# PRICING COMMODITY DERIVATIVES WITH BASIS RISK AND PARTIAL OBSERVATIONS

RENÉ CARMONA AND MICHAEL LUDKOVSKI

Abstract. We study the problem of pricing claims written on an over-the-counter energy contract. Because the underlying is illiquid, we work with an indifference pricing framework based on a liquid reference contract. Extending current convenience yield frameworks we propose a two-factor partially observed model for the benchmark asset. Moreover, we incorporate direct modeling of the unhedgeable basis. We then study the value function corresponding to utility pricing with exponential utility. After performing filtering this leads to an infinite-dimensional Hamilton-Jacobi-Bellman equation. We show that if the basis is totally independent, the indifference price of the claim is equal to its certainty equivalent. In the more interesting case where the basis depends on the unobserved factor we obtain a reduced-form expression for the price in terms of a conditional expectation. We show how to numerically compute this expectation using a Kalman or particle filter. Our basic model may be generalized to include nonlinear dynamics and further dependencies.

*Date*: July 2006.

*Key words and phrases.* commodity forwards, partial observations, indifference pricing, HJB equation, particle filtering.

## 1. Introduction

In this paper we study pricing of commodity contingent claims involving basis risk. The model is motivated by the industry practice of crosshedging over-the-counter (OTC) energy assets with a liquid reference contract. Typical examples are an option on a particular *local* natural gas forward hedged with the Nymex gas contract or a claim involving a specific grade of crude oil hedged with the Brent contract. To capture the economic intuition regarding asset dynamics we work with a two-factor model consisting of the benchmark forward contract and a stochastic drift factor that may be interpreted as the convenience yield. The second factor is unobserved and we explicitly incorporate learning via online filtering.

The considered market is incomplete due to the partial observations and the non-traded basis risk. As a result the no-arbitrage theory cannot be used; instead we apply indifference pricing based on exponential utility (Davis 2000, El Karoui and Rouge 2000, Musiela and Zariphopoulou 2004). The key advantage of exponential utility is that the resulting stochastic control problem may be linearized leading to explicit representation of the indifference price in terms of a Feynman-Kac type expectation. This representation involves the conditional distribution of the unobserved factor; in general this is an infinite-dimensional object that must be approximated with an appropriate filtering procedure. To obtain numerical results we describe a particle filter algorithm and provide a complete implementation to illustrate our methodology.

This paper has been inspired by the brief note of Lasry and Lions (1999) who point out that the wealth-invariance property of exponential utility carries over to models with partial observations. However, their report does not mention any applications and does not consider the case where the payoff depends on the unobserved. The closely related problem of indifference pricing with exponential utility and unhedgeable risks has been discussed in the fully observed setting by Sircar and Zariphopoulou (2005) and Becherer (2003). The particular case of cross-hedging and basis risk was studied by Davis (2000), Henderson (2002) and Monoyios (2004). More generally, our use of a latent stochastic factor is related to the series of papers by Runggaldier (Runggaldier 2004, and references therein) who has provided the general framework for filtering in financial models.

Our contribution to literature is three-fold. First, our work is a new application of utility based valuation to energy derivatives. Given that energy markets are highly incomplete and involve many non-traded features this methodology is a natural choice. Second, we emphasize the role of models with partially observed stochastic drift in commodity trading. Our approach extends the notion of the convenience yield. The convenience yield itself is an elusive concept, but it is clear that several factors are necessary to capture the forward contract dynamics. Thus, we write down a general two-factor model while remaining agnostic about the precise interpretation of the unobserved second factor. Finally, we demonstrate a new application of filtering techniques, especially particle filters, in finance. As opposed to standard cases where filtering is used for estimation, we employ the filter to actually *price* contingent claims.

From an applications point of view, our work belongs to the sequence of convenience yield models begun by Gibson and Schwartz (1990); see also (Schwartz 1997), Schwartz and Smith (2000), Casassus and Collin-Dufresne (2005). Like ours, these models exhibit a stochastic drift; however, unlike them, we work directly with the historical dynamics of the forward under the physical measure $\mathbb{P}$ and avoid specifying risk-neutral dynamics. We believe that this approach is advantageous for empirical fitting as we do not need to make additional assumptions regarding the form of the risk premium and can calibrate directly from historical data.

Finally, our analysis is connected to the problem of portfolio optimization with partial observations. Existing literature has concentrated on the Gaussian case where explicit computations are possible. In his early paper Lakner (1998) solved the classical Merton problem in this context using the Kalman filtering equations and the convex duality approach. Subsequently, this work was extended by Nagai (2000), Sekine (2003) and Brendle and Carmona (2005) to cover the general situation of correlated Gaussian observed and unobserved factors. We borrow from these methods to illustrate our results on a simple linear model. However, let us emphasize that our analysis is especially attractive for nonlinear models where Monte Carlo simulation is the only feasible approach.

The rest of the paper is organized as follows. In Section 2 we describe the financial setting of our problem and explain our pricing methodology. We then relate our model to existing proposals regarding commodity price dynamics. Section 3 recalls the filtering results we need and is followed by Section 4 which contains our key results on the resulting indifference prices. Section 5 explains how one may compute these prices and illustrates our findings with a numerical example. Finally, Section 6 concludes and outlines possible extensions to consider in the future.

## 2. Model Setup

In this section we describe the pricing model we use and the underlying financial motivation.

2.1. **Asset Dynamics.** We begin with the following model of asset dynamics. Let $\{\Omega, \mathcal{F}, \mathbb{P}\}$ be a complete probability space. Let $F_t$ be the value at time $t$ of a traded financial forward contract on the given commodity, and $X_t$ an *unobserved* stochastic drift factor. We then postulate

$$\begin{cases} dF_t = F_t \cdot \big( h(t, F_t, X_t)\, dt + \sigma(t, F_t)\, dW_t^1 \big), \\ dX_t = b(t, F_t, X_t)\, dt + a(t, F_t, X_t)\, dW_t^2, \end{cases} \tag{2.1}$$

with $W^1, W^2$ one-dimensional $\mathbb{P}$-Wiener processes with correlation $c$. Further motivation for (2.1) is provided in Section 2.3 below. Note that the diffusion coefficient of $F$ must not depend on $X$. Thus, this setup is inherently different from stochastic volatility models analyzed e.g. by Pham and Quenez (2001). On the other hand, the dynamics of unobserved $X$ can have generic dependence on the observable $F$.

We impose the following standing assumptions on the coefficients of (2.1):

**Assumption 1.**

- $h(t, f, x), b(t, f, x) \colon [0, T] \times \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ *are uniformly continuous and have bounded second derivatives.*
- $\sigma(t, f)$ *and* $a(t, f, x)$ *are uniformly continuous, have bounded third derivatives and are uniformly elliptic, that is* $\sigma^2(t, f) \geqslant \lambda$, $a^2(t, f, x) \geqslant \lambda$ *for all* $t$, $f$ *and* $x$, *for some constant* $\lambda > 0$.

Assumption 1 guarantees among other things that (2.1) has a unique strong solution.

Let $\mathcal{G}_t \triangleq \sigma\big\{(F_s, X_s) \colon 0 \leqslant s \leqslant t\big\}$ denote the natural filtration generated by the entire process, as contrasted with the observable filtration $\mathcal{F}_t \triangleq \sigma\{F_s \colon 0 \leqslant s \leqslant t\}$. In line with the usual predictability assumption, all our trading strategies will be required to be $\mathcal{F}_t$-adapted.

2.2. **A Linear Example.** For the sake of illustration, we focus on the following particular case of (2.1) throughout this paper. Let $Y_t \equiv \log F_t$ and take

$$\begin{cases} dY_t = \big(\mu - \frac{1}{2}\sigma^2 - X_t\big) dt + \sigma\, dW_t, \\ dX_t = \kappa(\theta - X_t)\, dt + c\alpha\, dW_t + \sqrt{1 - c^2}\alpha\, dW_t^{\perp}, \end{cases} \tag{2.2}$$

with $W^{\perp}$ a standard Wiener process independent of $W$. The equations in (2.2) emphasize the linearity of this setting.

The choice of an Ornstein-Uhlenbeck model for $X$ in (2.2) reflects the desire for a long-term mean-reversion that characterizes commodity markets (Fama and French 1988). Since in the long run commodities are consumption goods, we expect to achieve a supply-demand equilibrium. Thus, log-prices should be stationary. In (2.2) this is achieved by mean-reversion in $X$ coupled with a strong positive correlation between $X$ and $F$, $c \gg 0$. The feedback effect between $X$ and $F$ then causes weak mean-reversion in the forward price $F$.

Another advantage of (2.2) is that it permits several explicit computations. In particular, we have explicit formulas for the moments of $F_T$ if the initial distribution of $X_0$ is Gaussian.

**Proposition 1.** *Suppose $(F, X)$ follow satisfy (2.2) and the initial conditional distribution of $X_0$ is Gaussian $X_0 \sim \mathcal{N}(x_0, P_0)$. Then the moments of $F_T$ under $\mathbb{P}$ are given by:*

$$\mathbb{E}\left[(F_T)^\lambda\right] = F_0^\lambda \cdot \exp\left(\lambda(\frac{\mathrm{e}^{-\kappa T}-1)}{\kappa} x_0 + \lambda \cdot k_0\right), \qquad \forall \lambda > 0, \tag{2.3}$$

*where*

$$\begin{cases} k_0 = \displaystyle\int_0^T \Big[\mu + (\lambda - 1)\frac{\sigma^2}{2} + 2g_t\big(\lambda(c\sigma\alpha - P_t) + \kappa\theta\big) + 2\lambda g_t^2 \frac{(c\sigma\alpha - P_t)^2}{\sigma^2}\Big] dt, \\ g_t = \dfrac{1}{2\kappa}(\mathrm{e}^{\kappa(T-t)} - 1), \\ P_t = \displaystyle\int_0^t \Big[\alpha^2 - 2\kappa P_s - \frac{(c\sigma\alpha - P_s)^2}{\sigma^2}\Big] ds. \end{cases} \tag{2.4}$$

The proof of Proposition 1 is given in the Appendix. The equation for $P_t$ is of a Riccati type and has been well studied. It is known that $P_t$ is monotonic and converges to a limiting value.

2.3. **Financial Application.** A model of the form (2.1) arises in connection with pricing over-the-counter commodity derivatives. The commodity markets are characterized by their fragmented nature. Thus, there are only a few liquidly traded contracts existing along hundreds of similar but distinct over-the-counter products. The situation arises due to physical and/or geographic distinctions. For instance, there are dozens of grades of crude oil being produced in the world, but only the Brent North and West Texas Intermediate contracts are liquidly traded on the exchanges. Similarly, natural gas prices depend on the location where the gas is to be delivered, resulting in several hundred of geographic contracts. All of these are very illiquid and only traded in over-the-counter manner, so that direct hedging is impossible. To setup a hedge, the industry practice is to instead use a benchmark reference contract like the aforementioned North Sea Brent crude traded on the International Petroleum Exchange in London or the New York Mercantile Exchange (Nymex) Henry Hub gas; the corresponding spread between the claim of interest and the benchmark is termed *basis*. Thus, the attempted cross-hedge is inherently imperfect to the extent that the basis is non-traded and constitutes an additional source of risk.

The other industry practice is to avoid trading in the physical spot. Instead, nearly all trading is done via forwards or futures, often with financial settlement. Commodity spot markets are relatively illiquid and involve physical settlement which is inconvenient for financial trading. Moreover, trading in the physical asset requires dealing with *individual* storage costs of the agent, which might be different from the marginal storage costs reflected in the prices.

Given these two stylized facts, it is natural to advance the framework of Section 2.1 to analyze pricing of a claim on a particular local OTC contract $F^{loc}$. The market is incomplete and the riskiness of the claim is measured in terms of the benchmark forward $F$.

*Remark 1*: Instead of directly referring to the basis, an alternative is to write down a joint model for the OTC and benchmark contracts $(F^{loc}, F)$. Normally, one takes both processes to be diffusions with high correlation $c \approx 1$. However, this makes it difficult to guarantee that the spread $F^{loc} - F$ is bounded, which is economically desirable. We believe that the basis is a much more meaningful financial object than correlation and consequently isolate its effect in the ensuing mathematical analysis. Our approach is similar to the co-integration work of Duan and Pliska (2004); we refer to Carmona and Durrleman (2003), as well as Eydeland and Wolyniec (2003) for further discussion on modeling spreads in the energy markets.

2.4. **Pricing Framework.** Given the setting of Section 2.1, we are interested in pricing European derivatives $\tilde{\phi}(F^{loc}_T)$ on the OTC contract $F^{loc}$. In line with Section 2.3, $F^{loc}$ is non-traded and we take the point of view of an agent in the $F$-market. Consequently, we re-write $\tilde{\phi}(F^{loc}_T) = \phi(F_T, B)$ where the quantity $B$ is the basis corresponding to the spread between the commodity we are actually interested in —$F^{loc}$, and the traded contract $F$. Thus, the reader is invited to take $\phi(F_T, B) = \tilde{\phi}(F_T + B)$. The first parameter of $\phi$ can be used to absorb any dependence on $F$, so that without loss of generality $B$ is independent of $F_T$. The claim matures at time $T$; we assume that $T < \bar{T}$ where $\bar{T}$ is the maturity of the forward $F$. From now on we will only consider times $t \leq T$. For simplicity we also assume that $\phi\colon \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ is continuous and of linear growth in each parameter.

We distinguish four possibilities that together cover the entire spectrum of $F^{loc}$-contingent claims.

**Classification of Basis Risks.**

(a) $B \in \mathcal{F}_0$: *deterministic basis risk.*

(b) $B \perp\!\!\!\perp \mathcal{G}_T$: *totally independent basis risk.*

(c) $B \equiv B(X_T, Z)$, *where* $Z \perp\!\!\!\perp \mathcal{G}_T$: *basis risk is a noisy function of the unobserved factor at time* $T$.

(d) $B \equiv B_T$ *where* $(B_t)$ *is a third (correlated) observable stochastic factor complementing* (2.1).

As we will see below both cases (a) and (b) lead to trivial pricing while case (d) seems to be too hard. As a result, the most interesting situation is case (c) which we study in detail in Section 4 below.

We begin by remarking that case (a) is covered by the standard no-arbitrage theory, see e.g. Karatzas and Shreve (1998). Namely, any contingent claim of the form $\phi(F_T)$ can be perfectly replicated by a $(\mathcal{F}_t)$-measurable trading strategy, even when $X$ is unobserved. This counterintuitive result illustrates the power of continuous-time Girsanov transformations. Indeed, through a Girsanov change of measure the traded $F$ can be made into a local martingale under an equivalent martingale measure $\widetilde{\mathbb{P}}$. It can be easily checked that there exists a $\widetilde{\mathbb{P}}$-Wiener process $\widetilde{W}$ whose natural filtration is equal to the filtration generated by $F$ in (2.1). By the standard martingale representation theorem we conclude that any $\mathcal{F}_T$-measurable random variable can be written as a stochastic integral with respect to $F$. It follows that the unique no-arbitrage price of the claim $\phi(F_T)$ must equal its replication cost under $\widetilde{\mathbb{P}}$. We will return to the martingale measure $\widetilde{\mathbb{P}}$ in Section 3.

In cases (b)-(d) we have inherent incompleteness and the claim $\phi(F_T, B)$ cannot be replicated, either due to non-traded basis risk or due to lack of full information about $X$. Accordingly, replication arguments no longer apply and a whole range of prices for $\phi(F_T, B)$ are consistent with no-arbitrage. To avoid difficulties associated with super-replication we use indifference valuation, see e.g. Carmona (2006) for an overview. More precisely, assuming a subjective utility function for the buyer of the claim, we value $F^{loc}$-contingent claims based on the wealth-adjusted utility equivalent received by the agent that has access to the $F$-market. From a modeling point of view, this method focuses on the hedging strategy of the agent and results in a partially observed stochastic control problem.

2.5. **Utility Valuation.** Besides being exposed to the terminal payoff $\phi$, the agent performs portfolio optimization by dynamically rebalancing her asset holdings in the benchmark forward $F$ and the riskless bank account. For simplicity we assume that the interest rates are zero $r_t \equiv 0$. This disentangles the dynamics of the interest rates from the rest of the model and makes the effect of other parameters more transparent. If at time $t$ the agent invests $\pi_t$ dollars in the forward, then the corresponding wealth process $w^\pi$ satisfies

$$dw_t^\pi = h(t, F_t, X_t)\pi_t\, dt + \sigma(t, F_t)\pi_t\, dW_t, \qquad w_0^\pi = w_0. \tag{2.5}$$

Observe that because $F$ is a financial instrument the meaning of a self-financing strategy remains the same. On the contrary, if the agent attempted to trade in the *physical spot* contract, she would have to face storage/convenience costs which would require modification to the notion of self-financing strategies.

We denote by $\mathcal{A}_t^T$ the set of admissible portfolio strategies $\{\pi_s\}_{t\le s\le T}$ which consist of all square integrable $\mathbb{E}\{\int_t^T \pi_s^2 ds\} < \infty$, $(\mathcal{F}_t)$-adapted processes. Given $\mathcal{A}_t^T$, the main object of our analysis is the *value function* $V$ defined by

$$V^\phi(t, w, f, \xi) = \sup_{\pi\in\mathcal{A}_t^T} \mathbb{E}\Big[-\mathrm{e}^{-\gamma(w_T^{x,\pi}+\phi(F_T,B))}\Big| w_t = w, F_t = f, X_t \sim \xi\Big]. \tag{2.6}$$

The coefficient $\gamma > 0$ represents the degree of risk-aversion of the agent as measured by the exponential utility $U(x) = -\mathrm{e}^{-\gamma x}$. Above the initial value of $X_t$ is unknown, but we are given some initial distribution $\xi$.

The value function is the maximum expected utility to be derived from portfolio optimization and receiving the claim $\phi(F_T, B)$ given the specified initial conditions. Note that we do not require the agent to have positive wealth $w_t^\pi > 0$, and that all optimization is done under the objective measure $\mathbb{P}$.

The *buyer's* indifference price for claim $\phi$, $P = P^\phi(t, w, f, \xi)$ at time $t$ is finally defined by

$$V^\phi(t, w - P, f, \xi) \triangleq V^0(t, w, f, \xi), \tag{2.7}$$

where 0 represents the trivial claim paying nothing. The indifference price $P$ generally depends on all four state factors $(t, w, f, \xi)$. Intuitively, $P$ represents the decrease in initial wealth that just balances the increase in terminal utility from buying the derivative $\phi(F_T, B)$. This pricing mechanism can be shown to always assign the contingent claim a value that is consistent with no-arbitrage (El Karoui and Rouge 2000). Similarly, a seller's indifference price may be defined; in this paper we concentrate on the buyer's point of view in line with the financial application outlined in Section 2.3. The optimal hedging strategy of $\phi$ is then thought of as the difference between $\pi^{\phi,*}$ (if one exists) achieving the supremum in (2.6) and the $\pi^{0,*}$ corresponding to $V^0$.

The rest of the paper is devoted to studying (2.6) and (2.7). We will do so by applying the standard Hamilton-Jacobi-Bellman framework (Fleming and Soner 1993). However, we first need a Markovian system which means that we must replace $X_t$ by its conditional expectation given $\mathcal{F}_t$. This is known as the filtering problem and will be taken up in Section 3. Before we proceed, let us elaborate on the connection between (2.1-2.2) and other commodity models.

2.6. **Relationship to Convenience Yield Models.** Our use of a two-factor model is motivated by the widely accepted observation that one-factor models (Schwartz 1997) are not rich enough for a good fit to empirical forward prices. Traditionally, the second factor has been introduced as a stochastic drift of $F$ and was termed the (forward) *convenience yield*. The convenience yield reflects the physical nature of commodities by modifying the risk-free rate of return on the spot $S_t$ from $r_t$ to $r_t - X_t$. Indeed, physical ownership of the commodity carries an associated flow of services. On the one hand, the owner enjoys the benefit of direct access which is important if the asset is to be consumed. On the other hand, the decision to postpone consumption implies storage expenses. The convenience yield also has general equilibrium underpinnings through the theory of storage that was developed back in the 1950's (Brennan 1958).

In that context, our model (2.1) is just a generalized convenience yield model with unobserved convenience yield. A variety of models (Casassus and Collin-Dufresne 2005, Hilliard and Reis 1998, Gibson and Schwartz 1990, Schwartz 1997, Schwartz and Smith 2000) of the form (2.1), but with full observations, have been suggested in the literature. However, any finite dimensional factor model like (2.2) implies that the convenience yield can be fully recovered (or filtered modulo small noise disturbances like the bid-ask spread, as is assumed by Schwartz (1997)) from forward prices of varying maturity. For example, with two factors, we can use $X_t \approx -\log(F(t,T_2)/F(t,T_1))$, where $F(t,T_1)$ and $F(t,T_2)$ are the two closest futures contracts. Unfortunately, the implied convenience yield is highly unstable and inconsistent with the forward curve. Different forward contracts generate wildly different estimates, and our earlier empirical work (Carmona and Ludkovski 2004) strongly rejects the notion of implied $X_t$. Hence, we prefer to construct a partially observed model and focus on a single maturity $F_t = F(t,\bar{T})$. We have some anecdotal evidence that unobserved convenience yield models are also used in the industry.

To put (2.1) and the specific case (2.2) into a better perspective, let us directly compare it to other models advanced in the literature. Let us stress that we only compare the postulated asset dynamics; the actual pricing methodology is quite different. The existing models are usually stated directly in terms of risk-neutral dynamics so that pricing is done by simply computing discounted expectations of payoffs. However, this requires additional assumptions on the risk premia that we do not make.

We first take up the Gibson-Schwartz model (Gibson and Schwartz 1990) which can be considered as the epitome of convenience yield models. In this model it is postulated that under a martingale measure $\mathbb{Q}$, the spot price $S$ of the commodity evolves according to

$$\begin{cases} dS_t = (r - X_t)S_t\,dt + \sigma S_t\,d\bar{W}_t^1, \\ dX_t = (\kappa(\theta - X_t) - \lambda_X)\,dt + \alpha\,d\bar{W}_t^2. \end{cases} \tag{2.8}$$

Furthermore, Gibson and Schwartz (1990, p. 960) assume that the objective $\mathbb{P}$-dynamics are

$$\begin{cases} dS_t = \mu S_t\,dt + \sigma S_t\,dW_t^1, \\ dX_t = \kappa(\theta - X_t)\,dt + \alpha\,dW_t^2, \end{cases} \tag{2.9}$$

so that the historical spot is log-normal. It follows that the market price of risk for the spot is given by $\lambda_S = (r - \mu - X_t)/\sigma$, while the market price of risk for the convenience yield $\lambda_X$ is constant. By standard arbitrage arguments the forward $F$ is a $\mathbb{Q}$-martingale. Therefore, assuming the particular forms of risk premia above, the *historical* $F$-dynamics match the functional form of our equation (2.2).

The dynamics (2.2) are also similar to the Schwartz and Smith (2000) approach of a short-term and a long-term factor. The only difference is that instead of modeling the spot we model the forward. In this interpretation of (2.2), $(X_t - \theta)/\kappa$ represents the short-term fluctuations, while $\log F_t - X_t/\kappa$ represents the long-term trend of the reference underlying.

Let us mention that most existing models assume that the dynamics of $X$ are independent of $F$, which we believe is a significant restriction and is not present in our general case of (2.1). A case in point is the recent econometric study by Casassus and Collin-Dufresne (2005) who allow a linear dependence between $X$ and $F$ and document an improved empirical fit.

## 3. Filtering the Stochastic Drift

3.1. **Mathematical Preliminaries.** Below we will need to work with stochastic partial differential equations which require a bit of analytic machinery. In this section we summarize the notation and concepts we use. We denote by $C_0^\infty(\mathbb{R})$ (respectively $C_b^\infty(\mathbb{R})$) the space of infinitely-differentiable (resp. bounded) functions with compact support and by $H_\beta^k$ the weighted Sobolev spaces. Recall that for $\beta \geqslant 0$, the Hilbert space $H_\beta^k(\mathbb{R})$ is defined as the completion of

$\mathcal{C}_0^\infty(\mathbb{R})$ with respect to the norm

$$\|f\|_{k,\beta} = \sum_{|\alpha|\leqslant k} \Big( \int_{\mathbb{R}} \big(\partial_x^\alpha [(1+|x|^2)^{\beta/2} f(x)]\big)^2 dx \Big)^{1/2}.$$

$H_\beta^k$ can be thought of as the set of all measurable functions $f\colon \mathbb{R} \to \mathbb{R}$ such that $(1+|x|^2)^{\beta/2} f(x)$ has square-integrable derivatives up to order $k$. In particular, $H_0^0 = L^2$. We denote by $\langle f, g\rangle_{k,\beta}$ the inner product induced by the above norm (dropping the subscripts when the meaning is clear) and by $D$ the Fréchet derivative operator on $H_\beta^k$. Finally as usual, for any differential operator $\mathcal{L}$, $\mathcal{L}^*$ denotes its adjoint: $\langle \mathcal{L}f, g\rangle = \langle f, \mathcal{L}^* g\rangle$.

3.2. **Zakai Equation.** To compute the conditional distribution of $X_t$ given $\mathcal{F}_t$ we use the innovation process method following an excellent exposition in Bensoussan (1992). For notational clarity we suppress from now on all the dependencies on $t$. Define $\zeta_t$ by

$$d\zeta_t = -\zeta_t h(X_t, F_t) \frac{1}{\sigma(F_t)}\, dW_t, \qquad \zeta_0 = 1. \tag{3.1}$$

By Assumption 1 $\sigma(t, F_t)^{-1}$ is bounded, so that $\zeta_t$ is an exponential martingale with $\mathbb{E}[\zeta_t] = 1, \forall t \leqslant T$ (Bensoussan 1992, Lemma 4.1.1). Applying the Girsanov theorem we define a new probability measure $\widetilde{\mathbb{P}}$ via

$$\frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}}\Big|_{\mathcal{F}_t} \triangleq \zeta_t. \tag{3.2}$$

Then under $\widetilde{\mathbb{P}}$ there exists a Wiener process $\widetilde{W}$ such that

$$dF_t = \sigma(F_t) F_t\, d\widetilde{W}_t \qquad \text{and} \tag{3.3}$$

$$\begin{aligned} dX_t = &\Big( b(F_t, X_t) - \frac{ca(F_t, X_t)}{\sigma(F_t)} h(F_t, X_t) \Big)\, dt \\ &+ c\, a(F_t, X_t) \frac{1}{\sigma(F_t) F_t} dF_t + \sqrt{1-c^2} a(F_t, X_t)\, dW_t^\perp. \end{aligned} \tag{3.4}$$

Letting $d\widetilde{F}_t = \frac{1}{\sigma(F_t)F_t} dF_t$, $\widetilde{F}$ is another Wiener process under $\widetilde{\mathbb{P}}$. Moreover, $\widetilde{F}$ and $W^\perp$ are still *independent*, and the natural filtration of $\widetilde{F}$ coincides with the natural filtration of $W$: $\sigma\{\widetilde{F}_s : 0 \leqslant s \leqslant t\} \equiv \mathcal{F}_t$. The inverse $\eta_t \triangleq \frac{1}{\zeta_t} = \frac{d\mathbb{P}}{d\widetilde{\mathbb{P}}}\Big|_{\mathcal{F}_t}$ can be written as

$$\begin{aligned} \eta_t &= \exp\Big( \int_0^t \frac{h(F_s, X_s)}{\sigma(F_s)} dW_s + \frac{1}{2} \int_0^t \frac{h(F_s, X_s)^2}{\sigma^2(F_s)}\, ds \Big) \\ &= \exp\Big( \int_0^t \frac{h(F_s, X_s)}{\sigma^2(F_s) F_s}\, dF_s - \frac{1}{2} \int_0^t \frac{h(F_s, X_s)^2}{\sigma^2(F_s)}\, ds \Big). \end{aligned} \tag{3.5}$$

Consider a fixed function $f \in C_0^\infty(\mathbb{R})$ and let $p_t(f) \triangleq \widetilde{\mathbb{E}}[f(X_t)\eta_t|\mathcal{F}_t]$. To compute the conditional expectation $\Pi_t(f) \triangleq \mathbb{E}[f(X_t)|\mathcal{F}_t]$ we apply Bayes rule to obtain the Kallianpur-Striebel formula

$$\Pi_t(f) = \frac{\widetilde{\mathbb{E}}[f(X_t)\eta_t|\mathcal{F}_t]}{\widetilde{\mathbb{E}}[\eta_t|\mathcal{F}_t]} = \frac{p_t(f)}{p_t(1)}. \tag{3.6}$$

Assuming that $p_t(\cdot)$ possesses a smooth density $\rho_t(x)\,dx$, i.e.

$$\forall f \in \mathcal{C}_0^\infty(\mathbb{R}), \quad \widetilde{\mathbb{E}}\big[f(X_t)\eta_t\big|\,\mathcal{F}_t\big] = \int_{\mathbb{R}} \rho_t(x) f(x)\,dx = \langle \rho_t, f\rangle, \tag{3.7}$$

one can apply Itô's lemma to $d(\eta_t f(X_t))$ and use (3.4) to obtain that $\rho_t(x)$ must satisfy the adjoint Zakai equation (Bensoussan 1992, p. 112)

$$\begin{aligned} d\rho_t(x) = \Big(\frac{1}{2}\frac{\partial^2}{\partial x^2}\big\{a(F_t,x)\rho_t(x)\big\} - \frac{\partial}{\partial x}(b(F_t,x)\rho_t(x))\Big)\,dt \\ + \Big(h(F_t,x)F_t - \frac{\partial}{\partial x}(ca(F_t,x)\rho_t(x))\Big)d\widetilde{F}_t. \end{aligned} \tag{3.8}$$

Let

$$\mathcal{L}_X = \frac{1}{2}a^2(\cdot,x)\partial_{xx} + b(\cdot,x)\partial_x \tag{3.9}$$

denote the elliptic operator corresponding to the diffusion $X_t$ and define a first-order differential operator

$$\mathcal{S} \triangleq ca(\cdot,x)\partial_x + h(\cdot,x)\cdot \tag{3.10}$$

Then (3.8) can be re-written as the stochastic partial differential equation (SPDE)

$$d\rho_t(x) = \mathcal{L}_X^*\rho_t(x)\,dt + \mathcal{S}^*\rho_t(x)\,d\widetilde{F}_t. \tag{3.11}$$

To guarantee that (3.11) is well-posed we must work in the weighted Sobolev spaces $H_\beta^k$. To this end, we recall the following existence-uniqueness result for solutions of SPDEs.

**Lemma 1.** *(Gozzi and Şwiech (2000)) Let $\beta > 1/2$. Let $\mathcal{L}$ be the second order linear differential operator with domain $H_\beta^2$ defined in* (3.9) *and let $\mathcal{S}$ be the first order differential operator with domain $H_\beta^1$ defined in* (3.10)*. Suppose the coefficients satisfy Assumption 1, $c < 1$ and $\|\rho_0\|_{1,\beta}^2 < \infty$. Then there exists a unique strong solution $\rho_t \in H_\beta^0([0,T];\Omega,\mathcal{F}_t,\widetilde{\mathbb{P}})$ satisfying*

$$\rho_t = \rho_0 + \int_0^t \mathcal{L}^*\rho_s\,ds + \int_0^t \mathcal{S}^*\rho_s\,d\widetilde{F}_s.$$

*Moreover, there exists a constant $C$ such that*

$$\widetilde{\mathbb{E}}\|\rho_t\|^2_{1,\beta} \leqslant \|\rho_0\|^2_{1,\beta}(1+C\,t) \qquad \textit{and} \qquad \widetilde{\mathbb{E}}\int_0^T \|\rho_s\|^2_{2,\beta}\,ds \leqslant C\cdot\|\rho_0\|^2_{1,\beta}.$$

3.3. **Example: Filtering Gibson-Schwartz Model** (2.2)**.** For the particular case of linear model (2.2) the adjoint differential operators are

$$\begin{cases} \mathcal{L}^*_X(f)(x) = \big(\kappa(\theta-x)\,f'(x) - \kappa f(x)\big) + \dfrac{1}{2}\alpha^2 f''(x), \qquad \text{and} \\ \mathcal{S}^*(f)(x) = \mu - \dfrac{1}{2}\sigma^2 - x - c\alpha f'(x). \end{cases}$$

The change of measure is defined by

$$\frac{d\mathbb{P}}{d\widetilde{\mathbb{P}}}\Big|_{\mathcal{F}_t} \equiv \eta_t = \exp\Big(\int_0^t \frac{\mu-\frac{1}{2}\sigma^2 - X_s}{\sigma^2}\,dY_s - \frac{1}{2}\int_0^t \frac{(\mu-\frac{1}{2}\sigma^2-X_s)^2}{\sigma^2}ds\Big), \tag{3.12}$$

and the un-normalized density $\rho_t(x)$ satisfies

$$d\rho_t(x) = \Big[\frac{1}{2}\alpha^2\rho_t''(x) - \frac{\partial}{\partial x}\big(\kappa(\theta-x)\rho_t(x)\big)\Big]dt + \big(\mu - \frac{1}{2}\sigma^2 - x - c\alpha\rho_t'(x)\big)dY_t.$$

In fact, because Equation (2.2) is *linear*, one does not have to carry out the above general computations and may instead use the explicit Kalman filter (Harvey 1989). Assuming that the initial distribution is Gaussian $X_0 \sim \mathcal{N}(\hat{x}_0, P_0)$, it is known that $X$ is conditionally Gaussian for all times, $X_t|\mathcal{F}_t \sim \mathcal{N}(\hat{x}_t, P_t)$. The conditional mean $\hat{x}_t$ evolves according to

$$d\hat{x}_t = \kappa(\theta-\hat{x}_t)\,dt + \frac{(c\sigma\alpha - P_t)}{\sigma^2}\Big[d(Y_t) - (\mu - \frac{1}{2}\sigma^2 - \hat{x}_t)dt\Big], \tag{3.13}$$

while the conditional variance $P_t$ evolves as in (2.4). In other words, for any bounded function $f \in \mathcal{C}^\infty_0(\mathbb{R})$,

$$\mathbb{E}[f(X_t)|\,\mathcal{F}_t] = \int_{\mathbb{R}} f\big(\hat{x}_t + P_t^{1/2}\xi\big)\frac{e^{-\frac{1}{2}\xi^2}}{\sqrt{2\pi}}\,d\xi, \tag{3.14}$$

with $\hat{x}_t$ and $P_t$ given above. Moreover, $\langle\rho_t, f\rangle = \hat{\eta}_t\cdot\mathbb{E}[f(X_t)|\,\mathcal{F}_t]$ where

$$\hat{\eta}_t = \exp\Big(\int_0^t \frac{\mu-\frac{1}{2}\sigma^2-\hat{x}_s}{\sigma^2}\,dY_s - \frac{1}{2}\int_0^t \frac{(\mu-\frac{1}{2}\sigma^2-\hat{x}_s)^2}{\sigma^2}\,ds\Big). \tag{3.15}$$

## 4. Hamilton-Jacobi-Bellman Formulation

After these preliminaries we return to the main problem of computing (2.6). Let us give a brief preview of our strategy. Using the results of Section 3 we first replace the partially observed system (2.6) with a fully-observed one where the unobserved $X$ is substituted with its un-normalized conditional density $\rho(x)$. This leads to a standard, albeit infinite-dimensional, stochastic control formulation in equation (4.2) below. We then apply the formalism of the dynamic programming (DP) principle, obtaining a Hamilton-Jacobi-Bellman (HJB) equation in the Hilbert space $H^0_\beta$. Previous results imply that the value function $V$ is the unique viscosity solution of this HJB equation. Taking advantage of the wealth-invariance of the chosen exponential utility we guess a smooth candidate solution and apply the verification theorem to conclude that it is in fact the value function of (2.6). The candidate solution can be represented as a conditional expectation of a functional involving $\rho_T(x)$; this expression is succinct but requires a method of approximating $\rho_T$ to obtain numerical results. Accordingly, in Section 5 we take up this task using a particle filter technique.

Returning to the financial point of view let us now consider the case (b) in our classification of possible derivatives. Recall that this means that $B \perp\!\!\!\perp \mathcal{G}_T$ so that the basis at time $T$ is a totally independent random variable with some given distribution function $\mathbb{P}_b$. First, let us combine (3.7) with (2.6) to obtain

$$\begin{aligned}(4.1)\quad \mathbb{E}\Big[-\exp\big(-\gamma(w^\pi_T+\phi(F_T,B))\big)\Big] &= \widetilde{\mathbb{E}}\Big[-\exp\big(-\gamma(w^\pi_T+\phi(F_T,B))\big)\eta_T\Big]\\ &= \widetilde{\mathbb{E}}\Big[-\int_{-\infty}^{\infty}\exp\big(-\gamma(w^\pi_T+\phi(F_T,b))\big)d\mathbb{P}_b\int_{\mathbb{R}}\rho_T(x)dx\Big].\end{aligned}$$

In the last line the two terms factor since the terminal payoff is independent of $X_T$. By the Dynamic Programming principle it now follows that

$$(4.2)\qquad V^\phi(t,w,f,\xi) = \sup_{\pi\in\mathcal{A}^T_t}\widetilde{\mathbb{E}}\Big[-\int_{\mathbb{R}}\exp\big(-\gamma(w^\pi_T+\phi(F_T,b))\big)d\mathbb{P}_b\int_{\mathbb{R}}\rho_T(x)dx\,\big|\,\mathcal{F}_t\Big].$$

In equation (4.2) we have succeeded in reducing the partial observation problem to an equivalent problem with full observation, but at the expense of introducing the $H^0_\beta$-valued process $\rho_t$. The full state is now $(t,w_t,F_t,\rho_t)\in[0,T]\times\mathbb{R}^2\times H^0_\beta$. What we have is a degenerate case of the separation principle (Bensoussan 1992, Ch. 7): we have separated the problems of estimating the unobserved state and the proper utility maximization. Observe that the control only affects the wealth process $w^\pi$; however the dynamics of $w^\pi$ under $\widetilde{\mathbb{P}}$ are unaffected by $\rho$ so that the un-normalized conditional density $\rho_T$ only appears in (4.2) as a scaling factor.

By analogy with the finite dimensional case we formally expect that $V(t,w,f,\rho)$ satisfies the backward parabolic partial differential equation

$$\begin{aligned}(4.3)\quad V_t + \langle\mathcal{L}^*_X\rho, D_\rho(V)\rangle + \frac{1}{2}\langle\mathcal{S}^*\rho\cdot D_{\rho\rho}(V),\mathcal{S}^*\rho\rangle + \frac{1}{2}\sigma^2 f^2 V_{FF}&\\ +\sup_\pi\Big\{\sigma^2\pi f V_{wf} + \frac{1}{2}\sigma^2\pi^2 V_{ww} + \langle\mathcal{S}^*\rho\,,\sigma f D_\rho(V_F)+\sigma\pi D_\rho(V_w)\rangle\Big\} &= 0,\end{aligned}$$

with terminal condition $V(T, w, f, \xi) = -\langle \int e^{-\gamma(w+\phi(f,b))} d\mathbb{P}_b, \xi\rangle$. Here $D_\rho$ is the Fréchet derivative with respect to $\rho$ on $H^2_\beta$ and $\mathcal{L}, \mathcal{S}$ are defined in (3.9)-(3.10).

The following theorem was proved by Gozzi and Şwiech (2000, Theorem 5.4) and states that the value function is the unique solution of (4.3) in an appropriate sense.

**Proposition 2.** *(Gozzi and Şwiech) Let $\overline{\mathcal{A}_t^T}$ be the set of admissible relaxed controls, that is*

$$\overline{\mathcal{A}_t^T} = \{(\Omega, \mathcal{F}, \mathbb{P}, W, \pi),\ \pi \text{ is } \mathcal{F}_t^W\text{-adapted and square-integrable}\}.$$

*Then the value function $\overline{V} \in \mathcal{C}([0,T]\times\mathbb{R}^2\times H^0_\beta)$ minimizing (4.2) over $\overline{\mathcal{A}_t^T}$ is the unique viscosity solution of (4.3).*

We refer to Fleming and Soner (1993) for the rigorous discussion of viscosity solutions of PDEs. Further growth and continuity estimates on the value function can also be made, see Gozzi and Şwiech (2000). Note that in Proposition 2 the Wiener process $W$ is not given a priori but together with the set of admissible portfolios, a notion similar to weak solutions of SDEs.

The HJB equation (4.3) also provides an optimal hedging strategy. According to the maximum principle (Yong and Zhou 1999, Ch. 3), the optimal portfolio weights can be obtained by taking the pointwise supremum in the Hamiltonian term of the HJB equation. Thus, in (4.3), the optimal $\pi$ is

$$\pi_t^* = -\frac{F_t \cdot V_{Fw} + \langle \mathcal{S}^*\rho, D_\rho(V_w)\rangle}{\sigma(F_t) F_t V_{ww}}. \tag{4.4}$$

It would be useful to obtain a more computationally amenable expression for the second term which measures the sensitivity with respect to $\rho_t$. The financial implications of such a hedging strategy for $\phi(F_T, B)$ were studied by Monoyios (2004) in the case of a simple fully-observed model.

4.1. **Linearization with Exponential Utility.** In standard finite-dimensional settings it is well known (Davis 2000, Musiela and Zariphopoulou 2004, Sircar and Zariphopoulou 2005) that exponential utility leads to wealth invariance of the value function. In other words, the initial wealth $w$ trivially factors out of (4.3). Lasry and Lions (1999) showed that under sufficient regularity conditions this phenomenon still occurs for the infinite-dimensional setting of (4.1). Specifically, we guess that there exists a function $\psi(t, f, \rho)$ such that

$$V(t, w, f, \rho) = -e^{-\gamma(w+\psi(t,f,\rho))}. \tag{4.5}$$

Formally substituting (4.5) into (4.3) we obtain

$$\psi_t + \frac{1}{2}\sigma^2 f^2[\gamma\psi_F^2 + \psi_{FF}] + \langle\, \mathcal{L}_X^*\rho, D_\rho\psi\rangle + \frac{1}{2}\Big\langle\, \big(\gamma(D_\rho\psi)^2 + D_{\rho\rho}(\psi)\big)\mathcal{S}^*\rho, \mathcal{S}^*\rho\Big\rangle + \langle\, \mathcal{S}^*\rho, \gamma\sigma f D_\rho(\psi)\,\psi_F\rangle - \frac{\gamma}{2}\Big(\sigma f\psi_F + \langle\, \mathcal{S}^*\rho, D_\rho(\psi)\rangle\Big)^2 = 0. \tag{4.6}$$

After simple algebra, equation (4.6) linearizes to

$$\psi_t + \frac{1}{2}\sigma^2 f^2\psi_{FF} + \langle\, \mathcal{L}_X^*\rho, D_\rho(\psi)\rangle + \langle\, \mathcal{S}^*\rho, \sigma f D_\rho(\psi_F)\rangle + \frac{1}{2}\langle\, \mathcal{S}^*\rho D_{\rho\rho}(\psi)\ , \mathcal{S}^*\rho\rangle = 0. \tag{4.7}$$

We recognize in (4.7) the Kolmogorov partial differential equation for the joint diffusion $(F_t, \rho_t)$ (Da Prato and Zabczyk 1992). Consequently, we can apply the analogue of the Feynman-Kac formula to represent $\psi$ as a $\widetilde{\mathbb{P}}$-conditional expectation with respect to the observable filtration $(\mathcal{F}_t)$. It remains to determine the terminal condition. Starting with $-\mathrm{e}^{-\gamma(w+\psi(T,s,\rho))} = V(T,s,w,\rho)$ and using (4.1) we obtain

$$\begin{aligned} -\mathrm{e}^{-\gamma(w+\psi(T,f,\rho))} &= -\int_{\mathbb{R}} \mathrm{e}^{-\gamma(w+\phi(f,b))} d\mathbb{P}_b \int_{\mathbb{R}} \rho(x)dx \\ \iff \quad -\gamma(w+\psi(T,f,\rho)) &= -\gamma w + \log\int_{\mathbb{R}} \mathrm{e}^{-\gamma\phi(f,b)} d\mathbb{P}_b + \log\int_{\mathbb{R}}\rho(x)dx \\ \iff \quad \psi(T,f,\rho) &= -\frac{1}{\gamma}\log\int_{\mathbb{R}} \mathrm{e}^{-\gamma\phi(f,b)} d\mathbb{P}_b - \frac{1}{\gamma}\log\int_{\mathbb{R}}\rho(x)dx. \end{aligned} \tag{4.8}$$

The first term can be recognized as the certainty equivalent of the claim $\phi(f,B)$ under exponential utility. Let

$$\bar{\psi}(t,f,\xi) \triangleq -\widetilde{\mathbb{E}}\Big[\frac{1}{\gamma}\log\int \mathrm{e}^{-\gamma\phi(F_T,b)}d\mathbb{P}_b + \frac{1}{\gamma}\log\int_{\mathbb{R}}\rho_T(x)dx\Big|\, F_t = f, \rho_t = \xi\Big]. \tag{4.9}$$

Then by Theorem 9.17 in Da Prato and Zabczyk (1992), $\bar{\psi}$ is $C^{1,2,2}([0,T]\times\mathbb{R}_+\times H_\beta^0)$ and is a generalized solution of (4.7). In particular, it is the solution of (4.7) in the viscosity sense. Combining with Proposition 2 it follows that

$$V(t,w,f,\xi) = -\exp\Big\{-\gamma w + \widetilde{\mathbb{E}}\Big[\log\int \mathrm{e}^{-\gamma\phi(F_T,b)}d\mathbb{P}_b + \log\int_{\mathbb{R}}\rho_T(x)dx\Big|\, F_t = f, X_t\sim\xi\Big]\Big\}. \tag{4.10}$$

We see that the value function separates into the certainty equivalent of the claim $\phi$ plus the cost due to partial observations. We can rewrite the last term of (4.10) as $\log(\frac{d\mathbb{P}}{d\widetilde{\mathbb{P}}})$ after which it can be easily seen that its $\widetilde{\mathbb{P}}$-expectation is negative. This also demonstrates that it is square integrable and hence the expectation is well-defined. As expected, the agent is getting a smaller utility from buying the claim $\phi$ because she cannot observe $X$.

Because the additional cost imposed by the uncertainty in $X$ is independent of the given payoff $\phi$, there is cancellation when we apply the formula (2.7) for the indifference price. It follows that the utility-based price $P^\phi$ is

$$P^\phi = \widetilde{\mathbb{E}}\left[-\frac{1}{\gamma}\log\int e^{-\gamma\phi(F_T,b)}d\mathbb{P}_b\right], \tag{4.11}$$

which is the same as what one would obtain in a Black-Scholes world given a "totally unhedgeable" risk $B$ (Becherer 2003). Note the standard emergence (Henderson 2002) of the minimal martingale measure $\widetilde{\mathbb{P}}$ (which assigns zero risk premium to the non-traded X) as the market measure.

As an example, suppose the basis is additive: $\phi(F_T, B) = \tilde{\phi}(F_T) + B$. Then there is some constant $C$ such that $P^\phi(\cdot, t) = \widetilde{\mathbb{E}}[\tilde{\phi}(F_T)|\mathcal{F}_t] - C$ and a fixed cost is subtracted to cover the unhedgeable risk. This result is independent of the postulated model for the forward and the convenience yield, as long as the linearization in (4.7) occurs (note that in particular it covers the full-information setting as well). The fact that exponential utility leads to trivial indifference prices for models with stochastic drift seems to be known in folklore, but we have been unable to find a clear reference in the existing literature. We summarize this observation in the following

**Proposition 3.** *Consider a two-factor $(F_t, X_t)$ model where $F$ is the price of the underlying and $X$ is a stochastic drift factor. Let $\phi(F_T, B)$ be a European claim depending on the traded asset $F$ and some totally unhedgeable risk $B$. Then the indifference price of $\phi$ based on exponential utility is equal to the certainty equivalent of $\phi$ under the minimal martingale measure $\widetilde{\mathbb{P}}$. This holds true even if the second factor $X$ is unobserved.*

*Remark 3.* The HJB equation linearizes only in the one-dimensional case, when the entire system is driven by a univariate Wiener process. In particular, this excludes addition of further factors such as other liquid forwards.

4.2. **Basis Depending on the Convenience Yield.** We turn our attention to the third case (c) of our classification. This means that we now consider $B = B(X_T, Z)$, so that the basis is a noisy version of the unobserved factor. This might occur because the spreads tend to widen when the markets are tight and there is excess demand. In turn, tight supply means that the forward convenience yield is large as there is a strong benefit of holding the physical asset due to increased possibility of shortages. As a result, we see a potential link between the basis $B$ and the future convenience yield $X_T$. The situation is slightly unusual since we assume that

$X$ is unobserved, but nevertheless affects the terminal payoff. One should conclude that $X$ influences other quantities in the market and so the filtering can be improved if we enlarge our set of observables. However, in many cases the monitoring costs are high and any other over-the-counter quotes obtained are stale and/or highly inaccurate. Consequently, the agent may choose to altogether disregard all these data until final settlement. In effect, at time $T$ she obtains new information that was not available to her before.

A basic case is $B = aX_T + \epsilon$, where $a$ is a scaling constant and $\epsilon$ is an independent noise with a prescribed distribution $\mathbb{P}_\epsilon$. Hence, the basis is a noisy linear function of the convenience yield.

To price the claim $\phi(F_T, B)$ in this setting we can re-use the results of the previous section. Indeed, the HJB equation (4.3) remains unchanged and only the terminal condition is modified. Consequently, the guess (4.5) and ensuing linearization still occur. However, now we do *not* have the separability in (4.9). Repeating the computations we obtain

$$\psi(t,f,\xi) = \widetilde{\mathbb{E}}\Big[-\frac{1}{\gamma}\log\int_{\mathbb{R}}\int \mathrm{e}^{-\gamma\phi(F_T,ax+\epsilon)}d\mathbb{P}_\epsilon\,\rho_T(x)\,dx\Big|\,F_t = f, \rho_t = \xi\Big]. \tag{4.12}$$

The corresponding indifference price of $\phi$ involves a difference of two non-linear functionals of $\rho_T$. For example, for additive payoff $\phi(F_T, B) = \tilde{\phi}(F_T) + aX_T + \epsilon$, $\epsilon \sim \mathcal{N}(\mu_\epsilon, \sigma_\epsilon^2)$, $\epsilon \perp\!\!\!\perp \mathcal{G}_T$, the indifference price is given by

$$\begin{aligned} P^\phi(t,f,\xi) &= \frac{1}{\gamma}\widetilde{\mathbb{E}}\Big[\log\int_{\mathbb{R}}\rho_T(x)dx\Big|\mathcal{F}_t\Big] - \psi(t,f,\xi) \\ &= \widetilde{\mathbb{E}}\Big[\tilde{\phi}(F_T) - (\mu_\epsilon + \frac{\gamma\sigma_\epsilon^2}{2}) - \frac{1}{\gamma}\log\frac{\int\rho_T(x)\,dx}{\int \mathrm{e}^{-\gamma a x}\rho_T(x)\,dx}\Big|F_t = f, \rho_t = \xi\Big]. \end{aligned} \tag{4.13}$$

4.3. **Nonlinear Dynamics.** Because all the formulas in the preceding section are in terms of the infinitesimal generators of $(F, X)$, our results continue to hold when the dynamics (2.1) are not linear. One interesting case to consider is a local volatility model for the forward process, such as the constant elasticity of variance (CEV) extension of (2.2):

$$\begin{cases} dF_t = F_t(\mu - X_t)\,dt + \sigma F_t^{1+\beta}\,dW_t, \\ dX_t = \kappa(\theta - X_t)dt + c\alpha\,dW_t + \sqrt{1-c^2}\alpha\,dW_t^\perp. \end{cases} \tag{4.14}$$

The advantage of (4.14) is that the elasticity parameter $-1/2 \le \beta \le 1$ can be used to improve empirical fitting of the model. The price $F$ now enters the wealth dynamics

$$dw_t^\pi = (\mu - \pi X_t)\,dt + \pi\sigma F_t^\beta dW_t, \tag{4.15}$$

as well as the dynamics of $X_t$ under $\widetilde{\mathbb{P}}$, however our filtering analysis remains unchanged. We illustrate the behavior of (4.14) in Section 5.2.

## 5. Numerical Implementation and Results

In Sections 4.1 and 4.2 we obtained the indifference price and the value function corresponding to a claim $\phi(F_T, B)$ in terms of expectations of functionals of the un-normalized conditional density $\rho_T$ of $X_T$. To actually *compute* these expectations we employ Monte Carlo simulations. The basic idea is to simulate a large sample of random variables approximating $\rho_T$, to compute the desired functional of each sample and average the results. For linear models like (2.2) we can use the Kalman filter to filter $X$ exactly; recall that (3.13) explicitly gives the evolution of the two-dimensional sufficient statistic $(\hat{x}_t, P_t)$ in that case. Unfortunately, in general, such as with the CEV model (4.14), $\rho_T$ is an infinite-dimensional object and it is not immediately obvious how to efficiently approximate it. To this end we adapt the Zakai particle filter algorithm based on the description of Crišan et al. (1998). One advantage of this approach is that we directly compute the un-normalized density $\rho_T$. In contrast, if a Kalman filter is used then the true conditional distribution of $X_T$ under $\mathbb{P}$ is computed first, followed by a second step of approximating the Radon-Nikodym density $\eta_T$.

Given the form of (4.13) and (4.10), we work directly under the reference measure $\widetilde{\mathbb{P}}$. Under $\widetilde{\mathbb{P}}$ the forward price is Markov with respect to $(\mathcal{F}_t)$, so that it can be simulated independently of everything else. With an eye towards actual computer implementation, fix a time grid $t_k = k\Delta t$, with $t_0 = 0, t_M = T$. The forward process is then simulated using an Euler discretization of (3.3),

$$F_{t_k} = F_{t_{k-1}} + \sigma(t_{k-1}, F_{t_{k-1}})\sqrt{\Delta t}\,\epsilon_k, \qquad \epsilon_k \sim \mathcal{N}(0,1). \tag{5.1}$$

Given such $\bar{N}$ paths of $F$, we describe in the next section how to compute $\rho_{t_k}(x)$ along each path.

### 5.1. Particle Filtering for the Zakai Equation.

The main idea of particle filtering is to approximate the random measure $p_t$ of (3.6) by a pure point measure $A^N(t)$ which is an occupation measure of $N(t)$ particles $\{\alpha^i_t\}_{i=1}^{N(t)}$. Fix an auxiliary step parameter $\ell \in \mathbb{N}$ ($\ell$ is on the scale of $4-8$). The measure $A^N(t)$ is recursively updated every $\ell$ grid points (i.e. at each $t_{\ell \cdot k}$, $k = 1, \dots$) using a mutation/branching mechanism on the level of each particle.

As a first step, given an initial distribution $X_0 \sim \xi$, and a path $\{F_t\}_{t=k\Delta t}$ we construct a probability measure $A^N(0)$ by drawing $N$ i.i.d. samples from $\xi$:

$$A^N(0) = \frac{1}{N}\sum_{i=1}^{N} \delta_{\alpha_0^i}, \tag{5.2}$$

where $\alpha_0^i$ are independent and identically distributed random variables with common distribution $\xi$, and $\delta_x$ is the Dirac mass. During an interval $[t_{\ell\cdot k}, t_{\ell\cdot(k+1)})$, each particle $\alpha^i$ evolves independently according to the law of $X$ under $\widetilde{\mathbb{P}}$. Hence, for $t \in [t_{\ell\cdot k}, t_{\ell\cdot(k+1)})$,

$$\begin{aligned}\alpha_t^i = \alpha_{t_{\ell\cdot k}}^i &+ \int_{t_{\ell\cdot k}}^t \Big(b(F_s,\alpha_s^i) - \frac{c\,a(F_s,\alpha_s^i)}{\sigma(F_s)F_s} h(F_s,\alpha_s^i)\Big)\,ds \\ &+ \int_{t_{\ell\cdot k}}^t c\,a(F_s,\alpha_s^i)\frac{1}{\sigma(F_s)}dF_s + \int_{t_{\ell\cdot k}}^t \sqrt{1-c^2}a(F_s,\alpha_s^i)\,dW_s^i,\end{aligned} \tag{5.3}$$

with $A^N(t) = \frac{1}{N(t_{\ell\cdot k})}\sum_{i=1}^{N(t_{\ell\cdot k})}\delta_{\alpha_t^i}$, where $\{W^i\}$ is a collection of $N(t_{\ell\cdot k})$ independent Brownian motions.

At time $t_{\ell\cdot(k+1)}$ mutation occurs. Let (cf. (3.5))

$$\mu_k^i \triangleq \exp\Big(\int_{t_{\ell\cdot k}}^{t_{\ell(k+1)}} \frac{h(F_t,\alpha_t^i)}{\sigma(F_t)^2}\,dF_t - \frac{1}{2}\int_{t_{\ell\cdot k}}^{t_{\ell(k+1)}} \frac{h(F_t,\alpha_t^i)^2}{\sigma(F_t)^2}\,dt\Big). \tag{5.4}$$

Then each particle $\alpha_{t_{\ell\cdot(k+1)}}^i$ branches into either $\lfloor \mu_k^i \rfloor$ or $\lceil \mu_k^i \rceil$ offspring such that their expected number is precisely $\mu_k^i$. The branching of each particle is independent of all the others and only depends on the behavior of $F_t$ on $[t_{\ell\cdot k}, t_{\ell\cdot(k+1)})$. The new particles inherit the location of their parent. The fact that the possible number of offspring is as close to $\mu_k^i$ as possible is to reduce the variance of $A^N(t)$, cf. Crišan et al. (1998). We can now explain that the use of the auxiliary step parameter $\ell$ is to obtain a good approximation to the integrals of the form $\int_{t_{\ell\cdot k}}^{t_{\ell\cdot(k+1)}} g\,dF_s$ that appear in (5.3) and (5.4). On a computer these integrals are of course replaced with finite sums using the increments $(F_{(k+1)\Delta t} - F_{k\Delta t})$.

Since $\widetilde{\mathbb{E}}\left[\mu_k^i\right] = 1$, the expected number of particles always remains at their initial number, $\widetilde{\mathbb{E}}[N(t_{\ell\cdot(k+1)})] = \widetilde{\mathbb{E}}[N(t_{\ell\cdot k})] = N$. Moreover, it can be shown that with probability one the algorithm does not explode or die out. Let $M_F(\mathbb{R})$ be the space of finite measures on the real line with the topology of weak convergence. The results of Crišan (2003) show that for any $f$ continuous and bounded in $\mathbb{R}$, $\langle A^N(t), f\rangle$ is square integrable. Furthermore, if $N \to \infty$ and $\Delta t \to 0$ such that $N\sqrt{\Delta t} \to \infty$ (the number of particles grows quadratically in step size) then

$A^N(t)$ weakly converges to a measure $\tilde{p}(t) \in M_F(\mathbb{R})$ satisfying

$$\langle \tilde{p}(t), f \rangle = \langle \xi, f \rangle + \int_0^t \langle \tilde{p}(s), \mathcal{L}_X f \rangle \, ds + \int_0^t \langle \tilde{p}(s), \mathcal{S} f \rangle \, dF_s \quad \text{a.s.} \tag{5.5}$$

By the characterization theorem of Ocone and Kurtz (1988), we then have $\tilde{p}(t) = p_t$ of (3.6).

5.2. **Comparative Statics.** We have implemented both the Kalman filtering method using (3.15) and the Zakai particle filter using (5.4) and applied it to the indifference price formula (4.10). The results below should be seen as simple demonstrations of our methodology. We have not carried out thorough empirical fitting, and it would be interesting to select a parametric subset of (2.1) and carry out a full econometric testing of the resulting commodity forward dynamics.

We use the CEV "Gibson-Schwartz" model (4.14) with representative parameter values of $\mu = 0.06, \theta = X_0 = 0, \sigma = 0.4, \alpha = 0.5, \kappa = 2, \beta = 0, c = 0.7$ and risk-aversion $\gamma = 0.1$. Those values are meant to capture the high volatility of the forward convenience yield $X$, along with strong positive correlation to the forward $F$ itself. For simplicity, we take the mean level of $X$ to be zero. For the reader's convenience, we recall that the model (2.2) now says

$$\begin{cases} dF_t = (0.06 - X_t) F_t \, dt + 0.4 F_t \, dW_t, \\ dX_t = -2X_t \, dt + 0.5 \cdot (0.7 dW_t + \sqrt{0.51} dW_t^{\perp}). \end{cases}$$

The first numerical experiment that we perform is understanding the term-structure of the quantity $\widetilde{\mathbb{E}}[\log \int_{\mathbb{R}} \rho_T(x) dx]$ which is the utility-based cost of being unable to observe $X$. Note that in terms of the particle filter, $\widetilde{\mathbb{E}}[\log \int_{\mathbb{R}} \rho_T(x) dx] = \widetilde{\mathbb{E}}[\log\big(N(T)/N(0)\big)]$, where $N(T)$ is the number of particles at time $T$, as in Section 5.1. As Figure 1 illustrates, the utility adjustment is almost linear in time to maturity with a slight convexity in the beginning. Figure 1 also shows that decreasing the mean-reversion parameter $\kappa$, which increases the variance of $X_t$ and consequently of $\rho_t$, leads to a higher cost of partial observations since the uncertainty surrounding $X$ increases.

INSERT FIGURE 1 HERE

We next try to understand the implications of the formula (4.13) for the cost of uncertainty when the basis depends on $X$. In particular, we check the effect of various parameters of the $X_t$-dynamics, as well as the effect of time to maturity $T$.

Figure 4 shows the results for $T = 3, 6$ and $9$ months as we vary the initial mean $X_0 \sim \mathcal{N}(x_0, 0.1^2)$ and the elasticity parameter $\beta$. We see that the last term in (4.13) captures two effects coming from $X$. First, it provides direct adjustment for the derivative price since the final payoff depends on $X_T$. Thus, larger $x_0$ values increase the price as the expected payoff increases. Second, (4.13) provides a secondary adjustment based on the amount of uncertainty surrounding $X$. Accordingly, as time to maturity increases, the former influence of $x_0$ weakens due to the stationarity of the Ornstein-Uhlenbeck model (2.2) we assumed for $X$. From a different angle, if we take a smaller elasticity $\beta$ this decreases the volatility of $F$ which in turn decreases the uncertainty surrounding the drift $X$. As a result, the utility-based price in Figure 3 increases, similar to the effect with respect to $\kappa$ in Figure 1.

INSERT FIGURE 4 HERE

## 6. Conclusion & Extensions

We have introduced a new model for pricing claims written on an over-the-counter energy contract $F^{loc}$. Our model has two factors for the dynamics of the benchmark reference contract $F$ and includes explicit modeling of the basis risk. We have seen that in the cases where the basis is totally independent from $(F, X)$, the indifference price of an $F^{loc}$-contingent claim is simply the certainty equivalent of this claim under the minimal martingale measure $\widetilde{\mathbb{P}}$. In the more interesting case where the basis also depends on the unobserved factor we obtained the reduced-form expression (4.12) which may be explicitly computed using a Zakai particle filter algorithm. The most general case (d) would be to model the basis as a separate stochastic process, possibly correlated with $F$ and $X$. This would lead to an HJB equation similar to (4.3). However, as mentioned before, we do not know how to solve/linearize it if the number of observable state variables is bigger than one. Infinite-dimensional problems like (4.3) are very hard in general and few tools exist to tackle them.

The results of our numerical experiments demonstrate the feasibility of using partially observed models in commodity pricing. Use of partial observations removes model inconsistencies with respect to the forward curve. At the same time our approach is still consistent with existing spot/convenience models in the literature. Moreover, we can incorporate nonlinear dynamics which should allow for better econometric fit. To obtain a fully satisfactory model for empirical data, further extensions are likely to be necessary. For example, time-dependent parameters would surely be needed as most energy prices exhibit high degrees of seasonality. Moreover, the volatility must be time-dependent to incorporate the well-known Samuelson effect. The

Samuelson effect states that forward volatility $\sigma(F_t)$ increases as time to maturity decreases. Likewise, stochastic interest rates might need to be considered for long-dated contracts.

Another interesting direction to consider is the use of a *rolling* forward as the underlying asset, as in the case of commodity indices. That is, we could take $F_t = F(t, t+x)$ for a fixed time-to-maturity $x$. For instance, this could model the common (and controversial) practice of hedging with the nearest maturity contract, taking advantage of increased liquidity in the nearby forwards. We refer the reader to Neuberger (1999) for a more detailed analysis and discussion of associated rollover risk. Putting aside the business soundness of such strategies, the analytical difficulty lies in writing down the dynamics for $F(t, t+x)$. In general, this requires specifying the dynamics of all maturities $F(t, t')$ and causes the appearance of a non-trivial risk-neutral drift for $F_t$. Such an approach would lead us towards the universe of Heath-Jarrow-Morton term structure commodity models.

## Appendix A. Proof of Proposition 1

*Proof.* We follow in the footsteps of Nagai (2000) and Sekine (2003). Re-write $(F_T)^\lambda = (F_0)^\lambda \exp(\int_0^T \lambda\, dY_t)$. Then using (3.14) we change measures to $\widetilde{\mathbb{P}}$ after which it suffices to compute $\widetilde{\mathbb{E}}[L_T]$ where

$$L_t \triangleq \exp\Big(\int_0^t \Big\{\frac{\mu - \hat{x}_s - \frac{1}{2}\sigma^2}{\sigma^2} + \lambda\Big\} d(Y_s) - \frac{1}{2}\int_0^t \big(\mu - \hat{x}_s - \frac{1}{2}\sigma^2\big)^2 \frac{1}{\sigma^2}\, ds\Big). \tag{A.1}$$

We shall guess that $\widetilde{\mathbb{E}}[L_t]$ is an exponential of a linear function of the current best estimate $\hat{x}_0$. Accordingly, we look for $\chi_t$ such that $L_t e^{\chi_t}$ is a $\widetilde{\mathbb{P}}$-martingale and $\chi_t = 2\lambda g_t \cdot \hat{x}_t + \lambda k_t$ for some deterministic $g_t$ and $k_t$. Using (3.13) we compute

$$d\, e^{\chi_t} = e^{\chi_t}\big\{(2\lambda \hat{x}_t\, \dot{g}_t + \lambda \dot{k}_t)\, dt + 2\lambda g_t\, d\hat{x}_t + \frac{1}{2}(2\lambda g_t U_t)^2 \frac{1}{\sigma^2} dt\big\}.$$

Here $U_t \triangleq c\sigma\alpha - P_t$ and $\dot{g}, \dot{k}$ denote derivatives with respect to $t$. Using (3.12) and combining with (A.1) it follows that

$$\begin{aligned} d(L_t e^{\chi_t}) = L_t e^{\chi_t}\Big(&(\mu - \hat{x}_t - \frac{1}{2}\sigma^2 + \lambda\sigma^2)\frac{1}{\sigma^2} dY_t + \big(\frac{1}{2}\lambda^2\sigma^2 + \lambda(\mu - \hat{x}_t - \frac{1}{2}\sigma^2)\big)\, dt \\ &+ 2\lambda g_t \Big\{\frac{U_t}{\sigma^2} dY_t + (\kappa(\theta - \hat{x}_t) - (\mu - \hat{x}_t - \frac{1}{2}\sigma^2)\frac{U_t}{\sigma^2})\, dt\Big\} \\ &+ (2\lambda \hat{x}_t \dot{g}_t + \lambda \dot{k}_t)\, dt + \frac{1}{2}(2\lambda g_t U_t)^2 \frac{1}{\sigma^2}\, dt + 2\lambda g_t(\mu - \hat{x}_t - \frac{1}{2}\sigma^2 + \lambda\sigma^2)\frac{U_t}{\sigma^2}\, dt\Big). \end{aligned} \tag{A.2}$$

Next we pick $g_t$ and $k_t$ such that all the $dt$-drift terms in (A.2) disappear. This produces the following ordinary differential equations satisfied by $g_t$ and $k_t$:

$$\text{(A.3)} \qquad \begin{cases} \dot{g}_t - \dfrac{1}{2} - \kappa g_t = 0, \\ \dot{k}_t + \mu + 2\lambda g_t U_t + \dfrac{\lambda - 1}{2}\sigma^2 + 2\kappa\theta g_t + 2\lambda g_t^2 \dfrac{U_t^2}{\sigma^2} = 0. \end{cases}$$

For boundary conditions we take $g(T) = k(T) = 0$. Since $P_t$ is bounded, so is $g_t$ and $k_t$ and therefore $L_t e^{\chi_t}$ is indeed a true martingale. Consequently, $\widetilde{\mathbb{E}}[L_T] = \widetilde{\mathbb{E}}[L_T e^{\chi_T}] = L_0 e^{\chi_0} = e^{2\alpha g_0 \hat{x}_0 + \alpha k_0}$. Solving for $g_0, k_0$, in (A.3) we obtain the result of Proposition 1. □

## References


Becherer, D. (2003), 'Rational hedging and valuation of integrated risks under constant absolute risk aversion', *Insurance: Mathematics and Economics* **33**(1), 1–28.

Bensoussan, A. (1992), *Stochastic Control of Partially Observable Systems*, Cambridge University Press, Cambridge, UK.

Brendle, S. and Carmona, R. (2005), 'Hedging in partially observed markets'. Preprint, Princeton University.

Brennan, M. (1958), 'The supply of storage', *American Economic Review* **48**, 50–72.

Carmona, R. and Durrleman, V. (2003), 'Pricing and hedging spread options', *SIAM Review* **45(4)**, 627–685.

Carmona, R., ed. (2006), *Volume on Indifference Pricing*, Princeton University Press.

Carmona, R. and Ludkovski, M. (2004), Spot convenience yield models for energy assets, *in* 'Mathematics of Finance', Vol. 351 of *Contemporary Mathematics*, AMS, pp. 65–79.

Casassus, J. and Collin-Dufresne, P. (2005), 'Stochastic convenience yield implied from commodity futures and interest rates', *Journal of Finance* **60**(5), 2283–2331.

Crišan, D., Gaines, J. and Lyons, T. (1998), 'Convergence of a branching particle method to the solution of the Zakai equation', *SIAM Journal of Applied Mathematics* **58(5)**, 1568–1590.

Crišan, D. (2003), 'Exact rate of convergence for a branching particle approximation to the solution of the Zakai equation', *Annals of Probability* **31**(2), 693–718.

Da Prato, G. and Zabczyk, J. (1992), *Stochastic Equations in Infinite Dimensions*, Cambridge University Press, Cambridge, UK.

Davis, M. (2000), 'Optimal hedging with basis risk'. Preprint, Technical University of Vienna.

Duan, J. and Pliska, S. (2004), 'Option valuation with co-integrated asset prices', *Journal of Economic Dynamics and Control* **28**, 727–754.

El Karoui, N. and Rouge, R. (2000), 'Pricing via utility maximization and entropy', *Mathematical Finance* **10(2)**, 259–276.

Eydeland, A. and Wolyniec, K. (2003), *Energy and Power Risk Management: New Developments in Modeling, Pricing and Hedging*, John Wiley&Sons, Hoboken, NJ.

Fama, E. and French, K. (1988), 'Business cycles and the behaviour of metals prices', *Journal of Finance* **43**, 1075–1093.

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

Gibson, R. and Schwartz, E. S. (1990), 'Stochastic convenience yield and the pricing of oil contingent claims', *Journal of Finance* **45(3)**, 959–976.

Gozzi, F. and Şwiech, A. (2000), 'Hamilton Jacobi Bellman equations for the optimal control of the Duncan Mortensen Zakai equation', *Journal of Functional Analysis* **172**(2), 466–510.

Harvey, A. (1989), *Forecasting Structural Time Series Models and the Kalman Filter*, Cambridge University Press, Cambridge, UK.

Henderson, V. (2002), 'Valuation of claims on nontraded assets using utility maximization', *Mathematical Finance* **12(4)**, 351–373.

Hilliard, J. and Reis, J. (1998), 'Valuation of commodity futures and options under stochastic convenience yields, interest rates, and jump diffusions in the spot', *Journal of Financial and Quantitative Analysis* **33(1)**, 61–86.

Karatzas, I. and Shreve, S. (1998), *Methods of Mathematical Finance*, Applications of Mathematics, Vol 39, Springer-Verlag, New York, NY.

Lakner, P. (1998), 'Optimal trading strategy for an investor: the case of partial information', *Stochastic Processes and their Applications* **76**, 77–97.

Lasry, J.-M. and Lions, P.-L. (1999), 'Contrôle stochastique avec informations partielles et applications à la finance', *Comptes Rendus Academie de Sciences Paris, Série I* **328**, 1003–1010.

Monoyios, M. (2004), 'Performance of utility-based strategies for hedging basis risk', *Quantitative Finance* **4**, 245–255.

Musiela, M. and Zariphopoulou, T. (2004), 'A valuation algorithm for indifference prices in incomplete markets', *Finance & Stochastics* **8**(3), 399–414.

Nagai, H. (2000), Risk-sensitive dynamic asset management with partial information, *in* T. Hida and B. Rajput et al., eds, 'Stochastics in Finite and Infinite Dimension. A Volume in Honor of G. Kallianpur', Birkhauser, Boston, pp. 321–340.

Neuberger, A. (1999), 'Hedging long-term exposures with multiple short-term futures contracts', *Review of Financial Studies* **12**, 429–459.

Ocone, D. and Kurtz, T. (1988), 'Unique characterization of conditional distributions in nonlinear filtering', *Annals of Probability* **16**(1), 80–107.

Pham, H. and Quenez, M.-C. (2001), 'Optimal portfolio in partially observed stochastic volatility models', *Annals of Applied Probability* **11(1)**, 210–238.

Runggaldier, W. (2004), Estimation via stochastic filtering in financial market models, *in* 'Mathematics of Finance', Vol. 351 of *Contemporary Mathematics*, AMS.

Schwartz, E. (1997), 'The stochastic behavior of commodity prices: Implications for valuation and hedging', *Journal of Finance* **52(3)**, 922–973.

Schwartz, E. and Smith, J. (2000), 'Short-term variations and long-term dynamics in commodity prices', *Management Science* **46**(7), 893–911.

Sekine, J. (2003), Power-utility maximization for linear Gaussian factor model under partial information, Technical report, Osaka University.

Sircar, R. and Zariphopoulou, T. (2005), 'Bounds and asymptotic approximations for utility prices when volatility is random', *SIAM Journal of Control and Optimization* **43**(4), 1328–1353.

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

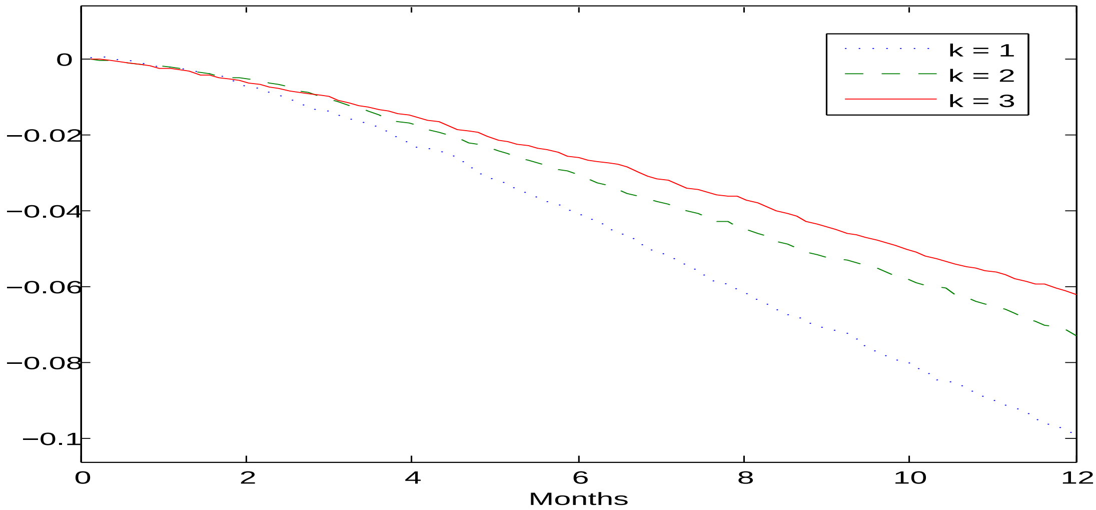


FIGURE 1. Term structure of $\log \int \rho_T(x)\,dx$ for different mean-reversion rates $\kappa$ in (2.2). The results were averaged over $\bar{N} = 10,000$ simulations of the particle filter based on (5.4) with $N = 500$ initial particles per run.

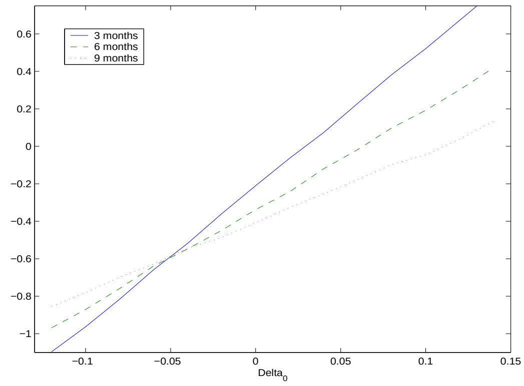


Figure 2 : Varying the initial mean $x_0$ of $X$.

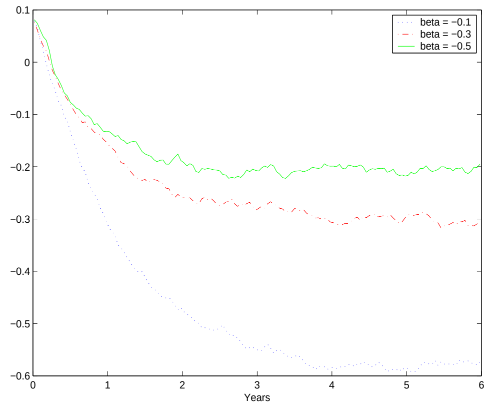


Figure 3 : Varying the CEV elasticity $\beta$.

FIGURE 4. Comparative Statics for the last term in (4.13). We use the particle filter from (5.4) to evaluate $\frac{1}{\gamma}(\log(\int \rho_T(x)\,dx - \log \int e^{-\gamma a x}\rho_T(x)\,dx)$ with risk aversion parameter $\gamma = 0.1$, noise parameter $a = 10$, and $\bar{N} = 10,000$ simulations.

DEPARTMENT OF OPERATIONS RESEARCH AND FINANCIAL ENGINEERING, ALSO WITH BENDHEIM CENTER FOR FINANCE, PRINCETON UNIVERSITY, PRINCETON, NJ 08544 USA

*E-mail address*: rcarmona@princeton.edu

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MICHIGAN, 530 CHURCH ST. ANN ARBOR, MI 48109 USA, PH: (734)763-3204 FAX: (734)763-0937

*E-mail address*: mludkov@umich.edu