
A MAXIMAL AFFINE STOCHASTIC VOLATILITY MODEL OF OIL PRICES

W. KEENER HUGHEN*

This study develops and estimates a stochastic volatility model of commodity prices that nests many of the previous models in the literature. The model is an affine three-factor model with one state variable driving the volatility and is maximal among all such models that are also identifiable. The model leads to quasi-analytical formulas for futures and options prices. It allows for time-varying correlation structures between the spot price and convenience yield, the spot price and its volatility, and the volatility and convenience yield. It allows for expected mean-reversion in the short term and for an increasing expected long-term price, and for time-varying risk premia. Furthermore, the model allows for the situation in which options' prices depend on risk not fully spanned by futures prices. These properties are desirable and empirically important for modeling many commodities, especially crude oil. © 2009 Wiley Periodicals, Inc. *Jrl Fut Mark* 30:101–133, 2010

INTRODUCTION

Energy derivatives markets have grown tremendously in recent years. Recent deregulation in the natural gas and electricity markets has sparked an explosion in the number and types of complex derivatives and financial structures offered by suppliers and speculators. The number of participants in the derivatives

*Correspondence author, University of North Carolina, Charlotte, 9201 University City Boulevard, Charlotte, North Carolina 28223. Tel: +704 687 7638, Fax: +704 687 6987, e-mail: whughen@unc.edu

Received 8 July 2008; Accepted 10 March 2009

■ *W. Keener Hughen is at the University of North Carolina, Charlotte, North Carolina.*

markets has also greatly increased. This is especially true in the crude oil market, which is the largest commodity market in the world.

Many commodities differ from financial assets in several ways. Oil, in particular, is often strongly backwardated, meaning that oil futures prices are often below spot prices, and the degree of backwardation is highly variable over time. Also, an economic argument can be made for prices to exhibit mean-reversion from both the supply and the demand sides. An abnormally high price should induce higher cost producers to enter the market, increasing supply and ultimately causing the price to decrease, and conversely, an abnormally low price will drive many producers to leave the market, decreasing supply. Similarly, higher prices may induce consumers to shift their behavior and become more energy efficient or seek alternative energy sources, and vice versa when prices are low. Assuming that the risk premium is not time-varying in a way that offsets the risk-neutral mean-reversion, this will result in the property that volatility of futures prices decreases with maturity. This property is known as the Samuelson effect and is empirically well documented.

Another important property displayed by many commodities, especially oil, is that the spot price is strongly heteroscedastic. The empirical evidence in support of stochastic volatility has been documented by a number of researchers. Duffie and Gray (1995) compute both realized and implied volatilities for various energy commodities including oil and find that the constant volatility hypothesis is rejected at the 95% confidence level and recommend the use of stochastic models for the volatility. Eydeland and Wolyniec (2003) document that oil futures markets have pronounced implied volatility smiles. Litzenberger and Rabinowitz (1995) find that oil price volatility is highly variable and that oil backwardation is positively related to the riskiness of the future spot price, in agreement with empirical predictions of their theoretical model. Routledge, Seppi, and Spatt (2000) note that the correlation structure between the spot price and the convenience yield should vary over time; this suggests a time-varying instantaneous correlation structure. Furthermore, Pindyck (2004) argues that volatility should be positively correlated with both spot price and convenience yield, so that on an average, an increase in volatility should raise both prices and the marginal value of storage in the short run.

Despite this evidence, many of the earlier commodity pricing models in the literature assume Gaussian dynamics for the spot price. This is most likely due to the tractability and analytic formulas for futures and other contingent claims prices afforded by Gaussian models. For example, the early one-factor models of Brennan and Schwartz (1985), Ross (1997), and Schwartz (1997) and the two-factor models of Gibson and Schwartz (1990), Gabillon (1995), Schwartz and Smith (2000), and Korn (2005) are Gaussian and in particular assume that the spot price has constant volatility. These models all have closed-form formulas

for futures and options prices. The two-factor models have stochastic convenience yield and are better able to reproduce both the Samuelson effect and the time-varying degree of backwardation for many commodities.

The Gaussian two-factor models have been extended in several ways. Schwartz (1997) and Casassus and Collin-Dufresne (2005) include stochastic interest rates along with spot price and convenience yield; the interest rate is assumed to follow an Ornstein–Uhlenbeck process, so that, again, each model is Gaussian. Cortazar and Naranjo (2006) propose a class of N -factor Gaussian models; when $N = 3$ their model includes stochastic convenience yield drift along with spot price and convenience yield. These three-factor models are shown to significantly improve the goodness-of-fit over the two-factor models, and principle component analyses (Clewlow & Strickland, 1999) show that three factors explain roughly 95% of the variation in futures prices.

Several studies have extended the Gaussian two-factor models to incorporate stochastic volatility. The two-factor model of Nielsen and Schwartz (2004) sets the volatility equal to a linear function of the convenience yield. Richter and Sorensen (2002) propose a three-factor model that includes a stochastic volatility state variable independent of spot price and convenience yield. Deng (1999) incorporates stochastic volatility to model the joint process followed by electricity and natural gas. Trolle and Schwartz (2008) propose an HJM-type model and conduct an extensive empirical investigation of volatility in the crude oil market. These stochastic volatility models, compared with their Gaussian counterparts, necessarily have more realistic volatility structures.

The model presented in this study is a three-factor affine stochastic volatility model for commodity prices that nests many of the models previously mentioned. In particular, it nests the one-factor models of Brennan and Schwartz (1985), Ross (1997), and Schwartz (1997); the two-factor models of Gibson and Schwartz (1990), Gabillon (1995), Schwartz and Smith (2000), and Korn (2005); and the stochastic volatility models of Nielsen and Schwartz (2004) and Richter and Sorensen (2002) (when there is no seasonality). Also, the Heston (1993) stochastic volatility model is the special case of the model with constant convenience yield and more restrictive volatility structure.

The model presented here is maximal among all identifiable three-factor affine models in which the volatility is driven by a single state variable and is therefore not itself a special case of a more general three-factor affine model. The model allows for time-varying correlation structures between the spot price and the convenience yield, the spot price and the volatility, and the volatility and the convenience yield. It allows for expected mean-reversion in the short term and for an increasing expected long-term price and it allows for the Samuelson effect and for time-varying degrees of backwardation. The model also leads to

quasi-analytical formulas for futures and options prices. Futures prices are computed by solving a system of ordinary differential equations whereas options prices are computed with a Fourier inversion approach.

Additionally, in contrast with all Gaussian models and the stochastic volatility model Richter and Sorensen (2002), the model presented here allows for the situation in which options prices depend on risk not fully spanned by futures prices. This feature is desirable because options need not be redundant securities. If options were redundant then they could be hedged by trading futures contracts only. In a recent empirical investigation, Trolle and Schwartz (2008) find strong evidence that crude oil options do depend on risk not fully spanned by futures, and they refer to this property as *unspanned stochastic volatility* (USV).

The techniques used to estimate the model are similar to those in the term structure literature, which use bond prices to estimate and fit the dynamics of the short rate. Formulas are derived for futures prices and then daily price data on crude oil futures are used to estimate the parameters and the state variables of the model. The estimation method used is Markov Chain Monte Carlo (MCMC). MCMC is well suited here because it is a unified estimation procedure, simultaneously estimating both parameters and the latent state variables. Furthermore it is based on conditional simulation, thereby avoiding any optimization or unconditional simulation; this is especially appealing because of the high dimensionality of the model.

Results indicate that this model improves over all one- and two-factor (nested) models previously investigated. The most general affine two-factor Gaussian model is also estimated using MCMC techniques, and the root mean squared error (RMSE) for the futures data is compared with the RMSE using the maximal stochastic volatility model. The RMSE is significantly lower using the maximal stochastic volatility model both in sample and out of sample. The model is also compared with the most general affine three-factor Gaussian model and the nested three-factor stochastic volatility model of Richter and Sorensen (2002). These two models are estimated, again using MCMC techniques, and the RMSEs are computed. The RMSE using the maximal stochastic volatility model is lower overall, though somewhat comparable to the RMSE using the three-factor Gaussian model. However, the models can have significantly different option pricing implications; in particular, Gaussian models cannot reproduce the implied volatility smiles observed for many commodities.

Additional out of sample analysis is performed in which the model-implied options prices are computed, using the estimates obtained with futures data, and compared to actual options prices. Neither model performs particularly well on this out of sample test. This suggests that volatility cannot be accurately inferred from futures prices, in agreement with the empirical findings of

Trolle and Schwartz (2008). Further evidence of USV in crude oil markets is provided by the low correlations between the estimated volatility state variable and various measures of realized volatility. Restrictions necessary to generate USV in the model are derived, and the model is re-estimated with these restrictions. The estimated volatility is very highly related with realized volatility, indicating that the USV version of the model may provide the best all around fit to both futures and options prices.

The remainder of the study is organized as follows. The next section presents and quantitatively describes the model. First, the risk-neutral dynamics are described and the model is shown to be the maximal model that is also identifiable among all affine three-factor models for which the volatility is driven by a single state variable. The model is completed by specifying the dynamics in the true, physical measure and it is shown that the model nests many of the previous models in the literature. Futures prices and options prices are then derived. The third section discusses the empirical methodology and gives a detailed description of the MCMC procedure. The results and comparisons with previous models are reported in the fourth section. The fifth section discusses USV, and the sixth section concludes.

THE MODEL

It is assumed that the spot price S_t is subject to three independent sources of risk, and there is given a probability space and a filtration generated by a standard Brownian motion W_t^* in \mathbb{R}^3 such that the spot price is an Ito process with respect to W_t^* . The underlying measure has the property that the price process of any derivative security is a Martingale after discounting and is referred to as the risk-neutral measure.

Following Gibson and Schwartz (1990), the (instantaneous net) convenience yield y_t may be viewed as the net dividend yield accruing to the marginal owner of the commodity. The total expected instantaneous return over an infinitesimal interval dt to the owner of the commodity is the sum of the expected instantaneous price return and the convenience yield y_t ; on the other hand, the total expected instantaneous return in the risk-neutral measure is equal to the short rate r_t . Therefore, r_t is equal to the sum of y_t and the risk-neutral expectation of the instantaneous price return, or

$$\frac{1}{S_t} E_t^*[dS_t] = (r_t - y_t)dt. \quad (1)$$

Equation (1) defines the convenience yield y_t . It will be convenient to work with the (instantaneous) cost of carry $\delta_t = r_t - y_t$ instead. Because futures prices are risk-neutral expectations of future spot prices, δ_t , is equal to the

slope of the log-futures price curve at maturity equal to zero and therefore would be observable if futures contracts with arbitrarily small maturity were traded.

Many of the earlier models assume at most two sources of risk, spanned by S_t and δ_t , and that these two processes have Gaussian dynamics. More recent empirical research provides strong evidence that volatility is stochastic for many commodities, including oil in particular, and several models have been proposed that allow for stochastic volatility. This study proposes a relatively simple, intuitive model that allows for stochastic volatility and nests many of these previous models.

Risk-Neutral Dynamics

Let V_t be the volatility of S_t

$$V_t = \frac{1}{S_t^2} E_t[dS_t^2].$$

With three sources of uncertainty, the dispersion term of the risk-neutral process for S_t is of the form $\Sigma_{t,11}dW_{t,1}^* + \Sigma_{t,12}dW_{t,2}^* + \Sigma_{t,13}dW_{t,3}^*$ for some functions $\Sigma_{t,11}, \Sigma_{t,12}, \Sigma_{t,13}$ of the state variables that satisfy

$$\Sigma_{t,11}^2 + \Sigma_{t,12}^2 + \Sigma_{t,13}^2 = V_t \tag{2}$$

by definition of volatility. Ito’s lemma implies the log spot price as follows:

$$d(\log S_t) = \left(\delta_t - \frac{1}{2} V_t \right) dt + \Sigma_{t,11}dW_{t,1}^* + \Sigma_{t,12}dW_{t,2}^* + \Sigma_{t,13}dW_{t,3}^*. \tag{3}$$

The drift $\delta_t - \frac{1}{2}V_t$ and instantaneous variance $\Sigma_{t,11}^2 + \Sigma_{t,12}^2 + \Sigma_{t,13}^2$ are both affine (linear) functions of δ_t and V_t . By assumption, $\log S_t, \delta_t,$ and V_t span the risks inherent in the prices of contingent claims on the commodity and these three state variables form a diffusion process with respect to the risk-neutral measure. Furthermore, it is assumed that both δ_t and V_t also have affine drifts, and the instantaneous covariance between any pair of the state variables $\log S_t, \delta_t,$ and V_t is an affine function of V_t only, so that the risk-neutral process for the state vector

$$X_t = \begin{bmatrix} \log S_t \\ \delta_t \\ V_t \end{bmatrix} \tag{4}$$

may be written

$$dX_t = [A^* + B^*X_t]dt + \Sigma_t dW_t^* \quad (5)$$

where

$$\Sigma_t \Sigma_t' = \Omega_0 + \Omega_1 V_t \quad (6)$$

for constant matrices Ω_0 and Ω_1 whose (1, 1) components are 0 and 1, respectively; the first component of A^* is 0, and the first row of B^* is (0, 1, $-1/2$).

Following the terminology introduced in Duffie and Kan (1996) for term structure models, models of the spot price of the form (5) in which $\log S_t$ and $\Sigma_t \Sigma_t'$ are affine functions of the state variables are referred to as affine models of the spot price. Three factor affine models in which the instantaneous variance is spanned by a single state variable as in (6) are referred to as $\mathbb{A}_1(3)$ models.

Maximality, Identifiability, and Admissibility

In this section, it is shown that the model defined by (4), (5), and (6) above is the most general $\mathbb{A}_1(3)$ model for the spot price whose parameters are all well defined, or identified. If the parameters are not well defined, then there will be no way to identify the parameters using prices of contingent claims on the spot. It is also shown that the model is maximal, in the sense that it is not nested in a more general $\mathbb{A}_1(3)$ model. However, some restrictions need to be imposed on the parameters to ensure admissibility so that the stochastic differential equation (SDE) (5) has a unique solution for any given initial value X_0 .¹

Suppose the spot price process S has the most general $\mathbb{A}_1(3)$ dynamics, so that there are parameters $\Theta = (a, b, A^*, B^*, \Omega_0, \Omega_1, c)$ and a process X in \mathbb{R}^3 for which the log of the spot price is an affine function of the state vector

$$\log S = a + b'X, \quad (7)$$

and X has affine drift and instantaneous variance

$$dX = (A^* + B^*X) dt + \Sigma dW^*, \quad (8)$$

where

$$\Sigma \Sigma' = \Omega_0 + \Omega_1 (c'X). \quad (9)$$

There are 30 free parameters: 1 in a , 3 in b , 3 in A^* , 9 in B^* , 6 in Ω_0 , 6 in Ω_1 , and because c is defined only up to a scalar multiple, 2 in c .

¹Other studies have investigated these issues in the context of affine interest rate models. These include the studies by Dai and Singleton (2000) and Collin-Dufresne, Goldstein, and Jones (2008).

These parameters are not well defined, i.e., they cannot all be identified using prices of contingent claims on S , because there is ambiguity in the choice of state variable X . Indeed, any invertible affine transformation of X , $\tilde{X} = \psi + \phi X$, will also have affine dynamics (7)–(9). Conversely, any transformation of X that has affine dynamics is necessarily an affine transformation of \mathbb{R}^3 . The group G of symmetries of the set of state variables with $\mathbb{A}_1(3)$ dynamics is therefore 12-dimensional and consists of pairs (ψ, ϕ) where ψ is an arbitrary vector in \mathbb{R}^3 and ϕ is an arbitrary invertible 3×3 matrix. Intuitively, this represents 12 degrees of ambiguity in the choice of state variable X , or equivalently, 12 degrees of ambiguity in the choice of parameter set Θ .

However, it may always be assumed that the first component of the state vector is the log of the spot price by replacing $a + b'X$ with X_1 . After making this substitution, $\Omega_0(1, 1) + \Omega_1(1, 1)c'X$ may be replaced with X_3 so that the instantaneous variance is a function of the third component of the state vector only and the third component is in fact the volatility of the first component, i.e., $\Sigma\Sigma' = \Omega_0 + \Omega_1 X_3$ where the $(1, 1)$ components of Ω_0 and Ω_1 are 0 and 1, respectively.

The group G_1 of symmetries of the state variables for which $X_1 = \log S$ and $X_3 = V$, the volatility of S , is the four-dimensional subgroup of G consisting of invertible affine transformations $\psi + \phi X$ of \mathbb{R}^3 for which the first and third components of ψ are zero and the first and third rows of ϕ are $(1, 0, 0)$ and $(0, 0, 1)$, respectively. By replacing $A_1^* + B_{11}^* X_1 + B_{12}^* X_2 + (B_{13}^* + \frac{1}{2}) X_3$ with X_2 , it may further be assumed that the risk-neutral drift of $\log S$ is $X_2 - \frac{1}{2} X_3$, i.e., it may be assumed that $X_2 = \delta$, the (instantaneous) cost of carry. This choice of $(\psi, \phi) \in G_1$ is uniquely determined; there is no more ambiguity in the choice of state variable.

Therefore, given the most general affine model satisfying Equations (7)–(9) there exists a unique invertible affine transformation of the state vector for which the model is of the form specified by (4), (5), and (6), where the $(1, 1)$ entries of the matrices Ω_0 and Ω_1 are 0 and 1, respectively. Because this choice of transformation is unique, the choice of state vector for which the drift and dispersion have these forms is unique and the group of symmetries is trivial. These model parameters are therefore well identified.

One fundamental issue for models whose dynamics are given by non-Gaussian SDEs is admissibility, i.e., whether there exists a unique solution to the SDE. Existence requires that the state vector X_t remain within a region in which the instantaneous variance $\Sigma\Sigma' = \Omega_0 + \Omega_1 V_t$ is positive semidefinite. Intuitively, the instantaneous variance must not allow X_t to move stochastically out of the positive semidefinite region, and the drift must not pull X_t out of the region as well. As Duffie and Kan (1996) point out, this imposes nontrivial restrictions on the parameters.

In order for V_t to be the volatility of S_t it must have a non-negative lower bound. When V_t approaches this lower bound, its instantaneous variance and its instantaneous covariance with the other two state variables must vanish otherwise V_t would drop below its lower bound with positive probability. Let ν be the negative of the infimum of V_t ; when $V_t = -\nu$, its instantaneous variance and covariances will vanish precisely when the third row (and column) of Ω_0 is equal to ν times the third row (and column) of Ω_1

$$\Omega_0 = \begin{bmatrix} 0 & s_{12} & \nu\sigma_{13} \\ s_{12} & s_{22} & \nu\sigma_{23} \\ \nu\sigma_{13} & \nu\sigma_{23} & \nu\sigma_{33} \end{bmatrix}, \Omega_1 = \begin{bmatrix} 1 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{bmatrix}. \quad (10)$$

Furthermore, the drift of V_t must be non-negative when V_t approaches its infimum else V_t could be driven below its lower bound. In particular, no other state variable can appear in the drift of V_t

$$\kappa_{31} = \kappa_{32} = 0, \nu \leq 0, \mu_3 - \kappa_{33}\nu \geq 0. \quad (11)$$

These parameter restrictions guarantee that the volatility state variable V_t has a well-defined non-negative infimum. Further restrictions in the form of inequality constraints on the parameters are needed to guarantee that the instantaneous variance $\Omega_0 + \Omega_1 V_t$ is positive semidefinite. Specifically, if it is positive semidefinite when $V_t = -\nu$ and if its smallest eigenvalue is nondecreasing in V_t for $V_t \geq -\nu$ then the instantaneous variance will remain positive semidefinite. Necessary and sufficient conditions on the parameters are that Ω_1 and $\Omega_0 - \Omega_1\nu$ are positive semidefinite.

Risk Premia

To empirically implement the model, or to forecast the state variables, a specification of the risk premium is needed. The risk premium is assumed to be affine in the state variables, and not merely constant as much of the previous empirical research assumes. Setting the risk premium to be an affine function of the state variables allows the risk premium to be time-varying while still maintaining the tractability of the affine framework. This in turn allows the state vector to have different strength of mean-reversion under the true physical and risk-neutral measures, and thus allows for different distortions of the term structures of expected future spot price under the two measures. This has important consequences for risk management, as demonstrated in Casassus and Collin-Dufresne (2005).

The state vector X_t is a solution to the risk-neutral SDE $dX = \mu^* dt + \Sigma dW^*$. If $\Lambda(X_t)$ is a process that satisfies

$$E^* \left[\exp \left(\int_0^T \Lambda'(X_t) dW_t^* - \frac{1}{2} \int_0^T \Lambda'(X_t) \Lambda(X_t) dt \right) \right] = 1 \tag{12}$$

in the risk-neutral measure \mathcal{Q} , then an equivalent measure P , referred to as the physical measure, may be defined by

$$P = \exp \left(\int_0^T \Lambda'(X_t) dW_t^* - \frac{1}{2} \int_0^T \Lambda'(X_t) \Lambda(X_t) dt \right) \mathcal{Q}.$$

By Girsanov's Theorem, the process $W_t = W_t^* - \int_0^t \Lambda(X_s) ds$ is a standard Brownian motion under the physical measure P and therefore X_t is also a solution, if a solution exists, to the SDE $dX = \mu dt + \Sigma dW$ where the drift of X_t in the physical measure is given by

$$\mu = \mu^* + \Sigma \Lambda.$$

Furthermore, the two measures are equivalent, which is necessary and sufficient for the model to admit no arbitrage.

In order for the SDE in the physical measure to have a solution, the parameters in μ and Σ must ensure that X_t remain within the region in which the instantaneous variance matrix is positive semidefinite. Theorem 7.19 in Lipster and Sharyaev (2004) shows that if in fact under either measure X_t remains in the positive-definite region, i.e., if the boundary of the positive semidefinite region is not accessible, then the market price of risk defined by

$$\Lambda = \Sigma^{-1}(\mu - \mu^*) \tag{13}$$

satisfies (12). Hence the two measures are equivalent and the model admits no arbitrage.

The risk premium is chosen to be an affine function of X_t so that μ is also affine: $\mu = A + BX$. The drift restrictions (11) are imposed on A and B to ensure the instantaneous variance matrix is positive semidefinite. Note that when the volatility state variable V_t equals its infimum, the instantaneous variance matrix $\Omega_0 + \Omega_1 V_t$ is positive semidefinite but not positive definite, i.e., its smallest eigenvalue is equal to zero. If the Ω_1 matrix is positive definite then the smallest eigenvalue is strictly increasing in V_t ; if in addition the volatility never attains its infimum then the smallest eigenvalue will be strictly positive. The Feller condition guarantees that V_t never attains its infimum; if it holds under both measures and Ω_1 is positive definite, then the two measures are equivalent.

The complete model is therefore

$$X_t = \begin{bmatrix} \log S_t \\ \delta_t \\ V_t \end{bmatrix} \tag{14}$$

$$dX_t = [A + BX_t]dt + \Sigma dW_t \tag{15}$$

$$A = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, B = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ 0 & 0 & b_{33} \end{bmatrix} \tag{16}$$

$$dX_t = [A^* + B^*X_t]dt + \Sigma dW_t^* \tag{17}$$

$$A^* = \begin{pmatrix} 0 \\ \mu_2 \\ \mu_3 \end{pmatrix}, B^* = \begin{bmatrix} 0 & 1 & -\frac{1}{2} \\ \kappa_{21} & \kappa_{22} & \kappa_{23} \\ 0 & 0 & \kappa_{33} \end{bmatrix} \tag{18}$$

$$\Sigma \Sigma' = \Omega_0 + \Omega_1 V_t \tag{19}$$

$$\Omega_0 = \begin{bmatrix} 0 & s_{12} & \nu \sigma_{13} \\ s_{12} & s_{22} & \nu \sigma_{23} \\ \nu \sigma_{13} & \nu \sigma_{23} & \nu \sigma_{33} \end{bmatrix}, \Omega_1 = \begin{bmatrix} 1 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{bmatrix}, \tag{20}$$

where Ω_1 is positive definite, $\Omega_0 - \Omega_1 \nu$ is positive semidefinite, $\nu \leq 0$, and

$$\mu_3 - \kappa_{33} \nu \geq \frac{1}{2} \sigma_{33}, a_3 - b_{33} \nu \geq \frac{1}{2} \sigma_{33}. \tag{21}$$

With this specification, the complete model has 24 parameters: 14 risk-neutral and 10 risk-premium parameters.

Nested Models

The maximal stochastic volatility nests many of the previous models in the literature, including all Gaussian affine one- and two-factor models. In the Appendix it is shown that every Gaussian affine two-factor model, i.e., every $\mathbb{A}_0(2)$ model, is a special case of the maximal stochastic volatility model in which the volatility is constant and may be written

$$dY_t = \left[\begin{pmatrix} 0 \\ \mu \end{pmatrix} + \begin{bmatrix} 0 & 1 \\ \kappa_1 & \kappa_2 \end{bmatrix} Y_t \right] dt + \Sigma dW^*, \tag{22}$$

$$dY_t = [A + BY_t]dt + \Sigma dW \tag{23}$$

where the first component of Y is the log of the spot price. For example, the Gibson and Schwartz (1990) model (and the Schwartz and Smith (2000) model, as well as the Gabillon (1995) model, after a change of variable) is the special case $\kappa_1 = 0$ and V is constant. The restriction $\kappa_1 = 0$ is equivalent to the assumption that the log spot price has a persistent component with respect to the risk-neutral measure. The model of Korn (2005) relaxes this restriction, and he shows this has important implications for the pricing and hedging of long-term derivatives contracts. Furthermore, all of these models assume the risk premium is constant.

Every Gaussian affine one-factor model, including those of Brennan and Schwartz (1985), Ross (1997), and Schwartz (1997), can be obtained as a special case of a Gaussian affine two-factor model where the convenience yield is an affine function of the log spot price, and thus Gaussian affine one-factor models are also nested within the maximal stochastic volatility model.

The maximal stochastic volatility model also nests all $\mathbb{A}_1(2)$ models of asset prices, including the Heston (1993) model and the model of Nielsen and Schwartz (2004). In the Appendix, it is shown that every $\mathbb{A}_1(2)$ model is a special case of the maximal stochastic volatility model in which the cost of carry is a linear function of the log spot price and the volatility and may be written

$$d \begin{pmatrix} \log S \\ V \end{pmatrix} = \left[\begin{pmatrix} \mu_1 \\ \mu_3 \end{pmatrix} + \begin{bmatrix} \kappa_{11} & \kappa_{13} \\ 0 & \kappa_{33} \end{bmatrix} \begin{pmatrix} \log S \\ V \end{pmatrix} \right] dt + \Sigma dW^*, \tag{24}$$

where

$$\Sigma \Sigma' = \begin{bmatrix} 0 & \nu \sigma_{13} \\ \nu \sigma_{13} & \nu \sigma_{33} \end{bmatrix} + \begin{bmatrix} 1 & \sigma_{13} \\ \sigma_{13} & \sigma_{33} \end{bmatrix} V.$$

The Heston (1993) model

$$d \begin{pmatrix} \log S \\ V \end{pmatrix} = \left[\begin{pmatrix} r - y \\ u \end{pmatrix} + \begin{bmatrix} 0 & -\frac{1}{2} \\ 0 & -\kappa \end{bmatrix} \begin{pmatrix} \log S \\ V \end{pmatrix} \right] dt + \sqrt{V} \begin{bmatrix} 1 & 0 \\ 0 & \sigma \end{bmatrix} \begin{pmatrix} dz_1^* \\ dz_2^* \end{pmatrix}, \tag{25}$$

$dz_1^* dz_2^* = \rho dt$ and the dividend yield y and interest rate r are constant, is the particular $\mathbb{A}_1(2)$ model corresponding to $\nu = 0$, $\kappa_{11} = 0$, and $\kappa_{13} = -1/2$. The model of Nielsen and Schwartz (2004) is

$$dS = (r - \delta)S dt + (b_1\delta + b_2)^{1/2} S dz_1^*$$

$$d\delta = (a - \kappa\delta)dt + \sigma(b_1\delta + b_2)^{1/2} dz_2^*$$

where $dz_1^* dz_2^* = \rho dt$. After making the change of variable $V = b_1\delta + b_2$ and rewriting the model in terms of $\log S$ and V , the model is the particular $\mathbb{A}_1(2)$ model corresponding to $\nu = 0$ and $\kappa_{11} = 0$ and therefore slightly generalizes Heston’s model. Both of these models restrict the risk premium to be a multiple of the volatility state variable; as discussed in “Risk Premia” section, this is an unnecessary restriction.

Richter and Sorensen (2002) develop an $\mathbb{A}_1(3)$ model in which the instantaneous variance matrix and the risk premium are both multiples of the volatility state variable. Their model allows for seasonality in the spot price by allowing the drift to be a periodic function of time. The time-homogeneous version of their model, that is, their model when there is no seasonality, is the special case of the maximal stochastic volatility model

$$\Omega_0 = 0, A = A^*, B = B^* + \begin{bmatrix} 0 & 0 & \lambda_1 \\ 0 & 0 & \lambda_2 \\ 0 & 0 & \lambda_3 \end{bmatrix},$$

where λ is constant.²

Futures and Options Prices

It is well known (for example, Duffie, 1996) that the time t futures price for delivery at time $T = t + \tau$ is the risk-neutral expectation of the future spot price

$$F_{t,\tau} = E_t^*[S_{t+\tau}].$$

By the law of iterated expectations the futures price for a fixed delivery time is a martingale in the risk-neutral measure, i.e., the risk-neutral drift of $F_{t,\tau}$ is equal to zero. This determines $F_{t,\tau}$ as the solution to the following partial differential equation:

$$\frac{\partial F}{\partial t} + \frac{\partial F}{\partial X}(A^* + B^*X) + \frac{1}{2}\text{trace}\left(\frac{\partial^2 F}{\partial X^2}(\Omega_0 + \Omega_1 V)\right) = 0 \tag{26}$$

with initial condition $F_{t,0} = S_t$.

²Although seasonality is an important feature of many commodities, especially agricultural commodities, and it would be straightforward to modify the maximal stochastic model to allow for this feature, this study focuses on crude oil that fluctuates very little with the seasons.

Following Duffie and Kan (1996) for affine term structure models, the solution to this differential equation is of the form

$$\log F_{t,\tau} = \alpha(\tau) + \beta(\tau)X_t; \tag{27}$$

substituting into (26) gives

$$\begin{aligned} \dot{\alpha} + \dot{\beta}X_t &= \beta A^* + \beta B^* X_t + \frac{1}{2} \text{trace}(\beta' \beta (\Omega_0 + \Omega_1 V_t)) \\ &= \beta A^* + \beta B^* X_t + \frac{1}{2} (\beta \Omega_0 \beta' + \beta \Omega_1 \beta' V_t). \end{aligned}$$

This equation must hold for all X ; comparing coefficients, we find

$$\dot{\alpha} = \beta A^* + \frac{1}{2} \beta \Omega_0 \beta' \tag{28}$$

$$\beta_1 = \beta B_{(1)}^* \tag{29}$$

$$\beta_2 = \beta B_{(2)}^* \tag{30}$$

$$\beta_3 = \beta B_{(3)}^* + \frac{1}{2} \beta \Omega_1 \beta' \tag{31}$$

where $B_{(i)}^*$ denotes the i th column of B^* . By the convergence property the initial values of α and β are

$$\alpha(0) = 0 \tag{32}$$

$$\beta(0) = (1, 0, 0). \tag{33}$$

The ordinary differential equations (28)–(31) together with the initial conditions (32) and (33) uniquely determine α and β and therefore uniquely determine the futures price. This system of equations is of the Riccati type and has no analytic solution in general. However, a numerical solution is always available once values for the parameters are chosen.

Options prices may be computed with the Fourier inversion approach introduced by Heston (1993) and later extended in Duffie, Pan, and Singleton (2000). Consider a European call option on a futures contract with delivery time T' and let $\tau = T' - t$ denote the time to maturity of the futures contract. The risk-neutral process followed by the futures price and the volatility state variable is

$$dF_t = F_t \beta(\tau) \Sigma_t dW_t^* \tag{34}$$

$$dV_t = (\mu_3 + \kappa_{33} V_t) dt + \vec{e}'_3 \Sigma_t dW_t^*, \tag{35}$$

where \vec{e}_3 is the vector $(0, 0, 1)$. For constant risk-free rate r , strike price K , expiry T , and current futures price F , the price of the option is given by

$$C = e^{-rT}(FP_1 - KP_2)$$

where, for $j = 0, 1$,

$$P_j = \frac{1}{2} - \frac{1}{\pi} \int_0^\infty \frac{1}{z} \text{Imag}[\exp(a_j(T) + b_j(T)V - iz \ln(F/K))] dz;$$

the functions $a_j(t)$ and $b_j(t)$ satisfy the Riccati ordinary differential equations

$$\begin{aligned} \dot{a}_j(t) &= \frac{1}{2}(j - iz)(j - iz - 1)\beta(\tau)\Omega_0\beta(\tau)' \\ &\quad + ((j - iz)\beta(\tau)\Omega_1\vec{e}_3\nu + \mu_3)b(t) + \frac{1}{2}\sigma_{33}\nu b^2(t), \\ \dot{b}_j(t) &= \frac{1}{2}(j - iz)(j - iz - 1)\beta(\tau)\Omega_0\beta(\tau)' \\ &\quad + ((j - iz)\beta(\tau)\Omega_1\vec{e}_3 + \kappa_{33})b(t) + \frac{1}{2}\sigma_{33}b^2(t), \end{aligned}$$

with initial conditions

$$a_j(0) = b_j(0) = 0.$$

EMPIRICAL METHODOLOGY

The model may be estimated using methodology analogous to that used in the term structure literature to estimate models of the short rate using bond prices. Given a time series of futures prices for contracts with various maturities, roughly speaking the risk-neutral parameters may be estimated from the cross-sectional information and the physical parameters may be estimated from the time-series information.

Data

The data consist of daily futures prices on Crude Oil as reported on the NYMEX from January 1990 through January 2008. There are six futures contracts available at each day in the sample whose maturities range from about 1 to 18 months. The last five years of data are saved for out of sample analyses, so the total number of daily-observed futures prices used for in sample estimation of the model is 19,548.

One complicating issue is that the actual time to maturity is not precisely determined. A futures contract on oil is for delivery at any time during the delivery month. Following common practice and noting that historically oil has almost always been in backwardation so it has been more beneficial to the short side to wait until the end of the delivery period to deliver, it will be assumed that maturity occurs at the end of the delivery month. Furthermore, the last day to trade a particular contract is the 3rd business day prior to the 25th of the month prior to the delivery month, so there are no futures prices available for contracts that mature in less than about five weeks. In particular, the true spot price, which is the futures price for a contract that matures immediately, is not observable.

MCMC Procedure

Let Θ denote the vector whose components are the parameters of the complete model, and let Y_t denote the 6-vector of log-futures prices at each time t (i.e., the n th component of Y_t is the log of the futures price of the n th futures contract at time t). Using the Euler discretization, the model may be rewritten in state space form

$$X_{t+\Delta t} = A\Delta t + (I + B\Delta t)X_t + \Sigma_t \sqrt{\Delta t} \varepsilon_{t+\Delta t} \quad (36)$$

$$Y_t = \alpha + \beta X_t + \varepsilon_t \quad (37)$$

where the time increment Δt is equal to one day, α is a 6-vector and β is a 6×3 matrix obtained from solving (28)–(33), and ε_t is a 6-vector of noise terms that represent measurement error in futures prices. Adding noise to (27) allows the model greater flexibility in fitting the data and is standard practice in both the commodity pricing and term structure literatures; these measurement errors are assumed to be serially and cross-sectionally uncorrelated and normally distributed.

The MCMC algorithm begins with an initial guess for the parameter vector Θ , $\Theta^{(0)}$. A state vector $X^{(1)}$ is then drawn from the conditional distribution $p(X|\Theta^{(0)}, Y)$, and in turn a new parameter vector $\Theta^{(1)}$ is drawn from the conditional distribution $p(\Theta|X^{(0)}, Y)$, and so on. This generates a Markov chain $\{\Theta^{(k)}, X^{(k)}\}$ whose limiting distribution is the posterior distribution $p(\Theta, X|Y)$ of interest; once the chain $\{\Theta^{(k)}, X^{(k)}\}$ has been run out far enough, the Monte Carlo method can then be used on these samples for numerical integration for parameter estimation. The key to MCMC is that the conditional distributions $p(\Theta, X|Y)$ and $p(X|\Theta, Y)$ are easier to characterize than the posterior distribution. Samples may be drawn from these distributions by using the Gibbs sampler and the Metropolis–Hastings algorithm.

There are a total of 30 parameters (14 risk-neutral parameters (nonzero entries of A^* , B^* , Ω_0 , and Ω_1), 10 risk-premium parameters (nonzero entries of A and B), and the six parameters in the variance matrix R of the pricing errors) and the time series of the state variable $\{X_t\}_{t=0}^T$ that needs to be estimated. The restrictions on these parameters are the Feller conditions $a_3 - b_{33}\nu \geq \sigma_{33}/2$ and $\mu_3 - \kappa_{33}\nu \geq \sigma_{33}/2$, and that $\Omega_0 - \Omega_1\nu$ is positive semi-definite and Ω_1 is positive definite.

Given an initial guess for the parameter vector $\Theta^{(0)}$, first the volatility state variable $V_t^{(1)}$ is drawn for $t = 0, \dots, T$. By the law of total probability, the conditional density for V_t for $t = 1, \dots, T - 1$ given all other state variables (including the volatility state variable at all other times), the futures prices, and the parameters is

$$\pi(V_t) \propto p(X_t|X_{t-1}, X_{t+1}, Y_t, \Theta) \propto p(Y_t|X_t, \Theta) p(X_{t+1}|X_t, \Theta) p(X_t|X_{t-1}, \theta). \quad (38)$$

The first kernel on the right-hand side is the conditional likelihood function of the futures price data on date t , viewed as a function of the random variable V_t , with the futures data, all other parameters, the values of the other state variables, and the volatility state variable at all other times treated as fixed (Equation (37)). The second and third kernels of $\pi(V_t)$ are Gaussian densities that describe the evolution of the state variables (Equation (36)).

The conditional density $\pi(V_t)$ as a function of V_t is not a recognizable distribution. The volatility state variable is drawn using independence Metropolis–Hastings with the Gaussian proposal density $p(Y_t|X_t, \Theta)p(X_t|X_{t-1}, \Theta)$. If a negative candidate for V_t is drawn, it is simply discarded and a new candidate is drawn.

The initial volatility V_0 has conditional density

$$\pi(V_0) \propto p(V_0)p(X_1|V_0, \Theta),$$

where $p(V_0)$ is the prior density, assumed to be Inverse Gamma. The initial volatility is drawn using random-walk Metropolis, with a fat-tailed symmetric proposal density. The terminal volatility V_T has Gaussian conditional density

$$\pi(V_T) \propto p(Y_T|X_T, \Theta)p(X_T|X_{T-1}, \Theta),$$

and thus may be drawn directly from a Gaussian distribution.

The other two state variables may be drawn using the Kalman filtering recursions. Their conditional density given the parameters and the volatility state variable is Gaussian with time-varying but deterministic variance. Therefore, the time series of $\log S$ and δ , including their initial values, may be drawn in a single block.

The parameter vector Θ is decomposed into four blocks: (A, B) , A^* , $(B^*, \Omega_0, \Omega_1)$ and R . The first block (A, B) has conditional density

$$\pi(A, B) \propto p(A, B)p(X|A, B, \Omega_1).$$

The conditional likelihood $p(X|A, B, \Omega_1)$ is a Gaussian function of (A, B) ; with a Gaussian prior density $p(A, B)$, draws of (A, B) are therefore obtained by directly sampling from a Gaussian distribution.

Similarly, because the risk-neutral drift term A^* appears in the futures price formula only and appears linearly (Equation (28)), the conditional likelihood $p(Y|A^*, B^*, \Omega_1, R, X)$ is a Gaussian function of A^* . With a Gaussian prior density $p(A^*)$, the conditional density

$$\pi(A^*) \propto p(A^*)p(Y|A^*, B^*, \Omega_1, R, X)$$

is also Gaussian, and draws of A^* are obtained by directly sampling from a Gaussian distribution.

The conditional densities of B^* , Ω_0 and Ω_1 are not Gaussian, as these parameters appear nonlinearly in the formulas for α and β . The conditional density $\pi(B^*, \Omega_0, \Omega_1)$ for this block of parameters is proportional to

$$\pi(B^*, \Omega_0, \Omega_1) \propto p(B^*)p(\Omega_0)p(\Omega_1)p(Y|B^*, \Omega_0, \Omega_1)p(X|\Omega_0, \Omega_1).$$

Because this is not a recognizable distribution function, random-walk Metropolis is used with a fat-tailed symmetric proposal density.

Finally, the parameters in the diagonal variance matrix R of the pricing error appear in the futures price formula only. The conditional density of each diagonal entry R_{mm} is

$$\pi(R_{mm}) \propto p(R_{mm})p(Y|R_{mm});$$

the likelihood $p(Y|R_{mm})$ is an Inverse Gamma density function of R_{mm} , and if the prior density $p(R_{mm})$ is also Inverse Gamma then each draw of R_{mm} for $n = 1, \dots, 6$ may be obtained by directly sampling from an Inverse Gamma distribution.

RESULTS

A Markov chain was generated of length 200,000 and the first 100,000 draws were discarded to negate the effects of initial conditions. To address concerns that the posterior draws are too highly autocorrelated, only one out of every ten iterations of the chain were saved, leaving 10,000 draws from the posterior distribution.

TABLE I
Descriptive Statistics on the Posterior Draws for the Parameters

<i>Parameter</i>	<i>Mean</i>	<i>Standard Deviation</i>	<i>5%</i>	<i>95%</i>
a_1	0.005	0.854	-1.428	1.399
a_2	0.006	0.881	-1.440	1.448
a_3	1.070	0.127	0.860	1.275
b_{11}	-0.080	0.266	-0.513	0.364
b_{12}	0.265	0.371	-0.345	0.875
b_{13}	2.005	0.800	0.684	3.347
b_{21}	-0.295	0.295	-0.774	0.192
b_{22}	-2.124	0.592	-3.107	-1.160
b_{23}	2.521	0.929	1.005	4.039
b_{33}	-4.605	0.598	-5.582	-3.596
v	-0.000	0.000	-0.001	-0.000
s_{12}	0.003	0.002	0.001	0.006
s_{22}	0.240	0.015	0.214	0.261
σ_{12}	-1.452	0.000	-1.452	-1.452
σ_{13}	0.103	0.000	0.103	0.103
σ_{22}	2.260	0.000	2.260	2.260
σ_{23}	-0.194	0.000	-0.194	-0.194
σ_{33}	0.161	0.000	0.161	0.161
μ_2	-0.541	0.009	-0.554	-0.523
μ_3	0.067	0.010	0.048	0.082
κ_{21}	-0.009	0.000	-0.010	-0.009
κ_{22}	-1.896	0.001	-1.898	-1.894
κ_{23}	2.718	0.002	2.715	2.721
κ_{33}	-0.113	0.001	-0.115	-0.112

Table I lists descriptive statistics (the means, standard deviations, and the 5- and 95-percentiles) on the posterior draws for the parameters using 1990–2003 data. As seen in the table, some parameters are estimated more precisely than others. In particular, the risk-neutral parameters are estimated with much greater precision than the drift parameters in the physical measure. This is because the physical drift parameters are estimated using filtered estimates of the state variables, whereas the risk-neutral parameters are estimated directly from the observed futures prices.

Plots of the estimated state variables over the period 1990–2003 are shown in Figures 1–3. The 10,000 draws are averaged to obtain the estimated time series for each state variable. The estimated spot price and the nearby futures price are shown in Figure 1. By the convergence property, the estimated spot price is very similar to the nearby futures price. The estimated cost of carry is shown in Figure 2; this state variable is negative much of the time indicating that the front end of the futures curve often slopes downward, consistent with previous findings that oil is often strongly backwardated. The (square root of

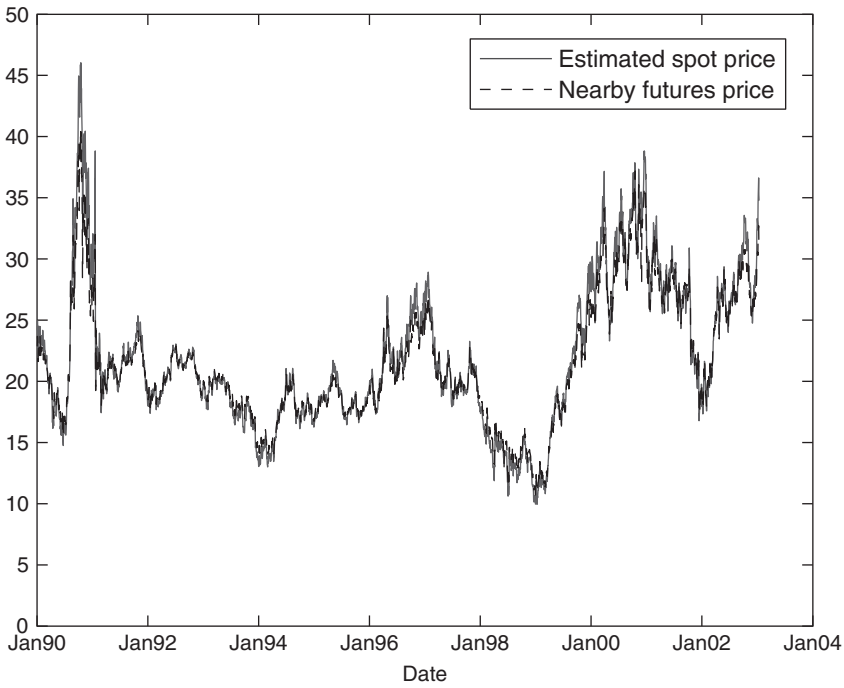


FIGURE 1

Estimated spot price and nearby futures price. The 10,000 draws of the spot price state variable are averaged to obtain the estimated spot price time series.

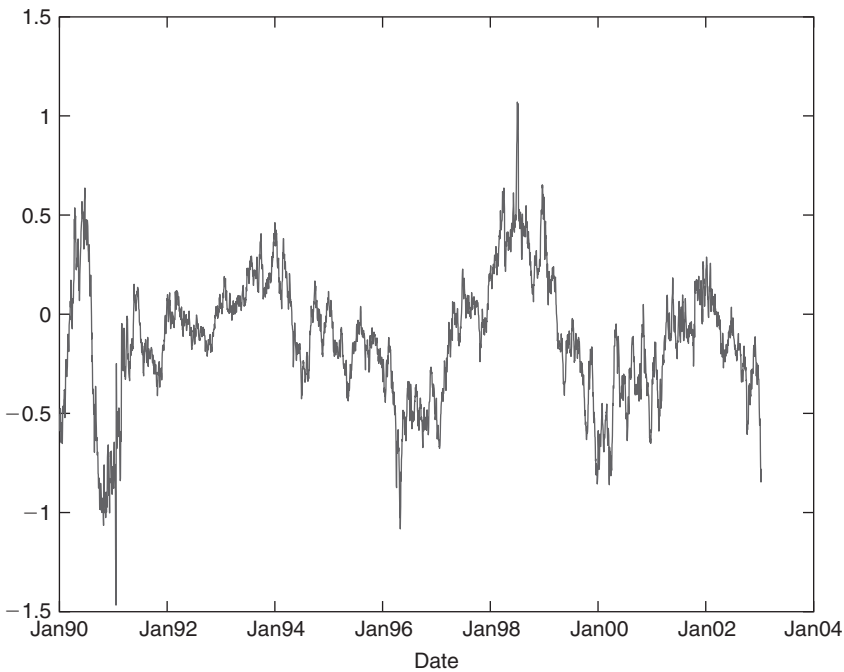


FIGURE 2

Estimated cost of carry. The 10,000 draws of the cost of carry state variable are averaged to obtain the estimated cost of carry time series.

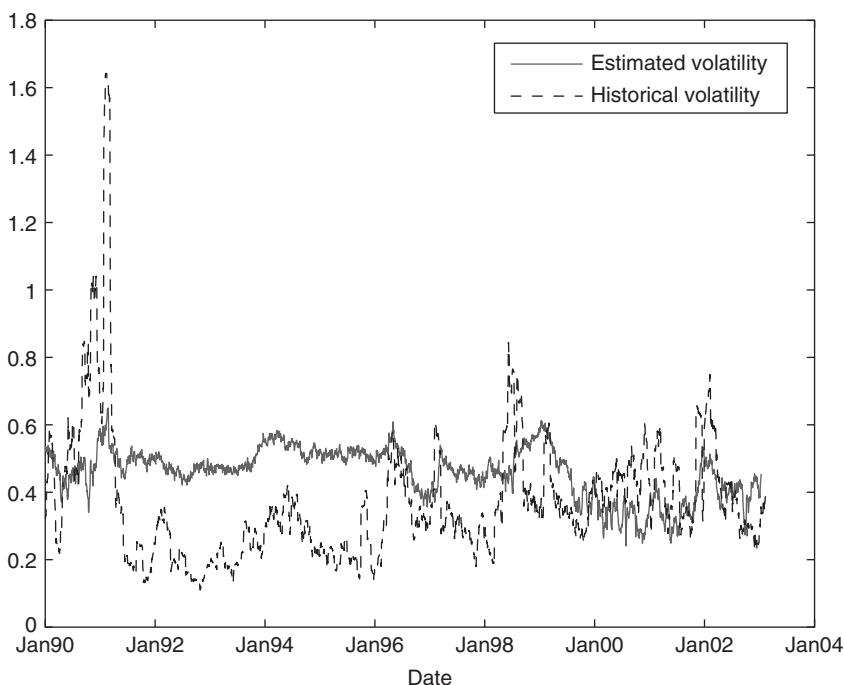


FIGURE 3

Estimated volatility and historical volatility. The 10,000 draws of the volatility state variable are averaged to obtain the estimated volatility time series. The historical volatility is computed using daily nearby futures prices with 30 day rolling windows.

the) estimated volatility state variable is shown in Figure 3. Also shown is the historical volatility, computed using the price of the nearby futures contract with rolling 30-day windows. The estimated volatility does not appear to be highly related to historical volatility. During the Gulf War, for example, the historical volatility peaked at over 160%, whereas the estimated volatility variable peaked at only about 65%. For the four years following the Gulf War, the estimated volatility variable is at least twice the historical volatility; at other times the estimated volatility is consistently below historical volatility. These observations suggest that volatility cannot be accurately inferred from futures prices, in agreement with the empirical findings of Trolle and Schwartz (2008). This is discussed further in fifth section.

For each draw $\{\theta^{(i)}, X^{(i)}\}$ of the parameters and state variables, the daily fitted log-futures prices and the corresponding RMSE of the observed daily log-futures prices were computed. For each draw $\{\theta^{(i)}, X^{(i)}\}$ the overall RMSE is the time-series average of the daily RMSEs. In this way, a distribution of overall RMSEs was constructed; the mean and standard deviation are reported in Table II. The results indicate that the model fits the futures data very well.

TABLE II

The Root Mean Squared Errors (RMSEs) for the Maximal Stochastic Volatility Model, $\mathbb{A}_0(2)$ Model, $\mathbb{A}_0(3)$ Model, and the Restricted Stochastic Volatility Model.

<i>Model</i>	<i>1990–2003 Data</i>		<i>2003–2008 Data</i>	
	<i>1990–2003 (IS)</i>	<i>2003–2008 (OOS)</i>	<i>1990–2003 (OOS)</i>	<i>2003–2008 (IS)</i>
Maximal SV	43.53 (2.05)	41.03 (1.82)	39.50 (1.45)	33.87 (1.38)
$\mathbb{A}_0(2)$	81.85 (1.16)	85.69 (1.25)	82.19 (0.92)	65.21 (0.87)
$\mathbb{A}_0(3)$	43.21 (1.96)	41.10 (1.84)	45.07 (1.37)	41.84 (1.38)
Restricted SV	50.01 (1.23)	45.33 (1.25)	48.86 (0.98)	46.40 (0.89)

Note. For each model, a distribution of RMSEs was obtained by computing the RMSE of the fitted log-futures prices with the actual log-futures prices, and the mean and standard deviation (in parentheses) are reported in basis points.

For example, the in sample RMSE is about 0.0044, which roughly corresponds a percentage error of 0.44% for the typical futures price. Note that the RMSE is very tightly centered around its mean as indicated by the relatively small standard deviation.

Comparisons with Other Models

The performance of the model is compared with several nested specifications and with the maximal $\mathbb{A}_0(3)$ model, the most general affine three-factor Gaussian model. As a first point of comparison, the maximal $\mathbb{A}_0(2)$ model was estimated; as discussed in “Nested Models” section, the models of Gibson and Schwartz (1990), Gabillon (1995), Schwartz and Smith (2000), and Korn (2005) are all special cases. Following an MCMC procedure as described in the Appendix, a Markov chain of length 200,000 was again generated, the first 100,000 draws were discarded, and only one out of every ten iterations were saved.

The mean and standard deviation of the distribution of the overall RMSEs using the maximal $\mathbb{A}_0(2)$ model is reported in Table II. For each of the two time periods, the mean RMSE using the maximal stochastic volatility model is about one-half the mean RMSE using the $\mathbb{A}_0(2)$ model. Because the standard deviations are so small for both models and the sample sizes are relatively large, the standard error is of the order 10^{-6} and thus the mean RMSE using the maximal stochastic volatility model is significantly lower, at virtually any level, than the mean RMSE using the $\mathbb{A}_0(2)$ model. This holds both in sample and out of sample. In this sense the maximal stochastic volatility model clearly outperforms the two-factor Gaussian model.

In light of earlier empirical research that suggests three factors provide a statistically significant improvement over two-factor models and the fact that the maximal stochastic volatility model has one more state variable and many more parameters, it is not surprising that the model gives a lower RMSE than does the two-factor $\mathbb{A}_0(2)$ model. The model is also compared with other three-factor models: the maximal $\mathbb{A}_0(3)$ Gaussian model and the (nested) restricted stochastic volatility model corresponding to the time-homogeneous version of the stochastic volatility model of Richter and Sorensen (2002).

It is shown in the Appendix that the maximal $\mathbb{A}_0(3)$ model has risk-neutral dynamics

$$dX = \left[\begin{pmatrix} 0 \\ 0 \\ \mu \end{pmatrix} + \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \kappa_1 & \kappa_2 & \kappa_3 \end{bmatrix} X \right] dt + \Sigma dW^*, \quad (39)$$

where $X_1 = \log S$ and $\Sigma\Sigma'$ is an arbitrary positive-definite matrix. If the risk premium is allowed to be a linear function of the state vector then the model has 10 risk-neutral and 12 risk-premium parameters, only 2 fewer than the number of parameters of the maximal stochastic volatility model.

This model is equivalent to both (a slightly extended version of) the Gaussian three-factor model of Cortazar and Naranjo (2006) and, under certain conditions, the model of Casassus and Collin-Dufresne (2005). The extended version of the three-factor Cortazar and Naranjo (2006) model may be written

$$\log S = (1, 1, 1)' \tilde{X}$$

$$d\tilde{X} = \left[\begin{pmatrix} u \\ 0 \\ 0 \end{pmatrix} + \begin{bmatrix} k_1 & 0 & 0 \\ 0 & k_2 & 0 \\ 0 & 0 & k_3 \end{bmatrix} \tilde{X} \right] dt + \tilde{\Sigma} dW^*,$$

where $\tilde{\Sigma}$ is an arbitrary constant positive-definite matrix, and the risk premium is an arbitrary linear function of the state vector. This model may be rotated to the form of the maximal $\mathbb{A}_0(3)$ model by applying the transformation

$$X = \begin{bmatrix} 0 \\ u \\ 0 \end{bmatrix} + \begin{bmatrix} 1 & 1 & 1 \\ k_1 & k_2 & k_3 \\ k_1^2 & k_2^2 & k_3^2 \end{bmatrix} \tilde{X},$$

and is therefore equivalent to the maximal $\mathbb{A}_0(3)$ model provided this transformation is invertible.

The model of Casassus and Collin-Dufresne (2005) is a Gaussian three-factor model that allows the instantaneous interest rate r to follow a Vasicek

process. In their model the state vector $(\log S, \gamma, r)$ may be written as a linear function of log-futures prices, which in turn may be written as a linear function of the state vector X of the maximal $\mathbb{A}_0(3)$ model. Consequently their model may be rotated to the form of the maximal $\mathbb{A}_0(3)$ model by applying the appropriate invertible linear transformation. If only futures prices and not bond prices are used to estimate their model, i.e., if r is not given its inherent economic interpretation, then their model is equivalent to the maximal $\mathbb{A}_0(3)$ model (and two of their parameters will not be identifiable).³

The maximal $\mathbb{A}_0(3)$ model and the restricted stochastic volatility model (corresponding to the nested time-homogeneous model of Richter and Sorensen, 2002) were estimated using MCMC procedures similar to those previously described, and the corresponding means and standard deviations of the distributions of the overall RMSEs are reported in Table II. As compared with the results using two-factor models, the RMSEs using each of these three-factor models are much more similar to the RMSEs using the maximal stochastic volatility model. In fact the RMSE using the maximal $\mathbb{A}_0(3)$ model is virtually indistinguishable from the RMSE using the maximal stochastic volatility model when estimated with 1990–2003 data. This result holds both in sample and out of sample. However, when estimated with 2003–2008 data, the RMSE using the maximal stochastic volatility model is noticeably lower than the RMSE using the maximal $\mathbb{A}_0(3)$ model, both in sample and out of sample.

Although the Gaussian $\mathbb{A}_0(3)$ model and the maximal stochastic volatility model both fit the futures data with similar accuracy as measured by RMSE, the two models can have significantly different options pricing implications. Gaussian models necessarily assume returns are normally distributed, whereas stochastic volatility models allow for return distributions that are non-symmetric and have fatter tails. This can cause computed option prices using the maximal stochastic volatility model to differ significantly from those using Gaussian models.

Options Pricing Implications

For the maximal stochastic volatility model, futures options prices may be computed using the Fourier inversion approach as described in “Futures and options Prices” section. For Gaussian models, futures prices are lognormally distributed and a standard result (Hull, 2005, for example) may be applied to compute options prices.

³Neither the maximal $\mathbb{A}_0(3)$ model nor the maximal stochastic volatility model necessarily assumes the instantaneous interest rate r is constant, only that its risk is spanned by the state vector X and hence by futures prices. However, when the model-implied options prices are computed in “Options Pricing Implications” section, it will be assumed for simplicity that r is constant.

It follows from Equations (34) and (35) that the instantaneous covariance of the futures price F_t and the volatility V_t is $F_t\beta(\tau)(\Omega_0 + \Omega_1V_t)\vec{e}_3$ for the maximal stochastic volatility model. Because this varies over time the model allows for a rich variety of implied volatility curves. Several studies, including Eydeland and Wolyniec (2003) and Trolle and Schwartz (2008), document that oil futures markets typically exhibit pronounced implied volatility smiles. The stochastic volatility model can generate a variety of implied volatility curves; Gaussian models, on the other hand, necessarily have constant implied volatility.

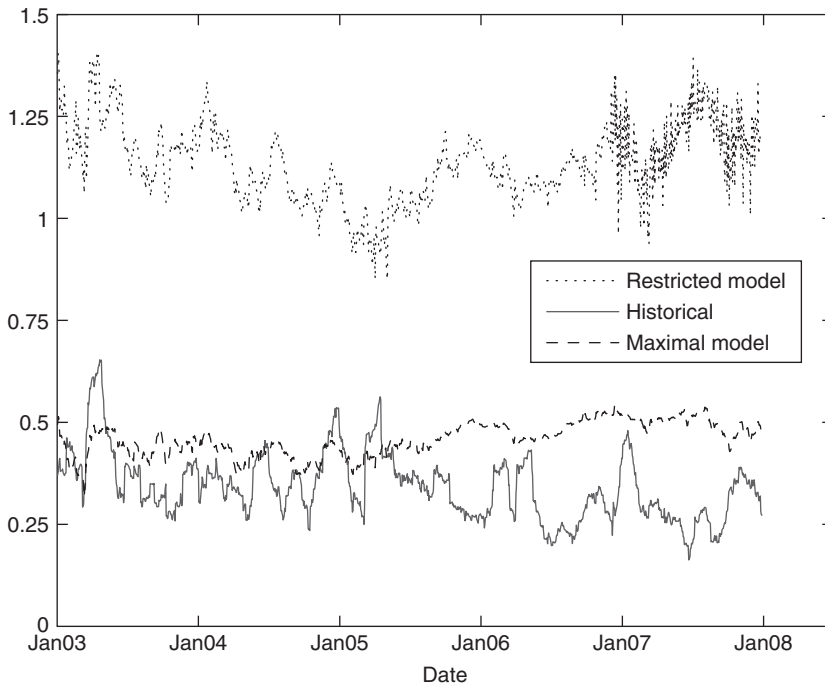
To measure the fit to actual options prices, options price data were obtained from NYMEX. The data consist of daily prices of call options on the nearby futures contract for the two-year period December 15, 2005–December 13, 2007. On each day of the sample, there are call prices corresponding to five strike prices, ranging in moneyness from about 0.8 to about 1.2. Those options with time to maturity less than two days or with price less than \$0.02 were discarded, so that the total number of observed options prices is 1,726.

For each model, the parameters and state variables estimated on 2003–2008 futures data were used to compute model-implied call prices for each day in the options data sample. The corresponding RMSE of the observed daily call prices were computed and the (overall) RMSE reported is the time-series average of the daily RMSE.

The results are not terribly impressive for any of the models. The maximal stochastic volatility model has RMSE 0.3606, only slightly lower than the RMSE 0.3667 using the maximal $\mathbb{A}_0(3)$ model. Interestingly, the best fit is obtained with the maximal $\mathbb{A}_0(2)$ model that has RMSE 0.2149, and the worst fit is obtained with the restricted stochastic volatility model that has RMSE 2.3587.

The most likely explanation for these results is found in Figure 4, which plots the estimated volatility state variables using the restricted stochastic volatility model and the maximal stochastic volatility model on 2003–2008 futures data. Also plotted is historical volatility over that time. Both the maximal and restricted volatility models overestimate realized volatility (as measured by historical volatility) over the two-year period December 15, 2005–December 13, 2007; however, the estimate using the maximal stochastic volatility model clearly more closely resembles historical volatility. The estimated volatility using the maximal $\mathbb{A}_0(2)$ model is 37.7%, much lower than the estimated volatility 54.2% using the maximal $\mathbb{A}_0(3)$ model, and much closer to the historical volatility average of 29.5% over the period December 15, 2005–December 13, 2007.

These are out of sample results, as only futures prices are used to estimate the parameters and state variables. If the models were estimated on both

**FIGURE 4**

Historical volatility and the estimated volatility using the maximal and restricted stochastic volatility models. For each model, the 10,000 draws of the volatility state variable are averaged to obtain the estimated volatility time series. The historical volatility is computed using daily nearby futures prices with 30 day rolling windows.

futures and options data together, the maximal stochastic volatility model would very likely fit the options data much better than the maximal $\mathbb{A}_0(3)$ simply because Gaussian models can generate only constant implied volatility curves. It would also very likely fit the options data much better than the restricted stochastic volatility model because of this model's restriction of the dynamics of the volatility state variable.

UNSPANNED STOCHASTIC VOLATILITY

The model's relatively poor fit to options data when estimated on futures data only suggests that options prices depend on risk not fully spanned by futures prices. In particular, if options prices are sensitive to volatility, then the high RMSE with options data suggests that the estimated volatility \hat{V} using futures prices is not highly related to the true volatility. This conjecture is further supported by the visual evidence presented in Figures 3 and 4 in which it appears that \hat{V} is not highly correlated with historical volatility V_{hist} .

In fact the correlation between \hat{V} and V_{hist} is only 0.0052 over the time period 1990–2003 and is -0.1609 over the time period 2003–2008. Similar results are obtained with GARCH estimates of realized volatility. The GARCH(1,1) model

$$\sigma_t^2 = \omega + a(\Delta \log S_{t-1})^2 + b\sigma_{t-1}^2$$

was estimated using the estimated spot price state variable; the resulting parameter estimates are $\omega = 1.10 \times 10^{-5}$, $a = 0.0963$, and $b = 0.8903$ over the time period 1990–2003 and $\omega = 2.84 \times 10^{-5}$, $a = 0.0528$, and $b = 0.8965$ over the time period 2003–2008. The correlation between \hat{V} and V_{GARCH} is 0.0602 over the time period 1990–2003 and is -0.1190 over the time period 2003–2008.

This evidence strongly supports the conjecture that volatility cannot be accurately inferred from futures prices. This agrees with the findings of Trolle and Schwartz (2008), who refer to this property as USV following the terminology introduced in Collin-Dufresne and Goldstein (2002) for term structure models.

The maximal stochastic volatility model does allow for USV, in contrast with all Gaussian $\mathbb{A}_0(N)$ models and with the restricted stochastic volatility model of Richter and Sorensen (2002). These models necessarily have the property that futures prices depend nontrivially on all the state variables and, consequently, imply that options prices cannot depend on any risk not fully spanned by futures prices. The conditions necessary for USV in the maximal stochastic volatility model are discussed below.

Futures prices are given by $\log F = \alpha + \beta_1 \log S + \beta_2 \delta + \beta_3 V$ where α and β solve the Riccati equations (28)–(33). Futures prices will not depend on the volatility state variable if and only if β_3 is identically zero. From Equation (31), this is equivalent to the condition that the function

$$Q = -\frac{1}{2}\beta_1 + \kappa_{23}\beta_2 + \frac{1}{2}\beta_1^2 + \sigma_{12}\beta_1\beta_2 + \frac{1}{2}\sigma_{22}\beta_2^2$$

is identically zero. Q is an analytic function of time to maturity τ and therefore will vanish everywhere if and only if all of its derivatives vanish at $\tau = 0$. Equations (29) and (30) may be used to compute the derivatives of Q up to any order. It is straightforward to show that the derivatives all vanish at $\tau = 0$ if and only if $\kappa_{23} + \sigma_{12} = \kappa_{21} = \sigma_{22} = 0$.

If $\sigma_{22} = 0$, the Ω_1 matrix will not be positive definite. To be well defined, the instantaneous variance matrix $\Omega_0 + \Omega_1 V_t$ must be positive definite for all $V_t > -\nu$. If Ω_1 and $\Omega_0 - \Omega_1 \nu$ are positive semidefinite and have disjoint null spaces, then $\Omega_0 + \Omega_1 V_t = \Omega_0 - \Omega_1 \nu + \Omega_1(V_t + \nu)$ will be positive definite for all $V_t > -\nu$ as required.

It follows that the model will exhibit USV if and only if

$$\kappa_{21} = \kappa_{23} = \sigma_{12} = \sigma_{22} = \sigma_{23} = 0, s_{22} > 0, \begin{bmatrix} -\nu & s_{12} \\ s_{12} & s_{22} \end{bmatrix} \geq 0, \begin{bmatrix} 1 & \sigma_{13} \\ \sigma_{13} & \sigma_{33} \end{bmatrix} > 0.$$

In particular, the model of Richter and Sorensen (2002), because it has $s_{22} = 0$, cannot exhibit USV.

With these restrictions in place, the model was re-estimated using 1990–2003 futures data. The RMSE is 0.0095, significantly higher than the RMSE using the unrestricted model and more similar to the RMSE using the maximal $\mathbb{A}_0(2)$ model. This result is congruous with the fact that the USV model and the $\mathbb{A}_0(2)$ model both maintain that the cross-sectional dynamics of the futures curve is driven by only two independent movements. The estimated volatility state variable, however, much more closely resembles both the historical volatility and the GARCH estimate of realized volatility and is plotted in Figure 5. The correlation between the USV estimated volatility and the GARCH estimate (historical volatility) is 0.8223 (0.7241, respectively).

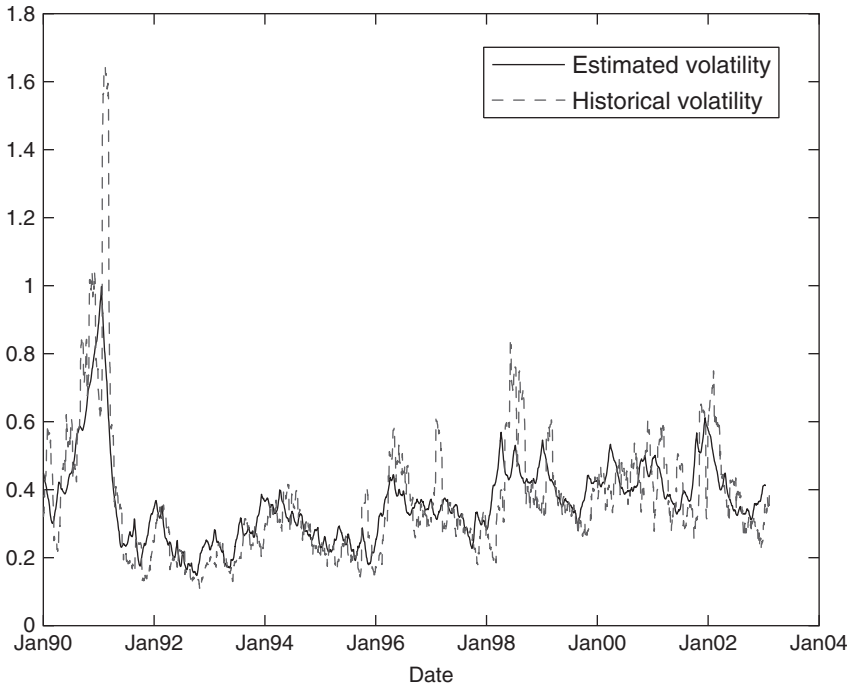


FIGURE 5

Historical volatility and the estimated volatility using the USV version of the maximal stochastic volatility model. The 10,000 draws of the volatility state variable are averaged to obtain the estimated volatility time series. The historical volatility is computed using daily nearby futures prices with 30 day rolling windows.

The USV version of the model is better able to estimate the true volatility. Given the analysis in “Options Pricing Implications” section, this suggests that the USV version would better fit options data; on the other hand, the USV version has fewer parameters and significantly restricts the risk-neutral dynamics of the cost of carry variable. The out of sample analysis of “Options Pricing Implications” section is not possible here because the risk-neutral parameters μ_3 and κ_{33} are not identifiable using futures data. However, even in the unlikely case that the volatility risk-premium is zero, the USV version of the model, when estimated on futures data only, has RMSE 0.1443 on the options data. This is considerably lower than the other models and suggests that the USV version may best explain the joint dynamics of futures and options prices.

CONCLUSION

The recent dramatic spikes in oil and other commodity prices that peaked in Summer 2008 has greatly magnified the importance of managing and hedging risk in these markets. Partial equilibrium models provide an internally consistent framework for pricing and hedging contingent claims on the underlying commodity; however, it is critical that the specification of the process for the underlying commodity is one that closely resembles the observed empirical features and agrees with theoretical insights.

This study has developed and estimated a model that is consistent with many of the observed features of the crude oil market. It nests many of the models previously investigated in the literature, and it is shown to be the maximal, identifiable model in its class. It has quasi-analytical formulas for futures and options prices. It allows for time-varying correlation structures between the state variables, mean-reversion in the short term and an increasing expected long-term price, the Samuelson effect, and for time-varying degrees of backwardation.

Furthermore, the model allows for USV, the situation in which options prices depend on risk not fully spanned by futures prices. This feature is theoretically important because options need not be redundant securities and therefore cannot be hedged by trading futures contracts only, and it has been shown to be empirically important for crude oil in particular.

One question that has not been addressed here is how well the model accounts for the joint dynamics of futures and option prices when estimated on both data sets. Given the significant improvement of three-factor models over two-factor models in fitting futures prices, better results would be expected to be obtained by incorporating both USV and three independent movements of the futures term structure. A natural extension is to relax the assumption that the instantaneous risk-free rate is spanned by the spot price, the convenience yield, and the volatility. This is left for future research.

APPENDIX

The Maximal $\mathbb{A}_0(N)$ Model

Consider an arbitrary $\mathbb{A}_0(N)$ model for the spot price S

$$\log S = a + b' \tilde{X}$$

where \tilde{X} has affine, Gaussian dynamics

$$d\tilde{X} = [\tilde{A}^* + \tilde{B}^* \tilde{X}]dt + \tilde{\Sigma}dW^*$$

for constant instantaneous variance matrix $\tilde{\Sigma}\tilde{\Sigma}'$. Set $X_1 = a + b' \tilde{X}$; for $n = 1, \dots, N - 1$, set \tilde{X}_{n+1} equal to the risk-neutral drift of X_n . As X is an affine transformation of \tilde{X} , it also has affine, Gaussian dynamics $dX = (A^* + B^*X)dt + \Sigma dW^*$, where A^* is the N -vector $(0, \dots, 0, \mu)$ and B^* is the $N \times N$ matrix whose N th row is arbitrary and whose n th row, for $n = 1, \dots, N - 1$, is the unit N -vector whose $(n + 1)$ th component is 1.

Futures prices are again of the form

$$\log F_{t,\tau} = \alpha(\tau) + \beta(\tau)X_t$$

where now α and β solve the ODE system

$$\dot{\alpha} = \beta A^* + \frac{1}{2}\beta \Sigma \Sigma' \beta'$$

$$\dot{\beta} = \beta B^*$$

with initial conditions $\alpha(0) = 0$ and $\beta(0) = (1, 0, \dots, 0)$.

If the risk premium is allowed to be a linear function of the state vector and not merely constant, the $\mathbb{A}_0(N)$ model may be recast in state space form as

$$dX = (A + BX) dt + \Sigma dW$$

$$Y = \alpha + \beta X + \varepsilon$$

where A and B are arbitrary, α and β solve the ODE system described above, and the measurement error ε is assumed to be serially and cross-sectionally uncorrelated and normally distributed.

Therefore the maximal $\mathbb{A}_0(N)$ may be estimated following an MCMC procedure as described in “MCMC procedure” section but modified somewhat. For instance, the full state vector may be drawn in a single block using the Kalman filter.

Proof That the Model Nests $\mathbb{A}_0(2)$ and $\mathbb{A}_1(2)$ Models

The previous section shows that the maximal $\mathbb{A}_0(2)$ model has risk-neutral dynamics

$$dX = \left[\begin{pmatrix} 0 \\ \mu \end{pmatrix} + \begin{bmatrix} 0 & 1 \\ \kappa_1 & \kappa_2 \end{bmatrix} X \right] dt + \Sigma dW^*, \tag{A1}$$

and the risk premium is an arbitrary linear function of the state vector. On the other hand, if the volatility V_t is constant, then the maximal stochastic volatility model reduces to

$$d \begin{pmatrix} \log S \\ \delta \end{pmatrix} = \left[\begin{pmatrix} -\frac{1}{2}V \\ \mu_2 \end{pmatrix} + \begin{bmatrix} 0 & 1 \\ \kappa_{21} & \kappa_{22} \end{bmatrix} \begin{pmatrix} \log S \\ \delta \end{pmatrix} \right] dt + \Sigma dW^*$$

where

$$\Sigma \Sigma' = \begin{bmatrix} V & s_{12} + \sigma_{12}V \\ s_{12} + \sigma_{12}V & s_{22} + \sigma_{22}V \end{bmatrix},$$

and the risk premium is again an arbitrary linear function of the state vector. Because the volatility V is constant, the individual parameters in the instantaneous variance $\Sigma \Sigma'$ are no longer well identified; it may be assumed $\sigma_{12} = \sigma_{22} = 0$, for instance. Finally, replace δ with $\delta - \frac{1}{2}V$ and μ_2 with $\mu_2 + \frac{1}{2}V\kappa_{22}$, so that the drift is of the same form as that of (A1).

Consider now the most general $\mathbb{A}_1(2)$ model: $\log S = a + b'X$, where

$$dX = [A^* + B^*X]dt + \Sigma dW^*$$

for

$$\Sigma \Sigma' = \Omega_0 + \Omega_1(c'X).$$

By the same reasoning as in the $\mathbb{A}_1(3)$ case, X_1 may be replaced with $a + b'X$ and then X_2 may be replaced with the resulting $\Omega_0(1, 1) + \Omega_1(1, 1)c'X$, and therefore it may be assumed $X_1 = \log S$ and X_2 is the volatility of S . With this transformation the risk-neutral dynamics of the state vector become

$$d \begin{pmatrix} \log S \\ V \end{pmatrix} = \left[\begin{pmatrix} \mu_1 \\ \mu_3 \end{pmatrix} + \begin{bmatrix} \kappa_{11} & \kappa_{13} \\ \kappa_{31} & \kappa_{33} \end{bmatrix} \begin{pmatrix} \log S \\ V \end{pmatrix} \right] dt + \Sigma dW^*, \tag{A2}$$

where

$$\Sigma \Sigma' = \begin{bmatrix} 0 & v\sigma_{13} \\ v\sigma_{13} & v\sigma_{33} \end{bmatrix} + \begin{bmatrix} 1 & \sigma_{13} \\ \sigma_{13} & \sigma_{33} \end{bmatrix} V,$$

and admissibility requires $\kappa_{31} = 0$. If the Feller condition is imposed, the transformed Ω_1 is positive definite, and the transformed $\Omega_0 - \Omega_1\nu$ is positive semidefinite, then the risk premium may be assumed to be an arbitrary linear function of the state vector with the caveat that the admissibility condition $b_{31} = 0$ is also imposed.

On the other hand, if the market for contingent claims on the commodity has only two sources of risk spanned by S_t and V_t , and if δ_t is a linear function of $\log S_t$ and V_t , say

$$\delta_t = \mu_1 + \kappa_{11} \log S_t + (\kappa_{13} + \frac{1}{2})V_t,$$

then the maximal stochastic volatility model reduces precisely to (A2). In particular, if $\delta_t = \delta$ is constant then the drift of $\log S_t$ is equal to $\delta - \frac{1}{2}V_t$.

BIBLIOGRAPHY

- Brennan, M., & Schwartz, E. (1985). Evaluating natural resource investments. *Journal of Business*, 58, 133–155.
- Casassus, J., & Collin-Dufresne, P. (2005). Stochastic convenience yield implied from commodity futures and interest rates. *Journal of Finance*, 60, 2283–2331.
- Clewlow, L., & Strickland, C. (1999). A multi-factor model for energy prices (working paper). Sydney, Australia: School of Finance and Economics, University of Technology.
- Collin-Dufresne, P., & Goldstein, R. (2002). Do bonds span the fixed-income markets? Theory and evidence for unspanned stochastic volatility. *Journal of Finance*, 57, 1685–1730.
- Collin-Dufresne, P., Goldstein, R., & Jones, C. (2008). Identification of maximal affine term structure models. *Journal of Finance*, 63, 743–795.
- Cortazar, G., & Naranjo, L. (2006). An N-factor Gaussian model of oil futures prices. *Journal of Futures Markets*, 26, 243–268.
- Deng, S. (1999). Stochastic models of energy commodity prices and their applications: Mean-reversion with jumps and spikes (working paper). Berkeley, CA: University of California Energy Institute, University of California.
- Dai, Q., & Singleton, K. (2000). Specification analysis of affine term structure models. *Journal of Finance*, 55, 1943–1978.
- Duffie, D. (1996). *Dynamic asset pricing theory*. Princeton: Princeton University Press.
- Duffie, D., & Gray, S. (1995). Volatility in energy prices. In R. Jameson (Ed.), *Managing energy price risk* (pp. 39–56). London, UK; Houston, TX: Risk Publications and Enron Capital and Trade Resources.
- Duffie, D., & Kan, R. (1996). A yield-factor model of interest rates. *Mathematical Finance*, 6, 379–406.
- Duffie, D., Pan, J., & Singleton, K. (2000). Transform analysis and option pricing for affine jump-diffusions. *Econometrica*, 68, 1343–1376.
- Eydeland, A., & Wolyniec, K. (2003). *Energy and power risk management*. New Jersey: Wiley.

- Gabillon, J. (1995). Analysing the forward curve. In R. Jameson (Ed.), *Managing energy price risk* (pp. 29–38). London, UK; Houston, TX: Risk Publications and Enron Capital and Trade Resources.
- Gibson, R., & Schwartz, E. (1990). Stochastic convenience yield and the pricing of oil contingent claims. *Journal of Finance*, 45, 959–976.
- Heston, S. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6, 327–343.
- Hull, J. (2005). *Options, futures, and other derivatives* (6th ed.). New Jersey: Prentice-Hall.
- Korn, O. (2005). Drift matters: An analysis of commodity derivatives. *Journal of Futures Markets*, 25, 211–241.
- Lipster, R., & Shiryaev, A. (2004). *Statistics of random processes* (2nd ed.). New York: Springer.
- Litzenberger, R., & Rabinowitz, N. (1995). Backwardation in oil futures markets: Theory and empirical evidence. *Journal of Finance*, 50, 1517–1545.
- Nielsen, M., & Schwartz, E. (2004). Theory of storage and the pricing of commodity claims. *Review of Derivatives Research*, 7, 5–24.
- Pindyck, R. (2004). Volatility and commodity price dynamics. *Journal of Futures Markets*, 24, 1029–1047.
- Richter, M., & Sorensen, C. (2002). Stochastic volatility and seasonality in commodity futures and options: The case of soybeans (working paper). Frederiksberg, Denmark: Copenhagen Business School.
- Ross, S. (1997). Hedging long run commitments: Exercises in incomplete market pricing. *Economic Notes by Banca Monte*, 26, 99–132.
- Routledge, B., Seppi, D., & Spatt, C. (2000). Equilibrium forward curves for commodities. *Journal of Finance*, 55, 1297–1338.
- Schwartz, E. (1997). The stochastic behavior of commodity prices: Implications for valuations and hedging. *Journal of Finance*, 52, 922–973.
- Schwartz, E., & Smith, J. (2000). Short-term variations and long-term dynamics in commodity prices. *Management Science*, 46, 894–911.
- Trolle, A., & Schwartz, E. (2009). Unspanned stochastic volatility and the pricing of commodity derivatives. *Review of Financial Studies*, forthcoming.