

Optimal hedging and parameter uncertainty

MICHAEL MONOYIOS[†]

Mathematical Institute, University of Oxford

[Received on 8 September 2006; accepted on 15 March 2007]

We explore the impact of drift parameter uncertainty in a basis risk model, an incomplete market in which a claim on a nontraded asset is optimally hedged using a correlated traded stock. Using analytic expansions for indifference prices and hedging strategies, we develop an efficient procedure to generate terminal hedging error distributions when the hedger has erroneous estimates of the drift parameters. These show that the effect of parameter uncertainty is occasionally benign, but often very destructive. In light of this, we develop a filtering approach in which the hedger updates her parameter estimates from observations of the asset prices, and we find an analytic solution to the hedger's combined filtering and control problem in the case that the drift of the traded asset is known with certainty.

1. Introduction

This paper examines the problem of drift parameter uncertainty on the optimal hedging of a claim in an incomplete market. We study a 'basis risk' model (Davis, 2006; Henderson, 2002; Monoyios, 2004b; Musiela & Zariphopoulou, 2004) in which a claim on a nontraded asset Y is optimally hedged using a correlated traded asset S , with correlation $\rho \in [-1, 1]$. Our contributions are threefold. First, we improve the analytic approach initiated in Monoyios (2004b) for generating the terminal hedging error distribution, associated with the utility-based hedging of the claim, over simulated asset price paths. Second, we use this approach to examine the impact of drift parameter misestimation on the terminal hedging error. Third, we propose a filtering approach to the parameter uncertainty problem, incorporating Bayesian learning into the drift parameter estimation.

In a complete market, such as the Black & Scholes (1973) (BS) setting, perfect hedging of a claim does not require estimation of the stock price drift μ in the stochastic differential equation (SDE) (1) for the stock S . As is well-known, estimation of the volatility σ is possible with reasonable confidence given sufficient data, but as discussed by Rogers (2001) and as we discuss in Section 3, it is virtually impossible to have confidence in an estimate of μ . This has serious implications for optimal investment rules such as the classical Merton (1969, 1971) policy, where knowledge of μ is needed to compute the optimal proportion of wealth to assign to the risky asset. By the same token, drift parameter uncertainty will also have a bearing on hedging strategies for derivatives in 'incomplete' markets when 'hedging' becomes synonymous with 'investment' since there is, in general, no preference-free way to perfectly replicate a claim.

In a lognormal basis risk model, it was shown in Monoyios (2004b) that an optimal strategy, based on exponential utility maximization, gave superior hedging performance than a 'naive' strategy (a BS-style hedge) that assumes the traded asset S is a good proxy for the nontraded asset Y (tacitly assuming that for high correlation, the approximation $\rho \approx 1$ is a good one). This was done by generating the distribution of hedging error associated with hedging the claim over many simulated asset price paths.

[†]Email: monoyios@maths.ox.ac.uk

With optimal hedging, the hedging error distribution was found to have, among other features, a higher median relative to that obtained with the BS-style strategy. This reflects the greater frequency of profits over losses generated by optimal hedging.

Unfortunately, however, the optimal strategy requires knowledge of the drift parameters of the diffusions S and Y , as well as the volatilities and correlation. The drift parameters are notoriously difficult to estimate with precision, as already remarked. In contrast, the naive strategy only requires knowledge of the volatilities of S and Y . Although estimation of volatility and correlation is itself not perfect, it can be improved with more observations much more rapidly than the drift estimation. For this reason, we focus on the severe problem of drift parameter uncertainty.

We investigate the effect of this parameter uncertainty on the hedging error distribution by supposing that the hedger misestimates the drift parameters and we look at the effect on the median hedging error distribution. This reveals that the effect of parameter uncertainty can indeed be severe. It can lead to large reductions in median hedging error distribution, depending on the size and direction of the misestimation. The direction of the misestimation can be such as to make the agent too aggressive, leading to selling the claim at too low a price and leading to frequent hedging losses. Conversely, if the sign of the estimation error is such as to make the agent more conservative, the relative frequency of hedging losses is reduced, but at the cost of selling the claim at too high a price, making the agent into an uncompetitive market maker.

We then initiate a filtering approach to this problem, allowing the agent to use observations of the asset prices to update her estimates of the drift parameters. This approach shows promise, though analytic results for indifference prices and hedging strategies are not generally available, as the dimension of the problem is increased by incorporating parameter uncertainty. However, we are able to develop an analytic approach in the special case when the agent is uncertain about the drift of the nontraded asset only. The effect of such Bayesian learning on the hedging error distribution will be a topic of future papers.

Partial information problems, in which agents do not have precise knowledge of drift parameters, have received some attention in the context of optimal investment problems such as the classical Merton (1969, 1971) problems, but this is the first attempt to examine such uncertainty in the context of optimal hedging of derivatives. Rogers (2001) investigated the disutility arising from parameter uncertainty versus that from discrete portfolio rebalancing for a Merton investor seeking to maximize expected utility of wealth or consumption (the theoretically optimal strategy requires continuous portfolio revision). Rogers finds that parameter uncertainty outweighs the effect of rebalancing the portfolio infrequently. Lakner (1995, 1998) uses a dual approach to examine the impact of drift parameter uncertainty on the Merton problem, while Brennan (1998) derives the asset price dynamics under the observation filtration and characterizes the optimal investment rule in terms of the solution of a Hamilton–Jacobi–Bellman (HJB) equation. Our approach is also based on stochastic control and a HJB equation, though it differs from Brennan’s in some significant ways. Even though we incorporate uncertainty in the drifts of two assets S and Y (Brennan deals with the Merton problem involving a single stock), we are able to treat the problem such that the resulting HJB equation has only one extra dimension compared to the full information case. Remarkably, in the case where the investor is uncertain only about the drift of Y , we are able to retain the same dimensionality as the full information problem. Future work will exploit these features to investigate in detail the effect on the hedging error distribution.

The rest of the paper is organized as follows: In Section 2, we outline the basis risk model. We recall previous results (Monoyios, 2004b) for the claim’s indifference price and optimal hedging strategy, and derive improved analytic formulae (compared to Monoyios, 2004b) for these objects, as well as a SDE for the ‘residual risk’ (or hedging error) process, allowing for an extremely efficient generation of the terminal hedging error distribution via simulation. We demonstrate the superiority of optimal over naive

hedging, given perfect knowledge of parameter values. In Section 3, we explore the impact of drift parameter uncertainty. We show that one needs hundreds of years of price data to have any confidence in drift parameter estimates, and we generate terminal hedging error distributions when the hedger misestimates the drifts of S and Y . Section 4 proposes a filtering algorithm to incorporate Bayesian learning into the drift parameter estimation, and Section 5 concludes.

2. Basis risk model

We consider a basis risk model, involving a traded asset S and a nontraded asset Y , following correlated log-Brownian motions, as in Davis (2006), Monoyios (2004b), Henderson (2002) and Musiela & Zariphopoulou (2004). On a complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} := (\mathcal{F}_t)_{0 \leq t \leq T}, P)$, the stock price process $S := (S_t)_{0 \leq t \leq T}$ is

$$dS_t = \mu S_t dt + \sigma S_t dB_t \quad (1)$$

and the nontraded asset price process $Y := (Y_t)_{0 \leq t \leq T}$ is

$$dY_t = \nu Y_t dt + \beta Y_t dW_t, \quad (2)$$

where μ, σ, ν and β are constants and the Brownian motions B and W have correlation $\rho \in [-1, 1]$:

$$dB_t dW_t = \rho dt, \quad W = \rho B + \sqrt{1 - \rho^2} Z,$$

where B and Z are independent Brownian motions.

An agent may trade the stock in a self-financing fashion, leading to the portfolio wealth process $X^\pi \equiv X := (X_t)_{0 \leq t \leq T}$ satisfying

$$dX_t = r X_t dt + \sigma \pi_t (\lambda dt + dB_t), \quad \lambda := (\mu - r)/\sigma,$$

where $r \geq 0$ is a constant rate of interest and $\pi := (\pi_t)_{0 \leq t \leq T}$ is the wealth in the stock, representing the agent's trading strategy. An admissible trading strategy is one satisfying $\int_0^T \pi_t^2 dt < \infty$ almost surely. Denote the set of such strategies by A .

A European claim on Y pays $h(Y_T)$ at time T . The claim cannot be perfectly replicated using a portfolio in S unless $|\rho| = 1$, so the market is incomplete. Suppose the agent takes a position in n claims at some time $t \leq T$. The utility-indifference approach to valuing such a position in the claim has been analysed by Davis (2006), Henderson (2002) and Musiela & Zariphopoulou (2004) among others. This strand of research culminated with Monoyios (2004b), who analysed the optimal hedging of the claim associated with the utility-indifference approach. This showed that optimal hedging produces a hedge profit and loss distribution that had a higher median hedging error than a naive strategy based on the $\rho \rightarrow 1$ limit of the optimal hedging formulae, and which therefore tacitly assumes that the traded asset is a good proxy for the nontraded asset. In this sense, optimal hedging was shown to be beneficial, even when the correlation is close to 1.

The major caveat to the above arguments is that the optimal strategies require knowledge of the drift parameters μ and ν of the assets, as well as the volatilities and correlation. In contrast, the naive strategies require only knowledge of the volatilities. This casts doubt on the true efficacy of the optimal schemes in the face of the severe parameter uncertainty associated with drift estimation.

We introduce the well-known minimal martingale measure Q^M of Föllmer & Schweizer (1991) which will feature in many of our formulae. The probability measure Q^M has density process with

respect to P given by

$$\frac{dQ^M}{dP} \Big|_{\mathcal{F}_t} = \mathcal{E}(-\lambda \cdot B)_t, \quad 0 \leq t \leq T.$$

Under Q^M , (S, Y) follow

$$\begin{aligned} dS_t &= rS_t dt + \sigma S_t dB_t^{Q^M}, \\ dY_t &= (v - \beta\rho\lambda)Y_t dt + \beta Y_t dW_t^{Q^M}, \end{aligned}$$

where B^{Q^M} and W^{Q^M} are correlated Brownian motions under Q^M . The discounted stock price S is a local Q^M -martingale (the drift of S is r), but this is not the case for the nontraded asset unless we have the perfect correlation case $\rho = 1$. In this case, Y is effectively a traded asset (as Y_t is then a function of S_t), so the Q^M -drift of Y is r . Therefore, given σ and β , in the $\rho = 1$ case the drifts are related by

$$\theta := \frac{v - r}{\beta} = \frac{\mu - r}{\sigma} =: \lambda.$$

In this case, the market becomes complete, and perfect hedging is possible. It is easy to show that with $\rho = 1$ so that $W = B$, we have

$$\begin{aligned} Y_t &= Y_0 \left(\frac{S_t}{S_0} \right)^{\beta/\sigma} e^{ct} \\ c &= \left(r + \frac{1}{2}\sigma\beta \right) \left(1 - \frac{\beta}{\sigma} \right). \end{aligned}$$

Let the claim price process be $v(t, Y_t)$, $0 \leq t \leq T$, where $v: [0, T] \times \mathbb{R}^+ \rightarrow \mathbb{R}^+$ is smooth enough to apply the Itô formula so that

$$dv(t, Y_t) = [v_t(t, Y_t) + \mathcal{A}^Y v(t, Y_t)]dt + \beta Y_t v_y(t, Y_t) dW_t,$$

where \mathcal{A}^Y is the generator of the process Y in (2). The replication conditions are

$$X_t = v(t, Y_t), \quad 0 \leq t \leq T, \quad dX_t = dv(t, Y_t).$$

Standard arguments then show that to perfectly hedge the claim, one must hold Δ_t shares of S at $t \in [0, T]$, given by

$$\Delta_t = \frac{\beta Y_t}{\sigma S_t} \frac{\partial v}{\partial y}(t, Y_t), \quad (3)$$

and the claim pricing function $v(t, y)$ satisfies

$$\begin{aligned} v_t(t, y) + (v - \beta\lambda)yv_y(t, y) + \frac{1}{2}\beta^2 y^2 v_{yy}(t, y) - rv(t, y) &= 0, \\ v(T, y) &= h(y). \end{aligned}$$

But with $\rho = 1$, $v - \beta\lambda = r$, so we get the BS partial differential equation (PDE), and

$$v(t, Y_t) = \text{BS}(t, Y_t),$$

where $\text{BS}(t, y)$ denotes the BS option pricing formula at time t with underlying asset price y .

2.1 Utility-indifference valuation and hedging

Now, suppose the correlation is not perfect so that the market is incomplete. We embed the problem in a utility-maximization framework in a manner that is by now classical. Let the agent have risk preferences expressed via the exponential utility function

$$U(x) = -\exp(-\alpha x), \quad x \in \mathbb{R}, \quad \alpha > 0.$$

The agent maximizes the expected utility of terminal wealth at time T with a random endowment of n units of claim pay-off:

$$J(t, x, y; \pi) = E[U(X_T + nh(Y_T)) | X_t = x, Y_t = y].$$

The value function is $u^{(n)}(t, x, y) \equiv u(t, x, y)$, defined by

$$u(t, x, y) := \sup_{\pi \in A} J(t, x, y; \pi), \quad (4)$$

$$u(T, x, y) = U(x + nh(y)). \quad (5)$$

Denote the optimal trading strategy that achieves the supremum in (4) by $\pi^* \equiv \pi^{*,n}$, and denote the optimal wealth process by $X^* \equiv X^{*,n}$.

We make the following assumption to ensure that we get a meaningful optimization problem in (4).

ASSUMPTION 1 The random endowment $nh(Y_T)$ is bounded below.

This assumption ensures that the maximum utility in (4) is well-defined, and means that we can cover the cases of short and long put positions and long call positions. The case of valuing a short call position in this framework is a topic for future research.

The following definitions of utility-based price and hedging strategy are now standard (see Monoyios, 2004a,b; Musiela & Zariphopoulou, 2004, for instance).

DEFINITION 1 (Indifference price) The indifference price per claim at $t \in [0, T]$, given $X_t = x, Y_t = y$ and $p(t, x, y) \equiv p^{(n)}(t, x, y)$, is defined by

$$u^{(n)}(t, x - np^{(n)}(t, x, y), y) = u^{(0)}(t, x, y).$$

We allow for possible dependence on t, x and y of $p^{(n)}$ in the above definition, but with exponential preferences it turns out that there is no dependence on x .

DEFINITION 2 (Optimal hedging strategy) The optimal hedging strategy for n units of the claim is $\pi^{(H)} := (\pi_t^{(H)})_{0 \leq t \leq T}$ given by

$$\pi_t^{(H)} := \pi_t^{*,n} - \pi_t^{*,0}, \quad 0 \leq t \leq T.$$

The solution to the optimization problem (4) is well-known, using the so-called distortion technique (Zariphopoulou, 2001). See Monoyios (2004b) for more details.

The HJB equation for the value function u is

$$u_t + rxu_x + \mathcal{A}^Y u - \frac{(\lambda u_x + \rho \beta y u_{xy})^2}{2u_{xx}} = 0.$$

The optimal trading strategy π^* is given by $\pi_t^* = \Pi^*(t, X_t^*, Y_t)$, where the function $\Pi^*: [0, T] \times \mathbb{R} \times \mathbb{R}^+$ is given by

$$\Pi^*(t, x, y) := - \left(\frac{\lambda u_x + \rho \beta y u_{xy}}{\sigma u_{xx}} \right). \quad (6)$$

We have the following well-known representation for the value function and indifference price.

PROPOSITION 1 (Henderson, 2002; Monoyios, 2004b; Musiela & Zariphopoulou, 2004) The value function $u \equiv u^{(n)}$ and indifference price $p \equiv p^{(n)}$, given $X_t = x$ and $Y_t = y$ for $t \in [0, T]$, are given by

$$\begin{aligned} u^{(n)}(t, x, y) &= -e^{-\alpha b(t, T)x - \frac{1}{2}\lambda^2(T-t)} [F(t, Y)]^{1/(1-\rho^2)}, \\ F(t, y) &= E^{\mathcal{Q}^M} [\exp(-\alpha(1-\rho^2)nh(Y_T)) | Y_t = y], \\ p^{(n)}(t, y) &= -\frac{1}{b(t, T)\alpha(1-\rho^2)n} \log F(t, y), \end{aligned} \quad (7)$$

where $b(t, T) := e^{r(T-t)}$.

The function $F(t, y)$ satisfies a linear PDE by virtue of the stochastic representation (7) and the Feynman–Kac theorem. The indifference pricing function $p(t, y) \equiv p^{(n)}(t, y)$ then satisfies

$$p_t + (v - \beta\rho\lambda)y p_y + \frac{1}{2}\beta^2 y^2 p_{yy} - rp - \frac{1}{2}\beta^2 y^2 n b(t, T)\alpha(1-\rho^2)(p_y)^2 = 0.$$

Given the above results, it is easy to show that the expression (6) for the optimal control loses its dependence on x and simplifies to

$$\Pi^*(t, y) := \frac{1}{\alpha\sigma b(t, T)} \left(\lambda + \frac{\rho\beta y}{1-\rho^2} \frac{F_y}{F} \right).$$

Then, applying Definition 2 gives the optimal hedging strategy for a position in n claims (see Monoyios, 2004b, for further details of this derivation).

PROPOSITION 2 The optimal hedging strategy for a position in n claims is to hold $\Delta_t^{(H)}$ shares at $t \in [0, T]$ given by

$$\Delta_t^{(H)} = -n\rho \frac{\beta}{\sigma} \frac{Y_t}{S_t} \frac{\partial p^{(n)}}{\partial y}(t, Y_t).$$

We note that if $n = 1$ and $\rho = 1$, we recover the perfect delta hedge (3), and that the claim price then satisfies the BS PDE.

2.2 The residual risk process

Suppose the agent trades n claims at time 0 for price $p^{(n)}(0, Y_0)$ per claim. The agent hedges this position over $[0, T]$ using the strategy $(\Delta_t^{(H)})_{0 \leq t \leq T}$. Her overall position has value process $V := (V_t)_{0 \leq t \leq T}$ given by $V_t = X_t^{(H)} + np^{(n)}(t, Y_t)$ so that

$$dV_t = dX_t^{(H)} + n dp^{(n)}(t, Y_t), \quad (8)$$

where $X^{(H)} = (X_t^{(H)})_{0 \leq t \leq T}$ is the value of the hedging portfolio in S , satisfying

$$\begin{aligned} dX_t^{(H)} &= \Delta_t^{(H)} dS_t + r(X_t^{(H)} - \Delta_t^{(H)} S_t) dt, \\ X_0^{(H)} &= -np^{(n)}(0, Y_0). \end{aligned}$$

Using this in (8) along with the Itô formula and the PDE satisfied by $p^{(n)}(t, y)$, we obtain

$$dV_t = rV_t dt + \frac{1}{2}\beta^2 n^2 b(t, T) \alpha(1 - \rho^2) Y_t^2 (p_y^{(n)})^2(t, Y_t) dt + \beta n \sqrt{1 - \rho^2} Y_t p_y^{(n)}(t, Y_t) dZ_t, \quad (9)$$

with $V_0 = 0$. We call V the residual risk (or hedging error) process. The term in dZ_t , orthogonal to the Brownian increments driving the stock price, is interpreted as the unhedgeable component of risk. If $\rho = 1$, we see that the process V becomes riskless, reflecting the fact that the market incompleteness disappears in this case.

2.3 Cumulant expansions

We are interested in analysing the distribution of the terminal hedging error V_T . This is not possible in closed form, so our approach is to use the SDE (9) to simulate V over many asset price histories and compute the distribution of terminal hedging error V_T . This is a similar approach to that in [Monoyios \(2004b\)](#), but the use of the SDE (9) makes the procedure more efficient than in [Monoyios \(2004b\)](#).

To simulate V via (9) efficiently, we use analytic approximations for $p(t, y)$ and $p_y(t, y)$, in the form of power series expansions in powers of $a := -\alpha(1 - \rho^2)n$. These arise from a Taylor expansion of the indifference pricing function

$$p^{(n)}(t, y) = \frac{1}{b(t, T)a} \log E^{Q^M}[\exp(ah(Y_T)) | Y_t = y]. \quad (10)$$

For a random variable X , recall that its cumulant generating function (CGF) ([Grimmett & Stirzaker, 2001](#); [Spanos, 1999](#)) is $\Psi_X(a) := \log E \exp(aX)$. Using linearity of the expectation operator, it is not hard to see that the CGF has a Taylor expansion of the form

$$\Psi_X(a) = \sum_{j=1}^{\infty} \frac{1}{j!} k_j(X) a^j,$$

where $k_j(X) \equiv k_j$ is the ‘ j th cumulant’ of X related to the j th central moment of X as described below. Writing

$$m_j(X) := E(X^j), \quad \mu_j(X) := E[(X - m_1)^j], \quad j \in \mathbb{N},$$

for the j th raw and central moments, it is not hard to show that

$$\begin{aligned} k_1(X) &= m_1(X), \\ k_2(X) &= \mu_2(X), \\ k_3(X) &= \mu_3(X), \\ k_4(X) &= \mu_4(X) - 3\mu_2^2(X), \\ k_5(X) &= \mu_5(X) - 10\mu_2(X)\mu_3(X), \end{aligned}$$

which can also be expressed in terms of the raw moments only: for $j \in \{2, 3, 4, 5\}$,

$$\begin{aligned} k_2 &= m_2 - m_1^2, \\ k_3 &= m_3 - 3m_1m_2 + 2m_1^3, \\ k_4 &= m_4 - 4m_1m_3 - 3m_2^2 + 12m_1^2m_2 - 6m_1^4, \\ k_5 &= m_5 - 5m_1m_4 - 10m_2m_3 + 20m_1^2m_3 + 30m_1m_2^2 - 60m_1^3m_2 + 24m_1^5, \end{aligned} \quad (11)$$

where the dependence on X is suppressed for brevity.

Since the pricing function (10) is proportional to the CGF of the pay-off under the minimal measure, we have the following analytic approximation for the indifference pricing function.

PROPOSITION 3 The indifference pricing function $p^{(n)}(t, y)$ has the power series expansion

$$p^{(n)}(t, y) = \frac{1}{b(t, T)} \sum_{j=1}^5 \frac{1}{j!} k_j(h(Y_T)) a^{j-1} + O(a^5), \quad (12)$$

where $a = -\alpha(1 - \rho^2)n$ and k_j is the j th cumulant of the pay-off under Q^M , conditional on $Y_t = y$. The expansion is valid for model parameters satisfying

$$E^{Q^M}[\exp(ah(Y_T)) | Y_t = y] \leq 2. \quad (13)$$

Proof. Expand the exponential inside the expectation in (10) as a Taylor series and use linearity of the expectation operator to give

$$p^{(n)}(t, y) = \frac{1}{ab(t, T)} \log \left(1 + \sum_{j=1}^{\infty} \frac{a^j}{j!} m_j(h(Y_T)) \right),$$

where $m_j(h(Y_T)) \equiv E^{Q^M}[h^j(Y_T) | Y_t = y]$. Then, (12) follows from applying the Taylor expansion of $\log(1 + x)$. This is valid for $|x| \leq 1$. In our case, this means we require

$$\sum_{j=1}^{\infty} \frac{a^j}{j!} m_j(h(Y_T)) \leq 1,$$

which is implied by (13), and the proof is complete. \square

Using these equations, we are able to produce an accurate perturbation series for $p^{(n)}(t, y)$, as a series of BS-type formulae, which can be differentiated term by term to give an analytic approximation for $p_y(t, y)$. The leading-order term in the price expansion is Davis' (1997) marginal price. The terms in the expansion depend ultimately on the moments $m_j := E^{Q^M}[h^j(Y_T) | Y_t = y]$, $j \in \mathbb{N}$, and (in the case of $p_y(t, y)$) on their partial derivatives $\partial m_j := \partial m_j / \partial y$, $j \in \mathbb{N}$. The formulae up to order a^4 are given below in the case of a put option on the nontraded asset. The formulae are extensions of those in Monoyios (2004b): first, the power series expansions for $p^{(n)}(t, y)$ and $p_y^{(n)}(t, y)$ are extended to higher order in a than in Monoyios (2004b); second, we develop a single succinct formula for the j th moment $m_j := E^{Q^M}[h^j(Y_T) | Y_t = y]$, $j \in \mathbb{N}$, and also for its derivative $\partial m_j := \partial m_j / \partial y$, $j \in \mathbb{N}$. These follow

from the fact that under Q^M , conditional upon $Y_t = y$, $\log Y_T$ is normally distributed. With $\mathcal{N}(M, \Sigma^2)$ denoting the normal probability law with mean M and variance Σ^2 , we have

$$\begin{aligned}\log Y_T &\sim \mathcal{N}\left(\log y + b - \frac{1}{2}v^2, v^2\right), \\ b &= (r - q)(T - t), \\ q &= r - (v - \beta\rho\lambda), \\ v^2 &= \beta^2(T - t).\end{aligned}\tag{14}$$

For the optimal hedging strategy, the explicit results are obtained by differentiating (12) with respect to y , giving the following corollary. Denote by $\partial\kappa_j$ the partial derivative of κ_j with respect to y :

$$\partial\kappa_j \equiv \frac{\partial\kappa_j}{\partial y},$$

where κ_j denotes any of m_j , μ_j and k_j .

COROLLARY 1 The partial derivative of the indifference price $p^{(n)}(t, y)$ with respect to y has the power series expansion

$$\frac{\partial p^{(n)}}{\partial y}(t, y) = \frac{1}{b(t, T)} \sum_{j=1}^5 \frac{1}{j!} (\partial k_j) a^{j-1} + O(a^5).$$

The partial derivatives of the cumulants are related to μ_j and $\partial\mu_j$ by

$$\begin{aligned}\partial k_1 &= \partial m_1, \\ \partial k_2 &= \partial \mu_2, \\ \partial k_3 &= \partial \mu_3, \\ \partial k_4 &= \partial \mu_4 - 6\mu_2 \partial \mu_2, \\ \partial k_5 &= \partial \mu_5 - 10(\mu_2 \partial \mu_3 + \mu_3 \partial \mu_2).\end{aligned}$$

For $j \in \{2, 3, 4, 5\}$, these relations may recast in terms of m_j and ∂m_j :

$$\begin{aligned}\partial k_2 &= \partial m_2 - 2m_1 \partial m_1, \\ \partial k_3 &= \partial m_3 - 3(m_1 \partial m_2 + m_2 \partial m_1 - 2m_1^2 \partial m_1), \\ \partial k_4 &= \partial m_4 - 4(m_1 \partial m_3 + m_3 \partial m_1) - 6m_2 \partial m_2 \\ &\quad + 12m_1(m_1 \partial m_2 + 2m_2 \partial m_1 - 2m_1^2 \partial m_1), \\ \partial k_5 &= \partial m_5 - 5(m_1 \partial m_4 + m_4 \partial m_1) - 10(m_2 \partial m_3 + m_3 \partial m_2) \\ &\quad + 20m_1(m_1 \partial m_3 + 2m_3 \partial m_1) + 30m_2(2m_1 \partial m_2 + m_2 \partial m_1) \\ &\quad - 60m_1^2(m_1 \partial m_2 + 3m_2 \partial m_1 - 2m_1^2 \partial m_1).\end{aligned}$$

The significance of the expansions is that they give easily computed closed-form approximations for the indifference price and optimal hedge. In the specific case of a put option, we have the following formulae for the raw moments of the pay-off under the minimal measure Q^M .

PROPOSITION 4 For a put option, $h(y) = (K - y)^+$, where $K > 0$ is the strike price, the j th moment $m_j := E^{Q^M}[h^j(Y_T)|Y_t = y]$, $j \in \mathbb{N}$, is given by

$$m_j = \sum_{\ell=0}^j \binom{j}{\ell} (-y)^\ell K^{(j-\ell)} \exp\left[\ell \left(b + \frac{1}{2}(\ell-1)v^2\right)\right] N(-d_1 - (\ell-1)v),$$

where $N(\cdot)$ denotes the standard cumulative normal distribution function, and where

$$d_1 = \frac{1}{v} \left[\log\left(\frac{y}{K}\right) + b + \frac{1}{2}v^2 \right],$$

$$b = (r - q)(T - t),$$

$$q = r - (v - \beta\rho\lambda),$$

$$v^2 = \beta^2(T - t).$$

Proof. For the put pay-off, we have, for $j \in \mathbb{N}$,

$$\begin{aligned} (h(Y_T))^j &= ((K - Y_T)^+)^j \\ &= (K - Y_T)^j I_{\{Y_T \leq K\}} \\ &= \sum_{\ell=0}^j \binom{j}{\ell} (-1)^\ell K^{(j-\ell)} Y_T^\ell I_{\{Y_T \leq K\}}, \end{aligned}$$

where $I_{\{Y_T \leq K\}}$ denotes the indicator function of the event $\{Y_T \leq K\}$. Given the lognormal distribution (14) of Y_T , it is easy to show that

$$E^{Q^M} \left[Y_T^\ell I_{\{Y_T \leq K\}} \mid Y_t = y \right] = y^\ell \exp\left(\ell \left(b + \frac{1}{2}(\ell-1)v^2\right)\right) N(-d_1 - (\ell-1)v),$$

from which the result follows. \square

PROPOSITION 5 Let $j \in \mathbb{N}$. For a put option pay-off, $h(y) = (K - y)^+$, ∂m_j is given by

$$\partial m_j = - \sum_{\ell=1}^j \binom{j}{\ell} (-y)^{(\ell-1)} K^{(j-\ell)} \exp\left(\ell \left(b + \frac{1}{2}(\ell-1)v^2\right)\right) \ell N(-d_1 - (\ell-1)v).$$

Proof. This is a straightforward (but lengthy) exercise in differentiation. \square

We now have a fast (instantaneous) and accurate (see [Monoyios, 2004b](#)) computation of the optimal price and hedging strategy. Note that the leading-order term in the expansion for the price is the marginal price, $\widehat{p}(t, y) = E^{Q^M}[h(Y_T)|Y_t = y]$, of the claim.

2.4 Optimal versus naive hedging

We wish to compare the hedging of the claim with the optimal strategy with the naive strategy (3) which takes S as a good proxy for Y . To this end, repeat the calculation leading to the residual risk SDE (9), but with the claim traded at the BS price $BS(0, Y_0)$ per claim and hedged using the $\rho \rightarrow 1$ limit of hedging formula (even though the true value of ρ is not equal to 1). This results in the naive hedging error process V^N following

$$dV_t^N = rV_t^N dt + n\beta Y_t(\theta - \lambda)v_y(t, Y_t)dt + n\beta Y_t v_y(t, Y_t)[(\rho - 1)dB_t + \sqrt{1 - \rho^2}dZ_t].$$

Once again, we note that this is not riskless, but becomes so if the true value of ρ is indeed 1. The naive trader hopes this proves a good approximation.

Suppose the agent sells a put option (so $n = -1$) on the nontraded asset. Figure 1 shows the optimal and naive hedging error distributions generated from 10,000 asset price histories for $\rho = 0.75$, $\alpha = 0.01$ with the other parameters as in Table 1. Summary statistics for the hedge error distributions, in Table 2, show that the optimal hedge error distribution has a higher mean, lower standard deviation and a higher median than the naive hedge error distribution. The increased median, in particular, shows how the relative frequency of profits over losses is increased when hedging optimally.

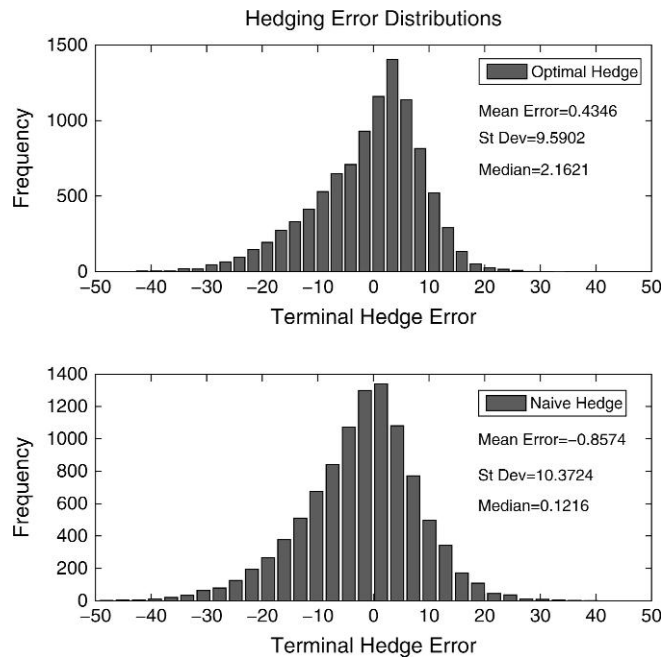


FIG. 1. Hedging error distributions over 10,000 simulated asset price histories. A short put position is hedged with $\rho = 0.75$ and $\alpha = 0.01$. The remaining parameters are as in Table 1. The put is sold optimally for price $p^{(-1)}(0, Y_0) = 13.14$ and naively for price $BS(0, Y_0) = 12.66$.

TABLE 1 *Model parameters for histograms in Fig. 1*

S_0	Y_0	K	r	μ	σ	ν	β	T
100	100	110	5%	13%	30%	10%	25%	1 year

TABLE 2 *Summary statistics for histograms in Fig. 1*

EV_T	$sd(V_T)$	$med(V_T)$	$\max(V_T)$	$\min(V_T)$
0.4346	9.5902	2.1621	33.40	-41.60
EV_T^N	$sd(V_T^N)$	$med(V_T^N)$	$\max(V_T^N)$	$\min(V_T^N)$
-0.8574	10.3724	0.1216	39.11	-47.99

3. Parameter uncertainty

The results above indicate that optimal hedging is beneficial, with the proviso that we know the parameters μ , σ , ν , β and ρ with certainty. This is a strong assumption, of course. The effect of drift parameter uncertainty is particularly severe. As Rogers (2001) mentions, one can achieve reasonable confidence in estimates of volatilities and correlation with a few years of data, but one needs hundreds of years of data to have any confidence in estimates of drift parameters.

It is worth spelling out some precise details on this that do not appear in Rogers (2001). For simplicity, let us take $r = 0$ and a stock price S following

$$dS_t = \sigma S_t(\lambda dt + dB_t),$$

where $\lambda = \mu/\sigma$, and consider an agent trying to infer the value of λ from observations of the share price. Assume (unrealistically, of course) for simplicity that the agent observes the stock returns continuously and that the volatility σ is known. The agent records the normalized returns

$$\frac{dS_t}{\sigma S_t} = \lambda dt + dB_t,$$

and uses these to estimate λ . Using observations over a time interval $[0, t]$, the best estimate of λ is $\bar{\lambda}(t)$ given by

$$\bar{\lambda}(t) = \frac{1}{t} \int_0^t \frac{dS_s}{\sigma S_s} = \lambda + \frac{B_t}{t}. \quad (15)$$

The estimator is normally distributed, $\bar{\lambda}(t) \sim N(\lambda, 1/t)$, so $(\bar{\lambda}(t) - \lambda)/(1/\sqrt{t})$ is a standard normal random variable. Hence, a 95% confidence interval for λ is

$$\left[\bar{\lambda}(t) - \frac{1.96}{\sqrt{t}}, \bar{\lambda}(t) + \frac{1.96}{\sqrt{t}} \right].$$

Suppose that the true parameter values are $\mu = 20\%$ per annum and $\sigma = 20\%$ per annum so that $\lambda = 1$. We ask, for how long do we have to observe the share price to be 95% certain that we know the value of λ to within 5% of its true value? That is, we require $|\bar{\lambda}(t) - \lambda| \leq 0.05$. This implies that

$$\bar{\lambda}(t) + \frac{1.96}{\sqrt{t}} - \left(\bar{\lambda}(t) - \frac{1.96}{\sqrt{t}} \right) = 0.1,$$

which gives $t \approx 1537$ years. This gives a measure of the severity of drift parameter uncertainty in lognormal models, and it is remarkable that (to the best of our knowledge) the above calculation does not appear in any of the standard financial mathematics texts.

In the rest of this paper, we shall restrict our attention to uncertainty in the drifts of the asset prices in the basis risk model, assuming that the agent has precise knowledge of the volatilities and the correlation.

3.1 Optimal hedging with erroneous drift parameters

In this section, we assess the impact of drift parameter misestimation by repeating the simulation experiments of Section 2.4 using erroneous values for the drift parameters μ and ν . That is, we assume that the agent computes the indifference price and optimal hedging strategy using her (incorrect) estimates of the parameters, but the simulated asset price histories are generated using the correct values of the parameters.

Figure 2 shows the optimal and naive hedging error distributions from hedging a short put position in the case when the agent's estimates of these parameters are 50% higher than the true values.

Summary statistics for the hedge error distributions, in Table 4, show that the optimal hedge is still superior to the naive hedge, though not to the extent seen earlier. (For comparison, Table 5 shows the hedging error statistics if the correct drift parameter values are used.) This can be traced to the fact that if the agent overestimates ν , she lowers the asking price for the put (relative to the asking price with the true model parameters), so the claim is viewed as less risky than it truly is. The effect of parameter

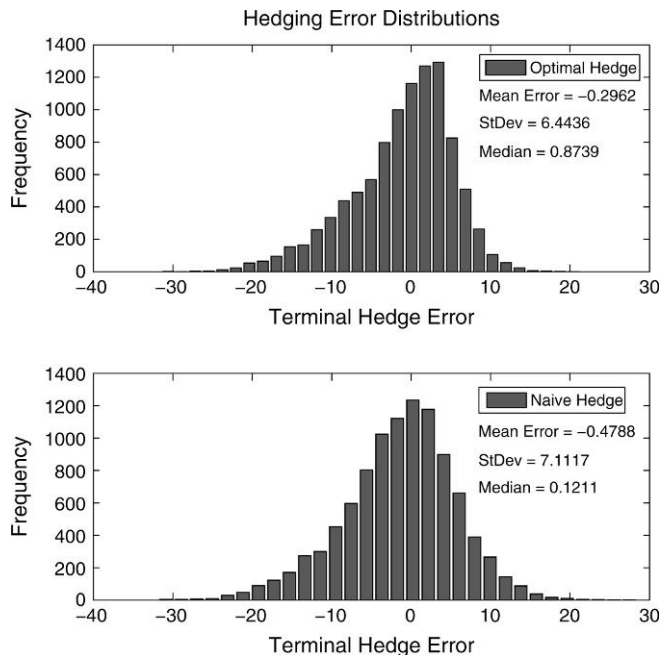


FIG. 2. Hedging error distributions over 10,000 simulated asset price histories when the agent overestimates the drifts μ and ν by 50%. A short put position is hedged with $\rho = 0.75$ and $\alpha = 0.01$. The remaining parameters (correct values) are as in Table 3. The agent optimally sells the put for price 10.81 (this price would be 11.45 using the true values of μ and ν) and naively for price $BS(0, Y_0) = 12.50$.

TABLE 3 *Model parameters for histograms in Fig. 2*

S_0	Y_0	K	r	μ	σ	ν	β	T
100	100	110	0	13%	18%	10%	15%	1 year

TABLE 4 *Summary statistics for histograms in Fig. 2*

EV_T	$sd(V_T)$	$med(V_T)$	$max(V_T)$	$min(V_T)$
-0.2962	6.4436	0.8739	20.44	-30.58
EV_T^N	$sd(V_T^N)$	$med(V_T^N)$	$max(V_T^N)$	$min(V_T^N)$
-0.4788	7.1117	0.1211	27.42	-30.88

TABLE 5 *Summary statistics for hedging a short put when the drifts μ and ν are correctly estimated with $\rho = 0.75$ and $\alpha = 0.01$. The remaining parameters are as in Table 3. The agent optimally sells the put for price 11.45 and naively for price $BS(0, Y_0) = 12.50$*

EV_T	$sd(V_T)$	$med(V_T)$	$max(V_T)$	$min(V_T)$
0.1923	6.5562	1.3779	21.91	-29.73
EV_T^N	$sd(V_T^N)$	$med(V_T^N)$	$max(V_T^N)$	$min(V_T^N)$
-0.5089	7.2561	0.0962	26.67	-32.86

misestimation in this case is relatively benign, mainly because the misestimation is in the same direction for both drifts μ and ν . As we shall see, this is not always the case.

Table 6 shows the results if the drifts are both underestimated by 50%. In this case, the optimal hedging appears to be even more beneficial than before because the agent perceives the claim to be riskier than it truly is, and raises the asking price of the put (relative to the asking price with the true model parameters).

Table 7 shows the results when the stock price drift μ is overestimated by 50% and the nontraded asset price drift ν is underestimated by 50%. The effect of parameter misestimation now becomes more significant. The agent believes that the claim is riskier than it truly is, and this effect is exacerbated by hedging with a traded asset which the agent believes is a less-effective hedging instrument than it really is. The result is that the agent significantly raises the price at which she sells the option (by about 50% over the price she would charge if she knew the true drift values), and this results in an optimal hedging error distribution that is much more favourable than the naive hedging distribution. Of course, this improvement has come at a cost of selling the claim at a very high price, so the agent becomes a very uncompetitive market maker who may not be able to sell the put at her chosen price.

Finally, Table 8 shows the results in the case that the stock price drift μ is underestimated by 50% and the nontraded asset price drift ν is overestimated by 50%. Now, the effect of parameter misestimation becomes truly destructive. The agent believes that the claim is 'less' risky than it truly is, so becomes overly aggressive, selling the claim for a low price, and this results in an optimal hedging

TABLE 6 *Summary statistics for hedging a short put when the drifts μ and ν are underestimated by 50% with $\rho = 0.75$ and $\alpha = 0.01$. The remaining parameters are as in Table 3. The agent optimally sells the put for price 12.11 (this price would be 11.45 using the true values of μ and ν) and naively for price $BS(0, Y_0) = 12.50$*

EV_T	$sd(V_T)$	$med(V_T)$	$\max(V_T)$	$\min(V_T)$
0.8916	6.4686	1.9822	23.01	-27.11
EV_T^N	$sd(V_T^N)$	$med(V_T^N)$	$\max(V_T^N)$	$\min(V_T^N)$
-0.3958	7.1714	0.2745	28.18	-30.02

TABLE 7 *Summary statistics for hedging a short put when the stock price drift μ is overestimated by 50% and the nontraded asset price drift ν is underestimated by 50% with $\rho = 0.75$ and $\alpha = 0.01$. The remaining parameters are as in Table 3. The agent optimally sells the put for price 18.27 (this price would be 11.45 using the true values of μ and ν) and naively for price $BS(0, Y_0) = 12.50$*

EV_T	$sd(V_T)$	$med(V_T)$	$\max(V_T)$	$\min(V_T)$
6.3905	6.6061	7.0682	27.71	-28.69
EV_T^N	$sd(V_T^N)$	$med(V_T^N)$	$\max(V_T^N)$	$\min(V_T^N)$
-0.5497	7.3022	0.1482	29.93	-38.36

TABLE 8 *Summary statistics for hedging a short put when the stock price drift μ is underestimated by 50% and the nontraded asset price drift ν is overestimated by 50% with $\rho = 0.75$ and $\alpha = 0.01$. The remaining parameters are as in Table 3. The agent optimally sells the put for price 6.03 (this price would be 11.45 using the true values of μ and ν) and naively for price $BS(0, Y_0) = 12.50$*

EV_T	$sd(V_T)$	$med(V_T)$	$\max(V_T)$	$\min(V_T)$
-4.3950	6.7976	-2.6025	14.55	-35.41
EV_T^N	$sd(V_T^N)$	$med(V_T^N)$	$\max(V_T^N)$	$\min(V_T^N)$
-0.5102	7.2510	0.0854	26.67	-32.86

error distribution that is much less favourable than the naive hedging distribution. This would lead to disastrous losses in practice.

Overall, we conclude that drift parameter misestimation is occasionally benign, but can be extremely destructive, depending on the sign of the misestimation. In Section 4, we propose a Bayesian learning algorithm that may help remedy this situation.

4. A filtering approach

In this section, we acknowledge uncertainty in drift parameters by taking them to be random variables with known prior distribution, and we shall require the agent's trading strategies to be adapted to the

observation filtration generated by the asset prices S and Y . We shall refer to this problem as one of ‘partial information’. A number of authors (Lakner, 1995, 1998; Brennan, 1998) have treated the classical Merton (1969, 1971) optimal investment problems under partial information. Our approach is similar though our problem is more complex. We show how it is possible to treat the problem as a Markovian one with only one extra dimension compared to the full information problem.

Furthermore, we are able to treat a special case, where the agent only has uncertainty in θ , in an analytic fashion. The implementation of the filtering solution to simulated and real data is left as a topic for future work. Here, in a preliminary foray into the filtering approach, we are concerned with showing the ideas behind the technique.

For simplicity, we take $r = 0$. The asset price SDEs are then

$$\begin{aligned} dS_t &= \sigma S_t(\lambda dt + dB_t), \\ dY_t &= \beta Y_t(\theta dt + dW_t), \end{aligned} \tag{16}$$

where $\lambda = \mu/\sigma$ and $\theta = v/\beta$. We shall suppose that the distributions of λ and θ are normal, with

$$\lambda \sim N(\lambda_0, v_0), \quad \theta \sim N(\theta_0, \gamma_0).$$

We suppose that the agent infers the values of λ_0 and v_0 (respectively, θ_0 and γ_0) by observing the values of S (respectively, Y) over some time interval $[-\tau, 0]$, and using classical estimation theory, as described by (15). The agent will then filter (update) her estimates of λ and θ from subsequent observations of

$$\begin{aligned} \zeta_t &:= \frac{1}{\sigma} \int_0^t \frac{dS_s}{S_s} = \lambda t + B_t, \\ \xi_t &:= \frac{1}{\beta} \int_0^t \frac{dY_s}{Y_s} = \theta t + W_t, \end{aligned} \tag{17}$$

over the hedging interval $[0, T]$. We can treat this as a standard Kalman filtering problem with the following solution.

THEOREM 1 Let $\mathbb{H} := (\mathcal{H}_t)_{0 \leq t \leq T}$ denote the observation filtration, with $\mathcal{H}_t = \mathcal{F}_t^\zeta \cup \mathcal{F}_t^\xi$, where $(\mathcal{F}_t^\zeta)_{0 \leq t \leq T}$ denotes the natural filtration of ζ (and similarly for ξ).

The problem with partial information can be reduced to one with full information, with the parameters λ and θ replaced by random parameters $\hat{\lambda}_t \equiv \hat{\lambda}(t, S_t)$ and $\hat{\theta}_t \equiv \hat{\theta}(t, Y_t)$ given by

$$\begin{aligned} \hat{\lambda}(t, S_t) &= \frac{\lambda_0 + v_0 \zeta_t}{1 + v_0 t}, \\ \hat{\theta}(t, Y_t) &= \frac{\theta_0 + \gamma_0 \xi_t}{1 + \gamma_0 t}, \end{aligned} \tag{18}$$

where

$$\begin{aligned} \zeta_t &= \frac{1}{\sigma} \log \left(\frac{S_t}{S_0} \right) + \frac{1}{2} \sigma t, \\ \xi_t &= \frac{1}{\beta} \log \left(\frac{Y_t}{Y_0} \right) + \frac{1}{2} \beta t. \end{aligned} \tag{19}$$

The asset price dynamics are then

$$dS_t = \sigma S_t(\hat{\lambda}(t, S_t)dt + d\widehat{B}_t), \quad (20)$$

$$dY_t = \beta Y_t(\hat{\theta}(t, Y_t)dt + d\widehat{W}_t), \quad (21)$$

where \widehat{W} and \widehat{B} are correlated Brownian motions in the observation filtration.

Proof. We give the analysis for the filtering of the stock price drift. The analysis is identical for the case of the nontraded asset.

Using the nomenclature of Øksendal (2003) for filtering theory, the ‘observation process’ is $\zeta = (\zeta_t)_{0 \leq t \leq T}$, described by

$$d\zeta_t = \lambda dt + dB_t.$$

The unobservable ‘signal process’ is λ , described by

$$d\lambda = 0.$$

Define the optimally filtered value of λ , the process $\hat{\lambda} = (\hat{\lambda}_t)_{0 \leq t \leq T}$ by the conditional expectation

$$\hat{\lambda}_t := E[\lambda | \mathcal{F}_t^\zeta], \quad \hat{\lambda}_0 = \lambda_0,$$

and denote the conditional variance of λ by the process $v = (v_t)_{0 \leq t \leq T}$, given by

$$v_t := E[(\lambda - \hat{\lambda}_t)^2 | \mathcal{F}_t^\zeta],$$

with initial value v_0 .

The ‘innovations process’ is $\widehat{B} = (\widehat{B}_t)_{0 \leq t \leq T}$, defined by

$$\widehat{B}_t := \zeta_t - \int_0^t \hat{\lambda}_s ds. \quad (22)$$

By classical filtering results (e.g. Lemma 6.2.6 in Øksendal, 2003), \widehat{B} is an \mathcal{F}_t^ζ -Brownian motion. Further, by the celebrated Kalman–Bucy filter (e.g. Theorem 6.2.8 in Øksendal, 2003), the optimal filter $\hat{\lambda}_t$ satisfies the SDE

$$d\hat{\lambda}_t = -v_t \hat{\lambda}_t dt + v_t d\zeta_t. \quad (23)$$

Moreover, v_t solves the Riccati equation

$$\frac{dv_t}{dt} = -v_t^2,$$

which has solution

$$v_t = \frac{v_0}{1 + v_0 t}.$$

Inserting this into (23) and solving the SDE for $\hat{\lambda}_t$ give (18). Moreover, $\zeta_t = \lambda t + B_t$ (from (17)), while the stock price SDE (16) has solution

$$S_t = S_0 \exp\left(-\frac{1}{2}\sigma^2 t + \sigma(\lambda t + B_t)\right) = S_0 \exp\left(-\frac{1}{2}\sigma^2 t + \sigma \zeta_t\right),$$

from which (19) follows.

Finally, using the definition (22) of the innovations process, we can write the SDE for the stock in terms of processes adapted to the observation filtration as follows:

$$dS_t = \sigma S_t d\tilde{\xi}_t = \sigma S_t (\hat{\lambda}_t dt + d\hat{B}_t),$$

which is (20). \square

Using Theorem 1, we can now use standard techniques to solve the agent's optimal hedging problem with asset price dynamics given by (20) and (21). The agent's value function is now

$$u(t, s, x, y) := \sup_{\pi \in A} J(t, s, x, y; \pi),$$

$$J(t, x, y; \pi) := E[U(X_T + nh(Y_T)) | S_t = s, X_t = x, Y_t = y],$$

$$u(T, s, x, y) = U(x + nh(y)),$$

where X is the wealth process described by $dX_t = (\pi_t/S_t)dS_t$.

Define the differential operator $\mathcal{A}^{(\mathcal{Q}^M)S,Y}$ as the generator of the 2D process S, Y in (20) and (21) under the minimal martingale measure:

$$\mathcal{A}^{(\mathcal{Q}^M)S,Y} f = \frac{1}{2} \sigma^2 s^2 f_{ss} + \beta y (\hat{\theta} - \rho \hat{\lambda}) f_y + \frac{1}{2} \beta^2 y^2 f_{yy} + \rho \sigma \beta s y f_{sy},$$

for any function $f(t, s, y)$. We then have the following result.

PROPOSITION 6 The value function $u(t, s, x, y)$ is given by

$$u(t, s, x, y) = -e^{-\alpha x} f(t, s, y), \quad (24)$$

where the function f satisfies

$$f_t + \mathcal{A}^{(\mathcal{Q}^M)S,Y} f - \frac{1}{2} \hat{\lambda}^2 f - \frac{(\rho \beta y f_y + \sigma s f_s)^2}{2f} = 0, \quad (25)$$

with $f(T, s, y) = e^{-\alpha nh(y)}$. The optimal trading strategy is $\pi = (\pi_t^*)_{0 \leq t \leq T}$ given by $\pi_t^* = \Pi^*(t, S_t, Y_t)$, where

$$\Pi^*(t, s, y) = \frac{\hat{\lambda}}{\sigma \alpha} + \frac{\rho \beta y f_y + \sigma s f_s}{\sigma \alpha f}. \quad (26)$$

Proof. The HJB equation for $u(t, s, x, y)$ is

$$u_t + \mathcal{A}^S u + \mathcal{A}^Y u + \rho \sigma \beta s y u_{sy} - \frac{(\hat{\lambda} u_x + \rho \beta y u_{xy} + \sigma s u_{sx})^2}{2u_{xx}} = 0,$$

where \mathcal{A}^S and \mathcal{A}^Y are the generators of the diffusions (20) and (21). The optimal trading strategy is $\pi^* = (\pi_t^*)_{0 \leq t \leq T}$ given by $\pi_t^* = \Pi^*(t, S_t, X_t^*, Y_t)$, where the optimal feedback control function $\Pi^*(t, s, x, y)$ is given by

$$\Pi^*(t, s, x, y) := - \left(\frac{\hat{\lambda} u_x + \rho \beta y u_{xy} + \sigma s u_{sx}}{\sigma u_{xx}} \right) \quad (27)$$

and X_t^* is the optimal wealth process satisfying $dX_t^* = (\pi_t^*/S_t)dS_t$. Now, separate out the dependence on initial wealth x as in (24). Then, the function f satisfies (25), and the expression (27) for the optimal control then simplifies to (26). \square

REMARK 1 In the above proof, we have implicitly supposed that a classical solution to the PDE (25) exists, to which the Itô formula can be applied. If this is the case, then optimality of the proposed solution for the value function follows from a verification theorem. In the absence of a classical solution to (25), techniques based on generalized solutions such as viscosity solutions would have to be used to demonstrate that the candidate solution is indeed optimal. This will be investigated in future research when a full solution to the optimal hedging problem under partial information will be sought.

We recognize the first term in (26) as the classical Merton strategy. The subsequent terms represent the additional hedging associated with (i) the randomness in Y (the risk from the claim), which includes uncertainty in the drift of Y , and (ii) the risk associated with uncertainty in the drift of S .

4.1 A special case with an analytic solution

We now consider the special case that the agent has precise knowledge of the stock price drift, but we retain uncertainty in the drift of the nontraded asset. Then, the dependence of the value function and optimal trading strategy on S disappears. The problem becomes similar to the full information problem and the distortion (Zariphopoulou, 2001) technique can be used to give the following solution for the value function and optimal trading strategy.

PROPOSITION 7 In the case where the agent has precise knowledge of the stock price drift, with uncertainty in the drift of the nontraded asset, the value function and indifference price are given by

$$\begin{aligned} u(t, x, y) &= -e^{-\alpha x - \frac{1}{2}\lambda^2(T-t)} [H(t, y)]^{1/(1-\rho^2)}, \\ H(t, y) &= E^{Q^M} [\exp(-\alpha(1-\rho^2)nh(Y_T)) | Y_t = y], \\ p(t, y) &= -\frac{1}{\alpha(1-\rho^2)n} \log H(t, y), \end{aligned}$$

where Q^M denotes the minimal martingale measure for the market (20) and (21), under which Y follows

$$dY_t = \beta(\hat{\theta}(t, Y_t) - \rho\lambda)Y_t dt + \beta Y_t d\widehat{W}_t^{Q^M},$$

and \widehat{W}^{Q^M} is a Q^M -Brownian motion.

In this special case, the problem solution becomes remarkably similar to the full information solution with the proviso that the market price of risk of the nontraded asset, $\theta = \nu/\beta$, is replaced by the random process $\theta(t, Y_t)$. In particular, the optimal hedging strategy for the claim will also be given by a formula similar to that in Proposition 2.

It is feasible that analytic formulae for the indifference price and optimal hedging strategy can be derived in this case. This is in progress, as is a numerical solution of the general partial information model, based upon finite-difference solution of the PDE (25). The aim will be to determine if, with partial information and Bayesian updating of the drift parameter estimates, the optimal hedging strategy is indeed superior to the BS-style strategy that does not require estimation of the drift parameters.

5. Conclusions

We have developed an efficient approach, based on enhanced analytic approximations for indifference prices and optimal hedging strategies, for examining the impact of drift parameter uncertainty on the

optimal hedging of a claim on a nontraded asset. This showed that the effect of parameter misestimation can be highly destructive, calling into question whether optimal hedging really is an improvement over a BS-style hedge that does not rely on drift parameter estimation. We then developed a filtering approach to incorporate Bayesian learning into the drift parameter estimation, and we were able to reduce this problem to one in which the HJB equation increased in dimensionality by only one dimension, compared to the full information case. Finally, we were able to treat analytically the special case where the agent has precise knowledge of the stock price drift, while being uncertain of the nontraded asset price drift. This approach shows promise, opening the possibility of incorporating parameter uncertainty into the optimal hedging program. The efficacy of such strategies will be the subject of forthcoming work, which will require efficient numerical solution of a 3D PDE. This will allow the optimal hedge to be tested over a large number of simulated asset price histories so that the terminal hedging error distribution can be computed. An empirical implementation over real data, as was done in the full information case in [Monoyios \(2004b\)](#), will also be a significant topic for future research.

Acknowledgements

This work was presented at the 2006 SIAM Conference on Financial Mathematics and Engineering, Boston. The author is grateful to participants in the conference, to seminar participants at Princeton University, University of Oxford, the Isaac Newton Institute for Mathematical Sciences and to two anonymous referees for helpful comments and suggestions.

REFERENCES

- BLACK, F. & SCHOLES, M. (1973) The pricing of options and corporate liabilities. *J. Polit. Econ.*, **81**, 637–659.
- BRENNAN, M. J. (1998) The role of learning in dynamic portfolio decisions. *Eur. Financ. Rev.*, **1**, 295–306.
- DAVIS, M. H. A. (1997) Option pricing in incomplete markets. *Mathematics of Derivative Securities* (M. A. H. Dempster & S. R. Pliska eds). Cambridge, UK: Cambridge University Press, pp. 216–226.
- DAVIS, M. H. A. (2006) Optimal hedging with basis risk. *From Stochastic Calculus to Mathematical Finance: The Shiryayev Festschrift* (Y. Kabanov, R. Lipster & J. Stoyanov eds). Berlin, Germany: Springer, pp. 169–188.
- FÖLLMER, H. & SCHWEIZER, M. (1991) Hedging of contingent claims under incomplete information. *Applied Stochastic Analysis* (M. H. A. Davis & R. J. Elliott eds), Stochastics Monographs, vol. 5. New York: Gordon & Breach, pp. 389–414.
- GRIMMETT, G. & STIRZAKER, D. (2001) *Probability and Random Processes*, 3rd edn. Oxford: Oxford University Press.
- HENDERSON, V. (2002) Valuation of claims on nontraded assets using utility maximization. *Math. Financ.*, **12**, 351–373.
- LAKNER, P. (1995) Utility maximization with partial information. *Stoch. Process. Appl.*, **56**, 247–273.
- LAKNER, P. (1998) Optimal trading strategy for an investor: the case of partial information. *Stoch. Process. Appl.*, **76**, 77–97.
- MERTON, R. C. (1969) Lifetime portfolio selection under uncertainty: the continuous-time case. *Rev. Econ. Stat.*, **51**, 247–257.
- MERTON, R. C. (1971) Optimum consumption and portfolio rules in a continuous-time model. *J. Econ. Theory*, **3**, 373–413 (erratum *ibid* **6**, 213–214, 1973).
- MONOYIOS, M. (2004a) Option pricing with transaction costs using a Markov chain approximation. *J. Econ. Dyn. Control*, **28**, 889–913.
- MONOYIOS, M. (2004b) Performance of utility-based strategies for hedging basis risk. *Quant. Financ.*, **4**, 245–255.

- MUSIELA, M. & ZARIPHOPOULOU, T. (2004) An example of indifference prices under exponential preferences. *Financ. Stoch.*, **8**, 229–239.
- ØKSENDAL, B. (2003) *Stochastic Differential Equations: An Introduction with Applications*, 6th edn. Berlin, Germany: Springer.
- ROGERS, L. C. G. (2001) The relaxed investor and parameter uncertainty. *Financ. Stoch.*, **5**, 131–154.
- SPANOS, A. (1999) *Probability Theory and Statistical Inference*. Cambridge: Cambridge University Press.
- ZARIPHOPOULOU, T. (2001) A solution approach to valuation with unhedgeable risks. *Financ. Stoch.*, **5**, 61–82.