

Utility Maximization Trading Two Futures with Transaction Costs*

Maxim Bichuch[†] and Steven Shreve[‡]

Abstract. An agent invests in two types of futures contracts, whose prices are possibly correlated arithmetic Brownian motions, and invests in a money market account with a constant interest rate. The agent pays a transaction cost for trading in futures proportional to the size of the trade. She also receives utility from consumption. The agent maximizes expected infinite-horizon discounted utility from consumption. We determine the first two terms in the asymptotic expansion of the value function in the transaction cost parameter around the known value function for the case of zero transaction cost. The method of solution when the futures are uncorrelated follows a method used previously to obtain the analogous result for one risky asset. However, when the futures are correlated, a new methodology must be developed. It is suspected in this case that the value function is not twice continuously differentiable, and this prevents application of the former methodology.

Key words. transaction costs, optimal control, asymptotic analysis, utility maximization

AMS subject classifications. 90A09, 91G10, 93E20

DOI. 10.1137/110853649

1. Introduction. We consider an agent seeking to optimally invest and consume in the presence of proportional transaction costs. The agent can trade in two kinds of futures contracts, driven by two possibly correlated Brownian motions. A futures contract is marked to market every day, and its value is thereby reset to zero. An agent holding a futures contract receives the daily differences in the futures prices. Because these absolute rather than relative changes in the futures price determine the cash flow that accrues from a position in futures, we model the futures price of the contract of type i as an arithmetic (rather than geometric) Brownian motion with drift and diffusion coefficients $\alpha_i, \sigma_i > 0$, respectively. The agent can also hold cash in the money market account with constant interest rate r , and she may consume.

A proportional transaction cost $\mu_i \lambda$ is charged for trading in the futures contracts of type i , where μ_i is a positive constant, $i = 1, 2$. The transaction costs and consumption are deducted from the money market account. The agent maximizes the expected discounted integral over $[0, \infty)$ of the utility of consumption, where the utility function is $U(C) \triangleq \frac{C^p}{p}$, $C \geq 0$, and $p \in (0, 1)$. We compute the coefficient of the leading term in the asymptotic expansion of the value function for positive λ about the known value function for the case $\lambda = 0$. This leading term is of order $\lambda^{2/3}$. The proof is divided into two cases: when the two futures prices are

*Received by the editors November 1, 2011; accepted for publication (in revised form) October 1, 2012; published electronically February 5, 2013.

<http://www.siam.org/journals/sifin/4/85364.html>

[†]Department of Operations Research and Financial Engineering, Princeton University, Princeton, NJ 08544 (mbichuch@princeton.edu).

[‡]Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA 15213 (shreve@andrew.cmu.edu). This author's research was partially supported by the National Science Foundation under grant DMS-0903475.

independent and when they are correlated.

This work generalizes that of Janeček and Shreve [13] by allowing the investor to trade in more than one risky asset. The proof technique in [13] is to reduce the dimensionality of the problem, write down an asymptotic expansion of the value function for the reduced problem, assume this function is a twice continuously differentiable solution of the corresponding Hamilton–Jacobi–Bellman (HJB) equation, and use this assumption to formally determine the leading coefficients in the asymptotic expansion. Then one can perturb both up and down the order- λ term in the asymptotic expansion to obtain functions that are provably super- and subsolutions of the HJB equation, and hence upper and lower bounds on the value function. These functions agree in the λ^0 and $\lambda^{2/3}$ terms, and hence identify the coefficients of these two terms.

A straightforward generalization of [13] works for the model of this paper when the futures prices are independent, and that is provided in section 9. (To save space, we omit the formal determination of the coefficients in the asymptotic expansion, which can be found in [5, section 3]. Here we begin the proof by writing down the functions corresponding to different values of K in (9.24) that we ultimately prove are super- and subsolutions of the HJB equation.) When the futures prices are correlated, the formal determination of the coefficients in the asymptotic expansion along the lines of [5, section 3] leads to inconsistent equations, which suggests that the second derivative of the value function is not continuous everywhere. Abandoning the methodology of [13] in this case, in section 10 we construct tight upper and lower bounds on the value function by considering an auxiliary problem having independent futures contracts. The construction of the lower bound uses the fact that the risky asset prices are arithmetic rather than geometric Brownian motions. In particular, the second-order operator \mathcal{L}_2 of (6.10) involves first and second derivatives in the radial direction only; see Case II of the proof of Lemma 10.6.

This paper is organized as follows. Section 2 sets out the model, notation, and all assumptions of the paper. The case of zero transaction cost has a solution similar to that due to Merton [17], [18] for investment in multiple geometric Brownian motion stock prices. The solution for arithmetic Brownian motions is provided in section 3. Similar to [17], [18], here the optimal solution is to hold the number of futures contracts equal to a constant vector, which we call the *Merton proportion*, times the value of the portfolio, and the value function is a positive constant times the utility function. We state the main result of the paper in section 4. Section 5 develops the HJB equation for the full problem. The continuity of the value function needed for the HJB characterization is proved in Appendix A. Section 6 reduces the dimensionality of the problem and develops the HJB equation for the reduced problem. This section includes the comparison theorem for super- and subsolutions of the HJB equation, whose proof is in Appendix B. Section 7 describes how to partition the solvency region for the reduced problem as a first step toward constructing super- and subsolutions of the HJB equation. Section 8 extends functions defined in the middle of the solvency region to other parts of the region and derives properties of the extensions. This method of extension is used in sections 9 and 10 to construct super- and subsolutions and thereby prove the main result reported in section 4.

For the case of a single risky asset, Magill and Constantinides [16] introduced transaction costs into Merton’s infinite-horizon discounted optimal consumption model [17], [18]. They

argued that in the presence of a proportional transaction cost, an agent should not trade when the state of the problem is in a region M containing the Merton proportion, and should trade just enough on the boundary of M to keep the state from exiting this region. Constantinides [6] numerically computed the effect of transaction costs on the value function for the problem with one risky asset, and observed that transaction costs have a “first-order effect on asset demand” and a “second-order effect on equilibrium asset return.” Davis and Norman [8] put this argument on a rigorous mathematical foundation. Shreve and Soner [21] used viscosity solution analysis to relax the assumptions in [8]. Because there is an explicit solution for the infinite-horizon discounted optimal consumption problem with zero transaction cost, but apparently no closed-form solution for positive transaction cost, a determination of the no-trading region M and the value function for the transaction cost problem by asymptotic analysis around the zero-transaction-cost problem has been developed. Heuristic analysis along these lines was provided by Whalley and Wilmott [23] and Atkinson and Al-Ali [2]. Rigorous calculations of the value function expansion were given by [21], [13], [14], the last one making rigorous the insight of Rogers [20]. These expansions make precise the numerical observations of [6].

The previous work pertains to problems with a single risky asset. For multiple independent risky assets, Akian, Menaldi, and Sulem [1] characterized the value function as the unique viscosity solution of the HJB equation and provided numerics. Atkinson and Mokkavesa [4] provided a perturbation analysis for independent assets. Correlated assets were treated by perturbation analysis by Atkinson and Ingpochai [3] for small correlations, and a numerical study for two correlated assets was performed by Muthuraman and Kumar [19]. These works provide strong evidence that in the case of multiple independent risky assets, the no-transaction region M is a rectangle, but if the assets are correlated, the region M is not even polyhedral.

Perturbation analysis is based on the assumption of “smooth pasting”; i.e., the value function is twice continuously differentiable across the boundary of M . This condition was verified in the case of one risky asset in [21]. However, as observed above, there is reason to doubt smooth pasting in models with correlated risky assets.

Goodman and Ostrov [12] take a duality approach to transaction costs and show that the quasi-steady state probability density of the optimal portfolio is uniform for a single stock but generally blows up when there are multiple stocks.

There is a related body of literature on growth-optimal portfolios in the presence of transaction costs; see, e.g., [9], [10], [11], [15], [22].

2. Model definition. An agent has three investment opportunities, a *money market account* with constant rate of interest r and *two types of futures contracts*. The futures price associated with the contract of type i is an arithmetic Brownian motion

$$F_i(t) = F_i(0) + \alpha_i t + \sigma_i B_i(t),$$

where each α_i is constant, each σ_i is a strictly positive constant, and (B_1, B_2) is a two-dimensional Brownian motion with correlation given by $\langle B_1, B_2 \rangle(t) = \rho t$, where ρ is a constant in $(-1, 1)$. The two-dimensional Brownian motion (B_1, B_2) is defined on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}(t)\}_{t \geq 0}, \mathbb{P})$. We use arithmetic rather than geometric Brownian motion to

describe futures prices because the cost of purchasing futures contracts is zero, and hence the absolute rather than the relative change in futures prices determines the payoff of these contracts as investment opportunities. Without changing the investment opportunities in this problem, we may replace one or both of the $F_i(t)$ by $-F_i(t) = -F_i(0) - \alpha_i t + \sigma_i(-B_i(t))$. We may also interchange the indices on the two arithmetic Brownian motions. By making these changes if necessary, we may and do assume almost without loss of generality that

$$(2.1) \quad \alpha_1 > 0, \quad 0 \leq \rho < 1.$$

To achieve (2.1) we need only that one of the two futures prices has nonzero drift. We shall need this condition only in section 10, where we analyze the case $\rho \neq 0$.

The agent chooses four adapted processes, $L_i, M_i, i = 1, 2$. The process $L_i(t)$ (respectively, $M_i(t)$) denotes the number of futures contracts of type i bought (respectively, sold) by time t , and is thus a nondecreasing, right-continuous process with initial condition $L_i(0) \geq 0$ (respectively, $M_i(0) \geq 0$) representing the number of futures contracts of time i bought (respectively, sold) at time 0. We adopt the convention $L_i(0-) = M_i(0-) = 0, i = 1, 2$. The agent begins with an initial position $Y_i(0-)$ in futures contracts of type i , and thus her *position in futures contracts* of type i at time t is

$$(2.2) \quad Y_i(t) = Y_i(0-) + L_i(t) - M_i(t), \quad t \geq 0.$$

The agent finances trading in futures by investing or borrowing at the interest rate r . There is a positive transaction cost for this trading. In particular, for $i = 1, 2$, there is a constant $\lambda_i = \mu_i \lambda$, where $\mu_i > 0$ and $\lambda > 0$ are constants, representing the *transaction cost per contract* incurred from trading in futures contracts of type i . The agent chooses an adapted nonnegative *rate of consumption process* C satisfying $\int_0^t C(u) du < \infty$ for all $t \geq 0$, and then the agent's *position in the money market account* evolves according to the equation

$$(2.3) \quad dX(t) = \sum_{i=1}^2 Y_i(t) dF_i(t) - \sum_{i=1}^2 \lambda_i (dL_i(t) + dM_i(t)) + (rX(t) - C(t)) dt.$$

We distinguish the agent's *initial position in the money market account* $X(0-)$ from the agent's position (after any transactions) at time zero. The latter is $X(0) = X(0-) - \sum_{i=1}^2 \lambda_i (L_i(0) + M_i(0))$.

This is an asymptotic study as $\lambda \downarrow 0$, and all other parameters, including $\mu_i, i = 1, 2$, are held constant. In particular, $\lambda_i = O(\lambda)$. In section 10, in order to analyze the case $\rho \neq 0$, in addition to (2.1) we will need the assumptions

$$(2.4) \quad \mu_2 \alpha_1 > \mu_1 \alpha_2, \quad \rho < \frac{\mu_2 \sigma_1}{\mu_1 \sigma_2}.$$

For much of the analysis it is convenient to write the trading and consumption processes as proportions of the money market position. As long as $X(u-)$ is bounded away from zero, $0 \leq u \leq t$, we may define

$$c(t) = \int_0^t \frac{C(u) du}{X(u)}, \quad \ell_i(t) = \int_0^t \frac{dL_i(u)}{X(u-)}, \quad m_i(t) = \int_0^t \frac{dM_i(u)}{X(u-)}, \quad i = 1, 2,$$

so that for $t \geq 0$,

$$(2.5) \quad dY_i(t) = X(t-)(d\ell_i(t) - dm_i(t)), \quad i = 1, 2,$$

$$(2.6) \quad dX(t) = \sum_{i=1}^2 Y_i(t)dF_i(t) - X(t-) \sum_{i=1}^2 \lambda_i(d\ell_i(t) + dm_i(t)) + X(t)(r - c(t))dt.$$

We define the *three-dimensional solvency region* to be the open set

$$(2.7) \quad \mathcal{D}_3 \triangleq \{(y_1, y_2, x) : x - \lambda_1|y_1| - \lambda_2|y_2| > 0\}.$$

We denote the closure of \mathcal{D}_3 by $\overline{\mathcal{D}}_3$ and the boundary by $\partial\mathcal{D}_3$. A policy (L_1, L_2, M_1, M_2, C) is *admissible* for the initial condition $(Y_1(0-), Y_2(0-), X(0-)) \in \overline{\mathcal{D}}_3$ if we have $(Y_1(t), Y_2(t), X(t)) \in \overline{\mathcal{D}}_3$ for all $t \geq 0$.

If (Y_1, Y_2, X) is outside $\overline{\mathcal{D}}_3$, then closing out the futures positions results in a strictly negative money market position; the agent is not permitted to enter this region. If $(Y_1, Y_2, X) \in \partial\mathcal{D}_3$, then the agent can immediately close out her futures positions, which will bring her money market position to zero. In fact, because not closing out the futures positions would result in a positive probability of exiting the closed solvency region, this is the only admissible action available (see [5, section 2, Remark 2.1] for the precise argument, or for the case of one risky asset, see [21]). We conclude that every admissible policy keeps (Y_1, Y_2, X) inside the open solvency region, where X is strictly positive, and we may use the representation (2.5), (2.6) rather than (2.2), (2.3) until the first time the boundary of the solvency region is hit, if it is ever hit. Upon hitting $\partial\mathcal{D}_3$, (Y_1, Y_2, X) jumps immediately to the origin and remains there. Thus we may characterize admissible policies in terms of the processes $(\ell_1, \ell_2, m_1, m_2, c)$ appearing in (2.5), (2.6), knowing that at the first time (Y_1, Y_2, X) reaches $\partial\mathcal{D}_3$, it jumps to the origin and remains there. We denote by $\mathcal{A}(y_1, y_2, x)$ the set of all such policies corresponding to the initial condition $(y_1, y_2, x) \in \overline{\mathcal{D}}_3$.

The agent has the *utility function* $U(C) = \frac{1}{p}C^p$, defined for $C \geq 0$, where p is a constant in $(0, 1)$. Let $\beta > 0$ be a constant *impatience factor*. The *value function* for the problem of expected infinite-horizon discounted consumption and investment is

$$(2.8) \quad v(y_1, y_2, x) \triangleq \sup_{(\ell_1, \ell_2, m_1, m_2, c) \in \mathcal{A}(y_1, y_2, x)} \mathbb{E} \int_0^\infty e^{-\beta t} U(c(t)X(t)) dt,$$

where, of course, we mean that X is given by (2.5), (2.6) with $Y_i(0-) = y_i$, $i = 1, 2$, and $X(0-) = x$, and we take $c(t)X(t) = 0$ after the first time that (Y_1, Y_2, X) reaches $\partial\mathcal{D}_3$. In particular,

$$(2.9) \quad v = 0 \text{ on } \partial\mathcal{D}_3.$$

We develop an asymptotic expansion in the small parameter $\lambda > 0$ for the function v .

Remark 2.1 (concavity of v). The function v is concave on $\overline{\mathcal{D}}_3$. In order to see this, let (L_1, L_2, M_1, M_2, C) be an admissible policy for $(y_1, y_2, x) \in \overline{\mathcal{D}}_3$ and let $(\widehat{L}_1, \widehat{L}_2, \widehat{M}_1, \widehat{M}_2, \widehat{C})$ be an admissible policy for $(\widehat{y}_1, \widehat{y}_2, \widehat{x}) \in \overline{\mathcal{D}}_3$. Because of the linearity of (2.2) and (2.3) and

the convexity of \mathcal{D}_3 , for any $\delta \in (0, 1)$, $\delta(L_1, L_2, M_1, M_2, C) + (1 - \delta)(\widehat{L}_1, \widehat{L}_2, \widehat{M}_1, \widehat{M}_2, \widehat{C})$ is admissible for $\delta(y_1, y_2, x) + (1 - \delta)(\widehat{y}_1, \widehat{y}_2, \widehat{x})$. The concavity of U then implies

$$\begin{aligned} & v(\delta y_1 + (1 - \delta)\widehat{y}_1, \delta y_2 + (1 - \delta)\widehat{y}_2, \delta x + (1 - \delta)\widehat{x}) \\ & \geq \mathbb{E} \int_0^\infty e^{-\beta t} U(\delta C(t) + (1 - \delta)\widehat{C}(t)) dt \\ & \geq \delta \mathbb{E} \int_0^\infty e^{-\beta t} U(C(t)) dt + (1 - \delta) \mathbb{E} \int_0^\infty e^{-\beta t} U(\widehat{C}(t)) dt. \end{aligned}$$

Maximizing the right-hand side over admissible policies for (y_1, y_2, x) and $(\widehat{y}_1, \widehat{y}_2, \widehat{x})$, we obtain the claimed concavity of v on $\overline{\mathcal{D}}_3$.

To make the notation more compact, we denote by $\vec{\alpha}$, $\vec{\lambda}$, and \vec{F} the two-dimensional column vectors of constants and processes

$$\vec{\alpha} \triangleq \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}, \quad \vec{\lambda} \triangleq \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix}, \quad \vec{F} \triangleq \begin{bmatrix} F_1 \\ F_2 \end{bmatrix},$$

and we use similar notation for other two-dimensional vectors of constants, variables, and processes. We define the nonsingular matrices

$$\Sigma \triangleq \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix}, \quad \mathbf{V} \triangleq \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}.$$

Finally, we define the vector

$$(2.10) \quad \vec{\theta} = \frac{1}{1-p} \mathbf{V}^{-1} \vec{\alpha}.$$

In order to avoid cases in the subsequent analysis, we make the following assumption.

Assumption 2.2. We assume throughout that $\vec{\theta}$ is in the open first quadrant.

The following assumption is essential.

Assumption 2.3. We assume throughout that

$$(2.11) \quad A \triangleq \frac{\beta - rp}{1-p} - \frac{p}{2(1-p)^2} \mathbf{V}^{-1} \vec{\alpha} \cdot \vec{\alpha}$$

is strictly positive, which, as we shall see in the next section, guarantees that the value function is finite for the problem with zero transaction cost λ , and hence the value function is finite for the less favorable problem with positive transaction cost.

We note for future reference that

$$(2.12) \quad \beta - rp - p\vec{\alpha} \cdot \vec{z} + \frac{1}{2}p(1-p)(\mathbf{V}\vec{z} \cdot \vec{z}) = (1-p) \left[A + \frac{1}{2}p(\mathbf{V}(\vec{z} - \vec{\theta}) \cdot (\vec{z} - \vec{\theta})) \right].$$

Remark 2.4. Notation such as \vec{z} will always be a two-dimensional column vector when used in matrix calculations such as in (2.12), and its transpose will be denoted by \vec{z}^T . For example, $\mathbf{V}\vec{z} \cdot \vec{z} = \vec{z}^T \mathbf{V}\vec{z}$. To save space when we want to specify the components of a column vector \vec{z} , we abuse notation and write $\vec{z} = (z_1, z_2)$.

3. Zero transaction cost ($\lambda = 0$). The model of section 2 makes sense even if $\lambda = 0$, and since in this case we can instantaneously change positions in the futures without cost, the value function is a function of the position in the money market alone. The solvency region is $\mathcal{D}_3 = \{(y_1, y_2, x) : x > 0\}$. Equation (2.3) simplifies to

$$(3.1) \quad dX(t) = \vec{Y}(t) \cdot d\vec{F}(t) + (rX(t) - C(t))dt.$$

In particular, X is continuous. The processes Y_i of (2.2) are right continuous with left-hand limits and have finite variation on each finite interval of time. We remove the finite-variation condition and obtain a more favorable problem whose value function is still a function of the position in the money market alone. This value function is

$$(3.2) \quad \sup_{Y_1, Y_2, C} \mathbb{E} \left[\int_0^\infty e^{-\beta t} U(C(t)) dt \mid X(0) = x \right], \quad x \geq 0,$$

where the supremum is over all adapted processes Y_1, Y_2 and nonnegative adapted processes C that ensure that $X(t)$ given by (3.1) is defined and nonnegative for all $t \geq 0$.

The argument by Merton [17], [18] for investing in multiple geometric Brownian motions is easily adapted to this situation to show that the value function (3.2) is

$$(3.3) \quad v_0(x) = \frac{1}{p} A^{p-1} x^p, \quad x \geq 0,$$

and the optimal portfolio and consumption policy are

$$(3.4) \quad \vec{Y}(t) = X(t)\vec{\theta}, \quad C(t) = AX(t), \quad t \geq 0,$$

where X is the money market position process obtained by using this policy. The argument is brief, and for the sake of completeness, we include it here.

It is straightforward to verify that v_0 given by (3.3) satisfies the HJB equation

$$(3.5) \quad \min_{\vec{y} \in \mathbb{R}^2, c \geq 0} \left[\beta v_0(x) - (\vec{\alpha} \cdot \vec{y} + rx - c)v_0'(x) - \frac{1}{2}(\mathbf{V}\vec{y} \cdot \vec{y})v_0''(x) - U(c) \right] = 0$$

for all $x > 0$. Indeed, the minimizing \vec{y} in (3.5) is $\frac{(p-1)v_0'(x)}{v_0''(x)}\vec{\theta} = x\vec{\theta}$ and the minimizing c is $(v_0'(x))^{1/(p-1)} = Ax$. Substitution of these values into the left-hand side of (3.5) and some simple algebra results in the equality

$$(3.6) \quad \beta v_0(x) - (\vec{\alpha} \cdot \vec{\theta} + r - A)xv_0'(x) - \frac{1}{2}(\mathbf{V}\vec{\theta} \cdot \vec{\theta})x^2v_0''(x) - U(Ax) = 0.$$

In particular, v_0 satisfies the HJB inequality

$$(3.7) \quad \beta v_0(x) - (\vec{\alpha} \cdot \vec{y} + rx - c)v_0'(x) - \frac{1}{2}(\mathbf{V}\vec{y} \cdot \vec{y})v_0''(x) - U(c) \geq 0, \quad \vec{y} \in \mathbb{R}^2, c \geq 0.$$

Inequality (3.7) together with Itô's formula applied to $e^{-\beta t}X(t)$ shows that for any adapted processes Y_1 and Y_2 that are right continuous with left limits, and for any consumption process

C satisfying $\int_0^t C(u)du < \infty$ for all $t \geq 0$,

$$(3.8) \quad \begin{aligned} 0 &\leq e^{-\beta(t \wedge \tau_n)} v_0(X(t \wedge \tau_n)) \\ &\leq v_0(X(0)) - \int_0^{t \wedge \tau_n} e^{-\beta u} U(C(u)) du + \int_0^{t \wedge \tau_n} e^{-\beta u} v_0'(X(u)) \Sigma \vec{Y}(u) \cdot d\vec{B}(u), \end{aligned}$$

where $\tau_n \triangleq \inf \{t \geq 0 : X(t) \leq 1/n\}$. This implies that the local martingale $v_0(X(0)) + \int_0^{t \wedge \tau_n} e^{-\beta u} v_0'(X(u)) \Sigma \vec{Y}(u) \cdot d\vec{B}(u)$ is nonnegative and hence is a supermartingale. Taking expectations in (3.8), we thus obtain $\mathbb{E} \int_0^{t \wedge \tau_n} e^{-\beta u} U(C(u)) du \leq v_0(X(0))$. Using the monotone convergence theorem as $n \rightarrow \infty$ and $t \rightarrow \infty$, and then maximizing over Y_1, Y_2 , and C , we conclude that

$$(3.9) \quad \sup_{Y_1, Y_2, C} \mathbb{E} \int_0^\infty e^{-\beta t} U(C(t)) dt \leq v_0(X(0)).$$

To obtain equality in (3.9), we use the policy given by (3.4) and use (3.6) in place of (3.7) to write

$$(3.10) \quad \begin{aligned} &e^{-\beta(t \wedge \tau_n)} v_0(X(t \wedge \tau_n)) \\ &= v_0(X(0)) - \int_0^{t \wedge \tau_n} e^{-\beta u} U(C(u)) du + \int_0^{t \wedge \tau_n} e^{-\beta u} v_0'(X(u)) \Sigma \vec{Y}(u) \cdot d\vec{B}(u) \end{aligned}$$

in place of (3.8). For this policy,

$$X(t) = X(0) \exp \left[\left(\vec{\alpha} \cdot \vec{\theta} + r - A - \frac{1}{2} \mathbf{V} \vec{\theta} \cdot \vec{\theta} \right) t + \Sigma \vec{\theta} \cdot \vec{B}(t) \right],$$

which satisfies $\mathbb{E}[X^{2p}(t)] = X^{2p}(0) e^{p(2\vec{\alpha} \cdot \vec{\theta} + 2r - 2A + \mathbf{V} \vec{\theta} \cdot \vec{\theta})t}$, and so the Itô integral in (3.10) is a martingale. Taking expectations and then limits in (3.10), we obtain

$$\frac{1}{p} A^{p-1} \lim_{t \rightarrow \infty} \lim_{n \rightarrow \infty} \mathbb{E} [e^{-\beta(t \wedge \tau_n)} X^p(t \wedge \tau_n)] + \mathbb{E} \int_0^\infty e^{-\beta t} U(C(t)) dt = v_0(X(0)).$$

But (2.12) implies $e^{-\beta t} X^p(t) = X^p(0) e^{-At} M(t)$, where $M(t) = \exp[p \Sigma \vec{\theta} \cdot \vec{B}(t) - \frac{1}{2} p^2 (\mathbf{V} \vec{\theta} \cdot \vec{\theta}) t]$ is a square-integrable martingale. For each $t \geq 0$, $\{X^p(0) e^{-A(t \wedge \tau_n)} M(t \wedge \tau_n)\}_{n=1}^\infty$ is a uniformly integrable sequence of random variables. For the process X under consideration, τ_n satisfies $\lim_{n \rightarrow \infty} \tau_n = \infty$, and hence

$$\lim_{n \rightarrow \infty} \mathbb{E} [e^{-\beta(t \wedge \tau_n)} X^p(t \wedge \tau_n)] = \mathbb{E} [e^{-\beta t} X^p(t)] = X^p(0) e^{-At},$$

which has limit zero as $t \rightarrow \infty$ because of Assumption 2.3. This concludes the proof of equality in (3.9).

4. Main result. The main result of this paper is the following theorem.

Theorem 4.1. *Under Assumptions 2.2 and 2.3 and conditions (2.1) and (2.4), the value function (2.8) for the problem with positive λ satisfies*

$$v(y_1, y_2, x) = v_0(x) - \underline{\gamma} \lambda^{2/3} x^p + O(\lambda),$$

where the absolute value of the $O(\lambda)$ term is bounded by λ times a constant that is independent of $\lambda > 0$ and (y_1, y_2, x) so long as this triple is in a compact subset of the solvency region \mathcal{D}_3 . The parameter $\underline{\gamma}$ is defined in (10.6) below.

The proof of Theorem 4.1 follows the statement of Theorem 10.1. In section 9 we construct a subsolution and a supersolution for the HJB equation when $\rho = 0$. These are consequently lower and upper bounds, respectively, for v . The subsolution and supersolution we construct differ by $O(\lambda)$ and hence determine v up to order λ . For the case $\rho \neq 0$, in section 10 we create an auxiliary problem with two types of futures contracts that are independent. We use this auxiliary problem and the result already obtained for independent futures to construct lower and upper bounds for v . Again, these bounds differ by $O(\lambda)$. The constructions suggest nearly (up to order λ) optimal policies, but in order not to further lengthen an already long paper, we do not rigorously construct these policies and prove that they are nearly optimal.

5. HJB equation for $\lambda > 0$. Because the agent can increase her position by size $\eta_1 > 0$ in type-one futures by reducing the money market account balance by $\lambda_1 \eta_1$, thereby moving in the $[1, 0, -\lambda_1]$ direction in (y_1, y_2, x) in \mathcal{D}_3 , a move in this direction cannot increase the value function. In other words, $v_1 - \lambda_1 v_x \leq 0$, where v_1 denotes the partial derivative of v with respect to y_1 . Moreover, in any region in which it is optimal for the agent to increase her position in type-one futures, $v_1 - \lambda_1 v_x = 0$. Analogously, $v_2 - \lambda_2 v_x \leq 0$, and in any region in which it is optimal for the agent to increase her position in type-two futures, $v_2 - \lambda_2 v_x = 0$. Furthermore, the agent can decrease her position by size $\eta_1 > 0$ in type-one futures by reducing the money market account balance by $\lambda_1 \eta_1$, which leads to $-v_1 - \lambda_1 v_x \leq 0$, with equality holding in any region in which it is optimal to decrease the position in type-one futures. Also, $-v_2 - \lambda_2 v_x \leq 0$, with equality holding in any region in which it is optimal to decrease the position in type-two futures. In the remaining region, where it is optimal to consume but not change the futures positions, the value function should satisfy the partial differential equation $\mathcal{L}_3 v - \tilde{U}(v_x) = 0$, where \mathcal{L}_3 is the linear differential operator

$$(5.1) \quad \mathcal{L}_3 \psi \triangleq \beta \psi - (rx + \vec{\alpha} \cdot \vec{y}) \psi_x - \frac{1}{2} (\mathbf{V} \vec{y} \cdot \vec{y}) \psi_{xx},$$

and \tilde{U} , obtained by optimizing over consumption, is

$$(5.2) \quad \tilde{U}(\tilde{c}) = \sup_{c \geq 0} (U(c) - c\tilde{c}) = \frac{1-p}{p} \tilde{c}^{p/(p-1)}, \quad \tilde{c} > 0.$$

(The Fenchel–Legendre transform of $-U$ is $\tilde{U}(-\tilde{c})$.)

The following theorem makes these considerations precise. For this theorem, we denote by $C^{1,1,2}(\mathcal{D}_3)$ the set of functions ψ defined and continuous on $\overline{\mathcal{D}_3}$ whose partial derivatives ψ_1 , ψ_2 , and ψ_{xx} are defined and continuous on \mathcal{D}_3 .

Theorem 5.1. *When λ is strictly positive, the value function v defined by (2.8) is continuous on $\overline{\mathcal{D}}_3$ defined by (2.7) and is a viscosity solution of the HJB equation*

$$(5.3) \quad \min [\mathcal{L}_3 v - \tilde{U}(v_x), -v_1 + \lambda_1 v_x, -v_2 + \lambda_2 v_x, v_1 + \lambda_1 v_x, v_2 + \lambda_2 v_x] = 0.$$

This means that for every $(y_1, y_2, x) \in \mathcal{D}_3$ and every function $\psi \in C^{1,1,2}(\mathcal{D}_3)$ agreeing with v at (y_1, y_2, x) , the inequality $\psi \geq v$ (respectively, $\psi \leq v$) on \mathcal{D}_3 implies

$$(5.4) \quad \min[\mathcal{L}_3 \psi - \tilde{U}(\psi_x), -\psi_1 + \lambda_1 \psi_x, -\psi_2 + \lambda_2 \psi_x, \psi_1 + \lambda_1 \psi_x, \psi_2 + \lambda_2 \psi_x]$$

is nonpositive (respectively, nonnegative) at (y_1, y_2, x) .

The proof of continuity of v on $\overline{\mathcal{D}}_3$ is in Appendix A. The proof that v is a viscosity solution of (5.3) is standard and is not given here. The details for this particular model are presented in [5, Appendix A, Theorem A.12], which is a parallel of the proof found in [21].

6. Reduction of dimension. Because \mathcal{D}_3 is a cone and every term in (2.5) and (2.6) is scaled by $X(t-)$ or $X(t)$, we have $\mathcal{A}(ay_1, ay_2, ax) = \mathcal{A}(y_1, y_2, x)$ whenever $a > 0$. It follows that

$$(6.1) \quad v(ay_1, ay_2, ax) = a^p v(y_1, y_2, x), \quad (y_1, y_2, x) \in \overline{\mathcal{D}}_3, \quad a \geq 0.$$

We note also that $(y_1, y_2, x) \in \mathcal{D}_3$ if and only if $x > 0$ and $(\frac{y_1}{x}, \frac{y_2}{x})$ is in the *two-dimensional solvency region*

$$(6.2) \quad \mathcal{D}_2 \triangleq \{(z_1, z_2) : 1 - \lambda_1 |z_1| - \lambda_2 |z_2| > 0\}.$$

We define

$$(6.3) \quad u(z_1, z_2) = v(z_1, z_2, 1), \quad (z_1, z_2) \in \overline{\mathcal{D}}_2,$$

and note that because of (6.1),

$$(6.4) \quad v(y_1, y_2, x) = x^p u\left(\frac{y_1}{x}, \frac{y_2}{x}\right), \quad (y_1, y_2, x) \in \overline{\mathcal{D}}_3 \setminus \{(0, 0, 0)\}.$$

Thus, to determine v it suffices to determine u . From (2.9) we have

$$(6.5) \quad u = 0 \text{ on } \partial\mathcal{D}_2.$$

We need to transform the HJB equation (5.3) into the new variables. Let φ be a C^2 function defined on \mathcal{D}_2 , and set

$$(6.6) \quad \psi(y_1, y_2, x) = x^p \varphi\left(\frac{y_1}{x}, \frac{y_2}{x}\right), \quad (y_1, y_2, x) \in \mathcal{D}_3.$$

We use the notation $z_i = \frac{y_i}{x}$, $\varphi_i = \frac{\partial}{\partial z_i} \varphi$, $\varphi_{i,j} = \frac{\partial^2}{\partial z_i \partial z_j} \varphi$. We further adopt the notation

$$\vec{z} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}, \quad \nabla \varphi(\vec{z}) = \begin{bmatrix} \varphi_1(\vec{z}) \\ \varphi_2(\vec{z}) \end{bmatrix}, \quad \nabla^2 \varphi(\vec{z}) = \begin{bmatrix} \varphi_{11}(\vec{z}) & \varphi_{12}(\vec{z}) \\ \varphi_{21}(\vec{z}) & \varphi_{22}(\vec{z}) \end{bmatrix}.$$

It is straightforward to calculate

$$(6.7) \quad \psi_x(y_1, y_2, x) = x^{p-1} [p\varphi(\vec{z}) - \nabla\varphi(\vec{z}) \cdot \vec{z}],$$

$$(6.8) \quad \psi_{xx}(y_1, y_2, x) = x^{p-2} [p(p-1)\varphi(\vec{z}) + 2(1-p)\nabla\varphi(\vec{z}) \cdot \vec{z} + \nabla^2\varphi(\vec{z})\vec{z} \cdot \vec{z}],$$

$$(6.9) \quad \psi_i(y_1, y_2, x) = x^{p-1}\varphi_i(\vec{z}), \quad i = 1, 2.$$

Consequently, $\mathcal{L}_3\psi(y_1, y_2, x) = x^p\mathcal{L}_2\varphi(\vec{z})$, where

$$(6.10) \quad \begin{aligned} \mathcal{L}_2\varphi(\vec{z}) \triangleq & \left[\beta - rp - p\vec{\alpha} \cdot \vec{z} + \frac{1}{2}p(1-p)(\mathbf{V}\vec{z} \cdot \vec{z}) \right] \varphi(\vec{z}) \\ & + (r + \vec{\alpha} \cdot \vec{z} - (1-p)(\mathbf{V}\vec{z} \cdot \vec{z}))\nabla\varphi(\vec{z}) \cdot \vec{z} - \frac{1}{2}(\mathbf{V}\vec{z} \cdot \vec{z})(\nabla^2\varphi(\vec{z})\vec{z} \cdot \vec{z}). \end{aligned}$$

For $i = 1, 2$, we define the first-order operators

$$(6.11) \quad \mathcal{B}_i\varphi(\vec{z}) \triangleq \lambda_i p\varphi(\vec{z}) - \varphi_i(\vec{z}) - \lambda_i \nabla\varphi(\vec{z}) \cdot \vec{z},$$

$$(6.12) \quad \mathcal{S}_i\varphi(\vec{z}) \triangleq \lambda_i p\varphi(\vec{z}) + \varphi_i(\vec{z}) - \lambda_i \nabla\varphi(\vec{z}) \cdot \vec{z}.$$

We may write the first-order operators in (5.3) applied to ψ as

$$(6.13) \quad -\psi_i(y_1, y_2, x) + \lambda_i \psi_x(y_1, y_2, x) = x^{p-1}\mathcal{B}_i\varphi(\vec{z}),$$

$$(6.14) \quad \psi_i(y_1, y_2, x) + \lambda_i \psi_x(y_1, y_2, x) = x^{p-1}\mathcal{S}_i\varphi(\vec{z}).$$

Recalling the discussion at the beginning of section 5, the reader can verify that in the region in which it is optimal to buy (respectively, sell) type- i futures, we would expect to have $\mathcal{B}_i u = 0$ (respectively, $\mathcal{S}_i u = 0$).

Definition 6.1. *A continuous function w defined on $\overline{\mathcal{D}}_2$ is a viscosity subsolution of the two-variable HJB equation*

$$(6.15) \quad \min [\mathcal{L}_2 w - \tilde{U}(pw - \nabla w \cdot \vec{z}), \mathcal{B}_1 w, \mathcal{B}_2 w, \mathcal{S}_1 w, \mathcal{S}_2 w] = 0$$

if for every point $\vec{z} \in \mathcal{D}_2$ and for every C^2 function φ defined on \mathcal{D}_2 and agreeing with φ at \vec{z} , $\varphi \geq w$ on \mathcal{D}_2 implies that

$$(6.16) \quad \min[\mathcal{L}_2\varphi - \tilde{U}(p\varphi - \nabla\varphi \cdot \vec{z}), \mathcal{B}_1\varphi, \mathcal{B}_2\varphi, \mathcal{S}_1\varphi, \mathcal{S}_2\varphi]$$

is nonpositive at \vec{z} . We say w is a viscosity supersolution of (6.15) if for every point $\vec{z} \in \mathcal{D}_2$ and every C^2 function φ defined on \mathcal{D}_2 and agreeing with φ at \vec{z} , $\varphi \leq w$ on \mathcal{D}_2 implies that the expression in (6.16) is nonnegative at \vec{z} . We say w is a viscosity solution of (6.15) if it is both a viscosity subsolution and a viscosity supersolution.

We have a counterpart to Theorem 5.1 for the two-variable function u .

Theorem 6.2. *When λ is strictly positive, the function u defined by (6.3) is continuous on \mathcal{D}_2 defined by (6.2) and is a viscosity solution of (6.15).*

Proof. Because v and u are related by (6.3), continuity of u on its domain $\overline{\mathcal{D}}_2$ follows from continuity of v on its domain $\overline{\mathcal{D}}_3$. If (z_1, z_2) is in \mathcal{D}_2 and φ is a C^2 function defined on \mathcal{D}_2 , agreeing with u at (z_1, z_2) , and dominating u on \mathcal{D}_2 , then $\psi(z_1, z_2, 1)$ defined by (6.6) is a

C^2 function defined on \mathcal{D}_3 , agreeing with v at $(z_1, z_2, 1)$, and dominating v on \mathcal{D}_3 . Theorem 5.1 implies that the expression (5.4) is nonpositive at (z_1, z_2, x) , and because of (6.7)–(6.9), (6.13), and (6.14), this implies that the expression in (6.16) is nonpositive at (z_1, z_2) . Hence, u is a viscosity subsolution of (6.15). The viscosity supersolution proof is the same. ■

Theorem 6.3 (comparison). *Assume $\lambda > 0$. Let the continuous functions w^+ and w^- defined on $\overline{\mathcal{D}}_2$ be a viscosity supersolution and a viscosity subsolution, respectively, of (6.15). If $w^+ \geq w^- = 0$ on $\partial\mathcal{D}_2$ and w^+ is strictly positive on \mathcal{D}_2 , then $w^+ \geq w^-$ on $\overline{\mathcal{D}}_2$.*

The proof of Theorem 6.3 is provided in Appendix B. Because u is both a supersolution and a subsolution of (6.15), satisfies (6.5), and is strictly positive on \mathcal{D}_2 , Theorem 6.3 has an immediate corollary.

Corollary 6.4. *Let w^\pm be as in Theorem 6.3. Then $w^- \leq u \leq w^+$ on $\overline{\mathcal{D}}_2$. In particular, u is the unique continuous function defined on \mathcal{D}_2 that vanishes on $\partial\mathcal{D}_2$, is strictly positive in \mathcal{D}_2 , and is a viscosity solution of (6.15).*

We close this section with an observation we will need in section 10.

Lemma 6.5. *We have $u(\vec{0}) = u(\vec{\theta}) + O(\lambda)$, where $\vec{0}$ denotes the origin in \mathbb{R}^2 .*

Proof. By selling θ_1 type-1 futures contracts and θ_2 type-2 futures contracts, the agent can move instantaneously from the position $(\theta_1, \theta_2, 1)$ to $(0, 0, 1 - \vec{\lambda} \cdot \vec{\theta})$ in \mathcal{D}_3 . Therefore, because of (6.4),

$$u(\vec{\theta}) = v(\theta_1, \theta_2, 1) \geq v(0, 0, 1 - \vec{\lambda} \cdot \vec{\theta}) = (1 - \vec{\lambda} \cdot \vec{\theta})^p v(0, 0, 1) = u(\vec{0}) + O(\lambda).$$

On the other hand, by purchasing $\theta_1/(1 + \vec{\lambda} \cdot \vec{\theta})$ type-1 and $\theta_2/(1 + \vec{\lambda} \cdot \vec{\theta})$ type-2 futures contracts, the agent can move instantaneously from the position $(0, 0, 1)$ to $\vec{\xi} \triangleq (\theta_1/(1 + \vec{\lambda} \cdot \vec{\theta}), \theta_2/(1 + \vec{\lambda} \cdot \vec{\theta}), 1/(1 + \vec{\lambda} \cdot \vec{\theta}))$ in \mathbb{R}^3 . Therefore,

$$u(\vec{0}) = v(0, 0, 1) \geq v(\vec{\xi}) = \left(\frac{1}{1 + \vec{\lambda} \cdot \vec{\theta}} \right)^p v(\theta_1, \theta_2, 1) = u(\vec{\theta}) + O(\lambda). \quad \blacksquare$$

7. Partitioning the solvency region. The left-hand side of the HJB equation (6.15) is the minimum of five terms. We conjecture but do not prove that the solvency region is partitioned into nine corresponding regions (see Figure 1) with the properties set out in this section. This section is for orientation purposes only and is not part of the proof development of the paper.

There should be an (open) middle region M in which $\mathcal{L}_2 u - \tilde{U}(pu - \nabla u \cdot \vec{z}) = 0$. The point $\vec{\theta}$ is inside the middle region M , and since $\vec{\theta}$ is in the open first quadrant by assumption, for sufficiently small $\lambda > 0$ we expect the region M to be a subset of the open first quadrant. When (Z_1, Z_2) , where $Z_i \triangleq \frac{Y_i}{X}$, $i = 1, 2$, is in this region, it is optimal to consume but not transact. Transactions occur at the boundaries of M to prevent (Z_1, Z_2) from exiting \overline{M} , the closure of M .

In addition, there should be an (open) “southern” region S , and when (Z_1, Z_2) is in this region, it is optimal to increase the position in futures contracts of type two, so that (Z_1, Z_2) moves along the ray emanating from $-(0, 1/\lambda_2)$ until it reaches ∂M . Indeed, suppose we begin with (Y_1, Y_2, X) at the point (y_1^0, y_2^0, x^0) so that (Z_1, Z_2) is at the point (z_1^0, z_2^0) , where $z_i^0 = y_i^0/x^0$. If we increase the position in type-two futures by $t > 0$, we move to the new point $(y_1^0, y_2^0 + t, x^0 - \lambda_2 t)$ in \mathbb{R}^3 , and in the two variables, we move to the point

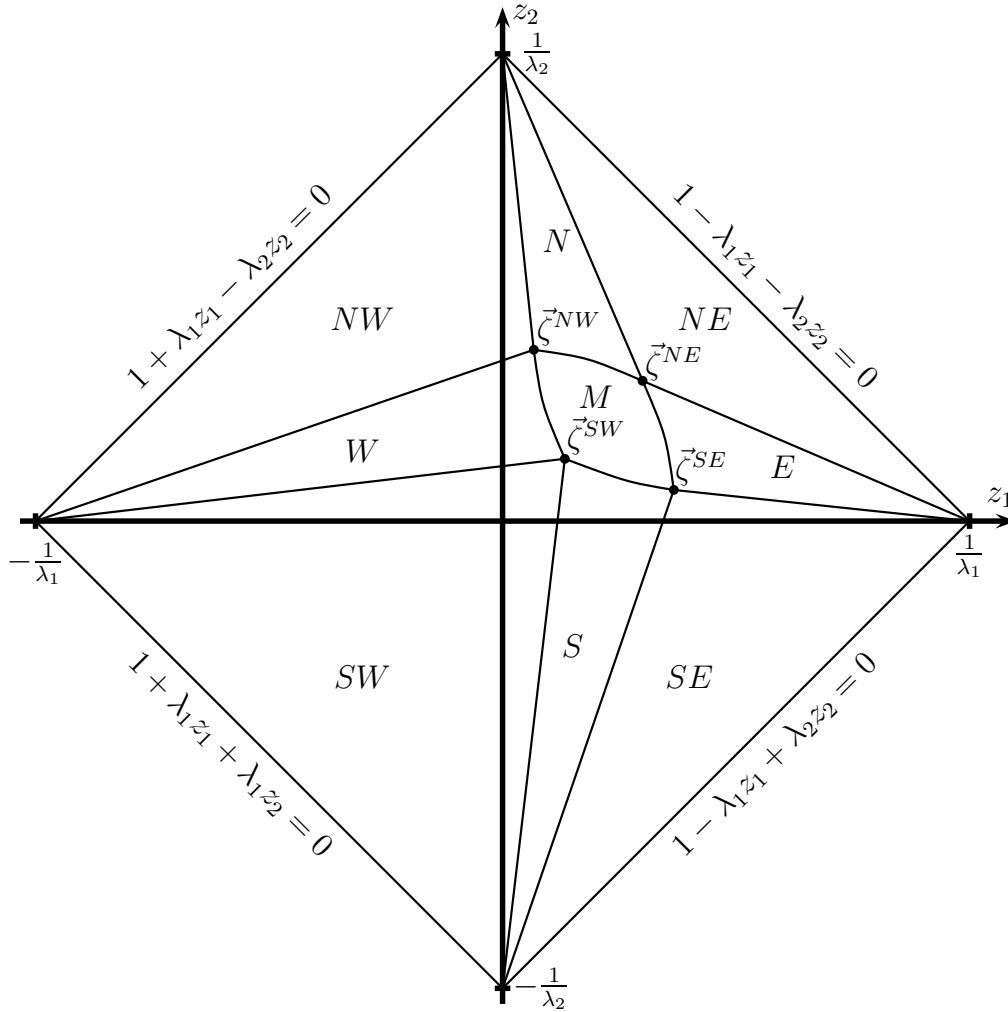


Figure 1. Solvency region.

$(z_1, z_2) = (y_1^0/(x^0 - \lambda_2 t), (y_2^0 + t)/(x^0 - \lambda_2 t))$. A bit of algebra confirms that this point is on the line passing through $(0, -1/\lambda_2)$ and (z_1^0, z_2^0) , whose equation is

$$(7.1) \quad \frac{z_1}{z_1^0} = \frac{1 + \lambda_2 z_2}{1 + \lambda_2 z_2^0}.$$

Similarly, there is an (open) “eastern” region in which it is optimal to sell type-one futures, moving in \mathcal{D}_2 along the ray emanating from $(1/\lambda_1, 0)$ until ∂M is reached. In the (open) “southeastern” region, it is optimal to simultaneously buy type-two futures and sell type-one futures, moving along the ray in \mathcal{D} that connects the initial position with the “corner” of M where S , E , and SE meet. We call this corner $\vec{\zeta}^{SE}$.

In W one should buy type-one futures, and in N one should sell type-two futures. In

SW one should simultaneously buy type-one and type-two futures, in NW one should simultaneously buy type-one and sell type-two futures, and in NE one should simultaneously sell type-one and type-two futures. In these three cases, the transactions indicated should move the state to the “nearest corner” of M . If the initial position is outside \overline{M} , then there should be an initial transaction that moves the state process (Z_1, Z_2) to ∂M as described above, and thereafter the state process should remain in \overline{M} , with the agent transacting just enough on ∂M to achieve this.

In the region S , in which it is optimal to buy type-two futures, we expect to have $\mathcal{B}_2 u = 0$. In the region SE , in which it is optimal to simultaneously buy type-two futures and sell type-one futures, we expect to have $\mathcal{B}_2 u = 0$ and $\mathcal{S}_1 u = 0$.

We do not actually construct this partition and the function u . Instead, we consider the case where the futures prices are uncorrelated ($\rho = 0$), and we construct a subsolution and a supersolution of (6.15) by setting up a partition of \mathcal{D}_2 of the form in Figure 1, constructing a function inside \overline{M} and using the appropriate first-order equations for each of the outer regions to extend this function to $\mathcal{D}_2 \setminus \overline{M}$. The case $\rho \neq 0$ requires a more subtle analysis based on an auxiliary problem in which the two types of futures are uncorrelated. We thus use the solution of the problem when $\rho = 0$ as a tool to treat the problem when $\rho \neq 0$. Because we need to extend a function defined on a middle region \overline{M} to $\mathcal{D}_2 \setminus \overline{M}$ more than once, we provide that construction generically in the next section.

8. Extension to $\mathcal{D}_2 \setminus \overline{M}$. In this section we assume that \mathcal{D}_2 has been partitioned into nine regions, as shown in Figure 1. These are not necessarily the regions discussed in section 7; i.e., they are not necessarily related to the optimal policy or the value function u . We assume that the nonempty open middle region M is a subset of the open first quadrant and is bounded by the graphs of four C^1 functions:

$$(8.1) \quad \text{southern boundary: } \partial_S M \triangleq \{(\zeta_1, \zeta_2) : \zeta_2 = f_S(\zeta_1), \zeta_1^{SW} \leq \zeta_1 \leq \zeta_1^{SE}\},$$

$$(8.2) \quad \text{eastern boundary: } \partial_E M \triangleq \{(\zeta_1, \zeta_2) : \zeta_1 = f_E(\zeta_2), \zeta_2^{SE} \leq \zeta_2 \leq \zeta_2^{NE}\},$$

$$(8.3) \quad \text{northern boundary: } \partial_N M \triangleq \{(\zeta_1, \zeta_2) : \zeta_2 = f_N(\zeta_1), \zeta_1^{NW} \leq \zeta_1 \leq \zeta_1^{NE}\},$$

$$(8.4) \quad \text{western boundary: } \partial_W M \triangleq \{(\zeta_1, \zeta_2) : \zeta_1 = f_W(\zeta_2), \zeta_2^{SW} \leq \zeta_2 \leq \zeta_2^{NW}\}.$$

All other boundaries in Figure 1 are straight lines. We further assume that the functions $f_S, f_E, f_N,$ and f_W satisfy the conditions

$$(8.5) \quad f'_S(\zeta_1) < \frac{1 + \lambda_2 f_S(\zeta_1)}{\lambda_2 \zeta_1}, \quad \zeta_1^{SW} \leq \zeta_1 \leq \zeta_1^{SE},$$

$$(8.6) \quad f'_E(\zeta_2) > \frac{-1 + \lambda_1 f_E(\zeta_2)}{\lambda_1 \zeta_2}, \quad \zeta_2^{SE} \leq \zeta_2 \leq \zeta_2^{NE},$$

$$(8.7) \quad f'_N(\zeta_1) > \frac{-1 + \lambda_2 f_N(\zeta_1)}{\lambda_2 \zeta_1}, \quad \zeta_1^{NW} \leq \zeta_1 \leq \zeta_1^{NE},$$

$$(8.8) \quad f'_W(\zeta_2) < \frac{1 + \lambda_1 f_W(\zeta_2)}{\lambda_1 \zeta_2}, \quad \zeta_2^{SW} \leq \zeta_2 \leq \zeta_2^{NW}.$$

Remark 8.1. In sections 8–10, we repeatedly begin with a function defined on \overline{M} , extend it to $\overline{\mathcal{D}_2}$, and develop its properties in the nonmiddle regions $S, SE, E, NE, N, NW, W,$

and SW . The proofs follow the same pattern. We prove the desired properties in S and SE only, and we provide key equations to assist the reader in developing the proof for E . The proof for W can be obtained by interchanging the coordinates in the proof given for S , and the proof in N can be obtained by interchanging the coordinates in E .

Given a point \vec{z} in $\overline{S} \cap \mathcal{D}_2$, there is a unique point $\vec{\zeta} \in \partial_S M$ lying on the line passing through \vec{z} and the southern vertex $(0, -1/\lambda_2)$ of \mathcal{D}_2 , as we now show. This line intersects $\partial_S M$ at any point $\vec{\zeta}$ satisfying $\zeta_1^{SW} \leq \zeta_1 \leq \zeta_1^{SE}$ and (cf. (7.1))

$$\frac{1 + \lambda_2 f_S(\zeta_1)}{\zeta_1} = \frac{1 + \lambda_2 z_2}{z_1}.$$

To solve this equation for ζ_1 , we must invert the function $\zeta_1 \mapsto (1 + \lambda_2 f_S(\zeta_1))/\zeta_1$. The derivative of this function is strictly negative for $\zeta_1^{SW} \leq \zeta_1 \leq \zeta_1^{SE}$ because of (8.5), and hence this function has a C^1 inverse

$$g_S: \left[\frac{1 + \lambda_2 f_S(\zeta_1^{SE})}{\zeta_1^{SE}}, \frac{1 + \lambda_2 f_S(\zeta_1^{SW})}{\zeta_1^{SW}} \right] \xrightarrow{\text{onto}} [\zeta_1^{SW}, \zeta_1^{SE}].$$

Because $\vec{z} \in \overline{S}$, $\frac{1 + \lambda_2 z_2}{z_1}$ is in the domain of g_S , and the point we seek is given by $\zeta_1 = g_S(\frac{1 + \lambda_2 z_2}{z_1})$, $\zeta_2 = f_S(\zeta_1)$. The mapping from \vec{z} to $\vec{\zeta}$ defined this way is C^1 .

If \vec{z} is in $\overline{E} \cap \mathcal{D}_2$, to find $\vec{\zeta} \in \partial_E M$ lying on the line through \vec{z} and the eastern vertex $(\frac{1}{\lambda_1}, 0)$, we must solve the equation $(1 - \lambda_1 f_E(\zeta_2))/\zeta_2 = (1 - \lambda_1 z_1)/z_2$ for ζ_2 . We proceed as before.

Remark 8.2. For future reference, we record here the conclusion that if $\vec{z} \in \overline{S} \cap \mathcal{D}_2$, then there is a unique point $\vec{\zeta} \in \partial_S M$ satisfying the equations

$$(8.9) \quad \frac{z_1}{\zeta_1} = \frac{1 + \lambda_2 z_2}{1 + \lambda_2 \zeta_2}, \quad \zeta_2 = f_S(\zeta_1).$$

We shall abuse notation, using $\vec{\zeta}$ to denote both the point in $\partial_S M$ and the mapping on $\overline{S} \cap \mathcal{D}_2$. The mapping $\vec{\zeta}$ is a C^1 function on S , and its first derivatives have continuous extensions to $\overline{S} \cap \mathcal{D}_2$. The corresponding equations for \vec{z} in $\overline{E} \cap \mathcal{D}_2$ and $\vec{\zeta} \in \partial_E M$ are

$$(8.10) \quad \frac{z_2}{\zeta_2} = \frac{1 - \lambda_1 z_1}{1 - \lambda_1 \zeta_1}, \quad \zeta_1 = f_E(\zeta_2).$$

Finally, we define a mapping on $\vec{\zeta}$ on $\overline{SE} \cap \mathcal{D}_2$ to be identically $\vec{\zeta}^{SE}$. We define analogous mappings for the other regions.

If \vec{z} lies in two regions, the two mappings defined at \vec{z} agree, and hence we can use the single symbol $\vec{\zeta}$ for both these mappings. For example, if $\vec{z} \in (\overline{S} \cap \overline{SE}) \cap \mathcal{D}_2$, then $\vec{\zeta} = \vec{\zeta}^{SE}$. Using this fact and the first equation in (8.9), one can verify (8.11). We have

$$(8.11) \quad \frac{z_1}{\zeta_1^{SE}} = \frac{1 + \lambda_2 z_2}{1 + \lambda_2 \zeta_2^{SE}} = \frac{1 - \lambda_1 z_1 + \lambda_2 z_2}{1 - \lambda_1 \zeta_1^{SE} + \lambda_2 \zeta_2^{SE}}, \quad \vec{z} \in (\overline{S} \cap \overline{SE}) \cap \mathcal{D}_2,$$

$$(8.12) \quad \frac{z_2}{\zeta_2^{SE}} = \frac{1 - \lambda_1 z_1}{1 - \lambda_1 \zeta_1^{SE}} = \frac{1 - \lambda_1 z_1 + \lambda_2 z_2}{1 - \lambda_1 \zeta_1^{SE} + \lambda_2 \zeta_2^{SE}}, \quad \vec{z} \in (\overline{E} \cap \overline{SE}) \cap \mathcal{D}_2.$$

The verification of (8.12) uses the first equation in (8.10). The second equation in each of (8.11) and (8.12) is a consequence of the first equation in these displays.

Definition 8.3. Let \mathcal{O} be an open subset of \mathbb{R}^2 , and let $\widehat{\mathcal{O}}$ be a set satisfying $\mathcal{O} \subset \widehat{\mathcal{O}} \subset \overline{\mathcal{O}}$. We say that a function w is C^k on $\widehat{\mathcal{O}}$ if w is C^k on \mathcal{O} and the partial derivatives of w up to order k have continuous extensions to $\widehat{\mathcal{O}}$. We call these continuous extensions the derivatives defined from inside \mathcal{O} , and we denote them by $w_i^{\mathcal{O}}$, $w_{i,j}^{\mathcal{O}}$, etc.

Theorem 8.4 (function extension). Let $\overline{\mathcal{D}}_2$ be partitioned into nine regions, as described in this section. For $\vec{z} \in \mathcal{D}_2 \setminus M$, let $\vec{\zeta}$ denote the point on ∂M defined in Remark 8.2. Let w^M be defined and C^1 on \overline{M} , and extend w^M to $\overline{\mathcal{D}}_2$ by the formula

$$(8.13) \quad w(\vec{z}) = \begin{cases} w^M(\vec{z}) & \text{if } \vec{z} \in \overline{M}, \\ \left(\frac{1 + \lambda_2 z_2}{1 + \lambda_2 \zeta_2} \right)^p w^M(\vec{\zeta}) = \left(\frac{z_1}{\zeta_1} \right)^p w^M(\vec{\zeta}) & \text{if } \vec{z} \in \overline{S} \cap \mathcal{D}_2, \\ \left(\frac{1 - \lambda_1 z_1 + \lambda_2 z_2}{1 - \lambda_1 \zeta_1^{SE} + \lambda_2 \zeta_2^{SE}} \right)^p w^M(\vec{\zeta}^{SE}) & \text{if } \vec{z} \in \overline{SE} \cap \mathcal{D}_2, \\ \left(\frac{1 - \lambda_1 z_1}{1 - \lambda_1 \zeta_1} \right)^p w^M(\vec{\zeta}) = \left(\frac{z_2}{\zeta_2} \right)^p w^M(\vec{\zeta}) & \text{if } \vec{z} \in \overline{E} \cap \mathcal{D}_2, \\ \left(\frac{1 - \lambda_1 z_1 - \lambda_2 z_2}{1 - \lambda_1 \zeta_1^{NE} - \lambda_2 \zeta_2^{NE}} \right)^p w^M(\vec{\zeta}^{NE}) & \text{if } \vec{z} \in \overline{NE} \cap \mathcal{D}_2, \\ \left(\frac{1 - \lambda_2 z_2}{1 - \lambda_2 \zeta_2} \right)^p w^M(\vec{\zeta}) = \left(\frac{z_1}{\zeta_1} \right)^p w^M(\vec{\zeta}) & \text{if } \vec{z} \in \overline{N} \cap \mathcal{D}_2, \\ \left(\frac{1 + \lambda_1 z_1 - \lambda_2 z_2}{1 + \lambda_1 \zeta_1^{NW} - \lambda_2 \zeta_2^{NW}} \right)^p w^M(\vec{\zeta}^{NW}) & \text{if } \vec{z} \in \overline{NW} \cap \mathcal{D}_2, \\ \left(\frac{1 + \lambda_1 z_1}{1 + \lambda_1 \zeta_1} \right)^p w^M(\vec{\zeta}) = \left(\frac{z_2}{\zeta_2} \right)^p w^M(\vec{\zeta}) & \text{if } \vec{z} \in \overline{W} \cap \mathcal{D}_2, \\ \left(\frac{1 + \lambda_1 z_1 + \lambda_2 z_2}{1 + \lambda_1 \zeta_1^{SW} + \lambda_2 \zeta_2^{SW}} \right)^p w^M(\vec{\zeta}^{SW}) & \text{if } \vec{z} \in \overline{SW} \cap \mathcal{D}_2. \end{cases}$$

(Note that \vec{z} can fall into more than one case in (8.13), but the definition of $w(\vec{z})$ is unambiguous because of (8.11), (8.12), and their analogues for the other six straight-line boundaries inside \mathcal{D}_2 .) Define $w = 0$ on $\partial \mathcal{D}_2$. Then w is continuous on $\overline{\mathcal{D}}_2$, w is C^1 in each of the non-middle open regions, and the first derivatives of w have continuous extensions to the closures of each of these regions intersected with \mathcal{D}_2 (although these extensions may differ for different regions), and

$$(8.14) \quad \mathcal{B}_2 w^S = 0 \text{ on } \overline{S} \cap \mathcal{D}_2, \quad \mathcal{S}_1 w^E = 0 \text{ on } \overline{E} \cap \mathcal{D}_2,$$

$$(8.15) \quad \mathcal{S}_2 w^N = 0 \text{ on } \overline{N} \cap \mathcal{D}_2, \quad \mathcal{B}_1 w^W = 0 \text{ on } \overline{W} \cap \mathcal{D}_2,$$

$$(8.16) \quad \mathcal{B}_2 w^{SE} = \mathcal{S}_1 w^{SE} = 0 \text{ on } \overline{SE} \cap \mathcal{D}_2,$$

$$(8.17) \quad \mathcal{S}_1 w^{NE} = \mathcal{S}_2 w^{NE} = 0 \text{ on } \overline{NE} \cap \mathcal{D}_2,$$

$$(8.18) \quad \mathcal{S}_2 w^{NW} = \mathcal{B}_1 w^{NW} = 0 \text{ on } \overline{NW} \cap \mathcal{D}_2,$$

$$(8.19) \quad \mathcal{B}_1 w^{SW} = \mathcal{B}_2 w^{SW} = 0 \text{ on } \overline{SW} \cap \mathcal{D}_2.$$

In addition, we have the following implications:

$$(8.20) \quad \mathcal{B}_2 w^M = 0 \text{ on } \partial_S M \implies w_i^S(\vec{z}) = \left(\frac{z_1}{\zeta_1}\right)^{p-1} w_i^M(\vec{\zeta}), \quad i = 1, 2, \quad \vec{z} \in \overline{S} \cap \mathcal{D}_2,$$

$$(8.21) \quad \mathcal{S}_1 w^M = 0 \text{ on } \partial_E M \implies w_i^E(\vec{z}) = \left(\frac{z_2}{\zeta_2}\right)^{p-1} w_i^M(\vec{\zeta}), \quad i = 1, 2, \quad \vec{z} \in \overline{E} \cap \mathcal{D}_2,$$

$$(8.22) \quad \mathcal{S}_2 w^M = 0 \text{ on } \partial_N M \implies w_i^N(\vec{z}) = \left(\frac{z_1}{\zeta_1}\right)^{p-1} w_i^M(\vec{\zeta}), \quad i = 1, 2, \quad \vec{z} \in \overline{N} \cap \mathcal{D}_2,$$

$$(8.23) \quad \mathcal{B}_1 w^M = 0 \text{ on } \partial_W M \implies w_i^W(\vec{z}) = \left(\frac{z_2}{\zeta_2}\right)^{p-1} w_i^M(\vec{\zeta}), \quad i = 1, 2, \quad \vec{z} \in \overline{W} \cap \mathcal{D}_2,$$

$$\mathcal{B}_2 w^M(\vec{\zeta}^{SE}) = \mathcal{S}_1 w^M(\vec{\zeta}^{SE}) = 0$$

$$\implies w_i^{SE}(\vec{z}) = \left(\frac{1 - \lambda_1 z_1 + \lambda_2 z_2}{1 - \lambda_1 \zeta_1^{SE} + \lambda_2 \zeta_2^{SE}}\right)^{p-1} w_i^M(\vec{\zeta}^{SE}), \quad i = 1, 2, \quad \vec{z} \in \overline{SE} \cap \mathcal{D}_2,$$

$$(8.24) \quad \text{and } w \text{ is } C^1 \text{ at } \vec{\zeta}^{SE},$$

$$\mathcal{S}_1 w^M(\vec{\zeta}^{NE}) = \mathcal{S}_2 w^M(\vec{\zeta}^{NE}) = 0$$

$$\implies w_i^{NE}(\vec{z}) = \left(\frac{1 - \lambda_1 z_1 - \lambda_2 z_2}{1 - \lambda_1 \zeta_1^{NE} - \lambda_2 \zeta_2^{NE}}\right)^{p-1} w_i^M(\vec{\zeta}^{NE}), \quad i = 1, 2, \quad \vec{z} \in \overline{NE} \cap \mathcal{D}_2,$$

$$(8.25) \quad \text{and } w \text{ is } C^1 \text{ at } \vec{\zeta}^{NE},$$

$$\mathcal{S}_2 w^M(\vec{\zeta}^{NW}) = \mathcal{B}_1 w^M(\vec{\zeta}^{NW}) = 0$$

$$\implies w_i^{NW}(\vec{z}) = \left(\frac{1 + \lambda_1 z_1 - \lambda_2 z_2}{1 + \lambda_1 \zeta_1^{NW} - \lambda_2 \zeta_2^{NW}}\right)^{p-1} w_i^M(\vec{\zeta}^{NW}), \quad i = 1, 2, \quad \vec{z} \in \overline{NW} \cap \mathcal{D}_2,$$

$$(8.26) \quad \text{and } w \text{ is } C^1 \text{ at } \vec{\zeta}^{NW},$$

$$\mathcal{B}_1 w^M(\vec{\zeta}^{SW}) = \mathcal{B}_2 w^M(\vec{\zeta}^{SW}) = 0$$

$$\implies w_i^{SW}(\vec{z}) = \left(\frac{1 + \lambda_1 z_1 + \lambda_2 z_2}{1 + \lambda_1 \zeta_1^{SW} + \lambda_2 \zeta_2^{SW}}\right)^{p-1} w_i^M(\vec{\zeta}^{SW}), \quad i = 1, 2, \quad \vec{z} \in \overline{SW} \cap \mathcal{D}_2$$

$$(8.27) \quad \text{and } w \text{ is } C^1 \text{ at } \vec{\zeta}^{SW}.$$

Consequently, if

$$(8.28) \quad \mathcal{B}_2 w^M = 0 \text{ on } \partial_S M, \quad \mathcal{S}_1 w^M = 0 \text{ on } \partial_E M, \quad \mathcal{S}_2 w^M = 0 \text{ on } \partial_N M, \quad \mathcal{B}_1 w^M = 0 \text{ on } \partial_W M,$$

then w is C^1 on \mathcal{D}_2 , and hence we may omit the superscripts on the partial derivatives w_i in (8.14)–(8.27). Furthermore, under condition (8.28),

$$(8.29) \quad \mathcal{B}_2 w = 0 \text{ on } (\overline{SW} \cup \overline{S} \cup \overline{SE}) \cap \mathcal{D}_2, \quad \mathcal{S}_1 w = 0 \text{ on } (\overline{SE} \cup \overline{E} \cup \overline{NE}) \cap \mathcal{D}_2,$$

$$(8.30) \quad \mathcal{S}_2 w = 0 \text{ on } (\overline{NW} \cup \overline{N} \cup \overline{NE}) \cap \mathcal{D}_2, \quad \mathcal{B}_1 w = 0 \text{ on } (\overline{SW} \cup \overline{W} \cup \overline{NW}) \cap \mathcal{D}_2.$$

If, in addition,

$$(8.31) \quad \min [\mathcal{B}_1 w, \mathcal{B}_2 w, \mathcal{S}_1 w, \mathcal{S}_2 w] = 0 \text{ on } \partial M,$$

then

$$(8.32) \quad \min [\mathcal{B}_1 w, \mathcal{B}_2 w, \mathcal{S}_1 w, \mathcal{S}_2 w] = 0 \text{ on } \mathcal{D}_2 \setminus M.$$

Finally, if condition (8.28) holds and w^M is C^2 on \overline{M} , then w is C^2 on each of the nonmiddle open regions and the second derivatives of w have continuous extensions to the closure of each of these regions intersected with \mathcal{D}_2 .

Proof. It is immediate from its definition that w is continuous on $\overline{\mathcal{D}_2}$. It is also immediate from Remark 8.2 and the definition of w that in each of the eight nonmiddle open regions, w is C^1 and its first derivatives in each region have limits as the boundaries of the region are approached, except for the boundary of \mathcal{D}_2 , where the derivatives explode. Therefore, it suffices to prove (8.14)–(8.19) in the interior of each of the indicated regions.

We prove the first equality in (8.14). Let $\vec{z} \in S$ be given. Then there is a point $\vec{\zeta} \in \partial_S M$ so that \vec{z} lies on the line segment connecting $(0, -1/\lambda_2)$ with $\vec{\zeta}$. We parameterize this line as $z_1(t) = t\lambda_2 z_1$, $z_2(t) = -(1/\lambda_2) + t(1 + \lambda_2 z_2)$. We use the second equation in the second line of (8.13) to compute the directional derivative of w along this line, evaluated at \vec{z} , which is $\frac{d}{dt} w(\vec{z}(t))|_{t=1/\lambda_2} = p\lambda_2 w(\vec{z})$. But this directional derivative evaluated at \vec{z} is also

$$z'_1(1/\lambda_2)w_1(\vec{z}) + z'_2(1/\lambda_2)w_2(\vec{z}) = \lambda_2 z_1 w_1(\vec{z}) + (1 + \lambda_2 z_2)w_2(\vec{z}).$$

Setting these two equal we obtain $\mathcal{B}_2 w(\vec{z}) = 0$.

For $\vec{z} \in E$, we parameterize the line through \vec{z} connecting $(1/\lambda_1, 0)$ with $\zeta \in \partial_E M$ by $z_1(t) = 1/\lambda_1 - t(1 - \lambda_1 z_1)$, $z_2(t) = t\lambda_1 z_2$. We use the second equation in the fourth line of (8.13), compute the directional at \vec{z} by two methods, and obtain $\mathcal{S}_1 w = 0$.

For $z \in SE$, we simply compute the partial derivatives in the third line of (8.13) and verify that (8.16) holds.

We prove the implications (8.20), (8.21), and (8.24). Again, it suffices to consider \vec{z} in the interior of each set. Assume the hypothesis of (8.20). For $\vec{z} \in S$, we compute

$$(8.33) \quad \begin{aligned} w_1(\vec{z}) &= \frac{\partial}{\partial z_1} \left[\left(\frac{1 + \lambda_2 z_2}{1 + \lambda_2 \zeta_2} \right)^p w^M(\vec{\zeta}) \right] \\ &= \left(\frac{z_1}{\zeta_1} \right)^p \left[-\frac{\lambda_2 p}{1 + \lambda_2 \zeta_2} w^M(\vec{\zeta}) \frac{\partial \zeta_2}{\partial z_1} + w_1^M(\vec{\zeta}) \frac{\partial \zeta_1}{\partial z_1} + w_2^M(\vec{\zeta}) \frac{\partial \zeta_2}{\partial z_1} \right], \end{aligned}$$

where we have made the substitution (8.9) after computing the derivative. We use the hypothesis of (8.20) in the form

$$(8.34) \quad -\frac{\lambda_2 p}{1 + \lambda_2 \zeta_2} w^M(\vec{\zeta}) + w_2^M(\vec{\zeta}) = -\frac{\lambda_2 \zeta_1}{1 + \lambda_2 \zeta_2} w_1^M(\vec{\zeta})$$

to simplify this, obtaining

$$(8.35) \quad w_1(\vec{z}) = \left(\frac{z_1}{\zeta_1} \right)^p w_1^M(\vec{\zeta}) \left[\frac{\partial \zeta_1}{\partial z_1} - \frac{\lambda_2 \zeta_1}{1 + \lambda_2 \zeta_2} \cdot \frac{\partial \zeta_2}{\partial z_1} \right].$$

However, (8.9) implies

$$(8.36) \quad \frac{\lambda_2}{1 + \lambda_2 z_2} = \frac{\partial}{\partial z_1} \left(\frac{\lambda_2 z_1}{1 + \lambda_2 z_2} \right) = \frac{\partial}{\partial z_1} \left(\frac{\lambda_2 \zeta_1}{1 + \lambda_2 \zeta_2} \right) = \frac{\lambda_2}{1 + \lambda_2 \zeta_2} \left[\frac{\partial \zeta_1}{\partial z_1} - \frac{\lambda_2 \zeta_1}{1 + \lambda_2 \zeta_2} \cdot \frac{\partial \zeta_2}{\partial z_1} \right],$$

and thus

$$(8.37) \quad \frac{\partial \zeta_1}{\partial z_1} - \frac{\lambda_2 \zeta_1}{1 + \lambda_2 \zeta_2} \cdot \frac{\partial \zeta_2}{\partial z_1} = \frac{1 + \lambda_2 \zeta_2}{1 + \lambda_2 z_2} = \frac{\zeta_1}{z_1}.$$

Substitution of this equation into (8.35) gives us the formula in (8.20) for w_1^S .

For $\vec{z} \in S$, we compute

$$(8.38) \quad \begin{aligned} w_2(\vec{z}) &= \frac{\partial}{\partial z_2} \left[\left(\frac{z_1}{\zeta_1} \right)^p w^M(\vec{\zeta}) \right] \\ &= \left(\frac{z_1}{\zeta_1} \right)^p \left[-\frac{p}{\zeta_1} w^M(\vec{\zeta}) \frac{\partial \zeta_1}{\partial z_2} + w_1^M(\vec{\zeta}) \frac{\partial \zeta_1}{\partial z_2} + w_2^M(\vec{\zeta}) \frac{\partial \zeta_2}{\partial z_2} \right]. \end{aligned}$$

We again use the hypothesis of (8.20), this time in the form

$$(8.39) \quad -\frac{p}{\zeta_1} w(\vec{\zeta}) + w_1^M(\vec{\zeta}) = -\frac{1 + \lambda_2 \zeta_2}{\lambda_2 \zeta_1} w_2^M(\vec{\zeta}),$$

to simplify, obtaining

$$(8.40) \quad w_2(\vec{z}) = \left(\frac{z_1}{\zeta_1} \right)^p w_2^M(\vec{\zeta}) \left[\frac{\partial \zeta_2}{\partial z_2} - \frac{1 + \lambda_2 \zeta_2}{\lambda_1 \zeta_1} \cdot \frac{\partial \zeta_1}{\partial z_2} \right].$$

However, (8.9) implies

$$(8.41) \quad \frac{1}{z_1} = \frac{\partial}{\partial z_2} \left(\frac{1 + \lambda_2 z_2}{\lambda_2 z_1} \right) = \frac{\partial}{\partial z_2} \left(\frac{1 + \lambda_2 \zeta_2}{\lambda_2 \zeta_1} \right) = \frac{1}{\zeta_1} \left[\frac{\partial \zeta_2}{\partial z_2} - \frac{1 + \lambda_2 \zeta_2}{\lambda_2 \zeta_1} \cdot \frac{\partial \zeta_1}{\partial z_2} \right],$$

and thus

$$(8.42) \quad \frac{\partial \zeta_2}{\partial z_2} - \frac{1 + \lambda_2 \zeta_2}{\lambda_2 \zeta_1} \cdot \frac{\partial \zeta_1}{\partial z_2} = \frac{\zeta_1}{z_1}.$$

Substitution into (8.40) gives us the formula in (8.20) for w_2^S .

Before turning to (8.21), we observe that we can take the limits in (8.33) and (8.38) as $\vec{z} \in S$ approaches $\vec{\zeta}^{SE}$. If, instead of the hypothesis of (8.20), we assume only $\mathcal{B}_2 w^M(\vec{\zeta}^{SE}) = 0$, we have (8.34) and (8.39) at $\vec{\zeta} = \vec{\zeta}^{SE}$, and we may substitute these into (8.33) and (8.38), respectively. The result of this is to obtain the conclusion of (8.20) at the point $\vec{z} = \vec{\zeta}^{SE}$. Similarly, if we assume only $\mathcal{S}_1 w^M(\vec{\zeta}^{SE}) = 0$, then the argument we give for (8.21) below can be modified to show that the conclusion of (8.21) holds at the point $\vec{z} = \vec{\zeta}^{SE}$. In other words, under the hypothesis of (8.24), we have for $i = 1, 2$,

$$(8.43) \quad w_i^S(\vec{\zeta}^{SE}) = w_i^M(\vec{\zeta}^{SE}) = w_i^E(\vec{\zeta}^{SE}).$$

Hypothesis (8.21) implies $-(p/\zeta_2)w^M(\vec{\zeta}) + w_2^M(\vec{\zeta}) = ((1 - \lambda_1 \zeta_1)/(\lambda_1 \zeta_2))w_1^M(\vec{\zeta})$ for $\vec{z} \in E$. Thus $w_1(\vec{z}) = (z_2/\zeta_2)^p w_1^M(\vec{\zeta}) [\partial \zeta_1 / \partial z_1 + ((1 - \lambda_1 \zeta_1)/(\lambda_1 \zeta_2)) \partial \zeta_2 / \partial z_1]$. However, (8.10) implies $-1/z_2 = -(1/\zeta_2) [\partial \zeta_1 / \partial z_1 + ((1 - \lambda_1 \zeta_1)/(\lambda_1 \zeta_2)) \partial \zeta_2 / \partial z_1]$, and this yields the formula in (8.21) for w_1^E . We next use the hypothesis of (8.21) in the form $((\lambda_1 p)/(1 - \lambda_1 \zeta_1))w^M(\vec{\zeta}) + w_1^M(\vec{\zeta}) =$

$((\lambda_1 \zeta_2)/(1 - \lambda_1 \zeta_1))w_2^M(\vec{\zeta})$ in order to compute $w_2(\vec{z}) = (z_2/\zeta_2)^p w_2^M(\vec{\zeta})[\partial\zeta_2/\partial z_2 + ((\lambda_1 \zeta_2)/(1 - \lambda_1 \zeta_2))\partial\zeta_1/\partial z_2]$. We simplify this using (8.10) and the formula derived from (8.10), $\lambda_1/(1 - \lambda_1 z_1) = (\lambda_1/(1 - \lambda_1 \zeta_1))[\partial\zeta_2/\partial z_2 + (\lambda_1 \zeta_2/(1 - \lambda_1 \zeta_1))\partial\zeta_1/\partial z_2]$. This yields the formula in (8.21) for $w_2^E(\vec{z})$.

We next compute the derivatives w_1 and w_2 in SE under the hypothesis of (8.24), namely, $\mathcal{B}_2 w^M(\vec{\zeta}^{SE}) = \mathcal{S}_1 w^M(\vec{\zeta}^{SE}) = 0$. These equations can be written as the system

$$\begin{aligned} \lambda_2 \zeta_1^{SE} w_1^M(\vec{\zeta}^{SE}) + (1 + \lambda_2 \zeta_2^{SE}) w_2^M(\vec{\zeta}^{SE}) &= \lambda_2 p w^M(\vec{\zeta}^{SE}), \\ -(1 - \lambda_1 \zeta_1^{SE}) w_1^M(\vec{\zeta}^{SE}) + \lambda_1 \zeta_2^{SE} w_2^M(\vec{\zeta}^{SE}) &= \lambda_1 p w^M(\vec{\zeta}^{SE}). \end{aligned}$$

This leads to the derivative formulas

$$-\frac{1}{\lambda_1} w_1^M(\vec{\zeta}^{SE}) = \frac{1}{\lambda_2} w_2^M(\vec{\zeta}^{SE}) = \frac{p}{1 - \lambda_1 \zeta_1^{SE} + \lambda_2 \zeta_2^{SE}} w^M(\vec{\zeta}^{SE}).$$

Using these formulas and the fact that $\vec{\zeta} = \vec{\zeta}^{SE}$ for all $\vec{z} \in \overline{SE}$, we can differentiate directly in (8.13) to obtain the formulas for w_i^{SE} , $i = 1, 2$, in (8.24).

Under the hypothesis of (8.24), we have obtained for $i = 1, 2$ the formula $w_i^{SE}(\vec{\zeta}^{SE}) = w_i^M(\vec{\zeta}^{SE})$. Combining this with (8.43), we conclude that w is C^1 at $\vec{\zeta}^{SE}$.

Under assumption (8.28), the hypotheses of (8.20)–(8.27) are satisfied. Equations (8.11), (8.12), and their counterparts for other boundaries show that the derivative formulas in (8.20)–(8.27) agree on the boundaries of the regions in which they are specified. Hence, w is C^1 on \mathcal{D}_2 , and (8.29)–(8.30) follow from (8.14)–(8.19).

We next show that (8.29)–(8.31) imply (8.32). Following Remark 8.1, we check this only in $\overline{S} \cap \mathcal{D}_2$, $\overline{E} \cap \mathcal{D}_2$, and $\overline{SE} \cap \mathcal{D}_2$. The first equation in (8.29) implies $pw(\vec{z}) - \nabla w(\vec{z}) \cdot \vec{z} = \frac{1}{\lambda_2} w_2(\vec{z})$ for $\vec{z} \in \overline{S} \cap \mathcal{D}_2$. Using this and (8.20), we see that for $\vec{z} \in \overline{S} \cap \mathcal{D}_2$,

$$\begin{aligned} \mathcal{B}_1 w(\vec{z}) &= \frac{\lambda_1}{\lambda_2} w_2(\vec{z}) - w_1(\vec{z}) = \left(\frac{z_1}{\zeta_1}\right)^{p-1} \left[\frac{\lambda_1}{\lambda_2} w_2(\vec{\zeta}) - w_1(\vec{\zeta}) \right], \\ \mathcal{S}_1 w(\vec{z}) &= \frac{\lambda_1}{\lambda_2} w_2(\vec{z}) + w_1(\vec{z}) = \left(\frac{z_1}{\zeta_1}\right)^{p-1} \left[\frac{\lambda_1}{\lambda_2} w_2(\vec{\zeta}) + w_1(\vec{\zeta}) \right], \\ \mathcal{S}_2 w(\vec{z}) &= 2w_2(\vec{z}) = 2 \left(\frac{z_1}{\zeta_1}\right)^{p-1} w_2(\vec{\zeta}). \end{aligned}$$

Substituting $\vec{\zeta}$ into these equations, we obtain formulas for $\mathcal{B}_1 w(\vec{\zeta})$, $\mathcal{S}_1 w(\vec{\zeta})$, and $\mathcal{S}_2 w(\vec{\zeta})$, and we then see that for $\vec{z} \in \overline{S} \cap \mathcal{D}_2$,

$$(8.44) \quad \mathcal{B}_1 w(\vec{z}) = \left(\frac{z_1}{\zeta_1}\right)^{p-1} \mathcal{B}_1 w(\vec{\zeta}), \quad \mathcal{S}_1 w(\vec{z}) = \left(\frac{z_1}{\zeta_1}\right)^{p-1} \mathcal{S}_1 w(\vec{\zeta}), \quad \mathcal{S}_2 w(\vec{z}) = \left(\frac{z_1}{\zeta_1}\right)^{p-1} \mathcal{S}_2 w(\vec{\zeta}).$$

But (8.31) implies that $\mathcal{B}_1 w(\vec{\zeta}) \geq 0$, $\mathcal{S}_1 w(\vec{\zeta}) \geq 0$, and $\mathcal{S}_2 w(\vec{\zeta}) \geq 0$. We conclude from (8.44) and the first equation in (8.14) that (8.32) holds on $\overline{S} \cap \mathcal{D}_2$.

The second equation in (8.29) implies $pw(\vec{z}) - \nabla w(\vec{z}) \cdot \vec{z} = -\frac{1}{\lambda_1} w_1(\vec{z})$ for $\vec{z} \in \overline{E} \cap \mathcal{D}_2$. Using (8.21) we see that $\mathcal{S}_2 w(\vec{z}) = (z_2/\zeta_2)^{p-1} \mathcal{S}_2 w(\vec{\zeta}) \geq 0$, $\mathcal{B}_1 w(\vec{z}) = (z_2/\zeta_2)^{p-1} \mathcal{B}_1 w(\vec{\zeta}) \geq 0$,

$\mathcal{B}_2 w(\vec{z}) = (z_2/\zeta_2)^{p-1} \mathcal{B}_2 w_1(\vec{\zeta}) \geq 0$. This and the second equation in (8.14) imply that (8.32) holds on $\overline{E} \cap \mathcal{D}_2$.

The proof that (8.32) holds on $\overline{SE} \setminus \partial \mathcal{D}_2$ follows a similar argument. For \vec{z} in this set, we showed in (8.16) that $\mathcal{B}_2 w(\vec{z}) = 0$. Using this and (8.24), we obtain in place of (8.44) the formulas

$$\begin{aligned} \mathcal{B}_1 w(\vec{z}) &= \left(\frac{1 - \lambda_1 z_1 + \lambda_2 z_2}{1 - \lambda_1 \zeta_1^{SE} + \lambda_2 \zeta_2^{SE}} \right)^{p-1} \mathcal{B}_1 w(\vec{\zeta}^{SE}), \\ \mathcal{S}_i w(\vec{z}) &= \left(\frac{1 - \lambda_1 z_1 + \lambda_2 z_2}{1 - \lambda_1 \zeta_1^{SE} + \lambda_2 \zeta_2^{SE}} \right)^{p-1} \mathcal{S}_i w(\vec{\zeta}^{SE}), \quad i = 1, 2. \end{aligned}$$

Therefore, (8.31) implies that (8.32) holds on $\overline{SE} \setminus \partial \mathcal{D}_2$.

Finally, the mapping from \vec{z} to $\vec{\zeta}$ is C^1 in each of the open sets S , SE , E , NE , N , NW , W , and SW , and the derivative of this mapping has a continuous extension to the closure of each of these sets (see Remark 8.2). If w^M is C^2 on \overline{M} , then from the first-derivative formulas in (8.20)–(8.27) we conclude that the second derivative of w is defined on each of these sets and has a continuous extension to the closure of each of these sets intersected with \mathcal{D}_2 . \blacksquare

Remark 8.5. Because, for $\vec{z} \in S$, we have $\frac{\partial \zeta_2}{\partial z_i} = f'_S(\zeta_1) \frac{\partial \zeta_1}{\partial z_i}$, $i = 1, 2$, we obtain from (8.37), (8.9), and (8.42) the formulas

$$(8.45) \quad \frac{\partial \zeta_1}{\partial z_1} = \left(\frac{1 + \lambda_2 \zeta_2}{1 + \lambda_2 z_2} \right) \cdot \frac{1 + \lambda_2 \zeta_2}{1 + \lambda_2 \zeta_2 - \lambda_2 \zeta_1 f'_S(\zeta_1)} = 1 + O(\lambda), \quad \vec{z} \in S,$$

$$(8.46) \quad \frac{\partial \zeta_1}{\partial z_2} = - \left(\frac{1 + \lambda_2 \zeta_2}{1 + \lambda_2 z_2} \right) \cdot \frac{\lambda_2 \zeta_1}{1 + \lambda_2 \zeta_2 - \lambda_2 \zeta_1 f'_S(\zeta_1)} = -\lambda_2 \zeta_1 + O(\lambda^2), \quad \vec{z} \in S.$$

Note that the denominator $1 + \lambda_2 \zeta_2 - \lambda_2 \zeta_1 f'_S(\zeta_1)$ in (8.45) and (8.46) is strictly positive by (8.5) and hence bounded below for $\vec{\zeta} \in \partial_S M$. For E , the analogous formulas are

$$(8.47) \quad \frac{\partial \zeta_2}{\partial z_2} = \left(\frac{1 - \lambda_1 \zeta_1}{1 - \lambda_1 z_1} \right) \cdot \frac{1 - \lambda_1 \zeta_1}{1 - \lambda_1 \zeta_1 + \lambda_1 \zeta_2 f'_E(\zeta_2)} = 1 + O(\lambda), \quad \vec{z} \in E,$$

$$(8.48) \quad \frac{\partial \zeta_2}{\partial z_1} = \left(\frac{1 - \lambda_1 \zeta_1}{1 - \lambda_1 z_1} \right) \cdot \frac{\lambda_1 \zeta_2}{1 - \lambda_1 \zeta_1 + \lambda_1 \zeta_2 f'_E(\zeta_2)} = \lambda_1 \zeta_2 + O(\lambda^2), \quad \vec{z} \in E.$$

We close this section with a theorem that provides sufficient conditions for a function of the type appearing in Theorem 8.4 to be a viscosity subsolution or viscosity supersolution of (6.15).

Theorem 8.6. *Suppose \mathcal{D}_2 is partitioned into finitely many disjoint open sets $\mathcal{O}_1, \dots, \mathcal{O}_n$ so that $\overline{\mathcal{D}_2} = \cup_{k=1}^n \overline{\mathcal{O}_k}$. We assume that for every $\vec{z} \in \mathcal{D}_2$, there is a line segment in the radial direction with one end point at \vec{z} that is entirely in one of the sets $\overline{\mathcal{O}_k}$. Suppose $w \in C(\overline{\mathcal{D}_2}) \cap C^1(\mathcal{D}_2)$, $w = 0$ on $\partial \mathcal{D}_2$, and w is C^2 in each $\overline{\mathcal{O}_k} \cap \mathcal{D}_2$ (Definition 8.3). If, for $k = 1, \dots, n$,*

$$(8.49) \quad \min [\mathcal{L}_2 w - \tilde{U}(pw - \nabla w \cdot \vec{z}), \mathcal{B}_1 w, \mathcal{B}_2 w, \mathcal{S}_1 w, \mathcal{S}_2 w] \leq 0 \text{ on } \overline{\mathcal{O}_k} \cap \mathcal{D}_2$$

(where (8.49) is evaluated in \mathcal{O}_k and also on $\partial\mathcal{O}_k \cap \mathcal{D}_2$, in the latter case using the continuous extension to $\overline{\mathcal{O}_k} \cap \mathcal{D}_2$ of the second derivatives of w defined in \mathcal{O}_k), then w is a viscosity subsolution of (6.15). If, for $k = 1, \dots, n$,

$$(8.50) \quad \min [\mathcal{L}_2 w - \tilde{U}(pw - \nabla w \cdot \vec{z}), \mathcal{B}_1 w, \mathcal{B}_2 w, \mathcal{S}_1 w, \mathcal{S}_2 w] \geq 0 \text{ on } \overline{\mathcal{O}_k} \cap \mathcal{D}_2,$$

then w is a viscosity supersolution of (6.15).

Proof. We assume (8.49) and prove that w is a subsolution of (6.15). The proof that (8.50) implies the supersolution property for w is analogous.

Let $\vec{z} \in \mathcal{D}_2$ be given. Let φ be a C^2 function defined on \mathcal{D}_2 that agrees with w at \vec{z} and dominates w on \mathcal{D}_2 . Then $\nabla\varphi(\vec{z}) = \nabla w(\vec{z})$. Set $\vec{z}(\alpha) = (1 + \alpha)\vec{z}$. By assumption, there is a k and $\varepsilon > 0$ such that $\overline{\mathcal{O}_k}$ contains either the line segment $\{\vec{z}(\alpha) : -\varepsilon \leq \alpha \leq 0\}$ or the line segment $\{\vec{z}(\alpha) : 0 \leq \alpha \leq \varepsilon\}$. For specificity, we consider the latter case. The function $f(\alpha) = \varphi(\vec{z}(\alpha)) - w(\vec{z}(\alpha))$ is C^2 on $[0, \varepsilon]$ and attains its minimum at $\alpha = 0$. Therefore, $0 \leq f''(0) = \nabla^2\varphi(\vec{z})\vec{z} \cdot \vec{z} - \nabla^2 w^{\mathcal{O}_k}(\vec{z})\vec{z} \cdot \vec{z}$, where $\nabla^2 w^{\mathcal{O}_k}$ denotes the second derivative of w extended by continuity from \mathcal{O}_k to $\overline{\mathcal{O}_k} \cap \mathcal{D}_2$. From the definition (6.10) of \mathcal{L}_2 , we see that (8.49) now implies the nonpositivity of (6.16) at \vec{z} . ■

9. The case $\rho = 0$. We first prove the special case of our main result, Theorem 4.1, under the assumption that $\rho = 0$. When $\rho = 0$, we do not need conditions (2.1) and (2.4). We state the result for the function u of two variables defined by (6.3).

Theorem 9.1. *Under Assumptions 2.2 and 2.3 and $\rho = 0$, the value function u for the problem with positive λ satisfies*

$$(9.1) \quad u(\vec{\theta}) = \frac{1}{p} A^{p-1} - \gamma \lambda^{2/3} + O(\lambda).$$

The constant γ in (9.1) is

$$(9.2) \quad \gamma \triangleq \frac{3}{2} A^{p-2} (\mathbf{V}\vec{\theta} \cdot \vec{\theta}) \sum_{i=1}^2 \frac{\mu_i \theta_i^2}{\nu_i},$$

where

$$(9.3) \quad \nu_i = \sqrt[3]{\frac{12\mu_i \theta_i^2}{(1-p)\sigma_i^2} (\mathbf{V}\vec{\theta} \cdot \vec{\theta})}, \quad i = 1, 2.$$

The remainder of this section is the proof of Theorem 9.1. It is divided into several steps. The idea of the proof is to construct two functions, a subsolution and a supersolution of (6.15). We construct each function by partitioning the solvency region as in Figure 1, defining the function in M , extending the function via Theorem 8.4, and verifying that the function is a subsolution or a supersolution using Theorem 8.6.

9.1. Partitioning the solvency region. We need two partitions of the solvency region, corresponding to a subsolution and a supersolution of (6.15). We create both partitions simultaneously. Let B be a positive constant and K a real constant to be chosen later. We

shall in fact choose K to be negative to get a subsolution and positive to get a supersolution. With ν_1 and ν_2 defined by (9.3) and

$$(9.4) \quad -\frac{1}{2}\nu_1\lambda^{1/3} \leq \delta_1 \leq \frac{1}{2}\nu_1\lambda^{1/3}, \quad -\frac{1}{2}\nu_2\lambda^{1/3} \leq \delta_2 \leq \frac{1}{2}\nu_2\lambda^{1/3},$$

define

$$(9.5) \quad h_i(\delta_i) \triangleq \mu_i \left[\frac{3}{2}\delta_i^2\lambda^{2/3} - \frac{1}{\nu_i^2}\delta_i^4 + \frac{3}{2}B\delta_i^2\lambda^{4/3} \right] = O(\lambda^{4/3})$$

so that

$$(9.6) \quad h_i'(\delta_i) = \mu_i \left[3\delta_i\lambda^{2/3} - \frac{4}{\nu_i^2}\delta_i^3 + 3B\delta_i\lambda^{4/3} \right] = O(\lambda),$$

$$(9.7) \quad h_i''(\delta_i) = \mu_i \left[3\lambda^{2/3} - \frac{12}{\nu_i^2}\delta_i^2 + 3B\lambda^{4/3} \right] = O(\lambda^{2/3}).$$

We further define

$$(9.8) \quad H(\delta_1, \delta_2) \triangleq \sum_{i=1}^2 \frac{1}{\nu_i} [-ph_i(\delta_i) + (\delta_i + \theta_i)h_i'(\delta_i)] = O(\lambda)$$

so that

$$(9.9) \quad \frac{\partial}{\partial \delta_i} H(\delta_1, \delta_2) = \frac{1}{\nu_i} [(1-p)h_i'(\delta_i) + (\delta_i + \theta_i)h_i''(\delta_i)] = O(\lambda^{2/3}).$$

Finally, we set

$$(9.10) \quad G_S(\delta_1, \delta_2) \triangleq \lambda_2 - p\gamma A^{1-p}\lambda_2\lambda^{2/3} + pK A^{1-p}\lambda_2\lambda + \frac{h_2'(\delta_2)}{\nu_2} + \lambda_2 H(\delta_1, \delta_2),$$

$$(9.11) \quad G_E(\delta_1, \delta_2) \triangleq \lambda_1 - p\gamma A^{1-p}\lambda_1\lambda^{2/3} + pK A^{1-p}\lambda_1\lambda - \frac{h_1'(\delta_1)}{\nu_1} + \lambda_1 H(\delta_1, \delta_2),$$

$$(9.12) \quad G_N(\delta_1, \delta_2) \triangleq \lambda_2 - p\gamma A^{1-p}\lambda_2\lambda^{2/3} + pK A^{1-p}\lambda_2\lambda - \frac{h_2'(\delta_2)}{\nu_2} + \lambda_2 H(\delta_1, \zeta_2),$$

$$(9.13) \quad G_W(\delta_1, \delta_2) \triangleq \lambda_1 - p\gamma A^{1-p}\lambda_1\lambda^{2/3} + pK A^{1-p}\lambda_1\lambda + \frac{h_1'(\delta_1)}{\nu_1} + \lambda_1 H(\delta_1, \delta_2).$$

We define the boundary functions of the middle region M appearing in (8.1)–(8.4) by

$$\begin{aligned} f_S(\zeta_1) &= F_S(\zeta_1 - \theta_1) + \theta_2, & f_E(\zeta_2) &= F_E(\zeta_2 - \theta_2) + \theta_1, \\ f_N(\zeta_1) &= F_N(\zeta_1 - \theta_1) + \theta_2, & f_W(\zeta_2) &= F_W(\zeta_2 - \theta_2) + \theta_1, \end{aligned}$$

where F_S , F_E , F_N , and F_W are defined implicitly by the formulas

$$(9.14) \quad G_S(\delta_1, F_S(\delta_1)) = 0, \quad G_E(F_E(\delta_2), \delta_2) = 0,$$

$$(9.15) \quad G_N(\delta_1, F_N(\delta_1)) = 0, \quad G_W(F_W(\delta_2), \delta_2) = 0.$$

In particular, we are using the change of variables $\delta_i = \zeta_i - \theta_i$ for $i = 1, 2$. We can make these definitions because of the following lemma.

Lemma 9.2. *For sufficiently small $\lambda > 0$, there exist C^1 functions $F_S(\delta_1)$, $F_E(\delta_2)$, $F_N(\delta_1)$, and $F_W(\delta_2)$, defined for (δ_1, δ_2) satisfying (9.4), so that (9.14) and (9.15) are satisfied. These functions are “nearly constant” in the sense that for (δ_1, δ_2) satisfying (9.4) and for any constant $\xi > \sqrt{\frac{2}{3}p\gamma A^{1-p} + B}$,*

$$\begin{aligned} -\frac{1}{2}\nu_2\lambda^{1/3} &< F_S(\delta_1) < -\frac{1}{2}\nu_2\lambda^{1/3}(1 - \xi\lambda^{1/3}), \\ \frac{1}{2}\nu_1\lambda^{1/3}(1 - \xi\lambda^{1/3}) &< F_E(\delta_2) < \frac{1}{2}\nu_1\lambda^{1/3}, \\ \frac{1}{2}\nu_2\lambda^{1/3}(1 - \xi\lambda^{1/3}) &< F_N(\delta_1) < \frac{1}{2}\nu_2\lambda^{1/3}, \\ -\frac{1}{2}\nu_1\lambda^{1/3} &< F_W(\delta_2) < -\frac{1}{2}\nu_1\lambda^{1/3}(1 - \xi\lambda^{1/3}). \end{aligned}$$

In addition, $f_S(\zeta_1)$, $f_E(\zeta_2)$, $f_N(\zeta_1)$, and $f_W(\zeta_2)$ are C^1 and satisfy (8.5)–(8.8) and

$$(9.16) \quad f'_S(\zeta_1) = O(\lambda^{1/3}), \quad f'_E(\zeta_2) = O(\lambda^{1/3}), \quad f'_N(\zeta_1) = O(\lambda^{1/3}), \quad f'_W(\zeta_2) = O(\lambda^{1/3})$$

as long as $\delta_1 = \zeta_1 - \theta_1$ and $\delta_2 = \zeta_2 - \theta_2$ satisfy (9.4).

Proof. Following Remark 8.1 we provide the proofs for F_S and F_E . Let δ_1 satisfy (9.4), let $\xi \in \mathbb{R}$ be given, and set $\delta_2^* = -\frac{1}{2}\nu_2\lambda^{1/3}(1 - \xi\lambda^{1/3})$. Then

$$\begin{aligned} G_S(\delta_1, \delta_2^*) &= \mu_2\lambda - p\gamma A^{1-p}\mu_2\lambda^{5/3} \\ &\quad + \frac{\mu_2}{\nu_2}\delta_2^* [3\lambda^{2/3} - \lambda^{2/3}(1 - 2\xi\lambda^{1/3} + \xi^2\lambda^{2/3}) + 3B\lambda^{4/3}] + O(\lambda^2) \\ &= \frac{3}{2}\mu_2 \left(\xi^2 - \frac{2}{3}p\gamma A^{1-p} - B \right) \lambda^{5/3} + O(\lambda^2), \end{aligned}$$

and the $O(\lambda^2)$ term is uniform in δ_1 . The $O(\lambda^{5/3})$ term in the last line is negative if $\xi = 0$ and is positive if ξ exceeds $\sqrt{\frac{2}{3}p\gamma A^{1-p} + B}$. Let us fix $\xi > \sqrt{\frac{2}{3}p\gamma A^{1-p} + B}$. Then for sufficiently small $\lambda > 0$ there is some $F_S(\delta_1)$ in the interval $(-\frac{1}{2}\nu_2\lambda^{1/3}, -\frac{1}{2}\nu_2\lambda^{1/3}(1 - \xi\lambda^{1/3}))$ for which $G_S(\delta_1, F_S(\delta_1)) = 0$. But for δ_1 and δ_2 satisfying (9.4),

$$(9.17) \quad \frac{\partial}{\partial \delta_2} G_S(\delta_1, \delta_2) = \frac{h_2''(\delta_2)}{\nu_2} + O(\lambda^{5/3}) \geq \frac{3B\mu_2}{\nu_2} \lambda^{4/3} + O(\lambda^{5/3}) > 0$$

for sufficiently small $\lambda > 0$, uniformly in δ_1 and δ_2 . Hence, the zero of $G_S(\delta_1, \cdot)$ is unique. Furthermore, because G_S is C^1 and (9.17) holds, the implicit function theorem implies that F_S is also C^1 . In addition, the fact that $(\partial/\partial \delta_1)G_S(\delta_1, \delta_2) = O(\lambda^{5/3})$ and relation (9.17) show that

$$F'_S(\delta_1) = -\frac{\frac{\partial}{\partial \delta_1} G_S(\delta_1, \delta_2)}{\frac{\partial}{\partial \delta_2} G_S(\delta_1, \delta_2)} \Big|_{\delta_2 = F_S(\delta_1)} = O(\lambda^{1/3}).$$

It follows that

$$(9.18) \quad f'_S(\zeta_1) = O(\lambda^{1/3}), \quad \theta_1 - \frac{1}{2}\nu_1\lambda^{1/3} \leq \zeta_1 \leq \theta_1 + \frac{1}{2}\nu_1\lambda^{1/3},$$

and thus (8.5) is satisfied for ζ_1 in the interval specified in (9.18).

Again let $\xi \in \mathbb{R}$ be given, but now define $\delta_1^* = \frac{1}{2}\nu_1\lambda^{1/3}(1 - \xi\lambda^{1/3})$ and compute $G_E(\delta_1^*, \delta_2) = \frac{3}{2}\mu_1(\xi^2 - \frac{2}{3}p\gamma A^{1-p} - B)\lambda^{5/3} + O(\lambda^2)$. We proceed as above to produce $F_E(\delta_2) \in (\frac{1}{2}\nu_1\lambda^{1/3}(1 - \xi\lambda^{1/3}), \frac{1}{2}\nu_1\lambda^{1/3})$ such that $G_E(F_E(\delta_2), \delta_2) = 0$. For δ_1 and δ_2 satisfying (9.4), $(\partial/\partial\delta_1)G_E(\delta_1, \delta_2) \leq -\frac{3B\mu_1}{\nu_1}\lambda^{4/3} + O(\lambda^{5/3}) < 0$ for sufficiently small $\lambda > 0$, uniformly in δ_1 and δ_2 . Because $(\partial/\partial\delta_2)G_E(\delta_1, \delta_2) = O(\lambda^{5/3})$, we must have $F'_E(\delta_2) = O(\lambda^{1/3})$. Hence $f'_E(\zeta_2) = O(\lambda^{1/3})$, $\theta_2 - \frac{1}{2}\nu_2\lambda^{1/3} \leq \zeta_2 \leq \theta_2 + \frac{1}{2}\nu_2\lambda^{1/3}$, which implies (8.6) for ζ_2 in the range just specified. \blacksquare

Remark 9.3. Because of the inequalities in Lemma 9.2, the graph in the (δ_1, δ_2) -plane of the continuous function $F_S(\delta_1)$, $-\frac{1}{2}\nu_1\lambda^{1/3} \leq \delta_1 \leq \frac{1}{2}\nu_1\lambda^{1/3}$, must intersect the graph in the (δ_1, δ_2) -plane of the continuous function $F_E(\delta_2)$, $-\frac{1}{2}\nu_2\lambda^{1/3} \leq \delta_2 \leq \frac{1}{2}\nu_2\lambda^{1/3}$. The intersection point must be unique because the derivatives of both F_S and F_E are of order $\lambda^{1/3}$, and it must lie in $(\frac{1}{2}\nu_1\lambda^{1/3}(1 - \xi\lambda^{1/3}), \frac{1}{2}\nu_1\lambda^{1/3}) \times (-\frac{1}{2}\nu_2\lambda^{1/3}, -\frac{1}{2}\nu_2\lambda^{1/3}(1 - \xi\lambda^{1/3}))$. We denote the intersection point by $\vec{\delta}^{SE} = (\delta_1^{SE}, \delta_2^{SE})$. We then define $\vec{\zeta}^{SE} = \vec{\delta}^{SE} + \vec{\theta}$. Similarly, we choose intersection points and define $\vec{\zeta}^{NE}$, $\vec{\zeta}^{NW}$, and $\vec{\zeta}^{SW}$. This completes the construction of the boundary functions described by (8.1)–(8.4) for the middle region M so that (8.5)–(8.8) are satisfied.

We have constructed M to be a proper subset of the rectangle

$$(9.19) \quad R(\lambda) \triangleq \left(\theta_1 - \frac{1}{2}\nu_1\lambda^{1/3}, \theta_1 + \frac{1}{2}\nu_1\lambda^{1/3} \right) \times \left(\theta_2 - \frac{1}{2}\nu_2\lambda^{1/3}, \theta_2 + \frac{1}{2}\nu_2\lambda^{1/3} \right).$$

We fix $\xi > \sqrt{\frac{2}{3}p\gamma A^{1-p} + B}$ and define four “perimeter sets” of $R(\lambda)$ by

$$\begin{aligned} P_S &\triangleq \left\{ \vec{z} \in R : \theta_2 - \frac{1}{2}\nu_2\lambda^{1/3} < z_2 < \theta_2 - \frac{1}{2}\nu_2\lambda^{1/3}(1 - \xi\lambda^{1/3}) \right\}, \\ P_E &\triangleq \left\{ \vec{z} \in R : \theta_1 + \frac{1}{2}\nu_1\lambda^{1/3}(1 - \xi\lambda^{1/3}) < z_1 < \theta_1 + \frac{1}{2}\nu_1\lambda^{1/3} \right\}, \\ P_N &\triangleq \left\{ \vec{z} \in R : \theta_2 + \frac{1}{2}\nu_2\lambda^{1/3}(1 - \xi\lambda^{1/3}) < z_2 < \theta_2 + \frac{1}{2}\nu_2\lambda^{1/3} \right\}, \\ P_W &\triangleq \left\{ \vec{z} \in R : \theta_1 - \frac{1}{2}\nu_1\lambda^{1/3} < z_1 < \theta_1 - \frac{1}{2}\nu_1\lambda^{1/3}(1 - \xi\lambda^{1/3}) \right\}. \end{aligned}$$

We have shown that for sufficiently small $\lambda > 0$,

$$(9.20) \quad (\zeta_1, f_S(\zeta_1)) \in P_S, \quad f'_S(\zeta_1) = O(\lambda^{1/3}), \quad \zeta_1^{SW} \leq \zeta_1 \leq \zeta_1^{SE},$$

$$(9.21) \quad (f_E(\zeta_2), \zeta_2) \in P_E, \quad f'_E(\zeta_2) = O(\lambda^{1/3}), \quad \zeta_2^{SE} \leq \zeta_2 \leq \zeta_2^{NE},$$

$$(9.22) \quad (\zeta_1, f_N(\zeta_1)) \in P_N, \quad f'_N(\zeta_1) = O(\lambda^{1/3}), \quad \zeta_1^{NW} \leq \zeta_1 \leq \zeta_1^{NE},$$

$$(9.23) \quad (f_W(\zeta_2), \zeta_2) \in P_W, \quad f'_W(\zeta_2) = O(\lambda^{1/3}), \quad \zeta_2^{SW} \leq \zeta_2 \leq \zeta_2^{NW}.$$

In other words, the boundaries of M are almost straight lines, and the region M almost coincides with the rectangle $R(\lambda)$.

The boundaries of M depend on the constant K appearing in the definitions of G_S , G_E , G_N , and G_W . However, the perimeter sets P_S , P_E , P_N , and P_W do not, and the $O(\lambda^{1/3})$ terms appearing in (9.20)–(9.23) depend only on bounds that are uniform over these perimeter sets, and hence they do not depend on K .

Finally, we define the eight nonmiddle regions of \mathcal{D}_2 , as shown in Figure 1.

9.2. Construction of w^M on \mathcal{D}_2 . We define a C^2 function w^M on the rectangle $R(\lambda)$ defined by (9.19) whose restriction to M will play the role of w^M in Theorem 8.4 by the formula

$$(9.24) \quad w^M(\vec{z}) \triangleq \frac{1}{p}A^{p-1} - \gamma\lambda^{2/3} + K\lambda - \sum_{i=1}^2 \frac{1}{\nu_i}A^{p-1}h_i(z_i - \theta_i), \quad \vec{z} \in R(\lambda),$$

where the constant K is chosen to satisfy conditions specified in the proofs of Lemmas 9.8 and 9.9 below. We may restrict w^M to \overline{M} and then extend this restricted function $w^M|_{\overline{M}}$ to $\overline{\mathcal{D}_2}$, as described in Theorem 8.4. We call the extended function w . Note that $w = w^M$ on \overline{M} , but w does not agree with w^M on $R(\lambda) \setminus \overline{M}$.

Lemma 9.4. *The function $w^M|_{\overline{M}}$ satisfies (8.28) and (8.31). Consequently, w is C^1 on \mathcal{D}_2 , its first derivatives are given by (8.20)–(8.27), and w satisfies (8.29), (8.30), and (8.32). For sufficiently small $\lambda > 0$, the function w also satisfies*

$$(9.25) \quad \min [\mathcal{B}_1 w, \mathcal{B}_2 w, \mathcal{S}_1 w, \mathcal{S}_2 w] \geq 0 \text{ on } \overline{M}.$$

Proof. Using the notation $\vec{z} = \vec{\delta} + \vec{\theta}$, it is straightforward to verify that for $\vec{z} \in R$,

$$(9.26) \quad \begin{aligned} \mathcal{B}_2 w^M(\vec{z}) &= A^{p-1}G_S(\vec{\delta}), & \mathcal{S}_1 w^M(\vec{z}) &= A^{p-1}G_E(\vec{\delta}), \\ \mathcal{S}_2 w^M(\vec{z}) &= A^{p-1}G_N(\vec{\delta}), & \mathcal{B}_1 w^M(\vec{z}) &= A^{p-1}G_W(\vec{\delta}). \end{aligned}$$

Equations (9.14) and (9.15) imply (8.28). Once we have proved (9.25), then (8.28) will imply (8.31).

It remains to prove (9.25). We show that $\mathcal{B}_2 w \geq 0$ and $\mathcal{S}_1 w \geq 0$ on M ; the proofs of the other two inequalities in (9.25) are analogous. In fact, we prove the stronger result that

$$\begin{aligned} \mathcal{B}_2 w^M(\vec{z}) &\geq 0, & \vec{z} \in R(\lambda) \text{ such that } z_2 &\geq f_S(z_1), \\ \mathcal{S}_1 w^M(\vec{z}) &\geq 0, & \vec{z} \in R(\lambda) \text{ such that } z_1 &\leq f_E(z_2). \end{aligned}$$

Since $\mathcal{B}_2 w^M(\vec{z}) = 0$ when $\vec{z} \in R(\lambda)$ and $z_2 = f_S(z_1)$, and $\mathcal{S}_1 w^M(\vec{z}) = 0$ when $\vec{z} \in R(\lambda)$ and $z_1 = f_E(z_2)$, for this it will suffice to show that

$$\frac{\partial}{\partial z_2} \mathcal{B}_2 w^M(\vec{z}) \geq 0, \quad \frac{\partial}{\partial z_1} \mathcal{S}_1 w^M(\vec{z}) \leq 0, \quad \vec{z} \in R(\lambda).$$

Using the first equality in (9.26), we have from (9.17) that

$$\frac{\partial}{\partial z_2} \mathcal{B}_2 w^M(\vec{z}) = A^{p-1} \frac{\partial}{\partial \delta_2} G_S(\vec{\delta}) \geq \frac{3B\mu_2}{\nu_2} A^{p-1} \lambda^{4/3} + O(\lambda^{5/3}),$$

which is strictly positive for sufficiently small $\lambda > 0$, uniformly in $\vec{z} \in R$. Similarly,

$$\frac{\partial}{\partial z_1} \mathcal{S}_1 w^M(\vec{z}) = A^{p-1} \frac{\partial}{\partial \delta_1} G_E(\vec{\delta}) \leq -\frac{3B\mu_1}{\nu_1} A^{p-1} \lambda^{4/3} + O(\lambda^{5/3}). \quad \blacksquare$$

Remark 9.5 (derivative estimates of w on ∂M). It is straightforward to compute the first derivatives of w^M on the perimeter sets P_S , P_E , P_N , and P_W . Using (9.24) we obtain

$$(9.27) \quad w_2^M = A^{p-1} \lambda_2 + O(\lambda^{4/3}), \quad w_1^M = O(\lambda) \text{ on } P_S,$$

$$(9.28) \quad w_1^M = -A^{p-1} \lambda_1 + O(\lambda^{4/3}), \quad w_2^M = O(\lambda) \text{ on } P_E,$$

$$w_2^M = -A^{p-1} \lambda_2 + O(\lambda^{4/3}), \quad w_1^M = O(\lambda) \text{ on } P_N,$$

$$w_1^M = A^{p-1} \lambda_1 + O(\lambda^{4/3}), \quad w_2^M = O(\lambda) \text{ on } P_W.$$

The $O(\lambda)$ and $O(\lambda^{4/3})$ terms in these formulas do not depend on K .

The second partial derivatives of w may not exist on ∂M , but we can compute the second partials of w^M inside the perimeter sets, and on ∂M these coincide with the second partials of w on computed from inside M . Direct computation reveals

$$(9.29) \quad w_{2,2}^M = O(\lambda), \quad w_{1,1}^M = O(\lambda^{2/3}) \text{ on } P_S \cup P_N,$$

$$(9.30) \quad w_{1,1}^M = O(\lambda), \quad w_{2,2}^M = O(\lambda^{2/3}) \text{ on } P_E \cup P_N,$$

$$(9.31) \quad w_{1,2}^M = 0 \text{ on } R(\lambda).$$

In particular, for $i, j \in \{1, 2\}$,

$$(9.32) \quad w_{i,j}^M(\vec{\zeta}^{SE}) = O(\lambda), \quad w_{i,j}^M(\vec{\zeta}^{NE}) = O(\lambda), \quad w_{i,j}^M(\vec{\zeta}^{NW}) = O(\lambda), \quad w_{i,j}^M(\vec{\zeta}^{SW}) = O(\lambda).$$

The $O(\lambda)$ and $O(\lambda^{2/3})$ terms in (9.29)–(9.32) do not depend on K .

We derive additional estimates on the pure second partials of w^M on the boundaries of M . The first equation in (8.28), $\mathcal{B}_2 w = 0$, evaluated along the southern boundary of M is

$$\lambda_2 p w(\zeta_1, f_S(\zeta_1)) - \lambda_2 \zeta_1 w_1(\zeta_1, f_S(\zeta_1)) - (1 + \lambda_2 f_S(\zeta_1)) w_2(\zeta_1, f_S(\zeta_1)) = 0$$

for $\zeta_1^{SW} \leq \zeta_1 \leq \zeta_1^{SE}$. Differentiation yields (recall (9.31))

$$(9.33) \quad \begin{aligned} & \lambda_2(p-1)w_1(\zeta_1, f_S(\zeta_1)) + \lambda_2(p-1)w_2(\zeta_1, f_S(\zeta_1))f'_S(\zeta_1) - \lambda_2 \zeta_1 w_{1,1}^M(\zeta_1, f_S(\zeta_1)) \\ & = (1 + \lambda_2 f_S(\zeta_1))w_{2,2}^M(\zeta_1, f_S(\zeta_1))f'_S(\zeta_1). \end{aligned}$$

From (9.27) and the second equation in (9.29), we see that the left-hand side of (9.33) is $O(\lambda^{5/3})$. In conclusion,

$$(9.34) \quad w_{2,2}^M(\zeta_1, f_S(\zeta_1))f'_S(\zeta_1) = O(\lambda^{5/3}), \quad \zeta_1^{SW} \leq \zeta_1 \leq \zeta_1^{SE}.$$

The second equation in (8.28), $\mathcal{S}_1 w = 0$, evaluated along the eastern boundary of M is

$$\lambda_1 p w(f_E(\zeta_2), \zeta_2) + (1 - \lambda_1 f_E(\zeta_2))w_1(f_E(\zeta_2), \zeta_2) - \lambda_1 \zeta_2 w_2(f_E(\zeta_2), \zeta_2) = 0$$

for $\zeta_2^{SE} \leq \zeta_2 \leq \zeta_2^{NE}$. Differentiation yields

$$(9.35) \quad w_{1,1}^M(f_E(\zeta_2), \zeta_2) f'_E(\zeta_2) = O(\lambda^{5/3}), \quad \zeta_2^{SE} \leq \zeta_2 \leq \zeta_2^{NE}.$$

The analogous equations on the other two boundaries of M are (see Remark 8.1)

$$(9.36) \quad w_{2,2}^M(\zeta_1, f_N(\zeta_1)) f'_N(\zeta_1) = O(\lambda^{5/3}), \quad \zeta_1^{NW} \leq \zeta_1 \leq \zeta_1^{NE},$$

$$(9.37) \quad w_{1,1}^M(f_W(\zeta_2), \zeta_2) f'_W(\zeta_2) = O(\lambda^{5/3}), \quad \zeta_2^{SW} \leq \zeta_2 \leq \zeta_2^{NW}.$$

The $O(\lambda^{5/3})$ terms appearing in (9.34)–(9.37) are independent of K . \blacksquare

9.3. The second partial derivatives of w outside M . We compute second partial derivatives of w in the regions S , E , and SE . In $\bar{S} \cap \mathcal{D}_2$, the first partial derivatives w_i are given by (8.20) with \vec{z} and $\vec{\zeta}$ related by (8.9). From this formula, using (9.31), we compute

$$\begin{aligned} w_{1,1}^S(\vec{z}) &= \frac{\partial}{\partial z_1} \left[\left(\frac{1 + \lambda_2 z_2}{1 + \lambda_2 \zeta_2} \right)^{p-1} w_1(\vec{\zeta}) \right] \\ &= \left(\frac{1 + \lambda_2 z_2}{1 + \lambda_2 \zeta_2} \right)^{p-1} \left[\frac{\lambda_2(1-p)}{1 + \lambda_2 \zeta_2} f'_S(\zeta_1) w_1(\vec{\zeta}) + w_{1,1}^M(\vec{\zeta}) \right] \frac{\partial \zeta_1}{\partial z_1}, \quad \vec{z} \in \bar{S} \cap \mathcal{D}_2. \end{aligned}$$

In light of the second equality in (9.20), (9.27), (9.29), and (8.45), we see that

$$(9.38) \quad w_{1,1}^S(\vec{\zeta}) = w_{1,1}^M(\vec{\zeta}) + O(\lambda^{5/3}), \quad \vec{\zeta} \in \partial_S M.$$

Recalling (9.31), we next compute

$$\begin{aligned} w_{1,2}^S(\vec{z}) &= \frac{\partial}{\partial z_2} \left[\left(\frac{z_1}{\zeta_1} \right)^{p-1} w_1(\vec{\zeta}) \right] \\ &= \left(\frac{z_1}{\zeta_1} \right)^{p-1} \left[\frac{1-p}{\zeta_1} w_1(\vec{\zeta}) + w_{1,1}^M(\vec{\zeta}) \right] \frac{\partial \zeta_1}{\partial z_2}, \quad \vec{z} \in \bar{S} \cap \mathcal{D}_2, \end{aligned}$$

and (9.27), (9.29), and (8.46) imply

$$(9.39) \quad w_{1,2}^S(\vec{\zeta}) = O(\lambda^{5/3}), \quad \vec{\zeta} \in \partial_S M.$$

Finally,

$$\begin{aligned} w_{2,2}^S(\vec{z}) &= \frac{\partial}{\partial z_2} \left[\left(\frac{z_1}{\zeta_1} \right)^{p-1} w_2(\vec{\zeta}) \right] \\ (9.40) \quad &= \left(\frac{z_1}{\zeta_1} \right)^{p-1} \left[\frac{1-p}{\zeta_1} w_2(\vec{\zeta}) + w_{2,2}^M(\vec{\zeta}) f'_S(\zeta_1) \right] \frac{\partial \zeta_1}{\partial z_2}, \quad \vec{z} \in \bar{S} \cap \mathcal{D}_2, \end{aligned}$$

and (9.27), (9.29), the second equality in (9.20), and (8.46) imply

$$(9.41) \quad w_{2,2}^S(\vec{\zeta}) = -(1-p)A^{p-1}\lambda_2^2 + O(\lambda^{7/3}), \quad \vec{\zeta} \in \partial_S M.$$

In $\overline{E} \cap \mathcal{D}_2$, the first partial derivatives w_i are given by (8.21) with \vec{z} and $\vec{\zeta}$ related by (8.10). Therefore

$$(9.42) \quad w_{1,1}^E(\vec{z}) = \left(\frac{z_2}{\zeta_2} \right)^{p-1} \left[\frac{1-p}{\zeta_2} w_1(\vec{\zeta}) + w_{1,1}^M(\vec{\zeta}) f'_E(\zeta_2) \right] \frac{\partial \zeta_2}{\partial z_1}, \quad \vec{z} \in \overline{E} \cap \mathcal{D}_2.$$

Using (9.28), (9.30), (9.21), and (8.48) we obtain

$$(9.43) \quad w_{1,1}^E(\vec{\zeta}) = -(1-p)A^{p-1}\lambda_1^2 + O(\lambda^{7/3}), \quad \zeta \in \partial_E M.$$

We also have

$$(9.44) \quad w_{1,2}^E(\vec{z}) = \left(\frac{z_2}{\zeta_1} \right)^{p-1} \left[\frac{1-p}{\zeta_2} w_2(\vec{\zeta}) + w_{2,2}^M(\vec{\zeta}) \right] \frac{\partial \zeta_2}{\partial z_1}, \quad \vec{z} \in \overline{E} \cap \mathcal{D}_2.$$

Using (9.28), (9.30), and (8.48) we obtain

$$(9.45) \quad w_{1,2}^E(\vec{\zeta}) = O(\lambda^{5/3}), \quad \vec{\zeta} \in \partial_E M.$$

Finally,

$$(9.46) \quad w_{2,2}^E(\vec{z}) = \left(\frac{1 - \lambda_1 z_1}{1 - \lambda_1 \zeta_1} \right)^{p-1} \left[\frac{\lambda_1(p-1)}{1 - \lambda_1 \zeta_1} f'_E(\zeta_2) w_2(\vec{\zeta}) + w_{2,2}^M(\vec{\zeta}) \right] \frac{\partial \zeta_2}{\partial z_2}, \quad \vec{z} \in \overline{E} \cap \mathcal{D}_2.$$

Using (9.28), (9.21), (9.30), and (8.47) we obtain

$$(9.47) \quad w_{2,2}^E(\vec{\zeta}) = w_{2,2}^M(\vec{\zeta}) + O(\lambda^{5/3}), \quad \vec{\zeta} \in \partial_E M.$$

We may take \vec{z} in (9.42), (9.44), and (9.46) to be in $\overline{SE} \cap \overline{E} \cap \mathcal{D}_2$ so that ζ in these formulas becomes $\vec{\zeta}^{SE}$. We see then that for $i, j \in \{1, 2\}$,

$$(9.48) \quad w_{i,j}^E(\vec{z}) = O(\lambda), \quad \vec{z} \in \overline{SE} \cap \overline{E} \cap \mathcal{D}_2,$$

where the $O(\lambda)$ term is uniform in \vec{z} as long as \vec{z} is bounded away from $\partial \mathcal{D}_2$; see Remark 9.10 below in this regard. An analogous proof shows that $w_{i,j}^S = O(\lambda)$ in compact subsets of $\overline{SE} \cap \overline{S} \cap \mathcal{D}_2$.

In \overline{SE} we begin with the formula (8.24) and compute for $\vec{z} \in \overline{SE} \cap \mathcal{D}_2$ that

$$(9.49) \quad w_{1,1}^{SE}(\vec{z}) = \frac{(1-p)\lambda_1}{1 - \lambda_1 \zeta_1^{SE} + \lambda_2 \zeta_2^{SE}} \left(\frac{1 - \lambda_1 z_1 + \lambda_2 z_2}{1 - \lambda_1 \zeta_1^{SE} + \lambda_2 \zeta_2^{SE}} \right)^{p-2} w_1(\vec{\zeta}^{SE}),$$

$$(9.50) \quad w_{1,2}^{SE}(\vec{z}) = -\frac{(1-p)\lambda_2}{1 - \lambda_1 \zeta_1^{SE} + \lambda_2 \zeta_2^{SE}} \left(\frac{1 - \lambda_1 z_1 + \lambda_2 z_2}{1 - \lambda_1 \zeta_1^{SE} + \lambda_2 \zeta_2^{SE}} \right)^{p-2} w_1(\vec{\zeta}^{SE}),$$

$$(9.51) \quad w_{2,2}^{SE}(\vec{z}) = -\frac{(1-p)\lambda_2}{1 - \lambda_1 \zeta_1^{SE} + \lambda_2 \zeta_2^{SE}} \left(\frac{1 - \lambda_1 z_1 + \lambda_2 z_2}{1 - \lambda_1 \zeta_1^{SE} + \lambda_2 \zeta_2^{SE}} \right)^{p-2} w_2(\vec{\zeta}^{SE}).$$

We then use (9.28) twice and (9.27) to conclude that

$$(9.52) \quad w_{1,1}^{SE}(\vec{\zeta}^{SE}) = -\frac{1-p}{1-\lambda_1\zeta_1^{SE} + \lambda_2\zeta_2^{SE}} A^{p-1}\lambda_1^2 + O(\lambda^{7/3}),$$

$$(9.53) \quad w_{1,2}^{SE}(\vec{\zeta}^{SE}) = \frac{1-p}{1-\lambda_1\zeta_1^{SE} + \lambda_2\zeta_2^{SE}} A^{p-1}\lambda_1\lambda_2 + O(\lambda^{7/3}),$$

$$(9.54) \quad w_{2,2}^{SE}(\vec{\zeta}^{SE}) = -\frac{1-p}{1-\lambda_1\zeta_1^{SE} + \lambda_2\zeta_2^{SE}} A^{p-1}\lambda_2^2 + O(\lambda^{7/3}).$$

The $O(\lambda)$, $O(\lambda^{5/3})$, and $O(\lambda^{7/3})$ in (9.38), (9.39), (9.41), (9.43), (9.45), (9.47), (9.48), and (9.52)–(9.54) do not depend on K .

Remark 9.6. Returning to (9.40), we use (8.46) and (8.9) to write

$$w_{2,2}^S(\vec{z}) = -\left(\frac{z_1}{\zeta_1}\right)^{p-2} \left[\frac{1-p}{\zeta_1} w_2(\vec{\zeta}) + w_{2,2}^M(\vec{\zeta}) f'_S(\zeta_1) \right] \cdot \frac{\lambda_2\zeta_1}{1 + \lambda_2\zeta_2 - \lambda_2\zeta_1 f'_S(\zeta_1)}$$

for $\vec{z} \in \bar{S} \cap \mathcal{D}_2$. Replacing \vec{z} by $\vec{\zeta} \in \partial_S M$, we obtain

$$w_{2,2}^S(\vec{\zeta}) = -\left[\frac{1-p}{\zeta_1} w_2(\vec{\zeta}) + w_{2,2}^M(\vec{\zeta}) f'_S(\zeta_1) \right] \cdot \frac{\lambda_2\zeta_1}{1 + \lambda_2\zeta_2 - \lambda_2\zeta_1 f'_S(\zeta_1)},$$

and substituting this back into the original equation, we conclude that

$$(9.55) \quad w_{2,2}^S(\vec{z}) = \left(\frac{z_1}{\zeta_1}\right)^{p-2} w_{2,2}^S(\vec{\zeta}), \quad \vec{z} \in \bar{S} \cap \mathcal{D}_2.$$

Arguing in the same way from (9.42), (8.48), and (8.10), we conclude that

$$(9.56) \quad w_{1,1}^E(\vec{z}) = \left(\frac{z_2}{\zeta_2}\right)^{p-2} w_{1,1}^E(\vec{\zeta}), \quad \vec{z} \in \bar{E} \cap \mathcal{D}_2.$$

9.4. $\mathcal{L}_2 w - \tilde{U}(pw - \nabla w \cdot \vec{z})$ in \bar{M} .

Lemma 9.7. Fix $a > 0$. For $0 \leq b < a$, the function \tilde{U} defined by (5.2) satisfies

$$\tilde{U}(a-b) = \frac{1-p}{p} a^{p/(p-1)} + ba^{1/(p-1)} + O(b^2).$$

Proof. A Taylor series expansion of $f(x) = x^{1/(p-1)}$ around $x = a$ yields

$$(a-b)^{1/(p-1)} = a^{1/(p-1)} + \frac{b}{1-p} a^{(2-p)/(p-1)} + O(b^2).$$

This implies that

$$\begin{aligned} \tilde{U}(a-b) &= \frac{1-p}{p} (a-b) \cdot (a-b)^{1/(p-1)} \\ &= \frac{1-p}{p} (a-b) \left[a^{1/(p-1)} + \frac{b}{1-p} a^{(2-p)/(p-1)} + O(b^2) \right], \end{aligned}$$

and the result follows by simplification of this last expression. \blacksquare

Lemma 9.8. *The function w defined in section 9.2 is C^2 on the closed set \overline{M} and for sufficiently large K satisfies*

$$(9.57) \quad \mathcal{L}_2 w^M - \tilde{U}(pw - \nabla w \cdot \vec{z}) \geq 0$$

on this set. For sufficiently small (i.e., negative) K , the reverse inequality holds.

Proof. On the set $\overline{M} \subset R$, the function w agrees with the function w^M defined by (9.24), and the second derivative limits of w on ∂M in (9.57) are understood to be those from inside M . From (9.5)–(9.7), for $\vec{z} \in \overline{M}$ we have

$$(9.58) \quad w(\vec{z}) = \frac{1}{p}A^{p-1} - \gamma\lambda^{2/3} + K\lambda + O(\lambda^{4/3}),$$

$$(9.59) \quad \nabla w(\vec{z}) \cdot \vec{z} = O(\lambda), \quad \nabla^2 w^M(\vec{z})\vec{z} \cdot \vec{z} = O(\lambda^{2/3}),$$

where the coefficients in the $O(\lambda^{4/3})$, $O(\lambda)$, and $O(\lambda^{2/3})$ terms in (9.58)–(9.59) depend on the model parameters but not on the constant K appearing in (9.58). We write $pw(\vec{z}) - \nabla w(\vec{z}) \cdot \vec{z} = a - b$ by setting $a = A^{p-1}$ and $b = p\gamma\lambda^{2/3} - pK\lambda + O(\lambda)$ and use Lemma 9.7 and (9.58) to obtain

$$(9.60) \quad \begin{aligned} \tilde{U}(pw(\vec{z}) - \nabla w(\vec{z}) \cdot \vec{z}) &= \frac{1-p}{p}A^p + A[p\gamma\lambda^{2/3} - pK\lambda + O(\lambda)] \\ &= (1-p)Aw(\vec{z}) + \gamma A\lambda^{2/3} - KA\lambda + O(\lambda). \end{aligned}$$

Together with these equations, (6.10) and (2.12) imply

$$(9.61) \quad \begin{aligned} &\mathcal{L}_2 w^M(\vec{z}) - \tilde{U}(pw(\vec{z}) - \nabla w(\vec{z}) \cdot \vec{z}) \\ &= \frac{1}{2}p(1-p)(\mathbf{V}(\vec{z} - \vec{\theta}) \cdot (\vec{z} - \vec{\theta}))w(\vec{z}) - \frac{1}{2}(\mathbf{V}\vec{z} \cdot \vec{z})(\nabla^2 w^M(\vec{z})\vec{z} \cdot \vec{z}) \\ &\quad - \gamma A\lambda^{2/3} + KA\lambda + O(\lambda), \quad \vec{z} \in \overline{M}, \end{aligned}$$

where the $O(\lambda)$ term is independent of K . For $\vec{z} \in \overline{M}$, $\mathbf{V}(\vec{z} - \vec{\theta}) \cdot (\vec{z} - \vec{\theta}) = O(\lambda^{2/3})$, and hence

$$(9.62) \quad \begin{aligned} \frac{1}{2}p(1-p)(\mathbf{V}(\vec{z} - \vec{\theta}) \cdot (\vec{z} - \vec{\theta}))w(\vec{z}) &= \frac{1}{2}(1-p)A^{p-1}\mathbf{V}(\vec{z} - \vec{\theta}) \cdot (\vec{z} - \vec{\theta}) + O(\lambda^{4/3}) \\ &= \frac{1}{2}(1-p)A^{p-1} \sum_{i=1}^2 \sigma_i^2 (z_i - \theta_i)^2 + O(\lambda^{4/3}). \end{aligned}$$

Because $\vec{z} = \vec{\theta} + O(\lambda^{1/3})$, the second equation in (9.59) implies

$$\begin{aligned}
& -\frac{1}{2}(\mathbf{V}\vec{z} \cdot \vec{z})(\nabla^2 w^M(\vec{z})\vec{z} \cdot \vec{z}) \\
&= -\frac{1}{2}(\mathbf{V}\vec{\theta} \cdot \vec{\theta})(\nabla^2 w^M(\vec{z})\vec{z} \cdot \vec{z}) + O(\lambda) \\
&= \frac{1}{2}(\mathbf{V}\vec{\theta} \cdot \vec{\theta}) \sum_{i=1}^2 \frac{1}{\nu_i} A^{p-1} z_i^2 h_i''(z_i - \theta_i) + O(\lambda) \\
&= \frac{1}{2}(\mathbf{V}\vec{\theta} \cdot \vec{\theta}) \sum_{i=1}^2 \frac{1}{\nu_i} A^{p-1} \theta_i^2 h_i''(z_i - \theta_i) + O(\lambda) \\
&= \frac{1}{2} A^{p-1} (\mathbf{V}\vec{\theta} \cdot \vec{\theta}) \sum_{i=1}^2 \frac{\mu_i \theta_i^2}{\nu_i} \left[3\lambda^{2/3} - \frac{12}{\nu_i^2} (z_i - \theta_i)^2 \right] + O(\lambda) \\
&= \frac{3}{2} A^{p-1} (\mathbf{V}\vec{\theta} \cdot \vec{\theta}) \sum_{i=1}^2 \frac{\mu_i \theta_i^2}{\nu_i} \lambda^{2/3} - 6 A^{p-1} (\mathbf{V}\vec{\theta} \cdot \vec{\theta}) \sum_{i=1}^2 \frac{\mu_i \theta_i^2}{\nu_i^3} (z_i - \theta_i)^2 + O(\lambda) \\
(9.63) \quad &= \gamma A \lambda^{2/3} - \frac{1}{2} (1-p) A^{p-1} \sum_{i=1}^2 \sigma_i^2 (z_i - \theta_i)^2 + O(\lambda),
\end{aligned}$$

where we have used (9.2) and (9.3) in the last step. Substituting (9.62) and (9.63) into (9.61), we obtain

$$(9.64) \quad \mathcal{L}_2 w^M(\vec{z}) - \tilde{U}(pw(\vec{z}) - \nabla w(\vec{z}) \cdot \vec{z}) = KA\lambda + O(\lambda),$$

where the $O(\lambda)$ term is independent of K . By choosing $K > 0$ sufficiently large, we can ensure that $KA\lambda + O(\lambda) \geq 0$. By choosing K sufficiently small (i.e., negative), we can ensure the reverse inequality. ■

9.5. $\mathcal{L}_2 w - \tilde{U}(pw - \nabla w \cdot \vec{z})$ in $\mathcal{D}_2 \setminus M$.

Lemma 9.9. *The function w defined in section 9.2 is C^2 on each of the eight nonmiddle sets $\overline{S} \cap \mathcal{D}_2, \overline{SE} \cap \mathcal{D}_2, \dots, \overline{SW} \cap \mathcal{D}_2$ and for sufficiently large K satisfies*

$$(9.65) \quad \mathcal{L}_2 w - \tilde{U}(pw - \nabla w \cdot \vec{z}) \geq 0$$

on each of these sets.

Proof. By Lemma 9.8, w is C^2 on \overline{M} , and by Lemma 9.4, it also satisfies (8.28). Theorem 8.4 implies that w is C^2 on each of the eight nonmiddle sets.

Following Remark 8.1, we prove (9.65) only on $\overline{S} \cap \mathcal{D}_2, \overline{E} \cap \mathcal{D}_2$, and $\overline{SE} \cap \mathcal{D}_2$. We first prove (9.65) on $\partial_S M$. According to Lemma 9.8, for sufficiently large K ,

$$(9.66) \quad \mathcal{L}_2 w^M(\vec{\zeta}) - \tilde{U}(pw(\vec{\zeta}) - \nabla w(\vec{\zeta}) \cdot \vec{\zeta}) \geq 0, \quad \vec{\zeta} \in \partial M.$$

Because w is C^1 , the only difference between this inequality and the inequality we wish to prove for $\vec{\zeta} \in \partial_S M$,

$$(9.67) \quad \mathcal{L}_2 w^S(\vec{\zeta}) - \tilde{U}(pw(\vec{\zeta}) - \nabla w(\vec{\zeta}) \cdot \vec{\zeta}) \geq 0,$$

is in the second-order term in the operator \mathcal{L}_2 . In (9.66) this term is

$$-\frac{1}{2}(\mathbf{V}\vec{\zeta} \cdot \vec{\zeta})(\nabla^2 w^M(\vec{\zeta})\vec{\zeta} \cdot \vec{\zeta}) = -\frac{1}{2}(\mathbf{V}\vec{\zeta} \cdot \vec{\zeta})w_{1,1}^M(\vec{\zeta})\zeta_1^2 + O(\lambda),$$

where we have used (9.29) and (9.31). As observed in Remark 9.5, the $O(\lambda)$ term appearing in this equation is independent of K . In (9.67) this term is

$$-\frac{1}{2}(\mathbf{V}\vec{\zeta} \cdot \vec{\zeta})(\nabla^2 w^S(\vec{\zeta})\vec{\zeta} \cdot \vec{\zeta}) = -\frac{1}{2}(\mathbf{V}\vec{\zeta} \cdot \vec{\zeta})w_{1,1}^M(\vec{\zeta})\zeta_1^2 + O(\lambda^{5/3}),$$

where we have used (9.38), (9.39), and (9.41). These terms differ by $O(\lambda)$, so (9.64) implies

$$\begin{aligned} & \mathcal{L}_2 w^S(\vec{\zeta}) - \tilde{U}(pw(\vec{\zeta}) - \nabla w(\vec{\zeta}) \cdot \vec{\zeta}) \\ &= \mathcal{L}_2 w^M(\vec{\zeta}) - \tilde{U}(pw(\vec{\zeta}) - \nabla w(\vec{\zeta}) \cdot \vec{\zeta}) + O(\lambda) \\ (9.68) \quad &= KA\lambda + O(\lambda), \quad \zeta \in \partial_S M, \end{aligned}$$

where the $O(\lambda)$ term is independent of K . Increasing K if necessary, we can ensure that (9.67) holds on $\partial_S M$.

We extend the inequality (9.67) from $\partial_S M$ to the rest of $\bar{S} \cap \mathcal{D}_2$. We begin with the inequality $\mathcal{B}_2 w = 0$, which holds in $\bar{S} \cap \mathcal{D}_2$ (Lemma 9.4). We first write this equality, compute the partial derivative with respect to z_1 , and then compute the partial derivative with respect to z_2 to obtain the three equations

$$(9.69) \quad \lambda_2 pw(\vec{z}) - w_2(\vec{z}) - \lambda_2 \nabla w(\vec{z}) \cdot \vec{z} = 0,$$

$$(9.70) \quad \lambda_2(p-1)w_1(\vec{z}) - w_{1,2}^S(\vec{z}) - \lambda_2 \nabla w_1^S(\vec{z}) \cdot \vec{z} = 0,$$

$$(9.71) \quad \lambda_2(p-1)w_2(\vec{z}) - w_{2,2}^S(\vec{z}) - \lambda_2 \nabla w_2^S(\vec{z}) \cdot \vec{z} = 0.$$

Multiplying (9.69) by $1-p$, multiplying (9.70) by z_1 , multiplying (9.71) by $z_2 - (1/\lambda_2)$, summing the three resulting equations, and then dividing by λ_2 , we obtain

$$(9.72) \quad -p(1-p)w(\vec{z}) + 2(1-p)\nabla w(\vec{z}) \cdot \vec{z} + \nabla^2 w^S(\vec{z})\vec{z} \cdot \vec{z} = \frac{1}{\lambda_2^2}w_{2,2}^S(\vec{z}), \quad \vec{z} \in \bar{S} \cap \mathcal{D}_2.$$

Equations (9.69) and (9.72) permit us to write

$$\begin{aligned} & \mathcal{L}_2 w^S(\vec{z}) - \tilde{U}(pw(\vec{z}) - \nabla w(\vec{z}) \cdot \vec{z}) \\ &= \beta w(\vec{z}) - (r + \vec{\alpha} \cdot \vec{z})(pw(\vec{z}) - \nabla w(\vec{z}) \cdot \vec{z}) \\ & \quad - \frac{1}{2}(\mathbf{V}\vec{z} \cdot \vec{z})(-p(1-p)w(\vec{z}) + 2(1-p)\nabla w(\vec{z}) \cdot \vec{z} + \nabla^2 w^S(\vec{z})\vec{z} \cdot \vec{z}) \\ & \quad - \tilde{U}(pw(\vec{z}) - \nabla w(\vec{z}) \cdot \vec{z}) \\ (9.73) \quad &= \beta w(\vec{z}) - (r + \vec{\alpha} \cdot \vec{z})\frac{1}{\lambda_2}w_2(\vec{z}) - \frac{1}{2}(\mathbf{V}\vec{z} \cdot \vec{z})\frac{1}{\lambda_2^2}w_{2,2}^S(\vec{z}) - \tilde{U}\left(\frac{1}{\lambda_2}w_2(\vec{z})\right). \end{aligned}$$

The region \bar{S} is the union of line segments connecting points $\vec{\zeta} \in \partial_S M$ with $(0, -1/\lambda_2)$. We parameterize these segments as

$$(9.74) \quad z_1(t) = \frac{\zeta_1}{1 + \lambda_2 \zeta_1 t}, \quad z_2(t) = \frac{\zeta_2 - \zeta_1 t}{1 + \lambda_2 \zeta_1 t}, \quad t \geq 0,$$

so that $1 + \lambda_2 z_2(t) = (1 + \lambda_2 \zeta_2)/(1 + \lambda_2 \zeta_1 t)$ and (cf. (8.9))

$$(9.75) \quad \frac{z_1(t)}{1 + \lambda_2 z_2(t)} = \frac{\zeta_1}{1 + \lambda_2 \zeta_2}.$$

We set $\vec{z}(t) = (z_1(t), z_2(t))$ and note that $\vec{z}(0) = \vec{\zeta}$ and $\vec{z}(\infty) = (0, -1/\lambda_2)$ (recall the abuse of notation of Remark 2.4). We show that

$$(9.76) \quad J_S(t) \triangleq \frac{1}{z_1^p(t)} \left[\beta w(\vec{z}(t)) - (r + \vec{\alpha} \cdot \vec{z}(t)) \frac{1}{\lambda_2} w_2(\vec{z}(t)) - \frac{1}{2} (\mathbf{V} \vec{z}(t) \cdot \vec{z}(t)) \frac{1}{\lambda_2^2} w_{2,2}^S(\vec{z}(t)) - \tilde{U} \left(\frac{1}{\lambda_2} w_2(\vec{z}(t)) \right) \right]$$

is nondecreasing in t . From (9.67) and (9.73), we have $J_S(0) \geq 0$. Once we show that J_S is nondecreasing, we can conclude that the expression in (9.73) is nonnegative in $\bar{S} \cap \mathcal{D}_2$.

Note first from (5.2) that

$$(9.77) \quad \frac{1}{z_1^p(t)} \tilde{U} \left(\frac{1}{\lambda_2} w_2(\vec{z}(t)) \right) = \tilde{U} \left(\frac{w_2(\vec{z}(t))}{\lambda_2 z_1^{p-1}(t)} \right).$$

Observe next from (8.13), (8.20), and (9.55) that

$$(9.78) \quad \frac{w(\vec{z}(t))}{z_1^p(t)} = \frac{w(\vec{\zeta})}{\zeta_1^p}, \quad \frac{w_2(\vec{z}(t))}{z_1^{p-1}(t)} = \frac{w_2(\vec{\zeta})}{\zeta_1^{p-1}}, \quad \frac{w_{2,2}^S(\vec{z}(t))}{z_1^{p-2}(t)} = \frac{w_{2,2}^S(\vec{\zeta})}{\zeta_1^{p-2}}.$$

Therefore

$$J_S(t) = \frac{\beta w(\vec{\zeta})}{\zeta_1^p} - \frac{1}{z_1(t)} (r + \vec{\alpha} \cdot \vec{z}(t)) \frac{w_2(\vec{\zeta})}{\lambda_2 \zeta_1^{p-1}} - \frac{1}{2 z_1^2(t)} (\mathbf{V} \vec{z}(t) \cdot \vec{z}(t)) \frac{w_{2,2}^S(\vec{\zeta})}{\lambda_2^2 \zeta_1^{p-2}} - \tilde{U} \left(\frac{w_2(\vec{\zeta})}{\lambda_2 \zeta_1^{p-1}} \right).$$

We examine the terms that depend on t . Using the definitions of $z_i(t)$, we have

$$(9.79) \quad \frac{1}{z_1(t)} (r + \vec{\alpha} \cdot \vec{z}(t)) = (r \lambda_2 - \alpha_2) t + \frac{r + \vec{\alpha} \cdot \vec{\zeta}}{\zeta_1},$$

$$(9.80) \quad \frac{1}{2 z_1^2(t)} (\mathbf{V} \vec{z}(t) \cdot \vec{z}(t)) = \frac{1}{2} \left[\sigma_1^2 + \sigma_2^2 \left(\frac{\zeta_2}{\zeta_1} - t \right)^2 \right].$$

Thus, $J_S(t)$ is a quadratic function of t , and because of (9.41), the coefficient of the t^2 term is positive for sufficiently small $\lambda > 0$. It remains only to show that $J_S'(0) \geq 0$. Using (9.40), (9.34), (8.46), (9.27), the fact that $\zeta \in P_S$, and the definition $\theta_2 = \frac{\alpha_2}{(1-p)\sigma_2^2}$ when $\rho = 0$, we

have

$$\begin{aligned}
J'_S(0) &= -(r\lambda_2 - \alpha_2) \frac{w_2(\vec{\zeta})}{\lambda_2 \zeta_1^{p-1}} + \sigma_2^2 \frac{\zeta_2 w_{2,2}^S(\vec{\zeta})}{\lambda_2^2 \zeta_1^{p-1}} \\
&= -(r\lambda_2 - \alpha_2) \frac{w_2(\vec{\zeta})}{\lambda_2 \zeta_1^{p-1}} - \frac{\sigma_2^2 \zeta_2}{\lambda_2^2 \zeta_1^{p-1}} \left[\frac{1-p}{\zeta_1} w_2(\vec{\zeta}) + O(\lambda^{5/3}) \right] (\lambda_2 \zeta_1 + O(\lambda^2)) \\
&= \frac{w_2(\vec{\zeta})}{\lambda_2 \zeta_1^{p-1}} [\alpha_2 - r\lambda_2 - (1-p)\sigma_2^2 \zeta_2] + O(\lambda^{2/3}) \\
&= \frac{1}{\zeta_1^{p-1}} (A^{p-1} + O(\lambda^{1/3})) [\alpha_2 - (1-p)\sigma_2^2 \zeta_2] + O(\lambda^{2/3}) \\
&= \frac{1}{\zeta_1^{p-1}} (A^{p-1} + O(\lambda^{1/3})) \left[\alpha_2 - (1-p)\sigma_2^2 \left(\theta_2 - \frac{1}{2} \nu_2 \lambda^{1/3} \right) \right] + O(\lambda^{2/3}) \\
(9.81) \quad &= \frac{1}{2\zeta_1^{p-1}} A^{p-1} (1-p)\sigma_2^2 \nu_2 \lambda^{1/3} + O(\lambda^{2/3}),
\end{aligned}$$

which is positive, uniformly for $\vec{\zeta} \in \partial_S M$, for sufficiently small $\lambda > 0$.

When $\zeta \in \partial_E M$, we have from (9.30), (9.31), and (9.43), (9.45), and (9.47) that

$$\begin{aligned}
-\frac{1}{2} (\mathbf{V}\vec{\zeta} \cdot \vec{\zeta}) (\nabla^2 w^M(\vec{\zeta}) \vec{\zeta} \cdot \vec{\zeta}) &= -\frac{1}{2} (\mathbf{V}\vec{\zeta} \cdot \vec{\zeta}) w_{2,2}^M(\vec{\zeta}) \zeta_2^2 + O(\lambda), \\
-\frac{1}{2} (\mathbf{V}\vec{\zeta} \cdot \vec{\zeta}) (\nabla^2 w^E(\vec{\zeta}) \vec{\zeta} \cdot \vec{\zeta}) &= -\frac{1}{2} (\mathbf{V}\vec{\zeta} \cdot \vec{\zeta}) w_{2,2}^M(\vec{\zeta}) \zeta_2^2 + O(\lambda^{5/3}).
\end{aligned}$$

These terms differ by $O(\lambda)$, and (9.64) implies, increasing K if necessary, that

$$(9.82) \quad \mathcal{L}_2 w^E(\vec{\zeta}) - \tilde{U}(pw(\vec{\zeta}) - \nabla w(\vec{\zeta}) \cdot \vec{\zeta}) \geq 0, \quad \vec{\zeta} \in \partial_E M.$$

We extend (9.82) to the rest of $\bar{E} \cap \mathcal{D}_2$. We begin with the inequality $\mathcal{S}_1 w = 0$, which holds on $\bar{E} \cap \mathcal{D}_2$ (Lemma 9.4). We first write this inequality, compute the partial derivative with respect to z_1 , and then compute the partial derivative with respect to z_2 to obtain the three equations

$$(9.83) \quad \lambda_1 p w(\vec{z}) + w_1(\vec{z}) - \lambda_1 \nabla w(\vec{z}) \cdot \vec{z} = 0,$$

$$(9.84) \quad \lambda_1 (p-1) w_1(\vec{z}) + w_{1,1}^E(\vec{z}) - \lambda_1 \nabla w_1^E(\vec{z}) \cdot \vec{z} = 0,$$

$$(9.85) \quad \lambda_1 (p-1) w_2(\vec{z}) + w_{1,2}^E(\vec{z}) - \lambda_1 \nabla w_2^E(\vec{z}) \cdot \vec{z} = 0.$$

Multiplying (9.83) by $1-p$, multiplying (9.84) by $z_1 + (1/\lambda_1)$, multiplying (9.85) by z_2 , summing the three resulting equations, and then dividing by λ_1 , we obtain

$$(9.86) \quad -p(1-p)w(\vec{z}) + 2(1-p)\nabla w(\vec{z}) \cdot \vec{z} + \nabla^2 w^E(\vec{z}) \vec{z} \cdot \vec{z} = \frac{1}{\lambda_1^2} w_{1,1}^E(\vec{z}), \quad \vec{z} \in \bar{E} \cap \mathcal{D}_2.$$

Therefore, in place of (9.73), we have

$$\begin{aligned}
&\mathcal{L}_2 w^E(\vec{z}) - \tilde{U}(pw(\vec{z}) - \nabla w(\vec{z}) \cdot \vec{z}) \\
(9.87) \quad &= \beta w(\vec{z}) + (r + \vec{\alpha} \cdot \vec{z}) \frac{1}{\lambda_1} w_1(\vec{z}) - \frac{1}{2} (\mathbf{V}\vec{z} \cdot \vec{z}) \frac{1}{\lambda_1^2} w_{1,1}^E(\vec{z}) - \tilde{U} \left(-\frac{1}{\lambda_1} w_1(\vec{z}) \right).
\end{aligned}$$

The region \overline{E} is the union of line segments connecting points $\vec{\zeta} \in \partial_E M$ with $(1/\lambda_1, 0)$. We parameterize these segments as $z_1(t) = (\zeta_1 + \zeta_2 t)/(1 + \lambda_1 \zeta_2 t)$, $z_2(t) = \zeta_2/(1 + \lambda_1 \zeta_2 t)$ and show that

$$J_E(t) \triangleq \frac{1}{z_2^p(t)} \left[\beta w(\vec{z}(t)) + (r + \vec{\alpha} \cdot \vec{z}(t)) \frac{1}{\lambda_1} w_1(\vec{z}(t)) - \frac{1}{2} (\mathbf{V} \vec{z}(t) \cdot \vec{z}(t)) \frac{1}{\lambda_1^2} w_{1,1}^E(\vec{z}(t)) - \tilde{U} \left(-\frac{1}{\lambda_1} w_1(\vec{z}(t)) \right) \right]$$

is nondecreasing in t . From (9.82) and (9.87), we have $J_E(0) \geq 0$. From (8.13), (8.21), and (9.56), we have

$$J_E(t) = \frac{\beta w(\vec{\zeta})}{\zeta_2^p} + \frac{1}{z_2(t)} (r + \vec{\alpha} \cdot \vec{z}(t)) \frac{w_1(\vec{\zeta})}{\lambda_1 \zeta_2^{p-1}} - \frac{1}{2z_2^2(t)} (\mathbf{V} \vec{z}(t) \cdot \vec{z}(t)) \frac{w_{1,1}^E(\vec{\zeta})}{\lambda_1^2 \zeta_2^{p-2}} - \tilde{U} \left(-\frac{w_1(\vec{\zeta})}{\lambda_1 \zeta_2^{p-1}} \right).$$

The terms that depend on t are

$$\begin{aligned} \frac{1}{z_2(t)} (r + \vec{\alpha} \cdot \vec{z}(t)) &= (r\lambda_1 + \alpha_1)t + \frac{r + \vec{\alpha} \cdot \vec{\zeta}}{\zeta_2}, \\ \frac{1}{2z_2^2(t)} (\mathbf{V} \vec{z}(t) \cdot \vec{z}(t)) &= \frac{1}{2} \left[\sigma_1^2 \left(\frac{\zeta_1}{\zeta_2} + t \right)^2 + \sigma_2^2 \right]. \end{aligned}$$

Thus, $J_E(t)$ is a quadratic function of t , and because of (9.43), the coefficient of the t^2 term is positive. It remains to show that $J'_E(0) \geq 0$. Using (9.42), (9.35), (8.48), and (9.28), one can show that $J'_E(0) = A^{p-1}(1-p)\sigma_1^2\nu_1\lambda^{1/3}/(2\zeta_2^{p-1}) + O(\lambda^{2/3})$, where we also use the fact that $\vec{\zeta} \in P_E$. For sufficiently small $\lambda > 0$, $J'_E(0) \geq 0$.

Finally, we establish (9.65) in $\overline{SE} \cap \mathcal{D}_2$. We first establish this inequality on the boundary between $\overline{SE} \cap \mathcal{D}_2$ and $\overline{E} \cap \mathcal{D}_2$. For this we repeat the proof above that (9.65) holds on \overline{E} , replacing $\vec{\zeta}$ in that argument by $\vec{\zeta}^{SE}$, so that $\vec{z}(t)$ traverses the boundary between $\overline{SE} \cap \mathcal{D}_2$ and $\overline{E} \cap \mathcal{D}_2$. This is possible because (9.83) holds on \overline{SE} (Lemma 9.4), and the counterparts of (9.82) and (9.86),

$$(9.88) \quad \mathcal{L}_2 w^{SE}(\vec{\zeta}^{SE}) - \tilde{U}(pw(\vec{\zeta}^{SE}) - \nabla w(\vec{\zeta}^{SE}) \cdot \vec{\zeta}^{SE}) \geq 0, \\ -p(1-p)w(\vec{z}) + 2(1-p)\nabla w(\vec{z}) \cdot \vec{z} + \nabla^2 w^{SE}(\vec{z})\vec{z} \cdot \vec{z} = \frac{1}{\lambda_1^2} w_{1,1}^{SE}(\vec{z}), \quad \vec{z} \in \overline{SE} \cap \mathcal{D}_2,$$

(9.89)

also hold, as we now show. Inequality (9.88) follows from (9.64), using the estimates (9.32) and (9.52)–(9.54), where it may again be necessary to increase K . Lemma 9.4 implies not only that (9.83) holds on \overline{SE} , but also that (9.84) and (9.85) hold there when we replace $w_{i,j}^E$ by $w_{i,j}^{SE}$ in these equations. We obtain (9.89) from these equations just as we obtained (9.86). We conclude that

$$(9.90) \quad \mathcal{L}_2 w^{SE}(\vec{z}) - \tilde{U}(pw(\vec{z}) - \nabla w(\vec{z}) \cdot \vec{z}) \geq 0, \quad \vec{z} \in \overline{SE} \cap \overline{E} \cap \mathcal{D}_2.$$

Lemma 9.4 implies that (9.69)–(9.71) and hence (9.72) hold in $\overline{SE} \cap \mathcal{D}_2$ (where we replace $w_{i,j}^S$ by $w_{i,j}^{SE}$ in these equations). Let $\vec{z} \in \overline{SE} \cap \mathcal{D}_2$ be given. The point \vec{z} is on a line segment connecting $(0, -1/\lambda_2)$ with a point we call $\vec{\zeta}$ in $\overline{SE} \cap \overline{E} \cap \mathcal{D}_2$. We parameterize this line segment by (9.74), and, analogously to (9.76), we define

$$J_{SE}(t) \triangleq \frac{1}{z_1^p(t)} \left[\beta w(\vec{z}(t)) - (r + \vec{\alpha} \cdot \vec{z}(t)) \frac{1}{\lambda_2} w_2(\vec{z}(t)) - \frac{1}{2} (\mathbf{V} \vec{z}(t) \cdot \vec{z}(t)) \frac{1}{\lambda_2^2} w_{2,2}^{SE}(\vec{z}(t)) - \tilde{U} \left(\frac{1}{\lambda_2} w_2(\vec{z}(t)) \right) \right].$$

From (9.90), we have $J_{SE}(0) \geq 0$. As in the proof of (9.65) in S , we need only to show that J_{SE} is nondecreasing.

Just as before, we have (9.77). Set $\eta = \frac{-\lambda_1 \zeta_1}{1 - \lambda_1 \zeta_1 + \lambda_2 \zeta_2}$. From (9.75), we have $\frac{z_1(t)}{\zeta_1} = \frac{1 + \lambda_2 z_2(t)}{1 + \lambda_2 \zeta_2}$, and hence

$$(9.91) \quad \frac{z_1(t)}{\zeta_1} = \eta \frac{z_1(t)}{\zeta_1} + (1 - \eta) \frac{1 + \lambda_2 z_2(t)}{1 + \lambda_2 \zeta_2} = \frac{1 - \lambda_1 z_1(t) + \lambda_2 z_2(t)}{1 - \lambda_1 \zeta_1 + \lambda_2 \zeta_2}.$$

From (8.13) applied to both $\vec{z}(t)$ and $\vec{\zeta}$, we have

$$w(\vec{z}(t)) = \left(\frac{1 - \lambda_1 z_1(t) + \lambda_2 z_2(t)}{1 - \lambda_1 \zeta_1^{SE} + \lambda_2 \zeta_2^{SE}} \right)^p w(\zeta^{SE}),$$

$$w(\vec{\zeta}) = \left(\frac{1 - \lambda_1 \zeta_1 + \lambda_2 \zeta_2}{1 - \lambda_1 \zeta_1^{SE} + \lambda_2 \zeta_2^{SE}} \right)^p w(\zeta^{SE}),$$

and these equations together with (9.91) imply the first equation in (9.78). Applying (8.27) to both $\vec{z}(t)$ and ζ and arguing in the same way, we obtain the second equation in (9.78). Applying (9.51) to both $\vec{z}(t)$ and ζ , we obtain the analogue

$$\frac{w_{2,2}^{SE}(\vec{z}(t))}{z_1^{p-2}(t)} = \frac{w_{2,2}^{SE}(\vec{\zeta})}{\zeta_1^{p-2}}$$

of the third equation in (9.78). Using the above equations we obtain the formula

$$J_{SE}(t) = \frac{\beta w(\vec{\zeta})}{\zeta_1^p} - \frac{1}{z_1(t)} (r + \vec{\alpha} \cdot \vec{z}(t)) \frac{w_2(\vec{\zeta})}{\lambda_2 \zeta_1^{p-1}} - \frac{1}{2 z_1^2(t)} (\mathbf{V} \vec{z}(t) \cdot \vec{z}(t)) \frac{w_{2,2}^{SE}(\vec{\zeta})}{\lambda_2^2 \zeta_1^{p-2}} - \tilde{U} \left(\frac{w_2(\vec{\zeta})}{\lambda_2 \zeta_1^{p-1}} \right).$$

The terms that depend on t satisfy (9.79) and (9.80), and because of (9.27) and (9.51), the coefficient of the t^2 term is positive for sufficiently small $\lambda > 0$. It remains only to show that $J'_{SE}(0) \geq 0$. The analogue of the first equality in (9.81) is

$$(9.92) \quad J'_{SE}(0) = -(r\lambda_2 - \alpha_2) \frac{w_2(\vec{\zeta})}{\lambda_2 \zeta_1^{p-1}} + \sigma_2 \frac{\zeta_2 w_{2,2}^{SE}(\vec{\zeta})}{\lambda_2^2 \zeta_1^{p-1}}.$$

Substitution of (8.24) and (9.51) into (9.92) yields

$$J'_{SE}(0) = \frac{(1 - \lambda_1 \zeta_1 + \lambda_2 \zeta_2)^{p-2}}{\lambda_2 \zeta_1^{p-1} (1 - \lambda_1 \zeta_1^{SE} + \lambda_2 \zeta_2^{SE})^{p-1}} [(\alpha_2 - r\lambda_2)(1 - \lambda_1 \zeta_1 + \lambda_2 \zeta_2) - \zeta_2 \sigma_2^2 (1 - p)] w_2(\vec{\zeta}^{SE}).$$

The positivity of $w_2(\vec{\zeta}^{SE})$ is ensured by (9.27). Because $\vec{\zeta}$ is on the line segment connecting $(1/\lambda_1, 0)$ and $\vec{\zeta}^{SE}$, we have (cf. (8.10)) $1 - \lambda_1 \zeta_1 = (\zeta_2/\zeta_2^{SE})(1 - \lambda_1 \zeta_1^{SE})$. It thus suffices to establish the positivity of

$$\begin{aligned} & (\alpha_2 - r\lambda_2)(1 - \lambda_1 \zeta_1 + \lambda_2 \zeta_2) - \zeta_2 \sigma_2^2 (1 - p) \\ &= \frac{\zeta_2}{\zeta_2^{SE}} [(\alpha_2 - r\lambda_2)(1 - \lambda_1 \zeta_1^{SE} + \lambda_2 \zeta_2^{SE}) - \sigma_2^2 (1 - p) \zeta_2^{SE}] \\ &= \frac{\zeta_2}{\zeta_2^{SE}} [\alpha_2 - \sigma_2^2 (1 - p) \zeta_2^{SE} + O(\lambda)] \\ &= \frac{\zeta_2}{\zeta_2^{SE}} \left[\alpha_2 - \sigma_2^2 (1 - p) \left(\theta_2 - \frac{1}{2} \nu_2 \lambda^{1/3} + O(\lambda^{2/3}) \right) \right] \\ &= \frac{\zeta_2}{\zeta_2^{NE}} \left[\frac{1}{2} \sigma_2^2 (1 - p) \nu_2 \lambda^{1/3} + O(\lambda^{2/3}) \right], \end{aligned}$$

which is positive for sufficiently small $\lambda > 0$. ■

9.6. Conclusion of the proof of Theorem 9.1. Since w^M is strictly positive on \overline{M} for sufficiently small λ , in fact, is $\frac{1}{p} A^{p-1} - \gamma \lambda^{2/3} + O(\lambda)$ on this set, the function w is strictly positive on \mathcal{D}_2 . Also, w is equal to zero on $\partial \mathcal{D}_2$.

We review the proof to this point to conclude that for sufficiently large positive K , say $K = K^+ > 0$, w satisfies (8.50) on each of the nine regions into which we have partitioned \mathcal{D}_2 . For the region \overline{M} , this is Lemmas 9.4 and 9.8. For the other regions, this is Lemma 9.9 and equality (8.32), which follows from the fact that w satisfies (8.31) (see Lemma 9.4). We have partitioned the solvency region \mathcal{D}_2 into nine regions, as shown in Figure 1. The boundaries of these regions are straight lines, except for the four boundaries of M , which are nearly horizontal or vertical because of (9.16). Therefore, for every $\vec{z} \in \mathcal{D}_2$, there is a line segment in the radial direction with one end point at \vec{z} that is entirely in the closure of one of the nine regions. Theorem 8.6 implies that w is a viscosity supersolution of (6.15). When $K = K^+$, we denote w by w^+ .

We have also shown that for sufficiently small K , say $K = K^- < 0$, w satisfies (8.49) in each of the nine regions into which we have partitioned \mathcal{D}_2 . For the region \overline{M} , this is Lemma 9.8. For the other regions, this is equality (8.32). Theorem 8.6 implies that w is a viscosity subsolution of (6.15). When $K = K^-$, we denote w by w^- .

Theorem 6.3 implies that $w^- \leq u \leq w^+$ on \mathcal{D} , and in particular,

$$\frac{1}{p} A^{p-1} - \gamma \lambda^{2/3} + K^- \lambda + O(\lambda^{4/3}) \leq u(\vec{\theta}) \leq \frac{1}{p} A^{p-1} - \gamma \lambda^{2/3} + K^+ \lambda + O(\lambda^{4/3}).$$

This is Theorem 9.1.

Remark 9.10. Let us fix a compact set $C \subset \mathbb{R}^2$ such that C contains $R(1)$ defined by (9.19). For $0 < \lambda \leq 1$, $R(\lambda)$ contains M , and w defined by (9.24) is equal to $\frac{1}{p}A^{p-1} - \gamma\lambda^{2/3} + O(\lambda)$ on M , and thus $u(\vec{\theta})$ differs from u on the boundary of M by $O(\lambda)$. The solvency region \mathcal{D}_2 depends on λ , expanding to cover all of \mathbb{R}^2 as λ decreases to zero. For small enough $\lambda > 0$, $C \subset \mathcal{D}_2$ and it is apparent from the extension formula (8.13) that w on $C \setminus M$ differs from its value on the boundary of M by $O(\lambda)$. Therefore, the conclusion (9.1) of Theorem 9.1 can be extended to assert that

$$(9.93) \quad u(\vec{z}) = \frac{1}{p}A^{p-1} - \gamma\lambda^{2/3} + O(\lambda), \quad \vec{z} \in C,$$

where $C \subset \mathbb{R}^2$ is an arbitrary compact set containing $R(1)$, and the $O(\lambda)$ term in (9.93) is uniform over C . In fact, we do not need the condition that C contains $R(1)$ since any compact subset of \mathbb{R}^2 is a subset of a compact subset containing $R(1)$.

10. The case $\rho \neq 0$. In this section we prove the main result, Theorem 4.1, under the assumption $\rho \neq 0$. As in section 9, we first state the result for the function u of two variables defined by (6.3); see Theorem 10.1 below. Following Theorem 10.1, we provide the proof of Theorem 4.1.

10.1. Auxiliary problem. In order to exploit our analysis of the case $\rho = 0$ contained in Theorem 9.1, we define an auxiliary problem in which the tradable assets are the risk-free asset with rate of return r , the first type of futures contract, and a fund holding both types of futures contracts. In particular, we define the two risky price processes

$$(10.1) \quad \underline{F}_1(t) = F_1(t), \quad \underline{F}_2(t) = -\frac{\rho\sigma_2}{\sigma_1}F_1(t) + F_2(t).$$

In vector notation, $\vec{F}(t) = \mathbf{D}\vec{F}(t)$, where

$$(10.2) \quad \mathbf{D} \triangleq \begin{bmatrix} 1 & 0 \\ -\frac{\rho\sigma_2}{\sigma_1} & 1 \end{bmatrix}.$$

Whereas $\frac{d}{dt}\langle \vec{F}, \vec{F} \rangle(t) = \mathbf{V}$, we have

$$(10.3) \quad \frac{d}{dt}\langle \vec{F}, \vec{F} \rangle = \underline{\mathbf{V}} \triangleq \mathbf{D}\mathbf{V}\mathbf{D}^T = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & (1 - \rho^2)\sigma_2^2 \end{bmatrix},$$

and hence \underline{F}_1 and \underline{F}_2 are independent Gaussian processes. The counterparts of the drift vector $\vec{\alpha}$ and the desired position vector $\vec{\theta}$ for the market with risky assets \underline{F}_1 and \underline{F}_2 and risk-free rate r are (cf. (2.10))

$$(10.4) \quad \underline{\vec{\alpha}} = \mathbf{D}\vec{\alpha}, \quad \underline{\vec{\theta}} = \frac{1}{1-p}\underline{\mathbf{V}}^{-1}\underline{\vec{\alpha}} = (\mathbf{D}^T)^{-1}\vec{\theta}.$$

In order to state the following theorem, we also define $\underline{\sigma}_1 = \sigma_1$, $\underline{\sigma}_2 = \sigma_2\sqrt{1-\rho^2}$, $\underline{\mu}_1 = \mu_1$, and $\underline{\mu}_2 = -\frac{\rho\sigma_2}{\sigma_1}\mu_1 + \mu_2$. Because of the second equality in (2.4), $\underline{\mu}_2$ is strictly positive. We also define $\underline{\lambda}_i = \underline{\mu}_i\lambda$, $i = 1, 2$.

Theorem 10.1. *Assume (2.1) and (2.4). Under Assumptions 2.2 and 2.3, the value function u for the problem with positive λ satisfies*

$$(10.5) \quad u(\vec{\theta}) = \frac{1}{p}A^{p-1} - \underline{\gamma}\lambda^{2/3} + O(\lambda).$$

The constant $\underline{\gamma}$ in (10.5) is

$$(10.6) \quad \underline{\gamma} \triangleq \frac{3}{2}A^{p-2}(\mathbf{V}\vec{\theta} \cdot \vec{\theta}) \sum_{i=1}^2 \frac{\mu_i \theta_i^2}{\underline{\nu}_i},$$

where

$$(10.7) \quad \underline{\nu}_i = \sqrt[3]{\frac{12\mu_i \theta_i^2}{(1-p)\sigma_i^2}(\mathbf{V}\vec{\theta} \cdot \vec{\theta})}, \quad i = 1, 2.$$

Remark 10.2. It is straightforward to verify that $\mathbf{V}\vec{\theta} \cdot \vec{\theta} = \mathbf{V}\vec{\theta} \cdot \vec{\theta}$.

Before providing the proof of Theorem 10.1, we consider its consequences. As in Remark 9.10, we see that for an arbitrary compact set $C \subset \mathbb{R}^2$,

$$(10.8) \quad u(\vec{z}) = \frac{1}{p}A^{p-1} - \underline{\gamma}\lambda^{2/3} + O(\lambda), \quad z \in C,$$

where the $O(\lambda)$ term in (10.8) is uniform over C . Note also that if $\rho = 0$, (10.5) and (10.8) reduce to (9.1) and (9.93), respectively, so Theorem 9.1 is a special case of Theorem 10.1.

Proof of Theorem 4.1. Assume that (y_1, y_2, x) is in a compact subset of the (open) solvency region \mathcal{D}_3 . Then x is bounded and bounded away from zero, so $\vec{z} \triangleq (\frac{y_1}{x}, \frac{y_2}{x})$ is in a compact set. Multiplying (10.8) by x^p and using (3.3) and (6.4), we obtain Theorem 4.1. ■

We prove Theorem 10.1 by first providing an upper bound for $u(\vec{\theta})$ in section 10.2. In subsequent sections we obtain a lower bound. Both these bounds are of the form of the right-hand side of (10.5), which implies that (10.5) holds.

10.2. Upper bound.

Lemma 10.3. *Under the assumptions of Theorem 10.1,*

$$(10.9) \quad u(\vec{\theta}) \leq \frac{1}{p}A^{p-1} - \underline{\gamma}\lambda^{2/3} + O(\lambda).$$

Proof. Consider the auxiliary problem with risky price processes given by (10.1). Suppose that for $i = 1, 2$, trading in the process with price \underline{F}_i incurs proportional transaction cost $\underline{\lambda}_i \triangleq \frac{\mu_i}{x}\lambda$. We verify that this problem satisfies Assumptions 2.2 and 2.3, i.e., that $\vec{\theta}$ is in the first quadrant and that

$$(10.10) \quad \frac{\beta - rp}{1-p} - \frac{p}{2(1-p)^2} \mathbf{V}^{-1} \vec{\alpha} \cdot \vec{\alpha}$$

is strictly positive. But $\mathbf{V}^{-1} \vec{\alpha} \cdot \vec{\alpha} = (\mathbf{D}\mathbf{V}\mathbf{D}^T)^{-1} \mathbf{D}\vec{\alpha} \cdot \mathbf{D}\vec{\alpha} = \mathbf{V}^{-1} \vec{\alpha} \cdot \vec{\alpha}$, and hence the expression in (10.10) is A defined by (2.11), and Assumption 2.3 for the auxiliary problem follows from Assumption 2.3 for the original problem.

The first equality in the formula for $\vec{\theta}$ in (10.4) shows that $\underline{\theta}_1 = \alpha_1 / ((1-p)\sigma_1^2)$, and this is positive because of (2.1). The second equality in the formula for $\vec{\theta}$ implies $\underline{\theta}_2 = \theta_2$, which is positive because of Assumption 2.2 for the original problem. Hence, Assumption 2.2 for the auxiliary problem holds.

In light of Lemma 6.5, it suffices to prove (10.9) with $u(\vec{0})$ replacing $u(\vec{\theta})$ on the left-hand side. We first apply Theorem 9.1 and Lemma 6.5 to the auxiliary problem to conclude that its two-dimensional value function \underline{u} satisfies

$$(10.11) \quad \underline{u}(\vec{0}) = \frac{1}{p}A^{p-1} - \underline{\gamma}\lambda^{2/3} + O(\lambda).$$

To show (10.9) it thus suffices to show that the auxiliary problem is “more favorable” than the original problem. In particular, we show that given a policy for the original problem that is admissible for the initial condition $(0, 0, 1)$, there is a corresponding policy for the auxiliary problem that is admissible for this initial condition and that results in the same expected utility.

Let $C, L_i, M_i, i = 1, 2$, be an admissible policy in the original problem for the initial condition $Y_1(0-) = 0, Y_2(0-) = 0, X(0-) = 1$. Then Y_1, Y_2, X given by (2.2) and (2.3) are in the closure of the solvency region \mathcal{D}_3 given by (2.7) at all times, i.e.,

$$(10.12) \quad X(t) - \lambda_1|Y_1(t)| - \lambda_2|Y_2(t)| \geq 0, \quad t \geq 0.$$

Define

$$\underline{C} = C, \quad \underline{L}_1 = L_1 + \frac{\rho\sigma_2}{\sigma_1}L_2, \quad \underline{M}_1 = M_1 + \frac{\rho\sigma_2}{\sigma_1}M_2, \quad \underline{L}_2 = L_2, \quad \underline{M}_2 = M_2,$$

and use this policy in the auxiliary problem with initial condition $\underline{Y}_1(0-) = 0, \underline{Y}_2(0-) = 0, \underline{X}(0-) = 1$. Then at each time $t \geq 0$,

$$\begin{aligned} \underline{Y}_1(t) &= L_1(t) - M_1(t) + \frac{\rho\sigma_2}{\sigma_1}(L_2(t) - M_2(t)) = Y_1(t) + \frac{\rho\sigma_2}{\sigma_1}Y_2(t), \\ \underline{Y}_2(t) &= L_2(t) - M_2(t) = Y_2(t). \end{aligned}$$

The corresponding auxiliary problem money market position satisfies (cf. (2.3))

$$\begin{aligned} d\underline{X}(t) &= \sum_{i=1}^2 \underline{Y}_i(t) d\underline{F}_i(t) - \sum_{i=1}^2 \underline{\lambda}_i(d\underline{L}_i(t) + d\underline{M}_i(t)) + (r\underline{X}(t) - \underline{C}(t)) dt \\ &= \sum_{i=1}^2 Y_i(t) dF_i(t) - \sum_{i=1}^2 \lambda_i(dL_i(t) + dM_i(t)) + (rX(t) - C(t)) dt. \end{aligned}$$

Because \underline{X} satisfies the stochastic differential equation (2.3) determining X , and $\underline{X}(0-) = X(0-) = 1$, we have $\underline{X}(t) = X(t)$ for all $t \geq 0$. The three-dimensional solvency region for the auxiliary problem is (cf. (2.7))

$$(10.13) \quad \underline{\mathcal{D}}_2 \triangleq \{(y_1, y_2, x : x - \lambda_1|y_2| - \lambda_2|y_2| > 0\}.$$

The triangle inequality implies

$$\underline{X}(t) - \underline{\lambda}_1 |\underline{Y}_1(t)| - \underline{\lambda}_2 |\underline{Y}_2(t)| \geq X(t) - \lambda_1 |Y_1(t)| - \lambda_2 |Y_2(t)|,$$

and admissibility of \underline{C} , \underline{L}_i , \underline{M}_i , $i = 1, 2$, in the auxiliary problem for the initial condition $(0, 0, 1)$ follows from (10.12). Because $C = \underline{C}$, the expected utility accumulated by the two policies in their respective problems coincides. ■

Remark 10.4. We note in the proof of Lemma 10.3 that if L_1 and M_2 increase simultaneously while M_1 and L_2 remain constant, which might happen for the optimal policy in the original problem, then \underline{L}_1 and \underline{M}_1 increase simultaneously. But it is strictly suboptimal to buy and sell the same risky asset simultaneously. In this situation we can construct a policy that does strictly better in the auxiliary problem than the policy constructed in the proof of Lemma 10.3, and hence it is possible that the value function in the auxiliary problem strictly dominates the value function in the original problem. In particular, it is not straightforward to use the auxiliary problem to obtain the reverse of inequality (10.9).

10.3. Partitioning the solvency region for the lower bound. In this section we obtain the reverse of inequality (10.9) by constructing a subsolution for the original problem. This subsolution is built on the observation that the second derivative term in the operator \mathcal{L}_2 defined by (6.10) is the second directional derivative in the radial direction. Furthermore, trading in either the positive or negative radial direction does not call for buying one type of contract while selling another type, and hence the problem noted in Remark 10.4 does not arise. As in section 9, we do not actually construct a nearly optimal policy, but rather construct a nearly optimal expected utility by partitioning \mathcal{D}_2 into regions similarly to Figure 1 and then constructing a function piecemeal in these regions.

We modify the construction in section 9.1 to first partition the solvency region $\underline{\mathcal{D}}_2$ for the auxiliary problem. Let B be a fixed positive constant and K a negative constant to be chosen later. Still using the notation of the auxiliary problem, for

$$(10.14) \quad -\frac{1}{2}\underline{\nu}_1\lambda^{1/3} \leq \underline{\delta}_1 \leq \frac{1}{2}\underline{\nu}_1\lambda^{1/3}, \quad -\frac{1}{2}\underline{\nu}_2\lambda^{1/3} \leq \underline{\delta}_2 \leq \frac{1}{2}\underline{\nu}_2\lambda^{1/3},$$

in place of (9.5) we define

$$(10.15) \quad \underline{h}_i(\underline{\delta}_i) \triangleq \underline{\mu}_i \left[\frac{3}{2}\underline{\delta}_i^2\lambda^{2/3} - \frac{1}{\underline{\nu}_i^2}\underline{\delta}_i^4 + \frac{3}{2}B\underline{\delta}_i^2\lambda^{4/3} \right] = O(\lambda^{4/3})$$

so that the analogues of (9.6) and (9.7) hold. We further define (cf. (9.8))

$$\underline{H}(\underline{\delta}_1, \underline{\delta}_2) \triangleq \sum_{i=1}^2 \frac{1}{\underline{\nu}_i} \left[-p\underline{h}_i(\underline{\delta}_i) + (\underline{\delta}_i + \underline{\theta}_i)\underline{h}'_i(\underline{\delta}_i) \right] = O(\lambda)$$

so that the analogue of (9.9) holds. Finally, we set

$$\begin{aligned} \underline{G}_S(\underline{\delta}_1, \underline{\delta}_2) &\triangleq \lambda_2 - p\underline{\gamma}A^{1-p}\lambda_2\lambda^{2/3} + pKA^{1-p}\lambda_2\lambda + \frac{\underline{h}'_2(\underline{\delta}_2)}{\underline{\nu}_2} + \lambda_2\underline{H}(\underline{\delta}_1, \underline{\delta}_2) + \frac{\rho\sigma_2\underline{h}'_1(\underline{\delta}_1)}{\sigma_1\underline{\nu}_1}, \\ \underline{G}_E(\underline{\delta}_1, \underline{\delta}_2) &\triangleq \lambda_1 - p\underline{\gamma}A^{1-p}\lambda_1\lambda^{2/3} + pKA^{1-p}\lambda_1\lambda - \frac{\underline{h}'_1(\underline{\delta}_1)}{\underline{\nu}_1} + \lambda_1\underline{H}(\underline{\delta}_1, \underline{\delta}_2), \end{aligned}$$

$$\begin{aligned}\underline{G}_N(\underline{\delta}_1, \underline{\delta}_2) &\triangleq \lambda_2 - p\underline{\gamma}A^{1-p}\lambda_2\lambda^{2/3} + pKA^{1-p}\lambda_2\lambda - \frac{h'_2(\underline{\delta}_2)}{\underline{\nu}_2} + \lambda_2\underline{H}(\underline{\delta}_1, \underline{\delta}_2) - \frac{\rho\sigma_2 h'_1(\underline{\delta}_1)}{\sigma_1\underline{\nu}_1}, \\ \underline{G}_W(\underline{\delta}_1, \underline{\delta}_2) &\triangleq \lambda_1 - p\underline{\gamma}A^{1-p}\lambda_1\lambda^{2/3} + pKA^{1-p}\lambda_1\lambda + \frac{h'_1(\underline{\delta}_1)}{\underline{\nu}_1} + \lambda_1\underline{H}(\underline{\delta}_1, \underline{\delta}_2).\end{aligned}$$

The definitions of \underline{G}_E and \underline{G}_W are parallel to the definitions (9.11) and (9.13) of G_E and G_W . The definitions of \underline{G}_S and \underline{G}_N are not parallel to the definitions of G_S and G_N in (9.10) and (9.12), but we have the relationships

$$(10.16) \quad \underline{G}_S(\underline{\delta}_1, \underline{\delta}_2) - \frac{\rho\sigma_2}{\sigma_1}\underline{G}_W(\underline{\delta}_1, \underline{\delta}_2) = \underline{\mu}_2(\lambda - p\underline{\gamma}A^{1-p}\lambda^{5/3}) + \frac{h'_2(\underline{\delta}_2)}{\underline{\nu}_2} + O(\lambda^2),$$

$$(10.17) \quad \underline{G}_N(\underline{\delta}_1, \underline{\delta}_2) - \frac{\rho\sigma_2}{\sigma_1}\underline{G}_E(\underline{\delta}_1, \underline{\delta}_2) = \underline{\mu}_2(\lambda - p\underline{\gamma}A^{1-p}\lambda^{5/3}) - \frac{h'_2(\underline{\delta}_2)}{\underline{\nu}_2} + O(\lambda^2).$$

Lemma 10.5. *For sufficiently small $\lambda > 0$, there exist C^1 functions $\underline{F}_E(\underline{\delta}_2)$ and $\underline{F}_W(\underline{\delta}_2)$, so that*

$$(10.18) \quad \underline{G}_E(\underline{F}_E(\underline{\delta}_2), \underline{\delta}_2) = 0, \quad \underline{G}_W(\underline{F}_W(\underline{\delta}_2), \underline{\delta}_2) = 0, \quad -\frac{1}{2}\underline{\nu}_2\lambda^{1/3} \leq \underline{\delta}_2 \leq \frac{1}{2}\underline{\nu}_2\lambda^{1/3}.$$

These functions are “nearly constant” in the sense that for $\underline{\delta}_2$ as specified in (10.18) and for any constant $\xi > \sqrt{\frac{2}{3}p\underline{\gamma}A^{1-p} + B}$,

$$\begin{aligned}\frac{1}{2}\underline{\nu}_1\lambda^{1/3}(1 - \xi\lambda^{1/3}) &< \underline{F}_E(\underline{\delta}_2) < \frac{1}{2}\underline{\nu}_1\lambda^{1/3}, \\ -\frac{1}{2}\underline{\nu}_1\lambda^{1/3} &< \underline{F}_W(\underline{\delta}_2) < -\frac{1}{2}\underline{\nu}_1\lambda^{1/3}(1 - \xi\lambda^{1/3}).\end{aligned}$$

In addition, $\underline{f}_E(\underline{\zeta}_2) \triangleq \underline{F}_E(\underline{\zeta}_2 - \underline{\theta}_2) + \underline{\theta}_1$ and $\underline{f}_W(\underline{\zeta}_2) \triangleq \underline{F}_W(\underline{\zeta}_2 - \underline{\theta}_2) + \underline{\theta}_1$ are C^1 and satisfy (8.6) and (8.8) and

$$(10.19) \quad \underline{f}'_E(\underline{\zeta}_2) = O(\lambda^{1/3}), \quad \underline{f}'_W(\underline{\zeta}_2) = O(\lambda^{1/3})$$

as long as $\underline{\delta}_2 = \underline{\zeta}_2 - \underline{\theta}_2$ is as specified in (10.18).

Proof. Because the definitions of \underline{G}_E and \underline{G}_W are parallel to the definitions of G_E and G_W , $\underline{\mu}_1 = \mu_1$, and $\underline{\nu}_1 = \nu_1$, the proof of Lemma 9.2 applies. \blacksquare

To construct a middle region \underline{M} in $\underline{\mathcal{D}}_2$, we first locate the southwest corner. We fix $\xi > \sqrt{\frac{2}{3}p\underline{\gamma}A^{1-p} + B}$. We set $\underline{\delta}_2 = -\frac{1}{2}\underline{\nu}_2\lambda^{1/3}(1 - \xi\lambda^{1/3})$ and use (10.16) and (10.18) to write

$$\underline{G}_S(\underline{F}_W(\underline{\delta}_2), \underline{\delta}_2) = \underline{\mu}_2(\lambda - p\underline{\gamma}A^{1-p}\lambda^{5/3}) + \frac{h'_2(\underline{\delta}_2)}{\underline{\nu}_2} + O(\lambda^2).$$

Writing $\underline{\delta}_2$ as

$$\underline{\delta}_2 = -\frac{1}{2}\underline{\nu}_2\lambda^{1/3} \left(1 - \lambda^{1/3} \sqrt{\frac{2}{3}p\underline{\gamma}A^{1-p} + B + \eta} \right),$$

so that

$$\frac{4\underline{\delta}_2^2}{\underline{\nu}_2^2} = \lambda^{2/3} - 2\lambda\sqrt{\frac{2}{3}p\underline{\gamma}A^{1-p} + B + \eta} + \frac{2}{3}p\underline{\gamma}A^{1-p}\lambda^{4/3} + B\lambda^{4/3} + \eta\lambda^{4/3},$$

we compute

$$(10.20) \quad \underline{G}_S(\underline{F}_W(\underline{\delta}_2), \underline{\delta}_2) = \frac{3}{2}\underline{\mu}_2\eta\lambda^{5/3} + O(\lambda^2).$$

The $O(\lambda^2)$ term in (10.20) is uniform in $\eta \in [-B, B]$, for sufficiently small $\lambda > 0$ the expression in (10.20) is negative for $\eta = -B$ and positive for $\eta = B$, and \underline{G}_S and \underline{F}_W are continuous. Hence, there must exist $\eta \in [-B, B]$ and corresponding $\underline{\delta}_2$ such that $\underline{G}_S(\underline{F}_W(\underline{\delta}_2), \underline{\delta}_2) = 0$. We denote this value of $\underline{\delta}_2$ by $\underline{\delta}_2^{SW}$, and we define $\underline{\delta}_1^{SW} = \underline{F}_{SW}(\underline{\delta}_2^{SW})$, $\vec{\delta}^{SW} = (\underline{\delta}_1^{SW}, \underline{\delta}_2^{SW})$. In conclusion, by increasing ξ to equal $\sqrt{\frac{2}{3}p\underline{\gamma}A^{1-p} + 2B}$ if necessary, we have

$$(10.21) \quad -\frac{1}{2}\underline{\nu}_1\lambda^{1/3} < \underline{\delta}_1^{SW} < -\frac{1}{2}\underline{\nu}_1\lambda^{1/3}(1 - \xi\lambda^{1/3}),$$

$$(10.22) \quad -\frac{1}{2}\underline{\nu}_2\lambda^{1/3} < \underline{\delta}_2^{SW} < -\frac{1}{2}\underline{\nu}_2\lambda^{1/3}(1 - \xi\lambda^{1/3}),$$

$$(10.23) \quad \underline{G}_W(\vec{\delta}^{SW}) = 0, \quad \underline{G}_S(\vec{\delta}^{SW}) = 0.$$

We locate the northeast corner of \underline{M} by a similar argument. In particular, we use (10.17) and (10.18) to write

$$\underline{G}_N(\underline{\delta}_1, \underline{F}_E(\underline{\delta}_1)) = \underline{\mu}_2(\lambda - p\underline{\gamma}A^{1-p}\lambda^{5/3}) - \frac{h'_2(\underline{\delta}_2)}{\underline{\nu}_2} + O(\lambda^2),$$

we write $\underline{\delta}_2$ as

$$\underline{\delta}_2 = \frac{1}{2}\underline{\nu}_2\lambda^{1/3} \left(1 - \lambda^{1/3} \sqrt{\frac{2}{3}p\underline{\gamma}A^{1-p} + B + \eta} \right),$$

and we obtain $\underline{G}_N(\underline{\delta}_1, \underline{F}_E(\underline{\delta}_1)) = \frac{3}{2}\underline{\mu}_2\eta\lambda^{5/3} + O(\lambda^2)$ in place of (10.20). In conclusion, we find $\vec{\delta}^{NE} = (\underline{\delta}_1^{NE}, \underline{\delta}_2^{NE})$ such that

$$(10.24) \quad \frac{1}{2}\underline{\nu}_1\lambda^{1/3}(1 - \xi\lambda^{1/3}) < \underline{\delta}_1^{NE} < \frac{1}{2}\underline{\nu}_1\lambda^{1/3},$$

$$(10.25) \quad \frac{1}{2}\underline{\nu}_2\lambda^{1/3}(1 - \xi\lambda^{1/3}) < \underline{\delta}_2^{NE} < \frac{1}{2}\underline{\nu}_2\lambda^{1/3},$$

$$(10.26) \quad \underline{G}_E(\vec{\delta}^{NE}) = 0, \quad \underline{G}_N(\vec{\delta}^{NE}) = 0.$$

The points $\vec{\delta}^{SW}$ and $\vec{\delta}^{NE}$ are plotted in Figure 2, and the graphs of \underline{F}_W and \underline{F}_E are shown. For sufficiently small $\lambda > 0$, both the line connecting $-\vec{\theta}$ with $\vec{\delta}^{NE}$ and the line connecting $-\vec{\theta}$ with $\vec{\delta}^{SW}$ have positive slopes, and the slope of the former is strictly greater than the slope of the latter, as we now show. The first inequality in (2.4) implies that

$$\underline{\mu}_2\underline{\alpha}_1 = \left(-\frac{\rho\sigma_2}{\sigma_1}\mu_1 + \mu_2 \right) \alpha_1 > \mu_1 \left(-\frac{\rho\sigma_2}{\sigma_1}\alpha_1 + \alpha_2 \right) = \underline{\mu}_1\underline{\alpha}_2.$$

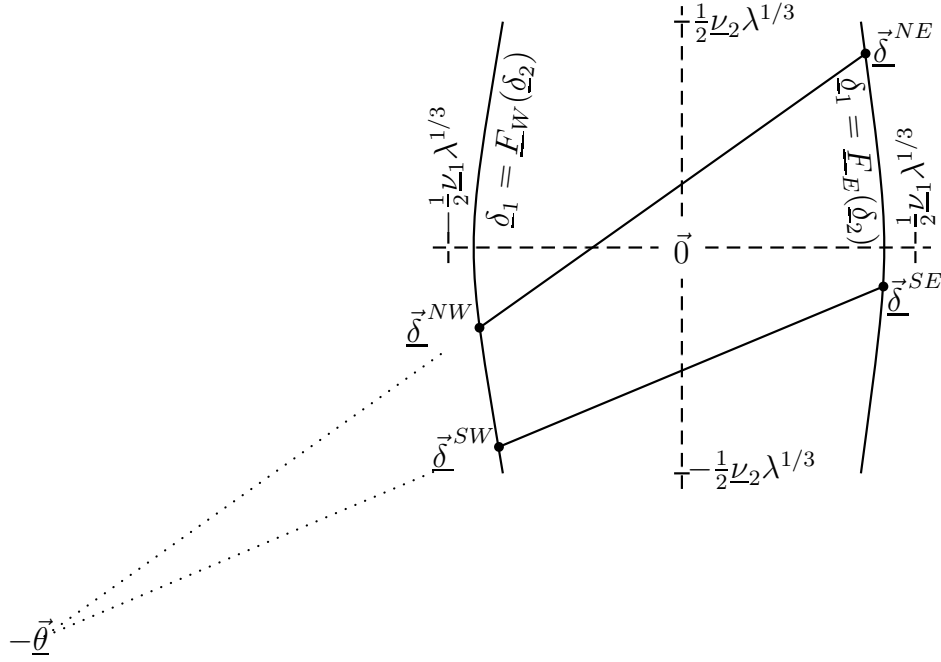


Figure 2. Construction of \underline{M} .

The second equation in (10.4) shows that $\underline{\alpha}_i = (1-p)\underline{\sigma}_i^2\underline{\theta}_i$, $i = 1, 2$, and hence $\underline{\sigma}_1^2\underline{\theta}_1\underline{\mu}_2 > \underline{\sigma}_2^2\underline{\theta}_2\underline{\mu}_1$. This is equivalent to $\underline{\theta}_1\sqrt[3]{\underline{\mu}_2\underline{\theta}_2^2/\underline{\sigma}_2^2} > \underline{\theta}_2\sqrt[3]{\underline{\mu}_1\underline{\theta}_1^2/\underline{\sigma}_1^2}$, which in turn implies $\underline{\theta}_1\underline{\nu}_2 > \underline{\theta}_2\underline{\nu}_1$. We conclude that

$$(10.27) \quad \frac{\underline{\theta}_2 + \frac{1}{2}\underline{\nu}_2\lambda^{1/3}}{\underline{\theta}_1 + \frac{1}{2}\underline{\nu}_1\lambda^{1/3}} > \frac{\underline{\theta}_2 - \frac{1}{2}\underline{\nu}_2\lambda^{1/3}}{\underline{\theta}_1 - \frac{1}{2}\underline{\nu}_1\lambda^{1/3}}.$$

Because of (10.21), (10.22), (10.24), and (10.25), the expressions on the left- and right-hand sides of (10.27) are the slopes, to order of accuracy $O(\lambda^{1/3})$, of the lines in question.

As indicated in Figure 2, the line passing through $-\underline{\theta}$ and $\underline{\delta}^{NE}$ intersects the graph of the continuous function \underline{F}_W . We select an intersection point and denote it by $\underline{\delta}^{NW}$. Similarly, the line passing through $-\underline{\theta}$ and $\underline{\delta}^{NW}$ intersects the graph of the continuous function \underline{F}_E . We select an intersection point and denote it by $\underline{\delta}^{SE}$.

Let us now change from the variables $\underline{\delta}$ to the variables \underline{z} by translating by $\underline{\theta}$, i.e., defining $\underline{z} = \underline{\delta} + \underline{\theta}$, so that $\underline{0}$ maps to $\underline{\theta}$ and $-\underline{\theta}$ maps to $\underline{0}$. The set \underline{M} is the image of the open bounded region in Figure 2 under this change of variables. Its corners are $\underline{\zeta}^{SW} = \underline{\delta}^{SW} + \underline{\theta}$, $\underline{\zeta}^{SE} = \underline{\delta}^{SE} + \underline{\theta}$, $\underline{\zeta}^{NE} = \underline{\delta}^{NE} + \underline{\theta}$, and $\underline{\zeta}^{NW} = \underline{\delta}^{NW} + \underline{\theta}$. It has two straight-line boundaries, both of which are segments of rays emanating from the origin. The other two boundaries are

the graphs of the functions

$$(10.28) \quad \underline{f}_W(\zeta_2) \triangleq \underline{F}_W(\zeta_2 - \theta_2) + \theta_1, \quad \zeta_2^{SW} \leq \zeta_2 \leq \zeta_2^{NW},$$

$$(10.29) \quad \underline{f}_E(\zeta_2) \triangleq \underline{F}_E(\zeta_2 - \theta_2) + \theta_1, \quad \zeta_2^{SE} \leq \zeta_2 \leq \zeta_2^{NE}.$$

Finally, we reverse the transformation used to obtain the auxiliary problem. In particular, with \mathbf{D} defined by (10.2), we make the change of variables

$$(10.30) \quad \vec{z} = \mathbf{D}^T \underline{z},$$

which we write componentwise as $(z_1, z_2) = (\underline{z}_1 - (\rho\sigma_2/\sigma_1)\underline{z}_2, \underline{z}_2)$ (see Remark 2.4). We set $M = \mathbf{D}^T \underline{M}$. Under this transformation, $\underline{\theta}$ maps to $\vec{\theta}$, which is assumed to be in the open first quadrant. Hence, for sufficiently small $\lambda > 0$, M is a subset of the open first quadrant. We denote by $\vec{\zeta}^{SW}$, $\vec{\zeta}^{SE}$, $\vec{\zeta}^{NE}$, and $\vec{\zeta}^{NW}$ the images under \mathbf{D}^T of $\underline{\zeta}^{SW}$, $\underline{\zeta}^{SE}$, $\underline{\zeta}^{NE}$, and $\underline{\zeta}^{NW}$, respectively. Since \mathbf{D}^T maps rays through the origin to rays through the origin, the boundary of M between $\vec{\zeta}^{SW}$ and $\vec{\zeta}^{SE}$ is a segment of a ray through the origin, as is the boundary of M between $\vec{\zeta}^{NW}$ and $\vec{\zeta}^{NE}$. Furthermore, $\zeta_2^{SW} = \underline{\zeta}_2^{SW} < \underline{\zeta}_2^{NW} = \zeta_2^{NW}$, and the boundary of M between $\vec{\zeta}^{SW}$ and $\vec{\zeta}^{NW}$ is the graph of the function

$$f_W(\zeta_2) \triangleq \underline{f}_W(\zeta_2) - \frac{\rho\sigma_2}{\sigma_1}\zeta_2, \quad \zeta_2^{SW} \leq \zeta_2 \leq \zeta_2^{NW}.$$

Similarly, the boundary of M between $\vec{\zeta}^{SE}$ and $\vec{\zeta}^{NE}$ is the graph of the function

$$f_E(\zeta_2) \triangleq \underline{f}_E(\zeta_2) - \frac{\rho\sigma_2}{\sigma_1}\zeta_2, \quad \zeta_2^{SE} \leq \zeta_2 \leq \zeta_2^{NE}.$$

We complete the partition of \mathcal{D}_2 into nine regions, as described in section 7 and shown in Figure 1. Note, however, that in contrast to Figure 1, here the region M has eastern and western boundaries that are nearly linear with negative slope rather than nearly constant and southern and northern boundaries that are segments of rays emanating from the origin. In particular, the hypothesis in Theorem 8.6 is satisfied that each point in \mathcal{D}_2 is the end point of a segment in the radial direction lying completely in the closure of one of the nine regions into which \mathcal{D}_2 has been partitioned. Because \underline{f}_W and \underline{f}_E satisfy analogues of (8.8) and (8.6), it is straightforward to verify that f_W and f_E also satisfy these inequalities.

10.4. Construction of a subsolution. Throughout this section, we use the change of variables (10.30) and the notation $\vec{z} = (z_1, z_2)$, $\underline{z} = (\underline{z}_1, \underline{z}_2)$. Define

$$(10.31) \quad \underline{w}(\underline{z}) = \frac{1}{p}A^{p-1} - \underline{\gamma}\lambda^{2/3} + K\lambda - \sum_{i=1}^2 \frac{1}{\underline{\nu}_i}A^{p-1}h_i(\underline{z}_i - \theta_i), \quad \underline{z} \in \overline{\underline{M}}.$$

According to Lemma 10.5, (10.28), and (10.29),

$$\underline{G}_W(\vec{z} - \vec{\theta}) = 0 \text{ for } \vec{z} \in \partial_W \underline{M}, \quad \underline{G}_E(\vec{z} - \vec{\theta}) = 0 \text{ for } \vec{z} \in \partial_E \underline{M}.$$

For $\vec{z} \in \underline{M}$,

$$\begin{aligned} \mathcal{B}_1 \underline{w}^M(\vec{z}) &= \lambda_1 p \underline{w}(\vec{z}) - (1 + \lambda_1 z_1) \underline{w}_1^M(\vec{z}) - \lambda_1 z_2 \underline{w}_2^M(\vec{z}) \\ &= A^{p-1} \left[\lambda_1 - p \underline{\gamma} A^{1-p} \lambda_1 \lambda^{2/3} + p K A^{1-p} \lambda_1 \lambda + \frac{h'_1(z_1 - \theta_1)}{\underline{\nu}_i} + \lambda_1 \underline{H}(\vec{z} - \vec{\theta}) \right] \\ &= A^{p-1} \underline{G}_W(\vec{z} - \vec{\theta}). \end{aligned}$$

Therefore,

$$(10.32) \quad \mathcal{B}_1 \underline{w}^M = 0 \text{ on } \partial_W \underline{M}.$$

Similarly, for $\vec{z} \in \underline{M}$,

$$\begin{aligned} \mathcal{S}_1 \underline{w}^M(\vec{z}) &= \lambda_1 p \underline{w}(\vec{z}) + (1 - \lambda z_1) \underline{w}_1^M(\vec{z}) - \lambda_1 z_2 \underline{w}_2^M(\vec{z}) \\ &= A^{p-1} \left[\lambda_1 - p \underline{\gamma} A^{1-p} \lambda_1 \lambda^{2/3} + p K A^{1-p} \lambda_1 \lambda - \frac{h'_1(z_1 - \theta_1)}{\underline{\nu}_1} + \lambda_1 \underline{H}(\vec{z} - \vec{\theta}) \right] \\ &= A^{p-1} \underline{G}_E(\vec{z} - \vec{\theta}). \end{aligned}$$

Therefore,

$$(10.33) \quad \mathcal{S}_1 \underline{w}^M = 0 \text{ on } \partial_E \underline{M}.$$

In addition, the second equation in (10.23) implies

$$(10.34) \quad \mathcal{B}_2 \underline{w}^M(\vec{\zeta}^{SW}) = A^{p-1} \left[\underline{G}_S(\vec{\zeta}^{SW} - \vec{\theta}) - \frac{\rho \sigma_2 h'_1(\zeta_1^{SW} - \theta_1)}{\sigma_1 \underline{\nu}_1} \right] = \frac{\rho \sigma_2}{\sigma_1} \underline{w}_1^M(\vec{\zeta}^{SW}).$$

Similarly, the second equation in (10.26) implies

$$(10.35) \quad \mathcal{S}_2 \underline{w}^M(\vec{\zeta}^{NE}) = A^{p-1} \left[\underline{G}_N(\vec{\zeta}^{NE} - \vec{\theta}) + \frac{\rho \sigma_2 h'_1(\zeta_1^{NE} - \theta_1)}{\sigma_1 \underline{\nu}_1} \right] = -\frac{\rho \sigma_2}{\sigma_1} \underline{w}_1^M(\vec{\zeta}^{NE}).$$

We recall (10.30) and define

$$(10.36) \quad w(\vec{z}) = \underline{w}((\mathbf{D}^T)^{-1} \vec{z}) = \underline{w} \left(z_1 + \frac{\rho \sigma_2}{\sigma_1} z_2, z_2 \right), \quad \vec{z} \in \overline{M},$$

so that

$$(10.37) \quad \nabla \underline{w}^M(\vec{z}) = \mathbf{D} \nabla w^M(\vec{z}), \quad \nabla^2 \underline{w}^M(\vec{z}) = \mathbf{D} \nabla^2 w^M(\vec{z}) \mathbf{D}^T.$$

It is straightforward to calculate from (10.32)–(10.35) that

$$(10.38) \quad \mathcal{B}_1 w^M = 0 \text{ on } \partial_W M, \quad \mathcal{S}_1 w^M = 0 \text{ on } \partial_E M,$$

$$(10.39) \quad \mathcal{B}_2 w^M(\vec{\zeta}^{SW}) = 0, \quad \mathcal{S}_2 w^M(\vec{\zeta}^{NE}) = 0.$$

We extend w to $\overline{\mathcal{D}}_2$ by Theorem 8.4. The extended function satisfies (8.14)–(8.19).

Lemma 10.6. *The function w defined on M by (10.36) and extended to $\underline{\mathcal{D}}_2$ by Theorem 8.4 is a subsolution of (6.15) on $\overline{\mathcal{D}}_2$.*

Proof. Let $\bar{z}^0 \in \mathcal{D}_2$ be given, and let φ be a C^2 function defined on \mathcal{D}_2 , agreeing with w at \bar{z}^0 , and dominating w on \mathcal{D}_2 . We must show that the expression in (6.16) is nonpositive.

Case I: $\bar{z}^0 \in M$. Using (10.3), (10.4), (10.30), and (10.37), it is straightforward to verify that $\bar{\alpha} \cdot \bar{z} = \underline{\alpha} \cdot \bar{z}$, $\mathbf{V}\bar{z} \cdot \bar{z} = \mathbf{V}\underline{z} \cdot \bar{z}$, $\nabla w^M(\bar{z}) \cdot \bar{z} = \nabla \underline{w}^M(\bar{z}) \cdot \bar{z}$, and $\nabla^2 w^M(\bar{z})\bar{z} \cdot \bar{z} = \nabla^2 \underline{w}^M(\bar{z})\bar{z} \cdot \bar{z}$. Therefore,

$$\begin{aligned} \mathcal{L}_2 w(\bar{z}) &= \left[\beta - rp - p\bar{\alpha} \cdot \bar{z} + \frac{1}{2}p(1-p)(\mathbf{V}\bar{z} \cdot \bar{z}) \right] w(\bar{z}) \\ &\quad + (r + \bar{\alpha} \cdot \bar{z} - (1-p)(\mathbf{V}\bar{z} \cdot \bar{z}))\nabla w(\bar{z}) \cdot \bar{z} - \frac{1}{2}(\mathbf{V}\bar{z} \cdot \bar{z})(\nabla^2 w(\bar{z})\bar{z} \cdot \bar{z}) \\ &= \left[\beta - rp - p\underline{\alpha} \cdot \bar{z} + \frac{1}{2}p(1-p)(\mathbf{V}\underline{z} \cdot \bar{z}) \right] \underline{w}(\bar{z}) \\ &\quad + (r + \underline{\alpha} \cdot \bar{z} - (1-p)(\mathbf{V}\underline{z} \cdot \bar{z}))\nabla \underline{w}(\bar{z}) \cdot \bar{z} - \frac{1}{2}(\mathbf{V}\underline{z} \cdot \bar{z})(\nabla^2 \underline{w}(\bar{z})\bar{z} \cdot \bar{z}). \end{aligned}$$

Similarly,

$$\tilde{U}(pw(\bar{z}) - \nabla w(\bar{z}) \cdot \bar{z}) = \tilde{U}(p\underline{w}(\bar{z}) - \nabla \underline{w}(\bar{z}) \cdot \bar{z}).$$

We apply Lemma 9.8 to the function \underline{w} to conclude that for sufficiently small (i.e., negative) K the inequality

$$(10.40) \quad \mathcal{L}_2 w - \tilde{U}(pw - \nabla w \cdot \bar{z}) \leq 0 \text{ on } M$$

holds. But w is C^2 in M and $\varphi - w$ attains its minimum over M at \bar{z}^0 . This implies that $\nabla w(\bar{z}^0) = \nabla \varphi(\bar{z}^0)$, and

$$(10.41) \quad \nabla^2 w(\bar{z}^0)\bar{z}^0 \cdot \bar{z}^0 \leq \nabla^2 \varphi(\bar{z}^0)\bar{z}^0 \cdot \bar{z}^0.$$

It follows from the definition of \mathcal{L}_2 and (10.40) that

$$\mathcal{L}_2 \varphi(\bar{z}^0) - \tilde{U}(p\varphi(\bar{z}^0) - \nabla \varphi(\bar{z}^0) \cdot \bar{z}^0) \leq \mathcal{L}_2 w(\bar{z}^0) - \tilde{U}(pw(\bar{z}^0) - \nabla w(\bar{z}^0) \cdot \bar{z}^0) \leq 0.$$

Case II: $\bar{z}^0 \in (\partial_N M \cup \partial_S M) \setminus \{\bar{\zeta}^{SW}, \bar{\zeta}^{SE}, \bar{\zeta}^{NE}, \bar{\zeta}^{NW}\}$. We use continuity to extend (10.40) to

$$(10.42) \quad \mathcal{L}_2 w^M - \tilde{U}(pw^M - \nabla w^M \cdot \bar{z}) \leq 0 \text{ on } \overline{M}.$$

We define $\partial_N^\circ M \triangleq \partial_N M \setminus \{\bar{\zeta}^{NW}, \bar{\zeta}^{NE}\}$ and $\partial_S^\circ M \triangleq \partial_S M \setminus \{\bar{\zeta}^{SW}, \bar{\zeta}^{SE}\}$. The function w might not be C^2 on $\partial_N^\circ M$ and $\partial_S^\circ M$, but because these boundaries are segments of rays emanating from the origin, w has two continuous derivatives in the radial direction on these boundaries. But $\nabla w(\bar{z}^0) \cdot \bar{z}^0$ and $\nabla^2 w(\bar{z}^0)\bar{z}^0 \cdot \bar{z}^0$ are the first and second derivatives of w in the radial direction, and because these are the only derivatives of w appearing on the left-hand side of (10.42), we may repeat the argument of Case I, using (10.42) in place of (10.40).

Case III: $\vec{z}^0 \in (\partial_W M \cup \partial_E M) \setminus \{\vec{\zeta}^{SW}, \vec{\zeta}^{SE}, \vec{\zeta}^{NE}, \vec{\zeta}^{NW}\}$. The derivative formulas (8.21) and (8.23) coupled with (10.38) imply that w is C^1 at \vec{z}^0 . Depending on the location of \vec{z}^0 , we obtain from (10.38) that either $\mathcal{B}_1 w(\vec{z}^0) = 0$ or $\mathcal{S}_1 w(\vec{z}^0) = 0$. By assumption on φ , $\nabla w(\vec{z}^0) = \nabla \varphi(\vec{z}^0)$, and hence either $\mathcal{B}_1 \varphi(\vec{z}^0) = 0$ or $\mathcal{S}_1 \varphi(\vec{z}^0) = 0$. In either case, (6.16) is nonpositive.

Case IV: $\vec{z}^0 \in \{\vec{\zeta}^{SW}, \vec{\zeta}^{NE}\}$. We treat the case $\vec{z}^0 = \vec{\zeta}^{SW}$ only. From the first equations in (10.38) and (10.39) and implication (8.27) in Theorem 8.4, we conclude that w is C^1 at $\vec{\zeta}^{SW}$. Hence, $\nabla \varphi(\vec{\zeta}^{SW}) = \nabla w(\vec{\zeta}^{SW})$, and the first equation in either (10.38) or (10.39) shows that (6.16) is nonpositive.

Case V: $\vec{z}^0 \in \{\vec{\zeta}^{SE}, \vec{\zeta}^{NW}\}$. We treat the case $\vec{\zeta}^{SE}$ only. We parameterize the line connecting $(1/\lambda_1, 0)$ with $\vec{\zeta}^{SE}$ by

$$\vec{z}(t) = (z_1(t), z_2(t)) = (1/\lambda_1 - t(1 - \lambda_1 \zeta_1^{SE}), t\lambda_1 \zeta_2^{SE})$$

so that $\vec{z}(1/\lambda_1) = \vec{\zeta}^{SE}$. The direction vector of this line $\vec{z}'(t) = (-1 + \lambda_1 \zeta_1^{SE}, \lambda_1 \zeta_2^{SE})$ approaches $(-1, 0)$ as $\lambda \downarrow 0$. On the other hand, equation (10.19) implies that

$$f'_E(\zeta_2) = \underline{f}_E(\zeta_2) - \frac{\rho\sigma_2}{\sigma_1} = -\frac{\rho\sigma_2}{\sigma_1} + O(\lambda^{1/3})$$

is bounded away from $\pm\infty$, so tangents to the boundary $\partial_E M$ are bounded away from horizontal. Likewise, the direction vector of $\partial_S M$, which is $\vec{\zeta}^{SE}$, converges to $\vec{\theta}$ as $\lambda \downarrow 0$, and hence this is also bounded away from horizontal. It follows that for sufficiently small $\lambda > 0$, the line $\vec{z}(t)$ crosses from \overline{E} into M at $t = 1/\lambda_1$.

The directional derivative of w in the $\vec{z}'(t)$ direction, computed at $\vec{\zeta}^{SE}$ from inside M , is $(-1 + \lambda_1 \zeta_1^{SE})w_1^M(\vec{\zeta}^{SE}) + \lambda_1 \zeta_2^{SE} w_2^M(\vec{\zeta}^{SE})$, and because of (10.33), this is equal to $\lambda_1 p w(\vec{\zeta}^{SE})$. According to (8.13), this directional derivative computed from inside E is

$$\frac{d}{dt} \left(\frac{z_2(t)}{\zeta_2^{SE}} \right)^p \Big|_{t=1/\lambda_1} w^M(\vec{\zeta}^{SE}) = p\lambda_1 w^M(\vec{\zeta}^{SE}),$$

and hence the directional derivative is continuous across the boundary point $\vec{\zeta}^{SE}$, i.e., $\vec{z}'(1/\lambda_1) \cdot \nabla w^M(\vec{\zeta}^{SE}) = \vec{z}'(1/\lambda_1) \cdot \nabla w^E(\vec{\zeta}^{SE})$. The function $\varphi(\vec{z}(t)) - w(\vec{z}(t))$ is minimized at $t = 1/\lambda_1$, and hence

$$\vec{z}'(1/\lambda_1) \nabla \varphi(\vec{\zeta}^{SE}) = \vec{z}'(1/\lambda_1) \nabla w^M(\vec{\zeta}^{SE}) = p\lambda_1 w(\vec{\zeta}^{SE}) = p\lambda_1 \varphi(\vec{\zeta}^{SE}).$$

This equation shows that $\mathcal{S}_1 \varphi(\vec{\zeta}^{SE}) = 0$, and hence (6.16) is nonpositive at $\vec{\zeta}^{SE}$.

Case VI: $\vec{z}^0 \in (\overline{E} \cap \mathcal{D}_2) \setminus \partial_E M$. Equation (8.14) implies that $\mathcal{S}_1 w^E(\vec{z}^0) = 0$. If $\vec{z}^0 \in E \cap \mathcal{D}_2$, where w is C^1 and hence $\nabla \varphi(\vec{z}^0) = \nabla w(\vec{z}^0)$, this implies that (6.16) is nonpositive. Suppose $\vec{z}^0 \in \partial E \cap \mathcal{D}_2$. We have already dealt with $\partial_E M$, so we need only to treat the case that \vec{z}^0 lies in the interior of the line segment connecting $(0, 1/\lambda_2)$ with $\vec{\zeta}^{SE}$ or with $\vec{\zeta}^{NE}$. We treated the former case in Case V, there showing that (6.16) is nonpositive at $\vec{\zeta}^{SE}$. The proof at \vec{z}^0 is simpler because it is not necessary to prove that the directional derivative is continuous at \vec{z}^0 . The proof if \vec{z}^0 lies in the interior of the line segment connecting $(0, 1/\lambda_2)$ with $\vec{\zeta}^{NE}$ is the same.

Case VII: $\bar{z}^0 \in (\bar{S} \cap \mathcal{D}_2) \setminus \partial_S M$. This proceeds as in Case VI.

Case VIII: $\bar{z}^0 \in SE$. Starting with the equation

$$(10.43) \quad w(\bar{z}) = \left(\frac{1 - \lambda_1 z_1 + \lambda_2 z_2}{1 - \lambda_1 \zeta_1^{SE} + \lambda_2 \zeta_2^{SE}} \right)^p w(\bar{\zeta}^{SE})$$

obtained from (8.13), we compute for $\bar{z}^0 \in SE$ that

$$(10.44) \quad \mathcal{B}_2 w(\bar{z}^0) = \lambda_2 p w(\bar{z}^0) - \lambda_2 z_1^0 w_1(\bar{z}^0) - (1 + \lambda_2 z_2^0) w_2(\bar{z}^0) = 0.$$

From (10.43) we see that w is C^1 in SE , and hence $\nabla \varphi(\bar{z}^0) = w(\bar{z}^0)$. It follows that $\mathcal{B}_2 \varphi(\bar{z}^0) = 0$, and hence the expression in (6.16) is nonpositive. ■

Appendix A. Continuity of v . This appendix proves the continuity of v claimed in Theorem 5.1 For this proof, we need the following lower bound.

Lemma A.1. *The constant $C_* \triangleq \frac{\beta - rp}{1-p}$ is positive, and*

$$(A.1) \quad v(y_1, y_2, x) \geq \frac{1}{p} C_*^{p-1} (x - \lambda_1 |y_1| - \lambda_2 |y_2|)^p, \quad (y_1, y_2, x) \in \bar{\mathcal{D}}_3.$$

Proof. Positivity of C_* follows from the definition (2.11) of A , Assumption 2.3 that $A > 0$, and the positive semidefiniteness of \mathbf{V}^{-1} .

In light of (2.9), to prove (A.1) we need only to consider $(y_1, y_2, x) \in \mathcal{D}_3$. For initial condition $(y_1, y_2, x) \in \mathcal{D}_3$, we consider the policy that liquidates the futures positions at time zero, so that $X(0) = x - \lambda_1 |y_1| - \lambda_2 |y_2| > 0$, invests solely in the money market account thereafter, and consumes at a constant proportional rate $c > 0$. Then $X(t) = X(0)e^{(r-c)t}$ for all $t \geq 0$. The resulting expected utility is $\frac{X^p(0)c^p}{p} \int_0^\infty e^{-(\beta - rp + cp)t} dt$. We obtain the lower bound

$$v(y_1, y_2, x) \geq \frac{c^p}{p(\beta - rp + cp)} (x - \lambda_1 |y_1| - \lambda_2 |y_2|)^p,$$

valid for all $c > 0$. Maximizing the right-hand side over $c > 0$, we obtain (A.1). ■

Theorem A.2 (continuity of v). *When $\lambda > 0$, the value function v is continuous on $\bar{\mathcal{D}}_3$.*

Proof. A concave function defined on a convex set is continuous on the interior of its domain. Thus, because of Remark 2.1, v is continuous on \mathcal{D}_3 . It remains to prove the continuity of v at the boundary of \mathcal{D}_3 , where it is equal to zero.

The boundary of \mathcal{D}_3 has four faces, which we denote by

$$F_{\pm\pm} \triangleq \{(y_1, y_2, x) : \lambda_1 |y_1| + \lambda_2 |y_2| = x, \pm y_1 \geq 0, \pm y_2 \geq 0\}.$$

These faces intersect at the origin, where the continuity of v follows from the bound

$$(A.2) \quad 0 \leq v(y_1, y_2, x) \leq v_0(x)$$

for all $(y_1, y_2, x) \in \bar{\mathcal{D}}_3$, and $v_0(x) = \frac{1}{p} A^{p-1} x^p$, introduced in section 3, is the value function corresponding to $\lambda = 0$. We prove the continuity on $F_{++}^\times \triangleq F_{++} \setminus \{(0, 0, 0)\}$. The proof of the continuity on the other three faces is analogous.

Let δ be a number in $(0, 1)$ to be chosen later (see (A.6) below), and define

$$H_\delta \triangleq \{(\vec{y}, x) \in \mathcal{D}_3 : (1 - \delta)x < \vec{\lambda} \cdot \vec{y} < x\}.$$

The boundary of H_δ comprises four parts: (1) the origin, (2) the set F_{++}^\times , (3) the ‘‘inner boundary’’

$$\partial_i H_\delta \triangleq \{(\vec{y}, x) \in \mathcal{D}_3 : (1 - \delta)x = \vec{\lambda} \cdot \vec{y}\},$$

and the ‘‘side boundary’’

$$\partial_s H_\delta \triangleq \{(\vec{y}, x) \in \partial \mathcal{D}_3 : (1 - \delta)x \leq \vec{\lambda} \cdot \vec{y} \leq x, \lambda_1|y_1| + \lambda_2|y_2| = x, y_1 y_2 < 0\}.$$

We define the constant

$$M_\delta \triangleq \sup \{v(y_1, y_2, 1) : (y_1, y_2, 1) \in \partial_i H_\delta\},$$

which is finite because of the upper bound (A.2). Lemma A.1 implies that $M_\delta \geq \frac{1}{p} C_*^{p-1} \delta^p$ (set $y_1 = (1 - \delta)/\lambda_1$, $y_2 = 0$ so that $(1, y_1, y_2) \in \partial_i H_\delta$). We next define A_δ by the equation $\frac{1}{p} A_\delta^{p-1} \delta^p = M_\delta$ so that $0 < A_\delta \leq C_*$. Finally, we set

$$(A.3) \quad \psi(y_1, y_2, x) = \begin{cases} \frac{1}{p} A_\delta^{p-1} (x - \lambda_1 y_1 - \lambda_2 y_2)^p & \text{if } (y_1, y_2, x) \in \overline{H}_\delta, \\ v(y_1, y_2, x) & \text{if } (y_1, y_2, x) \in \overline{\mathcal{D}}_3 \setminus \overline{H}_\delta \end{cases}$$

and subsequently show that ψ dominates v on H_δ . From inside \mathcal{D}_3 , F_{++}^\times can be approached only through H_δ . Because ψ is nonnegative and continuous on H_δ , and ψ has limit zero as F_{++}^\times is approached, this will establish the continuity of v on F_{++}^\times .

On $F_{++} \cup \partial_s H_\delta$, v is zero and ψ is nonnegative. On the remaining part of the boundary of H_δ , namely $\partial_i H_\delta$,

$$\psi(y_1, y_2, x) = \frac{1}{p} A_\delta^{p-1} \delta^p x^p = M_\delta x^p \geq x^p v\left(\frac{y_1}{x}, \frac{y_2}{x}, 1\right) = v(y_1, y_2, x),$$

where we have used (6.1) in the last step. Consequently, $v \geq \psi$ on ∂H_δ , and indeed,

$$(A.4) \quad \psi \geq v \text{ on } \overline{\mathcal{D}}_3 \setminus H_\delta.$$

Using the notation (5.1) and (5.2), we observe that in H_δ

$$(A.5) \quad \begin{aligned} & \mathcal{L}_3 \psi(y_1, y_2, x) - \tilde{U}(\psi_x(y_1, y_2, x)) \\ &= A_\delta^{p-1} (x - \vec{\lambda} \cdot \vec{y})^{p-2} \left[\frac{\beta}{p} (x - \vec{\lambda} \cdot \vec{y})^2 - (rx + \vec{\alpha} \cdot \vec{y})(x - \vec{\lambda} \cdot \vec{y}) \right. \\ & \quad \left. + \frac{1}{2} (1 - p) \mathbf{V} \vec{y} \cdot \vec{y} - \frac{1-p}{p} A_\delta (x - \vec{\lambda} \cdot \vec{y})^2 \right] \\ & \geq A_\delta^{p-1} (x - \vec{\lambda} \cdot \vec{y})^{p-2} \\ & \quad \times \left[r(x - \vec{\lambda} \cdot \vec{y})^2 - (rx + \vec{\alpha} \cdot \vec{y})(x - \vec{\lambda} \cdot \vec{y}) + \frac{1}{2} (1 - p) \mathbf{V} \vec{y} \cdot \vec{y} \right], \end{aligned}$$

where we have used the inequality $A_\delta \leq C_*$ in the last step. We bound the right-hand side of (A.5) from below. If $(y_1, y_2, x) \in H_\delta$ and $y_1 < 0$, then $y_2 > 0$. In this case, the inequality $\lambda_1|y_1| + \lambda_2|y_2| < x$ defining \mathcal{D}_3 and the inequality $\vec{\lambda} \cdot \vec{y} > (1 - \delta)x$ defining H_δ imply $\lambda_1|y_1| < \frac{1}{2}\delta x$. Similarly, if $(y_1, y_2, x) \in H_\delta$ and $y_2 < 0$, then $\lambda_2|y_2| < \frac{1}{2}\delta x$. If $y_1 \geq 0$, then regardless of the sign of y_2 , $(y_1, y_2, x) \in H_\delta$ implies $\lambda_1 y_1 \leq x - \lambda_2 y_2 \leq (1 + \frac{1}{2}\delta)x$. The same result holds for y_2 . In other words,

$$\max\{\lambda|y_1|, \lambda_2|y_2|\} \leq \left(1 + \frac{1}{2}\delta\right)x \leq 2x, \quad (y_1, y_2, x) \in H_\delta.$$

Therefore, for $(\vec{y}, x) \in H_\delta$,

$$rx + \vec{\alpha} \cdot \vec{y} \leq rx + \max\left[\frac{\alpha_1}{\lambda_1}, \frac{\alpha_2}{\lambda_2}\right](\lambda_1|y_1| + \lambda_2|y_2|) \leq \left(r + 4 \max\left[\frac{\alpha_1}{\lambda_1}, \frac{\alpha_2}{\lambda_2}\right]\right)x$$

and

$$(rx + \vec{\alpha} \cdot \vec{y})(x - \vec{\lambda} \cdot \vec{y}) \leq \delta \left(r + 4 \max\left[\frac{\alpha_1}{\lambda_1}, \frac{\alpha_2}{\lambda_2}\right]\right)x^2.$$

Let e_2 denote the smallest eigenvalue of \mathbf{V} , which is positive. For $(\vec{y}, x) \in H_\delta$,

$$\begin{aligned} \mathbf{V}\vec{y} \cdot \vec{y} &\geq e_2 \vec{y} \cdot \vec{y} \\ &\geq e_2 \min\left[\frac{1}{\lambda_1^2}, \frac{1}{\lambda_2^2}\right](\lambda_1^2 y_1^2 + \lambda_2^2 y_2^2) \\ &\geq \frac{1}{2}e_2 \min\left[\frac{1}{\lambda_1^2}, \frac{1}{\lambda_2^2}\right](\vec{\lambda} \cdot \vec{y})^2 \\ &\geq \frac{1}{2}e_2 \min\left[\frac{1}{\lambda_1^2}, \frac{1}{\lambda_2^2}\right](1 - \delta)^2 x^2. \end{aligned}$$

Therefore, the term in square brackets in the last line of (A.5) is bounded below by $(-\kappa_1\delta + \kappa_2(1 - \delta)^2)x^2$, where the positive constants κ_1 and κ_2 do not depend on δ . We choose $\delta \in (0, 1)$ so that this expression is positive, i.e., so that

$$(A.6) \quad \mathcal{L}\psi(y_1, y_2, x) - \tilde{U}(\psi_x(y_1, y_2, x)) \geq 0, \quad (y_1, y_2, x) \in H_\delta.$$

In the HJB equation (5.3) there are four first-order terms corresponding to the four possible “pure” trades, trades that involve only one risky asset and the money market account. Suppose that starting from $(y_1, y_2, x) \in H_\delta$, the agent increases or decreases her position in type-one futures by the amount $\eta_1 \geq 0$ and increases or decreases her position in type-two futures by the amount $\eta_2 \geq 0$. Then her position in the three assets changes to $(y_1 \pm \eta_1, y_2 \pm \eta_2, x - \lambda_1 \eta_1 - \lambda_2 \eta_2)$. Because

$$(x - \lambda_1 \eta_1 - \lambda_2 \eta_2) - \lambda_1(y_1 \pm \eta_1) - \lambda_2(y_2 \pm \eta_2) \leq x - \lambda_1 y_1 - \lambda_2 y_2,$$

the change in ψ satisfies

$$(A.7) \quad \begin{aligned} \psi(y_1 \pm \eta_1, y_2 \pm \eta_2, x - \lambda_1 \eta_1 - \lambda_2 \eta_2) - \psi(y_1, y_2, x) &\leq 0, \\ (y_1, y_2, x) \in H_\delta, (y_1 \pm \eta_1, y_2 \pm \eta_2, x - \lambda_1 \eta_1 - \lambda_2 \eta_2) &\in \overline{H}_\delta. \end{aligned}$$

This shows that at all points in H_δ , the derivatives of ψ in the directions in which the agent can trade are nonpositive.

The value function v has a property like (A.7), as discussed at the beginning of section 5. Specifically, for $\eta_1 \geq 0$, $\eta_2 \geq 0$, v satisfies

$$\begin{aligned} v(y_1 \pm \eta_1, y_2 \pm \eta_2, x - \lambda_1 \eta_1 - \lambda_2 \eta_2) - v(y_1, y_2, x) &\leq 0, \\ (y_1, y_2, x) \in \overline{\mathcal{D}}_3, (y_1 \pm \eta_1, y_2 \pm \eta_2, x - \lambda_1 \eta_1 - \lambda_2 \eta_2) &\in \overline{\mathcal{D}}_3. \end{aligned}$$

Combining this with (A.7), using the definition (A.3) of ψ outside \overline{H}_δ and the inequality $\psi \geq v$ on ∂H_δ (see (A.4)), we conclude that (A.7) can be strengthened to

$$(A.8) \quad \begin{aligned} \psi(y_1 \pm \eta_1, y_2 \pm \eta_2, x - \lambda_1 \eta_1 - \lambda_2 \eta_2) - \psi(y_1, y_2, x) &\leq 0, \\ (y_1, y_2, x) \in H_\delta, (y_1 \pm \eta_1, y_2 \pm \eta_2, x - \lambda_1 \eta_1 - \lambda_2 \eta_2) &\in \overline{\mathcal{D}}_3. \end{aligned}$$

For each positive integer $n > \frac{1}{\delta}$, define

$$K_n \triangleq \left\{ (\vec{y}, x) \in \mathcal{D}_3 : \left(1 - \delta + \frac{1}{n}\right) x \leq \vec{\lambda} \cdot \vec{y} \leq x \left(1 - \frac{1}{n}\right), \frac{1}{n} < x < n \right\},$$

a compact subset of H_δ . Let (L_1, L_2, M_1, M_2, C) be an admissible policy with initial condition $(y_1, y_2, x) \in K_n$, let Y_1, Y_2 , and X be given by (2.2) and (2.3), and define

$$\tau_n = \inf \{t \geq 0 : (\vec{Y}(t), X(t)) \notin K_n\}.$$

Itô's formula and inequality (A.8) imply

$$\begin{aligned} &e^{-\beta(t \wedge \tau_n)} \psi(\vec{Y}(t \wedge \tau_n), X(t \wedge \tau_n)) \\ &= \psi(y_1, y_2, x) \\ &+ \int_0^{t \wedge \tau_n} e^{-\beta s} \left[-\mathcal{L}\psi(\vec{Y}(s), X(s)) ds - C(s) \psi_x(\vec{Y}(s), X(s)) ds \right. \\ &\quad + \psi_x(\vec{Y}(s), X(s)) \Sigma \vec{Y}(s) \cdot d\vec{B}(s) \\ &\quad + \sum_{i=1}^2 \left(\psi_i(\vec{Y}(s), X(s)) - \lambda_i \psi_x(\vec{Y}(s), X(s)) \right) dL_i^c(s) \\ &\quad \left. + \sum_{i=1}^2 \left(-\psi_i(\vec{Y}(s), X(s)) - \lambda_i \psi_x(\vec{Y}(s), X(s)) \right) dM_i^c(s) \right] \\ &+ \sum_{0 \leq s \leq t \wedge \tau_n} e^{-\beta s} \left[\psi(\vec{Y}(s) + \Delta \vec{L}(s) - \Delta M(s), X(s) - \vec{\lambda} \cdot \Delta \vec{L}(s) - \vec{\lambda} \cdot \Delta M(s)) \right. \\ &\quad \left. - \psi(\vec{Y}(s), X(s)) \right] \\ &\leq \psi(y_1, y_2, x) + \int_0^{t \wedge \tau_n} e^{-\beta s} \left[-\mathcal{L}\psi(\vec{Y}(s), X(s)) + \tilde{U}(\psi_x(\vec{Y}(s), X(s))) \right] ds \\ &\quad - \int_0^{t \wedge \tau_n} e^{-\beta s} U(C(s)) ds + \int_0^{t \wedge \tau_n} e^{-\beta s} \psi_x(\vec{Y}(s), X(s)) \Sigma \vec{Y}(s) \cdot d\vec{B}(s). \end{aligned}$$

Taking expectations and using (A.6) we obtain

$$(A.9) \quad \mathbb{E} \left[e^{-\beta(t \wedge \tau_n)} \psi(\vec{Y}(t \wedge \tau_n), X(t \wedge \tau_n)) \right] \leq \psi(y_1, y_2, x) - \mathbb{E} \int_0^{t \wedge \tau_n} e^{-\beta s} U(C(s)) ds.$$

Letting $t \rightarrow \infty$, using Fatou's lemma on the left-hand side and the monotone convergence theorem on the right-hand side of (A.9), we see that

$$(A.10) \quad \mathbb{E} \left[e^{-\beta \tau_n} \mathbb{I}_{\{\tau_n < \infty\}} \psi(\vec{Y}(\tau_n), X(\tau_n)) \right] \leq \psi(y_1, y_2, x) - \mathbb{E} \int_0^{\tau_n} e^{-\beta s} U(C(s)) ds.$$

As $n \rightarrow \infty$, τ_n converges up to a limit, which we call τ_∞ . Fatou's lemma and the monotone convergence theorem applied to (A.10) result in

$$\mathbb{E} \left[e^{-\beta \tau_\infty} \mathbb{I}_{\{\tau_\infty < \infty\}} \psi(\vec{Y}(\tau_\infty)-, X(\tau_\infty)-) \right] \leq \psi(y_1, y_2, x) - \mathbb{E} \int_0^{\tau_\infty} e^{-\beta s} U(C(s)) ds.$$

But on the set $\{\tau_\infty < \infty\}$, $(Y(\tau_\infty)-, X(\tau_\infty)-)$ is in ∂H_δ , where ψ dominates v (see (A.4)). Furthermore, v does not increase along jumps caused by trades. Therefore,

$$\begin{aligned} \mathbb{E} \left[e^{-\beta \tau_\infty} \mathbb{I}_{\{\tau_\infty < \infty\}} v(\vec{Y}(\tau_\infty), X(\tau_\infty)) \right] &\leq \mathbb{E} \left[e^{-\beta \tau_\infty} \mathbb{I}_{\{\tau_\infty < \infty\}} v(\vec{Y}(\tau_\infty)-, X(\tau_\infty)-) \right] \\ &\leq \mathbb{E} \left[e^{-\beta \tau_\infty} \mathbb{I}_{\{\tau_\infty < \infty\}} \psi(\vec{Y}(\tau_\infty)-, X(\tau_\infty)-) \right]. \end{aligned}$$

By the dynamic programming principle, $v(y_1, y_2, x)$ is the supremum over admissible policies of $\mathbb{E}[e^{-\beta \tau_\infty} \mathbb{I}_{\{\tau_\infty < \infty\}} v(\vec{Y}(\tau_\infty), X(\tau_\infty))] + \mathbb{E} \int_0^{\tau_\infty} e^{-\beta s} U(C(s)) ds$, and we have just shown that this is dominated by $\psi(y_1, y_2, x)$.

Having thus shown that $v \leq \psi$ on K_n , we use the observation that $\cup_{n=1}^\infty K_n = H_\delta$ to assert that $v \leq \psi$ on H_δ . This is what we set out to prove. ■

Appendix B. Comparison theorem. This appendix provides the proof of Theorem 6.3. We follow the proof of Theorem 3.3 of [7], modifying it in order to account for the fact that \tilde{U} given by (5.2) has limit infinity as its argument goes down to zero, and hence condition (3.14) in [7] is not satisfied by the function H given by (B.1) below.

We have defined the concept of a viscosity solution of the two-variable HJB equation (6.15) in Definition 6.1. An equivalent definition is used in [7], which we state as Theorem B.1 below. For this we need some notation.

We denote by \mathfrak{S} the set of symmetric 2×2 matrices. We denote a generic element of $\mathcal{D}_2 \times \mathbb{R} \times \mathbb{R}^2 \times \mathfrak{S} \rightarrow \mathbb{R}$ by $(\vec{z}, s, \vec{q}, \mathbf{X})$, and we define functions S_i and B_i for $i = 1, 2$ and H_i for $i = 1, 2, 3, 4, 5$ on this set by the formulas (where we omit variables not involved in the

definition of the function)

$$\begin{aligned}
B_i(\vec{z}, s, \vec{q}) &= \lambda_i p s - q_i - \lambda_i \vec{q} \cdot \vec{z}, \quad i = 1, 2, \\
S_i(\vec{z}, s, \vec{q}) &= \lambda_i p s + q_i - \lambda_i \vec{q} \cdot \vec{z}, \quad i = 1, 2, \\
H_1(\vec{z}, s) &= (1 - p) \left[A + \frac{1}{2} p (\mathbf{V}(\vec{z} - \vec{\theta}) \cdot (\vec{z} - \vec{\theta})) \right] s, \\
H_2(\vec{z}, \vec{q}) &= (r + \alpha \cdot \vec{z} - (1 - p)(\mathbf{V} \vec{z} \cdot \vec{z}) \vec{q} \cdot \vec{z}, \\
H_3(\vec{z}, \mathbf{X}) &= -\frac{1}{2} (\mathbf{V} \vec{z} \cdot \vec{z}) (\mathbf{X} \vec{z} \cdot \vec{z}), \\
H_4(\vec{z}, s, \vec{q}) &= \begin{cases} -\tilde{U}(ps - \vec{q} \cdot \vec{z}) & \text{if } ps - \vec{q} \cdot \vec{z} > 0, \\ \infty & \text{otherwise,} \end{cases} \\
H_5(\vec{z}, s, \vec{q}, \mathbf{X}) &= H_1(\vec{z}, s) + H_2(\vec{z}, \vec{q}) + H_3(\vec{z}, \mathbf{X}) + H_4(\vec{z}, s, \vec{q}).
\end{aligned}$$

Finally, we set

$$(B.1) \quad H(\vec{z}, s, \vec{q}, \mathbf{X}) \triangleq \min [H_5(\vec{z}, s, \vec{q}, \mathbf{X}), B_1(\vec{z}, s, \vec{q}), B_2(\vec{z}, s, \vec{q}), S_1(\vec{z}, s, \vec{q}), S_2(\vec{z}, s, \vec{q})].$$

Using (2.12), we see that the expression (6.16) is $H(\vec{z}, \varphi(\vec{z}), \nabla \varphi(\vec{z}), \nabla^2 \varphi(\vec{z}))$.

Let w be a continuous function defined on \mathcal{D}_2 , and let \vec{z}^* be an element of \mathcal{D}_2 . We say that $(\vec{q}, \mathbf{X}) \in \mathbb{R}^2 \times \mathfrak{S}$ is a *second-order superjet* of w at \vec{z}^* if

$$(B.2) \quad w(\vec{z}) \leq w(\vec{z}^*) + \vec{q} \cdot (\vec{z} - \vec{z}^*) + \frac{1}{2} \mathbf{X}(\vec{z} - \vec{z}^*) \cdot (\vec{z} - \vec{z}^*) + o(\|\vec{z} - \vec{z}^*\|)$$

for all \vec{z} in an open neighborhood of \vec{z}^* , and we say (\vec{q}, \mathbf{X}) is a *second-order subjet* if the opposite of inequality (B.2) holds in an open neighborhood of \vec{z}^* . We denote the set of all second-order superjets (respectively, subjets) of w at \vec{z}^* by $J^{2,+}w(\vec{z}^*)$ (respectively, $J^{2,-}w(\vec{z}^*)$). As observed in [7],

$$(B.3) \quad \begin{aligned} J^{2,+}w(\vec{z}^*) &= \{(\nabla \varphi(\vec{z}^*), \nabla^2 \varphi(\vec{z}^*)) : \varphi \in C^2(\mathcal{D}_2), \varphi(\vec{z}^*) = w(\vec{z}^*), \varphi \geq w \text{ on } \mathcal{D}_2\}, \\ J^{2,-}w(\vec{z}^*) &= \{(\nabla \varphi(\vec{z}^*), \nabla^2 \varphi(\vec{z}^*)) : \varphi \in C^2(\mathcal{D}_2), \varphi(\vec{z}^*) = w(\vec{z}^*), \varphi \leq w \text{ on } \mathcal{D}_2\}. \end{aligned}$$

We further define

$$\begin{aligned} \overline{J}^{2,\pm}w(\vec{z}^*) &\triangleq \{(\vec{q}, \mathbf{X}) \in \mathbb{R}^2 \times \mathfrak{S} : \exists (\vec{z}_n, \vec{q}_n, \mathbf{X}_n) \in \mathcal{D}_2 \times \mathbb{R}^2 \times \mathfrak{S} \text{ such that} \\ &\quad (\vec{q}_n, \mathbf{X}_n) \in J^{2,\pm}w(\vec{z}_n) \text{ and } (\vec{z}_n, \vec{q}_n, \mathbf{X}_n) \rightarrow (\vec{z}^*, \vec{q}, \mathbf{X})\}. \end{aligned}$$

The conditions of the following theorem are taken as the definition of viscosity subsolution and supersolution in [7], and Remark 2.3 of [7] shows that these conditions are equivalent to Definition 6.1 used in this paper.

Theorem B.1. *A continuous function w defined on $\overline{\mathcal{D}_2}$ is a viscosity subsolution of (6.15) if and only if $H(\vec{z}^*, w(\vec{z}^*), \vec{q}, \mathbf{X}) \leq 0$ for every $\vec{z}^* \in \mathcal{D}_2$ and every $(\vec{q}, \mathbf{X}) \in J^{2,+}w(\vec{z}^*)$. The function w is a viscosity supersolution of (6.15) if and only if $H(\vec{z}^*, w(\vec{z}^*), \vec{q}, \mathbf{X}) \geq 0$ for every $\vec{z}^* \in \mathcal{D}_2$ and every $(\vec{q}, \mathbf{X}) \in J^{2,-}w(\vec{z}^*)$.*

Proof of Theorem 6.3. We define the symmetric positive semidefinite matrix

$$\mathbf{\Gamma}(\vec{z}) = \frac{1}{2}(\mathbf{V}\vec{z} \cdot \vec{z}) \begin{bmatrix} z_1^2 & z_1 z_2 \\ z_1 z_2 & z_2^2 \end{bmatrix}$$

and

$$\mathbf{\Gamma}^{\frac{1}{2}}(\vec{z}) = \begin{cases} \sqrt{\frac{\mathbf{V}\vec{z} \cdot \vec{z}}{2\vec{z} \cdot \vec{z}}} \begin{bmatrix} z_1^2 & z_1 z_2 \\ z_1 z_2 & z_2^2 \end{bmatrix} & \text{if } \vec{z} \neq 0, \\ \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} & \text{if } \vec{z} = 0 \end{cases}$$

so that $(\mathbf{\Gamma}^{\frac{1}{2}}(\vec{z}))^2 = \mathbf{\Gamma}(\vec{z})$ and $H_3(\vec{z}, \mathbf{X}) = -\frac{1}{2}\text{trace}(\mathbf{\Gamma}(\vec{z})\mathbf{X})$. Because $(\mathbf{V}\vec{z} \cdot \vec{z})/(\vec{z} \cdot \vec{z})$ is bounded from above and away from zero and because both $z_i^2(\partial/\partial z_j)\sqrt{(\mathbf{V}\vec{z} \cdot \vec{z})/(\vec{z} \cdot \vec{z})}$ and $z_i z_j(\partial/\partial z_j)\sqrt{(\mathbf{V}\vec{z} \cdot \vec{z})/(\vec{z} \cdot \vec{z})}$ are bounded on $\overline{\mathcal{D}}_2$ for $i, j \in \{1, 2\}$, $\mathbf{\Gamma}^{\frac{1}{2}}$ is Lipschitz on $\overline{\mathcal{D}}_2$.

For $\delta > 0$, define $\mathcal{O}_\delta \triangleq \{(\vec{z}, s, \vec{q}) \in \mathcal{D}_2 \times [0, \infty) \times \mathbb{R}^2 : ps - \vec{q} \cdot \vec{z} \geq \delta\}$ and $\mathcal{O} \triangleq \cup_{\delta > 0} \mathcal{O}_\delta$. On each \mathcal{O}_δ , H_4 is Lipschitz continuous.

We prove the comparison $w^+ \geq w^-$ on \mathcal{D}_2 by contradiction. Recall that $\overline{\mathcal{D}}_2$ is compact, w is continuous, and on $\partial\mathcal{D}_2$, $w^+ \geq w^- = 0$. Assume that $w^+ \geq w^-$ does not hold everywhere on $\overline{\mathcal{D}}_2$. Let $\vec{z}^0 \in \mathcal{D}_2$ be a point where $\delta_1 > 0$, the maximum of $w^- - w^+$ over $\overline{\mathcal{D}}_2$, is achieved. We use the variable doubling technique, penalizing the doubling. In particular, we maximize $w^-(\vec{z}) - w^+(\vec{\zeta}) - \frac{\eta}{2}\|\vec{z} - \vec{\zeta}\|^2$ over $\overline{\mathcal{D}}_2 \times \overline{\mathcal{D}}_2$ for each parameter $\eta > 0$. To this end, we define

$$M_\eta \triangleq \max_{(\vec{z}, \vec{\zeta}) \in \overline{\mathcal{D}}_2 \times \overline{\mathcal{D}}_2} \left\{ w^-(\vec{z}) - w^+(\vec{\zeta}) - \frac{\eta}{2}\|\vec{z} - \vec{\zeta}\|^2 \right\},$$

and we denote by $(\vec{z}^\eta, \vec{\zeta}^\eta)$ a point where the maximum is attained. This maximum satisfies $M_\eta \geq w^-(\vec{z}^0) - w^+(\vec{z}^0) = \delta_1 > 0$. From Lemma 3.1 of [7] (see Proposition 3.7 of [7] for the proof), we have

$$(B.4) \quad \lim_{\eta \rightarrow \infty} \eta \|\vec{z}^\eta - \vec{\zeta}^\eta\|^2 = 0.$$

Using the fact that $w^- = 0$ on $\partial\mathcal{D}_2$, we choose a compact set $K \subset \mathcal{D}_2$ so that $w^- < \delta_1$ on $\overline{\mathcal{D}}_2 \setminus K$. But $w^-(\vec{z}^\eta) \geq M_\eta \geq \delta_1$, and thus \vec{z}^η is in K for every $\eta > 0$. Using (B.4) and enlarging K if necessary, we can further guarantee that $\vec{\zeta}^\eta$ is also in K for all sufficiently large η . Finally, for all $\vec{z} \in K$ the quantity $1 - \lambda_1|z_1| - \lambda_2|z_2|$ is uniformly positive and bounded. By assumption, w^+ is strictly positive on K . Therefore, we have for some $\delta_2 > 0$ that for all sufficiently large η ,

$$(B.5) \quad \frac{pw^+(\vec{\zeta}^\eta)}{1 - \lambda_1|\zeta_1^\eta| - \lambda_2|\zeta_2^\eta|} \geq \delta_2.$$

We now apply Theorem 3.2 of [7] in the manner discussed in [7] immediately following the theorem. We conclude that \mathfrak{S} contains matrices \mathbf{X}^η and \mathbf{Y}^η such that

$$(B.6) \quad (\eta(\vec{z}^\eta - \vec{\zeta}^\eta), \mathbf{X}^\eta) \in \overline{\mathcal{J}}^{2,+} w^-(\vec{z}^\eta), \quad (\eta(\vec{z}^\eta - \vec{\zeta}^\eta), \mathbf{Y}^\eta) \in \overline{\mathcal{J}}^{2,-} w^+(\vec{\zeta}^\eta)$$

and

$$(B.7) \quad -3\eta \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \leq \begin{bmatrix} \mathbf{X}^\eta & \mathbf{0} \\ \mathbf{0} & -\mathbf{Y}^\eta \end{bmatrix} \leq 3\eta \begin{bmatrix} \mathbf{I} & -\mathbf{I} \\ -\mathbf{I} & \mathbf{I} \end{bmatrix},$$

where \mathbf{I} is the 2×2 identity matrix. From Theorem B.1, relation (B.6), and the continuity of H , it follows that

$$(B.8) \quad H(\bar{z}^\eta, w^-(\bar{z}^\eta), \eta(\bar{z}^\eta - \bar{\zeta}^\eta), \mathbf{X}^\eta) \leq 0 \leq H(\bar{\zeta}^\eta, w^+(\bar{\zeta}^\eta), \eta(\bar{z}^\eta - \bar{\zeta}^\eta), \mathbf{Y}^\eta).$$

We next argue that for η sufficiently large,

$$(B.9) \quad (\bar{\zeta}^\eta, w^+(\bar{\zeta}^\eta), \eta(\bar{z}^\eta - \bar{\zeta}^\eta)) \in \mathcal{O}_{\delta_2}.$$

We fix η for which (B.5) holds and use the second membership in (B.6) to choose a sequence $(\bar{\zeta}^{\eta,n}, \bar{q}^{\eta,n}, \mathbf{Y}^{\eta,n})$ converging to $(\bar{\zeta}^\eta, \eta(\bar{z}^\eta - \bar{\zeta}^\eta), \mathbf{Y}^\eta)$ and such that $(\bar{q}^{\eta,n}, \mathbf{Y}^{\eta,n}) \in J^{2,-} w^+(\bar{\zeta}^{\eta,n})$ for every n . Equation (B.3) implies that for each n there exists a function $\varphi^{\eta,n} \in C^2(\mathcal{D}_2)$ such that $\varphi^{\eta,n}(\bar{\zeta}^{\eta,n}) = w^+(\bar{\zeta}^{\eta,n})$, $\varphi^{\eta,n} \leq w^+$ on \mathcal{D}_2 , $\bar{q}^{\eta,n} = \nabla \varphi^{\eta,n}(\bar{\zeta}^{\eta,n})$, and $\mathbf{Y}^{\eta,n} = \nabla^2 \varphi^{\eta,n}(\bar{\zeta}^{\eta,n})$. Because w^+ is a viscosity supersolution of (6.15), Definition 6.1 implies that $\mathcal{B}_i \varphi^{\eta,n}(\bar{\zeta}^{\eta,n}) \geq 0$ and $\mathcal{S}_i \varphi^{\eta,n}(\bar{\zeta}^{\eta,n}) \geq 0$ for $i = 1, 2$. Multiplying either the inequality $\mathcal{B}_1 \varphi^{\eta,n}(\bar{\zeta}^{\eta,n}) \geq 0$ or $\mathcal{S}_1 \varphi^{\eta,n}(\bar{\zeta}^{\eta,n}) \geq 0$ by $|\zeta_1^{\eta,n}|$, depending on whether $\zeta_1^{\eta,n}$ is positive or negative, we obtain

$$(B.10) \quad \lambda_1 |\zeta_1^{\eta,n}| p \varphi^{\eta,n}(\bar{\zeta}^{\eta,n}) - \varphi_1^{\eta,n}(\bar{\zeta}^{\eta,n}) \zeta_1^{\eta,n} - \lambda_1 |\zeta_1^{\eta,n}| \nabla \varphi^{\eta,n}(\bar{\zeta}^{\eta,n}) \cdot \bar{\zeta}^{\eta,n} \geq 0.$$

Multiplying either the inequality $\mathcal{B}_2 \varphi^{\eta,n}(\bar{\zeta}^{\eta,n}) \geq 0$ or $\mathcal{S}_2 \varphi^{\eta,n}(\bar{\zeta}^{\eta,n}) \geq 0$ by $|\zeta_2^{\eta,n}|$, depending on whether $\zeta_2^{\eta,n}$ is positive or negative, we similarly obtain

$$(B.11) \quad \lambda_2 |\zeta_2^{\eta,n}| p \varphi^{\eta,n}(\bar{\zeta}^{\eta,n}) - \varphi_2^{\eta,n}(\bar{\zeta}^{\eta,n}) \zeta_2^{\eta,n} - \lambda_2 |\zeta_2^{\eta,n}| \nabla \varphi^{\eta,n}(\bar{\zeta}^{\eta,n}) \cdot \bar{\zeta}^{\eta,n} \geq 0.$$

Summing (B.10) and (B.11) and rearranging terms, we obtain

$$p w^+(\bar{\zeta}^{\eta,n}) - q^{\eta,n} \cdot \bar{\zeta}^{\eta,n} \geq \frac{p w^+(\bar{\zeta}^{\eta,n})}{1 - \lambda_1 |\zeta_1^{\eta,n}| - \lambda_2 |\zeta_2^{\eta,n}|}.$$

Passing to the limit and using (B.5) we conclude that

$$(B.12) \quad p w^+(\bar{\zeta}^\eta) - \eta(\bar{z}^\eta - \bar{\zeta}^\eta) \cdot \bar{\zeta}^\eta \geq \frac{p w^+(\bar{\zeta}^\eta)}{1 - \lambda_1 |\zeta_1^\eta| - \lambda_2 |\zeta_2^\eta|} \geq \delta_2,$$

which is (B.9). We further conclude from (B.4) and (B.12) that

$$p w^+(\bar{\zeta}^\eta) - \eta(\bar{z}^\eta - \bar{\zeta}^\eta) \cdot \bar{z}^\eta \geq \frac{\delta_2}{2}$$

for all η sufficiently large. In other words, for sufficiently large η ,

$$(B.13) \quad (\bar{z}^\eta, w^+(\bar{\zeta}^\eta), \eta(\bar{z}^\eta - \bar{\zeta}^\eta)) \in \mathcal{O}_{\delta_2/2}.$$

We verify at the end of this proof that H satisfies the following two conditions.

Condition 1. There exists $\gamma > 0$ such that

$$\gamma(s-t) \leq H(\vec{z}, s, \vec{q}, \mathbf{X}) - H(\vec{z}, t, \vec{q}, \mathbf{X}) \quad \forall s \geq t \text{ and } (\vec{z}, s, \vec{q}), (\vec{z}, t, \vec{q}) \in \mathcal{D}_2 \times \mathbb{R} \times \mathbb{R}^2, \mathbf{X} \in \mathfrak{S}.$$

Condition 2. For each $\delta > 0$ and bounded set $C \subset \mathbb{R}$, there exists a function $\omega: [0, \infty] \rightarrow [0, \infty]$ with $\omega(0+) = 0$ such that

$$\begin{aligned} H(\vec{\zeta}, s, \eta(\vec{z} - \vec{\zeta}), \mathbf{Y}) - H(\vec{z}, s, \eta(\vec{z} - \vec{\zeta}), \mathbf{X}) &\leq \omega(\eta\|\vec{z} - \vec{\zeta}\|^2 + \|\vec{z} - \vec{\zeta}\|), \quad s \in C, \\ (\vec{z}, s, \eta(\vec{z} - \vec{\zeta})) \in \mathcal{O}_\delta, (\vec{\zeta}, s, \eta(\vec{z} - \vec{\zeta})) \in \mathcal{O}_\delta, \text{ and } \mathbf{X}, \mathbf{Y} \in \mathfrak{S} &\text{ satisfying (B.7).} \end{aligned}$$

Using Condition 1 in the second inequality below, using (B.8) in the fourth inequality, and using (B.9) and (B.13) to justify the use of Condition 2 in the fourth inequality, we write

$$\begin{aligned} \gamma\delta_1 &\leq \gamma(w^-(\vec{z}^\eta) - w^+(\vec{\zeta}^\eta)) \\ &\leq H(\vec{z}^\eta, w^-(\vec{z}^\eta), \eta(\vec{z}^\eta - \vec{\zeta}^\eta), \mathbf{X}^\eta) - H(\vec{z}^\eta, w^+(\vec{\zeta}^\eta), \eta(\vec{z}^\eta - \vec{\zeta}^\eta), \mathbf{X}^\eta) \\ &= [H(\vec{z}^\eta, w^-(\vec{z}^\eta), \eta(\vec{z}^\eta - \vec{\zeta}^\eta), \mathbf{X}^\eta) - H(\vec{\zeta}^\eta, w^+(\vec{\zeta}^\eta), \eta(\vec{z}^\eta - \vec{\zeta}^\eta), \mathbf{Y}^\eta)] \\ &\quad + [H(\vec{\zeta}^\eta, w^+(\vec{\zeta}^\eta), \eta(\vec{z}^\eta - \vec{\zeta}^\eta), \mathbf{Y}^\eta) - H(\vec{z}^\eta, w^+(\vec{\zeta}^\eta), \eta(\vec{z}^\eta - \vec{\zeta}^\eta), \mathbf{X}^\eta)] \\ &\leq \omega(\eta\|\vec{z}^\eta - \vec{\zeta}^\eta\|^2 + \|\vec{z}^\eta - \vec{\zeta}^\eta\|). \end{aligned}$$

We now let $\eta \rightarrow \infty$ and use (B.4) to obtain a contradiction. We conclude that $w^- \leq w^+$ everywhere on $\overline{\mathcal{D}}_2$.

It thus suffices to show that $H = \min[H_5, B_1, B_2, S_1, S_2]$ satisfies Conditions 1 and 2, and for that, it suffices to show that each function H_5 , B_1 , B_2 , S_1 , and S_2 satisfies these conditions. We consider first B_1 , for which Condition 1 is clearly satisfied because $B_1(\vec{z}, s, \vec{q}) - B_1(\vec{z}, t, \vec{q}) = \lambda_1 p(s-t)$. Furthermore, $B_1(\vec{\zeta}, s, \eta(\vec{z} - \vec{\zeta})) - B_1(\vec{z}, s, \eta(\vec{z} - \vec{\zeta})) = \lambda_1 \eta \|\vec{z} - \vec{\zeta}\|^2$ for all $(\vec{\zeta}, s, \eta(\vec{z} - \vec{\zeta}))$ and $(\vec{z}, s, \eta(\vec{z} - \vec{\zeta}))$ in $\mathcal{D}_2 \times \mathbb{R} \times \mathbb{R}^2$, and thus B_1 satisfies Condition 2. Similar calculations show that B_2 , S_1 , and S_2 satisfy Conditions 1 and 2.

The function H_5 is the sum of the four functions H_1 , H_2 , H_3 , and H_4 . Of these four, only H_1 and H_4 are functions of s , and H_4 is increasing in s because \tilde{U} is decreasing. Therefore, for $s \geq t$,

$$\begin{aligned} H_5(\vec{z}, s, \vec{q}, \mathbf{X}) - H_5(\vec{z}, t, \vec{q}, \mathbf{X}) &= H_1(\vec{z}, s) - H_1(\vec{z}, t) + H_4(\vec{z}, s, \vec{q}) - H_4(\vec{z}, t, \vec{q}) \\ &\geq (1-p)A(s-t), \end{aligned}$$

and hence H_5 satisfies Condition 1.

We turn to the verification of Condition 2 for H_5 , which we do by verifying Condition 2 for H_1 , H_2 , H_3 , and H_4 . We have the bound

$$\begin{aligned} H_1(\vec{\zeta}, s) - H_1(\vec{z}, s) &= \frac{1}{2}p(1-p)s \left[\mathbf{V}(\vec{\zeta} - \vec{\theta}) \cdot (\vec{\zeta} - \vec{\theta}) - \mathbf{V}(\vec{z} - \vec{\theta}) \cdot (\vec{z} - \vec{\theta}) \right] \\ &= \frac{1}{2}p(1-p)s \left[\mathbf{V}(\vec{\zeta} - \vec{\theta}) \cdot (\vec{\zeta} - \vec{\theta}) - \mathbf{V}(\vec{\zeta} - \vec{\theta}) \cdot (\vec{z} - \vec{\theta}) \right] \\ &\quad + \frac{1}{2}p(1-p)s \left[\mathbf{V}(\vec{z} - \vec{\theta}) \cdot (\vec{\zeta} - \vec{\theta}) - \mathbf{V}(\vec{z} - \vec{\theta}) \cdot (\vec{z} - \vec{\theta}) \right] \\ &\leq \frac{1}{2}p(1-p)s \|\mathbf{V}\| (\|\vec{\zeta} - \vec{\theta}\| + \|\vec{z} - \vec{\theta}\|) \|\vec{\zeta} - \vec{z}\| \\ &\leq \text{Constant} \times \|\vec{\zeta} - \vec{z}\| \end{aligned}$$

because $\vec{\zeta}$ and \vec{z} are in the compact set $\overline{\mathcal{D}}_2$ and s is in the bounded set C . The argument that H_2 satisfies

$$H_2(\vec{\zeta}, \eta(\vec{z} - \vec{\zeta})) - H_2(\vec{z}, \eta(\vec{z} - \vec{\zeta})) \leq \text{Constant} \times (\eta \|\vec{z} - \vec{\zeta}\|^2 + \|\vec{z} - \vec{\zeta}\|)$$

is only slightly more complicated and is omitted.

Turning to H_3 , we have

$$\begin{aligned} H_3(\vec{\zeta}, \mathbf{Y}) - H_3(\vec{z}, \mathbf{X}) &= \frac{1}{2} \text{trace}(\mathbf{\Gamma}(\vec{z})\mathbf{X} - \mathbf{\Gamma}(\vec{\zeta})\mathbf{Y}) \\ &= \frac{1}{2} \text{trace}(\mathbf{\Gamma}^{\frac{1}{2}}(\vec{z})\mathbf{X}\mathbf{\Gamma}^{\frac{1}{2}}(\vec{z}) - \mathbf{\Gamma}^{\frac{1}{2}}(\vec{\zeta})\mathbf{Y}\mathbf{\Gamma}^{\frac{1}{2}}(\vec{\zeta})) \end{aligned}$$

because $\text{trace}(\mathbf{AB}) = \text{trace}(\mathbf{BA})$ for square matrices \mathbf{A} and \mathbf{B} . But when \mathbf{X} and \mathbf{Y} satisfy (B.7), the second inequality in (B.7) implies that

$$\begin{bmatrix} 3\eta\mathbf{I} - \mathbf{X} & -3\eta\mathbf{I} \\ -3\eta\mathbf{I} & 3\eta\mathbf{I} + \mathbf{Y} \end{bmatrix}$$

is positive semidefinite, which implies in turn that for every pair of two-dimensional vectors \vec{v} and \vec{w} ,

$$(B.14) \quad 0 \leq \begin{bmatrix} 3\eta\mathbf{I} - \mathbf{X} & -3\eta\mathbf{I} \\ -3\eta\mathbf{I} & 3\eta\mathbf{I} + \mathbf{Y} \end{bmatrix} \begin{bmatrix} \vec{v} \\ \vec{w} \end{bmatrix} \cdot \begin{bmatrix} \vec{v} \\ \vec{w} \end{bmatrix} = 3\eta\|\vec{v} - \vec{w}\|^2 - (\mathbf{X}\vec{v} \cdot \vec{v} - \mathbf{Y}\vec{w} \cdot \vec{w}).$$

We first apply inequality (B.14) with \vec{v} equal to the first column (which is also the first row) of $\mathbf{\Gamma}^{\frac{1}{2}}(\vec{z})$ and \vec{w} equal to the first column (which is also the first row) of $\mathbf{\Gamma}^{\frac{1}{2}}(\vec{\zeta})$ to conclude that the (1, 1) entry of $\mathbf{\Gamma}^{\frac{1}{2}}(\vec{z})\mathbf{X}\mathbf{\Gamma}^{\frac{1}{2}}(\vec{z}) - \mathbf{\Gamma}^{\frac{1}{2}}(\vec{\zeta})\mathbf{Y}\mathbf{\Gamma}^{\frac{1}{2}}(\vec{\zeta})$ is dominated by $3\eta L\|\vec{z} - \vec{\zeta}\|^2$, where L is the constant associated with the Lipschitz continuity of $\mathbf{\Gamma}^{\frac{1}{2}}(\cdot)$. We next apply (B.14) with \vec{v} equal to the second column of $\mathbf{\Gamma}^{\frac{1}{2}}(\vec{z})$ and \vec{w} equal to the second column of $\mathbf{\Gamma}^{\frac{1}{2}}(\vec{\zeta})$ to conclude that the (2, 2) entry of $\mathbf{\Gamma}^{\frac{1}{2}}(\vec{z})\mathbf{X}\mathbf{\Gamma}^{\frac{1}{2}}(\vec{z}) - \mathbf{\Gamma}^{\frac{1}{2}}(\vec{\zeta})\mathbf{Y}\mathbf{\Gamma}^{\frac{1}{2}}(\vec{\zeta})$ is dominated by $3\eta L\|\vec{z} - \vec{\zeta}\|^2$. Summing these two equalities, we see that

$$H_3(\vec{\zeta}, \mathbf{Y}) - H_3(\vec{z}, \mathbf{X}) = \frac{1}{2} \text{trace}(\mathbf{\Gamma}^{\frac{1}{2}}(\vec{z})\mathbf{X}\mathbf{\Gamma}^{\frac{1}{2}}(\vec{z}) - \mathbf{\Gamma}^{\frac{1}{2}}(\vec{\zeta})\mathbf{Y}\mathbf{\Gamma}^{\frac{1}{2}}(\vec{\zeta})) \leq 3\eta L\|\vec{z} - \vec{\zeta}\|^2$$

whenever \mathbf{X} and \mathbf{Y} satisfy (B.7). In other words, H_3 satisfies Condition 2.

Finally, assume $(\vec{\zeta}, s, \eta(\vec{z} - \vec{\zeta}))$ and $(\vec{z}, s, \eta(\vec{z} - \vec{\zeta}))$ are in \mathcal{O}_δ . The mean-value theorem implies

$$\begin{aligned} H_4(\vec{\zeta}, s, \eta(\vec{z} - \vec{\zeta})) - H_4(\vec{z}, s, \eta(\vec{z} - \vec{\zeta})) &= \tilde{U}(ps - \eta\vec{z} \cdot (\vec{z} - \vec{\zeta})) - \tilde{U}(ps - \eta\vec{\zeta} \cdot (\vec{z} - \vec{\zeta})) \\ &= -\tilde{U}'(\xi)\eta\|\vec{z} - \vec{\zeta}\|^2, \end{aligned}$$

where ξ is between $ps - \eta\vec{z} \cdot (\vec{z} - \vec{\zeta})$ and $ps - \eta\vec{\zeta} \cdot (\vec{z} - \vec{\zeta})$. Both these quantities are greater than or equal to δ , and hence $\xi \geq \delta$. Since \tilde{U}' is negative and increasing, $0 < -\tilde{U}'(\xi) \leq -\tilde{U}'(\delta)$, and consequently H_4 satisfies Condition 2. \blacksquare

REFERENCES

- [1] M. AKIAN, J. L. MENALDI, AND A. SULEM, *On an investment-consumption model with transaction costs*, SIAM J. Control Optim., 34 (1996), pp. 329–364.
- [2] C. ATKINSON AND B. AL-ALI, *On an investment-consumption model with transaction costs: An asymptotic analysis*, Appl. Math. Finance, 4 (1997), pp. 109–133.
- [3] C. ATKINSON AND P. INGPOCHAI, *The influence of correlation on multi-asset portfolio with transaction costs*, J. Comput. Finance, 10 (2007), pp. 53–96.
- [4] C. ATKINSON AND S. MOKKHAVESA, *Multi-asset portfolio optimization with transaction cost*, Appl. Math. Finance, 11 (2004), pp. 95–123.
- [5] M. BICHUCH, *Asymptotic Analysis for Optimal Investment and Consumption with Transaction Costs and Two Futures Contracts*, Ph.D. dissertation, Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA, 2010.
- [6] G. CONSTANTINIDES, *Capital market equilibrium with transaction costs*, J. Polit. Econ., 94 (1986), pp. 842–862.
- [7] M. G. CRANDALL, H. ISHII, AND P.-L. LIONS, *User’s guide to viscosity solutions of second-order partial differential equations*, Bull. Amer. Math. Soc. (N. S.), 27 (1992), pp. 1–67.
- [8] M. H. A. DAVIS AND A. NORMAN, *Portfolio selection with transaction costs*, Math. Oper. Res., 15 (1990), pp. 676–713.
- [9] J. F. DEWYNNE, S. HOWISON, S. L. LAW, AND C. F. LEE, *Correlated Multi-asset Portfolio Optimisation with Transaction Cost*, <http://arxiv.org/abs/0705.1949> (2009).
- [10] B. DUMAS AND E. LUCIANO, *An exact solution to a dynamic portfolio choice problem under transaction costs*, J. Finance, 46 (1991), pp. 577–595.
- [11] S. GERHOLD, J. MUHLE-KARBE, AND W. SCHACHERMAYER, *Asymptotics and duality for the Davis and Norman problem*, Stochastics, 84 (2012), pp. 625–641.
- [12] J. GOODMAN AND D. N. OSTROV, *Balancing small transaction costs with loss of optimal allocation in dynamic stock trading strategies*, SIAM J. Appl. Math., 70 (2010), pp. 1977–1998.
- [13] K. JANEČEK AND S. SHREVE, *Asymptotic analysis for optimal investment and consumption with transaction costs*, Finance Stoch., 8 (2004), pp. 181–206.
- [14] K. JANEČEK AND S. SHREVE, *Futures trading with transaction costs*, Illinois J. Math., 54 (2010), pp. 1239–1284.
- [15] J. KALLSEN AND J. MUHLE-KARBE, *On using shadow prices in portfolio optimization*, Ann. Appl. Probab., 20 (2010), pp. 1341–1358.
- [16] M. J. P. MAGILL AND G. M. CONSTANTINIDES, *Portfolio selection with transaction costs*, J. Econ. Theory, 13 (1976), pp. 245–263.
- [17] R. MERTON, *Optimum consumption and portfolio rules in a continuous-time case*, J. Econom. Theory, 3 (1971), pp. 373–413.
- [18] R. MERTON, *Erratum: “Optimum consumption and portfolio rules in a continuous-time model”* (J. Econom. Theory, 3 (1971), pp. 373–413), J. Econom. Theory, 6 (1973), pp. 213–214.
- [19] K. MUTHURAMAN AND S. KUMAR, *Multidimensional portfolio optimization with proportional transaction costs*, Math. Finance, 2 (2006), pp. 301–335.
- [20] L. C. G. ROGERS, *Why is the effect of proportional transaction costs $O(\delta^{2/3})$* , in Mathematics of Finance, Contemp. Math. 351, G. Yin and Q. Zhang, eds., AMS, Providence, RI, 2004, pp. 303–308.
- [21] S. SHREVE AND H. M. SONER, *Optimal investment and consumption with transaction costs*, Ann. Appl. Probab., 4 (1994), pp. 609–692.
- [22] M. TAKSAR, M. J. KLASS, AND D. ASSAF, *A diffusion model for optimal portfolio selection in the presence of brokerage fees*, Math. Oper. Res., 13 (1988), pp. 277–294.
- [23] A. E. WHALLEY AND P. WILMOTT, *An asymptotic analysis of an optimal hedging model for option pricing under transaction costs*, Math. Finance, 7 (1997), pp. 307–324.