Network for Studies on Pensions, Aging and Retirement

NETSPAR ACADEMIC SERIES

Dual Formulation of the Optimal Consumption Problem with Multiplicative Habit Formation

Thijs Kamma, Antoon Pelsser

DP 11/2022-034

Dual Formulation of the Optimal Consumption Problem with Multiplicative Habit Formation*

Thijs Kamma†
Dept. of Quantitative Economics
Maastricht University
NETSPAR

Antoon Pelsser
Dept. of Quantitative Economics
Maastricht University & University of Amsterdam
NETSPAR

November 2, 2022

Abstract

This paper provides a dual formulation of the optimal consumption problem with internal multiplicative habit formation. In this problem, the agent derives utility from the ratio of consumption to the internal habit component. Due to this multiplicative specification of the habit model, the optimal consumption problem is not strictly concave and incorporates irremovable path-dependency. As a consequence, standard Lagrangian techniques fail to supply a candidate for the corresponding dual formulation. Using Fenchel's Duality Theorem, we manage to identify a candidate formulation and prove that it satisfies strong duality. On the basis of this strong duality result, we develop an evaluation mechanism to measure the accuracy of analytical or numerical approximations to the optimal solutions.

Keywords: Fenchel duality, habit formation, life-cycle investment, stochastic optimal control, utility maximisation

JEL Classification: C61, D15, D53, D81, G11

*We are very grateful to An Chen, Tahir Choulli, Michel Vellekoop, Bas Werker and to the conference and seminar participants at the Netspar Pension Day (2021), Tilburg University (2021), Ulm University (2021), the International Pension Workshop (2022), the Winter Seminar on Mathematical Finance (2022), the 11th World Congress of the Bachelier Finance Society (2022), the 25th International Congress on Insurance: Mathematics and Economics (2022) and the 5th European Actuarial Journal Conference (2022), for their helpful comments and suggestions.

†Corresponding author. Mailing address: t.kamma@maastrichtuniversity.nl. PO Box 616, 6200 MD Maastricht, The Netherlands. Phone: +31 (0) 43 388 3695.

1 Introduction

Habit formation describes the phenomenon of an individual growing accustomed to a certain standard of living. In a financial context, this standard of living is dependent on a person's past decisions with regard to saving and consumption. Consuming more or less than a person-specific living standard may impact the utility levels of an individual, cf. Kahneman and Tversky (1979). It is therefore plausible that habit formation affects the current consumption behaviour of a person. To model and analyse the impact of habit-forming tendencies on this behaviour, a wide variety of studies have investigated optimal consumption problems that incorporate a habit level, representing the agent's living standard. These studies can be distinguished into two categories: (i) those that focus on additive habits, and (ii) those that concentrate on multiplicative habits.

We start by discussing the additive habits. In optimal consumption problems with additive habits, the utility-maximising individual draws utility from the *difference* between consumption and a habit level. The literature on these habits is pioneered by Constantinides (1990) and has been studied by e.g. Detemple and Zapatero (1991), Campbell and Cochrane (1999), Munk (2008), Muraviev (2011), and Yu (2015). Additive habit models typically employ arithmetic habit levels, which monotonically increase over time, cf. Detemple and Karatzas (2003), Bodie et al. (2004), and Polkovnichenko (2007). Furthermore, as most standard utility functions only admit strictly positive arguments, additive habit specifications force the agent to maintain consumption above the habit level. For this reason, the habit component is sometimes interpreted as a subsistence level, see e.g. Yogo (2008). This interpretation is sensible for exogenous habits. However, if we assume that habits are endogenous, the habit level depends on the individual's past decisions and becomes person-specific. Consequently, for endogenous habits, it is hard not to consider the habit component as a standard of living that increases over time.

Although individuals have a natural incentive to maintain consumption at least above their living standard, it is clear that additive habit models are too restrictive to be realistic. We attribute this restrictiveness to two main reasons. First of all, in practice, adverse changes in the financial circumstances can urge people to scale down consumption below the level to which they have become accustomed. Second, because of the latter phenomenon, an individual's standard of living may decrease over the course of a lifetime. To arrive at a more realistic model setup that manages to deal with the preceding two situations, the following two modifications can be made. As for the possibility of a declining standard of living, one can employ a geometric specification of the habit level, cf. Kozicki and Tinsley (2002), Corrado and Holly (2011), and van Bilsen et al. (2020a). Unlike the arithmetic habit levels, this geometric specification relies on the logarithmic transformation of consumption, and can therefore decrease over time. As for the possibility of scaling down consumption below the habit level, one can make use of multiplcative habit models.

We now continue with a discussion of the multiplicative habits. Optimal consumption problems with multiplicative habit formation assume that the utility-maximising individual derives utility from the *ratio* of consumption to a habit level. The specification of these habits dates back to Abel (1990), and has been economically advocated by Carroll (2000) and Carroll et al. (2000). Contrary to the additive case, consumption is in this multiplicative setup not constrained to achieve values above the habit level. Namely, since the ratio of consumption to the habit level is always strictly positive, it can be included as an argument in all standard utility functions. The multiplicative habit model consequently allows the agent to reduce consumption levels below the habit component. Furthermore, the multiplicative habit model endows the utility-maximising agent with a strong incentive to fix consumption near/above the habit level. This incentive is due to the fact that the utility function of the agent increases with the magnitude of the ratio.

When the habit level is endogenously determined (internal), standard solution techniques generally fail to solve optimal consumption problems with multiplicative habit formation in closed-form. Because of its dependence on past consumption decisions, the habit component gives rise to path-dependency in the objective function. This path-dependency is irremovable and cannot be handled in an analytical manner.[1] Due to the structure of multiplicative habits, the optimal consumption problem is not strictly concave. In general, non-concave optimisation problems are more difficult to solve than concave ones, see e.g. Chen et al. (2019). To be able to analyse the corresponding optimal solutions, the general approach is to fall back on (i) numerical routines, (ii) approximations or (iii) duality techniques. In a discrete-time setup, Fuhrer (2000) and Gomes and Michaelides (2003) employ numerical methods to analyse the internal multiplicative habit model. More recently, in a continuous-time setup, van Bilsen et al. (2020a) and Li et al. (2021) have made use of an approximation and numerical routines, respectively.[2] Although these studies provide valuable insights into the (optimal) solutions, they ignore potential benefits and insights from duality approaches. In fact, to the best of our knowledge, a dual formulation for the multiplicative habit model is not known.

In this paper, we provide a dual formulation of the optimal consumption problem with internal multiplicative habit formation. We derive this formulation in a continuous-time setup with power utility and a finite trading-horizon. The habit level of the utility-maximising individual is assumed to live by a geometric form. The conventional Lagrangian methods for obtaining dual formulations, e.g. those in Klein and Rogers (2007) and Rogers (2003, 2013), are unable to supply a dual for this multiplicative habit problem. Namely,

[1] This analytical intractability is unique to problems involving multiplicative habits. In case of additive habits, the path-dependency can be eliminated from the problem, cf. Schroder and Skiadas (2002).

[2] We exclusively mention studies that focus on the consumption problem with internal multiplicative habits. Problems involving external habit formation, see e.g. Carroll et al. (1997), Chan and Kogan (2002) and Gómez et al. (2009), do not pose issues when it comes to deriving optimal (duality) results. Martingale duality techniques, developed in the seminal contributions by Pliska (1986), Karatzas et al. (1987), and Cox and Huang (1989, 1991), suffice to analytically solve these consumption problems.

due to the fact that the problem is non-concave and involves path-dependency, the ordinary Legendre transform fails to establish the necessary conjugacy properties. Therefore, we resort to Fenchel's Duality Theorem and a change of variables to derive a dual formulation and prove that strong duality holds. Inspired by Bick et al. (2013) and Kamma and Pelsser (2022), we make use of this strong duality result to develop an evaluation mechanism, suitable for quantifying the accuracy of analytical or numerical approximations to the optimal solutions. This evaluation mechanism spawns a hard upper bound on the welfare losses associated with the approximations, and requires little to no numerical effort. For the approximation developed by van Bilsen et al. (2020a), we employ this mechanism to show that the corresponding welfare losses can be very small.

The remainder of the paper is organised as follows. Section 2 introduces the model setup and the optimal consumption problem. Section 3 presents our main result: the dual formulation. We divide this section into three parts. In the first part, we provide the dual and a rough sketch of its proof. In the second part, we address some technical features of the dual. In the third part, we comment on particular implications of the strong duality result. Subsequently, section 4 outlines the evaluation mechanism and provides some numerical results. Appendix A contains the proof of our main duality result.

2 Model Setup

In this section, we introduce the model setup. First, we lay out the financial market model. Second, we define the agent's wealth process. Third, we specify the agent's habit level. Fourth, we outline the optimal consumption problem.

2.1 Financial Market Model

Our financial market model is N-dimensional, defined in continuous-time, and based on the economic environments provided in Detemple and Rindisbacher (2010), van Bilsen et al. (2020a) and Kamma and Pelsser (2022). We define $T > 0$ as the finite trading or planning horizon, and $[0, T]$ as the corresponding trading interval. Moreover, we introduce the complete filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P})$. The components of this space live by their typical definitions, and its randomness is generated by an $\mathbb{R}^N$-valued standard Brownian motion, $\{W_t\}_{t\in[0,T]}$. As of now, all (in)equalities between random variables and stochastic processes are understood in a $\mathbb{P}$-a.s. or a $\mathrm{d}t \otimes \mathbb{P}$-a.e. sense.

The financial market, $\mathcal{M}$, contains a scalar-valued money market account and N risky assets that are represented by N semi-martingale processes. The money market account submits to the following ordinary differential equation (ODE):

$$\frac{\mathrm{d}B_t}{B_t} = r_t \mathrm{d}t, \;\; B_0 = 1. \tag{2.1}$$

Here, r_t represents the $\mathbb{R}$-valued instantaneous interest rate. We assume that r_t is $\mathcal{F}_t$-progressively measurable and fulfills $\int_0^T |r_t|\,\mathrm{d}t < \infty$. The price processes for the N risky assets (stocks) evolve according to the following stochastic differential equation (SDE) for all $i = 1, \ldots, N$:

$$\frac{\mathrm{d}S_{i,t}}{S_{i,t}} = \mu_{i,t}\mathrm{d}t + \sigma_{i,t}^{\top}\mathrm{d}W_t,\ S_{i,0} = 1, \tag{2.2}$$

where $\mu_{i,t}$ denotes the $\mathbb{R}$-valued instantaneous expected return on stock i and $\sigma_{i,t}$ the $\mathbb{R}^N$-valued vector containing the volatility processes for stock i, both of which are $\mathcal{F}_t$-progressively measurable. We postulate that $\int_0^T \|\mu_t\|_{\mathbb{R}^N}\,\mathrm{d}t < \infty$ and $\int_0^T \mathrm{Tr}\big(\sigma_t\sigma_t^{\top}\big)\mathrm{d}t < \infty$, in which $\mu_t \in \mathbb{R}^N$ has entries $\mu_{i,t}$, and $\sigma_t \in \mathbb{R}^{N\times N}$ rows $\sigma_{i,t}$, $i = 1, \ldots, N$. Observe here that $\|\cdot\|_{\mathbb{R}^N}$ denotes the N-dimensional Euclidean norm and that $\mathrm{Tr}\,(\cdot)$ represents the trace operator. To ensure invertibility of σ_t, we assume that σ_t is non-singular.

Due to the absence of trading restrictions, this financial market is complete, i.e. all traded risks are hedgeable. Hence, by the fundamental theorem of asset pricing, as formulated by Delbaen and Schachermayer (1994), there must exist a unique equivalent martingale measure. Correspondingly, there must exist a unique state price density (SPD), $\{M_t\}_{t\in[0,T]}$. Define $\lambda_t := \sigma_t^{-1}(\mu_t - r_t 1_N)$ as the market price of risk, then M_t reads:

$$\frac{\mathrm{d}M_t}{M_t} = -r_t\mathrm{d}t - \lambda_t^{\top}\mathrm{d}W_t,\ M_0 = 1. \tag{2.3}$$

Note that $\{B_t\}_{t\in[0,T]}$ is selected as the numéraire quantity. We assume that $\{\lambda_t\}_{t\in[0,T]}$ satisfies $\mathbb{E}\big[\exp\big(\frac{1}{2}\int_0^T \|\lambda_s\|_{\mathbb{R}^N}^2\mathrm{d}s\big)\big] < \infty$, cf. Karatzas and Shreve (1991). Moreover, we postulate that $\{\lambda_t\}_{t\in[0,T]}$ is such that M_t and $\log M_t$ attain values in $L^2(\Omega\times[0,T])$.[3] The latter assumption is necessary to assure well-posedness of the dual formulation. In order to evaluate financial instruments in a risk-neutral fashion, one can make use of M_t. For example, M_tB_t and M_tS_t are both $\mathbb{P}$-martingales with respect to $\{\mathcal{F}_t\}_{t\in[0,T]}$.

2.2 Dynamic Wealth Process

In this environment, the agent is free to continuously select an investment and a consumption strategy over $[0,T]$. Specifically, the agent's wealth process, $\{X_t\}_{t\in[0,T]}$, is affected by two endogenous terms: (i) a process for the proportion of wealth that is allocated to the stock, $\{\pi_t\}_{t\in[0,T]}$, and (ii) a consumption process, $\{c_t\}_{t\in[0,T]}$. We assume that both preceding endogenous processes are $\mathcal{F}_t$-progressively measurable. Let us fix a deterministic initial endowment, $X_0 \in \mathbb{R}_+$. Then, the agent's wealth process is defined by:

$$\mathrm{d}X_t = X_t\left[\left(r_t + \pi_t^{\top}\sigma_t\lambda_t\right)\mathrm{d}t + \pi_t^{\top}\sigma_t\mathrm{d}W_t\right] - c_t\mathrm{d}t, \tag{2.4}$$

[3] We define $L^p(\Omega\times[0,T];\mathbb{R}^n)$ as the standard Lebesgue space of all $\mathcal{F}_t$-progressively measurable functions, $f:\Omega\times[0,T]\to\mathbb{R}^n$, satisfying $\big(\int_{\Omega\times[0,T]}\|f_t\|_{\mathbb{R}^n}^p\,\mathbb{P}(\mathrm{d}t)\big)^{1/p} = \big(\mathbb{E}\big[\int_0^T\|f_t\|_{\mathbb{R}^n}^p\,\mathrm{d}t\big]\big)^{1/p} < \infty$. If $n = 1$, we drop the "$\mathbb{R}$"-notation from the definition of the L^p space.

Clearly, $\{c_t\}_{t\in[0,T]}$ is $\mathbb{R}_+$-valued and $\{\pi_t\}_{t\in[0,T]}$ is $\mathbb{R}^N$-valued. A trading-consumption pair, $\{c_t, \pi_t\}_{t\in[0,T]}$, is said to be *admissible* if it satisfies the following set of conditions: $X_t \geq 0$, $\int_0^T \pi_t^\top \sigma_t \sigma_t^\top \pi_t \mathrm{d}t < \infty$, $\int_0^T \left|\pi_t^\top \sigma_t \lambda_t + r_t X_t\right| \mathrm{d}t < \infty$, and $\log c_t \in L^2(\Omega \times [0,T])$. The set containing all admissible trading-consumption pairs is denoted by $\mathcal{A}_{X_0}$. Observe that the proportion of wealth that is allocated to the cash account can be recovered from $1 - \pi_t^\top 1_N$, where 1_N is an $\mathbb{R}^N$-valued vector containing only 1's. This specific proportion only plays a role through $\{\pi_t\}_{t\in[0,T]}$, due to which it can be excluded from the representation for $\{X_t\}_{t\in[0,T]}$. See e.g. Cuoco (1997) for a situation in which this is not the case.

2.3 Habit Level

The economic environment $\mathcal{M}$ consists of a utility-maximising agent who is internally habit-forming. As a consequence, the individual is in possession of a habit level, h_t at time $t \in [0,T]$. This habit level represents the level of consumption to which the agent has become accustomed. Naturally, h_t depends on the agent's preferences and his/her corresponding past consumption behaviour. Due to this dependence on past consumption decisions, the habit level constitutes an endogenous (internal) component. If h_t is exogenously determined ($\beta = 0$ below), the agent is externally habit-forming. By analogy with van Bilsen et al. (2020a) and references therein, we suppose that the logarithmic transformation of this habit level, h_t, is given by:

$$\mathrm{d}\log h_t = (\beta \log c_t - \alpha \log h_t)\,\mathrm{d}t, \;\; \log h_0 = 0. \tag{2.5}$$

The parameter $\beta \in \mathbb{R}_+$ expresses the relative importance of past consumption decisions in the specification of $\log h_t$. For large values of β, more weight is attached to these past consumption choices. For small values of β, the converse is true. The parameter $\alpha \in \mathbb{R}_+$ stands for the habit level's rate of depreciation. For small values of α, the habit level depends on past consumption decisions over a large time-horizon. For large values of α, the converse is true. We assume that $\alpha \geq \beta$ holds, for concavity purposes related to the optimal consumption problem. The limiting case $\alpha = \beta = 0$ results in $h_t = 1$ for all $t \in [0,T]$. Setting $\alpha = \beta = 0$ consequently recovers a model without habit formation.

We note that the solution to the ODE in (2.5) reads for all $t \in [0,T]$ as:

$$\log h_t = \beta \int_0^t e^{-\alpha(t-s)} \log c_s \mathrm{d}s. \tag{2.6}$$

Hence, the habit level lives by a geometric form. That is, $h_t = \exp\left\{\beta \int_0^t e^{-\alpha(t-s)} \log c_s \mathrm{d}s\right\}$ holds for all $t \in [0,T]$. In contrast with arithmetic habits, cf. Constantinides (1990) and van Bilsen et al. (2020b), this specification of h_t is not strictly increasing in time. As the geometric form consequently allows for decreases in h_t over $t \in [0,T]$, the interpretation of this habit component as a standard of living is more sensible. Ultimately, we observe that $\log h_t$ in (2.6) can be represented as follows: $\log h_t = \alpha \int_0^t e^{-\alpha(t-s)} \log c_s^{\beta/\alpha} \mathrm{d}s$, for all $t \in [0, T]$. This representation indicates that h_t can be interpreted as the geometric weighted moving average (GWMA) of transformed past consumption decisions, $\left\{c_s^{\beta/\alpha}\right\}_{s\in[0,t]}$. Clearly, if $\alpha = \beta \neq 0$, $\frac{c_t}{h_t}$ becomes a dimensionless quantity, and h_t reduces to the ordinary GWMA of (non-transformed) past consumption decisions, $\{c_s\}_{s\in[0,t]}$.

2.4 Optimal Consumption Problem

The habit-forming agent in $\mathcal{M}$ is at $t = 0$ in possession of a predetermined amount of cash, $X_0 \in \mathbb{R}_+$, and lives until $t = T$. Throughout the trading interval, $[0, T]$, this agent seeks to maximise expected lifetime utility from the ratio of consumption to the habit process by continuously selecting his/her consumption levels and corresponding portfolio weights. The habit-forming agent must determine these controls in agreement with the dynamic budget constraint in (2.4), such that the admissibility conditions are met. We assume that the preferences of the individual are characterised by the Von-Neumann-Morgenstern index: $\mathbb{E}\big[\int_0^T e^{-\delta t} U\left(c_t/h_t\right) \mathrm{d}t\big]$, cf. Detemple and Zapatero (1991). Consistent with this description, the agent faces the following problem:

$$
\begin{aligned}
\sup_{\{c_t,\pi_t\}_{t\in[0,T]}\in\mathcal{A}_{X_0}} \quad & \mathbb{E}\left[\int_0^T e^{-\delta t} U\left(\frac{c_t}{h_t}\right) \mathrm{d}t\right] \\
\text{s.t.} \quad & \mathrm{d}X_t = X_t\left[\left(r_t + \pi_t^\top \sigma_t \lambda_t\right) \mathrm{d}t + \pi_t^\top \sigma_t \mathrm{d}W_t\right] - c_t \mathrm{d}t, \\
& \mathrm{d}\log h_t = \left(\beta \log c_t - \alpha \log h_t\right) \mathrm{d}t, \quad h_0 = 1, \; X_0 \in \mathbb{R}_+.
\end{aligned} \tag{2.7}
$$

In this problem, $\delta \in \mathbb{R}_+$ represents the agent's time-preference parameter and $U : \mathbb{R}_+ \to \mathbb{R}$ denotes the agent's utility function. For simplicity, we assume that utility is given by an ordinary power (CRRA) function, $U(x) = \frac{x^{1-\gamma}}{1-\gamma}$ for all $x \in \mathbb{R}_+$. Here, γ defines the coefficient of relative risk-aversion. For purposes related to concavity of the optimisation problem, we fix $\gamma > 1$. We denote the first and second derivatives of U by U' and U'', respectively. The first derivative of U, U', is also known as marginal utility. By $I : \mathbb{R}_+ \to \mathbb{R}_+$, we denote the inverse of marginal utility ($I = U'^{-1}$). The preceding optimisation problem is not strictly concave and does not permit a derivation of the optimal controls in closed-form, because of the path-dependency induced by h_t. Namely, utility at time t depends, apart from on c_t alone, through h_t as well on $\{c_s\}_{s\in[0,t]}$.

3 Dual Formulation

In this section, we provide the main result of this paper (Theorem 3.1): the dual formulation of the optimal consumption problem in (2.7). We divide this section into three parts. First, we present the dual formulation and formalise that it satisfies strong duality, in Theorem 3.1. In the same part, we provide a rough sketch of the proof that culminates in the preceding strong duality result. Second, we present a set of three technical remarks related to the dual formulation. Third, we discuss certain implications of the strong duality result concerning the analytical structure of the optimal controls.

3.1 Main Result: Strong Duality

Theorem 3.1 contains the main result of this paper. Its statement formalises the fact that the optimal (dual) control problem in (3.1) and the optimal consumption problem in (2.7) are dual to each other, satisfying strong duality. First, we provide the theorem itself. Second, we comment on its corresponding proof.

Theorem 3.1. *Consider the optimal consumption problem in* (2.7) *and define the primal objective function:* $J(X_0, \{c_t, \pi_t\}) = \mathbb{E}\big[\int_0^T e^{-\delta t} U\big(\frac{c_t}{h_t}\big)\mathrm{d}t\big]$. *Furthermore, introduce the following concave conjugate:* $V(x) = \inf_{z\in\mathbb{R}}\{xz - (-e^{-z})\} = x - x\log x$, *for all* $x \in \mathbb{R}_+$. *Then, the dual formulation of the optimal consumption problem in* (2.7) *is given by:*

$$\begin{aligned}\inf_{\psi_t\in L^2(\Omega\times[0,T]),\eta\in\mathbb{R}_+} \mathbb{E}\Bigg[\int_0^T \Bigg\{ & e^{-\delta t}\frac{1}{1-\gamma}V\left(e^{\delta t}\psi_t\right) \\ & -\eta M_t V\left(\frac{\psi_t - \beta\mathbb{E}\left[\int_t^T e^{-\alpha(s-t)}\psi_s \mathrm{d}s \,\middle|\, \mathcal{F}_t\right]}{\eta M_t}\right)\Bigg\}\mathrm{d}t\Bigg] + \eta X_0.\end{aligned} \tag{3.1}$$

That is, suppose that $\mathcal{V}(X_0, \psi_t, \eta)$ *represents the dual objective function of* (3.1). *Then, the problems in* (2.7) *and* (3.1) *satisfy strong duality:*

$$\sup_{\{c_t,\pi_t\}_{t\in[0,T]}\in\mathcal{A}_{X_0}} J(X_0, \{c_t, \pi_t\}) = \inf_{\psi_t\in L^2(\Omega\times[0,T]),\eta\in\mathbb{R}_+} \mathcal{V}(X_0, \psi_t, \eta), \tag{3.2}$$

for all $X_0 \in \mathbb{R}_+$

Proof. The proof is given in Appendix A. □

Typically, the Legendre transform alone suffices to establish a strong duality result. However, due to the non-concavity and path-dependency of the objective of (2.7), the Legendre transform cannot be used to derive strong duality. Therefore, to prove Theorem 3.1, we apply a change of variables and employ Fenchel Duality, cf. Proposition A.1. This form of duality can be regarded as a generalisation of the Legendre result to problems involving path-dependent linear transformations of one of the control variables. On the basis of Fenchel Duality, deriving strong duality for problems (2.7) and (3.1) is straightforward. First, we re-express the primal problem (2.7) in terms of its static equivalent and $\log c_t$. Second, we use Fenchel Duality to demonstrate that strong duality holds for the static problem and $\inf_{\psi_t\in L^2(\Omega\times[0,T])}\mathcal{V}(X_0, \psi_t, \eta)$. Third and last, we resort to a technical argument (Lemma A.2) in order to extend this strong duality result to (2.7) and (3.1).

3.2 Technical Remarks

In this section, we address three technical aspects of the dual formulation in Theorem 3.1. The first two aspects touch upon the dual control variable, $\{\psi_t\}_{t\in[0,T]}$. As for the first aspect, we know that this control variable attains values in the "unconstrained" set $L^2(\Omega \times [0,T])$. However, the function V is defined over $\mathbb{R}_+$. Consequently, the dual forces the following two constraints upon $\{\psi_t\}_{t\in[0,T]}$: $\psi_t > 0$ and $\psi_t - \beta\mathbb{E}\big[\int_t^T e^{-\alpha(s-t)}\psi_s \mathrm{d}s \mid \mathcal{F}_t\big] > 0$ for all $t \in [0,T]$. The second aspect is closely related to the latter and concerns an alternative representation of the dual in (3.1). Suppose that we define a process $p_t = \psi_t - \beta\mathbb{E}\big[\int_t^T e^{-\alpha(s-t)}\psi_s \mathrm{d}s \mid \mathcal{F}_t\big]$ for all $t \in [0,T]$. Then, the latter identity can be regarded as a Volterra equation for ψ_t with the following solution: $\psi_t = p_t + \beta\mathbb{E}\big[\int_t^T e^{-[\alpha-\beta](s-t)} p_s \mathrm{d}s \mid \mathcal{F}_t\big]$ for all $t \in [0,T]$. Therefore, the dual formulation in (3.1) is identical to the following one:

$$\begin{aligned}\inf_{p_t\in L^2(\Omega\times[0,T]),\eta\in\mathbb{R}_+} \mathbb{E}\Bigg[\int_0^T \Bigg\{ & -\eta M_t V\left(\frac{p_t}{\eta M_t}\right) \\ & e^{-\delta t}\frac{1}{1-\gamma} V\left(e^{\delta t}\left\{p_t + \beta\mathbb{E}\left[\int_t^T e^{-[\alpha-\beta](s-t)} p_s \mathrm{d}s \,\middle|\, \mathcal{F}_t\right]\right\}\right)\Bigg\}\mathrm{d}t\Bigg] + \eta X_0.\end{aligned} \tag{3.3}$$

As in case of $\{\psi_t\}_{t\in[0,T]}$, it is clear that the the preceding dual formulation forces the following constraints upon $\{p_t\}_{t\in[0,T]}$: $p_t > 0$ and $p_t + \beta\mathbb{E}\big[\int_t^T e^{-[\alpha-\beta](s-t)} p_s \mathrm{d}s \mid \mathcal{F}_t\big] > 0$ for all $t \in [0,T]$. Although this re-definition of the dual control variable does not affect the dual optimality conditions, it implies slightly different duality relations. In the subsequent section 3.3, we address the nature of these relations in more detail.

The third and final aspect concerns the dual formulation for the model setup without habit formation. To recover this no-habit case, it suffices to fix $\alpha = \beta = 0$. Setting $\alpha = \beta = 0$ in the dual of Theorem 3.1 provides us with the following dual formulation:

$$\inf_{\psi_t\in L^2(\Omega\times[0,T]),\eta\in\mathbb{R}_+} \mathbb{E}\left[\int_0^T \left\{ e^{-\delta t}\frac{V\left(e^{\delta t}\psi_t\right)}{1-\gamma} - \eta M_t V\left(\frac{\psi_t}{\eta M_t}\right)\right\}\mathrm{d}t\right] + \eta X_0. \tag{3.4}$$

Here, the dual forces $\{\psi_t\}_{t\in[0,T]}$ to satisfy $\psi_t > 0$ for all $t \in [0,T]$. In line with the exclusion of h_t in the primal, the no-habit dual does not contain the $\mathbb{E}\big[\int_t^T e^{-\alpha(s-t)}\psi_s \mathrm{d}s \mid \mathcal{F}_t\big]$ term. The dual in (3.4) differs from the conventional one in e.g. Cvitanić and Karatzas (1992). Note that this is not troublesome, as the dual formulations for convex optimisation problems are not unique, cf. Rockafellar (2015). In fact, after inserting the optimal dual control, say ψ_t^{opt} (cf. Example 3.1), into (3.4), we find the conventional formulation. The aforementioned difference is attributable to the fact that this dual ensues from an application of Fenchel Duality. Namely, this notion of duality involves two convex conjugates instead of one. Moreover, it requires one to re-express the primal control as follows: $c_t = e^{-(-\log c_t)}$ for all $t \in [0,T]$. Due to these two features, the dual accommodates two functions that coincide with the concave conjugate of the exponential utility function ($x \mapsto -e^{-x}$).

3.3 Duality Relations

For convex optimisation problems, duality theory can be employed to disclose the relationship between the primal and dual controls, i.e. the duality relation. This duality relation infers how the primal controls analytically depend on the dual controls, and vice versa. The key characteristic of this relation is that it yields the optimal primal (dual) controls after insertion of the optimal dual (primal) controls (respectively). Therefore, the duality relation contains important information about the analytical structure of the optimal primal and dual variables. In addition to this, it provides an alternative view on the mechanisms that are involved with optimising the primal and dual problems. As the dual in (3.1) follows from Fenchel Duality rather than from the Legendre transform, its implied duality relations differ from the conventional ones. In fact, the duality relations[4] for the problems in (2.7) and (3.1) are for all $t \in [0,T]$ given by:

$$c_t^* = \frac{\psi_t - \beta \mathbb{E}\left[\int_t^T e^{-\alpha(s-t)} \psi_s \mathrm{d}s \;\middle|\; \mathcal{F}_t\right]}{\eta M_t} \quad \text{and} \quad h_t^* = c_t^* \left(e^{\delta t}\psi_t\right)^{\frac{1}{\gamma-1}}. \tag{3.5}$$

In a technical sense, the duality relation for consumption in (3.5), c_t^*, can be regarded as a specification of optimal consumption in some auxiliary (artificial) market. To ensure that consumption defined by this relation is admissible and optimal in the true market, the dual problem in (3.1) aims to characterise this identity for c_t^* in such a manner that it generates the habit level in (3.5), h_t^*. In an economic sense, we note that dual-implied consumption c_t^* is endowed with a "penalty term". Concretely, selecting high values for ψ_t at future dates, requires one to increase ψ_t today so as to arrive at similar utility levels. This mechanism inversely reflects the agent's viewpoint in the primal problem. Namely, if this agent selects high values for c_t today, via h_t, he/she is required to increase c_t even further to maintain similar utility levels. To obtain some insights into the role that ψ_t plays in minimising $\mathcal{V}$, we now conclude with Example 3.1.

Example 3.1. *Suppose that $\alpha = \beta = 0$. Then, minimisation of the dual results in:*

$$\psi_t^{\mathrm{opt}} = \left(\eta^{\mathrm{opt}} M_t\right) \left[e^{\delta t} \eta^{\mathrm{opt}} M_t\right]^{-\frac{1}{\gamma}}, \tag{3.6}$$

for all $t \in [0,T]$. Here, ψ_t^{opt} denotes the optimal dual process, and η^{opt} represents the corresponding dual-optimal constant. In particular, η^{opt} can be obtained from solving $\mathbb{E}\big[\int_0^T \frac{\psi_t^{\mathrm{opt}}}{\eta^{\mathrm{opt}} M_t} \mathrm{d}t\big] = X_0$ for η^{opt}. From the duality relations provided in (3.5), we know that ψ_t^{opt} should generate c_t^{opt} via $c_t^* = \frac{\psi_t - \beta\mathbb{E}\left[\int_t^T e^{-\alpha(s-t)}\psi_s \mathrm{d}s \,\middle|\, \mathcal{F}_t\right]}{\eta M_t}$. Using that $c_t^* = \frac{\psi_t}{\eta M_t}$ for

[4]These duality relations follow from the fact that the primal and dual objectives, in (2.7) and (3.1), are conjugate to each other. This concretely means that these expressions bind in the unique "point" outlined by (3.5), conditional on $\eta X_0 = \eta \mathbb{E}\big[\int_0^T c_t M_t \mathrm{d}t\big]$ being true. Note that the duality relation in (3.5) corresponds to the dual in (3.1). For the alternative, howbeit identical, representation in (3.3), the duality relations read: $c_t^* = \frac{p_t}{\eta M_t}$ and $h_t^* = c_t^*\big(e^{\delta t}\big\{p_t + \beta\mathbb{E}\big[\int_t^T e^{-[\alpha-\beta](s-t)} p_s \mathrm{d}s \mid \mathcal{F}_t\big]\big\}\big)^{\frac{1}{\gamma-1}}$ for all $t \in [0,T]$.

$\alpha = \beta = 0$, *we therefore find that optimal consumption is given by:*

$$c_t^{\text{opt}} = \frac{\psi_t^{\text{opt}}}{\eta^{\text{opt}} M_t} = \left(e^{\delta t} \eta^{\text{opt}} M_t\right)^{-\frac{1}{\gamma}}, \tag{3.7}$$

for all $t \in [0, T]$. *Employing the definition of* $I : \mathbb{R}_+ \to \mathbb{R}_+$, *we accordingly have that* $c_t^{\text{opt}} = I\left(e^{\delta t} \eta^{\text{opt}} M_t\right)$ *holds. Moreover, in the optimum characterised by* c_t^{opt} *and* ψ_t^{opt}, *the value for* $\eta^{\text{opt}} \in \mathbb{R}_+$ *is determined such that* $\mathbb{E}\left[\int_0^T c_t^{\text{opt}} M_t \mathrm{d}t\right] = X_0$ *holds. Hence, it is clear that* c_t^{opt} *in* (3.7) *coincides with optimal consumption in the no-habit case* ($\alpha = \beta = 0$). *See for instance Merton (1971) for a similar representation of* c_t^{opt}.

4 Approximations

In this section, we develop a duality-based mechanism for evaluating the accuracy of approximations to the optimal solutions of (2.7). We break this section down into three parts. First, we provide the general evaluation mechanism and comment on related technicalities. Second, we present the approximate solution proposed by van Bilsen et al. (2020a). In addition to this, we rely on the duality relation in (3.5) to develop a corresponding dual approximation. Third, we make use of the evaluation mechanism to numerically study the precision of these approximations.

4.1 Evaluation Mechanism

To quantify the accuracy of approximations to the optimal solutions of (2.7), we develop an evaluation mechanism. This mechanism is predicated on the evaluation techniques proposed in Bick et al. (2013) and Kamma and Pelsser (2022). These techniques make use of strong duality to note that any departure from the optimal primal and/or dual controls results in a duality gap. Concretely, the primal value function, $J\left(X_0, \{c_t, \pi_t\}\right)$, delivers a lower bound on the optimal dual value function, for each admissible pair $\{c_t, \pi_t\}_{t \in [0,T]} \in \mathcal{A}_{X_0}$. Likewise, the dual value function, $\mathcal{V}\left(X_0, \psi_t, \eta\right)$, spawns an upper bound on the optimal primal value function, for each feasible pair $(\eta, \psi_t) \in \mathbb{R}_+ \times \Psi$, where $\Psi := \left\{\psi_t \in L^2\left(\Omega \times [0, T]\right) \;\middle|\; \psi_t > 0,\ \psi_t > \beta \mathbb{E}\left[\int_t^T e^{-\alpha(s-t)} \psi_s \mathrm{d}s \;\middle|\; \mathcal{F}_t\right],\ \forall\, t \in [0, T]\right\}$. To be more precise, for all $X_0 \in \mathbb{R}_+$, $\{c_t, \pi_t\}_{t \in [0,T]} \in \mathcal{A}_{X_0}$ and $(\eta, \psi_t) \in \mathbb{R}_+ \times \Psi$, we have:

$$J\left(X_0, \{c_t, \pi_t\}\right) \leq \mathcal{V}\left(X_0, \psi_t, \eta\right). \tag{4.1}$$

Theorem 3.1 infers that the inequality in (4.1) binds if and only if $(c_t, \pi_t) = \left(c_t^{\text{opt}}, \pi_t^{\text{opt}}\right)$ and $(\eta, \psi_t) = \left(\eta^{\text{opt}}, \psi_t^{\text{opt}}\right)$ for all $t \in [0, T]$. Here, we employ the superscript "opt" to indicate that these concern the optimal primal/dual control variables. For the former reason, the difference between J and $\mathcal{V}$ grows, the farther $\{c_t, \pi_t\}_{t \in [0,T]}$ and/or (η, ψ_t) are situated from the optima. We can employ this observation to gauge the accuracy of particular approximations as follows. Suppose that $\{c'_t, \pi'_t\}_{t\in[0,T]} \in \mathcal{A}_{X_0}$ represents an arbitrary admissible trading-consumption pair and that $(\eta', \psi'_t) \in \mathbb{R}_+ \times \Psi$ specifies a feasible pair of dual controls. Then, $D(X_0) = \mathcal{V}(X_0, \psi'_t, \eta') - J(X_0, \{c'_t, \pi'_t\})$ characterises for all $X_0 \in \mathbb{R}_+$ the corresponding duality gap. As it is difficult to interpret the quantity $D(X_0) \in \mathbb{R}_+$, we calculate $\mathcal{C} \in \mathbb{R}_+$ such that the following equality holds:

$$J(X_0, \{c'_t, \pi'_t\}) = \mathcal{V}(X_0[1-\mathcal{C}], \psi'_t, \eta'). \tag{4.2}$$

Here, $\mathcal{C}$ can be interpreted as a fee that one pays to gain access to the optimal trading-consumption pair. In Bick et al. (2013) and Kamma and Pelsser (2022), a similar interpretation is used. Note that $\mathcal{C}$ grows with the magnitude of $D(X_0)$, and thus with the difference(s) between $\{c'_t, \pi'_t\}_{t\in[0,T]}$ and $\left\{c_t^{\text{opt}}, \pi_t^{\text{opt}}\right\}_{t\in[0,T]}$, as well as between (η', ψ'_t) and $\left(\eta^{\text{opt}}, \psi_t^{\text{opt}}\right)$. Note that this way of calculating the approximation error is numerically not demanding, as one does not require the optimal controls to obtain the error.

Suppose that we are in possession of an approximate trading-consumption pair, $\{c'_t, \pi'_t\}_{t\in[0,T]}$. Calculating the corresponding lower bounds (J) is straightforward. To see how we may obtain matching upper bounds ($\mathcal{V}$), we note that

$$c'_t = \frac{\psi'_t - \beta\mathbb{E}\left[\int_t^T e^{-\alpha(s-t)}\psi'_s \mathrm{d}s \;\middle|\; \mathcal{F}_t\right]}{\eta' M_t} \quad \text{and} \quad \widehat{c}'_t = \left(e^{\delta t}\psi'_t\right)^{\frac{1}{1-\gamma}}, \tag{4.3}$$

follows from the duality relations in (3.5). Here, we define $\widehat{c}'_t := \frac{c'_t}{h'_t}$ as approximate ratio consumption, where h'_t represents the corresponding approximate habit level. The first identity in (4.3) indicates that each admissible consumption strategy corresponds to dual control variable, ψ'_t. Similarly, the second identity shown in (4.3) demonstrates that each ratio consumption strategy implies an (analytical) expression for ψ'_t. One can then obtain a matching approximation to the dual constant, η', from: $\mathbb{E}\big[\int_0^T \frac{1}{\eta'}\big(\psi'_t - \beta\mathbb{E}\big[\int_t^T e^{-\alpha(s-t)}\psi'_s \mathrm{d}s \;\big|\; \mathcal{F}_t\big]\big)\mathrm{d}t\big] = X_0$ (conditional on $\psi'_t \in \Psi$). Since it could be the case that $\psi'_t \notin \Psi$, it may be necessary to project $\{\psi'_t\}_{t\in[0,T]}$ into Ψ, to ensure dual-feasibility of (η', ψ'_t). For this purpose, we introduce a projection operator $\text{proj}_\Psi : L^2(\Omega \times [0,T]) \to \Psi$ and re-define the primal-implied dual controls as follows:

$$\widehat{\psi}'_t = \text{proj}_\Psi(\psi'_t) \quad \text{and} \quad \widehat{\eta}' = \mathcal{R}^{-1}(X_0). \tag{4.4}$$

We define $\mathcal{R}^{-1} : \mathbb{R}_+ \to \mathbb{R}_+$ as the inverse of the following function: $\mathcal{R}(\eta) = \mathbb{E}\big[\int_0^T \frac{1}{\eta'}\big(\widehat{\psi}'_t - \beta\mathbb{E}\big[\int_t^T e^{-\alpha(s-t)}\widehat{\psi}'_s \mathrm{d}s \;\big|\; \mathcal{F}_t\big]\big)\mathrm{d}t\big]$. Clearly, $\left(\widehat{\eta}', \widehat{\psi}'_t\right)$ is feasible and generates an upper bound, $\mathcal{V}$, on the optimal value function. This enables us to determine $\mathcal{C}$ as in (4.2).

We would like to make three remarks. First, we observe that one can distil (feasible) dual controls from either c'_t or $\widehat{c}'_t$ in (4.3). The dual controls implied by c'_t differ from those implied by $\widehat{c}'_t$, except when $c'_t = c_t^{\text{opt}}$ and $\widehat{c}'_t = \widehat{c}_t^{\text{opt}}$. Second, we note that it is possible to obtain an analytical expression for ψ_t' from the duality relation for c_t' in (4.3). Similar to the alternative representation of the dual in (3.3), it suffices to identify the appropriate Volterra equation. Third and last, we stress that our evaluation mechanism does not need to be conceptually modified for the alternative dual in (3.3). The mere adjustment that has to be made is the analytical specification of the duality relations in (4.3).

4.2 Approximate Controls

Subsequently, we present the analytical approximation to optimal (ratio) consumption, $\widehat{c}_t^{\text{opt}}$, proposed by van Bilsen et al. (2020a). On the grounds of our evaluation mechanism, and the relations in (4.3), we develop a corresponding novel dual approximation. In the sequel, we do not pay attention to the trading strategy that hedges approximate consumption. Namely, here this trading strategy does not play a role in the specification of the approximate value function. The approximation of van Bilsen et al. (2020a) is based on a first-order Taylor expansion of the budget constraint in the static representation of (2.7), i.e. $\mathbb{E}\big[\int_0^T M_t\widehat{c}_t h_t \mathrm{d}t\big] \leq X_0$, around $\{\widehat{c}_t\}_{t\in[0,T]} = 1$. The motivation for such an expansion is that the habit level closely tracks optimal consumption. According to Theorem 3.1 in van Bilsen et al. (2020a), this approximation is for all $t \in [0,T]$ given by:

$$\widehat{c}_t' = \left(\eta' e^{\delta t} M_t \left\{1 + \beta \mathbb{E}\left[\int_t^T e^{-[\alpha-\beta](s-t)} \frac{M_s}{M_t} \mathrm{d}s \;\middle|\; \mathcal{F}_t\right]\right\}\right)^{-\frac{1}{\gamma}}. \tag{4.5}$$

Here, $\eta' \in \mathbb{R}_+$ is determined such that $\mathbb{E}\big[\int_0^T M_t\widehat{c}_t' h_t' \mathrm{d}t\big] = X_0$. Moreover, h_t' is the approximate habit level generated by $\{\widehat{c}_s'\}_{s\in[0,T]}$, i.e. $h_t' = e^{\beta\int_0^t e^{-[\alpha-\beta](t-s)} \log \widehat{c}_s' \mathrm{d}s}$.

Consistent with the outline of our evaluation mechanism, we must determine a corresponding dual approximation in order to measure the accuracy of $\widehat{c}_t'$ in (4.5). For this purpose, we must utilise either of the duality relations as shown in (4.3). Although the identity for $\widehat{c}_t'$ allows for an easy recovery of ψ_t', it may be the case that $\psi_t' \notin \Psi$. As we wish to avoid rigorous modifications enforced upon ψ_t' by a projection operator, we resort to the identity for c_t' instead. Note that $c_t' = \widehat{c}_t' h_t'$ for all $t \in [0,T]$. In the identity for c_t' in (4.3), we can recognize a clear Volterra equation. Its solution for ψ_t' can accordingly be formulated as: $\psi_t' = \eta M_t c_t' + \beta\mathbb{E}\big[\int_t^T e^{-[\alpha-\beta](s-t)} \eta M_s c_s' \mathrm{d}s \;\big|\; \mathcal{F}_t\big]$ for all $t \in [0,T]$. Hence,

$$\psi_t' = \eta' M_t c_t' \left(1 + \beta\mathbb{E}\left[\int_t^T e^{-[\alpha-\beta](s-t)} \frac{M_s c_s'}{M_t c_t'} \mathrm{d}s \;\middle|\; \mathcal{F}_t\right]\right), \tag{4.6}$$

characterises the approximation to ψ_t^{opt}, implied by the duality relation for c_t' in (4.3). Note that $\eta M_t c_t' > 0$ holds for all $t \in [0,T]$. As a consequence, we have that $\psi_t' > 0$ is true for all $t \in [0,T]$. Moreover, by construction, it is the case that $\psi_t' - \beta\mathbb{E}\left[\int_t^T e^{-\alpha(s-t)}\psi_s' \mathrm{d}s \;\middle|\; \mathcal{F}_t\right] = \eta' M_t c_t' > 0$ for all $t \in [0,T]$. Therefore, $\psi_t' \in \Psi$, implying that the approximation in (4.6) is dual-feasible. It should be noted that both approximations are truly optimal

	Coefficient of risk-aversion (γ)					Initial endowment (X_0)				
	6	8	10	12	14	8	9	10	11	12
$\mathcal{C}$ (%)	0.103	0.061	0.042	0.031	0.025	0.093	0.047	0.042	0.067	0.114
RSS	0.062	0.037	0.026	0.020	0.016	0.028	0.011	0.026	0.050	0.075
	Speed of mean-reversion ($\alpha = \beta$)					Time-preference (δ)				
	0.01	0.05	0.1	0.15	0.2	0.01	0.02	0.03	0.04	0.05
$\mathcal{C}$ (%)	0.000	0.010	0.042	0.095	0.167	0.048	0.045	0.042	0.039	0.036
RSS	0.003	0.014	0.026	0.037	0.048	0.027	0.026	0.026	0.026	0.026

Table 1. Upper bounds on welfare losses ($\mathcal{C}$). In the row denoted by $\mathcal{C}$, the table reports the upper bounds on the welfare losses corresponding to the approximate solution provided in equation (4.5). Additionally, in the row denoted by RSS, the table reports the "root" sum of squares (RSS) corresponding to this approximate solution. The welfare losses are calculated by solving (4.2) for $\mathcal{C}$, and are expressed in terms of percentages (%). The RSS is calculated as follows: $\text{RSS} = \mathbb{E}\big[\sum_{i=1}^{M}(c'_{t_i} - c''_{t_i})^2\big]^{\frac{1}{2}}$, where c'_{t_i} is approximate consumption for (4.5), and c''_{t_i} is approximate consumption implied by ψ'_t in (4.6) via the second duality relation in (4.3). Moreover, $t_1 < t_2 < \ldots < t_M$ represent M time-steps in the Euler scheme. The table reports $\mathcal{C}$ and RSS for different values of the four displayed parameters (γ, X_0, $\alpha = \beta$ and δ), under a baseline initialisation of the parameters. This baseline set is fixed as follows: $X_0 = 10$, $T = 10$, $\gamma = 10$, $\delta = 0.03$, $\alpha = \beta = 0.1$, $\mu = 0.05$, $r = 0.01$ and $\sigma = 0.2$. The results are based on $10,000$ simulations and an Euler scheme with 20 equidistant time-steps.

in the no-habit case ($\alpha = \beta = 0$). For the approximation to optimal consumption in (4.5), fixing $\alpha = \beta = 0$, yields that $\widehat{c}_t = c'_t = \left(e^{\delta t}\eta' M_t\right)^{-\frac{1}{\gamma}}$, which coincides with c_t^{opt} in (3.7). Similarly, letting $\alpha = \beta = 0$, the approximation to the optimal dual control in (4.6) becomes $\psi'_t = \eta' M_t c'_t = \eta' M_t \left[e^{\delta t}\eta' M_t\right]^{-\frac{1}{\gamma}}$, which coincides with (3.6). Observe that we have not distinguished between the primal and dual η', as these are equal.

4.3 Numerical Results

We evaluate the accuracy of the approximation shown in equation (4.5), using the evaluation mechanism of section 4.1. To this end, we set $N = 1$ in the market model, $\mathcal{M}$, and fix $r_t = r$, $\sigma_t = \sigma$, $\mu_t = \mu$, where r, σ and μ are constants. Based upon the parameter initialisation in van Bilsen et al. (2020a), we define: $X_0 = 10$, $T = 10$, $\gamma = 10$, $\delta = 0.03$, $\alpha = \beta = 0.1$, $\mu = 0.05$, $r = 0.01$ and $\sigma = 0.2$. In Table 1, we present the upper bounds on the welfare losses ($\mathcal{C}$) associated with the approximation, for different values of γ, X_0, $\alpha = \beta$ and δ. We compute the welfare losses from the equality displayed in (4.2), in which the value functions, J and $\mathcal{V}$, follow directly from the primal and dual approximations in (4.5) and (4.6), respectively. In addition to this, the table displays the "root" sum of squares (RSS) corresponding to the approximations. The RSS is calculated as follows: $\text{RSS} = \mathbb{E}\big[\sum_{i=1}^{M}(c'_{t_i} - c''_{t_i})^2\big]^{\frac{1}{2}}$, where c'_{t_i} is approximate consumption for (4.5), and c''_{t_i} is approximate consumption implied by ψ'_t in (4.6) via the second duality relation in (4.3). Here, $t_1 < t_2 < \ldots < t_M$ represent M time-steps in the Euler scheme. These results ensue from $10,000$ Monte-Carlo simulations and an Euler scheme with 20 equidistant time-steps.

Table 1 shows that the maximal welfare losses generated by the approximation in (4.5) vary between 0.000% and 0.167% of X_0. Bearing in mind that these numbers constitute upper bounds on the true errors, we can conclude that the approximation is near-optimal. This finding coincides with the results reported in van Bilsen et al. (2020a) and is confirmed by the magnitude of the values reported for RSS. In view of the fact that $c_t' \to c_t^{\text{opt}}$ as $\alpha = \beta \to 0$, it is clear why the table displays a positive relation between $\alpha = \beta$ and the magnitude of both $\mathcal{C}$ and the RSS. As for the same relations involving γ and δ, we can observe from (4.5) that increases in γ and δ result in $\widehat{c}_t$ attaining values closer to 1. That is, due to increases in γ and δ, the habit level tracks consumption more closely. Consequently, as shown in Table 1, approximate consumption in (4.5) becomes more accurate for increases in γ and δ, as c_t' is based on a Taylor expansion around $\{\widehat{c}_t\}_{t\in[0,T]} = 1$. To explain the relation between X_0 and both $\mathcal{C}$ and RSS, we note that $X_0 = \frac{1}{r}\left(1 - e^{-rT}\right) \approx T$ roughly implies that $\widehat{c}_t \approx 1$. The latter is a consequence of the budget constraint, $\mathbb{E}\big[\int_0^T M_t \widehat{c}_t h_t \mathrm{d}t\big] = X_0$. Hence, the performance of c_t' should decrease for values of X_0 and T that deviate from $X_0 = T$. This is shown in Table 1.[5]

A Proof of Theorem 3.1

To prove Theorem 3.1, we make use of Fenchel Duality as formalised in Theorem 4.3.3 of the textbook by Borwein and Zhu (2004). As this theorem lies at the heart of our proof, we provide its statement in the following proposition.

Proposition A.1. *Let $f : X \to \mathbb{R} \cup \{\infty\}$ and $g : Y \to \mathbb{R} \cup \{\infty\}$ be two continuous and convex functions. Additionally, introduce the bounded linear map $A : X \to Y$. Here, X and Y outline two Banach spaces. Then, the Fenchel problems are given by:*

$$\begin{aligned} p^* &= \inf_{x\in X} \{f(x) + g(Ax)\} \\ d^* &= \sup_{y^*\in Y} \{-f^*(A^* y^*) - g^*(-y^*)\}, \end{aligned} \tag{A.1}$$

and satisfy weak duality, $d^* \le p^*$. Here, f^* and g^* represent the convex conjugates of f and g, respectively, i.e. $f^*(x) = \sup_{z\in X}\{\langle x, z\rangle - f(z)\}$ and $g^*(y) = \sup_{z\in Y}\{\langle y, z\rangle - g(z)\}$, for all $x \in X^*$ and $y \in Y^*$. Note that X^* and Y^* are the dual spaces of X and Y, respectively. Moreover, A^* is the adjoint of A. Strong duality, i.e. $p^* = d^*$, holds if either of the following conditions is fulfilled:

(i) $0 \in \operatorname{core}(\operatorname{dom} g - A \operatorname{dom} f)$ *and f and g are both lower semi-continuous. Here,* core *stands for the algebraic interior, and* dom h *is given by* $\operatorname{dom} h = \{z \mid h(z) < \infty\}$

[5] We note that the running time for calculating $\mathcal{C}$ and RSS in Table 1 is effectively equal to zero. The mere required computational effort stems from the simulations. This is a significant advantage, as most evaluation procedures employ computationally demanding grid-search routines.

for any function h;

(ii) $A\, \mathrm{dom}\, f \cap \mathrm{cont}\, g \neq \phi$, *where* cont *are the points where the function is continuous.*

Moreover, if $|d^*| < \infty$ *holds, then the supremum in* (A.1) *is attained.*

Proof. See page 136 of Borwein and Zhu (2004). □

We continue by aligning the notation of our primal and dual problems with the notation of Proposition A.1. To this end, we start by deriving an alternative representation of (2.7). This alternative representation is based on the static formulation of optimal investment-consumption problems, due to Pliska (1986), Karatzas et al. (1987), and Cox and Huang (1989, 1991). We provide this formulation in the subsequent lemma.

Lemma A.2. *Define the following function:*

$$\mathcal{J}\left(X_0, -\log c_t, \eta\right) = \mathbb{E}\left[\int_0^T e^{-\delta t}\frac{e^{[1-\gamma](\log c_t - \log h_t)}}{1-\gamma})\mathrm{d}t\right] - \eta\mathbb{E}\left[\int_0^T e^{\log c_t} M_t \mathrm{d}t\right] + \eta X_0. \tag{A.2}$$

Then, for all $X_0 \in \mathbb{R}_+$*, the following optimisation problems are identical:*

$$\sup_{\{c_t,\pi_t\}_{t\in[0,T]}\in\mathcal{A}_{X_0}} J\left(X_0, \{c_t, \pi_t\}\right) = \inf_{\eta\in\mathbb{R}_+} \sup_{-\log c_t \in L^2(\Omega\times[0,T])} \mathcal{J}\left(X_0, -\log c_t, \eta\right). \tag{A.3}$$

Proof. By arguments similar to those that yield Lemma 2.2 in Cox and Huang (1989) and Proposition 7.3 in Cvitanić and Karatzas (1992), we know that $\{c_t, \pi_t\}_{t\in[0,T]} \in \mathcal{A}_{X_0}$ if and only if $\{c_t\}_{t\in[0,T]}$ satisfies $\mathbb{E}\left[\int_0^T c_t M_t \mathrm{d}t\right] \leq X_0$ and $\log c_t \in L^2\left(\Omega\times[0,T]\right)$. Therefore, maximisation of $J\left(X_0, \{c_t,\pi_t\}\right)$ over $\{c_t,\pi_t\}_{t\in[0,T]} \in \mathcal{A}_{X_0}$ is the same as maximisation of $J\left(X_0, \{c_t,\pi_t\}\right)$ over all $\log c_t \in L^2\left(\Omega\times[0,T]\right)$ such that $\mathbb{E}\left[\int_0^T c_t M_t \mathrm{d}t\right] \leq X_0$ holds. As a result, we are able to derive the following set of equations:

$$\begin{aligned}
&\sup_{\{c_t,\pi_t\}_{t\in[0,T]}\in\mathcal{A}_{X_0}} \mathbb{E}\left[\int_0^T e^{-\delta t} U\left(\frac{c_t}{h_t}\right)\mathrm{d}t\right] \\
&= \sup_{\log c_t\in L^2(\Omega\times[0,T])\ \text{s.t.}\ \mathbb{E}\left[\int_0^T c_t M_t \mathrm{d}t\right]\leq X_0} \mathbb{E}\left[\int_0^T e^{-\delta t} U\left(\frac{c_t}{h_t}\right)\mathrm{d}t\right] \\
&= \inf_{\eta\in\mathbb{R}_+}\left(\sup_{-\log c_t\in L^2(\Omega\times[0,T])}\left\{\mathbb{E}\left[\int_0^T e^{-\delta t} U\left(\frac{c_t}{h_t}\right)\mathrm{d}t\right] - \eta\mathbb{E}\left[\int_0^T c_t M_t \mathrm{d}t\right] + \eta X_0\right\}\right).
\end{aligned} \tag{A.4}$$

The last equality is a result of the following ingredients. First, we know that $c_t = X_0\epsilon\big(\mathbb{E}\big[\int_0^T M_t \mathrm{d}t\big]\big)^{-1}$ for $\epsilon \in (0,1)$ is a strictly feasible solution to the static formulation of the consumption problem. Second, we have that $h_t > 0$ and $c_t > 0$. Hence, $c_t = e^{\log c_t}$ and $h_t = e^{\log h_t}$. Using this, we derive that $\mathbb{E}\big[\int_0^T e^{-\delta t} U\big(\frac{c_t}{h_t}\big)\mathrm{d}t\big]$ is strictly concave in

$-\log c_t \in L^2(\Omega \times [0,T])$. Similarly, we have that $\eta\mathbb{E}\big[\int_0^T c_t M_t \mathrm{d}t\big]$ is strictly convex in $-\log c_t \in L^2(\Omega \times [0,T])$. Third, by concavity of U, and the fact that $\log c_t, \log h_t \in L^2(\Omega \times [0,T])$, it holds that $\mathbb{E}\big[\int_0^T e^{-\delta t} U\big(\frac{c_t}{h_t}\big)\mathrm{d}t\big] < \infty$. These properties validate the last equality, cf. Theorem 1 on page 217 of Luenberger (1997). The step from (A.4) to (A.3) is trivial using the definition of U, and that $c_t = e^{\log c_t}$ and $h_t = e^{\log h_t}$. □

To align our notation with the one of Proposition A.1, we should have:

$$d^* = \sup_{-\log c_t \in L^2(\Omega \times [0,T])} \mathcal{J}(X_0, -\log c_t, \eta). \tag{A.5}$$

Accordingly, in the nomenclature of the aforementioned proposition, we have that $y^* = -\log c_t$ and $Y = L^2(\Omega \times [0,T])$, which is a Banach space. Moreover, in terms of the functions f^* and g^*, and the mapping A, we must have the following:

$$\begin{aligned} -f^*(A^* y^*) &= \mathbb{E}\left[\int_0^T e^{-\delta t} \frac{e^{-[1-\gamma]A^*(-\log c_t)}}{1-\gamma} \mathrm{d}t\right] \\ -g^*(-y^*) &= -\eta \mathbb{E}\left[\int_0^T e^{\log c_t} M_t \mathrm{d}t\right] + \eta X_0, \end{aligned} \tag{A.6}$$

where the linear map A^* is given by:

$$A^*(-\log c_t) = -\log c_t + \beta \int_0^t e^{-\alpha(t-s)} \log c_s \mathrm{d}s. \tag{A.7}$$

Clearly, $A^* : L^2(\Omega \times [0,T]) \to L^2(\Omega \times [0,T])$. Therefore, by adjointess arguments, we must have that $A : L^2(\Omega \times [0,T]) \to L^2(\Omega \times [0,T])$, too.

According to the equations for f^* and g^* in (A.6), we ought to have:

$$\begin{aligned} f^*(x) &= -\mathbb{E}\left[\int_0^T e^{-\delta t} \frac{e^{-[1-\gamma]x_t}}{1-\gamma} \mathrm{d}t\right] \\ g^*(x) &= \eta \mathbb{E}\left[\int_0^T e^{x_t} M_t \mathrm{d}t\right] - \eta X_0. \end{aligned} \tag{A.8}$$

To obtain the definitions of f and g in (A.1), we recall that f^* and g^* are defined as the convex conjugates of f and g, according to: $f^*(x) = \sup_{z \in X} \{\langle x, z\rangle - f(z)\}$ and $g^*(y) = \sup_{z \in Y} \{\langle y, z\rangle - g(z)\}$, for all $x \in X^*$ and $y \in Y^*$. It is easy to show that the following definitions of f and g manage to satisfy the latter relations:

$$\begin{aligned} f(x) &= \mathbb{E}\left[\int_0^T e^{-\delta t} \frac{1}{1-\gamma} V\left(e^{\delta t} x_t\right) \mathrm{d}t\right] \\ g(x) &= -\mathbb{E}\left[\int_0^T \eta M_t V\left(\frac{x_t}{\eta M_t}\right) \mathrm{d}t\right] + \eta X_0. \end{aligned} \tag{A.9}$$

We observe that $X = Y = L^2(\Omega \times [0,T])$, and that the preceding definitions of

$f : X \to \mathbb{R} \cup \{\infty\}$ and $g : Y \to \mathbb{R} \cup \{\infty\}$ constitute two continuous and convex functions. Furthermore, we find that A, i.e. the adjoint of A^*, is given by the following linear mapping:

$$Ax_t = x_t - \beta \mathbb{E}\left[\int_t^T e^{-\alpha(s-t)} x_s \mathrm{d}s \,\middle|\, \mathcal{F}_t \right]. \tag{A.10}$$

Note here that:

$$\begin{aligned} \|Ax_t\|_{L^2(\Omega\times[0,T])} &\leq \|x_t\|_{L^2(\Omega\times[0,T])} + \beta \left\| \mathbb{E}\left[\int_t^T e^{-\alpha(s-t)} x_s \mathrm{d}s \,\middle|\, \mathcal{F}_t \right] \right\|_{L^2(\Omega\times[0,T])} \\ &\leq \|x_t\|_{L^2(\Omega\times[0,T])} + \beta \mathbb{E}\left[\int_0^T \mathbb{E}\left[\int_t^T e^{-2\alpha(s-t)} x_s^2 \mathrm{d}s \,\middle|\, \mathcal{F}_t \right] \mathrm{d}t \right]^{\frac{1}{2}} \\ &= \|x_t\|_{L^2(\Omega\times[0,T])} + \frac{1}{2}\frac{\beta}{\alpha} \left\| x_t \left(1 - e^{-2\alpha t}\right)^{\frac{1}{2}} \right\|_{L^2(\Omega\times[0,T])} \leq \frac{3}{2} \|x_t\|_{L^2(\Omega\times[0,T])}. \end{aligned} \tag{A.11}$$

The first inequality is due to the triangle inequality; the second inequality is a result of Hölder's inequality; the final inequality is trivial ($1 - e^{-2\alpha t} < 1$ for all $t \in [0, T]$). As a consequence of (A.11), we know that $A : X \to Y$ is a bounded linear map.

Considering Proposition A.1, we note that A dom $f \cap$ cont $g = \left(L^2\left(\Omega \times [0, T]\right) \cap \mathbb{R}\right) \cap \left(L^2\left(\Omega \times [0, T]\right) \cap \mathbb{R}_+\right) \neq \phi$. Hence, by Proposition A.1, we have strong duality, which finalises – via Lemma A.2 – the proof of Theorem 3.1:

$$d^* = \sup_{-\log c_t \in L^2(\Omega\times[0,T])} \mathcal{J}\left(X_0, -\log c_t, \eta\right) = p^* = \inf_{\psi_t \in L^2(\Omega\times[0,T])} \mathcal{V}\left(X_0, \psi_t, \eta\right). \tag{A.12}$$

References

Abel, A. B. (1990). Asset prices under habit formation and catching up with the Joneses. *The American Economic Review*, 80(2):38–42.

Bick, B., Kraft, H., and Munk, C. (2013). Solving constrained consumption–investment problems by simulation of artificial market strategies. *Management Science*, 59(2):485–503.

Bodie, Z., Detemple, J. B., Otruba, S., and Walter, S. (2004). Optimal consumption–portfolio choices and retirement planning. *Journal of Economic Dynamics and Control*, 28(6):1115–1148.

Borwein, J. M. and Zhu, Q. J. (2004). Techniques of variational analysis.

Campbell, J. Y. and Cochrane, J. H. (1999). By force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of Political Economy*, 107(2):205–251.

Carroll, C. D. (2000). Solving consumption models with multiplicative habits. *Economics Letters*, 68(1):67–77.

Carroll, C. D., Overland, J., and Weil, D. N. (1997). Comparison utility in a growth model. *Journal of Economic Growth*, 2(4):339–367.

Carroll, C. D., Overland, J., and Weil, D. N. (2000). Saving and growth with habit formation. *American Economic Review*, 90(3):341–355.

Chan, Y. L. and Kogan, L. (2002). Catching up with the Joneses: Heterogeneous preferences and the dynamics of asset prices. *Journal of Political Economy*, 110(6):1255–1285.

Chen, A., Hieber, P., and Nguyen, T. (2019). Constrained non-concave utility maximization: An application to life insurance contracts with guarantees. *European Journal of Operational Research*, 273(3):1119–1135.

Constantinides, G. M. (1990). Habit formation: A resolution of the equity premium puzzle. *Journal of Political Economy*, 98(3):519–543.

Corrado, L. and Holly, S. (2011). Multiplicative habit formation and consumption: A note. *Economics Letters*, 113(2):116–119.

Cox, J. C. and Huang, C.-F. (1989). Optimal consumption and portfolio policies when asset prices follow a diffusion process. *Journal of Economic Theory*, 49(1):33–83.

Cox, J. C. and Huang, C.-F. (1991). A variational problem arising in financial economics. *Journal of Mathematical Economics*, 20(5):465–487.

Cuoco, D. (1997). Optimal consumption and equilibrium prices with portfolio constraints and stochastic income. *Journal of Economic Theory*, 72(1):33–73.

Cvitanić, J. and Karatzas, I. (1992). Convex duality in constrained portfolio optimization. *Annals of Applied Probability*, 2(4):767–818.

Delbaen, F. and Schachermayer, W. (1994). A general version of the fundamental theorem of asset pricing. *Mathematische Annalen*, 300(3):463–520.

Detemple, J. B. and Karatzas, I. (2003). Non-addictive habits: optimal consumption-portfolio policies. *Journal of Economic Theory*, 113(2):265–285.

Detemple, J. B. and Rindisbacher, M. (2010). Dynamic asset allocation: Portfolio decomposition formula and applications. *The Review of Financial Studies*, 23(1):25–100.

Detemple, J. B. and Zapatero, F. (1991). Asset prices in an exchange economy with habit formation. *Econometrica: Journal of the Econometric Society*, pages 1633–1657.

Fuhrer, J. C. (2000). Habit formation in consumption and its implications for monetary-policy models. *American Economic Review*, 90(3):367–390.

Gomes, F. and Michaelides, A. (2003). Portfolio choice with internal habit formation: A life-cycle model with uninsurable labor income risk. *Review of Economic Dynamics*, 6(4):729–766.

Gómez, J.-p., Priestley, R., and Zapatero, F. (2009). Implications of keeping-up-with-the-Joneses behavior for the equilibrium cross section of stock returns: International evidence. *The Journal of Finance*, 64(6):2703–2737.

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–292.

Kamma, T. and Pelsser, A. (2022). Near-optimal asset allocation in financial markets with trading constraints. *European Journal of Operational Research*, 297(2):766–781.

Karatzas, I., Lehoczky, J. P., and Shreve, S. E. (1987). Optimal portfolio and consumption decisions for a small investor on a finite horizon. *SIAM Journal on Control and Optimization*, 25(6):1557–1586.

Karatzas, I. and Shreve, S. E. (1991). *Brownian Motion and Stochastic Calculus.* New York: Springer-Verlag.

Klein, I. and Rogers, L. C. (2007). Duality in optimal investment and consumption problems with market frictions. *Mathematical Finance*, 17(2):225–247.

Kozicki, S. and Tinsley, P. A. (2002). Dynamic specifications in optimizing trend-deviation macro models. *Journal of Economic Dynamics and Control*, 26(9-10):1585–1611.

Li, T., Wang, S., and Yang, J. (2021). Robust consumption and portfolio choices with habit formation. *Economic Modelling*, 98:227–246.

Luenberger, D. G. (1997). *Optimization by Vector Space Methods.* John Wiley & Sons.

Merton, R. C. (1971). Optimum consumption and portfolio rules in a continuous-time model. *Journal of Economic Theory*, 3(4):373–413.

Munk, C. (2008). Portfolio and consumption choice with stochastic investment opportunities and habit formation in preferences. *Journal of Economic Dynamics and Control*, 32(11):3560–3589.

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

Pliska, S. R. (1986). A stochastic calculus model of continuous trading: optimal portfolios. *Mathematics of Operations Research*, 11(2):371–382.

Polkovnichenko, V. (2007). Life-cycle portfolio choice with additive habit formation preferences and uninsurable labor income risk. *The Review of Financial Studies*, 20(1):83–124.

Rockafellar, R. T. (2015). *Convex Analysis*. Princeton university press.

Rogers, L. C. (2003). Duality in constrained optimal investment and consumption problems: a synthesis. In *Paris-Princeton Lectures on Mathematical Finance 2002*. Springer.

Rogers, L. C. (2013). *Optimal Investment*. Springer.

Schroder, M. and Skiadas, C. (2002). An isomorphism between asset pricing models with and without linear habit formation. *The Review of Financial Studies*, 15(4):1189–1221.

van Bilsen, S., Bovenberg, A. L., and Laeven, R. J. (2020a). Consumption and portfolio choice under internal multiplicative habit formation. *Journal of Financial and Quantitative Analysis*, 55(7):2334–2371.

van Bilsen, S., Laeven, R. J., and Nijman, T. E. (2020b). Consumption and portfolio choice under loss aversion and endogenous updating of the reference level. *Management Science*, 66(9):3927–3955.

Yogo, M. (2008). Asset prices under habit formation and reference-dependent preferences. *Journal of Business & Economic Statistics*, 26(2):131–143.

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