consumption is defined as $h = \{h_t, t \geq 0\}$:

$$h_t = h_0 \vee \sup_{s \leq t} c_s, \ t > 0 \tag{2.2}$$

and h_0 is the inherited running maximum level. It is naturally required that the wealth level is always non-negative under any admissible strategy to avoid bankruptcy. Thus we now introduce the admissible strategies.

Definition 2.1. Process $(c, \pi) = \{(c_t, \pi_t), \ t \geq 0\}$ is an admissible strategy of Problem (2.3) if it is adapted w.r.t $\{\mathcal{F}_t\}_{t \geq 0}$ with

$$\int_0^t c_s ds < \infty, \ a.s.,$$
$$\int_0^t \pi_s^2 ds < \infty, \ a.s.,$$
$$c_t \geq \lambda h_t, \ a.s.,$$
$$X_t > 0, \ a.s.,$$

for all $t \geq 0$, where $\{h_t, t \geq 0\}$ is the corresponding running maximum process given by (2.2) and $\{X_t, t \geq 0\}$ is the corresponding wealth process governed by SDE (2.1). We denote by $\mathcal{A}$ the set of admissible strategies.

The goal of the decision maker is to maximize his expected total discounted utility on infinite planning horizon $[0, \infty)$, which can be formulated as the optimization problem:

$$\sup_{(c, \pi) \in \mathcal{A}} \mathbb{E}_{x_0, h_0} \int_0^\infty e^{-\gamma t} U(c_t, h_t) dt, \tag{2.3}$$

where $\gamma > 0$ is the discount factor. The utility function $U(c, h)$ depends on both the consumption c and the running maximum h. Precisely,

$$U(c, h) = \begin{cases} \frac{1}{\beta_1}\left[1 - e^{-\beta_1(c - \alpha h)}\right], & \lambda h \leq c < \alpha h, \\ \frac{1}{\beta_2}\left[1 - e^{-\beta_2(c - \alpha h)}\right], & \alpha h \leq c \leq h, \end{cases}$$

where $0 \leq \lambda \leq \alpha < 1$, $\beta_1, \beta_2 > 0$. αh is the reference point where risk aversion increases as the consumption exceeds it. The utility is segmented with both segments taking the form of constant absolute risk aversion (abbr.CARA) utility. The absolute risk aversion above reference αh is β_2, while the absolute risk aversion below reference is β_1. The utility is concave and the marginal utilities at the two sides of the reference point αh are equal and finite.

Remark 1. When $\lambda = 0$, our model reduces to the model without drawdown constraint. When $\alpha = \lambda$ or $\beta_1 = \beta_2$, our model reduces to the model without reference αh and the corresponding risk aversion change. However, our model does not reduces to the model without reference αh if α approaches 1 from below. The reason lies in the fact that the updating of running maximum is under the leveled-up risk aversion coefficient β_2 when $\alpha < 1$ but under risk aversion coefficient β_1 when $\alpha = 1$ (the case $\alpha = 1$ actually reduces to the model without reference αh). In other words, the model is somewhat "not continuous" at $\alpha = 1$. Hence, we just consider the case $\alpha < 1$. Another two limiting cases $\beta_1 \to 0$ and $\beta_2 \to 0$ are discussed in Subsection 6.2.3.

Remark 2. Alternatively, we can consider more general reference point instead of αh. It turns out that the solution techniques we used are still applicable. Base on these techniques and tools, we give the results with an alternative reference in Appendix D, which is more general and reduces to the current model in a special case.

For convenience, we only deal with the case that $r = \gamma$ in this paper. For the general case, all the results are paralleled to the case $r = \gamma$ but more complicated. Interested readers can refer to Appendix E for more details.

At the end of this section, we provide a lemma which gives a necessary condition for an admissible strategy, which will be used later in Section 3.

Lemma 2.2. *For any $(c, \pi) \in \mathcal{A}$ (if $\mathcal{A} \neq \emptyset$), the corresponding wealth process $\{X_t, t \geq 0\}$ and running maximum process $\{h_t, t \geq 0\}$ must satisfy $X_t \geq \frac{\lambda}{r} h_t$, a.s. for $\forall t \geq 0$.*

Proof. Suppose that $(c, \pi) \in \mathcal{A}$ and there exists t_0 such that $\mathbb{P}(X_{t_0} < \frac{\lambda}{r} h_{t_0}) > 0$. Then there exists $\varepsilon > 0$ such that $\mathbb{P}(X_{t_0} < (1 - \varepsilon)\frac{\lambda}{r} h_{t_0}) > 0$. We only need to show that even for strategies with lowest consumption, i.e., $c_t = \lambda h_{t_0}$, $\forall t \geq t_0$, there exists $t_1 > t_0$ such that $\mathbb{P}(X_{t_1} < 0) > 0$. For strategies with $c_t = \lambda h_{t_0}$, $\forall t \geq t_0$, the dynamic of wealth becomes

$$dX_t = rX_t dt + \pi_t(\mu - r)dt + \pi_t \sigma dB_t - \lambda h_{t_0} dt, \ \forall t \geq t_0.$$

Solving this SDE, we obtain

$$X_t = e^{r(t-t_0)} X_{t_0} + \frac{\lambda h_{t_0}}{r}\left[1 - e^{r(t-t_0)}\right] + (\mu - r)e^{rt}\int_{t_0}^t e^{-ru}\pi_u du + \sigma e^{rt}\int_{t_0}^t e^{-ru}\pi_u dB_u, \ \forall t \geq t_0. \quad (2.4)$$

For $t = t_0 + \frac{1}{r}\ln(\frac{1}{\varepsilon})$, we have

$$\mathbb{P}\left(e^{r(t-t_0)} X_{t_0} + \frac{\lambda h_{t_0}}{r}\left[1 - e^{r(t-t_0)}\right] < 0\right)$$

$$= \mathbb{P}\left(X_{t_0} < (1 - \varepsilon)\frac{\lambda}{r} h_{t_0}\right)$$

$$> 0.$$

To handle the sum of last two terms, we introduce the probability measure $\tilde{\mathbb{P}}$ by

$$\left.\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}}\right|_{\mathcal{F}_t} := e^{-\frac{\mu - r}{\sigma} B_t - \frac{(\mu - r)^2}{2\sigma^2}t}, \ \forall t \geq 0.$$

8

Then $\{\tilde{B}_t := B_t + \frac{\mu-r}{\sigma}t, t \geq 0\}$ is standard Brownian motion under $\tilde{\mathbb{P}}$ and the sum of last two terms equals $\sigma e^{rt} \int_{t_0}^t e^{-ru}\pi_u d\tilde{B}_u$, which has zero expectation under $\tilde{\mathbb{P}}$. Hence

$$\tilde{\mathbb{P}}\Big(\sigma e^{rt} \int_{t_0}^t e^{-ru}\pi_u d\tilde{B}_u \leq 0\Big) > 0, \ \forall t \geq 0,$$

then

$$\mathbb{P}\left((\mu-r)e^{rt} \int_{t_0}^t e^{-ru}\pi_u du + \sigma e^{rt} \int_{t_0}^t e^{-ru}\pi_u dB_u\right)$$
$$= \mathbb{P}\Big(\sigma e^{rt} \int_{t_0}^t e^{-ru}\pi_u d\tilde{B}_u \leq 0\Big)$$
$$> 0, \ \forall t \geq 0.$$

Since $(c, \pi) \in \mathcal{A}$, we deduce that X_{t_0} and h_{t_0} are $\mathcal{F}_{t_0}$ measurable. Then the sum of first two terms and the sum of last two terms of (2.4) are independent. Therefore, choosing $t_1 = t_0 + \frac{1}{r}\ln(\frac{1}{\varepsilon})$, we have $\mathbb{P}(X_{t_1} < 0) > 0$ and the proof is complete. $\qquad\square$

3 Derivation of the Solution in Dual Form

In this section, we apply the martingale optimality principle to derive the HJB equation of Problem (2.3) and use the duality method to obtain the solution in dual form.

To begin with, the value function of Problem (2.3) is denoted by

$$V(x_0, h_0) \triangleq \sup_{(c,\pi)\in\mathcal{A}} \mathbb{E}_{x_0,h_0} \int_0^\infty e^{-\gamma t}U(c_t, h_t)dt.$$

Definition 3.1. An optimal strategy (c^*, π^*) of Problem (2.3) is an admissible strategy which satisfies

$$\mathbb{E}_{x_0,h_0} \int_0^\infty e^{-\gamma t}U(c_t^*, h_t^*)dt = V(x_0, h_0).$$

The martingale optimality principle shows that the process $\{\Gamma_t, t \geq 0\}$

$$\Gamma_t \triangleq e^{-\gamma t}V(X_t, h_t) + \int_0^t e^{-\gamma s}U(c_s, h_s)ds$$

is a local supermartingale for all admissible (c, π) and is a local martingale for the optimal (c^*, π^*). If the value function is smooth enough, applying the Itô's rule to $\{\Gamma_t, t \geq 0\}$, we derive the HJB equation of Problem (2.3) as follows[1]:

$$\begin{cases} \sup_{c\in[0,h],\pi\in\mathbb{R}} \left\{-\gamma V(x,h) + V_x(x,h)\big(rx + \pi(\mu-r) - c\big) + \frac{1}{2}V_{xx}(x,h)\sigma^2\pi^2 + U(c,h)\right\} = 0, \\ V_h(x,h) = 0 \text{ on } (x,h) \text{ s.t. } \dfrac{dh}{dt} \neq 0, \end{cases} \tag{3.1}$$

---

[1] For notational simplicity, we write x, h, c, π instead of X_t, h_t, c_t, π_t in (3.1) and (3.4). $\frac{dh}{dt}$ in the second line refers to $\frac{dh_t}{dt}$, which is the derivative of h_t in the sense of distribution. Heuristically, $\frac{dh}{dt} \neq 0$ means that the process $\{h_t\}$ strictly increases at the instant t.

and the optimal feedback form of π is

$$\pi^*_{\text{primal}}(x, h) = -\frac{\mu - r}{\sigma^2} \frac{V_x(x, h)}{V_{xx}(x, h)}.$$

As $\frac{x}{h} \to \frac{\lambda}{r}$, the optimal investment should decline to zero to avoid bankruptcy, we have

$$\lim_{\frac{x}{h} \to \frac{\lambda}{r}^+} \frac{V_x(x, h)}{V_{xx}(x, h)} = 0. \tag{3.2}$$

To solve HJB equation (3.1) with the boundary condition (3.2) based on the duality method, we introduce the conjugate of the value function as follows:

$$\tilde{V}(y, h) = \sup_{x \geq 0}\{V(x, h) - xy\}, \quad y > 0.$$

Then we have the duality transform

$$\begin{cases} x = -\tilde{V}_y(y, h), \\ V(x, h) = \tilde{V}(y, h) - y\tilde{V}_y(y, h), \\ V_x(x, h) = y, \\ V_{xx}(x, h) = -\frac{1}{\tilde{V}_{yy}(y, h)}, \\ V_h(x, h) = \tilde{V}_h(y, h), \end{cases} \tag{3.3}$$

and (3.1) is rewritten in dual form as follows:

$$\begin{cases} \sup_{c \in [0, h], \pi \in \mathbb{R}} \left\{ -\gamma \left[\tilde{V}(y, h) - y\tilde{V}_y(y, h) \right] + y \left[-r\tilde{V}_y(y, h) + \pi(\mu - r) - c \right] - \frac{1}{2\tilde{V}_{yy}(y, h)}\sigma^2\pi^2 \right. \\ \left. \qquad + U(c, h) \right\} = 0, \\ \tilde{V}_h(y, h) = 0 \text{ on } (y, h) \ s.t. \ \frac{dh}{dt} \neq 0. \end{cases} \tag{3.4}$$

For (3.4), the optimal feedback form $\pi^* = \pi^*(y, h)$ is

$$\pi^*(y, h) = \frac{(\mu - r)y\tilde{V}_{yy}(y, h)}{\sigma^2}. \tag{3.5}$$

And the optimal feedback form $c^* = c^*(y, h)$ maximizing $\bar{U}(c) \triangleq U(c, h) - cy$ on $[\lambda h, h]$ is

$$c^*(y, h) = \begin{cases} \lambda h, & e^{(\alpha - \lambda)\beta_1 h} \leq y, \\ -\frac{1}{\beta_1}\ln(y) + \alpha h, & 1 \leq y < e^{(\alpha - \lambda)\beta_1 h}, \\ -\frac{1}{\beta_2}\ln(y) + \alpha h, & e^{-(1-\alpha)\beta_2 h} \leq y < 1, \\ h, & 0 \leq y < e^{-(1-\alpha)\beta_2 h}. \end{cases} \tag{3.6}$$

For the region where $c^*(y, h) = h$, three sub cases need to be distinguished in order to apply the second equation of (3.4) to solve the HJB equation. The first case is that the current consumption just reaches the past consumption peak but does not update it; the second case is that the current consumption reaches the past consumption peak and updates it; the last case is that the current consumption exceeds the past consumption peak and forces the running maximum process to jump. The last case can only happen at $t = 0$ where the inherited running maximum level is lower but the initial wealth is abundant, which switches to the former two cases for $t > 0$. Hence we only

need to consider the first two cases for $t > 0$. The second equation of (3.4) refers to the second case where the running maximum h is updated, which instructs us to further separate the region according to the above different sub cases. Specifically, the running maximum is updated if and only if $\underset{c}{\text{argmax}}\{U(c,c) - cy\} \geq h$. Moreover, the running maximum h jumps if strict inequality holds. As $\underset{c}{\text{argmax}}\{U(c,c) - cy\} \geq h$ is equivalent to $y \leq (1-\alpha)e^{-(1-\alpha)\beta_2 h}$, we deduce that for any initial point (y_0, h_0) s.t. $y_0 < (1-\alpha)e^{-(1-\alpha)\beta_2 h_0}$, it will jump immediately to $\left(y_0, \frac{1}{(1-\alpha)\beta_2}\ln(\frac{1-\alpha}{y_0})\right)$ which is on the curve $y = (1-\alpha)e^{-(1-\alpha)\beta_2 h}$. Hence we only need to consider initial points (y_0, h_0) in the dual region where $y \geq (1-\alpha)e^{-(1-\alpha)\beta_2 h}$. Meanwhile, the second formula of (3.4) is equivalent to

$$V_h(y,h) = 0 \text{ for } y = (1-\alpha)e^{-(1-\alpha)\beta_2 h}. \tag{3.7}$$

By the duality transform (3.3), the constraint $x \geq \frac{\lambda h}{\gamma}$ is equivalent to $-\tilde{V}_y(y,h) \geq \frac{\lambda h}{\gamma}$. As such, the dual effective region $\mathcal{C}_d$ can be defined by

$$\mathcal{C}_d \triangleq \left\{(y,h)\big|y \geq (1-\alpha)e^{-(1-\alpha)\beta_2 h}, \ -\tilde{V}_y(y,h) \geq \frac{\lambda h}{\gamma}, \ h > 0\right\}.$$

Applying the duality transform again, the effective region $\mathcal{C}$ is

$$\mathcal{C} \triangleq \left\{(x,h)\big|V_x(x,h) \geq (1-\alpha)e^{-(1-\alpha)\beta_2 h}, \ x \geq \frac{\lambda h}{\gamma}, \ h > 0\right\}. \tag{3.8}$$

Using the first equation of (3.4), we obtain

$$-\gamma\tilde{V}(y,h) + \frac{(r-\mu)^2}{2\sigma^2}y^2\tilde{V}_{yy}(y,h) = -\tilde{U}(y,h), \tag{3.9}$$

where $\tilde{U}(y,h) = \underset{\lambda h \leq c \leq h}{\sup}\{U(c,h) - cy\}$ with

$$\tilde{U}(y,h) = \begin{cases} \frac{1}{\beta_1}(1 - e^{(\alpha-\lambda)\beta_1 h}) - \lambda hy, & e^{(\alpha-\lambda)\beta_1 h} \leq y, \\ \frac{1}{\beta_1}\left[1 - y + y\ln(y)\right] - \alpha hy, & 1 \leq y < e^{(\alpha-\lambda)\beta_1 h}, \\ \frac{1}{\beta_2}\left[1 - y + y\ln(y)\right] - \alpha hy, & e^{-(1-\alpha)\beta_2 h} \leq y < 1, \\ \frac{1}{\beta_2}\left[(1 - e^{-(1-\alpha)\beta_2 h}) - hy\right], & (1-\alpha)e^{-(1-\alpha)\beta_2 h} \leq y < e^{-(1-\alpha)\beta_2 h}. \end{cases}$$

Define $k = \frac{(r-\mu)^2}{2\sigma^2}$, $q_1 = \frac{k - \sqrt{k^2+4k\gamma}}{2k} < 0$, $q_2 = \frac{k + \sqrt{k^2+4k\gamma}}{2k} > 1$, then the general solution of (3.9) is

$$\tilde{V}(y,h) = \begin{cases} C_1(h)y^{q_1} + C_2(h)y^{q_2} - \frac{1}{\gamma}\lambda hy + \frac{1}{\gamma\beta_1}\left[1 - e^{(\alpha-\lambda)\beta_1 h}\right], & e^{(\alpha-\lambda)\beta_1 h} \leq y, \\ C_3(h)y^{q_1} + C_4(h)y^{q_2} + \frac{1}{\gamma\beta_1}\left[1 - y + y\ln(y)\right] + \frac{k}{\gamma^2\beta_1}y - \frac{1}{\gamma}\alpha hy, & 1 \leq y < e^{(\alpha-\lambda)\beta_1 h}, \\ C_5(h)y^{q_1} + C_6(h)y^{q_2} + \frac{1}{\gamma\beta_2}\left[1 - y + y\ln(y)\right] + \frac{k}{\gamma^2\beta_2}y - \frac{1}{\gamma}\alpha hy, & e^{-(1-\alpha)\beta_2 h} \leq y < 1, \\ C_7(h)y^{q_1} + C_8(h)y^{q_2} - \frac{1}{\gamma}hy + \frac{1}{\gamma\beta_2}\left[1 - e^{-(1-\alpha)\beta_2 h}\right], & (1-\alpha)e^{-(1-\alpha)\beta_2 h} \leq y < e^{-(1-\alpha)\beta_2 h}. \end{cases} \tag{3.10}$$

Based on the duality transform (3.3), the boundary condition (3.2) can be rewritten as follows:

$$\lim_{\tilde{V}_y(y,h) \to -\frac{\lambda h}{\gamma}^-} y\tilde{V}_{yy}(y,h) = 0. \tag{3.11}$$

Then we deduce from (3.11) and (3.10) that $\tilde{V}_y(y,h) \to -\frac{\lambda h}{\gamma} \Leftrightarrow y \to \infty$ and that $C_2(h) = 0$, $C_1(h) > 0$. Using (3.10), we obtain $-\tilde{V}_y(y,h) \geq \frac{\lambda h}{\gamma}$ and the dual effective region is simplified to

$$\mathcal{C}_d = \left\{(y,h)\big|y \geq (1-\alpha)e^{-(1-\alpha)\beta_2 h}, \ h > 0\right\}. \tag{3.12}$$

11

In addition, when $y = (1 - \alpha)e^{-(1-\alpha)\beta_2 h}$ and $h \to \infty$, the initial wealth $x \to \infty$ and the utility keeps near its maximum $\frac{1}{\beta_2}$ for infinitely long time. Thus the value function tends to $\frac{1}{\gamma\beta_2}$. This boundary condition can be expressed as

$$\lim_{h\to\infty} \left[\tilde{V}(y,h) - y\tilde{V}_y(y,h) \right]\Big|_{y=(1-\alpha)e^{-(1-\alpha)\beta_2 h}} = \frac{1}{\gamma\beta_2}.$$

The above boundary condition together with $C_2(h) = 0$, (3.7) and smooth-fit conditions

$$\tilde{V}(y+,h) = \tilde{V}(y-,h),$$
$$\tilde{V}_y(y+,h) = \tilde{V}_y(y-,h)$$

leads to

$$C_2(h) = 0, \quad C_4(h) = -\frac{k}{\gamma^2\beta_1}\frac{1-q_1}{q_2-q_1}e^{-(\alpha-\lambda)(q_2-1)\beta_1 h}, \tag{3.13}$$

$$C_6(h) = C_4(h) + \frac{k}{\gamma^2}\frac{\beta_2-\beta_1}{\beta_1\beta_2}\frac{1-q_1}{q_2-q_1}, \tag{3.14}$$

$$C_8(h) = C_6(h) + \frac{k}{\gamma^2\beta_2}\frac{1-q_1}{q_2-q_1}e^{(1-\alpha)(q_2-1)\beta_2 h}, \tag{3.15}$$

$$C_7(h) = \frac{(1-\alpha)^{q_2-q_1}\frac{k}{\gamma^2}\frac{1-q_1}{q_2-q_1}(\alpha-\lambda)(q_2-1)}{(1-\alpha)(q_2-q_1)\beta_2 + (\alpha-\lambda)(q_2-1)\beta_1}e^{-\left[(1-\alpha)(q_2-q_1)\beta_2+(\alpha-\lambda)(q_2-1)\beta_1\right]h}$$

$$+ (1-\alpha)^{q_2-q_1}\frac{k}{\gamma^2\beta_2}\frac{q_2-1}{q_2-q_1}e^{-(1-\alpha)(1-q_1)\beta_2 h}, \tag{3.16}$$

$$C_5(h) = C_7(h) - \frac{k}{\gamma^2\beta_2}\frac{q_2-1}{q_2-q_1}e^{-(1-\alpha)(1-q_1)\beta_2 h},$$

$$C_3(h) = C_5(h) - \frac{k}{\gamma^2}\frac{\beta_2-\beta_1}{\beta_1\beta_2}\frac{q_2-1}{q_2-q_1}, \tag{3.17}$$

$$C_1(h) = C_3(h) + \frac{k}{\gamma^2\beta_1}\frac{q_2-1}{q_2-q_1}e^{(\alpha-\lambda)(1-q_1)\beta_1 h}. \tag{3.18}$$

We can directly show

$$C_1(h) > 0, \quad C_4(h) < 0,$$
$$C_7(h) > 0, \quad C_8(h) > 0$$

and obtain the following order estimates of the coefficients $C_i(h)$, $1 \leq i \leq 8$, which will later be used in the proof of the verification theorem.

Lemma 3.2. *As $h \to \infty$,*

$$C_1(h) = \mathcal{O}(e^{(\alpha-\lambda)(1-q_1)\beta_1 h}),$$
$$C_3(h) = \mathcal{O}(1), \quad C_4(h) = \mathcal{O}(e^{-(\alpha-\lambda)(q_2-1)\beta_1 h}),$$
$$C_5(h) = \mathcal{O}(e^{-(1-\alpha)(1-q_1)\beta_2 h}), \quad C_6(h) = \mathcal{O}(1),$$
$$C_7(h) = \mathcal{O}(e^{-(1-\alpha)(1-q_1)\beta_2 h}), \quad C_8(h) = \mathcal{O}(e^{(1-\alpha)(q_2-1)\beta_2 h}).$$

4   Verification Theorem and Optimal Strategy

In this section, we establish the verification theorem and apply the duality transform to obtain the optimal strategy. First, we state the verification theorem which gives the optimal consumption and investment policy in dual form.

Theorem 4.1 (Verification Theorem). *For any $(x_0, h_0) \in \mathcal{C}$, where x_0 and h_0 are respectively the initial wealth and initial past spending maximum, and $\mathcal{C}$ is the effective region given by (3.8), the value function $V(x_0, h_0)$ can be attained by the optimal consumption and investment strategy given by*

$$(c^*, \pi^*) = \left\{ \left(c^*\big(Y_t(y^*), H_t^{\dagger}(y^*)\big), \pi^*\big(Y_t(y^*), H_t^{\dagger}(y^*)\big) \right), t \geq 0 \right\}$$

where $Y_t(\cdot)$ is given by

$$Y_t(y) \triangleq y e^{\gamma t} M_t$$

with $M \triangleq \left\{ M_t := e^{-(\gamma + \frac{(\mu - r)^2}{2\sigma^2})t - \frac{\mu - r}{\sigma} B_t}, t \geq 0 \right\}$ being the discounted state price density process, $\left\{ H_t^{\dagger}(\cdot), t \geq 0 \right\}$ is determined by

$$H_t^{\dagger}(y) \triangleq h_0 \vee \sup_{s \leq t} c^*\big(Y_s(y), H_s^{\dagger}(y)\big),$$

$y^* = y^*(x_0, h_0)$ is the unique solution of

$$\mathbb{E}_{x_0, h_0} \int_0^{\infty} c^*\big(Y_t(y), H_t^{\dagger}(y)\big) M_t dt = x_0 \tag{4.1}$$

and the feedback functions $c^*(\cdot, \cdot)$ and $\pi^*(\cdot, \cdot)$ are respectively given by

$$c^*(y, h) = \begin{cases} \lambda h, & e^{(\alpha - \lambda)\beta_1 h} \leq y, \\ -\frac{1}{\beta_1} \ln(y) + \alpha h, & 1 \leq y < e^{(\alpha - \lambda)\beta_1 h}, \\ -\frac{1}{\beta_2} \ln(y) + \alpha h, & e^{-(1-\alpha)\beta_2 h} \leq y < 1, \\ h, & (1 - \alpha) e^{-(1-\alpha)\beta_2 h} \leq y < e^{-(1-\alpha)\beta_2 h}. \end{cases} \tag{4.2}$$

$$\pi^*(y, h) = \frac{\mu - r}{\sigma^2} \begin{cases} \frac{r}{k}\left[C_1(h) y^{q_1 - 1} + C_2(h) y^{q_2 - 1} \right], & e^{(\alpha - \lambda)\beta_1 h} \leq y, \\ \frac{r}{k}\left[C_3(h) y^{q_1 - 1} + C_4(h) y^{q_2 - 1} \right] + \frac{1}{\gamma \beta_1}, & 1 \leq y < e^{(\alpha - \lambda)\beta_1 h}, \\ \frac{r}{k}\left[C_5(h) y^{q_1 - 1} + C_6(h) y^{q_2 - 1} \right] + \frac{1}{\gamma \beta_2}, & e^{-(1-\alpha)\beta_2 h} \leq y < 1, \\ \frac{r}{k}\left[C_7(h) y^{q_1 - 1} + C_8(h) y^{q_2 - 1} \right], & (1 - \alpha) e^{-(1-\alpha)\beta_2 h} \leq y < e^{-(1-\alpha)\beta_2 h}. \end{cases} \tag{4.3}$$

Proof. See Appendix A. $\qquad \square$

Now, to apply the dual transform to present the primal value function as well as the optimal consumption and investment policy with the primal variable, we need the following lemma in the dual transform, and its proof is given in Appendix C.

Lemma 4.2. $\tilde{V}_{yy}(y, h) > 0$ for $(y, h) \in \mathcal{C}_d$ and hence the inverse of $-\tilde{V}_y(\cdot, h)$ exists.

Let $f(\cdot, h)$ be the inverse of $-\tilde{V}_y(\cdot, h)$ based on Lemma 4.2, then, using dual transform (3.3), we have

$$\begin{cases} x = -\tilde{V}_y(y, h), \\ y = f(x, h). \end{cases} \tag{4.4}$$

As such, based on (4.4) and Lemma 4.2, the function $f(\cdot, h)$ is implicitly and uniquely determined by

$$x = -\tilde{V}_y\big(f(x, h), h\big). \tag{4.5}$$

Plugging (3.10) into (4.5), we obtain the following forms of $f(x, h)$ according to different regions:
(1) For $e^{(\alpha-\lambda)\beta_1 h} \le f(x, h)$, we have $f(x, h) = f_1(x, h)$ with $f_1(x, h)$ satisfying

$$x = -C_1(h)q_1\big[f_1(x, h)\big]^{q_1-1} - C_2(h)q_2\big[f_1(x, h)\big]^{q_2-1} + \frac{\lambda h}{\gamma}. \tag{4.6}$$

Based on Lemma 4.2, $V_{yy}(\cdot, h) > 0$, then the inequality $e^{(\alpha-\lambda)\beta_1 h} \le f(x, h)$ is equivalent to $x \le W_{\text{low}}(h)$ with

$$W_{\text{low}}(h) = -C_1(h)q_1 e^{-(\alpha-\lambda)(1-q_1)\beta_1 h} - C_2(h)q_2 e^{(\alpha-\lambda)(q_2-1)\beta_1 h} + \frac{\lambda h}{\gamma}. \tag{4.7}$$

(2) For $1 \le f(x, h) < e^{(\alpha-\lambda)\beta_1 h}$, $f(x, h) = f_2(x, h)$ with $f_2(x, h)$ satisfying

$$x = -C_3(h)q_1\big[f_2(x, h)\big]^{q_1-1} - C_4(h)q_2\big[f_2(x, h)\big]^{q_2-1} - \frac{1}{\gamma\beta_1}\ln\big[f_2(x, h)\big] - \frac{k}{\gamma^2\beta_1} + \frac{\alpha h}{\gamma}. \tag{4.8}$$

The inequality $1 \le f(x, h) < e^{(\alpha-\lambda)\beta_1 h}$ is equivalent to $W_{\text{low}}(h) < x \le W_{\text{ref}}(h)$, where

$$W_{\text{ref}}(h) = -C_3(h)q_1 - C_4(h)q_2 - \frac{k}{\gamma^2\beta_1} + \frac{\alpha h}{\gamma}. \tag{4.9}$$

(3) For $e^{-(1-\alpha)\beta_2 h} \le f(x, h) < 1$, $f(x, h) = f_3(x, h)$ with $f_3(x, h)$ satisfying

$$x = -C_5(h)q_1\big[f_3(x, h)\big]^{q_1-1} - C_6(h)q_2\big[f_3(x, h)\big]^{q_2-1} - \frac{1}{\gamma\beta_2}\ln\big[f_3(x, h)\big] - \frac{k}{\gamma^2\beta_2} + \frac{\alpha h}{\gamma}. \tag{4.10}$$

The inequality $e^{-(1-\alpha)\beta_2 h} \le f(x, h) < 1$ is equivalent to $W_{\text{ref}}(h) < x \le W_{\text{peak}}(h)$ with

$$W_{\text{peak}}(h) = -C_5(h)q_1 e^{(1-\alpha)(1-q_1)\beta_2 h} - C_6(h)q_2 e^{-(1-\alpha)(q_2-1)\beta_2 h} - \frac{k}{\gamma^2\beta_2} + \frac{h}{\gamma}. \tag{4.11}$$

(4) For $(1-\alpha)e^{-(1-\alpha)\beta_2 h} \le f(x, h) < e^{-(1-\alpha)\beta_2 h}$, $f(x, h) = f_4(x, h)$ with $f_4(x, h)$ satisfying

$$x = -C_7(h)q_1\big[f_4(x, h)\big]^{q_1-1} - C_8(h)q_2\big[f_4(x, h)\big]^{q_2-1} + \frac{h}{\gamma}. \tag{4.12}$$

The inequality $(1-\alpha)e^{-(1-\alpha)\beta_2 h} \le f(x, h) < e^{-(1-\alpha)\beta_2 h}$ is equivalent to $W_{\text{peak}}(h) < x \le W_{\text{updt}}(h)$ with

$$W_{\text{updt}}(h) = -C_7(h)q_1(1-\alpha)^{q_1-1}e^{(1-\alpha)(1-q_1)\beta_2 h} - C_8(h)q_2(1-\alpha)^{q_2-1}e^{-(1-\alpha)(q_2-1)\beta_2 h} + \frac{h}{\gamma}. \tag{4.13}$$

We summarize the forms of the primal value function as well as the optimal consumption and investment policy in terms of primal variable in the following Theorems 4.3 and 4.4:

Theorem 4.3. *For $(x,h) \in \mathcal{C}$, where $\mathcal{C}$ is the effective region given by (3.8), the value function of Problem (2.3) is*

$$V(x,h) = \begin{cases} C_1(h)\big[f_1(x,h)\big]^{q_1} + C_2(h)\big[f_1(x,h)\big]^{q_2} - \frac{1}{\gamma}\lambda h f_1(x,h) \\ \quad + \frac{1}{\gamma\beta_1}(1 - e^{(\alpha-\lambda)\beta_1 h}), \ \frac{\lambda h}{\gamma} \le x \le W_{\text{low}}(h), \\ C_3(h)\big[f_2(x,h)\big]^{q_1} + C_4(h)\big[f_2(x,h)\big]^{q_2} + \frac{1}{\gamma\beta_1}\Big[1 - f_2(x,h) + f_2(x,h)\ln\big[f_2(x,h)\big]\Big] \\ \quad + \frac{k}{\gamma^2\beta_1}f_2(x,h) - \frac{1}{\gamma}\alpha h f_2(x,h), \ W_{\text{low}}(h) < x \le W_{\text{ref}}(h), \\ C_5(h)\big[f_3(x,h)\big]^{q_1} + C_6(h)\big[f_3(x,h)\big]^{q_2} + \frac{1}{\gamma\beta_2}\Big[1 - f_3(x,h) + f_3(x,h)\ln\big[f_3(x,h)\big]\Big] \\ \quad + \frac{k}{\gamma^2\beta_2}f_3(x,h) - \frac{1}{\gamma}\alpha h f_3(x,h), \ W_{\text{ref}}(h) < x \le W_{\text{peak}}(h), \\ C_7(h)\big[f_4(x,h)\big]^{q_1} + C_8(h)\big[f_4(x,h)\big]^{q_2} - \frac{1}{\gamma}h f_4(x,h) \\ \quad + \frac{1}{\gamma\beta_2}\big[1 - e^{-(1-\alpha)\beta_2 h}\big], \ W_{\text{peak}}(h) < x \le W_{\text{updt}}(h), \end{cases}$$

(4.14)

where $W_{\text{low}}(h), W_{\text{ref}}(h), W_{\text{peak}}(h), W_{\text{updt}}(h)$ and $f_i(x,h), \ 1 \le i \le 4$ are given by (4.6)~(4.13).

Proof. Applying dual transform (3.3) and (4.4) yields

$$V(x,h) = \tilde{V}\big(f(x,h),h\big) + x f(x,h).$$

Plugging in (3.10), the desired result follows. □

Theorem 4.4. *For $(x_0,h_0) \in \mathcal{C}$, where $\mathcal{C}$ is the effective region given by (3.8), let $c^*_{\text{primal}}(\cdot,\cdot)$ and $\pi^*_{\text{primal}}(\cdot,\cdot)$ be the feedback functions in terms of primal variable given respectively by*

$$c^*_{\text{primal}}(x,h) = \begin{cases} \lambda h, & \frac{\lambda h}{\gamma} \le x \le W_{\text{low}}(h), \\ -\frac{1}{\beta_1}\ln\big[f_2(x,h)\big] + \alpha h, & W_{\text{low}}(h) < x \le W_{\text{ref}}(h), \\ -\frac{1}{\beta_2}\ln\big[f_3(x,h)\big] + \alpha h, & W_{\text{ref}}(h) < x \le W_{\text{peak}}(h), \\ h, & W_{\text{peak}}(h) < x \le W_{\text{updt}}(h), \end{cases}$$

$$\pi^*_{\text{primal}}(x,h) = \frac{\mu - r}{\sigma^2}\begin{cases} \frac{r}{k}\Big\{C_1(h)\big[f_1(x,h)\big]^{q_1-1} + C_2(h)\big[f_1(x,h)\big]^{q_2-1}\Big\}, & \frac{\lambda h}{\gamma} \le x \le W_{\text{low}}(h), \\ \frac{r}{k}\Big\{C_3(h)\big[f_2(x,h)\big]^{q_1-1} + C_4(h)\big[f_2(x,h)\big]^{q_2-1}\Big\} + \frac{1}{\gamma\beta_1}, & W_{\text{low}}(h) < x \le W_{\text{ref}}(h), \\ \frac{r}{k}\Big\{C_5(h)\big[f_3(x,h)\big]^{q_1-1} + C_6(h)\big[f_3(x,h)\big]^{q_2-1}\Big\} + \frac{1}{\gamma\beta_2}, & W_{\text{ref}}(h) < x \le W_{\text{peak}}(h), \\ \frac{r}{k}\Big\{C_7(h)\big[f_4(x,h)\big]^{q_1-1} + C_8(h)\big[f_4(x,h)\big]^{q_2-1}\Big\}, & W_{\text{peak}}(h) < x \le W_{\text{updt}}(h), \end{cases}$$

where $W_{\text{low}}(h), W_{\text{ref}}(h), W_{\text{peak}}(h), W_{\text{updt}}(h)$ and $f_i(x,h), \ 1 \le i \le 4$ are given by (4.6)~(4.13).

Then SDE

$$\begin{cases} dX_t = rX_t dt + \pi^*_{\text{primal}}(X_t, H^*_t)(\mu - r)dt + \pi^*_{\text{primal}}(X_t, H^*_t)\sigma dW_t - c^*_{\text{primal}}(X_t, H^*_t)dt, \\ X_0 = x_0 \end{cases}$$

with $H^*_t \triangleq h_0 \vee \sup\limits_{s \le t} c^*_{\text{primal}}(X_s, H^*_s)$ and $H^*_0 = h_0$, has a unique strong solution $\{X^*_t, \ t \ge 0\}$. The optimal consumption and investment strategy is

$$\Big\{\big(c^*_{\text{primal}}(X^*_t, H^*_t), \pi^*_{\text{primal}}(X^*_t, H^*_t)\big), \ t \ge 0\Big\}.$$

Proof. The proof is based on the following Lemmas 4.5 and 4.6. Then, as the proof is similar to that of Deng et al. (2022), we omit it here. □

Lemma 4.5. *The function f is C^1 within each sub-region of $\mathcal{C}$: $\frac{\lambda h}{\gamma} \le x \le W_{\text{low}}(h)$, $W_{\text{low}}(h) < x \le W_{\text{ref}}(h)$, $W_{\text{ref}}(h) < x \le W_{\text{peak}}(h)$, $W_{\text{peak}}(h) < x \le W_{\text{updt}}(h)$, and it is continuous at the boundary of $x = W_{\text{low}}(h)$, $x = W_{\text{ref}}(h)$, $x = W_{\text{peak}}(h)$. Moreover, we have*

$$
\begin{aligned}
f_x(x,h) &= -\frac{1}{\tilde{V}_{yy}(f,h)} \\
&= \begin{cases}
\frac{k}{r}\left\{-C_1(h)\big[f_1(x,h)\big]^{q_1-2} - C_2(h)\big[f_1(x,h)\big]^{q_2-2}\right\}^{-1}, & \frac{\lambda h}{\gamma} \le x \le W_{\text{low}}(h), \\[2mm]
\left(\frac{r}{k}\left\{-C_3(h)\big[f_2(x,h)\big]^{q_1-2} - C_4(h)\big[f_2(x,h)\big]^{q_2-2}\right\} - \frac{1}{\gamma\beta_1 f_2(x,h)}\right)^{-1}, & W_{\text{low}}(h) < x \le W_{\text{ref}}(h), \\[2mm]
\left(\frac{r}{k}\left\{-C_5(h)\big[f_3(x,h)\big]^{q_1-2} - C_6(h)\big[f_3(x,h)\big]^{q_2-2}\right\} - \frac{1}{\gamma\beta_2 f_3(x,h)}\right)^{-1}, & W_{\text{ref}}(h) < x \le W_{\text{peak}}(h), \\[2mm]
\frac{k}{r}\left\{-C_7(h)\big[f_4(x,h)\big]^{q_1-2} - C_8(h)\big[f_4(x,h)\big]^{q_2-2}\right\}^{-1}, & W_{\text{peak}}(h) < x \le W_{\text{updt}}(h),
\end{cases}
\end{aligned}
\tag{4.15}
$$

$$
f_h(x,h) = \tilde{V}_{yh}(f,h)f_x(x,h). \tag{4.16}
$$

Proof. The proof is similar to Lemma 5.6 in Deng et al. (2022) and omitted here. $\qquad\square$

Lemma 4.6. *The function c^*_{primal} is locally Lipschitz on $\mathcal{C}$ and the function π^*_{primal} is Lipschitz on $\mathcal{C}$.*

Proof. See Appendix C. $\qquad\square$

5 Numerical Analysis with Fixed Parameters

This section aims to illustrate and analyze some properties of the optimal policy and relevant boundaries by fixing the market parameters and numerically computing the results presented in Theorems 4.3 and 4.4.

For simplicity, we define the boundary of the lowest wealth level to satisfy the consumption constraint $c > \lambda h$ as

$$
W_{\text{bkrp}}(h) = \frac{\lambda h}{\gamma}.
$$

The effective region is then between the two boundaries $x = W_{\text{updt}}(h)$ and $x = W_{\text{bkrp}}(h)$. Using three boundaries $x = W_{\text{low}}(h), x = W_{\text{ref}}(h)$ and $x = W_{\text{peak}}(h)$, the effective region is further separated into four parts where the investor takes different strategies in consumption and portfolio selection due to different states of wealth and habit. The ineffective region is separated into two parts $\mathcal{C}_1^c$ and $\mathcal{C}_2^c$:

$$
\begin{aligned}
\mathcal{C}_1^c &\triangleq \big\{(x,h)\big|x < W_{\text{bkrp}}(h),\ h > 0\big\}, \\
\mathcal{C}_2^c &\triangleq \big\{(x,h)\big|x > W_{\text{updt}}(h),\ h > 0\big\}.
\end{aligned}
$$

$\mathcal{C}_1^c$ defines the region where the wealth is too low to maintain the lowest consumption level $c = \lambda h$. The other part $\mathcal{C}_2^c$ implies that the wealth is so high w.r.t the current running maximum level (it can only happen at time $t = 0$ in the optimal case) that it is optimal to consume at a level strictly higher than the running maximum and forces (x,h) to jump to $x = W_{\text{updt}}(h)$.
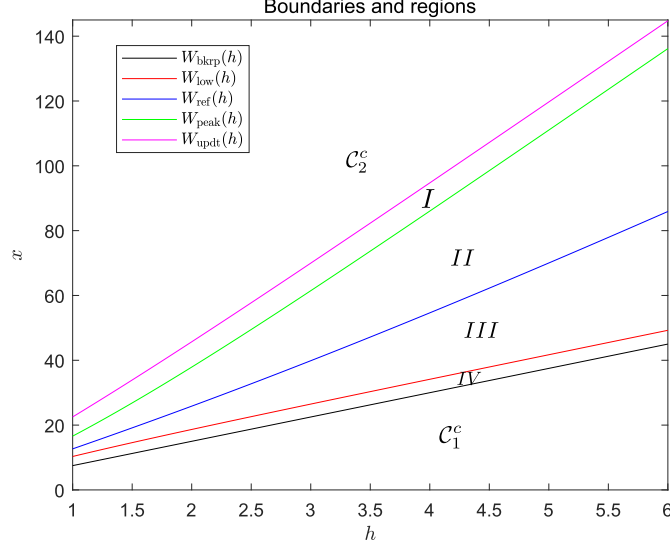
Figure 2: Boundary curves $x = W_{\text{bkrp}}(h), x = W_{\text{low}}(h), x = W_{\text{ref}}(h), x = W_{\text{peak}}(h)$ and $x = W_{\text{updt}}(h)$, different sub-regions I~IV (I:satisfactory, II:recovery, III:depression, IV:gloom) of the effective region and two sub-regions $\mathcal{C}_i^c, i = 1, 2$ of the ineffective region with $\lambda = 0.3$, $\alpha = 0.7$, $\beta_1 = 1$, $\beta_2 = 2$, $r = 0.04$, $\mu = 0.12$, $\sigma = 0.3$.

We fix parameters $\lambda = 0.3$, $\alpha = 0.7$, $\beta_1 = 1$, $\beta_2 = 2$, $r = 0.04$, $\mu = 0.12$, $\sigma = 0.3$ and compute all the boundaries that separate the ineffective region and the effective region and different sub-regions of the effective region. As shown in Figure 2, all boundaries are graphs of increasing functions w.r.t the variable h. For fixed running maximum level h, if the wealth x is so low that (x, h) fall into $\mathcal{C}_1^c$, the investor is too poor to sustain the lower bound consumption constraint; if the wealth is a little higher that (x, h) belongs to sub-region IV of the effective region, then the investor consumes at the lowest level set by the drawdown constraint; if the wealth is higher but not high enough (sub-region III), the investor chooses to consume at a higher level than the lowest level but no more than the reference point level αh where risk aversion increases; as the wealth increase and (x, h) enters sub-region II, the investor has enough wealth to consume above the reference point level αh but not enough to reach the running maximum level; with some more wealth than the former case (now in sub-region I), the investor is able the consume at the running maximum level but not wealthy enough to update it; if the investor has more wealth than $W_{\text{updt}}(h)$, then he may update the running maximum immediately by consuming above the historical running maximum level and causes (x, h) to jump onto the boundary $x = W_{\text{updt}}(h)$, where he consumes at the running maximum level and continuously updates it. Based on the aforementioned economic interpretation, we name the curve $W_{\text{updt}}(h)$ as "bliss" curve, and four sub-regions I-IV as "satisfactory", "recovery", "depression" and "gloom" region, respectively. As can be seen from the analysis below, such a division of state space helps to provide a structural description of both consumption and investment behaviour under our model.

The optimal consumption, shown in Figure 3, is non-decreasing in both the wealth x and the habit h. In satisfactory region and gloom region, the optimal consumption is indifferent with respect to x. That is, for the poorest people in effective region (gloom), they consume as little as possible, while for the wealthiest (satisfactory), they revisit their historical peak of consumption rate. However, in depression region and recovery region, increasing wealth will lead to an increase
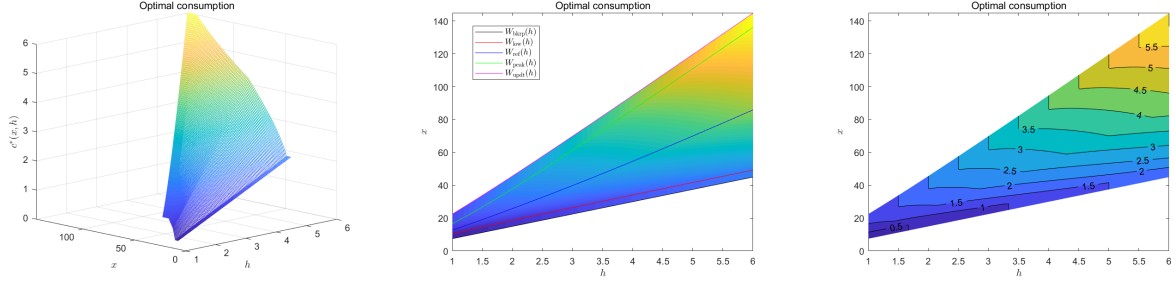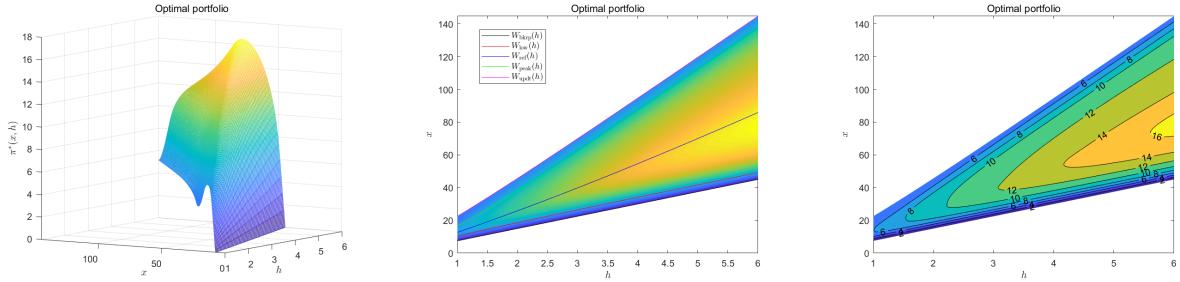
Figure 3: Optimal consumption with fixed parameters $\lambda = 0.3$, $\alpha = 0.7$, $\beta_1 = 1$, $\beta_2 = 2$, $r = 0.04$, $\mu = 0.12$, $\sigma = 0.3$. The left panel plots the optimal consumption as a function of wealth and consumption peak. The middle panel is the two-dimensional projection of the left panel with the boundaries shown. The right panel is the contour plot.

in optimal consumption and the increase is more substantial in the region with lower risk aversion (i.e., more substantial in depression region when $\beta_1 < \beta_2$, and in recovery region when $\beta_1 > \beta_2$). The above analysis suggests that increasing wealth causes one to consume more only in the following two cases: the first case is when the wealth is at least $W_{\mathrm{updt}}(h)$, he is so rich that he decides to consume more even at the cost of raising running maximum h; the other case is when his wealth is more than $W_{\mathrm{low}}(h)$ but less than $W_{\mathrm{peak}}(h)$. There are two sub-cases divided by whether the wealth is more than $W_{\mathrm{ref}}(h)$ in the second case, and the (marginal propensity to consume) MPC out of wealth is generally higher in the region with lower risk aversion.

Figure 4: Optimal portfolio with fixed parameters $\lambda = 0.3$, $\alpha = 0.7$, $\beta_1 = 1$, $\beta_2 = 2$, $r = 0.04$, $\mu = 0.12$, $\sigma = 0.3$. The left panel plots the optimal portfolio as a function of wealth and consumption peak. The middle panel is the two-dimensional projection of the left panel with the boundaries shown. The right panel is the contour plot.

The optimal portfolio is shown in Figure 4. The behavior of the optimal portfolio varies significantly in different regions. For fixed habit h, the optimal portfolio sees a dramatic increase with respect to the variable x in gloom and depression region where the risk aversion is low. However, once crossing $x = W_{\mathrm{ref}}(h)$ and the risk aversion shifting to the high level, increasing wealth causes the optimal portfolio to fall instead. The above result indicates that the change of risk aversion has an overwhelming impact on portfolio selection in our model. For those in gloom and depression region, earning money which increases his wealth stimulates him to invest more in risky assert; while for rich people in recovery and satisfactory region, as well as on bliss curve, the more he earns, the less he is willing to invest in risky assert. Similar conclusions can be obtained from the analysis of optimal proportion of wealth invested in risky assets, or economically, optimal portfolio allocation (see Figure 5).
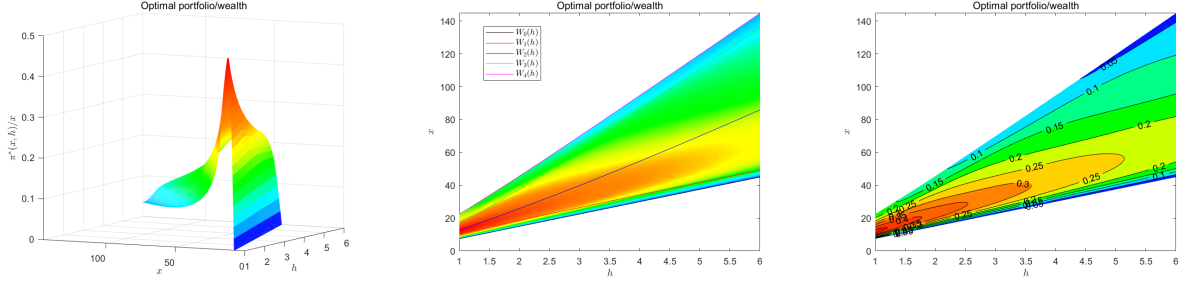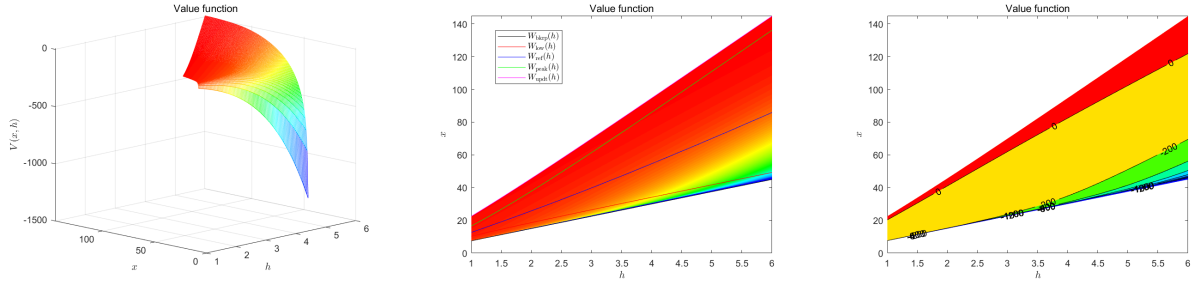
Figure 5: Optimal risky investment proportion with fixed parameters $\lambda = 0.3$, $\alpha = 0.7$, $\beta_1 = 1$, $\beta_2 = 2$, $r = 0.04$, $\mu = 0.12$, $\sigma = 0.3$. The left panel plots the optimal risky investment proportion as a function of wealth and consumption peak. The middle panel is the two-dimensional projection of the left panel with the boundaries shown. The right panel is the contour plot.

Figure 6: Value function with fixed parameters $\lambda = 0.3$, $\alpha = 0.7$, $\beta_1 = 1$, $\beta_2 = 2$, $r = 0.04$, $\mu = 0.12$, $\sigma = 0.3$. The left panel plots the value function as a function of wealth and consumption peak. The middle panel is the two-dimensional projection of the left panel with the boundaries shown. The right panel is the contour plot.

The value function shown in Figure 6, is increasing in wealth x and decreasing in habit h, which suggests that higher initial wealth and lower inherited past spending maximum result in higher optimal value for Problem (2.3). Meanwhile, for regions below $x = W_{\text{ref}}(h)$ where the risk aversion is low, especially for gloom region, the value function will fall dramatically due to a slight decrease in x or a slight increase in h. Nevertheless, for regions above $x = W_{\text{ref}}(h)$ where the risk aversion is high, the value function does not vary significantly as x and h vary. The above result indicates that one can live almost as pleasant as a millionaire with his initial wealth equal to $W_{\text{ref}}(h_0)$ where h_0 is his historical consumption peak. Another fact shown is that poor people are much more vulnerable to wealth shocks than the wealthy.

We are particularly interested in the decisions of consumption rate and risky investment proportion in terms of wealth, fixing a standard of living, i.e., the function $x \mapsto c^*(x, h)$ and $x \mapsto \pi^*(x, h)/x$, which we present in Figure 7. As can be seen, the sensitivity of the optimal consumption with respect to the variable x, or economically speaking, the marginal propensity to consume (MPC) out of wealth, generally decreases with a growing wealth. This is admitted in vast economic literature. However, a dedicated analysis (see Remarks 3-5) shows that while MPC out of wealth is indeed decreasing in lower part of the depression-recovery region and at the bliss curve, it is instead increasing in x in upper part of depression-recovery region. Besides, MPC out of wealth shrinks or swells by β_1/β_2 at $W_{\text{ref}}(h)$, which is itself an interesting fact[2]. This fact also indicates that the

---

[2]Different from other related papers, the marginal utility of consumption is continuous at the reference point
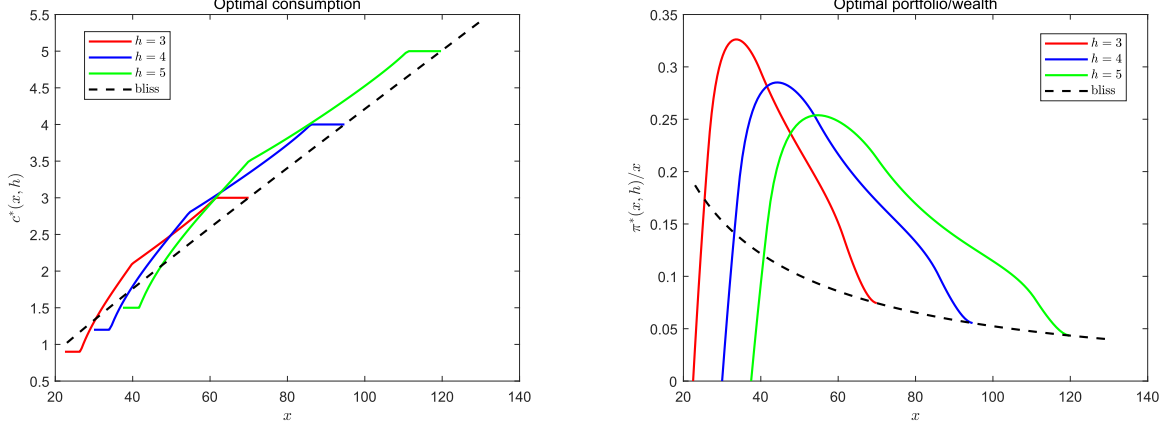
Figure 7: Optimal consumption and risky investment proportion in wealth with $h = 3, 4, 5$ and fixed parameters $\lambda = 0.3$, $\alpha = 0.7$, $\beta_1 = 1$, $\beta_2 = 2$, $r = 0.04$, $\mu = 0.12$, $\sigma = 0.3$. The bliss curve represents the optimal strategies when $x \geq W_{\text{updt}}(h)$, and h is updating on this curve.

MPC in recovery region can be globally lower or higher than that in depression region, and it is lower in our numerical result (left panel of Figure 7). See Figure 1 in the introduction for a more illustrative version of optimal consumption when fixing h.

Remark 3. Based on Theorem 4.4 and Lemma 4.5, we have for $W_{\text{low}}(h) < x \leq W_{\text{peak}}(h)$,

$$\frac{\partial c^*(x, h)}{\partial x} = \begin{cases} \frac{1}{\beta_1} \frac{1}{f_2(x,h)\tilde{V}_{yy}(f_2(x,h),h)}, & W_{\text{low}}(h) < x \leq W_{\text{ref}}(h), \\ \frac{1}{\beta_2} \frac{1}{f_3(x,h)\tilde{V}_{yy}(f_3(x,h),h)}, & W_{\text{ref}}(h) < x \leq W_{\text{peak}}(h). \end{cases}$$

$\frac{r}{k}(C_3(h) + C_4(h)) + \frac{1}{\gamma\beta_1} = \frac{r}{k}(C_5(h) + C_6(h)) + \frac{1}{\gamma\beta_2}$ leads to the continuity of $y\tilde{V}_{yy}(y, h)$ at $y = 1$. Hence $\frac{\partial c^*(x,h)}{\partial x}|_{x \to W_{\text{ref}}(h)^+} = \frac{\beta_1}{\beta_2} \frac{\partial c^*(x,h)}{\partial x}|_{x \to W_{\text{ref}}(h)^-}$, which indicates that the MPC out of wealth shrinks or swells by $\frac{\beta_1}{\beta_2}$ when exceeding $W_{\text{ref}}(h)$.

Remark 4. The bliss curve for optimal consumption is $c = W_{\text{updt}}^{-1}(x)$. Hence the bliss curve is concave if and only if $W_{\text{updt}}''(h) > 0$. Direct computation shows

$$W_{\text{updt}}''(h) = e^{-(1-\alpha)(q_2-1)\beta_2 h} \left[M_1 e^{-(\alpha-\lambda)(q_2-1)\beta_1 }h - M_2 \right]$$

where

$$M_1 = \frac{\frac{k}{\gamma^2\beta_1}(1-q_1)(q_2-1)^2(1-\alpha)^{q_2-1}\left[(1-\alpha)q_2\beta_2 + (\alpha-\lambda)(q_2-1)\beta_1\right]}{(1-\alpha)(q_2-q_1)\beta_2 + (\alpha-\lambda)(q_2-1)\beta_1} \left[(1-\alpha)\beta_2 + (\alpha-\lambda)\beta_1\right]^2,$$

$$M_2 = \frac{k}{\gamma^2} \frac{\beta_2 - \beta_1}{\beta_1\beta_2} \frac{1-q_1}{q_2-q_1} q_2(1-\alpha)^{q_2+1}(q_2-1)^2\beta_2^2.$$

Hence for $\beta_1 < \beta_2$, we have $M_1, M_2 > 0$ and thus there exists $\bar{h} = \frac{\ln(M_1) - \ln(M_2)}{(\alpha-\lambda)(q_2-1)\beta_1}$ such that the bliss curve is concave in x for $h \leq \bar{h}$. It implies that the MPC out of wealth decreases when $x \geq W_{\text{updt}}(h)$ and $h \leq \bar{h}$. With current parameters, the threshold $\bar{h}$ is approximately 6.6.

$c = \alpha h$, but we still document such an MPC shrink or MPC swell.

For $\beta_1 \geq \beta_2$, we have $M_1 > 0, M_2 \leq 0$ and thus the bliss curve is concave in x, which implies that the MPC out of wealth decreases when $x \geq W_{\text{updt}}(h)$.

Remark 5. With Lemma 4.2, Theorem 4.4 and Lemma 4.5, it can be shown that $\frac{\partial c^*(x,h)}{\partial x}$ is increasing (decreasing) in x if and only if $\pi^*(y, h)$ is increasing (decreasing) in y, where x and y are connected by $y = f(x, h)$. As for $e^{-(1-\alpha)\beta_2 h} \leq y < e^{(\alpha-\lambda)\beta_1 h}$,

$$\frac{\partial \pi^*(y,h)}{\partial y} = \frac{\mu - r}{\sigma^2} \frac{r}{k} \left\{ \begin{array}{l} C_3(h)(q_1 - 1)y^{q_1-2} + C_4(h)(q_2 - 1)y^{q_2-2}, \ 1 \leq y < e^{(\alpha-\lambda)\beta_1 h}, \\ C_5(h)(q_1 - 1)y^{q_1-2} + C_6(h)(q_2 - 1)y^{q_2-2}, \ e^{-(1-\alpha)\beta_2 h} \leq y < 1 \end{array} \right.$$

and

$$C_3(h)(q_1 - 1)e^{(q_1-2)(\alpha-\lambda)\beta_1 h} + C_4(h)(q_2 - 1)e^{(q_2-2)(\alpha-\lambda)\beta_1 h} = C_1(h)(q_1 - 1)e^{(q_1-2)(\alpha-\lambda)\beta_1 h} < 0,$$
$$C_3(h)(q_1 - 1) + C_4(h)(q_2 - 1) = C_5(h)(q_1 - 1) + C_6(h)(q_2 - 1),$$

we conclude that if

$$C_5(h)(q_1 - 1)e^{-(q_1-2)(1-\alpha)\beta_2 h} + C_6(h)(q_2 - 1)e^{-(q_2-2)(1-\alpha)\beta_2 h} > 0, \tag{5.1}$$

then there exists $\bar{x}(h) \in [W_{\text{low}}(h), W_{\text{peak}}(h)]$ such that the MPC out of wealth is decreasing for $x \in [W_{\text{low}}(h), \bar{x}(h)] \backslash \{W_{\text{ref}}(h)\}$ and increasing for $x \in [\bar{x}(h), W_{\text{peak}}(h)] \backslash \{W_{\text{ref}}(h)\}$ ($W_{\text{ref}}(h)$ needs to be excluded due to the MPC shrink or swell, see Remark 3). Specifically, $\bar{x}(h) = -\tilde{V}_y(\bar{y}(h), h)$ with $\bar{y}(h)$ given by

$$\bar{y}(h) = \left\{ \begin{array}{l} (-\frac{C_3(h)(q_1-1)}{C_4(h)(q_2-1)})^{\frac{1}{q_2-q_1}}, \text{ for } h \text{ s.t. } C_3(h)(q_1 - 1) + C_4(h)(q_2 - 1) > 0, \\ (-\frac{C_5(h)(q_1-1)}{C_6(h)(q_2-1)})^{\frac{1}{q_2-q_1}}, \text{ for } h \text{ s.t. } C_3(h)(q_1 - 1) + C_4(h)(q_2 - 1) \leq 0 \end{array} \right.$$

It is straight forward to verify numerically that (5.1) is satisfied for h in a reasonable range (say, $h \geq 0.2$ with the current parameters).

The threshold $\bar{x}(h)$ is above $W_{\text{ref}}(h)$ if and only if $C_3(h)(q_1-1)+C_4(h)(q_2-1) < 0$. We observe two cases: One case is that $\bar{x}(h)$ is always above $W_{\text{ref}}(h)$, which is the case for $\beta_1 \geq \beta_2$; the other case is that $\bar{x}(h)$ is above $W_{\text{ref}}(h)$ for h above certain threshold $\hat{h}$ and below $W_{\text{ref}}(h)$ for $h < \hat{h}$, which is the case for $\beta_1 < \beta_2$. Moreover, as the difference $\beta_2 - \beta_1$ decreases to zero, the threshold $\hat{h}$ increases to $+\infty$ and turns into the first case.

From the right panel of Figure 7, it is clear that for sampled h, risky investment proportion is a hump in variable x, and the peak is around $W_{\text{ref}}(h)$ (see also Figure 5) [3]. If we interpret the inverse of risky investment proportion as the so-called *implied relative risk aversion* (Jeon and Park (2020)), we find that it will be a smile in wealth: people with intermediate wealth level have lowest risk aversion, and hence have the highest risk tolerance. People with either very low or very high level of wealth are much more risk averse. This effect comes intuitively from our model settings. Poor people need to make their deposit above $W_{\text{bkrp}}(h)$ in order to satisfy the lowest consumption constraint, hence are very sensitive to risk. It is reasonable for them to keep most part of their wealth in safe assets. Rich people, on the other hand, have already been satisfied by the current level of consumption (or even continuously consume more and more) and they tend to avoid the risk of consumption declining to less than reference αh. A more illustrative version of optimal risky investment proportion when fixing h can be found in Figure 1.

---

[3]Numerical analysis shows that risky investment proportion is indeed a hump in x for h above a relatively small level. For h below that level (i.e, if h is extremely small), there is an apparent increase on the right end of the hump.
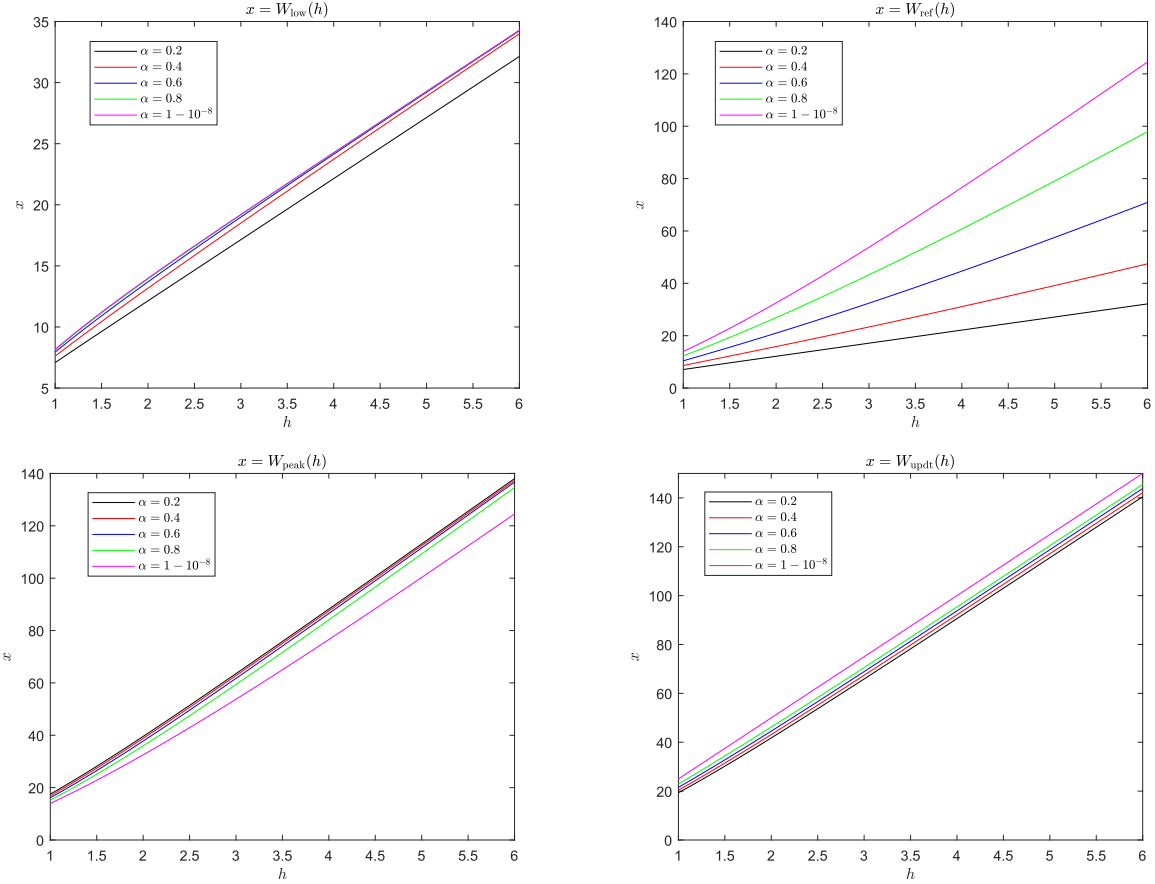
Figure 8: With fixed parameters $\lambda = 0.2$, $\beta_1 = 1$, $\beta_2 = 2$, $r = 0.04$, $\mu = 0.12$, $\sigma = 0.3$, the four boundaries varies with different α.

6 Sensitivity Analysis

6.1 Impact of λ and α on Thresholds

The parameter α determines the reference point αh where risk aversion changes. When α equals λ, our model reduces to the model without the risk aversion related reference point and the boundary $x = W_{\mathrm{ref}}(h)$ coincides with $x = W_{\mathrm{low}}(h)$. When α approaches 1 from below, the reference point approaches the running maximum and the boundary $x = W_{\mathrm{ref}}(h)$ coincides with $x = W_{\mathrm{peak}}(h)$. To figure out the consequences of the ratio α on different boundaries, we fix $\lambda = 0.2$, $\beta_1 = 1$, $\beta_2 = 2$, $r = 0.04$, $\mu = 0.12$, $\sigma = 0.3$, let α vary from λ to $1 - 10^{-8}$ and present the boundaries separately (we do not consider the boundary $W_{\mathrm{bkrp}}(h)$ since it does not depend on α). As shown in Figure 8, the boundary $x = W_{\mathrm{peak}}(h)$ moves downward as α increases, while the other three boundaries $x = W_{\mathrm{low}}(h), W_{\mathrm{ref}}(h)$ and $W_{\mathrm{updt}}(h)$ have the tendency to move upward as α increases. The above phenomenon illustrates that the criterion for the investor to be able to consume at the historical running maximum level lowers as α increases while other three criteria to enter a higher consumption level would generally raise up as α increases. The change of α is most influential on $x = W_{\mathrm{ref}}(h)$, which is the boundary of whether to consume more than αh and an important boundary in our analysis of both optimal portfolio and value function.

Another parameter of interest in our model is λ, which reflects the degree of the drawdown
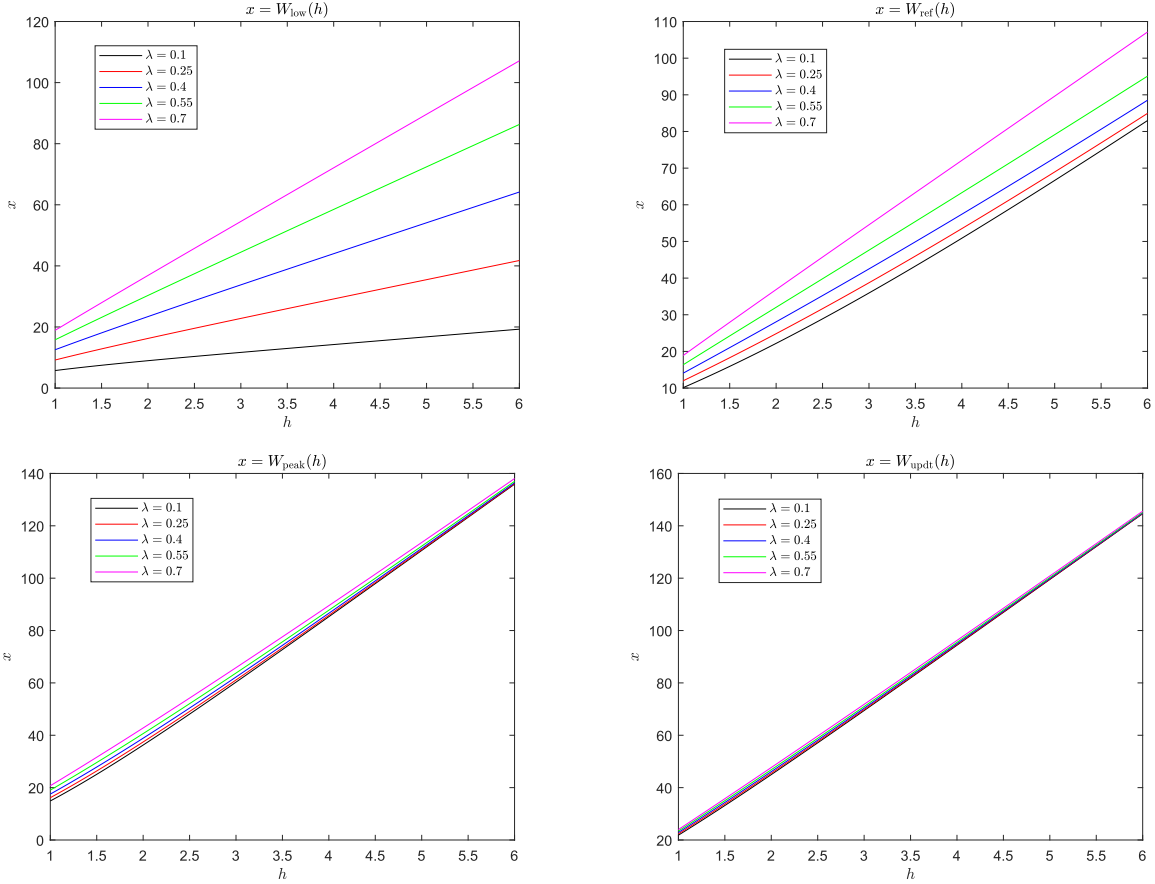
Figure 9: With fixed parameters $\alpha = 0.7$, $\beta_1 = 1$, $\beta_2 = 2$, $r = 0.04$, $\mu = 0.12$, $\sigma = 0.3$, the four boundaries varies with different λ.

constraint on consumption. The consumption level is allowed to be relatively lower with smaller λ. When λ equals 0, our model reduces to the model without drawdown constraint and the boundary $x = W_{\text{bkrp}}(h)$ coincides with the h axis. When λ equals α, our model reduces to the model without the risk aversion related reference point and the boundary $x = W_{\text{ref}}(h)$ coincides with $x = W_{\text{low}}(h)$. We fix $\alpha = 0.7$, $\beta_1 = 1$, $\beta_2 = 2$, $r = 0.04$, $\mu = 0.12$, $\sigma = 0.3$, let λ vary below α and graph the boundaries separately (boundary $x = W_{\text{bkrp}}(h)$ is not shown here since it is simply linear) in order to illustrate the impact of λ on the boundaries. As shown in Figure 9, all four thresholds are higher when λ increases. Among the four boundaries, λ is most influential on the boundary $x = W_{\text{low}}(h)$, which suggests that, with larger λ, more wealth is needed for getting rid of gloom. However, λ has a negligible effect on $x = W_{\text{peak}}(h)$ and $x = W_{\text{updt}}(h)$, indicating that whether to consume at the historical level and whether to update the historical level have almost no reliance on the degree of drawdown constraint.

6.2 Discussion for β_1 and β_2

In this subsection, we focus on the sensitivity analysis of the parameters β_1 and β_2, which are the risk aversion coefficients below and above the reference αh. The difference $\beta_2 - \beta_1$ reflects the magnitude of the risk aversion change over the reference. We mainly illustrate the results with $\beta_1 \leq \beta_2$. The differences in results with $\beta_1 \geq \beta_2$ are briefly discussed in 6.2.2. Two limiting cases
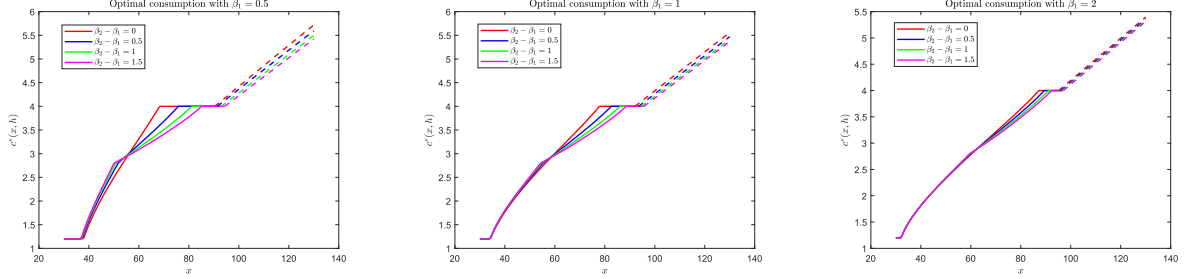
23

Figure 10: With fixed parameters $\alpha = 0.7$, $r = 0.04$, $\mu = 0.12$, $\sigma = 0.3$, $h = 4$, impact of $\beta_2 - \beta_1$ on optimal consumption with fixed β_1.
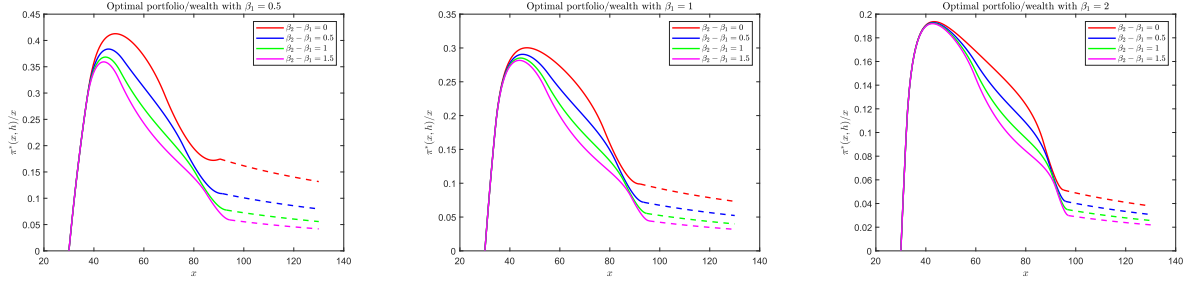

Figure 11: With fixed parameters $\alpha = 0.7$, $r = 0.04$, $\mu = 0.12$, $\sigma = 0.3$, $h = 4$, impact of $\beta_2 - \beta_1$ on optimal risky investment proportion with fixed β_1.

of interest are investigated in 6.2.3.

6.2.1 $\beta_1 \geq \beta_2$

In this part, we analyze the influence of $\beta_2 - \beta_1$ when $\beta_1 \geq \beta_2$.

To investigate the influence of $\beta_2 - \beta_1$, we have to fix either β_1 or β_2. One approach is fixing β_1, then $\beta_2 - \beta_1$ is the increase in risk aversion when the consumption exceeds the reference αh.

We first investigate the influence on optimal strategies. When $\beta_2 - \beta_1$ is enlarges, the main influence on consumption is the decelerating of consumption increase once across the boundary $x = W_{\text{ref}}(h)$ (see Figure 10), which delays the arrival of consumption peak $x = W_{\text{peak}}(h)$. The decelerating effect in consumption increase can be explained by the conservative consumption behavior due to higher risk aversion above the reference. Meanwhile, change of β_2 even influence the consumption in the region $x < W_{\text{ref}}(h)$, where only β_1 seems to be relative. This can be regarded as an important risk allocation behavior: people suppress the consumption when he can tolerate less risks to compensate the consumption elsewhere.

We now consider the risky assets allocation. As $\beta_2 - \beta_1$ enlarges, there is a decrease in risky investment proportion once over certain thresholds around $W_{\text{low}}(h)$. The amount of decrease varies with different wealth levels x and initial risk aversion β_1 (see Figure 11 for more details). It is reasonably expected that risky investment proportion should decrease due to higher risk aversion above the reference. Again, the decrease is not constrained to those wealth levels in recovery region ($x > W_{\text{ref}}(h)$). It instead occurs prior to the threshold $W_{\text{ref}}(h)$.

We also investigate the influence of $\beta_2 - \beta_1$ on the wealth thresholds. As is shown in Figure 12, the increase of $\beta_2 - \beta_1$ leads to higher threshold $W_{\text{peak}}(h)$, especially for small h. The influence on

24

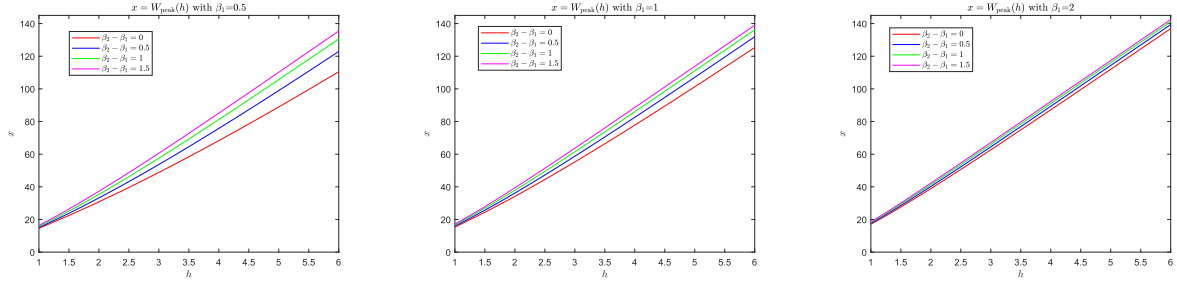other thresholds is negligible.

Figure 12: With fixed parameters $\alpha = 0.7$, $r = 0.04$, $\mu = 0.12$, $\sigma = 0.3$, impact of $\beta_2 - \beta_1$ on $x = W_{\text{peak}}(h)$ with fixed β_1.

Another approach to examine the effect of risk aversion change $\beta_2 - \beta_1$ is fixing β_2 and viewing $\beta_2 - \beta_1$ as the decrease in risk aversion when consumption falls below the reference αh.

As $\beta_2 - \beta_1$ enlarges, the decrease of consumption decelerates above $W_{\text{ref}}(h)$ but accelerates below $W_{\text{ref}}(h)$ (see Figure 13). The acceleration can be reasonably explained by aggressive consumption behavior due to lower risk aversion below the reference. Again, the effect of β_1 is not limited to $x < W_{\text{ref}}(h)$.

The risky investment proportion, on the other hand, sees a significant increase when $\beta_2 - \beta_1$ increases for wealth in the depression and recovery regions $[W_{\text{low}}(h), W_{\text{peak}}(h)]$, especially around $W_{\text{ref}}(h)$ (see Figure 14). This is due to lower risk aversion below the reference.

However, the change in $\beta_2 - \beta_1$ for fixed β_2 does not have a significant impact on wealth thresholds. We merely observe a decrease in $W_{\text{ref}}(h)$ for small values of β_2 as $\beta_2 - \beta_1$ enlarges (see Figure 15).
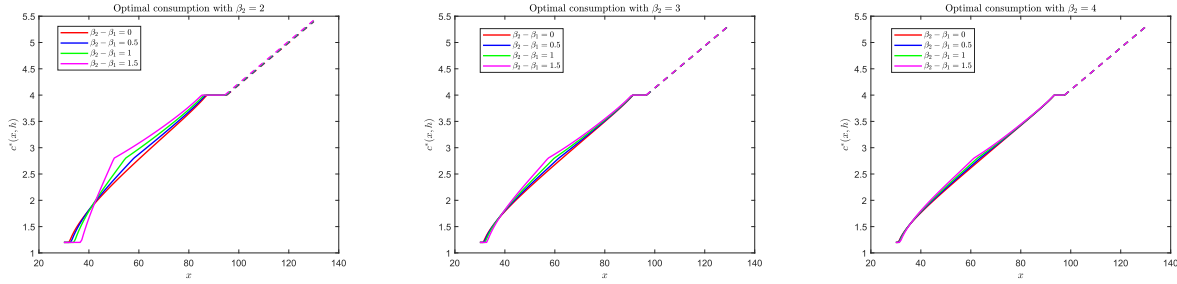

Figure 13: With fixed parameters $\alpha = 0.7$, $r = 0.04$, $\mu = 0.12$, $\sigma = 0.3$, $h = 4$, impact of $\beta_2 - \beta_1$ on optimal consumption with fixed β_2.

Remark 6. The influence of $\beta_2 - \beta_1$ can only be studied fixing either β_1 or β_2. The sensitivity analyses for the influence of $\beta_2 - \beta_1$ on the optimal controls as well as the thresholds are established for relatively small β_1 or β_2 (the fixed one). As numerical results illustrate, When the fixed β_i is sufficiently large, the optimal controls and the thresholds are nearly not affected by $\beta_2 - \beta_1$. This phenomenon can already be well observed in the right panels in Figure 10, Figure 12, Figure 13 and Figure 15.

A special case in our model is $\beta_2 - \beta_1 = 0$, which suggests that risk aversion does not change over the reference αh. In this case, W_{ref} just exists symbolically but has no economical significance.
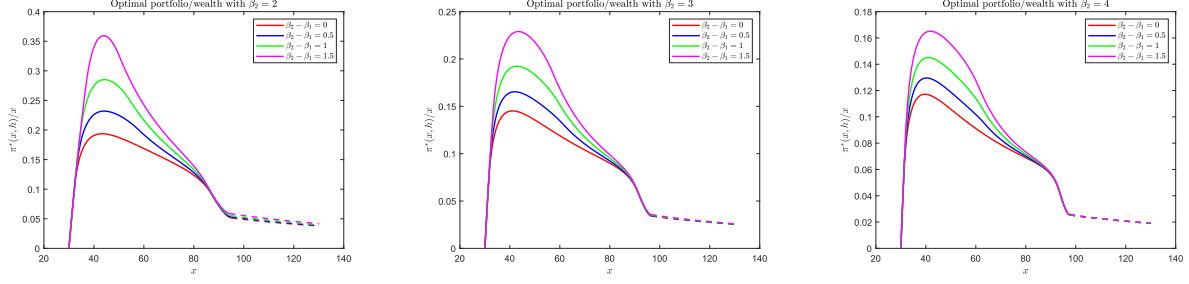
Figure 14: With fixed parameters $\alpha = 0.7$, $r = 0.04$, $\mu = 0.12$, $\sigma = 0.3$, $h = 4$, impact of $\beta_2 - \beta_1$ on optimal risky investment proportion with fixed β_2.
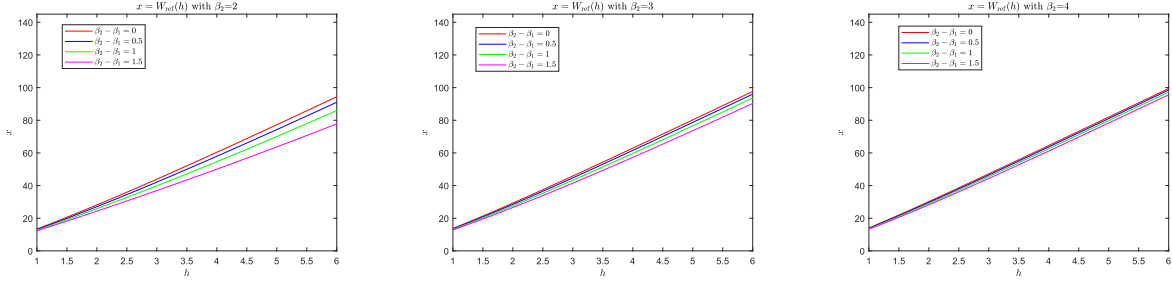

Figure 15: With fixed parameters $\alpha = 0.7$, $r = 0.04$, $\mu = 0.12$, $\sigma = 0.3$, impact of $\beta_2 - \beta_1$ on $x = W_{\mathrm{ref}}(h)$ with fixed β_2.

The depression region ($x \in [W_{\mathrm{low}}(h), W_{\mathrm{ref}}(h)]$) and recovery region ($x \in [W_{\mathrm{ref}}(h), W_{\mathrm{peak}}(h)]$) share strategies and one region would merge into another.

6.2.2 $\beta_1 > \beta_2$

In this part, we consider the case $\beta_1 > \beta_2$ and briefly discuss on the impact of $\beta_1 - \beta_2$ on optimal strategies and boundaries. The impact for $\beta_1 > \beta_2$ mainly corresponds to that for $\beta_1 < \beta_2$ and can be well interpreted. However, it is beyond our expectation to observe a new peak around $W_{\mathrm{peak}}(h)$ in risky investment proportion that overtakes the peak around $W_{\mathrm{ref}}(h)$.

The optimal consumption for $\beta_1 > \beta_2$ is still non-decreasing in x. The main difference occurs between $[W_{\mathrm{low}}(h), W_{\mathrm{peak}}(h)]$. As shown in Figure 16, when $\beta_1 > \beta_2$, the risk aversion is lower in the recovery region and the MPC out of wealth is thus generally higher in the recovery region, which results in an upward turn at $W_{\mathrm{ref}}(h)$ (it is a downward turn when $\beta_1 < \beta_2$). When fixing β_1, as $\beta_1 - \beta_2$ enlarges, the increase of consumption accelerates over $W_{\mathrm{ref}}(h)$, bringing the arrival of consumption peak $W_{\mathrm{peak}}(h)$ forward. When fixing β_2, as $\beta_1 - \beta_2$ enlarges, the decrease of consumption accelerates above $W_{\mathrm{ref}}(h)$ and decelerates below $W_{\mathrm{ref}}(h)$.

As shown in Figure 17, when $\beta_1 - \beta_2$ enlarges, the optimal investment proportion sees an apparent increase over certain threshold around $W_{\mathrm{low}}(h)$ for fixed β_1 while there is an apparent decrease for wealth in depression and recovery regions for fixed β_2. Besides, we observe a new peak around $W_{\mathrm{peak}}(h)$ that overtakes the peak around $W_{\mathrm{ref}}(h)$ (see the red line in Figure 17). It suggests that high proportion of risky investment is recommended for wealth in the recovery region, especially around $W_{\mathrm{ref}}(h)$ and $W_{\mathrm{peak}}(h)$.

The impact of increasing $\beta_1 - \beta_2$ on the boundaries coincides with the case $\beta_1 < \beta_2$: lower level
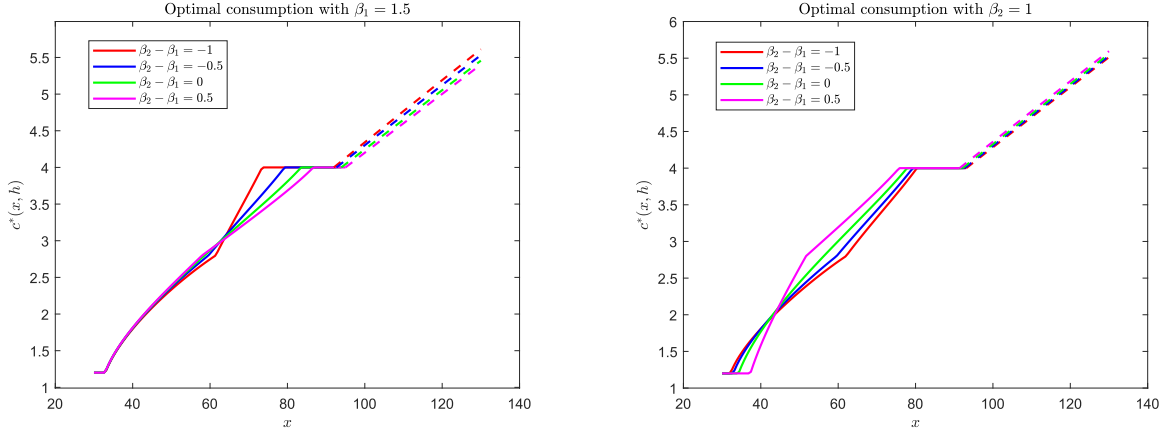
Figure 16: With fixed parameters $\alpha = 0.7$, $r = 0.04$, $\mu = 0.12$, $\sigma = 0.3$, $h = 4$, impact of $\beta_2 - \beta_1$ on optimal consumption with fixed $\beta_1 = 1.5$ (left) or fixed $\beta_2 = 1$ (right).
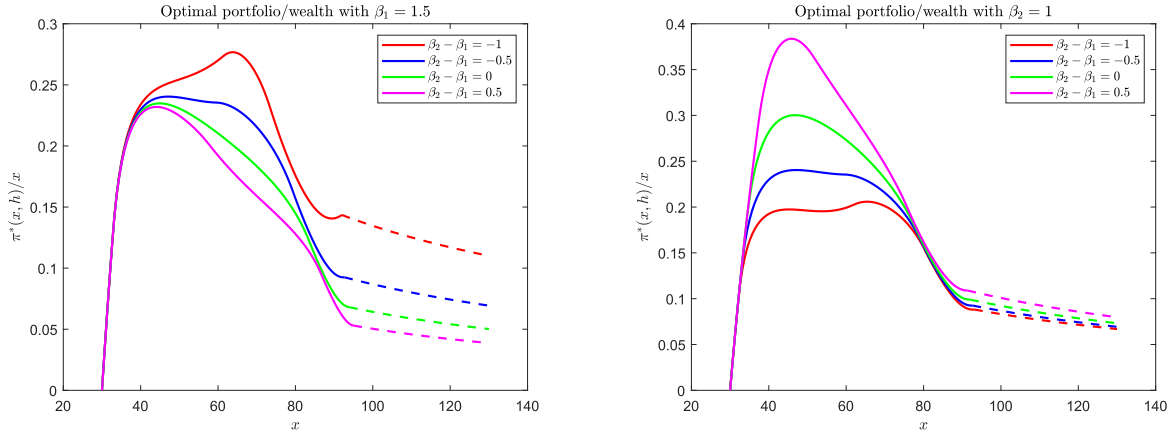

Figure 17: With fixed parameters $\alpha = 0.7$, $r = 0.04$, $\mu = 0.12$, $\sigma = 0.3$, $h = 4$, impact of $\beta_2 - \beta_1$ on optimal risky investment proportion with fixed $\beta_1 = 1.5$ (left) or fixed $\beta_2 = 1$ (right).

of $W_{\text{peak}}(h)$ with fixed β_1, higher level of $W_{\text{ref}}(h)$ with fixed β_2 and negligible effect on others.

6.2.3 Limiting cases

In this subsection, we briefly discuss two limiting cases when one of the β_i takes extreme value: fix β_1 and let $\beta_2 \to 0$; fix β_2 and let $\beta_1 \to 0$.

Fixing β_1 and letting $\beta_2 \to 0$ indicates that the agent becomes risk neutral when $c > \alpha h$. Since $C_1(h) \to \infty$ as $\beta_2 \to 0$, it is surprising to see that even the lowest constrained threshold $W_{\text{low}}(h)$ tends to infinity. The limiting consumption curve is a horizontal line $c \equiv \lambda h$. It suggests that the agent always consumes at the lowest constrained level $\frac{\lambda}{r}h$. He might be saving money from consuming as less as possible in order to aggressively invest in risk asset once he reaches $W_{\text{ref}}(h)$. However, the day never comes because $W_{\text{ref}}(h)$ tends to infinity as $\beta_2 \to 0$.

Fixing β_2 and letting $\beta_1 \to 0$, however, has completely different consequences. This limit corresponds to the case that the agent becomes risk neutral when $c < \alpha h$. A dedicated analysis (see Remark 7) shows that all the thresholds have finite limits as $\beta_1 \to 0$. In particular, $W_{\text{low}}(h)$

and $W_{\text{ref}}(h)$ share the same limit. It indicates that the agent never consumes between the lowest constrained level $\frac{\lambda h}{r}$ and the reference level αh in this limiting case. The result is similar to that of an S-shaped utility in Li et al. (2022) where the agent is risk seeking below the reference. However, in S-shaped utility, the optimal consumption jumps from 0 to a level that is strictly higher than reference point αh (see (3.25) in Li et al. (2022)), while our limiting optimal consumption jumps from the lowest level to exactly the reference point.

Remark 7. From (3.13)$\sim$(3.18), we have, as $\beta_1 \to 0$,

$$
\begin{aligned}
C_4(h) \to& -\infty, \\
C_6(h) \to& -\frac{k}{\gamma^2 \beta_2}\frac{1-q_1}{q_2-q_1} + \frac{k}{\gamma^2}\frac{1-q_1}{q_2-q_1}(\alpha-\lambda)(q_2-1)h, \\
C_8(h) \to& \frac{k}{\gamma^2 \beta_2}\frac{1-q_1}{q_2-q_1}(e^{(1-\alpha)(q_2-1)\beta_2 h}-1) + \frac{k}{\gamma^2}\frac{1-q_1}{q_2-q_1}(\alpha-\lambda)(q_2-1)h, \\
C_7(h) \to&(1-\alpha)^{q_2-q_1-1}(\alpha-\lambda)\frac{k}{\gamma^2}\frac{1-q_1}{q_2-q_1}\frac{q_2-1}{q_2-q_1}e^{-(1-\alpha)(q_2-q_1)\beta_2 h} \\
&+(1-\alpha)^{q_2-q_1}\frac{k}{\gamma^2 \beta_2}\frac{q_2-1}{q_2-q_1}e^{-(1-\alpha)(1-q_1)\beta_2 h}, \\
C_5(h) \to&(1-\alpha)^{q_2-q_1-1}(\alpha-\lambda)\frac{k}{\gamma^2}\frac{1-q_1}{q_2-q_1}\frac{q_2-1}{q_2-q_1}e^{-(1-\alpha)(q_2-q_1)\beta_2 h} \\
&+\left[(1-\alpha)^{q_2-q_1}-1\right]\frac{k}{\gamma^2 \beta_2}\frac{q_2-1}{q_2-q_1}e^{-(1-\alpha)(1-q_1)\beta_2 h}, \\
C_3(h) \to& -\infty, \\
C_1(h) \to&(1-\alpha)^{q_2-q_1-1}(\alpha-\lambda)\frac{k}{\gamma^2}\frac{1-q_1}{q_2-q_1}\frac{q_2-1}{q_2-q_1}e^{-(1-\alpha)(q_2-q_1)\beta_2 h} \\
&+(1-\alpha)^{q_2-q_1}\frac{k}{\gamma^2 \beta_2}\frac{q_2-1}{q_2-q_1}e^{-(1-\alpha)(1-q_1)\beta_2 h} \\
&+\frac{k}{\gamma^2 \beta_2}\frac{q_2-1}{q_2-q_1}+\frac{k}{\gamma^2}\frac{q_2-1}{q_2-q_1}(\alpha-\lambda)(1-q_1)h.
\end{aligned}
$$

Denote the finite limits of $C_i(h), i = 1, 5, 6, 7, 8$ by $C_i^L(h)$. Then we have, as $\beta_1 \to 0$,

$$
\begin{aligned}
W_{\text{low}}(h) \to& -C_1^L(h)q_1 + \frac{\lambda h}{\gamma}, \\
W_{\text{peak}}(h) \to& -C_5^L(h)q_1 e^{(1-\alpha)(1-q_1)\beta_2 h} - C_6^L(h)q_2 e^{-(1-\alpha)(q_2-1)\beta_2 h} - \frac{k}{\gamma^2 \beta_2} + \frac{h}{\gamma}, \\
W_{\text{updt}}(h) \to& -C_7^L(h)q_1(1-\alpha)^{q_1-1}e^{(1-\alpha)(1-q_1)\beta_2 h} - C_8^L(h)q_2(1-\alpha)^{q_2-1}e^{-(1-\alpha)(q_2-1)\beta_2 h} + \frac{h}{\gamma}.
\end{aligned}
$$

To show that $W_{\text{ref}}(h)$ tends to the same limit $-C_1^L(h)q_1 + \frac{\lambda h}{\gamma}$ as $W_{\text{low}}(h)$, we prove $W_{\text{ref}}(h) - W_{\text{low}}(h) \to 0$. In fact,

$$
\begin{aligned}
W_{\text{ref}}(h) - W_{\text{low}}(h) =& q_1 C_1(h) \left[e^{-(\alpha-\lambda)(1-q_1)\beta_1 h} - 1 \right] + \frac{(\alpha-\lambda)h}{\gamma} \\
& + \frac{k}{\gamma^2 \beta_1} \left[q_1 \frac{q_2 - 1}{q_2 - q_1} e^{(\alpha-\lambda)(1-q_1)\beta_1 h} + q_2 \frac{1 - q_1}{q_2 - q_1} e^{-(\alpha-\lambda)(q_2-1)\beta_1 h} - 1 \right] \\
\rightarrow & \frac{(\alpha-\lambda)h}{\gamma} + \frac{k}{\gamma^2} \left[q_1 \frac{q_2 - 1}{q_2 - q_1} (\alpha - \lambda)(1 - q_1)h - q_2 \frac{1 - q_1}{q_2 - q_1} (\alpha - \lambda)(q_2 - 1)h \right] \\
=& 0.
\end{aligned}
$$

7   Conclusion

We establish a new theoretical model focusing on the risky investment and consumption behavior of a sophisticated decision maker. We solve the optimal consumption and investment problem which maximizes the expected total discounted utility with running maximum related reference point and drawdown constraint. Mathematical analysis and computation illustrate that the optimal consumption and investment policy are of semi-explicit forms with five important thresholds classifying different ranks of people. Theoretical and numerical analysis of the solution and sensitivity analysis of the parameters are conducted as well. The results are of economic significance in the following aspects: the MPC out of wealth is generally decreasing but increasing with certain intermediate wealth levels, and it jumps inversely proportional to the risk aversion at the reference point; both DRRA and IRRA are possible and the implied relative risk aversion is roughly a smile in wealth; wealth shocks are more influential on the welfare of the poorer people. As a special feature of our model, risk aversion change results in significant changes in optimal strategies and the impact of local risk aversion change turns out to be global.

Data availability statement. Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

A   Proof of the Verification Theorem

Proof of Theorem 4.1. Define

$$
\hat{H}_t(y) \triangleq h_0 \vee \frac{1}{(1-\alpha)\beta_2} \ln\left(\frac{1-\alpha}{\inf\limits_{s \le t} Y_s(y)} \right).
$$

Then, for any $(x_0, h_0) \in \mathcal{C}$ and any $y > 0$, we have

$$
\mathbb{E}_{x_0, h_0} \int_0^\infty e^{-\gamma t} U(c_t, h_t) dt = \mathbb{E}_{x_0, h_0} \int_0^\infty e^{-\gamma t} \big(U(c_t, h_t) - Y_t(y)c_t \big) dt + y \mathbb{E}_{x_0, h_0} \int_0^\infty c_t M_t dt
$$

$$
\begin{aligned}
&\leq\ \mathbb{E}_{x_0,h_0} \int_0^\infty e^{-\gamma t} \tilde{U}\big(Y_t(y), H_t^\dagger(y)\big) dt + y x_0 \\
&=\ \mathbb{E}_{x_0,h_0} \int_0^\infty e^{-\gamma t} \tilde{U}\big(Y_t(y), \hat{H}_t(y)\big) dt + y x_0 \\
&=\ \tilde{V}(y, h_0) + y x_0, \quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\text{(A.1)}
\end{aligned}
$$

where the second, the third and the last line hold thanks to Lemma B.1, Lemma B.2 and Lemma B.4. And equality holds with $c_t = c^*(Y_t(y), H_t^\dagger(y))$ and $y = y^*$.

Using the explicit expressions of $c^*(\cdot, \cdot)$, $Y_t(\cdot)$ and $\hat{H}_t(\cdot)$, we know that $c^*\big(Y_t(y), \hat{H}_t(y)\big)$ is strictly decreasing in y with $\lim_{y \to 0^+} c^*\big(Y_t(y), \hat{H}_t(y)\big) = \infty$ and $\lim_{y \to \infty} c^*\big(Y_t(y), \hat{H}_t(y)\big) = \lambda h_0$. As such, there exists a unique y to solve

$$
\mathbb{E}_{x_0,h_0} \int_0^\infty c^*\big(Y_t(y), \hat{H}_t(y)\big) M_t dt = x_0.
$$

Hence, we deduce from Lemma B.2 that Eq.(4.1) has a unique solution y^*. Using (A.1) yields

$$
\inf_{y>0} \big\{ \tilde{V}(y, h_0) + y x_0 \big\} = \sup_{(c,\pi) \in \mathcal{A}} \mathbb{E}_{x_0,h_0} \int_0^\infty e^{-\gamma t} U(c_t, h_t) dt.
$$

To make the left hand side in which $\tilde{V}(\cdot, \cdot)$ is the solution of (3.9) equal the value function $V(x_0, h_0)$ which satisfies (3.1), $\tilde{V}(\cdot, \cdot)$ must satisfy (3.4). Recall that (3.4) leads to (3.9) for $c = c^*(y, h)$ and $\pi = \pi^*(y, h)$ where $c^*(\cdot, \cdot)$ is given by (3.6) which leads to (4.2) when restricted to $\mathcal{C}_d$ and $\pi^*(\cdot, \cdot)$ is given by (3.5). Plugging the expression of $\tilde{V}(y, h)$ and $q_i^2 - q_i = \frac{r}{k}$, $i = 1, 2$ into (3.5) leads to (4.3).

Thus, the dual of $\tilde{V}(\cdot, \cdot)$ given by (3.10) is actually the value function of (2.3) and that $\mathbb{E}_{x_0,h_0} \int_0^\infty e^{-\gamma t} U(c_t, h_t) dt$ attains its maximum at (c^*, π^*) given in Theorem 4.1. $\qquad\square$

B Auxiliary Lemmas for Theorem 4.1 and Their Proofs

The following four lemmas are needed in proving Theorem 4.1.

Lemma B.1. *The inequality holds in (A.1) and it becomes equality with $c_t = c^*\big(Y_t(y), H_t^\dagger(y)\big)$ and $y = y^*$.*

Lemma B.2. *For any $y > 0$ and any $t > 0$, we have $H_t^\dagger(y) = \hat{H}_t(y)$.*

Lemma B.3 (Transversality Condition)**.** *For any $y > 0$,*

$$
\lim_{T \to \infty} \mathbb{E}_{x_0,h_0} \Big[e^{-\gamma T} \tilde{V}\big(Y_T(y), \hat{H}_T(y)\big) \Big] = 0.
$$

Lemma B.4.

$$
\tilde{V}(y, h_0) = \mathbb{E}_{x_0,h_0} \int_0^\infty e^{-\gamma t} \tilde{U}\big(Y_t(y), \hat{H}_t(y)\big) dt.
$$

As the proofs of the first two lemmas are similar to the proofs of Lemma 5.2 and Lemma 5.3 in Deng et al. (2022), we omit them here and only prove Lemmas B.3 and B.4.

Proof of Lemma B.3. Based on the definition of $Y_t(\cdot)$ and $\hat{H}_t(\cdot)$, we have

$$\lim_{T\to\infty} Y_T(y) = \lim_{T\to\infty} y e^{-\frac{(\mu-r)^2}{2\sigma^2}T - \frac{\mu-r}{\sigma}B_T} = 0, \ a.s., \tag{B.1}$$

$$\lim_{T\to\infty} \hat{H}_T(y) = \lim_{T\to\infty} h_0 \vee \frac{1}{(1-\alpha)\beta_2} \ln(\frac{1-\alpha}{\inf\limits_{s\leq T} Y_s(y)}) = \infty, \ a.s.. \tag{B.2}$$

Using the expression of $\tilde{V}(\cdot,\cdot)$ yields

$$\lim_{T\to\infty} \mathbb{E}_{x_0,h_0}\left[e^{-\gamma T}\tilde{V}\big(Y_T(y), \hat{H}_T(y)\big)\right]$$

$$= \lim_{T\to\infty} \mathbb{E}_{x_0,h_0}\Bigg\{ e^{-\gamma T}\Big[C_7\big(\hat{H}_T(y)\big)Y_T(y)^{q_1} + C_8\big(\hat{H}_T(y)\big)Y_T(y)^{q_2} - \frac{1}{\gamma}\hat{H}_T(y)Y_T(y)$$

$$+ \frac{1}{\gamma\beta_2}\big(1 - e^{-(1-\alpha)\beta_2\hat{H}_T(y)}\big)\Big] 1_{\left\{(1-\alpha)e^{-(1-\alpha)\beta_2\hat{H}_T(y)}\leq Y_T(y)<e^{-(1-\alpha)\beta_2\hat{H}_T(y)}\right\}}$$

$$+ e^{-\gamma T}\Big[C_5\big(\hat{H}_T(y)\big)Y_T(y)^{q_1} + C_6\big(\hat{H}_T(y)\big)Y_T(y)^{q_2} + \frac{k}{\gamma^2\beta_2}Y_T(y)$$

$$+ \frac{1}{\gamma\beta_2}\big(1 - Y_T(y) + Y_T(y)\ln\big(Y_T(y)\big)\big) - \frac{1}{\gamma}\alpha\hat{H}_T(y)Y_T(y)\Big] 1_{\left\{e^{-(1-\alpha)\beta_2\hat{H}_T(y)}\leq Y_T(y)<1\right\}}\Bigg\}. \tag{B.3}$$

By Proposition 3.2, we have
$C_7\big(\hat{H}_T(y)\big) = \mathcal{O}(e^{-(1-\alpha)(1-q_1)\beta_2\hat{H}_T(y)})$, a.s. as $T\to\infty$, as such,

$$\mathbb{E}_{x_0,h_0}e^{-\gamma T}C_7\big(\hat{H}_T(y)\big)Y_T(y)^{q_1}1_{\left\{(1-\alpha)e^{-(1-\alpha)\beta_2\hat{H}_T(y)}\leq Y_T(y)<e^{-(1-\alpha)\beta_2\hat{H}_T(y)}\right\}}$$

$$= \mathcal{O}\left(e^{-\gamma T}\mathbb{E}_{x_0,h_0}e^{-(1-\alpha)(1-q_1)\beta_2\hat{H}_T(y)}Y_T(y)^{q_1}1_{\left\{(1-\alpha)e^{-(1-\alpha)\beta_2\hat{H}_T(y)}\leq Y_T(y)<e^{-(1-\alpha)\beta_2\hat{H}_T(y)}\right\}}\right)$$

$$= \mathcal{O}\left(e^{-\gamma T}\mathbb{E}_{x_0,h_0}Y_T(y)^{1-q_1+q_1}\right)$$

$$= \mathcal{O}\left(e^{-\gamma T}\mathbb{E}_{x_0,h_0}e^{-\frac{(\mu-r)^2}{2\sigma^2}T - \frac{\mu-r}{\sigma}B_T}\right)$$

$$= \mathcal{O}\left(e^{-\gamma T}\right), \quad \text{as } T\to\infty,$$

then

$$\lim_{T\to\infty} \mathbb{E}_{x_0,h_0}e^{-\gamma T}C_7\big(\hat{H}_T(y)\big)Y_T(y)^{q_1}1_{\left\{(1-\alpha)e^{-(1-\alpha)\beta_2\hat{H}_T(y)}\leq Y_T(y)<e^{-(1-\alpha)\beta_2\hat{H}_T(y)}\right\}} = 0. \tag{B.4}$$

Similarly,

$$\mathbb{E}_{x_0,h_0}e^{-\gamma T}C_8\big(\hat{H}_T(y)\big)Y_T(y)^{q_2}1_{\left\{(1-\alpha)e^{-(1-\alpha)\beta_2\hat{H}_T(y)}\leq Y_T(y)<e^{-(1-\alpha)\beta_2\hat{H}_T(y)}\right\}}$$

$$= \mathcal{O}\left(e^{-\gamma T}\mathbb{E}_{x_0,h_0}e^{(1-\alpha)(q_2-1)\beta_2\hat{H}_T(y)}Y_T(y)^{q_2}1_{\left\{(1-\alpha)e^{-(1-\alpha)\beta_2\hat{H}_T(y)}\leq Y_T(y)<e^{-(1-\alpha)\beta_2\hat{H}_T(y)}\right\}}\right)$$

$$= \mathcal{O}\left(e^{-\gamma T}\mathbb{E}_{x_0,h_0}Y_T(y)^{-(q_2-1)+q_2}\right)$$

$$= \mathcal{O}\left(e^{-\gamma T}\mathbb{E}_{x_0,h_0}e^{-\frac{(\mu-r)^2}{2\sigma^2}T - \frac{\mu-r}{\sigma}B_T}\right) = \mathcal{O}\left(e^{-\gamma T}\right), \quad \text{as } T\to\infty,$$

then

$$\lim_{T\to\infty} \mathbb{E}_{x_0,h_0} e^{-\gamma T} C_8\big(\hat{H}_T(y)\big) Y_T(y)^{q_2} 1_{\left\{(1-\alpha)e^{-(1-\alpha)\beta_2 \hat{H}_T(y)} \leq Y_T(y) < e^{-(1-\alpha)\beta_2 \hat{H}_T(y)}\right\}} = 0. \tag{B.5}$$

Using (B.2), as $T \to \infty$,

$$e^{-\gamma T} \frac{1}{\gamma \beta_2} \big(1 - e^{-(1-\alpha)\beta_2 \hat{H}_T(y)}\big) 1_{\left\{(1-\alpha)e^{-(1-\alpha)\beta_2 \hat{H}_T(y)} \leq Y_T(y) < e^{-(1-\alpha)\beta_2 \hat{H}_T(y)}\right\}} = \mathcal{O}\big(e^{-\gamma T}\big), \ a.s.,$$

from which we get

$$\lim_{T\to\infty} \mathbb{E}_{x_0,h_0} e^{-\gamma T} \frac{1}{\gamma \beta_2} \big(1 - e^{-(1-\alpha)\beta_2 \hat{H}_T(y)}\big) 1_{\left\{(1-\alpha)e^{-(1-\alpha)\beta_2 \hat{H}_T(y)} \leq Y_T(y) < e^{-(1-\alpha)\beta_2 \hat{H}_T(y)}\right\}} = 0. \tag{B.6}$$

By the same way as in deriving (B.4), we have

$$\lim_{T\to\infty} \mathbb{E}_{x_0,h_0} e^{-\gamma T} C_5\big(\hat{H}_T(y)\big) Y_T(y)^{q_1} 1_{\left\{e^{-(1-\alpha)\beta_2 \hat{H}_T(y)} \leq Y_T(y) < 1\right\}} = 0. \tag{B.7}$$

By Proposition 3.2, we have $C_6\big(\hat{H}_T(y)\big) = \mathcal{O}\big(1\big)$, a.s. as $T \to \infty$. The facts (B.1) and $q_2 > 1$ yield

$$\lim_{T\to\infty} \mathbb{E}_{x_0,h_0} e^{-\gamma T} C_6\big(\hat{H}_T(y)\big) Y_T(y)^{q_2} 1_{\left\{e^{-(1-\alpha)\beta_2 \hat{H}_T(y)} \leq Y_T(y) < 1\right\}} = 0. \tag{B.8}$$

and

$$\lim_{T\to\infty} \mathbb{E}_{x_0,h_0} e^{-\gamma T} \frac{k}{\gamma^2 \beta_2} Y_T(y) 1_{\left\{e^{-(1-\alpha)\beta_2 \hat{H}_T(y)} \leq Y_T(y) < 1\right\}} = 0. \tag{B.9}$$

Using (B.1) again and the fact $\lim_{y\to 0^+} (1 - y + y \ln y) = 1$, we obtain

$$\lim_{T\to\infty} \mathbb{E}_{x_0,h_0} e^{-\gamma T} \frac{1}{\gamma \beta_2} \Big[1 - Y_T(y) + Y_T(y) \ln \big(Y_T(y)\big)\Big] 1_{\left\{e^{-(1-\alpha)\beta_2 \hat{H}_T(y)} \leq Y_T(y) < 1\right\}} = 0. \tag{B.10}$$

Based on Girsanov's theorem, we have

$$e^{-\gamma T} \mathbb{E}_{x_0,h_0} \hat{H}_T(y) Y_T(y)$$
$$= \mathcal{O}\bigg(e^{-\gamma T} \mathbb{E}_{x_0,h_0}\Big[\sup_{s\leq T}\big\{\frac{(\mu-r)^2}{2\sigma^2}s + \frac{\mu-r}{\sigma}B_s\big\}e^{-\frac{(\mu-r)^2}{2\sigma^2}T - \frac{\mu-r}{\sigma}B_T}\Big]\bigg)$$
$$= \mathcal{O}\bigg(\frac{\mu-r}{\sigma} e^{-\gamma T}\big\{\sqrt{\frac{T}{2\pi}}e^{-\frac{(\mu-r)^2}{8\sigma^2}T} - \frac{\mu-r}{2\sigma}T\Phi(-\frac{\mu-r}{2\sigma}\sqrt{T})$$
$$+ \frac{\sigma}{\mu-r}\big[\Phi(\frac{\mu-r}{2\sigma}\sqrt{T}) - \Phi(-\frac{\mu-r}{2\sigma}\sqrt{T})\big]\big\}\bigg).$$

It follows that

$$\lim_{T\to\infty} \mathbb{E}_{x_0,h_0} e^{-\gamma T} \hat{H}_T(y) Y_T(y) = 0. \tag{B.11}$$

Thus, using (B.3)-(B.11), we have

$$\lim_{T\to\infty} \mathbb{E}_{x_0,h_0}\Big[e^{-\gamma T}\tilde{V}\big(Y_T(y), \hat{H}_T(y)\big)\Big] = 0.$$

$\square$

Proof of Lemma B.4. Applying Eq. (3.9) and Itô's rule, we obtain

$$
\begin{aligned}
d\Big\{e^{-\gamma t}\tilde{V}\big[Y_t(y),\hat{H}_t(y)\big]\Big\} = & -e^{-\gamma t}\tilde{U}\big[Y_t(y),\hat{H}_t(y)\big]dt \\
& -\frac{\mu-r}{\sigma}e^{-\gamma t}\tilde{V}_y\big[Y_t(y),\hat{H}_t(y)\big]Y_t(y)dB_t \\
& +e^{-\gamma t}\tilde{V}_h\big[Y_t(y),\hat{H}_t(y)\big]d\hat{H}_t(y).
\end{aligned}
\tag{B.12}
$$

Define the stopping times: $\forall n \geq 1$,

$$
\tau_n \triangleq \inf\Big\{t \geq 0\Big|Y_t(y) \geq n \text{ or } \hat{H}_t(y) \geq \frac{1}{(1-\alpha)\beta_2}\ln\big[(1-\alpha)n\big]\Big\}.
$$

It follows that $\lim_{n\to\infty}\tau_n = \infty$ and for $\forall\, n > \frac{1}{1-\alpha}e^{(1-\alpha)\beta_2 h_0}$, $\forall\, \kappa \geq 1$ and $\forall\, T > 0$,

$$
\begin{aligned}
\mathbb{E}_{x_0,h_0}1_{\{\tau_n \leq T\}} \leq & \,\mathbb{P}_{x_0,h_0}\Big(\big\{\sup_{t\in[0,T]}Y_t(y) \geq n\big\}\bigcup\big\{\sup_{t\in[0,T]}\hat{H}_t(y) \geq \frac{1}{(1-\alpha)\beta_2}\ln\big[(1-\alpha)n\big]\big\}\Big) \\
\leq & \,\mathbb{P}_{x_0,h_0}\big(\sup_{t\in[0,T]}Y_t(y) \geq n\big) + \mathbb{P}_{x_0,h_0}\big(\inf_{t\in[0,T]}Y_t(y) \leq \frac{1}{n}\big) \\
= & \,\mathbb{P}_{x_0,h_0}\big(\sup_{t\in[0,T]}Y_t(y) \geq n\big) + \mathbb{P}_{x_0,h_0}\big(\sup_{t\in[0,T]}Y_t(y)^{-1} \geq n\big) \\
\leq & \,n^{-2\kappa}\mathbb{E}_{x_0,h_0}\sup_{t\in[0,T]}Y_t(y)^{2\kappa} + n^{-2\kappa}\mathbb{E}_{x_0,h_0}\sup_{t\in[0,T]}Y_t(y)^{-2\kappa} \\
= & \,\mathcal{O}\big(n^{-2\kappa}(1+y^{2\kappa})e^{MT}\big)
\end{aligned}
\tag{B.13}
$$

for some constants M.

Integrating (B.12) from 0 to $T\wedge\tau_n$ and taking expectation on both sides, we obtain

$$
\begin{aligned}
\tilde{V}(y,h_0) = & \,\mathbb{E}_{x_0,h_0}\Big[e^{-\gamma T\wedge\tau_n}\tilde{V}\big(Y_{T\wedge\tau_n}(y),\hat{H}_{T\wedge\tau_n}(y)\big)\Big] \\
& + \mathbb{E}_{x_0,h_0}\int_0^{T\wedge\tau_n}e^{-\gamma t}\tilde{U}\big(Y_t(y),\hat{H}_t(y)\big)dt \\
& + \mathbb{E}_{x_0,h_0}\int_0^{T\wedge\tau_n}\frac{\mu-r}{\sigma}e^{-\gamma t}\tilde{V}_y\big(Y_t(y),\hat{H}_t(y)\big)Y_t(y)dB_t \\
& - \mathbb{E}_{x_0,h_0}\int_0^{T\wedge\tau_n}e^{-\gamma t}\tilde{V}_h\big(Y_t(y),\hat{H}_t(y)\big)d\hat{H}_t(y).
\end{aligned}
\tag{B.14}
$$

The four terms on the right hand side of (B.14) can be evaluated as follows:

The first term in (B.14) can be split into two parts:

$$
\begin{aligned}
& \mathbb{E}_{x_0,h_0}\Big[e^{-\gamma T\wedge\tau_n}\tilde{V}\big(Y_{T\wedge\tau_n}(y),\hat{H}_{T\wedge\tau_n}(y)\big)\Big] \\
& = \mathbb{E}_{x_0,h_0}\Big[e^{-\gamma T}\tilde{V}\big(Y_T(y),\hat{H}_T(y)\big)1_{\{T\leq\tau_n\}}\Big] \\
& + \mathbb{E}_{x_0,h_0}\Big[e^{-\gamma\tau_n}\tilde{V}\big(Y_{\tau_n}(y),\hat{H}_{\tau_n}(y)\big)1_{\{T>\tau_n\}}\Big].
\end{aligned}
\tag{B.15}
$$

Based on Lemma B.3, as $n \uparrow \infty$, the first term in (B.15) tends to $\mathbb{E}_{x_0,h_0}\left[e^{-\gamma T}\tilde{V}\left(Y_T(y), \hat{H}_T(y)\right)\right]$, which converges to zero as $T \uparrow \infty$. To handle the second term in (B.15), we first observe from the definition of τ_n that either

$$Y_{\tau_n}(y) \leq n, \ \hat{H}_{\tau_n}(y) = \frac{1}{(1-\alpha)\beta_2} \ln\left((1-\alpha)n\right)$$

or

$$Y_{\tau_n}(y) = n, \ \hat{H}_{\tau_n}(y) \leq \frac{1}{(1-\alpha)\beta_2} \ln\left((1-\alpha)n\right)$$

holds. In addition, it follows from the fact $\left(Y_{\tau_n}(y), \hat{H}_{\tau_n}(y)\right) \in \mathcal{C}_d$ that either

$$\frac{1}{n} \leq Y_{\tau_n}(y) \leq n, \ \hat{H}_{\tau_n}(y) = \frac{1}{(1-\alpha)\beta_2} \ln\left((1-\alpha)n\right)$$

or

$$Y_{\tau_n}(y) = n, \ 0 < \hat{H}_{\tau_n}(y) \leq \frac{1}{(1-\alpha)\beta_2} \ln\left((1-\alpha)n\right)$$

holds.

Applying Proposition 3.2 with (3.10), we obtain the order estimate of $\tilde{V}\left(Y_{\tau_n}(y), \hat{H}_{\tau_n}(y)\right)$ as follows:

For $\frac{1}{n} \leq Y_{\tau_n}(y) \leq n, \ \hat{H}_{\tau_n}(y) = \frac{1}{(1-\alpha)\beta_2} \ln\left[(1-\alpha)n\right]$, we have, for sufficiently large n,

$$\hat{H}_{\tau_n}(y) = \mathcal{O}\left(\ln n\right), \ e^{\hat{H}_{\tau_n}(y)} = \mathcal{O}\left(n^{\frac{1}{(1-\alpha)\beta_2}}\right).$$

If $(1-\alpha)e^{-(1-\alpha)\beta_2 \hat{H}_{\tau_n}(y)} \leq Y_{\tau_n}(y) < e^{-(1-\alpha)\beta_2 \hat{H}_{\tau_n}(y)}$, then we have, for sufficiently large n,

$$\tilde{V}\left(Y_{\tau_n}(y), \hat{H}_{\tau_n}(y)\right) = \mathcal{O}\left(n^{-q_1}\right).$$

If $e^{-(1-\alpha)\beta_2 \hat{H}_{\tau_n}(y)} \leq Y_{\tau_n}(y) < 1$, then we have, for sufficiently large n,

$$\tilde{V}\left(Y_{\tau_n}(y), \hat{H}_{\tau_n}(y)\right) = \mathcal{O}\left(n^{-q_1}\right).$$

If $1 \leq Y_{\tau_n}(y) < e^{(\alpha-\lambda)\beta_1 \hat{H}_{\tau_n}(y)}$, then we have, for sufficiently large n,

$$\tilde{V}\left(Y_{\tau_n}(y), \hat{H}_{\tau_n}(y)\right) = \mathcal{O}\left(n^{q_2}\right).$$

If $e^{(\alpha-\lambda)\beta_1 \hat{H}_{\tau_n}(y)} \leq Y_{\tau_n}(y)$, then we have, for sufficiently large n,

$$\tilde{V}\left(Y_{\tau_n}(y), \hat{H}_{\tau_n}(y)\right) = \mathcal{O}\left(n^{\frac{(\alpha-\lambda)\beta_1}{(1-\alpha)\beta_2}}\right).$$

For $Y_{\tau_n}(y) = n, \ 0 < \hat{H}_{\tau_n}(y) \leq \frac{1}{(1-\alpha)\beta_2} \ln\left[(1-\alpha)n\right]$, we have, for sufficiently large n, either

$$1 \leq Y_{\tau_n}(y) < e^{(\alpha-\lambda)\beta_1 \hat{H}_{\tau_n}(y)}$$

or

$$e^{(\alpha-\lambda)\beta_1 \hat{H}_{\tau_n}(y)} \leq Y_{\tau_n}(y).$$

If $1 \leq Y_{\tau_n}(y) < e^{(\alpha-\lambda)\beta_1 \hat{H}_{\tau_n}(y)}$, then we have, for sufficiently large n,

$$\tilde{V}\big(Y_{\tau_n}(y), \hat{H}_{\tau_n}(y)\big) = \mathcal{O}\big(n^{q_2}\big).$$

If $e^{(\alpha-\lambda)\beta_1 \hat{H}_{\tau_n}(y)} \leq Y_{\tau_n}(y)$, then we have, for sufficiently large n,

$$\tilde{V}\big(Y_{\tau_n}(y), \hat{H}_{\tau_n}(y)\big) = \mathcal{O}\big(n^{q_1 + (1-q_1)\frac{(\alpha-\lambda)\beta_1}{(1-\alpha)\beta_2}}\big).$$

In summary, we have $\tilde{V}\big(Y_{\tau_n}(y), \hat{H}_{\tau_n}(y)\big) = \mathcal{O}\big(n^{(-q_1)\vee q_2 \vee (1-q_1)\frac{(\alpha-\lambda)\beta_1}{(1-\alpha)\beta_2}}\big)$. Applying (B.13) with $\kappa \geq \frac{1}{2}\big((-q_1) \vee q_2 \vee (1-q_1)\frac{(\alpha-\lambda)\beta_1}{(1-\alpha)\beta_2}\big)$, we deduce that the second term in (B.15) converges to zero as $n \uparrow \infty$. Hence the first term in (B.14) tends to zero by first letting $n \uparrow \infty$ and then $T \uparrow \infty$. Because τ_n tends to ∞ as $n \to \infty$, based on monotone convergence theorem, we obtain that the second term in (B.14) tends to $\mathbb{E}_{x_0,h_0} \int_0^\infty e^{-\gamma t} \tilde{U}\big(Y_t(y), \hat{H}_t(y)\big) dt$, as $n \to \infty$ and $T \to \infty$. The third term in (B.14) vanishes because the integral is a martingale. If $\hat{H}_t(y)$ strictly increases, then $Y_t(y)$ must be strictly decreasing, hence $\big(Y_t(y), \hat{H}_t(y)\big)$ is on the boundary. Using the boundary condition, we have $\tilde{V}_h\big(Y_t(y), \hat{H}_t(y)\big) = 0$, as such, the last term in (B.14) vanishes. Thus, the proof follows. $\qquad\square$

C   Proofs of Other Results in Section 4

Proof of Lemma 4.2. If $e^{(\alpha-\lambda)\beta_1 h} \leq y$, then

$$\tilde{V}_{yy}(y,h) = C_1(h)\frac{r}{k}y^{q_1-2} + C_2(h)\frac{r}{k}y^{q_2-2}.$$

As $C_1(h) > 0$ and $C_2(h) = 0$, we have $\tilde{V}_{yy}(y,h) > 0$ for $e^{(\alpha-\lambda)\beta_1 h} \leq y$.

If $1 \leq y < e^{(\alpha-\lambda)\beta_1 h}$, then

$$y\tilde{V}_{yy}(y,h) = C_3(h)\frac{r}{k}y^{q_1-1} + C_4(h)\frac{r}{k}y^{q_2-1} + \frac{1}{\gamma\beta_1}.$$

Let $\psi(y) = y\tilde{V}_{yy}(y,h)$, then

$$\psi'(y) = C_3(h)\frac{r}{k}(q_1-1)y^{q_1-2} + C_4(h)\frac{r}{k}(q_2-1)y^{q_2-2}.$$

Noting that $C_4(h) < 0$, $\psi(y)$ is either increasing, decreasing or first increasing then decreasing, we only need to show $\psi(1) > 0$ and $\psi(e^{(\alpha-\lambda)\beta_1 h}) > 0$. Precisely, using $C_7(h) > 0$,

$$\begin{aligned}
\psi(1) =& C_3(h)\frac{r}{k} + C_4(h)\frac{r}{k} + \frac{1}{\gamma\beta_1} \\
=& \frac{r}{k}C_7(h) + \frac{1}{\gamma\beta_1}\frac{1-q_1}{q_2-q_1}\Big[1 - e^{-(\alpha-\lambda)(q_2-1)\beta_1 h}\Big] \\
& + \frac{1}{\gamma\beta_2}\frac{q_2-1}{q_2-q_1}\Big[1 - e^{-(1-\alpha)(1-q_1)\beta_2 h}\Big] \\
>& 0
\end{aligned}$$

and

35

$$\psi(e^{(\alpha-\lambda)\beta_1 h}) = C_3(h)\frac{r}{k}e^{(q_1-1)(\alpha-\lambda)\beta_1 h} + C_4(h)\frac{r}{k}e^{(q_2-1)(\alpha-\lambda)\beta_1 h} + \frac{1}{\gamma\beta_1}$$

$$= \frac{r}{k}C_7(h)e^{(q_1-1)(\alpha-\lambda)\beta_1 h} + \frac{1}{\gamma\beta_1}\frac{q_2-1}{q_2-q_1}\left[1 - e^{-(\alpha-\lambda)(1-q_1)\beta_1 h}\right]$$

$$+ \frac{1}{\gamma\beta_2}\frac{q_2-1}{q_2-q_1}\left[1 - e^{-(1-\alpha)(1-q_1)\beta_2 h}\right]e^{(q_2-1)(\alpha-\lambda)\beta_1 h}$$

$$> 0.$$

Thus $\tilde{V}_{yy}(y,h) > 0$ for $1 \leq y < e^{(\alpha-\lambda)\beta_1 h}$.

If $e^{-(1-\alpha)\beta_2 h} \leq y < 1$, then

$$y\tilde{V}_{yy}(y,h) = C_5(h)\frac{r}{k}y^{q_1-1} + C_6(h)\frac{r}{k}y^{q_2-1} + \frac{1}{\gamma\beta_2}$$

$$= \left[C_5(h)\frac{r}{k}y^{q_1-1} + \frac{1}{\gamma\beta_2}\frac{q_2-1}{q_2-q_1}\right] + \left[C_6(h)\frac{r}{k}y^{q_2-1} + \frac{1}{\gamma\beta_2}\frac{1-q_1}{q_2-q_1}\right].$$

For any fixed $h > 0$, if $C_5(h) \geq 0$, then $C_5(h)\frac{r}{k}y^{q_1-1} + \frac{1}{\gamma\beta_2}\frac{q_2-1}{q_2-q_1} > 0$; If $C_5(h) < 0$, then $C_5(h)\frac{r}{k}y^{q_1-1} + \frac{1}{\gamma\beta_2}\frac{q_2-1}{q_2-q_1}$ is increasing in y and

$$C_5(h)\frac{r}{k}y^{q_1-1} + \frac{1}{\gamma\beta_2}\frac{q_2-1}{q_2-q_1}$$

$$\geq C_5(h)\frac{r}{k}e^{(1-\alpha)(1-q_1)\beta_2 h} + \frac{1}{\gamma\beta_2}\frac{q_2-1}{q_2-q_1}$$

$$= \frac{(1-\alpha)^{q_2-q_1}\frac{1}{\gamma}\frac{1-q_1}{q_2-q_1}(\alpha-\lambda)(q_2-1)}{(1-\alpha)(q_2-q_1)\beta_2 + (\alpha-\lambda)(q_2-1)\beta_1}e^{-[(1-\alpha)(q_2-1)\beta_2+(\alpha-\lambda)(q_2-1)\beta_1]h}$$

$$+ (1-\alpha)^{q_2-q_1}\frac{1}{\gamma\beta_2}\frac{q_2-1}{q_2-q_1}$$

$$> 0.$$

Similarly, we have

$$C_6(h)\frac{r}{k}y^{q_2-1} + \frac{1}{\gamma\beta_2}\frac{1-q_1}{q_2-q_1} > 0.$$

It follows that $\tilde{V}_{yy}(y,h) > 0$ for $e^{-(1-\alpha)\beta_2 h} \leq y < 1$.

Finally, If $(1-\alpha)e^{-(1-\alpha)\beta_2 h} \leq y < e^{-(1-\alpha)\beta_2 h}$, then

$$\tilde{V}_{yy}(y,h) = C_7(h)\frac{r}{k}y^{q_1-2} + C_8(h)\frac{r}{k}y^{q_2-2}.$$

As $C_7(h) > 0$ and

$$C_8(h) = \frac{k}{\gamma^2\beta_1}\frac{1-q_1}{q_2-q_1}\left[1 - e^{-(\alpha-\lambda)(q_2-1)\beta_1 h}\right] + \frac{k}{\gamma^2\beta_2}\frac{1-q_1}{q_2-q_1}\left[e^{(1-\alpha)(q_2-1)\beta_2 h} - 1\right] > 0,$$

we have $\tilde{V}_{yy}(y,h) > 0$ for $(1-\alpha)e^{-(1-\alpha)\beta_2 h} \leq y < e^{-(1-\alpha)\beta_2 h}$. Thus the proof is complete. $\qquad\square$

Proof of Lemma 4.6. By Lemma 4.5, f is C^1 and continuous at the boundaries, as such, using the fact that $C_i(h)$, $1 \leq i \leq 8$, are C^1, we conclude that c^*_{primal} and π^*_{primal} given in Theorem 4.4 are locally Lipschitz on $\mathcal{C}$.

Now to prove that π^*_{primal} is Lipschitz, we only need to show that $\frac{\partial \pi^*_{\text{primal}}}{\partial x}$ and $\frac{\partial \pi^*_{\text{primal}}}{\partial h}$ are both bounded.

Step 1: $\frac{\partial \pi^*_{\text{primal}}}{\partial x}$ is bounded:

By definition of π^*_{primal} given in Theorem 4.4, we have

$$
\frac{\partial \pi^*_{\text{primal}}}{\partial x} = \frac{\mu - r}{\sigma^2}\frac{r}{k}
\begin{cases}
C_1(h)(q_1-1)\big[f_1(x,h)\big]^{q_1-2}\frac{\partial f_1(x,h)}{\partial x} \\
\quad + C_2(h)(q_2-1)\big[f_1(x,h)\big]^{q_2-2}\frac{\partial f_1(x,h)}{\partial x}, & \frac{\lambda h}{\gamma} \leq x \leq W_{\text{low}}(h), \\
C_3(h)(q_1-1)\big[f_2(x,h)\big]^{q_1-2}\frac{\partial f_2(x,h)}{\partial x} \\
\quad + C_4(h)(q_2-1)\big[f_2(x,h)\big]^{q_2-2}\frac{\partial f_2(x,h)}{\partial x}, & W_{\text{low}}(h) < x \leq W_{\text{ref}}(h), \\
C_5(h)(q_1-1)\big[f_3(x,h)\big]^{q_1-2}\frac{\partial f_3(x,h)}{\partial x} \\
\quad + C_6(h)(q_2-1)\big[f_3(x,h)\big]^{q_2-2}\frac{\partial f_3(x,h)}{\partial x}, & W_{\text{ref}}(h) < x \leq W_{\text{peak}}(h), \\
C_7(h)(q_1-1)\big[f_4(x,h)\big]^{q_1-2}\frac{\partial f_4(x,h)}{\partial x} \\
\quad + C_8(h)(q_2-1)\big[f_4(x,h)\big]^{q_2-2}\frac{\partial f_4(x,h)}{\partial x}, & W_{\text{peak}}(h) < x \leq W_{\text{updt}}(h).
\end{cases}
$$

Differentiating (4.6), we obtain

$$
\begin{aligned}
1 &= -C_1(h)q_1(q_1-1)\big[f_1(x,h)\big]^{q_1-2}\frac{\partial f_1(x,h)}{\partial x} \\
&\quad - C_2(h)q_2(q_2-1)\big[f_1(x,h)\big]^{q_2-2}\frac{\partial f_1(x,h)}{\partial x}.
\end{aligned}
\tag{C.1}
$$

It follows from (4.15) and (C.1) that, for $\frac{\lambda h}{\gamma} \leq x \leq W_{\text{low}}(h)$,

$$
\begin{aligned}
\frac{\partial \pi^*_{\text{primal}}}{\partial x} &= \frac{\mu - r}{\sigma^2}\left[(1-q_2) + \frac{r}{k}(q_1-q_2)C_1(h)\big[f_1(x,h)\big]^{q_1-2}\frac{\partial f_1(x,h)}{\partial x}\right] \\
&= \frac{\mu - r}{\sigma^2}\left[(1-q_2) + \frac{A_1(x,h)}{B_1(x,h)}\right],
\end{aligned}
$$

where

$$
\begin{aligned}
A_1(x,h) &= \frac{r}{k}(q_1-q_2)C_1(h)\left[f_1(x,h)\right]^{q_1-2}, \\
B_1(x,h) &= \frac{r}{k}\left\{-C_1(h)\big[f_1(x,h)\big]^{q_1-2} - C_2(h)\big[f_1(x,h)\big]^{q_2-2}\right\}.
\end{aligned}
$$

As $C_2(h) = 0$, $\frac{\partial \pi^*_{\text{primal}}}{\partial x}$ is constant for $\frac{\lambda h}{\gamma} \leq x \leq W_{\text{low}}(h)$. Differentiating (4.8), we obtain

$$
\begin{aligned}
1 &= -C_3(h)q_1(q_1-1)\big[f_2(x,h)\big]^{q_1-2}\frac{\partial f_2(x,h)}{\partial x} \\
&\quad - C_4(h)q_2(q_2-1)\big[f_2(x,h)\big]^{q_2-2}\frac{\partial f_2(x,h)}{\partial x} - \frac{1}{\gamma\beta_1 f_2(x,h)}\frac{\partial f_2(x,h)}{\partial x}.
\end{aligned}
\tag{C.2}
$$

Using (4.15) and (C.2), we have, for $W_{\text{low}}(h) \leq x \leq W_{\text{ref}}(h)$,

$$
\begin{aligned}
\frac{\partial \pi^*_{\text{primal}}}{\partial x} &= \frac{\mu - r}{\sigma^2} \left\{ (1 - q_2) + \frac{r}{k}(q_1 - q_2)C_3(h)\left[f_2(x,h)\right]^{q_1 - 2}\frac{\partial f_2(x,h)}{\partial x} \right. \\
&\qquad \left. + (1 - q_2)\frac{1}{\gamma\beta_1 f_2(x,h)}\frac{\partial f_2(x,h)}{\partial x} \right\} \\
&= \frac{\mu - r}{\sigma^2}\left[(1 - q_2) + \frac{A_2(x,h)}{B_2(x,h)} \right],
\end{aligned}
$$

where

$$
\begin{aligned}
A_2(x,h) &= \frac{r}{k}(q_1 - q_2)C_3(h)\left[f_2(x,h)\right]^{q_1 - 1} + (1 - q_2)\frac{1}{\gamma\beta_1}, \\
B_2(x,h) &= -\frac{r}{k}C_3(h)\left[f_2(x,h)\right]^{q_1 - 1} - \frac{r}{k}C_4(h)\left[f_2(x,h)\right]^{q_2 - 1} - \frac{1}{\gamma\beta_1}.
\end{aligned}
$$

As $e^{(\alpha - \lambda)\beta_1 h} > f_2(x,h) \geq 1$ for $W_{\text{low}}(h) \leq x \leq W_{\text{ref}}(h)$ and $C_3(h) = \mathcal{O}(1)$, there exists a constant $\bar{A}_2$ such that $|A_2(x,h)| \leq \bar{A}_2$ for $W_{\text{low}}(h) \leq x \leq W_{\text{ref}}(h)$. As $\tilde{V}_{yy}(y,h) > 0$, we know that $B_2(x,h) < 0$ for any $h \geq \bar{h}$. Using $C_4 < 0$ and $q_1 < 0 < 1 < q_2$ implies that $\frac{r}{k}C_3(h)y^{q_1 - 1} + \frac{r}{k}C_4(h)y^{q_2 - 1} + \frac{1}{\gamma\beta_1}$ as a function of y is either decreasing, or increasing, or first increasing then decreasing. Hence for $W_{\text{low}}(h) \leq x \leq W_{\text{ref}}(h)$,

$$
\begin{aligned}
-B_2(x,h) &\geq \min\left\{ \frac{r}{k}C_3(h) + \frac{r}{k}C_4(h) + \frac{1}{\gamma\beta_1}, \frac{r}{k}C_3(h)e^{-(1-q_1)(\alpha - \lambda)\beta_1 h} \right. \\
&\qquad \left. + \frac{r}{k}C_4(h)e^{(q_2 - 1)(\alpha - \lambda)\beta_1 h} + \frac{1}{\gamma\beta_1} \right\} \tag{C.3} \\
&> 0.
\end{aligned}
$$

Plugging the expressions of $C_3(h)$ and $C_4(h)$ in (3.13) and (3.17) into (C.3), we obtain that the expression in (C.3) is continuous in h with a limit $\frac{1}{\gamma\beta_1}\frac{q_2 - 1}{q_2 - q_1} > 0$ as $h \to \infty$. Hence there exists a constant $\bar{B}_2 > 0$ such that $-B_2(x,h) > \bar{B}_2 > 0$, i.e., $\frac{\partial \pi^*_{\text{primal}}}{\partial x}$ is bounded for $W_{\text{low}}(h) \leq x \leq W_{\text{ref}}(h)$. Similarly, $\frac{\partial \pi^*_{\text{primal}}}{\partial x}$ is also bounded for the rest two regions $W_{\text{ref}}(h) \leq x \leq W_{\text{peak}}(h)$ and $W_{\text{peak}}(h) \leq x \leq W_{\text{updt}}(h)$.

Step 2: $\frac{\partial \pi^*_{\text{primal}}}{\partial h}$ is bounded:

By definition of π^*_{primal}, denote $c_{11} \triangleq \frac{\mu - r}{\sigma^2}\frac{r}{k}$, we obtain

$$
\frac{\partial \pi^*_{\text{primal}}}{\partial h} = c_{11}
\begin{cases}
C_1(h)(q_1 - 1)\left[f_1(x,h)\right]^{q_1 - 2}\frac{\partial f_1(x,h)}{\partial h} + C_1'(h)\left[f_1(x,h)\right]^{q_1 - 1} \\
\quad + C_2(h)(q_2 - 1)\left[f_1(x,h)\right]^{q_2 - 2}\frac{\partial f_1(x,h)}{\partial h} + C_2'(h)\left[f_1(x,h)\right]^{q_2 - 1}, & \frac{\lambda h}{\gamma} \leq x \leq W_{\text{low}}(h), \\
C_3(h)(q_1 - 1)\left[f_2(x,h)\right]^{q_1 - 2}\frac{\partial f_2(x,h)}{\partial h} + C_3'(h)\left[f_2(x,h)\right]^{q_1 - 1} \\
\quad + C_4(h)(q_2 - 1)\left[f_2(x,h)\right]^{q_2 - 2}\frac{\partial f_2(x,h)}{\partial h} + C_4'(h)\left[f_2(x,h)\right]^{q_2 - 1}, & W_{\text{low}}(h) \leq x \leq W_{\text{ref}}(h), \\
C_5(h)(q_1 - 1)\left[f_3(x,h)\right]^{q_1 - 2}\frac{\partial f_3(x,h)}{\partial h} + C_5'(h)\left[f_3(x,h)\right]^{q_1 - 1} \\
\quad + C_6(h)(q_2 - 1)\left[f_3(x,h)\right]^{q_2 - 2}\frac{\partial f_3(x,h)}{\partial h} + C_6'(h)\left[f_3(x,h)\right]^{q_2 - 1}, & W_{\text{ref}}(h) \leq x \leq W_{\text{peak}}(h), \\
C_7(h)(q_1 - 1)\left[f_4(x,h)\right]^{q_1 - 2}\frac{\partial f_4(x,h)}{\partial h} + C_7'(h)\left[f_4(x,h)\right]^{q_1 - 1} \\
\quad + C_8(h)(q_2 - 1)\left[f_4(x,h)\right]^{q_2 - 2}\frac{\partial f_4(x,h)}{\partial h} + C_8'(h)\left[f_4(x,h)\right]^{q_2 - 1}, & W_{\text{peak}}(h) \leq x \leq W_{\text{updt}}(h).
\end{cases}
$$

Differentiating (4.6),

$$0 = -C_1(h)q_1(q_1-1)\big[f_1(x,h)\big]^{q_1-2}\frac{\partial f_1(x,h)}{\partial h} - C_1'(h)q_1\big[f_1(x,h)\big]^{q_1-1}$$
$$-C_2(h)q_2(q_2-1)\big[f_1(x,h)\big]^{q_2-2}\frac{\partial f_1(x,h)}{\partial h} - C_2'(h)q_2\big[f_1(x,h)\big]^{q_2-1} + \frac{\lambda}{\gamma}. \tag{C.4}$$

Then, using $C_2(h) = 0$, we have for $\frac{\lambda h}{\gamma} \le x \le W_{\text{low}}(h)$,

$$\frac{\partial \pi^*_{\text{primal}}}{\partial h} = c_{11}\frac{\lambda}{\gamma q_1}.$$

Hence $\frac{\partial \pi^*_{\text{primal}}}{\partial h}$ is constant for $\frac{\lambda h}{\gamma} \le x \le W_{\text{low}}(h)$.
Differentiating (4.8),

$$0 = -C_3(h)q_1(q_1-1)\big[f_2(x,h)\big]^{q_1-2}\frac{\partial f_2(x,h)}{\partial h} - C_3'(h)q_1\big[f_2(x,h)\big]^{q_1-1}$$
$$-C_4(h)q_2(q_2-1)\big[f_2(x,h)\big]^{q_2-2}\frac{\partial f_2(x,h)}{\partial h} - C_4'(h)q_2\big[f_2(x,h)\big]^{q_2-1}$$
$$-\frac{1}{\gamma\beta_1 f_2(x,h)}\frac{\partial f_2(x,h)}{\partial h} + \frac{\alpha}{\gamma}. \tag{C.5}$$

Using (C.5), we have for $W_{\text{low}}(h) \le x \le W_{\text{ref}}(h)$,

$$\frac{\partial \pi^*_{\text{primal}}}{\partial h} = c_{11}\left\{ C_3'(h)\big[f_2(x,h)\big]^{q_1-1}(1-\frac{q_1}{q_2}) + \frac{\alpha}{\gamma q_2}\right.$$
$$\left. +\Big[C_3(h)\big[f_2(x,h)\big]^{q_1-1}(q_1-q_2) - \frac{1}{\gamma\beta_1 q_2}\Big]\frac{1}{f_2}\frac{\partial f_2(x,h)}{\partial h}\right\}$$
$$= \frac{\mu-r}{\sigma^2}\frac{r}{k}\left\{ C_3'(h)\big[f_2(x,h)\big]^{q_1-1}(1-\frac{q_1}{q_2}) + \frac{\alpha}{\gamma q_2}\right.$$
$$\left. +\frac{k}{r}A_2(x,h)\frac{C_3'(h)q_1[f_2(x,h)]^{q_1-1} + C_4'(h)q_2[f_2(x,h)]^{q_2-1} - \frac{\alpha}{\gamma}}{B_2(x,h)}\right\}.$$

Thus, using the expression of $C_3(h)$ and $C_4(h)$ in (3.13) and (3.17), we obtain that $C_3'(h)$ and $C_4'(h)$ are both bounded. Moreover, we have $1 \le f_2(x,h) < e^{(\alpha-\lambda)\beta_1 h}$ for $W_{\text{low}}(h) < x \le W_{\text{ref}}(h)$. Applying the estimates of $A_2(x,h)$ and $B_2(x,h)$ again, we know that $\frac{\partial \pi^*_{\text{primal}}}{\partial h}$ is bounded for $W_{\text{low}}(h) < x \le W_{\text{ref}}(h)$. For the rest two regions $W_{\text{ref}}(h) \le x \le W_{\text{peak}}(h)$ and $W_{\text{peak}}(h) \le x \le W_{\text{updt}}(h)$, the proof is similar and omitted. $\qquad\square$

D   Results with general reference

We consider the alternative endogenous reference point $\alpha[\varphi(h)c + (1-\varphi(h))h]$, which is a fraction α of the convex combination of the current consumption and consumption peak. $\varphi(h)$ is the proportion assigned to current consumption (it is a function of h) and we assume that the proportion function φ is non-decreasing and smooth with values in $[0,1]$. The non-decreasing property suggests

that once the maximum is updated, its weight decreases. This assumption aims to capture the insight that, upon updating the consumption peak, the agent tends to put more emphasis on the current consumption c instead of the past peak. We further assume that $\varphi'(h)h + \varphi(h) \leq 1$ for all $h \geq h_0$, which implies the non-increasing property of the utility $U(c,h)$ w.r.t h.

Remark 8. $\varphi(h) = 0$ reduces to the case in the main part of the paper. $\varphi(h) = 1$ reduces to a non-habit model. A non-trivial choice of $\varphi(h)$ satisfying the mentioned assumptions might be the fractional function $\varphi(h) = \bar{\varphi}\frac{h}{h+\hat{h}}$ where constant $\hat{h} > 0$ is a benchmark level and $\bar{\varphi} \in [0,1]$ is a scaling constant.

The optimal dual feedback form $c^*(y,h)$ is replaced by

$$
c^*(y,h) = \begin{cases}
\lambda h, & 1 \vee (1 - \alpha\varphi(h))e^{[(\alpha-\lambda)-(1-\lambda)\alpha\varphi(h)]\beta_1 h} \leq y, \\
-\frac{1}{\beta_1}\frac{1}{1-\alpha\varphi(h)}\ln(\frac{y}{1-\alpha\varphi(h)}) + \alpha h\frac{1-\varphi(h)}{1-\alpha\varphi(h)}, & 1 \leq y < 1 \vee (1-\alpha\varphi(h))e^{[(\alpha-\lambda)-(1-\lambda)\alpha\varphi(h)]\beta_1 h}, \\
-\frac{1}{\beta_2}\frac{1}{1-\alpha\varphi(h)}\ln(\frac{y}{1-\alpha\varphi(h)}) + \alpha h\frac{1-\varphi(h)}{1-\alpha\varphi(h)}, & (1-\alpha\varphi(h))e^{-(1-\alpha)\beta_2 h} \leq y < 1, \\
h, & (1-\alpha)e^{-(1-\alpha)\beta_2 h} \leq y < (1-\alpha\varphi(h))e^{-(1-\alpha)\beta_2 h}.
\end{cases}
$$

We need the following assumption on the upper bound of $\varphi(h)$.

Assumption 1. $\varphi(\infty) < \frac{\alpha-\lambda}{\alpha(1-\lambda)}$.

This assumption is reasonable. On the one hand, we give a lower bound of the weight given to h in the reference point, so that consumption peak is always taken into consideration by the agent. On the other hand, under Assumption 1, for large h we always have $(1-\alpha\varphi(h))e^{[(\alpha-\lambda)-(1-\lambda)\alpha\varphi(h)]\beta_1 h} > 1$ so that every region in the expression of optimal consumption is not null, indicating that people with higher standard of living have more complicated behavior. Finally, it is interesting that Assumption 1 also serves as a convenient sufficient condition for verification theorem. See Remark 9.

The general solution to the dual HJB equation becomes

$$
\tilde{V}(y,h) = \begin{cases}
C_1(h)y^{q_1}+C_2(h)y^{q_2}-\frac{1}{\gamma}\lambda hy+\frac{1}{\gamma\beta_1}\left[1 - e^{[(\alpha-\lambda)-(1-\lambda)\alpha\varphi(h)]\beta_1 h}\right] \\
\quad, \ 1 \vee (1-\alpha\varphi(h))e^{[(\alpha-\lambda)-(1-\lambda)\alpha\varphi(h)]\beta_1 h} \leq y, \\
C_3(h)y^{q_1}+C_4(h)y^{q_2}+\frac{1}{\gamma\beta_1}(1-\frac{y}{1-\alpha\varphi(h)})+\frac{y}{\gamma\beta_1}\frac{1}{1-\alpha\varphi(h)}\ln(\frac{y}{1-\alpha\varphi(h)})-\frac{\alpha hy}{\gamma}\frac{1-\varphi(h)}{1-\alpha\varphi(h)} \\
\quad +\frac{k}{\gamma^2\beta_1}\frac{1}{1-\alpha\varphi(h)}y, \ 1 \leq y < 1 \vee (1-\alpha\varphi(h))e^{[(\alpha-\lambda)-(1-\lambda)\alpha\varphi(h)]\beta_1 h}, \\
C_5(h)y^{q_1}+C_6(h)y^{q_2}+\frac{1}{\gamma\beta_2}(1-\frac{y}{1-\alpha\varphi(h)})+\frac{y}{\gamma\beta_2}\frac{1}{1-\alpha\varphi(h)}\ln(\frac{y}{1-\alpha\varphi(h)})-\frac{\alpha hy}{\gamma}\frac{1-\varphi(h)}{1-\alpha\varphi(h)} \\
\quad +\frac{k}{\gamma^2\beta_2}\frac{1}{1-\alpha\varphi(h)}y, \ (1-\alpha\varphi(h))e^{-(1-\alpha)\beta_2 h} \leq y < 1, \\
C_7(h)y^{q_1}+C_8(h)y^{q_2}-\frac{1}{\gamma}hy+\frac{1}{\gamma\beta_2}\left[1-e^{-(1-\alpha)\beta_2 h}\right], \ (1-\alpha)e^{-(1-\alpha)\beta_2 h} \leq y < (1-\alpha\varphi(h))e^{-(1-\alpha)\beta_2 h}.
\end{cases}
$$

By the same approach, we obtain the expressions of $C_i(h), 1 \leq i \leq 8$ in the following two cases:

Case 1. For h such that $1 < (1-\alpha\varphi(h))e^{[(\alpha-\lambda)-(1-\lambda)\alpha\varphi(h)]\beta_1 h}$,

$$
C_2(h) = 0, \quad C_4(h) = -\frac{k}{\gamma^2\beta_1}\frac{1-q_1}{q_2-q_1}\frac{1}{(1-\alpha\varphi(h))^{q_2}}e^{-[(\alpha-\lambda)-(1-\lambda)\alpha\varphi(h)](q_2-1)\beta_1 h},
$$

$$
C_6(h) = C_4(h) + \frac{k}{\gamma^2}\frac{\beta_2-\beta_1}{\beta_1\beta_2}\frac{1-q_1}{q_2-q_1}\frac{1}{1-\alpha\varphi(h)} + \frac{1}{\gamma}\frac{\beta_2-\beta_1}{\beta_1\beta_2}\frac{q_1}{q_2-q_1}\frac{\alpha\varphi(h)}{1-\alpha\varphi(h)}
$$

$$
+ \frac{1}{\gamma}\frac{\beta_2-\beta_1}{\beta_1\beta_2}\frac{1-q_1}{q_2-q_1}\frac{1}{1-\alpha\varphi(h)}\ln(\frac{1}{1-\alpha\varphi(h)}),
$$

40

$$C_8(h) = C_6(h) + \frac{k}{\gamma^2 \beta_2} \frac{1-q_1}{q_2 - q_1} \frac{1}{(1-\alpha\varphi(h))^{q_2}} e^{(1-\alpha)(q_2-1)\beta_2 h},$$

$$C_7(h) = \frac{(1-\alpha)^{q_2-q_1} \frac{k}{\gamma^2} \frac{1-q_1}{q_2-q_1}(\alpha-\lambda)(q_2-1)}{(1-\alpha)(q_2-q_1)\beta_2 + (\alpha-\lambda)(q_2-1)\beta_1} e^{-\left[(1-\alpha)(q_2-q_1)\beta_2+(\alpha-\lambda)(q_2-1)\beta_1\right]h}$$
$$+ (1-\alpha)^{q_2-q_1} \frac{k}{\gamma^2 \beta_2} \frac{q_2-1}{q_2-q_1} e^{-(1-\alpha)(1-q_1)\beta_2 h},$$

$$C_5(h) = C_7(h) - \frac{k}{\gamma^2 \beta_2} \frac{q_2-1}{q_2-q_1} \frac{1}{(1-\alpha\varphi(h))^{q_1}} e^{-(1-\alpha)(1-q_1)\beta_2 h},$$

$$C_3(h) = C_5(h) - \frac{k}{\gamma^2} \frac{\beta_2-\beta_1}{\beta_1\beta_2} \frac{q_2-1}{q_2-q_1} \frac{1}{1-\alpha\varphi(h)} + \frac{1}{\gamma} \frac{\beta_2-\beta_1}{\beta_1\beta_2} \frac{q_2}{q_2-q_1} \frac{\alpha\varphi(h)}{1-\alpha\varphi(h)}$$
$$- \frac{1}{\gamma} \frac{\beta_2-\beta_1}{\beta_1\beta_2} \frac{q_2-1}{q_2-q_1} \frac{1}{1-\alpha\varphi(h)} \ln\left(\frac{1}{1-\alpha\varphi(h)}\right),$$

$$C_1(h) = C_3(h) + \frac{k}{\gamma^2 \beta_1} \frac{q_2-1}{q_2-q_1} \frac{1}{(1-\alpha\varphi(h))^{q_1}} e^{[(\alpha-\lambda)-(1-\lambda)\alpha\varphi(h)](1-q_1)\beta_1 h}.$$

In this case, as $\varphi(h) \in [0,1]$ and $e^{-[(\alpha-\lambda)-(1-\lambda)\alpha\varphi(h)]\beta_1 h} < 1 - \alpha\varphi(h) \le 1$, the estimates of $C_i(h), 1 \le i \le 8$ is the same as Section 3 except that the order estimate of $C_4(h)$ is replaced by $C_4(h) = \mathcal{O}(1)$.

Case 2. For h such that $1 \ge (1-\alpha\varphi(h))e^{[(\alpha-\lambda)-(1-\lambda)\alpha\varphi(h)]\beta_1 h}$ (the second region $1 \le y < 1 \vee (1-\alpha\varphi(h))e^{[(\alpha-\lambda)-(1-\lambda)\alpha\varphi(h)]\beta_1 h}$ is null), we have:

$$C_2(h) = 0,$$

$$C_6(h) = - \frac{1-q_1}{q_2-q_1} \frac{\lambda h}{\gamma} - \frac{q_1}{q_2-q_1} \frac{1}{\gamma\beta_1}\left(1 - e^{[(\alpha-\lambda)-(1-\lambda)\alpha\varphi(h)]\beta_1 h}\right) - \frac{1-q_1}{q_2-q_1} \frac{1}{\gamma\beta_2} \frac{1}{1-\alpha\varphi(h)} \ln\left(\frac{1}{1-\alpha\varphi(h)}\right)$$
$$+ \frac{q_1}{q_2-q_1} \frac{1}{\gamma\beta_2}\left(1 - \frac{1}{1-\alpha\varphi(h)}\right) + \frac{1-q_1}{q_2-q_1} \frac{\alpha h}{\gamma} \frac{1-\varphi(h)}{1-\alpha\varphi(h)} - \frac{1-q_1}{q_2-q_1} \frac{k}{\gamma^2\beta_2} \frac{1}{1-\alpha\varphi(h)},$$

$$C_8(h) = C_6(h) + \frac{k}{\gamma^2 \beta_2} \frac{1-q_1}{q_2-q_1} \frac{1}{(1-\alpha\varphi(h))^{q_2}} e^{(1-\alpha)(q_2-1)\beta_2 h},$$

$$C_7(h) = \frac{(1-\alpha)^{q_2-q_1} \frac{k}{\gamma^2} \frac{1-q_1}{q_2-q_1}(\alpha-\lambda)(q_2-1)}{(1-\alpha)(q_2-q_1)\beta_2 + (\alpha-\lambda)(q_2-1)\beta_1} e^{-\left[(1-\alpha)(q_2-q_1)\beta_2+(\alpha-\lambda)(q_2-1)\beta_1\right]h}$$
$$+ (1-\alpha)^{q_2-q_1} \frac{k}{\gamma^2 \beta_2} \frac{q_2-1}{q_2-q_1} e^{-(1-\alpha)(1-q_1)\beta_2 h},$$

$$C_5(h) = C_7(h) - \frac{k}{\gamma^2 \beta_2} \frac{q_2-1}{q_2-q_1} \frac{1}{(1-\alpha\varphi(h))^{q_1}} e^{-(1-\alpha)(1-q_1)\beta_2 h},$$

$$C_1(h) = C_5(h) + \frac{q_2-1}{q_2-q_1} \frac{\lambda h}{\gamma} - \frac{q_2}{q_2-q_1} \frac{1}{\gamma\beta_1}\left(1 - e^{[(\alpha-\lambda)-(1-\lambda)\alpha\varphi(h)]\beta_1 h}\right) + \frac{q_2-1}{q_2-q_1} \frac{1}{\gamma\beta_2} \frac{1}{1-\alpha\varphi(h)} \ln\left(\frac{1}{1-\alpha\varphi(h)}\right)$$
$$+ \frac{q_2}{q_2-q_1} \frac{1}{\gamma\beta_2}\left(1 - \frac{1}{1-\alpha\varphi(h)}\right) - \frac{q_2-1}{q_2-q_1} \frac{\alpha h}{\gamma} \frac{1-\varphi(h)}{1-\alpha\varphi(h)} + \frac{q_2-1}{q_2-q_1} \frac{k}{\gamma^2\beta_2} \frac{1}{1-\alpha\varphi(h)}.$$

Under Assumption 1, Case 2 does not happen for sufficiently large h so that the asymptotic estimates are not necessary. In Case 2, we still have

$$C_1(h) > 0, \quad C_7(h) > 0,$$
$$C_8(h) > 0.$$

Assumption 1 is actually a sufficient condition for $C_8(h) > 0$.

Remark 9. Under Assumption 1, the proof of $C_8(h) > 0$ in Case 2 is as follows: We write $C_8(h)$ as the sum of three terms given by

$$C_8(h) = \frac{k}{\gamma^2\beta_2} \frac{1-q_1}{q_2-q_1} \frac{1}{(1-\alpha\varphi(h))^{q_2}} [e^{(1-\alpha)(q_2-1)\beta_2 h} - (1-\alpha\varphi(h))^{q_2-1}]$$

$$+ \Big[\frac{1-q_1}{q_2-q_1}\frac{h}{\gamma}\frac{(\alpha-\lambda)-(1-\lambda)\alpha\varphi(h)}{1-\alpha\varphi(h)} - \frac{q_1}{q_2-q_1}\frac{1}{\gamma\beta_1}(1-e^{[(\alpha-\lambda)-(1-\lambda)\alpha\varphi(h)]\beta_1 h})\Big]$$

$$+ \Big[-\frac{1-q_1}{q_2-q_1}\frac{1}{\gamma\beta_2}\frac{1}{1-\alpha\varphi(h)}\ln\big(\frac{1}{1-\alpha\varphi(h)}\big) + \frac{q_1}{q_2-q_1}\frac{1}{\gamma\beta_2}\big(1-\frac{1}{1-\alpha\varphi(h)}\big)\Big].$$

The first term is positive due to $e^{(1-\alpha)(q_2-1)\beta_2 h} > 1 \geq (1-\alpha\varphi(h))^{q_2-1}$.

For the second term, it can be directly verified that

$$f(x) := \frac{1-q_1}{q_2-q_1}\frac{h}{\gamma}\frac{x}{1-\alpha\varphi(h)} - \frac{q_1}{q_2-q_1}\frac{1}{\gamma\beta_1}(1-e^{x\beta_1 h})$$

is increasing for $x \geq 0$. Thus $f((\alpha-\lambda)-(1-\lambda)\alpha\varphi(h)) > f(0) = 0$, which implies that the second term is positive.

For the last term, let

$$g(x) := -(1-q_1)x\ln(x) + q_1(1-x),$$

then one can directly show that $g(x)$ is decreasing for $x \in [1, \frac{1}{1-\alpha}]$. Hence $g(\frac{1}{1-\alpha\varphi(h)}) \geq g(1) = 0$ and the last term is non-negative.

Under all the aforementioned assumptions on φ, we can establish the verification theorem and apply duality to obtain the optimal strategy given in the following theorem.

Theorem D.1. *For $(x_0, h_0) \in \mathcal{C}$, where $\mathcal{C}$ is the effective region given by (3.8), let $c^*_{\mathrm{primal}}(\cdot, \cdot)$ and $\pi^*_{\mathrm{primal}}(\cdot, \cdot)$ be the feedback functions in terms of primal variable given respectively by*

$$c^*_{\mathrm{primal}}(x,h) = \begin{cases} \lambda h, & \frac{\lambda h}{\gamma} \leq x \leq W_{\mathrm{low}}(h) \wedge W_{\mathrm{ref}}(h), \\ -\frac{1}{\beta_1}\frac{1}{1-\alpha\varphi(h)}\ln\big(\frac{f_2(x,h)}{1-\alpha\varphi(h)}\big) + \alpha h\frac{1-\varphi(h)}{1-\alpha\varphi(h)}, & W_{\mathrm{low}}(h) \wedge W_{\mathrm{ref}}(h) < x \leq W_{\mathrm{ref}}(h), \\ -\frac{1}{\beta_2}\frac{1}{1-\alpha\varphi(h)}\ln\big(\frac{f_3(x,h)}{1-\alpha\varphi(h)}\big) + \alpha h\frac{1-\varphi(h)}{1-\alpha\varphi(h)}, & W_{\mathrm{ref}}(h) < x \leq W_{\mathrm{peak}}(h), \\ h, & W_{\mathrm{peak}}(h) < x \leq W_{\mathrm{updt}}(h). \end{cases}$$

$$\pi^*_{\mathrm{primal}}(x,h) = \frac{\mu-r}{\sigma^2}\begin{cases} \frac{r}{k}\Big\{C_1(h)\big[f_1(x,h)\big]^{q_1-1} + C_2(h)\big[f_1(x,h)\big]^{q_2-1}\Big\}, & \frac{\lambda h}{\gamma} \leq x \leq W_{\mathrm{low}}(h) \wedge W_{\mathrm{ref}}(h), \\ \frac{r}{k}\Big\{C_3(h)\big[f_2(x,h)\big]^{q_1-1} + C_4(h)\big[f_2(x,h)\big]^{q_2-1}\Big\} + \frac{1}{\gamma\beta_1}\frac{1}{1-\alpha\varphi(h)}, & W_{\mathrm{low}}(h) \wedge W_{\mathrm{ref}}(h) < x \leq W_{\mathrm{ref}}(h), \\ \frac{r}{k}\Big\{C_5(h)\big[f_3(x,h)\big]^{q_1-1} + C_6(h)\big[f_3(x,h)\big]^{q_2-1}\Big\} + \frac{1}{\gamma\beta_2}\frac{1}{1-\alpha\varphi(h)}, & W_{\mathrm{ref}}(h) < x \leq W_{\mathrm{peak}}(h), \\ \frac{r}{k}\Big\{C_7(h)\big[f_4(x,h)\big]^{q_1-1} + C_8(h)\big[f_4(x,h)\big]^{q_2-1}\Big\}, & W_{\mathrm{peak}}(h) < x \leq W_{\mathrm{updt}}(h). \end{cases}$$

where $f_i(x,h)$, $1 \leq i \leq 4$ are uniquely determined by

$$x = -C_1(h)q_1\big[f_1(x,h)\big]^{q_1-1} - C_2(h)q_2\big[f_1(x,h)\big]^{q_2-1} + \frac{\lambda h}{\gamma},$$

$$x = -C_3(h)q_1\big[f_2(x,h)\big]^{q_1-1} - C_4(h)q_2\big[f_2(x,h)\big]^{q_2-1} - \frac{1}{\gamma\beta_1}\frac{1}{1-\alpha\varphi(h)}\ln\Big[\frac{f_2(x,h)}{1-\alpha\varphi(h)}\Big]$$

$$- \frac{k}{\gamma^2\beta_1}\frac{1}{1-\alpha\varphi(h)} + \frac{\alpha h}{\gamma}\frac{1-\varphi(h)}{1-\alpha\varphi(h)},$$

$$x = -C_5(h)q_1\big[f_3(x,h)\big]^{q_1-1} - C_6(h)q_2\big[f_3(x,h)\big]^{q_2-1} - \frac{1}{\gamma\beta_2}\frac{1}{1-\alpha\varphi(h)}\ln\Big[\frac{f_3(x,h)}{1-\alpha\varphi(h)}\Big]$$

$$- \frac{k}{\gamma^2\beta_2}\frac{1}{1-\alpha\varphi(h)} + \frac{\alpha h}{\gamma}\frac{1-\varphi(h)}{1-\alpha\varphi(h)},$$

$$x = -C_7(h)q_1\big[f_4(x,h)\big]^{q_1-1} - C_8(h)q_2\big[f_4(x,h)\big]^{q_2-1} + \frac{h}{\gamma},$$

and $W_{\mathrm{low}}(h), W_{\mathrm{ref}}(h), W_{\mathrm{peak}}(h)$ *and* $W_{\mathrm{updt}}(h)$ *are given by*

$$W_{\mathrm{low}}(h) = -C_1(h)q_1\big[(1-\alpha\varphi(h))e^{[(\alpha-\lambda)-(1-\lambda)\alpha\varphi(h)]\beta_1 h}\big]^{q_1-1}$$

$$- C_2(h)q_2\big[(1-\alpha\varphi(h))e^{[(\alpha-\lambda)-(1-\lambda)\alpha\varphi(h)]\beta_1 h}\big]^{q_2-1} + \frac{\lambda h}{\gamma},$$

$$W_{\mathrm{ref}}(h) = -C_3(h)q_1 - C_4(h)q_2 - \frac{1}{\gamma\beta_1}\frac{1}{1-\alpha\varphi(h)}\ln\Big[\frac{1}{1-\alpha\varphi(h)}\Big]$$

$$- \frac{k}{\gamma^2\beta_1}\frac{1}{1-\alpha\varphi(h)} + \frac{\alpha h}{\gamma}\frac{1-\varphi(h)}{1-\alpha\varphi(h)},$$

$$W_{\mathrm{peak}}(h) = -C_5(h)q_1\big[(1-\alpha\varphi(h))e^{-(1-\alpha)\beta_2 h}\big]^{q_1-1} - C_6(h)q_2\big[(1-\alpha\varphi(h))e^{-(1-\alpha)\beta_2 h}\big]^{q_2-1}$$

$$+ \frac{h}{\gamma}\frac{1-\alpha}{1-\alpha\varphi(h)} - \frac{k}{\gamma^2\beta_2}\frac{1}{1-\alpha\varphi(h)} + \frac{\alpha h}{\gamma}\frac{1-\varphi(h)}{1-\alpha\varphi(h)},$$

$$W_{\mathrm{updt}}(h) = -C_7(h)q_1\big[(1-\alpha)e^{-(1-\alpha)\beta_2 h}\big]^{q_1-1} - C_8(h)q_2\big[(1-\alpha)e^{-(1-\alpha)\beta_2 h}\big]^{q_2-1} + \frac{h}{\gamma}.$$

Then SDE

$$\begin{cases} dX_t = rX_t dt + \pi^*_{\mathrm{primal}}(X_t, H_t^*)(\mu-r)dt + \pi^*_{\mathrm{primal}}(X_t, H_t^*)\sigma dW_t - c^*_{\mathrm{primal}}(X_t, H_t^*)dt, \\ X_0 = x_0 \end{cases} \tag{D.1}$$

with $H_t^* \triangleq h_0 \vee \sup\limits_{s\le t} c^*_{\mathrm{primal}}(X_s, H_s^*)$ *and* $H_0^* = h_0$, *has a unique strong solution* $\{X_t^*,\ t\ge 0\}$. *The optimal consumption and investment policy is*

$$\Big\{\big(c^*_{\mathrm{primal}}(X_t^*, H_t^*), \pi^*_{\mathrm{primal}}(X_t^*, H_t^*)\big),\ t\ge 0\Big\}.$$

Remark 10. It is interesting to notice that the depression region can possibly vanish under the current model. As has been mentioned, Assumption 1 ensures that this will not happen for large h. Specific characterizations of scenarios depends crucially on the form of φ. There are similar phenomena in Li et al. (2022), though due to completely different reasons.

Remark 11. The proofs of main results in this generalization are similar and thus omitted. The difference in proofs mainly lies in the proof of Lemma 4.2 and Lemma 4.6. In the proof of Lemma

4.2 for this generalization, we need to apply a similar decomposition as in Remark 9 to show that $\frac{r}{k}C_6(h)y^{q_2-1} + \frac{1}{\gamma\beta_2}\frac{1-q_1}{q_2-q_1}\frac{1}{1-\alpha\varphi(h)} > 0$ for $(1-\alpha\varphi(h))e^{-(1-\alpha)\beta_2 h} \leq y < 1$ in case $1 \geq (1 - \alpha\varphi(h))e^{[(\alpha-\lambda)-(1-\lambda)\alpha\varphi(h)]\beta_1 h}$. While in the proof of Lemma 4.6, we need to apply the boundedness of $\varphi(h)$ and $\varphi'(h)$ ($0 \leq \varphi'(h) < \frac{1-\varphi(h)}{h} \leq \frac{1-\frac{\alpha-\lambda}{\alpha(1-\lambda)}}{h_0}$) to show that $C_i'(h), 1 \leq i \leq 8$ are bounded.

E Results for $r \neq \gamma$

For $r \neq \gamma$, equation (3.9) should be replaced by

$$-\gamma\tilde{V}(y,h) + (\gamma-r)y\tilde{V}_y(y,h) + \frac{(r-\mu)^2}{2\sigma^2}y^2\tilde{V}_{yy}(y,h) = -\tilde{U}(y,h), \tag{E.1}$$

and $q_i, i = 1,2$ should be defined instead by $q_1 = \frac{k+r-\gamma-\sqrt{(k+r-\gamma)^2+4k\gamma}}{2k}$, $q_2 = \frac{k+r-\gamma+\sqrt{(k+r-\gamma)^2+4k\gamma}}{2k}$. It still holds that $q_1 < 0 < 1 < q_2$.

The general solution to (E.1) is

$$\tilde{V}(y,h) = \begin{cases} C_1(h)y^{q_1} + C_2(h)y^{q_2} - \frac{1}{r}\lambda hy + \frac{1}{\gamma\beta_1}\left[1 - e^{(\alpha-\lambda)\beta_1 h}\right], & e^{(\alpha-\lambda)\beta_1 h} \leq y, \\ C_3(h)y^{q_1}+C_4(h)y^{q_2}+\frac{1}{r\beta_1}y\ln(y)+\frac{\gamma-2r+k}{r^2\beta_1}y-\frac{\alpha h}{r}y+\frac{1}{\gamma\beta_1}, & 1 \leq y < e^{(\alpha-\lambda)\beta_1 h}, \\ C_5(h)y^{q_1}+C_6(h)y^{q_2}+\frac{1}{r\beta_2}y\ln(y)+\frac{\gamma-2r+k}{r^2\beta_2}y-\frac{\alpha h}{r}y+\frac{1}{\gamma\beta_2}, & e^{-(1-\alpha)\beta_2 h} \leq y < 1, \\ C_7(h)y^{q_1}+C_8(h)y^{q_2}-\frac{1}{r}hy+\frac{1}{\gamma\beta_2}\left[1-e^{-(1-\alpha)\beta_2 h}\right], & (1-\alpha)e^{-(1-\alpha)\beta_2 h} \leq y < e^{-(1-\alpha)\beta_2 h}. \end{cases}$$

We can obtain $C_i(h), 1 \leq i \leq 8$ in the same way as in Section 3.

$$C_2(h) = 0, \quad C_4(h) = \frac{1}{(q_2-q_1)\beta_1}\left[\frac{q_1}{\gamma} - \frac{1}{r} + \frac{\gamma-2r+k}{r^2}(q_1-1)\right]e^{-(\alpha-\lambda)(q_2-1)\beta_1 h},$$

$$C_6(h) = C_4(h) + \frac{\beta_2-\beta_1}{(q_2-q_1)\beta_1\beta_2}\left[-\frac{q_1}{\gamma} + \frac{1}{r} - \frac{\gamma-2r+k}{r^2}(q_1-1)\right],$$

$$C_8(h) = C_6(h) + \frac{1}{(q_2-q_1)\beta_2}\left[-\frac{q_1}{\gamma} + \frac{1}{r} - \frac{\gamma-2r+k}{r^2}(q_1-1)\right]e^{(1-\alpha)(q_2-1)\beta_2 h},$$

$$C_7(h) = \frac{1}{q_2-q_1}\frac{(1-\alpha)^{q_2-q_1}\left[-\frac{q_1}{\gamma}+\frac{1}{r}-\frac{\gamma-2r+k}{r^2}(q_1-1)\right](\alpha-\lambda)(q_2-1)}{(1-\alpha)(q_2-q_1)\beta_2 + (\alpha-\lambda)(q_2-1)\beta_1}e^{-\left[(1-\alpha)(q_2-q_1)\beta_2+(\alpha-\lambda)(q_2-1)\beta_1\right]h}$$

$$+ (1-\alpha)^{q_2-q_1}\frac{-\frac{q_1}{\gamma}+\frac{1}{r}-\frac{\gamma-2r+k}{r^2}(q_1-1)}{(1-q_1)\beta_2}\frac{q_2-1}{q_2-q_1}e^{-(1-\alpha)(1-q_1)\beta_2 h},$$

$$C_5(h) = C_7(h) + \frac{1}{(q_2-q_1)\beta_2}\left[-\frac{q_2}{\gamma} + \frac{1}{r} - \frac{\gamma-2r+k}{r^2}(q_2-1)\right]e^{-(1-\alpha)(1-q_1)\beta_2 h},$$

$$C_3(h) = C_5(h) + \frac{\beta_2-\beta_1}{(q_2-q_1)\beta_1\beta_2}\left[-\frac{q_2}{\gamma} + \frac{1}{r} - \frac{\gamma-2r+k}{r^2}(q_2-1)\right],$$

$$C_1(h) = C_3(h) + \frac{1}{(q_2-q_1)\beta_1}\left[\frac{q_2}{\gamma} - \frac{1}{r} + \frac{\gamma-2r+k}{r^2}(q_2-1)\right]e^{(\alpha-\lambda)(1-q_1)\beta_1 h}.$$

It can verified that $q_1 < \frac{\frac{\gamma-2r+k}{r^2}+\frac{1}{r}}{\frac{\gamma-2r+k}{r^2}+\frac{1}{\gamma}} < q_2$. In fact, an equivalent condition is $k\left(\frac{\frac{\gamma-2r+k}{r^2}+\frac{1}{r}}{\frac{\gamma-2r+k}{r^2}+\frac{1}{\gamma}}\right)^2 - (k+r-\gamma)\frac{\frac{\gamma-2r+k}{r^2}+\frac{1}{r}}{\frac{\gamma-2r+k}{r^2}+\frac{1}{\gamma}} - \gamma < 0$, which is equivalent(by direct computation) to the trivial inequality $-\frac{k^2}{\gamma r^2(\frac{\gamma-2r+k}{r^2}+\frac{1}{\gamma})^2} < 0$. Then we obtain from $q_1 < \frac{\frac{\gamma-2r+k}{r^2}+\frac{1}{r}}{\frac{\gamma-2r+k}{r^2}+\frac{1}{\gamma}} < q_2$ that

$$-\frac{q_1}{\gamma} + \frac{1}{r} - \frac{\gamma - 2r + k}{r^2}(q_1 - 1) > 0,$$

$$-\frac{q_2}{\gamma} + \frac{1}{r} - \frac{\gamma - 2r + k}{r^2}(q_2 - 1) < 0.$$

The difference of $C_i(h), 1 \le i \le 8$ between $r = \gamma$ and $r \ne \gamma$ is that certain positive coefficients such as $\frac{k}{\gamma^2}\frac{q_2-1}{q_2-q_1}$ in case $r = \gamma$ are replaced by other positive coefficients such as $\frac{1}{q_2-q_1}\left[\frac{q_2}{\gamma} - \frac{1}{r} + \frac{\gamma-2r+k}{r^2}(q_2 - 1)\right]$ in case $r \ne \gamma$. As a result, the estimates of $C_i(h), 1 \le i \le 8$ is completely the same as in Section 3 and the main results are similar. The optimal strategy is as follows.

Theorem E.1. *For $(x_0, h_0) \in \mathcal{C}$, where $\mathcal{C}$ is the effective region given by (3.8), let $c^*_{\mathrm{primal}}(\cdot, \cdot)$ and $\pi^*_{\mathrm{primal}}(\cdot, \cdot)$ be the feedback functions in terms of primal variable given respectively by*

$$c^*_{\mathrm{primal}}(x, h) = \begin{cases} \lambda h, & \frac{\lambda h}{\gamma} \le x \le W_{\mathrm{low}}(h), \\ -\frac{1}{\beta_1}\ln\left[f_2(x, h)\right] + \alpha h, & W_{\mathrm{low}}(h) < x \le W_{\mathrm{ref}}(h), \\ -\frac{1}{\beta_2}\ln\left[f_3(x, h)\right] + \alpha h, & W_{\mathrm{ref}}(h) < x \le W_{\mathrm{peak}}(h), \\ h, & W_{\mathrm{peak}}(h) < x \le W_{\mathrm{updt}}(h), \end{cases}$$

$$\pi^*_{\mathrm{primal}}(x, h) = \frac{\mu - r}{\sigma^2}\begin{cases} q_1(q_1-1)C_1(h)\left[f_1(x,h)\right]^{q_1-1} + q_2(q_2-1)C_2(h)\left[f_1(x,h)\right]^{q_2-1} \\ \quad, \ \frac{\lambda h}{\gamma} \le x \le W_{\mathrm{low}}(h), \\ q_1(q_1-1)C_3(h)\left[f_2(x,h)\right]^{q_1-1} + q_2(q_2-1)C_4(h)\left[f_2(x,h)\right]^{q_2-1} + \frac{1}{r\beta_1} \\ \quad, \ W_{\mathrm{low}}(h) < x \le W_{\mathrm{ref}}(h), \\ q_1(q_1-1)C_5(h)\left[f_3(x,h)\right]^{q_1-1} + q_2(q_2-1)C_6(h)\left[f_3(x,h)\right]^{q_2-1} + \frac{1}{r\beta_2} \\ \quad, \ W_{\mathrm{ref}}(h) < x \le W_{\mathrm{peak}}(h), \\ q_1(q_1-1)C_7(h)\left[f_4(x,h)\right]^{q_1-1} + q_2(q_2-1)C_8(h)\left[f_4(x,h)\right]^{q_2-1} \\ \quad, \ W_{\mathrm{peak}}(h) < x \le W_{\mathrm{updt}}(h), \end{cases}$$

where $f_i(x, h), \ 1 \le i \le 4$ are uniquely determined by

$$x = -C_1(h)q_1\left[f_1(x,h)\right]^{q_1-1} - C_2(h)q_2\left[f_1(x,h)\right]^{q_2-1} + \frac{\lambda h}{r},$$

$$x = -C_3(h)q_1\left[f_2(x,h)\right]^{q_1-1} - C_4(h)q_2\left[f_2(x,h)\right]^{q_2-1} - \frac{1}{r\beta_1}\ln[f_2(x,h)] - \frac{\gamma - r + k}{r^2\beta_1} + \frac{\alpha h}{r},$$

$$x = -C_5(h)q_1\left[f_3(x,h)\right]^{q_1-1} - C_6(h)q_2\left[f_3(x,h)\right]^{q_2-1} - \frac{1}{r\beta_2}\ln[f_3(x,h)] - \frac{\gamma - r + k}{r^2\beta_2} + \frac{\alpha h}{r},$$

$$x = -C_7(h)q_1\left[f_4(x,h)\right]^{q_1-1} - C_8(h)q_2\left[f_4(x,h)\right]^{q_2-1} + \frac{h}{r},$$

and $W_{\mathrm{low}}(h), W_{\mathrm{ref}}(h), W_{\mathrm{peak}}(h)$ and $W_{\mathrm{updt}}(h)$ are given by

$$W_{\mathrm{low}}(h) = -C_1(h)q_1 e^{-(\alpha-\lambda)(1-q_1)\beta_1 h} - C_2(h)q_2 e^{(\alpha-\lambda)(q_2-1)\beta_1 h} + \frac{\lambda h}{r},$$

$$W_{\mathrm{ref}}(h) = -C_3(h)q_1 - C_4(h)q_2 - \frac{\gamma - r + k}{r^2\beta_1} + \frac{\alpha h}{r},$$

$$W_{\text{peak}}(h) = - C_5(h)q_1e^{(1-\alpha)(1-q_1)\beta_2 h} - C_6(h)q_2e^{-(1-\alpha)(q_2-1)\beta_2 h} - \frac{\gamma - r + k}{r^2\beta_1} + \frac{h}{r},$$

$$W_{\text{updt}}(h) = - C_7(h)q_1(1-\alpha)^{q_1-1}e^{(1-\alpha)(1-q_1)\beta_2 h} - C_8(h)q_2(1-\alpha)^{q_2-1}e^{-(1-\alpha)(q_2-1)\beta_2 h} + \frac{h}{\gamma}.$$

Then SDE

$$\begin{cases} dX_t = rX_tdt + \pi^*_{\text{primal}}(X_t, H^*_t)(\mu - r)dt + \pi^*_{\text{primal}}(X_t, H^*_t)\sigma dW_t - c^*_{\text{primal}}(X_t, H^*_t)dt, \\ X_0 = x_0 \end{cases}$$

with $H^*_t \triangleq h_0 \vee \sup_{s \leq t} c^*_{\text{primal}}(X_s, H^*_s)$ *and* $H^*_0 = h_0$, *has a unique strong solution* $\{X^*_t,\ t \geq 0\}$. *The optimal consumption and investment policy is*

$$\left\{ \left(c^*_{\text{primal}}(X^*_t, H^*_t), \pi^*_{\text{primal}}(X^*_t, H^*_t)\right),\ t \geq 0 \right\}.$$

Here we just give the proof of Lemma 4.2 with $r \neq \gamma$. The proofs of other results with $r \neq \gamma$ are very similar to those with $r = \gamma$.

Proof of Lemma 4.2 with $r \neq \gamma$. If $e^{(\alpha-\lambda)\beta_1 h} \leq y$, then

$$\tilde{V}_{yy}(y, h) = C_1(h)q_1(q_1 - 1)y^{q_1-2} + C_2(h)q_2(q_2 - 1)y^{q_2-2}.$$

As $C_1(h) > 0$ and $C_2(h) = 0$, we have $\tilde{V}_{yy}(y, h) > 0$ for $e^{(\alpha-\lambda)\beta_1 h} \leq y$.
If $1 \leq y < e^{(\alpha-\lambda)\beta_1 h}$, then

$$y\tilde{V}_{yy}(y, h) = C_3(h)q_1(q_1 - 1)y^{q_1-1} + C_4(h)q_2(q_2 - 1)y^{q_2-1} + \frac{1}{r\beta_1}.$$

Let $\psi(y) = y\tilde{V}_{yy}(y, h)$, then

$$\psi'(y) = C_3(h)q_1(q_1 - 1)^2 y^{q_1-2} + C_4(h)q_2(q_2 - 1)^2 y^{q_2-2}.$$

Noting that $C_4(h) < 0$, $\psi(y)$ is either increasing, decreasing or first increasing then decreasing, we only need to show $\psi(1) > 0$ and $\psi(e^{(\alpha-\lambda)\beta_1 h}) > 0$. Precisely, using $C_7(h) > 0$ and the fact that

$$\frac{1}{r} = \frac{q_1(q_1 - 1)}{q_2 - q_1}\left[\frac{q_2}{\gamma} - \frac{1}{r} + \frac{\gamma - 2r + k}{r^2}(q_2 - 1)\right] + \frac{q_2(q_2 - 1)}{q_2 - q_1}\left[-\frac{q_1}{\gamma} + \frac{1}{r} - \frac{\gamma - 2r + k}{r^2}(q_1 - 1)\right],$$

we have

$$\begin{aligned} \psi(1) =& C_3(h)q_1(q_1 - 1) + C_4(h)q_2(q_2 - 1) + \frac{1}{r\beta_1} \\ =& q_1(q_1 - 1)C_7(h) + \frac{q_1(q_1 - 1)}{(q_2 - q_1)\beta_2}\left[\frac{q_2}{\gamma} - \frac{1}{r} + \frac{\gamma - 2r + k}{r^2}(q_2 - 1)\right]\left[1 - e^{-(1-\alpha)(1-q_1)\beta_2 h}\right] \\ &+ \frac{q_2(q_2 - 1)}{(q_2 - q_1)\beta_1}\left[-\frac{q_1}{\gamma} + \frac{1}{r} - \frac{\gamma - 2r + k}{r^2}(q_1 - 1)\right]\left[1 - e^{-(\alpha-\lambda)(q_2-1)\beta_1 h}\right] \\ >& 0 \end{aligned}$$

and

46

$$\psi(e^{(\alpha-\lambda)\beta_1 h}) = C_3(h)q_1(q_1-1)e^{(q_1-1)(\alpha-\lambda)\beta_1 h} + C_4(h)q_2(q_2-1)e^{(q_2-1)(\alpha-\lambda)\beta_1 h} + \frac{1}{r\beta_1}$$

$$= q_1(q_1-1)C_7(h)e^{(q_1-1)(\alpha-\lambda)\beta_1 h}$$
$$+ \frac{q_1(q_1-1)}{(q_2-q_1)\beta_2}\left[\frac{q_2}{\gamma} - \frac{1}{r} + \frac{\gamma - 2r + k}{r^2}(q_2-1)\right]\left[1 - e^{-(1-\alpha)(1-q_1)\beta_2 h}\right]e^{(q_1-1)(\alpha-\lambda)\beta_1 h}$$
$$+ \frac{q_1(q_1-1)}{(q_2-q_1)\beta_2}\left[\frac{q_2}{\gamma} - \frac{1}{r} + \frac{\gamma - 2r + k}{r^2}(q_2-1)\right]\left[1 - e^{-(\alpha-\lambda)(1-q_1)\beta_1 h}\right]$$
$$> 0.$$

Thus $\tilde{V}_{yy}(y,h) > 0$ for $1 \le y < e^{(\alpha-\lambda)\beta_1 h}$.

If $e^{-(1-\alpha)\beta_2 h} \le y < 1$, then

$$y\tilde{V}_{yy}(y,h) = C_5(h)q_1(q_1-1)y^{q_1-1} + C_6(h)q_2(q_2-1)y^{q_2-1} + \frac{1}{r\beta_2}$$
$$= q_1(q_1-1)\left\{C_5(h)y^{q_1-1} + \frac{1}{(q_2-q_1)\beta_2}\left[\frac{q_2}{\gamma} - \frac{1}{r} + \frac{\gamma - 2r + k}{r^2}(q_2-1)\right]\right\}$$
$$+ q_2(q_2-1)\left\{C_6(h)y^{q_2-1} + \frac{1}{(q_2-q_1)\beta_2}\left[-\frac{q_1}{\gamma} + \frac{1}{r} - \frac{\gamma - 2r + k}{r^2}(q_1-1)\right]\right\}.$$

For any fixed $h > 0$, if $C_5(h) \ge 0$, then $C_5(h)y^{q_1-1} + \frac{1}{(q_2-q_1)\beta_2}\left[\frac{q_2}{\gamma} - \frac{1}{r} + \frac{\gamma-2r+k}{r^2}(q_2-1)\right] > 0$; If $C_5(h) < 0$, then $C_5(h)y^{q_1-1} + \frac{1}{(q_2-q_1)\beta_2}\left[\frac{q_2}{\gamma} - \frac{1}{r} + \frac{\gamma-2r+k}{r^2}(q_2-1)\right]$ is increasing in y and

$$C_5(h)y^{q_1-1} + \frac{1}{(q_2-q_1)\beta_2}\left[\frac{q_2}{\gamma} - \frac{1}{r} + \frac{\gamma - 2r + k}{r^2}(q_2-1)\right]$$
$$\ge C_5(h)e^{(1-\alpha)(1-q_1)\beta_2 h} + \frac{1}{(q_2-q_1)\beta_2}\left[\frac{q_2}{\gamma} - \frac{1}{r} + \frac{\gamma - 2r + k}{r^2}(q_2-1)\right]$$
$$= C_7(h)e^{(1-\alpha)(1-q_1)\beta_2 h}$$
$$> 0.$$

Similarly, we have

$$C_6(h)y^{q_2-1} + \frac{1}{(q_2-q_1)\beta_2}\left[-\frac{q_1}{\gamma} + \frac{1}{r} - \frac{\gamma - 2r + k}{r^2}(q_1-1)\right] > 0.$$

It follows that $\tilde{V}_{yy}(y,h) > 0$ for $e^{-(1-\alpha)\beta_2 h} \le y < 1$.

Finally, If $(1-\alpha)e^{-(1-\alpha)\beta_2 h} \le y < e^{-(1-\alpha)\beta_2 h}$, then

$$\tilde{V}_{yy}(y,h) = C_7(h)q_1(q_1-1)y^{q_1-2} + C_8(h)q_2(q_2-1)y^{q_2-2}.$$

As $C_7(h) > 0$ and $C_8(h) > 0$, we have $\tilde{V}_{yy}(y,h) > 0$ for $(1-\alpha)e^{-(1-\alpha)\beta_2 h} \le y < e^{-(1-\alpha)\beta_2 h}$. Thus, the proof is complete. $\square$

References

Bahman Angoshtari, Erhan Bayraktar, and Virginia R. Young. Optimal investment and consumption under a habit-formation constraint. *SIAM Journal on Financial Mathematics*, 13(1): 321–352, 2022.

Luc Arrondel, Pierre Lamarche, and Frédérique Savignac. Wealth effects on consumption across the wealth distribution: Empirical evidence. 2015. URL http://ideas.repec.org/p/bfr/banfra/552.html.

T. Arun. The Merton problem with a drawdown constraint on consumption, 2012. URL https://arxiv.org/abs/1210.5205.

Don Bellante and Carole A. Green. Relative risk aversion among the elderly. *Review of Financial Economics*, 13(3):269–281, 2004.

Markus K. Brunnermeier and Stefan Nagel. Do wealth fluctuations generate time-varying risk aversion? Micro-evidence on individuals' asset allocation. *American Economic Review*, 98(3): 713–736, 2008.

David A. Chapman. Habit formation and aggregate consumption. *Econometrica*, 66(5):1223–1230, 1998.

Alain Cohn, Jan Engelmann, Ernst Fehr, and Michel André Maréchal. Evidence for countercyclical risk aversion: An experiment with financial professionals. *American Economic Review*, 105(2): 860–885, 2015.

Shuoqing Deng, Xun Li, Huyen Pham, and Xiang Yu. Optimal consumption with reference to past spending maximum. *Finance and Stochastics*, 26(2):217–266, 2022.

Jerome B. Detemple and Fernando Zapatero. Optimal consumption-portfolio policies with habit formation. *Mathematical Finance*, 2(4):251–274, 1992.

Philip H. Dybvig. Dusenberry's ratcheting of consumption: Optimal dynamic consumption and investment given intolerance for any decline in standard of living. *Review of Economic Studies*, 62(2):287–313, 1995.

Paolo Guasoni, Gur Huberman, and Dan Ren. Shortfall aversion. *Mathematical Finance*, 30(3): 869–920, 2020.

Luigi Guiso, Paola Sapienza, and Luigi Zingales. Time varying risk aversion. *Journal of Financial Economics*, 128(3):403–421, 2018.

Xuedong He and Linan Yang. Realization utility with adaptive reference points. *Mathematical Finance*, 29(2):409–447, 2019.

Junkee Jeon and Kyunghyun Park. Optimal retirement and portfolio selection with consumption ratcheting. *Mathematics and Financial Economics*, 14(3):353–397, 2020.

Hanqing Jin and Xun Yu Zhou. Behavioral portfolio selection in continuous time. *Mathematical Finance*, 18(3):385–426, 2008.

Xun Li, Xiang Yu, and Qinyi Zhang. Optimal consumption with loss aversion and reference to past spending maximum. 2022. URL http://arxiv.org/abs/2108.02648.

Robert C. Merton. Lifetime portfolio selection under uncertainty: The continuous time case. *The Review of Economics and Statistics*, 51(3):247–257, 1969.

Roman Muraviev. Additive habit formation: Consumption in incomplete markets with random endowments. *Mathematics and Financial Economics*, 5(2):67–99, 2011.

Ivan Paya and Peng Wang. Wealth fluctuations and investment in risky assets: The UK micro evidence on households asset allocation. *Journal of Empirical Finance*, 38:221–235, 2016.

Robert A. Pollak. Habit formation and dynamic demand functions. *The Journal of Political Economy*, 78(4):745–763, 1970.

Harl E. Ryder and Geoffrey M. Heal. Optimal growth with intertemporally dependent preferences. *The Review of Economic Studies*, 40(1):1–31, 1973.

Frederick W. Siegel and James P. Hoban. Relative risk aversion revisited. *The Review of Economics and Statistics*, 64(3):481–487, 1982.

Servaas Van Bilsen, Roger J.A. Laeven, and Theo E. Nijman. Consumption and portfolio choice under loss aversion and endogenous updating of the reference level. *Management Science*, 66(9): 3927–3955, 2020.

Jessica A. Wachter and Motohiro Yogo. Why do household portfolio shares rise in wealth? *Review of Financial Studies*, 23(11):3929–3965, 2010.

Xiang Yu. Utility maximization with addictive consumption habit formation in incomplete semi-martingale markets. *The Annals of Applied Probability*, 25(3):1383–1419, 2015.