Consumption-investment decisions with endogenous reference point and drawdown constraint

Zongxia Liang · Xiaodong Luo · Fengyi Yuan

Received: 31 March 2022 / Accepted: 24 April 2023 / Published online: 23 May 2023
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2023

Abstract
We study a consumption-investment decision problem related to the past spending maximum. In the problem, we consider two crucial consumption levels: the lowest constrained level and a reference level, and both levels are fractions of the past spending maximum. The decision-maker has different risk aversions on different sides of the reference level. We solve this stochastic control problem and derive semi-explicit forms of the value function, optimal consumption plan, and optimal investment strategy. We find five important wealth thresholds which are nonlinear functions of the past spending maximum. Based on numerical results and theoretical analysis, we also find that the model has significant economic implications. 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 varies inversely with risk aversion at the reference level; the implied relative risk aversion is roughly a smile in wealth; the welfare is much more vulnerable to wealth shocks when the reference level is not reached.

Keywords Consumer behavior · Past consumption peak · Drawdown constraint · Endogenous reference point · Stochastic control

JEL Classification C61 · G11 · G41

1 Introduction
Intuitively, the historical peak of consumption significantly impacts individual consumption decisions. For example, to consume below a specific ratio of the historical peak brings an
impulsion to “reclaim the past glory”. On the other hand, if consumption is forced to decline further to a level far below the historical peak, it becomes unbearable. In this case, people will try to increase their wealth (by financing, borrowing, or selling illiquid assets) at any cost to satisfy consumption at a specific (but low) ratio of the past peak. Therefore it is not surprising that there is literature studying the so-called drawdown constraint; see Dybvig [9] and Arun [2].

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

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

where $\lambda$ and $\alpha$ with $0 < \lambda < \alpha < 1$ are exogenous constants, representing two important thresholds of consumption level, and $\beta_1, \beta_2 > 0$ 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 level $\alpha h$. Besides, we consider a change in risk aversions on different sides of the reference level to reflect psychological effects when consumption exceeds or falls below the reference level. In economic literature, variations of individual risk aversions are extensively discussed, and empirical evidence has been found. On the one hand, based on both naturally occurring data and lab data, one finds that people become more risk averse when experiencing crisis or fear, see Cohn et al. [6], Guiso et al. [11] and references therein. In this paper, choosing $\beta_1 > \beta_2$, we model such a crisis as the declining 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-winning prospect theory, it is admitted that people become risk-seeking when the pay-offs fall into the loss domain. When incorporating such extreme gambling behavior into consumption decisions, somewhat radical optimal decisions are derived: people never consume between 0 (or the lowest constrained level) and the reference level; see Van Bilsen et al. [22] and Li et al. [15]. One possible reason is that consumption falling below a certain reference level is not generally treated as a loss but as bad luck or a temporary crisis. To incorporate the gambling effect into consumption decisions, we can conveniently choose $\beta_1 < \beta_2$ in our model. In this case, people are willing to take more risks (but are still risk-averse) when their consumption is in danger ($c < \alpha h$). Our main interests are investigating consumption and portfolio behaviors under the preference alteration at the reference level. The reference level itself, however, can be further generalized from the particular choice $\alpha h$. We consider the case when the reference level is a fraction of the convex combination between current consumption and consumption peak. We list the results of this generalization in Appendix D.

Our model leads to consumption and investment decisions with significant economic implications.

We find five important thresholds of wealth, depending on $h$ and denoted respectively by $W_{bkrp}(h)$, $W_{low}(h)$, $W_{ref}(h)$, $W_{peak}(h)$ and $W_{updt}(h)$, which are crucial to describe the derived consumption and investment decisions. For those with $x \geq W_{updt}(h)$, the best choice of consumption is to update the consumption peak to move to or maintain on the curve $x = W_{updt}(h)$. If $x < W_{bkrp}(h)$, the wealth of the agent is insufficient for keeping a consumption
rate $c \geq \lambda h$ considering her/his standard of living, hence it is not allowed in our discussion. The above results indicate that only $x \in [W_{\text{bkp}}(h), W_{\text{updt}}(h)]$ are of special interest, and they are said to be in the effective region. Given a standard of living, the optimal consumption rate is straightforward for people with the highest and lowest wealth levels: if $W_{\text{peak}}(h) < x \leq W_{\text{updt}}(h)$ (named as the satisfactory region), they will revisit the historical peak $c = h$; if $W_{\text{bkp}}(h) \leq x \leq W_{\text{low}}(h)$ (named as the gloom region), they choose to consume at the lowest level $c = \lambda h$. As for the region of intermediate wealth $W_{\text{low}}(h) < x \leq W_{\text{peak}}(h)$, using the reference threshold $W_{\text{ref}}(h)$, we divide it into two sub-regions: the depression region and the recovery region. We discover two important phenomena. First, although the (instantaneous) marginal propensity to consume (MPC) out of wealth $\frac{\partial c}{\partial x}$ is decreasing in the lower part of the depression-recovery region, we predict that it is increasing in the rest of depression-recovery region unless $h$ is sufficiently small. However, if the wealth level is $x = W_{\text{updt}}(h)$, called the bliss curve by us, the MPC out of wealth is decreasing again. Second, the MPC out of wealth jumps by a fixed proportion $\beta_1/\beta_2$ at the threshold $W_{\text{ref}}(h)$. This fact reveals one of the salient features of our model. It indicates that when their consumption crosses the reference level, people have lower or higher MPC according to their attitudes to crisis (gambling or stop-loss). See Fig. 2 in Sect. 5 for an illustration of optimal consumption decisions. By allowing different risk aversions on different sides of the reference level, we can explore consumption decisions in the limit case as $\beta_1$ or $\beta_2$ goes to 0. We find interesting connections between our model and the one in Li et al. [15]; see Sect. 6.2.3.

We also obtain the optimal investment strategy $\pi^*$, representing the amount of wealth invested in risky assets, and the optimal investment proportion $\pi^*/x$ is obtained as a by-product. Recall that in classical Merton’s problem, the optimal investment proportion is $\frac{\mu - \gamma}{\sigma^2}$ if relative risk aversion is $\gamma$. Therefore, the inverse of $\pi^*/x$ can be treated equivalently as the relative risk aversion, which we call implied relative risk aversion; see also Jeon and Park [13]. Keeping this in mind, we find that decreasing relative risk aversion (DRRA) and increasing relative risk aversion (IRRA) are both possible, even for a single agent. The implied relative risk aversion is roughly a $U$-shaped curve (smile) in variable $x$ with a trough around $W_{\text{ref}}(h)$, that is, the risky investment proportion is a hump with a peak around $W_{\text{ref}}(h)$. We further predict that this effect is more pronounced when $h$ is low. In economic literature, there has been a long-standing debate on how relative risk aversions vary with wealth, and evidence for both DRRA and IRRA is found; see Siegel and Hoban [21] and Bellante and Green [3]. We provide an explanation for this: the RRA can decrease in wealth because there is impulsion to get back to a higher consumption above $ah$ if $x$ is not large; the RRA can also increase in $x$ if $x$ has been enough for maintaining a satisfactory standard of living. As for the portfolio selection itself, although it is widely admitted that for macro data, the wealthier people tend to invest more proportion of their wealth in risky assets, there is no solid agreement on the same question in micro aspect. That is, what happens to risky investment if the wealth increases for a given household? Our model predicts that people proportionally reduce risky investment if their wealth grows, which is consistent with empirical studies (c.f. Brunnermeier and Nagel [4] and Paya and Wang [18]) or other possible alternative models (c.f. Wachter and Yogo [23]). Meanwhile, our model allows the opposite result when the wealth is not enough. Thus, we can explain the co-existence of both phenomena in some literature such as Brunnermeier and Nagel [4]. See Fig. 2 in Sect. 5 for an illustration of optimal risky assets allocation.

The literature most closely related to the present paper is Deng et al. [7]. In the aspect of reference point, we adopt the setting of Deng et al. [7] and rely on their solving techniques and other celebrated tools such as dynamic programming, duality theory, and region-wise solving method. Our choice of reference point is for simplicity. The set of solving techniques
we have used, inspired by Deng et al. [7], is widely applicable to other forms of reference points; see Appendix D. In addition to the reference point itself, our primary interests include investigating consumption and portfolio behaviors under the preference alteration at the reference point. Our model has several distinctive features, complementing the one proposed and solved in Deng et al. [7]. First of all, the preference change leads to an upward or downward turn of optimal consumption at the threshold $W_{\text{ref}}(h)$ instead of a relatively smooth curve in Deng et al. [7]. Moreover, we emphasize that the influences of the preference change on the optimal consumption choice are global. For example, adjusting the risk aversion from $\beta_2$ to $\beta_1 (< \beta_2)$ in the region $x > W_{\text{ref}}(h)$ can change the consumption behavior in the region $x < W_{\text{ref}}(h)$. To wit, knowing that her/his risk attitude will change when $x > W_{\text{ref}}(h)$, she/he chooses to adjust the consumption behavior before reaching the reference threshold $W_{\text{ref}}(h)$. Regarding the risky investment proportion, the preference change results in a significant increase or decrease with a wide range of wealth levels. The preference change, together with the drawdown constraint, leads to a different optimal investment proportion curve compared with Deng et al. [7]. Another advantage of the present model is that considering preference change highlights the importance of the wealth threshold $W_{\text{ref}}(h)$ that has been more or less neglected. For example, the risky investment proportion attains its maximum around $W_{\text{ref}}(h)$, and the value function has very different sensitivities to wealth shock on different sides of the curve $x = W_{\text{ref}}(h)$. The differences between our paper and Deng et al. [7], as well as other implications of our model, will become apparent in Sect. 5.

To sum up, the model studied in this paper combines three aspects of the economic and psychological background of the consumption and investment problem: (1) the reference point is modeled through the running maximum of past consumption; (2) a drawdown constraint is imposed on consumption; (3) risk aversion is changed at the reference point. We apply the solution method inspired by Deng et al. [7] and other celebrated tools. We derive the HJB equation of the problem, obtain the optimal consumption and portfolio strategies thanks to the duality theory, and numerically analyze the solution and sensitivity to parameters. From this simple and intuitive model, we find several interesting economic implications, such as MPC jump and RRA smile.

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

1.1 Related literature

Our model is based on a dynamic consumption-investment decision model, whose classical form dates back to Merton [16]. A wide range of literature extends this problem by taking habit formation into account; see Pollak [19], Detemple and Zapatero [8], and Chapman [5] for instance. The habit process and the form of preference are two essential components of habit formation. The habit process is a process whose value at time $t$ is determined by the consumption process up to time $t$. On the other hand, in habit formation, the decision maker’s utility depends on both the consumption process and the habit process, which describes how consumption habit affects current consumption decisions.

One can just set the habit process as the average of past consumption. In this direction, a more reasonable and flexible choice of the habit process is the so-called linear habit formation,
i.e., the weighted average of past consumption. Such a habit process has dominated the research in habit formation since early literature such as Ryder and Heal [20]. Recently, taking the running maximum process of past consumption as the habit process opens 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. This property brings about mathematical challenges because it is a singular control problem. We follow the running maximum habit formation model recently studied by Guasoni et al. [10] and Deng et al. [7]. However, our model considers more factors that influence the decision maker’s consumption and portfolio choices, including the constraint on consumption and the change in risk aversion.

How utility depends on the consumption $c$ and habit $h$ is another topic in habit formation. A typical case is that the utility is a function of $c - h$, suggesting that only excess consumption matters. A wide range of literature, for instance, Chapman [5], adopts such a habit formation preference setting. A more flexible model, in Deng et al. [7], requires the utility to depend on $c - \alpha h$, where $\alpha$ varies in $[0, 1]$. Guasoni et al. [10] uses a different approach by setting the utility to be a function of $\frac{c}{h}$, where $\alpha \in (0, 1)$. A common feature of these two choices is that a higher level of habit tends to reduce utility.

In addition, how consumption is constrained is also important when studying habit-related models. Usually, the constraint imposes a lower bound on consumption. In the extreme case, the consumption is prohibited from falling below the habit, which is termed as addictive habit formation; see Muraviev [17] and Yu [24] for linear habit, Dybvig [9] and Jeon and Park [13] for running maximum. In other literature, the consumption is constrained to be no less than a fraction of the habit, such as Arun [2] and Angoshtari et al. [1]. In this paper, we also impose such drawdown constraint, in which we require $c \geq \lambda h$, where $\lambda \in (0, \alpha)$. 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 [2] and Angoshtari et al. [1] obtain a threshold of wealth/habit ratio below which the agent chooses to consume at the lowest constrained level. We derive a similar phenomenon, but with a rather complicated threshold curve ($x = W_{\text{low}}(h)$) of wealth-habit pair instead of a simple ray.

As a more general topic, the utility with reference points has been widely studied in different economic problems, such as Jin and Zhou [14] and He and Yang [12]. Another standard model under this paradigm is the S-shaped utility developed by Kahneman and Tversky. For example, Li et al. [15] studies S-shaped utility in the context of running maximum habit formation. However, such non-concave utility results in an extreme strategy where the decision-maker never consumes between 0 and the reference point. Instead, we analyze a utility with different risk aversions on different sides of the reference point, but the utility is concave. As a result, we obtain a more reasonable optimal strategy where the optimal consumption varies from the lower bound $\lambda h$ to the running maximum $h$ continuously. Another related work is Van Bilsen et al. [22], which permits the agent to be risk averse in the loss domain (setting $\gamma_L > 1$ therein). This setting is consistent with ours, but they choose to work with CRRA functions.

## 2 Model formulation

The financial market consists of one risk-free asset and one risky asset in our model. The risk-free asset $\{S^0_t, t \geq 0\}$ and the risky asset $\{S^1_t, t \geq 0\}$ satisfy

$$
\begin{align*}
\begin{cases}
dS^0_t &= S^0_t r \, dt, \\
dS^1_t &= S^1_t [\mu \, dt + \sigma \, dB_t],
\end{cases}
\end{align*}
$$
where \( r > 0 \) is the constant interest rate, \( \mu \geq r \) is the expected return and \( \sigma > 0 \) is the volatility. \( B = \{ B_t, t \geq 0 \} \) is a standard Brownian motion on the filtered probability space \((\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})\) satisfying the usual conditions.

The decision-maker determines her/his dynamic spending rate \( c \triangleq \{ c_t, t \geq 0 \} \) and the dynamic amount of her/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 her/his wealth process \( \{ X_t, t \geq 0 \} \) follows the following stochastic differential equation (abbr. SDE):

\[
\begin{align*}
    dX_t &= rX_t dt + \pi_t(\mu - r) dt + \pi_t \sigma dB_t - c_t dt, \\
    X_0 &= x_0.
\end{align*}
\] (2.1)

Given the consumption process \( \{ c_t, t \geq 0 \} \), the running maximum process is defined by

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

and \( h_0 \) is the inherited running maximum level. The wealth level is naturally required to be non-negative under any admissible strategy to avoid bankruptcy. 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

\[
\begin{align*}
    \int_0^t c_s ds &< \infty, \quad a.s., \\
    \int_0^t \pi_s^2 ds &< \infty, \quad a.s., \\
    c_t &\geq \lambda h_t, \quad a.s., \\
    X_t &> 0, \quad a.s.,
\end{align*}
\]

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 satisfying SDE (2.1). We denote by \( \mathcal{A} \) the set of admissible strategies.

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

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

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

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

where \( 0 \leq \lambda \leq \alpha < 1, \beta_1, \beta_2 > 0 \), and \( \alpha h \) is the reference point. The utility takes the form of constant absolute risk aversion (abbr.CARA) on both sides of \( \alpha h \). The absolute risk aversion above the reference point \( \alpha h \) is \( \beta_2 \), whereas the absolute risk aversion below the reference point is \( \beta_1 \). The utility is concave, and the marginal utilities at the two sides of the reference point \( \alpha h \) 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 the risk aversion change. However, our model does not reduce to the model without reference point $\alpha h$ if $\alpha$ approaches 1 from below. The reason lies in the fact that the updating of running maximum is under the risk aversion coefficient $\beta_2$ when $\alpha < 1$ but under $\beta_1$ when $\alpha = 1$ (the case $\alpha = 1$ is the model without reference point $\alpha h$). In other words, the model is somewhat “not continuous” at $\alpha = 1$. Hence, we only consider the case $\alpha < 1$. Another two limiting cases $\beta_1 \to 0$ and $\beta_2 \to 0$ are discussed in Sect. 6.2.3.

Remark 2 Alternatively, we can consider a more general reference point instead of $\alpha h$. The techniques we used still work. Using these techniques and tools, we give the results with an alternative reference point in Appendix D.

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

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

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

Proof Suppose that $(c, \pi) \in A$ and there exists a $t_0$ such that $\mathbb{P}(X_{t_0} < \frac{\lambda}{r} h_{t_0}) > 0$. Then there exists a $\varepsilon > 0$ 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 the lowest consumption, i.e., $c_t = \lambda h_{t_0}$, $\forall t \geq t_0$, there exists a $t_1 > t_0$ such that $\mathbb{P}(X_{t_1} < 0) > 0$. For strategies with $c_t = \lambda h_{t_0}$, $\forall t \geq t_0$, the wealth process is determined by

$$dX_t = r X_t dt + \pi_t (\mu - r)dt + \pi_t \sigma dB_t - \lambda h_{t_0} dt, \quad \forall t \geq t_0. \quad (2.4)$$

Solving this SDE, we obtain

$$X_t = e^{r(t-t_0)} X_{t_0} + \frac{\lambda h_{t_0}}{r} \left[1 - e^{r(t-t_0)}\right] + (\mu - r) e^{rt} \int_{t_0}^{t} e^{-ru} \pi_u du$$

$$+ \sigma e^{rt} \int_{t_0}^{t} e^{-ru} \pi_u dB_u, \quad \forall t \geq t_0. \quad (2.4)$$

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 the last two terms in RHS of (2.4), we introduce the probability measure $\tilde{\mathbb{P}}$ by

$$d\tilde{\mathbb{P}} \bigg|_{\mathcal{F}_t} := e^{-\frac{\mu - r}{\sigma} B_t - \frac{(\mu - r)^2}{2\sigma^2} t}, \quad \forall t \geq 0.$$
Then \( \{ \tilde{B}_t := B_t + \mu \mathbb{P} - r \tilde{\sigma} b u, \quad t \geq 0 \} \) is a standard Brownian motion under \( \tilde{\mathbb{P}} \) and the sum of the last two terms equals \( \sigma \int_0^t e^{-ru} \pi_u d \tilde{B}_u \), which has zero expectation under \( \tilde{\mathbb{P}} \). Hence,

\[
\tilde{\mathbb{P}} \left( \sigma \int_0^t e^{-ru} \pi_u d \tilde{B}_u \leq 0 \right) > 0, \quad \forall t \geq 0.
\]

Consequently,

\[
\mathbb{P} \left( (\mu - r) \int_0^t e^{-ru} \pi_u du + \sigma \int_0^t e^{-ru} \pi_u dB_u \right) > 0, \quad \forall t \geq 0.
\]

Because \((c, \pi) \in \mathcal{A})\), we deduce that \(X_{t_0}^0\) and \(h_{t_0}\) are \( \mathcal{F}_{t_0} \) measurable. Thus the sum of the first two terms and the sum of the last two terms of (2.4) are independent. Choosing \( t_1 = t_0 + \frac{1}{r} \ln \left( \frac{1}{\varepsilon} \right) \), we have \( \mathbb{P}(X_{t_1} < 0) > 0 \) and thus the proof is completed. \( \square \)

### 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 theory to obtain the solution in dual form.

Denote the value function of Problem (2.3) 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 \((c^*, \pi^*)\) of Problem (2.3) is an admissible strategy satisfying

\[
\mathbb{E}_{x_0, h_0} \int_0^\infty e^{-\gamma t} U(c^*_t, h^*_t) dt = V(x_0, h_0).
\]

The martingale optimality principle indicates 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 \((c, \pi)\) and is also a local martingale for the optimal \((c^*, \pi^*)\). 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\(^1\):

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

\(^1\) For simplicity, we write \( x, h, c, \pi \) instead of \( X_t, h_t, c_t, \pi_t \) 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 \( \pi \) is
\[
\pi_{\text{primal}}^*(x, h) = -\frac{\mu - r}{\sigma^2} V_x(x, h).
\]

As \( \frac{\xi}{\eta} \to \frac{\lambda}{\tau} \), the optimal investment should decline to zero to avoid bankruptcy, i.e.,
\[
\lim_{\frac{\xi}{\eta} \to \frac{\lambda}{\tau}^+} \frac{V_x(x, h)}{V_{xx}(x, h)} = 0. \tag{3.2}
\]

To solve HJB equation (3.1) with the boundary condition (3.2) based on the duality theory, 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 dual transform
\[
\begin{cases}
x = -\tilde{V}_y(y, h), \\
V(x, h) = \tilde{V}(y, h) - y\tilde{V}_y(y, h), \\
V_x(x, h) = y, \\
V_{xx}(x, h) = -\frac{1}{V_{yy}(y, h)}, \\
V_h(x, h) = \tilde{V}_h(y, h),
\end{cases} \tag{3.3}
\]

and HJB equation (3.1) is rewritten in dual form as
\[
\begin{align*}
\sup_{c \in [0, h], \pi \in \mathbb{R}} & \left\{ -\gamma \left[ \tilde{V}(y, h) - y\tilde{V}_y(y, h) \right] + y \left[ -r\tilde{V}_y(y, h) + \pi(\mu - r) - c \right] \\
& - \frac{1}{2V_{yy}(y, h)} \sigma^2 y^2 + U(c, h) \right\} = 0, \\
\tilde{V}_h(y, h) &= 0 \text{ on } (y, h) \text{ s.t. } \frac{dh}{dt} \neq 0.
\end{align*} \tag{3.4}
\]

Using (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 \( \tilde{U}(c) \triangleq U(c, h) - cy \) on \([\lambda h, h]\) is
\[
c^*(y, h) = \begin{cases} 
\lambda h, & e^{(\alpha - \lambda)\beta_1 h} \leq y, \\
-\frac{1}{\beta_1} \ln(y) + \alpha h, & 1 \leq y < e^{(\alpha - \lambda)\beta_1 h}, \\
-\frac{1}{\beta_2} \ln(y) + \alpha h, & e^{-(1-\alpha)\beta_2 h} \leq y < 1, \\
h, & 0 \leq y < e^{-(1-\alpha)\beta_2 h}. \end{cases} \tag{3.6}
\]

Three sub-cases need to be distinguished when \( c^*(y, h) = h \). The first case is that the current consumption reaches the past consumption peak but does not update it; the second is that the current consumption updates it; the last is that the current consumption forces the running maximum process to jump. The last case only happens at \( t = 0 \), when the inherited running maximum level is low, but the initial wealth is abundant. It switches to the former two cases for \( t > 0 \). Thus 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 \( \arg\max_c \{U(c, c) - cy\} \geq h \). Moreover, the
running maximum \( h \) jumps if the strict inequality holds. As \( \text{argmax}_y \{ U(c, c) - cy \} \geq h \) 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. \( y_0 < (1 - \alpha) e^{-(1-\alpha)\beta_2 h} \), it will jump immediately to \( (y_0, \frac{1}{\gamma (1-\alpha)\beta_2} \ln \frac{1}{y_0}) \) which is on the curve \( y = (1 - \alpha) e^{-(1-\alpha)\beta_2 h} \). Thus we only need to consider initial points \((y_0, h_0)\) in the region \( y \geq (1 - \alpha) e^{-(1-\alpha)\beta_2 h} \). Meanwhile, the second equation of (3.4) is equivalent to

\[
V_h(y, h) = 0 \quad \text{for} \quad y = (1 - \alpha) e^{-(1-\alpha)\beta_2 h}.
\]  

(3.7)

By the dual 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 \( C_d \) is

\[
C_d \triangleq \left\{(y, h) \left| y \geq (1 - \alpha) e^{-(1-\alpha)\beta_2 h}, -\tilde{V}_y(y, h) \geq \frac{\lambda h}{\gamma}, h > 0 \right. \right\}.
\]

Applying the dual transform again, the effective region \( C \) is

\[
C \triangleq \left\{(x, h) \left| V_x(x, h) \geq (1 - \alpha) e^{-(1-\alpha)\beta_2 h}, x \geq \frac{\lambda h}{\gamma}, h > 0 \right. \right\}.
\]  

(3.8)

Using the first equation of (3.4), we obtain

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

(3.9)

where \( \tilde{U}(y, h) = \sup_{2\lambda h \leq e < \gamma h} \{ U(c, h) - cy \} \) is

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

(3.10)

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

\[
\tilde{V}(y, h) = \begin{cases}
C_1(h) y^{q_1} + C_2(h) y^{q_2} - \frac{1}{\gamma} \lambda y, & e^{(\alpha-\lambda)\beta_1 h} \leq y, \\
+ \frac{1}{\beta_1} [1 - e^{(\alpha-\lambda)\beta_1 h}],
\end{cases}
\]

\[
C_3(h) y^{q_1} + C_4(h) y^{q_2} + \frac{1}{\beta_1} [1 - y + y \ln(y)] + \frac{k}{\gamma^2} y - \frac{1}{\gamma} \lambda y, & 1 \leq y < e^{(\alpha-\lambda)\beta_1 h}, \\
C_5(h) y^{q_1} + C_6(h) y^{q_2} + \frac{1}{\beta_1} [1 - y + y \ln(y)] + \frac{k}{\gamma^2} y - \frac{1}{\gamma} \lambda y, & e^{-(1-\alpha)\beta_2 h} \leq y < 1,
\end{cases}
\]

(3.10)

Using the duality transform (3.3), the boundary condition (3.2) can be rewritten as

\[
\lim_{\tilde{V}_y(y, h) \to -\frac{\lambda h}{\gamma}} y \tilde{V}_{yy}(y, h) = 0.
\]  

(3.11)

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

\( \Rightarrow \) Springer
\[ C_d = \{ (y, h) \mid y \geq (1 - \alpha)e^{-(1-\alpha)\beta_2 h}, \ h > 0 \}. \tag{3.12} \]

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 \( \frac{1}{\gamma} \) for long time, and the value function tends to \( \frac{1}{\gamma} \). Thus we can express this boundary condition as

\[
\lim_{h \to \infty} \left[ \tilde{V}(y, h) - y \tilde{V}_y(y, h) \right] \bigg|_{y=(1-\alpha)e^{-(1-\alpha)\beta_2 h}} = \frac{1}{\gamma}. \tag{3.13}
\]

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

\[
\tilde{V}(y+, h) = \tilde{V}(y-, h), \quad \tilde{V}_y(y+, h) = \tilde{V}_y(y-, h)
\]

yields

\[
C_2(h) = 0, \quad C_4(h) = -\frac{k}{\gamma^2} \frac{1 - q_1}{\beta_1 q_2 - q_1} e^{-(\alpha-\lambda)(q_2-1)\beta_1 h}, \tag{3.14}
\]

\[
C_6(h) = C_4(h) + \frac{k}{\gamma^2} \frac{\beta_2 - \beta_1}{\beta_1 \beta_2 q_2 - q_1} \frac{1 - q_1}{q_2 - q_1}. \tag{3.15}
\]

\[
C_7(h) = \frac{(1 - \alpha)q_2 - q_1}{(1 - \alpha)(q_2 - q_1)} \frac{k}{\gamma^2} \frac{1 - q_1}{\beta_2 q_2 - q_1} \frac{1}{\beta_2 q_2 - q_1} e^{-(1-\alpha)(q_2-1)\beta_2 h} \nonumber \] \( + (1 - \alpha)q_2 - q_1 \frac{k}{\gamma^2} \frac{q_2 - 1}{\beta_2 q_2 - q_1} e^{-(1-\alpha)(1-q_1)\beta_2 h} \tag{3.16}
\]

\[
C_5(h) = C_7(h) - \frac{k}{\gamma^2} \frac{q_2 - 1}{\beta_2 q_2 - q_1} e^{-(1-\alpha)(1-q_1)\beta_2 h}, \tag{3.17}
\]

\[
C_3(h) = C_5(h) - \frac{k}{\gamma^2} \frac{\beta_2 - \beta_1}{\beta_1 \beta_2 q_2 - q_1} \frac{q_2 - 1}{q_2 - q_1}. \tag{3.18}
\]

We directly show

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

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

**Lemma 3.2** As \( h \to \infty \),

\[
C_1(h) = O(e^{(\alpha-\lambda)(1-q_1)\beta_1 h}), \quad C_3(h) = O(1), \quad C_4(h) = O(e^{-(\alpha-\lambda)(q_2-1)\beta_1 h}), \nonumber \] \( C_5(h) = O(e^{-(1-\alpha)(1-q_1)\beta_2 h}), \quad C_6(h) = O(1), \nonumber \] \( C_7(h) = O(e^{-(1-\alpha)(1-q_1)\beta_2 h}), \quad C_8(h) = O(e^{(1-\alpha)(q_2-1)\beta_2 h}). \)
4 Verification theorem and optimal strategy

In this section, we establish the verification theorem and apply the dual transform to obtain the optimal strategies. First, we state the verification theorem in dual form.

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

\[
(c^*, \pi^*) = \left\{ \left( c^*(Y_t(y^*), H_t^*(y^*)), \pi^*(Y_t(y^*), H_t^*(y^*)) \right) \mid t \geq 0 \right\},
\]

where \(Y_t(\cdot)\) is

\[
Y_t(y) \triangleq ye^{\gamma t} M_t
\]

with \(M \triangleq \left\{ M_t := e^{-(\gamma + (\alpha - \lambda)^2/2\sigma^2)t - \frac{\alpha - \lambda}{\sigma} B_t}, t \geq 0 \right\}\) being the discounted state price density process, \(H_t^*(\cdot), t \geq 0\) is

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

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

\[
\mathbb{E}_{x_0, h_0} \int_0^\infty c^*(Y_t(y), H_t^*(y)) M_t dt = x_0,
\]

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

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

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

**Proof** See Appendix A. \(\square\)

We need the following lemma about the dual transform to express the value function and optimal strategies in primal variables. The proof is in Appendix C.

**Lemma 4.2** \(\tilde{V}_{yy}(y, h) > 0\) for \((y, h) \in C_d\), and 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. Using dual transform (3.3), we have

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

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

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

Substituting (3.10) into (4.5), we obtain the following forms of $e(\alpha-\lambda)\beta_1h \leq f(x, h)$ according to different regions:

1. For $e(\alpha-\lambda)\beta_1h \leq f(x, h)$, we have $f(x, h) = f_1(x, h)$ with $f_1(x, h)$ satisfying

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

Based on Lemma 4.2, $V_{yy}(\cdot, h) > 0$, the inequality $e(\alpha-\lambda)\beta_1h \leq f(x, h)$ is equivalent to $x \leq W_{\text{low}}(h)$ with

$$W_{\text{low}}(h) = -C_1(h)q_1e^{-(\alpha-\lambda)(1-q_1)\beta_1h} - C_2(h)q_2e^{(\alpha-\lambda)(q_2-1)\beta_1h} + \frac{\lambda h}{\gamma}. \quad (4.7)$$

2. For $1 \leq f(x, h) < e(\alpha-\lambda)\beta_1h$, $f(x, h) = f_2(x, h)$ with $f_2(x, h)$ satisfying

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

The inequality $1 \leq f(x, h) < e(\alpha-\lambda)\beta_1h$ is equivalent to $W_{\text{low}}(h) < x \leq W_{\text{ref}}(h)$ with

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

3. For $e^{-(1-\alpha)\beta_2h} \leq f(x, h) < 1$, $f(x, h) = f_3(x, h)$ with $f_3(x, h)$ satisfying

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

The inequality $e^{-(1-\alpha)\beta_2h} \leq f(x, h) < 1$ is equivalent to $W_{\text{ref}}(h) < x \leq W_{\text{peak}}(h)$ with

$$W_{\text{peak}}(h) = -C_5(h)q_1e^{-(1-\alpha)(1-q_1)\beta_2h} - C_6(h)q_2e^{-(1-\alpha)(q_2-1)\beta_2h} - \frac{k}{\gamma^2\beta_2} + \frac{h}{\gamma}. \quad (4.11)$$

4. For $(1-\alpha)e^{-(1-\alpha)\beta_2h} \leq f(x, h) < e^{-(1-\alpha)\beta_2h}$, $f(x, h) = f_4(x, h)$ with $f_4(x, h)$ satisfying

$$x = -C_7(h)q_1\left[f_4(x, h)\right]^{q_1-1} - C_8(h)q_2\left[f_4(x, h)\right]^{q_2-1} + \frac{h}{\gamma}. \quad (4.12)$$
The inequality \((1 - \alpha)e^{-(1-\alpha)\beta_2 h} \leq f(x, h) < e^{-(1-\alpha)\beta_2 h}\) is equivalent to \(W_{\text{peak}}(h) < x \leq W_{\text{updt}}(h)\) with

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

In addition, we denote the lower bound for wealth by

\[
W_{\text{bkrp}}(h) = \frac{\lambda h}{\gamma}. \tag{4.14}
\]

In the following Theorems 4.3 and 4.4, we summarize the primal forms of the value function and the optimal strategies.

**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

\[
V(x, h) = \begin{cases} 
C_1(h)[f_1(x, h)]q_1 + C_2(h)[f_1(x, h)]q_2 \\
-\frac{1}{\gamma} \lambda hf_1(x, h) + \frac{1}{\gamma \mu_1}(1 - e^{(\alpha-\lambda)\beta_1 h}) + x f_1(x, h), & W_{\text{bkrp}}(h) \leq x \leq W_{\text{low}}(h), \\
C_3(h)[f_2(x, h)]q_1 + C_4(h)[f_2(x, h)]q_2 \\
+ \frac{1}{\gamma \mu_1} \ln \left[ f_2(x, h) \right] \\
+ \frac{k}{\gamma^2 \mu_1} f_2(x, h) - \frac{1}{\gamma} \alpha h f_2(x, h) + x f_2(x, h), & W_{\text{low}}(h) < x \leq W_{\text{ref}}(h), \\
C_5(h)[f_3(x, h)]q_1 + C_6(h)[f_3(x, h)]q_2 \\
+ \frac{1}{\gamma \mu_2} \ln \left[ f_3(x, h) \right] \\
+ \frac{k}{\gamma^2 \mu_2} f_3(x, h) - \frac{1}{\gamma} \alpha h f_3(x, h) + x f_3(x, h), & W_{\text{ref}}(h) < x \leq W_{\text{peak}}(h), \\
C_7(h)[f_4(x, h)]q_1 + C_8(h)[f_4(x, h)]q_2 - \frac{1}{\gamma} h f_4(x, h) \\
+ \frac{1}{\gamma \mu_2} \ln \left[ f_4(x, h) \right] \\
+ \frac{k}{\gamma^2 \mu_2} f_4(x, h) - \frac{1}{\gamma} \alpha h f_4(x, h) + x f_4(x, h), & W_{\text{peak}}(h) < x \leq W_{\text{updt}}(h),
\end{cases}
\]

where \(W_{\text{bkrp}}(h), W_{\text{low}}(h), W_{\text{ref}}(h), W_{\text{peak}}(h), W_{\text{updt}}(h)\) and \(f_i(x, h), 1 \leq i \leq 4\) are given by (4.6)–(4.14).

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

\[
V(x, h) = \tilde{V}(f(x, h), h) + x f(x, h).
\]

Using (3.10), the desired result follows. \qed

**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 variables:

\[
c^*_{\text{primal}}(x, h) = \begin{cases} 
\lambda h, & W_{\text{bkrp}}(h) \leq x \leq W_{\text{low}}(h), \\
-\frac{1}{\mu_1} \ln \left[ f_2(x, h) \right] + ah, & W_{\text{low}}(h) < x \leq W_{\text{ref}}(h), \\
-\frac{1}{\mu_2} \ln \left[ f_3(x, h) \right] + ah, & W_{\text{ref}}(h) < x \leq W_{\text{peak}}(h), \\
h, & W_{\text{peak}}(h) < x \leq W_{\text{updt}}(h),
\end{cases}
\]
\[ \pi_{\text{primal}}^*(x, h) = \frac{\mu - r}{\sigma^2} \]

Lemma 4.6

\[ \pi_{\text{primal}}^*(x, h) = \frac{\mu - r}{\sigma^2} \]

where \( W_{\text{bkpr}}(h), W_{\text{low}}(h), W_{\text{ref}}(h), W_{\text{peak}}(h), W_{\text{updt}}(h) \) and \( f_i(x, h), 1 \leq i \leq 4 \) are given by (4.6)–(4.14).

Then the SDE

\[
\begin{align*}
    dX_t &= rX_t dt + \pi_{\text{primal}}^*(X_t, H_t^*) (\mu - r) dt + \pi_{\text{primal}}^*(X_t, H_t^*) \sigma dW_t - c_{\text{primal}}^*(X_t, H_t^*) dt, \\
    X_0 &= x_0.
\end{align*}
\]

where \( H_t^* \equiv h_0 \vee \sup_{s \leq t} c_{\text{primal}}^*(X_s, H_s^*) \) and \( H_0^* = h_0 \) has a unique strong solution \( \{X_t^*, t \geq 0\} \). The optimal consumption and investment strategy is

\[
\left\{ (c_{\text{primal}}^*(X_t^*, H_t^*), \pi_{\text{primal}}^*(X_t^*, H_t^*)), t \geq 0 \right\}.
\]

Proof

The proof is based on the following Lemmas 4.5 and 4.6. As the proof is similar to that of Deng et al. [7], we omit it here.

Lemma 4.5

The function \( f \) is \( C^1 \) within each sub-region of \( \mathcal{C} \): \( W_{\text{bkpr}}(h) \leq x \leq W_{\text{low}}(h), W_{\text{low}}(h) < x \leq W_{\text{ref}}(h), W_{\text{ref}}(h) < x \leq W_{\text{peak}}(h), W_{\text{peak}}(h) < x \leq W_{\text{updt}}(h) \), and it is continuous at the boundary of \( x = W_{\text{low}}(h), x = W_{\text{ref}}(h), x = W_{\text{peak}}(h) \). Moreover,

\[
f_x(x, h) = -\frac{1}{V_{xy}(f, h)} \left\{ \begin{array}{l}
    \frac{r}{\kappa} \left[ -C_1(h) [f_1(x, h)]^{q_1 - 2} - C_2(h) [f_1(x, h)]^{q_2 - 2} \right]^{-1}, \quad W_{\text{bkpr}}(h) \leq x \leq W_{\text{low}}(h), \\
    \frac{r}{\kappa} \left[ -C_3(h) [f_2(x, h)]^{q_1 - 2} - C_4(h) [f_2(x, h)]^{q_2 - 2} \right]^{-1}, \quad W_{\text{low}}(h) < x \leq W_{\text{ref}}(h), \\
    \frac{r}{\kappa} \left[ -C_5(h) [f_3(x, h)]^{q_1 - 2} - C_6(h) [f_3(x, h)]^{q_2 - 2} \right]^{-1}, \quad W_{\text{ref}}(h) < x \leq W_{\text{peak}}(h), \\
    \frac{r}{\kappa} \left[ -C_7(h) [f_4(x, h)]^{q_1 - 2} - C_8(h) [f_4(x, h)]^{q_2 - 2} \right]^{-1}, \quad W_{\text{peak}}(h) < x \leq W_{\text{updt}}(h), \\
\end{array} \right.
\]

(4.16)

\[
f_h(x, h) = V_{yh}(f, h) f_x(x, h).
\]

(4.17)

Proof

The proof is similar to Lemma 5.7 in Deng et al. [7] and thus omitted here.

Lemma 4.6

The function \( c_{\text{primal}}^* \) is locally Lipschitz on \( \mathcal{C} \), and the function \( \pi_{\text{primal}}^* \) is Lipschitz on \( \mathcal{C} \).

Proof

See Appendix C.
We have obtained semi-explicit forms of the value function (in Theorem 4.3) and the optimal strategies (in Theorem 4.4). Now we introduce the following propositions on properties of the value function and the MPC out of wealth. The results are numerically illustrated in Sect. 5 (see Fig. 6 and the left panel of Fig. 7).

Proposition 4.7 The value function \( V(x, h) \) has the following properties:

(i) \( V(x, h) \) is increasing in \( x \) and decreasing in \( h \).

(ii) \( V_x(x, h) \) is decreasing in \( x \) with

\[
\lim_{x \downarrow W_{\text{bkrp}}(h)} V_x(x, h) = +\infty,
\]

\[
\lim_{x \uparrow W_{\text{updt}}(h)} V_x(x, h) = (1 - \alpha)e^{-(1 - \alpha)\beta_2 h},
\]

and \( V_x(x, h) > 1 \) for \( x < W_{\text{ref}}(h) \) whereas \( V_x(x, h) < 1 \) for \( x > W_{\text{ref}}(h) \).

**Proof** Proof of (i):

Since \( V(x, h) = \tilde{V}(f(x, h), h) + xf(x, h) \), we have

\[
V_x(x, h) = \tilde{V_y}(f(x, h), h) f_x(x, h) + f(x, h) + xf_x(x, h) \\
= -xf_x(x, h) + f(x, h) + xf_x(x, h) \\
= f(x, h) > 0.
\]

Hence \( V(x, h) \) is increasing in \( x \). To show it is decreasing in \( h \), we apply the martingale optimality principle and obtain as a byproduct of (3.1) that \( V_h(x, h) \leq 0 \). We only have to rule out the possibility that \( V_h(x, h) = 0 \) in some intervals. Using \( V(x, h) = \tilde{V}(f(x, h), h) + xf(x, h) \) again, we obtain \( V_h(x, h) = \tilde{V}_h(f(x, h), h) \). Thus \( V_h(x, h) = 0 \) cannot hold in any interval thanks to (3.10) and (3.13)–(3.18).

**Proof of (ii):**

The concave property of \( V(x, h) \) follows from the duality relation \( V_{xx}(x, h) = -\frac{1}{V_{yy}(y, h)} \) and Lemma 4.2. The rest follows by the duality relation \( V_x(x, h) = f(x, h) \). \( \square \)

Remark 3 Thanks to Proposition 4.7, we note that the value function (welfare) is very sensitive to wealth shocks when \( x \) is close to \( W_{\text{bkrp}} \) because \( V_x \) tends to infinity. On the other hand, \( V_x \) is very small when \( x \) is close to \( W_{\text{bkrp}}(h) \) because of the negative exponent. In addition, \( W_{\text{ref}}(h) \) is a critical point where \( V_x = 1 \). For numerical illustration, see Fig. 6 and related analysis in Sect. 5.

Proposition 4.8 The MPC out of wealth, i.e., \( \frac{\partial c^*(x, h)}{\partial x} \), has the following properties:

(i) The MPC out of wealth equals 0 for \( x \in [W_{\text{bkrp}}, W_{\text{low}}(h)] \cup [W_{\text{peak}}(h), W_{\text{updt}}(h)] \).

(ii) The MPC out of wealth shrinks or swells by \( \frac{\beta_1}{\beta_2} \) when wealth exceeds \( W_{\text{ref}}(h) \).

(iii) There exists a unique \( \tilde{h} > 0 \) solving

\[
C_5(h)(q_1 - 1)e^{-(q_1 - 2)(1 - \alpha)\beta_2 h} + C_6(h)(q_2 - 1)e^{-(q_2 - 2)(1 - \alpha)\beta_2 h} = 0. \tag{4.18}
\]

If \( h \leq \tilde{h} \), the MPC out of wealth is decreasing for \( x \in [W_{\text{low}}, W_{\text{peak}}(h)] \). If \( h > \tilde{h} \), there exists an \( \tilde{x}(h) \in [W_{\text{low}}(h), W_{\text{peak}}(h)] \) such that the MPC out of wealth is decreasing for \( x \in [W_{\text{low}}(h), \tilde{x}(h)] \) and increasing for \( x \in [\tilde{x}(h), W_{\text{peak}}(h)] \).
\[ \bar{h} = \frac{\ln(M_1) - \ln(M_2)}{(\alpha - \lambda)(q_2 - 1)\beta_1} \]

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

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

**Proof** See Appendix C. □

**Remark 4** Numerically, the \( \bar{h} \) in (iii) is relatively small. Say, \( \bar{h} \) is approximately 0.1 with parameters \( \lambda = 0.3, \alpha = 0.7, \beta_1 = 1, \beta_2 = 2, r = 0.04, \mu = 0.12, \sigma = 0.3 \). Meanwhile, the \( \bar{h} \) in (iv) is relatively large, e.g., \( \bar{h} \) is about 6.6 under the above parameters.

**Remark 5** As a distinctive result of our model, it is interesting that the MPC out of wealth can increase when \( x \) is sufficiently large. This only happens when \( \beta_1 < \beta_2 \), i.e., when considering the gambling-type risk attitude. However, as discussed in Remark 4, the threshold \( \bar{h} \) is large with reasonable model parameters, thus the increasing MPC only happens with extremely high wealth levels. Therefore we choose not to discuss this result in depth from an economic point of view.

## 5 Numerical analysis with fixed parameters

This section aims to illustrate and analyze the properties of the optimal strategies and relevant boundaries. Throughout this section, we fix parameters \( \lambda = 0.3, \alpha = 0.7, \beta_1 = 1, \beta_2 = 2, r = 0.04, \mu = 0.12, \sigma = 0.3 \).

We first investigate the boundaries. The effective region is between two boundaries \( x = W_{\text{updt}}(h) \) and \( x = W_{\text{bkrp}}(h) \). Using \( x = W_{\text{low}}(h), x = W_{\text{ref}}(h), \) and \( x = W_{\text{peak}}(h) \), the effective region is further divided into four parts; see sub-regions I–IV in Fig. 1. Such a state space division helps to provide a structural description of both consumption and investment behaviors under our model. In particular, the decision-maker keeps the consumption at the lowest constrained level in sub-region IV, and she/he revisits her/his past spending maximum in sub-region I. On the upper boundary \( x = W_{\text{updt}}(h) \) of sub-region I, the decision-maker continuously updates the running maximum of her/his consumption process. In sub-regions II and III, the decision-maker dynamically chooses her/his consumption according to her/his wealth \( x \) and standard of living \( h \); see Figs. 3 and 7. Based on the aforementioned economic interpretations, we name the curve \( W_{\text{updt}}(h) \) as the “bliss” curve, and four sub-regions I–IV as the “satisfactory”, “recovery”, “depression” and “gloom” regions, respectively. See Fig. 2
Fig. 1 Boundary curves \( x = W_{\text{bkrp}}(h), x = W_{\text{low}}(h), x = W_{\text{ref}}(h), x = W_{\text{peak}}(h) \) and \( x = W_{\text{updt}}(h) \), four different sub-regions I–IV (I: satisfactory, II: recovery, III: depression, IV: gloom) of the effective region, and two sub-regions \( C^c_i, i = 1, 2 \) of the ineffective region with \( \lambda = 0.3, \alpha = 0.7, \beta_1 = 1, \beta_2 = 2, r = 0.04, \mu = 0.12, \sigma = 0.3 \).

Fig. 2 Optimal consumption (left) and optimal risky investment proportion (right) when fixing an \( h \) for an illustration. We emphasize that the division is based on the \((x, h)\) pair, instead of the wealth only. As a result, it is possible that a poor person (with a low absolute level of wealth) is in the satisfactory region, while a person with more wealth falls into the gloom region.

On the other hand, the ineffective region is separated into two parts \( C^c_1 \) and \( C^c_2 \):

\[
C^c_1 = \{ (x, h) | x < W_{\text{bkrp}}(h), h > 0 \}, \\
C^c_2 = \{ (x, h) | x > W_{\text{updt}}(h), h > 0 \}.
\]

\( C^c_1 \) means that the wealth is too low to maintain the lowest consumption level \( c = \lambda h \). The other part \( C^c_2 \) implies that the wealth level is very high compared to the current standard of living (it only happens at time \( t = 0 \) under the optimal strategies). In this case, it is optimal to consume at a level strictly higher than the running maximum and force \((x, h)\) to jump to \( x = W_{\text{updt}}(h) \).
Next, we consider the optimal strategies. The optimal consumption, shown in Fig. 3, is non-decreasing in both the wealth $x$ and the habit $h$. The optimal consumption in the satisfactory and gloom regions is indifferent in $x$. People in the gloom region consume as little as possible, whereas those in the satisfactory region revisit their historical consumption peak. However, in the depression region and recovery region, increasing wealth will lead to an increase in optimal consumption. The increase is more substantial in the region with lower risk aversion (i.e., more substantial in the depression region when $\beta_1 < \beta_2$, and in the recovery region when $\beta_1 > \beta_2$). The above analysis suggests that an increase in wealth causes one to consume more in two cases. The first case is that, on the bliss curve, she/he decides to continuously update her/his maximal spending $h$; the other case is that her/his wealth is more than $W_{\text{low}}(h)$ but less than $W_{\text{peak}}(h)$. In the latter case, there are two sub-cases divided by whether the wealth is more than $W_{\text{ref}}(h)$, and the (marginal propensity to consume) MPC out of wealth is generally higher in the region with lower risk aversion.

The optimal portfolio is shown in Fig. 4. The behavior of the optimal portfolio varies significantly in different regions. For fixed habit $h$, the optimal portfolio dramatically increases with the variable $x$ in the gloom and depression regions where the risk aversion is low. However, once crossing $x = W_{\text{ref}}(h)$ and the risk aversion shifting to the high level, increasing wealth causes the optimal portfolio to fall. The above result indicates that the change in risk aversion has an overwhelming impact on portfolio selection in our model. For those in the gloom and depression regions, the increase in their wealth stimulates them to invest more in risky assets. In contrast, for people in recovery and satisfactory regions, as well as on the bliss curve, the more they earn, the less they are willing to invest in risky assets. Similar conclusions can be obtained from analyzing the optimal proportion of wealth invested in risky assets; see Fig. 5.

The value function shown in Fig. 6 is increasing in wealth $x$ and decreasing in habit $h$ (see Proposition 4.7(i)). The result suggests that higher initial wealth and lower 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 the gloom region, the value function will fall dramatically due to a slight decrease in $x$. Nevertheless, in regions above $x = W_{\text{ref}}(h)$, the value function is not sensitive to $x$. The results show that welfare is much

---

2 Each of Figs. 3, 4, 5 and 6 consists of three parts. The left panels are graphs of the corresponding two-variable functions. The middle panels are two-dimensional projections of the left panels, displayed as heat maps. The boundaries are marked on the middle panels of these figures. For a clearer presentation of the information in the middle panels, we put the corresponding contour maps in the right panels.
more vulnerable to wealth shocks below $W_{\text{ref}}(h)$. To understand this fact theoretically, see Proposition 4.7(ii) and Remark 3.

We are particularly interested in the decisions of consumption and risky investment proportion in terms of wealth, fixing a standard of living, i.e., the functions $x \mapsto c^*(x, h)$ and $x \mapsto \pi^*(x, h)/x$. We present them in Fig. 7.

From the left panel of Fig. 7, we note that the sensitivity of the optimal consumption to wealth, or economically speaking, the marginal propensity to consume (MPC) out of wealth, generally decreases with growing wealth, which is admitted in economic literature. This is indeed the case when $h$ is small; see Proposition 4.8(iii). When $h$ is large, a dedicated analysis [Proposition 4.8 (iii)–(iv)] shows that MPC out of wealth is still decreasing in the lower part
of the depression-recovery region and on the bliss curve. However, it is increasing in the upper part of the depression-recovery region. Besides, MPC out of wealth shrinks or swells by $\beta_1/\beta_2$ at $W_{\text{ref}}(h)$, which is itself an interesting fact; see Proposition 4.8(ii). For a more illustrative version of the optimal consumption when fixing $h$, see Fig. 2.

From the right panel of Fig. 7, it is clear that for the sampled $h$, the risky investment proportion is a hump in variable $x$. The peak is around $W_{\text{ref}}(h)$ (see also Fig. 5). If we interpret the inverse of risky investment proportion as the so-called *implied relative risk aversion* [13], we find that it will be a smile in wealth. That is to say, people with intermediate wealth levels relative to their consumption peak have the lowest risk aversion and the highest risk tolerance. People with either relatively low or relatively high levels of wealth are much more risk-averse. This effect comes intuitively from our model settings. People whose wealth level is low relative to their consumption peak must keep their deposit above $W_{\text{bkrp}}(h)$ to satisfy the lowest consumption constraint. Hence they are very sensitive to risk. Therefore, it is reasonable for them to keep most of their wealth in safe assets. On the other hand, people with relatively high wealth levels have already been satisfied by the current level of consumption (or they even continuously update their maximal spending). Thus, they tend to avoid the risk to keep consumption above the reference level $\alpha h$. An illustrative version of the optimal risky investment proportion when fixing $h$ can be found in Fig. 2.

6 Sensitivity analysis

6.1 Impact of $\lambda$ and $\alpha$ on thresholds

The parameter $\alpha$ determines the reference point $\alpha h$ where risk aversion changes. When $\alpha$ equals $\lambda$, our model reduces to the model without the risk aversion related reference point.

---

3 Different from other related papers, the marginal utility of consumption is continuous at the reference point $c = \alpha h$, but we still document such an MPC shrink or MPC swell. This fact also indicates that the MPC in the recovery region can be globally lower or higher than that in the depression region. It is lower in our numerical result (left panel of Fig. 7).

4 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.
and the boundary $x = W_{\text{ref}}(h)$ coincides with $x = W_{\text{low}}(h)$. When $\alpha$ approaches 1 from below, the reference point approaches the running maximum, and the boundary $x = W_{\text{ref}}(h)$ tends to $x = W_{\text{peak}}(h)$. To figure out the influence of $\alpha$ on different boundaries, we fix $\lambda = 0.2$, $\beta_1 = 1$, $\beta_2 = 2$, $r = 0.04$, $\mu = 0.12$, $\sigma = 0.3$, let $\alpha$ vary from $\lambda$ to $1 - 10^{-8}$ and present the boundaries separately. We do not consider the boundary $W_{\text{bkrp}}(h)$ because it does not depend on $\alpha$. As shown in Fig. 8, the boundary $x = W_{\text{peak}}(h)$ moves downward as $\alpha$ increases, whereas the other three boundaries $x = W_{\text{low}}(h)$, $W_{\text{ref}}(h)$ and $W_{\text{updt}}(h)$ have the tendency to move upward. The above phenomenon illustrates that, as $\alpha$ increases, less wealth is needed for the decision-maker to consume at the historical running maximum level. On the contrary, more wealth is needed to escape the gloom and depression regions. The change of $\alpha$ is the most influential on $x = W_{\text{ref}}(h)$, which is the boundary separating the depression and recovery regions.

Another parameter $\lambda$ reflects the drawdown constraint on consumption. The consumption level is allowed to be relatively lower with smaller $\lambda$. When $\lambda$ 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 $\lambda$ equals $\alpha$, our model reduces to the model without the 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 $\alpha$ vary below $\lambda$ and graph the boundaries separately (boundary $x = W_{\text{bkrp}}(h)$ is not shown here as it is simply linear) to illustrate the impact of $\lambda$ on the boundaries. As shown in Fig. 9, all four thresholds are higher when $\lambda$ increases. Among the
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 $\lambda$.

Four boundaries, $\lambda$ is most influential on the boundary $x = W_{low}(h)$, which suggests that, with larger $\lambda$, more wealth is needed to get rid of gloom. However, $\lambda$ has a negligible effect on $x = W_{peak}(h)$ and $x = W_{updt}(h)$, indicating that whether to consume at the historical level and whether to update the historical level have almost no dependence on the drawdown constraint.

6.2 Discussion for $\beta_1$ and $\beta_2$

In this subsection, we focus on the sensitivity analysis of the parameters $\beta_1$ and $\beta_2$ being the risk aversion coefficients below and above the reference level $\alpha h$. The difference $\beta_2 - \beta_1$ reflects the magnitude of the risk aversion changing at the reference. We mainly illustrate the results with $\beta_1 \leq \beta_2$. The differences with $\beta_1 > \beta_2$ are briefly discussed in Sect. 6.2.2. Two limiting cases of interest are also investigated in Sect. 6.2.3.

6.2.1 $\beta_1 \leq \beta_2$

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

To investigate the influence of $\beta_2 - \beta_1$, we have to fix either $\beta_1$ or $\beta_2$. One approach is fixing $\beta_1$, then $\beta_2 - \beta_1$ is the increase in risk aversion when the consumption exceeds the reference level $\alpha h$. 
We first investigate the influence on optimal strategies. When $\beta_2 - \beta_1$ enlarges, the consumption above the boundary $x = W_{\text{ref}}(h)$ (see Fig. 10) increases slower with wealth, which delays the arrival of consumption peak $x = W_{\text{peak}}(h)$. This decelerating effect can be explained by the conservative consumption behavior due to higher risk aversion above the reference. Meanwhile, the change of $\beta_2$ even influences the consumption in the region $x < W_{\text{ref}}(h)$, where only $\beta_1$ seems to be relative. We interpret this phenomenon as a risk allocation behavior: when they tolerate fewer risks, people suppress their consumption to subsidize consumption elsewhere.

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

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

Another approach to examining the effect of $\beta_2 - \beta_1$ is fixing $\beta_2$ and viewing $\beta_2 - \beta_1$ as the decrease in risk aversion when consumption falls below the reference level $\alpha h$.

As $\beta_2 - \beta_1$ enlarges, the consumption increases slower with wealth above $W_{\text{ref}}(h)$ but faster below it (see Fig. 13). The acceleration below $W_{\text{ref}}(h)$ can be reasonably explained by aggressive consumption behavior due to lower risk aversion. Again, the effect of $\beta_1$ is not limited to $x < W_{\text{ref}}(h)$ because such an aggressive decision need to be compensated by conservative behavior when $x > W_{\text{ref}}(h)$. 

 Springer
Fig. 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 $\beta_1$

Fig. 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 the optimal consumption with fixed $\beta_2$

Fig. 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 the optimal risky investment proportion with fixed $\beta_2$

On the other hand, for wealth in the depression and recovery regions $[W_{\text{low}}(h), W_{\text{peak}}(h)]$, the risky investment proportion has a significant increase when $\beta_2 - \beta_1$ enlarges, especially around $W_{\text{ref}}(h)$ (see Fig. 14). This increase is due to lower risk aversion below the reference. However, the change in $\beta_2 - \beta_1$ for fixed $\beta_2$ does not have a significant impact on wealth thresholds. We merely observe a decrease in $W_{\text{ref}}(h)$ for small values of $\beta_2$ as $\beta_2 - \beta_1$ enlarges (see Fig. 15).

Remark 6 The influence of $\beta_2 - \beta_1$ can only be studied by fixing either $\beta_1$ or $\beta_2$. 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 $\beta_1$ or $\beta_2$ (the fixed one). The numerical results illustrate that, when the fixed $\beta_i$ 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 Figs. 10, 12, 13, and 15.

A special case in our model is $\beta_2 - \beta_1 = 0$, which means that risk aversion does not change at the reference level $\alpha h$. In this case, $W_{\text{ref}}$ exists symbolically but has no economic
significance. The depression region ($x \in [W_{\text{low}}(h), W_{\text{ref}}(h)]$) and recovery region ($x \in [W_{\text{ref}}(h), W_{\text{peak}}(h)]$) share strategies and one region would merge into another one.

### 6.2.2 $\beta_1 > \beta_2$

In this part, we consider the case $\beta_1 > \beta_2$ and briefly discuss the impact of $\beta_1 - \beta_2$ on the 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, for risky investment proportion, it is beyond our expectation to observe a new peak around $W_{\text{peak}}(h)$ that overtakes the peak around $W_{\text{ref}}(h)$.

The optimal consumption for $\beta_1 > \beta_2$ is still non-decreasing in $x$. The main difference occurs in $[W_{\text{low}}(h), W_{\text{peak}}(h)]$. As shown in Fig. 16, when $\beta_1 > \beta_2$, the MPC out of wealth is generally higher in the recovery region, which results in an upward turn of consumption curve at $W_{\text{ref}}(h)$ (it is a downward turn when $\beta_1 < \beta_2$). When fixing $\beta_1$, as $\beta_1 - \beta_2$ enlarges, the consumption increases faster with wealth above $W_{\text{ref}}(h)$, bringing forward the arrival of consumption peak $W_{\text{peak}}(h)$. When fixing $\beta_2$, the conclusions are similar, but $W_{\text{peak}}(h)$ is larger due to a conservative strategy when $x < W_{\text{ref}}(h)$.

As shown in Fig. 17, when $\beta_1$ is fixed and $\beta_1 - \beta_2$ enlarges, the optimal investment proportion has an apparent increase over a certain threshold around $W_{\text{low}}(h)$. At the same time, for fixed $\beta_2$, the optimal investment proportion has a decrease in the depression and recovery regions. Besides, we observe a new peak around $W_{\text{peak}}(h)$ that overtakes the peak around...
Fig. 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 the optimal risky investment proportion with fixed $\beta_1 = 1.5$ (left) or fixed $\beta_2 = 1$ (right).

$W_{\text{ref}}(h)$ (see the red line in Fig. 17). It suggests that a high proportion of risky investment is recommended for wealth in the recovery region, especially around $W_{\text{ref}}(h)$ and $W_{\text{peak}}(h)$.

The impact of $\beta_1 - \beta_2$ on the boundaries coincides with the case $\beta_1 < \beta_2$: larger $\beta_1 - \beta_2$ leads to lower level of $W_{\text{peak}}(h)$ with fixed $\beta_1$, higher level of $W_{\text{ref}}(h)$ with fixed $\beta_2$, and it has negligible effects on the others. Therefore we decide not to present the related figures.

### 6.2.3 Limiting cases

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

Fixing $\beta_1$ and letting $\beta_2 \to 0$ indicates that the agent becomes risk neutral when $c > \alpha h$. Because $C_1(h) \to \infty$ 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. She/he might be saving money by consuming as little as possible to aggressively invest in risk assets once she/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 $\beta_2$ and letting $\beta_1 \to 0$, however, has completely different consequences. This limit corresponds to the agent becoming 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 $\lambda h$ and the reference level $\alpha h$. The result is similar to an S-shaped utility model in Li et al. [15] where the agent is risk-seeking below the reference. In S-shaped utility, the optimal consumption jumps from 0 to a level that is strictly higher than reference level $\alpha h$ (see (3.25) in Li et al. [15]). In contrast, our consumption jumps from the lowest level to exactly the reference level.

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

\[
C_4(h) \to -\infty,
\]

\[
C_6(h) \to -\frac{k}{\gamma^2 \beta_2} \frac{1 - q_1}{q_2 - q_1} + \frac{k}{\gamma^2} \frac{1 - q_1}{q_2 - q_1} (\alpha - \lambda)(q_2 - 1)h,
\]

\[
C_8(h) \to \frac{k}{\gamma^2 \beta_2} \frac{1 - q_1}{q_2 - q_1} (e^{(1 - \alpha)(q_2 - 1)\beta_2 h - 1}) + \frac{k}{\gamma^2} \frac{1 - q_1}{q_2 - q_1} (\alpha - \lambda)(q_2 - 1)h,
\]
To show that

\[ C_7(h) \to (1 - \alpha)^{g_2-q_1-1}(\alpha - \lambda) \frac{k}{\gamma^2} \frac{q_2 - 1}{q_2 - q_1} e^{-(1-\alpha)(q_2-q_1)\beta_2h} \]

\[ + (1 - \alpha)^{g_2-q_1} \frac{k}{\gamma^2 \beta_2} \frac{q_2 - 1}{q_2 - q_1} e^{-(1-\alpha)(1-q_1)\beta_2h}, \]

\[ C_5(h) \to (1 - \alpha)^{g_2-q_1-1}(\alpha - \lambda) \frac{k}{\gamma^2} \frac{q_2 - 1}{q_2 - q_1} e^{-(1-\alpha)(q_2-q_1)\beta_2h} \]

\[ + \left[(1 - \alpha)^{g_2-q_1-1} - 1\right] \frac{k}{\gamma^2 \beta_2} \frac{q_2 - 1}{q_2 - q_1} e^{-(1-\alpha)(1-q_1)\beta_2h}, \]

\[ C_3(h) \to -\infty, \]

\[ C_1(h) \to (1 - \alpha)^{g_2-q_1-1}(\alpha - \lambda) \frac{k}{\gamma^2} \frac{q_2 - 1}{q_2 - q_1} e^{-(1-\alpha)(q_2-q_1)\beta_2h} \]

\[ + (1 - \alpha)^{g_2-q_1} \frac{k}{\gamma^2 \beta_2} \frac{q_2 - 1}{q_2 - q_1} e^{-(1-\alpha)(1-q_1)\beta_2h} \]

\[ + \frac{k}{\gamma^2 \beta_2} \frac{q_2 - 1}{q_2 - q_1} + \frac{k}{\gamma^2} \frac{q_2 - 1}{q_2 - q_1} (\alpha - \lambda)(1-q_1)h. \]

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

\[ W_{\text{low}}(h) \to -C_i^L(h)q_1 + \frac{\lambda h}{\gamma}, \]

\[ W_{\text{peak}}(h) \to -C_5^L(h)q_1e^{(1-\alpha)(1-q_1)\beta_2h} - C_6^L(h)q_2e^{-(1-\alpha)(q_2-1)\beta_2h} - \frac{k}{\gamma^2 \beta_2} + \frac{h}{\gamma}, \]

\[ W_{\text{updh}}(h) \to -C_7^L(h)q_1(1 - \alpha)^{g_2-1}e^{(1-\alpha)(1-q_1)\beta_2h} \]

\[ - C_8^L(h)q_2(1 - \alpha)^{g_2-1}e^{-(1-\alpha)(q_2-1)\beta_2h} + \frac{h}{\gamma}. \]

To show that \( W_{\text{ref}}(h) \) has the same limit as \( W_{\text{low}}(h), \) we prove \( W_{\text{ref}}(h) - W_{\text{low}}(h) \to 0. \) In fact,

\[ W_{\text{ref}}(h) - W_{\text{low}}(h) = q_1 C_1(h) \left[ e^{-(\alpha-\lambda)(1-q_1)\beta_1h} - 1 \right] + \frac{(\alpha - \lambda)h}{\gamma} \]

\[ + \frac{k}{\gamma^2 \beta_1} \left[ q_1 \frac{q_2 - 1}{q_2 - q_1} e^{-(\alpha-\lambda)(1-q_1)\beta_1h} + q_2 \frac{1 - q_1}{q_2 - q_1} e^{-(\alpha-\lambda)(q_2-1)\beta_1h} - 1 \right] \]

\[ \to \frac{(\alpha - \lambda)h}{\gamma} + \frac{k}{\gamma^2} \left[ q_1 \frac{q_2 - 1}{q_2 - q_1} (\alpha - \lambda)(1-q_1)h - q_2 \frac{1 - q_1}{q_2 - q_1} (\alpha - \lambda)(q_2-1)h \right] \]

\[ = 0. \]

7 Conclusion

We establish a new theoretical model related to the past spending maximum, where the decision-maker has different risk attitudes on different sides of the reference level. Besides, we impose a drawdown constraint on consumption. Mathematical analysis and computation illustrate that the optimal consumption and investment policies are semi-explicit, with five crucial thresholds. Theoretical and numerical analysis of the solution and sensitivity analysis of the parameters are also conducted. 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 with risk aversion at the reference level; 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 when the reference level is not reached.

Acknowledgements The authors acknowledge the support from the National Natural Science Foundation of China (Grant No. 12271290 and No. 11871036). The authors also thank the members of the group of Actuarial Sciences and Mathematical Finance at the Department of Mathematical Sciences, Tsinghua University, for their feedback and useful conversations. We are also particularly grateful to the two anonymous referees and the Associated Editor whose suggestions greatly improve the manuscript’s quality.

Data availability 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

\[ \hat{H}_t(y) \triangleq h_0 \vee \frac{1}{(1-\alpha)\beta_2} \ln \left( \frac{1-\alpha}{\inf_{s \leq t} Y_s(y)} \right). \]

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

\[
\mathbb{E}_{x_0,h_0} \int_0^{\infty} e^{-\gamma t} U(c_t, h_t) dt = \mathbb{E}_{x_0,h_0} \int_0^{\infty} e^{-\gamma t} \left( U(c_t, h_t) - Y_t(y) c_t \right) dt \\
+ y \mathbb{E}_{x_0,h_0} \int_0^{\infty} c_t M_t dt \\
\leq \mathbb{E}_{x_0,h_0} \int_0^{\infty} e^{-\gamma t} \tilde{U}(Y_t(y), H_t(y)) dt + yx_0 \\
= \mathbb{E}_{x_0,h_0} \int_0^{\infty} e^{-\gamma t} \tilde{U}(Y_t(y), \hat{H}_t(y)) dt + yx_0 \\
= \tilde{V}(y, h_0) + yx_0,
\]

(A.1)

where the second, the third and the last line hold thanks to Lemmas B.1, B.2 and B.4. And equality holds with \(c_t = c^*(Y_t(y), \hat{H}_t(y))\) and \(y = y^*\).

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

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

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

\[
\inf_{y > 0} \left\{ \tilde{V}(y, h_0) + yx_0 \right\} = \sup_{(c, \pi) \in A} \mathbb{E}_{x_0,h_0} \int_0^{\infty} e^{-\gamma t} U(c_t, h_t) dt.
\]

Thus, the dual of \(\tilde{V}(\cdot, \cdot)\) 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 \((c^*, \pi^*)\) given in Theorem 4.1. \(\square\)
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 \( c_i = c^*(Y_t(y), H_t^i(y)) \) and \( y = y^* \).

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

Lemma B.3 (Transversality Condition) For any \( y > 0 \),
\[
\lim_{T \to \infty} \mathbb{E}_{x_0, h_0} \left[ e^{-\gamma T} \tilde{V}(Y_T(y), \hat{H}_T(y)) \right] = 0.
\]

Lemma B.4
\[
\tilde{V}(y, h_0) = \mathbb{E}_{x_0, h_0} \int_0^\infty e^{-\gamma t} \tilde{U}(Y_t(y), \hat{H}_t(y)) dt.
\]

As the proofs of the first two lemmas are similar to the proofs of Lemma 5.4 and Lemma 5.3 in Deng et al. [7], 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
\[
\lim_{T \to \infty} Y_T(y) = \lim_{T \to \infty} ye^{-\frac{(t-y)^2}{2\sigma^2}T - \frac{\mu - \sigma \gamma}{\sigma}BT} = 0, \quad a.s., \tag{B.1}
\]
\[
\lim_{T \to \infty} \hat{H}_T(y) = \lim_{T \to \infty} h_0 \vee \frac{1}{1 - \alpha \beta_2} \ln \left( \frac{1 - \alpha}{\inf_{s \leq T} Y_s(y)} \right) = \infty, \quad a.s. \tag{B.2}
\]
Using the expression of \( \tilde{V}(\cdot, \cdot) \) yields
\[
\lim_{T \to \infty} \mathbb{E}_{x_0, h_0} \left[ e^{-\gamma T} \tilde{V}(Y_T(y), \hat{H}_T(y)) \right] = \lim_{T \to \infty} \mathbb{E}_{x_0, h_0} \left\{ e^{-\gamma T} \left[ C_7(\hat{H}_T(y)) Y_T(y)^{q_1} + C_8(\hat{H}_T(y)) Y_T(y)^{q_2} - \frac{1}{\gamma^\prime} \hat{H}_T(y) Y_T(y) \right] \right. \\
+ \frac{1}{\gamma \beta_2^2} \left( 1 - e^{-(1-\alpha)\beta_2 \hat{H}_T(y)} \right) I \left\{ (1-\alpha) e^{-(1-\alpha)\beta_2 \hat{H}_T(y)} \leq Y_T(y) < e^{-(1-\alpha)\beta_2 \hat{H}_T(y)} \right\} \left. e^{-\gamma T} \left[ C_5(\hat{H}_T(y)) Y_T(y)^{q_1} + C_6(\hat{H}_T(y)) Y_T(y)^{q_2} + \frac{k}{\gamma^\prime \beta_2} Y_T(y) \right] \right. \\
+ \frac{1}{\gamma \beta_2^2} \left( 1 - Y_T(y) + Y_T(y) \ln (Y_T(y)) \right) - \frac{1}{\gamma^\prime} \alpha \hat{H}_T(y) Y_T(y) \right\} I \{ e^{-(1-\alpha)\beta_2 \hat{H}_T(y)} \leq Y_T(y) < 1 \} \} \] \tag{B.3}

By Proposition 3.2, we have \( C_7(\hat{H}_T(y)) = O(e^{-(1-\alpha)(1-q_1)\beta_2 \hat{H}_T(y)}) \), a.s. as \( T \to \infty \), as such,
\[
\mathbb{E}_{x_0, h_0} e^{-\gamma T} C_7(\hat{H}_T(y)) Y_T(y)^{q_1} I \left\{ (1-\alpha) e^{-(1-\alpha)\beta_2 \hat{H}_T(y)} \leq Y_T(y) < e^{-(1-\alpha)\beta_2 \hat{H}_T(y)} \right\} \] \]
then

\[
\lim_{T \to \infty} E_{x_0,h_0} e^{-\gamma T} C_7(\hat{H}_T(y)) Y_T(y) 1\{ (1-\alpha)e^{-(1-\alpha)\beta_2 \hat{H}_T(y)} \leq Y_T(y) < e^{-(1-\alpha)\beta_2 \hat{H}_T(y)} \} = 0.
\]

Similarly,

\[
E_{x_0,h_0} e^{-\gamma T} C_S(\hat{H}_T(y)) Y_T(y) 1\{ (1-\alpha)e^{-(1-\alpha)\beta_2 \hat{H}_T(y)} \leq Y_T(y) < e^{-(1-\alpha)\beta_2 \hat{H}_T(y)} \} = O(e^{-\gamma T}), \text{ as } T \to \infty,
\]

then

\[
\lim_{T \to \infty} E_{x_0,h_0} e^{-\gamma T} C_S(\hat{H}_T(y)) Y_T(y) 2 \{ (1-\alpha)e^{-(1-\alpha)\beta_2 \hat{H}_T(y)} \leq Y_T(y) < e^{-(1-\alpha)\beta_2 \hat{H}_T(y)} \} = 0.
\]

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

\[
e^{-\gamma T} \frac{1}{\gamma \beta_2} (1 - e^{-(1-\alpha)\beta_2 \hat{H}_T(y)}) 1\{ (1-\alpha)e^{-(1-\alpha)\beta_2 \hat{H}_T(y)} \leq Y_T(y) < e^{-(1-\alpha)\beta_2 \hat{H}_T(y)} \} = O(e^{-\gamma T}), \text{ a.s.},
\]

from which we get

\[
\lim_{T \to \infty} E_{x_0,h_0} e^{-\gamma T} \frac{1}{\gamma \beta_2} (1 - e^{-(1-\alpha)\beta_2 \hat{H}_T(y)}) 1\{ (1-\alpha)e^{-(1-\alpha)\beta_2 \hat{H}_T(y)} \leq Y_T(y) < e^{-(1-\alpha)\beta_2 \hat{H}_T(y)} \} = 0.
\]

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

\[
\lim_{T \to \infty} E_{x_0,h_0} e^{-\gamma T} C_5(\hat{H}_T(y)) Y_T(y) 1\{ e^{-(1-\alpha)\beta_2 \hat{H}_T(y)} \leq Y_T(y) < 1 \} = 0.
\]

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

\[
\lim_{T \to \infty} E_{x_0,h_0} e^{-\gamma T} C_6(\hat{H}_T(y)) Y_T(y) 1\{ e^{-(1-\alpha)\beta_2 \hat{H}_T(y)} \leq Y_T(y) < 1 \} = 0.
\]

and

\[
\lim_{T \to \infty} E_{x_0,h_0} e^{-\gamma T} \frac{k}{\gamma^2 \beta_2} Y_T(y) 1\{ e^{-(1-\alpha)\beta_2 \hat{H}_T(y)} \leq Y_T(y) < 1 \} = 0.
\]

Using (B.1) again and the fact \( \lim_{y \to y_0^+} (1 - y + y \ln y) = 1 \), we obtain

\[
\lim_{T \to \infty} E_{x_0,h_0} e^{-\gamma T} \frac{1}{\gamma \beta_2} \left[ 1 - Y_T(y) + Y_T(y) \ln (Y_T(y)) \right] 1\{ e^{-(1-\alpha)\beta_2 \hat{H}_T(y)} \leq Y_T(y) < 1 \} = 0.
\]
Based on Girsanov’s theorem, we have
\[ e^{-\gamma T} \mathbb{E}_{x_0, h_0} \hat{H}_T(y) Y_T(y) \]
\[ = \mathcal{O} \left( e^{-\gamma T} \mathbb{E}_{x_0, h_0} \left[ \sup_{s \leq T} \left\{ \left( \frac{\mu - r}{2\sigma^2} \right)^2 + s + \frac{\mu - r}{\sigma} B_s \right\} e^{-\frac{(\mu - r)^2}{2\sigma^2} T - \frac{\mu - r}{\sigma} B_T} \right] \right) \]
\[ = \mathcal{O} \left( \frac{\mu - r}{\sigma} e^{-\gamma T} \left\{ \sqrt{T} \mathbb{E} \left[ \frac{1}{2\pi} e^{-\frac{(\mu - r)^2}{8\sigma^2} T} - \frac{\mu - r}{2\sigma} T \Phi \left( -\frac{\mu - r}{2\sigma} \sqrt{T} \right) \right] \right\} \right) \]
\[ + \frac{\sigma}{\mu - r} \left[ \Phi \left( -\frac{r}{2\sigma} \sqrt{T} \right) - \Phi \left( -\frac{\mu - r}{2\sigma} \sqrt{T} \right) \right]. \]
It follows that
\[ \lim_{T \to \infty} \mathbb{E}_{x_0, h_0} e^{-\gamma T} \hat{H}_T(y) Y_T(y) = 0. \]  
(B.11)
Thus, using (B.3)–(B.11), we have
\[ \lim_{T \to \infty} \mathbb{E}_{x_0, h_0} e^{-\gamma T} \hat{V}(Y_T(y), \hat{H}_T(y)) = 0. \]
\[ \square \]

**Proof of Lemma B.4** Applying Eq. (3.9) and Itô’s rule, we obtain
\[ d \left\{ e^{-\gamma t} \hat{V}(Y_t(y), \hat{H}_t(y)) \right\} = -e^{-\gamma t} \hat{U}(Y_t(y), \hat{H}_t(y)) dt \]
\[ - \frac{\mu - r}{\sigma} e^{-\gamma t} \hat{V}_y(Y_t(y), \hat{H}_t(y)) Y_t(y) dB_t \]
\[ + e^{-\gamma t} \hat{V}_h(Y_t(y), \hat{H}_t(y)) d\hat{H}_t(y). \]  
(B.12)
Define the stopping times: \( \forall n \geq 1, \)
\[ \tau_n \triangleq \inf \left\{ t \geq 0 \mid Y_t(y) \geq n \text{ or } \hat{H}_t(y) \geq \frac{1}{(1 - \alpha)\beta_2} \ln \left( \frac{1}{1 - \alpha} \right)n \right\}. \]
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 \( \forall T > 0, \)
\[ \mathbb{E}_{x_0, h_0} 1_{\{ \tau_n \leq T \}} \leq \mathbb{P}_{x_0, h_0} \left( \left\{ \sup_{t \in [0, T]} Y_t(y) \geq n \right\} \cup \left\{ \sup_{t \in [0, T]} \hat{H}_t(y) \geq \frac{1}{(1 - \alpha)\beta_2} \ln \left( \frac{1}{1 - \alpha} \right)n \right\} \right) \]
\[ \leq \mathbb{P}_{x_0, h_0} \left( \sup_{t \in [0, T]} Y_t(y) \geq n \right) + \mathbb{P}_{x_0, h_0} \left( \inf_{t \in [0, T]} Y_t(y) \leq \frac{1}{n} \right) \]
\[ = \mathbb{P}_{x_0, h_0} \left( \sup_{t \in [0, T]} Y_t(y) \geq n \right) + \mathbb{P}_{x_0, h_0} \left( \sup_{t \in [0, T]} Y_t(y) \leq \frac{1}{n} \right) \]
\[ \leq n^{-2\kappa} \mathbb{E}_{x_0, h_0} \sup_{t \in [0, T]} Y_t(y)^{2\kappa} + n^{-2\kappa} \mathbb{E}_{x_0, h_0} \sup_{t \in [0, T]} Y_t(y)^{-2\kappa} \]
\[ = \mathcal{O}(n^{-2\kappa} (1 + y^{2\kappa}) e^{MT}) \]  
(B.13)
for some constants \( M. \)
Integrating (B.12) from 0 to \( T \wedge \tau_n \) and taking expectation on both sides, we obtain
\[ \hat{V}(y, h_0) = \mathbb{E}_{x_0, h_0} \left[ e^{-\gamma T \wedge \tau_n} \hat{V}(Y_{T \wedge \tau_n}(y), \hat{H}_{T \wedge \tau_n}(y)) \right] \]
If \( n \) holds. Based on Lemma B.3, as \( n \to \infty \), in addition, it follows from the fact that either

\[
\begin{align*}
Y_{\tau_n}(y) &\leq n, \quad \hat{H}_{\tau_n}(y) = \frac{1}{(1-\alpha)\beta_2} \ln \left( (1-\alpha)n \right) \\
Y_{\tau_n}(y) &\geq n, \quad \hat{H}_{\tau_n}(y) \leq \frac{1}{(1-\alpha)\beta_2} \ln \left( (1-\alpha)n \right)
\end{align*}
\]

holds. In addition, it follows from the fact \( Y_{\tau_n}(y), \hat{H}_{\tau_n}(y) \in C_d \) that either

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

or

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

holds.

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

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

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

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 \( n \),

\[
\tilde{V}(Y_{\tau_n}(y), \hat{H}_{\tau_n}(y)) = O(n^{-q_1}).
\]

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

\[
\tilde{V}(Y_{\tau_n}(y), \hat{H}_{\tau_n}(y)) = O(n^{-q_1}).
\]
If $1 \leq Y_{\tau_n}(y) < e^{(\alpha - \lambda)\beta_1} \hat{H}_{\tau_n}(y)$, then we have, for sufficiently large $n$,

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

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

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

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

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

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

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

In summary, we have $\tilde{V}(Y_{\tau_n}(y), \hat{H}_{\tau_n}(y)) = O(n^{(\alpha - q_1)\beta_1 \beta_2})$. Applying (B.13) with $\kappa \geq \frac{1}{2} \left( (\alpha - 1) \sqrt{q_2} \vee (1 - \alpha) \sqrt{1 - \alpha} \beta_1 \beta_2 \right)$, 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 $\tau_n$ tends to $\infty$ as $n \to \infty$, based on monotone convergence theorem, we obtain that the second term in (B.14) tends to $\mathbb{E}_{x_0,\hat{h}_0} \int_0^\infty e^{-\gamma t} \tilde{U}(Y_t(y), \hat{H}_t(y)) dt$, 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 $(Y_t(y), \hat{H}_t(y))$ is on the boundary. Using the boundary condition, we have $\tilde{V}_h(Y_t(y), \hat{H}_t(y)) = 0$, as such, the last term in (B.14) vanishes. Thus, the proof follows. \hfill $\Box$

### C Proofs of other results

#### Proof of Lemma 4.2

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

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

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

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

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

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 $C_7(h) > 0,$

$$\psi(1) = C_3(h) \frac{r}{k} + C_4(h) \frac{r}{k} + \frac{1}{\gamma \beta_1}$$

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

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

$$> 0$$

and

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

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

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

$$> 0.$$

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

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

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

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

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

$$> 0.$$
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

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

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.

\[ \square \]

**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 \(c_{\text{primal}}^*\) and \(\pi_{\text{primal}}^*\) given in Theorem 4.4 are locally Lipschitz on \(\mathcal{C}\).

Now to prove that \(\pi_{\text{primal}}^*\) 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 \(\pi_{\text{primal}}^*\) given in Theorem 4.4, we have

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

Differentiating (4.6), we obtain

\[ 1 = -C_1(h)(q_1 - 1)\left[f_1(x, h)\right]^{q_1-2} \frac{\partial f_1(x, h)}{\partial x} \\
- C_2(h)(q_2 - 1)\left[f_1(x, h)\right]^{q_2-2} \frac{\partial f_1(x, h)}{\partial x}. \] (C.1)

It follows from (4.16) and (C.1) that, for \(W_{\text{bkrp}}(h) \leq x \leq W_{\text{low}}(h),

\[ \frac{\partial \pi_{\text{primal}}^*}{\partial x} = \frac{\mu - r}{\sigma^2 k} \begin{cases} 
(1 - q_2) \frac{r}{k} (q_1 - q_2) C_1(h) \left[f_1(x, h)\right]^{q_1-2} \frac{\partial f_1(x, h)}{\partial x}, \\
(1 - q_2) \frac{A_1(x, h)}{B_1(x, h)}, 
\end{cases} \]

where

\[ A_1(x, h) = \frac{r}{k} (q_1 - q_2) C_1(h) \left[f_1(x, h)\right]^{q_1-2}, \]

\[ B_1(x, h) = \frac{r}{k} \left[-C_1(h)\left[f_1(x, h)\right]^{q_1-2} - C_2(h)\left[f_1(x, h)\right]^{q_2-2}\right]. \]
As $C_2(h) = 0$, $\frac{\partial \pi^*_\text{primal}}{\partial x}$ is constant for $W_{\text{brkp}}(h) \leq x \leq W_{\text{low}}(h)$. Differentiating (4.8), we obtain

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

Using (4.16) and (C.2), we have, for $W_{\text{low}}(h) \leq x \leq W_{\text{ref}}(h)$,

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

$$= \frac{\mu - r}{\sigma^2} \left[ (1 - q_2) + \frac{A_2(x, h)}{B_2(x, h)} \right],$$

where

$$A_2(x, h) = \frac{r}{k}(q_1 - q_2)C_3(h)[f_2(x, h)]^{q_1 - 1} + (1 - q_2) \frac{1}{\gamma \beta_1},$$

$$B_2(x, h) = -\frac{r}{k}C_3(h)[f_2(x, h)]^{q_1 - 1} - \frac{r}{k}C_4(h)[f_2(x, h)]^{q_2 - 1} - \frac{1}{\gamma \beta_1}.$$

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 $A_2$ such that $|A_2(x, h)| \leq A_2$ for $W_{\text{low}}(h) \leq x \leq W_{\text{ref}}(h)$. As $V_y(y, h) > 0$, we know that $B_2(x, h) < 0$ for any $h \geq \hat{h}$. $C_4 < 0$ and $q_1 < 0 < 1 < q_2$ imply that $\frac{r}{k}C_3(h)\gamma^{q_1 - 1} + \frac{r}{k}C_4(h)\gamma^{q_2 - 1} + \frac{1}{\gamma \beta_1}$ as a function of $h$ is either decreasing, or increasing, or first increasing then decreasing. Hence for $W_{\text{low}}(h) \leq x \leq W_{\text{ref}}(h)$,

$$-B_2(x, h) \geq \min \left\{ \frac{r}{k}C_3(h) + \frac{r}{k}C_4(h) + \frac{1}{\gamma \beta_1}, \frac{r}{k}C_3(h)e^{-(1 - q_1)(\alpha - \lambda)\beta_1 h} \right\}$$

$$> 0. \quad (C.3)$$

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, $\frac{\partial \pi^*_\text{primal}}{\partial h}$ 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 $\pi^*_{\text{primal}}$, denote $c_{11} \triangleq \frac{\mu - r}{\sigma^2} k$, we obtain

$$
\begin{align*}
\frac{\partial \pi^*_{\text{primal}}}{\partial h} &= c_{11} \begin{cases}
C_1(h)(q_1 - 1) \left[ f_1(x, h) \right]^{q_1-2} \frac{\partial f_1(x, h)}{\partial h} + C'_4(h) \left[ f_1(x, h) \right]^{q_1-1} \\
+ C_2(h)(q_2 - 1) \left[ f_1(x, h) \right]^{q_2-2} \frac{\partial f_1(x, h)}{\partial h} \\
+ C'_2(h) \left[ f_1(x, h) \right]^{q_2-1},
\end{cases} \\
W_{\text{bkrp}}(h) &\leq x \leq W_{\text{low}}(h), \\
C_3(h)(q_1 - 1) \left[ f_2(x, h) \right]^{q_1-2} \frac{\partial f_2(x, h)}{\partial h} + C'_3(h) \left[ f_2(x, h) \right]^{q_1-1} \\
+ C_4(h)(q_2 - 1) \left[ f_2(x, h) \right]^{q_2-2} \frac{\partial f_2(x, h)}{\partial h} \\
+ C'_4(h) \left[ f_2(x, h) \right]^{q_2-1},
\end{cases}
\end{align*}
$$

Differentiating (4.6),

$$
0 = -C_1(h)q_1(q_1 - 1) \left[ f_1(x, h) \right]^{q_1-2} \frac{\partial f_1(x, h)}{\partial h} - C'_1(h)q_1 \left[ f_1(x, h) \right]^{q_1-1} \\
- C_2(h)q_2(q_2 - 1) \left[ f_1(x, h) \right]^{q_2-2} \frac{\partial f_1(x, h)}{\partial h} - C'_2(h)q_2 \left[ f_1(x, h) \right]^{q_2-1} + \frac{\lambda}{\gamma}. \tag{C.4}
$$

Then, using $C_2(h) = 0$, we have for $W_{\text{bkrp}}(h) \leq x \leq W_{\text{low}}(h)$,

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

Hence $\frac{\partial \pi^*_{\text{primal}}}{\partial h}$ is constant for $W_{\text{bkrp}}(h) \leq x \leq W_{\text{low}}(h)$.

Differentiating (4.8),

$$
0 = -C_3(h)q_1(q_1 - 1) \left[ f_2(x, h) \right]^{q_1-2} \frac{\partial f_2(x, h)}{\partial h} - C'_3(h)q_1 \left[ f_2(x, h) \right]^{q_1-1} \\
- C_4(h)q_2(q_2 - 1) \left[ f_2(x, h) \right]^{q_2-2} \frac{\partial f_2(x, h)}{\partial h} - C'_4(h)q_2 \left[ f_2(x, h) \right]^{q_2-1} \\
- \frac{1}{\gamma \beta_1 f_2(x, h)} \frac{\partial f_2(x, h)}{\partial h} + \frac{\alpha}{\gamma}. \tag{C.5}
$$

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

$$
\begin{align*}
\frac{\partial \pi^*_{\text{primal}}}{\partial h} &= c_{11} \begin{cases}
C'_3(h) \left[ f_2(x, h) \right]^{q_1-1} \left( 1 - \frac{q_1}{q_2} \right) + \frac{\alpha}{\gamma q_2} \\
+ \left[ C_3(h) \left[ f_2(x, h) \right]^{q_1-1}(q_1 - q_2) - \frac{1}{\gamma \beta_1 q_2} \right] \frac{\partial f_2(x, h)}{\partial h}
\end{cases} \\
&= \frac{\mu - r}{\sigma^2} \frac{r}{k} \begin{cases}
C'_3(h) \left[ f_2(x, h) \right]^{q_1-1} \left( 1 - \frac{q_1}{q_2} \right) + \frac{\alpha}{\gamma q_2} \\
+ \frac{k}{r} A_2(x, h) \frac{C'_3(h)q_1 \left[ f_2(x, h) \right]^{q_1-1} + C'_4(h)q_2 \left[ f_2(x, h) \right]^{q_2-1} - \frac{\alpha}{\gamma}}{B_2(x, h)}
\end{cases}.
\end{align*}
$$
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 $1 \leq f_2(x, h) < e^{(\alpha - \lambda)\beta_1 h}$ for $W_{\text{low}}(h) < x \leq W_{\text{ref}}(h)$. Applying the estimates of $A_2(x, h)$ and $B_2(x, h)$ again, we know that $\frac{\partial^2 u}{\partial h^2}$ is bounded for $W_{\text{low}}(h) < x \leq W_{\text{ref}}(h)$. For the rest two regions $W_{\text{ref}}(h) \leq x \leq W_{\text{peak}}(h)$ and $W_{\text{peak}}(h) \leq x \leq W_{\text{upd}}(h)$, the proof is similar and omitted. □

**Proof of Proposition 4.8** Proof of (i) is trivial.

**Proof of (ii):**

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

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

$$
\frac{\xi}{\xi} [C_3(h) + C_4(h)] + \frac{1}{\gamma \beta_1} = \frac{\xi}{\xi} [C_3(h) + C_4(h)] + \frac{1}{\gamma \beta_2} \quad \text{leads to the continuity of } y \tilde{V}_{yy}(y, h)
$$

at $y = 1$. Hence $\frac{\partial e^*(x, h)}{\partial x} |_{x \rightarrow W_{\text{ref}}(h)^{+}} = \beta_1 \frac{\partial e^*(x, h)}{\partial x} |_{x \rightarrow W_{\text{ref}}(h)^{-}}$, i.e., the MPC out of wealth shrinks or swells by $\frac{\partial e}{\partial x}$ when wealth exceeds $W_{\text{ref}}(h)$.

**Proof of (iii):**

First, we claim that there exists unique $\tilde{h} > 0$ solving (4.18) and the left hand side of (4.18) is positive for $h > \tilde{h}$, negative for $0 \leq h < \tilde{h}$.

To prove the claim, define the left hand side of (4.18) as $I(h)$, i.e.,

$$
I(h) \equiv C_5(h)(q_1 - 1)e^{-(q_1 - 2)(1 - \alpha)\beta_2 h} + C_6(h)(q_2 - 1)e^{-(q_2 - 2)(1 - \alpha)\beta_2 h}.
$$

One can show that

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

where

$$
J(h) \equiv -\frac{(1 - \alpha)^{q_2 - q_1}(\alpha - \lambda)(1 - q_1)}{(1 - \alpha)(q_2 - q_1)\beta_2 + (\alpha - \lambda)(q_2 - 1)\beta_1} - \frac{1}{\beta_1}
$$

$$
+ \frac{\beta_2 - \beta_1}{\beta_1 \beta_2} e^{(\alpha - \lambda)(q_2 - 1)\beta_1 h}
$$

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

We have

$$
J'(h) = \frac{\beta_2 - \beta_1}{\beta_1 \beta_2}(\alpha - \lambda)(q_2 - 1)\beta_1 e^{(\alpha - \lambda)(q_2 - 1)\beta_1 h}
$$

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

with

$$
\lim_{h \rightarrow +\infty} J'(h) = +\infty.
$$

Hence $J(h)$ can either be increasing or first decreasing then increasing. From the facts

$$
J(0) = -(1 - \alpha)^{q_2 - q_1}\frac{(\alpha - \lambda)(1 - q_1) + (1 - \alpha)(q_2 - q_1) + (\alpha - \lambda)(q_2 - 1)\beta_1 \beta_2}{(1 - \alpha)(q_2 - q_1)\beta_2 + (\alpha - \lambda)(q_2 - 1)\beta_1} < 0,
$$
\[ \lim_{h \to +\infty} J(h) = +\infty, \]
we know that the claim follows.

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

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

Hence \( \frac{\partial \pi^*(y, h)}{\partial y} \) is a linear combination of two power functions of \( y \) for either \( 1 \leq y < e^{(\alpha-\lambda)\beta_1 h} \) or \( e^{-(1-\alpha)\beta_2 h} \leq y < 1 \). Because

\[
\frac{\partial \pi^*(y, h)}{\partial y} \bigg|_{y = e^{(\alpha-\lambda)\beta_1 h}} = C_3(h)(q_1 - 1)e^{(q_1-2)(\alpha-\lambda)\beta_1 h} + C_4(h)(q_2 - 1)e^{(q_2-2)(\alpha-\lambda)\beta_1 h} - C_1(h)(1 - q_1)e^{(q_1-2)(\alpha-\lambda)\beta_1 h} < 0,
\]

\[
\frac{\partial \pi^*(y, h)}{\partial y} \bigg|_{y = e^{-(1-\alpha)\beta_2 h}} = I(h) \begin{cases} < 0, & 0 \leq h < \bar{h}, \\ > 0, & h > \bar{h}, \end{cases}
\]

\[
\lim_{y \downarrow 1} \frac{\partial \pi^*(y, h)}{\partial y} = C_3(h)(q_1 - 1) + C_4(h)(q_2 - 1) = C_5(h)(q_1 - 1) + C_6(h)(q_2 - 1) = \lim_{y \downarrow 1} \frac{\partial \pi^*(y, h)}{\partial y},
\]

one concludes that if \( 0 \leq h < \bar{h} \), then \( \frac{\partial \pi^*(y, h)}{\partial y} < 0 \) for \( y \in \left[ e^{-(1-\alpha)\beta_2 h}, e^{(\alpha-\lambda)\beta_1 h} \right] \); if \( h > \bar{h} \), then \( \frac{\partial \pi^*(y, h)}{\partial y} < 0 \) for \( y \in (\bar{y}, e^{(\alpha-\lambda)\beta_1 h}) \), and \( \frac{\partial \pi^*(y, h)}{\partial y} > 0 \) for \( y \in \left[ e^{-(1-\alpha)\beta_2 h}, \bar{y} \right) \).

The above \( \bar{y} \) is the unique root of \( \frac{\partial \pi^*(y, h)}{\partial y} = 0 \) on \( \left[ e^{-(1-\alpha)\beta_2 h}, e^{(\alpha-\lambda)\beta_1 h} \right] \). Therefore, the property (iii) follows.

**Proof of property (iv):**

For \( x \geq W_{\text{upd}}(h) \), the running maximum \( h \) is updated and we have \( c = h = W_{\text{upd}}^{-1}(x) \). Hence the MPC out of wealth decreases if and only if the bliss curve \( W_{\text{upd}}^{-1}(x) \) is concave, or equivalently, \( W_{\text{upd}}''(h) \geq 0 \). Direct computation shows

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

For \( \beta_1 \geq \beta_2 \), we have \( M_1 > 0, M_2 \leq 0 \) and thus \( W_{\text{upd}}''(h) \geq 0 \). 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 \( W_{\text{upd}}''(h) \geq 0 \) for \( h \leq \bar{h} \). Therefore, the property (iv) follows.

\[ \square \]

**D A generalization of the reference point**

We consider the alternative endogenous reference point \( \alpha(\psi(h)c + (1 - \psi(h))h) \), which is a fraction \( \alpha \) 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 \( \varphi \) 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 \( \varphi(h) = \bar{\varphi} h / h + \hat{h} \) 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

\[
\begin{aligned}
&\begin{cases}
\lambda h, & 1 \vee (1 - \alpha \varphi(h))e^{[(\alpha - \lambda) - (1 - \lambda)\alpha\varphi(h)]\beta_1 h} \leq y, \\
-\frac{1}{\beta_1} \frac{1}{1 - \alpha \varphi(h)} \ln \left( \frac{y}{1 - \alpha \varphi(h)} \right) + \alpha h \frac{1 - \varphi(h)}{1 - \alpha \varphi(h)}, & 1 \leq y < 1 \vee (1 - \alpha \varphi(h))e^{[(\alpha - \lambda) - (1 - \lambda)\alpha\varphi(h)]\beta_1 h}, \\
-\frac{1}{\beta_2} \frac{1}{1 - \alpha \varphi(h)} \ln \left( \frac{y}{1 - \alpha \varphi(h)} \right) + \alpha h \frac{1 - \varphi(h)}{1 - \alpha \varphi(h)}, & (1 - \alpha \varphi(h))e^{-(1 - \alpha)\beta_2 h} \leq y < 1,
\end{cases}
\end{aligned}
\]

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
\[
\hat{V}(y, h) = \begin{cases} 
C_{1}(h)y^{q_{1}} + C_{2}(h)y^{q_{2}} - \frac{1}{2} \lambda hy & \\
+ \frac{1}{\gamma_{1}} \left( 1 - e^{(\alpha - \lambda) - (1 - \lambda) \alpha \varphi(h)) \beta_{1}h} \right), 
1 \vee (1 - \alpha \varphi(h))e^{(\alpha - \lambda) - (1 - \lambda) \alpha \varphi(h)) \beta_{1}h} \leq y, 
\end{cases}
\]

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 \( 1 < (1 - \alpha \varphi(h))e^{(\alpha - \lambda) - (1 - \lambda) \alpha \varphi(h)) \beta_{1}h} \),

\[
C_{2}(h) = 0, \quad C_{4}(h) = -\frac{k}{\gamma_{2}} \frac{1 - q_{1}}{\beta_{1} q_{2} - q_{1}} \frac{1}{1 - \alpha \varphi(h))q_{2}} e^{(\alpha - \lambda) - (1 - \lambda) \alpha \varphi(h))q_{2}(1 - \beta_{1}h),
\]

\[
C_{6}(h) = C_{4}(h) + \frac{k}{\gamma_{2}} \frac{\beta_{2} - \beta_{1}}{\beta_{1} \beta_{2}} \frac{1 - q_{1}}{q_{2} - q_{1}} \frac{1}{1 - \alpha \varphi(h))q_{2}} + \frac{1}{\gamma_{1}} \frac{\beta_{2} - \beta_{1}}{\beta_{2} q_{2} - q_{1}} \frac{1}{1 - \alpha \varphi(h))q_{2}} \frac{\alpha \varphi(h)}{q_{2} - q_{1}} \frac{1}{1 - \alpha \varphi(h))q_{2}},
\]

\[
C_{8}(h) = C_{6}(h) + \frac{k}{\gamma_{2}} \frac{1 - q_{1}}{q_{2} - q_{1}} \frac{1}{1 - \alpha \varphi(h))q_{2}} e^{(\alpha - \lambda) - (1 - \lambda) \alpha \varphi(h))q_{2}(1 - \beta_{2}h)},
\]

\[
C_{7}(h) = \frac{(1 - \alpha)q_{2} - q_{1}}{(1 - \alpha)(q_{2} - q_{1}) \beta_{2} + (\alpha - \lambda)(q_{2} - 1) \beta_{1}} e^{(\alpha - \lambda) - (1 - \lambda) \alpha \varphi(h))q_{2}((1 - \alpha)q_{2} - q_{1}) \beta_{2}h},
\]

\[
C_{5}(h) = C_{7}(h) - \frac{k}{\gamma_{2}} \frac{q_{2} - q_{1}}{q_{2} - q_{1}} \frac{1}{1 - \alpha \varphi(h))q_{2}} e^{(\alpha - \lambda) - (1 - \lambda) \alpha \varphi(h))q_{2}(1 - \beta_{2}h)},
\]

\[
C_{3}(h) = C_{5}(h) - \frac{k}{\gamma_{2}} \frac{\beta_{2} - \beta_{1}}{\beta_{1} \beta_{2}} \frac{q_{2} - q_{1}}{q_{2} - q_{1}} \frac{1}{1 - \alpha \varphi(h))q_{2}} + \frac{1}{\gamma_{1}} \frac{\beta_{2} - \beta_{1}}{\beta_{2} q_{2} - q_{1}} \frac{1}{1 - \alpha \varphi(h))q_{2}} \frac{\alpha \varphi(h)}{q_{2} - q_{1}} \frac{1}{1 - \alpha \varphi(h))q_{2}},
\]

\[
C_{1}(h) = C_{3}(h) + \frac{k}{\gamma_{2}} \frac{q_{2} - q_{1}}{q_{2} - q_{1}} \frac{1}{1 - \alpha \varphi(h))q_{2}} e^{(\alpha - \lambda) - (1 - \lambda) \alpha \varphi(h))q_{2}(1 - \beta_{1}h)},
\]

In this case, as \( \varphi(h) \in [0, 1] \) and \( e^{(\alpha - \lambda) - (1 - \lambda) \alpha \varphi(h))q_{2}h} < 1 - \alpha \varphi(h) \leq 1 \), the estimates of \( C_{i}(h) \), \( 1 \leq i \leq 8 \) is the same as Sect. 3 except that the order estimate of \( C_{4}(h) \) is replaced by \( C_{4}(h) = \mathcal{O}(1) \).

\( \quad \circled{S} \quad \) Springer
**Case 2** For $h$ such that $1 \geq (1 - \alpha \varphi(h)) e^{(\alpha - \lambda) - (\alpha - \lambda) \alpha \varphi(h)} \beta_1 h$ (the second region $1 \leq y < 1 \vee (1 - \alpha \varphi(h)) e^{(\alpha - \lambda) - (\alpha - \lambda) \alpha \varphi(h)} \beta_1 h$ is null), we have:

$$C_2(h) = 0,$$

$$C_6(h) = \frac{1 - q_1}{q_2 - q_1} \lambda h - \frac{1}{q_2 - q_1} \frac{1}{\gamma \beta_1} (1 - e^{(\alpha - \lambda) - (\alpha - \lambda) \alpha \varphi(h)} \beta_1 h)$$

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

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

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

$$C_1(h) = C_5(h) + \frac{q_2 - 1}{q_2 - q_1} \frac{1}{\gamma \beta_2} \frac{1}{1 - \alpha \varphi(h)} (1 - e^{(\alpha - \lambda) - (\alpha - \lambda) \alpha \varphi(h)} \beta_1 h)$$

$$C_9(h) = \sum_k \frac{1 - q_1}{q_2 - q_1} \frac{1}{\gamma^2 \beta_2} (1 - \alpha \varphi(h)) q_2 e^{(1 - \alpha)(q_2 - 1) \beta_2 h}$$

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

$$C_1(h) > 0, \quad C_7(h) > 0, \quad C_8(h) > 0.$$
For the second term, it can be directly verified that
\[ f(x) := \frac{1 - q_1}{q_2 - q_1} \frac{x}{1 - \alpha \varphi(h)} - \frac{q_1}{q_2 - q_1} \frac{1}{\gamma \beta_1} (1 - e^{x \beta_1 h}) \]
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 \( x \in [1, \frac{1}{1 - \alpha \varphi(h)}] \). 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 \( \varphi \), we can establish the verification theorem and apply duality to obtain the optimal strategy given in the following theorem.

**Theorem D.1** For \((x_0, h_0) \in C\), where \( C \) 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

\[
c^*_{\text{primal}}(x, h) = \left\{ \begin{array}{ll}
\frac{\partial}{\partial h} \left[ C_1(h) f_1(x, h) \right]^{q_1-1} + C_2(h) f_1(x, h) & \frac{\lambda h}{\gamma} \leq x \leq W_{\text{low}}(h) \wedge W_{\text{ref}}(h), \\
\frac{\partial}{\partial h} \left[ C_3(h) f_2(x, h) \right]^{q_2-1} + C_4(h) f_2(x, h) & W_{\text{low}}(h) \wedge W_{\text{ref}}(h) < x \leq W_{\text{ref}}(h), \\
\frac{\partial}{\partial h} \left[ C_5(h) f_3(x, h) \right]^{q_3-1} + C_6(h) f_3(x, h) & W_{\text{ref}}(h) < x \leq W_{\text{peak}}(h), \\
\frac{\partial}{\partial h} \left[ C_7(h) f_4(x, h) \right]^{q_4-1} + C_8(h) f_4(x, h) & W_{\text{peak}}(h) < x \leq W_{\text{updt}}(h),
\end{array} \right.
\]

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

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

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

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

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

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

\[
W_{\text{low}}(h) = -C_1(h) q_1 \left[ (1 - \alpha \varphi(h)) e^{(\alpha - \lambda) - (1 - \lambda) \alpha \varphi(h)} \right]^{q_1-1}.
\]
Then \( SDE \geq h \) large current model. As has been mentioned, Assumption 1 ensures that this will not happen for Remark 10.

It is interesting to notice that the depression region can possibly vanish under the r

For \( r \neq \gamma \), Eq. (3.9) should be replaced by

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

\[
W_{\text{ref}}(h) = - C_3(h)q_1 - C_4(h)q_2 - \frac{1}{\gamma \beta_1} \frac{1}{1 - \alpha \varphi(h)} \ln \left[ \frac{1}{1 - \alpha \varphi(h)} \right]
\]

\[
- \frac{k}{\gamma^2 \beta_1} \frac{1}{1 - \alpha \varphi(h)} + \frac{ah}{\gamma} \frac{1 - \varphi(h)}{1 - \alpha \varphi(h)}.
\]

\[
W_{\text{peak}}(h) = - C_5(h)q_1 \left[ (1 - \alpha \varphi(h)) e^{-(1 - \alpha) \beta_2 h} \right]^{q_1 - 1}
\]

\[
- C_6(h)q_2 \left[ (1 - \alpha \varphi(h)) e^{-(1 - \alpha) \beta_2 h} \right]^{q_2 - 1}
\]

\[
+ \frac{h}{\gamma} \frac{1 - \alpha \varphi(h)}{1 - \alpha \varphi(h)} - \frac{k}{\gamma^2 \beta_2} \frac{1}{1 - \alpha \varphi(h)} + \frac{ah}{\gamma} \frac{1 - \varphi(h)}{1 - \alpha \varphi(h)}.
\]

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

Then SDE

\[
\left\{ \begin{array}{l}
   dX_t = r X_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{array} \right.
\]

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

\[
\left\{ \left( c_{\text{primal}}^*(X_t^*, H_t^*), \pi_{\text{primal}}^*(X_t^*, H_t^*) \right), t \geq 0 \right\}.
\]

**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 \( \varphi \). There are similar phenomena in Li et al. [15], 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 Lemmas 4.2 and 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 \( \frac{1}{2} C_9(h) y^{q_2 - 1} + \frac{1}{\gamma \beta_2} \frac{1 - q_1}{q_2 - q_1} \frac{1}{1 - \alpha \varphi(h)} > 0 \) 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 \( \varphi'(h) \) \( (0 \leq \varphi'(h) < \frac{1 - \varphi(h)}{h} \leq \frac{1 - \varphi(h)}{h_0} \) to show that \( C_i'(h), 1 \leq i \leq 8 \) are bounded.

**E Results for \( r \neq \gamma \)**

For \( r \neq \gamma \), Eq. (3.9) should be replaced by

\[
- \gamma \bar{V}(y, h) + (\gamma - r) y \bar{V}_y(y, h) + \frac{(r - \mu)^2}{2\sigma^2} y^2 \bar{V}_{yy}(y, h) = - \bar{U}(y, h),
\]

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

\[
\bar{V}(y, h)
\]
coefficients such as
is completely the same as in Sect. 3 and the main results are similar. The optimal strategy is
\[ k_r = \frac{1}{\gamma r^2} \left( \frac{1}{\gamma} + \frac{\gamma - 2r + k}{r^2} \right) \]
\[ C_i(h) = C_i(h) y^{q_i} + C_4(h) y^{q_4} - \frac{1}{r} y + \frac{1}{\gamma r^2} \left[ 1 - e^{-(a-\lambda)(q_2-1)} \right] e^{(a-\lambda)\beta_1 h} \]
Theorem E.1 For \((x_0, h_0) \in C\), where \(C\) 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

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

and

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

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

\[
x = -C_1(h)q_1\left[ f_1(x, h) \right]^{q_1-1} - C_2(h)q_2\left[ f_1(x, h) \right]^{q_2-1} + \frac{\lambda h}{r}, \\
x = -C_3(h)q_1\left[ f_2(x, h) \right]^{q_1-1} - C_4(h)q_2\left[ f_2(x, h) \right]^{q_2-1} - \frac{1}{r \beta_1} \ln \left[ f_2(x, h) \right] \\
x = -C_5(h)q_1\left[ f_3(x, h) \right]^{q_1-1} - C_6(h)q_2\left[ f_3(x, h) \right]^{q_2-1} - \frac{1}{r \beta_2} \ln \left[ f_3(x, h) \right] \\
x = -C_7(h)q_1\left[ f_4(x, h) \right]^{q_1-1} - C_8(h)q_2\left[ f_4(x, h) \right]^{q_2-1} + \frac{h}{r},
\]

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

\[
W_{\text{low}}(h) = -C_1(h)q_1 e^{-(\alpha - \lambda)(1-q_1)\beta_1 h} - C_2(h)q_2 e^{(\alpha - \lambda)(q_2-1)\beta_1 h} + \frac{\lambda h}{r}, \\
W_{\text{ref}}(h) = -C_3(h)q_1 - C_4(h)q_2 - \frac{\gamma - r + k}{r^2 \beta_1} + \frac{\alpha h}{r}, \\
W_{\text{peak}}(h) = -C_5(h)q_1 e^{(1-\alpha)(1-q_1)\beta_2 h} - C_6(h)q_2 e^{-(1-\alpha)(q_2-1)\beta_2 h} - \frac{\gamma - r + k}{r^2 \beta_1} + \frac{h}{r}, \\
W_{\text{updt}}(h) = -C_7(h)q_1(1 - \alpha)^{q_1-1} e^{(1-\alpha)(1-q_1)\beta_2 h} \\
- C_8(h)q_2(1 - \alpha)^{q_2-1} e^{-(1-\alpha)(q_2-1)\beta_2 h} + \frac{h}{\gamma}.
\]

Then SDE

\[
\begin{cases} 
\frac{dX_t}{r X_t dt + \pi^*_\text{primal}(X_t, H_t^*) (\mu - r) dt + \pi^*_\text{primal}(X_t, H_t^*) \sigma dW_t - c^*_\text{primal}(X_t, H_t^*) dt, \\
X_0 = x_0
\end{cases}
\]
with $H_t^* \triangleq h_0 \vee \sup_{t \leq s} c_{\text{primal}}^*(X_s, H_s^*)$ and $H_0^* = h_0$, has a unique strong solution $\{X_t^*, \ t \geq 0\}$.

The optimal consumption and investment policy is

$$\left\{ \left( e_{\text{primal}}^*(X_t^*, H_t^*), \pi_{\text{primal}}^*(X_t^*, H_t^*) \right), \ t \geq 0 \right\}.$$  

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

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

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

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

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

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

$$> 0$$

and

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

$$= q_1(q_1 - 1)C_7(h)e^{(q_1 - 1)(\alpha - \lambda)\beta_1 h}$$

$$+ \frac{q_1(q_1 - 1)}{(q_2 - q_1)\beta_2} \left[ \frac{q_2}{\gamma} - \frac{1}{r} + \frac{\gamma - 2r + k}{r^2}(q_2 - 1) \right]$$

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

$$+ \frac{q_1(q_1 - 1)}{(q_2 - q_1)\beta_2} \left[ \frac{q_2}{\gamma} - \frac{1}{r} + \frac{\gamma - 2r + k}{r^2}(q_2 - 1) \right] \left[ 1 - e^{-(\alpha - \lambda)(1-q_1)\beta_1 h} \right]$$
Thus $\tilde{V}_{yy}(y, h) > 0$ for $1 \leq y < e^{(\alpha-\gamma)\beta_1 h}$.

If $e^{-(1-\alpha)\beta_2 h} \leq y < 1$, 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 $C_5(h) \geq 0$, 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

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

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

$$= C_7(h)e^{(1-\alpha)(1-q_1)\beta_2 h}$$

$$> 0.$$ 

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

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

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

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

\[\Box\]

References

1. Angoshtari, B., Bayraktar, E., Young, V.R.: Optimal investment and consumption under a habit-formation constraint. SIAM J. Financial Math. 13(1), 321–352 (2022)
2. Arun, T.: The Merton problem with a drawdown constraint on consumption. arXiv:1210.5205 (2012)
3. Bellante, D., Green, C.A.: Relative risk aversion among the elderly. Rev. Financial Econ. 13(3), 269–281 (2004)
4. Brunnermeier, M.K., Nagel, S.: Do wealth fluctuations generate time-varying risk aversion? Micro-evidence on individuals’ asset allocation. Am. Econ. Rev. 98(3), 713–736 (2008)
5. Chapman, D.A.: Habit formation and aggregate consumption. Econometrica 66(5), 1223–1230 (1998)
6. Cohn, A., Engelmann, J., Fehr, E., Maréchal, M.A.: Evidence for countercyclical risk aversion: an experiment with financial professionals. Am. Econ. Rev. 105(2), 860–885 (2015)
7. Deng, S., Li, X., Pham, H., Xiang, Yu.: Optimal consumption with reference to past spending maximum. Finance Stochast. 26(2), 217–266 (2022)
8. Detemple, J.B., Zapatero, F.: Optimal consumption-portfolio policies with habit formation. Math. Finance 2(4), 251–274 (1992)
9. Dybvig, P.H.: Dusenberry’s ratcheting of consumption: optimal dynamic consumption and investment given intolerance for any decline in standard of living. Rev. Econ. Stud. 62(2), 287–313 (1995)
10. Guasoni, P., Huberman, G., Ren, D.: Shortfall aversion. Math. Finance 30(3), 869–920 (2020)
11. Guiso, L., Sapienza, P., Zingales, L.: Time varying risk aversion. J. Financial Econ. 128(3), 403–421 (2018)
12. He, X., Yang, L.: Realization utility with adaptive reference points. Math. Finance 29(2), 409–447 (2019)
13. Jeon, J., Park, K.: Optimal retirement and portfolio selection with consumption ratcheting. Math. Financial Econ. 14(3), 353–397 (2020)
14. Jin, H., Zhou, X.Y.: Behavioral portfolio selection in continuous time. Math. Finance 18(3), 385–426 (2008)
15. Li, X., Yu, X., Zhang, Q.: Optimal consumption with loss aversion and reference to past spending maximum. arXiv:2108.02648 (2022)
16. Merton, R.C.: Lifetime portfolio selection under uncertainty: the continuous time case. Rev. Econ. Stat. 51(3), 247–257 (1969)
17. Muraviev, R.: Additive habit formation: consumption in incomplete markets with random endowments. Math. Financial Econ. 5(2), 67–99 (2011)
18. Paya, I., Wang, P.: Wealth fluctuations and investment in risky assets: the UK micro evidence on households asset allocation. J. Empir. Finance 38, 221–235 (2016)
19. Pollak, R.A.: Habit formation and dynamic demand functions. J. Political Econ. 78(4), 745–763 (1970)
20. Ryder, H.E., Heal, G.M.: Optimal growth with intertemporally dependent preferences. Rev. Econ. Stud. 40(1), 1–31 (1973)
21. Siegel, F.W., Hoban, J.P.: Relative risk aversion revisited. Rev. Econ. Stat. 64(3), 481–487 (1982)
22. Van Bilsen, S., Laeven, R.J.A., Nijman, T.E.: Consumption and portfolio choice under loss aversion and endogenous updating of the reference level. Manag. Sci. 66(9), 3927–3955 (2020)
23. Wachter, J.A., Yogo, M.: Why do household portfolio shares rise in wealth? Rev. Financial Stud. 23(11), 3929–3965 (2010)
24. Xiang, Y.: Utility maximization with addictive consumption habit formation in incomplete semimartingale markets. Ann. Appl. Probab. 25(3), 1383–1419 (2015)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.