PORTFOLIOS AND RISK PREMIA FOR THE LONG RUN

PAOLO GUASONI AND SCOTT ROBERTSON

Boston University

This paper develops a method to derive optimal portfolios and risk premia explicitly in a general diffusion model, for an investor with power utility and a long horizon. The market has several risky assets and is potentially incomplete. Investment opportunities are driven by, and partially correlated with, state variables which follow an autonomous diffusion. The framework nests models of stochastic interest rates, return predictability, stochastic volatility and correlation risk.

In models with several assets and a single state variable, long-run portfolios and risk premia admit explicit formulas up the solution of an ordinary differential equation, which characterizes the principal eigenvalue of an elliptic operator. Multiple state variables lead to a partial differential equation, which is solvable for most models of interest.

For each value of the relative risk aversion parameter, the paper derives the long-run portfolio, its implied risk premia and pricing measure, and their performance on a finite horizon.

Two applications to cross-sectional models with predictability, stochastic volatility and stochastic interest rates conclude.

AMS 2000 SUBJECT CLASSIFICATIONS: Primary 91B28, 62P05; Secondary 93E20.

KEYWORDS: Long Run, Portfolio Choice, Derivatives Pricing, Incomplete Markets.

INTRODUCTION

Investment opportunities change over time. Predictability[1] and heteroskedasticity in asset prices suggest that risk premia, in addition to interest rates, are time-varying, and driven by economic fundamentals. Fluctuations in investment opportunities drive a qualitative wedge between short-term and long-term investors. The former are rationally myopic, because short-term performance depends only on current market conditions. By contrast, long-term investors face strategic risk-return propositions, where the future evolution of the economy matters.

Dynamic portfolio choice and its pricing implications are central in asset pricing. Adverse shifts in investment opportunities can significantly affect welfare, especially at long horizons, and long-term investors hedge against such shifts, thereby departing from the static two-fund paradigm. The pricing implication is that intertemporal hedging portfolios carry risk premia in addition to the market equity premium, and multifactor models arise. Thus, the pricing of unspanned payoffs – such as derivatives in incomplete markets – should account for their potential as intertemporal hedging instruments.

Dynamic portfolio theory is technically complex. Only relatively simple models lead to explicit solutions, and even then their complexity hinders economic intuition. A major difficulty is the intertwined dependence of optimal portfolios and risk premia on both the investor's horizon, and current investment opportunities. The picture simplifies considerably in the limit of a long horizon,

[1]This paper benefited from the helpful comments of seminar participants at Cornell University, Hitotsubashi University, University of Michigan, Princeton University, University of Texas at Austin, AMS Meeting in San Diego, the Oberwolfach Workshop on Stochastic Analysis in Finance, and the Sixth Seminar on Stochastic Analysis at Ascona. Partially supported by the National Science Foundation under grants DMS-0532390 and DMS-0807994

[1]For predictability in stock returns, see Fama and French (1988); Campbell and Shiller (1988). For predictability in bond returns, see Fama and Bliss (1987); Cochrane and Piazzesi (2005)

while preserving its economic significance. The *long run* limit is a natural nontrivial benchmark for dynamic asset pricing, and contrasts the *short term* limit where intertemporal hedging disappears, and two-fund separation holds.

This paper develops a method to derive optimal portfolios and risk premia explicitly in a general diffusion model, for an investor with power utility and in the limit of a long horizon. The market has several risky assets and is potentially incomplete. Investment opportunities are driven by, and partially correlated with, state variables which follow an autonomous diffusion. The framework nests models of stochastic interest rates, return predictability, stochastic volatility and correlation risk.

The case for long-run analysis rests on tractability and accuracy. When a single state variable drives investment opportunities, an ordinary differential equation (17) identifies long-run portfolios and risk premia. This equation is furthermore linear if the R^2 of the regression of state shocks on excess returns is constant[2]. The general multivariate setting involves a quasilinear partial differential equation (28), which admits explicit solutions in several models of interest. The most prominent example is the usual regression model, where the state variables follow a VAR(1), expected returns are affine functions of state variables, and returns homoskedastic.

Long-run portfolios and risk premia are much simpler than their finite-horizon counterparts. At the portfolio level, the long-run features simple expressions for intertemporal hedging demands and their dependence on market parameters and risk aversion. At the pricing level, it yields martingale measures to price derivatives in incomplete markets. Thus, long-run policies are consistent with a goal of long-run performance which depends on risk aversion, but disregards the exact length of the long horizon. This approach has a longstanding tradition among institutional investors, who routinely refer to "long-term capital appreciation" or "aggressive long-term growth" as their official objectives, leaving an open-ended horizon.

The accuracy of long-run analysis stems from the finite-horizon bounds (21), which estimate the welfare loss of long-run policies on any horizon. *Long-run optimality* holds (Definition 5) when long-run policies are approximately optimal over long horizons.

For a single state variable with constant R^2, all long-run quantities depend on the principal eigenvalue and eigenfunction of the differential equation (17). In the terminology of Hansen and Scheinkman (2006), the principal eigenfunction drives the intertemporal hedging demand and the risk premia, the principal eigenvalue is the utility growth rate, and the finite-horizon bounds are different L^p norms of the transient component of the long-run stochastic discount factor. A similar nonlinear structure arises in the multivariate case.

Long-run optimality holds under joint parameter restrictions on risk aversion and asset-state dynamics. Although these restrictions depend on the specific model, a common economic interpretation emerges from applications. Long-run optimality *does not* hold only at the intersection of three extreme situations: *i*) high covariation of risk premia with state variables, *ii*) a nearly complete market, and *iii*) high risk aversion. Otherwise long-run optimality holds. Section 4 shows a calibration to the parameters of Barberis (2000), in a model where the dividend yield forecasts an equity index. Here, long-run optimality holds for risk-aversion less than 13.4.

The rest of the paper is organized as follows. After a brief literature review in the next section, Section 2 describes the model in detail, introducing notation. Section 3 contains the main result: a general method to obtain long-run policies in closed form. To improve readability, this section is divided into two parts: the first one covers models with a single state variable and constant R^2, discussing various implications. The second part contains the general multivariate case, including

[2] This additional assumption holds in the models of Kim and Omberg (1996), Brennan and Xia (2002), Wachter (2002), Chacko and Viceira (2005) among others.

the important case of the multivariate linear model, and discusses the connection with the usual Stochastic Control approach and the theory of Large Deviations. Section 4 derives long-run portfolios and risk premia explicitly in two models with cross-sectional predictability. The second model also features heteroskedasticity and stochastic interest rates. Section 5 concludes. All proofs are in the Appendix.

1. LITERATURE REVIEW

This paper cuts across several strands of literature, the main ones being dynamic portfolio choice and derivatives pricing in incomplete markets. Merton (1973) first recognized the intertemporal hedging motive in dynamic portfolio choice when investment opportunities are stochastic, and derived hedging demand implicitly in terms of the value function. Empirical evidence on stochastic investment opportunities accumulated mainly in the last two decades, giving new impetus to dynamic portfolio theory. Kim and Omberg (1996), Brennan and Xia (2002) and Wachter (2002) find analytical solutions for linear models, while Liu (2007) studies a class of quadratic models, calculating explicit solutions in certain cases. Campbell and Viceira (1999; 2001; 2002) propose approximate solutions based on the log-linearization of first-order conditions and budget constraints. Das and Sundaram (2000), Kogan and Uppal (2001) and Ferretti and Trojani (2005) propose different expansions of the value function which lead to analytical approximations.

The computational challenges are evident from the large number of numerical methods: discretizing the state space, value functions can be calculated with numerical integration Balduzzi and Lynch (1999), simulation Barberis (2000) or nonparametric regressions Brandt (1999). Brandt, Goyal, Santa-Clara and Stroud (2005) propose a simulation approach combined with dynamic programming. Detemple, Garcia and Rindisbacher (2003) use Malliavin Calculus to obtain portfolios as expectations, and calculate them with Monte Carlo simulations.

On the pricing side, Davis (1997) proposes option pricing by marginal utility in incomplete markets. This approach is essentially equivalent to the use of the *minimax martingale measure* of He and Pearson (1991) or the *least favorable completion* of Karatzas, Lehoczky, Shreve and Xu (1991). In the case of power utility, this measure is called q-optimal measure by Hobson (2004) and Henderson (2005) in the context of stochastic volatility models. In the limit of high relative risk aversion, the q-optimal measure reduces to the minimal entropy martingale measure (Grandits and Rheinländer, 2002), the pricing measure employed by an agent with exponential utility (Frittelli, 2000). This paper makes this pricing methodology tractable, by providing explicit solutions in the long-run limit.

Another related topic are asymptotic objectives. Dumas and Luciano (1991) employed a similar criterion for portfolio optimization under transaction costs, while Grossman and Zhou (1993) and Cvitanic and Karatzas (1995) used the same approach to tackle drawdown constraints. A series of papers by Bielecki and Pliska (1999; 2000), Fleming and Sheu (2000; 2002), Kuroda and Nagai (2002) and Nagai and Peng (2002) explores the similar criterion of *risk-sensistive control* in the context of linear diffusions. The asymptotic criterion in this paper crucially departs previous ones by examining both the primal (investment) and the dual (pricing) sides. This perspective yields a crisp economic interpretation, and provides estimates on finite-horizon performance.

2. MODEL

2.1. *Market*

Consider a financial market with a risk-free asset S^0 and n risky assets $S = (S^1, \dots, S^n)$. Investment opportunities (i.e. interest rates, expected returns and covariances) depend on k state variables $Y = (Y^1, \dots, Y^k)$, which capture the effect of economic fundamentals:

$$\frac{dS^0_t}{S^0_t} = r(Y_t)dt \tag{1}$$

$$\frac{dS^i_t}{S^i_t} = r(Y_t)dt + dR^i_t \qquad 1 \le i \le n \tag{2}$$

Cumulative excess returns $R = (R^1, \dots, R^n)$ and state variables follow the diffusion:

$$dR^i_t = \mu_i(Y_t)dt + \sum_{j=1}^{n} \sigma_{ij}(Y_t)dZ^j_t \qquad 1 \le i \le n \tag{3}$$

$$dY^i_t = b_i(Y_t)dt + \sum_{j=1}^{k} a_{ij}(Y_t)dW^j_t \qquad 1 \le i \le k \tag{4}$$

$$d\langle Z^i, W^j\rangle_t = \rho_{ij}(Y_t)dt \qquad 1 \le i \le n, 1 \le j \le k \tag{5}$$

where $Z = (Z^1, \dots, Z^n)$ and $W = (W^1, \dots, W^k)$ are multivariate Brownian Motions. This model provides a flexible framework which nests most diffusion models in Finance, and allows for price predictability, stochastic volatility, and correlation risk.

The law of (R, Y) determines the drifts b, μ and the covariation matrices $\Sigma = \sigma\sigma' = d\langle R, R\rangle_t/dt$, $A = aa' = d\langle Y, Y\rangle_t/dt$, and $\Upsilon = \sigma\rho a' = d\langle R, Y\rangle_t/dt$, where the prime sign denotes matrix transposition. By contrast, the matrices σ, a, ρ are identified only up orthogonal transformations.

Let $E \subset \mathbb{R}^k$ be an open connected set, and denote by $C^m(E, \mathbb{R}^d)$ the class of $\mathbb{R}^d$-valued continuous functions on E with m continuous derivatives. It is assumed $b \in C^1(E, \mathbb{R}^k)$, $\mu \in C^1(E, \mathbb{R}^n)$, $A \in C^2(E, \mathbb{R}^{k\times k})$, $\Sigma \in C^2(E, \mathbb{R}^{n\times n})$, $\Upsilon \in C^2(E, \mathbb{R}^{n\times k})$ and that the matrices $A(y)$, $\Sigma(y)$ are invertible for all $y \in E$. Denote by $\Omega = C([0, \infty), \mathbb{R}^{n+k})$ the space of continuous paths from $[0, \infty)$ to $\mathbb{R}^{n+k}$, endowed with the Borel σ-algebra $\mathcal{F}$. The following assumption ensures that $\mu, b, \Sigma, A, \Upsilon$ identify the law of Y:

Assumption 1 *For all $y \in E$, there exists a unique probability P^y on $(\Omega, \mathcal{F})$ such that the coordinate process (R, Y) satisfies (3)-(5) with initial condition $(R_0, Y_0) = (0, y)$, and $P^y(Y_t \in E$ for all $t \ge 0) = 1$.*[3]

An investor trades in the market according to a portfolio strategy $(\pi^i_t)^{1\le i\le n}_{t\ge 0}$, representing the proportions of wealth in each risky asset. Since the investor observes the state variables Y and the asset prices S (or equivalently, excess returns R), the portfolio π is adapted to (the augmentation

[3]The coordinate process (R, Y) satisfies (3)-(5) in the usual sense of the martingale problem of Stroock and Varadhan (2006). Set $u \in C^2(\mathbb{R}^n \times E, \mathbb{R})$, and define the generator of the process (R, Y):

$$Lu = (\mu\ b)'\nabla u + \frac{1}{2}\,\mathrm{tr}\left(\begin{pmatrix} \Sigma & \Upsilon \\ \Upsilon' & A \end{pmatrix} D^2 u\right)$$

A solution to the *martingale problem* is a family of probability measures $(P^y)_{y\in E}$ such that, $P^y(R_0 = 0, Y_0 = y) = 1$, and:

$$f(R_t, Y_t) - \int_0^t (Lf)(R_u, Y_u)du \qquad \text{is a } P^y \text{ martingale for all } f \in C_0^\infty(E)$$

Under the assumptions of regular coefficients made here (and even in greater generality), Theorem 6.1.7 in Stroock and Varadhan (2006) proves the existence of a local solution to the martingale problem. Assumption 1 requires the stronger property of existence and uniqueness of a global solution. This property does not follow from the regularity of coefficients alone, but nevertheless it holds in most models of interest.

of) the filtration generated by (R, Y), and R-integrable. The corresponding wealth process $(X_t^\pi)_{t\geq 0}$ follows:

$$\frac{dX_t^\pi}{X_t^\pi} = r(Y_t)dt + \pi_t' dR_t \tag{6}$$

Note that a positive initial capital $X_0 \geq 0$ implies a positive wealth at all times, i.e. $X_t^\pi \geq 0$ a.s. for all $t \geq 0$, which rules out doubling strategies.

The market defined by (1)-(5) is in general incomplete, as state-variable shocks dY_t are hedgeable only in part by the k mimicking portfolios $\Sigma^{-1}\Upsilon$ (each portfolio corresponds to a column of the matrix). The covariance matrix $\Upsilon'\Sigma^{-1}\Upsilon$ gauges the degree of incompleteness of the market, highlighting two extremes: *complete markets* for $\Upsilon'\Sigma^{-1}\Upsilon = A$, and *fully incomplete markets* for $\Upsilon = 0$.

Under market completeness, excess returns perfectly span state shocks. By contrast, full incompleteness entails excess returns orthogonal to state shocks. These features are dual of each other: since spanning is perfect in a complete market, tradeable assets embed all risk premia, and pricing is trivial (i.e. independent of preferences). Viceversa, in a fully incomplete market state shocks are unhedgeable, therefore investing is trivial (i.e. two-fund separation holds). When a single state variable is present, the scalar $\rho'\rho = \Upsilon'\Sigma^{-1}\Upsilon/A$ represents the multiple correlation coefficient, or R^2, in a regression of state variable shocks dY_t on the vector of excess returns dR_t, and measures market completeness on a scale from 0 (fully incomplete) to 1 (complete).

The representation of pricing rules employs the related concepts of stochastic discount factors and martingale measures.

Definition 2 *A stochastic discount factor is a strictly positive process $(M_t)_{t\geq 0}$ such that MS is a martingale, i.e.*

$$E[M_t S_t^i | \mathcal{F}_s] = M_s S_s^i \qquad \text{for all } 0 \leq s \leq t, 0 \leq i \leq n$$

A martingale measure is a probability Q, such that $Q|_{\mathcal{F}_t}$ and $P|_{\mathcal{F}_t}$ are equivalent for all $t \in [0, \infty)$, and the discounted prices S^i/S^0 (or equivalently, the excess returns R^i) are Q-martingales for all $1 \leq i \leq n$.

Martingale measures and stochastic discount factors are in a one-to-one correspondence through the relation $\frac{dQ}{dP}\big|_{\mathcal{F}_t} = S_t^0 M_t$, although their distinction is important in the present context of stochastic interest rates.

Proposition 3 *For any martingale measure Q, there exists an adapted integrable process $(\eta_t)_{t\geq 0}$ such that:*

$$\left.\frac{dQ}{dP}\right|_{\mathcal{F}_t} = \mathcal{E}\left(-\int_0^\cdot (\mu'\Sigma^{-1} + \eta'\Upsilon'\Sigma^{-1})\sigma dZ + \int_0^\cdot \eta' a dW\right)_t \tag{7}$$

where $\mathcal{E}(X)_t = \exp(X_t - \frac{1}{2}\langle X \rangle_t)$ denotes the stochastic exponential.

The process η has the interpretation of a vector of risk premia. In an incomplete market, the description of a pricing rule requires the arbitrary choice of risk premia η for the state-variable shocks dY_t, since absence of arbitrage alone does not price all payoffs. The symbols M^η and Q^η respectively denote the stochastic discount factor and martingale measure corresponding to η.

2.2. *Preferences*

The investor maximizes expected utility from terminal wealth at a long horizon T. In such context of long horizons, turnpike theorems[4] suggest that the CRRA (Constant Relative Risk Aversion) class of power utility functions determines the properties of long-run solutions:

$$U(x) = \frac{x^p}{p} \qquad p < 1, p \neq 0 \tag{8}$$

Denoting by E^y the expectation with respect to P^y, the finite-horizon utility maximization problem is:

$$\max_{\pi} \frac{1}{p} E^y_P \left[(X^\pi_T)^p \right] \tag{9}$$

Since power utility is homothetic, i.e. $U(cx) = c^p U(x)$, it suffices to consider the case $X_0 = 1$ of unit initial wealth. Logarithmic utility, excluded to simplify notation, corresponds to the limit $p \to 0$. In the following analysis, using the relative risk aversion parameter $\gamma = 1 - p$ would generate very inconvenient notation. To preserve economic intuition, recall that $p = 1 - \gamma$ and $q = p/(p-1) = 1 - \frac{1}{\gamma}$. Aggressive investors (less risk averse than logarithm) have $p > 0$ and $q < 0$, while conservative investors (more risk averse than logarithm) have $p < 0$ and $q > 0$. Risk aversion increases as q increases and as p decreases.

The martingale approach[5] to utility maximization relies on the duality between payoffs and stochastic discount factors. For power utility, the duality bound is an immediate consequence of Hölder's inequality:

LEMMA 4 *If a payoff $X \geq 0$ and a stochastic discount factor $M \geq 0$ satisfy $E[XM] \leq 1$, then:*

$$\frac{1}{p} E\left[X^p\right] \leq \frac{1}{p} E\left[M^q\right]^{1-p} \tag{10}$$

and equality holds if and only if $E[XM] = 1$, and for some $\alpha > 0$:

$$X^{p-1} = \alpha M \tag{11}$$

Equation (4) bounds the utility of any payoff by a moment of any stochastic discount factor, and viceversa. Thus, this bound plays for power utility the same role as the Hansen and Jagannathan (1991) bound for mean-variance preferences (i.e. for $p = 2$). The first-order condition (11) is the usual alignment of marginal utilities with state-price densities.

Consider a finite horizon T. Lemma 4 implies that a pair (π^T, η^T) of a portfolio π^T and risk premia η^T such that $X = X^{\pi^T}_T$ and $M = M^{\eta^T}_T$ is optimal if it satisfies (11). Then, denoting by $u_T(y)$ the value function (i.e. the maximal expected utility), the following equalities hold:

$$\frac{1}{p} E^y_P \left[(X^{\pi^T}_T)^p \right] = u_T(y) = \frac{1}{p} E^y_P \left[(M^{\eta^T}_T)^q \right]^{1-p} \tag{12}$$

hence π^T is the optimal portfolio, and the stochastic discount factor M^{η^T} identifies the pricing rule which makes an investor indifferent between buying and selling a small amount of any payoff, including unhedgeable ones.

[4]For turnpike theorems, see Leland (1972); Hakansson (1974); Huberman and Ross (1983); Cox and Huang (1992); Huang and Zariphopoulou (1999); Dybvig, Rogers and Back (1999)

[5]For the martingale approach, see Pliska (1986); Karatzas et al. (1987); Cox and Huang (1989); He and Pearson (1991); Kramkov and Schachermayer (1999)

2.3. *Long-Run Optimality*

In the Markov model defined by (1)-(5), stochastic control arguments show that the pair (π^T, η^T) achieving optimality is of the form $\pi^T(T-t,y)$ and $\eta^T(T-t,y)$. In other words, optimal policies depend on both state variables and the residual horizon. This joint dependence is the major source of intractability in portfolio choice and derivatives pricing problems.

However, as the horizon increases optimal policies seem to converge rapidly (as reported by Brandt (1999), Barberis (2000) and Wachter (2002)) to functions of state variables alone. These long-run limits $\pi(y)$, $\eta(y)$ are nearly optimal at long horizons, but become less and less optimal as the horizon approaches. At any finite horizon T, the duality bound (10) implies the inequalities:

$$\frac{1}{p}E_P^y\left[(X_T^\pi)^p\right] \leq u_T(y) \leq \frac{1}{p}E_P^y\left[(M_T^\eta)^q\right]^{1-p} \tag{13}$$

which reflect the welfare loss from using long-run policies not only at long horizons ($t \ll T$), but at all times ($t \in [0,T]$). A tangible measure of this loss is the certainty equivalent loss l_T, defined as the increase in the risk-free rate required to recover the utility loss:

$$\frac{1}{p}E_P^y\left[(e^{l_T T}X_T^\pi)^p\right] = u_T(y) \tag{14}$$

Substituting (14) into (13) yields an upper bound on the certainty equivalent l_T:

$$l_T \leq \frac{1}{p}\left(\frac{1}{T}\log E_P^y\left[(M_T^\eta)^q\right]^{1-p} - \frac{1}{T}\log E_P^y\left[(X_T^\pi)^p\right]\right) \tag{15}$$

This argument motivates the definition of a pair (π,η) as long-run optimal, when its certainty equivalent loss vanishes for long horizons:

DEFINITION 5 *A pair* $(\pi,\eta) \in C^1(E,\mathbb{R}^n)\times C^1(E,\mathbb{R}^k)$ *is* Long-Run Optimal *if:*

$$\lim_{T\to\infty}\frac{1}{T}\log E_P^y\left[(X_T^\pi)^p\right] = \lim_{T\to\infty}\frac{1}{T}\log E_P^y\left[(M^\eta)^q\right]^{1-p} = \lambda \tag{16}$$

where the limit λ is finite and does not depend on $y \in E$.

3. LONG RUN ANALYSIS

The presentation of main results is divided into two parts. The first part covers models with several assets but a single state variable, and constant multiple correlation (R^2). This setting, which is prevalent in the literature, simplifies the analysis in two ways. The presence of a single state variable reduces the calculation of long-run policies to the solution of an ordinary differential equation, and the assumption of a constant R^2 makes this equation linear.

The second part treats the general setting of several state variables and state-dependent covariances, leading to a quasilinear PDE which admits explicit solutions in models of interest. One important class is that of linear diffusions, where state-variables follow a continuous-time version of a VAR(1) model, returns are homoskedastic and expected returns are affine in state variables. Then, the PDE reduces to a system of algebraic Riccati equations.

In each part, long-run analysis takes place in two steps. The property of long-run optimality crucially depends on the second step, but the finite-horizon bounds in the first step are even more important for applications. The first step (Theorems 7 and 10 below) computes the finite-horizon performance of the long-run optimal "candidate" (π,η), which depends on the solution of the differential equations (17) and (28) respectively. By the duality bound (10), the equalities in (19)-(20) and (30)-(31) are respectively a lower and an upper bound on the finite-horizon value function, and hold regardless of the optimality of the candidate (π, η). The second step (Theorems 8 and 11 below) establishes a sufficient condition for long-run optimality, requiring that the bounds found in the first step converge at long horizons.

These theorems are simple, since they only rely on the local properties of the processes (R, Y). In particular, they avoid the knowledge of the transition density of Y, which may be very complicated if known at all. Recall that, although their dependence on y is omitted to alleviate notation, $b, \mu, \Sigma, \Upsilon, A, \phi$ and v are functions of the state variable y.

Since the one state variable, constant R^2 case is a particular realization of the general case, Theorems 7 and 8 follow directly from Theorems 10 and 11 respectively under the power transformation $\phi = \exp(v/\delta)$, for δ as in Theorem 7. Therefore, only the proofs of Theorems 10 and 11 are given in the Appendix. See the discussion after the statement of Theorem 10 for details.

3.1. *One State, Constant R-squared*

ASSUMPTION 6 *There is a single state variable ($k = 1$) and $\rho'\rho = \Upsilon'\Sigma^{-1}\Upsilon/A$ is constant (i.e. does not depend on the state y).*

THEOREM 7 (Finite-Horizon Bounds) *Let Assumption 6 hold, and set $\delta = \frac{1}{1-q\rho'\rho}$. Assume that:*

i) $\phi \in C^2(E, \mathbb{R}_{++})$ and $\lambda \in \mathbb{R}$ solve the linear ordinary differential equation:

$$\frac{A}{2}\ddot{\phi} + \left(b - q\Upsilon'\Sigma^{-1}\mu\right)\dot{\phi} + \frac{1}{\delta}\left(pr - \frac{q}{2}\mu'\Sigma^{-1}\mu - \lambda\right)\phi = 0 \tag{17}$$

ii) both the original and the auxiliary models:

$$(P)\ \begin{cases} dR_t = \mu dt + \sigma dZ_t \\ dY_t = bdt + adW_t \end{cases} \qquad (\hat{P})\ \begin{cases} dR_t = \frac{1}{1-p}\left(\mu + \delta\Upsilon\frac{\dot{\phi}}{\phi}\right)dt + \sigma d\hat{Z}_t \\ dY_t = \left(b - q\Upsilon'\Sigma^{-1}\mu + A\frac{\dot{\phi}}{\phi}\right)dt + ad\hat{W}_t \end{cases}$$

satisfy Assumption 1 under equivalent probabilities P and $\hat{P}$.

Then, the portfolio π and the risk premia η defined by:

$$\pi = \frac{1}{1-p}\Sigma^{-1}\left(\mu + \delta\Upsilon\frac{\dot{\phi}}{\phi}\right), \qquad \eta = \delta\frac{\dot{\phi}}{\phi} \tag{18}$$

satisfy the equalities:

$$E_P^y\left[(X_T^\pi)^p\right] = e^{\lambda T}\phi(y)^\delta E_{\hat{P}}^y\left[\phi(Y_T)^{-\delta}\right] \tag{19}$$

$$E_P^y\left[(M_T^\eta)^q\right]^{1-p} = e^{\lambda T}\phi(y)^\delta E_{\hat{P}}^y\left[\phi(Y_T)^{-\frac{\delta}{1-p}}\right]^{1-p} \tag{20}$$

Equations (19) and (20) provide lower and upper bounds on finite-horizon expected utility. Indeed, the duality inequality (10) yields:

$$\begin{aligned}\frac{1}{p}e^{\lambda T}\phi(y)^\delta E_{\hat{P}}^y\left[\phi(Y_T)^{-\delta}\right] &= \frac{1}{p}E_P^y\left[(X_T^\pi)^p\right] \leq u_T(y) \\ &\leq \frac{1}{p}E_P^y\left[(M_T^\eta)^q\right]^{1-p} = \frac{1}{p}e^{\lambda T}\phi(y)^\delta E_{\hat{P}}^y\left[\phi(Y_T)^{-\frac{\delta}{1-p}}\right]^{1-p}\end{aligned}$$

Combining (19) and (20) with (15) gives the central quantitative implication, that is an upper bound on the certainty equivalent loss l_T defined in the previous section:

$$l_T \leq \frac{1}{p}\left(\frac{1}{T}\log E^y_{\hat{P}}\left[\phi(Y_T)^{-\frac{\delta}{1-p}}\right]^{1-p} - \frac{1}{T}\log E^y_{\hat{P}}\left[\phi(Y_T)^{-\delta}\right]\right) \tag{21}$$

Closed-form expressions for this bound are available if the transition kernel of Y under $\hat{P}$ is known. At worst, Monte Carlo simulation is fast and trivial, involving two different powers of the same random variable $\phi(Y_T)$.

Theorem 7 reduces the long-run optimality (Definition 5) of (π, η) to the condition that the right-hand side in (21) converges to zero. Theorem 8 provides a criterion that covers most applications.

THEOREM 8 (Long-Run Optimality) *If, in addition to the assumptions of Theorem 7:*
i) for some $\bar{t} > 0$, the random variables $(Y_t)_{t\geq\bar{t}}$ are $\hat{P}$-tight;
ii) $\sup_{y\in E} F(y) < \infty$, where $F \in C(E, \mathbb{R})$ is defined as:

$$F = \begin{cases} \left(pr - \lambda - \frac{q}{2}\mu'\Sigma^{-1}\mu + \frac{1}{2}\delta(\delta-1)\left(\frac{\dot{\phi}}{\phi}\right)^2 A\right)\phi^{-\delta} & p < 0 \\ \left(pr - \lambda - \frac{q}{2}\mu'\Sigma^{-1}\mu - \frac{q}{2}\delta^2(1-\rho'\rho)\left(\frac{\dot{\phi}}{\phi}\right)^2 A\right)\phi^{-\frac{\delta}{1-p}} & 0 < p < 1 \end{cases} \tag{22}$$

Then the pair (π, η) in (18) is long-run optimal (Definition 5).

Testing assumptions i) and ii) is straightforward: ii) requires that the function F is bounded on the state space $E \subset \mathbb{R}$. To establish assumption i), it suffices to find some uniformly bounded "moment" ψ, which grows to infinity on the boundary of E. In other words, the following characterization of tightness holds:

LEMMA 9 *The family of random variables $(Y_t)_{t\geq\bar{t}}$ is tight if and only if there exists an increasing sequence of compact sets $(K_n)_{n\geq 1}$ such that $\cup_{n\geq 1} K_n = E$, and $\psi \in C(E, \mathbb{R}_{++})$ such that $\psi \geq n$ on $E \setminus K_n$ and $\sup_{t\geq\bar{t}} E\left[\psi(Y_t)\right] < \infty$.*

Theorem 7 (and its multivariate version Theorem 10) does not specify any boundary conditions for the solutions of the differential equation (17). Thus, the resulting finite horizon bounds hold for any possible solution pair (ϕ, λ). However, long-run optimality excludes all those solutions for which the certainty equivalent loss does not converge to zero. This condition alone identifies a unique pair (ϕ, λ) (up to positive multiples of ϕ) in most applications. Establishing sufficiently general assumptions under which long-run optimality leads to a unique pair (ϕ, λ) is a much more delicate issue, and is not treated here. [6]

3.1.1. *Long-Run Decomposition*

The bounds (19) and (20) decompose expected utility and its dual into a common "long-run" component $e^{\lambda T}$, and two "transient" components, in a close analogy to Hansen and Scheinkman

[6] The classical result in this direction is the Kreĭn and Rutman (1948) theorem, which states that every positive compact operator on a Banach lattice has a positive dominant eigenfunction. The application of the Krein-Rutman theorem requires a compact extension of the operator in (17). When Y is ergodic in the dynamics (26), there is a self-adjoint extension of this operator on $L^2(m)$, (m denoting the invariant probability), and therefore it is certainly closed. Compactness typically involves a Sobolev-type imbedding, which requires two crucial assumptions: a bounded set E and a uniformly elliptic operator. Unfortunately, virtually all models require an unbounded E, and singular diffusions (such as the square root diffusion) violate even uniformly ellipticity.

(2006). For a multiplicative functional N of a Markov process Y, they propose the decomposition:

$$N_t = \exp(\rho t)\frac{\varphi(y)}{\varphi(Y_t)}\hat{N}_t \tag{23}$$

where ρ and φ are respectively the principal eigenvalue and eigenfunction of the infinitesimal generator of Y, and $\hat{N}_t$ is a martingale. The bounds (19) and (20) yield similar expressions for terminal utilities and their dual counterparts:

$$(X_T^\pi)^p = e^{\lambda T}\frac{\phi(y)^\delta}{\phi(Y_T)^\delta}\frac{d\hat{P}}{dP} \qquad (M_T^\eta)^q = e^{\lambda T}\frac{\phi(y)^{\frac{\delta}{1-p}}}{\phi(Y_T)^{\frac{\delta}{1-p}}}\frac{d\hat{P}}{dP}$$

These decompositions are precisely of the form in (23), with the minor difference that "transient" components are powers of the principal eigenfunction ϕ, as opposed to the eigenfunction itself. Note also that λ and ϕ are the principal eigenvalue and eigenfunction of the operator in (17), which is not the generator of Y under either P or $\hat{P}$.

The use of quotes also hints at some caveats. Indeed, the interpretation of $e^{\lambda T}$ as a long-run component hinges on the condition that the $\hat{P}$-expectation of "transient components" has a less than exponential growth, which means that long-run optimality holds. This is not always the case: the examples in section 4 show how parameter restrictions are necessary even in the most common models.

3.1.2. *The Myopic Probability*

The bounds (19) and (20) in Theorem 7 and assumption *i*) in Theorem 8 depend on the equivalent probability $\hat{P}$, which plays a pivotal role in long-run analysis. $\hat{P}$ is neither the physical probability P, nor a risk-neutral probability. Instead, its interpretation becomes clear from its price dynamics.

Compare the original model, with price dynamics under P and power utility x^p/p, to the auxiliary model under $\hat{P}$ with logarithmic utility. The long-run optimal portfolio in the two models coincide. The first one is simply in (18), while the second one follows from the usual formula $\pi = \Sigma^{-1}\hat{\mu}$, where $\hat{\mu} = \frac{1}{1-p}\left(\mu + \delta\Upsilon\frac{\dot{\phi}}{\phi}\right)$ are the expected returns under $\hat{P}$. Thus, a long-horizon, power-utility investor under the probability P behaves exactly as a myopic (i.e. logarithmic) investor under $\hat{P}$.

3.1.3. *Long-Run Pricing*

For each value of the risk-aversion parameter $1-p$, the risk premia η in (18) deliver a pricing rule for derivative contracts involving the partially unhedgeable state variable Y. The martingale measure Q^η corresponding to the risk premia η is a long-run version of the *minimax martingale measure* of He and Pearson (1991), also called q-optimal measure by Hobson (2004) (in the case $p > 1$) and Henderson (2005). Its formal dynamics is:

$$\begin{cases} dR_t = \sigma d\tilde{Z}_t \\ dY_t = \left(b - \Upsilon'\Sigma^{-1}\mu + \left(A - \Upsilon'\Sigma^{-1}\Upsilon\right)\delta\frac{\dot{\phi}}{\phi}\right)dt + a d\tilde{W}_t \end{cases} \tag{24}$$

for some Brownian Motions $\tilde{Z}$ and $\tilde{W}$. Since this dynamics is distinct from that under P and $\hat{P}$, in general it is necessary to check its well-posedness, in the form of Assumption 1.

Observe that the drift of Y under the q-optimal measure has three components. The first term b is the drift under the original measure P. The second term $\Upsilon'\Sigma^{-1}\mu$ is the risk-neutral adjustment due to the correlation between the returns and the state shocks. The last term $\left(A - \Upsilon'\Sigma^{-1}\Upsilon\right)\delta\frac{\dot{\phi}}{\phi}$ accounts for preferences, which enter the equation through δ and ϕ.

3.1.4. *Complete and Fully Incomplete as Duals*

The formulas in (18) for optimal long-run policies illustrate the symmetric aspects of complete markets, where $A = \Upsilon'\Sigma^{-1}\Upsilon$, and fully incomplete markets, where $\Upsilon = 0$. In a complete market, price dynamics determines all risk premia, and the unhedgeable span is null. Thus, the pricing problem is trivial, as the dynamics in (24) becomes independent of the preference parameter p:

$$\begin{cases} dR_t = \sigma d\tilde{Z}_t \\ dY_t = \left(b - \Upsilon'\Sigma^{-1}\mu\right) dt + a d\tilde{W}_t \end{cases} \tag{25}$$

However, the investment problem is nontrivial, as the intertemporal component perfectly hedges the shifts in investment opportunities.

Conversely, in a fully incomplete market investment opportunities evolve independently of asset prices, hence intertemporal hedging is inaccessible, and myopic portfolios are optimal. However, a latent hedging motive remains present, and is responsible for the nonzero risk premia of unspanned payoffs, which may serve as hedging instruments.

In both cases, it is market dynamics, and not preferences, which make either the investment or the pricing problem trivial. By contrast, logarithmic preferences or constant investment opportunities remove the intertemporal hedging motive entirely, making both portfolios and risk premia trivial.

3.1.5. *Long-Run Risk-Return Tradeoff*

If the diffusion:

$$dY_t = (b - q\Upsilon'\Sigma^{-1}\mu)dt + a dW_t \tag{26}$$

is ergodic with invariant probability m, then equation (17) admits the interpretation of Euler-Lagrange equation of the variational problem:

$$\sup_{\int \phi^2 dm = 1} \int \left(\left(pr - \frac{q}{2}\mu'\Sigma^{-1}\mu \right) \phi^2 - \frac{\delta}{2} A \left(\dot{\phi} \right)^2 \right) dm \tag{27}$$

This objective function captures the *long-run risk-return tradeoff* among the three key long-run quantities: the risk-free rate r, the maximum squared Sharpe ratio $\mu'\Sigma^{-1}\mu$, and the intertemporal hedging motive in the last term. The relative importance of the three quantities depends on the risk-aversion parameter $1 - p$.

High risk aversion ($p \downarrow -\infty, q \uparrow 1, \delta \uparrow 1/(1 - \rho'\rho)$) implies a large negative weight for the interest-rate term pr, diverting the mass $\phi^2 dm$ away from states with a low interest rates. A highly risk-averse investor holds a large position in the risk-free rate, and wants to hedge heavily against low-interest states. Thus, the investor is mainly concerned with the reinvestment risk entailed by stochastic interest rates, and invests in risky assets primarily to hedge this variation, not for their higher return.

Risk-neutrality ($p \uparrow 1, q \downarrow -\infty, \delta \downarrow 0$) essentially eliminates concerns about both the interest rate and intertemporal hedging, emphasizing only states with maximal Sharpe ratio. Note however that solutions degenerate as $p \uparrow 1$, which reflects the usual explosion in leverage in the risk-neutral limit.

Logarithmic utility ($p = q = 0$, $\delta = 1$) makes the objective (27) trivial, as the first two terms vanish, and the last one is negative. The maximizer is $\phi = 1$, which attains the maximum value zero, hence portfolios are myopic and risk premia are zero.

3.2. *Several State Variables*

The differential equation governing the general case of several state variables with state dependent hedge ratios presents two difficulties. First, a multidimensional state space leads to a partial differential equation. Second, the equation involves a nonlinear term in the gradient, and therefore it is quasilinear.

THEOREM 10 *Assume that:*
i) $v \in C^2(E,\mathbb{R})$ *and* $\lambda \in \mathbb{R}$ *solve the quasilinear PDE:*

$$pr - \frac{q}{2}\mu'\Sigma^{-1}\mu + \frac{1}{2}\nabla v'\left(A - q\Upsilon'\Sigma^{-1}\Upsilon\right)\nabla v + \nabla v'\left(b - q\Upsilon'\Sigma^{-1}\mu\right) + \frac{1}{2}\operatorname{tr}\left(AD^2 v\right) = \lambda \tag{28}$$

ii) both the original and the auxiliary model:

$$(P)\begin{cases} dR_t = \mu dt + \sigma dZ_t \\ dY_t = bdt + adW_t \end{cases} \quad (\hat{P})\begin{cases} dR_t = \frac{1}{1-p}\left(\mu + \Upsilon\nabla v\right)dt + \sigma d\hat{Z}_t \\ dY_t = \left(b - q\Upsilon'\Sigma^{-1}\mu + (A - q\Upsilon'\Sigma^{-1}\Upsilon)\nabla v\right)dt + ad\hat{W}_t \end{cases}$$

satisfy Assumption 1 under equivalent probabilities P *and* $\hat{P}$.
Then, the pair (π,η) *given by:*

$$\pi = \frac{1}{1-p}\Sigma^{-1}\left(\mu + \Upsilon\nabla v\right), \qquad \eta = \nabla v \tag{29}$$

satisfies the equalities:

$$E_P^y\left[(X_T^\pi)^p\right] = e^{\lambda T + v(y)} E_{\hat{P}}^y\left[e^{-v(Y_T)}\right] \tag{30}$$

$$E_P^y\left[(M_T^\eta)^q\right]^{1-p} = e^{\lambda T + v(y)} E_{\hat{P}}^y\left[e^{-\frac{1}{1-p}v(Y_T)}\right]^{1-p} \tag{31}$$

Theorem 10 directly implies its one-dimensional version 7. With a single state variable, (28) reduces to the nonlinear ordinary differential equation:

$$pr - \frac{q}{2}\mu'\Sigma^{-1}\mu + \frac{1}{2}\frac{A}{\delta}\dot{v}^2 + \left(b - q\Upsilon'\Sigma^{-1}\mu\right)\dot{v} + \frac{1}{2}A\ddot{v} = \lambda$$

where $\delta = 1/(1-q\rho'\rho)$. If δ is a constant, the nonlinear term disappears through the change of variable $\phi = \exp(v/\delta)$, which is formally equivalent to the power transformation employed by Zariphopoulou (2001). Indeed, substituting $\dot{v} = \delta\dot{\phi}/\phi$ and $\ddot{v} = \delta\ddot{\phi}/\phi - \delta(\dot{\phi}/\phi)^2$ yields the linear equation (17).

THEOREM 11 *If, in addition to the assumptions of Theorem 10:*
i) for some $\bar{t} > 0$, *the random variables* $(Y_t)_{t\geq\bar{t}}$ *are* $\hat{P}$*-tight;*
ii) $\sup_{y\in E} F(y) < +\infty$, *where* $F \in C(E,\mathbb{R})$ *is defined by:*

$$F = \begin{cases} \left(pr - \lambda - \frac{q}{2}\mu'\Sigma^{-1}\mu + \frac{q}{2}\nabla v'\Upsilon'\Sigma^{-1}\Upsilon\nabla v\right)e^{-v} & p < 0 \\ \left(pr - \lambda - \frac{q}{2}\mu'\Sigma^{-1}\mu - \frac{q}{2}\nabla v'\left(A - \Upsilon'\Sigma^{-1}\Upsilon\right)\nabla v\right)e^{-\frac{1}{1-p}v} & 0 < p < 1 \end{cases}$$

Then the pair (π,η) *in* (18) *is long-run optimal (Definition 5).*

3.2.1. *Linear Diffusions*

The most common multivariate model is the linear diffusion[7], where the covariance matrices Σ, Υ, A are constant, and the drifts r, μ, b are affine functions of the state variable.

$$\begin{cases} dR_t = (\mu_0 + \mu_1 Y_t)dt + \sigma dZ_t \\ dY_t = -bY_t dt + a dW_t \\ r(Y_t) = r_0 + r_1' Y_t \end{cases} \tag{32}$$

where $\mu_0 \in \mathbb{R}^n$, $\mu_1 \in \mathbb{R}^{n\times k}$, $\sigma \in \mathbb{R}^{n\times n}$ $r_0 \in \mathbb{R}$, $r_1 \in \mathbb{R}^k$, $b \in \mathbb{R}^{k\times k}$, and $a \in \mathbb{R}^{k\times k}$. Under this model, state variables follow a multivariate Ornstein-Uhlenbeck process, and their discrete-sample dynamics is a centered VAR(1). For a linear diffusion, (28) admits a quadratic solution:

THEOREM 12 *For the model* (32)*, the equation* (28) *admits a solution of the form* $v(y) = v_0' y - \frac{1}{2} y' v_1 y$*, where* $v_0 \in \mathbb{R}^k$ *and* $v_1 \in \mathbb{R}^{k\times k}$*,* v_1 *symmetric, satisfy the algebraic equations:*

$$(v_1(A - q\Upsilon'\Sigma^{-1}\Upsilon) + (b + q\Upsilon'\Sigma^{-1}\mu_1)')v_0 - pr_1 + q(\mu_1' - v_1\Upsilon')\Sigma^{-1}\mu_0 = 0 \tag{33}$$

$$v_1(A - q\Upsilon'\Sigma^{-1}\Upsilon)v_1 + v_1(b + q\Upsilon'\Sigma^{-1}\mu_1) + (b + q\Upsilon'\Sigma^{-1}\mu_1)'v_1 - q\mu_1'\Sigma^{-1}\mu_1 = 0 \tag{34}$$

The corresponding utility growth rate is equal to:

$$\lambda = pr_0 - \frac{q}{2}\mu_0'\Sigma^{-1}\mu_0 + \frac{1}{2}v_0'(A - q\Upsilon'\Sigma^{-1}\Upsilon)v_0 - qv_0'\Upsilon'\Sigma^{-1}\mu_0 - \frac{1}{2}\operatorname{tr}(Av_1) \tag{35}$$

Furthermore, for $p < 0$ *and* $b + b'$ *strictly positive definite there exists a unique pair* v_0, v_1*, such that* $(Y_t)_{t\geq 0}$ *is* $\hat{P}$*-tight.*

Equation (34) is a quadratic equation in the unknown matrix v_1, and it belongs to the class of matrix Riccati equations, which arise in filtering theory and dynamical systems. It does not admit a closed-form solution in terms of matrix operations, but under the assumption $p < 0$, $b + b'$ strictly positive definite, the existence of solutions is guaranteed by Abou-Kandil, Freiling, Ionescu and Jank (2003, Lemma 2.4.1) and numerical techniques are widely available. Once the matrix v_1 is known, the linear equation (33) yields a unique solution for v_0, and the utility growth rate λ is quadratic in v_0 and linear in v_1.

Observe that Theorem 12 characterizes the candidate pair (π, η) and allows to find the finite-horizon bounds, but does not state its long-run optimality. This stronger property in fact holds only under parameter restrictions, as shown by the examples in section 4.

3.3. *Connections with Stochastic Control and Large Deviations*

3.3.1. *Stochastic Control*

Since the work of Merton (1969), most of the dynamic portfolio choice literature has employed stochastic optimal control as its main analytical tool. The relation between this paper and the stochastic control approach becomes clear by comparing equations (17) and (28) to the Hamilton-Jacobi-Bellman (HJB) equations of the utility maximization problem (9). Its value function $u(x, y, t)$ depends on the current wealth x, the current state y, and time t, and the homogeneity of power

[7] Linear diffusion models appear, for example, in Kim and Omberg (1996), Brennan, Schwartz and Lagnado (1997), Wachter (2002), Brennan and Xia (2002), Sangvinatsos and Wachter (2005).

utility entails that $u(x,y,t) = x^p e^{w(y,t)}/p$, thereby removing wealth from the reduced value function w. The corresponding HJB equation becomes:

$$-\frac{\partial w}{\partial t} = pr - \frac{q}{2}\mu'\Sigma^{-1}\mu + \frac{1}{2}\nabla w'\left(A - q\Upsilon'\Sigma^{-1}\Upsilon\right)\nabla w + \nabla w'\left(b - q\Upsilon'\Sigma^{-1}\mu\right) + \frac{1}{2}\operatorname{tr}\left(AD^2 w\right)$$

with the terminal condition $w(y,T) = 0$. Instead, the main PDE (28) is:

$$\lambda = pr - \frac{q}{2}\mu'\Sigma^{-1}\mu + \frac{1}{2}\nabla v'\left(A - q\Upsilon'\Sigma^{-1}\Upsilon\right)\nabla v + \nabla v'\left(b - q\Upsilon'\Sigma^{-1}\mu\right) + \frac{1}{2}\operatorname{tr}\left(AD^2 v\right)$$

In the former equation, the unknown function w depends on both time t and the state y, while v in the latter equation only depends on the state, although the constant λ is also unknown. Indeed, the former equation reduces to the latter under the parametric restriction:

$$w(t,y) = \lambda(T-t) + v(y)$$

This restriction improves on analytical tractability by reducing the dimension of the problem. The price of the tractability gain is that solutions of the time-homogeneous equation in general do not satisfy the boundary condition, and therefore are not exactly optimal at any time-horizon (except in the trivial case $v = 0$, arising with logarithmic utility or constant investment opportunities).

In this light, this paper provides the basic analytical tools to assess the approximate optimality of homogeneous solutions. Theorem 11 establishes the qualitative aspect of long-run optimality. At a quantitative level, Theorem 10 evaluates the accuracy of homogeneous solutions at any time horizon. From a theoretical standpoint, Theorem 11 plays a similar role to that of a verification theorem, which employs the estimates of Theorem 10.

A special case of equation (28) appears in the risk-sensitive control approach to optimal investment, initiated by Bielecki and Pliska (1999). In a linear diffusion model, they study the problem:

$$\max_{\pi} \liminf_{T\to\infty} \frac{1}{T} \log E[(X^\pi)^p] \tag{36}$$

where the supremum is taken over all progressively measurable strategies. Risk-sensitive control relies on control techniques to establish the existence and uniqueness to the homogeneous equation, then attempts to establish its optimality in the sense of (36). Fleming and Sheu (2000; 2002) carry out this program under the assumption that $|p|$ is small, that is if risk aversion is close enough to the logarithmic case. Long-run analysis, developed in this paper for general nonlinear models, sheds new light on this literature by characterizing finite-horizon performance. For example, Proposition 15 below relaxes the restriction of $|p|$ small to a necessary and sufficient condition, and explains the economic intuition behind it.

3.3.2. *Large Deviations*

Theorems 10 and 11 are closely related to the results of Donsker and Varadhan (1975; 1976; 1983) on Large Deviations of occupation times for continuous-time diffusions. Consider an ergodic Feller diffusion Y with generator L on some domain $\mathcal{D} \supset C^2(E,\mathbb{R})$, and denote by $M_1(E)$ the space of Borel probability measures on E. For a function $V \in C(E,\mathbb{R})$, and under certain joint conditions on Y and V, Donsker and Varadhan show that:

$$\lim_{T\to\infty} \frac{1}{T} \log E\left[\exp\left(\int_0^T V(Y_t)dt\right)\right] = \sup_{\mu\in M_1(E)}\left(\int_E V d\mu - I(\mu)\right) \tag{37}$$

For a one-dimensional diffusion with generator $Lu = \frac{1}{2}a^2\ddot{u} + b\dot{u}$ and invariant density $m(y)$ the *rate function* $I : M_1(E) \mapsto \mathbb{R}$ reduces to:

$$I(\mu) = \begin{cases} \frac{1}{2}\int_E a^2 \left(\dot{\psi}\right)^2 mdy & \text{if } \frac{d\mu}{dm} = \psi^2 \\ \infty & \text{otherwise} \end{cases}$$

Then, the following heuristic argument shows the relation between the Donsker and Varadhan (1975) theory and long-run optimality. Consider the terminal utility of a portfolio:

$$(X_T^\pi)^p = \exp\left(\int_0^T \left(pr + p\pi'\mu + \frac{1}{2}p(p-1)\pi'\Sigma\pi\right)dt\right)\mathcal{E}\left(\int_0^\cdot p\pi'\sigma dZ_t\right)_T$$

Now define P_π by setting $\frac{dP_\pi}{dP}$ equal to the stochastic exponential in the last term of this equation. It follows that:

$$E_P\left[(X_T^\pi)^p\right] = E_{P_\pi}\left[\exp\left(\int_0^T \left(pr + p\pi'\mu + \frac{1}{2}p(p-1)\pi'\Sigma\pi\right)dt\right)\right]$$

Thus, the Donsker-Varadhan asymptotics (37) yield:

$$\lim_{T\to\infty}\frac{1}{T}\log E_P\left[(X_T^\pi)^p\right] = \sup_{\int_E \psi^2 m_\pi = 1}\int_E \left(pr + p\pi'\mu + \frac{1}{2}p(p-1)\pi'\Sigma\pi - \frac{A}{2}\left(\frac{\dot{\psi}}{\psi}\right)^2\right)\psi^2 m_\pi \tag{38}$$

where m_π is the invariant density of Y under the probability P_π. To make the dependence on the portfolio π explicit, consider the change of variable $\psi^2 m_\pi = \phi^2 m$, where m is the invariant probability for Y evolving according to (26). This yields:

$$\frac{\dot{\psi}}{\psi} = \frac{\dot{\phi}}{\phi} + \frac{\dot{m}}{2m} - \frac{\dot{m}^\pi}{2m^\pi} = \frac{\dot{\phi}}{\phi} - \frac{qa\rho'\nu + pa\rho'\sigma'\pi}{a^2} \tag{39}$$

where $\nu = \sigma^{-1}\mu$. Substituting (39) into (38), the utility growth rate becomes the maximum over ϕ on the unit disc in $L^2(m)$ of the expression:

$$\int_E \phi^2 m\left(pr - \frac{1}{2}\left(a\frac{\dot{\phi}}{\phi} - q\rho'\nu\right)^2 + p\pi'\sigma\left(\left(1 - q\rho\rho'\right)\nu + a\frac{\dot{\phi}}{\phi}\rho\right) + \frac{1}{2}p(p-1)\pi'\sigma\left(1 - q\rho\rho'\right)\sigma'\pi\right) \tag{40}$$

The integrand is a quadratic function of π, and achieves its optimum at:

$$\pi = \frac{1}{1-p}\Sigma^{-1}\left(\mu + \delta\Upsilon\frac{\dot{\phi}}{\phi}\right) \tag{41}$$

Thus, subsituting (41) into (40), the utility growth rate reduces to:

$$\sup_{\int_E \phi^2 m = 1}\int_E \left(\left(pr - \frac{q}{2}\mu'\Sigma^{-1}\mu\right)\phi^2 - \frac{\delta}{2}A\left(\dot{\phi}\right)^2\right)mdy \tag{42}$$

The last expression is in fact (27), and its Euler equation is exactly the main differential equation (17). A similar reasoning on stochastic discount factors delivers the candidate long-run risk premia.

This argument, which explains the formal connection with Large Deviations, is suggestive but only heuristic. The main reason is that the Donsker-Varadhan asymptotics are correct under some delicate conditions, which may fail to hold even in the simplest models. The crucial observation is that, when the focus is on explicit formulas, a solution to (17) (or (28)) is available in the first place. Such a solution provides the finite-horizon bounds in Theorem 10, isolating the candidate exponential growth rate λ from the additional "transient" terms. Long-run optimality requires these terms to grow less than exponentially.

Further connections between Large Deviations and long-run investment problems exist in the literature. Pham (2003) proposes a Large Deviations criterion for the problem of outperforming a benchmark, and shows that it is essentially equivalent to a risk-sensitive control problem. Föllmer and Schachermayer (2007) give a definition of asymptotic arbitrage, and explore the relation with Large Deviations estimates of the maximal Sharpe ratio.

4. APPLICATIONS

This section brings to life the main results by deriving long-run portfolios and risk premia in two models with several risky assets and a single state variable. In both models expected returns are predictable. In addition, the second model features heteroskedasticity and stochastic interest rates. In both cases, the long-run optimal portfolios and risk premia have very simple expressions, in sharp contrast to the complexity of finite-horizon solutions, and in spite of the fact that the second model is not in the affine or quadratic class.

The parametric restrictions required by long-run optimality lead under both models to the same economic interpretation. Long-run optimality does not hold at the conjunction of three extreme situations: *i*) high covariation of risk premia with state shocks, *ii*) nearly complete markets, and *iii*) high risk-aversion. To understand this phenomenon, recall that long-run optimality essentially means that a time-homogenous strategy is approximately optimal on a sufficiently long time interval. Thus, the suboptimality of the long-run strategy in the latest part of the interval must lead to a small welfare loss. Since the myopic component of optimal portfolios is time-homogenous, only the intertemporal hedging is responsible for the welfare loss. All the above conditions concur to increase hedging demand. First, the covariation of risk premia is proportional to the leverage of the hedging portfolios $\Sigma^{-1}\Upsilon$. Second, intertemporal hedging is more attractive in a nearly complete market, where the tracking of state variables is almost exact. Third, intertemporal hedging is higher for more risk-averse investor, who reduce long-term risk at the expense of short-term return.

4.1. *Predictability*

The first model, a cross-sectional extension of the models in Kim and Omberg (1996) and Wachter (2002), is based on the linear diffusion (32) with several assets and a single state variable:

$$\begin{cases} dR_t = (\sigma\nu_0 + b\sigma\nu_1 Y_t)\,dt + \sigma dZ_t \\ dY_t = -\,bY_t dt + dW_t \\ d\langle R, Y\rangle_t = \rho dt \\ r(Y_t) = r_0 \end{cases}$$

where $\sigma \in \mathbb{R}^{n\times n}$; $\nu_0, \nu_1 \in \mathbb{R}^n$; $b, r_0 > 0$, and $\rho \in \mathbb{R}^n, 0 < \rho'\rho < 1$. In the main notation: $\mu(y) = \sigma\nu_0 + by\sigma\nu_1$, $\sigma(y) = \sigma$, $b(y) = -by$, $a(y) = 1$, and therefore $\Sigma = \sigma\sigma'$, $A = 1$, and $\Upsilon = \sigma\rho$. The general affine drift for R in (32) corresponds to $\nu_0 = \sigma^{-1}\mu_0$ and $\nu_1 = \frac{1}{b}\sigma^{-1}\mu_1$. The Riccati equation from (34) is

$$\delta^{-1}v_1^2 + 2b\left(1 + q\rho'\nu_1\right)v_1 - qb^2\nu_1'\nu_1 = 0$$

For $p < 0$, the solution v_1, v_0, λ from Theorem 12 is

$$v_1 = \delta b\left(\sqrt{\Theta} - \left(1 + q\rho'\nu_1\right)\right) \tag{43}$$

$$v_0 = q\delta\rho'\nu_0 - \frac{1}{\sqrt{\Theta}}\left(q\nu_1'\nu_0 + q\delta\rho'\nu_0\left(1 + q\rho'\nu_1\right)\right) \tag{44}$$

$$\lambda = pr_0 - \frac{1}{2}q\nu_0'\nu_0 + \frac{1}{2}\delta^{-1}v_0^2 - qv_0\rho'\nu_0 - \frac{1}{2}v_1 \tag{45}$$

where

$$\Theta = \left(1 + q\rho'\nu_1\right)^2 + \delta^{-1}q\nu_1'\nu_1 \tag{46}$$

The candidate long-run optimal pair (π, η), is affine in the state variable:

$$\begin{aligned} \pi(y) =& \frac{1}{1-p}\Sigma^{-1}\left(\mu(y) + v_0\sigma\rho - v_1 y\sigma\rho\right) \\ \eta(y) =& v_0 - v_1 y \end{aligned} \tag{47}$$

and the dynamics of (Y, R) under the candidate long-run martingale measure are:

$$\begin{cases} dR_t = \sigma dZ_t \\ dY_t = \left(-bY_t - \rho'\sigma^{-1}\mu + \left(1-\rho'\rho\right)\left(v_0 - v_1 Y_t\right)\right)dt + dW_t \end{cases} \tag{48}$$

This pair (π, η) is indeed long-run optimal, but only under a parameter restriction.

PROPOSITION 13 *Assume $p < 0$, (π, η) from (47) is long-run optimal if*

$$\left(1 - 2q\rho'\rho\right)\sqrt{\Theta} + \left(1 + q\rho'\nu_1\right) > 0 \tag{49}$$

In the case $\nu_1 = -\kappa\rho$ for $\kappa > 0$, which still nests the models of Kim and Omberg (1996) and Wachter (2002), the parameter restriction in (49) simplifies as follows:

COROLLARY 14 *Let $p < 0$ and $\nu_1 = -\kappa\rho$ for $\kappa \in \mathbb{R}$. If $0 < q\rho'\rho \leq 1/4$ then long-run optimality holds for all κ. For $1/4 < q\rho'\rho < 1$ long-run optimality holds if:*

$$\kappa < \frac{2}{4q\rho'\rho - 1} \tag{50}$$

Thus, long-run optimality requires a joint restriction on preferences (q) and price dynamics ($\rho'\rho$ and κ). First, since $q\rho'\rho < 1$, long-run optimality always holds if $\kappa < \frac{2}{3}$, that is if risk premia have low sensitivity with respect to state variables shocks. If this condition is not satisfied, long-run optimality still holds regardless of the level of incompleteness ($\rho'\rho$) if risk aversion is sufficiently low ($q < \frac{1}{4}$). Conversely, if the market is sufficiently incomplete ($\rho'\rho < \frac{1}{4}$), the restriction holds regardless of preferences. Hence, a violation of long-run optimality requires a high sensitivity of risk premia, high risk aversion, and a nearly complete market.

When long-run optimality fails, it does so at different scales, depending on parameters. Proposition 15 studies this phenomenon in the case $\kappa = 1$, which corresponds to a continuous time version of the model of Summers (1986):

PROPOSITION 15 *Assume $p < 0$ and $\kappa = 1$ from Corollary 14. Long-run optimality holds if $q\rho'\rho < \frac{3}{4}$. If $q\rho'\rho \geq \frac{3}{4}$, long-run optimality fails. In particular:*

i) if $q\rho'\rho > \frac{3}{4}$, there exists a finite T such that $\frac{1}{p}E[(X_T^\pi)^p] = -\infty$.

ii) if $q\rho'\rho = \frac{3}{4}$ and $\nu_0 = 0$, the certainty equivalent loss is bounded;

iii) if $q\rho'\rho = \frac{3}{4}$ and $\nu_0 \neq 0$, the certainty equivalent loss diverges to ∞.

4.1.1. *Calibration*

A calibration to a real data shows that long-run optimality holds for typical levels of risk aversion in the model with one asset and one state considered by Barberis (2000) and Wachter (2002), where the state variable represents the dividend yield, and the asset is an equity index. In the notation of this section, they use the set of parameter values (in monthly units) $\rho = -0.935$, $r = 0.14\%$, $\sigma = 4.36\%$, $\nu_0 = 0.0788$, $\kappa = 0.8944$, $b = 0.0226$. Then, condition (50) is satisfied for $p > -12.4$, that is for risk-aversion less than 13.4.

Figure 1 compares the finite-horizon performance of the long-run optimal portfolio to that of its myopic component. The plots show the estimates of the correponding upper bounds in (15): the myopic component prevails in the short term, but its performance progressively deteriorates as the horizon increases. The break-even horizon significantly increases with risk-aversion, passing from nine years for a risk-aversion of two, to twenty-three years for a risk-aversion of five. Also the magnitude of the certainty equivalent loss increases with risk aversion: the differences are within one percentage point for a risk-aversion of two, but increase to three precentage points for a risk-aversion of five.

Thus, in this calibration long-run optimal portfolios outperform their myopic components for long horizons, and long-run benefits increase as risk-aversion increases, but also the break-even horizon increases with risk-aversion. To understand this phenomenon, recall that long-run portfolios rebalance only in response to the state variable, ignoring that the horizon is approaching. The welfare cost of ignoring the horizon increases with risk-aversion, and even leads to the failure of long-run optimality for risk aversion greater than 13.4.

Note also that the conclusions in this section crucially depend on the numerical quantities involved. For example, the near perfect negative correlation between the dividend yield and the equity return il largely responsible for the slow convergence to zero the certainty equivalent of long-run portfolios, which in turns pushes the break-even horizon higher.

4.2. *Predictability and Heteroskedasticity*

The next model features a single state variable following the square-root diffusion of Feller (1951), which simultaneously affects the interest rate (Cox, Ingersoll and Ross, 1985), the volatilities of risky assets, and the Sharpe ratios. Note that the model is neither affine nor quadratic (due to the presence of the term with ν_0), and yet the long-run solution admits a simple expression.

$$(51)\qquad \begin{cases} dR_t = (\sigma\nu_0 + \sigma\nu_1 Y_t)\,dt + \sqrt{Y_t}\sigma dZ_t \\ dY_t = b\,(\theta - Y_t)\,dt + a\sqrt{Y_t}dW_t \\ d\langle R, Y\rangle_t = \rho dt \\ r(Y_t) = r_0 + r_1 Y_t \end{cases}$$

Here $\sigma \in \mathbb{R}^{n\times n}$; $\nu_0, \nu_1 \in \mathbb{R}^n$; $b, \theta, a, r_0, r_1 \in \mathbb{R}$ such that $b, \theta, a, r_1 > 0$; and $\rho \in \mathbb{R}^n$ such that $0 < \rho'\rho < 1$. In addition, the parametric restriction $c = b\theta - \frac{1}{2}a^2 > 0$ ensures that the state variable Y remains strictly positive (thereby satisfying Assumption 1 with $E = \mathbb{R}_{++}$). In the main notation: $\mu(y) = \sigma\,(\nu_0 + \nu_1 y)$, $\sigma(y) = \sqrt{y}\sigma$, $b(y) = b\,(\theta - y)$, $a(y) = a\sqrt{y}$, and therefore $\Sigma(y) = y\sigma\sigma'$, $A(y) = ay$, $\Upsilon(y) = y\sigma\rho a$. Guessing a form of the solution $v(y) = v_0 \log y + v_1 y$ the main ODE (28)

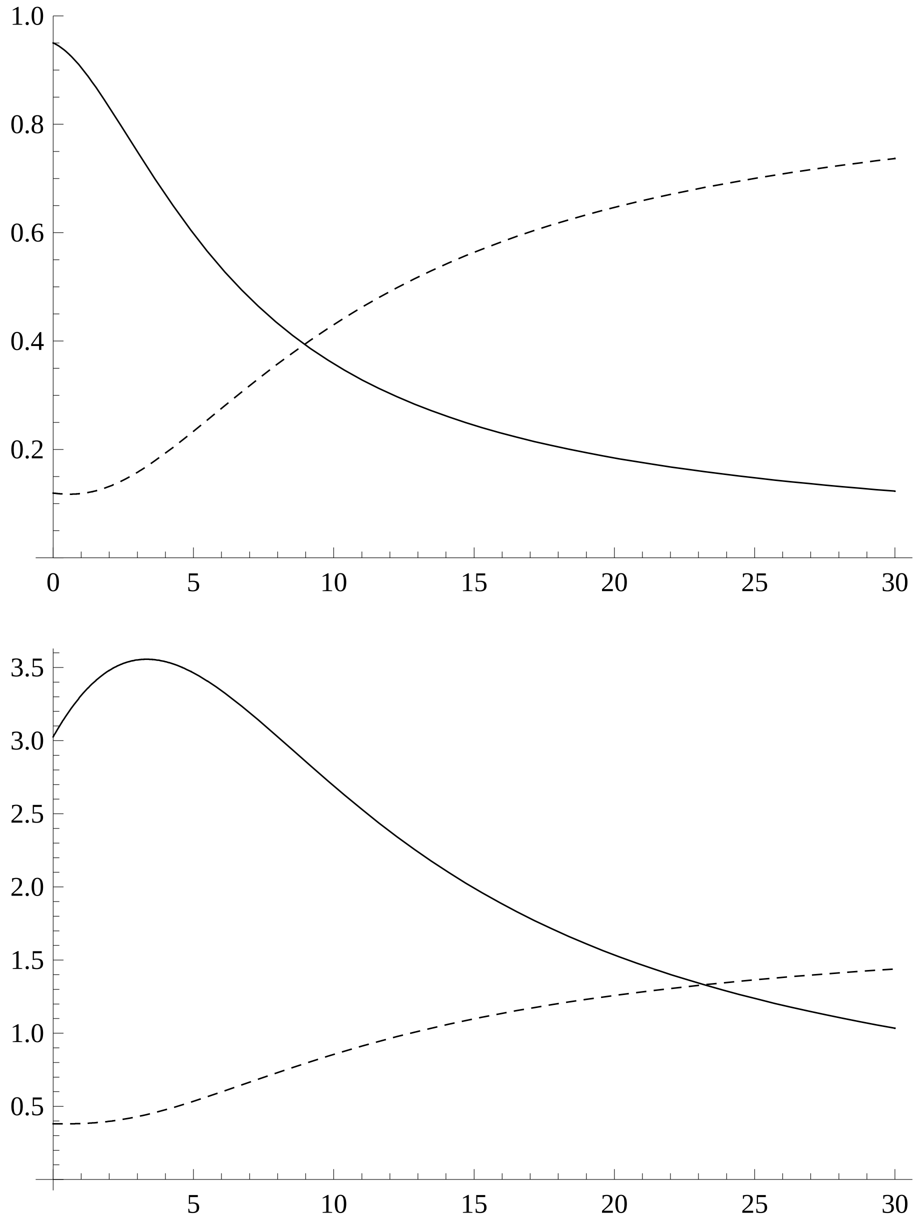

FIGURE 1.— The plots represent the annualized certainty equivalent loss bound (in percentage points) as a function of the horizon (in years) of the myopic component (dashed line) and long-run optimal (solid line) portfolios. Both plots are obtained by equation (15) setting η equal to the long-run optimal risk premium, and π equal to the lung-run optimal portfolio (solid line) and to the myopic portfolio $\frac{1}{1-p}\Sigma^{-1}\mu$ (dashed line). Risk Aversion is equal to two ($p=-1$, top) and to five ($p=-4$, bottom).

is:

$$pr_0 + pr_1 y - \frac{q}{2}\left(\nu_0' + y\nu_1'\right)\frac{1}{y}\left(\nu_0 + \nu_1 y\right) + \frac{1}{2}\left(\frac{v_0}{y} + v_1\right) a^2\delta^{-1} y \left(\frac{v_0}{y} + v_1\right) + \left(\frac{v_0}{y} + v_1\right)\left(b\theta - qa\rho'\nu_0 - \left(b + qa\rho'\nu_1\right) y\right) - \frac{1}{2}a^2\frac{v_0}{y} = \lambda$$

Setting the constant, linear and harmonic terms equal to zero leads to four candidate solutions, corresponding to any combination of signs in the terms $\pm\sqrt{\Theta}$ and $\pm\sqrt{\Lambda}$ below:

$$v_1 = \frac{\delta}{a^2}\left(b + qa\rho'\nu_1 \pm \sqrt{\Theta}\right)$$
$$v_0 = \frac{\delta}{a^2}\left(-\left(c - qa\rho'\nu_0\right) \pm \sqrt{\Lambda}\right)$$
$$\lambda = pr_0 - q\mu_0'\nu_1 + \frac{a^2}{\delta}v_0v_1 - v_0\left(b + qa\rho'\nu_1\right) + v_1\left(b\theta - qa\rho'\nu_0\right)$$

with

$$\Theta = \left(b + qa\rho'\nu_1\right)^2 + \frac{a^2}{\delta}\left(q\nu_1'\nu_1 - 2pr_1\right) \tag{52}$$
$$\Lambda = \left(c - qa\rho'\nu_0\right)^2 + \frac{a^2}{\delta}q\nu_0'\nu_0$$

When $p < 0, r_1 > 0$

$$\Theta > \left(b + qa\rho'\nu_1\right)^2 > 0 \tag{53}$$
$$\Lambda > \left(c - qa\rho'\nu_0\right)^2 > 0$$

Under $\hat{P}$, Y has again four possible dynamics:

$$dY_t = \left(\frac{1}{2}a^2 \pm \sqrt{\Lambda} \pm \sqrt{\Theta}Y_t\right)dt + a\sqrt{Y_t}dW_t$$

Now, since $(Y_t)_{t\geq 0}$ must be both strictly positive and tight, only the solution $-\sqrt{\Theta}$ and $\sqrt{\Lambda}$ is acceptable. Thus, the candidate optimizer is:

$$v_1 = \frac{\delta}{a^2}\left(b + qa\rho'\nu_1 - \sqrt{\Theta}\right)$$
$$v_0 = \frac{\delta}{a^2}\left(-\left(c - qa\rho'\nu_0\right) + \sqrt{\Lambda}\right) \tag{54}$$
$$\lambda = pr_0 - q\nu_0'\nu_1 + \frac{a^2}{\delta}v_0v_1 - v_0\left(b + qa\rho'\nu_1\right) + v_1\left(b\theta - qa\rho'\nu_0\right)$$

The candidate long-run optimal policies (π, η) are:

$$\pi(y) = \frac{1}{1-p}\Sigma^{-1}\left(\mu(y) + \sigma\rho a\left(v_0 + v_1 y\right)\right), \qquad \eta(y) = \frac{v_0}{y} + v_1$$

and the candidate long-run martingale measure is:

$$\begin{cases} dR_t = \sqrt{Y_t}\sigma dZ_t \\ dY_t = \left(b(\theta - Y_t) - qa(\rho'\nu_0 + \rho'\nu_1 Y_t) + a^2\left(1 - \rho'\rho\right)\left(v_0 + v_1 Y_t\right)dt + a\sqrt{Y_t}dW_t \end{cases}$$

Long-run optimality obtains under the following conditions:

PROPOSITION 16 *Assume that $p < 0$ and $r_1 > 0$. Then long-run optimality holds if:*

$$\left(1 - 2q\rho'\rho\right)\sqrt{\Lambda} + \left(c - qa\rho'\nu_0\right) > 0 \tag{55}$$
$$\left(1 - 2q\rho'\rho\right)\sqrt{\Theta} + \left(b + qa\rho'\nu_1\right) > 0$$

The main economic message of this parametric restriction is the same as in the previous example. Long-run optimality holds if either one of the following conditions is satisfied: the covariation of risk premia with state shocks is small ($a\rho'\nu_0, a\rho'\nu_1 \approx 0$), the market is sufficiently incomplete ($\rho'\rho \ll 1$) or risk aversion is low ($1 - p \ll \infty$).

5. CONCLUSION

The long-run limit yields a tractable, and yet nontrivial framework for dynamic portfolio choice and derivatives pricing in incomplete markets, leading to simple expressions for intertemporal hedging demand and unhedgeable risk premia. Long-run solutions are available in closed-form even in cases where their finite-horizon solutions are not, and finite-horizon performance of long-run solutions admits a characterization in terms of the long-run decomposition of Hansen and Scheinkman (2006).

Long-run optimality entails that the certainty equivalent loss of long-run policies vanishes for long horizons, and requires some joint restrictions on preferences and asset dynamics. It does not hold at the intersection of three extreme cases: risk premia highly covarying with state shocks, a nearly complete market, and high risk aversion. In all other cases, long-run optimality holds, and time-homogeneous portfolios are approximately optimal for long horizons.

Department of Mathematics and Statistics, 111 Cummington st, Boston MA 02215
`guasoni@bu.edu`; `scottrob@math.bu.edu`

APPENDIX A: PROOFS OF SECTION 2

Proof of Proposition 3: The Radon-Nykodim density process $D_t = \frac{dQ}{dP}|_{\mathcal{F}_t}$ is obviously a martingale, and the predictable representation property of the Brownian filtration (Revuz and Yor (1999, Theorem V.3.5)) implies the existence of two predictable processes H, K, such that $D_t = 1 + (H \cdot W)_t + (K \cdot B)_t$. Itô's formula further implies that $D = \mathcal{E}(h \cdot W)\mathcal{E}(k \cdot B)$, where $h = H/D$ and $k = K/D$ (Revuz and Yor (1999, Proposition VIII.1.6)). Again by Itô's formula, the condition that S^i/S^0 is a martingale for $1 \le i \le n$ implies that:

$$\rho h + \bar{\rho} k = -\sigma^{-1}\mu \quad \text{a.s. in } P \otimes dt \tag{56}$$

Recalling that $\bar{\rho}$ and A are invertible, define η by:

$$\eta_t = \left(A - \Upsilon'\Sigma^{-1}\Upsilon\right)^{-1}\left(a h_t + \Upsilon'\Sigma^{-1}\mu\right) \tag{57}$$

Then, (56) and (57) yield:

$$-\rho'\sigma'\left(\Sigma^{-1}\mu + \Sigma^{-1}\Upsilon\eta_t\right) + a'\eta_t = h_t \tag{58}$$

$$-\bar{\rho}'\sigma'\left(\Sigma^{-1}\mu + \Sigma^{-1}\Upsilon\eta_t\right) = k_t \tag{59}$$

and the claim (7) follows. To see the first identity in (58), left-multiply (57) by $a^{-1}\left(A - \Upsilon'\Sigma^{-1}\Upsilon\right)$. For for the second identity, observe that:

$$\begin{aligned}
-\bar{\rho}'\sigma'\left(\Sigma^{-1}\mu + \Sigma^{-1}\Upsilon\eta_t\right) - k_t &= (\bar{\rho})^{-1}\left(-\bar{\rho}\bar{\rho}'\sigma'\left(\Sigma^{-1}\mu + \Sigma^{-1}\Upsilon\eta_t\right) - \bar{\rho}k_t\right)\\
&= (\bar{\rho})^{-1}\left(-\bar{\rho}\bar{\rho}'\sigma'\left(\Sigma^{-1}\mu + \Sigma^{-1}\Upsilon\eta_t\right) + \rho h_t + \sigma^{-1}\mu\right)\\
&= (\bar{\rho})^{-1}\left(\rho\rho'\sigma^{-1}\mu - \rho a^{-1}\left(a h_t + a\rho'\sigma^{-1}\mu\right) + \rho h_t\right)\\
&= 0
\end{aligned}$$

where the first equality follows from (56), the second by (58), and the third by (57). *Q.E.D.*

Proof of Lemma 4: Denote by $q = \frac{p}{p-1}$. In the case $0 < p < 1$, Holder's inequality with $\tilde{p} = \frac{1}{p}$ and $\tilde{q} = \frac{\tilde{p}}{\tilde{p}-1} = \frac{1}{1-p}$ yields:

$$\begin{aligned}
E_P\left[X^p\right] &= E_P\left[(XM)^p M^{-p}\right] \le E_P\left[(XM)^{p\tilde{p}}\right]^{1/\tilde{p}} E_P\left[M^{-p\tilde{q}}\right]^{1/\tilde{q}}\\
&= E_P\left[XM\right]^{1/\tilde{p}} E_P\left[M^q\right]^{1-p} \le E_P\left[M^q\right]^{1-p}
\end{aligned}$$

because $E_P[XM] \le 1$, and the claim follows dividing by $p > 0$. If $p < 0$ ($0 < q < 1$) Holder's inequality with $\tilde{p} = \frac{1}{1-q}$, $\tilde{q} = \frac{1}{q}$ yields:

$$\begin{aligned}
E_P\left[M^q\right]^{1-p} &= E_P\left[(XM)^q X^{-q}\right]^{1-p} \le E_P\left[(XM)^{q\tilde{q}}\right]^{\frac{1-p}{\tilde{q}}} E_P\left[X^{-q\tilde{p}}\right]^{\frac{1-p}{\tilde{p}}}\\
&= E_P\left[XM\right]^{-p} E_P\left[X^p\right] \le E_P\left[X^p\right]
\end{aligned}$$

and the claim now follows dividing by $p < 0$. In both cases, the inequality becomes an equality when $E[XM] = 1$, and X^{p-1} is proportional to M. *Q.E.D.*

APPENDIX B: PROOFS OF SECTION 3

Proof of Theorem 10: Since the Brownian Motions Z and W are partially correlated, the following orthogonal decomposition holds:

$$dZ_t = \rho(Y_t)dW_t + \bar{\rho}(Y_t)dB_t \tag{60}$$

where $B = (B^1, \ldots, B^n)$ is a n-dimensional Brownian Motion independent of W, and the matrix $\bar{\rho}(y)$ is defined by the identity $(\rho\rho')(y) + (\bar{\rho}\bar{\rho}')(y) = I_n$.

The proofs of (30) and (31) are conceptually similar, although the algebra is different. Consider first (30).

$$\begin{aligned}(X_T^\pi)^p &= \exp\left(\int_0^T (r + \pi'\mu)\, dt\right)^p \mathcal{E}\left(\int_0^\cdot \pi'\sigma dZ_t\right)_T^p \\ &= \exp\left(\int_0^T \left(pr + p\pi'\mu + \frac{1}{2}p(p-1)\pi'\Sigma\pi\right) dt\right) \mathcal{E}\left(\int_0^\cdot p\pi'\sigma dZ_t\right)_T\end{aligned}$$

Substituting $\pi = \frac{1}{1-p}\Sigma^{-1}(\mu + \Upsilon\nabla v)$, and the decomposition $Z = \rho W + \bar{\rho}B$:

$$\begin{aligned}(X_T^\pi)^p = \exp&\left(\int_0^T \left(pr - \frac{q}{2}\mu'\Sigma^{-1}\mu + \frac{q}{2}\nabla v'\Upsilon'\Sigma^{-1}\Upsilon\nabla v\right) dt\right) \\ &\mathcal{E}\left(\int_0^\cdot -q\left(\Upsilon'\Sigma^{-1}\mu + \Upsilon'\Sigma^{-1}\Upsilon\nabla v\right)'(a')^{-1}dW_t - \int_0^\cdot q\left(\Sigma^{-1}\mu + \Sigma^{-1}\Upsilon\nabla v\right)'\sigma\bar{\rho}dB_t\right)_T\end{aligned}$$

From now on, denote by $H = \Upsilon'\Sigma^{-1}\Upsilon$. Dividing and multiplying by $\mathcal{E}\left(\int_0^\cdot \nabla v' a dW_t\right)_T$ yields:

$$\begin{aligned}(X_T^\pi)^p = \exp&\left(\int_0^T \left(pr - \frac{q}{2}\mu'\Sigma^{-1}\mu - q\nabla v'\Upsilon'\Sigma^{-1}\mu + \frac{1}{2}\nabla v'(A - qH)\nabla v\right) dt\right)\exp\left(-\int_0^T \nabla v' a dW_t\right) \\ &\mathcal{E}\left(\int_0^\cdot \left(-q\Upsilon'\Sigma^{-1}\mu + (A - qH)\nabla v\right)'(a')^{-1}dW_t - \int_0^\cdot q\left(\Sigma^{-1}\mu + \Sigma^{-1}\Upsilon\nabla v\right)'\sigma\bar{\rho}dB_t\right)_T\end{aligned} \tag{61}$$

Since P and $\hat{P}$ are equivalent, Girsanov's theorem implies that the above stochastic exponential coincides with $\frac{d\hat{P}}{dP}$. In addition, the P-Brownian Motion W and the $\hat{P}$-Brownian Motion $\hat{W}$ are related by:

$$\hat{W}_t = W_t - \int_0^t a^{-1}\left(-q\Upsilon'\Sigma^{-1}\mu + (A - qH)\nabla v\right) ds \tag{62}$$

Substituting (62) into (61) yields:

$$(X_T^\pi)^p = \frac{d\hat{P}}{dP}\exp\left(\int_0^T \left(pr - \frac{q}{2}\mu'\Sigma^{-1}\mu - \frac{1}{2}\nabla v'(A - qH)\nabla v\right) dt\right)\exp\left(-\int_0^T \nabla v' a d\hat{W}_t\right) \tag{63}$$

By Ito's rule, the following holds:

$$-\int_0^T \nabla v' a d\hat{W}_t = v(y) - v(Y_T) + \int_0^T \left(\nabla v'\left(b - q\Upsilon'\Sigma^{-1}\mu + (A - qH)\nabla v\right) + \frac{1}{2}\operatorname{tr}\left(AD^2 v\right)\right) dt \tag{64}$$

Substituting (64) into (63):

$$\begin{aligned}&(X_T^\pi)^p = \frac{d\hat{P}}{dP}\exp\left(v(y) - v(Y_T)\right) \\ &\exp\left(\int_0^T \left(pr - \frac{q}{2}\mu'\Sigma^{-1}\mu + \frac{1}{2}\nabla v'(A - qH)\nabla v + \nabla v'\left(b - q\Upsilon'\Sigma^{-1}\mu\right) + \frac{1}{2}\operatorname{tr}\left(AD^2 v\right)\right) dt\right)\end{aligned}$$

Since v solves (28) by assumption, the last term is equal to $e^{\lambda T}$, and taking the expectation yields (30). Consider now the equality (31).

$$(M_T^\eta)^q = \exp\left(-q\int_0^T r dt\right)\mathcal{E}\left(-\int_0^\cdot \left(\Sigma^{-1}\mu + \Sigma^{-1}\Upsilon\eta\right)'\sigma dZ_t + \int_0^\cdot \eta' a dW_t\right)_T^q$$

Again, the orthogonal decomposition $Z = \rho W + \bar{\rho}B$ and further simplifications yield:

$$\begin{aligned}(M_T^\eta)^q = \exp&\left(-q\int_0^T r dt\right) \\ &\mathcal{E}\left(\int_0^\cdot \left(-\Upsilon'\Sigma^{-1}\mu + (A - H)\eta\right)'(a')^{-1}dW_t - \int_0^\cdot \left(\Sigma^{-1}\mu + \Sigma^{-1}\Upsilon\eta\right)'\sigma\bar{\rho}dB_t\right)_T^q\end{aligned} \tag{65}$$

Then, use the equality $\mathcal{E}(X)^\alpha = \mathcal{E}(\alpha X)\exp\left(\frac{1}{2}\alpha(\alpha-1)\langle X\rangle\right)$, combined with the identity

$$\left\|a^{-1}\left(-\Upsilon'\Sigma^{-1}\mu + (A-H)\eta\right)\right\|^2 + \left\|\bar{\rho}'\sigma'\left(\Sigma^{-1}\mu + \Sigma^{-1}\Upsilon\eta\right)\right\|^2 = \mu'\Sigma^{-1}\mu + \eta'(A-H)\eta$$

to obtain:

$$\begin{aligned}(M_T^\eta)^q = &\exp\left((1-q)\int_0^T\left(pr - \frac{q}{2}\mu\Sigma^{-1}\mu - \frac{q}{2}\eta'(A-H)\eta\right)dt\right)\\ &\mathcal{E}\left(\int_0^\cdot q\left(-\Upsilon'\Sigma^{-1}\mu + (A-H)\eta\right)'(a')^{-1}dW_t - \int_0^\cdot q\left(\Sigma^{-1}\mu + \Sigma^{-1}\Upsilon\eta\right)'\sigma\bar{\rho}dB_t\right)_T\end{aligned}$$

Plugging $\eta = \nabla v$, and dividing and multiplying by $\mathcal{E}\left((1-q)\int_0^\cdot \nabla v' a dW_t\right)_T$ yields:

$$\begin{aligned}(M_T^\eta)^q = &\exp\left(-(1-q)\int_0^T \nabla v' a dW_t\right)\\ &\exp\left((1-q)\int_0^T\left(pr - \frac{q}{2}\mu'\Sigma^{-1}\mu + \frac{1}{2}\nabla v'(A-qH)\nabla v - q\nabla v'\Upsilon'\Sigma^{-1}\mu\right)dt\right)\\ &\qquad\mathcal{E}\left(\int_0^\cdot\left(-q\Upsilon'\Sigma^{-1}\mu + (A-qH)\nabla v\right)'(a')^{-1}dW_t - \int_0^\cdot q\left(\Sigma^{-1}\mu + \Sigma^{-1}\Upsilon\nabla v\right)'\sigma\bar{\rho}dB_t\right)\end{aligned}$$

At this point the argument mimics the one leading to (30): the stochastic exponential in the last line is equal to $\frac{d\hat{P}}{dP}$. Then, substituting (64) and (28) yields:

$$(M_T^\eta)^q = \frac{d\hat{P}}{dP}\exp\left(\frac{1}{1-p}(v(y) - v(Y_T))\right)e^{(1-q)\lambda T}$$

Taking the expectation, and raising it to the power $1-p$, (31) follows. *Q.E.D.*

The proof of Theorem 11 requires two lemmas:

LEMMA 17 *Let $\phi \in C(E;\mathbb{R}_{++})$, and $(\mu_T)_{T\geq 0}$ be a tight family of probability measures on $(E,\mathcal{B}(E))$. Then:*

(66) $$\liminf_{T\uparrow\infty}\frac{1}{T}\log\int\phi d\mu_T \geq 0$$

PROOF: By monotonicity, it is sufficient to prove that the limes infimum is zero for ϕ bounded. Let $(t_n)_{n\geq 1}$ be an increasing sequence satisfying $t_n \uparrow \infty$ and:

(67) $$\liminf_{T\uparrow\infty}\frac{1}{T}\log\int_E\phi d\mu_T = \lim_{n\to\infty}\frac{1}{t_n}\log\int_E\phi d\mu_{t_n}$$

Since the measures $(\mu_T)_{T\geq 0}$ are tight, they are relatively compact with respect the topology of weak convergence. Thus, up to a subsequence, there exists a probability measure μ on E such that:

$$\lim_{n\to\infty}\int_E\phi d\mu_{t_n} = \int_E\phi d\mu \in (0,\infty)$$

because ϕ is continuous, bounded and strictly positive. The thesis immediately follows:

$$\liminf_{T\uparrow\infty}\frac{1}{T}\log\int_E\phi d\mu_T = \lim_{n\to\infty}\frac{1}{t_n}\log\int_E\phi d\mu_{t_n} = \frac{\log\int_E\phi d\mu}{\lim_{n\to\infty}t_n} = 0$$

Q.E.D.

LEMMA 18 *Let P^y be a probability such that $Y_0 = y$ and*

$$dY_t = u(Y_t)dt + \sigma(Y_t)dW_t$$

and denote by $Lf = \nabla f'u + \frac{1}{2}\operatorname{tr}(\sigma\sigma' D^2 f)$. If $f \in C^2(E,\mathbb{R}_{++})$, then:

$$E_P^y\left[f(Y_T)\right] \leq f(y) + \left(0 \vee \sup_E Lf\right)T$$

PROOF: Let $f \in C^2_+(E)$. Itô's formula implies that:

$$f(Y_T) = f(y) + \int_0^T Lf(Y_t)dt + \int_0^T \nabla f' a dW_t$$

Let $(\tau_n)_{n\geq 1}$ be a reducing sequence of stopping times for the local martingale $\int_0^T \nabla f' a dW_t$. Then:

$$f(Y_{T\wedge\tau_n}) \leq f(y) + \left(0 \vee \sup_E (Lf)\right) T + \int_0^{T\wedge\tau_n} \nabla f' a dW_t$$

and therefore:

$$E^y_P\left[f\left(Y_{T\wedge\tau_n}\right)\right] \leq f(y) + \left(0 \vee \sup_E (Lf)\right) T$$

and the thesis follows by Fatou's lemma. *Q.E.D.*

PROOF OF THEOREM 11: It is clear that:

$$0 \leq \lim_{T\to\infty} \frac{1}{p}\left(\frac{1}{T}\log E^y_P\left[(M^\eta_T)^q\right]^{1-p} - \frac{1}{T}\log E^y_P\left[(X^\pi_T)^p\right]\right) \tag{68}$$
$$\leq \limsup_{T\to\infty}\frac{1}{pT}\log E^y_P\left[(M^\eta_T)^q\right]^{1-p} - \liminf_{T\to\infty}\frac{1}{pT}\log E^y_P\left[(X^\pi_T)^p\right]$$

Thus, it is sufficient to prove that both terms on the righ-hand side are zero. By equations (30) and (31), this conditions reduce to:

$$\limsup_{T\to\infty}\frac{1}{T}\log E^y_{\hat{P}}\left[\exp\left(-\frac{1}{1-p}v(Y_T)\right)\right] \leq 0 \tag{69}$$

$$\liminf_{T\to\infty}\frac{1}{T}\log E^y_{\hat{P}}\left[\exp\left(-v(Y_T)\right)\right] \geq 0 \tag{70}$$

for $p > 0$, and to:

$$\liminf_{T\to\infty}\frac{1}{T}\log E^y_{\hat{P}}\left[\exp\left(-\frac{1}{1-p}v(Y_T)\right)\right] \geq 0 \tag{71}$$

$$\limsup_{T\to\infty}\frac{1}{T}\log E^y_{\hat{P}}\left[\exp\left(-v(Y_T)\right)\right] \leq 0 \tag{72}$$

for $p < 0$.

The lower bounds (71) and (70) follow from the application of Lemma 17 to the functions $\phi = \exp\left(-\frac{1}{1-p}v\right)$ and $\phi = \exp\left(-v\right)$ respectively. For the upper bounds, first denote by:

$$Lf = \nabla f'\left(b - q\Upsilon'\Sigma^{-1}\mu + (A - qH)\nabla v\right) + \frac{1}{2}\operatorname{tr}\left(AD^2 f\right)$$

and observe that, for any $\alpha \in \mathbb{R}$:

$$L\left(e^{\alpha v}\right) = \alpha e^{\alpha v}\left(\nabla v'\left(b - q\Upsilon'\Sigma^{-1}\mu + (A - qH)\nabla v\right) + \frac{1}{2}\operatorname{tr}\left(AD^2 v\right) + \frac{1}{2}\alpha\nabla v' A\nabla v\right)$$
$$= \alpha e^{\alpha v}\left(\frac{1}{2}\nabla v'\left((1+\alpha)A - qH\right)\nabla v + \lambda - pr + \frac{q}{2}\mu'\Sigma^{-1}\mu\right)$$

where the second equality follows from (28). For $p < 0$, consider $\alpha = -1$:

$$L\left(e^{-v}\right) = e^{-v}\left(\frac{q}{2}\nabla v' H\nabla v - \lambda + pr - \frac{q}{2}\mu'\Sigma^{-1}\mu\right)$$

Assumption *ii)* implies that the right-hand side is bounded by some constant K, and Lemma 18 yields:

$$E^y_{\hat{P}}\left[e^{-v(Y_t)}\right] \leq e^{-v(y)} + (K \vee 0)\, T$$

and hence

$$\limsup_{T\to\infty}\frac{1}{T}\log E^y_{\hat{P}}\left[e^{-v(Y_t)}\right] \leq 0$$

Similarly, for $0 < p < 1$ consider $\alpha = -\frac{1}{1-p}$:

$$L\left(e^{-\frac{1}{1-p}v}\right) = \frac{1}{1-p}e^{-\frac{1}{1-p}v}\left(-\frac{q}{2}\nabla v'\left(A-H\right)\nabla v - \lambda + pr - \frac{q}{2}\mu'\Sigma^{-1}\mu\right)$$

Again, the right-hand side is bounded by K, and Lemma 18 yields:

$$E^y_{\hat{P}}\left[e^{-\frac{1}{1-p}v(Y_t)}\right] \leq e^{-\frac{1}{1-p}v(y)} + (K \vee 0)\,T$$

and the claim follows as in the previous case.

Q.E.D.

PROOF OF LEMMA 9: Suppose there exists a moment ψ satisfying the assumptions. Since ψ is non-negative and $\psi > n$ on $E \setminus K_n$, for any $t > \bar{t}$ Markov's inequality implies that:

$$P(Y_t \in E \setminus K_n) \leq \frac{1}{n}\sup_{t\geq\bar{t}} E\left[\psi(Y_t)\right]$$

and tightness follows from the compactness of K_n and the uniform bound on $E\left[\psi(Y_t)\right]$.

For the reverse implication, the tightness of $(Y_t)_{t\geq\bar{t}}$ implies that, there exists a sequence of compact sets $(A_n)_{n\geq1}$ such that

$$\sup_{t\geq\bar{t}} P(Y_t \in E \setminus A_n) \leq 2^{-n}$$

Now, consider another sequence of compact sets $(B_n)_{n\geq1}$ which satisfies $\cup_{n\geq1}B_{n\geq1} = E$ and $B_n \cap \overline{B_{n+1}} = \emptyset$. For example, set $B_n = \{x \in E : d(x, \partial E \geq) 1/n\}$. Then define the sequence of compact sets $(K_n)_{n\geq1}$ as $K_n = \cup_{B_i \not\supset A_n} B_i$. This sequence satisfies $\cup_n K_{n\geq1} = E$, $K_n \cap \overline{K_{n+1}} = \emptyset$ and $\sup_{t\geq\bar{t}} P(Y_t \in E \setminus K_n) \leq 2^{-n}$. By Urysohn's Lemma (Reed and Simon (1980, Theorem IV.7)) there exists continuous functions $\psi_n : E \mapsto [0,1]$ such that $\psi_n = 0$ on K_n and $\psi_n = 1$ on $E \setminus K_{n+1}$.

Now, set $\psi(x) = 1 + \sum_{n\geq1}\psi_n(x)$, and note that $\psi(x) = n + \psi_n(x)$ for $x \in K_{n+1} \cup (E \setminus K_n)$, hence ψ is continuous and $\psi \geq n$ on $E \setminus K_n$. Finally, for $t > \bar{t}$ the following holds:

$$\begin{aligned}E\left[\psi(Y_t)\right] &= E\left[\psi(Y_t)1_{K_1}(Y_t)\right] + \sum_{n=1}^{\infty} E\left[\psi(Y_t)1_{\{K_{n+1}\cup(E\setminus K_n)\}}(Y_t)\right] \\ &\leq P(Y_t \in K_1) + \sum_{n=2}^{\infty}(n+1)P(Y_t \notin K_n) \leq 1 + 2\sum_{n=1}^{\infty}(n+1)2^{-n} < \infty\end{aligned}$$

Q.E.D.

PROOF OF THEOREM 12: If there do exist $v_0 \in \mathbb{R}^k$ and $v_1 \in \mathbb{R}^{k\times k}$, v_1 symmetric satisfying (33) and (34) respectively then $v(y) = v_0'y - \frac{1}{2}y'v_1y$ will solve (28) with the constant λ from (35). Therefore it suffices to prove the existence and uniqueness of such a pair v_0, v_1 under the assumption of $p < 0$, $b + b'$ strictly positive definite. For a square matrix M write $M > 0$ if $M + M'$ is strictly positive definite and note that $M > 0$ if and only if each of its eigenvalues has strictly positive real part. The Riccati equation (34) admits the form

$$v_1\mathbf{B}\mathbf{B}'v_1 - v_1\mathbf{A} - \mathbf{A}'v_1 - \mathbf{C}'\mathbf{C} = 0$$

with

$$\mathbf{B} = \left(A - q\Upsilon'\Sigma^{-1}\Upsilon\right)^{1/2}; \mathbf{A} = -\left(b + q\Upsilon'\Sigma^{-1}\mu_1\right); \mathbf{C} = \sqrt{q}\sigma^{-1}\mu_1$$

where $\mathbf{B}$ is assumed to be the unique symmetric positive definite square root of $A - q\Upsilon'\sigma^{-1}\Upsilon$. $\mathbf{C}$ is a real valued matrix when $p < 0$. From Abou-Kandil, Freiling, Ionescu and Jank (2003, Lemma 2.4.1) if there exist two matrices $F_1 \in \mathbb{R}^{k\times k}$ and $F_2 \in \mathbb{R}^{n\times k}$ such that $\mathbf{A} - \mathbf{B}F_1 < 0$ and $\mathbf{A}' - \mathbf{C}'F_2 < 0$ then there exists a unique solution v_1 such that $\mathbf{A} - \mathbf{B}\mathbf{B}'v_1 < 0$. Since $b > 0$ and $p < 0$, the choice of $F_1 = -q\mathbf{B}^{-1}\Upsilon'\Sigma^{-1}\mu_1$ and $F_2 = -\sqrt{q}\rho a'$ suffices. The condition $\mathbf{A} - \mathbf{B}\mathbf{B}'v_1 < 0$ yields

$$-\left(\left(b + q\Upsilon'\Sigma^{-1}\mu_1\right) + \left(A - q\Upsilon'\Sigma^{-1}\Upsilon\right)v_1\right) < 0 \tag{73}$$

Therefore,

$$\left(v_1\left(A - q\Upsilon'\Sigma^{-1}\Upsilon\right) + \left(b + q\Upsilon'\Sigma^{-1}\mu_1\right)'\right) = \left(\left(b + q\Upsilon'\Sigma^{-1}\mu_1\right) + \left(A - q\Upsilon'\Sigma^{-1}\Upsilon\right)v_1\right)'$$

is invertible and v_0 from (33) is well defined. Furthermore, under $\hat{P}$, Y has the dynamics

$$dY_t = \left(-\left(\left(b + q\Upsilon'\Sigma^{-1}\mu_1\right) + \left(A - q\Upsilon'\Sigma^{-1}\Upsilon\right)v_1\right)Y_t - q\Upsilon'\Sigma^{-1}\mu_0 + \left(A - q\Upsilon'\Sigma^{-1}\Upsilon\right)v_0\right)dt + adW_t$$

For general k dimensional OU processes:

$$dZ_t = C(b - Z_t)dt + DdW_t; Z_0 = z$$

if $C > 0$ with $\lambda^* = \min_{i=1,\dots,k} \mathrm{Re}(\lambda_i) > 0$ where the minimum is taken over the eigenvalues $\{\lambda_i\}_{i=1,..,k}$ then

$$E\left[|Z_t|^2\right] \leq 2|b|^2 + 2k^2\|D\|^2_{H.S}\left(\frac{k}{\lambda^*} + k^2|z|^2\right)$$

and hence by (73) and Lemma 9 with $\psi(y) = |y|^2$, $(Y_t)_{t\geq 0}$ is tight under $\hat{P}$. *Q.E.D.*

APPENDIX C: PROOFS OF SECTION 4

PROOF OF PROPOSITION 13: In light of Theorem 12, it suffices to show that for $p < 0$ the following quantity is bounded as a function of y:

$$\left(pr_0 - \lambda - \frac{1}{2}q\left(\nu_0' + b\nu_1'y\right)\left(\nu_0 + b\nu_1 y\right) + \frac{1}{2}q\rho'\rho\left(v_0 - v_1 y\right)^2\right)e^{-v_0 y + \frac{1}{2}v_1 y^2} \tag{74}$$

Note that, for $p < 0$, Θ in (46) satisfies:

$$\Theta > \left(1 + q\rho'\nu_1\right)^2 \tag{75}$$

Therefore:

$$v_1 = b\delta\left(\sqrt{\Theta} - \left(1 + q\rho'\nu_1\right)\right) > 0$$

and (74) is bounded over $\mathbb{R}$ only if the quadratic term is negative:

$$\frac{1}{2}qv_1^2\rho'\rho - \frac{1}{2}qb^2\nu_1'\nu_1 < 0$$

But

$$\begin{aligned}\frac{1}{2}qv_1^2\rho'\rho - \frac{1}{2}qb^2\nu_1'\nu_1 &= \frac{1}{2}q\rho'\rho\delta^2b^2\left(\sqrt{\Theta} - \left(1 + q\rho'\nu_1\right)\right)^2 - \frac{1}{2}b^2\delta\left(\Theta - \left(1 + q\rho'\nu_1\right)^2\right)\\ &= \frac{1}{2}\delta^2b^2\left(\sqrt{\Theta} - \left(1 + q\rho'\nu_1\right)\right)\left(\left(2q\rho'\rho - 1\right)\sqrt{\Theta} - \left(1 + q\rho'\nu_1\right)\right)\end{aligned}$$

From (49), the quantity $(2q\rho'\rho - 1)\sqrt{\Theta} - (1 + q\rho'\nu_1)$ is negative, while $\frac{1}{2}\delta^2b^2\left(\sqrt{\Theta} - (1 + q\rho'\nu_1)\right)$ is positive by (75). Therefore, the leading quadratic term is negative and the result follows. *Q.E.D.*

PROOF OF LEMMA 14: When $\nu_1 = -\kappa\rho$ condition (49) reduces to:

$$\left(1 - 2q\rho'\rho\right)\left(1 + q\rho'\rho\left(\kappa^2 - 2\kappa\right)\right)^{1/2} + \left(1 - q\kappa\rho'\rho\right) > 0$$

Set $x = q\rho'\rho$ and consider the continuous function

$$f(x, \kappa) = (1 - 2x)\left(1 + x\left(\kappa^2 - 2\kappa\right)\right)^{1/2} + (1 - \kappa x)$$

on $0 < x < 1, \kappa \in \mathbb{R}$. On $0 < x \leq \frac{1}{4}$ this function has no zeros. Since $f(x, 0) = 2 - 2x > 0$ it follows that for $0 < q\rho'\rho \leq \frac{1}{4}$ any $\kappa \in \mathbb{R}$ leads to long-run optimality. On $\frac{1}{4} < x < 1$ this function is 0 along the curve $\kappa = \frac{2}{4x-1}$. For a fixed x and large positive κ

$$f(x, \kappa) \approx \kappa\left((1 - 2x)\sqrt{x} - x\right) < 0$$

and for large negative κ

$$f(x, \kappa) \approx |\kappa|\left((1 - 2x)\sqrt{x} + x\right) > 0$$

therefore for $\frac{1}{4} < q\rho'\rho < 1$ the restriction $\kappa < \frac{2}{4q\rho'\rho - 1}$ is necessary.

Q.E.D.

PROOF OF LEMMA 15: When $\kappa = 1$, by Lemma 14, long-run optimality holds for $0 < q\rho'\rho \le \frac{1}{4}$. For $\frac{1}{4} < q\rho'\rho < 1$ long-run optimality holds if $1 < \frac{2}{4q\rho'\rho - 1}$ or $q\rho'\rho < \frac{3}{4}$.

Consider now when $q\rho'\rho \ge \frac{3}{4}$ which is equivalent to $\delta \ge 4$ since $\delta = \frac{1}{1-q\rho'\rho}$. When $\kappa = 1$ the solution v_1, v_0 and Θ simplify considerably to $\Theta = \delta^{-1}$; $v_1 = b\left(\sqrt{\delta} - 1\right)$; and $v_0 = q\delta\rho'\nu_0$. Under $\hat{P}$, Y has the dynamics

$$dY_t = -\frac{b}{\sqrt{\delta}} Y_t dt + dW_t$$

For $Y_0 = y$, it follows that $Y_t \sim N(\mu_t, \sigma_t^2)$ with $\mu_t = y e^{-\frac{b}{\sqrt{\delta}}t}$, and $\sigma_t^2 = \frac{\sqrt{\delta}}{2b}\left(1 - e^{-2\frac{b}{\sqrt{\delta}}t}\right)$. Therefore,

$$E_{\hat{P}}^y\left[e^{-v(Y_t)}\right] = E\left[\exp\left(\mathbf{A}Y_t^2 + \mathbf{B}Y_t\right)\right]$$

where $\mathbf{A} = \frac{b}{2}\left(\sqrt{\delta} - 1\right)$, $\mathbf{B} = -q\delta\rho'\xi$. For $X \sim N(\mu, \sigma^2)$

$$E\left[e^{\mathbf{A}X^2 + \mathbf{B}X}\right] = \begin{cases} \left(1 - 2\mathbf{A}\sigma^2\right)^{-1/2} \exp\left(\left(1 - 2\mathbf{A}\sigma^2\right)^{-1}\left(\mu^2\mathbf{A} + \mu\mathbf{B} + \frac{1}{2}\sigma^2\mathbf{B}^2\right)\right) & \mathbf{A} < \frac{1}{2\sigma^2} \\ \infty & \mathbf{A} \ge \frac{1}{2\sigma^2} \end{cases} \tag{76}$$

Therefore, $E_{\hat{P}}^y\left[e^{-v(Y_t)}\right] < \infty$ if and only if $\frac{b}{2}\left(\sqrt{\delta} - 1\right) < \frac{1}{2\sigma_t^2}$. This condition reduces to

$$1 + \frac{\sqrt{\delta}}{2}\left(1 - \sqrt{\delta}\right)\left(1 - e^{-\frac{2b}{\sqrt{\delta}}t}\right) > 0 \tag{77}$$

Note that the left side of (77) is equal to 1 at $t = 0$ and monotonically decreasing in t for $\delta > 1$. Setting this expression equal to 0, and solving for t yields:

$$\hat{t} = -\frac{\sqrt{\delta}}{2b}\log\left(\frac{\sqrt{\delta}\left(\sqrt{\delta} - 1\right) - 2}{\sqrt{\delta}\left(\sqrt{\delta} - 1\right)}\right)$$

If $\sqrt{\delta}\left(\sqrt{\delta} - 1\right) > 2$ or equivalently, $\delta > 4$ then $\hat{t} > 0$ exists. This proves statement (i) in proposition 15.

If $\delta = 4$, the left side of (77) shows that:

$$1 + \frac{\sqrt{\delta}}{2}\left(1 - \sqrt{\delta}\right)\left(1 - e^{-\frac{2b}{\sqrt{\delta}}t}\right) = e^{-bt} > 0$$

for all $t > 0$ and hence

$$E_{\hat{P}}^y\left[e^{-v(Y_t)}\right] = \exp\left(\frac{b}{2}(t + y^2) - 4q\rho'\nu_0 e^{\frac{b}{2}t} y + \frac{8}{b}q^2(\rho'\nu_0)^2\left(1 - e^{-bt}\right)e^{bt}\right)$$

On the other hand, the condition $E_{\hat{P}}^y\left[e^{-\frac{1}{1-p}v(Y_t)}\right]^{1-p} < \infty$ is, for $\delta = 4$:

$$\frac{q}{2} + \frac{1}{2(1-p)}e^{-bt} > 0$$

which is true for all $t > 0$ because $\delta = 4$ only when $0 < q < 1$. Thus

$$E_{\hat{P}}^y\left[e^{-\frac{1}{1-p}v(Y_t)}\right]^{1-p} = \left(1 - (1-q)(1 - e^{-bt})\right)^{-\frac{1-p}{2}}$$
$$\times \exp\left(\frac{(1-p)\left(\frac{1}{2}b(1-q)y^2e^{-bt} - 4q(1-q)\rho'\nu_0 e^{-\frac{b}{2}t}y + \frac{8}{b}q^2(1-q)^2(\rho'\nu_0)^2(1 - e^{-bt})\right)}{(1 - (1-q)(1 - e^{-bt}))}\right)$$

If $\nu_0 = 0$, $E_{\hat{P}}^y\left[e^{-v(Y_t)}\right] \sim Ke^{\frac{b}{2}t}$ and $E_{\hat{P}}^y\left[e^{-\frac{1}{1-p}v(Y_t)}\right]^{1-p} \sim q^{-\frac{1-p}{2}}$ for large t, so the certainty equivalent loss is bounded by $-\frac{b}{2p}$, proving $ii)$ in Proposition 15.

If $\nu_0 \neq 0$, then $E_{\hat{P}}^y\left[e^{-v(Y_t)}\right] \sim e^{K_1 e^{bt}}$ and $E_{\hat{P}}^y\left[e^{-\frac{1}{1-p}v(Y_t)}\right]^{1-p} \sim q^{-\frac{1-p}{2}}e^{K_2}$ for large T, where K_1, K_2 are positive constants. In this case the certainty equivalent loss diverges to $-\infty$ with speed of the order of $\frac{K_1}{t}e^{bt}$. This proves $iii)$, and completes the proof. *Q.E.D.*

Proof of Proposition 16: Since v_0, v_1 satisfy (54), $v(y) = v_0 \log y + v_1 y$ solves (28). Under $\hat{P}$ the dynamics of Y are

$$dY_t = \sqrt{\Theta}\left(\frac{\sqrt{\Lambda}+\frac{1}{2}a^2}{\sqrt{\Theta}} - Y_t\right)dt + a\sqrt{Y_t}dW_t$$

It is now shown for any CIR process of the form

$$dY_t = \kappa\left(\theta - Y_t\right)dt + \xi\sqrt{Y_t}dW_t; Y_0 = y$$

with $\kappa, \theta, \xi > 0$ that $\{Y_t\}_{t\geq\bar{t}}$ is a tight family of random variables with $\bar{t} = \frac{\log 2}{\kappa}$. For each integer $n \geq 2$ consider the compact set $K_n = \left[\frac{1}{n}, n\right]$. From Markov's inequality:

$$P\left[Y_t \in K_n^c\right] \leq P\left[Y_t < \frac{1}{n}\right] + \frac{1}{n^2}E[Y_t^2]$$

$Y_t \overset{d}{=} \frac{\xi^2}{4\kappa}\left(1-e^{-\kappa t}\right)X$ where X is a non-central chi-square random variable with non-centrality parameter $\lambda = \frac{4\kappa e^{-\kappa t}y}{\xi^2\left(1-e^{-\kappa t}\right)}$ and $d = \frac{4\kappa\theta}{\xi^2}$ degrees of freedom. Therefore:

$$\begin{aligned}
E\left[Y_t^2\right] &= \frac{\xi^4}{16\kappa^2}\left(1-e^{-\kappa t}\right)^2 E\left[X^2\right] \\
&= \frac{\xi^4}{16\kappa^2}\left(1-e^{-\kappa t}\right)^2\left(\left(\frac{4\kappa\theta}{\xi^2} + \frac{4\kappa e^{-\kappa t}y}{\xi^2\left(1-e^{-\kappa t}\right)}\right)^2 + 2\left(\frac{4\kappa\theta}{\xi^2} + \frac{8\kappa e^{-\kappa t}y}{\xi^2\left(1-e^{-\kappa t}\right)}\right)\right) \\
&= \left(\left(1-e^{-\kappa t}\right)\theta + e^{-\kappa t}y\right)^2 + \left(1-e^{-\kappa t}\right)\left(\left(1-e^{-\kappa t}\right)\frac{\theta\xi^2}{2\kappa} + \frac{e^{-\kappa t}\xi^2 y}{\kappa}\right) \\
&\leq \theta^2 + 2\theta y + y^2 + \frac{\theta\xi^2}{2\kappa} + \frac{\xi^2 y}{\kappa}
\end{aligned}$$

and this is uniform in $t > 0$. As for $P[Y_t < \frac{1}{n}]$, the CDF for X is

$$P[X \leq x] = \sum_{j=0}^{\infty} e^{-\lambda/2}\frac{(\lambda/2)^j}{j!}\frac{1}{\Gamma(j+d/2)}\int_0^{x/2} s^{j+d/2-1}e^{-s}ds$$

For any $d > 0$ and integer $j \geq 0$

$$\int_0^{x/2} s^{j+d/2-1}e^{-s}ds \leq \frac{1}{j+d/2}\left(\frac{x}{2}\right)^{j+d/2} \leq \frac{2}{d}\left(\frac{x}{2}\right)^{j+d/2}$$

$$\Gamma\left(j+\frac{d}{2}\right) = \Gamma\left(\frac{d}{2}\right)\prod_{k=0}^{j-1}\left(k+\frac{d}{2}\right) \geq \Gamma\left(\frac{d}{2}\right)\left(\frac{d}{2}\right)^j$$

Therefore

$$P[X \leq x] \leq \frac{2}{d\Gamma(d/2)}\left(\frac{x}{2}\right)^{d/2} e^{-\frac{\lambda}{2}(1-x/d)}$$

And hence for $t \geq \bar{t}$ and $\beta = \frac{8\kappa}{\xi^2}$:

$$\begin{aligned}
P\left[Y_t \leq \frac{1}{n}\right] &= P\left[X \leq \frac{4\kappa}{n\xi^2(1-e^{\kappa t})}\right] \\
&\leq P\left[X \leq \frac{\beta}{n}\right] \\
&\leq \frac{2}{d\Gamma(d/2)}\left(\frac{\beta}{2n}\right)^{d/2} e^{-\frac{\lambda}{2}(1-\frac{\beta}{nd})}
\end{aligned}$$

Since $\lambda > 0$ for $t > 0$ and d does not depend upon t, it holds that $P\left[Y_t < \frac{1}{n}\right] \downarrow 0$ uniformly for $t \geq \bar{t}$ as $n \uparrow \infty$, proving tightness.

Therefore long-run optimality will follow if the quantity F for $p < 0$ from Theorem 11 is bounded over $(0, \infty)$. Specifying to this example, it is necessary to show

$$\left(pr_0 + pr_1 y - \lambda - \frac{1}{2}q\left(\nu_0' + y\nu_1\right)\frac{1}{y}\left(\nu_0 + y\nu_1\right) + \frac{1}{2}q\left(\frac{v_0}{y} + v_1\right)a^2\rho'\rho y\left(\frac{v_0}{y} + v_1\right)\right)e^{-(v_0\log y + v_1 y)} \tag{78}$$

is bounded on $y > 0$. This expression admits the form

$$\left(\mathbf{A} + \mathbf{B}y + \mathbf{C}y^2\right) y^{-v_0-1} e^{-v_1 y}$$

For v_0, v_1 from (54) by (53) it follows that $v_0 > 0, v_1 < 0$ and so (78) will follow only if $\mathbf{A} < 0, \mathbf{C} < 0$. As for $\mathbf{A}$

$$\begin{aligned}
\mathbf{A} &= \frac{1}{2} q a^2 \rho' \rho v_0^2 - \frac{1}{2} q \nu_0' \nu_0 \\
&= \frac{1}{2} q a^2 \rho' \rho \frac{\delta^2}{a^4} \left(\sqrt{\Lambda} - \left(b\theta - q a \rho' \nu_0 - \frac{1}{2} a^2 \right) \right)^2 - \frac{1}{2} \frac{\delta}{a^2} \left(\Lambda - \left(b\theta - q a \rho' \nu_0 - \frac{1}{2} a^2 \right)^2 \right) \\
&= \frac{1}{2} \frac{\delta^2}{a^2} \left(\sqrt{\Lambda} - \left(b\theta - q a \rho' \nu_0 - \frac{1}{2} a^2 \right) \right) \left(\left(2 q \rho' \rho - 1\right) \sqrt{\Lambda} - \left(b\theta - q a \rho' \nu_0 - \frac{1}{2} a^2 \right) \right)
\end{aligned}$$

From (55)

$$\left(2 q \rho' \rho - 1\right) \sqrt{\Lambda} - \left(b\theta - q a \rho' \nu_0 - \frac{1}{2} a^2 \right) < 0$$

Thus, $\mathbf{A} < 0$ since by (53)

$$\frac{1}{2} \frac{\delta^2}{a^2} \left(\sqrt{\Lambda} - \left(b\theta - q a \rho' \nu_0 - \frac{1}{2} a^2 \right) \right) > 0$$

As for $\mathbf{C}$

$$\begin{aligned}
\mathbf{C} &= \frac{1}{2} q a^2 \rho' \rho v_1^2 + p r_1 - \frac{1}{2} q \nu_1' \nu_1 \\
&= \frac{1}{2} q a^2 \rho' \rho \frac{\delta^2}{a^4} \left(\left(b + q a \rho' \nu_1\right) - \sqrt{\Theta} \right)^2 + \frac{1}{2} \frac{\delta}{a^2} \left(\left(b + q a \rho' \nu_1\right)^2 - \Theta \right) \\
&= \frac{1}{2} \frac{\delta^2}{a^2} \left(\left(b + q a \rho' \nu_1\right) - \sqrt{\Theta} \right) \left(\left(1 - 2 q \rho' \rho\right) \sqrt{\Theta} + \left(b + q a \rho' \nu_1\right) \right)
\end{aligned}$$

From (55)

$$\left(1 - 2 q \rho' \rho\right) \sqrt{\Theta} + \left(b + q a \rho' \nu_1\right) > 0$$

Thus, $\mathbf{C} < 0$, since by (53)

$$\frac{1}{2} \frac{\delta^2}{a^2} \left(\left(b + q a \rho' \nu_1\right) - \sqrt{\Theta} \right) < 0$$

Q.E.D.

REFERENCES

Abou-Kandil, H., Freiling, G., Ionescu, V. and Jank, G. (2003), *Matrix Riccati equations*, Systems & Control: Foundations & Applications, Birkhäuser Verlag, Basel. In control and systems theory.

Balduzzi, P. and Lynch, A. (1999), 'Transaction Costs and Predictability: Some Utility Cost Calculations', *Journal of Financial Economics* **52**(1), 47–78.

Barberis, N. (2000), 'Investing for the Long Run When Returns Are Predictable', *The Journal of Finance* **55**(1), 225–264.

Bielecki, T. R. and Pliska, S. R. (1999), 'Risk-sensitive dynamic asset management', *Appl. Math. Optim.* **39**(3), 337–360.

Bielecki, T. R. and Pliska, S. R. (2000), 'Risk sensitive asset management with transaction costs', *Finance Stoch.* **4**(1), 1–33.

Brandt, M. (1999), 'Estimating Portfolio and Consumption Choice: A Conditional Euler Equations Approach', *The Journal of Finance* **54**(5), 1609–1645.

Brandt, M., Goyal, A., Santa-Clara, P. and Stroud, J. (2005), 'A Simulation Approach to Dynamic Portfolio Choice with an Application to Learning About Return Predictability', *Review of Financial Studies* **18**(3), 831.

Brennan, M. J., Schwartz, E. S. and Lagnado, R. (1997), 'Strategic asset allocation', *J. Econom. Dynam. Control* **21**(8-9), 1377–1403. Computational financial modelling.

Brennan, M. and Xia, Y. (2002), 'Dynamic Asset Allocation under Inflation', *The Journal of Finance* **57**(3), 1201–1238.

Campbell, J. and Shiller, R. (1988), 'The dividend-price ratio and expectations of future dividends and discount factors', *Review of Financial Studies* **1**(3), 195–228.

Campbell, J. and Viceira, L. (1999), 'Consumption and Portfolio Decisions When Expected Returns Are Time Varying', *The Quarterly Journal of Economics* **114**(2), 433–495.

Campbell, J. and Viceira, L. (2001), 'Who Should Buy Long-Term Bonds?', *The American Economic Review* **91**(1), 99–127.

Campbell, J. and Viceira, L. (2002), *Strategic Asset Allocation: Portfolio Choice for Long-Term Investors*, Oxford University Press.

Chacko, G. and Viceira, L. (2005), 'Dynamic Consumption and Portfolio Choice with Stochastic Volatility in Incomplete Markets', *Review of Financial Studies* **18**(4), 1369–1402.

Cochrane, J. and Piazzesi, M. (2005), 'Bond risk premia', *The American economic review* **95**(1), 138–160.

Cox, J. C. and Huang, C.-f. (1989), 'Optimal consumption and portfolio policies when asset prices follow a diffusion process', *J. Econom. Theory* **49**(1), 33–83.

Cox, J. C. and Huang, C.-f. (1992), 'A continuous-time portfolio turnpike theorem', *J. Econom. Dynam. Control* **16**(3-4), 491–507.

Cox, J. C., Ingersoll, Jr., J. E. and Ross, S. A. (1985), 'A theory of the term structure of interest rates', *Econometrica* **53**(2), 385–407.

Cvitanic, J. and Karatzas, I. (1995), 'On portfolio optimization under drawdown constraints', *IMA Lecture Notes in Mathematics & Applications* **65**, 77–88.

Das, S. and Sundaram, R. (2000), A numerical algorithm for optimal consumption-investment problems, Technical report, Working Paper, Harvard University.

Davis, M. (1997), 'Option Pricing in Incomplete Markets', *Mathematics of Derivative Securities* .

Detemple, J., Garcia, R. and Rindisbacher, M. (2003), 'A Monte Carlo Method for Optimal Portfolios', *The Journal of Finance* **58**(1), 401–446.

Donsker, M. D. and Varadhan, S. R. S. (1975), 'Asymptotic evaluation of certain Markov process expectations for large time. I. II', *Comm. Pure Appl. Math.* **28**, 1–47; ibid. **28** (1975), 279–301.

Donsker, M. D. and Varadhan, S. R. S. (1976), 'Asymptotic evaluation of certain Markov process expectations for large time. III', *Comm. Pure Appl. Math.* **29**(4), 389–461.

Donsker, M. D. and Varadhan, S. R. S. (1983), 'Asymptotic evaluation of certain Markov process expectations for large time. IV', *Comm. Pure Appl. Math.* **36**(2), 183–212.

Dumas, B. and Luciano, E. (1991), 'An Exact Solution to a Dynamic Portfolio Choice Problem under Transactions Costs', *The Journal of Finance* **46**(2), 577–595.

Dybvig, P. H., Rogers, L. and Back, K. (1999), 'Portfolio turnpikes', *Review of Financial Studies* **12**(1).

Fama, E. and Bliss, R. (1987), 'The Information in Long-Maturity Forward Rates', *The American Economic Review* **77**(4), 680–692.

Fama, E. and French, K. R. (1988), 'Dividend yields and expected stock returns', *Journal of Financial Economics* **22**(1), 3–25.

Feller, W. (1951), 'Two singular diffusion problems', *Ann. of Math. (2)* **54**, 173–182.

Ferretti, R. and Trojani, F. (2005), General Analytical Solutions for Merton's-Type Consumption Investment Problems. Working Paper University of St. Gallen.

Fleming, W. H. and Sheu, S. J. (2000), 'Risk-sensitive control and an optimal investment model', *Math. Finance* **10**(2), 197–213. INFORMS Applied Probability Conference (Ulm, 1999).

Fleming, W. H. and Sheu, S. J. (2002), 'Risk-sensitive control and an optimal investment model. II', *Ann. Appl. Probab.* **12**(2), 730–767.

Föllmer, H. and Schachermayer, W. (2007), 'Asymptotic Arbitrage and Large Deviations', *Mathematics and Financial Economics* .

Frittelli, M. (2000), 'The minimal entropy martingale measure and the valuation problem in incomplete markets', *Math. Finance* **10**(1), 39–52.

Grandits, P. and Rheinländer, T. (2002), 'On the minimal entropy martingale measure', *Ann. Probab.* **30**(3), 1003–1038.

Grossman, S. and Zhou, Z. (1993), 'Optimal investment strategies for controlling drawdowns', *Mathematical Finance* **3**(3), 241–276.

Hakansson, N. H. (1974), 'Convergence to isoelastic utility and policy in multiperiod portfolio choice', *Journal of Financial Economics, 1* **1**, 201–224.

Hansen, L. and Jagannathan, R. (1991), 'Implications of Security Market Data for Models of Dynamic Economies', *The Journal of Political Economy* **99**(2), 225–262.

Hansen, L. and Scheinkman, J. (2006), 'Long Term Risk: An Operator Approach'.

He, H. and Pearson, N. D. (1991), 'Consumption and portfolio policies with incomplete markets and short-sale constraints: the infinite-dimensional case', *J. Econom. Theory* **54**(2), 259–304.

Henderson, V. (2005), ‘Analytical comparisons of option prices in stochastic volatility models’, *Math. Finance* **15**(1), 49–59.

Hobson, D. (2004), ‘Stochastic volatility models, correlation, and the q-optimal measure’, *Math. Finance* **14**(4), 537–556.

Huang, C.-f. and Zariphopoulou, T. (1999), ‘Turnpike behavior of long-term investments’, *Finance Stoch.* **3**(1), 15–34.

Huberman, G. and Ross, S. (1983), ‘Portfolio turnpike theorems, risk aversion, and regularly varying utility functions’, *Econometrica* **51**(5), 1345–1361.

Karatzas, I., Lehoczky, J. P. and Shreve, S. E. (1987), ‘Optimal portfolio and consumption decisions for a “small investor” on a finite horizon’, *SIAM J. Control Optim.* **25**(6), 1557–1586.

Karatzas, I., Lehoczky, J. P., Shreve, S. E. and Xu, G.-L. (1991), ‘Martingale and duality methods for utility maximization in an incomplete market’, *SIAM J. Control Optim.* **29**(3), 702–730.

Kim, T. and Omberg, E. (1996), ‘Dynamic Nonmyopic Portfolio Behavior’, *The Review of Financial Studies* **9**(1), 141–161.

Kogan, L. and Uppal, R. (2001), ‘Risk Aversion and Optimal Portfolio Policies in Partial and General Equilibrium Economies’, *NBER Working Paper* .

Kramkov, D. and Schachermayer, W. (1999), ‘The asymptotic elasticity of utility functions and optimal investment in incomplete markets’, *Ann. Appl. Probab.* **9**(3), 904–950.

Kreĭn, M. G. and Rutman, M. A. (1948), ‘Linear operators leaving invariant a cone in a Banach space’, *Uspehi Matem. Nauk (N. S.)* **3**(1(23)), 3–95.

Kuroda, K. and Nagai, H. (2002), ‘Risk-sensitive portfolio optimization on infinite time horizon’, *Stoch. Stoch. Rep.* **73**(3-4), 309–331.

Leland, H. (1972), On turnpike portfolios, *in* G. Szego and K. Shell, eds, ‘Mathematical Models in Investment and Finance’, North-Holland, Amsterdam.

Liu, J. (2007), ‘Portfolio Selection in Stochastic Environments’, *Review of Financial Studies* **20**(1), 1.

Merton, R. (1969), ‘Lifetime Portfolio Selection under Uncertainty: The Continuous-Time Case’, *The Review of Economics and Statistics* **51**(3), 247–257.

Merton, R. (1973), ‘An Intertemporal Capital Asset Pricing Model’, *Econometrica* **41**(5), 867–887.

Nagai, H. and Peng, S. (2002), ‘Risk-sensitive dynamic portfolio optimization with partial information on infinite time horizon’, *Ann. Appl. Probab.* **12**(1), 173–195.

Pham, H. (2003), ‘A large deviations approach to optimal long term investment’, *Finance Stoch.* **7**(2), 169–195.

Pliska, S. R. (1986), ‘A stochastic calculus model of continuous trading: optimal portfolios’, *Math. Oper. Res.* **11**(2), 370–382.

Reed, M. and Simon, B. (1980), *Methods of modern mathematical physics. I*, second edn, Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York. Functional analysis.

Revuz, D. and Yor, M. (1999), *Continuous martingales and Brownian motion*, Vol. 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*, third edn, Springer-Verlag, Berlin.

Sangvinatsos, A. and Wachter, J. (2005), ‘Does the Failure of the Expectations Hypothesis Matter for Long-Term Investors?’, *The Journal of Finance* **60**(1), 179–230.

Stroock, D. W. and Varadhan, S. R. S. (2006), *Multidimensional diffusion processes*, Classics in Mathematics, Springer-Verlag, Berlin. Reprint of the 1997 edition.

Summers, L. (1986), ‘Does the Stock Market Rationally Reflect Fundamental Values?’, *The Journal of Finance* **41**(3), 591–601.

Wachter, J. (2002), ‘Portfolio and consumption decisions under mean-reverting returns: An exact solution for complete markets’, *Journal of Financial and Quantitative Analysis* **37**(1), 63–91.

Zariphopoulou, T. (2001), ‘A solution approach to valuation with unhedgeable risks’, *Finance Stoch.* **5**(1), 61–82.