Modeling Commodity Spreads with Vector Autoregressions

Ted Kury
The Energy Authority ®
ICDSA 2007
June 1, 2007
Rule #1 of Pricing Models

Pricing models can offer valuable insight into the behavior of simple or complex markets
Rule #2 of Pricing Models

“Markets tend to be rational in the long run, but markets can stay irrational longer than you can stay solvent”
J.M. Keynes (attributed)
The Problem

• Model the prices of two closely-related commodities  
  – Natural gas at two different delivery points  
  – Interest rates  

• Prices usually tied together by some fundamental factor (e.g. transportation rates)  

• Capture not only the evolution of prices, but the relationship between prices
Natural Gas Time Series

Natural Gas Price History

$/MMBtu

HH Gas | TZ4 Gas | Spread

The Nation's Leader in Public Power Energy Trading
## Gas Price Characteristics

<table>
<thead>
<tr>
<th>%tile</th>
<th>Since 1994</th>
<th></th>
<th></th>
<th>Since 2000</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Henry Hub</td>
<td>Transco</td>
<td>Basis</td>
<td>Henry Hub</td>
<td>Transco</td>
<td>Basis</td>
</tr>
<tr>
<td>1st %tile</td>
<td>1.44</td>
<td>1.46</td>
<td>-0.20</td>
<td>2.03</td>
<td>2.10</td>
<td>-0.12</td>
</tr>
<tr>
<td>5th %tile</td>
<td>1.59</td>
<td>1.60</td>
<td>-0.06</td>
<td>2.41</td>
<td>2.46</td>
<td>-0.04</td>
</tr>
<tr>
<td>10th %tile</td>
<td>1.74</td>
<td>1.76</td>
<td>-0.03</td>
<td>2.82</td>
<td>2.86</td>
<td>-0.01</td>
</tr>
<tr>
<td>25th %tile</td>
<td>2.17</td>
<td>2.19</td>
<td>0.00</td>
<td>3.94</td>
<td>3.98</td>
<td>0.02</td>
</tr>
<tr>
<td>75th %tile</td>
<td>5.61</td>
<td>5.66</td>
<td>0.07</td>
<td>6.66</td>
<td>6.81</td>
<td>0.12</td>
</tr>
<tr>
<td>90th %tile</td>
<td>7.15</td>
<td>7.27</td>
<td>0.18</td>
<td>7.79</td>
<td>8.05</td>
<td>0.27</td>
</tr>
<tr>
<td>95th %tile</td>
<td>8.01</td>
<td>8.22</td>
<td>0.30</td>
<td>9.64</td>
<td>9.92</td>
<td>0.40</td>
</tr>
<tr>
<td>99th %tile</td>
<td>12.82</td>
<td>13.28</td>
<td>0.72</td>
<td>13.68</td>
<td>14.28</td>
<td>0.95</td>
</tr>
</tbody>
</table>
Presentation Outline

• Modeling Considerations
• Traditional Energy Models
  – Modeling Prices
  – Modeling Spreads
• Vector Autoregression Framework
Presentation Outline

• **Modeling Considerations**
• **Traditional Energy Models**
  – Modeling Prices
  – Modeling Spreads
• **Vector Autoregression Framework**
Modeling Considerations

- **Relative Prices**
  - Spread Option

- **Absolute and Relative Prices**
  - Barrier Option
  - Cash Flow at Risk of Forward Purchase or Sale
  - Absolute Product (Natural Gas) Cost
Presentation Outline

• Modeling Considerations
• Traditional Energy Models
  – Modeling Prices
  – Modeling Spreads
• Vector Autoregression Framework
The Usual Suspects

• Closed-Form Spread Option Formulas
• Geometric Brownian Motion Price Model
• Single Factor Mean Reversion Price Model
Traditional Spread Models

• Model such as Margrabe (1978)
• Derived from Black-Scholes, so it shares its assumptions
  – Lognormal price returns
  – Independent and identically distributed shocks
  – No transaction costs
Margrabe Valuation for Spread Options

- Similar to Black and Black-Scholes
  
  \[ w(x_1, x_2, t) = x_1 N(d_1) - x_2 N(d_2) \]

  where:

  \[ d_1 = \frac{\ln(x_1 / x_2) + \frac{1}{2} \sigma^2 t}{\sigma \sqrt{t}} \]

  \[ d_2 = d_1 - \sigma \sqrt{t} \]

  \[ \sigma^2 = \sigma_1^2 + \sigma_2^2 - 2 \rho \sigma_1 \sigma_2 \]
Geometric Brownian Motion Model (GBM)

• Use the log formulation since we’re modeling a trending series

\[ \ln S_t = \ln S_{t-1} + \varepsilon_t \]

where:

\[ \varepsilon_t \sim N(0, \sigma^2) \]
Price Evolution Under GBM

• Let’s just assume that $p$ is the log price
• At time 1
  \[ p_1 = p_0 + \varepsilon_1 \]
• At time 2
  \[ p_2 = [p_0 + \varepsilon_1] + \varepsilon_2 \]
• At time $t$, collecting the shock terms
  \[ p_t = p_0 + \sum_{i=1}^{t} \varepsilon_i \]
Price Variance Under GBM

• Variance of each individual shock term

\[ \text{Var}(\varepsilon_t) = \sigma^2 \]

• So the variance of \( t \) terms

\[ \text{Var}(p_t) = t \sigma^2 \]
Modeling the Spread Between Two Prices

• Assume two prices, $p$ and $q$, where

$$ q = p + \eta $$

and:

$$ \varepsilon_{pt} \sim N(0, \sigma_p^2) $$

$$ \varepsilon_{qt} \sim N(0, \sigma_q^2) $$

$$ \rho = \text{corr}(\varepsilon_p, \varepsilon_q) $$
Spread between GBM Prices

• So the basis at time $t$

$$q_t - p_t = p_0 + \eta_0 - p_0 + \sum_{i=1}^{t} \varepsilon_{qi} - \sum_{i=1}^{t} \varepsilon_{pi}$$

• or

$$q_t - p_t = \eta_0 + \sum_{i=1}^{t} \varepsilon_{qi} - \sum_{i=1}^{t} \varepsilon_{pi}$$

• With variance

$$Var(q_t - p_t) = t(\sigma_p^2 + \sigma_q^2 - 2\rho\sigma_p\sigma_q)$$
Single Factor Mean Reverting Model (SFMR)

- Framework of Pindyck (1999) and Schwartz (1997)

\[
\ln S_t = \ln S_{t-1} + \alpha (\mu - \ln S_{t-1}) + \varepsilon_t
\]

where:

\[
\varepsilon_t \sim N(0, \sigma^2)
\]

- \(\alpha\) is mean reversion rate
- \(\mu\) is log of the long run equilibrium price
Price Evolution Under SFMR

• Again, assuming that \( p \) is the log price
• At time 1, rearranging terms
  \[
p_1 = \alpha \mu + (1 - \alpha) p_0 + \epsilon_1
  \]
• At time 2
  \[
p_2 = \alpha \mu + (1 - \alpha)[\alpha \mu + (1 - \alpha) p_0 + \epsilon_1] + \epsilon_2
  \]
• At time \( t \), collecting terms
  \[
p_t = \alpha \mu \sum_{i=0}^{t-1} (1 - \alpha)^i + (1 - \alpha)^t p_0 + \sum_{i=1}^{t} \epsilon_i (1 - \alpha)^{t-i}
  \]
Price Variance Under SFMR

- Variance of each individual shock term
  \[ Var(\varepsilon_t) = \sigma^2 \]
- So the variance of \( t \) terms
  \[ Var(p_t) = \sum_{i=1}^{t} \sigma^2 (1 - \alpha)^{2(t-i)} \]
- Which reduces to
  \[ Var(p_t) = \frac{1 - (1 - \alpha)^{2t}}{1 - (1 - \alpha)^2} \sigma^2 \]
Variance Comparisons

Variance Growth Rates

- No Mean Reversion
- Slower Mean Reversion
- Faster Mean Reversion

The Nation's Leader in Public Power Energy Trading
Spread between SFMR Prices

- So the basis at time $t$

$$q_t - p_t = \left[ (1 - \alpha_p)^t - (1 - \alpha_q)^t \right] (\mu_p - p_0) + (1 - \alpha_q)^t \eta_0 + \left[ 1 - (1 - \alpha_q)^t \right] \eta$$

$$+ \sum_{i=0}^{t-1} \varepsilon_{qi} (1 - \alpha_q)^{t-i} - \sum_{i=0}^{t-1} \varepsilon_{pi} (1 - \alpha_p)^{t-i}$$

- With variance

$$Var(q_t - p_t) = \frac{1 - (1 - \alpha_p)^{2t}}{1 - (1 - \alpha_p)^2} \sigma_p^2 + \frac{1 - (1 - \alpha_q)^{2t}}{1 - (1 - \alpha_q)^2} \sigma_q^2 - 2 \rho \frac{\sqrt{1 - (1 - \alpha_p)^{2t}}}{1 - (1 - \alpha_p)^2} \sigma_p \sqrt{1 - (1 - \alpha_q)^{2t}} \sigma_q$$
Behavior of SFMR Expected Basis

Basis Evolution Under Different Mean Reversion Rates

- Price with Lower Mean Reversion Rate
- Price with Higher Mean Reversion Rate
- Basis with Bias
- Basis without Bias

The Nation's Leader in Public Power Energy Trading
Behavior of SFMR Shocks

Decay of SFMR Shocks

The Nation's Leader in Public Power Energy Trading
Behavior of SFMR Shocks

Decay of SFMR Shocks

- Slower Mean Reversion Rate
- Faster Mean Reversion Rate
- Difference
Behavior of SFMR Shocks

Decay of SFMR Shocks

Time

0.0000 0.1000 0.2000 0.3000 0.4000 0.5000 0.6000 0.7000 0.8000

0 3 6 9 12 15 18 21 24 27 30 33 36 39 42 45 48 51 54 57 60 63 66 69 72 75 78 81 84 87 90 93 96 99

Standard Normal Shock

0.0000 0.0100 0.0200 0.0300 0.0400 0.0500 0.0600 0.0700 0.0800

The Nation's Leader in Public Power Energy Trading
Traditional Energy Models

• Margrabe Spread Model
  – Shares assumptions, both good and bad, with Black-Scholes

• Geometric Brownian Motion
  – Fixed expected basis equal to today’s basis
  – Infinite variance of spread

• Single Factor Mean Reverting
  – Variable expected basis
  – Finite variance of spread, but different mean reversion rates can lead to much different decay rates
  – Difference in shocks can increase and may diverge from what is seen in reality
Presentation Outline

• Modeling Considerations
• Traditional Energy Models
  – Modeling Prices
  – Modeling Spreads
• Vector Autoregression Framework
Vector Autoregressions - The Better Mousetrap

• Vector autoregression framework allows greater flexibility
  – Established methodology
  – Robust diagnostic testing
  – Multiple methodologies to handle shocks

• Future path of prices depends on
  – Historical path of all modeled prices; and
  – Future path of other prices
Vector Autoregression Model (VAR)

• Models prices of goods that are close substitutes

\[ p_t = \alpha_1 + \beta_{11} p_{t-1} + \beta_{12} q_{t-1} + \varepsilon_{pt} \]
\[ q_t = \alpha_2 + \beta_{21} p_{t-1} + \beta_{22} q_{t-1} + \varepsilon_{qt} \]

where:

\[ \varepsilon_{pt} \sim N(0, \sigma_p^2) \]
\[ \varepsilon_{qt} \sim N(0, \sigma_q^2) \]
\[ \rho = \text{corr}(\varepsilon_p, \varepsilon_q) \]
Price Evolution under VAR

- Matrix representation

\[
\begin{bmatrix}
p_t \\
q_t \\
\end{bmatrix} = \begin{bmatrix}
\alpha_1 \\
\alpha_2 \\
\end{bmatrix} + \begin{bmatrix}
\beta_{11} & \beta_{12} \\
\beta_{21} & \beta_{22} \\
\end{bmatrix} \begin{bmatrix}
p_{t-1} \\
q_{t-1} \\
\end{bmatrix} + \begin{bmatrix}
\epsilon_{pt} \\
\epsilon_{qt} \\
\end{bmatrix}
\]

- Change notation to

\[P_t = A + BP_{t-1} + E_t\]

- Price at \(t\) in terms of \(P_0\)

\[P_t = B^t P_0 + \sum_{i=0}^{t-1} B^i (A + E_{t-i})\]
The Stability of the Weighting Matrix

• Given the weighting matrix

\[ B^1 = \begin{bmatrix} 0.726 & 0.253 \\ 0.171 & 0.822 \end{bmatrix} \]

• Subsequent powers are

\[ B^2 = \begin{bmatrix} 0.571 & 0.392 \\ 0.266 & 0.719 \end{bmatrix} \quad \quad B^3 = \begin{bmatrix} 0.482 & 0.467 \\ 0.316 & 0.659 \end{bmatrix} \]

\[ B^5 = \begin{bmatrix} 0.399 & 0.525 \\ 0.356 & 0.598 \end{bmatrix} \quad \quad B^{10} = \begin{bmatrix} 0.346 & 0.524 \\ 0.355 & 0.544 \end{bmatrix} \]
Sample VAR Simulation

VAR Simulation

$/MMBtu

05/01/07 05/03/07 05/05/07 05/07/07 05/09/07 05/11/07 05/13/07 05/15/07 05/17/07 05/19/07 05/21/07 05/23/07 05/25/07 05/27/07 05/29/07 05/31/07 06/02/07 06/04/07 06/06/07 06/08/07 06/10/07 06/12/07 06/14/07 06/16/07 06/18/07 06/20/07 06/22/07 06/24/07 06/26/07 06/28/07 06/30/07

Henry Hub  TZ4  Spread

The Nation's Leader in Public Power Energy Trading
Behavior of VAR Shocks

Decay of VAR Shocks

The Nation's Leader in Public Power Energy Trading
Behavior of VAR Shocks

Decay of VAR Shocks

The Nation's Leader in Public Power Energy Trading
Strengths of the VAR

- Construct paths for prices without associated forward curves
- Several well established tests to determine optimal number of lags
- Two methods to correlate price shocks
  - Actual correlation and normally distributed shocks
  - Resample actual historical shocks to pick up correlation and any non-normal distributions
Weaknesses of the VAR

• More rigorous process to determine parameters
• More diagnostic testing of model
  – Proper functional form
  – Stable system of equations
• Number of parameters grows quickly ($N^2L$) and can erode your degrees of freedom, so a larger data set may be required
Pipeline Management Risk Assessment

• Resource management problem
• Value of transportation capacity
• Risk depends on the price of gas at 3 hubs
• Most conservative test shows the need for 100 lags
Pipeline Model Simulation

Sample Model Iteration

$/MMBtu

Transco Zone 1  Transco Zone 3  Transco Zone 6  Z3-Z1 Spread  Z6-Z3 Spread

The Nation's Leader in Public Power Energy Trading
Summary

• Traditional models may not work well to model absolute price levels and commodity spreads
  – Infinite variance
  – Unstable mean spreads
  – Different rates of mean reversion can cause divergence over time

• Vector autoregression offers a flexible framework
  – Better captures price interactions
  – Derive future path for prices without forward curve
  – Handle non-normal and heteroscedastic shocks
Questions?

• Contact Info:
  Ted Kury
  tkury@teainc.org
  (904) 360-1444
References

• Margrabe, W., 1978, “The Value of an Option to Exchange One Asset for Another”, *Journal of Finance* 33:1 p. 177-186
