

FUTURES TRADING WITH TRANSACTION COSTS

KAREL JANEČEK AND STEVEN E. SHREVE

ABSTRACT. A model for optimal consumption and investment is posed whose solution is provided by the classical Merton analysis when there is zero transaction cost. A probabilistic argument is developed to identify the loss in value when a proportional transaction cost is introduced. There are two sources of this loss. The first is a loss due to “displacement” that arises because one cannot maintain the optimal portfolio of the zero-transaction-cost problem. The second loss is due to “transaction,” a loss in capital that occurs when one adjusts the portfolio. The first of these increases with increasing tolerance for departure from the optimal portfolio in the zero-transaction-cost problem, while the second decreases with increases in this tolerance. This paper balances the marginal costs of these two effects. The probabilistic analysis provided here complements earlier work on a related model that proceeded from a viscosity solution analysis of the associated Hamilton–Jacobi–Bellman equation.

1. Introduction

The underlying risky asset in this paper is a futures contract. Investing in futures is different from investing in stocks because the value of a futures contract is reset to zero by marking to market at the end of each trading day. With a stock, the share price determines the amount of capital an investor must commit to trade in the asset, and the *relative changes* in the share price

Received April 1, 2010; received in final form December 16, 2010.

The work of the second author was supported by the National Science Foundation under Grants DMS-04-04682 and DMS-09-03475. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. Part of this work was done while this author was visiting the Isaac Newton Institute for Mathematical Sciences, Cambridge, UK with support under EPSRC Grant N09176.

2010 *Mathematics Subject Classification.* 90A09, 60H30, 60G44.

determine the return on investment. A geometric Brownian motion model is frequently used for share prices so that these relative changes are normally distributed. With a futures contract, the investor does not commit capital to trade (the margin account that is set up is to guarantee credit-worthiness, not to pay a purchase price), and hence the *absolute changes* in the futures price determine the investor's profits and losses. To capture the preeminence of absolute changes in the futures price, we model the futures price as arithmetic rather than geometric Brownian motion. Our model for the futures price is thus

$$(1.1) \quad F(t) = F(0) + \alpha t + \sigma W(t),$$

where $F(0)$ and α are constants, σ is a positive constant, and W is a standard Brownian motion under a (physical) measure \mathbb{P} . We assume that $\alpha \neq 0$ in order to achieve a nontrivial solution. More precisely, in this paper we assume $\alpha > 0$; the results for $\alpha < 0$ are obtained by symmetry.

Consider an agent with initial capital $X(0) > 0$ who invests in a money market and takes positions in futures contracts on some asset or index. Let $X(t)$ denote the wealth of the agent at time t , all of which is held in a money market account with constant rate of interest $r > 0$. At each time t , the agent consumes at rate $C(t) \geq 0$ per unit time. In addition, the agent may take any long or short position in futures contracts by paying a small transaction cost $\lambda > 0$ times the size of the trade required to attain the position. In practice, entering, adjusting, or closing a futures position is costless except for money lost due to the bid-ask spread and other transaction fees. For large traders, these costs are proportional to trade size.

Consider a one-parameter class of utility functions defined for $C \geq 0$ by

$$(1.2) \quad U_p(C) = \begin{cases} \frac{1}{1-p} C^{1-p} & \text{if } p > 0, p \neq 1, \\ \log C & \text{if } p = 1. \end{cases}$$

For $p \geq 1$, we mean that $U_p(0) = -\infty$. Let $\beta > 0$ be a positive discount factor chosen so that

$$(1.3) \quad A(p) \triangleq \frac{\beta - r(1-p)}{p} - \frac{\alpha^2(1-p)}{2\sigma^2 p^2} > 0.$$

The value function for the agent's utility maximization problem is

$$(1.4) \quad v(x, y) \triangleq \sup \mathbb{E} \int_0^\infty e^{-\beta t} U_p(C(t)) dt,$$

where the supremum is taken over consumption and investment strategies that ensure that the agent is solvent at all times, that is, at each time the agent would have nonnegative wealth if he closed out his futures position.

This is an arithmetic Brownian motion version of the classical transaction cost problem posed by Magill and Constantinides [12], solved under restrictive assumptions by Davis and Norman [6], and thoroughly studied by Shreve and

Soner [17]. If λ were zero, this problem could be solved by the method due to Merton [14] and outlined in Section 3.4 below, and the optimal trading strategy would keep the position in futures divided by total wealth at the constant value

$$(1.5) \quad \bar{\theta} \triangleq \frac{\alpha}{\sigma^2 p}.$$

When λ is positive one should instead keep this ratio in an interval $[z_1^*, z_2^*]$, trading just enough to prevent the ratio from exiting the interval. This result has been obtained rigorously for the geometric Brownian motion model; see [6], [17]. The argument for the arithmetic Brownian motion model is not provided here, in part because it parallels the arguments in the cited papers. Our purpose is not to imitate those earlier works, which are based on analyses of the associated Hamilton–Jacobi–Bellman partial differential equation, but rather to take the form of the optimal solution as given and provide a purely probabilistic derivation, based on balancing the two costs discussed below, of the values of z_1^* and z_2^* . One cannot analytically solve for z_1^* and z_2^* , but it is possible to conduct an asymptotic analysis of these quantities. In this paper we use a probabilistic argument to show that $\bar{\theta} - z_1^*$ and $z_2^* - \bar{\theta}$ are of order $\lambda^{1/3}$, to determine the coefficients multiplying $\lambda^{1/3}$, and to estimate the loss in expected utility due to the positive transaction cost. This loss in utility is shown to be of order $\lambda^{2/3}$ and the coefficient multiplying $\lambda^{2/3}$ is also determined.

The first hint of the $O(\lambda^{1/3})$ result just reported appears in the appendix of [17]. A detailed but heuristic asymptotic analysis was carried out by Whalley and Wilmott [20]. A rigorous analysis based on viscosity sub- and supersolution arguments that determined the loss in utility and suggested but did not rigorously establish the location of z_1^* and z_2^* was conducted by Janeček and Shreve [7]. At the end of [7], a short but heuristic argument was provided for the main results of the paper. A more compelling heuristic argument was later developed by Rogers [16]. In both cases, the argument was built around the observation that there are two types of loss in the problem with positive transaction costs. The first is the loss due to *displacement*, a loss incurred because one cannot keep the ratio of position in risky asset to total wealth at the desired constant $\bar{\theta}$. The second is the loss due to paying the *transaction cost*. The loss due to displacement increases and the loss due to transaction decreases as the agent becomes more tolerant of departures from $\bar{\theta}$. By estimating these losses and equating the marginal losses, one discovers that z_1^* and z_2^* should differ from $\bar{\theta}$ by $O(\lambda^{1/3})$ and that the optimal expected utility in the problem with transaction cost $\lambda > 0$ is $O(\lambda^{2/3})$ less than the optimal expected utility in the problem with zero transaction cost. In this paper, we provide rigorous bounds on these losses and these bounds are sufficiently tight to enable us to determine the location of z_1^* and z_2^* up to order $\lambda^{1/3}$. More

precisely, under Assumption 4.1 below, this argument determines the highest order terms in the loss in value and in the location of z_1^* and z_2^* (Theorem 4.8). The argument in [16] provides a useful change of measure idea that is instrumental in developing the rigorous argument of this paper (see Sections 5.2 and 5.3).

In all the papers cited, the risky asset is a stock modeled as a geometric Brownian motion. In this paper, we take the risky asset to be a futures price processes modeled as an arithmetic Brownian motion. This removes some technicalities that occur when the agent has 100% of his wealth in the risky asset (see Remark 3.2). Otherwise, the two problems seem to be entirely parallel. We have chosen the arithmetic Brownian motion model in order to remove these technicalities and highlight the main features of the analysis.

Papers that perform asymptotic analysis on related transaction cost problems are [1], [9], [10]. Numerical treatments of transaction costs problems can be found in [2], [4], [15], [18], [19]. Analysis of finite-horizon problems can be found in [3], [5], [11].

2. The model

We return to the futures price process (1.1). Let L and M be nondecreasing, right-continuous processes with $L(0-) = M(0-) = 0$. We interpret $L(t)$ ($M(t)$) as the cumulative number of futures contracts bought (sold) by time t . The number of futures contracts owned by an agent at time t is

$$(2.1) \quad Y(t) = Y(0-) + L(t) - M(t).$$

The wealth $X(t)$ of the agent then evolves according to the equation

$$(2.2) \quad dX(t) = Y(t) dF(t) - \lambda(dL(t) + dM(t)) + rX(t) dt - C(t) dt.$$

So long as $X(u-) > 0$, $0 \leq u \leq t$, we may define $\ell(t) = \int_0^t \frac{dL(u)}{X(u-)}$, $m(t) = \int_0^t \frac{dM(u)}{X(u-)}$, and $c(t) = \int_0^t \frac{C(u) du}{X(u-)}$, and rewrite (2.1), (2.2) as

$$(2.3) \quad dY(t) = X(t-)(d\ell(t) - dm(t)),$$

$$(2.4) \quad dX(t) = Y(t)(\alpha dt + \sigma dW(t)) - \lambda X(t-)(d\ell(t) + dm(t)) + X(t)(r - c(t)) dt.$$

When ℓ and m are continuous, the ratio process $\theta(t) \triangleq Y(t)/X(t)$ satisfies

$$(2.5) \quad d\theta(t) = \theta(t)(-r + c(t) - \alpha\theta(t) + \sigma^2\theta^2(t)) dt - \sigma\theta^2(t) dW(t) + (1 + \lambda\theta(t)) d\ell(t) - (1 - \lambda\theta(t)) dm(t).$$

We require the agent to always have sufficient capital to close out the futures position and still be solvent. In other words, he must trade so that $(X(t), Y(t))$ stays in the closure $\bar{\mathcal{S}}$ of the solvency region

$$\mathcal{S} \triangleq \{(x, y); x + \lambda y > 0, x - \lambda y > 0\}.$$

By computing $d(X(t) + \lambda Y(t))$ and $d(X(t) - \lambda Y(t))$, one can see that if (X, Y) ever reaches the boundary $\partial\mathcal{S}$ of $\bar{\mathcal{S}}$, then to keep from exiting $\bar{\mathcal{S}}$, (X, Y) must jump to the origin and then the agent must make no further trades and must cease consumption. Hence, for purposes of the utility maximization problem described below, we only need to determine the optimal policy in the open region \mathcal{S} . In this region, the reformulation of (2.1), (2.2) as (2.3), (2.4) is legitimate because $\mathcal{S} \subset \{(x, y); x > 0\}$.

Let $(x, y) \in \mathcal{S}$ be given. Let ℓ and m be nondecreasing, right-continuous processes with $\ell(0-) = m(0-) = 0$, and let c be a nonnegative process. We say (ℓ, m, c) is *admissible at (x, y)* and write $(\ell, m, c) \in \mathcal{A}(x, y)$ provided that when we take $X(0-) = x$ and $Y(0-) = y$ and use ℓ, m and c in (2.3), (2.4), the resulting processes X and Y satisfy $(X(t), Y(t)) \in \bar{\mathcal{S}}$ for all $t \geq 0$. Note that because ℓ and m may jump at time zero, $X(0) = x - \lambda x(\ell(0) + m(0))$ and $Y(0) = y + x(\ell(0) - m(0))$. We shall see that except for a possible initial jump, the optimal ℓ and m for the utility maximization problem defined below are continuous.

We now define $v(x, y)$ by (1.4) for all $(x, y) \in \mathcal{S}$. The supremum in (1.4) is over $(\ell, m, c) \in \mathcal{A}(x, y)$. For $(x, y) \in \partial\mathcal{S}$, we necessarily have $(X(t), Y(t)) = (0, 0)$ for all $t \geq 0$, and hence define for $(x, y) \in \partial\mathcal{S}$,

$$v(x, y) = \begin{cases} 0 & \text{if } 0 < p < 1, \\ -\infty & \text{if } p \geq 1. \end{cases}$$

3. Properties of the value function

3.1. Homotheticity. For $\gamma > 0$, $\mathcal{A}(\gamma x, \gamma y) = \mathcal{A}(x, y)$, and when (ℓ, m, c) is chosen from this set, the pair of processes (X^γ, Y^γ) corresponding to the initial condition $(\gamma x, \gamma y)$ is the same as $(\gamma X, \gamma Y)$, where (X, Y) corresponds to the initial condition (x, y) . Because

$$U_p(c(t)X^\gamma(t)) = \begin{cases} \gamma^{1-p}U_p(c(t)X(t)) & \text{if } p > 0, p \neq 1, \\ \log \gamma + U_1(c(t)X(t)) & \text{if } p = 1, \end{cases}$$

v has the *homotheticity property* that for all $\gamma > 0$ and $(x, y) \in \bar{\mathcal{S}}$,

$$(3.1) \quad v(\gamma x, \gamma y) = \begin{cases} \gamma^{1-p}v(x, y) & \text{if } p > 0, p \neq 1, \\ v(x, y) + \frac{1}{\beta} \log \gamma & \text{if } p = 1. \end{cases}$$

From this homotheticity one can argue (see [6] or [17] for details in the geometric Brownian motion model) that the optimal policy when $(X(t), Y(t)) \in \mathcal{S}$ must depend on the ratio $Y(t)/X(t)$. In particular, there are two numbers $z_1^* = z_1^*(\lambda)$ and $z_2^* = z_2^*(\lambda)$ satisfying $-1/\lambda < z_1^* < z_2^* < 1/\lambda$ that define the *no-trade region* $NT \triangleq \{(x, y) \in \mathcal{S} : z_1^* < y/x < z_2^*\}$ (see Figure 1). If $-1/\lambda < Y(0-)/X(0-) < z_1^*$, the agent should immediately buy futures to

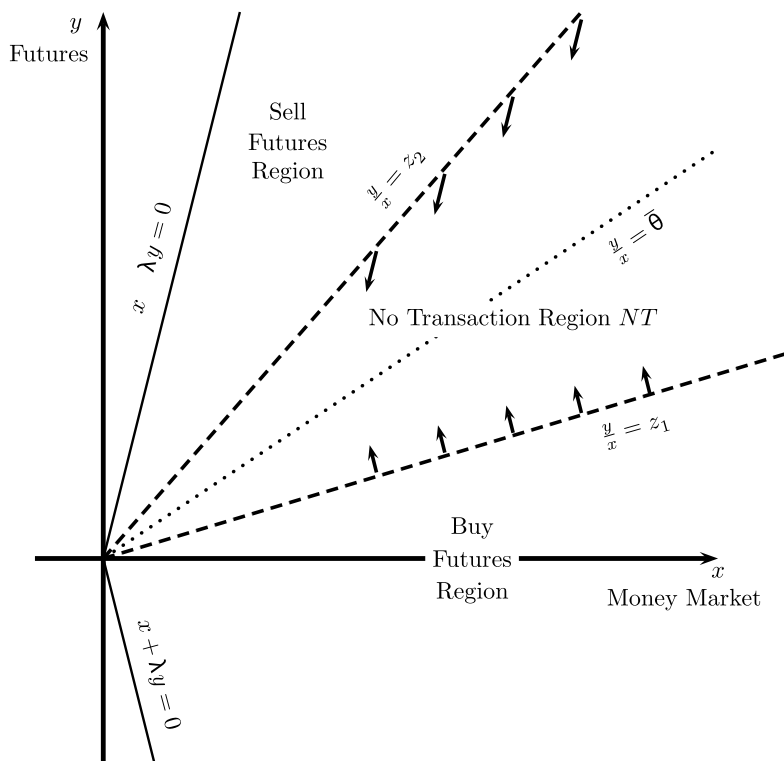


FIGURE 1. Solvency region.

bring $Y(0)/X(0)$ to z_1^* . If $z_2^* < Y(0-)/X(0-) < 1/\lambda$, the agent should immediately sell futures to bring $Y(0)/X(0)$ to z_2^* . In particular, $v(x, y)$ for $(x, y) \in \mathcal{S} \setminus NT$ can be specified in terms of v on the boundary $\{x > 0 : y/x = z_1^* \text{ or } y/x = z_2^*\}$ of NT by

$$v(x, y) = \begin{cases} v\left(\frac{x+\lambda y}{1+\lambda z_1^*}, \frac{z_1^*(x+\lambda y)}{1+\lambda z_1^*}\right) & \text{if } -\frac{1}{\lambda} < \frac{y}{x} \leq z_1^*, \\ v\left(\frac{x-\lambda y}{1-\lambda z_2^*}, \frac{z_2^*(x-\lambda y)}{1-\lambda z_2^*}\right) & \text{if } z_2^* \leq \frac{y}{x} < \frac{1}{\lambda}. \end{cases}$$

Once the pair (X, Y) is in \overline{NT} , the agent should trade only at the boundaries $y/x = z_2^*$ and $y/x = z_1^*$, and trade only enough to prevent (X, Y) from exiting \overline{NT} . In the open set NT , there should be consumption but no trading.

3.2. Homotheticity of type II. The futures trading setup has another useful property, which we call *homotheticity of type II*. Homotheticity of type II does not require that we have a utility function of the form (1.2).

THEOREM 3.1. *For any $(x, y) \in \mathcal{S}, \alpha \geq 0, \lambda \geq 0$, the value function satisfies*

$$(3.2) \quad v(x, y, \alpha, \sigma, \lambda) = v\left(x, ky, \frac{\alpha}{k}, \frac{\sigma}{k}, \frac{\lambda}{k}\right) \quad \forall k > 0,$$

where we have explicitly indicated the dependence of the value function on α and σ appearing in (1.1), (2.4), and on the transaction cost parameter λ .

Proof. The control (ℓ, m, c) is in $\mathcal{A}(x, y)$ with parameters α, σ, λ if and only if the control $(k\ell, km, c)$ is in $\mathcal{A}(x, ky)$ with parameters $\alpha/k, \sigma/k, \lambda/k$. Moreover, the Y process resulting from the control $(k\ell, km, c) \in \mathcal{A}(x, ky)$ is k times the Y process resulting from $(\ell, m, c) \in \mathcal{A}(x, y)$. The X processes are identical. The result follows. \square

REMARK 3.2. In the geometric Brownian motion stock model, when the agent who is faced with zero transaction cost would choose to invest 100% of his wealth in the stock ($\bar{\theta} = 1$), we have an anomalous case because the agent can maintain this position without trading. Because of this, the presence of a positive transaction cost λ reduces the value function by only $O(\lambda)$ rather than $O(\lambda^{2/3})$ (see Remark 1, p. 199 of [7]). One of the consequences of homotheticity of type II is that in the arithmetic Brownian motion futures model, the case $\bar{\theta} = 1$ has no special properties. Indeed, under the scaling of α, σ and λ implicit in (3.2), $\bar{\theta}$ is multiplied by k . Thus, the case of $\bar{\theta} = 1$ can be scaled into a case with $\bar{\theta} \neq 1$.

REMARK 3.3. For sufficiently small $k > 0$, the transaction cost parameter λ/k on the right-hand side of (3.2) can be arbitrarily large. If this transaction cost parameter exceeds one, the agent must pay for changing the bet size $Y(t)$ more than the size of the change. However, it can still be the case that an agent would want to increase the bet size because of high return α/k and small initial bet size. It might also be the case that the agent would want to reduce the bet size. In either case, the subsequent changes in $Y(t)$ are “marked to market” and affect the agent’s wealth $X(t)$ without incurring further transaction costs (see (2.4)).

REMARK 3.4. In the geometric Brownian motion model of [17], the authors show that the Merton proportion is inside the NT region for $\bar{\theta} < 1$ (see Theorem 11.2 and remarks on p. 675). For $\bar{\theta} > 1$, this is the case for sufficiently small transaction costs (see Theorem 2 in [7]), but $\bar{\theta}$ is outside the solvency region and hence outside NT for sufficiently high values of λ .

In the arithmetic model, the inclusion of $\bar{\theta}$ in NT and the relationship between $\bar{\theta}$ and 1 are not connected. Indeed, let us fix the parameters r, β and p . Then homotheticity of type II shows that there exist values for the parameters α, σ and λ for which $\bar{\theta} < 1$ and $\bar{\theta} \in NT$ if and only if there exist other values of these parameters such that $\bar{\theta} \geq 1$ and $\bar{\theta} \in NT$. Similarly, there exist values for α, σ , and λ such that $\bar{\theta} < 1$ and $\bar{\theta} \notin NT$ if and only if there

exist other values for these parameters such that $\bar{\theta} \geq 1$ and $\bar{\theta} \notin NT$. Finally, for any α and σ , there exists sufficiently large λ such that $\bar{\theta} \notin \mathcal{S}$, and thus $\bar{\theta} \notin NT$.

In this paper, we consider only parameter values for which $\bar{\theta}$ is in the interior of NT ; see Assumption 4.1.

3.3. Hamilton–Jacobi–Bellman (HJB) equation. The Hamilton–Jacobi–Bellman (HJB) equation for the model with $\lambda > 0$ is

$$(3.3) \quad \min \left\{ \beta v(x, y) - (rx + \alpha y)v_x(x, y) - \frac{1}{2}\sigma^2 y^2 v_{xx}(x, y) - \tilde{U}_p(v_x(x, y)), \right. \\ \left. \lambda v_x(x, y) - v_y(x, y), \lambda v_x(x, y) + v_y(x, y) \right\} = 0,$$

where $\tilde{U}_p : (0, \infty) \rightarrow \mathbb{R}$ is the *convex dual* (Legendre transform) of U_p :

$$(3.4) \quad \tilde{U}_p(\tilde{C}) \triangleq \max_{C>0} \{U_p(C) - C\tilde{C}\} = \begin{cases} \frac{p}{1-p} \tilde{C}^{(p-1)/p} & \text{if } p > 0, p \neq 1, \\ -\log \tilde{C} - 1 & \text{if } p = 1. \end{cases}$$

The maximizing C in (3.4) is $C = \tilde{C}^{-1/p}$.

It was shown in [17] that as long as the region NT does not contain the positive x -axis nor the positive y -axis, the value function for the geometric Brownian motion model is concave, twice continuously differentiable, and solves the appropriate HJB equation. Adapted to our case, in which $\bar{\theta}$ is strictly positive, those arguments show that for sufficiently small $\lambda > 0$ our value function $v(x, y)$ is concave, twice continuously differentiable, and satisfies the HJB equation (3.3) everywhere in \mathcal{S} . Because the focus of this paper is to obtain probabilistic estimates for the losses associated with positive values of λ , we do not present the necessary modifications of the lengthy analysis in [17] that justify these assertions.

For $\lambda > 0$, the minimum in (3.3) breaks down into three cases:

$$(3.5) \quad \beta v(x, y) - (rx + \alpha y)v_x(x, y) - \frac{1}{2}\sigma^2 y^2 v_{xx}(x, y) - \tilde{U}_p(v_x(x, y))$$

$$= 0 \quad \text{if } z_1^* \leq \frac{y}{x} \leq z_2^*,$$

$$(3.6) \quad \lambda v_x(x, y) - v_y(x, y)$$

$$= 0 \quad \text{if } -\frac{1}{\lambda} < \frac{y}{x} \leq z_1^*,$$

$$(3.7) \quad \lambda v_x(x, y) + v_y(x, y)$$

$$= 0 \quad \text{if } z_2^* \leq \frac{y}{x} < \frac{1}{\lambda}.$$

3.4. Zero transaction cost. If $\lambda = 0$, the problem with dynamics (2.3) and (2.4) is ill posed because the agent should keep $Y(t)/X(t)$ equal to the constant $\bar{\theta}$, and this is not possible when Y is of bounded variation and X

is not. Instead of (2.3) and (2.4), we let Y be a control variable and have a single state X with dynamics

$$(3.8) \quad dX(t) = Y(t)(\alpha dt + \sigma dW(t)) + X(t)(r - c(t)) dt.$$

The solvency region for the $\lambda = 0$ problem is $\{x : x > 0\}$. This is a classical problem that can be solved as in Merton [14]. The value function is

$$(3.9) \quad v_0(x) = \begin{cases} \frac{1}{1-p} A^{-p}(p) x^{1-p} & \text{if } p > 0, p \neq 1, \\ \frac{1}{\beta} \log \beta x + \frac{r-\beta}{\beta^2} + \frac{\alpha^2}{2\beta^2\sigma^2} & \text{if } p = 1, \end{cases}$$

which is finite for $x > 0$ because $A(p)$ given by (1.3) is assumed to be positive. The function $v_0(x)$ solves the HJB equation

$$(3.10) \quad \min_{y \in \mathbb{R}, c \geq 0} \left\{ \beta v_0(x) - (rx + \alpha y)v'_0(x) - \frac{1}{2} \sigma^2 y^2 v''_0(x) + cyv'_0(x) - U_p(cx) \right\} = 0.$$

The optimal ratio for y/x , found by minimizing over y in (3.10), is $\bar{\theta}$ given by (1.5). The optimal consumption level, found by minimizing over c in (3.10), is $A(p)$.

REMARK 3.5. The fact that $v_0(x) < \infty$ for $x > 0$ implies that the value function $v(x, y)$ for the less favorable problem with $\lambda > 0$ also satisfies $v(x, y) < \infty$ for $(x, y) \in \mathcal{S}$. Of course, $v(x, y) > -\infty$ for all $(x, y) \in \mathcal{S}$ because the agent can immediately trade to a zero position in futures and thereafter simply consume at rate $c = r$, which leaves X constant. We see in fact that on each compact subset of \mathcal{S} (\mathcal{S} corresponding to some λ_0), $v(x, y)$ is bounded uniformly over $\lambda \in (0, \lambda_0]$.

3.5. Initial estimates. The maximizing C in (3.4) when $\tilde{C} = v_x(x, y)$ is $C = (v_x(x, y))^{-1/p}$. We use the notation $C = cx$ (see, e.g., (2.4) and (1.4)), and the maximizing c is thus $\frac{1}{x}(v_x(x, y))^{-1/p}$. Because of the homotheticity (3.1), $v(x, y) = x^{1-p}v(1, \frac{y}{x})$ if $p \neq 1$ and $v(x, y) = v(1, \frac{y}{x}) + \frac{1}{\beta} \log x$ if $p = 1$, and hence

$$v_x(x, y) = \begin{cases} x^{-p}((1-p)v(1, \theta) - \theta v_y(1, \theta)) & \text{if } p \neq 1, \\ \frac{1}{x}(\frac{1}{\beta} - \theta v_y(1, \theta)) & \text{if } p = 1, \end{cases}$$

where $\theta = y/x$. For $z_1^* \leq \theta \leq z_2^*$, the maximizing c ,

$$(3.11) \quad c^*(\theta) = \begin{cases} ((1-p)v(1, \theta) - \theta v_y(1, \theta))^{-1/p} & \text{if } p \neq 1, \\ (\frac{1}{\beta} - \theta v_y(1, \theta))^{-1} & \text{if } p = 1, \end{cases}$$

is a function of θ . We take (3.11) to be the definition of $c^*(\theta)$ for all $\theta \in (-1/\lambda, 1/\lambda)$. This function is locally Lipschitz on $(-1/\lambda, 1/\lambda)$ because v is twice continuously differentiable.

PROPOSITION 3.6. *Let $[z_1, z_2]$ be a compact subinterval of \mathbb{R} which, for sufficiently small λ , contains z_1^* , z_2^* and $\bar{\theta}$. For $\theta \in [z_1, z_2]$, we have¹*

$$(3.12) \quad c^*(\theta) = \begin{cases} ((1-p)v(1, \bar{\theta}))^{-1/p} + O(\lambda) & \text{if } p \neq 1, \\ \beta + O(\lambda) & \text{if } p = 1, \end{cases}$$

and

$$(3.13) \quad v(1, \theta) = v(1, \bar{\theta}) + (\theta - \bar{\theta})O(\lambda).$$

Proof. From (3.6) and (3.7) we have

$$\lambda v_x(x, z_1x) - v_y(x, z_1x) = 0, \quad \lambda v_x(x, z_2x) + v_y(x, z_2x) = 0.$$

For $i = 1, 2$, the homotheticity $v(x, z_ix) = x^{1-p}v(1, z_i)$ for $p \neq 1$ or $v(x, z_ix) = v(1, z_i) + \frac{1}{\beta} \log x$ for $p = 1$ implies that

$$v_x(x, z_ix) + z_iv_y(x, z_ix) = \begin{cases} (1-p)x^{-p}v(1, z_i) & \text{if } p \neq 1, \\ \frac{1}{\beta x} & \text{if } p = 1. \end{cases}$$

We solve these equations for v_y :

$$v_y(x, z_1x) = \begin{cases} \frac{\lambda(1-p)x^{-p}}{1+\lambda z_1}v(1, z_1) & \text{if } p \neq 1, \\ \frac{\lambda}{(1+\lambda z_1)\beta x} & \text{if } p = 1, \end{cases}$$

$$v_y(x, z_2x) = \begin{cases} -\frac{\lambda(1-p)x^{-p}}{1-\lambda z_2}v(1, z_2) & \text{if } p \neq 1, \\ -\frac{\lambda}{(1-\lambda z_2)\beta x} & \text{if } p = 1. \end{cases}$$

Since v is concave, $v_y(x, \cdot)$ is decreasing, and this yields the bounds

$$v_y(x, z_2x) \leq v_y(x, y) \leq v_y(x, z_1x), \quad z_1x \leq y \leq z_2x.$$

Both bounds are $x^{-p}O(\lambda)$, so $v_y(1, \theta) = O(\lambda)$ for $z_1 \leq \theta \leq z_2$. Equation (3.13) follows immediately. A Taylor series expansion of (3.11) using (3.13) yields (3.12). □

REMARK 3.7. If $0 < p \leq 1$, then $pA(p) + (1-p)c^*(\theta) \geq pA(p)$, which is strictly positive. On the other hand, if $p > 1$, then $v(1, \bar{\theta}) \leq v_0(1) = \frac{1}{1-p}A^{-p}(p)$ and thus $((1-p)v(1, \bar{\theta}))^{-\frac{1}{p}} \leq A(p)$. It follows that for sufficiently small $\lambda_0 > 0$ and θ in an arbitrary compact subinterval of $(-1/\lambda_0, 1/\lambda_0)$,

$$pA(p) + (1-p)c^*(\theta) \geq A(p) + O(\lambda),$$

which is bounded away from zero as λ ranges over $(0, 1/\lambda_0]$.

¹ We mean by $O(\lambda)$ in (3.12) and (3.13) a term whose absolute value is bounded by λ times a constant that does not depend on θ in the compact subinterval $[z_1, z_2]$ nor on $\lambda \in (0, \varepsilon)$ for some $\varepsilon > 0$, although the bound may depend on z_1 and z_2 . See Remark 4.7 for a fuller discussion of the $O(\cdot)$ notation as it is used in this paper.

COROLLARY 3.8. *For sufficiently small $\lambda_0 > 0$, let $-1/\lambda_0 < z_1 < z_2 < 1/\lambda_0$, and let ν be a probability measure on $[z_1, z_2]$. Then for $\lambda \in (0, \lambda_0]$ and $y \in [z_1, z_2]$, we have*

$$v(1, y) = \int_{z_1}^{z_2} v(1, \theta)\nu(d\theta) + (z_2 - z_1)O(\lambda),$$

where the bound on the $O(\lambda)$ term depends on z_1 and z_2 but not on ν .

4. Main results

We want to estimate the difference in $v(x, y)$ given by (1.4) and $v_0(x)$ given by (3.9). We separate this difference into two parts, the loss due to transaction costs and the loss due to displacement, where ‘‘displacement’’ refers to the fact that in the problem with positive λ , we cannot keep $\theta(t)$ at $\bar{\theta}$. We then minimize the sum of these losses by equating marginal losses.

4.1. Decomposing the loss. In order not to digress into a lengthy analysis of the HJB equation, we assume rather than prove that there exists an optimal policy and it has the following form. This assumption is valid for nearly all choices of parameters in the geometric Brownian motion model (see [17]), and we conjecture that it holds for all choices of parameters satisfying (1.3) in the arithmetic Brownian motion model considered here.

ASSUMPTION 4.1. We denote the dependence of $z_i^* = z_i^*(\lambda)$ on λ . We assume that for $\lambda > 0$ sufficiently small, $0 < z_1^*(\lambda) < \bar{\theta} < z_2^*(\lambda)$ and there is a function $\varphi(\lambda)$ satisfying $\lim_{\lambda \downarrow 0} \varphi(\lambda) = 0$ and $z_2^*(\lambda) - z_1^*(\lambda) \leq \bar{\theta}\varphi(\lambda)$ for λ sufficiently small. Without loss of generality, we take $\varphi(\lambda) > O(\lambda^{1/3})$.

For the remainder of the paper, we consider only the case that the initial capital in the money market is $X(0) = 1$. We can do this without loss of generality because of homotheticity. For the computations below, we initially hold the consumption proportion rate c in (3.8) constant. We fix $c > 0$ so that it satisfies

$$(4.1) \quad pA(p) + (1 - p)c > 0.$$

We then obtain estimates that hold uniformly in c , provided that c is bounded and c and $pA(p) + (1 - p)c$ are bounded away from zero. If $0 < p \leq 1$, the second condition imposes no constraint on c .

We first set up a utility corresponding to *zero displacement and zero transaction cost*. To do this, we use $c(t) \equiv c$ and $Y(t) = \bar{\theta}X(t)$ in (3.8). We denote the resulting X process by X_0 , which is given by

$$(4.2) \quad X_0(t) = \exp \left\{ \left(r - c + \alpha\bar{\theta} - \frac{1}{2}\sigma^2\bar{\theta}^2 \right) t + \sigma\bar{\theta}W(t) \right\},$$

$$(4.3) \quad \mathbb{E}X_0^{1-p}(t) = \exp \left\{ (1 - p) \left(r - c + \frac{1}{2}\alpha\bar{\theta} \right) t \right\},$$

where we have used (1.5). One can further verify that

$$(4.4) \quad (1 - p) \left(r - c + \frac{1}{2} \alpha \bar{\theta} \right) = \beta - pA(p) - (1 - p)c.$$

Therefore, for $p \neq 1$, $\mathbb{E}X_0^{1-p}(t) = e^{(\beta - pA(p) - (1-p)c)t}$, whereas for $p = 1$, $\mathbb{E} \log X_0(t) = (r - c + \frac{\alpha^2}{2\sigma^2})t$. It is now straightforward to compute

$$(4.5) \quad u_0(c) \triangleq \mathbb{E} \int_0^\infty e^{-\beta t} U_p(cX_0(t)) dt = \begin{cases} \frac{c^{1-p}}{(1-p)(pA(p) + (1-p)c)} & \text{if } p \neq 1, \\ \frac{1}{\beta} \log c + \frac{r-c}{\beta^2} + \frac{\alpha^2}{2\beta^2\sigma^2} & \text{if } p = 1. \end{cases}$$

When $p > 1$, the expression on the right-hand side of (4.5) is negative because of (4.1). For all values of p , the expression on the right-hand side of (4.5) is maximized over c by $A(p)$, that is,

$$(4.6) \quad u_0(A(p)) = v_0(1).$$

We next set up a utility corresponding to *positive displacement and positive transaction cost*. To do this, we choose positive numbers w_1 and w_2 . We consider the value that can be achieved by trading just enough to keep the ratio of position in futures to wealth in money market inside the interval $[\bar{\theta}(1 - w_1), \bar{\theta}(1 + w_2)]$. Eventually, we will optimize over w_1 and w_2 .

Let $X_2(0) = 1$ and let $Y_2(0) = \theta_2(0)$, where $\theta_2(0)$ is a random variable independent of W and taking values in $[\bar{\theta}(1 - w_1), \bar{\theta}(1 + w_2)]$. If we took $(X_2(\cdot), Y_2(\cdot))$ to be the solution of (2.3) and (2.4) where $c(t)$ is some Lipschitz function $c(\theta_2(t))$ of $\theta_2(t) = Y_2(t)/X_2(t)$ and where $\ell = \ell_2$ and $m = m_2$ are the minimal continuous, nondecreasing processes such that

$$(4.7) \quad \theta_2(t) \triangleq Y_2(t)/X_2(t) \in [\bar{\theta}(1 - w_1), \bar{\theta}(1 + w_2)] \quad \forall t \geq 0,$$

then we would have $\ell_2(0) = m_2(0) = 0$, $X_2(\cdot)$, $Y_2(\cdot)$ and $\theta_2(\cdot)$ would be continuous, and (2.5) in this case would become

$$(4.8) \quad d\theta_2(t) = \theta_2(t)(-r + c(\theta_2(t)) - \alpha\theta_2(t) + \sigma^2\theta_2^2(t)) dt - \sigma\theta_2^2(t) dW(t) + (1 + \lambda\bar{\theta}(1 - w_1)) d\ell_2(t) - (1 - \lambda\bar{\theta}(1 + w_2)) dm_2(t).$$

We indeed take $\theta_2(\cdot)$ to be the solution of (4.8), leaving the choice of the distribution of $\theta_2(0)$ and the function $c(\cdot)$ open. However, for $X_2(\cdot)$, we fix a constant $c > 0$ satisfying (4.1) and let $X_2(\cdot)$ be the solution of the equation

$$(4.9) \quad dX_2(t) = X_2(t) \left[(r - c + \alpha\theta_2(t)) dt + \sigma\theta_2(t) dW(t) - \lambda(d\ell_2(t) + dm_2(t)) \right].$$

The value associated with X_2 is defined to be

$$(4.10) \quad u_2(c, c(\cdot), w_1, w_2) \triangleq \mathbb{E} \int_0^\infty e^{-\beta t} U_p(cX_2(t)) dt.$$

REMARK 4.2. We obtain estimates for $u_2(c, c(\cdot), w_1, w_2)$ that are uniform over $c(\cdot)$ (provided the class of $c(\cdot)$ considered is uniformly bounded, $pA(p) + (1 - p)c(\cdot)$ is uniformly bounded away from zero, and each $c(\cdot)$ in the class

varies by not more than $\kappa\lambda$ in $[\bar{\theta}(1 - w_1), \bar{\theta}(1 + w_2)]$, where the constant κ is uniform over the class) and uniform over c (provided that c and $pA(p) + (1 - p)c$ are bounded from above and away from zero). The two choices of $c(\cdot)$ that we will need to consider are $c(\cdot) = c^*(\cdot)$ given by (3.11) and $c(\cdot)$ equal to a constant c . The desired properties of $c^*(\cdot)$ follow from Remarks 3.5 and 3.7 and Proposition 3.6.

REMARK 4.3. If $c(\cdot)$ is $c^*(\cdot)$ given by (3.11) and if $\bar{\theta}(1 - w_1) = z_1^*$ and $\bar{\theta}(1 + w_2) = z_2^*$, then $\theta_2(t)$ given by (4.8) is the optimal portfolio proportion process, albeit with a random initial condition. We denote this process by θ^* , i.e.,

$$(4.11) \quad \theta^*(t) = \theta^*(t)(-r + c^*(\theta^*(t)) - \alpha\theta^*(t) + \sigma^2(\theta^*(t))^2) dt - \sigma(\theta^*(t))^2 dW(t) + (1 + \lambda z_1^*) d\ell^*(t) - (1 - \lambda z_2^*) dm^*(t),$$

where ℓ^* and m^* are the minimal continuous, nondecreasing processes such that $\theta^*(t)$ given by (4.11) stays in the interval $[z_1^*, z_2^*]$. If, in addition, we replace the constant c in (4.9) by $c^*(\theta^*(t))$ and call the resulting process X^* , i.e., $X^*(0) = 1$ and

$$(4.12) \quad dX^*(t) = X^*(t)[(r - c^*(\theta^*(t)) + \alpha\theta^*(t)) dt + \sigma\theta^*(t) dW(t) - \lambda(d\ell^*(t) + dm^*(t))],$$

then X^* is the optimal amount to be invested in the money market. In particular,

$$(4.13) \quad \mathbb{E}v(1, \theta^*(0)) = \mathbb{E} \int_0^\infty e^{-\beta t} U_p(c^*(\theta^*(t)) X^*(t)) dt.$$

Finally, we set up a utility for the intermediate situation of *positive displacement but zero transaction cost*. We define the process $X_1(\cdot)$ by setting $X_1(0) = 1$ and

$$(4.14) \quad dX_1(t) = X_1(t)[(r - c + \alpha\theta_2(t)) dt + \sigma\theta_2(t) dW(t)].$$

The process $\theta_2(\cdot)$ in (4.14) is the process determined by (4.8). The process X_1 does not incur transaction costs but it does incur a “displacement cost” because $\theta_2(t)$ is not identically equal to $\bar{\theta}$. We define the associated value

$$(4.15) \quad u_1(c, c(\cdot), w_1, w_2) \triangleq \mathbb{E} \int_0^\infty e^{-\beta t} U_p(cX_1(t)) dt.$$

The remainder of the paper develops the estimates reported in the following theorems. The proofs are deferred to Section 5.

THEOREM 4.4 (Transaction loss). *Let $w_1 > 0$ and $w_2 > 0$ be given and define $w \triangleq w_1 + w_2$. Then there exist positive constants C_1 and C_2 such that*

$$(4.16) \quad u_1(c, c(\cdot), w_1, w_2) - u_2(c, c(\cdot), w_1, w_2) \geq \max\{\min\{C_1\lambda w^{-1}, C_2\} + O(\lambda), 0\}.$$

Furthermore, if $\lambda/w = o(1)$, then

$$(4.17) \quad \begin{aligned} & u_1(c, c(\cdot), w_1, w_2) - u_2(c, c(\cdot), w_1, w_2) \\ &= \frac{c^{1-p}\sigma^2\bar{\theta}^3}{(pA(p) + (1-p)c)^2} \cdot \frac{\lambda}{w} + O(\lambda) + O(\lambda^2w^{-2}). \end{aligned}$$

THEOREM 4.5 (Displacement loss). *Let $w_1 > 0$ and $w_2 > 0$ be given and define $w \triangleq w_1 + w_2$. Let $\theta_2(0)$ have the distribution under \mathbb{P} corresponding to the stationary distribution of the solution to (5.56) below. Then*

$$(4.18) \quad \begin{aligned} & 0 \leq u_0(c) - u_1(c, c(\cdot), w_1, w_2) \\ &= \frac{c^{1-p}p\sigma^2\bar{\theta}^2(w_1^2 - w_1w_2 + w_2^2)}{6(pA(p) + (1-p)c)^2} + O(\lambda w^2) + O(w^3). \end{aligned}$$

Summing (4.17) and (4.18), we obtain the following corollary.

COROLLARY 4.6 (Total loss). *Under the hypotheses of Theorem 4.5, if $\lambda/w = o(1)$, then*

$$(4.19) \quad \begin{aligned} & 0 \leq u_0(c) - u_2(c, c(\cdot), w_1, w_2) \\ &= \frac{c^{1-p}\sigma^2\bar{\theta}^2}{(pA(p) + (1-p)c)^2} \left[\frac{\lambda\bar{\theta}}{w_1 + w_2} + \frac{p}{6}(w_1^2 - w_1w_2 + w_2^2) \right] \\ & \quad + O(\lambda) + O(w^3) + O(\lambda^2w^{-2}). \end{aligned}$$

REMARK 4.7. Constants appearing in the estimates in this work are permitted to depend on the model parameters r , α , σ and p , but not on λ , w_1 and w_2 , provided these are sufficiently small positive numbers. Constants also may not depend on t and ω . When we consider processes constrained to stay in an interval $[a, b]$, constants used in estimates may not depend on a and b . In some cases, to achieve this independence from a and b , we shall restrict attention to a and b for which $b - a$ is sufficiently small. Finally, the notation $O(1)$, $O(\lambda)$, $O(\lambda w^{-1})$, etc., is used to indicate any term whose absolute value is bounded by a constant times the argument appearing in the notation, so long as λ and w are sufficiently small (although terms like λw^{-1} might not be small). Moreover, $\lambda/w = o(1)$ means that $\lambda \downarrow 0$ and $w \downarrow 0$ in such a way that $\lambda/w \rightarrow 0$. In the case of (4.16)–(4.19), where c and $c(\cdot)$ appear in the relations, the constants and $O(\cdot)$ terms do not depend on c and $c(\cdot)$ when c ranges over a set of positive numbers for which c and $pA(p) + (1-p)c$ are bounded and bounded away from zero and $c(\cdot)$ ranges over a set of functions that are all bounded by the same bound, $pA(p) + (1-p)c(\cdot)$ is bounded away from zero by a bound independent of $c(\cdot)$, and each function in the set varies by no more than $O(\lambda)$ on compact subintervals (the properties enjoyed by $c^*(\cdot)$; see Remark 4.2).

4.2. Equating marginal losses. If we could ignore the $O(\cdot)$ terms in Corollary 4.6, in order to optimize over investment strategies we would minimize the convex function

$$(4.20) \quad g_\lambda(w_1, w_2) \triangleq \frac{\lambda\bar{\theta}}{w_1 + w_2} + \frac{p}{6}(w_1^2 - w_1w_2 + w_2^2)$$

appearing in (4.19). For future reference, we note that

$$(4.21) \quad \nabla g_\lambda(w_1, w_2) = \left[\begin{array}{c} -\frac{\lambda\bar{\theta}}{(w_1+w_2)^2} + \frac{p}{6}(2w_1 - w_2) \\ -\frac{\lambda\bar{\theta}}{(w_1+w_2)^2} + \frac{p}{6}(2w_2 - w_1) \end{array} \right],$$

$$(4.22) \quad \nabla^2 g_\lambda(w_1, w_2) = \frac{2\lambda\bar{\theta}}{(w_1 + w_2)^3} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} + \frac{p}{3} \begin{bmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{bmatrix}.$$

The minimum of g_λ is attained by

$$(4.23) \quad w_1(\lambda) = w_2(\lambda) \triangleq \left(\frac{3\lambda\bar{\theta}}{2p} \right)^{1/3},$$

so that $\lambda/(w_1(\lambda) + w_2(\lambda)) = o(1)$, the minimal value of g_λ is

$$(4.24) \quad g_\lambda(w_1(\lambda), w_2(\lambda)) = \theta^{2/3} \lambda^{2/3} \left(\frac{9p}{32} \right)^{1/3},$$

and substitution of this into the right-hand side of (4.19) results in

$$(4.25) \quad \begin{aligned} &u_2(c, c(\cdot), w_1(\lambda), w_2(\lambda)) \\ &= u_0(c) - \frac{c^{1-p} \sigma^2 \bar{\theta}^{8/3}}{(pA(p) + (1-p)c)^2} \left(\frac{9p}{32} \right)^{1/3} \lambda^{2/3} + O(\lambda). \end{aligned}$$

With $w_1(\lambda) = w_2(\lambda)$ given by (4.23), equation (4.25) is a direct consequence of Corollary 4.6.

If $p = 1$ and we ignore the $O(\lambda)$ term in (4.25) when maximizing over c , we find the maximal value at $A(1) = \beta$. Substitution into (4.25) yields (see (4.6))

$$(4.26) \quad u_2(\beta, \beta, w_1(\lambda), w_2(\lambda)) = v_0(1) - \frac{\sigma^2 \bar{\theta}^{8/3}}{A^{1+p}(p)} \left(\frac{9p}{32} \right)^{1/3} \lambda^{2/3} + O(\lambda).$$

The maximization of (4.25) over c is more difficult when $p \neq 1$, but we shall see (Lemma 5.13) that the maximizer is nearly $A(p)$. Substitution of this value of c into (4.25) leads to (4.26) even when $p \neq 1$.

Because the argument just given ignores the $O(\cdot)$ terms in Corollary 4.6 when maximizing over w_1 , w_2 and c , we cannot immediately assert that $u_2(A(p), A(p), w_1(\lambda), w_2(\lambda))$ is, up to $O(\lambda)$, the maximal utility that can be achieved in the problem with positive transaction cost λ . Our main result, Theorem 4.8 below, asserts that this is almost the case.

THEOREM 4.8 (Value function). *Under Assumption 4.1,*

$$(4.27) \quad v(1, \bar{\theta}) = v_0(1) - \frac{\sigma^2 \bar{\theta}^{8/3}}{A^{1+p}(p)} \left(\frac{9p}{32}\right)^{1/3} \lambda^{2/3} + O(\lambda^{5/6}),$$

$$(4.28) \quad z_i^*(\lambda) = w_i(\lambda) + O(\lambda^{5/12}), \quad i = 1, 2,$$

where we explicitly indicate the dependence of $z_i^* = z_i^*(\lambda)$ on $\lambda > 0$.

We note from Proposition 3.6 that so long as y lies in a compact subset of \mathbb{R} , we have $v(1, y) = v(1, \bar{\theta}) + O(\lambda)$, so (4.27) applies to $v(1, y)$ as well. Using homotheticity, we can extend the formula to $v(x, y)$.

5. Proofs

This section contains the proofs of Theorems 4.4, 4.5 and 4.8. To help the reader follow the flow of the argument, we provide diagrams showing the dependencies of the results in this section and Theorems 4.4, 4.5 and 4.8 and Corollary 4.6. We use $A \longrightarrow B$ to mean that the proof of B uses the result A . The argument leading to Corollary 4.6 is diagrammed in Figure 2. After obtaining Theorems 4.4 and 4.5 and Corollary 4.6, the argument leading to Theorem 4.8 is organized as indicated in Figure 3.

5.1. Local time estimates. The proofs of Theorems 4.4, 4.5 and 4.8 require estimates pertaining to the processes l_2 and m_2 appearing in (4.8). This section provides these.

Let $a, b \in \mathbb{R}$ be given with $a < b$. For $i = 1, 2$, let $f_i : [0, \infty) \rightarrow \mathbb{R}$ be a continuous function with $a \leq f_i(0) \leq b$. Let ℓ_i and m_i be the minimal nondecreasing functions such that

$$g_i(t) \triangleq f_i(t) + \ell_i(t) - m_i(t) \in [a, b] \quad \forall t \geq 0.$$

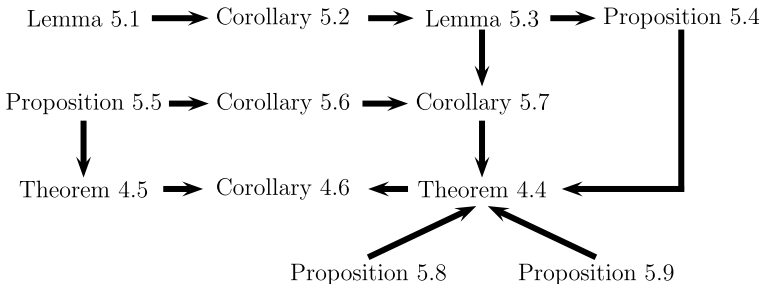


FIGURE 2. Dependencies to obtain Corollary 4.6.

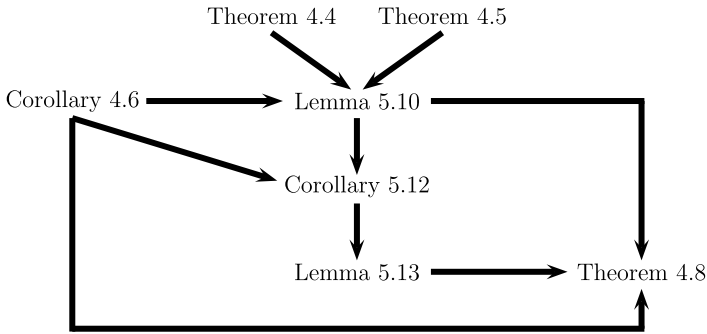


FIGURE 3. Dependencies to obtain Theorem 4.8.

The processes ℓ_i and m_i push only when g_i is at the boundary a or b , respectively. In other words, they satisfy

$$(5.1) \quad \ell_i(t) = \int_0^t \mathbb{I}_{\{g_i(s)=a\}} d\ell_i(s), \quad m_i(t) = \int_0^t \mathbb{I}_{\{g_i(s)=b\}} dm_i(s) \quad \forall t \geq 0.$$

Theorem 1.6 of [8] implies the following result.

LEMMA 5.1. Define $h \triangleq f_2 - f_1$ and assume that h is nondecreasing and $h(0) \geq 0$. Then for $t \geq 0$,

$$(5.2) \quad \ell_2(t) \leq \ell_1(t) \leq \ell_2(t) + h(t), \quad m_1(t) \leq m_2(t) \leq m_1(t) + h(t).$$

COROLLARY 5.2. In the context of Lemma 5.1, suppose $a \leq x \leq y \leq b$ and for some continuous function f with $f(0) = 0$, we have $f_1(t) = x + f(t)$ and $f_2(t) = y + f(t)$ for all $t \geq 0$. Then $\ell_2(t) \leq \ell_1(t) \leq \ell_2(t) + y - x$ and $m_1(t) \leq m_2(t) \leq m_1(t) + y - x$.

Let $a, b \in \mathbb{R}$ be given with $0 < b - a \leq 1$. Consider $\psi(\cdot)$ satisfying $\psi(0) \in [a, b]$ and

$$(5.3) \quad d\psi(t) = \mu(\psi(t)) dt + \sigma(\psi(t)) dW(t) + d\ell(t) - dm(t), \quad t \geq 0,$$

where W is a Brownian motion and $\mu(\cdot)$ and $\sigma(\cdot)$ are Lipschitz continuous functions defined on some compact interval I containing $[a, b]$. Here, $\ell(\cdot)$ and $m(\cdot)$ are the minimal nondecreasing processes such that $\psi(t) \in [a, b]$ for all $t \geq 0$. We define $\underline{\mu} \triangleq \min_{x \in I} \mu(x)$, $\bar{\mu} \triangleq \max_{x \in I} \mu(x)$, $\underline{\sigma} \triangleq \min_{x \in I} \sigma(x)$, $\bar{\sigma} \triangleq \max_{x \in I} \sigma(x)$, and we assume $\underline{\sigma} > 0$.

LEMMA 5.3. Let ψ be given by (5.3) with $\psi(0) \in [a, b]$, and assume that $\sigma(x) = 1$ for all x . Let $\psi_0(0) \in [a, b]$ be given and define $\psi_0(\cdot)$ by

$$(5.4) \quad \psi_0(t) = \psi_0(0) + W(t) + \ell_0(t) - m_0(t),$$

where $\ell_0(\cdot)$ and $m_0(\cdot)$ are the minimal nondecreasing processes such that $\psi_0(t) \in [a, b]$ for all $t \geq 0$. Then

$$\begin{aligned} \ell_0(t) - \bar{\mu}^+t - (b - a) &\leq \ell(t) \leq \ell_0(t) + \underline{\mu}^-t + (b - a), \\ m_0(t) - \underline{\mu}^-t - (b - a) &\leq m(t) \leq m_0(t) + \bar{\mu}^+t + (b - a). \end{aligned}$$

Proof. According to Corollary 5.2, a change of the initial condition in (5.4) by an amount less than or equal to $b - a$ changes the ℓ_0 and m_0 terms by no more than $b - a$. Therefore, it suffices to prove

$$(5.5) \quad \ell_0(t) - \bar{\mu}^+t \leq \ell(t) \leq \ell_0(t) + \underline{\mu}^-t, \quad m_0(t) - \underline{\mu}^-t \leq m(t) \leq m_0(t) + \bar{\mu}^+t$$

under the assumption $\psi_0(0) = \psi(0)$.

We prove the first inequality in (5.5); the others are similar. For this, we define $f(t) = \psi(0) + \int_0^t \mu(\psi(s)) ds + W(t)$. Then ℓ and m in (5.3) are the minimal nondecreasing processes for which $f + \ell - m \in [a, b]$. We set $f_0(t) = \psi(0) + W(t)$, so that ℓ_0 and m_0 appearing in (5.4) are the minimal nondecreasing processes for which $f_0 + \ell_0 - m_0 \in [a, b]$. If $\bar{\mu} \leq 0$, then $h \triangleq f_0 - f$ is nondecreasing, and the first inequality in (5.5) follows from the first inequality in (5.2). If $\bar{\mu} > 0$, then we also define $f_2(t) = \psi(0) + \bar{\mu}t + \sigma W(t)$, and denote by ℓ_2 and m_2 the minimal nondecreasing processes for which $f_2 + \ell_2 - m_2 \in [a, b]$. Now $f_2 - f$ and $f_2 - f_0$ are both nondecreasing. The first inequality in (5.2) implies $\ell_2 \leq \ell$ and the second implies $\ell_0(t) \leq \ell_2(t) + \bar{\mu}t$. Combining these, we again obtain the first inequality in (5.5). \square

PROPOSITION 5.4. *Let ψ be given by (5.3). For each positive integer k ,*

$$(5.6) \quad \mathbb{E}\ell^k(t) = O\left(\frac{(t + 1)^k}{(b - a)^k}\right), \quad \mathbb{E}m^k(t) = O\left(\frac{(t + 1)^k}{(b - a)^k}\right) \quad \forall t \geq 0.$$

Proof. We consider first the case that $[a, b] = [0, 1]$, $\mu(x) = 0$ and $\sigma(x) = 1$ for all $x \in [0, 1]$. We let $\psi(0)$ have the stationary distribution for this case (which happens to be uniform), so that the distribution of $\ell(n + 1) - \ell(n)$ is independent of $n = 0, 1, \dots$. We prove by induction that

$$(5.7) \quad \mathbb{E}\ell^k(n) \leq n^k \mathbb{E}\ell^k(1), \quad n = 1, 2, \dots$$

For $n = 1$, (5.7) holds. Assume (5.7) holds for some value of $n \geq 1$. Then

$$\begin{aligned} \mathbb{E}\ell^k(n + 1) &= \mathbb{E}[(\ell(n) + (\ell(n + 1) - \ell(n)))^k] \\ &= \sum_{i=0}^k \binom{k}{i} \mathbb{E}[\ell^i(n) (\ell(n + 1) - \ell(n))^{k-i}] \\ &\leq \sum_{i=0}^k \binom{k}{i} (\mathbb{E}[\ell^k(n)])^{\frac{i}{k}} \cdot (\mathbb{E}[(\ell(n + 1) - \ell(n))^k])^{\frac{k-i}{k}} \end{aligned}$$

$$\begin{aligned} &\leq \sum_{i=0}^k \binom{k}{i} n^i (\mathbb{E}[\ell^k(1)])^{\frac{i}{k}} \cdot (\mathbb{E}[\ell^k(1)])^{\frac{k-i}{k}} \\ &= \mathbb{E}\ell^k(1) \cdot \sum_{i=0}^k \binom{k}{i} n^i 1^{k-i} \\ &= (n+1)^k \mathbb{E}\ell^k(1). \end{aligned}$$

Since ℓ is nondecreasing, we have the first equality in (5.6) with $O((t+1)^k) = (t+1)^k \mathbb{E}\ell^k(1)$. We further have

$$(5.8) \quad \mathbb{E}[(\ell(t) + 1)^k] = 2^k O((t+1)^k) + 2^k = O((t+1)^k).$$

If $\psi(0)$ is a nonrandom initial condition in $[a, b]$, then Lemma 5.3 shows that $\ell(t)$ changes by no more than $b - a$, and (5.8) gives us (5.6) even in this case.

We now permit μ to be a Lipschitz continuous function on $[0, 1]$, but continue with the assumptions that $[a, b] = [0, 1]$ and $\sigma(x) = 1$ for all $x \in [0, 1]$. We obtain (5.6) for this case of doubly reflected Brownian motion with bounded drift on $[0, 1]$ from Lemma 5.3 and the case just considered.

For the case of general $[a, b]$ with $0 < b - a \leq 1$, general μ and σ , we define the time change $A(t) \triangleq \frac{1}{(b-a)^2} \int_0^t \sigma^2(\psi(u)) du$ for all $t \geq 0$, and its inverse $T(s) \triangleq A^{-1}(s)$, so that $B(s) \triangleq \frac{1}{b-a} \int_0^{T(s)} \sigma(\psi(u)) dW(u)$ is a Brownian motion. We note that $\underline{\sigma}^2 t / (b-a)^2 \leq A(t) \leq \bar{\sigma}^2 t / (b-a)^2$. We have

$$\begin{aligned} \varphi(s) &\triangleq \frac{1}{b-a} [\psi(T(s)) - a] \\ &= \varphi(0) + (b-a) \int_0^s \frac{\mu((b-a)\varphi(v) + a)}{\sigma^2((b-a)\varphi(v) + a)} dv + B(s) \\ &\quad + \frac{1}{b-a} \ell(T(s)) - \frac{1}{b-a} m(T(s)). \end{aligned}$$

The process φ is a doubly reflected Brownian motion on $[0, 1]$ with drift bounded below by $\underline{\mu}/\bar{\sigma}^2$ and above by $\bar{\mu}/\underline{\sigma}^2$. The processes $\frac{1}{b-a} \ell(T(s))$ and $\frac{1}{b-a} m(T(s))$ are the minimal nondecreasing processes that cause this reflection, and hence the case already considered implies

$$\frac{1}{(b-a)^k} \mathbb{E}\ell^k(T(s)) = O((s+1)^k), \quad \frac{1}{(b-a)^k} \mathbb{E}m^k(T(s)) = O((s+1)^k).$$

Replacing s by $A(t)$ and using the upper bound on $A(t)$, we obtain (5.6). \square

PROPOSITION 5.5. *Let ψ be given by (5.3). We assume $\psi(0)$ has the stationary distribution of the solution to (5.3) so that the marginal distribution of $\psi(t)$ does not depend on t , nor do $k_1 \triangleq \frac{1}{t} \mathbb{E}\ell(t)$ and $k_2 \triangleq \frac{1}{t} \mathbb{E}m(t)$. Let $f: [a, b] \rightarrow \mathbb{R}$ be twice continuously differentiable. We have*

$$(5.9) \quad \mathbb{E}f(\psi(t)) = k_2 g(b) - k_1 g(a),$$

where

$$(5.10) \quad g(x) \triangleq \frac{1}{h(x)} \int_{\bar{x}}^x \frac{2f(y)h(y)}{\sigma^2(y)} dy, \quad h(x) \triangleq \exp \left\{ \int_{\bar{x}}^x \frac{2\mu(y)}{\sigma^2(y)} dy \right\},$$

and $\bar{x} \in [a, b]$. Furthermore,

$$(5.11) \quad k_2 - k_1 = \mathbb{E}\mu(\psi(t)), \quad k_2h(a) = k_1h(b).$$

Proof. It is straightforward to verify that $\frac{1}{2}\sigma^2(x)g'(x) + \mu(x)g(x) = f(x)$. Let $G(x) = \int_{\bar{x}}^x g(y) dy$, and apply Itô's formula to obtain

$$G(\psi(t)) = G(\psi(0)) + \int_0^t f(\psi(u)) du + \int_0^t g(\psi(u))\sigma(\psi(u)) dW(u) + g(a)\ell(t) - g(b)m(t).$$

Taking expectations, we obtain (5.9). Equation (5.3) implies

$$\psi(t) = \psi(0) + \int_0^t \mu(\psi(u)) du + \int_0^t \sigma(\psi(u)) dW(u) + \ell(t) - m(t),$$

and taking expectations, we have the first part of (5.11). Finally, the function $H(x) = \int_{\bar{x}}^x \frac{1}{h(y)} dy$ satisfies $\frac{1}{2}\sigma^2(x)H''(x) + \mu(x)H'(x) = 0$, and applying Itô's formula to H , we obtain

$$H(\psi(t)) = H(\psi(0)) + \int_0^t H'(\psi(u))\sigma(\psi(u)) dW(u) + \frac{\ell(t)}{h(a)} - \frac{m(t)}{h(b)}.$$

Taking expectations, we obtain the second part of (5.11). □

COROLLARY 5.6. *Under the assumptions of Proposition 5.5, with $\mu(x) = 0$ and $\sigma(x) = 1$ for every x , we have $\mathbb{E}\ell(t) = \mathbb{E}m(t) = \frac{t}{2(b-a)}$.*

Proof. In this case, $h(x) = 1$ for every x and (5.11) implies $\mathbb{E}\ell(t) = \mathbb{E}m(t)$. Taking $f(y) = 1$ for every y and $\bar{x} = a$, we obtain the desired result from (5.9). □

COROLLARY 5.7. *Let ψ be given by (5.3), and assume that $\sigma(x) = 1$ for all x . Then for all $t \geq 0$, $\mathbb{E}\ell(t) = \frac{t}{2(b-a)} + O(b-a) + O(t)$ and $\mathbb{E}m(t) = \frac{t}{2(b-a)} + O(b-a) + O(t)$.*

Proof. If $\mu(\cdot)$ is identically zero and $\psi(0)$ is a random variable having the stationary distribution of $\psi(\cdot)$ on $[a, b]$, then Corollary 5.6 implies $\mathbb{E}\ell(t) = \frac{t}{2(b-a)}$. If $\psi(0)$ is a nonrandom initial condition in $[a, b]$ and $\mu(\cdot)$ is not identically zero, then Lemma 5.3 implies $|\mathbb{E}\ell(t) - \frac{t}{2(b-a)}| \leq b-a + (\bar{\mu}^+ \vee \underline{\mu}^-)t$. The proof for $m(t)$ is the same. □

PROPOSITION 5.8. *With $\psi(\cdot)$ as in (5.3) and with $0 < b - a \leq 1$, let γ_0, γ_1 and γ_2 be arbitrary positive constants. Then there exist constants γ_3, γ_4 and γ_5 depending only on $\gamma_0, \gamma_1, \gamma_2, \underline{\mu}, \bar{\mu}$, and $\bar{\sigma}$ (and not depending on a, b, λ or t) such that for all λ satisfying*

$$(5.12) \quad 0 < \lambda \leq \gamma_3 \wedge (\gamma_4(b - a)),$$

we have $\mathbb{E}e^{\gamma_1\lambda\ell(t)+\gamma_2\lambda m(t)} \leq \gamma_5e^{\gamma_0t}$ for all $t \geq 0$.

Proof. We first construct a positive convex solution $u(x)$ to the Hamilton–Jacobi–Bellman equation

$$(5.13) \quad \max_{\substack{\underline{\mu} \leq \mu \leq \bar{\mu} \\ \underline{\sigma} \leq \sigma \leq \bar{\sigma}}} \left\{ -\gamma_0 u(x) + \mu u'(x) + \frac{1}{2} \sigma^2 u''(x) + 1 \right\} = 0$$

with boundary conditions

$$(5.14) \quad u'(a) + \gamma_1 \lambda u(a) = 0, \quad u'(b) - \gamma_2 \lambda u(b) = 0.$$

In (5.14), λ is a positive number satisfying (5.12) with γ_3 and γ_4 to be chosen later. We seek a solution of the form

$$(5.15) \quad -\gamma_0 u(x) + \underline{\mu} u'(x) + \frac{1}{2} \underline{\sigma}^2 u''(x) + 1 = 0, \quad a \leq x \leq \delta,$$

$$(5.16) \quad -\gamma_0 u(x) + \bar{\mu} u'(x) + \frac{1}{2} \bar{\sigma}^2 u''(x) + 1 = 0, \quad \delta \leq x \leq b,$$

where $a < \delta < b$ and

$$(5.17) \quad u(\delta) = \min_{a \leq x \leq b} u(x) > 0, \quad u'(\delta) = 0.$$

A convex function satisfying (5.15)–(5.17) will satisfy (5.13) (recall $\underline{\sigma} > 0$).

From (5.15) and (5.16), we see that u must be given by

$$(5.18) \quad u(x) = \begin{cases} \frac{1}{\gamma_0} + A_+ e^{xp_+} + A_- e^{xp_-} & \text{if } a \leq x \leq \delta, \\ \frac{1}{\gamma_0} + B_+ e^{xq_+} + B_- e^{xq_-} & \text{if } \delta \leq x \leq b, \end{cases}$$

where $p_{\pm} \triangleq \frac{1}{\underline{\sigma}^2} (-\underline{\mu} \pm \sqrt{\underline{\mu}^2 + 2\underline{\sigma}^2 \gamma_0})$ and $q_{\pm} \triangleq \frac{1}{\bar{\sigma}^2} (-\bar{\mu} \pm \sqrt{\bar{\mu}^2 + 2\bar{\sigma}^2 \gamma_0})$. Note that p_+ and q_+ are strictly positive and p_- and q_- are strictly negative. In order for u to satisfy (5.14) and the smooth pasting conditions $u(\delta-) = u(\delta+)$ and $u'(\delta-) = 0 = u'(\delta+)$, we must have

$$(5.19) \quad A_+(p_+ + \gamma_1 \lambda) e^{ap_+} + A_-(p_- + \gamma_1 \lambda) e^{ap_-} + \frac{\gamma_1 \lambda}{\gamma_0} = 0,$$

$$(5.20) \quad B_+(q_+ - \gamma_2 \lambda) e^{bq_+} + B_-(q_- - \gamma_2 \lambda) e^{bq_-} - \frac{\gamma_2 \lambda}{\gamma_0} = 0,$$

$$(5.21) \quad A_+ e^{\delta p_+} + A_- e^{\delta p_-} - B_+ e^{\delta q_+} - B_- e^{\delta q_-} = 0,$$

$$(5.22) \quad p_+ A_+ e^{\delta p_+} + p_- A_- e^{\delta p_-} = 0,$$

$$(5.23) \quad q_+ B_+ e^{\delta q_+} + q_- B_- e^{\delta q_-} = 0.$$

Define

$$(5.24) \quad f(x) = p_+(p_- + \gamma_1\lambda)e^{-(x-a)p_-} - p_-(p_+ + \gamma_1\lambda)e^{-(x-a)p_+},$$

$$(5.25) \quad g(x) = -q_+(q_- - \gamma_2\lambda)e^{(b-x)q_-} + q_-(q_+ - \gamma_2\lambda)e^{(b-x)q_+}.$$

Then (5.19), (5.22) and (5.20), (5.23) imply

$$(5.26) \quad \begin{aligned} A_+ &= \frac{\gamma_1\lambda p_-}{\gamma_0 f(\delta)} e^{-\delta p_+}, \\ A_- &= -\frac{p_+}{p_-} e^{\delta(p_+ - p_-)} A_+ = -\frac{\gamma_1\lambda p_+}{\gamma_0 f(\delta)} e^{-\delta p_-}, \end{aligned}$$

$$(5.27) \quad \begin{aligned} B_+ &= \frac{\gamma_2\lambda q_-}{\gamma_0 g(\delta)} e^{-\delta q_+}, \\ B_- &= -\frac{q_+}{q_-} e^{\delta(q_+ - q_-)} B_+ = -\frac{\gamma_2\lambda q_+}{\gamma_0 g(\delta)} e^{-\delta q_-}. \end{aligned}$$

In order for (5.21) to hold, δ must satisfy

$$(5.28) \quad \frac{f(\delta)}{\gamma_1(p_+ - p_-)} = \frac{g(\delta)}{\gamma_2(q_+ - q_-)}.$$

To obtain a solution to this equation, we define

$$(5.29) \quad \gamma_3 \triangleq \frac{|p_-|}{2\gamma_1} \wedge \frac{q_+}{2\gamma_2}, \quad \gamma_4 \triangleq \frac{\gamma_0}{(\gamma_1 + \gamma_2)\bar{\sigma}^2}$$

and consider only λ satisfying (5.12). For such λ we have $p_- + \gamma_1\lambda < 0$ and $q_+ - \gamma_2\lambda > 0$ so $f'(x) < 0$ and $g'(x) > 0$ for $a \leq x \leq b$. Since

$$(5.30) \quad \frac{f(a)}{\gamma_1(p_+ - p_-)} = \lambda = \frac{g(b)}{\gamma_2(q_+ - q_-)},$$

there must exist a unique $\delta \in (a, b)$ satisfying (5.28). We need also to show that $f(\delta) < 0$ and $g(\delta) < 0$ so A_{\pm} and B_{\pm} are positive. This will establish the convexity and positivity of u . Denote by

$$\delta_1 = a + \frac{1}{p_+ - p_-} \log \frac{p_-(p_+ + \gamma_1\lambda)}{p_+(p_- + \gamma_1\lambda)}, \quad \delta_2 = b - \frac{1}{q_+ - q_-} \log \frac{q_+(q_- - \gamma_2\lambda)}{q_-(q_+ - \gamma_2\lambda)}$$

the unique solutions of $f(\delta_1) = 0, g(\delta_2) = 0$. Since $\log(1 + x) < x$ for $x > 0$,

$$\begin{aligned} \delta_1 &= a + \frac{1}{p_+ - p_-} \log \left(1 + \frac{(p_- - p_+)\gamma_1\lambda}{p_+(p_- + \gamma_1\lambda)} \right) < a - \frac{\gamma_1\lambda}{p_+(p_- + \gamma_1\lambda)}, \\ \delta_2 &= b - \frac{1}{q_+ - q_-} \log \left(1 + \frac{(q_- - q_+)\gamma_2\lambda}{q_-(q_+ - \gamma_2\lambda)} \right) > b + \frac{\gamma_2\lambda}{q_-(q_+ - \gamma_2\lambda)}. \end{aligned}$$

But (5.12) and (5.29) imply $p_- + \gamma_1\lambda \leq \frac{1}{2}p_- < 0$ and $q_+ - \gamma_2\lambda \geq \frac{1}{2}q_+ > 0$. Therefore,

$$\delta_2 - \delta_1 > b - a + \frac{2\gamma_2\lambda}{q_-q_+} + \frac{2\gamma_1\lambda}{p_+p_-} = (b - a) - \frac{\bar{\sigma}^2(\gamma_1 + \gamma_2)\lambda}{\gamma_0} \geq 0$$

by the fact that $\lambda \leq \gamma_4(b - a)$. Since $\delta_2 > \delta_1$, we have $\delta \in (\delta_1, \delta_2)$ and $f(\delta) < 0$ and $g(\delta) < 0$.

We now take the argument of u to be the process ψ of (5.3) and use (5.13) and (5.14) to obtain

$$\begin{aligned} & d[e^{-\gamma_0 t + \gamma_1 \lambda \ell(t) + \gamma_2 \lambda m(t)} u(\psi(t))] \\ & \leq e^{-\gamma_0 t + \gamma_1 \lambda \ell(t) + \gamma_2 \lambda m(t)} [-1 dt + \sigma(\psi(t)) u'(\psi(t)) dW(t)]. \end{aligned}$$

Integration yields

$$\begin{aligned} (5.31) \quad 0 & \leq e^{-\gamma_0 t + \gamma_1 \lambda \ell(t) + \gamma_2 \lambda m(t)} u(\psi(t)) \\ & \leq u(\psi(0)) - \int_0^t e^{-\gamma_0 s + \gamma_1 \lambda \ell(s) + \gamma_2 \lambda m(s)} ds \\ & \quad + \int_0^t e^{-\gamma_0 s + \gamma_1 \lambda \ell(s) + \gamma_2 \lambda m(s)} \sigma(\psi(s)) u'(\psi(s)) dW(s). \end{aligned}$$

We see that the Itô integral in (5.31) is bounded below and hence is a supermartingale. Taking expectations in (5.31) and using the fact that $0 < u(\delta) \leq u(\psi(t))$, we obtain $\mathbb{E}e^{\gamma_1 \lambda \ell(t) + \gamma_2 \lambda m(t)} \leq e^{\gamma_0 t} u(\psi(0)) / u(\delta)$ for all $t \geq 0$. It remains only to show that there is a constant γ_5 depending only on p_{\pm} , q_{\pm} , γ_0 , γ_1 , and γ_2 such that

$$(5.32) \quad \frac{u(x)}{u(\delta)} \leq \gamma_5 \quad \forall x \in [a, b].$$

Being convex, the function u attains its maximum over $[a, b]$ at either a or b . Thus, to prove (5.32), it suffices to obtain a positive lower bound on $\frac{u(\delta)}{u(a)}$ and $\frac{u(\delta)}{u(b)}$. We compute

$$\begin{aligned} \frac{u(\delta)}{u(a)} &= \frac{\frac{1}{\gamma_0} + A_+ e^{\delta p_+} + A_- e^{\delta p_-}}{\frac{1}{\gamma_0} + A_+ e^{a p_+} + A_- e^{a p_-}} \\ &= \frac{f(\delta) + \gamma_1 \lambda p_- - \gamma_1 \lambda p_+}{f(\delta) + \gamma_1 \lambda p_- e^{-(\delta-a)p_+} - \gamma_1 \lambda p_+ e^{-(\delta-a)p_-}} \\ &= 1 + \frac{\gamma_1 \lambda}{p_+ p_-} \left[\frac{p_+ (e^{-(\delta-a)p_-} - 1) - p_- (e^{-(\delta-a)p_+} - 1)}{e^{-(\delta-a)p_-} - e^{-(\delta-a)p_+}} \right] \\ &= 1 - \frac{\bar{\sigma}^2 \gamma_1 \lambda}{2\gamma_0} h_1(\delta - a), \end{aligned}$$

where

$$h_1(x) = \frac{p_+ (e^{-x p_-} - 1) - p_- (e^{-x p_+} - 1)}{e^{-x p_-} - e^{-x p_+}}.$$

We have $\lim_{x \downarrow 0} h_1(x) = 0$ and $\lim_{x \rightarrow \infty} h_1(x) = p_+$. Hence, $\gamma_6 \triangleq \sup_{x > 0} h_1(x)$ is finite and depends only on p_{\pm} and q_{\pm} . So long as $0 < \lambda \leq \frac{\gamma_0}{\bar{\sigma}^2 \gamma_1 \gamma_6}$, we have $\frac{u(\delta)}{u(a)} \geq \frac{1}{2}$. We reduce γ_3 given by (5.29) if necessary so that $\gamma_3 \leq \frac{\gamma_0}{\bar{\sigma}^2 \gamma_1 \gamma_6}$.

A similar computation shows that

$$\frac{u(\delta)}{u(b)} = 1 - \frac{\bar{\sigma}^2 \gamma_2 \lambda}{2\gamma_0} h_2(b - \delta),$$

where

$$h_2(x) = \frac{q_+(e^{xq_-} - 1) - q_-(e^{xq_+} - 1)}{e^{xq_+} - e^{xq_-}}.$$

We have $\lim_{x \downarrow 0} h_2(x) = 0$ and $\lim_{x \rightarrow \infty} h_2(x) = -q_-$. Hence, $\gamma_7 \triangleq \sup_{x>0} h_2(x)$ is finite and depends only on p_{\pm} and q_{\pm} . So long as $0 < \lambda \leq \frac{\gamma_0}{\bar{\sigma}^2 \gamma_2 \gamma_7}$, we have $\frac{u(\delta)}{u(a)} \geq \frac{1}{2}$. We reduce γ_3 if necessary so that $\gamma_3 \leq \frac{\gamma_0}{\bar{\sigma}^2 \gamma_2 \gamma_7}$. For λ satisfying (5.12), the bound (5.32) and hence the conclusion of the proposition holds with $\gamma_5 = 2$. □

PROPOSITION 5.9. *With $\psi(\cdot)$ as in (5.3), let $\gamma_0 > 0$, $\gamma_1 < 0$, and $\gamma_2 < 0$ be given. For $a, b \in \mathbb{R}$ with $b - a > 0$ and sufficiently small and $0 < \lambda \leq 1$,*

$$(5.33) \quad \mathbb{E} \int_0^\infty e^{-\gamma_0 t + \gamma_1 \lambda \ell(t) + \gamma_2 \lambda m(t)} dt \leq \frac{1}{\gamma_0} \left[1 + \frac{\lambda \underline{\sigma}^2}{\frac{(b-a)\gamma_0}{\gamma_1 \vee \gamma_2} - \lambda \underline{\sigma}^2 + O((b-a)^2)} \right].$$

Proof. We first construct a concave solution $u(x)$ to the Hamilton–Jacobi–Bellman equation (5.13) satisfying the boundary conditions (5.14). Instead of (5.15)–(5.17), here we seek a concave solution of the form

$$(5.34) \quad -\gamma_0 u(x) + \bar{\mu} u'(x) + \frac{1}{2} \underline{\sigma}^2 u''(x) + 1 = 0, \quad a \leq x \leq \delta,$$

$$(5.35) \quad -\gamma_0 u(x) + \underline{\mu} u'(x) + \frac{1}{2} \underline{\sigma}^2 u''(x) + 1 = 0, \quad \delta \leq x \leq b,$$

where $a < \delta < b$ and

$$(5.36) \quad u(\delta) = \max_{a \leq x \leq b} u(x), \quad u'(\delta) = 0.$$

A concave function satisfying (5.34)–(5.36) will satisfy (5.13).

From (5.34) and (5.35), we see that u must be given by (5.18), where now $p_{\pm} \triangleq \frac{1}{\underline{\sigma}^2} (-\bar{\mu} \pm \sqrt{\bar{\mu}^2 + 2\underline{\sigma}^2 \gamma_0})$ and $q_{\pm} \triangleq \frac{1}{\underline{\sigma}^2} (-\underline{\mu} \pm \sqrt{\underline{\mu}^2 + 2\underline{\sigma}^2 \gamma_0})$. Then p_+ and q_+ are strictly positive, p_- and q_- are strictly negative, and

$$(5.37) \quad p_+ p_- = q_+ q_- = -\frac{2\gamma_0}{\underline{\sigma}^2}.$$

In order for u to satisfy (5.14) and the smooth pasting conditions $u(\delta-) = u(\delta+)$ and $u'(\delta-) = u'(\delta+) = 0$, equations (5.19)–(5.23) must hold. These imply (5.26), (5.27), where f and g are defined by (5.24) and (5.25). In order for (5.21) to hold, δ must satisfy (5.28). However, in contrast to the proof of

Proposition 5.8, here we do not need to restrict λ in order to obtain a solution to this equation. Because γ_1 and γ_2 are negative,

$$\begin{aligned} f'(x) &= e^{-(x-a)p_+} p_+ p_- \left[-(p_- + \gamma_1 \lambda) e^{(x-a)(p_+ - p_-)} + p_+ + \gamma_1 \lambda \right] \\ &\leq e^{-(x-a)p_+} p_+ p_- \left[-(p_- + \gamma_1 \lambda) + p_+ + \gamma_1 \lambda \right] < 0, \\ g'(x) &= e^{(b-x)q_-} q_+ q_- \left[q_- - \gamma_2 \lambda - (q_+ - \gamma_2 \lambda) e^{(b-x)(q_+ - q_-)} \right] \\ &\geq e^{(b-x)q_-} q_+ q_- \left[q_- - \gamma_2 \lambda - (q_+ - \gamma_2 \lambda) \right] > 0. \end{aligned}$$

Since (5.30) holds, there must exist a unique $\delta \in (a, b)$ satisfying (5.28). Furthermore, $f(a) = \gamma_1 \lambda(p_+ - p_-)$ and $g(b) = \gamma_2 \lambda(q_+ - q_-)$ are both negative, so $f(\delta)$ and $g(\delta)$ are also negative. This shows that A_{\pm} and B_{\pm} are negative, so u is concave. We have solved (5.13), (5.14) for the case of positive γ_0 and negative γ_1 and γ_2 . Furthermore, our solution satisfies (5.34)–(5.36).

From (5.13) and (5.14), we obtain (5.31). Taking expectations and then letting $t \rightarrow \infty$ in (5.31), we obtain

$$(5.38) \quad \mathbb{E} \int_0^\infty e^{-\gamma_0 t + \gamma_1 \lambda \ell(t) + \gamma_2 \lambda m(t)} dt \leq u(\psi(0)) \leq u(\delta).$$

To complete the proof, it remains only to show that the right-hand side of (5.33) dominates $u(\delta)$.

We begin by observing that if $a < \delta \leq \frac{a+b}{2}$, then $f(\frac{a+b}{2}) \leq f(\delta)$, whereas, if $\frac{a+b}{2} \leq \delta < b$, then $g(\frac{a+b}{2}) \leq g(\delta)$. According to (5.18), (5.26), and (5.27)

$$(5.39) \quad u(\delta) = \frac{1}{\gamma_0} \left[1 - \frac{\gamma_1 \lambda(p_+ - p_-)}{f(\delta)} \right] = \frac{1}{\gamma_0} \left[1 - \frac{\gamma_2 \lambda(q_+ - q_-)}{g(\delta)} \right].$$

Because $-\frac{\gamma_1 \lambda(p_+ - p_-)}{f(\delta)}$ is negative, we increase this term by replacing $f(\delta)$ by a negative quantity with larger absolute value, that is, by a quantity smaller than $f(\delta)$. If $a < \delta \leq \frac{a+b}{2}$, we replace $f(\delta)$ by $f(\frac{a+b}{2})$ and obtain

$$(5.40) \quad u(\delta) \leq \frac{1}{\gamma_0} \left[1 - \frac{\gamma_1 \lambda(p_+ - p_-)}{f(\frac{a+b}{2})} \right].$$

If $\frac{a+b}{2} \leq \delta < b$, we replace $g(\delta)$ by $g(\frac{a+b}{2})$ in the last expression in (5.39) and instead obtain

$$(5.41) \quad u(\delta) \leq \frac{1}{\gamma_0} \left[1 - \frac{\gamma_2 \lambda(q_+ - q_-)}{g(\frac{a+b}{2})} \right].$$

According to (5.24), (5.37), and Taylor’s theorem,

$$\begin{aligned} f\left(\frac{a+b}{2}\right) &= p_+(p_- + \gamma_1 \lambda) e^{-\frac{1}{2}(b-a)p_-} - p_-(p_+ + \gamma_1 \lambda) e^{-\frac{1}{2}(b-a)p_+} \\ &= \left(\gamma_1 \lambda - \frac{(b-a)\gamma_0}{\sigma^2} \right) (p_+ - p_-) + O((b-a)^2), \end{aligned}$$

$$\begin{aligned}
 g\left(\frac{a+b}{2}\right) &= -q_+(q_- - \gamma_2\lambda)e^{\frac{1}{2}(b-a)q_-} + q_-(q_+ - \gamma_2\lambda)e^{\frac{1}{2}(b-a)q_+} \\
 &= \left(\gamma_2\lambda - \frac{(b-a)\gamma_0}{\sigma^2}\right)(q_+ - q_-) + O((b-a)^2).
 \end{aligned}$$

Substitution of these formulas into (5.40) and (5.41) shows that $u(\delta)$ is dominated by $1/\gamma_0$ times the maximum of

$$\left[1 + \frac{\gamma_1\lambda}{\frac{(b-a)\gamma_0}{\sigma^2} - \gamma_1\lambda + O((b-a)^2)}\right] \quad \text{and} \quad \left[1 + \frac{\gamma_2\lambda}{\frac{(b-a)\gamma_0}{\sigma^2} - \gamma_2\lambda + O((b-a)^2)}\right].$$

This is the right-hand side of (5.33), provided $b - a$ is sufficiently small. \square

5.2. Proof of Theorem 4.4. Solving (4.9) and (4.14), we see that

$$\begin{aligned}
 (5.42) \quad X_2(t) &= \exp\left\{\int_0^t \left(r - c + \alpha\theta_2(u) - \frac{1}{2}\sigma^2\theta_2^2(u)\right) du \right. \\
 &\quad \left. + \int_0^t \sigma\theta_2(u) dW(u) - \lambda(\ell_2(t) + m_2(t))\right\},
 \end{aligned}$$

$$\begin{aligned}
 (5.43) \quad X_1(t) &= \exp\left\{\int_0^t \left(r - c + \alpha\theta_2(u) - \frac{1}{2}\sigma^2\theta_2^2(u)\right) du \right. \\
 &\quad \left. + \int_0^t \sigma\theta_2(u) dW(u)\right\}.
 \end{aligned}$$

We consider first the case $p \neq 1$, for which we have

$$\begin{aligned}
 (5.44) \quad X_1^{1-p}(t) - X_2^{1-p}(t) &= X_1^{1-p}(t) [1 - e^{-\lambda(1-p)(\ell_2(t)+m_2(t))}] \\
 &= Z_2(t)\zeta(t) \exp\left\{(1-p) \int_0^t \left(r - c + \alpha\theta_2(u) - \frac{1}{2}p\sigma^2\theta_2^2(u)\right) du\right\},
 \end{aligned}$$

where

$$\begin{aligned}
 Z_2(t) &\triangleq \exp\left\{(1-p)\sigma \int_0^t \theta_2(u) dW(u) - \frac{1}{2}(1-p)^2\sigma^2 \int_0^t \theta_2^2(u) du\right\}, \\
 \zeta(t) &\triangleq 1 - e^{-\lambda(1-p)(\ell_2(t)+m_2(t))}.
 \end{aligned}$$

The right-hand side of (5.44) has the same sign as $\zeta(t)$, which is positive if $0 < p < 1$ and negative if $p > 1$. Regardless of whether $0 < p < 1$ or $p > 1$, $U_p(cX_1(t)) = \frac{c^{1-p}}{1-p} X_1^{1-p}(t) \geq \frac{c^{1-p}}{1-p} X_2^{1-p}(t) = U_p(cX_2(t))$, which implies

$$(5.45) \quad u_1(c, c(\cdot), w_1, w_2) - u_2(c, c(\cdot), w_1, w_2) \geq 0.$$

We introduce a Brownian motion \widetilde{W} under a probability measure $\widetilde{\mathbb{P}}$ and consider an auxiliary process $\widetilde{\theta}(\cdot)$ satisfying $\widetilde{\theta}(0) = \theta_2(0)$ and

$$(5.46) \quad d\widetilde{\theta}(t) = \widetilde{\theta}(t)(-r + c(\widetilde{\theta}(t)) - \alpha\widetilde{\theta}(t) + p\sigma^2\widetilde{\theta}^2(t)) dt - \sigma\widetilde{\theta}^2(t) d\widetilde{W}(t) \\ + (1 + \lambda\bar{\theta}(1 - w_1)) d\widetilde{\ell}(t) - (1 - \lambda\bar{\theta}(1 + w_2)) d\widetilde{m}(t),$$

where $\widetilde{\ell}(\cdot)$ and $\widetilde{m}(\cdot)$ are the minimal nondecreasing processes that keep $\widetilde{\theta}(\cdot)$ in the interval $[\bar{\theta}(1 - w_1), \bar{\theta}(1 + w_2)]$. Following (5.42)–(5.43), we introduce

$$\widetilde{X}_2(t) = \exp\left\{ \int_0^t \left(r - c + \alpha\widetilde{\theta}(u) - \frac{1}{2}\sigma^2\widetilde{\theta}^2(u) \right) du \right. \\ \left. + \int_0^t \sigma\widetilde{\theta}(u) d\widetilde{W}(u) - \lambda(\widetilde{\ell}(t) + \widetilde{m}(t)) \right\}, \\ \widetilde{X}_1(t) = \exp\left\{ \int_0^t \left(r - c + \alpha\widetilde{\theta}(u) - \frac{1}{2}\sigma^2\widetilde{\theta}^2(u) \right) du + \int_0^t \sigma\widetilde{\theta}(u) d\widetilde{W}(u) \right\}.$$

Then just as in (5.44), we have

$$\widetilde{X}_1^{1-p}(t) - \widetilde{X}_2^{1-p}(t) \\ = \widetilde{Z}_2(t)\widetilde{\zeta}(t) \exp\left\{ (1-p) \int_0^t \left(r - c + \alpha\widetilde{\theta}(u) - \frac{1}{2}p\sigma^2\widetilde{\theta}^2(u) \right) du \right\},$$

where

$$\widetilde{Z}_2(t) \triangleq \exp\left\{ (1-p)\sigma \int_0^t \widetilde{\theta}(u) d\widetilde{W}(u) - \frac{1}{2}(1-p)^2\sigma^2 \int_0^t \widetilde{\theta}^2(u) du \right\}, \\ \widetilde{\zeta}(t) = 1 - e^{-\lambda(1-p)(\widetilde{\ell}(t) + \widetilde{m}(t))}.$$

Because $\theta_2(\cdot)$ is bounded, Z_2 is a martingale. Fix $T > 0$ and define a new probability measure \mathbb{P}_2^T by $\frac{d\mathbb{P}_2^T}{d\mathbb{P}} = Z_2(T)$. Under \mathbb{P}_2^T , the process

$$W_2^T(t) \triangleq W(t) - (1-p)\sigma \int_0^t \theta_2(u) du, \quad 0 \leq t \leq T,$$

is a Brownian motion. We may rewrite (4.8) as

$$(5.47) \quad d\theta_2(t) = \theta_2(t)(-r + c(\theta_2(t)) - \alpha\theta_2(t) + p\sigma^2\theta_2^2(t)) dt \\ - \sigma\theta_2^2(t) dW_2^T(t) + (1 + \lambda\bar{\theta}(1 - w_1)) d\ell_2(t) \\ - (1 - \lambda\bar{\theta}(1 + w_2)) dm_2(t).$$

Comparing (5.47) and (5.46), we conclude that the four-dimensional process $(X_1(t), X_2(t), \zeta(t), \theta_2(t); 0 \leq t \leq T)$ has the same law under \mathbb{P}_2^T as the process $(\widetilde{X}_1(t), \widetilde{X}_2(t), \widetilde{\zeta}(t), \widetilde{\theta}(t); 0 \leq t \leq T)$ under $\widetilde{\mathbb{P}}$.

The term $\exp\{(1-p)\int_0^t(r-c+\alpha\theta_2(u)-\frac{1}{2}p\sigma^2\theta_2^2(u))du\}$ in (5.44) is nearly deterministic for small w_1 and w_2 . To exploit this fact, we define

$$\Delta(t) \triangleq (1-p)\int_0^t\left(\alpha(\theta_2(u)-\bar{\theta})-\frac{1}{2}p\sigma^2(\theta_2^2(u)-\bar{\theta}^2)\right)du$$

and the analogue

$$\tilde{\Delta}(t) \triangleq (1-p)\int_0^t\left(\alpha(\tilde{\theta}(u)-\bar{\theta})-\frac{1}{2}p\sigma^2(\tilde{\theta}^2(u)-\bar{\theta}^2)\right)du.$$

We consider only $w_1 > 0, w_2 > 0$ such that $w \triangleq w_1 + w_2 \leq 1$, and for such w_1, w_2 , there exists a constant k independent of w_1, w_2 such that

$$(5.48) \quad |\Delta(t)| \leq kwt, \quad |\tilde{\Delta}(t)| \leq kwt.$$

Let $t \geq 0$ be given and choose $T \geq t$. Using (1.5) and (4.4), we may write

$$(5.49) \quad \begin{aligned} \mathbb{E}X_1^{1-p}(t) - \mathbb{E}X_2^{1-p}(t) &= \mathbb{E}[Z_2(t)\zeta(t)e^{(\beta-pA(p)-(1-p)c)t+\Delta(t)}] \\ &= e^{(\beta-pA(p)-(1-p)c)t}\mathbb{E}_2^T[\zeta(t)e^{\Delta(t)}] \\ &= e^{(\beta-pA(p)-(1-p)c)t}\tilde{\mathbb{E}}[\tilde{\zeta}(t)e^{\tilde{\Delta}(t)}]. \end{aligned}$$

According to Taylor’s theorem,

$$(5.50) \quad \tilde{\zeta}(t) = \lambda(1-p)(\tilde{\ell}(t) + \tilde{m}(t)) - \frac{1}{2}\lambda^2(1-p)^2(\tilde{\ell}(t) + \tilde{m}(t))^2e^{\xi(t)},$$

where $\xi(t)$ is between 0 and $-\lambda(1-p)(\tilde{\ell}(t) + \tilde{m}(t))$. We introduce the time change $A(t) \triangleq \int_0^t\sigma^2\tilde{\theta}^4(u)du$, the inverse time change $T(s) \triangleq A^{-1}(s)$, and the $\tilde{\mathbb{P}}$ -Brownian motion $B(s) \triangleq -\int_0^{T(s)}\sigma\tilde{\theta}^2(u)d\tilde{W}(u)$. Defining $\psi(s) \triangleq \tilde{\theta}(T(s))$, we rewrite (5.46) as

$$d\psi(s) = \frac{1}{\sigma^2\psi^3(s)}(-r+c(\psi(s))-\alpha\psi(s)+p\sigma^2\psi^2(s))ds + dB(s) + \ell(s) - m(s),$$

where $\ell(s) \triangleq (1+\lambda\bar{\theta}(1-w_1))\tilde{\ell}(T(s))$ and $m(s) \triangleq (1-\lambda\bar{\theta}(1+w_2))\tilde{m}(T(s))$. Corollary 5.7 implies $\tilde{\mathbb{E}}\ell(s) = \frac{s}{2\theta w} + O(w) + O(s)$, and since

$$\frac{\ell(\sigma^2\bar{\theta}^4(1-w_1)^4t)}{1+\lambda\bar{\theta}(1-w_1)} \leq \tilde{\ell}(t) \leq \frac{\ell(\sigma^2\bar{\theta}^4(1+w_2)^4t)}{1+\lambda\bar{\theta}(1-w_1)},$$

we see that

$$(5.51) \quad \tilde{\mathbb{E}}[\tilde{\ell}(t)] = \frac{\sigma^2\bar{\theta}^3t}{2w} + O(\lambda tw^{-1}) + O(1) + O(t).$$

The same applies to $\tilde{m}(t)$, which leads to

$$(5.52) \quad \tilde{\mathbb{E}}[\tilde{\ell}(t) + \tilde{m}(t)] = \frac{\sigma^2\bar{\theta}^3t}{w} + O(\lambda tw^{-1}) + O(1) + O(t).$$

Let ε be a fixed positive constant and assume w_1 and w_2 are sufficiently small so that $kw < \varepsilon$. Then

$$\begin{aligned} & \int_0^\infty e^{-\varepsilon t} \tilde{\mathbb{E}} |(\tilde{\ell}(t) + \tilde{m}(t))(e^{\tilde{\Delta}(t)} - 1)| dt \\ & \leq \int_0^\infty e^{-\varepsilon t} \tilde{\mathbb{E}} [(\tilde{\ell}(t) + \tilde{m}(t))(e^{kw t} - 1)] dt \\ & = \int_0^\infty [e^{-\varepsilon t + kw t} - e^{-\varepsilon t}] \left[\frac{\sigma^2 \bar{\theta}^3 t}{w} + O(\lambda t w^{-1}) + O(1) + O(t) \right] dt \\ & = \left(\frac{\sigma^2 \bar{\theta}^3}{w} + O(\lambda w^{-1}) + O(1) \right) \left(\frac{1}{(\varepsilon - kw)^2} - \frac{1}{\varepsilon^2} \right) \\ & \quad + \left(\frac{1}{\varepsilon - kw} - \frac{1}{\varepsilon} \right) O(1) \\ & = O(1). \end{aligned}$$

It follows that (recall (4.1))

$$\begin{aligned} (5.53) \quad & \tilde{\mathbb{E}} \int_0^\infty e^{-(pA(p)+(1-p)c)t} (\tilde{\ell}(t) + \tilde{m}(t)) e^{\tilde{\Delta}(t)} dt \\ & = \tilde{\mathbb{E}} \int_0^\infty e^{-(pA(p)+(1-p)c)t} (\tilde{\ell}(t) + \tilde{m}(t)) dt \\ & \quad + \tilde{\mathbb{E}} \int_0^\infty e^{-(pA(p)+(1-p)c)t} (\tilde{\ell}(t) + \tilde{m}(t)) (e^{\tilde{\Delta}(t)} - 1) dt \\ & = \int_0^\infty e^{-(pA(p)+(1-p)c)t} \left(\frac{\sigma^2 \bar{\theta}^3 t}{w} + O(\lambda t w^{-1}) + O(1) + O(t) \right) dt \\ & \quad + O(1) \\ & = \frac{\sigma^2 \bar{\theta}^3}{(pA(p) + (1-p)c)^2 w} + O(\lambda w^{-1}) + O(1). \end{aligned}$$

Returning to (5.49) and using (5.50) and (5.53), we compute

$$\begin{aligned} & u_1(c, c(\cdot), w_1, w_2) - u_2(c, c(\cdot), w_1, w_2) \\ & = \frac{c^{1-p}}{1-p} \int_0^\infty e^{-\beta t} (\mathbb{E} X_1^{1-p}(t) - \mathbb{E} X_2^{1-p}(t)) dt \\ & = \frac{c^{1-p}}{1-p} \int_0^\infty e^{-(pA(p)+(1-p)c)t} \tilde{\mathbb{E}} [\tilde{\zeta}(t) e^{\tilde{\Delta}(t)}] dt \\ & = \lambda c^{1-p} \tilde{\mathbb{E}} \int_0^\infty e^{-(pA(p)+(1-p)c)t} (\tilde{\ell}(t) + \tilde{m}(t)) e^{\tilde{\Delta}(t)} dt \\ & \quad - \frac{1}{2} \lambda^2 (1-p) c^{1-p} \tilde{\mathbb{E}} \int_0^\infty e^{-(pA(p)+(1-p)c)t} (\tilde{\ell}(t) + \tilde{m}(t))^2 e^{\xi(t) + \tilde{\Delta}(t)} dt \end{aligned}$$

$$\begin{aligned}
 &= \frac{c^{1-p}\sigma^2\bar{\theta}^3}{(pA(p) + (1-p)c)^2} \cdot \frac{\lambda}{w} + O(\lambda^2w^{-1}) + O(\lambda) \\
 &\quad + \frac{1}{2}\lambda^2(p-1)c^{1-p}\tilde{\mathbb{E}} \int_0^\infty e^{-(pA(p)+(1-p)c)t} (\tilde{\ell}(t) + \tilde{m}(t))^2 e^{\xi(t)+\tilde{\Delta}(t)} dt.
 \end{aligned}$$

If $p > 1$, the last term is nonnegative, and we have

$$u_1(c, c(\cdot), w_1, w_2) - u_2(c, c(\cdot), w_1, w_2) \geq C_1\lambda w^{-1} + O(\lambda)$$

for some positive constant C_1 . Combining this with (5.45), we obtain (4.16). If $p > 1$ and $\lambda/w = o(1)$, then the hypotheses of Proposition 5.8 are satisfied by the process $\tilde{\theta}(\cdot)$ of (5.46) with $b - a = \bar{\theta}w$, $\gamma_1 = \gamma_2 = 2(p - 1)$, and $\gamma_0 > 0$ chosen to satisfy $-(pA(p) + (1 - p)c) + kw + \gamma_0 < 0$ (w sufficiently small). This proposition, together with Proposition 5.4 and Hölder’s inequality, implies

$$\begin{aligned}
 &\tilde{\mathbb{E}}[(\tilde{\ell}(t) + \tilde{m}(t))^2 e^{\xi(t)+\tilde{\Delta}(t)}] \\
 &\leq e^{kwt} \tilde{\mathbb{E}}[(\tilde{\ell}(t) + \tilde{m}(t))^2 e^{\xi(t)}] \\
 &\leq e^{kwt} (\tilde{\mathbb{E}}[(\tilde{\ell}(t) + \tilde{m}(t))^4])^{1/2} (\tilde{\mathbb{E}}[e^{2\xi(t)}])^{1/2} \\
 &= e^{kwt+\gamma_0 t} O((t+1)^2w^{-2}).
 \end{aligned}$$

It follows that

$$(5.54) \quad \lambda^2 \tilde{\mathbb{E}} \int_0^\infty e^{-(pA(p)+(1-p)c)t} (\tilde{\ell}(t) + \tilde{m}(t))^2 e^{\xi(t)+\tilde{\Delta}(t)} dt = O(\lambda^2w^{-2}),$$

and (4.17) is proved for the case $p > 1$.

If $0 < p < 1$, then $e^{\xi(t)} \leq 1$, so

$$\tilde{\mathbb{E}}[(\tilde{\ell}(t) + \tilde{m}(t))^2 e^{\xi(t)+\tilde{\Delta}(t)}] \leq e^{kwt} O((t+1)^2w^{-2}).$$

For w sufficiently small so that $-(pA(p) + (1 - p)c) + kw < 0$, we again have (5.54), which implies (4.17). The assumption $\lambda/w = o(1)$ is not needed in the proof of (4.17) in the case $0 < p < 1$.

To obtain (4.16) when $0 < p < 1$, we choose $\gamma_0 > pA(p) + (1 - p)c$, set

$$\gamma_1 = \frac{p-1}{1+\lambda\bar{\theta}(1-w_1)}, \quad \gamma_2 = \frac{p-1}{1-\lambda\bar{\theta}(1+w_2)}, \quad \underline{\sigma} = \sigma\bar{\theta}^2(1-w_1)^2,$$

and note that $\gamma_1 \vee \gamma_2 = \frac{p-1}{1+\lambda\bar{\theta}(1-w_1)}$. Recalling (5.46), we see that Proposition 5.9 implies for sufficiently small w that

$$\begin{aligned}
 &\tilde{\mathbb{E}} \int_0^\infty e^{-\gamma_0 t + \lambda(p-1)\tilde{\ell}(t) + \lambda(p-1)\tilde{m}(t)} dt \\
 &\leq \frac{1}{\gamma_0} + \frac{\lambda(p-1)\sigma^2\bar{\theta}^4(1-w_1)^4}{\gamma_0^2w(1+\lambda\bar{\theta}(1-w_1)) + \lambda(1-p)\gamma_0\sigma^2\bar{\theta}^4(1-w_1)^4 + O(w^2)}
 \end{aligned}$$

$$\begin{aligned} &\leq \frac{1}{\gamma_0} - \frac{\lambda(1-p)\sigma^2\bar{\theta}^4(1-w_1)^4}{2\max\{\gamma_0^2w(1+\lambda\bar{\theta}(1-w_1)) + O(w^2), \lambda(1-p)\gamma_0\sigma^2\bar{\theta}^4(1-w_1)^4\}} \\ &\leq \frac{1}{\gamma_0} - \frac{1}{2} \min\left\{\frac{(1-p)\sigma^2\bar{\theta}^4\lambda}{2\gamma_0^2w}, \frac{1}{\gamma_0}\right\} = \frac{1}{\gamma_0} - \min\{C'_1\lambda w^{-1}, C'_2\} \end{aligned}$$

for positive constants C'_1 and C'_2 . Because $0 < p < 1$, we have $\tilde{\zeta}(t) \geq 0$ and (5.48), (5.49) imply for $w > 0$ sufficiently small that $\mathbb{E}X_1^{1-p}(t) - \mathbb{E}X_2^{1-p}(t) \geq e^{(\beta-\gamma_0)t}\tilde{\mathbb{E}}\tilde{\zeta}(t)$. Therefore,

$$\begin{aligned} (5.55) \quad &u_1(c, c(\cdot), w_1, w_2) - u_2(c, c(\cdot), w_1, w_2) \\ &= \frac{c^{1-p}}{1-p} \int_0^\infty e^{-\beta t} (\mathbb{E}X_1^{1-p}(t) - \mathbb{E}X_2^{1-p}(t)) dt \\ &\geq \frac{c^{1-p}}{1-p} \int_0^\infty e^{-\gamma_0 t} \tilde{\mathbb{E}}\tilde{\zeta}(t) dt \\ &= \frac{c^{1-p}}{1-p} \tilde{\mathbb{E}} \left[\int_0^\infty e^{-\gamma_0 t} (1 - e^{\lambda(p-1)(\tilde{\ell}(t) + \tilde{m}(t))}) dt \right] \\ &\geq \min\{C_1\lambda w^{-1}, C_2\} \end{aligned}$$

for positive constants C_1 and C_2 . This combined with (5.45) yields (4.16).

If $p = 1$, then $\mathbb{P}_2^T = \mathbb{P}$. Let $t \geq 0$ be given and choose $T \geq t$. We observe from (5.42), (5.43), (5.51), and the counterpart of (5.51) for $\tilde{m}(t)$ that

$$\begin{aligned} \mathbb{E} \log X_1(t) - \mathbb{E} \log X_2(t) &= \lambda \mathbb{E} [\ell_2(t) + m_2(t)] \\ &= \lambda \mathbb{E}_2^T [\ell_2(t) + m_2(t)] \\ &= \lambda \tilde{\mathbb{E}} [\tilde{\ell}(t) + \tilde{m}(t)] \\ &= \frac{\lambda \sigma^2 \bar{\theta}^3 t}{w} + O(\lambda^2 t w^{-1}) + O(\lambda) + O(\lambda t), \end{aligned}$$

which is obviously nonnegative. Multiplying by $e^{-\beta t}$ and integrating from $t = 0$ to $t = \infty$, we obtain (4.17) once we recall that $A(1) = \beta$. Indeed, we obtain (4.17) with the term $O(\lambda^2 w^{-1})$ (a special case of $O(\lambda^2 w^{-2})$ in place of the term $O(\lambda^2 w^{-2})$), and this version of (4.17) yields (4.16).

5.3. Proof of Theorem 4.5. We introduce a Brownian motion \widehat{W} under a probability measure \mathbb{P} and consider the auxiliary process $\hat{\theta}(\cdot)$ satisfying

$$\begin{aligned} (5.56) \quad d\hat{\theta}(t) &= \hat{\theta}(t)(-r + c(\hat{\theta}(t)) - \alpha\hat{\theta}(t) + \sigma^2\hat{\theta}^2(t) - (1-p)\sigma^2\hat{\theta}^2(t)\bar{\theta}) dt \\ &\quad - \sigma\hat{\theta}^2(t) d\widehat{W}(t) + d\hat{\ell}(t) - d\hat{m}(t), \end{aligned}$$

where $\hat{\ell}(\cdot)$ and $\hat{m}(\cdot)$ are the minimal nondecreasing processes that keep $\hat{\theta}(\cdot)$ in the interval $[\bar{\theta}(1-w_1), \hat{\theta}(1+w_2)]$. We assume the initial condition $\hat{\theta}(0)$ has the stationary distribution of the solution to (5.56), so the distribution of $\hat{\theta}(t)$

under $\widehat{\mathbb{P}}$ does not depend on t . This stationary distribution exists because the drift term in (5.56) is bounded and the diffusion terms does not vanish, so the process is recurrent, and hence the conditions of Theorem 3.3 of [13] are satisfied.

Define the martingale $\overline{Z}(t) \triangleq \exp\{(1-p)\sigma\overline{\theta}W(t) - \frac{1}{2}(1-p)^2\sigma^2\overline{\theta}^2 t\}$. For fixed $T > 0$, define the probability measure $\overline{\mathbb{P}}^T$ by $\frac{d\overline{\mathbb{P}}^T}{d\mathbb{P}} = \overline{Z}(T)$, under which $\overline{W}^T(t) \triangleq W(t) - (1-p)\sigma\overline{\theta}t$, $0 \leq t \leq T$, is a Brownian motion and (4.8) becomes

$$(5.57) \quad d\theta_2(t) = \theta_2(t)(-r + c(\theta_2(t)) - \alpha\theta_2(t) + \sigma^2\theta_2^2(t) - (1-p)\sigma^2\theta_2^2(t)\overline{\theta}) dt - \sigma\theta_2^2(t) d\overline{W}^T(t) + d\ell(t) - dm(t),$$

where $\ell(t) = (1 + \lambda\overline{\theta}(1 - w_1))\ell_2(t)$, $m(t) = (1 - \lambda\overline{\theta}(1 + w_2))m_2(t)$. We assume $\theta_2(0)$ has the stationary distribution of the solution of (5.56), so that $(\theta_2(t); 0 \leq t \leq T)$ has the same law under $\overline{\mathbb{P}}^T$ as the process $(\widehat{\theta}(t); 0 \leq t \leq T)$ under $\widehat{\mathbb{P}}$. In particular,

$$(5.58) \quad \overline{\mathbb{E}}^T [(\theta_2(t) - \overline{\theta})^2] = \widehat{\mathbb{E}} [(\widehat{\theta}(t) - \overline{\theta})^2], \quad 0 \leq t \leq T.$$

We show that

$$(5.59) \quad \widehat{\mathbb{E}} [(\widehat{\theta}(t) - \overline{\theta})^2] = \frac{1}{3}\overline{\theta}^2(w_1^2 - w_1w_2 + w_2^2) + O(\lambda w^2) + O(w^3).$$

To establish (5.59) we appeal to Proposition 5.5 with $a = \overline{\theta}(1 - w_1)$, $b = \overline{\theta}(1 + w_2)$, $\sigma(x) = -\sigma x^2$ and $\mu(x) = x(-r + c(x) - \alpha x + \sigma^2 x^2 - (1-p)\sigma^2 x^2 \overline{\theta})$. Recall that we consider only functions $c(\cdot)$ that are bounded uniformly in λ and vary over $[\overline{\theta}(1 - w_1), \overline{\theta}(1 + w_2)]$ by no more than $O(\lambda)$ (see Remark 4.2). Therefore, for $y \in [\overline{\theta}(1 - w_1), \overline{\theta}(1 + w_2)]$, we have $\mu(y) = \mu(\overline{\theta}) + O(\lambda) + O(w)$ and $\sigma^2(y) = \sigma^2\overline{\theta}^4 + O(w)$. With $\overline{x} = \overline{\theta}$ in Proposition 5.5, for $\overline{\theta}(1 - w_1) \leq x \leq \overline{\theta}(1 + w_2)$, we have

$$\begin{aligned} \int_{\overline{\theta}}^x \frac{2\mu(y)}{\sigma^2(y)} dy &= \int_{\overline{\theta}}^x \left[\frac{2\mu(\overline{\theta})}{\sigma^2\overline{\theta}^4} + O(\lambda) + O(w) \right] dy \\ &= \frac{2\mu(\overline{\theta})}{\sigma^2\overline{\theta}^4}(x - \overline{\theta}) + O(\lambda w) + O(w^2), \\ h(x) &= 1 + \frac{2\mu(\overline{\theta})}{\sigma^2\overline{\theta}^4}(x - \overline{\theta}) + O(\lambda w) + O(w^2). \end{aligned}$$

Equations (5.11) imply $k_2 - k_1 = \mu(\overline{\theta}) + O(\lambda) + O(w)$ and

$$k_2 \left(1 - \frac{2\mu(\overline{\theta})w_1}{\sigma^2\overline{\theta}^3} + O(\lambda w) + O(w^2) \right) = k_1 \left(1 + \frac{2\mu(\overline{\theta})w_2}{\sigma^2\overline{\theta}^3} + O(\lambda w) + O(w^2) \right),$$

which yield

$$k_1 \left(\frac{2\mu(\overline{\theta})w}{\sigma^2\overline{\theta}^3} + O(\lambda w) + O(w^2) \right) = \mu(\overline{\theta}) + O(\lambda) + O(w),$$

and this implies

$$k_1 = \frac{\sigma^2 \bar{\theta}^3}{2w} + O(\lambda w^{-1}) + O(1), \quad k_2 = \frac{\sigma^2 \bar{\theta}^3}{2w} + O(\lambda w^{-1}) + O(1).$$

Following (5.10) with $f(y) = (y - \bar{\theta})^2$, we compute

$$\begin{aligned} g(x) &= \frac{1}{h(x)} \int_{\bar{\theta}}^x \frac{2(y - \bar{\theta})^2 h(y)}{\sigma^2(y)} dy \\ &= \frac{1}{1 + O(w)} \int_{\bar{\theta}}^x \frac{2(y - \bar{\theta})^2 (1 + O(w))}{\sigma^2 \bar{\theta}^4 + O(w)} dy \\ &= \left(\frac{2}{\sigma^2 \bar{\theta}^4} + O(w) \right) \int_{\bar{\theta}}^x (y - \bar{\theta})^2 dy \\ &= \frac{2}{3\sigma^2 \bar{\theta}^4} (x - \bar{\theta})^3 + O(w^4). \end{aligned}$$

According to (5.9), $\widehat{\mathbb{E}}[(\hat{\theta}(t) - \bar{\theta})^2] = k_2 g(\bar{\theta}(1 + w_2)) - k_1 g(\bar{\theta}(1 - w_1))$, which is (5.59).

We now consider the case $p \neq 1$. From (5.43), (4.3) and (1.5), we have

$$\begin{aligned} &\frac{X_1^{1-p}(t)}{\bar{Z}(t) \mathbb{E} X_0^{1-p}(t)} \\ &= \exp \left\{ (1-p) \int_0^t \left(\alpha(\theta_2(u) - \bar{\theta}) - \frac{1}{2} \sigma^2 \theta_2^2(u) + \frac{1}{2} p \sigma^2 \bar{\theta}^2 \right) du \right. \\ &\quad \left. + (1-p) \int_0^t \sigma \theta_2(u) dW(u) - (1-p) \sigma \bar{\theta} W(t) + \frac{1}{2} (1-p)^2 \sigma^2 \bar{\theta}^2 t \right\} \\ &= \exp \left\{ (1-p) \sigma \int_0^t (\theta_2(u) - \bar{\theta}) d\bar{W}^T(u) \right. \\ &\quad \left. - \frac{1}{2} (1-p) \sigma^2 \int_0^t (\theta_2(u) - \bar{\theta})^2 du \right\}. \end{aligned}$$

For arbitrary $t \geq 0$, we choose $T \geq t$ and have

$$\begin{aligned} (5.60) \quad \mathbb{E} X_1^{1-p}(t) &= \mathbb{E} X_0^{1-p}(t) \cdot \bar{\mathbb{E}}^T \exp \left\{ (1-p) \sigma \int_0^t (\theta_2(u) - \bar{\theta}) d\bar{W}^T(u) \right. \\ &\quad \left. - \frac{1}{2} (1-p) \sigma^2 \int_0^t (\theta_2(u) - \bar{\theta})^2 du \right\} \\ &= \mathbb{E} X_0^{1-p}(t) \cdot \widehat{\mathbb{E}} \exp \left\{ (1-p) \sigma \int_0^t (\hat{\theta}(u) - \bar{\theta}) d\widehat{W}(u) \right. \\ &\quad \left. - \frac{1}{2} (1-p) \sigma^2 \int_0^t (\hat{\theta}(u) - \bar{\theta})^2 du \right\}. \end{aligned}$$

Because

$$M(t) \triangleq \exp \left\{ (1-p)\sigma \int_0^t (\widehat{\theta}(u) - \bar{\theta}) d\widehat{W}(u) - \frac{1}{2}(1-p)^2\sigma^2 \int_0^t (\widehat{\theta}(u) - \bar{\theta})^2 du \right\}$$

is a $\widehat{\mathbb{P}}$ -martingale, for $0 < p < 1$,

$$\begin{aligned} (5.61) \quad & \widehat{\mathbb{E}} \exp \left\{ (1-p)\sigma \int_0^t (\widehat{\theta}(u) - \bar{\theta}) d\widehat{W}(u) - \frac{1}{2}(1-p)\sigma^2 \int_0^t (\widehat{\theta}(u) - \bar{\theta})^2 du \right\} \\ &= \widehat{\mathbb{E}} \left[M(t) \exp \left\{ -\frac{1}{2}p(1-p)\sigma^2 \int_0^t (\widehat{\theta}(u) - \bar{\theta})^2 du \right\} \right] \\ &\leq \widehat{\mathbb{E}} M(t) \\ &= 1, \end{aligned}$$

and (5.60) implies $\mathbb{E}X_1^{1-p}(t) \leq \mathbb{E}X_0^{1-p}(t)$. If $p > 1$, the inequality in (5.61) is reversed and $\mathbb{E}X_1^{1-p}(t) \geq \mathbb{E}X_0^{1-p}(t)$. Regardless of whether $0 < p < 1$ or $p > 1$, $\mathbb{E}U_p(cX_1(t)) = \frac{c^{1-p}}{1-p} \mathbb{E}X_1^{1-p}(t) \leq \frac{c^{1-p}}{1-p} \mathbb{E}X_0^{1-p}(t) = \mathbb{E}U_p(cX_0(t))$. The inequality in (4.18) follows from (4.15) and (4.5).

It remains to compute the $\widehat{\mathbb{E}}$ expectation on the right-hand side of (5.60). To simplify the notation, we set

$$I(t) \triangleq (1-p)\sigma \int_0^t (\widehat{\theta}(u) - \bar{\theta}) d\widehat{W}(u), \quad R(t) \triangleq -\frac{1}{2}(1-p)\sigma^2 \int_0^t (\widehat{\theta}(u) - \bar{\theta})^2 du,$$

so that the expectation we need to compute is

$$\begin{aligned} (5.62) \quad & \widehat{\mathbb{E}} [\exp(I(t) + R(t))] = \widehat{\mathbb{E}} \left[1 + I(t) + R(t) + \frac{1}{2}(I(t) + R(t))^2 \right] \\ &+ \widehat{\mathbb{E}} \sum_{n=3}^{\infty} \frac{1}{n!} (I(t) + R(t))^n. \end{aligned}$$

We first bound the remainder

$$(5.63) \quad \left| \widehat{\mathbb{E}} \sum_{n=3}^{\infty} \frac{1}{n!} (I(t) + R(t))^n \right| \leq \sum_{n=3}^{\infty} \frac{2^n}{n!} \widehat{\mathbb{E}} [|I(t)|^n + |R(t)|^n].$$

Because $\langle I \rangle(t) \leq k_3 w^2 t$, where $k_3 = (1-p)^2 \sigma^2 \bar{\theta}^2$, there is a Brownian motion \widehat{B} such that $\max_{0 \leq s \leq t} |I(s)| \leq \max_{0 \leq s \leq k_3 w^2 t} |\widehat{B}(s)|$. Doob's maximal martingale inequality implies that for integers $n \geq 2$,

$$\begin{aligned} \widehat{\mathbb{E}} \left[\max_{0 \leq s \leq k_3 w^2 t} |\widehat{B}(s)|^n \right] &\leq \left(\frac{n}{n-1} \right)^n \widehat{\mathbb{E}} [|\widehat{B}(k_3 w^2 t)|^n] \\ &= \left(\frac{n}{n-1} \right)^n k_3^{\frac{n}{2}} w^n t^{\frac{n}{2}} \widehat{\mathbb{E}} [|\widehat{B}(1)|^n]. \end{aligned}$$

It can be verified by integration by parts and induction that for $n \geq 1$,

$$\widehat{\mathbb{E}}[|\widehat{B}(1)|^{2n}] = \frac{(2n)!}{2^n n!}, \quad \widehat{\mathbb{E}}[|\widehat{B}(1)|^{2n+1}] = \sqrt{\frac{2}{\pi}} 2^n n!.$$

Because $(2^n n!)^2 \leq (2n+1)!$,

$$\begin{aligned} (5.64) \quad & \sum_{n=1}^{\infty} \frac{2^{2n+1}}{(2n+1)!} \widehat{\mathbb{E}}[|I(t)|^{2n+1}] \\ & \leq \sqrt{\frac{2}{\pi}} \sum_{n=1}^{\infty} \frac{2^{2n+1}}{(2n+1)!} \left(\frac{2n+1}{2n}\right)^{2n+1} k_3^{n+\frac{1}{2}} w^{2n+1} t^{n+\frac{1}{2}} 2^n n! \\ & \leq 3w \sqrt{\frac{2k_3 t}{\pi}} \sum_{n=1}^{\infty} \frac{1}{n!} \left(\frac{9}{2} k_3 w^2 t\right)^n \\ & = O(w^3 t^{\frac{3}{2}}) e^{O(w^2)t}. \end{aligned}$$

On the other hand,

$$\begin{aligned} (5.65) \quad & \sum_{n=2}^{\infty} \frac{2^{2n}}{(2n)!} \widehat{\mathbb{E}}[|I(t)|^{2n}] \\ & = \sum_{n=2}^{\infty} \frac{2^{2n}}{(2n)!} \left(\frac{2n}{2n-1}\right)^{2n} k_3^n w^{2n} t^n \frac{(2n)!}{2^n n!} \\ & \leq \sum_{n=2}^{\infty} \frac{1}{n!} \left(\frac{32}{9} k_3 w^2 t\right)^n \\ & = O(w^4 t^2) e^{O(w^2)t}. \end{aligned}$$

Finally,

$$(5.66) \quad \sum_{n=3}^{\infty} \frac{2^n}{n!} \widehat{\mathbb{E}}[|R(t)|^n] \leq \sum_{n=3}^{\infty} \frac{1}{n!} (|1-p|\sigma^2 \bar{\theta}^2 w^2 t)^n = O(w^6 t^3) e^{O(w^2)t}.$$

Combining (5.64)–(5.66), we have obtained the bound $O(w^3 t^{3/2} + w^4 t^2 + w^6 t^3) e^{O(w^2)t}$ on the expression in (5.63).

For the other terms in (5.62), we use (5.59) to compute

$$\begin{aligned} & \widehat{\mathbb{E}}I(t) = 0, \\ & \widehat{\mathbb{E}}R(t) = -\frac{1}{2}(1-p)\sigma^2 \int_0^t \widehat{\mathbb{E}}(\theta_2(u) - \bar{\theta})^2 du \\ & \quad = -\frac{1}{6}(1-p)\sigma^2 \bar{\theta}^2 (w_1^2 - w_1 w_2 + w_2^2)t + O(\lambda w^2 t) + O(w^3 t), \\ & \widehat{\mathbb{E}}I^2(t) = (1-p)^2 \sigma^2 \int_0^t \widehat{\mathbb{E}}(\theta_2(u) - \bar{\theta})^2 du \\ & \quad = \frac{1}{3}(1-p)^2 \sigma^2 \bar{\theta}^2 (w_1^2 - w_1 w_2 + w_2^2)t + O(\lambda w^2 t) + O(w^3 t), \end{aligned}$$

$$\widehat{\mathbb{E}}R^2(t) = O(w^4t^2),$$

$$|\widehat{\mathbb{E}}[I(t)R(t)]| \leq \widehat{\mathbb{E}}^{1/2}[I^2(t)]\widehat{\mathbb{E}}^{1/2}[R^2(t)] = O(w^3t^{\frac{3}{2}}).$$

From (5.60), (5.62) and the above estimates, we see that

$$(5.67) \quad \mathbb{E}X_1^{1-p}(t) = \mathbb{E}X_0^{1-p}(t) \left(1 - \frac{1}{6}p(1-p)\sigma^2\bar{\theta}^2(w_1^2 - w_1w_2 + w_2^2)t \right. \\ \left. + O(\lambda w^2t) + O(w^3t) \right. \\ \left. + O(w^3t^{\frac{3}{2}} + w^4t^2 + w^6t^3)e^{O(w^2)t} \right).$$

Recall from (4.3), (4.4) that $e^{-\beta t}\mathbb{E}X_0^{1-p}(t) = e^{-(pA(p)+(1-p)c)t}$. For sufficiently small w , the $O(w^2)$ term in (5.67) satisfies $-pA(p) - (1-p)c + O(w^2) < 0$ (we are still working under the condition (4.1)), and this implies (4.18):

$$u_1(c, c(\cdot), w_1, w_2) \\ = \frac{c^{1-p}}{1-p} \int_0^\infty e^{-\beta t} \mathbb{E}X_1^{1-p}(t) dt \\ = \frac{c^{1-p}}{1-p} \int_0^\infty e^{-\beta t} \mathbb{E}X_0^{1-p}(t) dt \\ - \frac{1}{6}p\sigma^2\bar{\theta}^2(w_1^2 - w_1w_2 + w_2^2)c^{1-p} \int_0^\infty te^{-(pA(p)+(1-p)c)t} dt \\ + O(\lambda w^2) + O(w^3) \\ = u_0(c) - \frac{p\sigma^2\bar{\theta}^2(w_1^2 - w_1w_2 + w_2^2)c^{1-p}}{6(pA(p) + (1-p)c)^2} + O(\lambda w^2) + O(w^3).$$

If $p = 1$, then $\overline{\mathbb{P}}^T = \mathbb{P}$ and (4.2), (5.43), the fact that $\alpha = \sigma^2\bar{\theta}$ (see (1.5)), and (5.58), (5.59) imply

$$\mathbb{E} \log X_0(t) - \mathbb{E} \log X_1(t) \\ = \frac{1}{2}\sigma^2 \int_0^t \widehat{\mathbb{E}}[(\widehat{\theta}(u) - \bar{\theta})^2] du \\ = \frac{1}{6}\sigma^2\bar{\theta}^2(w_1^2 - w_1w_2 + w_2^2)t + O(\lambda w^2t) + O(w^3t).$$

Multiplying by $e^{-\beta t}$ and integrating out t , we obtain (4.18) (recall $A(1) = \beta$).

5.4. Optimizing over the c , w_1 and w_2 . Recall the positive numbers $w_1(\lambda)$ and $w_2(\lambda)$ of (4.23) that minimize g_λ and satisfy (4.25).

LEMMA 5.10. *Let λ_0 be a positive constant, and let $x_1(\cdot)$ and $x_2(\cdot)$ be mappings from $(0, \lambda_0)$ into $(0, \infty]$ such that $\lim_{\lambda \downarrow 0} x_1(\lambda) = \lim_{\lambda \downarrow 0} x_2(\lambda) = 0$.*

Assume that for some $q \in (2/3, 1]$, we have

$$(5.68) \quad u_2(c, c(\cdot), w_1(\lambda), w_2(\lambda)) \leq u_2(c, c(\cdot), x_1(\lambda), x_2(\lambda)) + O(\lambda^q).$$

Then

$$(5.69) \quad x_i(\lambda) = O(\lambda^{1/3}), \quad x_i(\lambda) = w_i(\lambda) + O(\lambda^{q/2}), \quad i = 1, 2,$$

and

$$(5.70) \quad u_2(c, c(\cdot), w_1(\lambda), w_2(\lambda)) = u_2(c, c(\cdot), x_1(\lambda), x_2(\lambda)) + O(\lambda^{q/2+1/3}).$$

In this lemma, u_2 is computed under the assumption that $\theta_2(\cdot)$ has the stationary distribution of the processes $\hat{\theta}(\cdot)$ given by (5.56). The $O(\cdot)$ terms in (5.68)–(5.70) are uniform over the number c and the function $c(\cdot)$ within the class described by Remark 4.2.

Proof. Define $x(\lambda) \triangleq x_1(\lambda) + x_2(\lambda)$. We first show that

$$(5.71) \quad \liminf_{\lambda \downarrow 0} \frac{x(\lambda)}{\lambda^{1/3}} > 0.$$

If this were not the case, then we could choose a sequence $\lambda_n \downarrow 0$ and positive numbers $k_n \rightarrow \infty$ such that

$$(5.72) \quad \lambda_n^{1/3} \geq k_n x(\lambda_n) \quad \forall n.$$

From (4.25), (5.68), (4.18), (4.16), and (5.72) we would have

$$\begin{aligned} & \frac{c^{1-p} \sigma^2 \bar{\theta}^{8/3}}{(pA(p) + (1-p)c)^2} \left(\frac{9p}{32}\right)^{1/3} \\ &= \lambda_n^{-2/3} [u_0(c) - u_2(c, c(\cdot), w_1(\lambda_n), w_2(\lambda_n))] + O(\lambda_n^{1/3}) \\ &\geq \lambda_n^{-2/3} [u_0(c) - u_2(c, c(\cdot), x_1(\lambda_n), x_2(\lambda_n))] + O(\lambda_n^{q-2/3}) \\ &\geq \lambda_n^{-2/3} [u_1(c, c(\cdot), x_1(\lambda_n), x_2(\lambda_n)) - u_2(c, c(\cdot), x_1(\lambda_n), x_2(\lambda_n))] \\ &\quad + O(\lambda_n^{q-2/3}) \\ &\geq \min \left\{ \frac{C_1 \lambda_n^{1/3}}{x(\lambda_n)}, C_2 \lambda_n^{-2/3} \right\} + O(\lambda_n^{q-2/3}) \\ &\geq \min \{ C_1 k_n, C_2 \lambda_n^{-2/3} \} + O(\lambda_n^{q-2/3}). \end{aligned}$$

The last term has limit infinity as $n \rightarrow \infty$. This contradiction implies (5.71).

We next show that

$$(5.73) \quad \limsup_{\lambda \downarrow 0} \frac{x(\lambda)}{\lambda^{1/3}} < \infty.$$

If this were not the case, then we could choose a sequence $\lambda_n \downarrow 0$ and positive numbers $K_n \rightarrow \infty$ such that

$$(5.74) \quad x(\lambda_n) \geq K_n \lambda_n^{1/3} \quad \forall n.$$

From (4.25), (5.68), (4.16), and (4.18) we would have

$$\begin{aligned} & \frac{c^{1-p}\sigma^2\bar{\theta}^{8/3}}{(pA(p) + (1-p)c)^2} \left(\frac{9p}{32}\right)^{1/3} \\ &= \lambda_n^{-2/3} [u_0(c) - u_2(c, c(\cdot), w_1(\lambda_n), w_2(\lambda_n))] + O(\lambda_n^{1/3}) \\ &\geq \lambda_n^{-2/3} [u_0(c) - u_2(c, c(\cdot), x_1(\lambda_n), x_2(\lambda_n))] + O(\lambda_n^{q-2/3}) \\ &\geq \lambda_n^{-2/3} [u_0(c) - u_1(c, c(\cdot), x_1(\lambda_n), x_2(\lambda_n))] + O(\lambda_n^{q-2/3}) \\ &= \frac{c^{1-p}p\sigma^2\bar{\theta}^2}{6(pA(p) + (1-p)c)^2} \cdot \frac{x_1^2(\lambda_n) - x_1(\lambda_n)x_2(\lambda_n) + x_2^2(\lambda_n)}{\lambda_n^{2/3}} \\ &\quad + O(\lambda_n^{q-2/3}) + O(\lambda_n^{1/3}x^2(\lambda_n)) + O(\lambda_n^{-2/3}x^3(\lambda_n)). \end{aligned}$$

However, (5.71) implies that for some constant C ,

$$\lambda_n^{1/3}x^2(\lambda_n) = \frac{\lambda_n^{1/3}}{x(\lambda_n)}x^3(\lambda_n) \leq Cx^3(\lambda_n) \leq C\lambda_n^{-2/3}x^3(\lambda_n) = O(\lambda_n^{q-2/3}).$$

Furthermore,

$$x_1^2(\lambda) - x_1(\lambda)x_2(\lambda) + x_2^2(\lambda) = \frac{1}{4}x^2(\lambda) + \frac{3}{4}(x_1(\lambda) - x_2(\lambda))^2 \geq \frac{1}{4}x^2(\lambda).$$

Therefore,

$$\begin{aligned} & \frac{c^{1-p}\sigma^2\bar{\theta}^{8/3}}{(pA(p) + (1-p)c)^2} \left(\frac{9p}{32}\right)^{1/3} \\ &\geq \frac{1}{4} \left[\frac{c^{1-p}p\sigma^2\bar{\theta}^2}{6(pA(p) + (1-p)c)^2} + O(x(\lambda_n)) \right] \frac{x^2(\lambda_n)}{\lambda_n^{2/3}} + O(\lambda_n^{q-2/3}). \end{aligned}$$

This last expression has limit infinity as $n \rightarrow \infty$ because of (5.74), and this contradiction implies (5.73).

From (5.71), we see that $\lambda/x(\lambda) = o(1)$. From (5.71) and (5.73) we conclude that every cluster point of $\lambda^{-1/3}x(\lambda)$ is in $(0, \infty)$ and a cluster point exists. Let us call such a cluster point L , and passing to a subsequence if necessary, we assume without loss of generality that $L_1 \triangleq \lim_{n \rightarrow \infty} \lambda_n^{-1/3}x_1(\lambda_n)$ and $L_2 \triangleq \lim_{n \rightarrow \infty} \lambda_n^{-1/3}x_2(\lambda_n)$ exist and, of course, $L = L_1 + L_2$. Using the notation (4.20), the equality in (4.19) implies

$$\begin{aligned} & u_0(c) - u_2(c, c(\cdot), x_1(\lambda_n), x_2(\lambda_n)) \\ &= \frac{c^{1-p}\sigma^2\bar{\theta}^2}{(pA(p) + (1-p)c)^2} g_\lambda(x_1(\lambda_n), x_2(\lambda_n)) \\ &= \frac{c^{1-p}\sigma^2\bar{\theta}^2}{(pA(p) + (1-p)c)^2} g_1(\lambda_n^{-1/3}x_1(\lambda_n), \lambda_n^{-1/3}x_2(\lambda_n)) \lambda_n^{2/3}. \end{aligned}$$

Dividing this by $\lambda^{2/3}$ and taking the limit as $n \rightarrow \infty$, we now use (5.68) and (4.25), to obtain

$$\begin{aligned}
 (5.75) \quad & \frac{c^{1-p}\sigma^2\bar{\theta}^2}{(pA(p) + (1-p)c)^2} g_1(L_1, L_2) \\
 &= \lim_{n \rightarrow \infty} \frac{u_0(c) - u_2(c, c(\cdot), x_1(\lambda_n), x_2(\lambda_n))}{\lambda_n^{2/3}} \\
 &\leq \lim_{n \rightarrow \infty} \frac{u_0(c) - u_2(c, c(\cdot), w_1(\lambda_n), w_2(\lambda_n))}{\lambda_n^{2/3}} \\
 &= \frac{c^{1-p}\sigma^2\bar{\theta}^{8/3}}{(pA(p) + (1-p)c)^2} \left(\frac{9p}{32}\right)^{1/3}.
 \end{aligned}$$

But the minimal value of $g_1(L_1, L_2)$ over $L_1 \geq 0$ and $L_2 \geq 0$ such that $L_1 + L_2 = L \in (0, \infty)$, uniquely attained by

$$(5.76) \quad L_1 = L_2 = \left(\frac{3\bar{\theta}}{2p}\right)^{1/3}$$

(cf. (4.23)), is $\bar{\theta}^{2/3}(\frac{9p}{32})^{1/3}$. We conclude that the inequality in (5.75) is equality and (5.76) holds. Since this is the case for every cluster point of $\lambda_n^{-1/3}x_1(\lambda_n)$ and $\lambda_n^{-1/3}x_2(\lambda_n)$, then even without passing to a subsequence, we must have

$$(5.77) \quad \lim_{\lambda \downarrow 0} \frac{x_1(\lambda)}{\lambda^{1/3}} = \lim_{\lambda \downarrow 0} \frac{x_2(\lambda)}{\lambda^{1/3}} = \left(\frac{3\bar{\theta}}{2p}\right)^{1/3} = \frac{w_1(\lambda)}{\lambda^{1/3}} = \frac{w_2(\lambda)}{\lambda^{1/3}}.$$

This provides the first equality in (5.69).

We show that

$$(5.78) \quad \limsup_{\lambda \downarrow 0} \frac{1}{\lambda^{q/2}} [|x_1(\lambda) - w_1(\lambda)| + |x_2(\lambda) - w_2(\lambda)|] < \infty,$$

which is just a restatement of the second equality in (5.69). If this were not the case, there would exist a sequence $\lambda_n \downarrow 0$ and a sequence of positive numbers $K_n \rightarrow \infty$ such that

$$(5.79) \quad |x_1(\lambda_n) - w_1(\lambda_n)| + |x_2(\lambda_n) - w_2(\lambda_n)| \geq K_n \lambda_n^{q/2} \quad \forall n.$$

We observe from (4.22) that

$$\nabla^2 g_\lambda(w_1, w_2) \geq \frac{p}{3} \begin{bmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{bmatrix},$$

where inequality of matrices is in the sense of a positive semidefinite difference. The operator norm of $\nabla^2 g_\lambda$ thus satisfies

$$(5.80) \quad \|\nabla^2 g_\lambda(w_1, w_2)\|^2 \geq \frac{p^2}{9} \max_{x_1^2+x_2^2=1} [x_1, x_2] \begin{bmatrix} 1 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \frac{p^2}{6}.$$

For $0 \leq s \leq 1$ and $i = 1, 2$, set $y_i(s, \lambda_n) = sx_i(\lambda_n) + (1 - s)w_i(\lambda_n)$. Then

$$\begin{aligned} & \frac{d^2}{ds^2} g_{\lambda_n}(y_1(s, \lambda_n), y_2(s, \lambda_n)) \\ &= \begin{bmatrix} x_1(\lambda_n) - w_1(\lambda_n) \\ x_2(\lambda_n) - w_2(\lambda_n) \end{bmatrix}^{tr} \nabla^2 g_{\lambda_n}(y_1(s, \lambda_n), y_2(s, \lambda_n)) \begin{bmatrix} x_1(\lambda_n) - w_1(\lambda_n) \\ x_2(\lambda_n) - w_2(\lambda_n) \end{bmatrix} \\ &\geq \frac{p^2}{6} [(x_1(\lambda_n) - w_1(\lambda_n))^2 + (x_2(\lambda_n) - w_2(\lambda_n))^2]. \end{aligned}$$

Using the the fact that $\nabla g_{\lambda_n}(w_1(\lambda_n), w_2(\lambda_n)) = 0$, we integrate from $s = 0$ to $s = t$ to obtain

$$\frac{d}{dt} g_{\lambda_n}(y_1(t, \lambda_n), y_2(t, \lambda_n)) \geq \frac{p^2}{6} [(x_1(\lambda_n) - w_1(\lambda_n))^2 + (x_2(\lambda_n) - w_2(\lambda_n))^2] t.$$

A second integration, this time from $t = 0$ to $t = 1$, the equivalence of all norms in \mathbb{R}^2 , and (5.79) yield

$$\begin{aligned} (5.81) \quad & g_{\lambda_n}(x_1(\lambda_n), x_2(\lambda_n)) \\ &\geq g_{\lambda_n}(w_1(\lambda_n), w_2(\lambda_n)) \\ &\quad + \frac{p^2}{12} [(x_1(\lambda_n) - w_1(\lambda_n))^2 + (x_2(\lambda_n) - w_2(\lambda_n))^2] \\ &\geq g_{\lambda_n}(w_1(\lambda_n), w_2(\lambda_n)) \\ &\quad + K' (|x_1(\lambda_n) - w_1(\lambda_n)| + |x_2(\lambda_n) - w_2(\lambda_n)|)^2 \\ &\geq g_{\lambda_n}(w_1(\lambda_n), w_2(\lambda_n)) + K' K_n^2 \lambda_n^q \end{aligned}$$

for some constant $K' > 0$. From (5.81), (4.19), the fact that $x(\lambda_n) = O(\lambda^{1/3})$, (5.68), (4.25) and (4.24), we have

$$\begin{aligned} & \frac{c^{1-p} \sigma^2 \bar{\theta}^3}{(pA(p) + (1-p)c)^2} \left[\frac{g_{\lambda_n}(w_1(\lambda_n), w_2(\lambda_n))}{\lambda_n^{2/3}} + K' K_n^2 \lambda_n^{q-2/3} \right] \\ &\leq \frac{c^{1-p} \sigma^2 \bar{\theta}^3}{(pA(p) + (1-p)c)^2} \cdot \frac{g_{\lambda_n}(x_1(\lambda_n), x_2(\lambda_n))}{\lambda_n^{2/3}} \\ &= \lambda_n^{-2/3} [u_0(c) - u_2(c, c(\cdot), x_1(\lambda_n), x_2(\lambda_n))] + O(\lambda_n^{1/3}) \\ &\leq \lambda_n^{-2/3} [u_0(c) - u_2(c, c(\cdot), w_1(\lambda_n), w_2(\lambda_n))] + O(\lambda_n^{q-2/3}) \\ &\leq \frac{c^{1-p} \sigma^2 \bar{\theta}^2}{(pA(p) + (1-p)c)^2} \frac{g_{\lambda_n}(w_1(\lambda_n), w_2(\lambda_n))}{\lambda_n^{2/3}} + O(\lambda_n^{q-2/3}). \end{aligned}$$

Canceling the term involving g_{λ_n} on both sides of this equality, we obtain

$$K' K_n^2 \lambda_n^{q-2/3} \leq O(\lambda_n^{q-2/3}),$$

which is impossible because $K' K_n^2 \rightarrow \infty$. This shows that the second equality in (5.69) must hold.

Because $w_1(\lambda)$ is a positive constant times $\lambda^{1/3}$, the second inequality in (5.69) can be rewritten as $x_i(\lambda) = w_i(\lambda)(1 + O(\lambda^{q/2-1/3}))$, and hence

$$\begin{aligned} g_\lambda(x_1(\lambda), x_2(\lambda)) &= \frac{\lambda \bar{\theta}}{w_1(\lambda) + w_2(\lambda)} (1 + O(\lambda^{q/2-1/3})) \\ &\quad + \frac{p}{6} (w_1^2(\lambda) - w_1(\lambda)w_2(\lambda) + w_2^2(\lambda)) (1 + O(\lambda^{q/2-1/3})) \\ &= g_\lambda(w_1(\lambda), w_2(\lambda)) + O(\lambda^{q/2+1/3}). \end{aligned}$$

Equation (5.70) follows from Corollary 4.6. □

REMARK 5.11. We actually expect $u_2(c, c_2(\cdot), w_1, w_2)$ to be maximized by $(x_1(\lambda), x_2(\lambda))$ satisfying a slightly stronger version of (5.69), namely, $x_1(\lambda) = w_1(\lambda) + O(\lambda^{2/3})$ and $x_2(\lambda) = w_2(\lambda) + O(\lambda^{2/3})$, in which case we could replace (5.70) by $u_2(c, c_1(\cdot), w_1(\lambda), w_2(\lambda)) = u_2(c, c_2(\cdot), x_1(\lambda), x_2(\lambda)) + O(\lambda)$.

We may now optimize $u_2(c, c(\cdot), w_1, w_2)$ over $(w_1, w_2) \in (0, \infty)^2$.

COROLLARY 5.12. *Recall the function φ of Assumption 4.1. We have*

$$\begin{aligned} (5.82) \quad &\sup_{\substack{w_1 > 0, w_2 > 0 \\ w_1 + w_2 \leq \varphi(\lambda)}} u_2(c, c(\cdot), w_1, w_2) \\ &= u_0(c) - \frac{c^{1-p} \sigma^2 \bar{\theta}^{8/3}}{(pA(p) + (1-p)c)^2} \left(\frac{9p}{32}\right)^{1/3} \lambda^{2/3} + O(\lambda^{5/6}). \end{aligned}$$

Here u_2 is computed under the assumption that $\theta_2(\cdot)$ has the stationary distribution of the process $\hat{\theta}(\cdot)$ given by (5.56).

Proof. Because $O(\lambda)$ is a special case of $O(\lambda^{5/6})$, (4.25) implies

$$\begin{aligned} &\sup_{\substack{w_1 > 0, w_2 > 0 \\ w_1 + w_2 \leq \varphi(\lambda)}} u_2(c, c(\cdot), w_1, w_2) \\ &\geq u_2(c, c(\cdot), w_1(\lambda), w_2(\lambda)) \\ &= u_0(c) - \frac{c^{1-p} \sigma^2 \bar{\theta}^{8/3}}{(pA(p) + (1-p)c)^2} \left(\frac{9p}{32}\right)^{1/3} \lambda^{2/3} + O(\lambda^{5/6}). \end{aligned}$$

The reverse inequality follows from Lemma 5.10 with $q = 1$. □

Finally, we optimize over c . When $p = 1$, $A(p) + (1-p)c = A(1) = \beta$ and the maximal value in (5.82), attained by $c = A(1) = \beta$, is (see (4.5) and (4.6))

$$\begin{aligned} (5.83) \quad &u_0(A(1)) - \frac{\sigma^2 \bar{\theta}^{8/3}}{\beta^2} \left(\frac{9}{32}\right)^{1/3} \lambda^{2/3} + O(\lambda^{5/6}) \\ &= v_0(1) - \frac{\sigma^2 \bar{\theta}^{8/3}}{\beta^2} \left(\frac{9}{32}\right)^{1/3} \lambda^{2/3} + O(\lambda^{5/6}). \end{aligned}$$

For $p \neq 1$, we need the following lemma. Because $A(1) = \beta$, (5.83) is a special case of (5.84) below.

LEMMA 5.13. *Choose $a \in (0, A(p))$ and $b \in (A(p), \infty)$ if $0 < p \leq 1$ or $b \in (A(p), \frac{pA(p)}{p-1})$ if $p > 1$. Then*

$$\begin{aligned}
 (5.84) \quad & \sup_{c \in [a, b]} \sup_{\substack{w_1 > 0, w_2 > 0 \\ w_1 + w_2 \leq \varphi(\lambda)}} u_2(c, c(\cdot), w_1, w_2) \\
 &= \sup_{c \in [a, b]} \left[u_0(c) - \frac{c^{1-p} \sigma^2 \bar{\theta}^{8/3}}{(pA(p) + (1-p)c)^2} \left(\frac{9p}{32} \right)^{1/3} \lambda^{2/3} + O(\lambda^{5/6}) \right] \\
 &= v_0(1) - \frac{\sigma^2 \bar{\theta}^{8/3}}{A^{1+p}(p)} \left(\frac{9p}{32} \right)^{1/3} \lambda^{2/3} + O(\lambda^{5/6}).
 \end{aligned}$$

Proof. Because of (5.83), we only need to consider the case $p \neq 1$. To simplify notation, we denote $\eta = \sigma^2 \bar{\theta}^{8/3} (9p/32)^{1/3}$ and recall the definition (4.5) of u_0 to write

$$(5.85) \quad f(c) \triangleq u_0(c) - \frac{c^{1-p} \eta \lambda^{2/3}}{(pA(p) + (1-p)c)^2} = u_0(c) \left[1 - \frac{(1-p)\eta \lambda^{2/3}}{pA(p) + (1-p)c} \right].$$

We will show that

$$(5.86) \quad \sup_{c \in [a, b]} f(c) = v_0(1) - \frac{\eta \lambda^{2/3}}{A^{1+p}(p)} + O(\lambda).$$

Because c in the maximization in (5.84) is restricted to $[a, b]$, the $O(\lambda^{5/6})$ term in (5.84) is bounded by $\lambda^{5/6}$ times a constant independent of c , the left-hand side of (5.84) is equal to $(\sup_{c \in [a, b]} f(c)) + O(\lambda^{5/6})$, and (5.84) will follow.

We compute

$$f'(c) = \frac{c^{-p}}{(pA(p) + (1-p)c)^2} \left[p(A(p) - c) + \frac{(1-p)\eta \lambda^{2/3} ((1+p)c - pA(p))}{pA(p) + (1-p)c} \right].$$

For sufficiently small $\lambda > 0$, $f'(a) > 0$. The expression $\frac{(1+p)c - pA(p)}{pA(p) + (1-p)c}$ is increasing in c . If $0 < p < 1$, this expression is bounded and its derivative with respect to c is also bounded. Therefore, $\lim_{c \rightarrow \infty} f'(c) = -\infty$ and thus f' has a zero in $[a, \infty)$. For sufficiently small λ , the expression in square brackets is strictly decreasing and hence f' cannot have more than one zero. If $p > 1$, then $\lim_{c \uparrow \frac{pA(p)}{p-1}} \frac{(1+p)c - pA(p)}{pA(p) + (1-p)c} = \infty$, so $\lim_{c \uparrow \frac{pA(p)}{p-1}} f'(c) = -\infty$. In this case, the term in square brackets is strictly decreasing, so again f' has exactly one zero in $[a, \infty)$. In both cases, the zero of f' corresponds to a maximum value of f .

For $c \in [a, b]$,

$$f'(c) = \frac{c^{-p}[p(A(p) - c) + O(\lambda^{2/3})]}{(pA(p) + (1 - p)c)^2},$$

where the $O(\lambda^{2/3})$ term is bounded by a constant independent of $c \in [a, b]$ times $\lambda^{2/3}$. For $\varepsilon > 0$, $f'(A(p) - \varepsilon\lambda^{1/2})$ is positive and $f'(A(p) + \varepsilon\lambda^{1/2})$ is negative for sufficiently small $\lambda > 0$. Therefore, the zero of f' is of the form $A(p) + O(\lambda^{1/2})$. For sufficiently small $\lambda > 0$, this point is in $[a, b]$.

Because $u'_0(A(p)) = 0$, we can use (5.85) and a Taylor series expansion of u_0 around $A(p)$ to obtain

$$\begin{aligned} \sup_{c \in [a, b]} f(c) &= f(A(p) + O(\lambda^{1/2})) \\ &= u_0(A(p) + O(\lambda^{1/2})) \left[1 - \frac{(1 - p)\eta\lambda^{2/3}}{A(p) + O(\lambda^{1/2})} \right] \\ &= u_0(A(p)) - \frac{\eta\lambda^{2/3}}{A^{1+p}(p)} + O(\lambda). \end{aligned}$$

Equation (5.86) follows from (4.6). □

5.5. Proof of Theorem 4.8. According to Corollary 3.8, for (4.27) it suffices to prove

$$(5.87) \quad \int_{z_1^*}^{z_2^*} v(1, \theta) \, d\nu(\theta) = v_0(1) - \frac{\sigma^2 \bar{\theta}^{8/3}}{A^{1+p}(p)} \left(\frac{9p}{32} \right)^{1/3} \lambda^{2/3} + O(\lambda^{5/6})$$

for a distribution ν of our choosing. We begin by choosing $w_1^*(\lambda)$ and $w_2^*(\lambda)$ so that $z_1^*(\lambda) = \bar{\theta}(1 - w_1^*(\lambda))$ and $z_2^*(\lambda) = \bar{\theta}(1 + w_2^*(\lambda))$, where $z_1^*(\lambda)$ and $z_2^*(\lambda)$ are described in Section 3.1 and Assumption 4.1. We let $\theta^*(0)$ have the distribution described in Theorem 4.5, and in place of $c(\cdot)$ in (4.8) we use $c^*(\cdot)$, the optimal consumption process given by (3.11), which satisfies (3.12). We choose positive numbers $c_1(\lambda)$ and $c_2(\lambda)$ so that for some positive constant k ,

$$(5.88) \quad c^*(\theta) - k\lambda \leq c_1(\lambda) \leq c^*(\theta) \leq c_2(\lambda) \leq c^*(\theta) + k\lambda \quad \forall \theta \in [z_1^*(\lambda), z_2^*(\lambda)].$$

As indicated by the notation, $c_1(\lambda)$ and $c_2(\lambda)$ depend on λ but k does not. Despite their dependence on λ , $c_1(\lambda)$ and $c_2(\lambda)$ are bounded above and away from zero and $pA(p) + (1 - p)c_i(\lambda)$ is bounded away from zero, uniformly in λ ; see Remark 3.7. Therefore, we can choose a and b satisfying the conditions in Lemma 5.13 so that $a \leq c_1(\lambda) \leq c_2(\lambda) \leq b$ for all λ sufficiently small. A Taylor expansion of the function $f(x) = x^{1-p}$ around $x = 1$ shows that

$$(5.89) \quad \left(\frac{c_2(\lambda)}{c_1(\lambda)} \right)^{1-p} = 1 + f'(\xi) \frac{c_2(\lambda) - c_1(\lambda)}{c_1(\lambda)} = 1 + O(\lambda),$$

where we have used the fact that $c_1(\lambda)$ is bounded away from zero and also used (5.88).

We continue under the assumption $p \neq 1$. We use $c_1(\lambda)$ in (4.9) to define a process \widehat{X}_2 satisfying $\widehat{X}(0) = 1$ and

$$d\widehat{X}(t) = \widehat{X}(t) \left[(r - c_1(\lambda) + \alpha\theta^*(t)) dt + \sigma\theta^*(t) dW(t) - \lambda(d\ell^*(t) + dm^*(t)) \right],$$

where θ^* and ℓ^* and m^* are defined by (4.11) and text following (4.11). We take the distribution of $\theta^*(0)$ to be the one described in Theorem 4.5. Recalling the process X^* of (4.12), we compute

$$d \log \widehat{X}(t) - d \log X^*(t) = (c^*(\theta^*(t)) - c_1(\lambda)) dt \geq 0.$$

Since $X^*(0) = \widehat{X}(0) = 1$, we see that $\widehat{X}(t) \geq X^*(t)$ for all $t \geq 0$, almost surely. From (4.13) and using (5.89), we see that

$$\begin{aligned} (5.90) \quad \mathbb{E}v(1, \theta^*(0)) &\leq \mathbb{E} \int_0^\infty e^{-\beta t} U_p(c_2(\lambda) X^*(t)) dt \\ &\leq \left(\frac{c_2(\lambda)}{c_1(\lambda)} \right)^{1-p} \mathbb{E} \int_0^\infty e^{-\beta t} U_p(c_1(\lambda) \widehat{X}(t)) dt \\ &= \left(\frac{c_2(\lambda)}{c_1(\lambda)} \right)^{1-p} u_2(c_1(\lambda), c^*(\cdot), w_1^*(\lambda), w_2^*(\lambda)) \\ &= u_2(c_1(\lambda), c^*(\cdot), w_1^*(\lambda), w_2^*(\lambda)) (1 + O(\lambda)) \\ &\leq v_0(1) - \frac{\sigma^2 \bar{\theta}^{8/3}}{A^{1+p}(p)} \left(\frac{9p}{32} \right)^{1/3} \lambda^{2/3} + O(\lambda^{5/6}). \end{aligned}$$

To establish (5.87), it remains to prove the reverse of inequality (5.90). Let a and b be as in Lemma 5.13 and let $c \in [a, b]$ be given. Let $w_1 > 0$ and $w_2 > 0$ also be given. Let $\theta_2(t)$ be given by (4.8), where $c(\cdot) \equiv c$ and $\theta_2(0)$ has the distribution described in Theorem 4.5. Because $c(\cdot)$ in (4.8) matches c in (4.9), the policy that uses constant consumption proportion c and keeps $\theta_2(t)$ in $[\bar{\theta}(1 - w_1), \bar{\theta}(1 + w_2)]$ is feasible in the transaction cost problem with random initial condition $(1, \theta_2(0))$. This implies $u_2(c, c, w_1, w_2) \leq \mathbb{E}v(1, \theta_2(0))$. But Corollary 3.8 implies $\mathbb{E}v(1, \theta_2(0)) = \mathbb{E}v(1, \theta^*(0)) + O(\lambda)$. Consequently,

$$(5.91) \quad \sup_{\substack{w_1 > 0, w_2 > 0 \\ w_1 + w_2 \leq \varphi(\lambda)}} u_2(c, c, w_1, w_2) \leq \mathbb{E}v(1, \theta^*(0)) + O(\lambda).$$

When we maximize over $c \in [a, b]$, Lemma 5.13 gives us the reverse of inequality (5.90).

In the case that $p = 1$, we replace (5.90) by

$$\begin{aligned} \mathbb{E}v(1, \theta_2(0)) &\leq \mathbb{E} \int_0^\infty e^{-\beta t} \log(c_2(\lambda)X^*(t)) dt \\ &= \log \frac{c_2(\lambda)}{c_1(\lambda)} + \mathbb{E} \int_0^\infty e^{-\beta t} \log(c_1(\lambda)\widehat{X}(t)) dt \\ &= \log \frac{c_2(\lambda)}{c_1(\lambda)} + u_2(c_1(\lambda), c^*(\cdot), w_1^*, w_2^*) \\ &= u_2(c_1(\lambda), c^*(\cdot), w_1^*, w_2^*)(1 + O(\lambda)) \\ &\leq v_0(1) - \frac{\sigma^2 \bar{\theta}^{8/3}}{\beta^2} \left(\frac{9}{32}\right)^{1/3} \lambda^{2/3} + O(\lambda^{5/6}) \end{aligned}$$

and proceed as before. This completes the proof of (4.27).

The equality we have established in (5.90) is

$$u_2(c_1(\lambda), c^*(\cdot), w_1^*(\lambda), w_2^*(\lambda)) = v_0(1) - \frac{\sigma^2 \bar{\theta}^{8/3}}{A^{1+p}(p)} \left(\frac{9p}{32}\right)^{1/3} \lambda^{2/3} + O(\lambda^{5/6}).$$

This along with (4.25) and the second equality in (5.84) imply

$$\begin{aligned} &u_2(c_1(\lambda), c^*(\cdot), w_1(\lambda), w_2(\lambda)) \\ &= u_0(c_1(\lambda)) - \frac{c_1^{1-p}(\lambda)\sigma^2 \bar{\theta}^{8/3}}{(pA(p) + (1-p)c_1(\lambda))^2} \left(\frac{9p}{32}\right)^{1/3} \lambda^{2/3} + O(\lambda) \\ &\leq v_0(1) - \frac{\sigma^2 \bar{\theta}^{8/3}}{A^{1+p}(p)} \left(\frac{9p}{32}\right)^{1/3} \lambda^{2/3} + O(\lambda^{5/6}) \\ &= u_2(c_1(\lambda), c^*(\cdot), w_1^*(\lambda), w_2^*(\lambda)) + O(\lambda^{5/6}). \end{aligned}$$

Equation (4.28) follows from (5.69) in Lemma 5.10 with $q = 5/6$.

REFERENCES

[1] C. Atkinson and P. Wilmott, *Portfolio management with transaction costs: An asymptotic analysis of the Morton and Pliska model*, Math. Finance **5** (1995), 357–367.
 [2] G. M. Constantinides, *Capital market equilibrium with transaction costs*, J. Polit. Econ. **94** (1986), 842–862.
 [3] M. Dai, L. Jiang, P. F. Li and F. H. Yi, *Finite horizon optimal investment and consumption with transaction costs*, SIAM J. Control Optimization **48** (2009), 1134–1154. MR 2491593
 [4] M. Dai and Y. Zhong, *Penalty methods for continuous-time portfolio selection with proportional transaction costs*, J. Comp. Finance **13** (2010), 1–31. MR 2608988
 [5] M. Dai and F. H. Yi, *Finite horizon optimal investment with transaction costs: A parabolic double obstacle problem*, J. Diff. Equations **246** (2009), 1445–1469. MR 2488692
 [6] M. H. A. Davis and A. Norman, *Portfolio selection with transaction costs*, Math. Oper. Res. **15** (1990), 676–713. MR 1080472
 [7] K. Janeček and S. Shreve, *Asymptotic analysis for optimal investment and consumption with transaction costs*, Finance Stochast. **8** (2004), 181–206. MR 2048827

- [8] L. Kruk, J. Lehoczky, K. Ramanan and S. Shreve, *An explicit formula for the Skorohod map on $[0, a]$* , Ann. Probab. **35** (2007), 1740–1768. MR 2349573
- [9] R. Korn, *Portfolio optimisation with strictly positive transaction costs and impulse control*, Finance Stochast. **2** (1998), 85–114.
- [10] R. Korn and S. Laue, *Portfolio optimisation with transaction costs and exponential utility*, Stochastic processes and related topics (R. Buckdahn, H. J. Engelbert and M. Yor, eds.), Taylor & Francis, London, 2002. MR 1987316
- [11] H. Liu and M. Loewenstein, *Optimal portfolio selection with transaction costs and finite horizons*, Rev. Fin. Studies **15** (2002), 805–835.
- [12] M. J. Magill and G. M. Constantinides, *Portfolio selection with transaction costs*, J. Econ. Theory **13** (1976), 245–263. MR 0469196
- [13] G. Maruyama and H. Tanaka, *Some properties of one-dimensional diffusion processes*, Mem. Faculty Science, Kyusyu Univ., Ser. A **11** (1957), 117–141. MR 0097128
- [14] R. Merton, *Optimum consumption and portfolio rules in a continuous-time case*, J. Econ. Theory **3** (1971), 373–413 [Erratum **6** (1973), 213–214].
- [15] K. Muthuraman and S. Kumar, *Multidimensional portfolio optimization with proportional transaction costs*, Math. Finance **16** (2006), 301–335. MR 2212268
- [16] L. C. G. Rogers, *Why is the effect of proportional transaction costs $O(\delta^{2/3})$?* Mathematics of finance, AMS Contemporary Mathematics Series, vol. 351 (G. Yin and Q. Zhang, eds.), Amer. Math. Soc., Providence, RI, 2004, pp. 303–308. MR 2076549
- [17] S. Shreve and H. M. Soner, *Optimal investment and consumption with transaction costs*, Ann. Appl. Probab. **4** (1994), 609–692. MR 1284980
- [18] A. Tourin and T. Zariphopoulou, *Numerical schemes for investment models with singular transactions*, Computat. Econ. **7** (1994), 287–307. MR 1318095
- [19] A. Tourin and T. Zariphopoulou, *Viscosity solutions and numerical schemes for investment/consumption models with transaction costs*, Numerical methods in finance (L. C. G. Rogers and D. Talay, eds.), Isaac Newton Institute Publications, Cambridge Univ. Press, Cambridge, 1997. MR 1470518
- [20] A. E. Whalley and P. Wilmott, *An asymptotic analysis of an optimal hedging model for option pricing under transaction costs*, Math. Finance **7** (1997), 307–324. MR 1459062

KAREL JANEČEK, RSJ ALGORITHMIC TRADING, 118 00 PRAGUE, CZECH REPUBLIC AND DEPARTMENT OF PROBABILITY AND MATHEMATICAL STATISTICS, FACULTY OF MATHEMATICS AND PHYSICS, CHARLES UNIVERSITY PRAGUE, 186 75 PRAGUE, CZECH REPUBLIC

E-mail address: karel.janecek@rsj.cz

STEVEN E. SHREVE, DEPARTMENT OF MATHEMATICAL SCIENCES, CARNEGIE MELLON UNIVERSITY, PITTSBURGH, PA, 15213 USA

E-mail address: shreve@andrew.cmu.edu