FILTERING AND PORTFOLIO OPTIMIZATION WITH STOCHASTIC UNOBSERVED DRIFT IN ASSET RETURNS

JEAN-PIERRE FOUQUE*, ANDREW PAPANICOLAOU†, AND RONNIE SIRCAR‡

Abstract. We consider the problem of filtering and control in the setting of portfolio optimization in financial markets with random factors that are not directly observable. The example that we present is a commodities portfolio where yields on futures contracts are observed with some noise. Through the use of perturbation methods, we are able to show that the solution to the full problem can be approximated by the solution of a solvable HJB equation plus an explicit correction term.

Key words. Portfolio optimization, filtering, Hamilton-Jacobi-Bellman equation, asymptotic approximations.

Subject classifications. 91G20, 60G35, 35Q93, 35C20

Dedicated to George Papanicolaou in honor of his 70th birthday

1. Introduction

A central problem in Financial Mathematics concerns portfolio optimization: how to allocate capital in an optimal allocation between investment opportunities with differing risk characteristics. A logical objective for maximization is the *expected utility* of portfolio value at a future time, measured with respect to a stochastic model of market uncertainty and a concave increasing utility function, whose concavity models the notion of risk-aversion. A great success in analyzing this problem within a continuous time model driven by Brownian motion was the work of Merton [15] in 1969, which provides one of the few explicit solutions of a Hamilton-Jacobi-Bellman (HJB) PDE in stochastic control. In that work, the utility function U was (typically) a power function:

$$U(x) = \frac{x^{1-\gamma}}{1-\gamma}, \qquad \gamma > 0, \gamma \neq 1,$$

where γ is known as the constant of (relative) risk aversion; and the stocks were modeled as geometric Brownian motions which, in one-dimension, means that a stock price S_t evolves according to the stochastic differential equation (SDE)

$$\frac{dS_t}{S_t} = \mu\, dt + \sigma\, dW_t,$$

where W is a standard Brownian motion, μ is a constant expected growth rate, and σ is the constant volatility parameter. This is the same model used by Black & Scholes in their famous option pricing analysis [5].

Since Merton's work, there is a long line of research that relaxes some of the original assumptions in order to make the model more realistic, for instance by making σ and/or μ themselves stochastic, allowing for transaction costs, or incorporating the unobservability of the parameters. Here, we

*Department of Statistics & Applied Probability, University of California, Santa Barbara, CA 93106-3110, *fouque@pstat.ucsb.edu*. Work supported by NSF grant DMS-1107468.

†School of Mathematics & Statistics, The University of Sydney, NSW Australia *alpapani@maths.usyd.edu.au*. Work partially supported by NSF grant DMS-0739195.

‡ORFE Department, Princeton University, E-Quad, Princeton NJ 08544; *sircar@princeton.edu*. Work partially supported by NSF grant DMS-1211906.

analyze the Merton problem when the growth rate is an unobserved Gaussian process Y_t, whose level is estimated by filtering from observations of the stock price S. By incorporating a time scale separation in the fluctuations of Y, we can make progress using singular perturbation techniques.

1.1. A Motivating Application

Consider an index portfolio consisting of future contracts in several commodities. For instance, the S&P GSCI is comprised of 24 commodities from all commodity sectors - energy products, industrial metals, agricultural products, livestock products, and precious metals. It is typical for the major indices to establish weights for each commodity at the beginning of the year, and then let the weights fluctuate throughout the year as prices change. The rollover strategies for well-known indices vary significantly, but the criterion will depend on whether the market is in contango (e.g. non-energy products where the yield curve tends to be upward-sloping) or if the market is in backwardation (e.g. energy markets where the yield curve tends to be downward-sloping). Hence, the problem is to determine an optimal rollover strategy given the state of the yield curve. We shall consider the convenience yield (i.e. the portion of the commodity's yield that is not due to financing or cost of carry) to be observable only through noisy measurements from the market data.

Let $F_{t,T}$ denote a future contract at time $t \leq T$ with maturity at time T. The yield on this contract is

$$\mathbb{Y}_{t,T} = \frac{1}{T-t} \log\left(F_{t,T}/F_{t,t}\right)$$

where $F_{t,t}$ is the future contract expiring now and is equivalent to the 'fair' spot price. This quantity embodies three elements of the market: 1) the cost of financing the position, 2) the cost of storage 3) the convenience afforded to the party with direct access to the physical good. The last of these three points is referred to as *the convenience yield.* Whether or not there is a convenience yield is not exactly determinable from the yield curve. This is because the spot price is not something that really exists in most commodities markets. The widely-held view is that quantities such as the cost of storage and interest rate are exogenously inserted into the commodities market, but the spot price is a noisy estimate even in the most liquid markets (other than electricity markets). Hence, the true yield $\mathbb{Y}_{t,T}$ is latent and must be estimated - and will be particularly noisy for the contracts with short time-to-maturity. In fact, the state of contango and backwardation are also latent. In the presence of these latent (or unobserved) states, we say that the market has *partial information.* In this paper we consider a simplified market with just one commodity, with partial information being described by a hidden Markov model (HMM). In the most general setting we let S_t denote an observable vector of the prices on several commodities portfolios, and we let Y_t be a vector of latent variables for their stochastic rates of returns. Returns on these portfolios are given by

$$\frac{dS_t^i}{S_t^i} = Y_t^i \, dt + \sum_j \sigma_{ij} \, dW_t^i$$

where i is the vector's index, $\sigma\sigma^T$ is a covariance matrix, and W_t is a (vector) Brownian motion.

To make inferences on the state of the yield curve given the history $(S_u)_{u \leq t}$, we employ filtering to keep track of the posterior distribution of Y_t. Filtering is advantageous because it brings the uncertainty of the yield curve into the equations for pricing and hedging, whereas uncertainty is overlooked when we assume full information and insert a point estimator of Y_t (for instance using a moving average of past returns). A drawback of filtering is that it relies heavily on model parameters, and spurious posteriors may occur if there is model mis-specification, as the filter might be 'over-learning'. However, the presence of distinct time scales in the HMM (if there are any) can be exploited to avoid over-learning, because the filtering and control can be approximated with simpler formulas that rely on roughly-accurate statistical properties of the processes rather than overly precise dynamics. Model mis-specification may still be an issue, but not as much as it would have been if perturbation theory, particularly averaging, was not used.

1.2. Related Literature

The extensive literature on optimal investment problems is discussed in many books and survey articles, and we mention, for instance [19, 22] and [20]. Models where stochastic volatility is driven by multiscale *observable* factors were analyzed in [8, 12, 11, 9]. The full information portfolio problem with a mean reverting drift has been studied by [23] for the complete market (perfectly correlated) case. The portfolio with partial information has been studied in [21], with Markov chain switching in [2], and with the effects of discrete trading in [3]. Asymptotic (nonlinear) filtering has been addressed in [18] and in [17].

1.3. Results in this Paper

In this paper we will explore the problem of portfolio optimization in the presence of partial observation. The unobserved drift (or commodities yield) Y is a fast mean-revering Ornstein-Uhlenbeck process with time scale parameter $0 < \varepsilon \ll 1$, whose posterior distribution is obtained using a Kalman filter. The filter is then used as an input for a portfolio problem with the objective of maximizing expected terminal utility. We quantify the small-ε asymptotic behavior of the Kalman filter in Section 2, and calculate expansions of the optimized value function in powers of $\sqrt{\varepsilon}$ in Section 3. We present numerical examples in Section 4 to show how the partial information value function compares to that in the full information case, and we also explore some 'practical' portfolio strategies that are sub-optimal, but which do not require tracking the stochastic growth rate that is moving on a fast time scale.

2. Portfolio Returns & Filtering

We work with the following model for the returns on a traded asset with price S_t and stochastic growth rate Y_t on a fixed finite time horizon $[0, T]$:

$$\begin{aligned} \frac{dS_t}{S_t} &= Y_t\,dt + \sigma\,dW_t \qquad \text{(observed)}, \\ dY_t &= \frac{1}{\varepsilon}(\theta - Y_t)\,dt + \frac{\beta}{\sqrt{\varepsilon}}\,dB_t \quad \text{(latent/hidden)}, \end{aligned} \tag{2.1}$$

where W and B are Brownian motions with correlation coefficient $|\rho| < 1$: $\mathbb{E}\{dW_t dB_t\} = \rho\,dt$. The positive parameters σ, ε and β, as well as the long run mean drift level θ are considered to be known. We denote by $\mathcal{L}^\varepsilon$:

$$\mathcal{L}^\varepsilon \triangleq \left(\frac{\partial}{\partial t} + \frac{1}{2}\sigma^2 s^2 \frac{\partial^2}{\partial s^2} + ys\frac{\partial}{\partial s}\right) + \frac{1}{\sqrt{\varepsilon}}\sigma\beta\rho s\frac{\partial^2}{\partial s \partial y} + \frac{1}{\varepsilon}\mathcal{M},$$

where

$$\mathcal{M} \triangleq \frac{1}{2}\beta^2\frac{\partial^2}{\partial y^2} + (\theta - y)\frac{\partial}{\partial y}\ ,$$

so $\mathcal{L}^\varepsilon$ is $\frac{\partial}{\partial t}$ plus the infinitesimal generator of the system (2.1).

2.1. Filtering

Let $(\mathcal{F}_t)$ denote the filtration generated by observing S, defined by the σ-algebra $\mathcal{F}_t \triangleq \sigma\{S_u : u \leq t\}$, which is all the information available to the observer at time $t \leq T$. For any square-integrable function $g : \mathbb{R} \to \mathbb{R}$, the posterior moment is

$$\hat{g}_t \triangleq \mathbb{E}\{g(Y_t)|\mathcal{F}_t\} \qquad \forall t \leq T. \tag{2.2}$$

This is the linear Gaussian case for filtering, which means that the Kalman-Bucy filter applies. In particular, the conditional distribution $Y_t \mid \mathcal{F}_t$ is normal:

$$\mathbb{P}\left(Y_t \leq y \middle| \mathcal{F}_t\right) = \frac{1}{\sqrt{2\pi\Sigma^\varepsilon(t)}} \int_{-\infty}^{y} \exp\left(-\frac{1}{2}\left(\frac{v - \widehat{Y}_t}{\sqrt{\Sigma^\varepsilon(t)}}\right)^2\right) dv$$

where we have denoted

$$\widehat{Y}_t = \mathbb{E}\left\{Y_t \middle| \mathcal{F}_t\right\} \tag{2.3}$$

$$\Sigma^\varepsilon(t) = \mathbb{E}(Y_t - \widehat{Y}_t)^2. \tag{2.4}$$

The conditioning on $\mathcal{F}_t$ is dropped in (2.4) because $\widehat{Y}_t$ is a Gaussian projection onto the observations $(S_u)_{u\leq t}$, and hence the residual difference (i.e. $Y_t - \widehat{Y}_t$) is independent of the observations. Furthermore, we define the (normalized) innovations process as

$$\nu_t \triangleq \frac{1}{\sigma} \int_0^t \left(\frac{dS_u}{S_u} - \widehat{Y}_u du\right),$$

which turns out to be an $\mathcal{F}_t$-adapted Brownian motion. It follows that $\widehat{Y}_t$ and $\Sigma^\varepsilon(t)$ are solutions of the equations (see [6] or Chapter 3 of [1]):

$$d\widehat{Y}_t = \frac{1}{\varepsilon}\left(\theta - \widehat{Y}_t\right) dt + \left(\frac{\Sigma^\varepsilon(t)}{\sigma} + \frac{\beta\rho}{\sqrt{\varepsilon}}\right) d\nu_t, \tag{2.5}$$

$$\frac{d}{dt}\Sigma^\varepsilon(t) = -\frac{2}{\varepsilon}\left(\Sigma^\varepsilon(t) - \frac{\beta^2(1-\rho^2)}{2}\right) - \frac{2\beta\rho}{\sigma\sqrt{\varepsilon}}\Sigma^\varepsilon(t) - \left(\frac{\Sigma^\varepsilon(t)}{\sigma}\right)^2. \tag{2.6}$$

Equations (2.5) and (2.6) are the essential pieces of the Kalman-Bucy filter, but because they constitute a conditional or marginal distribution, we refer to $(\widehat{Y}_t, \Sigma^\varepsilon(t))$ as a *marginal Kalman filter*.

2.2. Filter Asymptotics

The posterior distribution of Y_t has a convenient limit wherein a dimension reduction takes place. As $\varepsilon \to 0$, $\widehat{Y}_t$ converges weakly to a normally distributed random variable and will forget the history of the observations. This is a consequence of ergodic theory, as we explain in the remainder of this subsection.

For simplicity, we take $\Sigma^\varepsilon(0) \equiv 0$[1]. The solution to (2.6) is given explicitly by

$$\Sigma^\varepsilon(t) = -\left(\frac{1 - e^{-td}}{1 - \frac{\alpha_-}{\alpha_+}e^{-td}}\right)\alpha_-,$$

[1] For general $\Sigma^\varepsilon(0) \geq 0$, the analysis holds with the same limits as $\varepsilon \to 0$. Verification for general $\Sigma^\varepsilon(0)$ requires some modification to the terms $\alpha_\pm$ and d.

where

$$\alpha_{\pm} = \frac{(\sigma^2+\sqrt{\varepsilon}\beta\rho\sigma) \pm \sqrt{(\sigma^2+\sqrt{\varepsilon}\beta\rho\sigma)^2 + \varepsilon\left(\beta\sigma\sqrt{1-\rho^2}\right)^2}}{\varepsilon}$$

$$d = \frac{2}{\varepsilon}\left((1+\sqrt{\varepsilon}\beta\rho/\sigma)^2 + \varepsilon\left(\frac{\beta\sqrt{1-\rho^2}}{\sigma}\right)^2\right)^{1/2}.$$

From L'Hôpital's rule, we have

$$\alpha_- \to -\frac{\beta^2(1-\rho^2)}{2} \qquad \text{and} \quad \alpha_+ \to \infty \qquad\qquad \text{as } \varepsilon \to 0,$$

so that

$$\Sigma^\varepsilon(t) \to \Sigma^{(0)} \triangleq \frac{\beta^2(1-\rho^2)}{2} \tag{2.7}$$

as $\varepsilon \to 0$. Furthermore, we can use the right-hand side of (2.6) to determine a $\sqrt{\varepsilon}$ expansion of $\Sigma^\varepsilon(t)$:

$$\Sigma^\varepsilon(t) = \left(1 - \sqrt{\varepsilon}\frac{\rho\beta}{\sigma}\right)\Sigma^{(0)} + o(\sqrt{\varepsilon}) \qquad \text{for } t > 0. \tag{2.8}$$

Notice the expansion in (2.8) has terms that do not depend on t. This is because the time dependence of $\Sigma^\varepsilon(t)$ decays exponentially with ε, but the expansion shows us that the location of the steady state of the solution to (2.6) has a $\sqrt{\varepsilon}$ perturbation. This expansion will be useful later.

An immediate consequence of the limiting behavior of $\Sigma^\varepsilon(t)$ is that the solution to (2.5) converges weakly,

$$\widehat{Y}_t = e^{-\Delta t/\varepsilon}\widehat{Y}_{t-\Delta t} + (1-e^{-\Delta t/\varepsilon})\theta + \int_{t-\Delta t}^t e^{-(t-u)/\varepsilon}\left(\frac{\Sigma^\varepsilon(u)}{\sigma} + \frac{\rho\beta}{\sqrt{\varepsilon}}\right)d\nu_u \tag{2.9}$$

for an arbitrarily small but fixed constant $\Delta t > 0$. From equation (2.9) we see that the information from $\mathcal{F}_{t-\Delta t}$ is contained entirely in $\widehat{Y}_{t-\Delta t}$, and we also see that $\widehat{Y}_t$'s dependence on the history up to time $t-\Delta t$ is decaying at an exponential rate. Hence, $\widehat{Y}_t$ becomes independent of $\mathcal{F}_{t-\Delta t}$ in the small-ε limit for any fixed $\Delta t > 0$, which tells us that the limiting posterior of Y_t does not depend on the observations, but instead will tend toward an invariant distribution. Indeed, from equation (2.9) we see that $\widehat{Y}_t$ has a weak limit that is normally distributed:

$$\widehat{Y}_t \Rightarrow N\left(\theta, \frac{\beta^2\rho^2}{2}\right) \qquad \text{as } \varepsilon \to 0,$$

where $N(a, b^2)$ denotes a normally distributed random variable with mean a and variance b^2. From a data processing point of view, this means that there is little information lost if ε is small, in which case the filter is still just as accurate with a short history of observations as it would have been with a long history. We should also point out that $\widehat{Y}_t$ has an ergodic theory,

$$\mathbb{E}\left|\frac{1}{t}\int_0^t \widehat{Y}_u du - \theta\right|^2 \to 0 \qquad \text{as } \varepsilon \to 0,$$

which can be deduced by integrating equation (2.9) in t and taking the limit in ε.

3. The Portfolio Problem with Partial Information

In the Merton problem, an investor allocates his capital dynamically over time between a risky stock and a riskless bank account to maximize expected utility of wealth at a fixed time horizon $T < \infty$. Let π_t denote the dollar amount invested in the stock at time t. The value X of his self-financing portfolio follows

$$dX_t = \pi_t \frac{dS_t}{S_t} + r(X_t - \pi_t)\, dt$$

where $r \geq 0$ is the risk-free rate, and the process (π_t) is the investor's control process adapted to $\mathcal{F}_t$, meaning he does not directly observe the drift process Y, and satisfying the admissibility condition $\mathbb{E} \int_0^T \pi_t^2\, dt < \infty$.

Given a smooth utility function $U(x)$ on $\mathbb{R}^+$ satisfying the "usual" Inada conditions (see, for instance, [13]), we define the value function,

$$\mathcal{V}^\varepsilon(t, x) = \sup_\pi \mathbb{E}\{U(X_T) | \{X_t = x\} \vee \mathcal{F}_t\}, \tag{3.1}$$

which is the portfolio problem with partial information. Notice that (3.1) is a non-Markovian formulation of the problem. However, the Kalman filter fully parameterizes the distribution of $Y_t|\mathcal{F}_t$, and (3.1) turns out to be a Markov control problem, meaning the value function is a deterministic function of $(X_t, \hat{Y}_t)$. From here forward we will take $r = 0$, to which case the problem can be reduced by a simple change of variable.

3.1. The HJB Equation

The value function for the full information case (i.e. when Y is observed) is based on the operator $\mathcal{L}^\varepsilon$ given in Section 2. For the partial information case, the control problem can be written in its Markovian formulation as follows:

$$\mathcal{V}^\varepsilon(t, x) = V^\varepsilon(t, x, y) \triangleq \sup_\pi \mathbb{E}\{U(X_T) | X_t = x, \widehat{Y}_t = y\}.$$

Furthermore, if it has sufficient regularity, the value function V^ε is the solution to an HJB equation,

$$V_t^\varepsilon + \frac{1}{2}\left(\frac{\Sigma^\varepsilon(t)}{\sigma} + \frac{\beta\rho}{\sqrt{\varepsilon}}\right)^2 V_{yy}^\varepsilon + \frac{(\theta - y)}{\varepsilon} V_y^\varepsilon + \max_\pi \left\{\frac{1}{2}\sigma^2\pi^2 V_{xx}^\varepsilon + \pi\left(y V_x^\varepsilon + \left(\Sigma^\varepsilon(t) + \frac{\sigma\beta\rho}{\sqrt{\varepsilon}}\right) V_{xy}^\varepsilon\right)\right\} = 0, \tag{3.2}$$

for all $t < T$, and with terminal condition $V^\varepsilon(T, x, y) = U(x)$ (see for instance [19, Chapter3]). In general, regularity results for this type of fully-nonlinear HJB PDE problem are not available. However, our approach here is to perturb around a case for which explicit solutions with sufficient regularity are known, that is, the case where Y_t constant. In addition, for some specific terminal conditions $U(x)$ (such as power utility), explicit solutions are available for stochastic Y_t (see [14]).

From (3.2) we see that the optimal strategy is given in feedback form by

$$\pi^* = -\frac{y V_x^\varepsilon}{\sigma^2 V_{xx}^\varepsilon} - \left(\frac{\Sigma^\varepsilon(t)}{\sigma^2} + \frac{\beta\rho}{\sqrt{\varepsilon}\sigma}\right) \frac{V_{xy}^\varepsilon}{V_{xx}^\varepsilon}. \tag{3.3}$$

We will see from the asymptotic expansions that we will construct in this section, that V_{xy}^ε is actually of order ε, and so π^* in equation (3.3) has a well-defined limit as ε goes to zero. Plugging

π^* into (3.2) we have the the following nonlinear PDE:

$$V_t^\varepsilon + \frac{1}{2}\left(\frac{\Sigma^\varepsilon(t)}{\sigma}\right)^2 V_{yy}^\varepsilon + \frac{1}{\sqrt{\varepsilon}}\frac{\beta\rho\Sigma^\varepsilon(t)}{\sigma}V_{yy}^\varepsilon + \frac{1}{\varepsilon}\mathcal{L}_0 V^\varepsilon - \frac{\left(\frac{y}{\sigma}V_x^\varepsilon + \left(\frac{1}{\sigma}\Sigma^\varepsilon(t) + \frac{\beta\rho}{\sqrt{\varepsilon}}\right)V_{xy}^\varepsilon\right)^2}{2V_{xx}^\varepsilon} = 0, \quad (3.4)$$

where

$$\mathcal{L}_0 \triangleq \frac{1}{2}\beta^2\rho^2\frac{\partial^2}{\partial y^2} + (\theta - y)\frac{\partial}{\partial y}. \quad (3.5)$$

Classical solutions to (3.4) exist for logarithmic and power utilities (see [6, 7]), but the small ε behavior allows us to expand the solution around a simpler constant growth rate classical Merton problem. We expand V^ε in powers of ε,

$$V^\varepsilon = V^{(0)} + \sqrt{\varepsilon}V^{(1)} + \varepsilon V^{(2)} + \varepsilon^{3/2}V^{(3)} + \dots, \quad (3.6)$$

insert into (3.4), and compare powers of ε.

The order ε^{-1} terms lead to

$$\mathcal{L}_0 V^{(0)} - \frac{\rho^2\beta^2}{2}\frac{\left(V_{xy}^{(0)}\right)^2}{V_{xx}^{(0)}} = 0,$$

which allows us to take $V^{(0)}$ to be constant in y: $V^{(0)} = V^{(0)}(t,x)$. Based on this choice for $V^{(0)}$ and the expansion of $\Sigma^\varepsilon(t)$ in (2.8), the nonlinear term in the HJB PDE can be expanded up to order $\sqrt{\varepsilon}$ as

$$\frac{\left(\frac{y}{\sigma}\left(V_x^{(0)} + \sqrt{\varepsilon}V_x^{(1)}\right) + \frac{\sqrt{\varepsilon}}{\sigma}\Sigma^{(0)}V_{xy}^{(1)} + \rho\beta\left(V_{xy}^{(1)} + \sqrt{\varepsilon}V_{xy}^{(2)}\right)\right)^2}{2V_{xx}^{(0)}}\left(1 - \sqrt{\varepsilon}\frac{V_{xx}^{(1)}}{V_{xx}^{(0)}}\right) + \dots,$$

in which there are no terms of order $\varepsilon^{-1/2}$. Therefore, there is no contribution by the nonlinear term at order $\varepsilon^{-1/2}$ and we have

$$\mathcal{L}_0 V^{(1)} = 0,$$

and so $V^{(1)}$ can also be chosen to be constant in y: $V^{(1)} = V^{(1)}(t,x)$.

Collecting order ε^0 terms leads us to:

$$V_t^{(0)} + \mathcal{L}_0 V^{(2)} - \frac{y^2}{2\sigma^2}\frac{\left(V_x^{(0)}\right)^2}{V_{xx}^{(0)}} = 0. \quad (3.7)$$

We now introduce some useful notation.

Definition 3.1. *The risk tolerance function based on the zeroth value function is defined as*

$$R^{(0)}(t,x) \triangleq -\frac{V_x^{(0)}(t,x)}{V_{xx}^{(0)}(t,x)}. \quad (3.8)$$

Also we define the operators

$$D_k \triangleq \left(R^{(0)}(t,x)\right)^k \frac{\partial^k}{\partial x^k} \tag{3.9}$$

and the linear operator

$$\mathcal{L}_{t,x,y} \triangleq \frac{\partial}{\partial t} + \frac{y^2}{2\sigma^2} D_2 + \frac{y^2}{\sigma^2} D_1. \tag{3.10}$$

REMARK 1. *In general, the value function $V^{(0)}(t,x)$ inherits the properties of the utility $U(x)$, specifically it is increasing and strictly concave in the wealth level x. The quantity $-\frac{U''}{U'}$ is classically known as the Arrow-Pratt measure of risk aversion (see for instance [10, Chapter1]). The risk tolerance given in Definition 3.1 is simply the reciprocal of this measure using the value function $V^{(0)}$.*

In particular, equation (3.7) can be written as

$$\mathcal{L}_0 V^{(2)} + \mathcal{L}_{t,x,y} V^{(0)} = 0, \tag{3.11}$$

because we can write the nonlinear term as

$$-\frac{\left(V_x^{(0)}\right)^2}{V_{xx}^{(0)}} = \frac{-V_x^{(0)}}{V_{xx}^{(0)}} V_x^{(0)} = D_1 V^{(0)} \quad \text{or} \quad -\frac{\left(V_x^{(0)}\right)^2}{V_{xx}^{(0)}} = -\left(\frac{-V_x^{(0)}}{V_{xx}^{(0)}}\right)^2 V_{xx}^{(0)} = -D_2 V^{(0)}. \tag{3.12}$$

3.2. The Zero-Order Term $\mathbf{V^{(0)}}$

Let μ denote the invariant density,

$$\mu(y) = \frac{1}{\sqrt{\pi\beta^2\rho^2}} e^{-\left(\frac{y-\theta}{\beta\rho}\right)^2}, \tag{3.13}$$

for which $\int \mu(y)\mathcal{L}_0 g(y)dy = 0$ for all $g \in C^2(\mathbb{R}) \cap L^2(\mu)$, and denote the average with respect to the invariant density μ as

$$\langle g \rangle \triangleq \int g(y)\mu(y)dy.$$

The Fredholm alternative leads us to the following proposition showing that $V^{(0)}$ is equal to the Merton value function with a constant Sharpe ratio:

PROPOSITION 3.2. *The zero-order term $V^{(0)}$ satisfies the PDE,*

$$V_t^{(0)} - \frac{1}{2}\bar{\lambda}^2 \frac{\left(V_x^{(0)}\right)^2}{V_{xx}^{(0)}} = 0, \qquad \textit{for } t < T, \tag{3.14}$$

$$V^{(0)}(T,x) = U(x), \tag{3.15}$$

with the squared Sharpe ratio given by

$$\bar{\lambda}^2 = \frac{\langle y^2 \rangle}{\sigma^2} = \frac{\beta^2\rho^2}{2\sigma^2} + \frac{\theta^2}{\sigma^2}. \tag{3.16}$$

The solution of (3.14) *and* (3.15) *is the Merton value function with effective Sharpe ratio* $\sqrt{\bar{\lambda}^2}$.

Proof. According to the Fredholm alternative, equation (3.11) has a solution for $V^{(0)}$ if and only if

$$\left\langle \mathcal{L}_{t,x,\cdot} V^{(0)} \right\rangle = 0. \tag{3.17}$$

Since we know $V^{(0)}$ is constant in y,

$$\left\langle \mathcal{L}_{t,x,\cdot} V^{(0)} \right\rangle = \left\langle \mathcal{L}_{t,x,\cdot} \right\rangle V^{(0)} = \mathcal{L}_{t,x} V^{(0)} = 0,$$

where from here forward we denote

$$\mathcal{L}_{t,x} \triangleq \left\langle \mathcal{L}_{t,x,\cdot} \right\rangle = \frac{\partial}{\partial t} + \frac{1}{2}\bar{\lambda}^2 D_2 + \bar{\lambda}^2 D_1, \tag{3.18}$$

and $\bar{\lambda}$ is defined in (3.16). This PDE for $V^{(0)}$ can be re-written as (3.14) using the expressions (3.12). $\square$

We will informally refer to $V^{(0)}$ as solving a nonlinear diffusion equation $\mathcal{L}_{t,x} V^{(0)} = 0$, and refer to the Sharpe ratio as a diffusion coefficient: a larger Sharpe ratio clearly means the value function $V^{(0)}$ diffuses faster, as can be shown easily using comparison principles for (3.14).

3.3. The First-Order Correction $\mathbf{V^{(1)}}$

From equation (3.11), we have

$$\begin{aligned} \mathcal{L}_0 V^{(2)} &= -\left(\mathcal{L}_{t,x,y} - \mathcal{L}_{t,x}\right) V^{(0)} \\ &= -\left(\frac{y^2}{\sigma^2} - \bar{\lambda}^2\right)\left(\frac{1}{2}D_2 + D_1\right) V^{(0)}, \end{aligned} \tag{3.19}$$

and it follows therefore that

$$V^{(2)} = -\phi(y)\left(\frac{1}{2}D_2 + D_1\right) V^{(0)} + C(t,x) \tag{3.20}$$

where ϕ is a solution to the Poisson equation

$$\mathcal{L}_0 \phi = \frac{y^2}{\sigma^2} - \bar{\lambda}^2, \tag{3.21}$$

and C is a constant (in y) of integration. Differentiating (3.20) with respect to x and y yields

$$V^{(2)}_{xy} = -\frac{1}{2}\phi'(y)\frac{\partial}{\partial x}\left(\frac{\left(V^{(0)}_x\right)^2}{V^{(0)}_{xx}}\right) \quad \text{and} \quad V^{(2)}_{yy} = -\frac{1}{2}\phi''(y)\frac{\left(V^{(0)}_x\right)^2}{V^{(0)}_{xx}},$$

which will be useful below.

In the nonlinear term of (3.4), collecting all the terms up to order $\sqrt{\varepsilon}$, gives

$$-\frac{1}{2}\left(\frac{y}{\sigma}\right)^2\frac{\left(V^{(0)}_x\right)^2}{V^{(0)}_{xx}} + \sqrt{\varepsilon}\left\{\frac{y}{\sigma}\left(\frac{y}{\sigma}V^{(1)}_x - \frac{1}{2}\rho\beta\phi'(y)\frac{\partial}{\partial x}\left(\frac{\left(V^{(0)}_x\right)^2}{V^{(0)}_{xx}}\right)\right)R^{(0)} + \frac{1}{2}\left(\frac{y}{\sigma}\right)^2\left(R^{(0)}\right)^2 V^{(1)}_{xx}\right\},$$

with $C(t,x)$ no longer appearing. Then collecting the order $\sqrt{\varepsilon}$ terms in (3.4) and using the expansion for $\Sigma^\varepsilon(t)$ in (2.8), we have

$$\mathcal{L}_0 V^{(3)} + V_t^{(1)} - \frac{\beta\rho}{\sigma}\Sigma^{(0)}\phi''(y)\frac{\left(V_x^{(0)}\right)^2}{V_{xx}^{(0)}}$$
$$+\left\{\frac{y}{\sigma}\left(\frac{y}{\sigma}V_x^{(1)} - \frac{1}{2}\rho\beta\phi'(y)\frac{\partial}{\partial x}\left(\frac{\left(V_x^{(0)}\right)^2}{V_{xx}^{(0)}}\right)\right)R^{(0)} + \frac{1}{2}\left(\frac{y}{\sigma}\right)^2\left(R^{(0)}\right)^2 V_{xx}^{(1)}\right\} = 0,$$

which can be rearranged to obtain

$$\mathcal{L}_0 V^{(3)} + \mathcal{L}_{t,x,y}V^{(1)} + \frac{\beta\rho}{\sigma}\Sigma^{(0)}\phi''(y)D_1V^{(0)} + \frac{\rho\beta y}{2\sigma}\phi'(y)D_1^2V^{(0)} = 0. \tag{3.22}$$

Applying the Fredholm alternative, the solvability condition for equation (3.22) is

$$\mathcal{L}_{t,x}V^{(1)} + \frac{\beta\rho}{\sigma}\Sigma^{(0)}\left\langle\phi''\right\rangle D_1V^{(0)} + \frac{\rho\beta}{2\sigma}\left\langle y\phi'\right\rangle D_1^2V^{(0)} = 0, \tag{3.23}$$

along with the terminal condition $V^{(1)}(T,x) = 0$.

3.4. Explicit Expression for $\mathbf{V^{(1)}}$

The following result is known from [4], and is used more recently in [16] and [9].

PROPOSITION 3.3. *The risk tolerance function $R^{(0)}$ satisfies Black's (fast diffusion) equation:*

$$R_t^{(0))} + \frac{1}{2}\bar{\lambda}^2\left(R^{(0)}\right)^2 R_{xx}^{(0)} = 0 \tag{3.24}$$

$$R^{(0)}(T,x) = -\frac{U'(x)}{U''(x)}, \tag{3.25}$$

where $\bar{\lambda}^2$ is defined in (3.16).

We give the derivation for completeness in Appendix A. The following Lemma is derived and used in [9].

LEMMA 3.4. *The operators $\mathcal{L}_{t,x}$ and D_1 commute when operating on smooth functions.*

The proof is given in [9, Lemma 2.3], and we give it again in Appendix A for completeness. Based on Proposition 3.3 and Lemma 3.4, we have the following explicit expression for the solution to (3.23):

PROPOSITION 3.5. *The correction term $V^{(1)}$ is given explicitly in terms of $V^{(0)}$ as*

$$V^{(1)}(t,x) = (T-t)\left(A_1^\rho D_1 + A_2^\rho D_1^2\right)V^{(0)}, \tag{3.26}$$

where A_1^ρ and A_2^ρ are group parameters given by

$$A_1^\rho = \frac{\beta^3\rho(1-\rho^2)}{2\sigma}\left\langle\phi''\right\rangle \qquad A_2^\rho = \frac{\beta\rho}{2\sigma}\left\langle y\phi'\right\rangle.$$

Proof. Based on Lemma 3.4, it is straightforward to verify that (3.26) is the solution to (3.23). The formula for A_1^ρ uses the formula in equation (2.7) in place of $\Sigma^{(0)}$. ∎

We can take the result in Proposition 3.5 even further, by taking advantage of the Gaussian moments in the invariant measure to compute A_1^ρ and A_2^ρ explicitly. The calculations given in Appendix A.3 result in the following formulae for the group parameters of Proposition 3.5:

$$A_1^\rho = -\frac{\beta^3\rho}{2\sigma^3}(1-\rho^2), \tag{3.27}$$

$$A_2^\rho = -\frac{\beta\rho}{4\sigma^3}\left(\beta^2\rho^2+4\theta^2\right). \tag{3.28}$$

4. Numerical Study: The Loss in Utility Due to Partial Information

The preceding sections presented an asymptotic value function expansion that is a useful tool for the portfolio problem with general utility because explicit solutions are not so easy to obtain when U is not one of the few utilities (e.g. power utility or exponential utility) for which there are explicit formulae. In fact, the partial information portfolio optimization with power utility was shown in [6] to have an explicit solution up to a system of Riccati equations. There are also the explicit solutions of [2] for partial information with power utility and a modulating Markov chain.

The class of utility functions that we consider in this section is a mixture of power utilities introduced in [9]:

$$U(x) = c_1\frac{x^{1-\gamma_1}}{1-\gamma_1} + c_2\frac{x^{1-\gamma_2}}{1-\gamma_2} \tag{4.1}$$

with $c_1, c_2 \geq 0$ and $\gamma_1 \geq \gamma_2 > 0$ with $\gamma_1, \gamma_2 \neq 1$. Our study will show how the full information case compares to the partial case, and will demonstrate how the presence of correlation, measured by ρ, drives the effects that market incompleteness and partial information have on the value function at order $\sqrt{\varepsilon}$. Note the the principle term $V^{(0)}$ depends on the correlation only through ρ^2 in the average Sharpe ratio $\bar{\lambda}^2$ in equation (3.16), whereas the correction term $V^{(1)}$ has coefficients A_1^ρ and A_2^ρ from equations (3.27) and (3.28), respectively, that are cubic in ρ. In contrast, in the full-informed case we will find that the principle term does not depend on ρ, and the correction depends on ρ linearly.

We will find that for $\varepsilon \ll 1$ and $\rho < 0$, a marginal increase in $|\rho|$ will result in a marginal increase in utility for both the fully informed investor and the partially informed. Empirical studies show that ρ is strongly negative, and so this is the case that we consider. The interpretation is the following: negative ρ is a stabilizing effect, wherein losses to the asset's value signal an increase in the expected future gains, particularly as ρ gets close to -1.

4.1. The Full Information Case

Full information is the case when $\mathcal{F}_t$ is generated by both Brownian motions (W, B). In this case, the HJB equation is as follows:

$$V_t^{\varepsilon,\text{full}} + \frac{1}{\varepsilon}\mathcal{L}_0^1 V^{\varepsilon,\text{full}} - \frac{\left(\frac{y}{\sigma}V_x^{\varepsilon,\text{full}} + \frac{\beta\rho}{\sqrt{\varepsilon}}V_{xy}^{\varepsilon,\text{full}}\right)^2}{2V_{xx}^{\varepsilon,\text{full}}} = 0 \tag{4.2}$$

$$V^{\varepsilon,\text{full}}(T,x,y) = U(x), \tag{4.3}$$

where $\mathcal{L}_0^1 \triangleq \frac{1}{2}\beta^2\frac{\partial^2}{\partial y^2} + (\theta - y)\frac{\partial}{\partial y}$, i.e. it is the operator defined in (3.5) but with the correlation parameter taken to such that $|\rho| = 1$. Following the same procedure as Section 3, we expand the solution to (4.2) in powers of ε,

$$V^{\varepsilon,\text{full}} = V^{(0),\text{full}} + \sqrt{\varepsilon}V^{(1),\text{full}} + \varepsilon V^{(2),\text{full}} + \ldots,$$

and based on a solvability condition (similar to (3.17)) we find the zero-order term $V^{(0)}(t,x)$ satisfies the PDE,

$$V_t^{(0),\mathrm{full}} - \frac{1}{2}\bar{\lambda}_1^2 \frac{\left(V_x^{(0),\mathrm{full}}\right)^2}{V_{xx}^{(0),\mathrm{full}}} = 0 \qquad \text{for } t < T, \tag{4.4}$$

$$V^{(0),\mathrm{full}}(T,x) = U(x), \tag{4.5}$$

where $\bar{\lambda}_1^2 = \frac{\beta^2}{2\sigma^2} + \frac{\theta^2}{\sigma^2}$ is the (squared) Sharpe ratio given in (3.16) evaluated at $\rho = 1$. Therefore, $V^{(0),\mathrm{full}}$ for full information corresponds to a Merton problem with a greater Sharpe ratio than $V^{(0)}$ for partial information, because $\bar{\lambda}_1^2 > \bar{\lambda}^2$ for any $|\rho| < 1$.

Continuing as was done in Section 3, we find the correction term $V^{(1)}$ to be given explicitly in terms of $V^{(0)}$ as

$$V^{(1),\mathrm{full}}(t,x) = -\frac{\beta\rho(T-t)}{4\sigma^3}\left(\beta^2 + 4\theta^2\right)(D_1^{\mathrm{f}})^2 V^{(0),\mathrm{full}}(t,x)\ , \tag{4.6}$$

with the operator $D_1^{\mathrm{f}} = R^{(0),\mathrm{full}}(t,x)\frac{\partial}{\partial x}$ where $R^{(0),\mathrm{full}}$ comes from $V^{(0),\mathrm{full}}$:

$$R^{(0,\mathrm{full})}(t,x) \triangleq -\frac{V_x^{(0,\mathrm{full})}(t,x)}{V_{xx}^{(0,\mathrm{full})}(t,x)}.$$

4.1.1. Comparison with Partial Information Formulas

The HJB equation in (4.2) and the partial information equation of (3.4) are almost the same, with the following two exceptions:

- $\mathcal{L}_0^1$ operates in (4.2) whereas $\mathcal{L}_0$ operates in (3.4), and
- $\Sigma^\varepsilon(t)$ is nowhere present in (4.2), which is intuitive because Y_t is observed under full information, meaning that there is no estimation error.

Comparing equations (4.4) and (4.6) with (3.14) and (3.26), respectively, it can be seen that

$$V^{(0),\mathrm{full}} = V^{(0)}\Big|_{|\rho|=1} \qquad V^{(1),\mathrm{full}} = |\rho|\ V^{(1)}\Big|_{|\rho|=1}\ .$$

4.2. Numerical Method

Numerical solutions for calculating $V^{(0)}$ and $V^{(1)}$ are based on solving Black's equation (3.24) for $R^{(0)}$. From (3.24), it can be determined that $V_x^{(0)}$ is given by the following formula:

$$V_x^{(0)}(t,x) = V_x^{(0)}(t,x_{\max})\exp\left(\int_x^{x_{\max}} \frac{1}{R^{(0)}(t,\xi)}\,d\xi\right), \tag{4.7}$$

where $x_{\max}$ is large enough to where we can used large-x asymptotics to invert $V_x^{(0)}(t,x_{\max})$ and $V^{(0)}(t,x_{\max})$. The recipe for $V^{(0)}$ is to solve for $R^{(0)}$ and then integrate over x,

$$V^{(0)}(t,x) = V^{(0)}(t,x_{\max}) - \int_x^{x_{\max}} V_x^{(0)}(t,\xi)\,d\xi.$$

We now describe our discretization scheme for numerically solving Black's equation in (3.24). Let $\Delta t > 0$ be a time step, let $\Delta x > 0$ be a spatial step, and let $x_{\max}$ be a very large (positive) number to mark the end of the wealth process's numerical domain. For $n \in \{0, 1, \ldots, (T-t)/\Delta t\}$ and $j \in \{0, 1, \ldots, x_{\max}/\Delta x\}$, we solve numerically at the discrete points $\{(t_n, x_j) : t_n = T - n\Delta t, x_j = j\Delta x\}$ such that the solution is defined as $R_j^n \approx R^{(0)}(t_n, x_j)$. The numerical scheme is the following:

$$\frac{R_j^n - R_j^{n+1}}{\Delta t} + \frac{1}{2}\bar{\lambda}^2 \left(R_j^n\right)^2 \frac{R_{j+1}^n - 2R_j^n + R_{j-1}^n}{(\Delta x)^2} = 0, \tag{4.8}$$

using the boundary conditions $R_0^n = 0$ and $R_{x_{\max}/\Delta x}^n = \frac{1}{\gamma_2}$. The large-$x$ asymptotics are the following:

$$V^{(0)}(t,x) \sim c_1 \frac{x^{1-\gamma_1}}{1-\gamma_1} g_{12}(t) + c_2 \frac{x^{1-\gamma_2}}{1-\gamma_2} g_2(t) \qquad \text{for } x \gg 0, \tag{4.9}$$

where

$$g_2(t) = \exp\left(\frac{1}{2}\bar{\lambda}^2 \frac{1-\gamma_2}{\gamma_2}(T-t)\right)$$
$$\text{and} \qquad g_{12}(t) = \exp\left(\frac{\bar{\lambda}^2(1-\gamma_1)}{\gamma_2^2}\left(\gamma_2 - \frac{\gamma_1}{2}\right)(T-t)\right).$$

Based on this the scheme in (4.8) and the the large-x asymptotics in (4.9), the order-zero term is solved.

To compute the correction term of $V^{(1)}$, some algebra applied to D_1, D_1^2, $V^{(0)}$, and $V_x^{(0)}$ leads to the following reformulation of the expression in equation (3.26):

$$V^{(1)}(t,x) = (T-t)\left(A_1^\rho R^{(0)} + A_2^\rho R^{(0)}(R_x^{(0)} - 1)\right) V_x^{(0)}(t,x),$$

where $R_x^{(0)}$ is obtained by numerical differentiation, and $V_x^{(0)}$ is obtained by numerically integrating (4.7) for $x < x_{\max}$ and by differentiating the asymptotic approximation in (4.9) for $x \geq x_{\max}$. For the numerics in this section, we take $x_{\max} = 20$ with $\Delta x = \frac{x_{\max}}{400}$, and $T = 1$ with $\Delta t = .99\frac{\gamma_2}{\bar{\lambda}^2 \Delta x}$.

4.3. Loss in Utility Due to Partial Information & the Effects of ρ

The effect of ρ on $V^{(0)}$ on the partial and full information cases is described as follows: $V^{(0),\text{full}}$ will diffuse significantly faster than $V^{(0),\text{partial}}$ when $|\rho| < 1$, and so there is an **information premium,**

$$V^{(0),\text{partial}}(t,x) < V^{(0),\text{full}}(t,x) \qquad \text{for } t < T,\ x \geq 0, \text{ and } |\rho| < 1.$$

This follows from the fact that $V^{(0),\text{full}}$ and $V^{(0),\text{partial}}$ are solutions to a Merton problem, with full information having a higher Sharpe ratio than partial, namely, $\bar{\lambda}^2 < \bar{\lambda}_1^2$. It could also be argued that diffusion in the zero-order terms happens at a faster rate under full information, and a comparison principle will yield the inequality.

The order-$\sqrt{\varepsilon}$ term is more nuanced in how it behaves with changes of ρ. For our experiments we consider cases where $\rho \leq 0$ for the following reason: this is a type of stabilizing effect, where negative returns on the asset suggest that a correction is due, and hence, expected returns will have a slight increase as returns decrease. Indeed, a highly negative correlation between returns and drift is part of the framework in [23]. Figures 4.1 and 4.2 show how full information and partial information can change (relative to each other) for changes in ρ, and with the remaining parameters being fixed as $\epsilon = .01$, $\theta = .05$, $\sigma = .2$, $\beta = .2$, and $T = 1$. Also plotted is the **certainty equivalent,**

$$CE(t,x) = U^{-1}(V^{(0)}(t,x) + \sqrt{\varepsilon}V^{(1)}(t,x)),$$

which tells us how much cash must be held in the risk-free bank account to make the investor utility indifferent to investing in the risky asset. These plots show us how ρ that is close to zero in absolute value results in very slow diffusion in the zero-order terms, due to the diffusion coefficient $\bar{\lambda}$ being monotone increasing in $|\rho|$. Indeed, in Figure 4.1 we see partial information value functions close to their terminal conditions and partial information CE's close to the diagonal when $\rho = -.3$, but in Figure 4.2 we see that these quantities have diffused further and shifted upward.

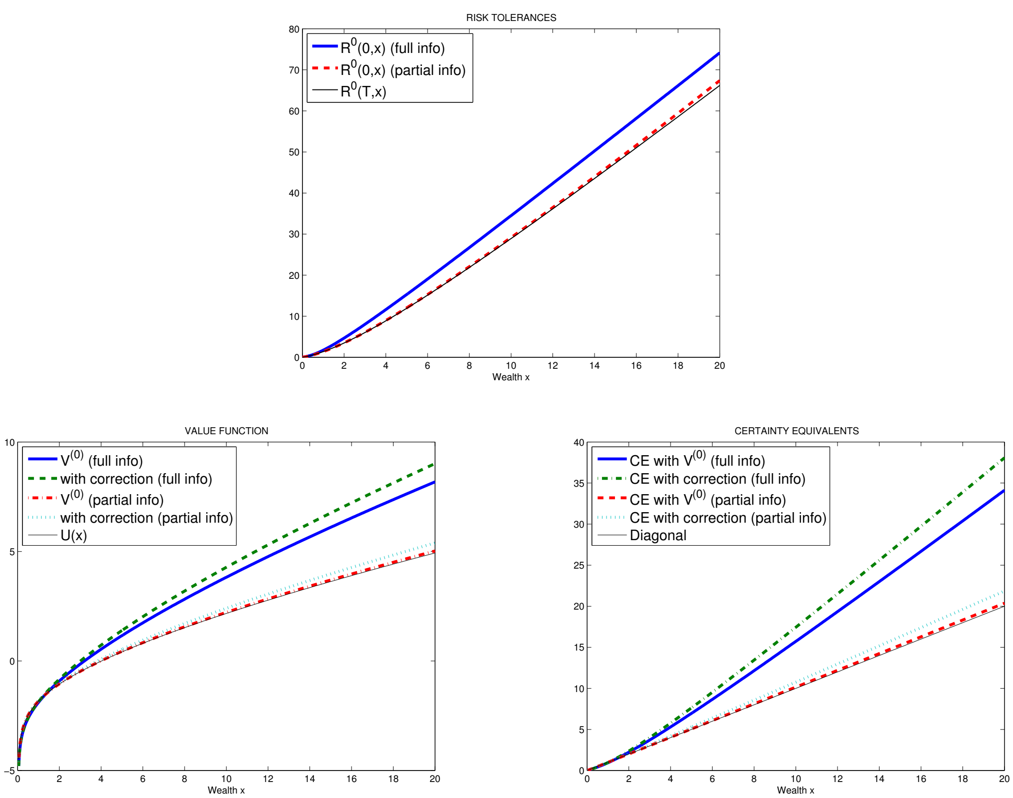

FIG. 4.1. Case $\rho = -.30$ and $\epsilon = .01$, with remaining parameters $\theta = .05$, $\sigma = .2$, $\beta = .2$, and $T = 1$.

4.4. The Practical Strategy

Recall the optimal strategy π_t^* in (3.3), and notice that it is the solution to an unconstrained optimization problem. It might be considered impractical to implement such a strategy because it will require the trades to track the fast motion in y. Alternatively, it would be more practical to perform a constrained optimization wherein the optimal strategy, call it $\bar{\pi}_t$, has an expansion with

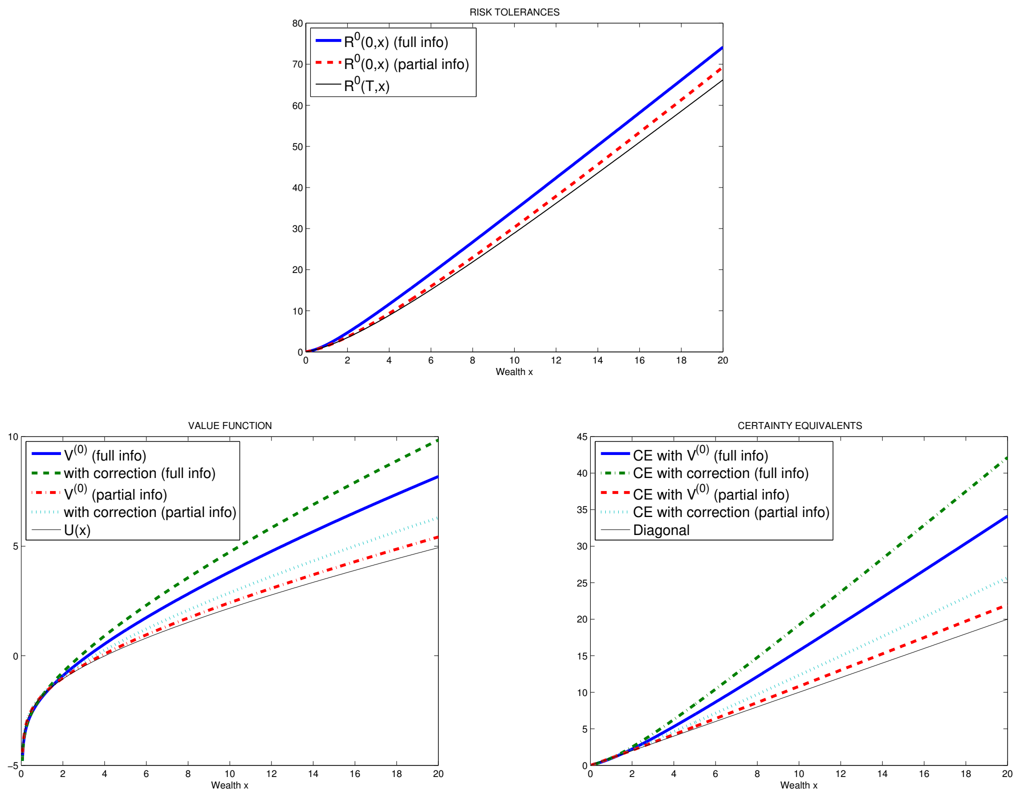

FIG. 4.2. Case $\rho = -.60$ and $\epsilon = .01$, with remaining parameters $\theta = .05$, $\sigma = .2$, $\beta = .2$, and $T = 1$.

a zero-order term that does not depend on y. We call this the practical strategy, but find there to be significant loss in utility because it has the effect of taking $\rho = 0$.

In terms of a PDE, the practical strategy yields a value function $\bar{V}^\varepsilon$ such that

$$\bar{V}^\varepsilon_t + \frac{1}{2}\left(\frac{\Sigma^\varepsilon(t)}{\sigma}\right)^2 \bar{V}^\varepsilon_{yy} + \frac{1}{\sqrt{\varepsilon}}\frac{\beta\rho\Sigma^\varepsilon(t)}{\sigma}\bar{V}^\varepsilon_{yy} + \frac{1}{\varepsilon}\mathcal{L}_0\bar{V}^\varepsilon$$
$$+\max_{\bar{\pi}}\left\{\frac{1}{2}\sigma^2\bar{\pi}^2\bar{V}^\varepsilon_{xx} + \bar{\pi}\left(y\bar{V}^\varepsilon_x + \frac{\Sigma^\varepsilon(t)\sqrt{\varepsilon}+\sigma\beta\rho}{\sqrt{\varepsilon}}\bar{V}^\varepsilon_{xy}\right)\right\} = 0, \tag{4.10}$$

where $\bar{\pi}$ is an element among the constrained set of strategies for which the zero-order term in the strategy expansion does not depend on y. In other words, we write the expansion of $\bar{\pi}_t$ in powers of ε:

$$\bar{\pi}_t = \bar{\pi}^{(0)}_t + \sqrt{\varepsilon}\bar{\pi}^{(1)}_t + \varepsilon\bar{\pi}^{(2)}_t + \ldots,$$

and assume a priori that the zero-order term does not depend on y. Furthermore, we expand the practical strategy's value function:

$$V^\varepsilon = \bar{V}^{(0)} + \sqrt{\varepsilon}\bar{V}^{(1)} + \varepsilon\bar{V}^{(2)} + \dots.$$

Inserting the expansions of $\bar{\pi}$ and $\bar{V}^\varepsilon$ into (4.10) (and recall from (2.8) that $\Sigma^\varepsilon(t)$ is regular in $\sqrt{\varepsilon}$), we find from the terms of order ε^{-1} that $\bar{V}^{(0)}$ is constant in y, we find from the terms of order $\varepsilon^{-1/2}$ that $\bar{V}^{(1)}$ is constant in y, and from the terms of order ε^0 we find the following equation:

$$\max_{\bar{\pi}^{(0)}} \left\{ \bar{V}_t^{(0)} + \mathcal{L}_0 \bar{V}^{(2)} + \frac{1}{2}\sigma^2 \left(\bar{\pi}^{(0)}\right)^2 \bar{V}_{xx}^{(0)} + \bar{\pi}^{(0)} y \bar{V}_x^{(0)} \right\} = 0. \tag{4.11}$$

Now, for the optimal $\bar{\pi}_t^{(0)}$ not depending on y, the solvability condition for this PDE (by the Fredholm alternative) is

$$\bar{V}_t^{(0)} + \frac{1}{2}\sigma^2 \left(\bar{\pi}_t^{(0)}\right)^2 \bar{V}_{xx}^{(0)} + \bar{\pi}_t^{(0)} \langle y \rangle \bar{V}_x^{(0)} = 0.$$

Hence, because $\langle y \rangle = \theta$, it follows that the practical solution's zero-order value function solves

$$\bar{V}_t^{(0)} + \max_{\bar{\pi}^{(0)}} \left\{ \frac{1}{2}\sigma^2 \left(\bar{\pi}^{(0)}\right)^2 \bar{V}_{xx}^{(0)} + \bar{\pi}^{(0)} \theta \bar{V}_x^{(0)} \right\} = 0, \tag{4.12}$$

which is the Merton value function with Sharpe ratio $\frac{\theta}{\sigma}$. In terms of the Sharpe ratio from the unconstrained problem,

$$\frac{\theta}{\sigma} < \sqrt{\bar{\lambda}^2} \qquad \text{for } \rho \neq 0.$$

Based on the numerics from Section 4.3, this means that there will be a significant loss in utility if the practical strategy is used with ρ away from zero.

Surprisingly, this calculation suggests that the best sub-optimal strategy that does not depend on tracking the fast moving Y or its filter $\hat{Y}$ is *not* the Merton strategy with the constant Sharpe ratio $\bar{\lambda}$, but instead to use the Sharpe ratio $\frac{\theta}{\sigma}$. In other words, this Sharpe ratio corresponds to the partial information case at $\rho = 0$, and that for non-zero ρ, tracking gives an order one utility enhancement.

5. Summary

We have explored a portfolio optimization problem involving a commodities market wherein the yield curve is only partially observed. This partial information portfolio problem requires us to include the Kalman filter for the yields in the conditioning for the value function. We calculated the small-ε asymptotics behaviour of the Kalman filter, and based on the the HJB equation we calculated the small-ε expansions of the optimised value function. The methodology is useful for problems involving general utility functions, because explicit formula for solutions to the HJB equation are not available. In the numerics we compared the partial information problem to the full information problem, and found there to be an information premium that depends in large part on the correlation parameter ρ. We also explored some practical strategies that do not trade the asset in fast time scales.

Appendix A. Derivations.

A.1. Proof of Proposition 3.3

Substituting $R^{(0)} = -V_x^{(0)}/V_{xx}^{(0)}$ from (3.8) into (3.14) and differentiating with respect to x yields

$$V_{tx}^{(0)} = \frac{\bar{\lambda}^2}{2}\left(R^{(0)}\right)^2 V_{xxx}^{(0)} + \bar{\lambda}^2 R^{(0)} R_x^{(0)} V_{xx}^{(0)}.$$

But from $R^{(0)}V_{xx}^{(0)} = -V_x^{(0)}$ we have $\left(R^{(0)}\right)^2 V_{xxx}^{(0)} = (R_x^{(0)}+1)V_x^{(0)}$, and so the above expression becomes

$$V_{tx}^{(0)} = \frac{\bar{\lambda}^2}{2}(R_x^{(0)}+1)V_x^{(0)} - \bar{\lambda}^2 R_x^{(0)} V_x^{(0)},$$

which gives

$$V_{tx}^{(0)} = -\frac{\bar{\lambda}^2}{2}(R_x^{(0)}-1)V_x^{(0)}. \tag{A.1}$$

Next, differentiating (3.8) with respect to t gives

$$R_t^{(0)} = -\frac{V_{tx}^{(0)}}{V_{xx}^{(0)}} + \frac{V_x^{(0)}}{(V_{xx}^{(0)})^2}V_{txx}^{(0)}. \tag{A.2}$$

Differentiating (A.1) with respect to x yields

$$V_{txx}^{(0)} = -\frac{\bar{\lambda}^2}{2}(R_{xx}^{(0)}-1)V_{xx}^{(0)} - \frac{\bar{\lambda}^2}{2}R_{xx}^{(0)}V_x^{(0)},$$

and substituting this and (A.1) into (A.2) yields (3.24).

A.2. Proof of Lemma 3.4

For any smooth $w(t,x)$, we compute

$$\begin{aligned}
D_2D_1w - D_1D_2w &= (R^{(0)})^2\frac{\partial^2}{\partial x^2}(R^{(0)}w_x) - R^{(0)}\frac{\partial}{\partial x}((R^{(0)})^2 w_{xx})\\
&= (R^{(0)})^2(R_{xx}^{(0)}w_x + 2R_x^{(0)}w_{xx} + R^{(0)}w_{xxx}) - R^{(0)}(2R^{(0)}R_x^{(0)}w_{xx} + (R^{(0)})^2 w_{xxx})\\
&= (R^{(0)})^2 R_{xx}^{(0)} w_x.
\end{aligned}$$

Then

$$\begin{aligned}
\mathcal{L}_{t,x}D_1w &= \left(\frac{\partial}{\partial t} + \frac{1}{2}\bar{\lambda}^2 D_2 + \bar{\lambda}^2 D_1\right)D_1 w\\
&= D_1\left(\frac{\partial}{\partial t} + \frac{1}{2}\bar{\lambda}^2 D_2 + \bar{\lambda}^2 D_1\right)w + \left(R_t + \frac{1}{2}\bar{\lambda}^2(R^{(0)})^2 R_{xx}^{(0)}\right)w_x\\
&= D_1\mathcal{L}_{t,x}w,
\end{aligned}$$

where we have used (3.24).

A.3. Derivation of Formulas (3.27)-(3.28)

From the definition of μ in (3.13) we can deduce the relation $\frac{1}{2}\beta^2\rho^2\mu' - (\theta - y)\mu = 0$, from which equation (3.21) can be equivalently written as

$$\frac{\beta^2\rho^2}{2}\left(\phi'\mu\right)' = \left(\frac{y^2}{\sigma^2} - \bar{\lambda}^2\right)\mu,$$

and consequently, since $\left(\frac{1}{\mu(y)}\right)' = \frac{2(y-\theta)}{\beta^2\rho^2}\frac{1}{\mu(y)}$, we have

$$\begin{aligned}
\phi'(y) &= \frac{2}{\beta^2\rho^2\mu(y)}\int_{-\infty}^{y}\left(\frac{u^2}{\sigma^2} - \bar{\lambda}^2\right)\mu(du) \\
\phi''(y) &= \frac{2}{\beta^2\rho^2}\left(\frac{y^2}{\sigma^2} - \bar{\lambda}^2\right) + \frac{2(y-\theta)}{\beta^2\rho^2}\phi'(y).
\end{aligned}$$

By definition we have $\left\langle \frac{y^2}{\sigma^2} - \bar{\lambda}^2 \right\rangle = 0$, and therefore, one is left with

$$\left\langle \phi'' \right\rangle = \frac{2}{\rho^2\beta^2}\left\langle y\phi' \right\rangle - \frac{2\theta}{\rho^2\beta^2}\left\langle \phi' \right\rangle.$$

Now, the average of ϕ' is

$$\begin{aligned}
\beta\rho\left\langle \phi' \right\rangle &= \frac{2}{\beta\rho}\left\langle \frac{1}{\mu}\int_{-\infty}^{y}\left(\frac{u^2}{\sigma^2} - \bar{\lambda}^2\right)\mu(du) \right\rangle \\
&= \frac{2}{\beta\rho}\int_{-\infty}^{\infty}\int_{-\infty}^{y}\left(\frac{u^2}{\sigma^2} - \bar{\lambda}^2\right)\mu(u)du dy \\
&= -\frac{2}{\beta\rho}\int_{-\infty}^{\infty} u\left(\frac{u^2}{\sigma^2} - \bar{\lambda}^2\right)\mu(u)du \qquad \text{(using integration by parts)} \\
&= -\frac{2}{\beta\rho\sigma^2}\left\langle y^3 \right\rangle + \frac{2\bar{\lambda}^2}{\beta\rho}\left\langle y \right\rangle \\
&= -\frac{2}{\beta\rho\sigma^2}\left(3\theta\frac{\beta^2\rho^2}{2} + \theta^3\right) + \frac{2\theta}{\beta\rho\sigma^2}\left(\frac{\beta^2\rho^2}{2} + \theta^2\right) \\
&= -\frac{2\theta\beta\rho}{\sigma^2},
\end{aligned}$$

and the average of $y\phi'$ is

$$\begin{aligned}
\beta\rho\left\langle y\phi'\right\rangle &= \frac{2}{\beta\rho}\left\langle \frac{y}{\mu}\int_{-\infty}^{y}\left(\frac{u^2}{\sigma^2}-\bar{\lambda}^2\right)\mu(du)\right\rangle \\
&= \frac{2}{\beta\rho}\int_{-\infty}^{\infty} y\int_{-\infty}^{y}\left(\frac{u^2}{\sigma^2}-\bar{\lambda}^2\right)\mu(u)dudy \\
&= -\frac{1}{\beta\rho}\int_{-\infty}^{\infty} u^2\left(\frac{u^2}{\sigma^2}-\bar{\lambda}^2\right)\mu(u)du \qquad \text{(using integration by parts)} \\
&= -\frac{1}{\beta\rho\sigma^2}\left\langle y^4\right\rangle + \frac{\bar{\lambda}^2}{\beta\rho}\left\langle y^2\right\rangle \\
&= -\frac{1}{\beta\rho\sigma^2}\left(3\frac{\beta^4\rho^4}{4}+6\theta^2\frac{\beta^2\rho^2}{2}+\theta^4\right)+\frac{1}{\beta\rho\sigma^2}\left(\frac{\beta^2\rho^2}{2}+\theta^2\right)^2 \\
&= -\frac{1}{\beta\rho\sigma^2}\left(\frac{\beta^4\rho^4}{2}+4\theta^2\frac{\beta^2\rho^2}{2}\right) \\
&= -\frac{\beta\rho}{2\sigma^2}\left(\beta^2\rho^2+4\theta^2\right),
\end{aligned}$$

where explicit expressions for $\left\langle y^k\right\rangle$ for $k=0,1,2,3,4$ are obtained because they are Gaussian moments.

REFERENCES

[1] A. Bain and D. Crisan. *Fundamentals of Stochastic Filtering*. Springer, New York, NY, 2009.

[2] N. Bäuerle and U. Rieder. Portfolio optimization with unobservable Markov-modulated drift process. *Journal of Applied Probability*, 42(2):362–378, 2005.

[3] N. Bäuerle, S. Urban, and L. Veraart. The relaxed investor with partial information. *SIAM Journal on Financial Mathematics*, 3(1):304–327, 2012.

[4] F. Black. Investment and consumption through time. Reprinted as "Individual investment and consumption under uncertainty" in Luskin, D. L. (Ed.), Portfolio Insurance: A guide to dynamic hedging. Wiley, pp. 207-225., 1968.

[5] F. Black and M. Scholes. The Pricing of Options and Corporate Liabilities. *J. Political Econ.*, 81:637–659, 1973.

[6] S. Brendle. Portfolio selection under incomplete information. *Stochastic Processes and their Applications*, 116(5):701 – 723, 2006.

[7] R. Carmona. From Markovian to partially observable models. In *Indifference Pricing*. Princeton University Press, 2009.

[8] J.-P. Fouque, G. Papanicolaou, R. Sircar, and K. Solna. *Multiscale Stochastic Volatility for Equity, Interest Rate, and Credit Derivatives*. Cambridge University Press, 2011.

[9] J.-P. Fouque, R. Sircar, and T. Zariphopoulou. Portfolio optimization & stochastic volatility asymptotics. 2013. Submitted for publication.

[10] C. Huang and R. Litzenberger. *Foundations of Financial Economics*. Prentice Hall, Englewood Cliffs, NJ, 1988.

[11] M. Jonsson and R. Sircar. Optimal investment problems and volatility homogenization approximations. In A. Bourlioux, M. Gander, and G. Sabidussi, editors, *Modern Methods in Scientific Computing and Applications*, volume 75 of *NATO Science Series II*, pages 255–281. Kluwer, 2002.

[12] M. Jonsson and R. Sircar. Partial hedging in a stochastic volatility environment. *Mathematical Finance*, 12(4):375–409, October 2002.

[13] I. Karatzas and S. Shreve. *Methods of Mathematical Finance*. Springer-Verlag, 1998.

[14] T. S. Kim and E. Omberg. Dynamic nonmyopic portfolio behavior. *Review of Financial Studies*, 9(1):141–61, 1996.

[15] R. C. Merton. Lifetime portfolio selection under uncertainty: the continous-time case. *Rev. Econom. Statist.*, 51:247–257, 1969.

[16] M. Musiela and T. Zariphopoulou. Portfolio choice under space-time monotone performance criteria. *SIAM Journal on Financial Mathematics*, 1:326–365, 2010.

[17] A. Papanicolaou and K. Spiliopoulos. Filtering the maximum likelihood for multiscale problems. 2013. Submitted for publication.

[18] J. Park, B. Rozovsky, and R. Sowers. Efficient nonlinear filtering of a singularly perturbed stochastic hybrid system. *LMS J. Computational Mathematics*, 14:254–270, 2011.

[19] H. Pham. *Continuous-time Stochastic Control and Optimization with Financial Applications*. Springer, 2009.

[20] L.C.G. Rogers. *Optimal Investment*. Springer, 2013.

[21] J. Sass and U. Haussmann. Optimizing the terminal wealth under partial information: The drift process as a continuous time Markov chain. *Finance and Stochastics*, 8(4):553–577, November 2004.

[22] W. Schachermayer. Introduction to the mathematics of financial markets. In Pierre Bernard, editor, *Lectures on Probability Theory and Statistics, Saint-Fleur summer school 2000*, number 1816 in Lecture Notes in Mathematics, pages 111–177. Springer Verlag, 2003.

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