A two-factor mean-reverting model

Commodity markets exhibit multi-factor behaviour as well as mean reversion. Building upon their previous paper, David Beaglehole and Alain Chebanier conclude the current Masterclass series by developing a two-factor mean-reverting model for crude oil that is then applied to various exotic derivatives valuation problems.

This article presents a two-factor model of commodity prices. The model has the properties of mean reversion, spot-dependent volatility and two-factor-ness. We first present the model and place it in the context of previous studies. We focus on the market for crude oil futures as traded on the New York Mercantile Exchange (Nymex). This contract has the symbol 'CL' and is sometimes referred to as West Texas Intermediate, which is one of the deliverable grades. Second, we examine the covariance matrix of some constant maturity interpolated Nymex CL futures contracts to see how well our proposed two-factor formulation fits the hedge ratios observed in the market, when compared with a one-factor model. For this, we apply a simple method of moments approach to calibrating the models. Third, we fit a one-factor non-skew, a two-factor non-skew, a one-factor skew and a two-factor skew model to options data from Nymex together with one market quote on a crude oil swaption, also based on the Nymex CL contract. Fourth, we examine the pricing implications for various derivatives that have been structured around crude oil. These instruments are: away-from-the-money swaptions, flex options, long-dated Asian-style straddles and time-spread options.

The two-factor skew model

Several authors have observed that for a number of commodities more than one factor is important in describing the changes in the future curve. These include Gabillon (1991), Cortazar & Schwartz (1994), Littenberger & Rabinowitz (1995), Schwartz (1997), Venetc (1997) and Hilliard & Reis (1998). This phenomenon has been noted for energy, gas, oil products and base metals. There is also an extensive literature on multi-factor models in the context of interest rates, as is summarised in Rebonato (1996). This multi-factorness of the forward curve has potential implications for hedging vanilla products and pricing exotic products.

Another issue, which affects the pricing and hedging of commodity derivatives, is the dependence of the local volatility of the commodity on its price level. This issue has been explored in the context of equity options by Derman (1994) and Dupire (1994). Neither of these approaches allowed for mean reversion in the underlying asset price. Other approaches have looked at stochastic volatility and jumps as an explanation for the skew in the volatility surface. Stein & Stein (1991) and Heston (1993) presented jump diffusion models with a skewed local volatility surface for gener- 

ics, but without considering mean reversion, or two factors with the extension that \( \eta_t \) follows the mean-reverting stochastic process:

\[
dt \eta_t = -\lambda_{\eta_t} \eta_t dt + \sigma_{\eta_t}(t) dW_{\eta_t}
\]

with the extension that \( \eta_t \) follows the mean-reverting stochastic process:

\[
dt \eta_t = -\lambda_{\eta_t} \eta_t dt + \sigma_{\eta_t}(t) dW_{\eta_t}
\]

where:

\[
dW_{\eta_t} = \rho_t dt
\]

and \( \mu_t \) is some function of time that is chosen to recover the forward curve. Note that here \( \eta_t \) can be thought of in the constant-parameter formulation as a perturbation of the long-term target rate around its mean value. Here, \( \sigma_{\eta_t}(t) \) is the diffusion volatility of \( \eta_t \) and \( W_t \) and \( W_{\eta_t} \) are two Brownian motions with correlation \( \rho_t \) which depends on time. Note that we can re-formulate this in terms of the long-term target value \( \theta_t \) by the substitution:

\[
\theta_t = \eta_t + \frac{\mu_t}{\lambda_t}
\]

In this case, we obtain:

\[
dt u_t = \left\{ \lambda_t (\theta_t - u_t) - \frac{1}{2} \sigma^2(u_t, t) \right\} dt + \sigma(u_t, t) dW_t
\]

and:

\[
d\theta_t = \lambda_{\theta_t} (\theta_t - \theta_t) dt + \sigma_{\theta_t}(t) dW_{\theta_t}
\]

where:

\[
\bar{\theta}_t = \frac{\mu_t}{\lambda_t} + \frac{1}{\lambda_{\theta_t}} \frac{d}{dt} \frac{\mu_t}{\lambda_t}
\]

is the target of the target rate.

This model inherits the skew behaviour of the earlier model, but now allows for two-factor-ness. We can further extend this model to consider a shifted log where:

\[
\mu_t = \ln \left( S_t + \frac{\beta}{1-\beta} F_t \right)
\]

Here \( \beta \) is the blend parameter and \( F_t \) is the commodity forward price today for delivery at \( t \). However, for the purposes of simplicity of exposition we shall use the unshifted lognormal form in this article. The advantage of the shifted lognormal form arises for markets where the structure of the volatility surface is close to this form. In a case like this, a shifted log model would give a function \( \sigma(u_t, t) \) with much less variation as a function of \( u_t \).

Model description. In an earlier article, Beaglehole & Chebanier (2002) present a one-factor skew model for commodities. Their model proposes the following process for the logarithm, \( u_t \), of the spot price, \( S_t \):

\[
dt u_t = \left\{ \lambda_t (\theta_t - u_t) - \frac{1}{2} \sigma^2(u_t, t) \right\} dt + \sigma(u_t, t) dW_t
\]

where \( \sigma(u_t, t) \) is the local diffusion volatility of the log commodity price, \( \lambda_t \) is the rate of mean reversion in the log commodity price, \( \theta_t \) is some long-term target rate and \( W_t \) is a Brownian motion. In this article, we have an equivalent formulation:

\[
du_t = \left\{ \lambda_t (\theta_t - u_t) - \frac{1}{2} \sigma^2(u_t, t) \right\} dt + \sigma(u_t, t) dW_t
\]

As in our earlier article, we choose to model the volatility function on any given time slice as being a continuously differentiable quadratic spline,
with piecewise linear interpolation in time. The above formulation results in the new stochastic differential equation (SDE):

\[ dx_t = \left[ \mu_t + \lambda_t \left( \eta_t - u_t \right) - \frac{1}{2} \sigma_t^2 (u_t) \right] dt + dW_t \]

Beaglehole & Chebanier (2002) detail the efficient calculation of the drift function. For convenience, we define:

\[ h(x, t) = \frac{\mu_t - \lambda_t \eta_t}{\sigma(u_t)} - \frac{1}{2} \frac{\partial \sigma(u_t)}{\partial u} \frac{\partial u_t}{\partial t} \]  

To solve this problem, we split the differential operator and solve sequentially on each variable. For this technique to work, it is important that the variables be locally uncorrelated so that the linear equations we end up solving are tridiagonal. If we can reduce the equations to tridiagonal form then we can solve them with a number of calculations that is proportional to the number of grid points for any given time node. The solution method is known as the alternating directions implicit method, and in particular the locally one directional variant of this. To orthogonalise variables, we define a new variable:

\[ z_t = \eta_t - \rho \sigma x_t \]

This variable is locally orthogonal to the \( x \) process. The SDE for this process is then:

\[ dz_t = \left[ -\left( \lambda_t + \rho \lambda_t \right) \right] \eta_t - \rho h(x, t) - \frac{d\rho}{dt} x_t \right] dt + \sigma_d dW_{t,j} \]

where:

\[ dW_{t,j} = 0 \]

\[ \sigma_d = \sqrt{1 - \rho^2 \sigma^2 (t)} \]

We could also simplify this by defining:

\[ f(x_t, z_t, t) = \frac{\lambda_t}{\sigma(u_t)} \left[ \eta_t - \sigma h(x, t) - \rho \frac{d\sigma}{dt} x_t \right] \]

which leads to the SDE formulation:

\[ dz_t = f(x_t, z_t, t) dt + \sigma_d dW_{t,j} \]

Also note that if we have boundary conditions determined by a continuous knock-out condition based on the prompt contract, then the solution region in \((x_t, z_t)\) is a rectangle in fact, the two-factor model is not much more difficult to implement than the one-factor model of Beaglehole & Chebanier (2002).

**Forward induction.** Any option pricing model needs to correctly price the current futures curve. In the previous sections, the relationship between \( \mu_t \) and the input forward curve was discussed, but we did not specify exactly how to derive \( \mu_t \) from the input forward curve. The basic approach is to first create a time axis. Given this time axis, iterate to find a piecewise constant function \( \mu_t \), which, when used in the backward induction, exactly recovers the values of all futures contracts. If you always perform a backward induction this would take a long time. Instead, you can perform a forward induction on the partial differential equation lattice. This starts with a probability density of 1.0 at the current prompt price and \( \eta \) value and then solves the 'forward' equation to propagate this discounted risk-neutral density forward on the lattice. The discounted futures price should then be the product of the solution to this equation with the prompt values on the lattice at the forward time point. You can keep the solution up to time \( t \) when iterating to find a drift that recovers the futures price to the next time node at time \( t+1 \).

It turns out that the discounted risk-neutral density function satisfies both the original equation and its forward equation. The pricing equation we are solving is:

\[ \frac{1}{2} G_{xx} + \frac{1}{2} \left( 1 - \rho^2 \right) \sigma_t^2 (u_t) G_x + \left( \frac{h(x, t)}{\sigma(u_t)} \right) G_z + \frac{\partial}{\partial t} G = 0 \]

while the forward equation is:

\[ \frac{1}{2} G_{xx} + \frac{1}{2} \left( 1 - \rho^2 \right) \sigma_t^2 (u_t) G_z + \left[ \frac{h(x', t')}{\sigma(u_{t'})} \right] G + \frac{\partial}{\partial t} G = 0 \]

The relationship between these two equations is discussed in Karatzas & Shreve (1991). Here, the variables \( x', z' \) and \( \eta \) denote the forward values of the corresponding variables \( t, x \) and \( \eta \) on the lattice. If we assume that the probability of reaching the upper and lower limits of our lattice are effectively zero, then we say we have 'homogeneous' boundary conditions. It turns out that solving the forward equation, for 'homogeneous' boundary conditions, corresponds to taking the transpose of the matrix operators that are usually applied for a single finite difference backward step. The advantage of forward induction is that the number of calculations is proportional to the total number of grid points in any cross-section of the lattice times the number of time points in the lattice. Without forward induction, the run time would be proportional to the total number of grid points in the lattice times the square of the number of time points in the lattice.

**Approximate centring.** We can further accelerate the process of fitting the forward curve by using an analytic approximation. Our objective is to obtain a sequence \( \mu_1, \mu_2, ..., \mu_n \) that recovers the futures prices at \( t_1, t_2, ..., t_n \) Since each forward solution of each iteration to estimate these coefficients takes the same time as the final solution of the backward equation, it is useful to derive an approximation that can be used in conjunction with, typically, a single solution for each time step of the forward equation. Our approach is to approximate the one-step ahead forward at each lattice point by using a solution that applies to a flat local volatility surface. Let \( \bar{E} \left[ \ln(S_{t+1}) \right] \) denote the expected value as of time \( t \) of \( \ln(S_{t+1}) \) under the local non-skew volatility process and:

\[ \bar{\text{VAR}}_E \left[ \ln(S_{t+1}) \right] \]

denote the corresponding variance. Specifically, we have that:

\[ \bar{E} \left[ \ln(S_{t+1}) \right] = e^{\lambda_t \left( \ln(S_{t+1}) - \lambda_t \ln(S_t) \right)} \lambda_t \eta_t + \left( \mu_t - \lambda_t \ln(S_t) \right) \exp \left( -\frac{\lambda_t}{\ln(S_t)} \right) \]

\[ \lambda_t \]

\[ \bar{E}_E \left[ \ln(S_{t+1}) \right] = \exp \left( \bar{E} \left[ \ln(S_{t+1}) \right] + \frac{1}{2} \bar{\text{VAR}}_E \left[ \ln(S_{t+1}) \right] \right) \]

The expression for:

\[ \bar{\text{VAR}}_E \left[ \ln(S_{t+1}) \right] \]

is indirectly available from the article by Beaglehole & Tenney (1991). Our approach is to take this expression and apply it at each time step. So if we perform backward induction we know the solution to the discounted risk-neutral density, or Green’s function, as a solution to the forward equation starting from an initial delta condition, then we can find out the approximate value for the next time interval of the \( \mu_t \) function. In particular, if the current solution for the Green’s function on the finite mesh for the \( k \)th value of \( x \) and the \( k \)th value of \( \eta \) is \( G(t_j, x_j, \eta_k; t_{k-1}, x_{k-1}, \eta_{k-1}) \), and the known futures price at time \( t_{k+1} \) is \( F_{t_{k+1}} \), then approximately:

\[ F_{t_{k+1}} \Delta \sum G(t_j, x_j, \eta_k; t_{k-1}, x_{k-1}, \eta_{k-1}) \psi \left( t_j, S(t_j, x_j), \eta_k, t_{k+1} \right) \]
From this equation we can back out the value of $\mu_t$. Also note that to some extent this approach is self-correcting since errors in previous time-step iterations are compensated for in the current time step.

**Solving for time-spread options.** Time-spread options can be valued on a lattice in much the same way as regular Asian options. Specifically, you solve backwards on multiple parallel grids, each representing a different value for the current average. A jump condition is applied at each discrete averaging point.

In a sense, a time-spread option is an example of a weighted Asian average option, where the weight on the first setting is one, the weighting on the last setting is minus one and the weighting on all other settings is zero. So, if we can solve the general problem of solving for an option on a weighted average, where the weight on the first setting is one, the weighting on all other settings is zero, we can solve for the time-spread option. We do this by solving on a set of parallel lattices, which each represents a separate value at time $t_f$ of the running average quantity:

$$Payout = \left( \sum_{i=0}^{n} w_i S_i - K \right)^n$$

where a reasonable distribution is chosen for the range of values that this average can take.

**An analysis of the covariance matrix of Nymex CL futures prices**

We took the historical changes of Nymex CL futures, interpolated to constant monthly maturities. These futures prices were taken out to remaining terms of 24 months and cover the time period from January 1995 to June 19, 2001. The first principal component accounts for 91.2% of the total variance in the futures, while the next principal component accounts for a further 7.1%. Thus the first two principal components account for 98.3% of all variation in the Nymex CL futures curve out to 24 months. The subsequent principal components have a much smaller contribution to variance, and we chose for this reason to exclude these from our model. In addition, it should be remembered that a three- or four-factor lattice model can be quite slow to run.

As a next step, we fitted a one-factor non-skew mean-reverting model to this covariance matrix of futures, by minimising the fit of this model to the covariance moments. Figure 1 shows the percentage errors in the covariance matrix. The values of the fit were $\lambda = 23\%$, $\sigma = 29\%$. Note that there are errors of as much as 40% of the original covariance. A rule of thumb in looking at percentage errors in covariance is that they are expected to be twice the percentage error of a fit to volatility.

After this, we fitted a two-factor non-skew mean-reverting model. In this case, we see that the largest percentage error is around 10%, and most errors are much smaller. The average percentage error is only 2%. Figure 2 shows the covariance error percentages. We obtained the following values from the fit: $\lambda = 152\%$, $\sigma = 37\%$, $\lambda_0 = 0\%$, $\sigma_0 = 21\%$ and $\rho = 57\%$.

This analysis justifies moving from a one-factor to a two-factor analysis of the market. It also brings into question whether more than two factors are needed. Note that a principal components analysis is more general than our restricted Markovian model. With a Markov model, you cannot specify an arbitrary factor loading for the first principal component since you are limited to exponentially decaying volatility factors. So even if the first principal component accounts for 91.2% of the variation in the market, this does not mean that an exponentially decaying factor can capture all this variance.

**Calibration approach**

First we use historical return data of crude oil futures to fix the mean reversion of the short rate, the mean reversion of the ‘target’ rate and the correlation between the spot rate and the target rate. We then fit a local volatility surface for the spot rate, using three or four parameters per unique option expiration. We actually use five options for each expiry, so the fit is a least squares fit. We also fit a single time-independent value for $\sigma_0$.

The calibration targets are a set of American-style options on Nymex CL and a European-style swaption quote from the over-the-counter market that exercises into a swap on Nymex CL. Once we have calibrated the full...
two-factor skew model, we then treat this calibrated model as the market and fit a one-factor non-skew, a one-factor skew and a two-factor non-skew model to the prices implied by this model. This is to ensure the greatest degree of consistency in the results. We were able to fit 30 options with an root mean squared error of 0.14 of a volatility point. All prices were within their bid-offer spreads after the calibration.

Figure 3 shows the implied local volatility surface for the spot rate that is implied by the fit of the two-factor skew model. The graph is fairly regular, showing a bias towards higher percentage volatilities at lower crude oil prices. Figure 4 displays the implied future at-the-money spot volatilities of the one-factor and two-factor non-skew models. As expected, the two-factor model displays higher spot volatility over time because it applies a much higher rate of mean reversion to changes in prompt futures commodity prices, implying a higher local volatility for the spot price.

We fit the one-factor skew model to seven options at each available expiry using seven parameters for each expiry. This is to ensure the widest agreement on the future distribution of spot.

### Pricing comparisons

Here, we compare the four calibrations obtained in the previous section by using them to price a range of exotic options that have been entered into in the context of crude oil or related fuels. In considering these differences, we should bear in mind that for long-dated American-style option on futures, the bid-offer for at-the-money options is around 0.5% lognormal implied volatility, widening to more like 1.0% lognormal implied volatility away-from-the-money.

#### European-style commodity swaptions

In all the previous calibrations, we fit to an at-the-money European-style swaption with 18 months to expiry and with an underlying tenor of one year, covering calendar year 2003. We did not attempt to calibrate the prices of swaptions at other than the at-the-money strikes. The question is then what these various models imply for the pricing of these other swaptions. Figure 5 illustrates the implied flat volatilities for a range of strikes. As expected, the non-skew one-factor and two-factor models display almost no departure from a flat implied volatility graph. There is a very small trend due to the fact that the forwards are not exactly scalar multiples of each other. The more dramatic difference, as expected, is for the two skew models. For both, there is a smile, but the smile is much greater for the one-factor case than for the two-factor case. The intuitive explanation is that in the two-factor case the swap rate is driven by both a skewed ‘spot’ commodity price and a non-skewed intermediate term target rate. Because the covariance between these rates is smaller in a two-factor model, the skew of the swap rate is less than that of the single contract rate option.

#### Flex options

A flex option is an option to exercise up to a specified fraction of a full strip of European-style options. In our examples, we look at a two-year flex option for which one or six of the available 24 monthly options can be exercised. The options are examined at different strike prices. In figure 6, we see the case of a flex option with only one available exercise ‘bullet’. Note that this is identical to a Bermudan-style option on the near contract. In this case, it is clear that there is a huge impact due to two-factorness. Even at-the-money, the difference is around six volatility points, while deep in-the-money the volatility difference can be as much as 15 volatility points. The values for skewed models are higher than for non-skewed-in-the-money, as expected. Deep out-of-the-money, the two-factor becomes much less important, with skew effects dominating.

The reason that two-factorness is so important for in-the-money flex options is that when the curve is steeply backwardated, the lower the strike of the flex option the more uncertainty as to which option will be exercised. This means that the option to switch between different exercise dates is more valuable, and also is driven partly by the changes in the shape of the forward curve. This is because the correlation between forwards is less in the two-factor model.
Figure 7 shows an interesting fact. Even with six out of 24 bullets being available, qualitatively the same picture applies, though now the graphs are somewhat closer together.

- **Long-dated Asian options.** Most options on fuels that are cash settled are based on an average of daily sets during a month. A less standard contract involves an option on the average of daily settings over, possibly, several years. We will consider a two-year structure with monthly averaging but a single exercise, specified with varying strikes. Figure 8 displays the implied flat volatility for the structure at different strikes. What we see here is that there is a relatively small difference of about half of a volatility point between the one- and two-factor models. The two-factor model gives a lower value because the correlation between settings is a little lower for this model. What is interesting is that the smiles that are derived from the one-factor skew and the two-factor skew models are very close to being shifts of each other. This is a very different picture than we find from looking at the smile in European-style swaptions. The intuition here is that the relevant volatilities are the term volatilities of spot, which are matched by both the one-factor and the two-factor calibrations.

- **Time-spread options.** A time-spread option is an option on the difference between the prompt month and next month futures contract prices. Typically, these trade in strips of options. We look at the value of each option in a strip going out to two years. Clearly we would expect a two-factor model to give a very different price to a one-factor model for these structures. Figure 9 shows the implied volatility of these time-spread options going out to two years. As expected, the two-factor models show a higher price for these options than the one-factor models. This is a consequence of the higher local volatilities already noted for the two-factor model and also the volatility of the forward curve being higher in the two-factor models. The skew models give a higher value to these options because they consider the full volatility surface, which is generally higher than the at-the-money volatility curve for both one- and two-factor models.

### Conclusions and future directions for research

We have established the importance in considering the two-factor skew of the futures curve for crude oil when pricing a variety of commodity exotics. In all the cases we considered, except long-dated Asian structures, two-factor skewness is a first-order consideration. This observation applies to many related traded commodities, particularly those linked to energy and oil products. The issues that need to be addressed further are stochastic volatility and jumps. To a large extent, having a highly mean-reverting short-term factor is analogous to having jumps, so stochastic volatility is probably the next most significant issue in the energy context.

If we take the process of this article and then assume that stochastic volatility applies to the volatility of the short rate then we would expect some of the smile in the local volatility surface to be flattened after recalibrating to the data. Since in all cases we fit to the same American-style options on futures prices we must look at the pricing implications for other derivatives. One category of derivatives that we have not discussed in this paper is barrier options. The values of barrier options are typically quite different under pure local volatility versus pure stochastic volatility models.

David Beaglehole is a managing director and Alain Chebanier is an associate at Deutsche Bank in North America. They wish to thank Mohan Rajagopalan of the commodities trading desk at Deutsche Bank in North America for many helpful comments.

### References

- Andersen L and J Andreasen, 2000
  - Jump-diffusion processes: volatility smile fitting and numerical methods for option pricing
  - Review of Derivatives Research 4, pages 231–262

- Beaglehole D and A Chebanier, 2002
  - Mean reversion with a smile
  - Risk April, pages 95–98

- Beaglehole D and M Tenney, 1991
  - General solutions to some interest rate contingent claim pricing problems
  - Journal of Fixed Income, September, pages 69–83

- Cortazar G and E Schwartz, 1994
  - The valuation of commodity contingent claims
  - Journal of Derivatives, summer, pages 27–39

- Derman E and I Kani, 1994
  - Riding on a smile
  - Risk February, pages 32–39

- Dupire B, 1994
  - Pricing with a smile
  - Risk January, pages 18–20

- Gabillet J, 1991
  - The term structure of oil futures prices
  - Institute of Energy Studies

- Heston S, 1993
  - A closed form solution for options with stochastic volatility with applications to bond and currency options
  - Review of Financial Studies 6(2), pages 327–343

- Hilliard J and J Reis, 1998
  - Valuation of commodity futures and options under stochastic convenience yield, interest rates and jump diffusions in the spot

- Karatzas I and S Shreve, 1991
  - Brownian motion and stochastic calculus
  - Springer-Verlag, second edition

- Litzenberger R and N Rabinowitz, 1995
  - Backwardation in oil futures markets: theory and empirical evidence
  - Journal of Finance 50(5), January, pages 1,517–1,545

- Merton R, 1976
  - Option pricing when underlying stock returns are discontinuous

- Rebonato R, 1996
  - Interest rate option models
  - John Wiley & Sons

- Schwartz E, 1997
  - The stochastic behavior of commodity prices: implications for valuation and hedging
  - Journal of Finance 52(3), July, pages 923–972

- Stein E and J Stein, 1991
  - Stock price distributions with stochastic volatility: an analytic approach
  - Review of Financial Studies 4(4), pages 727–752

- Veretete L and S Hodges, 1997
  - Modelling commodity futures spreads: an empirical study
  - Proceedings of the Tenth Annual European Futures Research Symposium, Chicago Board of Trade, September, pages 57–91