PORTFOLIOS AND RISK PREMIA FOR THE LONG RUN

BY PAOLO GUASONI AND SCOTT ROBERTSON

Boston University, Dublin City University, and Carnegie Mellon 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 quasilinear partial differential equation, which is solvable for many models of interest.

The paper derives the long-run optimal portfolio and the long-run optimal pricing measures depending on relative risk aversion, as well as their finite-horizon performance.

Introduction. Long-run asymptotics are a powerful tool to obtain explicit formulas in portfolio choice and derivatives pricing, but their use has been mostly ad hoc, in the absence of general results. 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 that follow an autonomous diffusion.

Investment and pricing problems share a reputation for mathematical complexity. This common trait is not an accident: the central message of duality theory\(^1\) is that the two problems are indeed equivalent, as state-price densities are proportional to the marginal utilities of optimal payoffs. In spite of this conceptual equivalence, portfolio choice and derivatives pricing have followed largely different strands of literature, each of them with its own terminology.

The portfolio choice literature focuses on finding the \textit{intertemporal hedging} component of optimal portfolios\(^2\). Long-run asymptotics have appeared in this literature under different names: the risk-

---

\(^1\)See for example Pliska (1986); Karatzas et al. (1987); Cox and Huang (1989); He and Pearson (1991); Kramkov and Schachermayer (1999), as well as Castañeda-Leyva and Hernández-Hernández (2005) for a setting similar to the one in this paper.

\(^2\)Kim and Omberg (1996); Brennan and Xia (2002); Wachter (2002); Munk and Sørensen (2004); Liu (2007), compute optimal portfolios explicitly for certain models.
sensitive control approach\textsuperscript{3}, turnpike results\textsuperscript{4}, and large deviations criteria\textsuperscript{5} are all efforts to achieve tractability by means of the long-run limit.

The derivatives pricing literature strives to identify martingale measures that are optimal in the sense of the minimax martingale measure of He and Pearson (1991) or the least favorable completion of Karatzas, Lehoczky, Shreve and Xu (1991). Power utility leads to the $q$-optimal measure (Hobson, 2004; Henderson, 2005), which embeds several other martingale measures\textsuperscript{6}: it reduces to the minimal measure for $q = 0$, to the minimal entropy measure for $q = 1$, and to the variance-optimal measure for $q = 2$.

The advantages of long-run asymptotics are their tractability and accuracy. Long-run portfolios and risk premia are much simpler than their finite-horizon counterparts, and allow explicit expressions even in cases for which the latter do not. In general, long-run policies are identified by the quasilinear partial differential equation (22), which admits explicit solutions in several models of interest. This equation reduces to an ordinary differential equation in the case of a single state variable, which is furthermore linear if the state variable has a constant correlation with excess returns. The accuracy of the long-run approach stems from the bounds (27), which estimate the duality gap at any horizon, and hence the potential departure from the unknown finite-horizon optimum. Long-run optimality holds (Definition 6) when long-run policies are approximately optimal over long horizons. The main result of this paper gives a sufficient condition for long-run optimality in a general multidimensional diffusion. Furthermore, this condition is sharp for certain models, and a calibration to the parameters estimated by Barberis (2000) shows that it is satisfied for reasonable levels of risk aversion.

Two duality insights are central to our results. First, the usual duality between payoffs and martingale densities extends to their stochastic logarithms, which are portfolios and risk premia. Second, long-run asymptotics become easier in a duality context, because candidate long-run risk-premia yield an upper bound on the maximal expected utility, and vice versa. This observation allows to overcome some difficulties arising in the verification theorems of the risk-sensitive control literature.

An important concept arising in long-run analysis is the myopic probability: a long-run investor with power utility under the original probability behaves like a logarithmic (or myopic) investor under the myopic probability. This probability plays an important role both for long-run analysis and for finite-horizon bounds, and its existence is crucial for the long-run optimality result.

The rest of the paper is organized as follows. Section 1 describes the model in detail, introducing notation. Section 2 contains the main result: a general method to obtain long-run policies in closed form. It also provides sufficient conditions, adapted from Kaise and Sheu (2006), for the existence of solutions to the associated ergodic Bellman equation. Section 3 discusses the various implications of these results for portfolio choice and derivatives pricing, and the connections with the Stochastic Control and Large Deviations approaches. Section 4 derives long-run portfolios and risk premia in two models of interest. The last one combines stochastic interest rates, drifts, and volatilities, and still admits simple closed form solutions. Section 5 concludes. All proofs are in the Appendix.

\textbf{1. Model.}


\textsuperscript{4}Leland (1972); Halanssone (1974); Huberman and Ross (1983); Cox and Huang (1992); Jin (1998); Huang and Zariphopoulos (1999); Dybvig, Rogers and Back (1999)

\textsuperscript{5}Pham (2003), Föllmer and Schachermayer (2007)

1.1. Market. Consider a financial market with a risk-free asset $S^0$ and $n$ risky assets $S = (S^1, \ldots, S^n)$. Investment opportunities (interest rates, expected returns and covariances) depend on $k$ state variables $Y = (Y^1, \ldots, Y^k)$, which model their change over time:

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

$$\frac{dS^i_t}{S^i_t} = r(Y_t)dt + dR^i_t \quad 1 \leq i \leq n$$

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

$$dR^i_t = \mu_i(Y_t)dt + \sum_{j=1}^n \sigma_{ij}(Y_t)dZ^j_t \quad 1 \leq i \leq n$$

$$dY^i_t = b_i(Y_t)dt + \sum_{j=1}^k a_{ij}(Y_t)dW^j_t \quad 1 \leq i \leq k$$

$$d\langle Z^i, W^j \rangle_t = \rho_{ij}(Y_t)dt \quad 1 \leq i \leq n, 1 \leq j \leq k$$

where $Z = (Z^1, \ldots, Z^n)$ and $W = (W^1, \ldots, W^k)$ are multivariate Brownian Motions. This model provides a flexible framework that nests most diffusion models in Finance, including the models of correlation risk considered by Buraschi, Porchia and Trojani (2010), in which $\rho$ is a function of a state variable.

The law of $(R, Y)$ determines the drifts $b, \mu$ and the covariance matrices $\Sigma = \sigma\sigma' = d\langle R, R \rangle_t/dt$, $A = aa' = d\langle Y, Y \rangle_t/dt$, and $\Upsilon = \sigma\rho\sigma' = d\langle R, Y \rangle_t/dt$, where the prime sign denotes matrix transposition. By contrast, the matrices $\sigma, a, \rho$ are identified only up orthogonal transformations. The market defined by (1)-(5) is in general incomplete, and 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 $\Upsilon = 0$.

Let $E \subseteq \mathbb{R}^k$ be an open connected set. Denote by $C^m(E, \mathbb{R}^d)$ (resp. $C^{m,\gamma}(E, \mathbb{R}^d)$) the class of $\mathbb{R}^d$-valued functions on $E$ with continuous (resp. locally $\gamma$-Hölder continuous) partial derivatives of $m^{\omega}$-order. The superscripts are dropped for $m = 0$ or $d = 1$, so that $C^{0,\gamma}(E, \mathbb{R}^1)$ is denoted by $C_\gamma(E, \mathbb{R})$. The following assumption prescribes that the coefficients $r, \mu, b, A, \Sigma$ and $\Upsilon$ are regular and nondegenerate:

**Assumption 1.** $r \in C_\gamma(E, \mathbb{R})$, $b \in C^{1,\gamma}(E, \mathbb{R}^k)$, $\mu \in C^{1,\gamma}(E, \mathbb{R}^n)$, $A \in C^{2,\gamma}(E, \mathbb{R}^{k \times k})$, $\Sigma \in C^{2,\gamma}(E, \mathbb{R}^{n \times n})$ and $\Upsilon \in C^{2,\gamma}(E, \mathbb{R}^{n \times k})$. The symmetric matrices $A$ and $\Sigma$ are positive definite for all $y \in E$.

The region $E$ is typically of the form $E = \mathbb{R}^d \times (0, \infty)^{k-d}$ for some $0 \leq d \leq k$, as in the case of multi-variate Ornstein-Uhlenbeck processes, Feller diffusions, or a combination thereof. Fernholz and Karatzas (2005) consider models in which $E$ is the open simplex in $\mathbb{R}^{n-1}$.

To construct the processes $(R, Y)$, let $\Omega = C([0, \infty), \mathbb{R}^{n+k})$ endowed with the topology of uniform convergence on bounded intervals. Let $\mathcal{B}$ be the Borel sigma algebra and let $(\mathcal{B}_t)_{t \geq 0}$ be the filtration generated by the coordinate process $X$ defined by $X_t(\omega) = \omega_t$ for $\omega \in \Omega$. For a second order differential operator $L$ of the form (6) below, a solution to the martingale problem for $L$ on $\mathbb{R}^n \times E$ is a family of Borel probability measures $(P^x)_{x \in \mathbb{R}^n \times E}$ on $(\Omega, \mathcal{B})$ such that, for each $x \in \mathbb{R}^n \times E$, i) $P^x(X_0 = x) = 1$, ii) $P^x(X_t \in \mathbb{R}^n \times E, \forall t \geq 0) = 1$, and iii) $(f(X_t) - f(X_0) - \int_0^t (Lf)(X_u)du; \mathcal{B}_t)$ is a $P^x$ martingale for all $f \in C^2_b(\mathbb{R}^n \times E)$. 


The next assumption ensures that $\mu, b, A, \Sigma$ and $\Upsilon$ identify the law of $(R,Y)$, $x = (r,y)$, with $r \in \mathbb{R}^n, y \in \mathbb{R}^k$, denotes the starting points of $R$ and $Y$.

**Assumption 2.** There exists a unique solution $(P^{(r,y)})_{r \in \mathbb{R}^n, y \in \mathbb{E}}$ to the martingale problem for:

$$L = \frac{1}{2} \sum_{i,j=1}^{n+k} \tilde{A}^{i,j}(x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{n+k} \tilde{b}^i(x) \frac{\partial}{\partial x_i}$$

$$\tilde{A} = \begin{pmatrix} \Sigma & \Upsilon \\ \Upsilon' & A \end{pmatrix}, \quad \tilde{b} = \begin{pmatrix} \mu \\ b \end{pmatrix}$$

Since $R_0 = 0$ for all the models considered in this paper, the measure $P^{(0,y)}$ in Assumption 2 is simply denoted as $P^y$. Henceforth, consider the filtration $(\mathcal{F}_t)_{t \geq 0}$ defined as $\mathcal{F}_t = \mathcal{B}_t$, that is the right continuous envelope of $\mathcal{B}_t$. Under Assumption 1, $f(X_t) - f(X_0) - \int_0^t (Lf)(X_u)du$ is a martingale also with respect to $\mathcal{F}_t$ (cf. Karatzas and Shreve (1991, Section 5.4)).

**Remark 3.** For consistency of notation, Assumption 2 involves the joint diffusion process $(R,Y)$. However, it is essentially an Assumption on the process $Y$. Indeed, if Assumption 1 holds, and if there is a unique solution $(Q^y)_{y \in \mathbb{E}}$ to the martingale problem for the operator

$$L^Y = \frac{1}{2} \sum_{i,j=1}^k A^{i,j}(y) \frac{\partial^2}{\partial y_i \partial y_j} + \sum_{i=1}^k b^i(y) \frac{\partial}{\partial y_i},$$

then Assumption 2 holds. To see this fact, first consider the martingale problem for the operator associated to $(Y,B)$, where $B$ is an $n$-dimensional standard Brownian Motion independent of $Y$. Then, write $Z = \rho W + \tilde{\rho}B$ where $\tilde{\rho}$ is a square root of $1 - \rho \rho'$, and define the integrals for $R$ in (3) accordingly.

1.2. **Trading Strategies.** An investor trades in the market according to a portfolio $\pi = (\pi_i)_{1 \leq i \leq n}$, representing the proportions of wealth in each risky asset. Since the investor observes the state variables $Y$ and the asset returns $R$, the portfolio $\pi$ is adapted to the filtration generated by $(R,Y)$, and is $R$-integrable. The corresponding wealth process $X^\pi = (X^\pi_t)_{t \geq 0}$ follows:

$$dX^\pi_t = r(Y_t)dt + \pi_t' dR_t$$

Note that a positive initial capital $X_0 \geq 0$ implies a positive wealth at all times, that is $X^\pi_t \geq 0$ a.s. for all $t \geq 0$, thereby ruling out doubling strategies (see for example Harrison and Pliska (1981)).

1.3. **Preferences.** The investor’s preferences are modeled using the power utility function:

$$U(x) = \frac{x^p}{p} \quad \text{for} \quad p < 1, p \neq 0$$

Denoting by $E^y_P$ the expectation with respect to $P^y$, the goal is to maximize expected utility from terminal wealth. With a finite horizon $T$, the problem is:

$$\max_\pi \frac{1}{T} E^y_P [(X^\pi_T)^p]$$

$R$-integrability means each of the integrals $\int_0^T \pi \mu dt, \int_0^T \pi \sigma dZ_t$ is well defined.
Since power utility is homothetic \((U(cx) = c^p U(x))\), it suffices to consider the case \(X_0 = 1\) of unit initial wealth. Henceforth, \(q\) denotes the conjugate exponent to \(p\):

\[
q := \frac{p}{p - 1}
\]

To preserve economic intuition, recall that \(p = 1 - \gamma\) and \(q = 1 - \frac{1}{\gamma}\), where \(\gamma\) is the investor’s relative risk aversion. Risk aversion increases as \(q\) increases and as \(p\) decreases, and logarithmic utility corresponds to the limit \(p \rightarrow 0\).

The martingale approach to utility maximization relies on the duality between final payoffs and pricing rules, represented by the related concepts of stochastic discount factors and martingale measures.

**Definition 4.** A stochastic discount factor is a strictly positive adapted process \(M = (M_t)_{t \geq 0}\), such that \(MS\) is a martingale:

\[
E^y_P [M_t S_t \mid \mathcal{F}_s] = M_s S_s^y \quad \text{for all } 0 \leq s \leq t, 0 \leq i \leq n
\]

A martingale measure is a probability \(Q\), such that \(Q|_{\mathcal{F}_i}\) and \(P^y|_{\mathcal{F}_i}\) 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 \(\frac{dQ}{dP} \mid_{\mathcal{F}_i} = S^i_t M_t\), although their distinction is important in the present context of stochastic interest rates. Except for a complete market, in which the martingale measure is unique, the description of a pricing rule requires the choice of unhedgeable risk premia \(\eta\). For any \((\mathcal{F}_i)_{t \geq 0}\) adapted, integrable process \(\eta\), a candidate (local) martingale measure \(Q^{\eta,y}\) is given by

\[
dQ^{\eta,y} / dP|_{\mathcal{F}_i} = Z^\eta_t, \quad \text{where}
\]

\[
Z^\eta_t = \mathcal{E} \left( -\int_0^t (\mu' \Sigma^{-1} + \eta' \Sigma^{-1}) \sigma dZ + \int_0^t \eta' dW \right)_t
\]

and where \(\mathcal{E}(X)_t = \exp(X_t - \frac{1}{2} \langle X \rangle_t)\). Clearly, \(Z\) must be a martingale for \(Q^{\eta,y}\) to be an equivalent local martingale measure. For such \(\eta\), let \(M^\eta\) denote the corresponding stochastic discount factor:

\[
M^\eta_t = \exp \left( -\int_0^t r dt \right) \mathcal{E} \left( -\int_0^t (\mu' \Sigma^{-1} + \eta' \Sigma^{-1}) \sigma dZ + \int_0^t \eta' dW \right)_t
\]

Note that for any strategy \(\pi\) and risk premia \(\eta\), by (8), the process \(X^\pi M^\eta\) is a super-martingale, even if the right hand side of (12) is a only a local martingale. For power utility, the following lemma applied to \(X = X_T^\pi\) and \(M = M^\eta_T\) for any \(T > 0\) shows that the duality bound is an immediate consequence of Hölder’s inequality and the super-martingale property of the process \(X^\pi M^\eta\).

**Lemma 5.** Let \(X,M\) be random variables on a probability space \((\Omega, \mathcal{F}, P)\) such that \(X,M > 0\) \(P\) almost surely and \(E_P [XM] \leq 1\). Then

\[
\frac{1}{p} E_P [X^p] \leq \frac{1}{p} E_P [M^\eta]^{1-p}
\]

and equality holds if and only if \(E_P [XM] = 1\) and, for some \(\alpha > 0\):

\[
X^{\alpha - 1} = \alpha M
\]
Equation (14) bounds the utility of any terminal wealth by a moment of any stochastic discount factor, and vice versa. The first-order condition (15) is the usual alignment of marginal utilities with state-price densities.

Consider a finite horizon $T$. Lemma 5 implies that a pair $(\pi^T, \eta^T)$ of a portfolio $\pi^T$ and risk premia $\eta^T$ such that $X = X_T^{\pi^T}$ and $M = M^{\eta^T}_T$ is optimal if it satisfies (15) and $E_P^y \left[ X_T^{\pi^T} M^{\eta^T}_T \right] = 1$. Denoting by $u_T(y)$ the value function, that is the maximal expected utility, the following equalities hold:

\begin{equation}
\frac{1}{p} E_P^y \left[ (X_T^{\pi^T})^p \right] = u_T(y) = \frac{1}{p} E_P^y \left[ (M_T^{\eta^T})^q \right]^{1-p}
\end{equation}

hence $\pi^T$ is the optimal portfolio, and the stochastic discount factor $M^{\eta^T}_T$ identifies the pricing rule that makes an investor indifferent between buying and selling a small amount of any payoff, including unhedgeable ones.

1.4. Long-Run Optimality. In the Markov model defined by (1)-(5), stochastic control arguments (see for example Pham (2002) among many others), show that the pair $(\pi^T, \eta^T)$ achieving optimality is of the form $\pi^T_t = \pi^T(T - t, Y_t)$ and $\eta^T_t = \eta^T(T - t, Y_t)$ for deterministic functions

$$\pi^T : [0, T] \times E \rightarrow \mathbb{R}^n \quad \eta^T : [0, T] \times E \rightarrow \mathbb{R}^k$$

Thus, 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.

Brandt (1999), Barberis (2000) and Wachter (2002) report that optimal policies converge rapidly to functions of state variables alone. Thus, the optimal policy for a long horizon $[0, T]$ is approximately equal to a time-homogeneous function for most of the interval, departing from it as the horizon $T$ approaches. The question is whether using a time-homogenous policy throughout the interval $[0, T]$ can be approximately optimal.

For any functions $\pi \in C(E, \mathbb{R}^n)$, $\eta \in C(E, \mathbb{R}^k)$ consider the portfolio $\pi = (\pi(Y_t))_{t \geq 0}$ and risk premia $\eta = (\eta(Y_t))_{t \geq 0}$. At any finite horizon $T$, the duality bound (14) implies that:

\begin{equation}
\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}
\end{equation}

The first inequality reflects the potential gap between the utility of the long-run portfolio and the value function. A tangible measure of this gap is the increase in the risk-free rate $l_T$ required to recover this loss, as to match the expected utility of the long-run optimal portfolio under the higher rate with the maximum expected utility at the regular rate. This is the certainty equivalent loss, defined as:

\begin{equation}
\frac{1}{p} E_P^y \left[ (e^{l_T T} X_T^\pi)^p \right] = u_T(y)
\end{equation}

Substituting (18) into (17) yields an upper bound on $l_T$:

\begin{equation}
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)
\end{equation}

This argument motivates the definition of a pair $(\pi, \eta)$ as long-run optimal when its certainty equivalent loss vanishes for long horizons.
**Definition 6.** A pair \((\pi, \eta) \in C(E, \mathbb{R}^n) \times C(E, \mathbb{R}^k)\) is long-run optimal if, for all \(y \in E\):

\[
\limsup_{T \to \infty} \frac{1}{p} \left( \frac{1}{T} \log E_P^y \left[ (M^\eta_T)^q \right]^{1-p} - \frac{1}{T} \log E_P^y \left[ (X^\pi_T)^p \right] \right) = 0
\]

Long-run optimality defined here is essentially equivalent to the criterion used by Grossman and Vila (1992) to solve portfolio choice problems with leverage constraints. Grossman and Zhou (1993) apply the same idea to drawdown constraints, and Dumas and Luciano (1991) to transaction costs. The *risk-sensitive control* literature studies a similar objective for multidimensional linear diffusions.

Definition 6 departs from the existing literature by examining both the primal (investment) and the dual (pricing) problems. This perspective yields verification theorems that are valid for general multidimensional diffusions, provides estimates on finite-horizon performance, and allows to identify the parameter sets for which long-run optimality holds.

Definition 6 allows another interpretation based on management fees: an investor with sufficiently long horizon prefers a long-run optimal portfolio to the optimal finite-horizon portfolio, if the long-run portfolio has slightly lower fees. The argument is straightforward: suppose that the portfolio \(\pi\) requires the payment of a (continuously compounded) fee \(\varphi\), while the finite-horizon portfolio \(\pi_T\) entails a higher fee \(\varphi_T \geq \varphi\). Accounting for such fees, the portfolio \(\pi\) has expected utility \(\frac{1}{p} E_P^y \left[ (X^\pi_T e^{-\varphi T})^p \right]\). However, by the bound (14) the finite-horizon portfolio \(\pi_T\) satisfies:

\[
\frac{1}{p} E_P^y \left[ (X^\pi_T e^{-\varphi T})^p \right] \leq \frac{1}{p} E_P^y \left[ (M^\eta_T)^q \right]^{1-p} e^{-p \varphi T}
\]

Hence, \(\pi\) is preferred to \(\pi_T\) when:

\[
\varphi_T - \varphi \geq \frac{1}{p} \left( \frac{1}{T} \log E_P^y \left[ (M^\eta_T)^q \right]^{1-p} - \frac{1}{T} \log E_P^y \left[ (X^\pi_T)^p \right] \right)
\]

For a long-run optimal pair \((\pi, \eta)\), the limit of the right-hand side is zero. Thus, for any minimal difference in fees, a long-run optimal portfolio is preferable for investors with sufficiently long horizons.

2. **Long-Run Analysis.** The construction of long-run optimal portfolios \(\pi\) and risk premia \(\eta\) takes place in two steps. In the first step, Theorem 7 computes the finite-horizon performance of the long-run optimal “candidates” \(\pi\) and \(\eta\). In the second step, Theorem 9 establishes a sufficient condition for long-run optimality, requiring that the bounds found in the first step converge at long horizons.

The candidate long-run optimal \(\pi\) and \(\eta\) crucially depend upon the solution of the quasi-linear partial differential equation (PDE) in (22), which acts as a long-run version of the Hamilton-Jacobi Bellman equation. Thus, Theorems 7 and 9 are akin to verification theorems of stochastic control theory, but for the asymptotic objective in Definition 6. An advantage of these results is that they only rely on the local properties of the processes \((R, Y)\), avoiding the knowledge of the transition density of \(Y\), which may be very complicated if known at all.

The second part of this section studies the existence of solutions to the ergodic Bellman equation in (22). Theorems 13, 18, and Proposition 15 below adapt the results of Kaise and Sheu (2006) to the present setting, and under some extra conditions in addition to the assumptions of Theorems 7 and 9. Their main message is that the quasi-linear PDE generally admits only one candidate for long run optimality. As shown with examples in Section 4, this candidate may or may not be long-run optimal.
2.1. Main Results. 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 \( y \).

**Theorem 7.** In addition to Assumptions 1 and 2, assume that:

i) \( v \in C^2(E, \mathbb{R}) \) and \( \lambda \in \mathbb{R} \) solve the ergodic HJB equation (cf. Section 3.2):

\[
pr - \frac{q}{2} \mu' \Sigma^{-1} \mu + \frac{1}{2} \nabla v' (A - q \Upsilon \Sigma^{-1} \Upsilon) \nabla v + \nabla v' (b - q \Upsilon \Sigma^{-1} \mu) + \frac{1}{2} \text{tr} (AD^2 v) = \lambda
\]

ii) there is a unique solution \( \left( \tilde{P}^y, y \right) \) to the martingale problem for

\[
\hat{L} = \frac{1}{2} \sum_{i,j=1}^{n+k} \tilde{A}^{i,j}(x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{n+k} \tilde{b}^i(x) \frac{\partial}{\partial x_i} \quad \hat{b} = \left( b - q \Upsilon \Sigma^{-1} \mu + (A - q \Upsilon \Sigma^{-1} \Upsilon) \nabla v \right)
\]

where \( \tilde{A} \) is as in (6).

Then, the pair \((\pi, \eta)\) given by:

\[
\pi = \frac{1}{1-p} \Sigma^{-1} (\mu + \Upsilon \nabla v), \quad \eta = \nabla v
\]

satisfies the equalities:

\[
E_P^y \left[ (X_T^\pi)^p \right] = e^{\lambda T + v(y)} E_P^y \left[ e^{-v(\Upsilon_T)} \right]
\]

\[
E_P^y \left[ (M_T^\eta)^{1-p} \right] = e^{\lambda T + v(y)} E_P^y \left[ e^{-\frac{1}{1-p} v(\Upsilon_T)} \right]^{1-p}
\]

**Remark 8.** Equations (25) and (26) provide lower and upper bounds on finite-horizon expected utility. Indeed, the duality inequality (14) yields:

\[
\frac{1}{p} e^{\lambda T + v(y)} E_P^y \left[ e^{-v(\Upsilon_T)} \right] = \frac{1}{p} E_P^y \left[ (X_T^\pi)^p \right] \leq u_T(y)
\]

\[
\leq \frac{1}{p} E_P^y \left[ (M_T^\eta)^{1-p} \right] = \frac{1}{p} e^{\lambda T + v(y)} E_P^y \left[ e^{-\frac{1}{1-p} v(\Upsilon_T)} \right]^{1-p}
\]

Combining (25) and (26) with (19) yields the central quantitative implication: an upper bound on the certainty equivalent loss:

\[
l_T \leq \frac{1}{p} \left( \frac{1}{T} \log E_P^y \left[ e^{-\frac{1}{1-p} v(\Upsilon_T)} \right]^{1-p} - \frac{1}{T} \log E_P^y \left[ e^{-v(\Upsilon_T)} \right] \right)
\]

Theorem 7 now reduces the long-run optimality (Definition 6) of \((\pi, \eta)\) to the condition that the right-hand side in (27) converges to zero. Theorem 9 below provides a criterion that covers most applications, and Proposition 25 below shows a model in which this criterion is sharp, in that it holds for all the parameter values for which long-run optimality holds.

**Theorem 9.** If, in addition to the assumptions of Theorem 7:

i) the random variables \((Y_t)_{t \geq 0}\) are \( \tilde{P}^y \)-tight\(^8\) in \( E \) for each \( y \in E \);

---

\(^8\)Recall that a family of \( E \)-valued random variables \((X_t)_{t \geq 0}\) is \( P \)-tight in \( E \) if the induced measures \((P \circ X_t^{-1})_{t \geq 0}\) form a tight family in \( M_1(E) \), the space of Borel measures on \( E \). Thus, \((X_t)_{t \geq 0}\) is \( P \)-tight in \( E \) if for each \( \varepsilon > 0 \) there exists a compact \( K_\varepsilon \subset E \) such that \( \sup_{t \geq 0} P(X_t \notin K_\varepsilon) \leq \varepsilon \).
ii) \( \sup_{y \in E} F(y) < +\infty \), where \( F \in C(E, \mathbb{R}) \) is defined as:

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

Then the pair \((\pi, \eta)\) in (24) is long-run optimal.

Section 4 shows how to check conditions i) and ii) in typical classes of models.

**Remark 10.** A sufficient condition for i) above to hold is that there exist a non-negative \( \psi \in C(E, \mathbb{R}) \) such that, for each \( n \), the level set \( K_n = \{ y \in E : \psi(y) \leq n \} \) is compact, and that \( \bar{M} = \sup_{t \geq 0} \mathbb{E}_t^\mu \left[ \psi(Y_t) \right] < \infty \). If such a \( \psi \) exists, then Markov’s inequality implies that, for each \( n \)

\[
\sup_{t \geq 0} \mathbb{P}^\psi(Y_t \in K_n) \leq \frac{1}{n} \sup_{t \geq 0} \mathbb{E}_t^\mu \left[ \psi(Y_t) \right] = \frac{\bar{M}}{n}
\]

Thus, \( \hat{P}^\psi \) tightness in \( E \) follows.

2.2. **Solutions to the Ergodic Bellman Equation.** This subsection provides conditions for the existence of a solution pair \((v, \lambda)\) to (22), such that the tightness condition in Theorem 9 holds. These results are obtained adapting the arguments in Kaise and Sheu (2006) to the present setting. Define \( \Lambda \) as the set of \( \lambda \in \mathbb{R} \) for which a solution \( v \) to (22) exists:

\[
(29) \quad \Lambda = \{ \lambda \in \mathbb{R} \mid \exists v \in C^{2,\gamma}(E, \mathbb{R}) \text{ solving } (22) \}
\]

\( \Lambda \) depends both on the region \( E \) and on the coefficients in the PDE (22). The foregoing results require the following assumption on the region \( E \), which holds in virtually all models in the literature:

**Assumption 11.** There exist \( y_0 \in E \), and a sequence of bounded open subsets \( E_n \subseteq E \), star-shaped\(^9\) with respect to \( y_0 \) and with a \( C^{2,\gamma} \) boundary, and strictly increasing to \( E \), in that \( E = \bigcup_{n=1}^\infty E_n \) and \( E_n \cap (E \setminus E_{n+1}) = \emptyset \).

This assumption is satisfied by any convex set \( E \) for which there is a convex function \( \psi \in C^{2,\gamma}(E, \mathbb{R}) \) such that \( \psi(y) \uparrow \infty \) as \( y \to \partial E \). In this case, it suffices to set \( E_n = \{ y \in E : \psi(y) < n \} \).

The next assumption requires that the potential \( pr - \frac{q}{2} \mu' \Sigma^{-1} \mu \) is bounded from above. When it does not hold, solutions to (22) may not exist (Pinsky, 1995, Chapter 4.5).

**Assumption 12.** \( \sup_{y \in E} (pr(y) - \frac{q}{2} \mu(y)' \Sigma(y)^{-1} \mu(y)) < \infty \).

Note that Assumption 12 always holds if \( p < 0 \), and the interest rate \( r \) is bounded from below, which is a typical situation in financial models. With Assumptions 1, 2 and 11, denote by \((R, Y)\) the coordinate process of the solution \((P^y)_{y \in E}\) of the martingale problem corresponding to the operator \( L \) from (6). Regarding the state variable \( Y \), the statement of existence results requires a few basic definitions in ergodic theory (see Pinsky (1995); Pinchover (1992) for more details).

\( Y \) is **transient** if \( P^y(Y \in E \setminus E_n \text{ for all } n \geq N(\omega)) = 1 \) for all \( n \geq 1 \) and \( y \in E \). \( Y \) is **recurrent** if \( P^x(\tau(\varepsilon, y) < \infty) = 1 \) for all \( x, y \in E \) and \( \varepsilon > 0 \), where \( \tau(\varepsilon, y) = \inf \{ t \geq 0 \mid |Y_t - y| \leq \varepsilon \} \). If \( Y \)

\(^9\)Recall that \( F \subseteq \mathbb{R}^k \) is **star shaped** for some \( x_0 \in F \) if for each \( x \in F \) the segment \( \{\alpha x_0 + (1 - \alpha)x; \ 0 \leq \alpha \leq 1\} \) is contained within \( F \). A convex set is star shaped with respect to any of its points.
is recurrent, there exists some $\tilde{\phi} > 0$ such that $
abla L \tilde{\phi} = 0$, where $L$ is the formal adjoint to $L$. $Y$ is positive recurrent, or ergodic, if $\int_E \tilde{\phi} dy < \infty$, and null recurrent otherwise. If $Y$ is positive recurrent and $\tilde{\phi}$ is normalized to be a probability density, then for all $y \in E$ and $f \in L^1(E, \tilde{\phi})$

$$\lim_{T \to \infty} E^y_P[f(Y_T)] = \int_E f \tilde{\phi} dx$$

If $Y$ is ergodic, then (30) implies that $Y$ is $P^y$-tight in $E$ for each $y \in E$, and for all $y \in E$ and $(t_n)_{n \geq 1} \uparrow \infty$, the measures $(P^y \circ Y_{t_n})^{-1}$ weakly converge to the measure with density $\tilde{\phi}$, which does not depend upon the starting point $y \in E$. With these definitions and results, the following theorem shows that there exists only one possible pair $(\lambda, v)$, solving (22), which can lead to long-run optimality:

**Theorem 13.** Let Assumptions 1, 11 and 12 hold. Then there exists $\lambda_c \in \mathbb{R}$ such that $\Lambda = [\lambda_c, \infty)$. Furthermore, $Y$ is $(\hat{P}^y)_{y \in E}$-transient for any $\hat{P}$ corresponding to a solution $(v, \lambda)$ of (22) with $\lambda > \lambda_c$.

Clearly, $(\hat{P}^y)_{y \in E}$ transience and $(\hat{P}^y)_{y \in E}$-tightness in $E$ are incompatible with one another. Furthermore, since ergodicity implies tightness, the question is whether the pair $(\lambda_c, v_c)$ makes $Y$ ergodic under $(\hat{P}^y)_{y \in E}$. The following results give conditions under which this is indeed the case. The first proposition is valid for a single state variable and constant correlations $\rho$. Then, (22) linearizes under a power transformation, and classical tests for transience and recurrence of linear diffusions apply. The second proposition considers the general multidimensional case, but under a stronger restriction on the potential term.

### 2.2.1. One State, Constant Correlations

**Assumption 14.** Let Assumptions 1 and 12 hold. Further, assume that:

1. $E = (\alpha, \beta)$ with $-\infty \leq \alpha < \beta \leq \infty$.
2. $\rho' \rho = Y'\Sigma^{-1} Y / A$ is constant.

Note that Assumption 11 is always satisfied for a single state variable. Set

$$\delta = \frac{1}{1 - q\rho' \rho}$$

The change of variable $\phi = \exp(v/\delta)$, essentially equivalent to the power transformation of Zariphopoulou (2001), reduces the quasi-linear ODE in (22) to the linear ODE:

$$\frac{A}{2} \ddot{\phi} + (b - qY'\Sigma^{-1} \mu) \dot{\phi} + \frac{1}{\delta} (V - \lambda) \phi = 0$$

Let $\lambda \in \Lambda$ and let $\phi \in C^{2, \gamma}(E, \mathbb{R})$ with $\phi > 0$ be a solution to (32) obtained by $\phi = \exp(v/\delta)$. Under $\hat{P}^y$, $Y$ has the dynamics

$$dY_t = \left(b - qY'\Sigma^{-1} \mu + A \frac{\dot{\phi}}{\phi}\right) dt + adW_t$$

Using Feller's test for explosions the following proposition (Pinsky, 1995, Corollary 5.1.11) gives sufficient conditions for $Y$ to be $P^y$-tight in $E$ for the candidate optimal pair $(\lambda_c, \phi_c)$.
PROPOSITION 15. Let Assumption 14 hold, let \((\lambda_c, v_c)\) be as in Theorem 13, and let \(\phi_c = \exp(v_c/\delta)\). Denote by:

\[
m_{\nu}(y) = \frac{1}{A(y)} \exp \left( \int_{y_0}^{y} \frac{2(b - q Y' \Sigma^{-1} \mu)}{\Lambda(z)} \, dz \right)
\]

where \(y_0 \in (\alpha, \beta)\). Then, the family of random variables \((Y_t)_{t \geq 0}\) is \(\hat{P}^y\)-tight in \(E\) if and only if:

\[
\int_{\alpha}^{y_0} \frac{1}{\phi_c^2 \Lambda_\nu} \, dy = \int_{y_0}^{\beta} \frac{1}{\phi_c^2 \Lambda_\nu} \, dy = \infty \quad \text{and} \quad \int_{\alpha}^{\beta} \phi_c^2 m_\nu \, dy < \infty
\]

2.2.2. The General Case. This subsection treats the general case of \(k\) state variables under the following assumption:

ASSUMPTION 16. There exists a function \(w \in C^{2,\gamma}(E, \mathbb{R})\) such that:

\[
\limsup_{n \uparrow \infty} E_{\nu} |w_{\nu} - \frac{q}{2} \mu' \Sigma^{-1} \mu' + \frac{1}{2} \nabla w' (A - q Y' \Sigma^{-1} Y) \nabla w + \nabla w' (b - q Y' \Sigma^{-1} \mu) + \frac{1}{2} \text{tr} (ADw^2) | = -\infty
\]

REMARK 17. For general regions \(E\), condition (36) plays a similar role as condition (A3) in Kaise and Sheu (2006). Assumption 16 is satisfied, for example, when

\[
\limsup_{n \uparrow \infty} E_{\nu} |w_{\nu} - \frac{q}{2} \mu' \Sigma^{-1} \mu| = -\infty
\]

In this case, \(w \equiv 1\) satisfies (36).

In this setting, the main existence criterion is the following:

THEOREM 18. Let Assumptions 1, 11, 12, and 16 hold, and let \((\lambda_c, v_c)\) be as in Theorem 13. Then \(v_c\) is unique up to an additive constant, and \((Y_t)_{t \geq 0}\) is \(\hat{P}^y\)-tight in \(E\) for all \(y \in E\).

Proposition 25 in Section 4 below shows that long run optimality may still fail, even when the tightness condition is satisfied for the pair \((\lambda_c, v_c)\). The reason is that, even if \(Y\) is ergodic under \((P^y)_{y \in E}\) with invariant density \(\tilde{c}\), the ergodic property in (30) may not hold for (25) and (26), because the functions

\[
\exp(-v(y)), \quad \exp(-(1-p)^{-1}v(y))
\]

therein may not be in \(L^1(E, \tilde{c})\). Thus, long run optimality requires additional assumptions, such as (28) in Theorem 9.

If the functions in (38) are in \(L^1(E, \tilde{c})\), then (30) yields additional information about the speed at which the certainty equivalent loss \(l_T\) converges to zero in the limit of a long horizon. The following proposition provides such a result, in the case of a single state variable. Recall that risk aversion is \(\gamma = 1 - p = 1/(1 - q)\). The main message is that long-run optimality may fail only for \(i)\) high risk aversion and highly incomplete market or \(ii)\) low risk aversion and near complete market. In particular, for \(-1 < q \leq 1/2\) (i.e risk aversion within 1/2 and 2), long-run optimality holds for any level of incompleteness. In addition, for \(q > 1/2\) (resp. \(q < -1\), long-run optimality holds if \(\rho' \rho < 1/2\) (resp. \(\rho' \rho > 1/2\)).

Section 4 takes up this issue in specific models, obtaining necessary and sufficient conditions. By contrast, the following sufficient condition holds under general assumptions, regardless of the model considered.
Proposition 19. Let Assumption 14 hold, and assume that $m_\nu$ in (34) satisfies $\int_E m_\nu dy < \infty$. If $(\lambda_c, v_c)$ are such that $Y$ is $(\hat{P}^\nu)_{y \in E}$-ergodic, then long-run optimality holds if

$$\rho' \rho \in \begin{cases} [0, \frac{1}{2q}] & \text{for } \frac{1}{2} < q < 1 \\ [0, 1] & \text{for } -1 \leq q \leq \frac{1}{2}, q \neq 0 \\ [1 + \frac{q}{2q}, 1] & \text{for } q < -1 \end{cases}$$

In such cases, there exists a constant $K > 0$ such that the certainty equivalent loss $l_T$ satisfies

$$0 \leq \limsup_{T \to \infty} Tl_T \leq K$$

The next corollary states an important special case, which does not even require the knowledge of the principal eigenfunction $\phi$, since $m_\nu$ only depends on the model parameters.

Corollary 20. Under Assumption 14, if $\int_E m_\nu dy < \infty$ and (37) are satisfied, then long run optimality holds for $q$ and $\rho' \rho$ satisfying (39).

3. Implications and Ramiﬁcations.

3.1. The Myopic Probability. The bounds (25) and (26) in Theorem 7 and assumption i) in Theorem 9 depend on the equivalent probability $\hat{P}^\nu$, which plays a pivotal role in long-run analysis. In general, $\hat{P}^\nu$ is neither the physical probability $P^\nu$, nor a risk-neutral probability. Instead, its interpretation becomes clear from its dynamics, which is (for $\hat{P}$-Brownian motions $\hat{Z}, \hat{W}$):

$$dR_t = \frac{1}{1-q} (\mu + Y \nabla v) dt + \sigma d\hat{Z}_t$$

$$dY_t = (b - q Y \Sigma^{-1} \mu + (A - q Y \Sigma^{-1} Y) \nabla v) dt + ad\hat{W}_t$$

Compare the original model, with price dynamics under $P^\nu$ and power utility $x^\nu/p$, to the auxiliary model under $\hat{P}^\nu$ with logarithmic utility. The long-run optimal portfolio in the two models coincide. The first one is simply in (24), while the second one follows from the usual formula $\pi = \Sigma^{-1} \hat{\mu}$, where $\hat{\mu} = \frac{1}{1-q} (\mu + Y \nabla v)$ are the expected returns under $\hat{P}^\nu$. Thus, a long-horizon, power-utility investor under the probability $P^\nu$ behaves exactly as a myopic (or logarithmic) investor under $\hat{P}^\nu$.

This observation shows that $\hat{P}^\nu$ corresponds to the long-horizon limit of the probability $P$ considered by Kramkov and Sirbu (2006a; 2006b; 2007) in finite horizon, in the context of sensitivity analysis pricing of option prices. Černý and Kallsen (2007) study mean-variance hedging for semimartingales, and obtain optimal strategies in terms of the predictable characteristics of asset prices under an opportunity neutral probability $P^*$, which is similar in spirit to $\hat{P}^\nu$, in that it reduces the mean-variance objective to a logarithmic utility objective.

3.2. Connections with 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 Theorems 7, 9 and the stochastic control approach becomes clear by comparing equation (22) to the Hamilton-Jacobi-Bellman (HJB) equations of the utility maximization problem (10). Its value function $u(x, y, t)$ depends on the current wealth $x$, the current state $y$, and time $t$. The homogeneity of power utility entails that $u(x, y, t) = x^\nu e^{u(y, t)}/p$, thereby removing wealth from the reduced value function $w$. The corresponding HJB equation becomes (see for example Pham (2002))

$$-\frac{\partial w}{\partial t} = pr - \frac{q}{2} \mu' \Sigma^{-1} \mu + \frac{1}{2} \nabla w' (A - q Y \Sigma^{-1} Y) \nabla w + \nabla w' (b - q Y \Sigma^{-1} \mu) + \frac{1}{2} \text{tr} (AD^2 w)$$
with the terminal condition \( w(y, T) = 0 \). Instead, the main PDE (22) is:

\[
\lambda = pr - \frac{q}{2} \mu' \Sigma^{-1} \mu + \frac{1}{2} \nabla v' (A - q \Upsilon' \Sigma^{-1} \Upsilon) \nabla v + \nabla v' (b - q \Upsilon' \Sigma^{-1} \Upsilon) + \frac{1}{2} \text{tr}(AD^2 v)
\]

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 \( \lambda \) is also unknown. Indeed, the former equation reduces to the latter under the restriction:

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

This restriction gains 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).

A special case of equation (22) 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:

\[
(42) \max \liminf_{T \to \infty} \frac{1}{T} \log E[(X^\pi_T)^p]
\]

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 (42). 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. The results in this paper, which apply to general nonlinear models, shed new light on this literature by characterizing finite-horizon performance. For example, Proposition 25 below relaxes the restriction of \(|p| \) small to a necessary and sufficient condition, and explains the economic intuition behind it.

3.3. Long-Run \( q \)-optimal measure. For each value of the risk-aversion parameter \( 1 - p \), the risk premia \( \eta \) in (24) deliver a pricing rule for derivatives involving the partially unhedgeable state variable \( Y \). The martingale measure \( Q^\eta \) corresponding to the risk premia \( \eta \) is a long-run version of the minimax martingale measure of He and Pearson (1991), called \( q \)-optimal measure by Hobson (2004) and Henderson (2005). Its formal dynamics is:

\[
(43)
\begin{align*}
\{ dR_t &= \sigma d\tilde{Z}_t \\
\dot{Y}_t &= (b - \Upsilon' \Sigma^{-1} \mu + (A - \Upsilon' \Sigma^{-1} \Upsilon) \nabla v) dt + a d\tilde{W}_t
\end{align*}
\]

for some Brownian Motions \( \tilde{Z} \) and \( \tilde{W} \). Since this dynamics is distinct from the one under \( P^y \) and \( \hat{P}^y \), in general it is necessary to check its well-posedness, in the form of Assumption 2.

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 \( P^y \). 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 \( (A - \Upsilon' \Sigma^{-1} \Upsilon) \nabla v \) accounts for preferences, which enter the equation through \( v \).

3.4. Complete and Fully Incomplete as Duals. The formulas in (24) highlight the symmetric aspects of complete markets, where \( A = \Upsilon' \Sigma^{-1} \Upsilon \) identically, and fully incomplete markets, where \( \Upsilon = 0 \). In a complete market the pricing problem is trivial, as the dynamics in (43) becomes independent of the preference parameter \( p \):
However, the investment problem is nontrivial, because the optimal portfolio includes a component that perfectly hedges the state variables.

Conversely, in a fully incomplete market myopic portfolios are optimal, because all portfolios evolve orthogonally to state variables. However, a latent hedging motive remains present, and generates nonzero risk premia for state variables, which have a potential 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 $\mu$ and $\sigma$ remove the intertemporal hedging motive entirely, making both problems trivial.

The quasi-linear ODE in (22) becomes linear under a power transformation in both the complete and fully incomplete cases, as in the one state variable case discussed in Section 2.2.1. In the complete case the power transformation $\phi = e^{(1-q)v}$ leads to the linear equation

$$\frac{1}{2} \text{tr}(AD^2\phi) + \nabla \phi' (b - qY\Sigma^{-1}\mu) + (1 - q) \left( pr - \frac{q}{2} \mu'\Sigma^{-1}\mu - \lambda \right) \phi = 0$$

In the fully incomplete case the power transformation $\phi = e^v$ leads to the linear equation

$$\frac{1}{2} \text{tr}(AD^2\phi) + \nabla \phi' b + \left( pr - \frac{q}{2} \mu'\Sigma^{-1}\mu - \lambda \right) \phi = 0$$

The criticality theory of Pinsky (1995) applies to these cases under multivariate restrictions similar to those given in Assumption 14, with $\delta = \frac{1}{1-q}$ in the complete case and $\delta = 1$ in the fully incomplete case. Furthermore, the multivariate results of Theorem 18 apply as long as (36) can be verified.

3.5. Long-Run Decomposition. The bounds (25) and (26) 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 (2009). For a multiplicative functional $N$ of a Markov process $Y$, they propose the decomposition:

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

where $\rho$ and $\varphi$ are respectively the principal eigenvalue and eigenfunction of the infinitesimal generator of $Y$, and $\hat{N}_t$ is a martingale. The bounds (25) and (26) yield similar expressions for terminal utilities and their dual counterparts:

$$(X_T^\pi)^p = e^{\lambda T} \frac{e^{v(y)}}{e^{v(Y_T)}} d\hat{P}^y$$

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

These decompositions are precisely of the form in (47), with the difference that on the dual side the transient components are powers of $e^v$, as opposed to $e^v$ itself. Note also that the operator in (22) is not the generator of $Y$ under either $P^y$ or $\hat{P}^y$, since it is non-linear, has non-zero potential and has a different drift. Further, the interpretation of $e^{\lambda T}$ as a long-run component hinges on the condition that the $\hat{P}^y$-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.

3.6. Large Deviations. Theorems 7 and 9 are closely related to the results of Donsker and Varadhan (1975; 1976; 1983) on Large Deviations of occupation times for diffusions. Though the results also hold in the multi-dimensional case of $k > 1$ state variables, the following discussion considers a single state variable for simplicity of notation.
Let $E = (\alpha, \beta)$ for $-\infty \leq \alpha < \beta \leq \infty$ and consider a diffusion $Y$ with generator $L^Y$ from (7) (with $k = 1$), assuming that the coefficients $A$ and $b$ are such that $Y$ is positive recurrent under $(P^y)_{y \in E}$. Let $m$ be the invariant measure, which has a density by (30). With a slight abuse of notation, let $m(dy) = m(y)dy$.

Denote by $M_1(\alpha, \beta)$ the space of Borel probability measures on $(\alpha, \beta)$. Under certain conditions on $Y$, Donsker and Varadhan show that, for all continuous bounded functions $V : (\alpha, \beta) \mapsto \mathbb{R}$ and all $y \in (\alpha, \beta)$:

$$\lim_{T \to \infty} \frac{1}{T} \log E_P^y \left[ \exp \left( \int_0^T V(Y_t)dt \right) \right] = \sup_{\mu \in M_1((\alpha, \beta))} \left( \int_\alpha^\beta Vd\mu - I(\mu) \right)$$

For $\mu \in M_1(\alpha, \beta)$ absolutely continuous with respect to $m$ (and hence Lebesgue measure), and with density $\mu(y)$ such that $\psi(y)^2 \equiv \mu(y)/m(y)$ satisfies certain regularity and decay conditions, the rate function $I(\mu)$ reduces to:

$$I(\mu) = \frac{1}{2} \int_\alpha^\beta A(\psi)^2 mdy$$

Using this representation for the rate function, 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 $\pi(Y_t)$ for some function $\pi : E \mapsto \mathbb{R}^n$:

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

Define $P_\pi^y$ by setting $\frac{dP_\pi^y}{dP^y}$ equal to the stochastic exponential in the last term of this equation. It follows that:

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

Assuming they may be applied under $P_\pi^y$, the Donsker and Varadhan asymptotics (48) yield:

$$\lim_{T \to \infty} \frac{1}{T} \log E_P^y [(X_T^\pi)^p]$$

$$= \sup_{\psi \in L_1^2(m_\pi)} \int_\alpha^\beta \left( \left( pr + p\pi' \mu + \frac{1}{2}p(p-1)\pi' \Sigma \pi \right) \psi^2 - \frac{1}{2} A(\psi)^2 \right) m_\pi dy$$

where $m_\pi$ is the invariant density of $Y$ under $P_\pi^y$ and $L_1^2(m_\pi)$ is the unit disc in $L^2(m_\pi)$. In a similar manner to (34) in Proposition 15, $m_\pi$ admits the formula

$$m_\pi(y) = \frac{1}{A(y)} \exp \left( \int_{y_0}^y \frac{2(b + pY'\pi)(z)}{A(z)} dz \right)$$

Here, $y_0$ is some interior point in $E$. To make the dependence on the portfolio $\pi$ explicit, the change of variable $\psi^2 m_\pi = \phi^2 m_\nu$ yields ($m_\nu$ is defined in (34)):

$$\frac{\dot{\psi}}{\psi} = \frac{\dot{\phi}}{\phi} + \frac{\dot{m}_\nu}{2 m_\nu} - \frac{\dot{m}_\pi}{2 m_\pi} = \frac{\dot{\phi}}{\phi} - A^{-1} (qY' \Sigma^{-1} \mu + pY' \pi)$$

Substituting (50) into (49), the utility growth rate becomes

$$\sup_{\phi \in L_1^2(m_\nu)} \int_\alpha^\beta \left( \pi' A \pi + \pi' B + C \right) \phi^2 m_\nu dy$$
where

\[ A = \frac{1}{2} p (p - 1) \left( \Sigma - q \psi A^{-1} \psi' \right) \]

\[ B = p \left( (1_n - q \psi A^{-1} \psi') \mu + \psi \psi' \right) \]

\[ C = pr - \frac{1}{2} A \left( \frac{\psi}{\phi} \right)^2 + q \psi' \Sigma^{-1} \psi \left( \frac{\psi}{\phi} \right) - \frac{1}{2} q^2 \psi' \Sigma^{-1} \psi A^{-1} \psi' \Sigma^{-1} \mu \]

The integrand is a quadratic function of \( \pi \), and achieves its optimum at:

\[ \pi = \frac{1}{1 - p} \Sigma^{-1} \left( \mu + \delta \psi \psi' \right) \]

Thus, substituting (52) into (51), the utility growth rate reduces to:

\[ \sup_{\phi \in L^2_{\alpha}(m_\psi)} \int_{\beta}^{\alpha} \left( \left( pr - \frac{q}{2} \psi' \Sigma^{-1} \mu \right) \phi^2 - \frac{\delta}{2} A \left( \phi \right)^2 \right) m_\psi dy \]

A similar reasoning on stochastic discount factors delivers the candidate long-run risk premia. The Euler-Lagrange equation associated to (53) is the ODE in (32). Thus, Large Deviations arguments act as a guide for producing the candidate long-run optimal policies.

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.

4. Applications. This section applies the main results to two models, assuming that the investor is more risk averse than the log investor \( (p < 0) \). In the first model, the state variables follow a multivariate Ornstein-Uhlenbeck process, which drives the drift of the return process. Under general conditions, this model admits a unique solution \( v \) to (22), leading to \( \hat{P}_y \) tightness. For a single state variable, long-run optimality is characterized in terms of precise parameter restrictions.

In the second model, interest rates, drifts and volatilities are stochastic. Each of these quantities is affine in a single common state variable, which follows a Feller diffusion. Although this model does not belong to the affine class, the long-run optimal portfolios and risk premia have very simple expressions.

The parametric restrictions required by long-run optimality lead under each single variate model to the same economic interpretation. Long-run optimality does not hold at the conjunction of three extreme situations: i) high covariation between risk premia and state variables, ii) nearly complete markets, and iii) high risk-aversion. To understand this phenomenon, recall that long-run optimality means that a time-homogenous strategy is approximately optimal on a long time interval. Thus, the sub-optimality of the long-run strategy in the latest part of the interval must lead to a small utility loss. Since the myopic component of the optimal finite-horizon portfolio is time-homogenous, any loss in utility is attributed to the intertemporal hedging component. All of the three extreme situations mentioned above concur to amplify the intertemporal hedging component. First, the covariation of risk premia is proportional to the hedging portfolios \( \Sigma^{-1} \psi \). Second, intertemporal hedging is more attractive in a nearly complete market, where state variables are almost replicable. Third, intertemporal hedging is higher for more risk-averse investors, who reduce long-term risk at the expense of short-term return.
4.1. Linear Diffusion. This is the most common multivariate model, with constant covariance matrices $\Sigma$, $\Upsilon$, $A$, and drifts $r, \mu, b$ that are affine functions of the state variable. The dynamics is:

$$
\begin{align*}
    dR_t &= (\mu_0 + \mu_1 Y_t) dt + \sigma dZ_t \\
    dY_t &= -b Y_t dt + adW_t \\
    d(R, Y)_t &= \rho dt \\
    r(Y_t) &= r_0 + r'_1 Y_t
\end{align*}
$$

(54)

where $\mu_0 \in \mathbb{R}^n$, $\mu_1 \in \mathbb{R}^{n \times k}$, $\sigma \in \mathbb{R}^{n \times n}$, $b \in \mathbb{R}^{k \times n}$, $a \in \mathbb{R}^{k \times k}$, $\rho \in \mathbb{R}^{n \times k}$, $r_0 \in \mathbb{R}$ and $r_1 \in \mathbb{R}$. Under this model, $E = \mathbb{R}^k$, and state variables follow a multivariate Ornstein-Uhlenbeck process. This setting is considered in most of the literature in risk sensitive control mentioned in the introduction, and it is also implicit in the use of vector autoregressions in the Econometrics literature. The coefficients in (54) satisfy the following

**Assumption 21.** All four matrices $\Sigma = \sigma \sigma'$, $\mu'_1 \mu_1$, $b + b'$ and $a$ are positive definite.

Assumption 21 implies that condition (37) holds for $p < 0$. Hence, by Theorem 18 (or by the results in Kaise and Sheu (2006)), there is a unique pair $(\lambda, v)$ solving (22) such that for each $y \in E$, $(Y_t)_{t \geq 0}$ is $\bar{P}^y$-tight in $E$. The next theorem shows that $v$ is in fact quadratic:

**Theorem 22.** Let Assumption 21 hold for the model in (54). For $p < 0$ equation (22) admits a unique solution $v$ such that for each $y \in \mathbb{R}^k$, $(Y_t)_{t \geq 0}$ is $\bar{P}^y$-tight in $\mathbb{R}^k$. The solution is of the form $v(y) = v'_0 y - \frac{1}{2} y' v_1 y$, where $v_0 \in \mathbb{R}^k$, $v_1 \in \mathbb{R}^{k \times k}$ are symmetric, and satisfy the algebraic equations:

$$
\begin{align}
    (v_1 (A - q \Upsilon' \Sigma^{-1} \Upsilon) + (b + q \Upsilon' \Sigma^{-1} \mu'_1) v_0 - pr_1 + q(\mu'_1 - v_1 \Upsilon') \Sigma^{-1} \mu_0 = 0 \\
    v_1 (A - q \Upsilon' \Sigma^{-1} \Upsilon) v_1 + v_1 (b + q \Upsilon' \Sigma^{-1} \mu_1) + (b + q \Upsilon' \Sigma^{-1} \mu_1)' v_1 - q \mu'_1 \Sigma^{-1} \mu_1 = 0
\end{align}
$$

(55, 56)

The corresponding utility growth rate equals to:

$$
\lambda = pr_0 - \frac{q}{2} \mu'_0 \Sigma^{-1} \mu_0 + \frac{1}{2} v'_0 (A - q \Upsilon' \Sigma^{-1} \Upsilon) v_0 - q v'_0 \Upsilon' \Sigma^{-1} \mu_0 - \frac{1}{2} \text{tr}(A v_1)
$$

Equation (56) 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 numerical techniques for obtaining the solution are available (see (Abou-Kandil et al., 2003, Chapter 2)). Once the matrix $v_1$ is known, the linear equation (55) yields a unique solution for $v_0$, and $\lambda$ is quadratic in $v_0$ and linear in $v_1$.

Observe that Theorems 22 and 7 characterize the candidate pair $(\pi, \eta)$, and allow to find the finite-horizon bounds, but do not address long-run optimality. This stronger property in fact holds only under parameter restrictions, and is now studied in detail for a single state. In this case, the linear diffusion yields an extension of the models in Kim and Omberg (1996) and Wachter (2002):

$$
\begin{align*}
    dR_t &= (\sigma v_0 + b \sigma v_1 Y_t) dt + \sigma dZ_t \\
    dY_t &= -b Y_t dt + dW_t \\
    d(R, Y)_t &= \rho dt \\
    r(Y_t) &= r_0
\end{align*}
$$

(58)

The constants are the same as in (54), except that here $\mu_0 = \sigma v_0$, $\mu_1 = b \sigma v_1$ where $v_0, v_1 \in \mathbb{R}^n$, for ease of notation. Note that $a = 1$ and $r_1 = 0$. The Riccati equation from (56) is

$$
\delta^{-1} v_1^2 + 2 b (1 + q \sigma' \nu_1) v_1 - q \rho^2 \nu_1^2 v_1 = 0
$$
for \( \delta = 1/(1 - q\rho') \). Under Assumption 21, for \( p < 0 \), the solution \( v_1, v_0, \lambda \) from Theorem 22 is

\begin{align}
 v_1 &= \delta b \left( \sqrt{\Theta} - (1 + q\rho' \nu_1) \right) \\
 v_0 &= q\delta\rho' \nu_0 - \frac{1}{\sqrt{\Theta}} (qv'_1 \nu_0 + q\delta\rho' \nu_0 (1 + q\rho' \nu_1)) \\
 \lambda &= pr_0 - \frac{1}{2} qv'_0 \nu_0 + \frac{1}{2} \delta^{-1} v_0^2 - qv_0\rho' \nu_0 - \frac{1}{2} v_1
\end{align}

where

\( \Theta = (1 + q\rho' \nu_1)^2 + \delta^{-1} qv'_1 \nu_1 \)

The candidate long-run optimal pair \((\pi, \eta)\) is affine in the state variable:

\begin{align}
\pi(y) &= \frac{1}{1 - p} \Sigma^{-1} (\mu(y) + v_0\sigma\rho - v_1 y \sigma\rho) \\
\eta(y) &= v_0 - v_1 y
\end{align}

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

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

This pair \((\pi, \eta)\) is indeed long-run optimal, but only under a parameter restriction.

**Proposition 23.** Let Assumption 21 hold, and let \( p < 0 \), \((\pi, \eta)\) from (63) is long-run optimal if

\( (1 - 2q\rho') \sqrt{\Theta} + (1 + q\rho' \nu_1) > 0 \)

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 (65) simplifies as follows:

**Corollary 24.** Let Assumption 21 hold. For \( p < 0 \) 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 \( \kappa \). For \( 1/4 < q\rho' \rho < 1 \) long-run optimality holds if:

\( \kappa < \frac{2}{4q\rho' \rho - 1} \)

Thus, long-run optimality requires a joint restriction on preferences \((q)\) and price dynamics \((\rho' \rho \) and \( \kappa \). First, since \( q\rho' \rho < 1 \), long-run optimality always holds if \( \kappa < \frac{2}{4} \), that is if risk premia have low covariation with changes in state variables. 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. The next proposition studies this phenomenon in the case \( \kappa = 1 \), which corresponds to a continuous time version of the model of Summers (1986):
PROPOSITION 25. Let Assumption 21 hold. For $p < 0$ and $\kappa = 1$ from Corollary 24, long-run optimality holds if $qp' \rho < \frac{3}{4}$. If $qp' \rho \geq \frac{3}{4}$, long-run optimality fails. In particular:

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

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

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

4.1.1. Calibration. A calibration to 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). 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 (66) 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 the one of its myopic component. The plots show the estimates of the corresponding upper bounds in (19): the myopic component prevails in the short run, 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 percentage points for a risk-aversion of five.

4.2. Stochastic drifts, volatilities, and interest rates. 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 their drifts. Note that the model is neither affine nor quadratic (due to the presence of the term with $\nu_0$), and yet the long-run solution admits a simple expression.

\begin{align}
\left\{\begin{array}{l}
\frac{dR_t}{\sqrt{Y_t}} = (\sigma \nu_0 + \sigma \nu_1 Y_t) \, dt + \sqrt{Y_t} \sigma dZ_t \\
\frac{dY_t}{\sqrt{Y_t}} = b(\theta - Y_t) \, dt + a \sqrt{Y_t} dW_t \\
\frac{d(R, Y)_t}{\sqrt{Y_t}} = \rho dt \\
r(Y_t) = r_0 + r_1 Y_t
\end{array}\right. \\
(67)
\end{align}

where $\sigma \in \mathbb{R}^{n \times n}$; $\nu_0, \nu_1 \in \mathbb{R}^n$; $b, \theta, a \in \mathbb{R}$; $\rho \in \mathbb{R}^n$ and $r_0, r_1 \in \mathbb{R}$. The parameters satisfy the following:

ASSUMPTION 26. $b, \theta, a, r_1 \geq 0$, and $b \theta > \frac{1}{2} a^2$.

Assumption 26 ensures that the state variable $Y$ remains strictly positive, thereby satisfying Assumption 2 with $E = (0, \infty)$. Guessing a form of the solution $v(y) = v_0 \log y + v_1 y$, the main ODE (22) becomes:

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

Multiplying the above equation by $y$ and setting the constant, linear and quadratic terms to zero leads to four candidate solutions, corresponding to any combination of signs in the terms $\pm \sqrt{\Theta}$ and...
Fig 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 (19) setting \( \eta \) equal to the long-run optimal risk premium, and \( \pi \) equal to the long-run optimal portfolio (solid line) and to the myopic portfolio \( \frac{1}{\eta p} \Sigma^{-1} \mu \) (dashed line). Risk Aversion is equal to two \( (p = -1, \text{ top}) \) and to five \( (p = -4, \text{ bottom}) \).

\[
\pm \sqrt{\Lambda} \text{ below:}
\]

\[
v_1 = \frac{\delta}{a^2} \left( b + qa \rho' \nu_1 \pm \sqrt{\Theta} \right)
\]

\[
v_0 = \frac{\delta}{a^2} \left( - (c - qa \rho' \nu_0) \pm \sqrt{\Lambda} \right)
\]

\[
\lambda = pr_0 - qv_0' \nu_1 + \frac{a^2}{\delta} v_0 v_1 - v_0 (b + qa \rho' \nu_1) + v_1 (b \theta - qa \rho' \nu_0)
\]
with
\[ c = b\theta - \frac{1}{2}a^2 \]
\[ \Theta = (b + qap\nu_1)^2 + \frac{a^2}{\delta} (q\nu_1'\nu_1 - 2pr_1) \]
\[ \Lambda = (c - qap\nu_0)^2 + \frac{a^2}{\delta} q\nu_0\nu_0 \]

When \( p < 0, r_1 > 0 \)
\[ \Theta > (b + qap\nu_1)^2 > 0 \]
\[ \Lambda > (c - qap\nu_0)^2 > 0 \]

Under \( \hat{P} \), \( Y \) has again four possible dynamics:
\[ dy_t = \left( \frac{1}{2}a^2 \pm \sqrt{\Lambda} \pm \sqrt{\Theta} Y_t \right) dt + a\sqrt{Y_t}dW_t \]

The choice of \(-\sqrt{\Theta}\) and \(\sqrt{\Lambda}\) ensures that \( Y \) satisfies Assumption 2 under \( \hat{P}^y \) and is \( \hat{P}^y \)-tight in \((0, \infty)\) for each \( y \in (0, \infty) \). The latter statement follows by the positivity of \( \Theta, \Lambda \), see Pinsky (1995, Corollary 5.1.11) and the discussion immediately after Assumption 12. Thus, the candidate optimizer is:
\[ v_1 = \frac{\delta}{a^2} \left( b + qap\nu_1 - \sqrt{\Theta} \right) \]
\[ v_0 = \frac{\delta}{a^2} \left( - (c - qap\nu_0) + \sqrt{\Lambda} \right) \]
\[ \lambda = pr_0 - q\nu_0'\nu_1 + \frac{a^2}{\delta} v_0 v_1 - v_0 (b + qap\nu_1) + v_1 (b\theta - qap\nu_0) \]

The candidate long-run optimal policies \((\pi, \eta)\) are:
\[ \pi(y) = \frac{1}{1 - p} \Sigma^{-1} (\mu(y) + \sigma\rho a (v_0 + v_1 y)), \quad \eta(y) = \frac{v_0}{y} + v_1 \]

and the candidate long-run martingale measure is:
\[ \left\{ \begin{array}{l}
\sigma dZ_t \\
\sigma dt
\end{array} \right\}
\[ dy_t = (b(\theta - Y_t) - q\rho'\nu_0 + \rho'\nu_1 Y_t) + a^2 (1 - \rho^2) (v_0 + v_1 Y_t) dt + a\sqrt{Y_t}dW_t \]

Long-run optimality obtains under the following conditions:

**Proposition 27.** Let Assumption 26 hold. For \( p < 0 \), long-run optimality holds if:
\[ (1 - 2q\rho') \sqrt{\Lambda} + (c - qap\nu_0) > 0 \]
\[ (1 - 2q\rho') \sqrt{\Theta} + (b + qap\nu_1) > 0 \]

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 \((a\rho'v_0, a\rho'v_1 \approx 0)\), the market is sufficiently incomplete \((\rho^2 \approx 1)\) or risk aversion is low \((1 - p \ll \infty)\).
5. Conclusion. Long-run analysis is a tractable, and yet nontrivial framework for dynamic portfolio choice and derivatives pricing in incomplete markets, and yields simple expressions for portfolios and risk premia. Long-run solutions admit closed-form even in cases in which finite-horizon policies do not, and the finite-horizon performance of long-run solutions has a simple expression.

Long-run optimality entails that the certainty equivalent loss 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. Otherwise, long-run optimality holds, and time-homogeneous portfolios are approximately optimal for long horizons.

APPENDIX A: PROOFS OF SECTION 1

Proof of Lemma 5. Denote by \( q = \frac{p}{p-1} \). In the case \( 0 < p < 1 \), Holder’s inequality with \( \tilde{p} = \frac{1}{p} \) and \( \tilde{q} = \frac{\tilde{p}}{1-\tilde{p}} = \frac{1}{1-q} \) yields:

\[
E_P \left[ X^p \right] = E_P \left[ (XM)^p M^{-p} \right] \leq E_P \left[ (XM)^{\tilde{p}} M^{-\tilde{p}} \right]^{1/\tilde{q}}
\]

\[
= E_P \left[ XM \right]^{1/\tilde{p}} E_P \left[ M^{\tilde{q}} \right]^{-p} \leq E_P \left[ M^q \right]^{-p}
\]

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

\[
E_P \left[ M^q \right]^{-p} = E_P \left[ (XM)^q X^{-q} \right]^{1-p} \leq E_P \left[ (XM)^{\tilde{q}} X^{\tilde{q}} \right]^{1-p} E_P \left[ X^{-q} \right]^{-\tilde{q}}
\]

\[
= E_P \left[ XM \right]^{-p} E_P \left[ X^p \right] \leq E_P \left[ X^p \right]
\]

and the claim now follows dividing by \( p < 0 \). In both cases, the inequality becomes an equality when \( E_P \left[ XM \right] = 1 \), and \( X^{p-1} \) is proportional to \( M \). \( \square \)

APPENDIX B: PROOFS OF SECTION 2

Proof of Theorem 7. Since the Brownian Motions \( Z \) and \( W \) are partially correlated, the following orthogonal decomposition holds (see the discussion following Assumption 2)

\[
dZ_t = \rho(Y_t)dW_t + \bar{\rho}(Y_t)dB_t
\]

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 \( (\rho \bar{\rho}')(y) + (\bar{\rho} \bar{\rho}')(y) = I_n \). For \( t \geq 0 \), define the process \( D \) by:

\[
D_t = E \left[ \int_0^t \left( -qY' \Sigma^{-1} \mu + (A - qY' \Sigma^{-1} Y) \nabla v \right)' (a')^{-1} dW_t - \int_0^t q \left( \Sigma^{-1} \mu + \Sigma^{-1} Y v \right)' \sigma d\tilde{B}_t \right]_t
\]

By assumption, the operators associated to the models \( P \) and \( \hat{P} \) satisfy Assumption 2. Since Assumption 1 also holds, Cheridito, Filipović and Yor (2005, Theorem 2.4, Remark 2.4.2) implies that for each \( y \in E \) and \( t \geq 0 \), \( P^y \) and \( \hat{P}^y \) are equivalent on \( \mathcal{F}_t \) with

\[
\frac{d\hat{P}^y}{dP^y} \bigg|_{\mathcal{F}_t} = D_t
\]
Thus, \((D_t, \mathcal{F}_t)_{t \geq 0}\) is a \(P^\eta\)-martingale. With this notation, it suffices to prove that given the solution pair \((v, \lambda)\) to (22), with \(\pi, \eta\) as in (24), the following \(P^\eta\) almost sure identities hold:

\[
(X_T^\pi)^P = e^{\lambda T + v(y) - v(Y_T)} D_T
\]

(75)

\[
(M_T^\eta)^\eta = e^{\frac{1}{2} \pi(\lambda T + v(y) - v(Y_T))} D_T
\]

(76)

Indeed, if (75) and (76) hold, then (25) and (26) follow by taking expectations with respect to \(P^\eta\). Consider first (75). Passing to logarithms, it suffices to prove that:

\[
p \log X_T^\pi - \log D_T = \lambda T + v(y) - v(Y_T)
\]

(77)

The first term on the left-hand side of (77) is:

\[
p \log X_T^\pi = \int_0^T \left( pr + p\pi' \mu - \frac{1}{2} p\pi' \Sigma \pi \right) dt + \int_0^T p\pi' \sigma dZ_t
\]

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

\[
p \log X_T^\pi = \int_0^T \left( pr - \frac{1}{2} q(1 + q) \mu' \Sigma^{-1} \mu - q^2 \mu' \Sigma^{-1} \Sigma^{-1} \nabla \Sigma^{-1} \nabla v + \frac{1}{2} q(1 - q) \nabla v' \Sigma^{-1} \nabla v \right) dt
\]

\[
- q \int_0^T (\mu + \nabla v)' \Sigma^{-1} \sigma \rho dW_t - q \int_0^T (\mu + \nabla v)' \Sigma^{-1} \bar{\rho} dB_t
\]

The second term in the left-hand side of (77) follows from (73):

\[
\log D_T = \int_0^T \left( - \frac{1}{2} q^2 \mu' \Sigma^{-1} \mu + q(1 - q) \mu' \Sigma^{-1} \nabla v - \frac{1}{2} \nabla v' (A + (q^2 - 2q) \Sigma^{-1} A) \nabla v \right) dt
\]

\[
+ \frac{1}{2} \int_0^T \nabla v' a - q (\mu + \nabla v)' \Sigma^{-1} \sigma \rho dW_t - q \int_0^T (\mu + \nabla v)' \Sigma^{-1} \bar{\rho} dB_t
\]

Subtracting (79) from (78) yields:

\[
p \log X_T^\pi - \log D_T = \int_0^T \left( pr - \frac{1}{2} q \mu' \Sigma^{-1} \mu - q \mu' \Sigma^{-1} \nabla v + \frac{1}{2} \nabla v' (A - q \Sigma^{-1} A) \nabla v \right) dt - \int_0^T \nabla v' a dW_t
\]

Now, Itô’s formula allows to substitute:

\[
- \int_0^T \nabla v' a dW_t = v(y) - v(Y_T) + \int_0^T \nabla v' b dt + \frac{1}{2} \int_0^T \text{tr}(AD^2 v) dt
\]

(80)

and the claim (77) follows by recalling that \((v, \lambda)\) solves (22).

Consider now the equality (76). Again, by taking logarithms it suffices to show that:

\[
q \log M_T^\eta - \log D_T = \frac{1}{1-p} (\lambda T + v(y) - v(Y_T))
\]

(81)

The first term in the left-hand side is equal to (plugging \(\eta = \nabla v\) and \(Z = \rho W + \bar{\rho} B\)):

\[
q \log M_T^\eta = -q \int_0^T \rho dt + q \log \mathcal{E} \left( - \int_0^T (\Sigma^{-1} \mu + \Sigma^{-1} \Sigma \eta)' \sigma dZ_t + \int_0^T \eta' a dW_t \right)_T
\]

\[
= -q \int_0^T \left( r + \frac{1}{2} \mu' \Sigma^{-1} \mu + \frac{1}{2} \nabla v' (A - \nabla v' \Sigma^{-1} \mu \nabla v + \frac{1}{2} \nabla v' \Sigma^{-1} \Sigma^{-1} \nabla v \right) dt
\]

\[
+ q \int_0^T \left( \nabla v' a - (\mu + \nabla v)' \Sigma^{-1} \sigma \rho dW_t - q \int_0^T (\mu + \nabla v)' \Sigma^{-1} \bar{\rho} dB_t
\]
and subtracting (79) yields:

\[ q \log M^p_T - \log D_T = \frac{1}{1-p} \int_0^T \left( pr - \frac{1}{2} q \mu' \Sigma^{-1} \mu - q \mu' \Sigma^{-1} \mathcal{Y} \nabla v + \frac{1}{2} \nabla v' \left( A - q \mathcal{Y}' \Sigma^{-1} \mathcal{Y} \right) \nabla v \right) dt \]

\[- \frac{1}{1-p} \int_0^T \nabla v' \, dW_t \]

As in the previous case, (81) now follows by substituting (80) and recalling that \((v, \lambda)\) solves (22).

The proof of Theorem 9 requires two lemmas:

**Lemma 28.** Let \( \phi \in C(E; \mathbb{R}) \), \( \phi > 0 \), and let \((\mu_T)_{T \geq 0}\) be a tight family of probability measures on \((E, \mathcal{B}(E))\). Then:

\[
\liminf_{T \to \infty} \frac{1}{T} \log \int E \phi d\mu_T \geq 0
\]

**Proof.** By monotonicity, it suffices to prove the result for \( \phi \) bounded. Let \((t_n)_{n \geq 1}\) be an increasing sequence satisfying \( t_n \uparrow \infty \) and:

\[
\liminf_{T \to \infty} \frac{1}{T} \log \int E \phi d\mu_T = \lim_{n \to \infty} \frac{1}{t_n} \log \int_E \phi d\mu_{t_n}
\]

Since \( \phi \) is bounded above, the limit in (83) is non-positive. Since the measures \((\mu_T)_{T \geq 0}\) are tight, they are relatively compact with respect to the topology of weak convergence. Thus, up to a subsequence, there exists a probability measure \( \mu \) on \( E \) such that

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

because \( \phi \) is continuous, bounded and positive. Thus, for any \( 0 < \varepsilon < \int_E \phi d\mu \), there is an \( n_\varepsilon \) such that \( n \geq n_\varepsilon \) implies

\[
\frac{1}{t_n} \log \int_E \phi d\mu_{t_n} \geq \frac{1}{t_n} \log \left( \int_E \phi d\mu - \varepsilon/2 \right)
\]

Hence, taking \( n \uparrow \infty \) it follows that the limit in (83) is indeed 0.

**Lemma 29.** Let \((Q^y)_{y \in E}\) be a solution to the martingale problem for the operator \( L \) on \( E \), where:

\[ L = \frac{1}{2} \sum_{i,j=1}^k \Sigma(y)^{ij} \frac{\partial^2}{\partial y_i \partial y_j} + \sum_{i=1}^k \mu(y)^i \frac{\partial}{\partial y_i} \]

If \( f \in C^2(E, \mathbb{R}) \), \( f > 0 \), then \( E_Q^y [f(Y_T)] \leq f(y) + (0 \vee \sup_E L f) T \).

**Proof.** Let \( f \in C^2(E, \mathbb{R}) \), \( f > 0 \). Since \((Q^y)_{y \in E}\) solves the martingale problem for \( L \) on \( E \), the process

\[
f(Y_T) - \int_0^T (Lf)(Y_t) dt
\]
is a local martingale under $Q^y$. Let $(\tau_n)_{n\geq 1}$ be a reducing sequence of stopping times for this local martingale. Then

$$E^y_Q [f (Y_{T \wedge \tau_n})] = f(y) + E^y_Q \left[ \int_0^{T \wedge \tau_n} (L f)(Y_t) dt \right] \leq f(y) + \left( 0 \vee \sup_E (L f) \right) T$$

and the thesis follows by Fatou's lemma, taking $n \uparrow \infty$.

**Proof of Theorem 9.** Applying the inequality in (17) to equations (25) and (26) from Theorem 7 gives

$$\begin{align*}
0 \leq & \liminf_{T \to \infty} \frac{1}{p} \left( \frac{1}{T} \log E^y_P [\left( (M^y_T)^{\alpha v} \right)^{1-p} - \frac{1}{T} \log E^y_P [(X_T^y)^p] \right) \\
& \leq \limsup_{T \to \infty} \frac{1}{pT} \log E^y_P [\left( (M^y_T)^{\alpha v} \right)^{1-p} - \frac{1}{T} \log E^y_P [(X_T^y)^p] \\
& = \limsup_{T \to \infty} \frac{1}{pT} \log E^y_P \left[ e^{-\frac{1}{1-p} v(Y_T)} \right] - \liminf_{T \to \infty} \frac{1}{pT} \log E^y_P \left[ e^{-v(Y_T)} \right]
\end{align*}$$

Thus, it is sufficient to prove for $0 < p < 1$ that

$$\begin{align*}
\limsup_{T \to \infty} \frac{1}{pT} \log E^y_P \left[ e^{-\frac{1}{1-p} v(Y_T)} \right] & \leq 0 \\
\liminf_{T \to \infty} \frac{1}{pT} \log E^y_P \left[ e^{-v(Y_T)} \right] & \geq 0
\end{align*}$$

and for $p < 0$ that

$$\begin{align*}
\liminf_{T \to \infty} \frac{1}{pT} \log E^y_P \left[ e^{-\frac{1}{1-p} v(Y_T)} \right] & \geq 0 \\
\limsup_{T \to \infty} \frac{1}{pT} \log E^y_P \left[ e^{-v(Y_T)} \right] & \leq 0
\end{align*}$$

for $p < 0$. The lower bounds (87) and (86) follow from the application of Lemma 28 to the functions $\phi = \exp \left( -\frac{1}{1-p} v \right)$ and $\phi = \exp (-v)$ respectively. For the upper bounds, first denote by:

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

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

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

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

where the second equality follows from (22). For $p < 0$, consider $\alpha = -1$:

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

Assumption ii) implies that the right-hand side is bounded by some constant $K$, and Lemma 29 yields:

$$E^y_P [e^{-v(Y_T)}] \leq e^{-v(y)} + (K \vee 0) T$$
and hence
\[ \limsup_{T \to \infty} \frac{1}{T} \log E_P^y \left[ e^{-\lambda(Y_T)} \right] \leq 0 \]
Similarly, for \( 0 < p < 1 \) consider \( \alpha = -\frac{1}{1-p} \):
\[ L \left( e^{-\frac{1}{1-p}v} \right) = \frac{1}{1-p} e^{-\frac{1}{1-p}v} \left( -\frac{q}{2} \nabla v' (A - Y' \Sigma^{-1} Y) \nabla v - \lambda + pr - \frac{q}{2} \mu' \Sigma^{-1} \mu \right) \]
Again, the right-hand side is bounded by \( K \), and Lemma 29 yields:
\[ E_P^y \left[ e^{-\frac{1}{1-p}v(Y_T)} \right] \leq e^{-\frac{1}{1-p}v(y)} + (K \vee 0) T \]
and the claim follows as in the previous case. \( \square \)

The proofs of Theorems 13 and 18 are obtained by adapting the arguments in Kaise and Sheu (2006) to the present setting. Because the structure of these proofs remains the same, the arguments are not repeated in detail, focusing instead on the necessary modifications.

Henceforth, all references in italics point to Kaise and Sheu (2006). For ease of notation, set:
\[ \hat{A} = A - q Y' \Sigma^{-1} Y, \quad \hat{b} = b - q Y' \Sigma^{-1} \mu, \quad V = pr - \frac{q}{2} \mu' \Sigma^{-1} \mu \]
and define the quasilinear operator \( M \) acting on \( f \in C^2(E) \) by
\[ Mf = \frac{1}{2} \text{tr}(AD^2f) + \frac{1}{2} \nabla \hat{A} \nabla f + \hat{b} \nabla f + V \]
so that (22) becomes \( Mv = \lambda \). The following results carry over immediately from Kaise and Sheu (2006) and Ladyzhenskaya and Ural’tseva (1968) with only the obvious notational changes:

**Lemma 30** (Lemma 2.4 and the discussion following). Let Assumption 1 hold. Let \( \lambda \in \mathbb{R} \). Then
i) If, for each \( N \) there exists a \( v_N \in C^{2, \gamma}(E_N, \mathbb{R}) \) satisfying \( Mv_N = \lambda \) in \( E_N \), then there exists a \( v \in C^{2, \gamma}(E, \mathbb{R}) \) satisfying \( Mv = \lambda \).
ii) If \( \lambda_m \to \lambda \) and for each \( m \) there is a \( v_m \in C^{2, \gamma}(E, \mathbb{R}) \) solving \( Mv_m = \lambda_m \), then there is a \( v \in C^{2, \gamma}(E, \mathbb{R}) \) solving \( Mv = \lambda \).

**Remark 31.** Note that Lemma 2.4 requires the uniform ellipticity of diffusion matrices, in the form of conditions \((A1), (A2)\). Under Assumption 1, these conditions are satisfied on each \( E_N \) by continuity and positivity. In the present setting, the conclusion of Lemma 2.4 is that, if \( v_N \in C^{2, \gamma}(E_N, \mathbb{R}) \) solves \( Mv_N = \lambda \) in \( E_N \), then for each \( n \) there exist constants \( A_n, B_n, C_n > 0 \) such that for \( N > 2n \)
\[ \sup_{E_N} |\nabla v_N| \leq \max \left\{ A_n + B_n \lambda, C_n \right\} \]

**Lemma 32** (Theorem 8.4 in Ladyzhenskaya and Ural’tseva (1968)). Let Assumptions 1 and 11 hold. For any \( N \), if there exists functions \( f_1, f_2 \in C^{2, \gamma}(E_{N+1}, \mathbb{R}) \cap C(\bar{E}_{N+1}, \mathbb{R}) \) satisfying \( Mf_1 > \lambda \) and \( Mf_2 < \lambda \) in \( E_N \), then there exists a function \( v_N \in C^{2, \gamma}(E_N, \mathbb{R}) \) satisfying \( Mv_N = \lambda \).

**Proof of Theorem 13.** In Kaise and Sheu (2006) Theorem 13 is split into two Theorems: Theorem 2.6 shows that \( \Lambda = [\lambda_c, \infty) \), while Theorem 3.2 shows that \( Y \) is \( (\hat{P}^y)_{y \in E} \)-transient for \( \lambda > \lambda_c \). The following arguments show that the conclusions of both Theorem 2.6 and Theorem 3.2 remain valid.

Regarding Theorem 2.6, given Lemmas 30 and 32, it suffices to prove that
(A) There exists a $\lambda_0 \in \mathbb{R}$ such that for each $N$ there is a $v_N \in C^{2,\gamma}(E_N, \mathbb{R})$ satisfying $Mv_N = \lambda_0$ in $E_N$.

(B) The set $\{ \lambda \in \mathbb{R} : \exists v \in C^{2,\gamma}(E, \mathbb{R})$ satisfying $Mv \geq \lambda \}$ is unbounded from above.

(C) $\lambda_c \equiv \inf \Lambda > -\infty$.

Indeed, if (A) holds true then Lemma 30 yields a solution $v \in C^{2,\gamma}(E, \mathbb{R})$ to $Mv = \lambda_0$ proving $\Lambda \neq \emptyset$. Now, let $\lambda \in \Lambda$, let $v \in C^{2,\gamma}(E, \mathbb{R})$ satisfy $Mv = \lambda$ and let $\lambda > \lambda$. If (B) holds, then there is a $f_1 \in C^{2,\gamma}(E, \mathbb{R})$ satisfying $Mf_1 > \lambda$. Thus, with $f_2 = v$, Lemma 32 applies for any $N$ and hence by Lemma 30 it follows that $\lambda \in \Lambda$. Thus $\lambda \in \Lambda \Rightarrow [\lambda, \infty) \subseteq \Lambda$. By (C) it follows that $\Lambda$ is bounded from below and by Lemma 30 it follows that $\Lambda$ is closed. Thus, $\Lambda = [\lambda_c, \infty)$ is the desired result.

Therefore, it remains to prove (A), (B) and (C). First, note that with $a$, $\sigma$ the unique symmetric, positive definite square roots of $A$ and $\Sigma$ respectively, it follows that

$$\hat{A} = a (1_k - q\rho') a = (1 - q)aa + qa (1_k - \rho') a$$

By construction $1_n - \rho\rho' \geq 0$ and hence $1_k - \rho\rho \geq 0$. Thus, setting

$$c = \begin{cases} 1 - q & p < 0 \\ 1 & 0 < p < 1 \end{cases} \quad \tilde{c} = \begin{cases} 1 & p < 0 \\ 1 - q & 0 < p < 1 \end{cases}$$

it follows that $c, \tilde{c} > 0$ are such that on $E$

$$cA \leq \hat{A} \leq \tilde{c}A$$

Define the linear operators $L^c$ and $L^\tilde{c}$ acting on $f \in C^2(E)$ by

$$L^c f = \frac{1}{2} \text{tr}(AD^2 f) + \tilde{b}'\nabla f + cV f, \quad L^\tilde{c} f = \frac{1}{2} \text{tr}(AD^2 f) + \tilde{b}'\nabla f + \tilde{c}V f$$

Let $\lambda^*$ and $\tilde{\lambda}^*$ denote the generalized principal eigenvalue for $L^c$ and $L^\tilde{c}$ on $E$ (Pinsky, 1995, Chapter 4.3). Assumptions 1, 11 and 12 imply that $\lambda^*, \tilde{\lambda}^* \in \mathbb{R}$. By (91), it follows that

$$\frac{1}{cg} L^c g \leq Mf \leq \frac{1}{cg} L^\tilde{c} \tilde{g}$$

where $g = e^{cf}$ and $\tilde{g} = e^{\tilde{cf}}$. For any $\beta > \max \{ \lambda^*, \tilde{\lambda}^*, 1 \}$, let $\lambda = \frac{\beta}{2} \left( \frac{1}{c} + \frac{1}{\tilde{c}} \right)$. By construction there exist $g, \tilde{g} \in C^{2,\gamma}(E)$, $g, \tilde{g} > 0$ satisfying $L^c g = \beta g$ and $L^\tilde{c} \tilde{g} = \beta \tilde{g}$ in $E$. Set $f_1 = \frac{1}{\tilde{c}} \log \tilde{g}$. By (92), it follows that on $E$

$$Mf_1 - \lambda \geq \frac{1}{cg} L^c g - \lambda = \frac{\beta}{2} \left( \frac{1}{c} - \frac{1}{\tilde{c}} \right) > 0$$

Similarly, setting $f_2 = \frac{1}{c} \log g$ and using (92) it follows that on $E$

$$Mf_2 - \lambda \leq \frac{1}{cg} L^\tilde{c} \tilde{g} - \lambda = \frac{\beta}{2} \left( \frac{1}{\tilde{c}} - \frac{1}{c} \right) < 0$$

Therefore, (A) holds by first applying Lemma 32 and then Lemma 30 i) for $\lambda_0 = \lambda$. Furthermore, since $\beta$ can be taken arbitrarily large, (B) holds as well. Regarding (C), let $\lambda \in \Lambda$, and let $v$ be the associated function solving $Mv = \lambda$ and let $g = e^{cv}$. By (92)

$$\lambda = \sup_E (Mv) \geq \frac{1}{c} \sup_E \left( \frac{L^c g}{g} \right) \geq \frac{1}{c} \inf_{g \geq 0} \sup_E \left( \frac{L^c g}{g} \right) = \frac{\lambda^*}{c}$$

where the last equality follows by (Pinsky, 1995, Theorem 4.4.5). Thus, $\lambda_c > -\infty$.

Regarding Theorem 3.2 note that by (91), Remark 2.1 holds for $c, \tilde{c}$ and the argument in Lemma 3.1 carries over exactly, up to obvious changes in notation. 

\[\square\]
Before proving Theorem 18, the following definitions and results are needed from the theory of Large Deviations for occupancy times of diffusions. Let Assumptions 1, 11 and 12 hold. To make the dependence upon \( \lambda \) specific, for \( \lambda \in \Lambda \) let \( \{\hat{P}^y\}_{y \in E} \) be the measure in Theorem 7 and let \( \hat{L}^\lambda \) be the operator associated to \( \{\hat{P}^y\}_{y \in E} \). Let \( M_1(E) \) denote the space of Borel probability measures on \( E \). Define the function \( I^\lambda : M_1(E) \to \mathbb{R} \) by

\[
I^\lambda(\mu) = -\inf_{u \in \mathcal{U}} \int_E \frac{\hat{L}^\lambda u}{u} \, d\mu \quad \text{where} \quad \mathcal{U} = \left\{ u \in C^2(E, \mathbb{R}) \mid u(x) \geq \varepsilon_u > 0, \frac{\hat{L}^\lambda u}{u} \text{ bounded} \right\}
\]

It is clear that \( I^\lambda \) is non-negative (\( u = 1 \)) and lower semi-continuous with respect to the weak topology on \( M_1(E) \) (\( \hat{L}^\lambda u / u \) is bounded). Set \( \tau_n = \inf \{ t \geq 0 : Y_t \in E_n^\ast \} \) and \( \tau = \lim_{n \to \infty} \tau_n \). Denote by \( \mu_T \) the occupation measure for \( Y \) on \( \{ T < \tau \} \), which satisfies:

\[
\mu_T(B) = \frac{1}{T} \int_0^T 1_{Y_s \in B} \, ds, \quad \text{for all} \ T < \tau \text{ and all Borel} \ B \subset E
\]

For compact \( K \subset M_1(E) \) it follows that for all \( y \in E \) (Donsker and Varadhan, 1976, Section 7)

\[
\limsup_{T \to \infty} \frac{1}{T} \log \hat{P}^y(\mu_T \in K, T < \tau) \leq -\inf_{\mu \in K} I^\lambda(\mu)
\]

Furthermore, the following facts hold:

**Lemma 33.** Let Assumptions 1, 11 and 12 hold. If there exists a \( \mu^* \in M_1(E) \), such that \( I^\lambda(\mu^*) = 0 \) then for all \( y \in E \), \( \hat{P}^y(\tau = \infty) = 1 \), and \( \mu^* \) possesses a \( C^{2,\gamma}(E, \mathbb{R}) \) density \( g^* \), such that \( \hat{L}^\lambda g^* = 0 \), where \( \hat{L}^\lambda \) is the formal adjoint to \( \hat{L}^\lambda \).

**Proof of Lemma 33.** That \( \hat{P}^y(\tau = \infty) = 1 \) for all \( y \in E \) follows by repeating the argument in Lemma 3.5 up to, and through the point where \( \hat{P}^y(\tau = \infty) = 1 \) for \( \mu^* \) a.e. \( y \in E \). By (Pinsky, 1995, Theorem 1.15.1) the same conclusion can be extended to all \( y \in E \).

As for the second statement, the argument in Lemma 3.6 can be repeated up to, and including, the point where it is shown that \( I^\lambda(\mu^*) = 0 \) implies \( \int_E \hat{L}^\lambda w(\mu^*) \, dx = 0 \) for all \( w \in C^\infty(E, \mathbb{R}) \). Thus, the conclusions follow from (Pinsky, 1995, page 181). In particular, \( \int_E g^* \, dy = 1 < \infty \).  

**Proof of Theorem 18.** As in the proof of Theorem 13, the proof of Theorem 18 adapts the results in Kaise and Sheu (2006) to the present setting, now extending Theorem 3.7 and Theorem 3.8. Theorem 3.7 yields that for \( \lambda = \lambda_c \), \( Y \) is \( \{\hat{P}^y\}_{y \in E} \)-positive recurrent, while Theorem 3.8 states that the solution \( v_c \) corresponding to \( \lambda_c \) is unique up to an additive constant. Theorem 3.8 carries over with only notational changes, in the light of (91) and Lemma 30.

Theorem 3.7 is composed of three parts. The first part (Proposition 3.3) states that if \( Y \) is transient under \( \{\hat{P}^y\}_{y \in E} \), then there exists \( \alpha > 0 \) such that for all \( f \in C_0(E), f \geq 0, y \in E, \) and \( T \) large enough

\[
E_{\hat{P}^y} [f(Y_T), T < \tau] \leq C(y)e^{-\alpha T}
\]

The second part uses the first part to show that \( Y \) is recurrent under \( \{\hat{P}^{\lambda_c,y}\} \). The third part states that \( Y \) is actually positive recurrent. The argument used to prove the second part (recurrence) carries over unchanged, hence details are provided on the first and third parts.
As for the first part (Proposition 3.3) assume that $Y$ is transient under $(\hat{P}^y)_{y \in E}$. Consider the set $C_m = \{ \mu \in M_1(E) : \mu(E_l) \geq 1 - \delta_l \; \forall l \geq m \}$. To construct the constants $\delta_l$, set

$$U_0 = - \left( V + \frac{1}{2} \nabla w' \hat{A} \nabla w + \hat{b}' \nabla w + \frac{1}{2} \text{tr} (AD^2) \right)$$

for the function $w$ from Assumption 16. Note that by (36), $\lim_{n \to \infty} \inf_{y \in E \setminus E_n} U_0(y) = \infty$. Let $c$ be as in (91). Set $\beta_0 = \inf_{y \in E} c(U_0(y) + \lambda)$ and $\beta_l = \inf_{y \in E \setminus E_l} c(U_0(y) + \lambda)$. Let $M > 0$, and set

$$\delta_l = \frac{M + |\beta_0|}{|\beta_0| + \beta_l}$$

Since $\beta_l \uparrow \infty$, it follows that $\delta_l \downarrow 0$ and hence the set $C_m$ is relatively compact (weak topology) for $m$ large enough, so that $\beta_m > 0$ and $\delta_m < 1$. Set $\bar{\phi} = \exp(c(w - v))$ where $v$ is such that $Mv = \lambda$. Then, repeating the arguments in Proposition 3.3 it follows that for any $f \in C_0(E)$ (equation (3.22))

$$E^y_P[f(Y_T), T < \tau] \leq \sup_E |f| \hat{P}^y \mu_T \in C_m, T < \tau] + \sup_E |f/\bar{\phi}| \bar{\phi}(y)e^{-Mt}$$

By (93) it follows, after taking $M \uparrow \infty$ that

$$\limsup_{T \to \infty} \frac{1}{T} \log E^y_{\hat{P}^\lambda} [f(Y_T), T < \tau] \leq - \inf_{\mu \in \tilde{C}_m} I^\lambda(\mu)$$

Now, since $I^\lambda$ is lower semi-continuous and $\tilde{C}_m$ is compact, $- \inf_{\mu \in \tilde{C}_m} I^\lambda(\mu) = -I^\lambda(\mu^*)$ for some $\mu^* \in \tilde{C}_m$.

It is now shown that $I^\lambda(\mu^*) > 0$. Suppose, by contradiction, that $I^\lambda(\mu^*) = 0$ and that $Y$ is transient under $(\hat{P}^\lambda)_{y \in E}$. Since $I^\lambda(\mu^*) = 0$, Lemma 33 implies that:

a) $\hat{P}^\lambda[y = \infty] = 1$ for all $y \in E$.

b) with $\hat{L}^\lambda$ denoting the adjoint operator to $\hat{L}^\lambda$, there exists a $C^{2,\gamma}(E, \mathbb{R})$ positive function $g^*$ such that $\hat{L}^\lambda g^* = 0$ and $\int_E g^*(x) dx = 1$.

However, Pinsky (1995, Corollary 4.9.4) implies that if $Y$ is transient under $(\hat{P}^\lambda)_{y \in E}$ and there exists some $\bar{\phi} \in C^{2,\gamma}(E, \mathbb{R}), \bar{\phi} > 0$ which satisfies $\hat{L}^\lambda \bar{\phi} = 0$, then $\int_E \bar{\phi} dy < \infty$ implies $\hat{P}^\lambda[y < \infty] > 0$ for all $y \in E$. This conclusion contradicts a) above. Thus, $I^\lambda(\mu^*) > 0$ and, in view of (95), the inequality in (94) holds for any $0 < \alpha < I^\lambda(\mu^*)$ and large enough $T$.

To show the third part (positive recurrence for $Y$ under $(\hat{P}^\lambda)_{y \in E}$), the same steps as in part one can be repeated to obtain (95), as these steps do not require that $Y$ is transient. Now, if $\inf_{\mu \in \tilde{C}_m} I^\lambda(\mu) > 0$, then for all $f \in C_0(E)$, (94) implies that

$$\int_0^\infty E^y_{\hat{P}^\lambda} [f(Y_T), T < \tau] dT < \infty$$

in which case $Y$ is transient under $(\hat{P}^\lambda)_{y \in E}$ (Pinsky, 1995, Chapter 4.2). But part two implies that $Y$ is recurrent. Thus, there is a $\mu^* \in \tilde{C}_m$ such that $I^\lambda(\mu^*) = 0$. Since $\mu^*$ is a probability measure, ergodicity follows by Lemma 33.

Proof of Proposition 19. The invariant density for $Y$ under $(\hat{P}^y)_{y \in E}$ is $\bar{\phi} = \phi_c^2 m_\nu$ where $\phi_c = \exp(v_c/\delta)$ and $m_\nu$ is from (34). Equations (25) and (26) from Theorem 7 become

$$E^y_P[(X_T^y)^p] = e^{\lambda T + v(y)} E^y_P \left[ \phi_c(Y_T)^{-\delta} \right]$$

$$E^y_P[(M_T^y)]^{1-p} = e^{\lambda T + v(y)} E^y_P \left[ \phi_c(Y_T)^{-\delta} \right]^{1-p}$$

Proof of Proposition 19. The invariant density for $Y$ under $(\hat{P}^y)_{y \in E}$ is $\bar{\phi} = \phi_c^2 m_\nu$ where $\phi_c = \exp(v_c/\delta)$ and $m_\nu$ is from (34). Equations (25) and (26) from Theorem 7 become

$$E^y_P[(X_T^y)^p] = e^{\lambda T + v(y)} E^y_P \left[ \phi_c(Y_T)^{-\delta} \right]$$

$$E^y_P[(M_T^y)]^{1-p} = e^{\lambda T + v(y)} E^y_P \left[ \phi_c(Y_T)^{-\delta} \right]^{1-p}$$
Since $\int_E m_\nu dy < \infty$, $\int_E \phi_c^2 m_\nu dy < \infty$ it follows that

$$\int_E \phi_c^2 \phi_c^2 m_\nu dy < \infty \quad \text{and} \quad \int_E \phi_c^d \phi_c^2 m_\nu dy < \infty$$

provided that $2 - \delta > 0$ and $2 - \frac{\delta}{1-\rho} > 0$. These conditions are equivalent to those in (39). Thus, by the ergodic result (30) it holds that

$$\lim_{T \to \infty} E_y^y \left[ \phi_c(Y_T)^-\delta \right] = \int_E \phi_c^-\delta \phi_c^2 m_\nu dy \equiv K_1$$

and long run optimality follows. Furthermore, in light of (19), the conclusion in (40) follows with $K = \frac{1}{p} \log(K_2/K_1)$.

**APPENDIX C: PROOFS OF SECTION 4**

**Proof of Theorem 22.** If $v_0 \in \mathbb{R}^k$ and $v_1 \in \mathbb{R}^{k \times k}$, $v_1$ symmetric solve (55) and (56) respectively then $v(y) = v_0 y - \frac{1}{2} y v_1 y$ solves (22) for $\lambda$ from (57). Under Assumption 21, the condition (37) holds, hence Theorems 13 and 18 imply that, if for $v$ the associated process $Y$ is $\hat{P}^y$-tight in $\mathbb{R}^k$ for each $y \in \mathbb{R}^k$, then $v$ is the desired solution. The Riccati equation (56) admits the form

$$v_1 B B' v_1 - v_1 A - A' v_1 - C' C = 0$$

with

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

where $B$ is assumed to be the unique symmetric positive definite square root of $A - q Y' \Sigma^{-1} Y$. $C$ is a real valued matrix when $p < 0$. For a real valued square matrix $M$, write $M > 0$ if $M$ is strictly positive definite. If $M > 0$, then the real part of each of its eigenvalues is strictly positive. To see this, let $x, \lambda$ such that $M x = \lambda x$. Then (with $\bar{x}$ denoting the complex conjugate of $x$)

$$0 < \bar{x}' (M + M') x = \bar{x}' \lambda x + \bar{x}\lambda' x = 2|x|^2 \text{Re} (\lambda)$$

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 $A - B F_1 < 0$ and $A' - C' F_2 < 0$, then there exists a unique solution $v_1$ such that $A - B B' v_1 < 0$. Since $b > 0$ and $p < 0$, the choice of $F_1 = -q B^{-1} Y' \Sigma^{-1} \mu_1$ and $F_2 = -\sqrt{q} \sigma a'$ suffices. The condition $A - B B' v_1 < 0$ yields

$$-((b + q Y' \Sigma^{-1} \mu_1) + (A - q Y' \Sigma^{-1} Y) v_1) < 0$$

Therefore,

$$\left( v_1 (A - q Y' \Sigma^{-1} Y) + (b + q Y' \Sigma^{-1} \mu_1) \right)' = ((b + q Y' \Sigma^{-1} \mu_1) + (A - q Y' \Sigma^{-1} Y) v_1)' > 0$$

has eigenvalues with strictly positive real part, and is invertible. Thus $v_0$ from (55) is well defined. It remains to prove that $(Y_t)_{t \geq 0}$ is $\hat{P}^y$ tight in $\mathbb{R}^k$. Under $\hat{P}^y$, $Y$ has the dynamics

$$dY_t = -((b + q Y' \Sigma^{-1} \mu_1) + (A - q Y' \Sigma^{-1} Y) v_1) Y_t - q Y' \Sigma^{-1} \mu_0 + (A - q Y' \Sigma^{-1} Y) v_0) dt + dW_t$$

imsart-aap ver. 2010/04/27 file: longrun_1103.tex date: March 2, 2011
Setting
\[
D = \left( b + q\Sigma^{-1}\mu_1 \right) + \left( A - q\Sigma^{-1}Y \right) v_1 \\
E = D^{-1} \left( -q\Sigma^{-1}\mu_0 + \left( A - q\Sigma^{-1}Y \right) v_0 \right)
\]
the dynamics takes the form
\[
dY_t = D(E - Y_t)dt + dW_t
\]
where \( D \in \mathbb{R}^{k \times k} \) is such that \( D > 0 \) and \( E \in \mathbb{R}^k \). For \( Z_t = Y_t - E \) it follows that
\[
dZ_t = -DZ_tdt + dW_t
\]
Since
\[
|Y_t|^2 \leq 2|E|^2 + 2|Z_t|^2
\]
is suffices to show that \((Z_t)_{t\geq0} \) is \( \hat{P}^y \) tight in \( \mathbb{R}^k \), which follows because any compact set is contained in a closed ball around the origin, and by (97) if \( Z_t \) is in a closed ball around 0, then so is \( Y_t \). Using the methods derived in Bhattacharya (1978) it follows that under \( \hat{P}^y \), \( Z \) is positive recurrent. To show this, let \( \lambda^* \) and \( \lambda_\ast \) denote the maximum and minimum eigenvalues of \( A = aa' \) and let \( \theta_\ast \) denote the minimum eigenvalue of \( D + D' \). Since \( D > 0 \) and by Assumption 21, \( A > 0 \) it follows that \( \lambda^* > 0 \) and \( \theta_\ast > 0 \). Furthermore
\[
\inf_{x:|x|=1} x'Ax = \lambda^* \quad \inf_{x:|x|=1} x'Dx = \theta_\ast
\]
Thus, with the notation of Pinsky (1995, Theorem 6.2), \( \alpha(r) = \lambda^* \) and
\[
\beta(r) = \sup_{x:|x|=1} \frac{\text{tr}(A) - x'Ax - 2r^2x'Dx}{x'Ax} \leq \frac{\text{tr}(A) - \lambda^* - 2r^2\theta_\ast}{\lambda^*}
\]
where the last inequality follows for \( r \) large enough so that the numerator is negative. Therefore, Pinsky (1995, Theorem 6.2) applies and \( Z \) is positive recurrent.

Proof of Proposition 23. In light of Theorems 22 and 9, it suffices to show that for \( p < 0 \) the following quantity is bounded as a function of \( y \):
\[
(pr_0 - \lambda - \frac{1}{2}q(\nu_0' + b\nu_1y)(\nu_0' + b\nu_1y) + \frac{1}{2}q\rho'\rho(\nu_0 - \nu_1y)^2) e^{-\nu_0y + \frac{1}{2}\nu_1y^2}
\]
Note that, for \( p < 0 \), \( \Theta \) in (62) satisfies:
\[
\Theta > \left( 1 + q\rho'\nu_1 \right)^2
\]
Therefore:
\[
\nu_1 = b\delta \left( \sqrt{\Theta} - (1 + q\rho'\nu_1) \right) > 0
\]
and (98) is bounded over \( \mathbb{R} \) only if the quadratic term is negative:
\[
\frac{1}{2}q\nu_1^2\rho - \frac{1}{2}qb^2\nu_1'\nu_1 < 0
\]
But
\[ \frac{1}{2} qv^2_0 \rho' \rho - \frac{1}{2} q b^2 \nu_1 v_1 = \frac{1}{2} q \rho' \rho^2 b^2 \left( \sqrt{\Theta} - (1 + q \rho' \nu_1) \right)^2 - \frac{1}{2} b^2 \delta \left( \Theta - (1 + q \rho' \nu_1)^2 \right) \]
\[ = \frac{1}{2} \delta^2 b^2 \left( \sqrt{\Theta} - (1 + q \rho' \nu_1) \right) \left( (2q \rho' - 1) \sqrt{\Theta} - (1 + q \rho' \nu_1) \right) \]

From (65), the quantity \((2q \rho' - 1) \sqrt{\Theta} - (1 + q \rho' \nu_1)\) is negative, while \(\frac{1}{2} \delta^2 b^2 \left( \sqrt{\Theta} - (1 + q \rho' \nu_1) \right)\) is positive by (99). Therefore, the leading quadratic term is negative and the result follows.

\[ \Box \]

**Proof of Corollary 24.** When \(\nu_1 = -\kappa \rho\) condition (65) reduces to:
\[ (1 - 2q \rho' \rho) (1 + q \rho' \rho (\kappa^2 - 2 \kappa))^{1/2} + (1 - q \kappa \rho' \rho) > 0 \]

Set \(x = q \rho' \rho\) and consider the continuous function
\[ f(x, \kappa) = (1 - 2x) (1 + x (\kappa^2 - 2 \kappa))^{1/2} + (1 - \kappa x) \]
on \(0 < x < 1, \kappa \in \mathbb{R}\). For a fixed \(0 < x < 1\) consider the implicit equation for \(\kappa\) obtained by setting \(f(x, \kappa) = 0\). For \(0 < x \leq \frac{1}{4}\) one can show there are no \(\kappa \in \mathbb{R}\) such that \(f(x, \kappa) = 0\). For \(\frac{1}{4} < x < 1\), \(f(x, \kappa) = 0\) only along the curve \(\kappa = \frac{2}{4x - 1}\). For a fixed \(x\) and large positive \(\kappa\)
\[ f(x, \kappa) \approx \kappa \left( (1 - 2x) \sqrt{x} - x \right) < 0 \]
and for large negative \(\kappa\)
\[ f(x, \kappa) \approx |\kappa| \left( (1 - 2x) \sqrt{x} + x \right) > 0 \]
therefore, plugging back in \(q \rho' \rho\) for \(x\), for \(\frac{1}{4} < q \rho' \rho < 1\) the restriction \(\kappa < \frac{2}{4q \rho' \rho - 1}\) is necessary.

\[ \Box \]

**Proof of Proposition 25.** When \(\kappa = 1\), by Corollary 24, long-run optimality holds for \(0 < q \rho' \rho \leq \frac{1}{4}\). 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 \(q \rho' \rho \geq \frac{3}{4}\), which is equivalent to \(\delta \geq 4\) since \(\delta = \frac{1}{1 - q \rho' \rho}\). When \(\kappa = 1\) the solution \(v_1, v_0\) and \(\Theta\) simplify considerably to \(\Theta = \delta^{-1}; v_1 = b \left( \sqrt{\delta} - 1 \right);\) and \(v_0 = q \delta' \nu_0\). 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 \(\mu_t = ye^{-\frac{b}{\sqrt{\delta}} t}\), and \(\sigma_t^2 = \frac{\delta}{2b} \left( 1 - e^{-\frac{2b}{\sqrt{\delta}} t} \right)\).
Therefore,
\[ E^y_{\hat{P}} \left[ e^{-v(Y_t)} \right] = E \left[ \exp \left( A Y_t^2 + B Y_t \right) \right] \]
where \(A = \frac{b}{2} \left( \sqrt{\delta} - 1 \right), B = -q \delta' \nu_0\). For \(X \sim N(\mu, \sigma^2)\)
\[ E \left[ e^{AX^2 + BX} \right] = \left\{ (1 - 2A \sigma^2)^{-1/2} \exp \left( (1 - 2A \sigma^2)^{-1} \left( \mu^2 A + \mu B + \frac{1}{2} \sigma^2 B^2 \right) \right) \right\}_{\infty} \]
\[ A < \frac{1}{2\sigma^2} \]
\[ A \geq \frac{1}{2\sigma^2} \]
Therefore, \(E^y_{\hat{P}} \left[ e^{-v(Y_t)} \right] < \infty\) if and only if \(\frac{b}{2} \left( \sqrt{\delta} - 1 \right) < \frac{1}{2\sigma_t^2}\). This condition reduces to
\[ 1 + \frac{\sqrt{\delta}}{2} \left( 1 - \sqrt{\delta} \right) \left( 1 - e^{-\frac{2b}{\sqrt{\delta}} t} \right) > 0 \]
Note that the left side of (101) 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:

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

If $\sqrt{\delta} (\sqrt{\delta} - 1) > 2$ or equivalently, $\delta > 4$ then $\hat{t} > 0$ exists. This proves statement (i) in Proposition 25. If $\delta = 4$, the left side of (101) reduces to $e^{-bt} > 0$ for all $t > 0$. Thus

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

On the other hand for $\delta = 4$, it follows that $\mathbb{E}_\hat{P}^y \left[ e^{-\frac{1}{1-p}v(Y_t)} \right]^{1-p} < \infty$ if

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

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

$$
\mathbb{E}_\hat{P}^y \left[ e^{-\frac{1}{1-p}v(Y_t)} \right]^{1-p} = \left( 1 - (1 - q)(1 - e^{-bt}) \right)^{-\frac{1-p}{2}} \exp \left( (1 - p) \left( \frac{b}{2} (1-q) y^2 e^{-bt} - 4q(1-q)\rho'\nu_0 e^{bt} y + \frac{8}{b} q^2 (1-q)^2 (\rho' \nu_0)^2 (1 - e^{-bt}) \right) \right)
$$

If $\nu_0 = 0$, $\mathbb{E}_\hat{P}^y \left[ e^{-v(Y_t)} \right] \sim Ke^{bt}$ and $\mathbb{E}_\hat{P}^y \left[ e^{-\frac{1}{1-p}v(Y_t)} \right]^{1-p} \sim q^{-\frac{1-p}{2}}$ for large $t$, so the certainty equivalent loss is bounded by $-\frac{b}{2p}$, proving ii) in Proposition 25.

If $\nu_0 \neq 0$, then $\mathbb{E}_\hat{P}^y \left[ e^{-v(Y_t)} \right] \sim e^{K_1 e^{bt}}$ and $\mathbb{E}_\hat{P}^y \left[ e^{-\frac{1}{1-p}v(Y_t)} \right]^{1-p} \sim q^{-\frac{1-p}{2}} e^{K_2}$ 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.

**Proof of Proposition 27.** Since $\nu_0, \nu_1$ satisfy (70), $v(y) = \nu_0 \log y + \nu_1 y$ solves (22). Under $\hat{P}$ the dynamics of $Y$ are

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

and, as mentioned in Section 4.2, the positivity of $\Theta, \Lambda$ give that $Y$ is $\hat{y}^t$ tight in $(0, \infty)$ for each $y \in (0, \infty)$.

Therefore long-run optimality will follow, if the quantity $F$ for $p < 0$ from Theorem 9 is bounded over $(0, \infty)$. Specifying to this example, it is necessary to show that

$$
\left( p\rho_0 + p\rho_1 y - \lambda - \frac{1}{2} q (\nu'_0 + y\nu_1) \right) \left( \nu_0 + y\nu_1 \right) + \frac{1}{2} q \left( \frac{\nu_0}{y} + \nu_1 \right) a^2 \rho'_y y \left( \frac{\nu_0}{y} + \nu_1 \right) e^{-(\nu_0 \log y + \nu_1 y)}
$$

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

$$
(A + By + Cy^2) y^{-\nu_0 - 1} e^{-\nu_1 y}
$$
For \( v_0, v_1 \) from (70), by (69) it follows that \( v_0 > 0, v_1 < 0 \) and so (102) will follow only if \( A < 0, C < 0 \). As for \( A \)

\[
A = \frac{1}{2} qa^2 \rho' \rho v_0^2 - \frac{1}{2} q v_0' v_0
\]

\[
= \frac{1}{2} qa^2 \rho' \rho \frac{\delta^2}{a^4} \left( \sqrt{\Lambda} - \left( b \theta - qa \rho' v_0 - \frac{1}{2} a^2 \right) \right)^2 - \frac{1}{2} \frac{\delta}{a^2} \left( \Lambda - \left( b \theta - qa \rho' v_0 - \frac{1}{2} a^2 \right)^2 \right)
\]

\[
= \frac{1}{2} \frac{\delta^2}{a^2} \left( \sqrt{\Lambda} - \left( b \theta - qa \rho' v_0 - \frac{1}{2} a^2 \right) \right) \left( 2qa \rho' - 1 \right) \sqrt{\Lambda} - \left( b \theta - qa \rho' v_0 - \frac{1}{2} a^2 \right)
\]

From (71)

\[
(2qa \rho' - 1) \sqrt{\Lambda} - \left( b \theta - qa \rho' v_0 - \frac{1}{2} a^2 \right) < 0
\]

Thus, \( A < 0 \) since by (69)

\[
\frac{1}{2} \frac{\delta^2}{a^2} \left( \sqrt{\Lambda} - \left( b \theta - qa \rho' v_0 - \frac{1}{2} a^2 \right) \right) > 0
\]

As for \( C \)

\[
C = \frac{1}{2} qa^2 \rho' \rho v_1^2 + pr_1 - \frac{1}{2} q v_1' v_1
\]

\[
= \frac{1}{2} qa^2 \rho' \rho \frac{\delta^2}{a^4} \left( (b + qa \rho' v_1) - \sqrt{\Theta} \right)^2 + \frac{1}{2} \frac{\delta}{a^2} \left( (b + qa \rho' v_1)^2 - \Theta \right)
\]

\[
= \frac{1}{2} \frac{\delta^2}{a^2} \left( (b + qa \rho' v_1) - \sqrt{\Theta} \right) \left( (1 - 2qa \rho') \sqrt{\Theta} + (b + qa \rho' v_1) \right)
\]

From (71)

\[
(1 - 2qa \rho') \sqrt{\Theta} + (b + qa \rho' v_1) > 0
\]

Thus, \( C < 0 \), since by (69)

\[
\frac{1}{2} \frac{\delta^2}{a^2} \left( (b + qa \rho' v_1) - \sqrt{\Theta} \right) < 0
\]

\[
\square
\]

REFERENCES


Dynam. Control 22(7), 1001–1026.