Dynamic Portfolio Choice and Risk Aversion

Costis Skiadas*

September 2, 2005
Working Paper No. 331
prepared for the *Handbook of Financial Engineering*

1 Introduction

This chapter analyzes the optimal consumption-portfolio choice of a risk-averse agent, with emphasis on the modeling of risk aversion given a stochastic investment opportunity set. The main part of the analysis is based on Schroder and Skiadas (2003), hereafter abbreviated to SS03. A novel contribution of this chapter is a decision theoretic development of the notions of source-dependent first or second order risk aversion that are implicit in the utility representations of SS03. These ideas unify, at least in the context of continuous information, standard notions of risk aversion with some models of ambiguity aversion or robustness that have recently received considerable attention in the literature. The dynamic portfolio methodology presented should, however, also be of interest to readers only concerned with conventional source-independent risk aversion in a dynamic setting.

Following Merton's (1969, 1971) seminal work, most papers on dynamic portfolio choice assume that the investor maximizes expected time-additive state-independent utility, that we refer to as "additive utility" for the purposes of this discussion. While additive utility is adequate in models with i.i.d. returns, it is well-known (see, for example, Epstein (1992)) that it is overly restrictive in more general stochastic settings arising in asset pricing models with predictability, stochastic volatility, or varying risk premia. For example, we will argue that any two (continuous) additive utilities that imply identical preferences over deterministic plans must be ordinally equivalent, and therefore equally risk averse. In this chapter, we consider utility functions for which risk aversion can be adjusted without changing the utility value of deterministic plans.

In order to focus on the role of risk aversion and limit the chapter's length, our setting will be restricted to one in which information is revealed continuously by a finite set of Brownian motions, there is a single nondurable and infinitely divisible consumption good, the planning horizon is exogenous, a consumption nonnegativity constraint does not bind, and there are no transaction costs or trading constraints. Moreover, while markets can be incomplete, they are sufficiently complete so that the investor's endowed income stream is tradeable. The last section will point

*Kellogg School of Management, Department of Finance, Northwestern University, 2001 Sheridan Road, Evanston, IL 60208. I am grateful to Mark Schroder for the many years of collaborative research on which this article is based, as well as corrections to this chapter. I also thank Darrell Duffie, Flavio de Andrade, Ali Lazrak, Jared Williams, and Jacob Sagi for valuable feedback. The latest version of this article can be found at www.kellogg.nwu.edu/faculty/skiadas/home.htm

to extensions in the literature relaxing various combinations of these assumptions (typically at the cost of other restrictions).

In its simplest form, the utility function we will adopt is the Duffie and Epstein (1992) continuous-information limit of the recursive utility of Kreps and Porteus (1978), which includes the widely used homothetic recursive utility specification of Epstein and Zin (1989) (a special case of which is expected discounted power or logarithmic utility[1]). In the Kreps-Porteus formulation, current utility is computed as a function of current consumption and a von Neumann-Morgenstern (1944) certainty equivalent of the continuation utility. Given sufficient smoothness, the classic analysis of small risks of Arrow (1965, 1970) and Pratt (1964) implies that the certainty equivalent can be replaced by a quadratic criterion in the continuous-information limit, which is the reason why some elements of the original single-period mean-variance portfolio analysis of Markowitz (1952) survive in a continuous-information setting with Duffie-Epstein utility. Assuming constant relative risk aversion, the optimal portfolio is a weighted sum of an instantaneously mean-variance efficient portfolio and a hedging portfolio that accounts for the stochastic nature of the investment opportunity set (and vanishes in the case of i.i.d. instantaneous returns).

An extension of Duffie-Epstein utility we will consider allows risk aversion to depend on the source of risk. For example, investors have been documented to show a preference toward investing in the familiar: domestic stocks, firms whose products are familiar, local firms, one's employer's stock.[2] The well-known experimental findings of Ellsberg (1961), and a large literature following it, show that subjects prefer to bet on risk sources to which probabilities can be more unambiguously assigned, a phenomenon known as ambiguity aversion.[3] One can think of risk as reflecting not only the risk that is conditional on the assumed model of the risk source, but also uncertainty about the model's validity, which is itself too difficult to model. Since model risk can vary with the source of risk, it is useful to consider source-dependent risk aversion. With this motivation, we will extend the Kreps-Porteus recursion by letting the certainty equivalent be a function of the entire vector of continuation utilities attributable to each Brownian motion separately. The locally-quadratic analysis under Duffie-Epstein utility extends to this case, but with a different coefficient of risk aversion assigned to each source of risk.

As shown in Skiadas (2003), Duffie-Epstein utility includes the "robust" specifications of Anderson, Hansen, and Sargent (2000), Hansen, Sargent, Turmuhambetova, and Williams (2001) and Maenhout (1999). Similarly, a simple extension of the argument in Skiadas (2003) shows that the criterion of Uppal and Wang (2003) is equivalent to a form of recursive utility with source dependent risk aversion (included in the "quasi-quadratic proportional aggregator" specification of SS03). The multiple-prior expressions of these authors suggest a robustness interpretation of risk aversion. Conversely, their robustness interpretation of multiple-prior formulations can be thought of as conventional risk aversion in the context of recursive-utility. To avoid this semantic redundancy, in this chapter we define formally only risk aversion, and we think of robustness or ambiguity aversion as an informal motivation for calibrating risk aversion toward a given source of risk.

Another way in which the Duffie-Epstein representation will be extended relates to the distinction between first and second order risk aversion made in a static setting by Segal and Spivak (1990). The Arrow-Pratt analysis, and by extension the Duffie-Epstein limit of Kreps-Porteus util-

[1] Under some regularity, a homothetic additive utility is necessarily the additive special case of Epstein-Zin utility. On the other hand, Epstein-Zin utility is only a parametric special case of the much broader class of homothetic Duffie-Epstein utilities.

[2] Daniel, Hirshleifer, and Teoh (2002) survey such psychological biases in asset markets.

[3] The view of ambiguity aversion as a form of risk aversion is further supported by the arguments of Klibanoff, Marinacci, and Mukerji (2002).

ity, relies on the smoothness of the von Neumann-Morgenstern certainty equivalent, an assumption for which there is no compelling decision-theoretic justification. Smooth expected utility implies local risk-neutrality, meaning that an investor should be willing to undertake any actuarially favorable investment in sufficiently small scale, and should be unwilling to perfectly insure a sufficiently small risk at actuarially unfavorable terms. We will consider a source-dependent extension of the Kreps-Porteus utility with non-smooth certainty equivalent for which these conclusions are invalidated, and we will derive corresponding optimal trading strategy expressions that highlight the relationship between first-order risk aversion and portfolio holdings.

Motivated by the notion of ambiguity aversion, Epstein and Schneider (2003) formulated a multiple prior utility, whose continuous-information limit was studied by Chen and Epstein (2002). Consistent with the view of ambiguity aversion as being a form of risk aversion, the Chen-Epstein "κ-ignorance" formulation is mathematically equivalent to a case of the above mentioned extension of Duffie-Epstein utility with source-dependent first-order risk aversion. More generally, Lazrak and Quenez (2003) analyzed the properties of a utility that is defined as a solution to a general Backward Stochastic Differential Equation (BSDE), and includes the Chen-Epstein formulation. Lazrak and Quenez made the important observation that comparative risk aversion can depend on the "direction" of risk, which leads to the interpretation of the Chen-Epstein notion of ambiguity aversion as plain risk aversion. Complementing the Lazrak-Quenez analysis, this chapter provides a heuristic decision-theoretic foundation of their proposed utility form, that we will refer to simply as "recursive utility." The more specific models of risk aversion discussed above correspond to special functional forms of recursive utility.

Following the development of SS03, optimality conditions will first be derived for general concave recursive utilities, as a system of forward-backward stochastic differential equations (FBSDEs). The forward component of the system is the wealth process, which follows the investor's budget equation, and the backward components are the utility and shadow-price-of-wealth processes. The FBSDE system uncouples if the problem is scale-invariant (with respect to wealth). Combining scale-invariance with the various types of risk aversion discussed above, we will be able to formulate some interesting optimal trading strategy expressions, in terms of the solution to a single BSDE. Moreover, we will give some examples of preferences and stochastic investment opportunity sets for which the BSDE of the optimality conditions takes a quadratic form whose solution can be reduced to a tractable ODE system. A parallel theory based on translation-invariant recursive utility (which generalizes expected discounted exponential utility) can be found in Schroder and Skiadas (2005a), and is briefly discussed in the final section.

Merton approached the dynamic optimal portfolio selection problem using the Hamilton-Jacobi-Bellman equation of optimal control theory, modern expositions of which are given by Fleming and Soner (1993) and Yong and Zhou (1999). Examples of solutions with Epstein-Zin utility using this method are Giovannini and Weil (1989), Svensson (1989), Obstfeld (1994), Zariphopoulou and Tiu (2002), and Chacko and Viceira (forthcoming). Cox and Huang (1989) and Karatzas, Lehoczky, and Shreve (1987) rederived the Merton solution by using the state-price density property of marginal utilities at the optimum, in a way that relied on utility additivity. This "utility gradient approach" was generalized to include recursive utilities in Duffie and Skiadas (1994), Schroder and Skiadas (1999), El Karoui, Peng, and Quenez (2001), and Schroder and Skiadas (2003, 2005), and is the method adopted in this chapter. (An alternative dynamic programming derivation of the scale-invariant solutions is outlined in SS03.) While some further leads to the literature will be given in the final section, this chapter is not intended as a literature survey, and no attempt has been made to be comprehensive. Monographs or collected papers on dynamic portfolio choice include Merton

(1990), Korn (1997), Sethi (1997), Karatzas and Shreve (1998), Gollier (2001), and Campbell and Viceira (2002). An overview of the econometrics of portfolio choice is given by Brandt (forthcoming).

The mathematical background for this chapter is outlined in the appendices of Duffie (2001), and is covered in detail by Karatzas and Shreve (1988). Less widely known are the more recent mathematical tools of BSDEs and FBSDEs, a general perspective on which can be found in the expositions of El Karoui, Peng, and Quenez (1997) and Ma and Yong (1999).

The remainder of this chapter is organized in five sections. Section 2 sets up the problem and characterizes the optimum with minimal restrictions on preferences over consumption plans. Section 3 develops recursive utility, and the associated optimality conditions. Section 4 motivates some more specialized recursive utility forms, representing the various types of risk aversion introduced above. Section 5 formulates optimality conditions for these special recursive utility forms, assuming utility homotheticity. Section 6 concludes with an outline of several extensions.

2 Optimality and State Pricing

This section introduces the stochastic setting, the investor problem, and the basic optimality verification argument in terms of the state price density property of a utility supergradient density. The section imposes only minimal preference assumptions, and will be specialized to recursive utility in subsequent sections. The essential tool of linear BSDEs will be introduced in the context of state pricing. The section concludes with a discussion of the inadequacy of additive utility as a representation of risk aversion, which motivates the discussion of recursive utility to follow.

2.1 Dynamic Investment Opportunity Set

Uncertainty is represented by the probability space $(\Omega, \mathcal{F}, P)$, on which is defined a d-dimensional Brownian motion $B = \left(B^1, \ldots B^d\right)'$ over a finite time-horizon $[0, T]$. As with every vector in this chapter, B is a column vector, and the prime denotes transposition. Information is represented by the (augmented) filtration $\{\mathcal{F}_t : t \in [0, T]\}$ generated by the Brownian motion B. Intuitively, we think of an information tree whose time-t nodes or *spots* correspond to the possible paths of B up to time t. A *time-t spot* is therefore a continuous function of the form $\omega^t : [0, t] \to \mathbb{R}^d$. Conditional expectation given time-t information, $\mathcal{F}_t$, is denoted E_t. Similarly, covariance (variance) given $\mathcal{F}_t$ is denoted cov_t (var_t). We assume that $\mathcal{F} = \mathcal{F}_T$, and therefore $E_T[x] = x$ for every random variable x.

A *process* in this chapter is by definition a stochastic process that is progressively measurable with respect to $\{\mathcal{F}_t\}$. For any process x, we think of the time-t value x_t (alternatively denoted $x(t)$) as a function of the realized spot ω^t. In heuristic explanations (that ignore issues regarding null sets) we will write $x[\omega^t]$ to express this dependence. Given any subset S of some Euclidean space, we let $\mathcal{L}(S)$ denote the set of processes of the form $x : \Omega \times [0, T] \to S$. For any integer p, typically $p = 1$ or 2, we define the set $\mathcal{L}_p(S)$ of all $x \in \mathcal{L}(S)$ such that $\int_0^T |x_t|^p \, dt < \infty$ with probability one (where $|\cdot|$ denotes Euclidean norm).

We consider a financial market allowing instantaneous default-free borrowing and lending at a continuously-compounded rate given by the process r. A dollar can be invested from time t to time $t + dt$ earning interest $r_t dt$, which is risk-free in the sense that $\mathrm{Var}_t[r_t dt] = 0$, but whose value depends on time-t information. For expositional simplicity, r is assumed bounded (although this assumption is violated in some later applications). The rest of the market consists of trading in m risky assets, whose instantaneous cumulative *excess* returns are represented by the m-dimensional process $R = \left(R^1, \ldots, R^m\right)'$, in the sense that a dollar invested at time t in risky asset i is worth $1 + r_t dt + dR_t^i$ at time $t + dt$.

We assume that R is an Ito process with dynamics

$$dR_t = \mu_t^R dt + \sigma_t^{R\prime} dB_t, \quad \mu^R \in \mathcal{L}_1\left(\mathbb{R}^m\right), \quad \sigma^R \in \mathcal{L}_2\left(\mathbb{R}^{d\times m}\right). \tag{1}$$

There is, therefore, one column of σ^R for every risky asset, and one row for every component of the Brownian motion B. The investment opportunity set is defined by the triple $\left(r, \mu^R, \sigma^R\right)$, whose value can vary from spot to spot. We think of (1) as an instantaneous linear factor model, where

$$\mu_j^R(t)\, dt = E_t\left[dR_t^j\right] \quad \text{and} \quad \sigma_{ij}^R(t)\, dt = \operatorname{cov}_t\left[dB_t^i, dR_t^j\right], \quad i = 1, \ldots, d, \quad j = 1, \ldots, m.$$

Since $E_t[dB_t] = 0$ and $E_t[dB_t dB_t'] = I dt$ (where I is an identity matrix) the conditional variance-covariance matrix of dR_t is

$$E_t\left[\left(dR_t - E_t[dR_t]\right)\left(dR_t - E_t[dR_t]\right)'\right] = \sigma_t^{R\prime}\sigma_t^R dt.$$

A time-t *allocation* is an $\mathcal{F}_t$-measurable random vector $\psi_t = \left(\psi_t^1, \ldots, \psi_t^m\right)'$, where ψ_t^i represents the proportion of wealth invested at time t in risky asset i, with the remaining non-consumed wealth invested risk-free. Negative proportions indicate short positions. The choice of a time-t allocation can depend on time-t information, and therefore we think of ψ_t as a function of the realized time-t spot. A dollar invested at time t according to allocation ψ_t is worth

$$1 + r_t dt + \psi_t' dR_t = 1 + \left(r_t + \psi_t' \mu_t^R\right) dt + \left(\sigma_t^R \psi_t\right)' dB_t$$

at time $t + dt$. The vector $\sigma_t^R \psi_t$ represents the *risk profile* of the allocation ψ_t, since it specifies the loadings of the instantaneous excess return $\psi_t' dR_t$ on the instantaneous factors dB_t.

If the column span of σ^R is $\mathbb{R}^d$ at all times then the market is complete, in the sense that every risk profile is feasible through some allocation at all times. We do *not* assume that the market is complete, allowing the rank of σ^R to be less than d. While the market can be effectively complete even if σ^R drops rank, we will consider applications in which market incompleteness is a binding constraint. We will not allow, however, the rank of σ^R to vary from spot to spot, and we assume that at no spot of the information tree are any of the assets redundant over an infinitesimal time interval. This is the economic content of the following condition, assumed throughout:

Asset non-redundancy: *The columns of σ^R, corresponding to the m risky assets, are everywhere linearly independent, and therefore $m \leq d$.*

As a consequence of this assumption, the $m \times m$ instantaneous variance-covariance rate matrix $\sigma^{R\prime}\sigma^R$ is everywhere invertible. If σ_t is a risk profile attainable through the allocation ψ_t, meaning that $\sigma_t^R \psi_t = \sigma_t$, then ψ_t is the unique allocation with this property, and is given by

$$\psi_t = \left(\sigma_t^{R\prime}\sigma_t^R\right)^{-1} \sigma_t^{R\prime}\sigma_t. \tag{2}$$

The traditional portfolio analysis of Markowitz (1952) can be applied conditionally spot-by-spot on the information tree. Selecting an allocation ψ_t results in an instantaneous excess return with conditional mean and variance

$$E_t\left[\psi_t' dR_t\right] = \psi_t' \mu_t^R dt \quad \text{and} \quad \operatorname{var}_t\left[\psi_t' dR_t\right] = \psi_t' \sigma_t^{R\prime}\sigma_t^R \psi_t \; dt.$$

Let μ_t be any $\mathcal{F}_t$-measurable random variable. Minimizing $\text{var}_t\left[\psi_t' dR_t\right]$ subject to the constraint $E_t\left[\psi_t' dR_t\right] = \mu_t dt$ results in an allocation of the form

$$\psi_t = k_t \left(\sigma_t^{R\prime}\sigma_t^R\right)^{-1}\mu_t^R,$$

for some $\mathcal{F}_t$-measurable random variable k_t that depends on the choice of μ_t. We call an allocation of this form *instantaneously minimum-variance efficient*. The corresponding squared conditional instantaneous Sharpe ratio is maximized, and is given by

$$\frac{E_t\left[\psi_t' dR_t\right]^2}{\text{var}_t\left[\psi_t' dR_t\right]} = \mu_t^{R\prime}\left(\sigma_t^{R\prime}\sigma_t^R\right)^{-1}\mu_t^R dt. \tag{3}$$

2.2 Strategies, Utility, and Optimality

An optimal investment strategy is one that finances a consumption plan for which there exists no other consumption plan that is both more desirable and feasible. In this subsection we formalize this notion, while placing minimal restrictions on investor preferences.

We let $\mathcal{H}$ denote the Hilbert space of every $x \in \mathcal{L}(\mathbb{R})$ such that $E\left[\int_0^T x_t^2 dt + x_T^2\right] < \infty$, with the inner product

$$(x|y) = E\left[\int_0^T x_t y_t dt + x_T y_T\right], \quad x, y \in \mathcal{H}.$$

The set of *strictly positive*[4] elements of $\mathcal{H}$ is $\mathcal{H}_{++} = \mathcal{H} \cap \mathcal{L}(\mathbb{R}_{++})$. The element of $\mathcal{H}$ that is identically equal to one is denoted $\mathbf{1}$.

We postulate a convex cone $\mathcal{C} \subseteq \mathcal{H}_{++}$ of *consumption plans* such that $\mathbf{1} \in \mathcal{C}$, and for any $x \in \mathcal{H}$ and $y, z \in \mathcal{C}$, $y \leq x \leq z$ implies $x \in \mathcal{C}$. For any $c \in \mathcal{C}$ and time $t < T$, we interpret c_t as the time-t consumption rate, while c_T represents a terminal lump-sum consumption or bequest. In a typical application, $\mathcal{C}$ is specified by some integrability restriction required for a utility function to be well-defined. The strict positivity of consumption plans reflects our implicit assumption that a consumption nonnegativity constraint is nonbinding. In later sections, the positivity of optimal consumption will be enforced by assuming infinite marginal utility at zero.

We consider an investor with initial wealth $w_0 > 0$ and no subsequent income. (This includes the case of an endowed income stream as long as it can be traded.) A *consumption strategy* is any process $\rho \in \mathcal{L}_1(\mathbb{R}_{++})$ such that $\rho_T = 1$. For $t < T$, we interpret ρ_t as the investor's consumption rate as a proportion of time-t wealth, while the convention $\rho_T = 1$ is used below to express the assumption that final wealth equals terminal consumption. A *trading strategy* is any process $\psi \in \mathcal{L}(\mathbb{R}^m)$ such that $\psi'\mu^R \in \mathcal{L}_1(\mathbb{R})$ and $\sigma^R\psi \in \mathcal{L}_2\left(\mathbb{R}^d\right)$, with ψ_t representing a time-t allocation. A *strategy* is a pair (ρ, ψ) of a consumption strategy and a trading strategy.

The *wealth process* W *generated* by a strategy (ρ, ψ) is defined through the *budget equation*

$$\frac{dW_t}{W_t} = (r_t - \rho_t)\, dt + \psi_t' dR_t, \quad W_0 = w_0. \tag{4}$$

The consumption plan c is *financed* by the strategy (ρ, ψ) if $c = \rho W$ (meaning that $c_t = \rho_t W_t$ for every time t, and therefore $c_T = W_T$). A consumption plan is *feasible* if it can be financed by some strategy.

[4] More precisely, any two processes x, y such that $(x - y|x - y) = 0$ are identified as elements of $\mathcal{H}$. A *strictly positive* element of $\mathcal{H}$ is one that can be identified in this way with a process in $\mathcal{L}(\mathbb{R}_{++})$.

The investor's problem is to select a feasible consumption plan that is optimal. To define optimality, we introduce utility functions. We say that the investor *prefers* plan b to plan a at spot ω^t if, conditionally on the realization of ω^t, an agent with plan a as the status quo would switch to plan b if presented with the opportunity to do so at no cost. The investor is *indifferent* between two plans if neither plan is preferred to the other.

We are going to measure utility concretely by taking as a unit the consumption plan $\mathbf{1}$. We assume throughout that the investor prefers more consumption to less, and therefore, given any scalars α, β such that $\beta > \alpha > 0$, the agent prefers $\beta\mathbf{1}$ to $\alpha\mathbf{1}$ at every spot. We further assume that, given any consumption plan c and spot ω^t, there exists a (necessarily unique) scalar α such that, conditionally on the realization of spot ω^t, the agent is indifferent between plans c and $\alpha\mathbf{1}$. We call this value of α the *spot-ω^t cardinal utility* of c, and denote it $U(c)[\omega^t]$. Specifying a value at every spot of the information tree defines the *cardinal utility process* $U(c)$ of plan c. We note that, by definition, $U_T(c) = c_T$.

Another preference assumption we adopt is that if the investor is indifferent between a and a' and between b and b', then the investor prefers b to a if and only if the investor prefers b' to a'. Applying this condition with $a' = U(a)[\omega^t]\mathbf{1}$ and $b' = U(b)[\omega^t]\mathbf{1}$, we conclude that, conditionally on the realization of spot ω^t, the investor prefers plan b to plan a if and only if $U(b)[\omega^t] > U(a)[\omega^t]$. The investor's objective at spot ω^t is therefore to select the feasible consumption plan c of maximum spot-ω^t utility $U(c)[\omega^t]$.

Utility maximization at every spot can be an inconsistent objective, since the investor may have an incentive to deviate at some spot from a strategy selected at an earlier spot. We exclude this possibility by assuming the following key condition throughout.

Dynamic Consistency: *Suppose two consumption plans a and b are equal up to a stopping time τ, and $P[U_\tau(b) \geq U_\tau(a)] = 1$. Then $U_0(b) \geq U_0(a)$, with the inequality being strict if it is also the case that $P[U_\tau(b) > U_\tau(a)] > 0$.*

Suppose time-zero utility is maximized by the strategy (ρ, ψ), which finances the consumption plan c, and generates the wealth process W. Then there cannot exist a stopping time τ and trading strategy $(\tilde{\rho}, \tilde{\psi})$, that finances consumption plan $\tilde{c}$ and generates a wealth process $\tilde{W}$, such that $W_\tau = \tilde{W}_\tau$, $P[U_\tau(\tilde{c}) \geq U_\tau(c)] = 1$, and $P[U_\tau(\tilde{c}) > U_\tau(c)] > 0$. Otherwise, by dynamic consistency, the strategy that starts as (ρ, ψ) and switches to $(\tilde{\rho}, \tilde{\psi})$ at time τ would result in higher time-zero utility than $U_0(c)$, contradicting the time-zero optimality of strategy (ρ, ψ).

Dynamic consistency justifies the following definition of optimality in terms of the single time-zero utility function $U_0 : \mathcal{C} \to \mathbb{R}$.

Definition 1 *The consumption plan c is* optimal *if it is feasible and there exists no feasible consumption plan $\tilde{c}$ such that $U_0(\tilde{c}) > U_0(c)$. A strategy (ρ, ψ) is* optimal *if it finances an optimal consumption plan. Finally, a consumption or trading strategy is* optimal *if it is part of an optimal strategy.*

A function $\tilde{U}_0 : \mathcal{C} \to \mathbb{R}$ is *ordinally equivalent* to $U_0 : \mathcal{C} \to \mathbb{R}$ if $\tilde{U}_0 = f \circ U_0$ for some strictly increasing function $f : \mathbb{R} \to \mathbb{R}$. We call such a function $\tilde{U}_0$ an *ordinal utility* representation of the investor's time-zero preferences. Optimality of a plan c is not affected if the cardinal utility U_0 in the above definition is replaced by any ordinally equivalent utility. A property of U_0 is *ordinal* if it is also true of any utility that is ordinally equivalent to U_0. Dynamic consistency is an example of an ordinal property.

Throughout this chapter, we fix a utility function $U_0 : \mathcal{C} \to \mathbb{R}$ that can be either ordinal or cardinal, the distinction being made only where relevant. In addition we assume:

Monotonicity: For any $c, c + x \in \mathcal{C}$, $0 \neq x \geq 0$ implies $U_0(c + x) > U_0(c)$.

Concavity: For all $c^0, c^1 \in \mathcal{C}$, $\alpha \in (0, 1)$ implies $U_0\left(\alpha c^1 + (1 - \alpha) c^0\right) \geq \alpha U_0\left(c^1\right) + (1 - \alpha) U_0\left(c^0\right)$.

Monotonicity is an ordinal property, while concavity is not. For cardinal utility, concavity can be thought of as an expression of a preference for consumption smoothing. Later we will introduce the important class of scale-invariant problems in which U_0 is assumed to have the additional ordinal property of homotheticity. A cardinal utility is homothetic if and only if it is homogeneous of degree one, in which case concavity is equivalent to the ordinal property of quasiconcavity.

Let (ρ, ψ) be a candidate optimal strategy that generates the wealth process W and finances the consumption plan $c = \rho W$. We will verify the optimality of c by constructing a utility supergradient density at c that is also a state price density at c. These notions are defined below.

Definition 2 *(a) A process $\pi \in \mathcal{H}$ is a* state price density at c *if $(\pi|x) \leq 0$ for any $x \in \mathcal{H}$ such that $c + x$ is a feasible consumption plan.*
(b) A process $\pi \in \mathcal{H}$ is a supergradient density *of U_0 at c if $U_0(c + x) \leq U_0(c) + (\pi|x)$ for every $x \in \mathcal{H}$ such that $c + x \in \mathcal{C}$.*

Interpreting $(\pi|x)$ as a present value of x, the state-price density property states that there is no feasible incremental consumption plan relative to c that has positive present value. A supergradient density can be thought of as a generalized notion of marginal utility. Since U_0 is assumed (strictly) increasing and concave, any supergradient density of U_0 is necessarily strictly positive. Given a reference plan, the state-price density property depends on the market opportunities and not on preferences, while the supergradient density property depends on preferences and not on the market opportunities.

The following observation is the basis for the optimality verification arguments in this chapter.

Proposition 3 *Suppose c is a feasible consumption plan, and $\pi \in \mathcal{H}_{++}$ is a supergradient density of U_0 at c that is also a state price density at c. Then the plan c is optimal.*

Proof. If $c + x \in \mathcal{C}$ is feasible, then $U_0(c + x) \leq U_0(c) + (\pi|x) \leq U_0(c)$. ■

Remark 4 *We will not discuss the necessity of optimality conditions in this chapter. A simple partial converse to the above proposition is given in SS03.*

2.3 State Price Dynamics and linear BSDEs

In order to apply the optimality verification argument of Proposition 3, we study below the dynamics of a state price density. In the process we introduce the mathematical tool of a linear Backward Stochastic Differential Equation (BSDE), which plays a basic role in this chapter and asset pricing theory in general.

The key to understanding the state price density dynamics is the following notion of risk pricing:

Definition 5 *A* market-price-of-risk process *is any process $\eta \in \mathcal{L}_2\left(\mathbb{R}^d\right)$ such that*

$$\mu^R = \sigma^{R\prime}\eta. \tag{5}$$

Recalling the linear-factor-model interpretation (1), the above equation can be thought of as (exact) factor pricing, with η_t^i representing the time-t price of instantaneous linear factor dB_t^i. Since σ^R is assumed everywhere full-rank, a market-price-of-risk process is unique if and only if $m = d$.

The existence of a market-price-of-risk process is implied by the absence of arbitrage opportunities. While a rigorous statement and proof of this claim can be found in Karatzas and Shreve (1998), it is worth recalling the essential idea. In an arbitrage-free market there cannot be an instantaneously riskless allocation with positive instantaneous excess returns; that is,

$$\sigma_t^R \psi_t = 0 \quad \text{implies} \quad \psi_t' \mu_t^R = 0. \tag{6}$$

The existence of a market price of risk process is the dual equivalent to (6). Clearly, (5) implies (6). Conversely, we define the orthogonal decomposition $\mu_t^R = \sigma_t^{R\prime}\eta_t + \varepsilon_t$, where $\sigma_t^R \varepsilon_t = 0$. If (6) holds, then $\varepsilon_t' \mu_t^R = 0$, and therefore $\varepsilon_t'\varepsilon_t = \varepsilon_t'\mu_t^R = 0$, proving that $\mu_t^R = \sigma_t^{R\prime}\eta_t$.

Suppose that the process $\pi \in \mathcal{H}_{++}$ follows the dynamics

$$\frac{d\pi_t}{\pi_t} = -r_t dt - \eta_t' dB_t, \quad t \in [0, T], \tag{7}$$

for some market-price-of-risk process η. We will argue that π is a state-price density at any given consumption plan satisfying an integrability condition.

Consider any strategy (ρ, ψ), generating the wealth process W, and financing the consumption plan $c = \rho W$. Letting $\Sigma = W\sigma^R \psi$ in the budget equation (4) and using the assumption $\mu^R = \sigma^{R\prime}\eta$ results in

$$dW_t = -\left(c_t - r_t W_t - \eta_t' \Sigma_t\right) dt + \Sigma_t' dB_t, \quad W_T = c_T. \tag{8}$$

This is a *linear BSDE*. The Ito process W *solves* the BSDE if (8) is satisfied for some $\Sigma \in \mathcal{L}_2\left(\mathbb{R}^d\right)$. Given the solution W, the corresponding $\Sigma \in \mathcal{L}_2\left(\mathbb{R}^d\right)$ is uniquely determined (by the uniqueness of Ito representations) and therefore we can also think of a solution as being the pair $(W, \Sigma) \in \mathcal{L}_1(\mathbb{R}) \times \mathcal{L}_2\left(\mathbb{R}^d\right)$. Nonlinear BSDEs are introduced in the following section, where it is explained that a BSDE is essentially a backward recursion on the information tree. For the linear case, the backward recursion interpretation is suggested by a present value formula given in the lemma below. Even though the symbols have specific meanings in this context, the lemma is stated in a way that applies to a general linear BSDE.

Lemma 6 *Suppose W solves BSDE* (8) *for some $c \in \mathcal{H}$, $r \in \mathcal{L}_1(\mathbb{R})$, and $\eta \in \mathcal{L}_2(\mathbb{R})$. Let $\pi \in \mathcal{H}_{++}$ be any strictly positive process with dynamics* (7).

(a) If $W \in \mathcal{L}(\mathbb{R}_+)$, then

$$W_t \geq \frac{1}{\pi_t} E_t \left[\int_t^T \pi_s c_s ds + \pi_T c_T \right], \quad t \in [0, T]. \tag{9}$$

(b) If $E\left[\sup_t \pi_t W_t\right] < \infty$, then

$$W_t = \frac{1}{\pi_t} E_t \left[\int_t^T \pi_s c_s ds + \pi_T c_T \right], \quad t \in [0, T]. \tag{10}$$

Proof. Suppose (W, Σ) satisfies (8). Integration by parts gives $d(\pi W) = -\pi c dt + \;...dB$. Let $\{\tau_n\}$ be an increasing sequence of stopping times converging to T almost surely, and such that the

...dB term stopped at τ_n is a martingale. Integrating the last equation from t to T, and applying the operator E_t on both sides, we find

$$\pi_t W_t = E_t \left[\int_t^{\tau_n} \pi_s c_s ds + \pi_{\tau_n} W_{\tau_n} \right].$$

If $W \geq 0$, we can take the limit as $n \to \infty$ and apply Fatou's lemma to conclude (9). If $E[\sup_t \pi_t W_t] < \infty$, then we can apply dominated convergence to conclude (10). ■

Remark 7 *Conversely, if W is given by* (10)*, then W solves BSDE* (8)*. This can be shown by rearranging* (10)*, and using integration by parts and a martingale representation theorem.*

In our context, where W is the wealth process generated by a strategy financing the consumption plan c, the above lemma implies the state-price-density property of π :

Proposition 8 *Suppose $\pi \in \mathcal{H}_{++}$ follows the dynamics* (7) *for a market-price-of-risk process η. If $E[\sup_t \pi_t W_t] < \infty$, then π is a state price density at c.*

Proof. Suppose $c + x$ is a feasible consumption plan. By Lemma 6, $\pi_0 w_0 \geq (\pi | c + x)$ and $\pi_0 w_0 = (\pi | c)$. Therefore, $(\pi | x) \leq 0$. ■

Remark 9 *The necessity of condition* (7) *for an Ito process $\pi \in \mathcal{H}_{++}$ to be a state price density at c is shown, under some regularity assumptions, in SS03, where the characterization is also extended to allow for trading constraints. For example, necessity follows if $\mathcal{C} = \mathcal{H}_{++}$ and $c \in \mathcal{C}$ is continuous.*

In Lemma 6, we saw that the linear term $rW + \eta'\Sigma$ in BSDE (8) corresponds to stochastic discounting in the present value formula (10). Alternatively, the two terms can be interpreted separately, with rW corresponding to temporal discounting and $\eta'\Sigma$ corresponding to a change of measure. To see how, we define, given any $\eta \in \mathcal{L}_2(\mathbb{R}^d)$, the processes ξ^η and B^η by

$$\frac{d\xi_t^\eta}{\xi_t^\eta} = -\eta_t' dB_t, \quad \xi_0^\eta = 1, \quad \text{and} \quad dB_t^\eta = dB_t + \eta_t dt, \quad B_0^\eta = 0. \tag{11}$$

We recall that ξ^η is a positive supermartingale, and is a martingale if and only if $E\xi_T^\eta = 1$. Suppose $\eta \in \mathcal{L}_2(\mathbb{R}^d)$ is such that ξ^η is a martingale. In this case an equivalent-to-P probability measure P^η, with expectation operator E^η, is well-defined through the change-of-measure formula $E^\eta[x] = E[\xi_T x]$ (or $dP^\eta/dP = \xi_T$). By Girsanov's theorem, B^η is standard Brownian motion under P^η. The linear BSDE (8) can equivalently be stated as

$$dW_t = -(c_t - r_t W_t)\, dt + \Sigma_t' dB_t^\eta, \quad W_T = c_T.$$

Lemma 6 and Remark 7 imply that if $E^\eta \left[\sup_t \exp\left(-\int_0^t r_\tau d\tau \right) |W_t| \right] < \infty$, then W solves BSDE (8) if and only if

$$W_t = E_t^\eta \left[\int_t^T e^{-\int_t^s r_\tau d\tau} c_s ds + e^{-\int_t^T r_\tau d\tau} c_T \right]. \tag{12}$$

Equation (12) is the familiar risk-neutral-pricing version of the present-value formula (10), stating that financial wealth is equal to the present value of the future cash flow that this wealth finances. In a Markovian setting, such a present value can be computed (under some regularity) in terms of a corresponding PDE solution, sometimes referred to as the Feynman-Kac solution (see Duffie (2001)). The PDE form can be derived by writing W as a function of time and the underlying Markov state, applying Ito's lemma, and matching terms with the linear BSDE (8). This type of construction applies more generally to BSDEs, and can be used to characterize optimal portfolios, as will be outlined for a class of scale-invariant solutions in subsection 5.

2.4 Expected Time-Additive Utility and What's Wrong with It

Having understood the structure of state price dynamics, which is unrelated to preferences, we turn our attention to the utility side. Our objective is to specify some utility functional structure that properly captures a notion of risk aversion, and then compute the supergradient density dynamics. Combining the latter with the state price dynamics will result in optimality conditions.

A widely used functional form of the time-zero utility function $U_0 : \mathcal{C} \to \mathbb{R}$ is

$$U_0(c) = E\left[\int_0^T e^{-\beta t} u(c_t)\, dt + e^{-\beta T} \upsilon(c_T)\right], \tag{13}$$

for some $\beta \in \mathbb{R}$ and concave increasing functions $u, \upsilon : \mathbb{R}_{++} \to \mathbb{R}$. The more concave u is, the more risk averse the utility. An advantage of this specification is that a supergradient density can be computed separately at each spot, simplifying the investor problem, at least under complete markets. For example, suppose that (13) holds with $u = \upsilon$, the derivative u' exists and maps $\mathbb{R}_{++}$ onto $\mathbb{R}_{++}$, and the optimal consumption plan c satisfies $u'(c) \in \mathcal{H}_2$. It is straightforward to check that the process $e^{-\beta t} u'(c_t)$ is a supergradient density of U_0 at c. If the market is complete ($m = d$), then there exists a unique state price density π with $\pi_0 = 1$, given by the dynamics (7) with $\eta = \sigma^{R\prime -1} \mu^R$. The optimal consumption is $c_t = u'^{-1}\left(e^{\beta t} k \pi_t\right)$, where the scalar $k > 0$ is selected so that $(\pi | c) = w_0$. The corresponding wealth process W is given by the present value formula (10). If $dW/W = ...dt + \sigma' dB$, then we have seen that the corresponding optimal trading strategy ψ is given by (2). This is essentially the analysis of the Merton problem by Cox and Huang (1989) and Karatzas, Lehoczky, and Shreve (1987). (Later in this chapter, we will see that much of the simplicity of the above argument is lost if markets are incomplete.)

We argue that, despite its popularity, the time-additive utility specification (13) is fundamentally flawed as a representation of risk aversion, which is a good reason for investing some effort in studying recursive utility. Emphasizing the temporal aspect of consumption, we focus in the remainder of this section on preferences over consumption plans with fixed terminal consumption or bequest, and we therefore assume that (13) holds with $\upsilon = 0$. We show below that in this case the investor's preferences over deterministic choices determine, up to ordinal equivalence, the investor's entire utility function, and in particular the investor's risk aversion. On the other hand, we will see that with recursive utility two investors can have identical preference in a deterministic environment, and yet one investor can be more risk averse than the other.

The main result below utilizes the following standard uniqueness result from additive representation theory. A proof can be found in Narens (1985) or Wakker (1989).

Lemma 10 *For any integer $N > 1$, and each $i \in \{1, 2\}$, suppose the function $F^i : \mathbb{R}^N_{++} \to \mathbb{R}$ has the additive structure $F^i(x_1, \ldots, x_N) = \sum_{n=1}^N f^i_n(x_n)$, $x \in \mathbb{R}^N_{++}$, where $f^i_n : \mathbb{R}_{++} \to \mathbb{R}$, $n = 1, \ldots, N$. Suppose also that F^1 and F^2 are ordinally equivalent, meaning that $F^1(x) \geq F^1(y)$ if and only if $F^2(x) \geq F^2(y)$. Then there exists $\alpha \in \mathbb{R}_{++}$ and $\beta \in \mathbb{R}^N$ such that $f^1_n = \alpha f^2_n + \beta_n$, $n = 1, \ldots, N$.*

Proposition 11 *For each* $i \in \{1,2\}$*, suppose the utility function* $U_0^i : \mathcal{C} \to \mathbb{R}$ *takes the form* $U_0^i(c) = E\int_0^T v^i(t, c_t)\,dt$*, where* $v^i : [0,T] \times \mathbb{R}_{++} \to \mathbb{R}$ *is continuous. Suppose also that, for any deterministic consumption plans*[5] $a, b \in \mathcal{C}$*,* $U^1(a) \geq U^1(b)$ *if and only if* $U^2(a) \geq U^2(b)$*. Then the utility functions* U^1 *and* U^2 *are ordinally equivalent on the entire space* $\mathcal{C}$*.*

Proof. After replacing U^i with $U^i - U^i(\mathbf{1})$, we can and do assume that $v^i(t,1) = 0$ for all t. Given any integer $N > 1$, we define the time intervals $J^n = [(n-1)T/N, nT/N)$, $n = 1, \ldots, N$, partitioning $[0,T)$. Let D_N be the set of deterministic plans of the form $\sum_{n=1}^N x_n 1_{J^n}$. Since U_0^1 and U_0^2 order the elements of D_N the same, we can apply the above lemma with $f_n^i(x) = \int_{J^n} v^i(t,x)\,dt$ to conclude that, for some $\alpha \in \mathbb{R}_{++}$ and all n,

$$\int_{J^n} v^1(t,x)\,dt = \alpha \int_{J^n} v^2(t,x)\,dt. \tag{14}$$

Repeating the above argument with $2N$ in place of N results in the same relationship, with the same constant α, since D_N can be embedded into D_{2N}. For any $x > 0$, we can therefore take a sequence of intervals $\{J^n : n = 1, 2, \ldots\}$ containing x, whose length converges to zero and (14) holds for all n. Dividing both sides of (14) by the length of J_n and taking the limit as $n \to \infty$, we conclude that $v^1(t,x) = \alpha v^2(t,x)$. ■

The limitation of additive utility in capturing risk aversion is illustrated in the following variant of what seems to be a folklore example (which I learned from Duffie and Epstein (1992)).[6]

Example 12 *Suppose that* $T = 100$ *and* $U_0(c) = E\left[\int_0^{100} v(t, c_t)\,dt\right]$ *for some continuous function* $v : [0,100] \times \mathbb{R}_{++} \to \mathbb{R}$*. The plans* a *and* b *are defined by*

$$a_t = 1 + 1{,}000 \times 1_{\{t>1,\ B_1>0\}} \quad \text{and} \quad b_t = 1 + 1{,}000 \times \sum_{n=1}^{99} 1_{\{1+n \geq t > n,\ B_n - B_{n-1} > 0\}}.$$

While $Ea_t = Eb_t$ *for all* t*, one could reasonably argue that plan* b *is less risky than plan* a*. Yet, it is straightforward to check that* $U_0(a) = U_0(b)$*.*

3 Recursive Utility

In this section, we motivate and define recursive utility, and we derive its basic properties. By combining a computation of the utility supergradient dynamics with last section's state price dynamics, we will obtain optimality conditions under recursive utility as a FBSDE system. Finally, we will introduce homothetic recursive utility and its role in uncoupling the FBSDE system of the optimality conditions. Continuous-time recursive utility was first defined and analyzed by Duffie and Epstein (1992), who imposed some special structure that is useful in modeling risk aversion. Following Lazrak and Quenez (2003), we adopt a broader definition of recursive utility as a solution to a general BSDE. In the following section, we will see that the broader definition of recursive utility allows for interesting models of risk aversion that go beyond the Duffie-Epstein specification.

[5] We call a process x *deterministic* if x_t is $\mathcal{F}_0$-measurable for every time t.

[6] In their introduction, Duffie and Epstein (1992) give another example of the limitation of additive utility that is based on the notion of preferences for the timing of resolution of uncertainty of Kreps and Porteus (1978). The notion was extended in Skiadas (1998) in terms of preferences over pairs of consumption plans and information streams (filtrations). Additivity relates to the non-dependence of utility on the filtration argument.

3.1 Recursive Utility and BSDEs

We begin with a heuristic derivation from general principles of recursive utility. The argument should also help clarify the interpretation of a BSDE as a continuous-time representation of a backward recursion on an information tree.

We consider a dynamic utility function $U : \mathcal{C} \to \mathbb{R}$, where $U(c)$ is an Ito process for every $c \in \mathcal{C}$. In addition to our earlier assumptions of dynamic consistency, monotonicity, and concavity, we impose the following simplifying restriction.

Irrelevance of Past or Unrealized Consumption: For any consumption plans a and b, any time $t \leq T$, and any event $A \in \mathcal{F}_t$, if $a = b$ on[7] $A \times [t, T]$, then $U(a) = U(b)$ on $A \times [t, T]$.

This assumption is not an essential aspect of a recursive utility structure, but serves as a natural benchmark in an analysis whose main focus is risk aversion. Together with dynamic consistency, it implies that, for any consumption plan c and times $t < u$, the restriction of $U_t(c)$ on a time-t event A can be expressed as a function of the restriction of c on $A \times [t, u]$ and the restriction of $U_u(c)$ on A. More formally, we can show:[8]

Lemma 13 *Given any times $t < u \leq T$ and event $A \in \mathcal{F}_t$, suppose that the consumption plans a and b are equal on $A \times [t, u]$ and $U_u(a) = U_u(b)$ on A. Then $U_t(a) = U_t(b)$ on A.*

Proceeding heuristically, we apply the above functional relationship with the time-event (t, A) corresponding to a single spot ω^t and $u = t + dt$, where dt is an infinitesimal time-interval. Fixing any $c \in \mathcal{C}$, we let $U = U(c)$. Given the instantaneous factor decomposition

$$U_{t+dt} = m_t + \Sigma_t' dB_t, \quad \text{where}$$
$$m_t = E_t\left[U_{t+dt}\right] \quad \text{and} \quad \Sigma_t^i = \operatorname{cov}_t\left[U_{t+dt}, dB_t^i\right], \quad i = 1, \ldots, d, \tag{15}$$

we obtain the functional restriction

$$U_t = \Phi(t, c_t, m_t, \Sigma_t), \tag{16}$$

for some (possibly state-dependent) function $\Phi : \Omega \times [0, T] \times \mathbb{R}_{++} \times \mathbb{R}^{1+d} \to (0, \infty)$ that is adapted to the underlying information structure. Utility monotonicity and concavity heuristically imply[9] that $\Phi(\omega, t, c, m, \Sigma)$ is increasing in (c, m) and concave in (c, m, Σ). Given U_{t+dt}, equation (16), with m_t and Σ_t defined in (15), is used to compute U_t. Equation (16) is therefore a heuristic backward recursion on the information tree, which determines the entire utility process U given the terminal value U_T.

[7] This means that the indicator of $\{(\omega, u) : \omega \in A,\ u \in [t, T],\ a(\omega, u) \neq b(\omega, u)\}$ is zero as an element of $\mathcal{H}$.

[8] **Proof:** Let $D = A \cap \{U_t(a) > U_t(b)\}$, and define the stopping times $\sigma = t1_D + T1_{\Omega \setminus D}$ and $\tau = u1_D + T1_{\Omega \setminus D}$. We define the plan a' (resp. b') to be equal to a (resp. b) on $D \times [t, T]$, and some arbitrary plan c outside $D \times [t, T]$. Since $a = a'$ on $[\sigma, T]$, we have $U(a) = U(a')$ on $[\sigma, T]$, and therefore $U_\tau(a') = U_\tau(a)$ a.s. Analogously, $U_\tau(b') = U_\tau(b)$ a.s., and therefore $U_\tau(a') = U_\tau(b')$ a.s. Since a' and b' are equal up to the stopping time τ, dynamic consistency implies $U_0(a') = U_0(b')$. On the other hand, a' and b' are equal up to σ, $U_\sigma(a') > U_\sigma(b')$ on D, and $U_\sigma(a') = U_\sigma(b')$ on $\Omega \setminus D$. If $P(D) > 0$, then dynamic consistency would imply $U_0(a') > U_0(b')$, a contradiction. Therefore $P(D) = 0$. This shows $U_t(a) \leq U_t(b)$ on A. The reverse inequality is true by symmetry.

[9] The idea is that the dependence of $\Phi(\omega, t, c, m, \Sigma)$ on (c, m, Σ) is through the pair (c, U), where $U = m + \Sigma' dB$. One can heuristically identify (c, U) with a plan that is equal to c at spot ω^t (corresponding to (ω, t)), takes the value U on $[t + dt, T]$ conditionally on spot ω^t having occured, and it takes, say, the value one at all remaining spots. Utility monotonicity and concavity over the set of such plans translates to the corresponding properties for $\Phi(\omega, t, \cdot)$.

To formulate a rigorous version of the utility recursion, we assume that the function F, called an (infinitesimal) aggregator, is implicitly defined, at any state ω and time $t < T$, by

$$\mu = -F(\omega, t, c, U, \Sigma) \iff U = \Phi(\omega, t, c, U + \mu dt, \Sigma). \tag{17}$$

By monotonicity of Φ in the conditional mean argument, there is at most one value μ satisfying the right-hand side equation in (17), and therefore F is uniquely determined given Φ. Moreover, the monotonicity and concavity properties of Φ imply that $F(\omega, t, c, U, \Sigma)$ is increasing in c and concave in (c, U, Σ). (If Φ is strictly concave in m it also follows[10] that F is decreasing in U. We will not need to assume this condition, although it is helpful in verifying technical regularity conditions.) We use the notation $U_T = F(T, c_T)$ to express the dependence of terminal utility on terminal consumption (which is the identity for cardinal utility).

Assuming the Ito decomposition $dU = \mu dt + \Sigma' dB$, and therefore $m = U + \mu dt$, recursion (16) is equivalent to the drift restriction $\mu_t = -F(t, c_t, U_t, \Sigma_t)$, resulting in the utility dynamics

$$dU_t = -F(t, c_t, U_t, \Sigma_t)\, dt + \Sigma_t' dB_t, \quad U_T = F(T, c_T). \tag{18}$$

Equation (18) is a BSDE to be solved jointly in the (adapted) processes pair (U, Σ). The function $f(\omega, t, y, z) = F(\omega, t, c(\omega, t), y, z)$, is known as the BSDE *driver*. We say that the Ito process U *solves* BSDE (18) if there exists a (necessarily unique) $\Sigma \in \mathcal{L}_2(\mathbb{R}^d)$ such that (18) is satisfied.

Example 14 (Expected Discounted Utility) *In the above heuristic argument, suppose*

$$\begin{aligned}\Phi(\omega, t, c, m, \Sigma) &= u(\omega, t, c)\, dt + m \exp(-\beta(\omega, t)\, dt), \\ F(\omega, t, c, U, \Sigma) &= u(\omega, t, c) - \beta(\omega, t) U, \quad t < T, \quad F(\omega, T, c) = u(\omega, T, c).\end{aligned}$$

By Lemma 6, under a regularity assumption, the solution BSDE (18) *is*

$$U_t = E_t\left[\int_t^T \exp\left(-\int_t^s \beta_\tau d\tau\right) u(s, c_s)\, ds + \exp\left(-\int_t^T \beta_\tau d\tau\right) u(T, c_T)\right].$$

Initial BSDE existence and uniqueness results, based on the type of Lipschitz-growth assumptions on the driver familiar from SDE theory, were first obtained by Pardoux and Peng (1990) and Duffie and Epstein (1992). (An improved version of the Pardoux-Peng argument can be found in El Karoui, Peng, and Quenez (1997).) These conditions are violated in our main homothetic application to follow, which includes the widely used Epstein-Zin utility (a continuous-time version of the recursive utility parametrization used in Epstein and Zin (1989)). Existence, uniqueness and basic properties for continuous-time Epstein-Zin utility were shown in Appendix A of Schroder and Skiadas (1999). BSDE theory has been further developed by Hamadene (1996), Lepeltier and Martin (1997, 1998, 2001) , Kobylanski (2000), and others. Issues of existence and uniqueness will not be further addressed in this chapter.

Given the above motivation, we now formally define the utility class used in the main part of this chapter. We assume that utility takes values in an open interval $I_U \subseteq \mathbb{R}$, which is equal to $\mathbb{R}_{++}$ for cardinal utility. Utility processes will be assumed to be members of a linear subspace $\mathcal{U} \subseteq \mathcal{L}(I_U)$, taken as a primitive. We assume throughout that every $U \in \mathcal{U}$ is an Ito process and satisfies $E\left[\sup_t U_t^2\right] < \infty$. Below we define a *dynamic utility*, meaning that an entire utility process $U(c)$ is assigned to a plan c. Later we will verify that dynamic consistency is satisfied, and is therefore sufficient to maximize time-zero utility.

[10] To see that, make a plot of $\Phi(\omega, t, c, U + \mu dt, \Sigma)$ as a function of U. The concave graph intersects the 45-degree line at U. As μ increases, the graph moves up and the intersection with the 45-degree line moves to the right.

Definition 15 *An (increasing in consumption and concave)* aggregator *is a progressively measurable function of the form* $F : \Omega \times [0,T] \times \mathbb{R}_{++} \times I_U \times \mathbb{R}^d \to (0,\infty)$ *satisfying:*

1. $F(\omega, t, c, U, \Sigma)$ *is strictly increasing in* c *and concave in* (c, U, Σ).
2. $F(\omega, T, c, U, \Sigma)$ *does not depend on* (U, Σ)*, and is therefore denoted* $F(\omega, T, c)$.

The function $U : \mathcal{C} \to I_U$ *is* recursive utility *with aggregator function* F *if, for any* $c \in \mathcal{C}$, $U(c)$ *solves BSDE* (18) *uniquely in* $\mathcal{U}$*. The aggregator* F *is* deterministic *if it does not depend on the state variable. The recursive utility* U *is* state-independent *if the corresponding aggregator* F *is deterministic, and for any deterministic plan* c, $U(c)$ *is the unique deterministic element of* $\mathcal{U}$ *solving the ODE* $dU_t = -F(t, c_t, U_t, 0)\, dt,\ U_T = F(T, c_T)$.

Remark 16 (Aggregator and Beliefs) *Suppose that* U *is recursive utility with aggregator* F*, and the process* $b \in \mathcal{L}(\mathbb{R}^d)$ *is (for simplicity) bounded. Consider the modified aggregator*

$$F^b(\omega, t, c, U, \Sigma) = F(\omega, t, c, U, \Sigma) + b(\omega, t)' \Sigma.$$

Recalling the notation in (11)*, we note that*

$$dU_t = -F^b(t, c_t, U_t, \Sigma_t)\, dt + \Sigma_t' dB_t^b \quad \textit{and} \quad dR_t = \left(\mu_t^R - \sigma_t^{R\prime} b_t\right) dt + \sigma_t^{R\prime} dB_t^b.$$

Since B^b *is Brownian motion under the probability* P^b *(where* $dP^b/dP = \xi_T^b$*), an investor with prior* P^b *still assesses the same risk profile* σ^R*, but believes that the instantaneous expected returns are* $\mu^R - \sigma^{R\prime} b$*. A solution method to the investor's problem for* $b = 0$ *extends to any value of* b *after the formal substitution* $\left(P, B, \mu^R\right) \to \left(P^b, B^b, \mu^R - \sigma^{R\prime} b\right)$.

3.2 Some Basic Properties of Recursive Utility

In this subsection we derive, under regularity assumptions, some basic properties of recursive utility. We first verify dynamic consistency, monotonicity, concavity, and the irrelevance of past or unrealized alternatives. We then discuss comparative risk aversion, and finally we compute the dynamics of a utility supergradient density.

The following notation will be useful. For any function of the form $f : \Omega \times [0,T] \times S \to \mathbb{R}$, where S is a convex subset of some Euclidean space X, we define the superdifferential notation:

$$\partial f(\omega, t, s) = \left\{\delta \in X : f(\omega, t, s + h) \leq f(\omega, t, s) + \delta' h \text{ for all } h \in X \text{ such that } s + h \in S\right\}.$$

Given any processes $d \in \mathcal{L}(X)$ and $x \in \mathcal{L}(S)$, the notation $d \in \partial f(x)$ means that the indicator function of all (ω, t) such that $d(\omega, t) \notin \partial f(\omega, t, x(\omega, t))$ is the zero element of $\mathcal{H}$. Given any $d = (a, b) \in \mathcal{L}_1(\mathbb{R}) \times \mathcal{L}_2(\mathbb{R}^d)$, we let $\mathcal{E}(d)$ or $\mathcal{E}(a, b)$ denote the stochastic exponential with dynamics

$$\frac{d\mathcal{E}_t(a, b)}{\mathcal{E}_t(a, b)} = a_t dt + b_t' dB_t, \quad \mathcal{E}_0(a, b) = 1.$$

The key to deriving properties of recursive utility is the validity of the so-called comparison principle, stated below in terms of the progressively measurable functions $f^i : \Omega \times [0,T] \times I_U \times \mathbb{R}^d \to \mathbb{R}$, $i = 0, 1$.

Condition 17 (Comparison Principle) *For each* $i \in \{0,1\}$, *suppose* $\left(U^i, \Sigma^i\right) \in \mathcal{U} \times \mathcal{L}\left(\mathbb{R}^d\right)$ *solves the BSDE*

$$dU_t^i = -f^i\left(t, U_t^i, \Sigma_t^i\right)dt + \Sigma_t^{i\prime} dB, \quad t \in [0,T], \quad U_T^i \text{ given.}$$

Given stopping times σ, τ *such that* $\sigma \leq \tau$ *a.s., suppose also that*[11]

$$f^0\left(t, U^1, \Sigma^1\right) \leq f^1\left(t, U^1, \Sigma^1\right) \text{ on } [\sigma, \tau] \quad \text{and} \quad U_\tau^0 \leq U_\tau^1 \text{ a.s.}$$

Then $U_\sigma^0 \leq U_\sigma^1$ *a.s. Assuming further that* $P\left[U_\tau^0 < U_\tau^1\right] > 0$, *then* $P\left[U_\sigma^0 < U_\sigma^1\right] > 0$.

A comparison lemma (or stochastic Gronwall-Bellman inequality in the language of Duffie and Epstein) imposes sufficient regularity restrictions for the comparison principle to hold. Various comparison lemmas are given in the BSDE literature referenced earlier. We show below an apparently new version whose applicability relies on our concavity assumption.

Lemma 18 (Comparison Lemma) *The comparison principle (Condition 17) holds if there exists some* $d \in \mathcal{L}_1(\mathbb{R}) \times \mathcal{L}_2\left(\mathbb{R}^d\right)$ *such that* $d \in \partial f^0\left(U^1, \Sigma^1\right)$ *a.e. and* $E\left[\sup_t \mathcal{E}_t(d)^2\right] < \infty$.

Proof. Notationally suppressing the arguments (ω, t), we define the processes $y = U^1 - U^0 \in \mathcal{U}$, $z = \Sigma^1 - \Sigma^0$, and $p = f^1\left(U^1, \Sigma^1\right) - f^0\left(U^1, \Sigma^1\right)$, and we note that

$$dy = -\left(f^0\left(U^1, \Sigma^1\right) - f^0\left(U^0, \Sigma^0\right) + p\right)dt + z'dB.$$

Let $d = (d_U, d_\Sigma)$ be as in the lemma's statement, and define the process $q = f^0\left(U^1, \Sigma^1\right) - f^0\left(U^0, \Sigma^0\right) - (d_U y + d_\Sigma' z)$. Then the above dynamics for y can be restated as

$$dy = -\left(\delta + d_U y + d_\Sigma' z\right)dt + z'dB,$$

where $\delta = p + q$. Our assumptions imply that $\delta \geq 0$ on $[\sigma, \tau]$ and $y_\tau \geq 0$ a.s. Arguing as in the proof of Lemma 6, we can select a sequence of stopping times $\{\tau_n\}$ such that $\tau_n \uparrow \tau$ a.s. and

$$\mathcal{E}_\sigma(d)\, y_\sigma = E_\sigma\left[\int_\sigma^{\tau_n} \mathcal{E}_t(d)\, \delta_t dt + \mathcal{E}_{\tau_n}(d)\, y_{\tau_n}\right] \geq E_\sigma\left[\mathcal{E}_{\tau_n} y_{\tau_n}\right] \quad \text{a.s.}$$

We recall that, by assumption, $y \in \mathcal{U}$ implies $E\left[\sup_t y_t^2\right] < \infty$. Letting $n \to \infty$, and using dominated convergence, it follows that $y_\sigma \geq 0$ a.s. ■

Next we introduce a regularity condition that will allow us to apply the comparison lemma to derive the utility properties we are interested in. The reader who wants to skip technicalities can read "regular" as meaning "we can apply the comparison principle where we have to."

Given any aggregator F and $c \in \mathcal{C}$, we use the notation $F^c(\omega, t, y, z) = F(\omega, t, c(\omega, t), y, z)$. We call an aggregator F *regular* if, given any $(c, U) \in \mathcal{C} \times \mathcal{U}$ with $dU = \ldots\, dt + \Sigma' dB$, there exists $d \in \mathcal{L}_1(\mathbb{R}) \times \mathcal{L}_2\left(\mathbb{R}^d\right)$ such that $d \in \partial F^c(U, \Sigma)$ a.e. and $E\left[\sup_t \mathcal{E}_t(d)^2\right] < \infty$. For example, suppose F is differentiable and $F_U \leq 0$ (which follows from the strict concavity of Φ in equation 16). In this case, regularity of F becomes an integrability restriction on F_Σ, which is satisfied if F_Σ is bounded. Boundedness of F_Σ is usually too strong an assumption, however, and confirming regularity is more challenging.

[11] For processes x, y, we say that $x \leq y$ on $[\sigma, \tau]$ if the indicator function of the set of all (ω, t) such that $x(\omega, t) > y(\omega, t)$ and $\sigma(\omega) \leq t \leq \tau(\omega)$ is zero as an element of $\mathcal{H}$.

Proposition 19 *A recursive utility with a regular aggregator is dynamically consistent, monotonically increasing, concave, and satisfies the irrelevance of past or unrealized alternatives condition.*

Proof. Suppose U is recursive utility with aggregator F, and $c^0, c^1 \in \mathcal{C}$. We use the notation $U^i = U\left(c^i\right)$ and $dU^i = \ldots dt + \Sigma^{i\prime} dB$. To show monotonicity, suppose $c^1 \geq c^0$. The comparison lemma with $f^i = F^{c^i}$ implies $U^1 \geq U^0$. To show concavity, we fix any $\alpha \in (0,1)$ and define the notation $x^\alpha = (1-\alpha)\, x^0 + \alpha x^1$. Notationally suppressing the arguments (ω, t), we define the process $p = F\left(c^\alpha, U^\alpha, \Sigma^\alpha\right) - (1-\alpha)\, F\left(c^0, U^0, \Sigma^0\right) - \alpha F\left(c^1, U^1, \Sigma^1\right)$, and note that

$$dU^\alpha = -\left(F\left(c^\alpha, U^\alpha, \Sigma^\alpha\right) - p\right) dt + \Sigma^\alpha dB, \quad U_T^\alpha = F\left(T, c_T^\alpha\right) - p_T.$$

The concavity assumption on F implies that $p \geq 0$. Apply the comparison lemma with $f^1 = F^{c^\alpha}$ and $f^0 = F^{c^\alpha} - p$, we conclude that $U\left(c^\alpha\right) \geq U^\alpha$, confirming concavity. The remaining claims are left as an exercises in the application of the comparison lemma. ■

A state-independent recursive utility with aggregator F ranks deterministic plans in a way determined by the function $(t, c, U) \mapsto F\left(t, c, U, 0\right)$, while the dependence of F on Σ can be used to adjust risk aversion without affecting the utility of deterministic plans. The formal statement of this property is based on the following partial order of utility functions:

Comparative Risk Aversion: A utility function $U_0^1 : \mathcal{C} \to \mathbb{R}$ is *more risk averse* than a utility function $U_0^2 : \mathcal{C} \to \mathbb{R}$ if

- For any deterministic plans $a, b \in \mathcal{C}$, $U_0^1(a) \geq U_0^1(b) \iff U_0^2(a) \geq U_0^2(b)$.
- For any $c \in \mathcal{C}$ and deterministic $\bar{c} \in \mathcal{C}$, $U_0^2(\bar{c}) \geq U_0^2(c) \implies U_0^1(\bar{c}) \geq U_0^1(c)$.

Remark 20 *If U_0^1 and U_0^2 are cardinal utilities, then U_0^1 is more risk averse than U_0^2 if and only if $U_0^1(c) = U_0^2(c)$ for every deterministic plan c, and $U_0^1(c) \leq U_0^2(c)$ for every plan c.*

Proposition 21 *Suppose that, for $i \in \{1, 2\}$, U^i is state-independent recursive utility with aggregator F^i, and F^1 is regular. If $F^1(t, c, U, 0) = F^2(t, c, U, 0)$ and $F^1(t, c, U, \Sigma) \leq F^2(t, c, U, \Sigma)$ for all (t, c, U, Σ), then U_0^1 is more risk-averse than U_0^2.*

Proof. By definition, $F^1(t, c, U, 0) = F^2(t, c, U, 0)$ implies that $U^1(c) = U^2(c)$ for every deterministic plan c. The proof is completed using the comparison lemma. ■

Finally, we derive a utility supergradient density expression for recursive utility, which will be key in establishing optimality conditions.

Proposition 22 *Suppose U is recursive utility with aggregator F such that $F_c, F_U \in \mathcal{L}_1(\mathbb{R})$ and $F_\Sigma \in \mathcal{L}_2\left(\mathbb{R}^d\right)$ satisfy*

$$\left(F_c, F_U, F_\Sigma\right) \in \partial F\left(c, U, \Sigma\right) \tag{19}$$

and $E\left[\sup_t \mathcal{E}_t\left(F_U, F_\Sigma\right)^2\right] < \infty$. Let the process π be defined by

$$\pi = \mathcal{E}\left(F_U, F_\Sigma\right) F_c.$$

Provided it belongs to $\mathcal{H}$, the process π is a supergradient density of U_0 at c.

Proof. Assuming $c+x\in\mathcal{C}$, we define $\delta=U(c+x)-U(c)$, $\Delta=\Sigma(c+x)-\Sigma(c)$, and $p=F(c,U,\Sigma)+F_c x+F_U\delta+F_\Sigma'\Delta-F(c+x,U+\delta,\Sigma+\Delta)\geq 0$, where the last inequality follows from the assumed condition (19). The BSDEs for $U(c+h)$ and $U(c)$ imply the linear BSDE

$$d\delta=-\left(F_c x+F_U\delta+F_\Sigma'\Delta-p\right)dt+\Delta' dB,\quad \delta_T=F_c(T)\,x_T-p_T.$$

The comparison lemma implies $\delta_0\leq(\pi|x)$, where $\pi=\mathcal{E}(F_U,F_\Sigma)\,F_c$. ■

3.3 Optimality under Recursive Utility

Proposition 3 verifies the optimality of a feasible consumption plan c based on the existence of a process that is both a utility supergradient density at c and a state price density at c. Specializing this argument to recursive utility, in this subsection, we apply Ito's lemma to the supergradient density expression of Proposition 22, and we use the state price dynamics of Proposition 8 to derive sufficient optimality conditions for recursive utility as a FBSDE system.

We fix a reference recursive utility $U:\mathcal{C}\to I_U$ with aggregator F, relative to which optimality is defined. By definition, $F(\omega,t,\cdot)$ is concave but not necessarily differentiable. In the following section, we will see that nonsmoothness of $F(\omega,t,c,U,\Sigma)$ in (U,Σ) is useful in modeling first-order risk aversion. On the other hand, we will have no use for nonsmoothness of F in the consumption argument, and we therefore assume the existence of the corresponding partial derivative F_c. In addition, we finesse the issue of a consumption nonnegativity constraint by the usual trick of making marginal utility go to infinity near zero. Finally, we assume that marginal utility converges to zero as consumption goes to infinity. These assumptions and some associated notation are summarized below, and are adopted for the remainder of this chapter's main part.

Regularity Assumptions and Notation: *The partial derivative F_c exists everywhere and the (strictly decreasing) function $F_c(\omega,t,\cdot,U,\Sigma)$ maps $(0,\infty)$ onto $(0,\infty)$, for any (ω,t,U,Σ). The function $\mathcal{I}:\Omega\times[0,T]\times(0,\infty)\times I_U\times\mathbb{R}^d\to(0,\infty)$ is therefore well-defined implicitly by*

$$F_c(\omega,t,\mathcal{I}(\omega,t,\lambda,U,\Sigma),U,\Sigma)=\lambda,\quad \lambda\in(0,\infty).$$

The superdifferential of F with respect to (U,Σ) is defined by

$$\partial_{U,\Sigma}F(\omega,t,c,U,\Sigma)=\left\{(a,b)\in\mathbb{R}\times\mathbb{R}^d:(F_c(\omega,t,c,U,\Sigma),a,b)\in\partial F(\omega,t,c,U,\Sigma)\right\}.$$

We fix a reference strategy (ρ,ψ), generating the wealth process W, and financing the consumption plan $c=\rho W$. To formulate sufficient conditions for the optimality of c, we define the strictly positive process

$$\lambda_t=F_c(t,c_t,U_t,\Sigma_t),\tag{20}$$

which is equivalent to

$$c_t=\mathcal{I}(t,\lambda_t,U_t,\Sigma_t).\tag{21}$$

By the usual envelope-type argument of microeconomics, if c is optimal then λ_t represents the shadow price of the time-t wealth constraint. Intuitively, given an optimal plan, the additional utility obtained by infinitesimally increasing wealth at some spot is the same whether the agent reoptimizes at the new wealth level or whether the agent just consumes all additional wealth on the spot (over an interval dt). Although we will not need to formalize this intuition, it will be helpful to keep in mind the interpretation of λ as a shadow-price-of-wealth process. The dynamics of λ are denoted

$$\frac{d\lambda_t}{\lambda_t} = \mu_t^{\lambda} dt + \sigma_t^{\lambda\prime} dB_t. \tag{22}$$

We know that (under regularity assumptions) $\pi = \mathcal{E}\lambda$ is a supergradient density at c, where $\mathcal{E} = \mathcal{E}(F_U, F_\Sigma)$ is computed as in Proposition 22. Integration by parts gives

$$\frac{d\pi}{\pi} = \left(F_U + \mu^{\lambda} + \sigma^{\lambda\prime} F_\Sigma\right) dt + \left(F_\Sigma + \sigma^{\lambda}\right)' dB.$$

To ensure that π is also a state-price density at c, we match terms with the dynamics of Proposition 8, resulting in the restrictions:

$$r = -\left(F_U + \mu^{\lambda} + \sigma^{\lambda\prime} F_\Sigma\right), \quad \eta = -\left(F_\Sigma + \sigma^{\lambda}\right), \quad \mu^R = \sigma^{R\prime}\eta.$$

We round up the optimality conditions by combining the above restrictions with the utility and wealth dynamics, as well equations (21) and (22) :

Condition 23 (Optimality Conditions for Recursive Utility) *The trading strategy ψ and the processes $\left(U, \Sigma, \lambda, \sigma^{\lambda}, W\right) \in \mathcal{U} \times \mathcal{L}_2\left(\mathbb{R}^d\right) \times \mathcal{L}\left(\mathbb{R}_{++}\right) \times \mathcal{L}_2\left(\mathbb{R}^d\right) \times \mathcal{L}\left(\mathbb{R}_{++}\right)$ solve*

$$\begin{aligned} dU &= -F\left(\mathcal{I}\left(\lambda, U, \Sigma\right), U, \Sigma\right) dt + \Sigma' dB, \quad U_T = F\left(T, W_T\right), \\ \frac{d\lambda}{\lambda} &= -\left(r + F_U + \sigma^{\lambda\prime} F_\Sigma\right) dt + \sigma^{\lambda\prime} dB, \quad \lambda_T = F_c\left(T, W_T\right), \\ dW &= \left(W\left(r + \psi'\mu^R\right) - \mathcal{I}\left(\lambda, U, \Sigma\right)\right) dt + W\psi'\sigma^{R\prime} dB, \quad W_0 = w_0, \\ \mu^R + \sigma^{R\prime}&\left(F_\Sigma + \sigma^{\lambda}\right) = 0, \quad \left(F_U, F_\Sigma\right) \in \left(\partial_{U,\Sigma} F\right)\left(\mathcal{I}\left(t, \lambda, U, \Sigma\right), U, \Sigma\right). \end{aligned}$$

Proposition 24 *Suppose Condition 23 holds, and let $c_t = \mathcal{I}\left(t, \lambda_t, U_t, \Sigma_t\right)$ and $\rho_t = c_t / W_t$. If $c \in \mathcal{C}$, $\pi = \mathcal{E}\left(F_U, F_\Sigma\right)\lambda \in \mathcal{H}$, and $E\left[\sup_t \pi_t W_t\right] < \infty$, then the strategy (ρ, ψ) is optimal, it generates the wealth process W, and it finances the consumption plan c, whose utility process is U.*

Proof. The dynamics of W can be used to verify that (ρ, ψ) finances c with wealth process W. By Proposition 22, π is a utility supergradient density at c. By Proposition 8, π is also a state price density at c. By Proposition 3, c is optimal. The dynamics of U show that $U = U(c)$. ∎

Remark 25 *In SS03 the above conditions are extended to include convex trading constraints, and a necessity argument is given for a smooth aggregator under some regularity assumptions. The case of no intermediate or no terminal consumption is essentially the same as above, omitting the appropriate consumption arguments in the formulation.*

Condition 23 is a FBSDE system. The wealth dynamics are computed recursively forward in time, starting with $W_0 = w_0$, and the dynamics of (U, λ) are computed recursively backward on the information tree, starting with their terminal values. The forward and backward components are coupled. In the following subsection we will introduce scale-invariance as a way of uncoupling this FBSDE system. In a Markovian setting, a PDE version of the FBSDE system can be obtained as in Ma, Protter, and Yong (1994). The construction is outlined in SS03, as well as later in this chapter for a more special class of homothetic recursive utilities.

3.4 Homothetic Recursive Utility

The utility function $U_0 : \mathcal{C} \to \mathbb{R}$ is *homothetic* (or *scale-invariant*) if for any $c^1, c^2 \in \mathcal{C}$,

$$U_0\left(c^1\right) = U_0\left(c^2\right) \quad \text{implies} \quad U_0\left(kc^1\right) = U_0\left(kc^2\right) \text{ for all } k \in (0, \infty).$$

If U_0 is homothetic and cardinal, then[12] it is homogeneous (of degree one). For recursive utility, with $I_U = (0, \infty)$, homogeneity of U_0 is implied by (and is essentially equivalent to) homogeneity of the aggregator with respect to the utility argument; that is, an aggregator of the form

$$F(\omega, t, c, U, \Sigma) = UG\left(\omega, t, \frac{c}{U}, \frac{\Sigma}{U}\right), \quad F(\omega, T, c) = c, \tag{23}$$

for some function $G : \Omega \times [0, T] \times (0, \infty) \times \mathbb{R}^d \to \mathbb{R}$ that we call a *proportional aggregator.*

Assuming the above aggregator form, suppose that (ρ, ψ) is an optimal strategy with corresponding wealth process W and utility process U. Recalling the interpretation of the process λ as the sensitivity of the optimal time-t utility value on time-t wealth, the homogeneity of the utility function implies that

$$U = \lambda W. \tag{24}$$

The intuition behind this relationship is straightforward. Suppose that at some spot ω^t the investor with unit wealth finds a consumption plan $\bar{c}$ optimal, resulting in the spot-ω^t optimal utility value $\lambda[\omega^t]$. If the same investor's wealth at ω^t were instead $W[\omega^t]$ units of account, then, by the homogeneity of preferences and the budget equation, the investor would find the consumption plan $W[\omega^t]\,\bar{c}$ optimal at ω^t, resulting in an optimal utility value $U[\omega^t] = \lambda[\omega^t]\,W[\omega^t]$. In other words, the optimization problem at every spot is a scaled version of the unit-wealth version of the problem.

Equation (24) allows us to reduce the optimality conditions to a single BSDE for λ. The general form of this BSDE can be found in SS03. Rather than dealing with the general case here, we will instead consider, in Section 5, optimality under some more special proportional aggregator functional forms that are motivated by the models of risk aversion of the following section.

We close this section with an example of a proportional aggregator specification, under which the optimal consumption strategy is a given process, for any investment opportunity set.

Example 26 (A Robustly Optimal Consumption Strategy) *Let the aggregator F be given by equation* (23) *for a proportional aggregator of the form*

$$G(t, c, \sigma) = \beta(t) \log(c) + H(t, \sigma), \quad t < T, \tag{25}$$

where β is any strictly positive and (for simplicity) bounded process. While the optimal trading strategy depends on the specification of H and the investment opportunity set, we argue below that the optimal consumption strategy is independent of both, and is simply equal to β. Ignoring technical details, we assume the sufficiency and necessity of the optimality conditions and the existence of an optimum. To see the essential part of the argument, suppose (ρ, ψ) is an optimal strategy, with corresponding wealth process W, consumption plan $c = \rho W$, utility process $U = ...dt + \Sigma' dB$, and shadow-price-of-wealth process λ. Using equation (24)*, we observe that*

$$\lambda_t = F_c(t, c, U, \Sigma) = \beta_t \frac{U_t}{c_t} = \beta_t \frac{W_t}{c_t} \frac{U_t}{W_t} = \frac{\beta_t}{\rho_t} \lambda_t.$$

Therefore, $\rho = \beta$, independently of the investment opportunity set. A more complete and general version of this argument is given in SS03 (where trading constraints are allowed).

[12] Proof: For any $c \in C$ and $k \in (0, \infty)$, $U_0(c) = U_0(U_0(c)\,\mathbf{1})$ implies $U_0(kc) = U_0(kU_0(c)\,\mathbf{1}) = kU_0(c)$.

4 Modeling Risk Aversion

This section formulates some concrete representations of possibly source-dependent second or first order risk aversion in the context of recursive utility. These representations will be used in the following section to derive optimal trading strategy formulas that help clarify the relationship between risk aversion and portfolio allocations.

4.1 Conditional Certainty Equivalents

Let us recall the essential intuition of a recursive utility formulation, captured by Lemma 13. We fix a consumption plan c with cardinal utility process $U = U(c)$. For any time-t spot ω^t, the corresponding utility value $U_t[\omega^t]$ can be computed as a function of ω^t, the immediate consumption $c[\omega^t]\,dt$, and the restriction of the random variable U_{t+dt} to the spot ω^t, which we denote $U_{t+dt}[\omega^t]$.

In this section, we assume that the functional dependence of U_t on U_{t+dt} enters through the conditional certainty-equivalent $\nu_t(U_{t+dt})$, an $\mathcal{F}_t$-measurable random variable such that $\nu_t(U_{t+dt})[\omega^t]$ depends on U_{t+dt} only through its restriction $U_{t+dt}[\omega^t]$, and is the identity if $U_{t+dt}[\omega^t]$ is constant. The value of $\nu_t(U_{t+dt})$ is a conditional certainty equivalent in the sense that, conditionally on the spot ω^t and immediate consumption $c[\omega^t]\,dt$, the investor is indifferent between the continuation of the plan c and a constant consumption rate of $\nu_t(U_{t+dt})[\omega^t]$ over the entire remaining period $[t+dt, T]$. Under this assumption, we can write the heuristic recursion for the utility process $U = U(c)$ as

$$U_t = \phi\left(t, dt, c_t, \nu_t\left(U_{t+dt}\right)\right), \tag{26}$$

where ϕ can be spot-dependent. The dependence of ϕ on the recursion interval dt is important in the approximation argument that follows. We further assume that ϕ has continuous partial derivatives ϕ_{dt}, ϕ_c, and ϕ_U. Since preferences are increasing, ϕ_c and ϕ_U are strictly positive.

In the following three subsections we derive the functional form of the aggregator F for various specifications of the certainty equivalent ν. In each formulation, the conditional certainty equivalent has a local representation in terms of the instantaneous linear factor structure (15) of U_{t+dt} that takes the form

$$\nu_t\left(U_{t+dt}\right) = U_t + \left(\mu_t - \mathcal{A}\left(t, U_t, \Sigma_t\right)\right) dt, \tag{27}$$

for some function $\mathcal{A}$ satisfying $\mathcal{A}(t, U, 0) = 0$. The function $\mathcal{A}$ represents the risk aversion implicit in ν. Recalling equations (16) and (26), and using a first-order Taylor expansion of ϕ, we obtain

$$\begin{aligned} U_t &= \Phi\left(t, c_t, U_t + \mu_t dt, \Sigma_t\right) = \phi\left(t, dt, c_t, U_t + \left(\mu_t - \mathcal{A}\left(t, U_t, \Sigma_t\right)\right) dt\right) \\ &= U_t + \left[\phi_{dt}\left(t, 0, c_t, U_t\right) + \phi_U\left(t, 0, c_t, U_t\right)\left(\mu_t - \mathcal{A}\left(t, U_t, \Sigma_t\right)\right)\right] dt. \end{aligned}$$

Using the definition of F in terms of Φ in (17) and the last equation, we obtain the aggregator functional form

$$F\left(\omega, t, c, U, \Sigma\right) = f\left(\omega, t, c, U\right) - \mathcal{A}\left(\omega, t, U, \Sigma\right), \tag{28}$$

where $f(\omega, t, c, U) = \phi_{dt}(\omega, t, 0, c, U) / \phi_U(\omega, t, 0, c, U)$.

Suppose now that ϕ and $\mathcal{A}$ are state independent, and therefore f and F are also state independent. If the plan c is deterministic, then $\Sigma = 0$ and (since $\mathcal{A}(t, U, 0) = 0$) utility can be computed in terms of the aggregator section $F(t, c, U, 0) = f(t, c, U)$. The function f (or ϕ) therefore determines the investor's preferences over deterministic choices. By Proposition 21, given f, the larger $\mathcal{A}$ is the more risk averse the investor. This hierarchical separation of preferences toward deterministic choices and risk aversion can also be seen directly in the recursive form (26). If c is deterministic, then so is U, and therefore $\nu_t(U_{t+dt}) = U_{t+dt}$. This shows that utility over deterministic plans is determined by ϕ. Given ϕ, increasing $\mathcal{A}$ decreases the conditional certainty equivalent value and therefore U_t, resulting in more risk averse utility.

The key behavioral restriction introduced by the assumption of the recursive form (26) is that, given the agent's preferences over deterministic choices, the agent's risk aversion at a spot ω^t, represented by $\mathcal{A}(U,\Sigma)[\omega^t]$, does not depend on the amount $c[\omega^t]\,dt$ consumed at time t. This separation of current consumption and risk aversion is reflected in the separable representation (28).

Homothetic utility with an aggregator of the form (28) is obtained by further imposing the functional restriction (23). In this case, the proportional aggregator G takes the functional form

$$G(\omega,t,c,\sigma) = g(\omega,t,c) - \mathcal{R}(\omega,t,\sigma), \tag{29}$$

where $g(\omega,t,c) = f(\omega,t,c,1)$ and $\mathcal{R}(\omega,t,\sigma) = \mathcal{A}(\omega,t,1,\sigma)$. The argument $\sigma = \Sigma/U$ represents utility risk measured relative to the utility level U.

4.2 The Duffie-Epstein Limit of Kreps-Porteus Utility

The first specialization of the aggregator form (28) we consider results from the continuous-time formulation of Kreps and Porteus (1978) utility due to[13] Duffie and Epstein (1992). In this formulation, the conditional certainty equivalent ν_t is defined by

$$u(\nu_t(U_{t+dt})) = E_t[u(U_{t+dt})], \tag{30}$$

for some concave, twice continuously differentiable utility function $u : I_U \to \mathbb{R}$. We denote the corresponding coefficient of absolute risk aversion by

$$a(U) = -\frac{u''(U)}{u'(U)}.$$

In the current context, the classic Arrow (1965, 1970) and Pratt (1964) approximation of expected utility for small risks can be expressed through Ito's lemma as

$$u(U_{t+dt}) = u(U_t) + u'(U_t)\,dU_t + \frac{1}{2}u''(U_t)(dU_t)^2.$$

Given the instantaneous linear factor structure (15) of U_{t+dt}, the above results in

$$E_t[u(U_{t+dt})] = u(U_t) + \left(u'(U_t)\,\mu_t + \frac{1}{2}u''(U_t)\,\Sigma_t'\Sigma_t\right)dt.$$

Using the first-order Taylor expansion $u(\nu_t(U_{t+dt})) = u(U_t) + u'(U_t)(\nu_t(U_{t+dt}) - U_t)$ and simplifying results in the certainty-equivalent expression (27) with the quadratic risk-aversion component

$$\mathcal{A}(\omega,t,U,\Sigma) = \frac{1}{2}a(U)\,\Sigma'\Sigma.$$

[13] In fact, Duffie-Epstein utilities are obtained as the continuous-time limit of a broader class of discrete-time utilities than the Kreps-Porteus class, since the investor's certainty equivalent over continuation utility need only be von-Neumann-Morgenstern in an approximate local sense. It is sufficient for our purposes, however, to think of Duffie-Epstein utility as (sufficiently smooth) continuous-time Kreps-Porteus utility.

The corresponding aggregator (28) takes *Duffie-Epstein* form:

$$F(\omega,t,c,U,\Sigma) = f(\omega,t,c,U) - \frac{1}{2}a(U)\Sigma'\Sigma. \tag{31}$$

We refer to Duffie and Epstein (1992) for further analysis of this utility form. For example, they show that there is always an ordinally equivalent utility version with the same recursive representation but $a = 0$. The latter restriction can be analytically helpful, but minimizes the usefulness of the hierarchical separation of choice over deterministic plans and risk aversion of Proposition 21. If $a = 0$ and $F = f$ is linear in U, as in Example 14, then one obtains time-additive expected discounted utility. As discussed in the Introduction (and shown in Skiadas, 2003), the Duffie-Epstein specification also includes the type of "robust" criteria used by Anderson, Hansen, and Sargent (2000), Hansen, Sargent, Turmuhambetova, and Williams (2001), and Maenhout (1999).

Homothetic Duffie-Epstein utility is obtained if the aggregator takes the homogenous form (23), for a proportional aggregator of the form

$$G(\omega,t,c,\sigma) = g(\omega,t,c) - \frac{1}{2}\gamma\sigma'\sigma, \quad \text{for some } \gamma \in \mathbb{R}_+,$$

a special case of (29). Assuming g is state-independent, utility over deterministic plans is entirely determined by g. Given g, the higher γ is, the higher the investor's relative risk aversion.

Example 27 *The continuous-time version of Epstein-Zin utility is obtained by letting*

$$g(\omega,t,c) = \beta\frac{c^{1-\delta}-1}{1-\delta},$$

for some $\beta \in \mathbb{R}_{++}$ and $\delta \in \mathbb{R}_+$. For $\delta = 1$, the function $\left(x^{1-\delta}-1\right)/(1-\delta)$ is interpreted as $\log(x)$ *(the limit as δ approaches one). The parameters (β,δ) determine choice over deterministic plans, while, for any given (β,δ), increasing the coefficient γ makes the investor more risk averse. A time-additive ordinally equivalent utility is obtained if and only if $\gamma = \delta$, in which case*

$$\frac{1}{\beta}\frac{U_t(c)^{1-\gamma}-1}{1-\gamma} = E_t\left[\int_t^T e^{-\beta(s-t)}\frac{c_s^{1-\gamma}-1}{1-\gamma}ds\right]. \tag{32}$$

4.3 Source-Dependent Risk Aversion

We noted in the Introduction that it is of interest to consider risk aversion that can depend on the source of risk, for example, as an expression of aversion to ambiguity associated with a given source of risk. With a version of Proposition 21, Lazrak and Quenez (2003) made the important observation that the functional dependence of a general aggregator $F(t,c,U,\Sigma)$ on Σ allows the modeling of risk-aversion that varies with the direction of risk. Since Σ represents loadings to instantaneous linear factors, such directional risk aversion can be interpreted as source-dependent risk aversion. SS03 and Schroder and Skiadas (2005a) derived optimal trading strategies for special functional aggregator forms that allow source-dependent risk aversion. These functional forms are motivated below and the following subsection.

We begin with a simple extension of Duffie-Epstein utility that allows for source-dependent risk aversion, where each Brownian motion is viewed as a separate source of risk. In the Duffie-Epstein formulation, the certainty equivalent (30) is applied to the aggregate risky continuation utility $U_{t+dt} = U_t + \mu_t dt + \Sigma_t' dB_t$. Here we assume that the investor perceives and worries about the individual risk terms $\Sigma_t^1 dB_t^1, \ldots, \Sigma_t^d dB_t^d$ separately, since they represent exposure to different sources of risk. We model this by postulating a twice continuously differentiable function $u : \mathbb{R}^{1+d} \to \mathbb{R}$ such that the time-t conditional certainty equivalent in the recursive specification (26) is defined by

$$u\left(\nu_t\left(U_{t+dt}\right), 0, \ldots, 0\right) = E_t\left[u\left(U_t + \mu_t dt, \Sigma^1 dB^1, \ldots, \Sigma^d dB^d\right)\right]. \tag{33}$$

The first and second order partial derivatives of $u\left(x_0, x_1, \ldots, x_d\right)$ with respect to x_i are denoted u_i and u_{ii}, respectively. The absolute risk aversion coefficient with respect to the i^{th} risk source is defined by

$$a^i\left(U\right) = -\frac{u_{ii}\left(U, 0, \ldots, 0\right)}{u_0\left(U, 0, \ldots, 0\right)}. \tag{34}$$

We also define the diagonal matrix $A\left(U\right) = \operatorname{diag}\left[a^1\left(U\right), \ldots, a^d\left(U\right)\right]$. Applying Ito's lemma and taking conditional expectations results in

$$E_t\left[u\left(U_t + \mu_t dt, \Sigma^1 dB^1, \ldots, \Sigma^d dB^d\right)\right] = u\left(U_t, 0, \ldots, 0\right) + u_0\left(U_t, 0, \ldots, 0\right)\left(\mu_t - \frac{1}{2}\Sigma_t' A\left(U_t\right)\Sigma_t\right) dt.$$

Similarly, we have the first-order Taylor expansion

$$u\left(\nu_t\left(U_{t+dt}\right), 0, \ldots, 0\right) = u\left(U_t, 0, \ldots, 0\right) + u_0\left(U_t, 0, \ldots, 0\right)\left(\nu_t\left(U_{t+dt}\right) - U_t\right).$$

Matching the lasts two expressions and simplifying results in the certainty-equivalent expression (27), and corresponding aggregator (28), with the quadratic risk-aversion component

$$\mathcal{A}\left(\omega, t, U, \Sigma\right) = \frac{1}{2}\Sigma' A\left(U\right)\Sigma.$$

The Duffie-Epstein case is obtained if $a^i = a$ for all i.

Remark 28 *A simple extension is obtained if the Brownian motion in the above formulation is replaced by a new Brownian motion $\bar{B}$, where $d\bar{B}$ is obtained by a possibly spot-dependent rotation of dB. More formally, we assume there exists some process $\Phi \in \mathcal{L}\left(\mathbb{R}^{d\times d}\right)$ such that, for all t, Φ_t is orthogonal ($\Phi_t'\Phi_t = I$) and $d\bar{B}_t = \Phi_t dB_t$. In this case, $U_{t+dt} = U_t + \mu_t dt + \bar{\Sigma}_t' d\bar{B}_t$, where $\bar{\Sigma}_t = \Phi_t \Sigma_t$. The above argument results in the aggregator form* (28) *with*

$$\mathcal{A}\left(t, U_t, \Sigma_t\right) = \frac{1}{2}\bar{\Sigma}_t' A\left(U_t\right)\bar{\Sigma}_t = \frac{1}{2}\Sigma_t'\Phi_t' A\left(U_t\right)\Phi_t\Sigma_t.$$

In the Duffie-Epstein case, this aggregator is identical to the one derived above. With source-dependent risk aversion, however, the aggregator form changes with a Brownian motion rotation.

Combining the above representation with the homothetic specification (23) results in a proportional aggregator of the form (29), where $\mathcal{R}$ is a quadratic form. This specification includes the Uppal and Wang (2003) "robust" criterion, as can be shown by a simple extension of the argument in Skiadas (2003).

4.4 First-Order Risk Aversion

Consider an investor who maximizes expected von Neuman-Morgenstern (vNM) utility in a single period setting. If one were to zoom in a very small area of the graph of the vNM utility, one would see a straight line. This means that an investor is essentially risk-neutral toward the addition of sufficiently small risks to a given wealth level. As an implication, such an investor would seek some exposure to all investment opportunities of positive expected excess return, and would not completely insure a source of risk in actuarially unfavorable terms. These conclusions extend to the recursive utility formulations of the last two sections, as will become clear in the following section. In reality, we observe that investors routinely do not participate in investment opportunities with positive Sharpe ratios, and they do pay actuarially unfavorable premia to completely insure some sources of risk (for example, against loss of individual items of negligible value relative to total wealth). While such behavior can relate to a number of issues, we focus here on a certainty-equivalent specification exhibiting first-order risk aversion in the sense of Segal and Spivak (1990).

In a static expected-utility setting, first-order risk aversion amounts to introducing a kink of the vNM utility around the given wealth level, hence removing local risk neutrality. Since a risk-averse vNM utility can have at most countably many kinks, the approach seems problematic. If one keeps track of different sources of risk, however, as in the source-dependent certainty equivalent introduced above, this problem does not arise. As in the last section, we assume the recursive utility specification (26) with the source-dependent certainty equivalent specification in (33), except that the function u in (33) is now replaced with the function

$$\hat{u}\left(x_0, x_1, \ldots, x_d\right) = u\left(x_0, x_1, \ldots, x_d\right) - \sum_{i=1}^{d} \delta^i\left(x_0\right)\left|x_i\right|.$$

We assume that each δ^i is differentiable and nonnegative valued, and that u is exactly as in the last section. Noting that $\hat{u}(U, 0, \ldots, 0) = u(U, 0, \ldots 0)$, it follows that the conditional certainty equivalent ν_t is defined by

$$u\left(\nu_t\left(U_{t+dt}\right), 0, \ldots, 0\right) = E_t\left[u\left(U_t + \mu_t dt, \Sigma^1 dB^1, \ldots, \Sigma^d dB^d\right)\right] - \sum_{i=1}^{d} E_t\left[\delta^i\left(U_t + \mu_t dt\right)\left|\Sigma_t^i dB_t^i\right|\right].$$

The left-hand side and the first term of the right-hand side in the above equation are computed exactly as in the last section. To compute the last term, we first note that $E_t\left|dB_t^i\right| = \sqrt{2/\pi}\, dt$ (a consequence of the fact that dB_t^i is normally distributed with zero mean and variance dt). By the usual Ito calculus, $\delta^i\left(U_t + \mu_t dt\right) = \delta^i\left(U_t\right) + \delta^{i\prime}\left(U_t\right)\mu_t dt$ and $dt^2 = 0$, and therefore

$$E_t\left[\delta^i\left(U_t + \mu_t dt\right)\left|\Sigma_t^i dB_t^i\right|\right] = u_0\left(U, 0, \ldots, 0\right)\kappa^i\left(U_t\right)\left|\Sigma_t^i\right| dt,$$

where

$$\kappa^i(U) = \frac{\delta^i(U)\sqrt{2/\pi}}{u_0(U, 0, \ldots, 0)}.$$

Combining the above calculations in the certainty equivalent definition and simplifying, we derive the certainty-equivalent expression (27), and corresponding aggregator (28), with

$$\mathcal{A}(\omega, t, U, \Sigma) = \kappa(U)'\left|\Sigma\right| + \frac{1}{2}\Sigma' A(U)\Sigma.$$

where $\kappa(U) = \left(\kappa^1(U), \ldots, \kappa^d(U)\right)'$ and $|\Sigma| = \left(\left|\Sigma^1\right|, \ldots, \left|\Sigma^d\right|\right)'$. Combining the above representation with the homothetic specification (23) results in a proportional aggregator of the form (29), where $\mathcal{R}$ takes a functional form obtained by setting $U = 1$ in the expression for $\mathcal{A}$ above. We revisit the homothetic case in the following section, where the effect of first-order risk aversion on portfolio choice is discussed.

We have noted that source-dependent (second-order) risk aversion can reflect uncertainty about the risk-premium model associated with a given source of risk, a claim formalized through an equivalence with a robust-control type criterion. The same interpretation applies to source-dependent first-order risk aversion. For example, if each κ^i is a constant above, the resulting formulation is equivalent to the "κ-ignorance" multiple-prior example of Chen and Epstein (2002). A related convex duality argument can be found in El Karoui, Peng, and Quenez (2001).

5 Scale-Invariant Solutions

In this section we study optimal strategies under the homothetic case of last section's utilities, thus taking advantage of the simplifications of scale invariance introduced in subsection 3.4, as well as specific risk aversion parameterizations.

We assume throughout that $I_U = (0, \infty)$ and that the investor has recursive utility with an aggregator of the homogenous form (23). The corresponding proportional aggregator is further restricted to be of the form $G(\omega, t, c, \sigma) = g(t, c) - \mathcal{R}(\omega, t, \sigma)$, where g is state-independent[14] and $\mathcal{R}(\omega, t, 0) = 0$. The function g entirely determines choice over deterministic plans, while increasing $\mathcal{R}$ increases the investor's risk aversion, without changing the investor's choice over deterministic plans. The function g must also reflect our earlier restrictions on the aggregator F and its partial derivative F_c. We summarize our assumptions in the following condition, which is imposed throughout this section:

Condition 29 *The function $g : [0, T] \times (0, \infty) \to \mathbb{R}$ is such that, for every time t, $g(t, \cdot)$ is differentiable and concave, and its derivative, $g_c(t, \cdot)$, exists and maps $(0, \infty)$ onto $(0, \infty)$. For any $c \in \mathcal{C}$, the utility process $U = U(c)$ is strictly positive and solves, uniquely in $\mathcal{U}$, the BSDE*

$$\frac{dU_t}{U_t} = -\left(g\left(t, \frac{c_t}{U_t}\right) - \mathcal{R}(t, \sigma_t)\right) dt + \sigma_t' dB_t, \quad U_T = c_T. \tag{35}$$

Finally, g is sufficiently regular so that, for any deterministic plan c, the ordinary differential equation $dU/U = -g(t, c/U)\, dt$, $U_T = 1$, has a unique deterministic solution U in $\mathcal{U}$.

To state the simplified optimality conditions under the above specification we define the functions $\mathcal{I}^g, g^* : [0, T] \times (0, \infty) \to (0, \infty)$ by

$$g_c(t, \mathcal{I}^g(t, \lambda)) = \lambda \quad \text{and} \quad g^*(t, \lambda) = \max_{c \in \mathbb{R}_{++}} (g(t, c) - \lambda c) = g(t, \mathcal{I}^g(t, \lambda)) - \mathcal{I}^g(t, \lambda)\, \lambda.$$

As discussed in subsection 3.4, since utility is homogeneous of degree-one, at the optimum, the utility process U, the wealth process W, and the shadow-price-of-wealth process λ are related by $U = \lambda W$. If c is the corresponding optimal plan, the condition $\lambda = F_c(t, c, U, \Sigma)$ and a simple calculation give the optimal consumption strategy as $\rho_t = \lambda_t \mathcal{I}^g(t, \lambda_t)$. The BSDE for λ and optimal trading strategy depend on the specification of the relative risk-aversion function $\mathcal{R}$.

[14] Allowing g to be state-dependent entails essentially notational changes in the optimality conditions, but we forgo this generality for simplicity.

5.1 Smooth Quasi-Quadratic Proportional Aggregator

The first specification we consider includes the homothetic version of last section's models of risk aversion with a smooth aggregator. The case of first-order risk aversion will be treated at the end of this section. Up to that point, we assume:

Condition 30 (Smooth Quasi-Quadratic Proportional Aggregator) *Condition 29 holds with*

$$\mathcal{R}(\omega, t, c, \sigma) = \frac{1}{2}\sigma' Q(\omega, t)\,\sigma, \tag{36}$$

for some bounded $Q : \Omega \times [0, T] \to \mathbb{R}^{d\times d}$*, where* $Q(\omega, t)$ *is symmetric positive definite for all* (ω, t)*.*

In terms of the risk aversion function $A(U)$ of subsection 4.3, $Q = A(1)$, and therefore Q can be thought of as a relative risk aversion matrix. In the Duffie-Epstein case, $Q = \gamma I$, where the γ is a coefficient of relative risk aversion, common to all sources of risk. If Q is diagonal, then its i^{th} diagonal element corresponds to relative risk aversion toward risk generated by the i^{th} Brownian motion. Remark 28 leads us to consider non-diagonal positive definite specifications of Q.

For each (ω, t), we define the function $\mathcal{Q}(\omega, t, \cdot) : \mathbb{R}^d \to \mathbb{R}$ as the following quadratic, where the argument z is a d-dimensional column vector:

$$\mathcal{Q}(t, z) = z' Q_t z - \left(\mu_t^R - \sigma_t^{R\prime}(Q_t - I)\,z\right)' \left(\sigma_t^{R\prime} Q_t \sigma_t^R\right)^{-1} \left(\mu_t^R - \sigma_t^{R\prime}(Q_t - I)\,z\right).$$

Under Condition 30, the optimality conditions (Condition 23) reduce to the following steps:

1. The shadow-price-of-wealth process λ solves the BSDE:

$$\frac{d\lambda_t}{\lambda_t} = -\left(r_t + g^*(t, \lambda) - \frac{1}{2}\mathcal{Q}\left(t, \sigma^\lambda\right)\right) dt + \sigma_t^{\lambda\prime} dB_t, \quad \lambda_T = 1. \tag{37}$$

2. Given the solution $\left(\lambda, \sigma^\lambda\right)$ from step one, the optimal strategy is

$$\rho_t = \lambda_t \mathcal{I}^g(t, \lambda_t) \quad \text{and} \quad \psi_t = \left(\sigma_t^{R\prime} Q_t \sigma_t^R\right)^{-1} \left(\mu_t^R - \sigma_t^{R\prime}(Q_t - I)\,\sigma_t^\lambda\right). \tag{38}$$

3. Given the strategy (ρ, ψ), the corresponding wealth process W is computed from the budget equation. The optimal consumption plan is $c = \rho W$, and its utility process is $U(c) = \lambda W$.

The proof of this claim (given in SS03) is a matter of direct calculation using the specific aggregator form, and the key homogeneity condition $U = \lambda W$. The optimal trading strategy expression can be derived as follows, where, for any strictly positive Ito process x, the notation σ^x is defined by $dx/x = ...dt + \sigma^{x\prime} dB$. The budget equation implies that $\sigma^W = \sigma^R \psi$, and $U = \lambda W$ implies that $\sigma^U = \sigma^\lambda + \sigma^W$. We have also seen that the supergradient density expression $\pi = \mathcal{E}\lambda$ leads to the restriction $\mu^R + \sigma^{R\prime}\left(F_\Sigma + \sigma^\lambda\right) = 0$. Using the fact that $F_\Sigma = -Q\sigma^U = -Q\left(\sigma^\lambda + \sigma^R \psi\right)$, the optimal trading strategy expression in (38) follows.

The above optimal trading strategy can deviate from an instantaneously mean-variance efficient solution for two possible reasons. One is source dependence of risk aversion (reflected in Q) and the other is the term involving σ^λ which arises from the dynamic stochastic nature of the investment opportunity set. Two special cases in which instantaneous mean-variance efficiency is recovered are given in the following examples. For expositional simplicity, we informally identify optimality with the first-order conditions of optimality.

Example 31 (Deterministic Investment Opportunity Set and Risk Aversion) *Suppose the processes r, μ^R, σ^R, and Q are all deterministic. Then the above three steps simplify significantly by setting $\sigma^\lambda = 0$. That is, λ is a deterministic process solving the ODE*

$$\frac{d\lambda_t}{\lambda_t} = -\left(r_t + g^*(t,\lambda) + \frac{1}{2}\mu_t^{R\prime}\left(\sigma_t^{R\prime}Q_t\sigma_t^R\right)^{-1}\mu_t^R\right)dt, \quad \lambda_T = 1.$$

Since λ is deterministic and g is assumed state-independent, the optimal consumption strategy $\rho = \lambda\mathcal{I}^g(\lambda)$ is also deterministic. The optimal trading strategy is $\psi = \left(\sigma^{R\prime}Q\sigma^R\right)^{-1}\mu^R$.

Suppose further that $Q = \gamma I$ for some deterministic process γ. Then the optimal trading strategy $\psi = \gamma^{-1}\left(\sigma^{R\prime}\sigma^R\right)^{-1}\mu^R$ is instantaneously mean-variance efficient, does not depend on g, and is therefore the same as for the case of time-additive power expected utility (considered by Merton (1971)). On the other hand, λ and the optimal consumption strategy depend on the specification of g. The investment opportunity set enters the dynamics of λ through the maximum squared conditional Sharpe ratio of equation (3)*.*

Example 32 (Robustly Mean-Variance Efficient Optimal Trading Strategies) *Even under a stochastic investment opportunity set, the instantaneously mean-variance efficient strategy $\psi = \left(\sigma^{R\prime}\sigma^R\right)^{-1}\mu^R$ is optimal if $Q = I$ (the identity matrix). Moreover, for $Q = I$, the investment opportunity set enters the BSDE for λ only through λ and the maximum squared instantaneous Sharpe ratio of equation* (3)*. Combining time additivity with the assumption $Q = I$ implies that g is logarithmic, while without time-additivity g is unrestricted. A discrete-time example of this type was first given by Giovannini and Weil (1989). The construction is further extended in Example 34 below.*

As noted earlier, in a Markovian setting, a BSDE is characterized (under some regularity) by a corresponding PDE. The argument is outlined below for the BSDE (37) satisfied by λ.

Example 33 (Markovian Solutions) *Given is some underlying n-dimensional Markov process Z, uniquely solving the SDE*

$$dZ = a(t,Z)\,dt + b(t,Z)'\,dB, \quad Z_0 = z_0,$$

for some $z_0 \in \mathbb{R}^n$ and functions $a : [0,T] \times \mathbb{R}^n \to \mathbb{R}^n$ and $b : [0,T] \times \mathbb{R}^n \to \mathbb{R}^{d\times n}$. With some convenient abuse of notation, we assume that $r_t = r(t,Z_t)$, $\eta_t = \eta(t,Z_t)$, and $\delta_t = \delta(t,Z_t)$, for some functions $r, \delta : [0,T] \times \mathbb{R}^n \to \mathbb{R}$ and $\eta : [0,T] \times \mathbb{R}^n \to \mathbb{R}^d$. We conjecture that λ can be written as a function of time and the Markov state that is smooth enough to apply Ito's lemma. With the usual abuse of notation, we write $\lambda(\omega,t) = \lambda(t, Z(\omega,t))$. Omitting the argument (t, Z_t), and with subscripts of λ denoting partial derivatives, Ito's lemma implies:

$$d\lambda = \left(\lambda_t + \lambda_z' a + \frac{1}{2}tr\left[b\lambda_{zz}b'\right]\right)dt + \lambda_z' b' dB.$$

Comparing the above dynamics to BSDE (37) *suggests that $\sigma^\lambda = b\lambda_z/\lambda$ and λ solves the PDE*

$$r + g^*(\lambda) - \frac{1}{2}\mathcal{Q}\left(b\frac{\lambda_z}{\lambda}\right) + \frac{\lambda_t}{\lambda} + a'\frac{\lambda_z}{\lambda} + \frac{1}{2}tr\left[b\frac{\lambda_{zz}}{\lambda}b'\right] = 0, \quad \lambda(T,\cdot) = 1,$$

where r and $\mathcal{Q}$ are viewed as functions of the underlying Markov state, in the same notational pattern used earlier for λ. Reversing the above steps, one can construct a solution to BSDE (37) *from a solution to the above PDE.*

5.2 Relating Complete and Incomplete Market Solutions

Continuing with the assumption of a smooth quasiquadratic proportional aggregator (Condition 30), we consider some connections between complete and incomplete market solutions. (We are of course leaving out the most important type of market incompleteness: undiversifiable income risk. A tractable class of problems dealing with the latter is outlined in the final section.)

We recall that m is the number of traded risky assets. Given any matrix A with n rows, where $n \geq m$, we use the block matrix notation:

$$A = \begin{bmatrix} A_M \\ A_N \end{bmatrix}, \quad A_M \in \mathbb{R}^m, \quad A_N \in \mathbb{R}^{n-m},$$

For example, $R = R_M$, $\mu^R = \mu^R_M$, and $\sigma^{R\prime} = \left[\sigma^{R\prime}_M, \sigma^{R\prime}_N\right]$.

While the solution summarized by BSDE (37) is valid for incomplete markets, the role of nonmarketed uncertainty becomes clearer after passing to a new Brownian motion that generates the same filtration as B and separates marketed and nonmarketed uncertainty. Informally, at each spot, the linear span of $dR_M - \mu^R_M dt$ can be obtained as the linear span of the first m elements of a rotated version of dB at the given spot. This transformation (stated formally in SS03) corresponds to the type of spot-by-spot Brownian motion rotation of Remark 28, which preserves the quasiquadratic proportional aggregator structure. We therefore lose no generality in assuming that

$$dR_M = \mu^R_M dt + \sigma^{R\prime}_M dB_M \quad \text{and} \quad \sigma^R_N = 0. \tag{39}$$

For the remainder of this section, we assume the normalized return structure (39), and we think of M and N as sets of indices corresponding to marketed and nonmarketed uncertainty, respectively. The processes r, μ^R_M, and σ^R_M need not be adapted to the filtration generated by B_M.

A market-price-of-risk process in this context takes the form

$$\eta = \begin{bmatrix} \eta_M \\ \eta_N \end{bmatrix}, \quad \text{where} \quad \eta_M = \left(\sigma^{R\prime}_M\right)^{-1} \mu^R_M. \tag{40}$$

The process η_M represents the *price of marketed risk*, while the unrestricted process η_N represents the *price of nonmarketed risk.* The latter parameterizes the set of every state price density π consistent with the given market:

$$\pi = \pi^M \xi^{\eta_N}, \quad \text{where} \quad \frac{d\pi^M}{\pi^M} = -r dt - \eta'_M dB_M \quad \text{and} \quad \frac{d\xi^{\eta_N}}{\xi^{\eta_N}} = -\eta'_N dB_N.$$

If π is a state price density that is also a utility supergradient density at an optimum, then the corresponding η_N reflects the shadow price of nonmarketed risk, in the following sense: Consider a hypothetical market completion in which risk generated by dB_N is priced by η_N. In such a market, the investor would find it optimal to not trade risk generated by dB_N, since the original incomplete-markets strategy would still be optimal. Since the original strategy need not be optimal under any other choice of η_N, the incomplete-markets optimal utility is the minimum of optimal utilities over all market completions (parameterized by η_N). This connection between complete and incomplete market solutions is illustrated more concretely in Example 35 below, and extends to more general convex constraints (see Cvitanić and Karatzas (1992) and Karatzas and Shreve (1998) for the case of time-additive expected utility, and SS03 and Appendix A of Schroder and Skiadas (2005b) for the case of recursive utility).

For expositional simplicity, in the remainder of this section we further assume that the relative risk aversion process Q assumes the block diagonal structure

$$Q = \begin{bmatrix} Q_{MM} & 0 \\ 0 & Q_{NN} \end{bmatrix}, \tag{41}$$

where $Q_{MM} \in \mathcal{L}(\mathbb{R}^{m \times m})$ and $Q_{NN} \in \mathcal{L}\left(\mathbb{R}^{(d-m)\times(d-m)}\right)$. In this context, the BSDE (37) of the first-order conditions satisfied by λ holds with

$$\mathcal{Q}\left(\sigma^{\lambda}\right) = \sigma_N^{\lambda\prime} Q_{NN} \sigma_N^{\lambda} + 2\left(\eta_M + \sigma_M^{\lambda}\right)' \sigma_M^{\lambda} - \left(\eta_M + \sigma_M^{\lambda}\right)' Q_{MM}^{-1} \left(\eta_M + \sigma_M^{\lambda}\right). \tag{42}$$

The corresponding optimal trading strategy is

$$\psi_M = \left(Q_{MM} \sigma_M^R\right)^{-1} \left(\eta_M - \left(Q_{MM} - I_{MM}\right) \sigma_M^{\lambda}\right),$$

where I_{MM} is the $m \times m$ identity matrix.

Example 34 (Mean-Variance Efficiency) *If $Q_{MM} = I_{MM}$, then ψ_M is instantaneously mean-variance efficient, which extends Example 32.*

Example 35 (Fictitious Market Completion and Duality) *Consider the above incomplete-market setting, with the normalized return dynamics (39), where $m < d$, and the block-diagonal Q in (41). Suppose that $\left(\lambda, \sigma^{\lambda}\right)$ solves the BSDE of the optimality conditions, (ρ, ψ_M) is the corresponding optimal strategy, and U is the corresponding optimal utility process.*

Given any choice of a nonmarketed-price-of-risk process η_N, we consider the complete market obtained by introducing $d - m$ fictitious assets, whose cumulative excess return process $R_N = \left(R^{m+1}, \ldots, R^d\right)'$ follows the dynamics $dR_N = \eta_N dt + dB_N$. The unique market-price-of-risk process in this fictitious complete market is given by (40). We let U^{η_N} denote the corresponding complete-market optimal utility process. Simple algebra shows that if one makes the specific selection

$$\eta_N = \left(Q_{NN} - I_{NN}\right) \sigma_N^{\lambda},$$

then $\left(\lambda, \sigma^{\lambda}\right)$ satisfies the BSDE of the optimality conditions in the fictitious complete market defined by this choice of η_N. Moreover, the corresponding optimal strategy in the fictitious complete market is $\left(\rho, (\psi_M, \psi_N)'\right)$, where (ρ, ψ_M) is the incomplete-market optimal strategy and $\psi_N = 0$. In other words, the above specification of η_N prices nonmarketed risk so that the investor finds it optimal to not trade the fictitious assets at all. As a consequence $U = U^{\eta_N}$. For any other choice of η_N, the strategy $(\rho, (\psi_M, 0))$ need not be optimal in the fictitious complete market defined by η_N, and therefore $U \leq U^{\eta_N}$.

A different type of connection between incomplete and complete market solutions is given in the following example (which is generalized in SS03). A particular case of the example shows that if the investor has the time-additive expected power utility (32) with $\gamma \in (0, 2)$, then the solution to the investor's problem in an incomplete market is equivalent (in a sense clarified below) to the solution of the complete market problem obtained by pricing nontraded uncertainty risk-neutrally, and setting the investor's relative risk aversion toward nonmarketed uncertainty to $1/(2 - \gamma)$. The original additive-utility problem with incomplete markets is therefore equivalent to a complete-markets problem with recursive utility.

Example 36 (Market-Incompletess and Source-Dependent Risk Aversion) *We further specialize the quasiquadratic form* (36) *of the proportional aggregator by assuming that*

$$Q = \gamma I, \quad \text{where } \gamma \in (0, 2).$$

In Example 27 we saw that this class includes cases of Epstein-Zin utility, as well as time-additive expected discounted power utility. Let (ρ, ψ_M) *be an incomplete-markets optimal strategy, with corresponding shadow-price-of-wealth process* λ*, wealth process* W*, and utility process* U*.*

We complete the market by introducing fictitious assets that are priced risk-neutrally; that is, the price-of-nonmarketed risk process is zero $(\eta_N = 0)$*. We let the corresponding excess return dynamics be given by* $R_N = B_N$*. In the resulting fictitious complete market, we consider the optimal strategy, not of the original investor, but rather of a fictitious investor whose proportional aggregator is*

$$\bar{G}(t, c, \sigma) = g(t, c) - \frac{1}{2}\left(\gamma \sigma_M' \sigma_M + \frac{1}{2-\gamma} \sigma_N' \sigma_N\right).$$

In other words, the fictitious investor's relative risk aversion toward nonmarketed risk is modified from γ *to* $1/(2-\gamma)$*. Let* $(\bar{\rho}, \bar{\psi})$ *be the optimal strategy of the fictitious investor in the fictitious complete market, and let* $\bar{\lambda}$*,* $\bar{W}$ *and* $\bar{U}$ *be the corresponding shadow-price-of wealth, wealth, and utility processes. Comparing optimality conditions, we observe that*

$$\bar{\lambda} = \lambda, \quad \bar{\rho} = \rho, \quad \bar{\psi}_M = \psi_M, \quad \text{and} \quad \frac{\bar{W}_t}{W_t} = \frac{\bar{U}_t}{U_t} = \exp\left(\int_0^t \bar{\psi}_N{}' dB_N\right).$$

The incomplete market solution can therefore be immediately recovered from the fictitious complete-market solution. This is true even though the specification of the fictitious-investor preferences does not depend on market prices!

5.3 Solutions Based on Quadratic BSDEs

We have seen that for homothetic recursive utility the optimality conditions reduce to a single BSDE satisfied by the shadow-price-of-wealth process λ. In this section, we specify a quasiquadratic proportional aggregator for which the BSDE satisfied by $\log(\lambda)$ takes a quadratic form. Under further assumptions on the dynamics of the investment opportunity set, the solution to this BSDE can be expressed as a quadratic function of the state, with deterministic coefficients that solve an ODE system. This type of solution is familiar from a class of term-structure modes surveyed by Duffie (forthcoming) and Piazzesi (2005). In risk-neutral pricing the relevant BSDE is linear. Our application extends the solution method to quadratic BSDEs, where the quadratic term reflects risk aversion. For expositional simplicity, we outline below only some examples, referring to SS03 for a more general treatment. Extensions with unpredictable jumps are given in Schroder and Skiadas (2005b). Examples of this type of solution can also be found in Kim and Omberg (1996), Chacko and Viceira (forthcoming), Schroder and Skiadas (1999), Liu (2001), and Wachter (2002).

Continuing with the assumption of a homothetic recursive utility, in this section we further specialize the proportional aggregator to be of the form

$$G(t, c, \sigma) = \alpha + \beta \log(c) - \frac{\gamma}{2}\sigma'\sigma, \tag{43}$$

for some constants $\alpha \in \mathbb{R}$ and $\beta, \gamma \in \mathbb{R}_{++}$. The parameters (α, β) determine preferences over deterministic choices. Given (α, β), the parameter γ adjusts risk aversion.

Remark 37 *The treatment in SS03 allows possible source-dependent risk aversion, and parameters α and β that are processes, the latter deterministic. The above specification for $\beta = 0$ results in a utility that is ordinally equivalent to expected power utility for terminal consumption. Even though we have not covered the case of no intermediate consumption, essentially the same analysis applies.*

Recalling Example 26, the optimal strategy for the above utility specification is

$$\rho = \beta \quad \text{and} \quad \psi = \frac{1}{\gamma}\left(\sigma^{R\prime}\sigma^{R}\right)^{-1}\left(\mu^{R} - (\gamma - 1)\,\sigma^{R\prime}\sigma^{\lambda}\right).$$

A myopic solution results for $\gamma = 1$, corresponding to time-additive logarithmic utility (the intersection of Examples 26 and 32). The solution also simplifies if the investment opportunity set is deterministic, in which case $\sigma^{\lambda} = 0$ (Example 31). To compute the optimal strategy with a stochastic investment opportunity set and $\gamma \neq 1$, we need to determine $\left(\lambda, \sigma^{\lambda}\right)$ by solving BSDE (37). Making the convenient change of variables

$$\ell_t = \log(\lambda_t),$$

we note the above condition implies that $g^*(t, \lambda) = \alpha - \beta + \beta \log(\beta) - \beta \ell_t$. Direct computation then shows that BSDE (37) can be written as the quadratic BSDE in ℓ :

$$d\ell_t = -\left(p_t - \beta \ell_t + h_t' \sigma_t^{\ell} + \frac{1}{2}\sigma_t^{\ell\prime} H_t \sigma_t^{\ell}\right) dt + \sigma_t^{\ell\prime} dB, \quad \ell_T = 0, \tag{44}$$

where

$$p = r + \alpha - \beta + \beta \log(\beta) + \frac{1}{2\gamma}\mu^{R\prime}\left(\sigma^{R\prime}\sigma^{R}\right)^{-1}\mu^{R},$$

$$h = \frac{1-\gamma}{\gamma}\sigma^{R}\left(\sigma^{R\prime}\sigma^{R}\right)^{-1}\mu^{R}, \quad H = (1-\gamma)\left[I + \frac{1-\gamma}{\gamma}\sigma^{R}\left(\sigma_t^{R\prime}\sigma^{R}\right)^{-1}\sigma^{R\prime}\right].$$

A general set of conditions under which the above quadratic BSDE can be reduced to an ODE is given in SS03. We only consider here two representative examples. As in the last section, we assume the normalization $dR = \mu^R dt + \sigma_M^{R\prime} dB_M$, and therefore the price-of-marketed-risk process is $\eta_M = \left(\sigma_M^{R\prime}\right)^{-1}\mu^R$. We outline the form of the solution below, leaving the details as an exercise.

Example 38 *Given is an underlying n-dimensional state vector Z following the dynamics*

$$dZ = (\mu - \theta Z)\, dt + \Sigma' dB,$$

for some $\mu \in \mathbb{R}^n$, $\Sigma \in \mathbb{R}^{d \times n}$, and $\theta \in \mathbb{R}^{n \times n}$. The short rate process and price of marketed risk are assumed to be given by

$$r = C_0^r + C_1^{r\prime} Z + \frac{1}{2} Z' C_2^r Z \quad \text{and} \quad \eta_M = C_0^{\eta} + C_1^{\eta\prime} Z,$$

where the coefficients C_i^r and C_i^{η} are all constants of conforming dimensions. In this case, we conjecture a solution to BSDE (44) of the form

$$\ell_t = C_0(t) + C_1(t)' Z_t + \frac{1}{2} Z' C_2(t) Z_t,$$

where the $C_i(t)$ are deterministic differentiable processes. Applying Ito's lemma to the above conjectured expression, collecting terms and comparing to the corresponding coefficients of BSDE (44) confirms that such a solution indeed solves BSDE (44), provided the coefficients C_i solve a system of ordinary differential equations.

Example 39 *We modify the above example by assuming the dynamics*

$$dZ = (\mu - \theta Z)\, dt + \Sigma' \mathit{diag}\left(\sqrt{\upsilon + VZ}\right) dB,$$
$$r = C_0^r + C_1^{r\prime} Z \quad \mathit{and} \quad \eta_M = \mathit{diag}\left(\sqrt{\upsilon_M + V_{M*} Z_t}\right) \varphi,$$

where $diag(x)$ denotes the diagonal matrix with x on the diagonal, $\sqrt{x}$ denotes the vector with i^{th} element $\sqrt{x_i}$, and $\mu \in \mathbb{R}^n$, $\Sigma \in \mathbb{R}^{d\times n}$, $\theta \in \mathbb{R}^{n\times n}$, $C_0^r \in \mathbb{R}$, $C_1^r \in \mathbb{R}^n$, $\upsilon = [\upsilon_M', \upsilon_N']' \in \mathbb{R}^d$, $V = [V_{M*}', V_{N*}']' \in \mathbb{R}^{d\times n}$, $\varphi \in \mathbb{R}^m$. In this case, the conjectured solution takes the form

$$\ell_t = C_0(t) + C_1(t)' Z,$$

where again the $C_i(t)$ are differentiable deterministic processes. Arguing as in the last example, one obtains an ODE solved by C_1 alone, and another ODE (which uses the solution of the first ODE) that is satisfied by C_0. Given the pair (C_0, C_1) satisfying the ODE pair, the above affine expression defines a solution to BSDE (44).

5.4 Solutions with First-Order Risk Aversion

The final set of scale-invariant solutions we consider utilizes the kinked proportional aggregator of Section 4.4, representing source-dependent first-order risk aversion. More specifically, we assume the following condition, using the notation

$$|x|' = (|x_1|, \ldots, |x_d|), \quad x \in \mathbb{R}^d.$$

Condition 40 (Quasi-Quadratic Proportional Aggregator) *Condition 29 holds with*

$$\mathcal{R}(\omega, t, c, \sigma) = \kappa(\omega, t)' |\sigma| + \frac{1}{2}\sigma' Q(\omega, t)\, \sigma,$$

*for some bounded processes $\kappa : \Omega \times [0, T] \to \mathbb{R}^d$ and $Q : \Omega \times [0, T] \to \mathbb{R}^{d\times d}$, where $Q(\omega, t)$ is **diagonal** and positive definite for all (ω, t).*

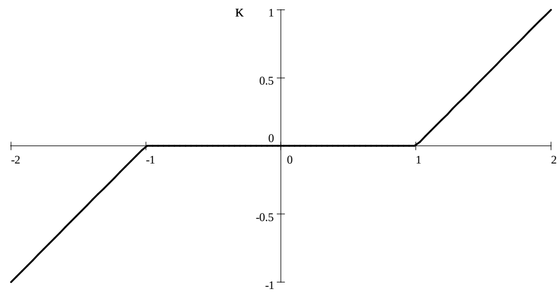
We adopt the notation and return normalization of Section 5.2. In particular, the excess return dynamics and the marketed-price-of-risk process are

$$dR = \mu^R dt + \sigma_M^{R\prime} dB_M \quad \text{and} \quad \eta_M = \left(\sigma_M^{R\prime}\right)^{-1} \mu_M^R.$$

To formulate optimality conditions in this setting, we define, for any $\kappa \in \mathbb{R}_+$, the *collar function*

$$K(\alpha; \kappa) = \min\{\max\{0, \alpha - \kappa\}, \alpha + \kappa\}, \quad \alpha \in \mathbb{R},$$

plotted below for $\kappa = 1$:



The collar function will be applied to vectors coordinate by coordinate:

$$K(\alpha;\kappa) = (K(\alpha_1;\kappa_1), \ldots, K(\alpha_m;\kappa_m))', \quad \text{for any } \alpha \in \mathbb{R}^m \text{ and } \kappa \in \mathbb{R}^m_+.$$

The BSDE for λ in this case is of the same form as in the smooth-quasi-quadratic case, except that $\mathcal{Q}$ is replaced by the function $\mathcal{K} : \Omega \times [0,T] \times \mathbb{R}^d \to \mathbb{R}$ defined by

$$\mathcal{K}(z) = 2\kappa_N' |z_N| + z_N' Q_{NN} z_N + 2(\eta_M + z_M)' z_M - K(\eta_M + z_M;\kappa_M)' Q_{MM}^{-1} K(\eta_M + z_M;\kappa_M).$$

Expression (42) for $\mathcal{Q}$ is recovered if one sets $\kappa = 0$.

With the above notation and return normalization in place, the optimality conditions (Condition 23) under Condition 40 reduce to the following steps:

1. The shadow-price-of-wealth process λ solves the BSDE:

$$\frac{d\lambda_t}{\lambda_t} = -\left(r_t + g^*(t,\lambda_t) - \frac{1}{2}\mathcal{K}\left(t,\sigma_t^\lambda\right)\right) dt + \sigma_t^{\lambda\prime} dB_t, \quad \lambda_T = 1.$$

2. Given the solution $\left(\lambda, \sigma^\lambda\right)$ from step one, the optimal strategy is (omitting time indices):

$$\rho = \lambda \mathcal{I}^g(\lambda) \quad \text{and} \quad \psi = \left(\sigma_M^R\right)^{-1}\left[Q_{MM}^{-1} K\left(\eta_M + \sigma_M^\lambda;\kappa_M\right) - \sigma_M^\lambda\right].$$

3. Given the strategy (ρ,ψ), the corresponding wealth process W is computed from the budget equation. The optimal consumption plan is $c = \rho W$, and its utility process is $U(c) = \lambda W$.

The proof of the above claim (given in SS03) is a matter of direct calculation using the specific aggregator form, and the key homogeneity condition $U = \lambda W$. The latter also implies that if $dU/U = \ldots dt + \sigma' dB$, then

$$\sigma_M = (Q_{MM})^{-1} K\left(\eta_M + \sigma_M^\lambda;\kappa_M\right).$$

Consequently, for any $i \in M$, σ^i vanishes whenever $\eta_M^i + \sigma_M^{\lambda i} \in [-\kappa_i, \kappa_i]$. Such perfect hedging of utility risk with respect to some source of risk is not encountered with second-order risk aversion alone. The following example (from SS03) extends Section 5.3 of Chen and Epstein (2002). Further examples of non-participation as an expression of source-dependent first-order risk aversion can be found in Epstein and Miao (2000) and SS03.

Example 41 (Deterministic Investment Opportunity Set) *Suppose that r, μ^R, σ^R, κ, and Q are all deterministic. Then the solution simplifies by setting $\sigma^\lambda = 0$. In particular, the optimal trading strategy is $\psi = \left(Q_{MM}\sigma_M^R\right)^{-1} K(\eta_M; \kappa_M)$. Let us further assume, for simplicity, that σ_{MM}^R is diagonal with positive diagonal. For any $i \in M$, $\psi_i = 0$ when $\eta_i \in [-\kappa_i, +\kappa_i]$; the agent will not participate in the market for risk i, unless its instantaneous expected return relative to its risk is sufficiently far from zero. This type of solution can be combined with different belief specifications, as in Remark 16, to obtain a richer set of optimal portfolio holdings. In particular, adding the term $b'\sigma$ to the proportional aggregator, for some (bounded) process b, means that the investor believes the market price of risk process to be $\eta - b$, rather than η, and therefore the investor will not participate in the market for risk source i if $\eta_i \in [b_i - \kappa_i, b_i + \kappa_i]$. If we further assume that $b_i = -\kappa_i$, then the optimal holding of asset i is $\psi_i = Q_{ii}^{-1}\mu_i^R / \left(\sigma_{ii}^R\right)^2$ when $\mu_i^R > 0$ (just as with $\kappa = b = 0$), but the agent will only short asset i when $\mu_i^R < -2\kappa_i\sigma_{ii}^R$. In other words, in this case, the optimal portfolio is identical to the Merton solution for positive expected excess returns, yet it is optimal for the investor to not go short for sufficiently small negative expected returns.*

6 Extensions

This section concludes with two direct extensions of the main chapter's material, followed by a list of further topics and related references.

6.1 Convex Trading Constraints

We outline an extension of this chapter's arguments to include convex trading constraints, referring to SS03 for details. Examples of analysis of the Merton problem with constraints based on the Hamilton-Jacobi-Bellman approach include Zariphopoulou (1994) and Vila and Zariphopoulou (1997). Convex duality with trading constraints and additive utility is studied by He and Pearson (1991), Karatzas, Lehoczky, Shreve, and Xu (1991) (incomplete markets); Shreve and Xu (1992) (short-sale constraints); and Cvitanić and Karatzas (1992) (convex constraints). Related discussions with recursive preferences can be found in El Karoui, Peng, and Quenez (2001), SS03, and Schroder and Skiadas (2005b). We will not discuss duality here. Also not discussed here are constraints that prevent the investor to borrow against future income, which is the focus of He and Pagès (1993), El Karoui and Jeanblanc-Picquè (1998), and Detemple and Serrat (2003).

We consider this chapter's setting with the additional constraint that the investor's trading strategy must be valued in some given convex set $K \subseteq \mathbb{R}^m$ at all times. For example, $K = \mathbb{R}_+^m$ represents the impossibility of short-selling. The definition of a feasible cash flow now includes the requirement that it can be financed by a K-valued trading strategy. We let $\delta_K(\varepsilon) = \sup\{k'\varepsilon_t : k \in K\}$ denote the support function of K.

We fix a feasible strategy (ρ, ψ) financing the consumption plan c. Given our new notion of feasibility, the definition of a state-price density at c is the same as before. The smaller the set K, the smaller the set of feasible incremental cash flows, and therefore the larger the set of state price densities at c. Under some regularity, state price dynamics are characterized in SS03 as

$$\frac{d\pi_t}{\pi_t} = -\left(r_t + \delta_K(\varepsilon_t)\right)dt - \eta_t' dB_t, \quad \varepsilon_t = \mu_t^R - \sigma_t^{R\prime}\eta_t, \quad \psi_t'\varepsilon_t = \delta_K(\varepsilon_t).$$

Proposition 3 still applies here, so combining the above dynamics with those of a utility supergradient density results in sufficient optimality conditions as a constrained FBSDE system.

As in the unconstrained case, scale invariance results in the uncoupling of the forward and backward components of the BSDE system. For example, consider a scale-invariant recursive utility with the smooth quasi-quadratic proportional aggregator of Condition 30. As shown in SS03, in this case the optimality conditions can be written as the constrained BSDE:

$$\frac{d\lambda_t}{\lambda_t} = -\left(r_t + \delta_K(\varepsilon_t) + g^*(t,\lambda_t) - \frac{1}{2}\sigma_t^{\lambda\prime} Q_t \sigma_t^\lambda + \frac{1}{2}\psi_t' \sigma_t^{R\prime} Q_t \sigma_t^R \psi_t\right) dt + \sigma_t^{\lambda\prime} dB_t, \quad \lambda_T = 1,$$

$$\psi_t = \left(\sigma_t^{R\prime} Q_t \sigma_t^R\right)^{-1} \left(\mu_t^R - \varepsilon_t - \sigma_t^{R\prime}(Q_t - I)\sigma_t^\lambda\right) \in K, \quad \psi_t'\varepsilon_t = \delta_K(\varepsilon_t).$$

Example 42 *Under Condition 30, a particularly simple expression for the optimal trading strategy is obtained if* $K = \{k \in \mathbb{R}^m : \alpha \le l'k \le \beta\}$ *where* $l \in \mathbb{R}^m$ *and* α *and* β *are valued in* $[-\infty, +\infty]$. *The case of a short-sale constraint on asset* i *corresponds to* $\alpha = 0$, $\beta = \infty$, *and* l *a vector of zeros except for a one in the* i^{th} *position. The case of a cap on the proportion of wealth borrowed, possibly combined with a limit on short sales as a fraction of wealth, corresponds to letting* l *be a vector of ones. We assume that* K *is nonempty, and define*

$$\psi_t^* = A_t\left(\mu_t^R - \sigma_t^{R\prime}(Q_t - I)\sigma_t^\lambda\right), \quad A_t = \left(\sigma_t^{R\prime} Q_t \sigma_t^R\right)^{-1}.$$

The above expression gives the optimal trading strategy as a function of σ^λ *in the unconstrained case* $(\alpha = -\infty,\ \beta = \infty)$. *The (constrained) optimal trading strategy* ψ *and process* ε *in the dynamics of* λ *are are given by*

$$\psi_t = \psi_t^* - A_t\varepsilon_t, \quad \varepsilon_t = -\left(l'A_t l\right)^{-1} l\left(\min\left\{\max\left\{l'\psi_t^*, \alpha\right\}, \beta\right\} - l'\psi_t^*\right). \tag{45}$$

6.2 Translation-Invariant Formulations and Nontradeable Income

A parallel theory to this chapter's scale-invariance argument is based on a notion of translation invariance in a setting that allows for a nontradeable income stream. This type of formulation is familiar in the subclass of problems with expected discounted exponential utility and Gaussian dynamics, as, for example, in Svensson and Werner (1993) and Musiela and Zariphopoulou (2004)). We outline below a formulation with recursive utility, which is a special case of Schroder and Skiadas (2005a,b).

We modify our earlier setting by assuming that the investor is endowed, in addition to the initial wealth w_0, with an income stream e, a possibly nontradeable cash flow. Consumption in this subsection is allowed to take negative values, and financial wealth can vanish. The representation of portfolios in terms of proportions of wealth is therefore unsuitable in our new setting. We correct this by defining a trading plan to be a process $\phi \in \mathcal{L}(\mathbb{R}^m)$, where ϕ_t^i represents a dollar amount invested in asset i at time t. The dollar amount invested in the money market at time t is $W_t - \sum_{i=1}^m \phi_t^i$, where W_t represents total time-t financial wealth (excluding e). Ignoring some integrability requirements, a *plan* is a triple (c, ϕ, W) of a consumption plan c, a trading plan ϕ, and a wealth process W. The plan (c, ϕ, W) is *feasible* if it satisfies the budget equation:

$$W_0 = w, \quad dW_t = (r_t W_t + e_t - c_t)\,dt + \phi_t' dR_t, \quad c_T = W_T + e_T. \tag{46}$$

The derivation and form of the optimality conditions as a FBSDE system in this setting is similar to this chapter's main analysis, as explained in Schroder and Skiadas (2005a,b).

We place restrictions on the market and preferences in terms of a strictly positive (bounded) cash flow γ, that is fixed throughout. On the market side, we assume there is a tradeable fund that generates γ as a dividend stream. We refer to this fund as the "γ-fund," and we let Γ and ϱ be its value process and trading plan, respectively. The γ-fund budget equation is

$$d\Gamma_t = (r_t\Gamma_t - \gamma_t)\,dt + \varrho_t' dB_t, \quad \Gamma_T = 1.$$

We also define the corresponding dividend-yield process δ and allocation process κ, by[15]

$$\delta_t = \frac{\gamma_t}{\Gamma_t} \quad \text{and} \quad \kappa_t = \frac{\varrho_t}{\Gamma_t}.$$

For example, if r and γ are deterministic, the γ-fund can be implemented entirely through the money market, meaning that $\varrho = 0$. More generally, one can assume that risky asset one is a share in the γ-fund, and therefore $\kappa_t = (1, 0, \ldots, 0)$, $\varrho^1 = \Gamma$, and $\varrho^j = 0$ for $j \in \{2, \ldots, m\}$.

On the preference side, we assume that the investor's time-zero utility function is *translation-invariant with respect to* γ, meaning that, for any consumption plans a and b,

$$U_0(a) = U_0(b) \quad \text{implies} \quad U_0(a + k\gamma) = U_0(b + k\gamma) \text{ for all } k \in \mathbb{R}.$$

If utility is normalized so that the investor is indifferent between the consumption plan c and the consumption plan $U_0(c)\,\gamma$, the above property can equivalently be stated as quasilinearity with respect to γ; that is, $U_0(c + k\gamma) = U_0(c) + k$ for any consumption plan c and scalar k. For recursive utility, the latter restriction is essentially equivalent to the BSDE form:

$$dU_t = -G\left(t, \frac{c_t}{\gamma_t} - U_t, \Sigma_t\right) dt + \Sigma_t' dB_t, \quad U_T = c_T, \tag{47}$$

for a possibly state-dependent function G that we call an *absolute aggregator*. For concreteness, we combine the above representation with our earlier formulation of possibly source-dependent second-order risk aversion, resulting in the quasi-quadratic absolute aggregator specification

$$G(t, x, \Sigma) = g(t, x) - \frac{1}{2}\Sigma' Q_t \Sigma. \tag{48}$$

In the remainder of this subsection, we assume this absolute aggregator form, with the same restrictions on g and Q as in Condition 30.

Example 43 (Expected Discounted Exponential Utility) *Let β be any (say bounded) process, and suppose the utility process V of the plan c is well-defined by*

$$V_t = E_t\left[\int_t^T -\exp\left(-\int_t^s \beta_u du - \frac{c_s}{\gamma_s}\right) ds - \exp\left(-\int_t^T \beta_u du - \frac{c_T}{\gamma_T}\right)\right].$$

Then the ordinally equivalent utility process $U_t = -\log(-V_t)$ solves BSDE (47) *with the absolute aggregator* (48)*, where $Q(\omega, t) = 1$ and $g(\omega, t, x) = \beta(\omega, t) - \exp(-x)$.*

[15] In Schroder and Skiadas (2005a) ϱ was set equal to a constant for simplicity, but the analysis applies essentially unchanged with ϱ being a more general trading plan as assumed here. A more relevant simplification is obtained if κ is constant, as in Schroder and Skiadas (2005b).

Analogously to the scale-invariance argument, translation-invariance with respect to γ uncouples the FBSDE of the first-order conditions. Intuitively, if the agent's problem is solved at some information spot at a given financial wealth level, it is also solved at all wealth levels, since the agent can always invest any additional wealth to the γ-fund while preserving optimality.

More specifically, at the optimum, the utility process U, the wealth process W, and the shadow-price-of-wealth process λ, are related by

$$U_t = \frac{1}{\Gamma_t}\left(Y_t + W_t\right) \quad \text{and} \quad \lambda_t = \frac{1}{\Gamma_t},$$

where the process Y solves the quadratic BSDE

$$dY_t = -\left(e_t + p_t - r_t Y_t + \Sigma_t^{Y\prime} h_t + \frac{1}{2}\Sigma_t^{Y\prime} H_t \Sigma_t^Y\right) dt + \Sigma_t^{Y\prime} dB_t, \quad Y_T = e_T,$$

with

$$\begin{aligned}
p &= \Gamma g^*(\delta) + \frac{\Gamma}{2}\left(\mu^R - \sigma_t^{R\prime}\sigma_t^R\kappa\right)'\left(\sigma^{R\prime}Q\sigma^R\right)^{-1}\left(\mu^R - \sigma_t^{R\prime}\sigma_t^R\kappa\right), \\
h &= -\sigma^R\kappa - Q\sigma^R\left(\sigma^{R\prime}Q\sigma^R\right)^{-1}\left(\mu^R - \sigma_t^{R\prime}\sigma_t^R\kappa\right), \\
H &= \frac{1}{\Gamma}\left(Q\sigma^R\left(\sigma^{R\prime}Q\sigma^R\right)^{-1}\sigma^{R\prime}Q - Q\right).
\end{aligned}$$

The optimal plan trading plan ϕ and consumption plan c can be written as

$$\phi = \phi^0 + U\varrho, \quad c = \gamma U + \gamma g_x^{-1}\left(\frac{1}{\Gamma}, \frac{\sigma^R\phi^0 + \Sigma^Y}{\Gamma}\right),$$
$$\text{where} \quad \phi^0 = \left(\sigma^{R\prime}Q\sigma^R\right)^{-1}\left(\Gamma\mu^R - \sigma^{R\prime}\left(\sigma^R\varrho + Q\Sigma^Y\right)\right).$$

Just as with the quadratic BSDE case of the scale-invariant formulation, for a certain class of price dynamics the solution reduces to an ODE system. We refer to Schroder and Skiadas (2005a,b) for details, further examples, as well as extensions outlined below.

6.3 Other Directions

We conclude with a list of selected related topics and a sample of associated references that can be consulted for a large number of further related references. Naturally, important areas are not mentioned, including a large body of empirical work, as well as numerical studies, including simulation methods.

Transaction Costs: The Merton analysis has been extended to include proportional transaction costs by Davis and Norman (1990), Shreve and Soner (1994), Liu and Loewenstein (2002), and others. Grossman and Laroque (1990) and Cuoco and Liu (2000) studied problems in which transaction costs apply to changes in the stock of a durable good. Proportional transaction costs preserve scale invariance, motivating the use of expected discounted power utility in the above papers. Fixed transaction costs on the other hand destroy scale invariance. For this reason existing analytically tractable formulations with fixed transaction costs are based on translation invariance, so far only with additive exponential utility, as in Vayanos (1998) and Liu (2004). Optimality conditions with both proportional and fixed transaction costs with i.i.d. returns are given in Oksendal and Sulem (2002). The above is only a sample of a large number of papers with some form of transaction costs. I am not aware of any related literature that uses recursive utility.

Nontradeable Income: We have seen that simplifications in the optimality conditions with nontradeable income result in the translation-invariant formulation. The latter excludes, for example, any notion of constant relative risk aversion, and is unrealistic as a model of personal investment decisions. More general models of nontradeable income must deal with a fully coupled FBSDE system. The Merton problem with nontradeable income and additive utility has been analyzed in terms of the Hamilton-Jacobi-Bellman approach by Duffie and Zariphopoulou (1993), Duffie, Fleming, Soner, and Zariphopoulou (1997), and Koo (1998). Related theoretical results with nontradeable income and additive utilities include Cuoco (1997), Kramkov and Schachermeyer (1999, 2003), Cvitanić, Schachermayer, and Wang (2001), and Hugonnier and Kramkov (2002).

Endogenous labor supply and retirement: Bodie, Merton, and Samuelson (1992) and Bodie, Detemple, Otruba, and Walter (2004), among others, have analyzed the lifetime consumption-portfolio problem with endogenous labor supply. The endogenous decision to retire has recently been revisited by Dybvig and Liu (2005) and Farhi and Panageas (2005). Recursive utility formulations in this area are yet to be developed.

Habit Formation: Early asset pricing models with preferences incorporating habit formation include Sundaresan (1989), Constantinides (1990), and Detemple and Zapatero (1991). Duffie and Skiadas (1994) introduced recursive utility with habit formation, and computed the corresponding utility gradient density. The latter can be used to formulate optimality conditions as a FBSDE system, similarly to this chapter's analysis. For the case of linear habit formation (or durability), Schroder and Skiadas (2002) gave formulas for transforming state price dynamics and optimal strategies that reduce the investor's problem with habit formation to one without habit formation. This technique can be used to mechanically translate this chapter's solutions under complete markets (or under incomplete markets with a deterministic short rate process) to corresponding solutions that incorporate linear habit formation.

Discontinuous Information: Merton's original work includes examples of discontinuous information generated by Poisson jumps. The extension of Merton's work to Lévy type processes using the Hamilton-Jacobi-Bellman approach is presented in the monograph by Øksendal and Sulem (2005). This chapter's arguments are extended in Schroder and Skiadas (2005b) so that the filtration is generated by Brownian motions as well as quite general marked point processes. The above references provide links to several other papers on this topic.

Nonlinear Wealth Dynamics: Cuoco and Cvitanić (1998), El Karoui, Peng, and Quenez (2001), and Schroder and Skiadas (2005b) characterize optimality with wealth dynamics that can allow nonlinearities reflecting, for example, market impact, or differential borrowing and lending rates, or certain limited forms of taxation. The last reference includes the extension of this chapter's scale/translation invariance arguments in this case.

References

ANDERSON, E., L. HANSEN, AND T. SARGENT (2000): "Robustness, Detection and the Price of Risk," working paper, Dept. of Economics, University of Chicago.

ARROW, K. (1965): *Aspects of the Theory of Risk Bearing*. Yrjo Jahnssonin Saatio, Helsinki.

——— (1970): *Essays in the Theory of Risk Bearing*. North Holland, London.

BODIE, Z., J. B. DETEMPLE, S. OTRUBA, AND S. WALTER (2004): "Optimal Consumption-Portfolio Choices and Retirement Planning," *Journal of Economic Dynamics and Control*, 28, 1115–1148.

BODIE, Z., R. C. MERTON, AND W. F. SAMUELSON (1992): "Labor Supply Flexibility and Portfolio Choice in a Life Cycle Model," *Journal of Economic Dynamics and Control*, 16, 427–49.

BRANDT, M. W. (forthcoming): "Portfolio Choice Problems," in Y. Ait-Sahalia and L.P. Hansen (eds.), Handbook of Financial Econometrics, Elsevier - North Holland, New York, NY.

CAMPBELL, J., AND L. VICEIRA (2002): *Strategic Asset Allocation*. Oxford University Press, New York.

CHACKO, G., AND L. VICEIRA (forthcoming): "Dynamic Consumption and Portfolio Choice with Stochastic Volatility in Incomplete Markets," *Review of Financial Studies*.

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

CONSTANTINIDES, G. M. (1990): "Habit Formation: A Resolution of the Equity Premium Puzzle," *Journal of Political Economy*, 98, 519–543.

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

CUOCO, D. (1997): "Optimal Consumption and Equilibrium Prices with Portfolio Constraints and Stochastic Income," *Journal of Economic Theory*, 72, 33–73.

CUOCO, D., AND J. CVITANIĆ (1998): "Optimal Consumption Choices for a Large Investor," *Journal of Economic Dynamics and Control*, 22, 401–436.

CUOCO, D., AND H. LIU (2000): "Optimal Consumption of a Divisible Durable Good," *Journal of Economic Dynamics and Control*, 24, 561–613.

CVITANIĆ, J., AND I. KARATZAS (1992): "Convex Duality in Constrained Portfolio Optimization," *The Annals of Applied Probability*, 2, 767–818.

CVITANIĆ, J., W. SCHACHERMAYER, AND H. WANG (2001): "Utility Maximization in Incomplete Markets with Random Endowment," *Finance and Stochastics*, 5, 259–272.

DANIEL, K., D. HIRSHLEIFER, AND S. H. TEOH (2002): "Investor Psychology in Capital Markets: Evidence and Policy Implications," *Journal of Monetary Economics*, 49, 139–209.

DAVIS, M., AND A. R. NORMAN (1990): "Portfolio Selection with Transaction Costs," *Mathematics of Operations Research*, 15, 676–713.

DETEMPLE, J., AND A. SERRAT (2003): "Dynamic Equilibrium with Liquidity Constraints," *Review of Financial Studies*, 16, 597–629.

DETEMPLE, J., AND F. ZAPATERO (1991): "Asset Prices in an Exchange Economy with Habit Formation," *Econometrica*, 59, 1633–1657.

DUFFIE, D. (2001): *Dynamic Asset Pricing Theory*. Princeton University Press, Princeton, New Jersey, third edn.

DUFFIE, D. (forthcoming): "Credit Risk Modeling with Affine Processes," *Journal of Banking and Finance*.

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

DUFFIE, D., V. FLEMING, M. SONER, AND T. ZARIPHOPOULOU (1997): "Hedging in Incomplete Markets with HARA Utility," *Journal of Economic Dynamics and Control*, 21, 753–782.

DUFFIE, D., AND C. SKIADAS (1994): "Continuous-Time Security Pricing: A Utility Gradient Approach," *Journal of Mathematical Economics*, 23, 107–131.

DUFFIE, D., AND T. ZARIPHOPOULOU (1993): "Optimal Investment with Undiversifiable Income Risk," *Mathematical Finance*, 3, 135–148.

DYBVIG, P. H., AND H. LIU (2005): "Lifetime Consumption and Investment: Retirement and Constrained Borrowing," working paper, Olin School of Business, Washington University, St. Louis, MO.

EL KAROUI, N., AND M. JEANBLANC-PICQUÈ (1998): "Optimization of Consumption with Labor Income," *Finance and Stochastics*, 2, 409–440.

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

——— (2001): "A Dynamic Maximum Principle for the Optimization of Recursive Utilities Under Constraints," *Annals of Applied Probability*, 11, 664–693.

ELLSBERG, D. (1961): "Risk, Ambiguity, and the Savage Axioms," *Quarterly Journal of Economics*, 75, 643–669.

EPSTEIN, L. (1992): "Behavior under Risk: Recent Developments in Theory and Applications," in *Advances in Economic Theory*, edited by J.-J. Laffont, Cambridge University Press, Cambridge, U.K.

EPSTEIN, L., AND J. MIAO (2000): "A Two-Person Dynamic Equilibrium under Ambiguity," working paper, Dept. of Economics, Univ. of Rochester, Rochester, NY.

EPSTEIN, L., AND M. SCHNEIDER (2003): "Recursive Multiple Priors," *Journal of Economic Theory*, 113, 1–31.

Epstein, L., and S. Zin (1989): "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 57, 937–969.

Farhi, E., and S. Panageas (2005): "Saving and Investing for Early Retirement: A Theoretical Analysis," working paper, dept. of Economics, MIT, Cambridge, MA.

Fleming, W. H., and H. M. Soner (1993): *Controlled Markov Processes and Viscosity Solutions*. Springer, New York, NY.

Giovannini, A., and P. Weil (1989): "Risk Aversion and Intertemporal Substitution in the Capital Asset Pricing Model," NBER working paper No. 2824, Cambridge, MA.

Gollier, C. (2001): *The Economics of Risk and Time*. MIT Press, Cambridge, MA.

Grossman, S. J., and G. Laroque (1990): "Asset Pricing and Optimal Portfolio Choice in the Presence of Illiquid Durable Consumption Goods," *Econometrica*, 58, 25–51.

Hamadene, S. (1996): "Équations Différentielles Stochastiques Rétrogrades: Le Cas Localement Lipschitzien," *Ann. Inst. Henri Poincaré*, 32, 645–659.

Hansen, L., T. Sargent, G. Turmuhambetova, and N. Williams (2001): "Robustness and Uncertainty Aversion," working paper, Dept. of Economics, University of Chicago.

He, H., and H. Pagès (1993): "Labor Income, Borrowing Constraints and Equilibrium Asset Prices: A Duality Approach," *Economic Theory*, 3, 663–696.

He, H., and N. Pearson (1991): "Consumption and Portfolio Policies with Incomplete Markets and Short-Sale Constrains: The Infinite Dimensional Case," *Journal of Economic Theory*, 54, 259–304.

Hugonnier, J., and D. Kramkov (2002): "Optimal Investment with Random Endowments in Incomplete Markets," working paper, HEC Montreal and Carnegie Mellon University.

Karatzas, I., J. Lehoczky, and S. Shreve (1987): "Optimal Portfolio and Consumption Decisions for a 'Small Investor' on a Finite Horizon," *SIAM Journal of Control and Optimization*, 25, 1557–1586.

Karatzas, I., J. Lehoczky, S. Shreve, and G. Xu (1991): "Martingale and Duality Methods for Utility Maximization in an Incomplete Market," *SIAM Journal of Control and Optimization*, 29, 702–730.

Karatzas, I., and S. Shreve (1988): *Brownian Motion and Stochastic Calculus*. Springer Verlag, New York.

Karatzas, I., and S. Shreve (1998): *Methods of Mathematical Finance*. Springer, New York.

Kim, T., and E. Omberg (1996): "Dynamic Nonmyopic Portfolio Behavior," *Review of Financial Studies*, 9, 141–162.

Klibanoff, P., M. Marinacci, and S. Mukerji (2002): "A Smooth Model of Decision Making Under Ambiguity," working paper, MEDS, Kellogg School of Management, Northwestern University.

Kobylanski, M. (2000): "Backward Stochastic Differential Equations and Partial Differential Equations with Quadratic Growth," *The Annals of Probability*, 28, 558–602.

Koo, H. (1998): "Nontraded Assets in Incomplete Markets," *Mathematical Finance*, 8, 49–65.

Korn, R. (1997): *Optimal Portfolios*. World Scientific, River Edge, NJ.

Kramkov, D., and W. Schachermayer (1999): "The Asymptotic Elasticity of Utility Functions and Optimal Investment in Incomplete Markets," *Annals of Applied Probability*, 9, 904–950.

——— (2003): "Necessary and Sufficient Conditions in the Problem of Optimal Investment in Incomplete Markets," *Annals of Applied Probability*, 13, 1504–1516.

Kreps, D., and E. Porteus (1978): "Temporal Resolution of Uncertainty and Dynamic Choice Theory," *Econometrica*, 46, 185–200.

Lazrak, A., and M. C. Quenez (2003): "A Generalized Stochastic Differential Utility," *Mathematics of Operations Research*, 28, 154–180.

Lepeltier, J.-P., and J. San Martín (1997): "Backward Stochastic Differential Equations with Continuous Coefficient," *Statistics and Probability Letters*, 32, 425–430.

——— (1998): "Existence for BSDE with Superlinear-Quadratic Coefficient," *Stochastics and Stochastics Reports*, 63, 227–240.

——— (2002): "On the Existence or Non-Existence of Solutions for Certain Backward Stochastic Differential Equations," *Bernoulli*, 8, 123–137.

Liu, H. (2004): "Optimal Consumption and Investment with Transaction Costs and Multiple Risky Assets," *Journal of Finance*, 59, 289–338.

Liu, H., and M. Loewenstein (2002): "Optimal Portfolio Selection with Transaction Costs and Finite Horizons," *Review of Financial Studies*, 15, 805–835.

Liu, J. (2001): "Portfolio Selection in Stochastic Environments," working paper, UCLA.

Ma, J., P. Protter, and J. Yong (1994): "Solving Forward-Backward Stochastic Differential Equations Explicitly – A Four Step Scheme," *Probability Theory and Related Fields*, 98, 339–359.

Ma, J., and J. Yong (1999): *Forward-Backward Stochastic Differential Equations and Their Applications*. Springer-Verlag, Berlin Heidelberg, Germany.

Maenhout, P. (1999): "Robust Portfolio Rules and Asset Pricing," working paper, INSEAD.

Markowitz, H. (1952): "Portfolio Selection," *Journal of Finance*, 7, 77–91.

Merton, R. (1969): "Lifetime Portfolio Selection under Uncertainty: The Continuous Time Case," *Review of Economics and Statistics*, 51, 247–257.

——— (1971): "Optimum Consumption and Portfolio Rules in a Continuous-Time Model," *Journal of Economic Theory*, 3, 373–413, Erratum 6 (1973): 213–214.

——— (1990): *Continuous Time Finance*. Blackwell, Malden, MA.

Musiela, M., and T. Zariphopoulou (2004): "An Example of Indifference Prices under Exponential Preferences," *Finance and Stochastics*, 8, 229–239.

Narens, L. (1985): *Abstract Measurement Theory.* MIT Press, Cambridge, MA.

Obstfeld, M. (1994): "Risk-Taking, Global Diversification, and Growth," *American Economic Review*, 84, 1310–1329.

Oksendal, B., and A. Sulem (2002): "Optimal Consumption and Portfolio with Fixed and Proportional Transaction Costs," *SIAM Journal of Control and Optimization*, 40, 1765–1790.

Øksendal, B., and A. Sulem (2005): *Applied Stochastic Control of Jump Diffusions.* Springer, New York.

Pardoux, E., and S. Peng (1990): "Adapted Solution of a Backward Stochastic Differential Equation," *Systems and Control Letters*, 14, 55–61.

Piazzesi, M. (2005): "Affine Term Structure Models," in *Handbook of Financial Econometrics.* Elsevier North Holland, Amsterdam, New York.

Pratt, J. W. (1964): "Risk Aversion in the Small and in the Large," *Econometrica*, 32, 122–136.

Schroder, M., and C. Skiadas (1999): "Optimal Consumption and Portfolio Selection with Stochastic Differential Utility," *Journal of Economic Theory*, 89, 68–126.

——— (2002): "An Isomorphism Between Asset Pricing Models with and Without Linear Habit Formation," *Review of Financial Studies*, 15, 1189–1221.

——— (2003): "Optimal Lifetime Consumption-Portfolio Strategies under Trading Constraints and Generalized Recursive Preferences," *Stochastic Processes and Their Applications*, 108, 155–202.

——— (2005a): "Lifetime Consumption-Portfolio Choice under Trading Constraints and Nontradeable Income," *Stochastic Processes and their Applications*, 115, 1–30.

——— (2005b): "Optimality and State Pricing in Constrained Financial Markets with Continuous and Discontinuous Information," Working Paper 327, Dept. of Finance, Kellogg School of Management, Northwestern University.

Segal, U., and A. Spivak (1990): "First-Order versus Second-Order Risk Aversion," *Journal of Economic Theory*, 51, 111–125.

Sethi, S. (1997): *Optimal Consumption and Investment with Bankruptcy.* Kluwer Academic Publishers, Norwell, MA.

Shreve, S., and H. M. Soner (1994): "Optimal Investment and Consumption with Transaction Costs," *Annals of Applied Probability*, 4, 609–692.

Shreve, S., and G. Xu (1992): "A Duality Method for Optimal Consumption and Investment under Short-Selling Prohibition: I, General Market Coefficients; and II, Constant Market Coefficients," *Annals of Applied Probability*, 2, 87–112 and 314–328.

Skiadas, C. (1998): "Recursive Utility and Preferences for Information," *Economic Theory*, 12, 293–312.

——— (2003): "Robust Control and Recursive Utility," *Finance and Stochastics*, 7, 475–489.

Sundaresan, S. (1989): "Intertemporally Dependent Preferences and the Volatility of Consumption and Wealth," *Review of Financial Studies*, 2, 73–89.

Svensson, L. (1989): "Portfolio Choice with Non-Expected Utility in Continuous Time," *Economic Letters*, 30, 313–317.

Svensson, L., and I. Werner (1993): "Nontraded Assets in Incomplete Markets: Pricing and Portfolio Choice," *European Economic Review*, 37, 1149–1168.

Uppal, R., and T. Wang (2003): "Model Misspecification and Under-Diversification," *Journal of Finance*.

Vayanos, D. (1998): "Transaction Costs and Asset Prices: A Dynamic Equilibrium Model," *Review of Financial Studies*, 11, 1–58.

Vila, J.-L., and T. Zariphopoulou (1997): "Optimal Consumption and Portfolio Choice with Borrowing Constraints," *Journal of Economic Theory*, 77, 402–431.

von Neumann, J., and O. Morgenstern (1944): *Theory of Games and Economic Behavior*. Princeton University Press, Princeton, NJ.

Wachter, J. (2002): "Portfolio and Consumption Decisions under Mean-Reverting Returns: An Exact Solution in Complete Markets," *Journal of Financial and Quantitative Analysis*, 37, 63–91.

Wakker, P. P. (1989): *Additive Representations of Preferences*. Kluwer, Dordrecht, The Netherlands.

Yong, J., and X. Y. Zhou (1999): *Stochastic Controls: Hamiltonian Systems and HJB Equations*. Springer-Verlag, New York.

Zariphopoulou, T. (1994): "Consumption-Invenstment Models with Constraints," *SIAM Journal on Control and Optimization*, 32, 59–85.

Zariphopoulou, T., and C.-I. Tiu (2002): "Optimization in Incomplete Markets with Recursive Utility," working paper, University of Texas, Austin.