MATHEMATICAL FINANCE
An International Journal of Mathematics, Statistics and Financial Economics

Mathematical Finance, Vol. 26, No. 1 (January 2016), 86–121

LINKED RECURSIVE PREFERENCES AND OPTIMALITY

SHLOMO LEVENTAL AND SUMIT SINHA

Department of Statistics and Probability, Michigan State University

MARK SCHRODER

Eli Broad College of Business, Department of Finance, Michigan State University

We study a class of optimization problems involving linked recursive preferences in a continuous-time Brownian setting. Such links can arise when preferences depend directly on the level or volatility of wealth, in principal–agent (optimal compensation) problems with moral hazard, and when the impact of social influences on preferences is modeled via utility (and utility diffusion) externalities. We characterize the necessary first-order conditions, which are also sufficient under additional conditions ensuring concavity. We also examine applications to optimal consumption and portfolio choice, and applications to Pareto optimal allocations.

KEY WORDS: BSDE, recursive preferences, translation-invariant preferences, team contract, Pareto optimality, behavioral contract theory, optimization.

1. INTRODUCTION

We study a class of optimization problems involving linked recursive preferences in a continuous-time Brownian setting. Such links can arise when preferences depend directly on the level or volatility of wealth, in principal–agent (optimal compensation) problems with moral hazard, and when the impact of social influences on preferences is modeled via utility (and utility diffusion) externalities. We characterize the necessary first-order conditions (FOCs), which are also sufficient under additional conditions ensuring concavity. We also examine applications to optimal consumption and portfolio choice, and applications to Pareto optimal allocations.

The optimization problems we study all reduce to maximizing a linear combination of a multidimensional backward stochastic differential equation (BSDE) system. This BSDE system was proposed by El Karoui, Peng, and Quenez (1997) as an extension of Duffie and Epstein's (1992) stochastic differential utility (SDU). Lazrak and Quenez (2003) show, in the single-agent (one-dimensional) case, that this recursive specification allows considerable flexibility in separately modeling risk aversion and intertemporal substitution, and unifies many preference classes including SDU and multiple-prior formulations (Maenhout 1999; Anderson, Hansen, and Sargent 2000; Chen and Epstein 2002). We show that the multidimensional analog can be used to model preferences in which wealth or wealth diffusion enter the aggregators, preferences in which social influences are modeled through the dependence of each agent's aggregator on utility or risk levels (or utility diffusions) of other agents, and preferences in principal/agent

This paper is dedicated to Stephanie Schroder, who edited earlier versions of this paper. We benefited from the many helpful suggestions of the anonymous referees and the Associate Editor.
Manuscript received September 2011; final revision received April 2013.

DOI: 10.1111/mafi.12047

problems in which moral hazard induces dependence of the principal's aggregator on the agent's utility and/or utility diffusion.[1] The setup is general enough to model principal–agent problems under relative income concerns (see Goukasian and Wan 2010) or relative utility concerns.

Another contribution of our paper is that we define a multidimensional extension of the translation-invariant (TI) class of BSDEs, introduced by Schroder and Skiadas (2005) as an extension of time-additive exponential utility. We show that the solution to this class of BSDEs can simplify to the solution of a single unlinked BSDE and a system of pure forward equations. Furthermore, the solution method simplifies and easily generalizes, and the conditions for sufficiency are relaxed compared to the general case. The simplification of the solution for this class is illustrated in Example 6.3 that solves the optimal consumption/portfolio problem with homothetic preferences and direct utility for wealth.

Our solution method is based on an extension of the utility-gradient approach originating in Cox and Huang (1989) and Karatzas, Lehoczky, and Shreve (1987) for additive utilities, and extended by Skiadas (1992) and Duffie and Skiadas (1994) to recursive preferences. Our general optimization result (Theorem 2.4 below) can be viewed as the natural multidimensional extension of theorem 4.2 of El Karoui, Peng, and Quenez (2001). They derive a maximum principle for the optimal consumption/portfolio problem of a single agent with recursive utility and nonlinear wealth dynamics. They formulate their problem in terms of BSDEs for utility and wealth and obtain first-order conditions (FOCs) in terms of two adjoint processes, which represent utility and wealth gradient densities.[2] We consider a general system of linked BSDEs and obtain FOCs in terms of a system of linked adjoint processes.

The outline of the paper is as follows. In Section 2, we present the general optimization problem and characterize the FOCs of the solution. We obtain the FOCs by computing the utility gradient and supergradient densities for the linked recursive utility system, and deriving an extension of the Kuhn–Tucker theorem. In Section 3, we define the TI BSDE system and present a simple sufficiency proof and solution method using dynamic programming. We also solve a principal–agent team-contract problem to illustrate the tractability. Section 4 applies our result to Pareto optimality under linked recursive preferences, and illustrates with a quadratic-aggregator example, and an example presenting some general results with TI aggregators. In Section 5, we examine the optimal consumption/portfolio choice of an agent with consumption, utility, and utility-diffusion externalities. Example 5.2 illustrates by solving in closed form a symmetric equilibrium with consumption and linear-utility externalities, which extends the consumption externality model of Galí (1994). Section 6 examines the optimal consumption/portfolio choice problem with direct dependence of the utility aggregator on wealth. Example 6.2 illustrates with the case of an aggregator depending on only a linear combination of

[1]See, for example, Schattler and Sung (1993) who pioneer the FOC approach, in which the dependence of the principal's utility on agent diffusion becomes clear. For a characterization of optimality as a system of forward–backward SDEs(FBSDEs) in principal–agent problems in more general settings, see Levental, Schroder, and Sinha (2011), who focus on recursive agent and principal preferences. For FBSDE characterizations in the context of principal–agent problems see also Cvitanic and Zhang (2012) who examine a variety of hidden-action hidden-type problems and solve the system explicitly in some specific cases.

[2]See also Cadenillas and Karatzas (1995), who develop a stochastic maximum principle for a model with linear dynamics and a convex cost criterion (which is applied to the optimal consumption/investment problem with a linear wealth equation and concave additive utility function), and obtain an explicit solution for the adjoint process.

consumption and wealth, and Example 6.3 with homothetic preferences. Finally, Appendix A contains proofs of the main results, Appendix B develops the key concepts and an extension of the Kuhn–Tucker theorem, and Appendix C contains derivations for some of the examples.

2. GENERAL MAXIMIZATION PRINCIPLE

2.1. Setting

All uncertainty is generated by a d-dimensional standard Brownian motion B over the finite time horizon $[0, T]$, supported by a probability space $(\Omega, \mathcal{F}, P)$. All processes appearing in this paper are assumed to be progressively measurable with respect to the augmented filtration $\{\mathcal{F}_t : t \in [0, T]\}$ generated by B. For any subset $V \subseteq \mathbb{R}^n$ (respectively, $V \subseteq \mathbb{R}^{n\times m}$), let $\mathcal{L}(V)$ denote the set of V-valued progressively measurable processes, and, for any $p \in \mathbb{R}$, denote

$$\mathcal{L}_p(V) = \left\{ x \in \mathcal{L}(V)\colon\ E\left[\int_0^T \|x_t\|^p dt\right] < \infty \right\},$$

where $\|x_t\|^2 = x_t' x_t$ (respectively, $\operatorname{trace}(x_t' x_t)$). We also use the following subspace of $\mathcal{L}_p(V)$ defined by

$$\bar{\mathcal{L}}_p(V) = \left\{ x \in \mathcal{L}_p(V)\colon\ E[\|x_T\|^p] < \infty \right\}.$$

It is well known that $\bar{\mathcal{L}}_2(V)$ with $V \subseteq \mathbb{R}^n$ is an inner product space with inner product defined by

$$(x|y) = E\left[\int_0^T x_t' y_t dt + x_T' y_T\right], \qquad x, y \in \bar{\mathcal{L}}_2(V).$$

The qualification "almost surely" is omitted throughout.

There are $N \geq 1$ agents. The set of *consumption plans* is $\mathcal{C} = \bar{\mathcal{L}}_2(\boldsymbol{C}^k)$, for some open interval $\boldsymbol{C} \subset \mathbb{R}$, and $k \leq N$. We interpret c_t, $t < T$, as a length-k vector of consumption rates (each dimension valued in $\boldsymbol{C}$), and c_T as a vector of lump-sum terminal consumption.

For any $c \in \mathcal{C}$, we define the $\mathbb{R}^N$-valued utility process Z as part of the pair $(Z, \Sigma) \in \bar{\mathcal{L}}_2(\mathbb{R}^N) \times \mathcal{L}_2(\mathbb{R}^{N\times d})$ that solves the BSDE

$$dZ_t(c) = -\Phi(t, c_t, Z_t, \Sigma_t)dt + \Sigma_t dB_t, \quad Z_T = \bar{\Phi}(c_T). \tag{2.1}$$

The function $\Phi : \Omega \times [0, T) \times \boldsymbol{C}^k \times \mathbb{R}^N \times \mathbb{R}^{N\times d} \to \mathbb{R}^N$ is called the *aggregator* and is $\mathcal{P} \times \mathcal{B}(\boldsymbol{C}^k \times \mathbb{R}^{N+N\times d})/\mathcal{B}(\mathbb{R}^N)$ measurable, where $\mathcal{P}$ is the predictable σ-field on $\Omega \times [0, T]$. The terminal-utility aggregator $\bar{\Phi} : \Omega \times \boldsymbol{C}^k \to \mathbb{R}^N$ is $\mathcal{F}_T \times B(\boldsymbol{C}^k)/\mathcal{B}(\mathbb{R}^N)$ measurable. We let $\Phi^i(t)$ and Σ_t^i denote the ith row of $\Phi(t)$ and $\Sigma(t)$, respectively. Let $\Phi_c \in \mathbb{R}^{N\times k}$ denote the matrix with typical element $\Phi_c^{ij} = \partial\Phi^i/\partial c^j$; $\Phi_{Z^j} \in \mathbb{R}^N$ the vector with typical element $\Phi_{Z^j}^i = \partial\Phi^i/\partial Z^j$; and $\Phi_{\Sigma^m} \in \mathbb{R}^{N\times d}$ the matrix with typical element $\Phi_{\Sigma^m}^{ij} = \partial\Phi^i/\partial\Sigma^{mj}$.

Unless stated otherwise, we assume throughout the following regularity condition on the aggregators.

CONDITION 2.1. *a)* $\Phi^i(\omega, t, c, Z, \Sigma)$ *and* $\bar{\Phi}^i(\omega, c)$ *have continuous and uniformly bounded derivatives with respect to* (c, Z, Σ) *and* c*, respectively, for* $i = 1, \ldots, N$*, and all* $(\omega, t, c, Z, \Sigma) \in \Omega \times [0, T) \times \boldsymbol{C}^k \times \mathbb{R}^N \times \mathbb{R}^{N\times d}$.

b) $\|\Phi^i_c(\omega, t, c, Z, \Sigma)\| > 0$, $\|\bar{\Phi}^i_c(\omega, c)\| > 0$, *for* $i = 1, \ldots, N$*, and all* $(\omega, t, c, Z, \Sigma) \in \Omega \times [0, T) \times \boldsymbol{C}^k \times \mathbb{R}^N \times \mathbb{R}^{N\times d}$.

c) $\Phi(t, 0, 0, 0) \in \mathcal{L}_2(\mathbb{R}^N)$.

d) $E(\|\bar{\Phi}(c_T)\|^2) < \infty$.

Condition 2.1 implies the existence of a unique pair $(Z_t, \Sigma_t) \in \bar{\mathcal{L}}_2(\mathbb{R}^N) \times \mathcal{L}_2(\mathbb{R}^{N\times d})$ that solves (2.1) for each $c \in \mathcal{C}$.

REMARK 2.2. Condition 2.1 is formulated to simplify the presentation. More general conditions for existence and uniqueness of solutions to (2.1) are available, for example, in El Karoui et al. (1997), and Briand and Confortola (2008); the latter also provides more general conditions for differentiability of solutions of (2.1). In particular,[3] we can relax the requirement of uniformly bounded $\|\Phi_c\|$ and $\|\bar{\Phi}_c\|$ and instead impose a growth condition on $\|\Phi_c\|$ together with a continuity requirement and stronger integrability condition on the consumption process c. Then, the key Lemmas 2.6 and 2.7 below, as well as Theorem 2.4, still hold.[4]

In addition to Condition 2.1, we sometimes impose the following condition (for concavity and therefore sufficiency of the FOCs).

CONDITION 2.3. *a)* $\Phi^i(\omega, t, \cdot)$ *and* $\bar{\Phi}^i(\omega, \cdot)$ *are concave functions for all* $i \in \{1, \ldots, N\}$ *and* $(\omega, t) \in \Omega \times [0, T)$.

b) $\Phi^j_{Z^i}(\omega, t, c, Z, \Sigma) \geq 0$ *and* $\Phi^j_{\Sigma^i}(\omega, t, c, Z, \Sigma) = 0$ *for all* $i, j \in \{1, \ldots, N\}$, $i \neq j$*, and* $(\omega, t, c, Z, \Sigma) \in \Omega \times [0, T) \times \boldsymbol{C}^k \times \mathbb{R}^N \times \mathbb{R}^{N\times d}$.[5]

Throughout the paper, we use $\mathbf{1}$ to denote a vector of ones, and $\boldsymbol{e}_i$ to a vector with one in the ith position and zeros elsewhere (in each case, the vector length is obvious from the context).

2.2. Optimization Problem

Fixing some nonzero weights $\beta \in \mathbb{R}^N_+$, the problem is

$$\max_{c\in\mathcal{C}} \beta' Z_0(c) \text{ subject to } Z_0(c) \geq K \tag{2.2}$$

(i.e., $Z^i_0(c) \geq K^i$, $i = 1, \ldots, N$), where $K^i \in \mathbb{R} \cup -\infty$ (to allow for void constraints).

The FOCs depend on a system of adjoint processes, which appear in the utility gradient expression given in Lemma 2.6 below. For any $c \in \mathcal{C}$, and solution (Z, Σ) to (2.1), we define the adjoint process ε, with some initial value $\varepsilon_0 \in \mathbb{R}^N$, as the solution to the SDE

[3]See, in Briand and Confortola (2008), hypothesis A1 (iv) and the regularity condition imposed on X in proposition 3.2.

[4]We thank an anonymous referee for pointing this out. See the derivation of Example 4.5 for an application in which we relax the assumption of bounded $\|\Phi_c\|$ and $\|\bar{\Phi}_c\|$.

[5]This condition is essentially the same as the quasi-monotonicity assumption (assumption 2.5) in Wu and Xu (2009). As shown in the proof of Lemma 2.7, $\Phi^j_{Z^i}(t) \geq 0$ and $\Phi^j_{\Sigma^i}(t) = 0$ are coarse conditions that ensure nonnegativity of the adjoint process. Concavity of the aggregators together with nonnegativity of the adjoints implies concavity.

(see the notation for derivatives in Section 2.1)

$$(2.3)\quad d\varepsilon_t^i = \sum_{j=1}^{N} \varepsilon_t^j \Phi_{Z^i}^j(t, c_t, Z_t, \Sigma_t)dt + \sum_{j=1}^{N} \varepsilon_t^j \Phi_{\Sigma^i}^j(t, c_t, Z_t, \Sigma_t)' dB_t, \quad i = 1, \ldots, N;$$

which can be written more compactly as

$$d\varepsilon_t^i = \varepsilon_t' \Phi_{Z^i}(t, c_t, Z_t, \Sigma_t)dt + \varepsilon_t' \Phi_{\Sigma^i}(t, c_t, Z_t, \Sigma_t)dB_t, \quad i = 1, \ldots, N.$$

Condition 2.1 implies existence and uniqueness of ε, and $\varepsilon \in \bar{\mathcal{L}}_2(\mathbb{R}^N)$.

The solution to (2.2) is characterized in the following theorem.

THEOREM 2.4 (Maximum Principle). *Suppose $c \in \mathcal{C}$ and (Z, Σ) solve the BSDE (2.1) and ε solves the SDE (2.3).*

a) (Necessity) If c solves the problem (2.2), then there is some $\kappa \in \mathbb{R}_+^N$ such that

$$\varepsilon_0 = \beta + \kappa, \quad \varepsilon_t' \Phi_c(t, c_t, Z_t, \Sigma_t) = 0, \quad t < T, \quad \varepsilon_T' \bar{\Phi}_c(c_T) = 0,$$

$$(2.4)\qquad \kappa'\{Z_0(c) - K\} = 0, \quad Z_0(c) \geq K.$$

b) (Sufficiency) If Condition 2.3 holds and (2.4) is satisfied, then c is optimal.

The necessity part of the proof is based on applying a variation of the Kuhn–Tucker theorem (see Appendix B) to a suitable gradient density of the BSDEs. The following definition of the gradient of a functional is standard (see Skiadas 1992):

DEFINITION 2.5. Let $\nu : \mathcal{C} \to \mathbb{R}$. The process $\pi \in \bar{\mathcal{L}}_2(\mathbb{R}^k)$ is a *gradient density* of ν at $c \in \mathcal{C}$ if

$$(\pi | h) = \lim_{\alpha \downarrow 0} \frac{\nu(c + \alpha h) - \nu(c)}{\alpha} \quad \text{for all } h \text{ such that } c + \alpha h \in \mathcal{C} \text{ for some } \alpha > 0.$$

A key calculation is given in the following lemma, which gives the utility gradient density of any linear functional of Z_0.

LEMMA 2.6. *Suppose $c \in \mathcal{C}$ and (Z, Σ) solve the BSDE (2.1). For any initial value $\varepsilon_0 \in \mathbb{R}^N$, the utility gradient density of $\varepsilon_0' Z_0$ at c is given by*

$$\pi_t = \varepsilon_t' \Phi_c(t, c_t, Z_t, \Sigma_t), \quad t < T, \quad \pi_T = \varepsilon_T' \bar{\Phi}_c(c_T),$$

where ε satisfies the SDE (2.3).

Proof. See Appendix A. □

The next lemma is the main step in proving the sufficiency part of Theorem 2.4.

LEMMA 2.7. *Suppose $c \in \mathcal{C}$ and (Z, Σ) solve the BSDE (2.1) and ε solves the SDE (2.3) for some $\varepsilon_0 \in \mathbb{R}_+^N$. If Condition 2.3 holds, then*

$$(2.5)\quad \varepsilon_0'\{Z_0(c + h) - Z_0(c)\} \leq E\left(\int_0^T \varepsilon_t' \Phi_c(t, c_t, Z_t, \Sigma_t)h_t dt + \varepsilon_T' \bar{\Phi}_c(c_T)h_T\right),$$

for all h such that $c + h \in \mathcal{C}$. Furthermore, $Z_0^i(\cdot)$ is concave for $i = 1, \ldots, N$; and if $\Phi_c, \bar{\Phi}_c \geq 0$, then $Z_0^i(\cdot)$ is nondecreasing.

Proof. See Appendix A. □

We now use Theorem 2.4 to sketch a characterization of the optimum as the solution to a FBSDE system. Furthermore, a reduction in dimensionality is attained, which can be useful for small N.

Assume that ε^i is strictly positive for some $i \in \{1, 2, \dots, N\}$; by relabeling we obtain strict positivity of ε^1. Define λ_t by

$$\lambda_t = \left(1, \frac{\varepsilon_t^2}{\varepsilon_t^1}, \dots, \frac{\varepsilon_t^N}{\varepsilon_t^1}\right)', \quad t \in [0, T]. \tag{2.6}$$

The FOCs (2.4) imply $\lambda_t' \Phi_c(t, c_t, Z_t, \Sigma_t) = 0$, $t < T$, and $\lambda_T' \bar{\Phi}_c(c_T) = 0$. Assume that we can invert the FOCs to solve for consumption. That is, there exists functions $\varphi : \Omega \times [0, T] \times \mathbb{R}^N \times \mathbb{R}^{N\times d} \to \boldsymbol{C}^k$ and $\bar{\varphi} : \Omega \times \mathbb{R}^N \to \boldsymbol{C}^k$ satisfying $\lambda_t' \Phi_c(t, \varphi(t, \lambda_t, Z_t, \Sigma_t), Z_t, \Sigma_t) = 0$ and $\lambda_T' \bar{\Phi}_c(\bar{\varphi}(\lambda_T)) = 0$, respectively. Applying Ito's Lemma to compute the dynamics of λ^i, we can express the FOCs (2.4) as a FBSDE system for (Z, Σ, λ):

$$\begin{aligned}
dZ_t(c) &= -\Phi(t, c_t, Z_t, \Sigma_t)dt + \Sigma_t dB_t, \quad Z_T = \bar{\Phi}(\bar{\varphi}(\lambda_T)), \\
d\lambda_t^i &= \lambda_t' \left\{\Phi_{Z^i}(t) - \lambda_t^i \Phi_{Z^1}(t) + \left(\lambda_t^i \Phi_{\Sigma^1}(t)\Phi_{\Sigma^1}'(t) - \Phi_{\Sigma^i}(t)\Phi_{\Sigma^1}'(t)\right)\lambda_t\right\} dt \\
&\quad + \lambda_t' \left\{\Phi_{\Sigma^i}(t) - \lambda_t^i \Phi_{\Sigma^1}(t)\right\} dB_t, \\
\lambda_0^i &= (\beta^i + \kappa^i)/(\beta^1 + \kappa^1), \quad i = 2, \dots, N, \\
c_t &= \varphi(t, \lambda_t, Z_t, \Sigma_t), \\
0 &= \kappa' \{Z_0(c) - K\}, \quad Z_0(c) \geq K, \quad \kappa \geq 0.
\end{aligned} \tag{2.7}$$

3. TI BSDES

In this section, we examine an aggregator form which leads to a particularly tractable solution. We show that when $k = N - 1$, the solution reduces to solving a single unlinked backward equation, followed by a system of $N - 1$ forward SDEs. Furthermore, we use a dynamic-programming argument that relaxes the sufficiency conditions of Theorem 2.4.

We assume throughout this section that $\boldsymbol{C} = \mathbb{R}$ (and therefore $\mathcal{C} = \bar{\mathcal{L}}_2(\mathbb{R}^k)$). We also assume a TI aggregator, defined as follows:

DEFINITION 3.1. A TI aggregator takes the form

$$\Phi(t, c, Z, \Sigma) = \psi(t, Mc - Z, \Sigma), \quad \bar{\Phi}(c) = Mc + \zeta,$$

for some $\psi : \Omega \times [0, T] \times \mathbb{R}^N \times \mathbb{R}^{N\times d} \to \mathbb{R}^N$, where $M \in \mathbb{R}^{N\times k}$ is assumed to satisfy rank $(M) = k$, and $\zeta = (\zeta^1, \dots, \zeta^N)' \in \mathcal{F}_T$ is assumed to satisfy $E\|\zeta\|^2 < \infty$.

We can interpret ζ as supplemental lump-sum terminal consumption in addition to terminal component of the control c (say, a lump-sum endowment; an intermediate endowment can be included via the ω argument of ψ).

A key property of the TI class is the (easily verified) quasi-linear property

$$Z_t(c + \alpha) = Z_t(c) + M\alpha, \quad t \in [0, T], \quad \text{for all } \alpha \in \mathbb{R}^k. \tag{3.1}$$

Special cases of the TI class are common in finance. Example 6.3 below shows that homothetic preferences and the standard present value operator (that is, the budget equation considered as a BSDE) are both within the TI class after transforming to logs.

The purpose of the following example is to show that additive exponential preferences are also in the TI class. Obviously, the aggregator (3.3) violates Condition 2.1(a), and so Theorem 3.8 below does not directly apply. Nevertheless, Appendix C contains a proof of existence and uniqueness of (3.2) and therefore the BSDE (2.1) with the TI aggregator (3.3).

EXAMPLE 3.2 (Additive exponential). Consider the two-agent ($N = k = 2$) system of utility functions:

$$z_t^i = -E_t\left\{\int_t^T \exp(-a^{i\prime} Mc_s)\left|z_s^j\right|^{-a_j^i} ds + \exp(-(M^i c_T + \zeta^i))\right\}, \quad i, j \in \{1, 2\}, \quad i \neq j, \tag{3.2}$$

where M^i denotes the ith row of M; we assume that $|c^i|$ and $|\zeta^i|$ are uniformly bounded; and $a^i = (a_1^i, a_2^i)'$ with $a_j^i \geq -1$ and the normalization $a_i^i = 1$. Agent i's utility is increasing (decreasing) in agent j's utility if $a_j^i < 0$ ($a_j^i > 0$). If $a^i = \boldsymbol{e}_i$ and $M^i = \gamma \boldsymbol{e}_i$, for some $\gamma \in \mathbb{R}_{++}$, then agent i's utility is standard additive exponential with coefficient of absolute risk aversion γ, and no consumption or utility externalities.[6]

The ordinally equivalent utility $Z_t^i = -\ln(-z_t^i)$ then satisfies the BSDE (2.1) with the TI aggregator

$$\psi^i(t, Mc - Z, \Sigma) = -\exp(-a^{i\prime}(Mc - Z)) - \frac{1}{2}\Sigma^{i\prime}\Sigma^i, \quad i \in \{1, 2\}. \tag{3.3}$$

We show below that the matrix M plays a key role in determining the properties of the solution. A specification that arises in principal–agent problems and Pareto efficiency problems is given in the following example.

EXAMPLE 3.3. Suppose each agent's aggregator depends only on their own consumption and utility level (but possibly on the diffusions of others), and total consumption is given by some stochastic process $C \in \bar{\mathcal{L}}(\mathbb{R})$. Let $\mathcal{C} = \bar{\mathcal{L}}_2(\mathbb{R}^{N-1})$. Defining $c_t = (c_t^1, \ldots, c_t^{N-1})'$ as the first $N-1$ agents' consumption (and so $k = N-1$), then $c_t^N = C_t - \mathbf{1}'c_t$, where $\mathbf{1}$ denotes a vector of ones. Then, M is obtained by stacking a $\mathbb{R}^{(N-1)\times(N-1)}$ identity matrix I_{N-1} on top of an $\mathbb{R}^{(N-1)}$-valued row vector of -1s:

$$M = \begin{pmatrix} I_{N-1} \\ -\mathbf{1}' \end{pmatrix}. \tag{3.4}$$

The agents' aggregators in this simple case take the form

$$\begin{aligned} dZ_t^i &= -f^i\big(t, c_t^i - Z_t^i, \Sigma_t\big)dt + \Sigma_t^{i\prime} dB_t, \quad Z_T^i = c_T^i + \zeta^i, \quad i = 1, \ldots, N-1, \\ dZ_t^N &= -f^N\big(t, C_t - \mathbf{1}'c_t - Z_t^N, \Sigma_t\big)dt + \Sigma_t^{N\prime} dB_t, \quad Z_T^N = C_T - \mathbf{1}'c_T + \zeta^N, \end{aligned}$$

[6]The preferences in Goukasian and Wan (2010) correspond to $a^i = e_i$ for each i, but M nondiagonal (that is, consumption externalities but no utility externalities).

for some $f^i : \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R}^{N\times d} \to \mathbb{R}$. For example, in a class of principal–agent problems with a single principal (let agent-N represent the principal here) and agents $1, \ldots, N-1$, the principal's problem is to choose the pay processes c to maximize the utility of the principal's (i.e., $\beta = \boldsymbol{e}_N$) cash-flow process $C - \mathbf{1}'c$, which is what remains of C after the agents are paid, subject to the "participation constraint" $Z_0^i \geq K^i$, $i = 1, \ldots, N-1$ (and $K^N = -\infty$). The agents' actions/efforts change the measure, which adds a dependence of each aggregator on the other agents' utility diffusion processes. Example 3.9 below finds an explicit solution for the case of Brownian C and quadratic diffusion penalties.

The key to the tractability of the TI class is the next lemma that shows that the adjoint processes ε must lie in the null space of M'. When $k = N-1$, this implies that λ defined in (2.6) is a constant vector.

LEMMA 3.4. *In the TI class, at the optimum c, we have*

$$\varepsilon_t' M = 0, \quad t \in [0, T], \tag{3.5}$$

where ε_t satisfies (2.3) with $\varepsilon_0 = \beta + \kappa$.

Proof. On the one hand, letting $\varepsilon_0 = \beta + \kappa$, Theorem 2.4 implies

$$\lim_{\alpha \downarrow 0} \frac{\varepsilon_t'\{Z_t(c + \alpha v) - Z_t(c)\}}{\alpha} = 0.$$

On the other hand, by quasi-linearity (3.1), the left-hand side equals $\varepsilon_t' M v$. Therefore, $\varepsilon_t' M v = 0$ for every $v \in \mathbb{R}^k$ and $t \in [0, T]$. □

Lemma 3.4 implies the following necessary condition on M for an optimum to exist, which we assume throughout the rest of this section.[7]

CONDITION 3.5.

$$v' M = 0 \textit{ for some } v \geq \beta. \tag{3.6}$$

REMARK 3.6. By Condition 3.5, we can choose some $\hat{v} \geq \beta$ satisfying $\hat{v}' M = 0$; let $\hat{\alpha} = \min\{\alpha \geq 0 : \alpha \hat{v} \geq \beta\}$, and define $v = \hat{\alpha}\hat{v}$ (i.e., we choose the "smallest" v satisfying (3.6)). Supposing a solution to problem (2.2) exists, we get that $v^i > \beta^i$ implies that ith constraint is binding. In Example 3.3, with M given by (3.4) and $\beta = \boldsymbol{e}_N$, then $v = \mathbf{1}$ and the first $N-1$ constraints must therefore (be nonvoid and) bind at any solution.

It follows from the next lemma that the existence of a unique solution to problem (2.2) implies that in the TI class, at most $N-k$ constraints in (2.2) are nonbinding at the optimum. Note that the lemma applies even with void constraints, implying that there cannot be a unique solution with more than $N-k$ void constraints.

LEMMA 3.7. *Under the TI aggregators class, if a solution to (2.2) exists, and if more than $N-k$ constraints are nonbinding, the solution is not unique.*

Proof. See Appendix A. □

[7]By Farkas' Lemma, Condition 3.5 rules out the existence of any (Pareto improving) fixed consumption increment $x \in \mathbb{R}^k$ satisfying $Mx \in \mathbb{R}_+^k$ and $\beta' Mx > 0$; that is, by quasi-linearity (3.1), an increment that reduces no agent's utilty but strictly improves at least one agent's utilty.

The main result of this section is Theorem 3.8 below, which provides a sufficiency proof together with a method for constructing a solution to the problem (2.2) under TI aggregators. We first provide a brief sketch of the solution method based on the results of Theorem 2.4 and Lemma 3.4. But we will see that the conditions required for the proof of Theorem 3.8 are considerably weaker than the sufficiency conditions for Theorem 2.4.

We assume throughout the rest of this section that $k = N - 1$, in which case a unique solution implies at most one nonbinding constraint; the key simplification in this case is that the null space of M', in which ε_t lies, is one dimensional (that is, λ is a constant vector). In light of Lemma 3.7, we rearrange the equations so that $K^i \in \mathbb{R}$, $i = 1, \dots, N-1$, and $K^N = -\infty$. Therefore, if a unique optimum exists, the first $N-1$ constraints bind.

Choose v as in Remark 3.6. Defining

$$Y_t = v' Z_t, \quad x_t = Mc_t - Z_t, \quad t \in [0, T], \tag{3.7}$$

we have the FOC condition, from (2.4) and (3.5),

$$v' \psi_x(t, x_t, \Sigma_t) M = 0, \quad t \in [0, T]. \tag{3.8}$$

Assuming invertibility, the $N-1$ conditions in (3.8), and the identity $Y_t = -v' x_t$, which is implied by the identities in (3.7), together imply

$$x_t = \phi(t, Y_t, \Sigma_t), \quad t \in [0, T],$$

for some $\phi : \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R}^{N \times d} \to \mathbb{R}^N$.

A constant λ also implies a zero diffusion term in (2.7):

$$\sum_{j=1}^{N} v^j \left\{ \psi^j_{\Sigma^i}(t, \phi(t, Y_t, \Sigma_t), \Sigma_t) - \left(\frac{v^i}{v^1} \right) \psi^j_{\Sigma^1}(t, \phi(t, Y_t, \Sigma_t), \Sigma_t) \right\} = 0,$$
$$i = 2, \dots, N. \tag{3.9}$$

The restrictions (3.9) together with the identity $v' \Sigma_t = \Sigma_t^{Y\prime}$ can, assuming invertibility, allow us to obtain $\Sigma_t = \theta(t, Y_t, \Sigma_t^Y)$ for some $\theta : \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^{N \times d}$. We then solve the BSDE for (Y, Σ^Y):

$$dY_t = -v' \psi\left(t, \phi\left(t, Y_t, \theta\left(t, Y_t, \Sigma_t^Y\right)\right), \theta\left(t, Y_t, \Sigma_t^Y\right)\right) dt + \Sigma_t^{Y\prime} dB_t, \quad Y_T = v' \zeta. \tag{3.10}$$

The solution for (Y, Σ^Y) gives us the diffusion coefficients of Z. We solve the forward equations corresponding to the binding constraints:

$$\begin{aligned} dZ_t^i &= -\psi^i\left(t, \phi\left(t, Y_t, \theta\left(t, \Sigma_t^Y\right)\right), \theta\left(t, Y_t, \Sigma_t^Y\right)\right) dt + \Sigma_t^{i\prime} dB_t, \\ Z_0^i &= K^i, \quad i = 1, \dots, N-1. \end{aligned}$$

For a vector Z, we denote by $Z^{(-i)}$ the vector with ith element removed, and for a matrix M, we denote by $M^{(-i)}$ the matrix with the ith row removed. We solve for optimal consumption $\hat{c}$ from $Z^{(-N)}$ in (3.13) below.

The solution method is made more transparent and simple in the following theorem.

THEOREM 3.8. *Suppose, for all $t \in [0, T]$,*

$$(\hat{x}_t, \hat{\Sigma}_t) = \arg \max_{(x,\Sigma)\in\mathbb{R}^N\times\mathbb{R}^{N\times d}} v'\psi(t, x, \Sigma) \tag{3.11}$$

subject to

$$v'x = -Y_t, \quad v'\Sigma = \Sigma_t^{Y\prime},$$

and (Y, Σ^Y) uniquely solves the BSDE

$$dY_t = -v'\psi(t, \hat{x}_t, \hat{\Sigma}_t)dt + \Sigma_t^{Y\prime} dB_t, \quad Y_T = v'\zeta. \tag{3.12}$$

Then, the optimal policy is

$$\hat{c}_t = (M^{(-N)})^{-1}\big(Z_t^{(-N)} + \hat{x}_t^{(-N)}\big), \quad t \in [0, T), \quad \hat{c}_T = \big(M^{(-N)}\big)^{-1}\big(Z_T^{(-N)} - \zeta^{(-N)}\big), \tag{3.13}$$

where $Z^{(-N)}$ solves the forward SDE system

$$dZ_t^i = -\psi^i(t, \hat{x}_t, \hat{\Sigma}_t)dt + \hat{\Sigma}_t^i dB_t, \quad Z_0^i = K^i, \quad i = 1, \ldots, N-1. \tag{3.14}$$

Furthermore, the optimal objective function is $\beta' Z_0(\hat{c}) = Y_0 - (v - \beta)'K$.

Proof. See Appendix A. □

Constructing a solution essentially amounts to maximizing a linear combination of drift terms. The solution $(\hat{x}_t, \hat{\Sigma}_t)$, if it exists, is shown in the Appendix to be a Lipschitz continuous function of (Y_t, Σ_t^Y); the existence and uniqueness of the BSDE solution (Y, Σ^Y) follows from standard results. The approach extends easily to include constraints on the controls, as in Example 6.3 below, which solves a constrained portfolio choice problem, and in the next example, as technical device to satisfy Condition 2.1 (we impose the constraint $\|\Sigma_t\| \leq n$, for some sufficiently large $n > 0$, to obtain Lipschitz continuity of the aggregator with respect to Σ_t).

The following example illustrates the construction of a solution in a principal–agent problem. The solution departs from the assumption of additive exponential preferences in Koo, Shim, and Sung (2008) (additive exponential corresponds to exponential $h^i(t, \cdot)$ and scalar matrix Q^i, $i = 1, \ldots, N$). However, we consider only the case of additive impact of agent effort on the Brownian drift. Similar to the result of Koo et al. (2008), in which terminal pay is linear in the terminal “outcome process,” we obtain optimal pay linear in the Brownian motion. Pay is linear in the cash flow (as in Holmstrom and Milgrom 1987) under scalar risk-aversion and effort penalties, Q^i and q^i, all i. We deal with the violation of Lipschitz continuity by obtaining a closed-form solution to (3.12) with constant diffusion, which yields optimality within the space of uniformly bounded diffusion processes.

EXAMPLE 3.9 (Team contract). We specialize the principal–agent setting in Example 3.3. We assume that the aggregate cash flow is Brownian: $C_t = \sigma' B_t$, $t \in [0, T]$. The principal (“agent” N) pays the $N - 1$ agents the consumption plans $c = (c^1, \ldots, c^{N-1})$, consuming $C - \mathbf{1}'c$ herself. The matrix M is defined in (3.4), and therefore $v = \mathbf{1}$. Assuming quadratic penalties for effort and risk, and a principal with infinite elasticity of intertemporal substitution,[8] the utility processes of the agents and principal ("agent" N) under the optimal effort plans of the agents satisfy (see the Appendix for details), for $i = 1, \dots, N-1$,

$$\text{(3.15)}\quad dZ_t^i = -\left\{ h^i\big(t, x_t^i\big) - \frac{1}{2}\Sigma_t^{i\prime} Q^i \Sigma_t^i + \Sigma_t^{i\prime} \sum_{j=1}^{N-1} q^j \Sigma_t^j \right\} dt + \Sigma_t^{i\prime} dB_t, \quad Z_T^i = c_T^i,$$

$$\text{(3.16)}\quad dZ_t^N = -\left\{ \alpha + \beta\big(C_t + x_t^N\big) - \frac{1}{2}\Sigma_t^{N\prime} Q^N \Sigma_t^N + \Sigma_t^{N\prime} \sum_{j=1}^{N-1} q^j \Sigma_t^j \right\} dt + \Sigma_t^{N\prime} dB_t,$$

$$Z_T^N = C_T - \mathbf{1}' c_T,$$

where $x_t^i = c_t^i - Z_t^i$ and $x_t^N = -\mathbf{1}' c_t - Z_t^N = -Y_t - \sum_{i=1}^{N-1} x_t^i$; the $d \times d$ matrices Q^i and q^i are all symmetric and positive definite; the h^i are deterministic functions of (t, x_t^i) with $h_x^i(t, x)$ bounded and continuous for all (t, x), and $\int_0^T h^i(t, 0)^2 dt < \infty$; and $\alpha \in \mathbb{R}$, $\beta \in \mathbb{R}_{++}$. The dependence of each aggregator on the agent j's utility diffusion results because agent j's effort shifts the drift of the Brownian motion, and optimal effort by agent j is $q^j \Sigma_t^j$.

Theorem 3.8, together with additive separability of the aggregators, implies that the optimal controls $\{(\hat{x}^i, \hat{\Sigma}^i); i = 1, \dots, N-1\}$ are obtained by maximizing the sum of all the braced terms in (3.16) after substituting $\Sigma_t^N = \Sigma_t^Y - \sum_{j=1}^{N-1} \Sigma_t^j$. Defining

$$\text{(3.17)}\qquad \eta_t = \max_{\{x_t^1, \dots, x_t^{N-1}\}} \sum_{j=1}^{N-1} \left\{ h^j\big(t, x_t^j\big) - \beta x_t^j \right\},$$

then the optimal controls $\{\hat{x}_t^1, \dots, x_t^{N-1}\}$ are the maximizing arguments (assuming they exist). The optimal diffusion controls are

$$\text{(3.18)}\qquad \hat{\Sigma}_t^i = w^i \Sigma_t^Y, \quad i = 1, \dots, N-1.$$

where[9]

$$w^i = (Q^i)^{-1} \left\{ Q^N \left(I + \sum_{j=1}^{N-1} (Q^j)^{-1} Q^N \right)^{-1} \left\{ I - \sum_{j=1}^{N-1} (Q^j)^{-1} q^j \right\} + q^i \right\}.$$

The assumption of Brownian C yields the following solution to the BSDE (3.12):

$$\text{(3.19)}\qquad Y_t = \sigma' B_t + \int_t^T e^{-\beta(s-t)} \left\{ \alpha + \eta_s - \frac{1}{2}\sigma' Q^Y \sigma \right\} ds,$$

[8]This assumption yields a drift in the BSDE (C12) that is affine in Y, and therefore a simple closed-form solution with Brownian cash flows. Alternatively, the same affine form is obtained if h^i is affine in x^i for any agent i.

[9]With identical agents, $Q^i = Q$ and $q^i = q$ for all i, this simplifies to $w^i = \{Q + (N-1)Q^N\}^{-1}(Q^N + q)$.

where

$$Q^Y = Q^N \left(1 - \sum_{j=1}^{N-1} w^j\right) - \sum_{j=1}^{N-1} q^j w^j. \tag{3.20}$$

For each agent i, we substitute $\{\hat{\Sigma}_t^j = w^k \sigma;\ j = 1, \ldots, N-1\}$ into the equation for Z^i in (3.16), now treated as a forward equation with initial value $Z_0^i = K^i$, to yield terminal consumption, whose only stochastic component is $\sigma' w^i B_T$. From the solution to Z^i, we obtain intermediate consumption: $c_t^i = \hat{x}_t^i + Z_t^i$, $t \in [0, T)$, whose only stochastic part is $\sigma' w^i B_t$ (because $\hat{x}_t^i$ is deterministic).

4. PARETO OPTIMALITY UNDER LINKED RECURSIVE UTILITY

In this section, we use Theorem 2.4 to characterize Pareto optimal allocations with linked recursive preferences. It is convenient to fix the consumption space $\mathcal{C} = \bar{\mathcal{L}}_2(\mathbb{R}_{++}^N)$, as well as the aggregate consumption process $C \in \bar{\mathcal{L}}_2(\mathbb{R}_{++})$.[10]

DEFINITION 4.1. A consumption plan $c \in \mathcal{C}$ is called feasible if $\sum_{i=1}^N c_t^i = C_t$, $t \in [0, T]$. A feasible allocation c is Pareto optimal if there is no feasible allocation $\tilde{c}$ such that $Z_0^i(c) \leq Z_0^i(\tilde{c})$, $1 \leq i \leq N$, with strict inequality for at least one i.

For any nonzero set of weights $\beta \in \mathbb{R}_+^N$ we say (as in Duffie, Geoffard, and Skiadas 1994) that the consumption plan c is β-efficient if it solves the following optimization problem:

$$\max_{c \in \mathcal{C}} \beta' Z_0(c) \ \text{ subject to } \sum_{i=1}^N c_t^i = C_t, \quad t \in [0, T]. \tag{4.1}$$

It is well known that under monotonicity and concavity of the utility functions $Z_0^i(\cdot)$, Pareto optimality is equivalent β-efficiency. In the unlinked case, with each agent i's aggregator, a function of only i's consumption, utility and diffusion, concavity of the aggregator, and monotonicity of the aggregator in consumption imply concavity and monotonicity of $Z_0^i(\cdot)$ (see Duffie and Epstein 1992 for the SDU case). In the linked case, however, current comparison theorems impose the additional restrictions in Condition 2.3 to obtain these properties.

With sufficient conditions for concavity of $Z^i(\cdot)$, the usual separating hyperplane argument shows that Pareto optimality implies β-efficiency.

PROPOSITION 4.2. *Suppose Condition 2.3 holds. If c is Pareto optimal, then c is also β-efficient for some $\beta \in \mathbb{R}_+^N$.*

The converse of the proposition is trivial if $\beta \in \mathbb{R}_{++}^N$. However, if $\beta \in \mathbb{R}_+^N$ but not strictly positive, we must again assume Condition 2.3.

Necessary and sufficient conditions for β-efficiency are obtained from Theorem 2.4.

PROPOSITION 4.3. *Suppose $c \in \mathcal{C}$ and (Z, Σ) solve the BSDE (2.1) and ε solves the SDE (2.3).*

[10]In the application of Theorem 2.4, we will substitute the feasiblity constraint and let $k = N - 1$.

a) If c is β-efficient, then

$$\varepsilon_t'\Phi_{c^1}(t,c_t,Z_t,\Sigma_t)=\varepsilon_t'\Phi_{c^2}(t,c_t,Z_t,\Sigma_t)=\cdots=\varepsilon_t'\Phi_{c^N}(t,c_t,Z_t,\Sigma_t),\quad t<T,$$
$$\varepsilon_T'\bar{\Phi}_{c^1}(c_T)=\varepsilon_T'\bar{\Phi}_{c^2}(c_T)=\cdots=\varepsilon_T'\bar{\Phi}_{c^N}(c_T),\tag{4.2}$$

where $\varepsilon_0=\beta$.

b) Let $\beta\in\mathbb{R}_+^N\setminus 0$. If Condition 2.3 holds and (4.2) is satisfied, then c is β-efficient.

Proof. We apply Theorem 2.4(a) (with $k=N-1$) after substituting $c^N=C-\sum_{i=1}^{N-1}c^i$. Then, $\varepsilon_t'\Phi_c(t)=0$ in (2.4) is equivalent to

$$\varepsilon_t'\frac{\partial\Phi(t,c_t,Z_t,\Sigma_t)}{\partial c^i}=\varepsilon_t'\{\Phi_{c^i}(t,c_t,Z_t,\Sigma_t)-\Phi_{c^N}(t,c_t,Z_t,\Sigma_t)\}=0,$$
$$i\in\{1,2,\ldots,N-1\},\tag{4.3}$$

and similarly for the terminal aggregator. □

COROLLARY 4.4. *Suppose, for each i, that Φ^i depends on c^i (the agent's own consumption) but not on other agents' consumption. Then, Proposition 4.3 holds with condition (4.2) replaced by*

$$\varepsilon_t^1\Phi_{c^1}^1(t,c_t^1,Z_t,\Sigma_t)=\varepsilon_t^2\Phi_{c^2}^2(t,c_t^2,Z_t,\Sigma_t)=\cdots=\varepsilon_t^N\Phi_{c^N}^N(t,c_t^N,Z_t,\Sigma_t),\quad t<T.$$
$$\varepsilon_t^1\bar{\Phi}_{c^1}^1(c_T)=\varepsilon_t^2\bar{\Phi}_{c^2}^2(c_T)=\cdots=\varepsilon_t^N\bar{\Phi}_{c^N}^N(c_T).\tag{4.4}$$

The following example obtains a β-efficient allocation for a simple quadratic aggregator.[11] The violation of Lipschitz continuity is managed by finding a solution within the space of uniformly bounded consumption processes (see the Appendix for details and the extension to unbounded aggregate consumption using Remark 2.2).

EXAMPLE 4.5 (Quadratic aggregator). We depart from the assumption of strictly positive consumption and let $\mathcal{C}=\bar{\mathcal{L}}((-n,n)^N)$ for some sufficiently large n. Suppose $\beta\in\mathbb{R}_{++}^N$, and aggregate consumption, C, is uniformly bounded.[12] Let agent i's aggregator, $i=1,\ldots,N$, be

$$\Phi^i(t,c,Z)\frac{1}{2}(c-p)'Q^i(c-p)q^{i\prime}Z,\quad t<T,\quad \bar{\Phi}^i(c)\frac{1}{2}(c-p)'Q^i(c-p),$$

where $Q^i=\operatorname{diag}(\bar{q}_1^i,\ldots,\bar{q}_N^i)$ is positive definite, $q^i=(q_1^i,\ldots,q_N^i)'\in\mathbb{R}^N$, $p\in\mathbb{R}^N$, and $q_j^i\geq 0$ for all $i\neq j$. The adjoint processes satisfy the linear SDE

$$d\varepsilon_t^i=\sum_{j=1}^N\varepsilon_t^j q_i^j\,dt,\quad \varepsilon_0=\beta,\quad i=1,\ldots,N.\tag{4.5}$$

[11] In this example and Example 5.2 below, we assume a linear dependence of the aggregator on Z to obtain a simple expression for ε. If we generalize to the additively separable form

$$\Phi^i(t,c,Z,\Sigma)=-\frac{1}{2}(c-p)'Q^i(c-p)+g^i(t,Z,\Sigma)$$

(and analogously for Example 5.2), we obtain the same expression for c, but an adjoint ε, from equation (2.3), with coefficients that depend on the solution of the BSDE (2.1).

[12] The derivation in the Appendix shows that the assumption of uniformly bounded C can be relaxed.

Defining $\alpha_i(t) = \sum_{j=1}^{N} \varepsilon_t^j \bar{q}_i^j$, Proposition 4.3(a) gives the FOCs

$$\alpha_1(t)\big(c_t^1 - p^1\big) = \cdots = \alpha_N(t)\big(c_t^N - p^N\big), \quad t \in [0, T].$$

Then, $c \in \mathcal{C}$ is β-efficient if and only if[13]

$$(4.6) \quad c^i(t) - p^i = \left(\alpha_i(t) \sum_{j=1}^{N} \frac{1}{\alpha_j(t)}\right)^{-1} \big(C_t - \mathbf{1}'p\big), \quad i = 1, \ldots, N, \quad t \in [0, T].$$

We show in the next example that the problem of β-efficiency under TI preferences (and no positivity constraint on consumption) results in either no (finite) solution or an infinite number of solutions.

EXAMPLE 4.6 (TI preferences). We relax the requirement of strictly positive consumption and let $\mathcal{C} = \bar{\mathcal{L}}_2(\mathbb{R}^N)$ and $C \in \bar{\mathcal{L}}_2(\mathbb{R})$. We now apply the results in Section 3 to the $k = N - 1$ dimensional control $c^{(-N)}$ (the first $N - 1$ elements of c). If each agent's aggregator depends only on his/her own consumption, then M is given by (3.4), but we make no such assumption here. There are two cases:

a) If $\beta' M \neq 0$, then no solution to (4.1) exists. This is seen by applying the quasi-linearity property (3.1) to $\alpha = M'\beta$ to get $\beta' Z_t(c^{(-N)} + M'\beta) = \beta' Z_t(c^{(-N)}) + \|M'\beta\|^2$ for any $c \in \mathcal{C}$. Thus, no allocation can be β-efficient.

b) If $\beta' M = 0$, then $\beta' Z_t(c^{(-N)} + \alpha) = \beta' Z_t(c^{(-N)})$ for all $\alpha \in \mathbb{R}^{N-1}$, and so if a solution exists it cannot be unique. A solution exists if there is a solution to (3.11). This is shown by imposing $(N - 1)$ arbitrary finite constraints on all but agent i, letting $K^i = -\infty$, and applying Theorem 3.8 to construct a solution. The optimum, $Y_0 = \beta' Z_0(\hat{c})$, is independent of these constraints (note that $\beta' M = 0$ implies that $v = \beta$). We therefore get an infinite number of solutions to the β-efficiency problem, one for each arbitrary set of constraints.

5. OPTIMAL CONSUMPTION WITH EXTERNALITIES

In this section, we examine the optimal consumption and portfolio problem of an agent when each agent's utility is affected by the consumption, utility, and risk levels (as measured by utility diffusion) of the other agents. Although the consumption processes of the other agents are given, the consumption process chosen by agent i can impact the utility processes of the other agents, which feeds back into agent i's aggregator, making the problem nonstandard. Example 5.2 extends the results of Galí (1994) and Gomez, Priestley, and Zapatero (2009), which examine symmetric equilibria with consumption externalities and identical agents, by adding linear utility externalities and obtaining a simple solution. We follow the example by showing how symmetric equilibria with consumption and linear-utility externalities can be obtained from a single-agent (with

[13]More generally, dropping the assumption of diagonal Q^is, and letting Q_i^j denote ith row of Q^j, Proposition 4.3(a) gives the FOCs

$$\sum_{j=1}^{N} \varepsilon_t^j Q_1^j (c_t - p) = \cdots = \sum_{j=1}^{N} \varepsilon_t^j Q_N^j (c_t - p), \quad t \in [0, T].$$

If, for all t, these $N - 1$ inequalities together with the constraint $\mathbf{1}'c_t = C_t$ have a solution $c \in \mathcal{C}$, then c is the unique optimal plan in $\mathcal{C}$.

no externalities) optimal consumption problem with modified prices under more general preferences.[14]

Throughout we assume that the dimension of the consumption plan of all agents is N, but the dimension of the control of agent i is one. Each agent i trades in a complete securities market, which contains a money-market security with short-rate process $r \in \mathcal{L}(\mathbb{R})$, and a set of d risky assets. We denote by $\phi^i \in \mathcal{L}(\mathbb{R}^d)$ the trading plan of agent i with ϕ_t^i representing the vector of time t market values of the risky asset investments. Let $\mu^R \in \mathcal{L}_1(\mathbb{R}^d)$ denote the excess (above the riskless rate) instantaneous expected returns process of the risky assets, and $\sigma^R \in \mathcal{L}_2(\mathbb{R}^{d\times d})$ the returns diffusion process, which is assumed to be invertible for all (ω, t). The planned consumption, trading, and wealth for agent i are feasible if $c \in \mathcal{C}$ and the usual *budget equation* is satisfied:

$$dW_t^i = \left(W_t^i r_t + \phi_t^{i\prime}\mu_t^R - c_t^i\right)dt + \phi_t^{i\prime}\sigma_t^{R\prime}dB_t, \quad c_T^i = W_T^i, \tag{5.1}$$

as well as the integrability conditions (the latter is to rule out doubling-type strategies)

$$\int_0^T \left(\left|\phi_s^{i\prime}\mu_s^R\right| + \phi_s^{i\prime}\sigma^{R\prime}\sigma^R\phi_s^i\right)ds < \infty, \quad E\left(\sup_{t\in[0,T]}\left\{\max\left(0, -W_t^i\right)\right\}^2\right) < \infty.$$

We can view the wealth process (5.1) as a forward equation, starting at an initial wealth level w_0^i with the terminal lump-sum balance W_T consumed at T; or we can define agent i's wealth process $W^i = W^i(c^i)$ as part of the pair (W^i, σ^i) solving the BSDE

$$dW_t^i = \left(W_t^i r_t + \sigma_t^{i\prime}\eta_t - c_t^i\right)dt + \sigma_t^{i\prime}dB_t, \quad c_T^i = W_T^i; \tag{5.2}$$

where $\eta_t = (\sigma_t^{R\prime})^{-1}\mu_t^R$ is the market price of risk, and the trading strategy financing c^i is $\phi_t^i = (\sigma_t^R)^{-1}\sigma_t^i$. Thus, $W_t^i(c^i)$ represents the time t cost of financing $\{c_s^i; s \in [t, T]\}$. We assume that there is a unique *state-price density* $\pi \in \bar{\mathcal{L}}_2(\mathbb{R}_{++})$ satisfying

$$\frac{d\pi_t}{\pi_t} = -r_t dt - \eta_t' dB_t, \quad \pi_0 = 1; \tag{5.3}$$

therefore, $W_t^i(c^i) = \frac{1}{\pi_t}E_t(\int_t^T c_s^i\pi_s ds + \pi_T c_T^i)$ for every $c^i \in \bar{\mathcal{L}}_2(\mathbb{R})$ (see El Karoui et al. 1997). By linearity, it follows that π is the gradient density of $W_0^i(c^i) = (\pi|c^i)$.

Agent i's problem is to choose a consumption process c^i to maximize utility subject to the wealth constraint, taking as given c^{-i}, the consumption processes of the other agents:

$$\max_{c^i: c\in\mathcal{C}} Z_0^i(c^i, c^{-i}) \text{ subject to } (\pi|c^i) \le w_0^i, \tag{5.4}$$

where $Z_0(c)$ is the initial utility specified by (2.1). The problem is nonstandard because of the possible dependence of agent i's aggregator, $\Phi^i(t)$, on $\{Z_t^j, \Sigma_t^j;\ j \neq i\}$. A perturbation in i's consumption plan can affect the other agent's utility and utility-diffusion processes, which, in turn, indirectly impacts agent i's utility process.

[14]Because the state-price density must be modified, the single agent is not a "representative agent" in the usual sense.

We can adapt Theorem 2.4 to the problem as follows:

COROLLARY 5.1. *Suppose $c \in \mathcal{C}$ and (Z, Σ) solve the BSDE (2.1) and ε solves the SDE (2.3).*

a) (Necessity) If c solves the problem (5.4), then there is some $\kappa \in \mathbb{R}_+$ such that[15]

$$\Phi_{c^i}(t, c_t, Z_t, \Sigma_t)\varepsilon_t = \kappa\pi_t, \quad t < T, \quad \bar{\Phi}_{c^i}(c_T)\varepsilon_T = \kappa\pi_T \tag{5.5}$$
$$\varepsilon_0 = \boldsymbol{e}_i, \quad \kappa\left\{w_0^i - (\pi\,|\,c^i)\right\} = 0.$$

b) (Sufficiency) If Condition 2.3 holds and (5.5) is satisfied, then c is optimal.

If the ith aggregator Φ^i depends only on Z_t^i and Σ_t^i (but can depend on the vector c_t), then $\varepsilon^j = 0$ for $j \neq i$, and the FOC reduces to the standard result for generalized recursive utility (see, for example, El Karoui et al. 2001 and, in the SDU case, Duffie and Skiadas 1994):

$$\Phi_{c^i}^i\left(t, c_t, Z_t^i, \Sigma_t^i\right)\varepsilon_t^i = \kappa\pi_t, \quad t < T, \quad \bar{\Phi}_{c^i}^i(c_T)\varepsilon_T^i = \kappa\pi_T,$$
$$\text{where} \quad \frac{d\varepsilon_t^i}{\varepsilon_t^i} = \Phi_{Z^i}^i\left(t, c_t, Z_t^i, \Sigma_t^i\right)dt + \Phi_{\Sigma^i}^i\left(t, c_t, Z_t^i, \Sigma_t^i\right)' dB_t, \quad \varepsilon_0^i = 1. \tag{5.6}$$

The following example applies Corollary 5.1 to compute a symmetric equilibrium with consumption and linear utility externalities.

EXAMPLE 5.2 (Symmetric equilibria with externalities). We build on the models of Abel (1990), Galí (1994), and Gomez et al. (2009), and suppose two ($k = 2$) identical agents, labeled a and b, with aggregators (the previous papers assume no utility externality: $\beta = 0$)

$$\Phi^i(t, c, Z) = \frac{1}{1-\gamma}(c^i)^{1-\gamma}(c^j)^{\mu} + \alpha_t Z^i + \beta_t Z^j, \quad t < T, \quad \bar{\Phi}^i(t, c) = \frac{1}{1-\gamma}(c^i)^{1-\gamma}(c^j)^{\mu},$$

$i, j \in \{a, b\}, i \neq j$, and where $\alpha \in \mathcal{L}(\mathbb{R})$ and $\beta \in \mathcal{L}(\mathbb{R}_+)$ are bounded; and $0 < \mu < \gamma < 1$. Because $\mu > 0$, the marginal utility of i's consumption is increasing in the consumption of the other agent (consistent with "catching up with the Joneses"), motivating agent i to consume more. When $\beta > 0$, the agent's utility is increasing in the utility of the other. The state-price density π in (5.3) is assumed to satisfy $\pi \in \bar{\mathcal{L}}_{4/(\mu-\gamma)}(\mathbb{R}_{++})$. To deal with the violation of Lipschitz continuity, we modify our consumption set to $\mathcal{C} = \bar{\mathcal{L}}_4(\boldsymbol{C}^k)$ and use a limiting argument (see the Appendix for the derivation). The adjoint system for agent a is $\varepsilon_t = \exp(\int_0^t \alpha_s ds)\tilde{\varepsilon}_t$ where $\tilde{\varepsilon}_t = (\tilde{\varepsilon}_t^a, \tilde{\varepsilon}_t^b)'$ is given by[16]

$$\tilde{\varepsilon}_t = \begin{pmatrix} \cosh\left(\int_0^t \beta_s ds\right) \\ \sinh\left(\int_0^t \beta_s ds\right) \end{pmatrix}, \quad t \in [0, T]$$

[15]Recall that $\boldsymbol{e}_i$ is a length N vector with one in the ith position and zeros elsewhere.

[16]It is easy to confirm that

$$d\tilde{\varepsilon}_t = \begin{pmatrix} 0 & \beta_t \\ \beta_t & 0 \end{pmatrix}\tilde{\varepsilon}_t dt, \quad \tilde{\varepsilon}_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

(the adjoint system for agent b has the elements reversed).[17] The FOCs for agents a and b are from Corollary 5.1,

$$\left(c_t^i\right)^{-\gamma}\left(c_t^j\right)^{\mu}\tilde{\varepsilon}_t^i + \frac{\mu}{1-\gamma}\left(c_t^j\right)^{1-\gamma}\left(c_t^i\right)^{\mu-1}\tilde{\varepsilon}_t^j = \exp\left(-\int_0^t \alpha_s ds\right)\kappa^i \pi_t, \quad i, j \in \{a, b\}, \quad i \neq j.$$

A symmetric Nash equilibrium, with $c^a = c^b = \hat{c}$, is therefore given in closed form by

$$\hat{c}_t = \left(\kappa \exp\left(-\int_0^t \alpha_s ds\right) \frac{\pi_t}{\tilde{\varepsilon}_t^a + \frac{\mu}{1-\gamma}\tilde{\varepsilon}_t^b}\right)^{\frac{1}{\mu-\gamma}}, \quad t \in [0, T], \tag{5.7}$$

with κ chosen to exhaust the budget: $(\pi\,|\,\hat{c}) = w^a = w^b$. The symmetric Nash equilibrium with consumption and utility externalities can therefore be obtained from the standard additive single-agent solution (with $\beta = \mu = 0$) by making two modifications:

(a) Replace the coefficient of relative risk aversion γ by $\gamma - \mu$ (this was originally shown by Galí 1994 in the discrete-time case).
(b) Multiply the subjective discount factor $\exp(\int_0^t \alpha_s ds)$ by $(\tilde{\varepsilon}_t^a + \frac{\mu}{1-\gamma}\tilde{\varepsilon}_t^b)$.

Because $(\tilde{\varepsilon}_t^a + \frac{\mu}{1-\gamma}\tilde{\varepsilon}_t^b)$ is monotonically increasing over time, optimal consumption in the Nash equilibrium with utility externality is more deferred relative to the standard case with no externality. Thus, the symmetric equilibrium with consumption and utility externalities would appear as a single-agent equilibrium without externalities, but with reduced risk aversion and diminished (possibly negative) subjective discounting of future consumption.

Alternatively, we can obtain both the utility process and optimal consumption for the two-agent symmetric equilibrium by solving the optimal consumption problem of a single agent with aggregator

$$\psi(t, \hat{c}, z) = \frac{1}{1-\gamma}\hat{c}^{1+\mu-\gamma} + (\alpha_t + \beta_t)z, \quad t < T, \quad \bar{\psi}(T, \hat{c}) = \frac{1}{1-\gamma}\hat{c}^{1+\mu-\gamma}$$

(note that $\psi(t, \hat{c}, z) = \Phi^i(t, \{\hat{c}, \hat{c}\}, \{z, z\}, \Sigma))$. The FOC (5.5) can then be written

$$\psi_c(t, \hat{c}_t, Z_t)\exp\left(\int_0^t [\alpha_s + \beta_s]\, ds\right) = \kappa\hat{\pi}_t, \quad \text{where}$$
$$\hat{\pi}_t = \left(1 + \frac{\mu}{1-\gamma}\right)\exp\left(\int_0^t \beta_s ds\right)\frac{\pi_t}{\tilde{\varepsilon}_t^a + \frac{\mu}{1-\gamma}\tilde{\varepsilon}_t^b},$$

which is the FOC (5.6) for the single-agent problem with modified state-price density $\hat{\pi}$.

The results of Example 5.2, establishing a mapping between symmetric Nash equilibria with utility/consumption externalities, and single-agent solutions with no externalities, can be extended to more general preferences. Suppose two identical agents with consumption and linear-utility externalities:

$$\Phi^i(t, c, Z, \Sigma) = \phi(t, \{c^i, c^j\}, Z^i, \Sigma^i) + \beta_t Z_t^j, \quad t < T, \tag{5.8}$$
$$\bar{\Phi}^i(c) = \bar{\phi}(\{c^i, c^j\}) \quad i, j \in \{a, b\}, \quad i \neq j,$$

[17] Following up on Footnote 2, sufficiency of the FOC holds when the assumption of nonnegative β is relaxed to $\int_0^t \beta_s ds \geq 0$ for $t \in [0, T]$.

for some $\phi : \Omega \times [0, T] \times \boldsymbol{C}^2 \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ and $\bar{\phi} : \Omega \times \boldsymbol{C}^2 \to \mathbb{R}$. The dependence of Φ^i on c^j captures the consumption externality, and the $\beta_t Z_t^j$ term represents the utility externality. We also suppose the derivatives of the aggregators with respect to own and the other's consumption are proportional at equal consumption:[18]

$$\text{(5.9)} \quad \phi_{c[1]}(t, \{k, k\}, z, \sigma) = \eta_t \phi_{c[2]}(t, \{k, k\}, z, \sigma) \quad \text{for all } t \in [0, T];\ k, z \in \mathbb{R};\ \sigma \in \mathbb{R}^d,$$

for some $\eta_t \in \mathcal{L}(\mathbb{R}_{++})$. Finally, define the single-agent aggregators $\psi : \Omega \times [0, T] \times \boldsymbol{C} \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ and $\bar{\psi} : \Omega \times \boldsymbol{C} \to \mathbb{R}$ by

$$\text{(5.10)} \qquad \psi(t, \hat{c}, z, \sigma) = \phi(t, \{\hat{c}, \hat{c}\}, z, \sigma) + \beta_t z, \quad \bar{\psi}(\hat{c}) = \bar{\phi}(\{\hat{c}, \hat{c}\}).$$

For a *given* consumption process $\hat{c} \in \mathcal{C}$, the solution to the individual-agent BSDE (2.1) and the solution to the symmetric equilibrium for either agent, (Z^i, Σ^i), $i \in \{a, b\}$, will be the same. We will use (Z^a, Σ^a) for the common solution.

We now show the modification to state prices to match the optimal consumption processes. From (5.6), the single-agent (no externalities) FOC is

$$\psi_c\left(t, \hat{c}_t, Z_t^a, \Sigma_t^a\right) \varepsilon_t = \kappa \pi_t, \quad t < T, \quad \bar{\psi}_c(\hat{c}_T)\, \varepsilon_T = \kappa \pi_T,$$

where $$\frac{d\varepsilon_t}{\varepsilon_t} = \beta_t + \psi_Z\left(t, \hat{c}_t \mathbf{1}, Z_t^a, \Sigma_t^a\right) dt + \psi_\Sigma\left(t, \hat{c}_t \mathbf{1}, Z_t^a, \Sigma_t^a\right)' dB_t, \quad \varepsilon_0 = 1.$$

Using the assumed property (5.9), Corollary 5.1 implies the following FOC for the symmetric equilibrium:[19]

$$\psi_c(t, \hat{c}_t, Z_t^a, \Sigma_t^a)\varepsilon_t = \kappa \hat{\pi}_t, \quad t < T, \quad \bar{\psi}_c(t, \hat{c}_T)\varepsilon_T = \kappa \hat{\pi}_T,$$

where $$\hat{\pi}_t = (1 + \eta_t) \exp\left(\int_0^t \beta_s ds\right) \left(\frac{\pi_t}{\tilde{\varepsilon}_t^a + \eta_t \tilde{\varepsilon}_t^b}\right), \quad t \in [0, T].$$

Thus, given κ, the symmetric equilibrium with linear externality (assuming property (5.9)) can be solved from a single-agent optimal consumption problem by replacing the state-price density π with $\hat{\pi}$. To determine κ, however, the budget under the original state prices must be exhausted.

6. OPTIMAL PORTFOLIO WITH DIRECT UTILITY FOR WEALTH

We consider the portfolio maximization problem of a single agent in complete markets with an aggregator that depends on current wealth. Special cases include St-Amour (2005), which examines time-additive HARA utility, but with a linear combination of consumption and wealth replacing consumption in the aggregator; and Ingersoll (2011),

[18] $\phi_{c[i]}$ denotes the derivative with respect to ith element of c.

[19] This is equivalent to

$$\phi_{c^a}\left(t, \bar{c}_t \mathbf{1}, Z_t^a, \Sigma_t^a\right) \left\{\tilde{\varepsilon}_t^a + \eta_t \tilde{\varepsilon}_t^b\right\} \varepsilon_t^0 = \kappa \pi_t,$$

where $\varepsilon_t^0 = \exp(-\int_0^t \beta_s ds)\varepsilon_t$; or equivalently,

$$\frac{d\varepsilon_t^0}{\varepsilon_t^0} = \phi_{Z^a}\left(t, \bar{c}_t \mathbf{1}, Z_t^a, \Sigma_t^a\right) dt + \phi_{\Sigma^a}\left(t, \bar{c}_t \mathbf{1}, Z_t^a, \Sigma_t^a\right)' dB_{t,}, \quad \varepsilon_0^0 = 1.$$

who examines (in discrete time) an aggregator that is a function of wealth and consumption. We first present the FOCs in the general recursive case, and consider some specializations. Example 6.2 considers the case where consumption and wealth enter the aggregator as a linear combination, as in St-Amour (2005), but with a general aggregator also dependent on utility and utility diffusion, and shows that the solution can be obtained by solving the problem without wealth dependence after modifying the short-rate process. Example 6.3 solves the optimal consumption/portfolio problem for a general homothetic class of recursive utility with wealth dependence and constrained trading.

We let $N = 2$ (two BSDEs) and $k = 1$, but for notational clarity, let Z be one-dimensional and introduce the additional BSDE (5.2) representing the agent's wealth. The (scalar-valued) utility satisfies the BSDE

$$dZ_t(c) = -\Phi(t, c_t, Z_t, W_t, \Sigma_t)\,dt + \Sigma_t dB_t, \quad Z_T = \bar{\Phi}(c_T),$$

where $\Phi : \Omega \times [0, T] \times \boldsymbol{C} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ and $\bar{\Phi} : \Omega \times \boldsymbol{C} \to \mathbb{R}$. The agent's problem is to maximize utility subject to the budget constraint:

$$\max_{c \in \mathcal{C}} Z_0(c) \text{ subject to } (\pi | c) \le w_0. \tag{6.1}$$

We solve the problem with the following corollary to Theorem 2.4:

COROLLARY 6.1. *a) (Necessity) If $c \in \mathcal{C}$ solves the problem (6.1), then there is some $\kappa \in \mathbb{R}_+$ such that*

$$\begin{aligned} \Phi_c(t, c_t, Z_t, W_t, \Sigma_t) = \varepsilon_t^2/\varepsilon_t^1, \quad t < T, \quad \bar{\Phi}_c(c_T) = \varepsilon_T^2/\varepsilon_T^1, \\ \kappa\{(\pi|c) - w_0\} = 0, \quad (\pi|c) \le w_0, \end{aligned} \tag{6.2}$$

where (Z, Σ) solves the BSDE (2.1), and $\varepsilon = (\varepsilon^1, \varepsilon^2)$ solves the SDE system[20]

$$\begin{aligned} \frac{d\varepsilon_t^1}{\varepsilon_t^1} &= \Phi_Z(t, c_t, Z_t, W_t, \Sigma_t)\,dt + \Phi_\Sigma(t, c_t, Z_t, W_t, \Sigma_t)\,dB_t, \quad \varepsilon_0^1 = 1, \\ d\varepsilon_t^2 &= -\left\{\varepsilon_t^2 r_t + \varepsilon_t^1 \Phi_W(t, c_t, Z_t, W_t, \Sigma_t)\right\} dt - \varepsilon_t^2 \eta_t' dB_t, \quad \varepsilon_0^2 = \kappa. \end{aligned}$$

b) (Sufficiency) Assume that $\Phi(t, \cdot)$ and $\bar{\Phi}(\cdot)$ are concave functions and $\Phi_W(t) \ge 0$, for $t < T$. If (6.2) holds, then c is optimal.

The first adjoint process, ε^1, is the standard one for unlinked recursive utility as given in (5.6) (with $i = 1$). The dynamics of the second adjoint process, ε^2, are the same as the state-price density π in (5.3) after adjusting the short rate for the incremental impact of wealth on the aggregator, which is accomplished by replacing the short rate r with $r + \Phi_W \varepsilon^1/\varepsilon^2$ (assuming $\varepsilon^2 > 0$). Just as with a higher interest rate, $\Phi_W > 0$ has the effect of deferring more consumption to the future, reducing current consumption and increasing wealth.

[20]Note the reversal in the sign of Φ_W, which follows because we apply Theorem 2.4 to $-W$ to get the correct inequality constraint.

EXAMPLE 6.2. Suppose the aggregator depends only on a linear combination of consumption and wealth:[21]

$$\Phi(t, c_t, Z_t, W_t, \Sigma_t) = \psi(t, x_t, Z_t, \Sigma_t), \quad t < T,$$

where $x_t = c_t + \delta W_t$ for some $\delta \geq 0$, and $\psi(t, \cdot)$ is a concave function. Then, the FOC (necessary and sufficient)

$$\psi_x(t, x_t, Z_t, \Sigma_t) = \varepsilon_t^2/\varepsilon_t^1, \quad t < T, \quad \bar{\Phi}_c(c_T) = \varepsilon_T^2/\varepsilon_T^1.$$

Substituting $\Phi_W(t) = \delta\psi_x(t) = \delta\varepsilon_t^2/\varepsilon_t^1$, the dynamics of ε_t^2 simplify to

$$\frac{d\varepsilon_t^2}{\varepsilon_t^2} = -(r_t + \delta)dt - \eta_t' dB_t, \quad \varepsilon_0^2 = \kappa,$$

which are the same as the dynamics of the state-price density π defined in (5.3), but with an interest rate of $r + \delta$ instead of r.[22] The optimal consumption problem is therefore the same as in the case of recursive utility without wealth dependence, but with the interest rate changed from r to $r + \delta$, and the budget constraint changed from $(\pi|c) \leq w_0\pi_0$ to $(\varepsilon^2|x) \leq w_0\varepsilon_0^2$.

The final example uses Theorem 3.8 in Section 3 to solve the optimal consumption and portfolio problem with homothetic preferences.

EXAMPLE 6.3 (Homothetic wealth-dependent utility). Suppose the homothetic specification

$$\frac{dU_t}{U_t} = -\left\{g\left(t, \frac{c_t}{U_t}, \frac{W_t}{U_t}\right) + q\left(t, \sigma_t^U\right)\right\} dt + \sigma_t^{U\prime} dB_t, \quad U_T = C_T, \tag{6.3}$$

where $g : \Omega \times [0, T] \times \mathbb{R}_{++}^2 \to \mathbb{R}$, $q : \Omega \times [0, T] \times \mathbb{R}^d \to \mathbb{R}$, $g(t, \cdot)$ and $q(t, \cdot)$ are concave, and g satisfies the Inada conditions in the c_t/U_t argument. The special case with no dependence on W_t/U_t is examined in Schroder and Skiadas (2003). Epstein-Zin utility corresponds to the power form of g in (6.6) below with $\alpha = 0$ (no dependence on W_t/U_t) and $q(t, \sigma_t^U) = \delta\sigma_t^{U\prime}\sigma_t^U$ for some $\delta > 0$. We also relax the assumption of complete markets.

Defining the investment proportion process ψ by $\psi_t^i = \phi_t^i/W_t$, and the consumption-to-wealth ratio $\rho_t = c_t/W_t$, the budget equation (5.1) can be written

$$\frac{dW_t}{W_t} = \left(r_t - \rho_t + \psi_t'\mu_t^R\right) dt + \psi_t'\sigma_t^{R\prime} dB_t, \quad W_T = C_T. \tag{6.4}$$

Because the utility and wealth aggregators fall within the TI class (after transforming W, U, and c to logs),[23] the problem can be solved using the dynamic programming

[21] Note that this specification does not fall within the TI class, which requires the time t aggregator to depend only on a linear combination of (c_t, W_t, Z_t), and requires the terminal utility be a linear in c_T.

[22] That is, $\varepsilon_t^2 = e^{-\delta t}\kappa\pi_t$, $t \in [0, T]$. An alternative approach to the problem, used in St-Amour (2005), is to apply an isomorphism along the lines of Schroder and Skiadas (2002).

[23] Let $Z_t = [\ln(U_t), -\ln(W_t)]'$ and apply a logarithmic transformation to consumption. Then $M = (1, -1)'$, $v = (1, 1)'$ and $Y_t = v'Z_t = \ln(\lambda_t)$. As stated in Section 3, the dynamic programming approach in Theorem 3.8 extends easily to additional controls, such as the constrainted portfolio choice we introduce here.

approach of Section 3. We impose possible trading restrictions by assuming that $\psi_t \in K$, $t \in [0, T]$, for some convex set K. The homothetic form implies $U_t = \lambda_t W_t$ for some λ satisfying

$$\frac{d\lambda_t}{\lambda_t} = \mu_t^\lambda dt + \sigma_t^{\lambda\prime} dB_t, \quad \lambda_T = 1.$$

The optimality condition (from Theorem 3.8) is

$$-\mu_t^\lambda = \max_{x>0,\ y\in K} \left\{ r_t - x + y' \left(\mu_t^R + \sigma_t^{R\prime}\sigma_t^\lambda\right) + g\left(t, \frac{x}{\lambda_t}, \frac{1}{\lambda_t}\right) + q\left(t, \sigma_t^\lambda + \sigma_t^R y\right) \right\}, \tag{6.5}$$

with the optimal (ρ_t, ψ_t) representing the maximizing arguments. The additive separability of the aggregator implies that we can separately solve for ρ_t and ψ_t:

$$\begin{aligned} \hat{\rho}_t &= \arg\max_{x>0} \left\{ g\left(t, \frac{x}{\lambda_t}, \frac{1}{\lambda_t}\right) - x \right\}, \\ \hat{\psi}_t &= \arg\max_{y\in K} \left\{ y' \left(\mu_t^R + \sigma_t^{R\prime}\sigma_t^\lambda\right) + q\left(t, \sigma_t^\lambda + \sigma_t^R y\right) \right\}. \end{aligned}$$

Given these solutions, we obtain μ_t^λ as a function of σ_t^λ, and then solve for the BSDE for $(\lambda, \sigma^\lambda)$ to complete the solution.

If g has the power form

$$g(t, z_1, z_2) = \begin{cases} \beta_t \dfrac{1}{1-\gamma} \left\{ \left(z_1 z_2^\alpha\right)^{\frac{1-\gamma}{1+\alpha}} - 1 \right\} & \text{if } 1 \neq \gamma > 0, \\ \beta_t \left\{ \dfrac{1}{1+\alpha} \ln\left(\dfrac{c_t}{U_t}\right) + \dfrac{\alpha}{1+\alpha} \ln\left(\dfrac{W_t}{U_t}\right) \right\} & \text{if } \quad \gamma = 1, \end{cases} \tag{6.6}$$

with $\alpha + \gamma > 0$, we get

$$\hat{\rho}_t = \left(\frac{\beta_t}{1+\alpha} \lambda_t^{\gamma-1} \right)^{\left(\frac{1+\alpha}{\alpha+\gamma}\right)}.$$

In the log case ($\gamma = 1$), this simplifies to $\hat{\rho}_t = \beta_t/(1+\alpha)$, which is invariant to λ (and therefore invariant to the dynamics of μ^R, r, and σ^R), and decreasing in α, reflecting a desire to postpone consumption and increase wealth when more weight is placed on wealth in the aggregator.

APPENDIX A: PROOFS OMITTED FROM THE TEXT

A.1 Proof of Lemma 2.6.

We start by defining the following notation. Let $x \in \mathbb{R}^{n^2\times d}$ and $y \in \mathbb{R}^n$. That is, x consists of n blocks of $n \times d$ matrices, which we denote x_i, $i = 1, \ldots, N$. We define the product between x and y as

$$(x, y) = \{\text{trace}(x_i' y);\ i = 1, \ldots, N\}.$$

For the convenience of the reader, we present without proof a well-known result on linear multidimensional BSDEs.

PROPOSITION A.1. Let $\beta \in \mathcal{L}(\mathbb{R}^{n\times n})$ and $\gamma \in \mathcal{L}(\mathbb{R}^{n^2\times d})$ both be uniformly bounded, $\varphi \in \mathcal{L}_2(\mathbb{R}^n)$ and $E(\|\xi\|^2) < \infty$. Then, the BSDE

$$-dY_t = (\varphi_t + \beta_t Y_t + (\gamma_t, Z_t))dt - Z_t dB_t, \quad Y_T = \xi,$$

has a unique solution (Y, Z) in $\bar{\mathcal{L}}_2(\mathbb{R}^n) \times \mathcal{L}_2(\mathbb{R}^{n\times d})$. Furthermore Y_t satisfies

$$\Gamma_t' Y_t = E_t\left(\Gamma_T'\xi + \int_t^T \Gamma_s'\varphi_s ds\right),$$

where, for any $a \in \mathbb{R}^n$, Γ_t is the $\mathbb{R}^n$-dimensional adjoint process defined by the forward linear SDE

$$d\Gamma_t^i = \sum_{j=1}^n \Gamma_t^j \beta_t^{ji} dt + \sum_{j=1}^n \Gamma_t^j \Pi_i^j(t) dB_t, \quad \Gamma_0 = a,$$

and where $\Pi_i(t)$ is the $n \times d$ matrix defined by concatenating the ith row from each $n \times d$ block in γ_t (consisting of n blocks).

Proof. See, for example, El Karoui et al. (1997). □

Proof of Lemma 2.6. Let h be a process such that $c + h \in \mathcal{C}$. Since $\mathcal{C}$ is convex, we have $c + \alpha h \in \mathcal{C}$ for any constant $\alpha \in [0, 1]$. Let $(Z^\alpha, \Sigma^\alpha)$ be the solution to the BSDE (2.1) corresponding to $c + \alpha h$. By the results on BSDEs in Briand and Confortola (2008), the derivative $(\partial Z, \partial\Sigma)$ of $(Z^\alpha, \Sigma^\alpha)$ with respect to α is given by the solution of following BSDE:

$$\begin{aligned} -d\partial Z_t &= (\Phi_c(t)h_t + \Phi_Z(t)\partial Z_t + (\Phi_\Sigma(t), \partial\Sigma_t))dt - (\partial\Sigma_t)dB_t, \\ \partial Z_T &= \bar{\Phi}_c(T, c_T)h_T. \end{aligned} \tag{A.1}$$

To get the exact form of $(\partial Z, \partial\Sigma)$, we use the adjoint process $\varepsilon_t \in \mathbb{R}^N$ presented as a solution of the equation (2.3). Observe Condition 2.1 implies existence and uniqueness of ε. By Proposition (A.1) (by Briand and Confortola 2008, we have $E[esssup_{t\in[0,T]}\|Z_t\|^2] < \infty$ and equation (2.3) is a linear SDE), the solution $(\partial Z, \partial\Sigma)$ of equation (A.1) is given by

$$\varepsilon_t'\partial Z_t = E_t\left(\int_t^T \varepsilon_s'\Phi_c(s, c_s, Z_s, \Sigma_s)h_s ds + \varepsilon_T'\bar{\Phi}_c(c_T)h_T\right), \quad t \in [0, T]. \quad \square$$

A.2 Proof of Lemma 2.7

We will first start by proving the nonnegativity of ε, the solution to the SDE (2.3).

LEMMA A.2. *Assume $\Phi_{Z^i}^j(t) \geq 0$ and $\Phi_{\Sigma^i}^j(t) = 0$ for all $t \in [0, T]$, $i \neq j$, $i, j \in \{1, 2, \ldots, N\}$. Then, ε_t^i, the solution of equation (2.3), satisfies $\varepsilon_t^i \geq 0$, if $\varepsilon_0^i \geq 0$, $i = 1, \ldots, N$, $0 \leq t \leq T$.*

Proof. Consider the SDE

$$d\eta_t^i = \eta_t^i\Phi_{Z^i}^i(t, c_t, Z_t, \Sigma_t)dt + \eta_t^i\Phi_{\Sigma^i}^i(t, c_t, Z_t, \Sigma_t)'dB_t, \quad i = 1, \ldots, N.$$

Under Condition 2.1 and the assumptions for the lemma, we can easily check that the conditions of theorem 1.1 of Geib and Manthey (1994) are satisfied. It follows from their theorem that $\varepsilon_0^i \geq \eta_0^i$ implies $\varepsilon_t^i \geq \eta_t^i$, $i = 1, \ldots, N$. Therefore, by selecting $\varepsilon_0^i = \eta_0^i \geq 0$, we get

$$\varepsilon_t^i \geq \eta_t^i \geq 0, \quad t \in [0, T], \quad i = 1, \ldots, N.$$

□

Proof of the Gradient Inequality:

Using the nonnegativity of ε, we now prove (2.5). Let $c, c + h \in \mathcal{C}$, and let $(Z(c), \Sigma(c))$ and $(Z(c+h), \Sigma(c+h))$ denote the solutions to equation (2.1). Define $\Delta = Z(c+h) - Z(c)$ and $\delta = \Sigma(c+h) - \Sigma(c)$. Using integration by parts, we have

$$\begin{aligned} d(\varepsilon_t' \Delta_t) = {} & -\varepsilon_t' \{\Phi_Z(t, c_t + h_t, Z_t + \Delta_t, \Sigma_t + \delta_t) - \Phi(t, c_t, Z_t, \Sigma_t) - \Phi_Z(t) Z_t \\ & - (\Phi_\Sigma(t), \Sigma_t)\}\, dt + M_t \end{aligned}$$

for some local martingale M, where we use the abbreviation for derivatives $\Phi_i(t) = \Phi_i(t, c_t, Z_t, \Sigma_t)$ for $i \in \{c, Z, \Sigma\}$. By concavity,

$$\Phi(t, c_t + h_t, Z_t + \Delta_t, \Sigma_t + \delta_t) - \Phi(t, c_t, Z_t, \Sigma_t) - \Phi_c(t) h_t - \Phi_Z(t) Z_t - (\Phi_\Sigma(t), \Sigma_t) \leq 0.$$

Let $\{\tau_n : n = 1, 2, \ldots\}$ be an increasing sequence of stopping times such that $\tau_n \to T$, and M stopped at τ_n, $M_{t \wedge \tau_n}$, is a martingale. Integrating and taking expectation, we get (note that the ε is nonnegative by Lemma A.2)

$$\varepsilon_0' \Delta_0 \leq E\left(\int_0^{\tau_n} \varepsilon_t' \Phi_c(t) h_t \, dt + \varepsilon_{\tau_n}' \Delta_{\tau_n} \right).$$

Because $\Phi_Z(t)$ and $\Phi_\Sigma(t)$ are uniformly bounded, we get $E[\sup_{t \leq T} \|\varepsilon_t\|^2] < \infty$. Using a similar argument as in Proposition 2.1 from El Karoui et al. (1997), we get $E[\sup_{t \leq T} \|Z_t\|^2] < \infty$. Therefore, $E[(\sup_{t \leq T} \|\varepsilon_t\|) \cdot (\sup_{t \leq T} \|Z_t\|)] < \infty$. Letting $n \to \infty$ and interchanging limit and expectation, we get

$$\varepsilon_0' \Delta_0 \leq E\left(\int_0^T \varepsilon_t' \Phi_c(t) h_t \, dt + \varepsilon_T' \bar{\Phi}_c(c_T) h_T \right),$$

using also the concavity of $Z^i(T) = \bar{\Phi}^i(\cdot)$ for each i.

Proof of Concavity: Let $c_a, c_b \in \mathcal{C}$, and for some $\alpha \in (0, 1)$ define $c = \alpha c_a + (1 - \alpha) c_b$. Let (Z, Σ) denote the solution to the BSDE (2.1) at c. The gradient inequality (2.5) implies, for $i \in \{a, b\}$,

$$\begin{aligned} \varepsilon_0' \{Z_0(c_i) - Z_0(c)\} \leq E\Bigg(& \int_0^T \varepsilon_t' \Phi_c(t, c_t, Z_t, \Sigma_t) \\ & \times \{c_i(t) - c(t)\}\, dt + \varepsilon_T' \bar{\Phi}_c(c_T) \{c_i(t) - c(T)\} \Bigg). \end{aligned}$$

Taking a weighted average (with weights α and $(1 - \alpha)$) implies

$$\varepsilon_0' \{\alpha Z_0(c_a) + (1 - \alpha) Z_0(c_b) - Z_0(c)\} \leq 0.$$

Because $\varepsilon_0 \in \mathbb{R}^N_+$ is arbitrary, it follows that $Z^i_0()$ is concave for $i = 1, \ldots, N$.

Proof of Monotonicity: Suppose $h_t \leq 0$, $t \in [0, T]$, and $\Phi_c, \bar{\Phi}_c \geq 0$. Then, the gradient inequality (2.5) implies

$$\varepsilon_0' \{Z_0(c+h) - Z_0(c)\} \leq E\left(\int_0^T \varepsilon_t' \Phi_c(t, c_t, Z_t, \Sigma_t) h_t dt + \varepsilon_T' \bar{\Phi}_c(c_T) h_T\right) \leq 0.$$

Because $\varepsilon_0 \in \mathbb{R}^N_+$ is arbitrary, it follows that $Z^i_0()$ is increasing for $i = 1, \ldots, N$.

A.3 Proof of Theorem 2.4

a) We will use Corollary B.10 from Appendix B with $-G = Z_0(c) - K$, and $-f = \beta' Z_0(c)$ to prove the necessity part. By Definition B.1(d) in Appendix B, c is a regular point of $\{c \in \mathcal{C} : G(c) \leq 0\}$ if $G(c) + \delta G(c; h) < 0$ for some h such that $c + h \in \mathcal{C}$, and where $\delta G(c; h)$ is the Gateaux derivative of G at c in the direction $h \in \mathcal{C}$. By Lemma 2.6, we see that in our case, the condition for a regular point is satisfied if for every initial value $\varepsilon_0 = \boldsymbol{e}_i$, $i = 1, \ldots, N$, there exists an h with $c + h \in \mathcal{C}$ and

$$E\left(\int_0^T \varepsilon_t' \Phi_c(t, c_t, Z_t, \Sigma_t) h_t dt + \varepsilon_T' \bar{\Phi}_c(c_T) h_T\right) > 0. \tag{A.2}$$

Obviously since $\varepsilon_0 = \boldsymbol{e}_i$, it follows that the solution ε of the SDE (2.3) is not identically zero. From the definition of $\mathcal{C}$, it is obvious that there exists some distinct process $d \in \mathcal{C}$ such that $c_t \neq d_t$ for all $t \in [0, T]$. Furthermore, because $\mathcal{C}$ is extended convex (see Definition B.7 in Appendix B), there exists a strictly positive process δ such that $\alpha c + (1 - \alpha) d \in \mathcal{C}$ for each process α satisfying $-\delta \leq \alpha \leq 1 + \delta$. Define $h_1 = -\delta c + (1+\delta) d - c = (1+\delta)(d - c)$ and $h_2 = -\delta d + (1+\delta) c - c = \delta(c - d)$ (so that $c + h_1, c + h_2 \in \mathcal{C}$) and note that h_1 and h_2 have opposite directions. We can choose h_t from $\{h_1(t), h_2(t)\}$, depending on the sign of $\varepsilon_t' \Phi_c(t, c_t, Z_t, \Sigma_t)$ (or $\varepsilon_T' \bar{\Phi}_c(c_T)$ at time T) so that (A.2) is satisfied. Therefore, c is a regular point.

By Corollary B.10, if c is optimal, then there exists a $\kappa \in \mathbb{R}^N_+$ such that the Gateaux derivative of $\beta' Z_0(c) + \kappa' \{Z_0(c) - K\}$ in the direction of h is 0 for all h such that $c + h \in \mathcal{C}$ and $\kappa' \{Z_0(c) - K\} = 0$. Using Lemma 2.6 to compute the Gateaux derivative, or utility gradient density, for $(\beta + \kappa)' Z_0(c) - \kappa' K$, we therefore get

$$E\left(\int_0^T \varepsilon_t' \Phi_c(t, c_t, Z_t, \Sigma_t) h_t dt + \varepsilon_T' \Phi_c(T, c_T) h_T\right) = 0, \quad t \in [0, T],$$
$$\forall h \text{ such that } c + h \in \mathcal{C},$$

where ε solves the SDE (2.3) with initial value $\varepsilon_0 = \beta + \kappa$.

Because the above statement is true $\forall h$ such that $c + h \in \mathcal{C}$, we have $\varepsilon_t' \Phi_c(t, c_t, Z_t, \Sigma_t) = 0$, $t < T$, and $\varepsilon_T' \bar{\Phi}_c(c_T) = 0$ (otherwise we get a contradiction using h constructed as above). This completes the necessity proof.

b) Lemma (2.7), $\varepsilon_0 = \beta + \kappa$ and (2.4) together imply

$$\beta' \{Z_0(c+h) - Z_0(c)\} \leq \kappa' \{Z_0(c) - Z_0(c+h)\} = \kappa' \{K - Z_0(c+h)\}$$

(the equality follows from complementary slackness). It follows immediately that $\beta' Z_0(c+h) > \beta' Z_0(c)$ implies $\kappa'\{K - Z_0(c+h)\} > 0$, and therefore a violation of at least one constraint.

A.4 Proof of Lemma 3.7

Let $\hat{c}$ be a solution to (2.2) and suppose exactly $n \leq N$ constraints in (2.2) are nonbinding; without loss of generality, assume that these correspond to $i = 1, \ldots, n$. That is, $Z_t^i(\hat{c}) > K^i$, $i = 1, \ldots, n$; and $Z_t^i(\hat{c}) = K^i$, $i = n+1, \ldots, N$. Decompose $M = [M^{a\prime}, M^{b\prime}]'$ where $M^a \in \mathbb{R}^{n\times k}$ and $M^b \in \mathbb{R}^{(N-n)\times k}$; similarly, decompose $Z = [Z^{a\prime}, Z^{b\prime}]'$, $\beta = [\beta^{a\prime}, \beta^{b\prime}]'$, $K = [K^{a\prime}, K^{b\prime}]'$, and $\varepsilon_0 = [\varepsilon_0^{a\prime}, \varepsilon_0^{b\prime}]'$. Defining $\tilde{M} \in \mathbb{R}^{(N-n+1)\times k}$ by $\tilde{M} = \binom{\beta^{a\prime} M^a}{M^b}$, then nonuniqueness is implied if rank $(\tilde{M}) < k$. This follows because then there is a $x \in \mathbb{R}^k$ satisfying $\tilde{M}x = 0$, and therefore (using (3.1) $Z_t^b(\hat{c} + \alpha x) - Z_t^b(\hat{c}) = \alpha M^b x = 0$ and $\beta^{a\prime} Z_t^a(\hat{c} + \alpha x) - \beta^{a\prime} Z_t^a(\hat{c}) = \alpha\beta^{a\prime} M^a x = 0$. Choose $\alpha \in \mathbb{R}$ sufficiently small that $Z_t^a(\hat{c} + \alpha x) = Z_t^a(\hat{c}) + \alpha M^a x \geq K^a$. From $\varepsilon_0^a = \beta^a$ (from (2.4) and the supposition that the first n constraints are nonbinding), we get that $\varepsilon_0' M = 0$ (which is implied by optimality of $\hat{c}$ and (3.5)) is equivalent to $(1, \varepsilon_0^{b\prime})\tilde{M} = 0$, which implies rank $(\tilde{M}) \leq N - n$. Therefore, $n > N - k$ implies rank $(\tilde{M}) < k$ and therefore nonuniqueness.

A.5 Proof of Theorem 3.8

We first prove an envelope-theorem-type result that implies Lipschitz continuity of the drift of Y under uniform Lipschitz continuity of $\psi^i(\omega, t, x, \Sigma)$ in (x, Σ) for all i. Note that the uniform Lipschitz condition is weaker than the assumption of uniformly bounded derivatives of ψ^i assumed in Condition (2.1).

LEMMA A.3. *If $\psi^i(t, \cdot)$ is uniformly Lipschitz for $i = 1, \ldots, N$, then $\mu^Y(t, \cdot)$ defined by*

$$\mu^Y(\omega, t, Y, \Sigma^Y) = \max_{(x,\Sigma)\in\mathbb{R}^N\times\mathbb{R}^{N\times d}} v'\psi(\omega, t, x, \Sigma) \quad \text{subject to } v'x = -Y \text{ and } v'\Sigma = \Sigma^{Y\prime},$$

(A.4)

is uniformly Lipschitz.

Proof. For simplicity of notation, we will omit the (ω, t) arguments. By the uniform Lipschitz property for ψ^i, there exists a $C_1 \in \mathbb{R}_+$ such that, for each $i = 1, \ldots, N$,

$$|\psi^i(\tilde{x}, \tilde{\Sigma}) - \psi^i(x, \Sigma)| \leq C_1\{\|\tilde{x} - x\| + \|\tilde{\Sigma} - \Sigma\|\}$$
$$\text{for all } (\omega, t) \in \Omega \times [0, T],\ \tilde{x}, x \in \mathbb{R}^N \text{ and } \tilde{\Sigma}, \Sigma \in \mathbb{R}^{N\times d}.$$

Fix (ω, t), and choose any (Y, Σ^Y) and $(\tilde{Y}, \tilde{\Sigma}^Y)$ (both in $\mathbb{R} \times \mathbb{R}^d$) and suppose the maximizing arguments in (A.4) are (x, Σ) and $(\tilde{x}, \tilde{\Sigma})$, respectively (both are in $\mathbb{R}^N \times \mathbb{R}^{N\times d}$). Denote

$$\mu^Y(Y, \Sigma^Y) = v'\psi(x, \Sigma), \quad \mu^Y(\tilde{Y}, \tilde{\Sigma}^Y) = v'\psi(\tilde{x}, \tilde{\Sigma})$$

(and note that $v'x = -Y_t$, $v'\tilde{x} = -\tilde{Y}_t$, $v'\Sigma = \Sigma_t^{Y\prime}$, $v'\tilde{\Sigma} = \tilde{\Sigma}_t^{Y\prime}$). Choose i corresponding to some $v_i > 0$. Because (x, Σ) maximizes (A.4) for (Y, Σ^Y) and $(\tilde{x}, \tilde{\Sigma})$ for $(\tilde{Y}, \tilde{\Sigma}^Y)$, we have

$$\mu^Y(\tilde{Y}, \tilde{\Sigma}^Y) = v'\psi(\tilde{x}, \tilde{\Sigma}) \geq v'\psi\left(x + \frac{1}{v_i}\mathbf{e}_i\{Y - \tilde{Y}\}, \Sigma + \frac{1}{v_i}\mathbf{e}_i\{\Sigma^{Y\prime} - \tilde{\Sigma}^{Y\prime}\}\right),$$
$$\mu^Y(Y, \Sigma^Y) = v'\psi(x, \Sigma) \geq v'\psi\left(\tilde{x} + \frac{1}{v_i}\mathbf{e}_i\{\tilde{Y} - Y\}, \tilde{\Sigma} + \frac{1}{v_i}\mathbf{e}_i\{\tilde{\Sigma}^{Y\prime} - \Sigma^{Y\prime}\}\right).$$

By Lipschitz continuity, there exists a constant $C_2 \in \mathbb{R}_+$, independent (ω, t), such that

$$\begin{aligned} v'\psi & \left(x + \frac{1}{v_i}\mathbf{e}_i\{Y - \tilde{Y}\}, \Sigma + \frac{1}{v_i}\mathbf{e}_i\{\Sigma^{Y\prime} - \tilde{\Sigma}^{Y\prime}\}\right) \\ & \geq v'\psi(x, \Sigma) - C_2\{|\tilde{Y} - Y| + \|\tilde{\Sigma}^{Y\prime} - \Sigma^{Y\prime}\|\}, \\ v'\psi & \left(\tilde{x} + \frac{1}{v_i}\mathbf{e}_i\{\tilde{Y} - Y\}, \tilde{\Sigma} + \frac{1}{v_i}\mathbf{e}_i\{\tilde{\Sigma}^{Y\prime} - \Sigma^{Y\prime}\}\right) \\ & \geq v'\psi(\tilde{x}, \tilde{\Sigma}) - C_2\{|\tilde{Y} - Y| + \|\tilde{\Sigma}^{Y\prime} - \Sigma^{Y\prime}\|\}. \end{aligned}$$

Combining the results yields

$$|\mu^Y(Y, \Sigma^Y) - \mu^Y(\tilde{Y}, \tilde{\Sigma}^Y)| \leq C_2\{\|\tilde{Y} - Y\| + \|\tilde{\Sigma}^{Y\prime} - \Sigma^{Y\prime}\|\}. \qquad \square$$

We now prove Theorem 3.8. Suppose $(\hat{x}, \hat{\Sigma})$ solve (3.11) and let (Y, Σ^Y) solve the BSDE (3.12). The existence and uniqueness of the solution is implied by Lemma 3.7 . Let $Z^{(-N)}$ and $\hat{c}$ be computed as in (3.14) and (3.13), and define $Z^N = (Y - v^{(-N)\prime}Z^{(-N)})/v^N$. It is straightforward to confirm that $(Z, \hat{\Sigma})$, so constructed solves the BSDE system

$$dZ_t = -\psi(t, M\hat{c}_t - Z_t, \hat{\Sigma}_t)dt + \hat{\Sigma}_t dB_t, \quad Z_T = M\hat{c}_T + \zeta. \tag{A.5}$$

Now consider any $\tilde{c} \in \mathcal{C}$ and let $(\tilde{Z}, \tilde{\Sigma})$ denote the solution to the BSDE (A.5), with $(\tilde{Z}, \tilde{\Sigma}, \tilde{c})$ replacing $(Z, \hat{\Sigma}, \hat{c})$, and define $\tilde{x}_t = M\tilde{c}_t - \tilde{Z}_t$. Defining $\tilde{Y}_t = v'\tilde{Z}_t$ and $\tilde{\Sigma}_t^{Y\prime} = v'\tilde{\Sigma}_t$, and letting $\tilde{\Sigma}_t^{(-N)}$ denote the first $N-1$ rows of $\tilde{\Sigma}_t$, then $(\tilde{Y}, \tilde{\Sigma}^Y)$ solves the BSDE

$$\begin{aligned} d\tilde{Y}_t = & -\left\{v'\psi\left(t, \tilde{x}_t, \tilde{\Sigma}_t^{(-N)}, \left\{\tilde{\Sigma}_t^{Y\prime} - \sum_{i=1}^{N-1} v^i \tilde{\Sigma}^i\right\} \Big/ v^N\right)\right\} dt \\ & + \tilde{\Sigma}_t^{Y\prime} dB_t, \quad \tilde{Y}_T = v'\zeta. \end{aligned} \tag{A.6}$$

By (3.11), we have

$$v'\psi(t, \hat{x}_t, \hat{\Sigma}_t) = p_t + v'\psi\left(t, \tilde{x}_t, \tilde{\Sigma}_t^{(-N)}, \left\{\Sigma_t^{Y\prime} - \sum_{i=1}^{N-1} v^i \tilde{\Sigma}^i\right\} \Big/ v^N\right), \quad t \in [0, T],$$

for some nonnegative process p, and therefore

$$\begin{aligned} dY_t = & -\left\{p_t + v'\psi\left(t, \tilde{x}_t, \tilde{\Sigma}_t^{(-N)}, \left\{\Sigma_t^{Y\prime} - \sum_{i=1}^{N-1} v^i \tilde{\Sigma}^i\right\} \Big/ v^N\right)\right\} dt \\ & + \Sigma_t^{Y\prime} dB_t, \quad Y_T = v'\zeta. \end{aligned} \tag{A.7}$$

The comparison lemma of El Karoui et al. (1997) applied to (A.7) and (A.6) implies $Y_0 \geq \tilde{Y}_0$ and therefore (because constraints $1, \ldots, N-1$ are binding) $Z_0^N(\hat{c}) \geq Z_0^N(\tilde{c})$. Because this holds for all $\tilde{c} \in \mathcal{C}$, $\hat{c}$ must be optimal.

APPENDIX B: VARIATIONS OF THE KUHN–TUCKER THEOREM

In this Appendix, we prove variations of the Kuhn–Tucker theorem. The method of the proof is similar to the proof of the Kuhn–Tucker theorem in section 1.4 of Luenberger (1969). We start by proving the theorem for general normed spaces and then specialize to collection of processes. We will need the following definitions for the theorem:

DEFINITION B.1. Let B and Z be normed spaces, $X \subseteq B$, $G : X \to Z$ and $f : X \to \mathbb{R}$.
a) X is called "extended convex" if $\forall x_1, x_2 \in X$, there is $\delta = \delta(x_1, x_2) > 0$ so that

$$\alpha x_1 + (1-\alpha)x_2 \in X \quad \text{for all} - \delta \leq \alpha \leq 1 + \delta.$$

b) Denote $H_x \equiv \{h \in B : x + h \in X\}$, $x \in X$. Let X be extended convex and let $x \in X$, $h \in H_x$. We denote the *Gateaux derivative* of G in the direction h as

$$\delta G(x, h) = \lim_{\alpha \to 0} \frac{G(x + \alpha h) - G(x)}{\alpha}.$$

Observe that since X is extended convex, we do not have to restrict α to be positive in the definition of $\delta G(x, h)$. If $\delta G(x, h)$ exists for all $h \in H_x$, we say that G is Gateaux differentiable at x.
c) The functional $\delta f(x, \cdot) : H_x \to \mathbb{R}$ will be called the *supergradient* for f at x if

$$f(x + h) - f(x) \leq \delta f(x, h), \qquad \forall h \in H_x.$$

d) Let Z contain a convex cone A; namely, $\alpha x + \beta y \in A$, for each $\alpha, \beta > 0$, $x, y \in A$. Also, we denote $x \geq y$ (respectively, $x > y$) if $x - y \in A$ (respectively, $x - y \in \text{interior}(A)$) and A is a positive cone. The point $x_0 \in X$ is said to be a *regular point* of $\{G(x) \leq 0\}$ if $G(x_0) \leq 0$ and there is $h \in H_{x_0}$ so that $G(x_0) + \delta G(x_0, h) < 0$, where $\delta G(x_0, h)$ is the Gateaux derivative of G at x_0. Finally, we denote $z^* \geq 0$ for any $z^* \in Z^*$(Z^* is the dual space of Z) if $z^*[x] \geq 0$ for each $x \in A$.

REMARK B.2. (a) H_x is convex (respectively, extended convex) if X is convex (respectively, extended convex), but it is not necessarily a linear subspace of B. For a mapping $T : H_x \to Y$, where Y is a vector space, to be linear simply means

$$T(ah_1 + bh_2) = aT(h_1) + bT(h_2), \quad \text{for all } a, b \in \mathbb{R},\ h_1, h_2, ah_1 + bh_2 \in H_x.$$

If $\delta G(x, \cdot)$ (respectively, $\delta f(x, \cdot)$) is linear in this sense, we say that G has a linear Gateaux derivative (respectively, linear supergradient) at x.
(b) We can define the supergradient of G in the same way as we defined for f, and we can define $x_0 \in X$ to be a regular point in the same way as in Definition B.1(d) with the understanding that $\delta G(x_0, h)$ stands for the supergradient of G.

Next, we will define a weaker concept than extended convex that will be useful in certain cases (see the concept of extended convex for processes in Definition B.7).

DEFINITION B.3. We will say that X, a subset of a vector space B, is "weakly extended convex with respect to a collection of functions $F = \{f : B \to \mathbb{R}^d\}$" if $f(X)$ is extended convex in $\mathbb{R}^d$ for each $f \in F$.

EXAMPLE B.4 (Extended Convex). Let B be a vector space whose elements are all the real-valued functions defined on the set A (say). Let $C = \{f \in B : f(x) > 0, x \in A\}$. Obviously, $C \subset B$ is convex. A simple condition for C to be extended convex is:

$$0 < \inf_{x \in A}\{f(x)\} \leq \sup_{x \in A}\{f(x)\} < \infty, \quad \text{for all } f \in C.$$

THEOREM B.5 (Kuhn–Tucker). *Let B be a normed space, Z a normed space that contains a positive cone P with nonempty interior, $X \subset B$ extended convex, $G : X \to Z$ and $f : X \to \mathbb{R}$. Let $x_0 \in X$ satisfy $G(x_0) \leq 0$ and $f(x_0) = \min_{\{x \in X: G(x) \leq 0\}}\{f(x)\}$. Assume:*

(i) G has a linear Gateaux derivative and f has a linear supergradient at x_0.

(ii) x_0 is a regular point of $\{G(x) \leq 0\}$.

Then there is $z^* \in Z^*$, $z^* \geq 0$ such that

$$\delta f(x_0, h) + z^*[\delta G(x_0, h)] = 0, h \in H_{x_0}, \quad z^*[G(x_0)] = 0.$$

Proof. In the space $W = \mathbb{R} \times Z$, define the sets

$$A = \{(r, z) : r \geq \delta f(x_0; h), z \geq G(x_0) + \delta G(x_0; h) \text{ for some } h \in H_{x_0}\},$$
$$D = \{(r, z) : r \leq 0, z \leq 0\}.$$

The set D is obviously convex. Letting (r_1, z_1) and $(r_2, z_2) \in A$, and using the fact that the supergradient of f and the Gateaux derivative of G are linear, we have $(\lambda r_1 + (1-\lambda)r_2, \lambda z_1 + (1-\lambda)z_2) \in A$ where $0 \leq \lambda \leq 1$. Therefore, A is also convex. The set D contains interior points because P does. If $(r, z) \in A$, with $r < 0$ and $z < 0$, then there exists $h \in H_{x_0}$ such that

$$\delta f(x_0; h) < 0, \quad G(x_0) + \delta G(x_0; h) < 0.$$

The point $G(x_0) + \delta G(x_0; h)$ is the center of some sphere of radius ρ contained in the negative cone in Z. Then, for $0 < \alpha < 1$, the point $\alpha(G(x_0) + \delta G(x_0; h))$ is the center of an open sphere of radius $\alpha\rho$ contained in negative cone; hence so is the point $(1-\alpha)G(x_0) + \alpha(G(x_0) + \delta G(x_0; h)) = G(x_0) + \alpha\delta G(x_0; h)$. By the definition of the Gateaux derivative, we have

$$\| G(x_0 + \alpha h) - G(x_0) - \alpha\delta G(x_0; h) \| = o(\alpha).$$

Therefore, $G(x_0 + \alpha h) < 0$ for sufficiently small α. By definition of supergradient for f, we have $f(x_0 + h) - f(x_0) \leq \delta f(x_0; h)$. Then, the supposition $\delta f(x_0; h) < 0$ implies $f(x_0 + h) < f(x_0)$. This contradicts the optimality of x_0; therefore, A contains no interior points of D.

By theorem 3, section 5.12 (from Luenberger 1969), there is a hyperplane separating A and D. Hence, there are r_0, z^*, and δ such that

$$r_0 r + z^*[z] \geq \gamma \quad \text{for all } (r, z) \in A,$$
$$r_0 r + z^*[z] \leq \gamma \quad \text{for all } (r, z) \in D.$$

Because $(0, 0)$ belongs to both A (choose $h = 0$) and D, we have $\gamma = 0$. It follows at once that $r_0 \geq 0$ and $z^* \geq 0$ (otherwise, you can choose $(r, z) \in D$ with one component 0 and the other negative, resulting in a contradiction). Furthermore, $r_0 > 0$ because of the existence of $h \in H_{x_0}$ such that $G(x_0) + \delta G(x_0; h) < 0$ (by regularity of x_0). By scaling, we can assume without loss of generality that $r_0 = 1$.

From the separation property, we have for all $h \in H_{x_0}$

$$\delta f(x_0; h) + z^*[G(x_0) + \delta G(x_0; h)] \geq 0$$

(because trivially $(\delta f(x_0; h), [G(x_0) + \delta G(x_0; h)]) \in A$). Setting $h = 0$ gives $z^*[G(x_0)] \geq 0$ but $G(x_0) \leq 0$, $z^* \geq 0$ implies $z^*[G(x_0)] \leq 0$ and hence $z^*[G(x_0)] = 0$. We conclude that

$$\delta f(x_0; h) + z^*[\delta G(x_0, h)] \geq 0, \quad \forall h \in H_{x_0}. \tag{B.1}$$

For any $h \in H_{x_0}$, extended convexity of X implies that there exists $\gamma > 0$ such that $x_0 \pm \gamma h \in X$; that is, $\pm\gamma h \in H_{x_0}$. By linearity of the supergradient and Gateaux derivative with respect to h, we have

$$\pm\gamma\{\delta f(x_0; h) + z^*[\delta G(x_0, h)]\} = \delta f(x_0; \pm\gamma h) + z^*[\delta G(x_0, \pm\gamma h)].$$

The last equation together with equation (B.8) gives

$$\delta f(x_0; h) + z^*[\delta G(x_0, h)] = 0. \qquad \square$$

COROLLARY B.6. *Using a proof similar to that of Theorem B.5, we can replace assumption (i) of that theorem by one of the following assumptions:*
a) G has a linear Gateaux derivative and f has a linear Gateaux derivative at x_0.
b) G has a linear supergradient and f has a linear Gateaux derivative at x_0.
c) G has a linear supergradient and f has a linear supergradient at x_0.

Next, we will prove below a corollary to Theorem B.5 for processes. We will start with a variation of the definition of extended convex that is appropriate to a collection of stochastic processes.

DEFINITION B.7. A collection of vector-valued stochastic processes X will be called extended convex if for all $x_1, x_2 \in X$, there is a process $\delta = \delta(\omega, t; x_1, x_2) > 0$ such that for each process $\alpha = \alpha(\omega, t)$ that satisfies $-\delta \leq \alpha \leq 1 + \delta$, we have

$$\alpha x_1 + (1 - \alpha)x_2 \in X.$$

Observe that the concept just defined is essentially a special case of weakly extended convex (see Definition B.3) with respect to the functionals $\Omega \times [0, T]$. Namely, each (ω, t) represents a functional given by: $x \to x(\omega, t)$, for each process $x \in X$.

REMARK B.8. An example of an extended convex set X is the following:

$$X = \{x(t, \omega) \in A_{t,\omega}, 0 \leq t \leq T, \omega \in \Omega\} \cap \bar{\mathcal{L}}_2\left(\mathbb{R}^k\right),$$

where $A_{t,\omega}$ is a convex and open set in $\mathbb{R}^k$ containing $\mathbb{R}^k_{++}$.

We note that the set of consumption plans $\mathcal{C}$ is a special case of the remark.

We next define the concepts of *supergradient density* and *gradient density* for real-valued functionals defined on a collection of processes.

DEFINITION B.9. Let $X \subseteq \bar{\mathcal{L}}_2(\mathbb{R}^n)$ be an extended convex subset of processes. Let $\nu : X \to \mathbb{R}$ be a functional. For any $x_0 \in X$,

(a) The process $\pi \in \bar{\mathcal{L}}_2(\mathbb{R}^n)$ is a *supergradient density* of ν at x_0 if

$$\nu(x_0 + h) - \nu(x_0) \leq (\pi | h) \quad \text{for all } h \in H_{x_0}.$$

(b) $\pi \in \bar{\mathcal{L}}_2(\mathbb{R}^n)$ is a *gradient density* of ν at x_0 if

$$(\pi | h) = \lim_{\alpha \downarrow 0} \frac{\nu(x_0 + \alpha h) - \nu(x_0)}{\alpha} \quad \text{for all } h \in H_{x_0}.$$

COROLLARY B.10. *Let $X \subseteq \bar{\mathcal{L}}_2(\mathbb{R}^n)$ be an extended convex set of processes. Let $G : X \to \mathbb{R}^m$ and $f : X \to \mathbb{R}$. Let $x_0 \in X$ satisfy $G(x_0) \leq 0$ and $f(x_0) = \min_{\{x \in X:\, G(x) \leq 0\}}\{f(x)\}$. Assume:*

(i) G has a gradient density, and f has a supergradient density or a gradient density at x_0.

(ii) x_0 is a regular point of $\{G(x) \leq 0\}$.

Then, there is a $z^* \in \mathbb{R}^m_+$ such that

$$\delta f(x_0, h) + z^{*\prime}[\delta G(x_0, h)] = 0, \quad \textit{all } h \in H_{x_0},$$
$$z^{*\prime}[G(x_0)] = 0.$$

Proof. Recall that all the processes that we deal with are assumed to be progressively measurable. We will prove the corollary under the assumption that f has a supergradient density at x_0. The proof goes exactly as Theorem B.5 through equation (B.8). Fix some $h \in H_{x_0}$. Because X is extended convex, there exists a $\gamma \in \mathcal{L}(\mathbb{R}_{++})$ such that $x_0 \pm \gamma h \in X$. For each $\varepsilon > 0$, define $A(\varepsilon) = \{(\omega, t) : \gamma(\omega, t) > \varepsilon\}$. Because 1_A is a progressively measurable process and $\varepsilon 1_{A(\varepsilon)} < \gamma$, we have $x_0 \pm \varepsilon h 1_{A(\varepsilon)} \in X$. As in the proof of Theorem B.5, we get (from equation (B.1))

$$\delta f(x_0; h 1_{A(\varepsilon)}) + z^{*'}[\delta G(x_0, h 1_{A(\varepsilon)})] = 0.$$

Since the gradient and supergradient densities are continuous in the increment and $h 1_{A(\varepsilon)} \to h$ in $\mathcal{L}_2(\mathbb{R}^n)$ as $\varepsilon \to 0$, we have

$$\delta f(x_0; h) + z^{*\prime}[\delta G(x_0, h)] = 0.$$

The proof is similar when we assume that f has gradient density at x_0. □

APPENDIX C: DERIVATIONS FOR SOME EXAMPLES

C.1 Derivation of Example 3.2

We first show existence of a solution to (3.2). Consider, for any $\delta > 0$, the BSDE system

$$dz^i_t = \exp(-a^{i\prime} M c_t)\left(\left|z^j_t\right| \vee \delta\right)^{-a^i_j} dt - z^i_t \Sigma^{i\prime}_t dB, \quad z^i_T = -\exp\left(-\left(M^i c_T + \zeta^i\right)\right),$$

$$(\text{C.1}) \quad i \in \{1, 2\}, \quad j \neq i.$$

The assumptions of uniform boundedness and $a^i_j \geq -1$ for all $j \neq i$ imply that the aggregator satisfies the conditions for existence and uniqueness in Pardoux and Peng (1990). It follows from equation (3.2) and the fact that $\exp(-(M^i c_T + \zeta^i)) > \varepsilon$ for some $\varepsilon > 0$, and the aggregator is positive, that the solution must satisfy $|z^i_t| \geq \varepsilon$ for all i and t; therefore, if we choose $\delta \in (0, \varepsilon)$, we obtain the same solution (and it must be unique) to the BSDE system

$$dz^i_t = \exp\left(-a^{i\prime} Mc_t\right) \left|z^j_t\right|^{-a^i_j} dt - z^i_t \Sigma^{i\prime}_t dB_t, \quad z^i_T = -\exp\left(-\left(M^i c_T + \zeta^i\right)\right),$$
$$(\text{C.2}) \qquad i \in \{1, 2\}, \quad j \neq i.$$

Pardoux and Peng (1990) imply that $z^i \Sigma^i$ is square integrable for each i, and therefore z satisfies (3.2).

Ito's Lemma applied to $Z^i_t = -\ln(-z^i_t)$, together with $|z^j_t|^{-a^i_j} = \exp(a^i_j Z^i_t)$ imply that the (C10) is equivalent to

$$dZ^i_t = -\psi^i(t, Mc_t - Z_t, \Sigma_t)dt + \Sigma^{i\prime}_t dB_t, \quad Z^i_T = M^i c_T + \zeta^i,$$

with ψ^i defined by (3.3).

C.2 Derivation of Example 3.9

We first sketch the derivation of (3.16). Each agent i's effort process e^i_t increases the drift of the Brownian motion (by either changing the probability measure, or directly changing the drift). Agent effort is assumed noncontractible,[24] and each agent considers only the impact of effort on own utility. Given pay process $c^i \in \mathcal{C}$, agent i's utility function, given the effort processes of the other agents, under optimal agent-i effort is[25]

$$dZ^i_t = -\max_{e^i_t} \left\{ h^i\left(t, x^i_t\right) - \frac{1}{2}\Sigma^{i\prime}_t Q^{Ri} \Sigma^i_t - \frac{1}{2} e^{i\prime}_t Q^{ei} e^i_t + \Sigma^{i\prime}_t \sum_{j=1}^{N-1} e^j_t \right\} dt + \Sigma^{i\prime}_t dB_t,$$
$$(\text{C.3}) \qquad Z^i_T = c^i_T.$$

The function h^i primarily governs intertemporal substitution, and the quadratic penalty terms model risk-aversion and effort disutility. The penalty matrices Q^{Ri}, $Q^{ei} \in \mathbb{R}^{d\times d}$ are assumed positive definite. Agent i's optimal effort is $\hat{e}^i_t = (Q^{ei})^{-1} \Sigma^i_t$, and therefore agent utility under optimal effort satisfies

$$dZ^i_t = -\left\{ h^i\left(t, x^i_t\right) - \frac{1}{2}\Sigma^{i\prime}_t Q^i \Sigma^i_t + \Sigma^{i\prime}_t \sum_{j=1}^{N-1} q^j \Sigma^j_t \right\} dt + \Sigma^{i\prime}_t dB_t,$$

[24] More precisely, the principal's controls are not adapted to the natural filtration of the effort processes.

[25] Using Briand and Hu (2008), a sufficient restriction on the effort processes to obtain existence and uniqueness for each agent's utility BSDE is the Novikov condition

$$E\left\{ \exp\left(\frac{1}{2} \int_0^T \left\| \sum_{j=1}^{N-1} e^j_t \right\|^2 dt \right) \right\} < \infty.$$

(C.4) $$Z^i_T = c^i_T,$$

where $Q^i = Q^{Ri} + (Q^{ei})^{-1}$ and $q^i = (Q^{ei})^{-1}$, $i = 1, \dots, N-1$.

The principal receives the cash flow process C, out of which the agent salaries are paid. The principal's utility BSDE (Z^N, Σ^N) is also affected by agent effort, though she exerts no effort herself:

$$dZ^N_t = -\left\{ h^N\left(t, C_t + x^N_t\right) - \frac{1}{2}\Sigma^{N\prime}_t Q^N \Sigma^N_t + \Sigma^{N\prime}_t \sum_{j=1}^{N-1} q^j \Sigma^j_t \right\} dt + \Sigma^{N\prime}_t dB_t,$$

(C.5) $$Z^N_T = C_T - \mathbf{1}' c_T.$$

To obtain a simple closed-form solution, we assume the affine form of h^N in (3.16) (infinite elasticity of intertemporal substitution). Using Theorem 3.8, and substituting the constraint $1' x_t = -Y_t$, optimal intertemporal pay minus utility is obtained as the maximizing argument of

$$\max_{\{x^1_t, \dots, x^{N-1}_t\}} h^N\left(t, C_t - Y_t - \sum_{j=1}^{N-1} x^j_t\right) + \sum_{j=1}^{N-1} h^j\left(t, x^{Uj}_t\right) = \alpha + \beta(C_t - Y_t) + \eta_t,$$

where η is defined in (3.17). The optimal diffusion controls are (substituting $\Sigma^Y_t = \Sigma^N_t + \sum_{j=1}^{N-1} \Sigma^j_t$)

(C.6) $$\{\hat{\Sigma}^1_t, \dots, \hat{\Sigma}^{N-1}_t\} = \arg \max_{\{\Sigma^1_t, \dots, \Sigma^{N-1}_t\}} -\frac{1}{2}\left(\Sigma^Y_t - \sum_{j=1}^{N-1}\Sigma^j_t\right) Q^N \left(\Sigma^Y_t - \sum_{j=1}^{N-1}\Sigma^j_t\right) - \frac{1}{2}\sum_{j=1}^{N-1}\Sigma^{j\prime}_t Q^j \Sigma^j_t + \Sigma^{Y\prime}_t \sum_{j=1}^{N-1} q^j \Sigma^j_t,$$

which has the FOC (necessary and sufficient)

$$Q^i \Sigma^i_t = Q^N \left(\Sigma^Y_t - \sum_{j=1}^{N-1}\Sigma^j_t\right) + q^i \Sigma^Y_t, \quad i = 1, \dots, N-1.$$

Defining

$$\bar{Q}^a = \sum_{j=1}^{N-1} (Q^j)^{-1} Q^N, \quad \bar{Q}^b = \sum_{j=1}^{N-1} (Q^j)^{-1} q^j$$

and summing over i, we get

$$\sum_{i=1}^{N-1} \Sigma^i_t = (I + \bar{Q}^a)^{-1}(\bar{Q}^a + \bar{Q}^b)\Sigma^Y_t,$$

and therefore, the optimal diffusion controls are

$$\hat{\Sigma}^i_t = (Q^i)^{-1}\{Q^N (I + \bar{Q}^a)^{-1}\{I - \bar{Q}^b\} + q^i\}\Sigma^Y_t.$$

The BSDE (3.10) for (Y, Σ^Y) is then

$$\text{(C.7)}\quad dY_t = -\left\{\alpha + \eta_t + \beta(C_t - Y_t) - \frac{1}{2}\Sigma_t^{Y\prime} Q^Y \Sigma_t^Y\right\} dt + \Sigma_t^{Y\prime} dB_t, \qquad Y_T = C_T,$$

where Q^Y is defined in (3.20). We can confirm the quadratic term $-\frac{1}{2}\Sigma_t^{Y\prime} Q^Y \Sigma_t^Y$ by writing the right side of (C11) as

$$-\frac{1}{2}\left(\Sigma_t^Y - \sum_{j=1}^{N-1}\hat{\Sigma}_t^j\right) Q^N \left(\Sigma_t^Y - \sum_{j=1}^{N-1}\hat{\Sigma}_t^j\right) - \frac{1}{2}\sum_{j=1}^{N-1}\hat{\Sigma}_t^{j\prime}\left(Q^j\hat{\Sigma}_t^j - q_t^j\Sigma_t^Y\right)$$
$$+\frac{1}{2}\hat{\Sigma}_t^{Y\prime}\sum_{j=1}^{N-1} q^j\hat{\Sigma}_t^j = -\frac{1}{2}\Sigma_t^{Y\prime}\left\{Q^N - \sum_{j=1}^{N-1}(Q^N + q^j)w^j\right\}\Sigma_t^Y,$$

where the equality follows from using the FOC to replace $Q^i\hat{\Sigma}_t^i - q^i\Sigma_t^Y$ with $Q^N(\Sigma_t^Y - \sum_{j=1}^{N-1}\hat{\Sigma}_t^j)$.

Finally, it is easy to confirm that (3.19) (with $\Sigma_t^Y = \sigma$) solves the BSDE (C12).

C.3 Derivation of Example 4.5

We first show that a uniform bound on $|C|$ implies a uniform bound on the candidate $|c^i|$ in (4.6). It is well known that (4.5) has a unique solution, and the assumption $q_j^i \geq 0$ for all $i \neq j$ (recall that $\varepsilon_0 = \beta \in \mathbb{R}_{++}^N$) implies that $\varepsilon_t^i > 0$ for all i and t. Define $q^{\max} = \max_{i,j}\{q_i^j\}$, $q^{\min} = \min_{i,j}\{q_i^j\}$, and $\Lambda_t = \sum_{j=1}^N \varepsilon_t^j$. Then,

$$Nq^{\min}\Lambda_t dt \leq d\Lambda_t \leq Nq^{\max}\Lambda_t dt$$

implies

$$\exp(Nq^{\min}t)\Lambda_0 \leq \Lambda_t \leq \exp(Nq^{\max}t)\Lambda_0.$$

It follows that the αs are uniformly bounded above and away from zero (note that $\bar{q}_j^i > 0$ for all i, j):

$$\min_{i,j}\{\bar{q}_i^j\}\exp(Nq^{\min}t)\Lambda_0 \leq \alpha_i(t) \leq \max_{i,j}\{\bar{q}_i^j\}\exp(Nq^{\max}t)\Lambda_0,$$

and therefore, there exists some $m \in \mathbb{R}_+$ such that

$$\text{(C.8)}\qquad \left(\alpha_i(t)\sum_{j=1}^N \frac{1}{\alpha_j(t)}\right)^{-1} \leq m.$$

Therefore, the candidate $|c^i|$ is uniformly bounded.

The proof is completed by applying Proposition 4.3 to $\mathcal{C} = \bar{\mathcal{L}}((-n, n)^N)$ for any sufficiently large n.

More generally, we can relax the uniformly bounded C along the lines of Remark 2.2. Defining

$$\mathcal{C} = \left\{ c \in \mathcal{L}\left(\boldsymbol{C}^k\right) : c \text{ continuous and } E\left(\sup_{t\in[0,T]} \|c_t\|^4\right) < \infty \right\},$$

then $C \in C$ implies that c given by (5.1) is optimal within $\mathcal{C}$.

C.4 Derivation of Example 5.2

Note that the generalization discussed in Remark 2.2 does not help with this example because Φ_c here blows up at zero consumption.

Define $\hat{c}$ by the scalar-valued process (5.7), with κ chosen to exhaust the budget. Because $\hat{c} \in \mathcal{C}$ is budget feasible, we have (the arguments of Z^a are a's and b's consumption processes, respectively)

$$\max_{\{c\in\mathcal{C}:(\pi|c)=w^a\}} Z^a(c,\hat{c}) \geq Z^a(\hat{c},\hat{c}). \tag{C.9}$$

Theorem 2.1 in El Karoui et al. (1997) implies uniqueness and existence of $Z^a(c,\hat{c})$ for any $c \in \mathcal{C}$.

Now define $\hat{c}_n$ as the solution (5.7) with the same κ, but with π_t replaced with $\pi_n(t) = \min(\pi_t, n)$, $t \in [0, T]$. Note that $\hat{c}_n$ is bounded from below and decreases monotonically to $\hat{c}$ as $n \to \infty$. Also, the initial wealth w_n required to finance $\hat{c}_n$, given by $w_n = (\pi_n|\hat{c}_n)$, satisfies $w_n \downarrow w_0$ as $n \uparrow \infty$.

LEMMA C.1. *Given prices π_n and initial wealth w_n, consumption $\hat{c}_n$ is optimal for a given consumption $\hat{c}_n$ by b:*

$$Z^a(\hat{c}_n,\hat{c}_n) = \max_{\{c\in\mathcal{C}:(\pi_n|c)\leq w_n\}} Z^a(c,\hat{c}_n), \qquad n > 0.$$

Proof. For any $m > 0$, define the consumption space $\mathcal{C}_m = \bar{\mathcal{L}}_4((\frac{1}{m},\infty)^N)$. Because the aggregators are Lipschitz in consumption within $\mathcal{C}_m$, we apply Theorem 2.4 to obtain

$$Z^a(\hat{c}_n,\hat{c}_n) = \max_{\{c\in\mathcal{C}_m:(\pi_n|c)\leq w_n\}} Z^a(c,\hat{c}_n)$$

for any sufficiently large m. Concavity of $Z^i()$, $i \in \{a, b\}$ (from Lemma 2.7)[26] implies that $\hat{c}_n$ is also optimal within the larger set $\mathcal{C}$. To show this, suppose $\tilde{c} \in \mathcal{C}$ satisfies $(\pi_n|\tilde{c}) \leq w_n$ and $Z^a(\tilde{c},\hat{c}_n) > Z^a(\hat{c}_n,\hat{c}_n)$. Then, $c^\lambda = \lambda\hat{c}_n + (1-\lambda)\tilde{c}$ satisfies $c^\lambda \geq \lambda/m$ and $Z^a(c^\lambda,\hat{c}_n) \geq \lambda Z^a(\hat{c}_n,\hat{c}_n) + (1-\lambda)Z^a(\tilde{c},\hat{c}_n)$. But this is a contradiction because c^λ is also budget feasible and uniformly bounded away from zero. □

Using the above lemma, together with monotonicity of Z^a in agent b's consumption,[27] and the fact that the budget-feasible set of consumption process under prices π_n and

[26]Despite the absence of bounded Φ_c and $\bar{\Phi}_c$ for all $c \in \mathcal{C}$, the proof of concavity from Lemma 2.7 still holds in this application because Φ_c and $\bar{\Phi}_c$ on the right side of the gradient inequalities are both evaluated at $c^\lambda \geq \lambda/m$ and the BSDE solution at c^λ.

[27]Similar to the concavity result above, the proof of $Z^a(\hat{c}_n,\hat{c}_n) \geq Z^a(\hat{c}_n,\hat{c})$ from Lemma 2.7 still applies because Φ_c and $\bar{\Phi}_c$ on the right side of the gradient inequalities are both evaluated at $\hat{c}_n$ and the BSDE solution at $\hat{c}_n$.

initial wealth w_n is strictly larger than under π and w^a (because $(\pi|c) \leq w^a$ implies $(\pi_n|c) \leq w_n$) yields

$$Z^a(\hat{c}_n, \hat{c}_n) \geq \max_{\{c\in\mathcal{C}:(\pi|c)\leq w^a\}} Z^a(c, \hat{c}) \geq Z^a(\hat{c}, \hat{c}).$$

Noting that

$$Z^a(\hat{c}_n, \hat{c}_n) = \frac{1}{1-\gamma} E\left(\int_0^T e^{\int_0^t[\alpha_s+\beta_s]ds}\hat{c}_n(t)^{1+\mu-\gamma}\,dt + e^{\int_0^T[\alpha_s+\beta_s]ds}\hat{c}_n(t)^{1+\mu-\gamma}\right),$$

we apply the monotone convergence theorem to get

$$\lim_{n\to\infty} Z^a(\hat{c}_n, \hat{c}_n) = Z^a(\hat{c}, \hat{c}).$$

Therefore, $\max_{\{c\in\mathcal{C}:(\pi|c)\leq w^a\}} Z^a(c, \hat{c}) = Z^a(\hat{c}, \hat{c})$. The same argument applies to agent b, which establishes the Nash equilibrium.

REFERENCES

ABEL, A. (1990): Asset Prices under Habit Formation and Catching Up with the Joneses, *Am. Econ. Rev. Papers Proc.* 80, 38–42.

ANDERSON, E., L. HANSEN, and T. SARGENT (2000): Robustness, Detection and the Price of Risk, Working Paper, Department of Economics, University of Chicago.

BRIAND, P., and F. CONFORTOLA (2008): Differentiability of Backward Stochastic Differential Equations in Hilbert Spaces with Monotone Generators, *App. Math. Optim.* 57(2), 149–176.

BRIAND, P., and Y. HU (2008): Quadratic BSDEs with Convex Generators and Unbounded Terminal Conditions, *Probab. Theory Related Fields* 141, 543–567.

CADENILLAS, A., and I. KARATZAS (1995): The Stochastic Maximum Principle for Linear, Convex Optimal Control with Random Coefficients, *Siam J. Control Optim.* 33, 590–624.

CHEN, Z., and L. EPSTEIN (2002): Ambiguity, Risk, and Asset Returns in Continuous Time, *Econometrica* 70, 1403–1443.

COX, J., and C.-F. HUANG (1989): Optimal Consumption and Portfolio Policies When Asset Prices Follow a Diffusion Process, *J. Econ. Theory* 49, 33–83.

CVITANIC, J., and J. ZHANG (2012): *Contract Theory in Continuous-Time Models*, Heidleberg, New York, Dordrecht, London: Springer Finance.

DUFFIE, D., and L. G. EPSTEIN (1992): Stochastic Differential Utility, *Econometrica* 60, 353–394.

DUFFIE, D., P.-Y. GEOFFARD, and C. SKIADAS (1994): Efficient and Equilibrium Allocations with Stochastic Differential Utility, *J. Math. Econ.* 23, 133–146.

DUFFIE, D., and C. SKIADAS (1994): Continuous-Time Security Pricing: A Utility Gradient Approach, *J. Math. Econ.* 23, 107–131.

EL KAROUI, N., S. PENG, and M.-C. QUENEZ (1997): Backward Stochastic Differential Equations in Finance, *Math. Finance* 7, 1–71.

EL KAROUI, N., S. PENG, and M.-C. QUENEZ (2001): A Dynamic Maximum Principle for the Optimization of Recursive Utilities under Constraints, *Ann. Appl. Probab.* 11, 664–693.

GALÍ, J. (1994): Keeping Up with the Joneses: Consumption Externalities, Portfolio Choice, and Asset Prices, *J. Money Credit Bank.* 26, 1–8.

GEIB, C., and R. MANTHEY (1994): Comparison Theorems for Stochastic Differential Equations in Finite and Infinite Dimensions, *Stoch. Process. Appl.* 53, 23–35.

GOMEZ, J. P., R. PRIESTLEY, and F. ZAPATERO (2009): Implications of Keeping Up with the Joneses Behavior for the Equilibrium Cross Section of Stock Returns: International Evidence, *J. Finance* 64, 2703–2737.

GOUKASIAN, L., and X. WAN (2010): Optimal Incentive Contracts under Relative Income Concerns, *Math. Financ. Econ.* 4, 57–88.

HOLMSTROM, B., and P. MILGROM (1987): Aggregation and Linearity in the Provision of Intertemporal Incentives, *Econometrica* 55, 303–328.

INGERSOLL, J. E. (2011): *Time-Additive Consumtion-Wealth Utility*, Working Paper, Yale School of Management, International Center for Finance.

KARATZAS, I., J. LEHOCZKY, and S. SHREVE (1987): Optimal Portfolio and Consumption Decisions for a "Small Investor" on a Finite Horizon, *SIAM J. Control Optim.* 25, 1557–1586.

KOO, H. K., G. SHIM, and J. SUNG (2008): Optimal Multi-Agent Performance Measures for Team Contracts, *Math. Finance* 18(4), 649–667.

LAZRAK, A., and M. C. QUENEZ (2003): A Generalized Stochastic Differential Utility, *Math. Oper. Res.* 28, 154–180.

LEVENTAL, S., M. SCHRODER, and S. K. SINHA (2011): The Continuous-Time Principal-Agent Problem with Moral Hazard and Recursive Preferences, Working Paper, Michigan State University.

LUENBERGER, D. G. (1969): *Optimization by Vector Space Methods*, New York: Wiley.

MAENHOUT, P. (1999): Robust Portfolio Rules and Asset Pricing, Working Paper, INSEAD.

PARDOUX, E., and S. PENG (1990): Adapted Solution of a Backward Stochastic Differential Equation, *Syst. Control Lett.* 14, 55–61.

SCHATTLER, H., and J. SUNG (1993): The First-Order Approach to the Continuous-Time Principal-Agent Problem with Exponential Utility, *J. Econ. Theory* 61, 331–371.

SCHRODER, M., and C. SKIADAS (2002): An Isomorphism between Asset Pricing Models with and without Linear Habit Formation, *Rev. Financ. Stud.* 15, 1189–1221.

SCHRODER, M., and C. SKIADAS (2003): Optimal Lifetime Consumption-Portfolio Strategies under Trading Constraints and Generalized Recursive Preferences, *Stoch. Process. Appl.* 108, 155–202.

SCHRODER, M., and C. SKIADAS (2005): Lifetime Consumption-Portfolio Choice under Trading Constraints and Nontradeable Income, *Stoch. Process. Appl.* 115, 1–30.

SKIADAS, C. (1992): *Advances in the Theory of Choice and Asset Pricing*, PhD thesis, Stanford University.

ST-AMOUR, P. (2005): *Direct Preference for Wealth in Aggregate Household Portfolios, FAME Research Paper Series rp136*, International Center for Financial Asset Management and Engineering.

WU, Z., and M. XU (2009): Comparison Theorems for Forward Backward SDEs, *Stat. Probab. Lett.* 79, 426–435.