SIAM J. FINANCIAL MATH.
Vol. 1, pp. 366–395

Merton Problem with Taxes: Characterization, Computation, and Approximation*

Imen Ben Tahar†, H. Mete Soner‡, and Nizar Touzi§

Abstract. We formulate a computationally tractable extension of the classical Merton optimal consumption-investment problem to include the capital gains taxes. This is the continuous-time version of the model introduced by Dammon, Spatt, and Zhang [*Rev. Financ. Stud.*, 14 (2001), pp. 583–616]. In this model the tax basis is computed as the average cost of the stocks in the investor's portfolio. This average rule introduces only one additional state variable, namely the tax basis. Since the other tax rules such as the first in first out rule require the knowledge of all past transactions, the average model is computationally much easier. We emphasize the linear taxation rule, which allows for tax credits when capital gains losses are experienced. In this context wash sales are optimal, and we prove it rigorously. Our main contributions are a first order explicit approximation of the value function of the problem and a unique characterization by means of the corresponding dynamic programming equation. The latter characterization builds on technical results isolated in the accompanying paper [I. Ben Tahar, H. M. Soner, and N. Touzi, *SIAM J. Control Optim.*, 46 (2007), pp. 1779–1801]. We also suggest a numerical computation technique based on a combination of finite differences and the Howard iteration algorithm. Finally, we provide some numerical results on the welfare consequences of taxes and the quality of the first order approximation.

Key words. optimal consumption and investment in continuous time, transaction costs, capital gains taxes, finite differences

AMS subject classifications. 91B28, 35B37, 35K20

DOI. 10.1137/080742178

1. Introduction.

Since the seminal papers of Merton [26, 27], there has been extensive literature on the problem of optimal consumption and investment decision in financial markets subject to imperfections. We refer the reader to Cox and Huang [10] and Karatzas, Lehoczky, and Shreve [21] for the case of incomplete markets, Cvitanić and Karatzas [11] for the case of portfolio constraints, and Constantinides and Magill [9], Davis and Norman [13], Shreve and Soner [28], Barles and Soner [3], and Duffie and Sun [15] for the case of transaction costs.

However, the problem of taxes on capital gains received limited attention, although taxes represent a much higher percentage than transaction costs in real securities markets. Compared to ordinary income, capital gains are taxed only when the investor sells the security, allowing for a deferral option. One may think that the taxes on capital gains have an appreciable impact on an individual's consumption and investment decisions. Indeed, under taxation

*Received by the editors December 1, 2008; accepted for publication (in revised form) March 8, 2010; published electronically May 26, 2010. This research was part of the Chair *Financial Risks* of the *Risk Foundation* sponsored by Société Générale, the Chair *Derivatives of the Future* sponsored by the Fédération Bancaire Française, and the Chair *Finance and Sustainable Development* sponsored by EDF and Calyon.

http://www.siam.org/journals/sifin/1/74217.html

†CEREMADE, Université Paris Dauphine, 75775 Cedex 16, Paris, France (imen@ceremade.dauphine.fr).

‡Sabanci University, Istanbul 34956, Turkey (hmsoner@ethz.ch). This author's research was partly supported by the King Saud University D.S.F.P and by the Turkish Technical and Scientific Council, TÜBITAK.

§CMAP, Ecole Polytechnique, 91128 Palaiseau Cedex, Paris, France (nizar.touzi@polytechnique.edu).

of capital gains, the portfolio rebalancing implies additional charges, therefore altering the available wealth for future consumption. This possibly induces a depreciation of consumption opportunities compared to a tax-free market. On the other hand, since taxes are paid only when embedded capital gains are actually realized, the investor may choose to defer the realization of capital gains and liquidate his/her position in case of a capital loss, particularly when the tax code allows for tax credits.

The first relevant work in the previous literature is due to Constantinides [8], who shows that the investment and consumption decisions are separable and that the optimal strategy consists in realizing losses and deferring gains. These results rely heavily on the possibility of short selling the risky asset. Since capital gains realizations are observed in real securities markets, the subsequent literature considers the problem under the no-short-sales constraint.

In a multiperiod context, many challenging difficulties appear because of the path dependency of the problem. The taxation code specifies the basis to which the price of a security has to be compared in order to evaluate the capital gains (or losses). The tax basis is defined as either (i) the specific purchase price of the asset to be sold, (ii) the purchase price of a freely chosen share held in the portfolio (of course the number of chosen shares must be more than the ones to be sold), or (iii) the weighted average of past purchase prices. In some countries, investors can choose any one of the above definitions of the tax basis.

A deterministic model with the above definition (i) of the tax basis, together with the *first in first out* priority rule for the stock to be sold, is introduced by Jouini, Koehl, and Touzi [20, 19]. An existence result is proved, and the first order conditions of optimality are derived under some conditions. However, the numerical complexity due to the path dependency of the problem is not solved in the context of this model.

A financial model with the above definition (ii) of the taxation rule was considered by Dybvig and Koo [16] in the context of a four-period binomial model. Some numerical progress was achieved later by DeMiguel and Uppal [14], who were able to consider more periods in the binomial model and/or more stocks. This numerical progress is limited, as these authors were not able to go beyond 10 periods in the single-asset framework.

The taxation rule (iii), where the tax basis is the weighted average of past purchase prices, was first considered by Dammon, Spatt, and Zhang [12] in the context of a binomial model with short-sales constraints and the linear taxation rule. The average tax basis is actually used in Canada. Dammon, Spatt, and Zhang [12] considered the problem of maximizing the expected discounted utility from future consumption and provided a numerical analysis of this model based on the dynamic programming principle. The important technical feature of this model is that the path dependency of the problem is seriously reduced, as the dynamics of the tax basis is Markov. This implies a significant advantage of this model in comparison to [16]. This advantage was further justified by DeMiguel and Uppal [14], who provided numerical evidence that the certainty equivalent loss from using the average tax basis (iii) instead of the exact tax basis (ii) is typically less than 1% for a large choice of parameter values. The analysis of [12] was further extended to the multiasset framework by Gallmeyer, Kaniel, and Tompaidis [18].

In this paper, we formulate a continuous-time version of the Dammon–Spatt–Zhang utility maximization problem under capital gains taxes. Our model is similar to that of Leland [24], who instead considered the problem of minimizing the tracking error to some benchmark index.

The financial market consists of a tax-free riskless asset and a risky one. The holdings in the risky asset are subject to the no-short-sales constraint, and the total wealth is restricted by the no-bankruptcy condition. The risky asset is subject to taxes on capital gains. As in [12], the tax basis is defined as the weighted average of past purchase prices, and the taxation rule is linear, thus allowing for tax credits. However, we differ from [12] by considering an infinite horizon problem, as our main goal is to provide analytical tools for this class of problems. In particular, our model does not allow for tax forgiveness at death. Clearly, one should keep this difference in mind when interpreting results for investors with a short horizon.

The investor preferences are described by a power utility with a constant relative risk aversion factor. This assumption is needed only to reduce the computational complexity of the problem. However, with this reduction explicit descriptions of the value function and the optimal strategy are still not available.

This model enables us to *rigorously* prove several interesting properties observed in practice. Although these results are sometimes intuitively clear, their proofs require careful analysis and the use of the tractability of the model. The first of these results is the optimality of the *wash sales*. Namely, in Proposition 3.5 we prove that it is always optimal to realize capital losses whenever the tax basis exceeds the spot price. This property is observed in practice and is stated and embedded directly in the definition of the tax basis in [12]. We also prove the continuity of the value function (and even Lipschitz continuity, up to a change of variables). We recall that, in the tax-free models of [26, 9, 13, 28], this property follows from the obvious concavity of the value function. Under capital gains taxes, the concavity argument fails, and the numerical results of section 6 suggest that the value function is indeed not concave!

The first main result of this paper is to provide an explicit approximation of the value function which follows from an upper and a lower bound proved in section 4. In view of the absence of closed form solutions, such an approximation is useful for understanding the model better. Although this explicit approximation holds for small interest rate and tax parameters, our numerical experiments indicate that this approximation is satisfactory with realistic values of interest rate and tax parameters, as it leads to a relative error within 10%. These findings are reported in section 6. This first order approximation allows one to draw the following observations:

- The lower bound is derived as the limit of the value implied by a sequence of strategies which mimics the Merton optimal strategy in a Merton-type fictitious frictionless financial market with tax-deflated drift and volatility coefficients. The risk premium of this fictitious financial market is smaller than that of the original market. So, even if the optimal strategy in our problem is not available in explicit form, our first order expansion is accompanied by an explicit strategy which achieves "the first order maximal utility value."
- In a situation of a capital loss, our first order approximation is increasing in the tax rate. For small interest rate and tax parameters, the advantage taken from an initial tax credit is never compensated by the increase of tax over the lifetime horizon. This is in agreement with Cadenillas and Pliska [7], who found that "sometimes investors are better off with a positive tax rate."
- Finally, the investment component of this approximation sequence exhibits a smaller exposition to the risky asset. This is in line with the risk premium puzzle highlighted

by Mehra and Prescott [25]. However, one should note that this model is only a partial equilibrium model and that the level of the equity premium is determined by general equilibrium considerations.

Our analysis of the optimal consumption-investment problem relies on a numerical approach based on dynamic programming and partial differential equations. Therefore, the second main result of this paper is a characterization of the value function as the limit (uniformly on compact subsets) of an approximating sequence defined by a slight perturbation of the "natural" dynamic programming equation of our problem. The financial interpretation of our perturbation is, on the one hand, to introduce a small transaction cost parameter and, on the other hand, to modify simultaneously the taxation rule when the tax basis approaches the critical point zero. Our analysis relies on the technical results in our accompanying paper [5], which shows that the perturbed dynamic programming equation has a unique continuous viscosity solution within the class of polynomially growing functions.

Finally, based on our dynamic programming characterization, we suggest a numerical approximation method combining finite differences with the Howard iterations. Unfortunately, we have no theoretical convergence result for our algorithm. Indeed, establishing such convergence results for Hamilton–Jacobi–Bellman equations corresponding to singular control problems is an open question in numerical analysis, and the existing results, based on the monotone scheme method of Barles and Souganidis [4], are restricted to the bounded control context; see Bonnans and Zidani [6], Krylov [22, 23], Barles and Jakobsen [2], and Fahim, Touzi, and Warin [17]. This difficulty was already observed in the related literature on transaction costs; see Akian, Menaldi, and Sulem [1] and Tourin and Zariphopoulou [31]. Following the latter papers, we therefore concentrate our effort on realizing the empirical convergence of the algorithm. The numerical scheme is implemented to obtain the qualitative behavior of the solution and to understand the welfare consequences of the taxation. In particular, the numerical approximation of the optimal strategy displays a bang-bang behavior, as expected in our singular control problem. As in the transaction cost context of [9, 13, 28], the state space is partitioned into three regions: the *no-transaction region* NT, the *buy region* B, and the *sell region* S; but in contrast with the transaction cost framework these regions are not cones.

Notation. For a domain $\mathbf{D}$ in $\mathbb{R}^n$, we denote by $\mathrm{USC}(\mathbf{D})$ (resp., $\mathrm{LSC}(\mathbf{D})$) the collection of all upper semicontinuous (resp., lower semicontinuous) functions from $\mathbf{D}$ to $\mathbb{R}$. The set of continuous functions from $\mathbf{D}$ to $\mathbb{R}$ is denoted by $\mathrm{C}^0(\mathbf{D}) := \mathrm{USC}(\mathbf{D}) \cap \mathrm{LSC}(\mathbf{D})$. For a parameter $\delta > 0$, we say that a function $f : \mathbf{D} \longrightarrow \mathbb{R}$ has δ-polynomial growth if

$$\sup_{x \in \mathbf{D}} \frac{|f(x)|}{1 + |x|^\delta} < \infty.$$

We finally denote $\mathrm{USC}_\delta(\mathbf{D}) := \{f \in \mathrm{USC}(\mathbf{D}) : f \text{ has } \delta\text{-polynomial growth}\}$. The sets $\mathrm{LSC}_\delta(\mathbf{D})$ and $\mathrm{C}^0_\delta(\mathbf{D})$ are defined similarly.

2. Consumption-investment models with capital gains taxes.

2.1. The financial assets. Throughout this paper, we consider a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$, endowed with a standard scalar Brownian motion $W = \{W_t, 0 \le t\}$, and we denote by $\mathbb{F}$ the $\mathbb{P}$-completion of the natural filtration of the Brownian motion. We consider a financial market consisting of one bank account with constant interest rate $r > 0$ and one risky asset whose price process evolves according to the Black–Scholes model:

$$dP_t \;=\; P_t\left[(r+\theta\sigma)dt \;+\; \sigma\, dW_t\right], \tag{2.1}$$

where $\theta > 0$ is a constant risk premium, and $\sigma > 0$ is a constant volatility parameter. The positivity restriction on the risk premium coefficient ensures that positive investment in the risky asset is interesting.

2.2. Taxation rule on capital gains. The sales of the stock are subject to *taxes on capital gains*. The amount of tax to be paid for each sale of risky asset, at time t, is computed by comparison of the current price P_t to an index B_t defined as the weighted average price of the shares purchased by the investor up to time t. When $P_t \geq B_t$, i.e., the current price of the risky asset is greater than the weighted average price, the investor would *realize a capital gain* by selling the risky asset. Similarly, when $P_t \leq B_t$, the sale of the risky asset corresponds to the *realization of a capital loss*.

In order to better explain the definition of the tax basis B, we provide the following example derived from the official Canadian tax code.

Table 1 reports transactions performed by an individual on shares of STU Ltd and how the tax basis of the individual changes over time.

Table 1
Extracted from Capital Gains 2007*; http://www.cra.gc.ca.*

Transaction	Price P (dollars)	Number of shares (unitless)	Portfolio composition (unitless)	Tax basis B (dollars)
Purchase at t_1	15.00	100	100 : \$15.00/share	**15.00**
Purchase at t_2	20.00	150	100 : \$15.00/share	
			150 : \$20.00/share	**18.00**
Sale at t_3	-	200	20 : \$15.00/share	
		$=\frac{4}{5}(100+150)$	30 : \$20.00/share	**18.00**
Purchase at t_4	21.00	350	20 : \$15.00/share	
			30 : \$20.00/share	
			350 : \$21.00/share	**20.625**

Just after a sale transaction, the tax basis is not changed. However, sales do alter the tax basis starting from the date of the next purchase. Notice, however, that the tax basis is affected only by the number of shares sold and not by the sale price.

The sale of a unit share of stock at some time t is subject to the payment of an amount of tax computed according to the tax basis of the portfolio at time t. In this paper, we consider a linear taxation rule, i.e., this amount of tax is given by

$$\ell(P_t - B_t) := \alpha\,(P_t - B_t)\,, \tag{2.2}$$

where $\alpha \in [0, 1)$ is a constant tax rate coefficient. Our interest is of course in the case $\alpha > 0$. When the tax basis is smaller than the spot price, the investor realizes a capital gain. Then, by selling one unit of risky asset at the spot price P_t, the amount of tax to be paid is $\alpha(P_t - B_t)$.

When the tax basis is larger than the spot price, the investor receives the tax credit $\alpha(B_t - P_t)$ for each unit of asset sold at time t.

Remark 2.1. In practice, the realized capital losses are deduced from the total amount of taxes that the investor has to pay, and the annual deductible capital losses amount may be limited by the tax code. In our model, we follow Dammon, Spatt, and Zhang [12] by adopting the simplifying assumption that capital losses are credited immediately without any limit.

Remark 2.2. Our definition of the tax basis B is slightly different from that of Dammon, Spatt, and Zhang [12], who set the tax basis to be equal to the spot price whenever the average purchase price exceeds the current price. This does not affect the results, as Proposition 3.5 shows that wash sales are optimal.

2.3. Consumption-investment strategies. We denote by X_t the (cash) position on the bank, Y_t the amount invested in the risky assets, and

$$K_t \;:=\; B_t \frac{Y_t}{P_t}, \quad t \geq 0, \tag{2.3}$$

the position on the risky asset account evaluated at the basis price. The trading in risky assets is subject the no-short-sales constraint

$$Y_t \;\geq\; 0 \quad \mathbb{P}\text{-a.s.} \quad \text{for all} \quad t \geq 0, \tag{2.4}$$

and the position of the investor is required to satisfy the solvency condition

$$Z_t := X_t + Y_t - \ell\,(P_t - B_t)\,\frac{Y_t}{P_t} \;=\; X_t + (1-\alpha)Y_t + \alpha K_t \;\geq\; 0 \quad \mathbb{P}\text{-a.s.}; \tag{2.5}$$

i.e., the after-tax liquidation value of the portfolio is nonnegative at any point in time.

Trading on the financial market is described by means of the transfers between the two investment opportunities defined by two $\mathbb{F}$-adapted, right-continuous, and nondecreasing processes $L = \{L_t,\, t \geq 0\}$ and $M = \{M_t,\, t \geq 0\}$ with $L_{0^-} = M_{0^-} = 0$. The amount transferred from the bank to the nonrisky asset account at time t is given by dL_t and corresponds to a purchase of risky asset. The amount transferred from the risky asset account to the bank at time t, corresponding to a sale of risky asset, is given by $Y_{t-}dM_t$ and is expressed in terms of proportions of the total holdings in risky asset as in the example of Table 1.

To force the short-sales constraint (2.4) to hold, we restrict the jumps of M by

$$\Delta M_t \;\leq\; 1 \quad \text{for} \quad t \geq 0 \quad \mathbb{P}\text{-a.s.} \tag{2.6}$$

With these notations, the evolution of the wealth on the risky asset account is given by

$$dY_t \;=\; Y_t \frac{dP_t}{P_t} \;+\; dL_t \;-\; Y_{t-}dM_t, \tag{2.7}$$

and, by definition of the tax basis B and (2.3), we have

$$dK_t = dL_t - K_{t-}dM_t. \tag{2.8}$$

Observe that the contribution of the sales in the dynamics of K_t is evaluated at the basis price. For any given initial condition (Y_{0-}, K_{0-}), (2.7)–(2.8) define a unique $\mathbb{F}$-adapted process (Y, K) taking values in $\mathbb{R}^2_+$, the nonnegative orthant of $\mathbb{R}^2$.

In addition to the trading activities, the investor consumes in continuous time at the rate $C = \{C_t,\, t \geq 0\}$. Here, C is an $\mathbb{F}$-progressively measurable process with

$$C \geq 0 \quad \text{and} \quad \int_0^T C_t dt < \infty \quad \mathbb{P}\text{-a.s.} \quad \text{for all} \quad T > 0. \tag{2.9}$$

Then, the bank component of the wealth process satisfies the dynamics

$$\begin{aligned} dX_t &= (rX_t - C_t)\, dt - dL_t + Y_{t-} dM_t - \ell\,(P_t - B_{t-})\, \frac{Y_{t-} dM_t}{P_t} \\ &= (rX_t - C_t)\, dt - dL_t + [(1-\alpha)Y_{t-} + \alpha K_{t-}]\, dM_t. \end{aligned} \tag{2.10}$$

Since the processes Y and K have been defined previously, the above dynamics uniquely defines an $\mathbb{F}$-adapted process X valued in $\mathbb{R}$ for any given initial condition X_{0-}.

Remark 2.3. In Dammon, Spatt, and Zhang [12], the nonrisky asset is also subject to a constant proportional taxation rule. This is obviously caught by our model by interpreting r as the after-tax instantaneous interest rate.

For later use, we report the dynamics of the corresponding liquidation value process defined in (2.5), which follows from (2.7), (2.8), (2.9), and (2.10):

$$dZ_t = (rZ_t - C_t)\, dt + (1-\alpha)\, Y_t \left(\frac{dP_t}{P_t} - rdt \right) - r\alpha K_t dt. \tag{2.11}$$

Definition 2.1. (i) *A consumption investment strategy is a triple of $\mathbb{F}$-adapted processes $\nu = (C, L, M)$, where C satisfies (2.9), L and M are nondecreasing and right continuous, $L_{0-} = M_{0-} = 0$, and the jumps of M satisfy (2.6).*

(ii) *Given an initial condition $s = (x, y, k) \in \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}_+$ and a consumption-investment strategy ν, we denote by $S^{s,\nu} = (X^{s,\nu}, Y^{s,\nu}, K^{s,\nu})$ the unique strong solution of (2.7), (2.8), (2.9), and (2.10) with initial condition $S^{s,\nu}_{0-} = s$.*

(iii) *Given an initial condition $s = (x, y, k) \in \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}_+$, a consumption-investment strategy ν is said to be s-admissible if the corresponding state process $S^{s,\nu}$ satisfies the no-bankruptcy constraint (2.5). We shall denote by $\mathcal{A}(s)$ the collection of all s-admissible consumption-investment strategies.*

The admissibility conditions imply that the process $S^{s,\nu}$ is valued in the closure $\bar{\mathcal{S}}$ of

$$\mathcal{S} = \left\{ (x, y, k) \in \mathbb{R}^3 \,:\, x + (1-\alpha)y + \alpha k > 0\,,\ y > 0\,,\ k > 0 \right\}. \tag{2.12}$$

We partition the boundary of $\mathcal{S}$ into $\partial\mathcal{S} = \partial^z\mathcal{S} \cup \partial^y\mathcal{S} \cup \partial^k\mathcal{S}$ with

$$\partial^y\mathcal{S} := \left\{ (x, y, k) \in \bar{\mathcal{S}} \,:\, y = 0 \right\}, \quad \partial^k\mathcal{S} := \left\{ (x, y, k) \in \bar{\mathcal{S}} \,:\, k = 0 \right\}$$

and

$$\partial^z\mathcal{S} = \left\{ (x, y, k) \in \bar{\mathcal{S}} \,:\, z := x + (1-\alpha)y + \alpha k = 0 \right\}.$$

2.4. The consumption-investment problem. The investor preferences are characterized by a power utility function with constant relative risk aversion coefficient $1-p \in (0,1)$:

$$U(c) := \frac{c^p}{p}, \quad c \geq 0, \quad \text{for some} \quad p \in (0,1).$$

The restriction on the relative risk aversion coefficient to $(0,1)$ allows us to simplify the analysis of this paper, as the boundary condition on $\partial^z \mathcal{S}$ is easily obtained; see Proposition 3.2. However, several of our results hold for a general parameter $p < 0$, and we will indicate whenever it is the case.

For every initial data $s \in \bar{\mathcal{S}}$ and any admissible strategy $\nu \in \mathcal{A}(s)$, we introduce the consumption-investment criterion

$$J_T(s,\nu) := \mathbb{E}\left[\int_0^T e^{-\beta t} U(C_t)dt + e^{-\beta T} U(Z_T^{s,\nu})\mathbf{1}_{\{T<\infty\}}\right], \quad T \in \mathbb{R}_+ \cup \{+\infty\}. \tag{2.13}$$

The consumption-investment problem is defined by

$$V(s) := \sup_{\nu \in \mathcal{A}(s)} J_\infty(s,\nu), \quad s \in \bar{\mathcal{S}}. \tag{2.14}$$

We shall assume that the parameters r, θ, σ, p, and β satisfy the condition

$$\bar{c}(r,\theta) := \frac{\beta - pr}{1-p} - \frac{p\theta^2}{2(1-p)^2} > 0, \tag{2.15}$$

which has been pointed out as a sufficient condition for the finiteness of the value function in the context of a financial market without taxes in [26] and [28].

2.5. Review of the tax-free model. In this section, we briefly review the solution of the consumption-investment problem when the financial market is free from taxes on capital gains. The properties of the corresponding value function are going to be useful to state relevant bounds for the maximal utility achieved in a financial market with taxes.

In the classical formulation of the tax-free consumption-investment problem [26], the investment control variable is described by means of a unique process π which represents the proportion of wealth invested in risky assets at each time, and the consumption process C is expressed as a proportion c of the total wealth:

$$d\bar{Z}_t = \bar{Z}_t\left[(r - c_t)dt + \pi_t \sigma(\theta dt + dW_t)\right]. \tag{2.16}$$

In this context, a consumption-investment admissible strategy is a pair of adapted processes (c,π) such that c is nonnegative and

$$\int_0^T c_t dt + \int_0^T |\pi_t|^2 dt < \infty \quad \mathbb{P}\text{-a.s.} \quad \text{for all} \quad T > 0.$$

We shall denote by $\bar{\mathcal{A}}$ the collection of all such consumption-investment strategies. For every initial condition $z \geq 0$ and strategy $(c,\pi) \in \bar{\mathcal{A}}$, there is a unique strong solution to (2.16) that we denote by $\bar{Z}^{z,c,\pi}$. The frictionless consumption-investment problem is

$$\bar{V}(z) := \sup_{(c,\pi)\in\bar{\mathcal{A}}} \mathbb{E}\left[\int_0^\infty e^{-\beta t} U\left(c_t \bar{Z}_t^{z,c,\pi}\right) dt\right]. \tag{2.17}$$

Theorem 2.1 (see [26]). *Let condition* (2.15) *hold. Then, for all* $z \geq 0$,

$$\bar{V}(z) = \bar{c}(r,\theta)^{p-1}\,\frac{z^p}{p},$$

and the constant consumption-investment strategy $\bar{c}(r,\theta)$, $\bar{\pi}(\sigma,\theta) := \frac{\theta}{(1-p)\sigma}$ *is optimal.*

Remark 2.4. The reduction of the model of section 2 to the frictionless case, i.e., $\alpha = 0$, does not alter the value function. This can be proved by approximating any investment strategy by a sequence of bounded variation strategies. However, the investment strategies in our formulation are constrained to have bounded variation. This is needed because sales and purchases have different impacts on the bank component of the wealth process (2.10). Since the Merton optimal strategy is well known to be unique and has unbounded variation, it follows that existence fails to hold in our formulation.

3. First properties of the value function. We first show that the optimal consumption-investment problem under taxes reduces to the Merton problem when the interest rate is zero. Let $\bar{c}(r,\theta)$ be as in Theorem 2.1.

Proposition 3.1. *For* $r = 0$, $V(s) = \frac{\bar{c}(0,\theta)^{p-1}}{p}\,(x + (1-\alpha)y + \alpha k)^p$, $s = (x,y,k) \in \bar{\mathcal{S}}$.

Proof. Notice that the optimal consumption-investment problem (2.14) can be expressed equivalently in terms of the state processes (Z,Y,K) instead of (X,Y,K), and observe that the interest rate parameter r is involved only in the dynamics of Z. When $r = 0$, the dynamics of the process Z in (2.11)

$$dZ_t = -C_t dt + (1-\alpha)\,Y_t \frac{dP_t}{P_t}$$

is independent of the tax basis K_t. Since the dynamics of Y in (2.7) is independent of K, it follows that the value function does not depend on the variable k. Next, for $Z_t > 0$, defining $c_t := C_t/Z_t$ and $\pi_t := (1-\alpha)Y_t/Z_t$, we see that the solution Z of the above equation is the same as the solution $\bar{Z}^{z,c,\pi}$ of (2.16). In view of the state constraint $Z \geq 0$, our state dynamics in the context $r = 0$ is then equivalent to (2.16). Then, the only difference between the control problem V and the corresponding Merton problem $\bar{V}$ is the class of admissible trading strategies, which does not induce any difference on the value function; see Remark 2.4. ■

The argument in the above proof clearly does not involve the specific nature of the utility function. Therefore an analogous result holds for *any utility function*.

Since the tax basis is not inflated by the interest rate r, for nonzero values of r the tax basis plays a role in the solution. This explains the importance of the point $r = 0$.

We next discuss the value function on the boundary of the state space $\mathcal{S}$. Observe that there is no a priori information on the boundary components $\partial^y\mathcal{S}$ and $\partial^k\mathcal{S}$. This is one source of difficulty in the numerical part of this paper, as this state constraint problem needs special treatment; see [5].

Proposition 3.2. *For every* $s \in \partial^z\mathcal{S}$, *we have* $V(s) = 0$.

Proof. Let s be in $\partial^z\mathcal{S}$, and let ν be in $\mathcal{A}(s)$. By the definition of the set admissible controls, the process $Z^{s,\nu}$ is nonnegative. By Itô's lemma, together with the nonnegativity of C and K and the nondecrease of L, this provides

$$0 \;\leq\; e^{-rt}Z_t^{s,\nu} \leq (1-\alpha)\int_0^t e^{-ru}Y_u^{s,\nu}\sigma\left[\theta du + dW_u\right].$$

Let $\mathbb{Q}$ be the probability measure equivalent to $\mathbb{P}$ under which the process $\{\theta u + W_u,\, u \geq 0\}$ is a Brownian motion. The process appearing on the right-hand side of the last inequality is a $\mathbb{Q}$-supermartingale because it is a nonnegative $\mathbb{Q}$-local martingale. By taking expected values under $\mathbb{Q}$, it then follows from the last inequalities that $Z^{s,\nu} = Y^{s,\nu} = K^{s,\nu} = C = L \equiv 0$. We have then proved that, for $s \in \partial^z \mathcal{S}$, any admissible strategy $\nu = (C, L, M) \in \mathcal{A}(s)$ is such that $C = L \equiv 0$, implying that $V(s) = 0$. ■

Proposition 3.3. *The value function V is nondecreasing with respect to each of the variables x, y, and k. Moreover, for $(x, y, k) \in \bar{\mathcal{S}}$ with $z := x + (1-\alpha)y + \alpha k > 0$,*

$$V(x,y,k) \;=\; z^p\, \mathcal{V}\left(\frac{y}{z}, \frac{k}{z}\right), \quad \textit{where} \quad \mathcal{V}(\xi,\zeta) := V\left(1-(1-\alpha)\xi - \alpha\zeta, \xi, \zeta\right). \tag{3.1}$$

Proof. 1. The monotonicity property with respect to x, y, and k follows immediately from the dynamics of the problem and the bound $\Delta M \leq 1$.

2. Let $\nu = (C, L, M)$ be an arbitrary strategy in $\mathcal{A}(s)$, and define the strategy $\nu' := (\delta C, \delta L, M)$. We easily verify that $S^{\delta s, \nu'} = \delta S^{s,\nu} \in \bar{\mathcal{S}}$, which implies that ν' is in $\mathcal{A}(\delta s)$, and therefore

$$V(\delta s) \geq \mathbb{E}\left[\int_0^\infty e^{-\beta u} U(\delta C_u) du\right] \;=\; \delta^p J_\infty(\delta, \nu),$$

where the last equality follows from the homogeneity property of the utility function U. By the arbitrariness of ν in $\mathcal{A}(s)$, this shows that $V(\delta s) \geq \delta^p V(s)$.

3. By writing $V(s) = V\left(\delta^{-1}\delta s\right) \geq \delta^{-p} V\left(\delta s\right)$, it follows from the previous step that we in fact have $V(\delta s) = \delta^p V(s)$, and the required result follows immediately from this homotheticity property. ■

In the absence of taxes on capital gains, i.e., $\alpha = 0$, it is easy to deduce from the concavity of U that the value function V is concave and therefore continuous. The numerical results exhibited in section 6 reveal that this property is no longer valid when $\alpha > 0$. The proof of the following continuity result is obtained by first reducing the continuity problem to the ray $\{(x, 0, 0),\, x \in \mathbb{R}_+\}$. This is achieved by means of a comparison result in the sense of viscosity solutions. Then the continuity on the latter ray is proved by a direct argument.

Proposition 3.4. *The function $\mathcal{V}$ of* (3.1) *is Lipschitz continuous on $\bar{\mathcal{S}}$.*

Proof. See Appendix A. ■

We now show that it is always worth realizing capital losses whenever the tax basis exceeds the spot price of the risky asset. In other words, given $s = (x, y, k) \in \bar{\mathcal{S}}$, every admissible strategy $\nu \in \mathcal{A}(s)$, with $K_\tau^{s,\nu} > Y_\tau^{s,\nu}$ (i.e., $B_\tau^{s,\nu} > P_\tau$) for some stopping time τ, can be improved strictly by realizing the capital loss on the entire portfolio at time τ. This property is observed in practice and is known as a *wash sale*. It was stated in [12] and embedded directly in the definition of the tax basis. This result is *independent of the choice of the utility function.*

Proposition 3.5. *Consider some $s \in \bar{\mathcal{S}}$ and $\nu = (C, L, M) \in \mathcal{A}(s)$. Assume that $K_\tau^{s,\nu} > Y_\tau^{s,\nu}$ a.s. for some finite stopping time τ. Then, there exists an admissible strategy $\tilde{\nu} = \left(\tilde{C}, \tilde{L}, \tilde{M}\right) \in \mathcal{A}(s)$ such that, for any utility function,*

$$Y^{\tilde{\nu}} = Y^{\nu}, \quad \Delta\tilde{M} - \Delta M = \mathbf{1}_{\{\tau\}}, \quad \textit{and} \quad J_\infty(s, \tilde{\nu}) \;>\; J_\infty(s, \nu);$$

i.e., a wash sale is optimal.

Proof. We organize the proof in two steps.

1. Set $(L', M') := (L, M) + (Y_\tau, 1)(1 - \Delta M_\tau)1_{t\geq\tau}$. We shall prove that $\nu' = (C, L', M') \in \mathcal{A}(s)$ and that the resulting state process satisfies

$$Y^{s,\nu'} = Y^{s,\nu}, \quad Z^{s,\nu'} \geq Z^{s,\nu}, \quad K^{s,\nu'} \leq K^{s,\nu} \quad \text{a.s.} \quad \text{and} \quad Z_t^{\nu'} > Z_t^{\nu} \quad \text{a.s.} \quad \text{on} \quad \{t > \tau\}.$$

To see this, observe that since ν and ν' differ only by the jump at the stopping time τ, and $\Delta Y_\tau^{\nu'} = \Delta Y_\tau^{\nu}$, we have

$$Y^{s,\nu'} = Y^{s,\nu}, \quad \left(Z_t^{s,\nu'}, K_t^{s,\nu'}\right) = (Z_t^{s,\nu}, K_t^{s,\nu}) \quad \text{for} \quad t < \tau \quad \text{and} \quad Z_\tau^{s,\nu'} = Z_\tau^{s,\nu}$$

by the continuity of the process Z. Observe that the newly defined strategy ν' consists in selling out the whole portfolio at time τ as $\Delta M'_\tau = 1$. Hence $K_\tau^{s,\nu'} = Y_\tau^{s,\nu'} = Y_\tau^{s,\nu}$, and we compute directly from (2.8) that

$$K_t^{s,\nu'} - K_t^{s,\nu} = (Y_\tau^{s,\nu} - K_\tau^{s,\nu})\, e^{-M_t^c + M_\tau^c} \prod_{\tau < u \leq t} (1 - \Delta M_u) \;>\; 0 \quad \text{for} \quad t \geq \tau,$$

since $Y_\tau^{s,\nu} - K_\tau^{s,\nu} < 0$. By (2.11), this provides

$$e^{-rt}\left(Z_t^{s,\nu'} - Z_t^{s,\nu}\right) = -r\alpha \int_\tau^t e^{-ru}\left(K_u^{s,\nu'} - K_u^{s,\nu}\right) du \;>\; 0 \quad \text{for} \quad t > \tau,$$

which shows that $Z^{s,\nu'} \geq 0$ and $\nu' \in \mathcal{A}(s)$.

2. Define the strategy $\tilde{\nu} = (\tilde{C}, \tilde{L}, \tilde{M})$ by

$$\tilde{C}_t \;:=\; C_t \;+\; \xi\left(Z_t^{s,\tilde{\nu}} - Z_t^{s,\nu}\right) 1_{t\geq\tau} \quad \text{and} \quad \left(\tilde{L}, \tilde{M}\right) \;:=\; \left(L', M'\right), \tag{3.2}$$

where ξ is an arbitrary positive constant. Observe that $(Y^{s,\tilde{\nu}}, K^{s,\tilde{\nu}}) = (Y^{s,\nu'}, K^{s,\nu'})$, and $Z_t^{s,\tilde{\nu}} = Z_t^{s,\nu'} = Z_t^{s,\nu}$ for $t \leq \tau$. In particular, $K^{s,\tilde{\nu}} - K^{s,\nu} = K^{s,\nu'} - K^{s,\nu} \leq 0$. Set $\Delta K := K^{s,\tilde{\nu}} - K^{s,\nu}$ and $\Delta Z := Z^{s,\tilde{\nu}} - Z^{s,\nu}$. In order to check the admissibility of the strategy $\tilde{\nu}$, we directly compute that

$$\begin{aligned} e^{-r(t-\tau)}\Delta Z_t &= \Delta Z_\tau - r\alpha \int_\tau^t e^{-r(u-\tau)}\Delta K_u du + \xi \int_\tau^t e^{-r(u-\tau)}\Delta Z_u du \\ &\geq \xi \int_\tau^t e^{-r(u-\tau)}\Delta Z_u du. \end{aligned}$$

By the Gronwall inequality, this implies that $Z_t^{\tilde{\nu}} > Z_t^{\nu}$ on $\{t > \tau\}$, and therefore $\tilde{C} > C$ on $\{t > \tau\}$ with positive Lebesgue$\otimes P$ measure. Hence $J_\infty(s; \tilde{\nu}) > J_\infty(s; \nu)$. ∎

Remark 3.1. It follows from the previous proposition that $V(x, y, k) = V(x + y + \alpha(k - y), 0, 0)$ whenever $k > y$. Then, we may restrict our analysis of the value function to the set $\{(x, y, k) \in \bar{\mathcal{S}} : k \leq y\}$. We could not find any benefit from this reduction. Even the numerical implementation is not simplified by this domain restriction because we have no natural boundary condition on $\{k = y\}$. We therefore continue our analysis of the value function V on the total domain $\bar{\mathcal{S}}$.

Remark 3.2. The previous proposition highlights the difficulty in characterizing a solution of the optimal consumption-investment problem under taxes. The optimality of wash sales suggests that the optimal trading strategy has a local time type of behavior. We have no theoretical result to support this intuition. Our very limited information on the regularity of the value function is the main obstacle in developing a verification argument similar to that of Davis and Norman [13].

4. The first order approximation. In this section, we provide upper and lower bounds for the value function. The upper bound expresses that there is no way for the investor to take advantage of tax credits in order to do better than in the tax-free financial market, and this holds for *any utility function*. Our derivation of the lower bound explicitly uses the power utility but *without any restriction* on the risk aversion factor p. These bounds will be used in order to obtain a first order approximation result for $p < 1$.

Proposition 4.1. *For any $s = (x, y, k) \in \bar{\mathcal{S}}$, we have $V(s) \leq \bar{V}(x + (1-\alpha)y + \alpha k)$.*

Proof. Let $s = (x, y, k)$ be in $\bar{\mathcal{S}}$. Consider some consumption-investment strategy $\nu = (C, L, M)$ in $\mathcal{A}(s)$. Define a consumption-investment strategy $\tilde{\nu} = (C, (1-\alpha)L, M)$, and denote by $(\tilde{X}, \tilde{Y})$ the corresponding tax-free bank and risky asset account processes with the initial endowment $(x + \alpha k,\, (1-\alpha)y)$. Clearly, $\tilde{Y} = (1-\alpha)Y^{s,\nu} \geq 0$. To see that $\tilde{\nu}$ is admissible in the tax-free financial market, observe that the process $\tilde{Z} := \tilde{X} + \tilde{Y}$ satisfies

$$\tilde{Z}_{0-} - Z^{s,\nu}_{0-} = 0 \quad \text{and} \quad \tilde{Z}_t - Z^{s,\nu}_t \geq e^{rt} \int_0^t e^{-ru} r\alpha K^{s,\nu}_u du \geq 0,$$

so that $\tilde{Z}^{s,\nu} \geq Z^{s,\nu} \geq 0$. Hence, $\bar{V}(x + (1-\alpha)y + \alpha k) \geq J_\infty(s, \nu)$; see Remark 2.4. The required result follows from the arbitrariness of $\nu \in \mathcal{A}(s)$. ∎

Proposition 4.2. *For $s = (x, y, k)$ in $\bar{\mathcal{S}}$ and $z = x + (1-\alpha)y + \alpha k$, there exists a sequence of admissible strategies $(\nu^n)_{n\geq 1} \subset \mathcal{A}(s)$ such that*

$$V(s) \;\geq\; \bar{c}\left(r, \tilde{\theta}^\alpha\right)^{p-1} \frac{z^p}{p} \;=\; \lim_{n\to\infty} J_\infty\left(s, \nu^n\right), \quad \textit{where} \quad \tilde{\theta}^\alpha \;:=\; \theta - \frac{r\alpha}{\sigma(1-\alpha)};$$

i.e., the value function of the Merton frictionless problem with the smaller risk premium $\tilde{\theta}^\alpha$ can be approached as close as possible in the context of the financial market with taxes.

This result is proved by producing a sequence of admissible strategies $(C_n, L_n, M_n)_{n\geq 1} \subset \mathcal{A}(s)$ which approximates Merton's value function with the smaller risk premium $\tilde{\theta}^\alpha$. To give an intuitive justification of this result, we rewrite (2.11) as

$$dZ_t \;=\; (rZ_t - C_t)\, dt \;+\; Y_t\, \tilde{\sigma}^\alpha \left(dW_t + \tilde{\theta}^\alpha dt\right) \;+\; r\alpha\, (Y_t - K_t)\, dt, \tag{4.1}$$

where $\tilde{\theta}^\alpha$ is defined as in the statement of Proposition 4.2 and $\tilde{\sigma}^\alpha := (1-\alpha)\sigma$. Compare the above $\bar{Z}$ dynamics to (2.16) with modified parameters $(\tilde{\sigma}^\alpha, \tilde{\theta}^\alpha)$ and with $c_t = C_t/\bar{Z}_t$, $\pi_t = Y_t/\bar{Z}_t$. The only difference is the term $r\alpha(Y - K)$. However, in view of Proposition 3.5, we expect this term to be nonnegative for the optimal strategy (if it exists). This hints that the liquidation value process $\bar{Z}$ (with the above choices C and Y) is larger than the wealth process in the fictitious tax-free financial market with a modified risk premium. This formally justifies the inequality of Proposition 4.2.

The proof reported in Appendix B exhibits an explicit sequence of strategies which mimics the optimal consumption-investment strategy in the Merton frictionless model while keeping the difference $Y - K$ small or, equivalently, the tax basis close to the spot price of the risky asset.

Remark 4.1. Let $b := r + \theta\sigma$ be the instantaneous mean return coefficient in our financial market. Then, the modified risk premium $\tilde{\theta}^\alpha$ can be easily interpreted in terms of the modified volatility coefficient $\sigma^\alpha = (1-\alpha)\sigma$ and a similarly modified instantaneous mean return coefficient $b^\alpha := (1-\alpha)b$ as $\tilde{\theta}^\alpha = (b^\alpha - r)/\sigma^\alpha$. This fictitious financial market with such modified coefficients corresponds to the situation where the investor is forced to realize the capital gains or losses, at each time t, before adjusting the portfolio.

Propositions 4.1 and 4.2 provide the following bounds on the value function V:

$$\frac{\bar{c}(r,\tilde{\theta}^\alpha)^{p-1}}{p}(x+(1-\alpha)y+\alpha k)^p \le V(x,y,k) \le \frac{\bar{c}\,(r,\theta)^{p-1}}{p}(x+(1-\alpha)y+\alpha k)^p, \tag{4.2}$$

where $\tilde{\theta}^\alpha$ is defined as in the statement of Proposition 4.2, and $\bar{c}$ is defined as in Theorem 2.1. Observe that $\tilde{\theta}^\alpha = \theta$ whenever $\alpha = 0$ or $r = 0$. Therefore, we might expect that these bounds are tight for small interest rate or tax parameters.

Corollary 4.1. *For $s = (x,y,k) \in \mathcal{S}$, we have*

$$V(s) = V^{\text{app}}(s) + \text{o}(\alpha + r),$$

where $\text{o}(\xi)$ is a function on $\mathbb{R}$ with $\text{o}(\xi)/\xi \longrightarrow 0$ as $\xi \to 0$, and

$$V^{\text{app}}(s) := \frac{\bar{c}(0,\theta)^{p-1}}{p}\left(1 + \frac{rp}{\bar{c}(0,\theta)}\right)(x+y)^p + \alpha\,\bar{c}(0,\theta)^{p-1}(k-y)(x+y)^{p-1}.$$

Proof. It is sufficient to observe that the bounds on the value function V in (4.2) are smooth functions with the identical partial gradient with respect to (r,α) at the origin. This follows from the fact that $(\partial\tilde{\theta}^\alpha/\partial\alpha) = (\partial\tilde{\theta}^\alpha/\partial r) = 0$ at $(r,\alpha) = (0,0)$. ■

Remark 4.2. Observe that the function $\bar{c}$ defined in Theorem 2.1 is increasing in the r variable. Then, the above first order expansion shows that the value function V is also increasing in the interest rate variable (for small interest rate and tax parameters). This is intuitively clear, as the larger interest rate provides the investor a better opportunity set.

The dependence of the value function on the tax rate α is more complex, and it depends on the initial position of the tax basis. If the initial tax basis is larger than the spot price, i.e., in a situation of capital gain loss, the investor takes immediate advantage of the tax credit, as stated in Proposition 3.5, and the value function V is increasing in α (for small α). In the opposite situation, i.e., when the initial tax basis is smaller than the spot price, the value function is decreasing in α. Finally, when the initial tax basis coincides with the spot price, the value function is not sensitive to the tax rate in the first order.

This variation of the value function (up to the first order) in terms of the tax rate α is somehow surprising. Indeed, in a capital loss situation, an increase of the tax parameter implies two opposing results:

- an increase of the tax credit is received initially by the agent;
- a larger amount of tax is paid during the infinite lifetime of the agent.

Our first order expansion shows that, for small interest rate and tax parameters, the increase of initial tax credit is never compensated by the increase of tax over the infinite lifetime. This is in agreement with Cadenillas and Pliska [7], who found that "sometimes investors are better off with a positive tax rate."

The same reasoning also shows that when there are no initial embedded capital gains (i.e., when $y = k = 0$) the effect of the tax parameter is only second order.

Remark 4.3. Since the lower bound in (4.2) has the same first order Taylor expansion as the value function V, we can view the corresponding strategy as nearly optimal. From the discussion following Proposition 4.2, the portfolio allocation defining the lower bound is by definition an approximation of the constant portfolio allocation

$$\bar{\pi}(\tilde{\sigma}^\alpha, \tilde{\theta}^\alpha) \;=\; \frac{1}{(1-p)\sigma^2}\left[\frac{b}{1-\alpha} - \frac{r}{(1-\alpha)^2}\right],$$

where $b := \sigma\theta + r$ is the instantaneous mean return of the risky asset. Direct computation shows that $\bar{\pi}(\tilde{\sigma}^\alpha, \tilde{\theta}^\alpha) \le \bar{\pi}(\sigma, \theta)$ if and only if $r \ge (1-\alpha)\,(b - (1-\alpha)(b-r))$. Using the data set of Dammon, Spatt, and Zhang [12] ($r = 6\%$, $b = 9\%$, $\alpha = 36\%$), we see that $\bar{\pi}(\tilde{\sigma}^\alpha, \tilde{\theta}^\alpha) \le \bar{\pi}(\sigma, \theta)$. Since this nearly optimal strategy exhibits a smaller exposition to the risky asset, the presence of taxes on capital gains contributes to explaining the equity premium puzzle highlighted by Mehra and Prescott [25].

Notice that this observation is in contradiction with the numerical results of Dammon, Spatt, and Zhang [12], who found that the exposition to the risky asset is increased by the presence of taxes. This is due to the fact that the bank account in their model is also subject to taxes with the same tax rate as for the risky asset, which implies that the optimal portfolio strategy in the forced realization case is given by

$$\hat{\pi}^\alpha \;=\; \frac{b(1-\alpha) - r(1-\alpha)}{(1-p)\sigma^2(1-\alpha)^2} = \frac{\bar{\pi}(\sigma,\theta)}{1-\alpha},$$

which is increasing in α.

5. Characterization by the dynamic programming equation.

The chief goal of this section is to provide a characterization of V by means of a second order partial differential equation for which we shall provide a numerical solution in the subsequent section. Unfortunately, we are unable to obtain a characterization of V by the corresponding dynamic programming equation. Therefore, we shall exhibit a consistent approximation V^ε as the unique solution of an approximating second order partial differential equation.

For s in $\bar{\mathcal{S}}$ and $\nu = (C, L, M)$ in $\mathcal{A}$, the jumps of the state processes S are given by

$$\Delta S_t^{s,\nu} \;=\; -\Delta L_t\,\mathbf{g^b} \;-\; \Delta M_t\left[(1-\alpha)Y_{t-}^{s,\nu} + \alpha K_{t-}^{s,\nu}\right]\mathbf{g^s}\left(S_{t-}^{s,\nu}\right),$$

where the vector fields $\mathbf{g^b}$ and $\mathbf{g^s}(x,y,k)$ are defined by

$$\mathbf{g^b} := \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix} \quad \text{and} \quad \mathbf{g^s}(s) := \begin{pmatrix} -1 \\ \frac{1}{1-\alpha} \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \frac{-\alpha}{1-\alpha} \\ 1 \end{pmatrix} \frac{k}{(1-\alpha)\,y + \alpha k}\,\mathbf{1}_{(y,k)\neq 0}.$$

The value function of our consumption-investment problem is formally expected to solve the corresponding dynamic programming equation:

$$\min\left\{-\mathcal{L}v\;,\;\mathbf{g^b}\cdot Dv\;,\;\mathbf{g^s}\cdot Dv\right\}\;=\;0\quad\text{on}\quad\bar{\mathcal{S}}\setminus\partial^z\mathcal{S}\quad\text{and}\quad v=0\quad\text{on}\quad\partial^z\mathcal{S},\tag{5.1}$$

where $\mathcal{L}$ is the second order differential operator defined by

$$\mathcal{L}v\;=\;-\beta v+rxv_x+byv_y+\frac{1}{2}\sigma^2y^2v_{yy}+\tilde{U}(v_x),\quad\text{with}\quad\tilde{U}(q)\;=\;\sup_{c\geq0}\left(U(c)-cq\right).$$

Observe that we have no information on the regularity of the value function V; hence we cannot prove that V is a classical solution to (5.1). Moreover, the value function V is known only on the boundary $\partial^z\mathcal{S}$ (see Proposition 3.2), but there is no possible knowledge of V on $\partial^y\mathcal{S}\cup\partial^k\mathcal{S}$. We then need to use the notion of viscosity solutions which allows for a weak formulation of solutions to partial differential equations and boundary conditions.

Because $\mathbf{g^s}$ is not locally Lipschitz continuous, it is not clear that there is a unique characterization of the value function V as the constrained viscosity solution of (5.1). Due to this technical difficulty, we isolated in the accompanying paper [5] some viscosity results for a slightly modified equation. The objective of this section is to build on these results in order to characterize the value function V as a limit of an approximating function defined as the unique viscosity solution of a conveniently perturbed equation.

For every $\varepsilon>0$, we define the function

$$f^\varepsilon(x,y,k)\;:=\;1\wedge\left(\frac{k}{\varepsilon z}-1\right)^+,\quad\text{with}\quad z:=x+(1-\alpha)y+\alpha k,$$

together with the approximation of $\mathbf{g^b}$ and $\mathbf{g^s}$:

$$\mathbf{g}^{\mathbf{b}}_\varepsilon:=\begin{pmatrix}1+\varepsilon\\-1\\-1\end{pmatrix}\quad\text{and}\quad\mathbf{g}^{\mathbf{s}}_\varepsilon(x,y,k):=\mathbf{g^s}\left(x,y,kf^\varepsilon(s)\right)$$

for $s\in\mathcal{S}\setminus\partial^z\mathcal{S}$. Notice that $\mathbf{g}^{\mathbf{s}}_\varepsilon$ is locally Lipschitz continuous on $\bar{\mathcal{S}}\setminus\partial^z\mathcal{S}$, and $\mathbf{g}^{\mathbf{s}}_\varepsilon(s)=\mathbf{g^s}(s)$ whenever $k\geq2\varepsilon z$. The main result of this section provides a characterization of the value function V by means of the approximating equation:

$$\min\left\{-\mathcal{L}v\;,\;\mathbf{g}^{\mathbf{b}}_\varepsilon\cdot Dv\;,\;\mathbf{g}^{\mathbf{s}}_\varepsilon\cdot Dv\right\}\;=\;0\quad\text{on}\quad\bar{\mathcal{S}}\setminus\partial^z\mathcal{S}\quad\text{and}\quad v\;=\;0\quad\text{on}\quad\partial^z\mathcal{S}.\tag{5.2}$$

We first recall the notion of a constrained viscosity solution first introduced in [29, 30]. For a locally bounded function $u:\bar{\mathcal{S}}\longrightarrow\mathbb{R}$, we shall use the classical notation in viscosity theory for the corresponding upper semicontinuous and lower semicontinuous envelopes:

$$u^*(s)\;:=\;\limsup_{\mathcal{S}\ni s'\to s}u(s')\quad\text{and}\quad u_*(s)\;:=\;\liminf_{\mathcal{S}\ni s'\to s}u(s').$$

Definition 5.1. (i) *A locally bounded function u is a constrained viscosity subsolution of (5.2) if $u^*\leq0$ on $\partial^z\mathcal{S}$, and for all $s\in\bar{\mathcal{S}}\setminus\partial^z\mathcal{S}$ and $\varphi\in C^2(\bar{\mathcal{S}})$ with $0=(u^*-\varphi)(s)=\max_{\bar{\mathcal{S}}\setminus\partial^z\mathcal{S}}(u^*-\varphi)$ we have* $\min\left\{-\mathcal{L}\varphi,\,\mathbf{g}^{\mathbf{b}}_\varepsilon\cdot D\varphi,\,\mathbf{g}^{\mathbf{s}}_\varepsilon\cdot D\varphi\right\}\leq0$.

(ii) *A locally bounded function* u *is a constrained viscosity supersolution of* (5.2) *if* $u_* \geq 0$ *on* $\partial^z \mathcal{S}$, *and for all* $s \in \mathcal{S}$ *and* $\varphi \in C^2(\mathcal{S})$ *with* $0 = (u_* - \varphi)(s) = \min_{\mathcal{S}}(u_* - \varphi)$ *we have* $\min\left\{-\mathcal{L}\varphi,\, \mathbf{g}^{\mathbf{b}}_\varepsilon \cdot D\varphi,\, \mathbf{g}^{\mathbf{s}}_\varepsilon \cdot D\varphi\right\} \geq 0$.

(iii) *A locally bounded function* u *is a constrained viscosity solution of* (5.2) *if it is a constrained viscosity subsolution and supersolution.*

In the above definition, there is no boundary value assigned to the value function on $\partial^y \mathcal{S} \cup \partial^k \mathcal{S}$. Instead, the subsolution property holds on this boundary. Notice that the supersolution property is satisfied only in the interior of the domain $\mathcal{S}$.

Theorem 5.1. *For each* $\varepsilon > 0$, *the boundary value problem* (5.2) *has a unique constrained viscosity solution* V^ε *in the class* $C^0_p\left(\bar{\mathcal{S}}\right)$. *Moreover, the following hold:*

(i) *the family* $(V^\varepsilon)_{\varepsilon>0}$ *is nonincreasing and converges to the value function* V *uniformly on compact subsets of* $\bar{\mathcal{S}}$ *as* $\varepsilon \searrow 0$;

(ii) *for every* $s \in \bar{\mathcal{S}}$ *and* $\delta \geq 0$, *we have* $V^\varepsilon(\delta s) = \delta^p V^\varepsilon(s)$.

Proof. The existence of V^ε as the unique constrained viscosity solution of (5.2) in the class $C^0_p\left(\bar{\mathcal{S}}\right)$ is shown in Theorem 3.2 of [5], where we introduced the value function $v^{\varepsilon,\lambda}$ of a consumption-investment problem with transaction costs $\lambda > 0$ and an ε-modified taxation rule near the ray $\{(x,0,0),\, x \in \mathbb{R}_+\}$. Here $V^\varepsilon = v^{\varepsilon,\varepsilon}$.

We next use Theorem 5.1 of [5], which provides a stochastic control representation of $v^{\varepsilon,\lambda}$. In particular, it follows directly from this representation that $v^{\varepsilon,\lambda}(\delta s) = \delta^p v^{\varepsilon,\lambda}(s)$ for all $\delta \geq 0$. This homotheticity property is obviously inherited by $V^\varepsilon = v^{\varepsilon,\varepsilon}$, as announced in (ii).

We now prove (i) in the next two steps.

Step 1. The monotonicity of the sequence $(V^\varepsilon)_{\varepsilon>0}$ is inherited from the nonincrease of the sequence $\left(v^{\varepsilon,\lambda}\right)_{\varepsilon>0}$, proved in Proposition 6.2 of [5], together with the decrease of the sequence $\left(v^{\varepsilon,\lambda}\right)_{\lambda>0}$. It follows from this monotonicity property that $V^0 := \lim_{\varepsilon\to 0} V^\varepsilon$ is well defined. Now observe that $v^{\varepsilon,\lambda} \leq V^\varepsilon \leq v^{0,\varepsilon} \leq V$ for every $\lambda \geq \varepsilon$. Then $\lim_{\lambda\to 0}\lim_{\varepsilon\to 0} v^{\varepsilon,\lambda} \leq V^0 \leq V$. By Proposition 6.3 of [5], this implies that

$$\lim_{\lambda\to 0} v^\lambda \leq v^{0,\varepsilon} \leq V, \quad \text{where} \quad v^\lambda := \lim_{\varepsilon\to 0} v^{\varepsilon,\lambda} \tag{5.3}$$

is the value function of the optimal consumption-investment problem with taxes and proportional transaction cost $\lambda > 0$. In order to conclude that $V^0 = V$, it remains to show that

$$\lim_{\lambda\to 0} v^\lambda = V.$$

Step 2. For fixed $s \in \bar{\mathcal{S}}$, let $\nu^n = (C^n, L^n, M^n)$ be such that

$$V(s) - \frac{1}{n} \leq J_\infty(s, \nu^n) \quad \text{for all} \quad n \geq 1.$$

Denote by $Z^{\lambda,n}$ the after-tax liquidation value process with consumption-investment strategy ν^n in a financial market subject to taxes and constant proportional transaction cost parameter λ. See Appendix A for the precise formulation. We also denote by Z^n the corresponding after-tax liquidation value process without transaction costs. Then, it follows immediately from the dynamics of these processes that

$$Z^{\lambda,n}_t = Z^n_t - \lambda L^n_t \quad \text{for all} \quad t \geq 0.$$

Then the stopping times

$$\tau^\lambda := \inf\left\{t > 0 : \; Z^{\lambda,n} \le \frac{1}{2} Z_t^n\right\} \quad \text{satisfy} \quad \tau^\lambda \longrightarrow \infty \quad \text{as} \quad \lambda \searrow 0 \quad \mathbb{P}\text{-a.s.}$$

Define the strategies $\tilde{\nu}^n$ by

$$\tilde{\nu}^n := \nu^n \mathbf{1}_{[0,\tau^\lambda)} + \left(0, L^n_{\tau^\lambda -}, M^n_{\tau^\lambda -} + 1\right) \mathbf{1}_{[\tau^\lambda,\infty)}.$$

Clearly, $\tilde{\nu}^n$ is admissible for the problem with transaction costs, i.e., $\tilde{\nu}^n \in \mathcal{A}^\lambda(s)$ in the notation of Appendix A. Then

$$v^\lambda(s) \;\ge\; J_\infty(s, \tilde{\nu}^n) \;=\; \mathbb{E}\left[\int_0^{\tau^\lambda} U(C_t^n) dt\right] \longrightarrow \mathbb{E}\left[\int_0^\infty U(C_t^n) dt\right] \;\ge\; V(s) - \frac{1}{n},$$

where we use the monotone convergence theorem. By the arbitrariness of n and (5.3), this shows that $V^0 = V$. Finally the convergence holds uniformly on compact subsets by the monotonicity of (V^ε) and the continuity of the limit V. ■

6. Numerical estimate for V. We have stated in the previous section that the value function V is approximated by the functions $(V^\varepsilon)_{\varepsilon>0}$, where, for each $\varepsilon > 0$, V^ε can be computed as the unique viscosity solution of the boundary value problem (5.2). In this section, we provide a numerical estimate for V, based on a numerical scheme for (5.2). Unfortunately, we have no theoretical convergence result for our algorithm. Indeed, as discussed in the introduction, establishing convergence results for Hamilton–Jacobi–Bellman equations corresponding to singular control problems is an open question in numerical analysis. We therefore follow previous related works such as [1] and [31] by concentrating our effort on realizing the empirical convergence of the algorithm.

Following Akian, Menaldi, and Sulem [1], we adopt a numerical scheme based on the finite difference discretization and the classical Howard algorithm. For the convenience of the reader we briefly describe it hereafter, and we refer the reader to [1] for a detailed discussion.

6.1. Change of variables and reduction of the state dimension. By the homotheticity property of V^ε (Theorem 5.1(ii)) we have for $s = (x, y, k) \in \bar{\mathcal{S}} \setminus \partial^z \mathcal{S}$ and $z := x + [(1-\alpha)y + \alpha k]$

$$V^\varepsilon(s) \;=\; z^p \, \mathcal{V}^\varepsilon\left(\frac{y}{z}, \frac{k}{z}\right), \quad \text{where} \quad \mathcal{V}^\varepsilon(\xi_1, \xi_2) \;:=\; V^\varepsilon\left(1 - (1-\alpha)\xi_1 - \alpha\xi_2, \xi_1, \xi_2\right).$$

Next, for a vector $\xi \in \mathbb{R}^2_+$, we define the vector $\zeta \in [0,1)^2$ by

$$\zeta_i \;:=\; \xi_i/(1+\xi_i), \quad i = 1,2, \quad \text{and} \quad \Psi^\varepsilon(\zeta) \;:=\; \mathcal{V}^\varepsilon(\xi).$$

This reduces the domain of $\mathcal{V}^\varepsilon$ from $\mathbb{R}^2_+$ to the bounded domain $[0,1)^2$. By changing variables, it is immediately checked that Ψ^ε is a continuous constrained viscosity solution on $[0,1) \times [0,1)$ of

$$\text{(6.1)} \qquad \min_{a \in \mathbb{A}} \left\{ \beta(a)\Psi^\varepsilon(\zeta) - \sum_{i=1}^2 b_i(a,\zeta) \cdot D_i \Psi^\varepsilon(\zeta) - \frac{1}{2} \sum_{i,j=1}^2 \eta_{ij}(a,\zeta) D^2_{ij} \Psi^\varepsilon(\zeta) - g(a) \right\} \;=\; 0,$$

where the control set $\mathbb{A}$ and the expressions of β, $(b_i)_{i=1,2}$, $(\eta_{i,j})_{i,j=1,2}$, and g are obtained by immediate calculation.

6.2. Finite differences for (6.1). We adopt a classical finite difference discretization in order to obtain a numerical scheme for (6.1). Let N be a positive integer, and set $h := \frac{1}{N}$, the finite difference step; we set $e_1 := (1,0)$ and $e_2 := (0,1)$, and we define the uniform grid $\bar{\mathcal{S}}_h := [0,1]^2 \cap (h\mathbb{Z})^2$. We denote by $\zeta^h := (\zeta_1^h, \zeta_2^h)$ a point of the grid $\bar{\mathcal{S}}_h$, and we set $\mathcal{S}_h := (0,1) \times [0,1) \cap (h\mathbb{Z})^2$. In order to define a discretization of (6.1), we approximate the partial derivatives of Ψ^ε by the corresponding backward and forward finite differences

$$\begin{aligned} b_i(a,\zeta)\partial_i \Psi^\varepsilon(\zeta) &\approx \begin{cases} b_i(a,\zeta) D_i^+ \Psi^\varepsilon(\zeta) & \text{if } \; b_i(a,\zeta) \geq 0, \\ b_i(a,\zeta) D_i^- \Psi^\varepsilon(\zeta) & \text{if } \; b_i(a,\zeta) < 0, \end{cases} \\ \partial_{ii}\Psi^\varepsilon(\zeta) &\approx D_i^2 \Psi^\varepsilon(\zeta), \\ \eta_{ij}(a,\zeta)\partial_{ij}\Psi^\varepsilon(\zeta) &\approx \begin{cases} \eta_{ij}(a,\zeta) D_{ij}^+ \Psi^\varepsilon(\zeta) & \text{if } \; \eta_{ij}(a,\zeta) \geq 0, \\ \eta_{ij}(a,\zeta) D_{ij}^- \Psi^\varepsilon(\zeta) & \text{if } \; \eta_{ij}(a,\zeta) < 0, \end{cases} \end{aligned}$$

where the finite difference operators are defined for $i \neq j \in \{1,2\}$ by

$$\begin{aligned} D_i^+ \Psi^\varepsilon(\zeta) &= \frac{\Psi^\varepsilon(\zeta + he_i) - \Psi^\varepsilon(\zeta)}{h}, \quad D_i^- \Psi^\varepsilon(\zeta) = \frac{\Psi^\varepsilon(\zeta) - \Psi^\varepsilon(\zeta - he_i)}{h}, \\ D_i^2 \Psi^\varepsilon(\zeta) &= \frac{\Psi^\varepsilon(\zeta + he_i) - 2\Psi^\varepsilon(\zeta) + \Psi^\varepsilon(\zeta - he_i)}{h^2}, \\ D_{ij}^\pm \Psi^\varepsilon(\zeta) &= \frac{1}{2h^2}\left\{2\Psi^\varepsilon(\zeta) + \Psi^\varepsilon(\zeta + he_i \pm he_j) + \Psi^\varepsilon(\zeta - he_i \mp he_j) \right. \\ &\qquad \left. - \Psi^\varepsilon(\zeta + he_i) - \Psi^\varepsilon(\zeta - he_i) - \Psi^\varepsilon(\zeta + he_j) - \Psi^\varepsilon(\zeta - he_j)\right\}. \end{aligned}$$

In order to compute these differences at every point of $\mathcal{S}_h$, we extend Ψ^ε as follows:

$$\Psi^\varepsilon\left(\zeta_0^h\right) = \Psi^\varepsilon\left(\zeta_0^h + he_1\right), \quad \Psi^\varepsilon\left(\zeta_1^h\right) = \Psi^\varepsilon\left(\zeta_1^h - he_1\right)$$

for $\zeta_0^h \in \{0\} \times [0,1]$ and $\zeta_1^h \in \{1\} \times [0,1]$ and

$$\Psi^\varepsilon\left(\zeta_0^h - he_2\right) = \Psi^\varepsilon\left(\zeta_0^h\right), \quad \Psi^\varepsilon\left(\zeta_1^h\right) = \Psi^\varepsilon\left(\zeta_1^h - he_2\right)$$

for $\zeta_0^h \in [0,1] \times \{0\}$ and $\zeta_1^h \in [0,1] \times \{1\}$. This provides a system of $(N-1)N$ nonlinear equations with the $(N-1)N$ unknowns $\Psi_h^\varepsilon(\zeta^h)$, $\zeta^h \in \mathcal{S}_h$:

$$\min_{a \in \mathbb{A}} \{A_h^a \Psi_h^\varepsilon - g(a)\} = 0. \tag{6.2}$$

6.3. The classical Howard algorithm. To solve (6.2) we adopt the classical Howard algorithm, which can be described as follows:

$$\begin{aligned} \text{Step 0:} \quad & \text{start from an initial value for the control } \; a^0 \in \mathbb{A}, \\ & \Psi_h^0 \; \text{ solution of } \; A_h^{a^0}\varphi - g(a^0) \; = \; 0, \\ \text{Step } k+1,\, k \geq 0\text{:} \quad & \text{find } \; a^{k+1} \; \in \; \operatorname{argmin}_{a \in \mathbb{A}} \left\{A_h^a \Psi_h^k - g(a)\right\}, \\ & \Psi_h^{k+1} \; \text{ solution of } \; A_h^{a^{k+1}}\varphi - g(a^{k+1}) \; = \; 0. \end{aligned}$$

7. Numerical results.

We implement the above numerical algorithm with the following parameters:

$$p = 0.3, \quad \sigma = 0.3, \quad \text{and} \quad \beta = 0.1.$$

We also fix the instantaneous mean return of the risky asset to

$$b := \theta\sigma + r = 0.11.$$

Our numerical experiments showed that by taking a finite difference step $1/20 \le h \le 1/40$ our algorithm converges within a reasonable computation time: convergence error $|\Psi_h^{k+1} - \Psi_h^k| \sim 10^{-6}$, computation time $\sim$ 1–5 minutes.

7.1. Accuracy of the first order Taylor expansion.

It is clear that one should not expect this algorithm to be reliable near the boundary of the grid. However, realistic initial points are far from the boundary, and we expect the error to be small for these points. The main purpose of this subsection is to examine the accuracy of the first order approximation for different sets of parameters r and α:

$$r \in \{0.001\,,\, .01\,,\, .07\} \quad \text{and} \quad \alpha \in \{.001\,,\, .01\,,\, .05\,,\, .1\,,\, .2\,,\, .3\,,\, .36\}\,.$$

Figure 1 plots the mean relative error between the results of the first order expansion and the numerical algorithm over all points of the grid:

$$\frac{1}{N(N-1)} \sum_{i,j} \frac{\left|\mathcal{V}_\varepsilon^h\left(\zeta_{ij}^h\right) - \mathcal{V}^{\rm app}\left(\zeta_{ij}^h\right)\right|}{\mathcal{V}^{\rm app}\left(\zeta_{ij}^h\right)},$$

where $N(N-1)$ is the total number of points in the grid, $\mathcal{V}_\varepsilon^h$ is the approximation of $\mathcal{V}_\varepsilon$ obtained by our numerical scheme, and

$$\mathcal{V}^{\rm app}(\xi_1, \xi_2) := V^{\rm app}(1 - (1-\alpha)\xi_1 - \alpha\xi_2, \xi_1, \xi_2)\,.$$

As expected, the relative error is zero at the origin and increases when the values of the parameters r and α increase. The error size is large due to the boundary effects.

Indeed, in Figure 2 we focus our attention on a region away from the boundary and concentrate on $(y,k) \in [0,1]^2$ (i.e., a region with small initial embedded capital gains). We observe that the average relative error is remarkably small and is of the order of 4% for realistic values of r and α. This figure is our main numerical result, as it shows the reasonable accuracy of the first order Taylor approximation $V^{\rm app}$ of the value function V.

7.2. Welfare analysis.

In view of Remark 4.3, an ε-maximizing strategy is given by the constant portfolio allocation $\bar{\pi}^\alpha$ and the constant consumption-wealth ratio $\bar{c}(r, \tilde{\theta}^\alpha)$. The expected utility realized by following this approximating strategy corresponds to the lower bound $\tilde{V}(z) = \bar{c}(r, \tilde{\theta}^\alpha) z^p/p$ of Proposition 4.2.

In order to compare this approximating strategy to the optimal one, we report in Figures 3 and 4 the welfare cost, z^* such that $V(1-(1-\alpha)\xi_1 - \alpha\xi_2, \xi_1, \xi_2) = \tilde{V}(1+z^*)$, with the following parameters:

$$p = .3, \quad \beta = .1, \quad b := r + \theta\sigma = .11, \quad \sigma = .3, \quad \text{and} \quad r = .07.$$

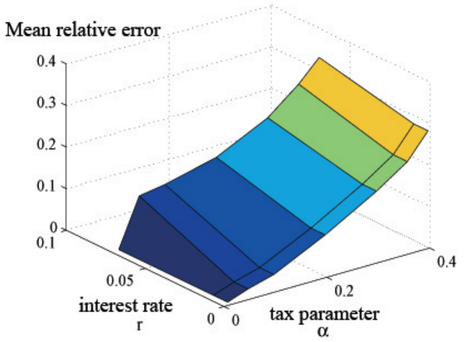

Figure 1. *Mean relative error on* $[0, 10]^2$.

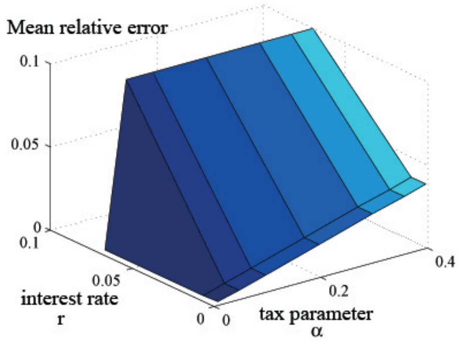

Figure 2. *Mean relative error on* $[0, 1]^2$.

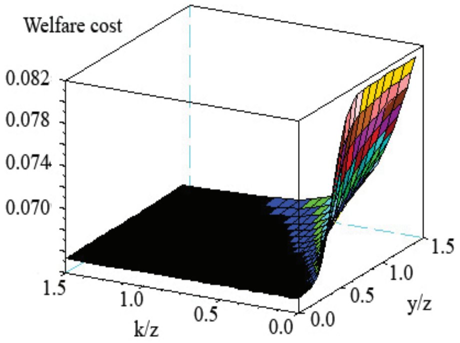

Figure 3. *Welfare cost for* $\alpha = 0.20$.

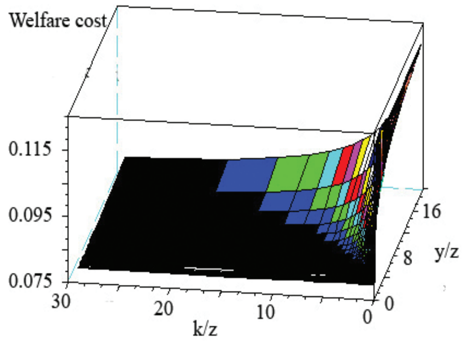

Figure 4. *Welfare cost for* $\alpha = 0.36$.

The welfare cost is nonincreasing with respect to the tax basis and remains relatively small for reasonable values of the parameters α: it reaches a maximum of 8% for $\alpha = 0.2$ and of 12% for $\alpha = 0.36$.

7.3. Optimal consumption-investment strategies. Throughout this subsection we implement our numerical algorithm with the following parameters: $p = 0.3$, $\beta = 0.1$, $b := r + \theta\sigma = 0.11$, $\sigma = 0.3$, and $r = .07$.

The tax-free model. For $\alpha = 0.0$, our algorithm produces the well-known results of the Merton frictionless model. Given the above values of the parameters, Merton's optimal strategy is given by $\bar{\pi} = 0.6349$ and $\bar{c} = 0.1074$.

Figure 5 reports the numerical solution for the function $\mathcal{V}_\varepsilon^h$. We verify that the function $\mathcal{V}_\varepsilon^h$ in this tax-free context does not depend on the variable ξ_2, so that the value function V_ε^h does not depend on the k component. We also see that the value function is concave. Figure 6 reports the optimal investment strategy and produces the expected partition of the state space into three regions:

- The region of no transaction (NT) corresponds to positions such that the proportion of wealth allocated to the risky asset $y/(x+y)$ is equal to $\bar{\pi}$. In this region no position adjustment is considered by the investor.
- The Sell region is where the investor immediately sells risky assets so as to attain the region NT by moving along the ray $(1, -1)$.

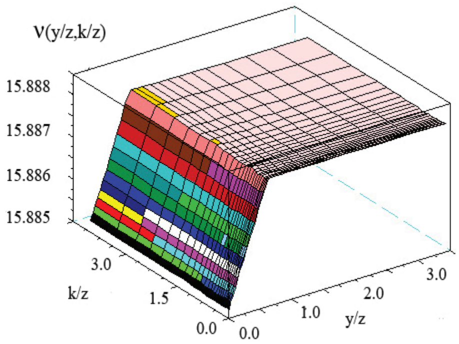

Figure 5. $\mathcal{V}_\varepsilon^h$ *for* $\alpha = 0.0$.

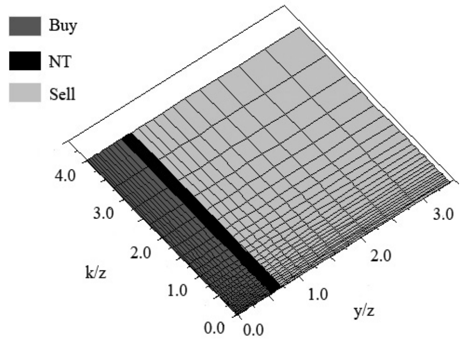

Figure 6. *Partition for* $\alpha = 0.0$.

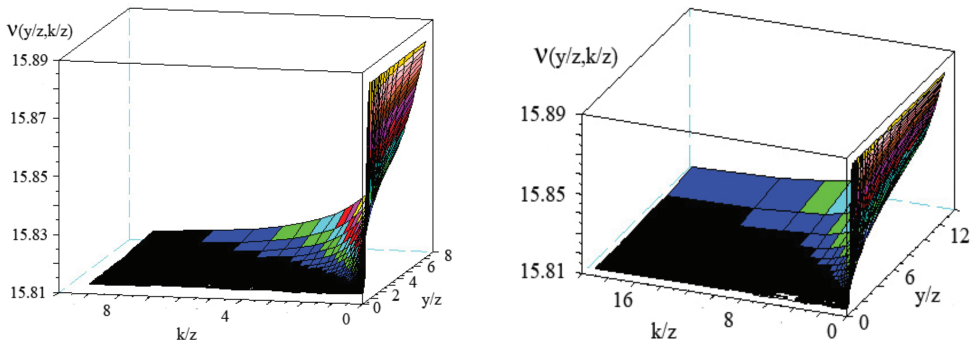

Figure 7. $\mathcal{V}_\varepsilon^h$ *for* $\alpha = 0.20$.

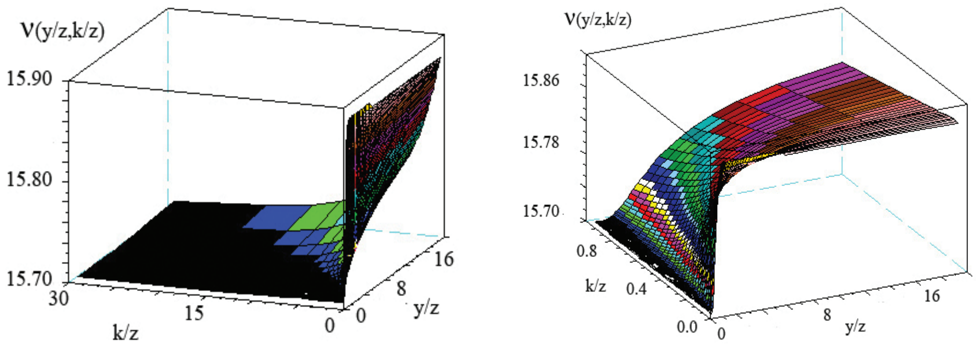

Figure 8. $\mathcal{V}_\varepsilon^h$ *for* $\alpha = 0.36$.

- The Buy region is where the investor immediately purchases risky assets so as to attain the region NT by moving along the ray $(-1, 1)$.

We numerically verify again that this partition is independent of the variable ξ_2.

The value function approximation with taxes. We next concentrate on the case where the tax coefficient is positive. Figures 7 and 8 report the numerical solution for the function $\mathcal{V}_\varepsilon^h$ for $\alpha = 0.2$ and 0.36. The main observation out of these numerical results is that, for a positive tax parameter, the value function is no longer concave. This surprising feature leads

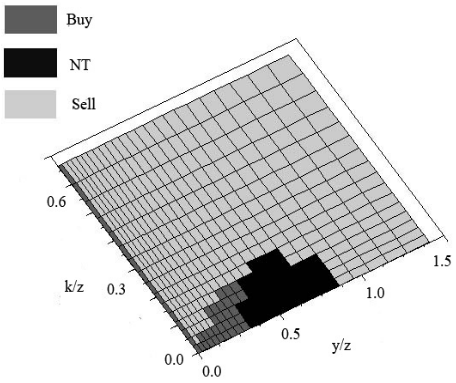

Figure 9. *Partition for* $\alpha = 0.20$.

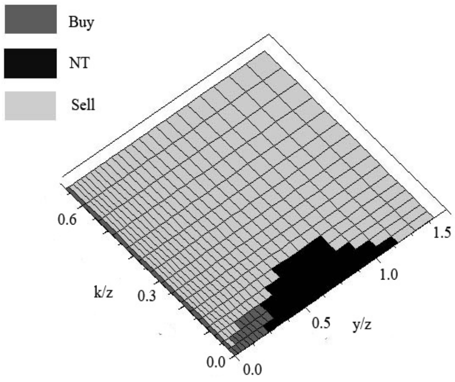

Figure 10. *Partition for* $\alpha = 0.36$.

to mathematical difficulties, as we had to derive the dynamic programming equation without any a priori knowledge that the value function is continuous.

Optimal investment strategy under taxes. Figures 9 and 10 show that, for positive α, the domain is again partitioned into three nonintersecting regions:

- The no-transaction region NT is where no portfolio adjustment is performed by the optimal investor.
- The Sell region is where the investor immediately sells risky assets so as to attain the region NT by moving towards the origin along the ray $((1-\alpha)y + \alpha k, -y, -k) = -[(1-\alpha)y + \alpha k]\mathbf{g}^{\mathbf{s}}$.
- The Buy region is where the investor immediately purchases risky assets so as to attain the region NT by moving along the ray $(-1, 1, 1) = -\mathbf{g}^{\mathbf{b}}$.

For positive α, the boundaries of the no-transaction region depend on the tax basis. The range of the proportion of wealth allocated to the risky asset, (y/z), for which no-transaction is optimal, is very sensible to the values of the tax basis (k/z). Indeed, we observe that the Buy region is limited from the left side by the *wash-sales* region which is part of the Sell region, exactly according to the statement of Proposition 3.5.

We also observe that, for small values of the k variable, the no-transaction region NT contains the Merton optimal portfolio proportions $\bar{\pi}(\sigma, \theta)$ and $\bar{\pi}(\tilde{\sigma}^\alpha, \tilde{\theta}^\alpha)$ corresponding, respectively, to our financial market and to the fictitious financial market with modified parameters.

Optimal consumption strategy under taxes. Figures 11 and 12 report the consumption-wealth ratio for $\alpha = .2$ and $.36$. We notice that this ratio depends on the value of the basis as well as on proportion of wealth allocated to the risky asset. Moreover, in the presence of taxes, on each point of the grid this ratio is higher than Merton's optimal consumption-wealth ratio.

Appendix A. Proof of Proposition 3.4. In order to prove the continuity of V, we follow [5] by introducing the approximation v^λ defined as the value function of the control problem

$$v^\lambda(s) := \sup_{\nu \in \mathcal{A}^\lambda(s)} J^\lambda(s, \nu), \quad \text{where} \quad J^\lambda(s, \nu) := \mathbb{E}\left[\int_0^\infty e^{-\beta t} U(C_t) dt\right], \tag{A.1}$$

and $\mathcal{A}^\lambda(s)$ is the collection of all $\mathbb{F}$-adapted processes $\nu = (C, L, M)$, where C satisfies (2.9),

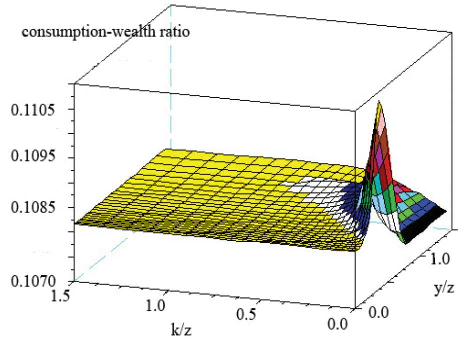

Figure 11. *Consumption for* $\alpha = 0.20$.

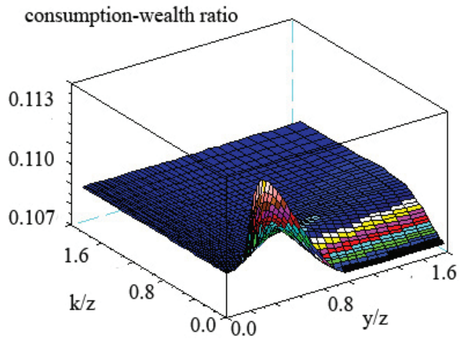

Figure 12. *Consumption for* $\alpha = 0.36$.

L and M are nondecreasing and right continuous, $L_{0-} = M_{0-} = 0$, the jumps of M satisfy (2.6), and the process $S^{s,\nu} = (X^{s,\nu}, Y^{s,\nu}, K^{s,\nu})$ defined by $S^{s,\nu}_{0-} = s$ and the dynamics (2.7), (2.8),

$$dX_t = (rX_t - C_t)\,dt - (1+\lambda)dL_t + [(1-\alpha)Y_{t-} + \alpha K_{t-}]\,dM_t$$

takes values in $\bar{\mathcal{S}}$.

The above control problem corresponds to an optimal consumption investment problem with capital gains taxes and proportional transaction cost $\lambda > 0$ on purchased risky assets. Clearly,

$$v^\lambda \searrow V \quad \text{as} \quad \lambda \searrow 0. \tag{A.2}$$

For later use, we recall the following results from [5].

Theorem A.1. *For $\lambda \geq 0$, the function v^λ is a constrained viscosity solution of*

$$\min\left\{-\mathcal{L}v^\lambda; \mathbf{g}^{\mathbf{b}}_\lambda \cdot Dv^\lambda; \mathbf{g}^{\mathbf{s}} \cdot Dv^\lambda\right\} = 0 \quad on \quad \bar{\mathcal{S}} \setminus \partial^z S \quad and \quad v^\lambda = 0 \quad on \quad \partial^z \mathcal{S}. \tag{A.3}$$

Theorem A.2. *For $\lambda > 0$, let u be an upper semicontinuous viscosity subsolution of (A.3) and v be a lower semicontinuous viscosity supersolution of (A.3), with $(u-v)^+ \in \mathrm{USC}_p(\bar{\mathcal{S}})$. Assume further that $(u-v)(x,0,0) \leq 0$ for all $x \geq 0$. Then $u \leq v$ on $\bar{\mathcal{S}}$.*

We first need to prove the continuity of v^λ.

Lemma A.1. *The function v^λ is continuous on $\bar{\mathcal{S}}$.*

Proof. By Proposition 4.1, the semicontinuous envelopes $v^{\lambda*}$ and v^λ_* satisfy the polynomial growth condition $\left(v^{\lambda*} - v^\lambda_*\right)^+ \in \mathrm{USC}_p(\bar{\mathcal{S}})$. We also know from Theorem A.1 that they are, respectively, a constrained subsolution and supersolution of (5.1). We now claim that

$$\left(v^\lambda_* - v^{\lambda*}\right)(x,0,0) = 0 \quad \text{for all} \quad x \geq 0, \tag{A.4}$$

so that $v^\lambda_* \geq v^{\lambda*}$ by the comparison result of Theorem A.2, and therefore $v^\lambda_* = v^{\lambda*}$ since the reverse inequality holds by definition.

It remains to prove (A.4). Notice that for all $s = (x,y,k) \in \bar{\mathcal{S}}$ and $z := x + (1-\alpha)y + \alpha k$

$$v^\lambda(z,0,0) \;\leq\; v^\lambda(s) \;\leq\; v^\lambda(z+y,0,0)\,. \tag{A.5}$$

Before proving these inequalities, let us complete the proof of $v_*^\lambda = v^{\lambda *}$ on $\{(x,0,0) : x \geq 0\}$. For an arbitrary $x \in \mathbb{R}_+$, let $\{s_n = (x_n, y_n, k_n),\ n \geq 1\}$, $\{s'_n = (x'_n, y'_n, k'_n),\ n \geq 1\}$ be two sequences in $\bar{\mathcal{S}}$ such that

$$s_n, s'_n \xrightarrow[n\to\infty]{} (x,0,0), \quad v^\lambda(s_n) \xrightarrow[n\to\infty]{} v_*^\lambda(x,0,0), \quad \text{and} \quad v^\lambda(s'_n) \xrightarrow[n\to\infty]{} v^{\lambda *}(x,0,0).$$

By (A.5), together with the homotheticity property of Proposition 3.3, we see that

$$\begin{aligned} v^\lambda(s'_n) \ &\leq\ v^\lambda(z'_n + y'_n, 0, 0) = \left(z'_n + y'_n\right)^p v^\lambda(1,0,0), \\ v^\lambda(s_n) \ &\geq\ v^\lambda(z_n, 0, 0) = (z_n)^p\, v^\lambda(1,0,0), \end{aligned}$$

where $z_n = x_n + (1-\alpha)y_n + \alpha k_n$ and $z'_n = x'_n + (1-\alpha)y'_n + \alpha k'_n$. Letting $n \to \infty$ in the above inequalities and recalling that $z_n, z'_n + y'_n \to x$, we get the required result.

We now turn to the proof of (A.5).

- The left-hand side of (A.5) holds since for each consumption-investment strategy $\nu = (C, L, M) \in \mathcal{A}(z,0,0)$ the strategy $\bar{\nu} := \nu + \{1 - \Delta M_0\}\,(0,0,1) \in \mathcal{A}(s)$.
- The right-hand side of (A.5) holds since for each $\nu = (C, L, M) \in \mathcal{A}(s)$ the strategy $\bar{\nu} := \nu + \{y(1-\Delta M_0)\}\,(0,1,0) \in \mathcal{A}(\bar{s})$, where $\bar{s} := (z+y, 0, 0)$.

Indeed, since ν and $\bar{\nu}$ differ only by the jump at time $t=0$, the dynamics of the state processes $S^{s,\nu}$ and $S^{\bar{s},\bar{\nu}}$ are such that $Y_0^{s,\nu} = Y_0^{\bar{s},\bar{\nu}}$ and therefore $Y_t^{s,\nu} = Y_t^{\bar{s},\bar{\nu}}$ for $t \geq 0$, and

$$\begin{aligned} K_t^{\bar{s},\bar{\nu}} - K_t^{s,\nu} &= \left(K_0^{\bar{s},\bar{\nu}} - K_0^{s,\nu}\right) e^{-M_t^c} \prod_{0\leq s\leq t} (1 - \Delta M_s) \\ &\leq \left(K_0^{\bar{s},\bar{\nu}} - K_0^{s,\nu}\right)^+ \ =\ (y-k)^+(1-\Delta M_0). \end{aligned}$$

Then the corresponding liquidation value processes $Z^{s,\nu}$ and $Z^{\bar{s},\bar{\nu}}$ are such that

$$\begin{aligned} Z_t^{\bar{s},\bar{\nu}} - Z_t^{s,\nu} &= e^{rt}\left\{ Z_0^{\bar{s},\bar{\nu}} - Z_0^{s,\nu} - \alpha \int_0^t e^{-rs}\left(K^{s,\nu} - K^{\bar{s},\bar{\nu}}\right)_s ds \right\} \\ &\geq e^{rt}\left\{ Z_0^{\bar{s},\bar{\nu}} - Z_0^{s,\nu} - (K_0^{s,\nu} - K_0^{\bar{s},\bar{\nu}})^+ \right\} \\ &= e^{rt}\left\{ y - (y-k)^+(1-\Delta M_0) \right\} \ \geq\ 0. \end{aligned}$$

It follows that $Z^{\bar{s},\nu} \geq 0$; hence $\bar{\nu} \in \mathcal{A}(\bar{s})$. ■

Proof of Proposition 3.4. Since $\left(v^\lambda\right)_{\lambda>0}$ is a nonincreasing sequence of continuous functions (Lemma A.1) converging to V as $\lambda \searrow 0$, it follows that the function V is lower semicontinuous.

Let $\mathcal{V}$ be the lower semicontinuous function defined on $\mathbb{R}_+^2$ by

$$\mathcal{V}(\xi,\zeta) := V\left(1 - (1-\alpha)\xi + \alpha\zeta, \xi, \zeta\right), \tag{A.6}$$

so that

$$V(x,y,k) \ =\ z^p\, \mathcal{V}\left(\frac{y}{z}, \frac{k}{z}\right), \quad \text{with} \quad z = x + (1-\alpha)y + \alpha k, \tag{A.7}$$

by the homotheticity property of V stated in Proposition 3.3. By Theorem A.1, we have $\mathbf{g^b} \cdot DV \geq 0$ and $\mathbf{g^s} \cdot DV \geq 0$ in the viscosity sense. By a direct change of variables, this implies that $\mathcal{V}$ is a lower semicontinuous viscosity supersolution of the equation

$$p\mathcal{V} - (\xi\mathcal{V}_\xi + \zeta\mathcal{V}_\zeta) - \varepsilon^{-1}\left(\mathcal{V}_\xi + \mathcal{V}_\zeta\right) \ \geq\ 0 \quad \text{and} \quad \xi\mathcal{V}_\xi + \zeta\mathcal{V}_\zeta \ \geq\ 0.$$

Also, from the monotonicity of V in x, y, and k, it follows that $\mathcal{V}$ is a lower semicontinuous viscosity supersolution of the equation

$$p\mathcal{V} - (\xi\mathcal{V}_\xi + \zeta\mathcal{V}_\zeta) + \min\left\{0\ ,\ \frac{1}{1-\alpha}\mathcal{V}_\xi\ ,\ \frac{1}{\alpha}\mathcal{V}_\zeta\right\}\ \geq\ 0.$$

Observe that $\mathcal{V}$ is bounded as a consequence of the upper bound provided in Proposition 4.1. We then deduce from the above viscosity supersolution properties that $-|\nabla\mathcal{V}| \geq -A$ on $(0,\infty)^2$, in the viscosity sense, for some constant A. Hence $\mathcal{V}$ is Lipschitz continuous. ∎

Appendix B. Proof of Proposition 4.2.

Preliminaries and notation. For $s \in \bar{\mathcal{S}}$ and $\nu \in \mathcal{A}(s)$, the process $Z^{s,\nu}$ is defined by the initial condition $Z_0^{s,\nu} = z := x + (1-\alpha)y + \alpha k$ and the dynamics

$$dZ_t^{s,\nu} = (rZ_t^{s,\nu} - C_t)\,dt + Y_t^{s,\nu}\tilde{\sigma}^\alpha\left(\tilde{\theta}^\alpha dt + dW_t\right) + r\alpha Y_t^{s,\nu}\left(1 - \frac{B_t^{s,\nu}}{P_t}\right)dt,$$

where

$$\tilde{\sigma}^\alpha\ :=\ (1-\alpha)\sigma \quad \text{and} \quad \tilde{\theta}^\alpha\ :=\ \theta - \frac{r\alpha}{\tilde{\sigma}^\alpha}.$$

Our purpose is to show that the value function V outperforms the maximal utility achieved in a frictionless financial market consisting of one bank account with the constant interest rate r and one risky asset with price process P^α given by

$$dP_t^\alpha\ =\ P_t^\alpha\left[rdt + \tilde{\sigma}^\alpha(\tilde{\theta}^\alpha dt + dW_t)\right] \quad \text{and} \quad \tilde{P}_0^\alpha\ =\ P_0.$$

From Theorem 2.1, the solution of the optimal consumption-investment problem with price process P^α is given by the constant controls $\bar{c}^\alpha := \bar{c}(r, \tilde{\theta}^\alpha)$ and $\bar{\pi}^\alpha := \bar{\pi}(\tilde{\sigma}^\alpha, \tilde{\theta}^\alpha)$ and the corresponding optimal wealth process is defined by

$$\bar{Z}^\alpha = z \quad \text{and} \quad d\bar{Z}_t^\alpha\ =\ \bar{Z}^\alpha\left[(r - \bar{c}^\alpha)dt + \bar{\pi}^\alpha\tilde{\sigma}^\alpha\left(\tilde{\theta}^\alpha dt + dW_t\right)\right].$$

In order to prove the required result, we shall fix an arbitrary maturity $T > 0$ and construct a sequence of admissible strategies $\hat{\nu}^{T,n}$ such that

$$V(s) \geq \lim_{n\to\infty} J_T\left(s, \hat{\nu}^{T,n}\right)\ =\ \mathbb{E}\left[\int_0^T e^{-\beta t}U\left(\bar{c}\bar{Z}_t^\alpha\right)dt\right].$$

Then, the required result follows by sending T to infinity in this inequality.

A sequence of strategies tracking the Merton optimal policy. Let $T > 0$ be a fixed maturity. We construct a sequence of consumption-investment strategies $\hat{\nu}^{T,n,k}$ by forcing the tax basis B to be close to the spot price and by *tracking* Merton's optimal strategy, i.e., keeping the proportion of wealth invested in the risky asset and the proportion of wealth dedicated for consumption

$$\pi_t\ :=\ \frac{Y_t}{Z_t}\mathbf{1}_{\{Z_t\neq 0\}} \quad \text{and} \quad c_t\ :=\ \frac{C_t}{Z_t}\mathbf{1}_{\{Z_t\neq 0\}}, \quad 0 \leq t \leq T,$$

close to the pair $(\bar{\pi}^\alpha, \bar{c}^\alpha)$.

To do this, we define a convenient sequence $(\nu^{T,n,m})_{n,m\geq 1} := (C^{T,n,m}, L^{T,n,m}, M^{T,n,m})_{n,m\geq 1}$ for all $s = (x, y, k) \in \bar{\mathcal{S}}$. We shall denote by

$$\left(X^{T,n,m}, Y^{T,n,m}, K^{T,n,m}\right) = \left(X^{\nu^{T,n,m}}, Y^{\nu^{T,n,m}}, K^{\nu^{T,n,m}}\right)$$

the corresponding state processes and by $Z^{T,n,m} = Z^{\nu^{T,n,m}}$ the corresponding after-tax liquidation value process. For all integers $n \geq 1$ and $m \geq 1$, the consumption-investment strategy $\nu^{T,n,m}$ is defined as follows:

1. The consumption strategy is defined by $C_t^{T,n,m} := \bar{c}^{\alpha} Z_t^{T,n,m}$ for $0 \leq t \leq T$. The investment strategy is piecewise constant:

$$dL_t^{T,n,m} = dM_t^{T,n,m} = 0 \quad \text{for all} \quad t \in (0, T] \setminus \{\tau_j^{T,n,m}, j \leq m\},$$

where the sequence of stopping times $(\tau_j^{T,n,m})_{j\leq m}$ is defined in step 3.

2. At time 0, set $\Delta L_0^{T,n,m} := \bar{\pi}^{\alpha} z$ and $\Delta M_0^{T,n,m} := 1$, so that

$$K_0^{T,n,m} = Y_0^{T,n,m}, \quad \pi_0^{T,n,m} = \bar{\pi}^{\alpha}, \quad \text{and} \quad Z_0^{T,n,m} = z.$$

3. We now introduce the sequence of stopping times $\tau_j^{T,n,m}$ as the hitting times of the pair process $(\pi^{T,n,m}, \frac{K^{T,n,m}}{Y^{T,n,m}})$ of some barrier close to $(\bar{\pi}, 1)$. Set

$$\tau_0^{T,n,m} := 0 \quad \text{and} \quad \tau_j^{T,n,m} := T \wedge \tau_j^{\pi} \wedge \tau_j^{B} \quad \text{for} \quad 1 \leq j \leq m,$$

where

$$\tau_j^{\pi} := \inf\left\{t \geq \tau_{j-1}^{T,n,m} \ : \ |\pi_t^{T,n,m} - \bar{\pi}^{\alpha}| \ > \ n^{-1}\bar{\pi}^{\alpha}\right\},$$
$$\tau_j^{B} := \inf\left\{t \geq \tau_{j-1}^{T,n,m} \ : \ \left|1 - \frac{K^{T,n,m}}{Y^{T,n,m}}\right| \ > \ n^{-1}\right\}.$$

4. Finally, we specify the jumps $\left(\Delta L^{T,n,m}, \Delta M^{T,n,m}\right)$ at each time $\tau_j^{T,n,m}$, $j \geq 1$, by

$$\Delta L_t^{T,n,m} := \bar{\pi}^{\alpha} Z_t^{T,n,m} \quad \text{and} \quad \Delta M_t^{T,n,m} := 1 \quad \text{for} \quad t \in \{\tau_j^{T,n,m}, j < m\},$$

so that

$$\pi_t^{T,n,m} = \bar{\pi}^{\alpha} \quad \text{and} \quad B_t^{T,n,m} = P_t \quad \text{for} \quad t \in \{\tau_j^{T,n,m}, j < m\}.$$

And at time $t = \tau_m^{T,n,m}$, set $\Delta L_t^{T,n,m} := 0$ and $\Delta M_t^{T,n,m} = 1$, so that all the wealth is transferred to the bank:

$$Y_t^{T,n,m} = K_t^{T,n,m} = 0 \quad \text{and} \quad X_t^{T,n,m} = Z_{\tau_m^{T,n,m}}^{T,n,m} e^{r(T-\tau_m^{T,n,m})} \quad \text{for} \quad \tau_m^{T,n,m} \leq t \leq T.$$

Lemma B.1. *For each $n \geq 1$, the sequence $\left(\tau_m^{T,n,m}\right)_{m\geq 0}$ converges to T $\mathbb{P}$-a.s.*

Proof. Let $n \geq 1$. In order to alleviate the notation, we shall denote $\tau_m := \tau_m^{T,n,m}$, $\pi := \pi^{T,n,m}$, and $(Y, Z, K) := \left(Y^{T,n,m}, Z^{T,n,m}, K^{T,n,m}\right)$.

1. In a first step, we shall prove that $\tau_m^\pi \to \infty$ a.s. For sufficiently large n, observe that our sequence of stopping times $(\tau_m^\pi)_{m\geq 1}$ is equivalently defined by

$$\tau_{m+1}^\pi := \inf\left\{t \geq \tau_m \;:\; |\tilde{\pi}_t - \bar{\pi}^\alpha| \geq \bar{\pi}^\alpha/n\right\},$$

where $\tilde{\pi}$ is defined by

$$\tilde{\pi}_t = \pi_{t\wedge\theta}, \quad \text{with} \quad \theta := \inf\left\{t \geq \tau_m \;:\; |\pi_t - \bar{\pi}^\alpha| \geq \frac{\bar{\pi}^\alpha}{2} \text{ and } \left|\frac{K_t}{Y_t} - 1\right| \geq \frac{1}{2}\right\}.$$

Notice that $d\tilde{\pi}_t = f(u)du + g(u)dW_u$, $t \geq \tau_m$, with uniformly bounded processes f and g. We then estimate

$$\begin{aligned}
\mathbb{P}\left[\tau_{m+1}^\pi \leq \tau_m + \frac{1}{m}\right] &\leq \mathbb{P}\left[\sup_{\tau_m \leq t \leq \tau_m + \frac{1}{m}} |\tilde{\pi}_t - \bar{\pi}^\alpha| \geq \bar{\pi}^\alpha n^{-1}\right] \\
&\leq \left(\frac{n}{\bar{\pi}^\alpha}\right)^4 \mathbb{E}\left[\left(\sup_{\tau_m \leq t \leq \tau_m + \frac{1}{m}} |\tilde{\pi}_t - \bar{\pi}^\alpha|\right)^4\right] \\
&\leq C\left(\frac{1}{m^4} + \frac{1}{m^2}\right) \quad \text{for some} \quad \text{positive constant } C,
\end{aligned}$$

where we used the Chebyshev and the Burkholder–Davis–Gundy inequalities.

2. Following the same reasoning, we prove that $\mathbb{P}\left[\tau_{m+1}^B \leq \tau_m + \frac{1}{m}\right] \leq C\left(\frac{1}{m^4} + \frac{1}{m^2}\right)$, and we conclude that $\sum_{m<\infty} \mathbb{P}\left[\tau_{m+1} \leq \tau_m + \frac{1}{m}\right] < \infty$. Then, by the Borel–Cantelli lemma, $\mathbb{P}\left[\limsup_m \{\tau_{m+1} \leq \tau_m + \frac{1}{m}\}\right] < 0$, which implies that $\lim_{m\to\infty} \tau_m = \infty$. ∎

Lemma B.2. *For each integer n, we have $\nu^{T,n,m} \in \mathcal{A}(s)$.*

Proof. By (4.1), we have

$$dZ_t^{T,n,m} = Z_t^{T,n,m}\left[(r - \bar{c}^\alpha)\,dt + \pi_t^{T,\nu,m}\tilde{\sigma}^\alpha\left(\tilde{\theta}^\alpha dt + dW_t\right) + r\alpha\pi_t^{T,n,m}\left(1 - B_t^{T,n}\right)dt\right].$$

Also, $0 < \left(1 - n^{-1}\right)\bar{\pi}^\alpha \leq \pi_t^{T,n,m} \leq \left(1 + n^{-1}\right)\bar{\pi}^\alpha$. In particular, the process $\pi^{T,n,m}$ is bounded, so that the above dynamics implies that the process $Z^{T,n,m}$ is positive, and $Y_t^{T,n,m} = \pi_t^{T,n,m} Z_t^{T,n,m} > 0$ $\mathbb{P}$-a.s. ∎

The convergence result.

Lemma B.3. *There is a constant A depending on T such that*

$$\mathbb{E}\left[\sup_{0\leq t\leq T}\left|Z_t^{T,n,m} - \bar{Z}_t^\alpha\right|^2\right] \leq \left(n^{-2} + \mathbb{E}[T - \tau_m^{T,n,m}]\right) Ae^{AT}.$$

Proof. By definition of the sequence of consumption-investment strategies $\left(\nu^{T,n,m}\right)$, we have

$$\sup_{0\leq t\leq T\wedge\tau^{T,n,m}}\left|\pi_t^{T,n} - \bar{\pi}^\alpha\right| \;\leq\; \frac{1}{n}\bar{\pi}^\alpha \quad \text{and} \quad \sup_{0\leq t\leq T\wedge\tau^{T,n,m}}\left|1 - \frac{K_t^{T,n,m}}{Y_t^{T,n,m}}\right| \;\leq\; \frac{1}{n}. \tag{B.1}$$

By direct computation, we decompose the difference $Z^{T,n,m} - \bar{Z}^{\alpha}$ into

$$D_t \;:=\; Z_t^{T,n,m} - \bar{Z}_t^{\alpha} = F_t + G_t + H_t,$$

where

$$F_t := \int_0^t D_t \left[(r - \bar{c}^{\alpha})du + \pi_u^{T,n,m} \left(\tilde{\sigma}^{\alpha}\tilde{\theta}^{\alpha}du + \alpha r \left(1 - \frac{K_u^{T,n,m}}{Y_u^{T,n,m}} \right) du + \tilde{\sigma}^{\alpha} dW_u \right) \right],$$

$$G_t := \int_0^t \bar{Z}_u^{\alpha}\tilde{\sigma}^{\alpha} \left(\pi_t^{T,n,m} - \bar{\pi}^{\alpha} \right) \left(\theta^{\alpha} du + dW_u \right),$$

$$H_t := \alpha r \int_0^t \pi_u^{T,n,m} \bar{Z}_u^{\alpha} \left(1 - \frac{K_u^{T,n,m}}{Y_u^{T,n,m}} \right) du.$$

In the subsequent calculation, A will denote a generic (T-dependent) constant whose value may change from line to line. We shall also denote $V_t^* := \sup_{0 \le u \le t} |V_u|$ for all processes $(V_t)_t$.

We first start by estimating the first component F. Observe that the process $\pi^{T,n,m}$ is bounded by $2\bar{\pi}$. Then

$$|F_t|^2 \le A \int_0^t |D_u^*|^2 \, du + 2 \left(\int_0^t D_u \pi_u^{t,n,m} \tilde{\kappa} dW_u \right)^2.$$

By the Burkholder–Davis–Gundy inequality, this provides

$$\mathbb{E}\,|F_t^*|^2 \le A \int_0^t \mathbb{E}\,|D_u^*|^2 \, du.$$

By a similar calculation, it follows from the boundedness of $\pi^{T,n,m}$ and from (B.1) that

$$\mathbb{E}|G_t^*|^2 \;\le\; A \left(n^{-2} + \mathbb{E}[T - \tau_m^{T,n,m}] \right) \quad \text{and} \quad \mathbb{E}|H_t^*|^2 \;\le\; A \left(n^{-2} + \mathbb{E}[T - \tau_m^{T,n,m}] \right).$$

Collecting the above estimates, we see that

$$E|D_t^*|^2 \;\le\; A \left(n^{-2} + \mathbb{E}[T - \tau_m^{T,n,m}] \right) + K \int_0^t \mathbb{E}|D_u^*|^2 du \quad \text{for all} \quad t \le T,$$

and we obtain the required result by the Gronwall inequality. ■

B.1. Proof of Proposition 4.2. For $s = (x, y, k) \in \bar{\mathcal{S}}$, and $T > 0$,

$$\left| J_T(s, \nu^{T,n,m}) - \int_0^T e^{-\beta t} U\left(\bar{c}^{\alpha} \bar{Z}_t^{\alpha} \right) dt \right| = \left| \int_0^T e^{-\beta t} \left(U\left(\bar{c}^{\alpha} Z_t^{T,n,m} \right) - U\left(\bar{c}^{\alpha} \bar{Z}_t^{\alpha} \right) dt \right) \right|$$
$$\le A \int_0^T e^{-\beta t} \left| Z_t^{T,n,m} - \bar{Z}_t^{\alpha} \right|^p dt$$

for some positive constant A. Now, by Lemma B.1 and the estimate of Lemma B.3, it follows that

$$\lim_{n,m \to \infty} J_T(s, \nu^{T,n,m}) = \int_0^T e^{-\beta t} U\left(\bar{c}^{\alpha} \bar{Z}_t^{\alpha} \right) dt.$$

Since $V(s) \ge J_T(s, \nu^{T,n,m})$ for every $T > 0$, this implies that

$$V(s) \ge \lim_{T \to \infty} \int_0^T e^{-\beta t} U\left(\bar{c}^{\alpha} \bar{Z}_t^{\alpha} \right) dt \;=\; \gamma\left(r, \tilde{\theta}^{\alpha} \right) \frac{z^p}{p}. \qquad ■$$

Acknowledgments. We are grateful to Stathis Tompaidis for numerous discussions on the modeling issue of this paper. We have also benefited from interesting comments and discussions with Victor DeMiguel, Chris Rogers, Jose Scheinkman, and Raman Uppal.

REFERENCES

[1] M. Akian, J. L. Menaldi, and A. Sulem, *On an investment-consumption model with transaction costs*, SIAM J. Control Optim., 34 (1996), pp. 329–364.

[2] G. Barles and E. R. Jakobsen, *On the convergence rate of approximation schemes for Hamilton-Jacobi-Bellman equations*, M2AN Math. Model. Numer. Anal., 36 (2002), pp. 33–54.

[3] G. Barles and H. M. Soner, *Option pricing with transaction costs and a nonlinear Black-Scholes equation*, Finance Stoch., 2 (1998), pp. 369–397.

[4] G. Barles and P. E. Souganidis, *Convergence of approximation schemes for fully nonlinear equations*, Asymptotic Anal., 4 (1991), pp. 271–283.

[5] I. Ben Tahar, H. M. Soner, and N. Touzi, *The dynamic programming equation for the problem of optimal investment under capital gains taxes*, SIAM J. Control Optim., 46 (2007), pp. 1779–1801.

[6] J. F. Bonnans and H. Zidani, *Consistency of generalized finite difference schemes for the stochastic HJB equation*, SIAM J. Numer. Anal., 41 (2003), pp. 1008–1021.

[7] A. Cadenillas and S. Pliska, *Optimal trading of a security when there are taxes and transaction costs*, Finance Stoch., 3 (1999), pp. 137–165.

[8] G. M. Constantinides, *Capital market equilibrium with personal taxes*, Econometrica, 51 (1983), pp. 611–636.

[9] G. M. Constantinides and M. J. P. Magill, *Portfolio selection with transaction costs*, J. Econom. Theory, 13 (1976), pp. 245–263.

[10] J. Cox and C. F. Huang, *Optimal consumption and portfolio policies when asset prices follow a diffusion process*, J. Econom. Theory, 49 (1989), pp. 33–83.

[11] J. Cvitanić and I. Karatzas, *Convex duality in constrained portfolio optimization*, Ann. Appl. Probab., 2 (1992), pp. 767–818.

[12] R. M. Dammon, C. S. Spatt, and H. H. Zhang, *Optimal consumption and investment with capital gains taxes*, Rev. Financ. Stud., 14 (2001), pp. 583–616.

[13] M. H. A. Davis and A. R. Norman, *Portfolio selection with transaction costs*, Math. Oper. Res., 15 (1990), pp. 676–713.

[14] V. DeMiguel and R. Uppal, *Portfolio investment with the exact tax basis via nonlinear programming*, Management Sci., 51 (2005), pp. 277–290.

[15] D. Duffie and T. Sun, *Transaction costs and portfolio choice in a discrete-continuous-time setting*, J. Econom. Dynam. Control, 14 (1990), pp. 35–51.

[16] P. Dybvig and H. Koo, *Investment with Taxes*, Working paper, Washington University, St. Louis, MO, 1996.

[17] A. Fahim, N. Touzi, and X. Warin, *A Probabilistic Numerical Method for Fully Nonlinear Parabolic PDEs*, preprint, 2009.

[18] M. Gallmeyer, R. Kaniel, and S. Tompaidis, *Tax management strategies with multiple risky assets*, J. Finan. Econ., 80 (2006), pp. 243–291.

[19] E. Jouini, P.-F. Koehl, and N. Touzi, *Optimal investment with taxes: An existence result*, J. Math. Econom., 33 (2000), pp. 373–388.

[20] E. Jouini, P.-F. Koehl, and N. Touzi, *Optimal investment with taxes: An optimal control problem with endogenous delay*, Nonlinear Anal., 37 (1999), pp. 31–56.

[21] I. Karatzas, J. P. Lehoczky, and S. E. Shreve, *Optimal portfolio and consumption decisions for a "small investor" on a finite horizon*, SIAM J. Control Optim., 25 (1987), pp. 1557–1586.

[22] N. V. Krylov, *On the rate of convergence of finite-difference approximations for Bellman's equations*, St. Petersburg Math. J., 9 (1997), pp. 245–256.

[23] N. V. Krylov, *On the rate of convergence of finite-difference approximations for Bellman's equations with Lipschitz coefficients*, Appl. Math. Optim., 3 (2005), pp. 365–399.

[24] H. E. Leland, *Optimal Portfolio Management with Transaction Costs and Capital Gains Taxes*, preprint, 1999.
[25] R. Mehra and E. Prescott, *The equity premium: A puzzle*, J. Monet. Econ., 15 (1985), pp. 145–161.
[26] R. C. Merton, *Lifetime portfolio selection under uncertainty: The continuous-time model*, Rev. Econ. Statist., 51 (1969), pp. 247–257.
[27] R. C. Merton, *Optimum consumption and portfolio rules in a continuous-time model*, J. Econom. Theory, 3 (1971), pp. 373–413.
[28] S. E. Shreve and H. M. Soner, *Optimal investment and consumption with transaction costs*, Ann. Appl. Probab., 4 (1994), pp. 609–692.
[29] H. M. Soner, *Optimal control with state-space constraint* I, SIAM J. Control Optim., 24 (1986), pp. 552–561.
[30] H. M. Soner, *Optimal control with state-space constraint* II, SIAM J. Control Optim., 24 (1986), pp. 1110–1122.
[31] A. Tourin and T. Zariphopoulou, *Numerical schemes for investment models with singular transactions*, Comput. Econom., 7 (1994), pp. 287–307.