The Coskewness Premium: Where does it come from?

Sung June Pyun

University of Southern California

Dec 16th, 2014
Coskewness: $U'''' > 0$

(Kraus and Litzenberger, 1976; Harvey and Siddique, 2000)

$$E_t[r_{i,t+1}] = \text{Cov}_t(r_{i,t+1}, r_{m,t+1}) \lambda_m - \text{Cov}_t(r_{i,t+1}, \sigma^2_{m,t+1}) \lambda_c$$

Volatility risk: Inter-temporal decision (Ang et al 2006)

$$E_t[r_{i,t+1}] = \text{Cov}_t(r_{i,t+1}, r_{m,t+1}) \lambda_m - \text{Cov}_t(r_{i,t+1}, \Delta \sigma^2_{t+1}) \lambda_v$$

In today’s presentation:

- Are coskewness and volatility risk related? If yes, how?
- What is the dynamic relation between the premia?
Jump-diffusion Process

- Two sources of market risk: Jump risk + Diffusion risk
- Simultaneous and negatively correlated jumps in price and volatility
  - Allows stochastic vol-of-vol, solves volatility puzzle
  - Bad (good) news → higher (lower) variation, negative (positive) price jumps

I show: Coskewness \(\approx VR + TR\)

1) Volatility Risk (VR): exposure to volatility of diffusion
2) State-transition Risk (TR): exposure to stochastic jumps
   - Jumps lead to a state with different valuations and variances

* Merton 1976; Barndorff-Nielsen and Shephard 2003, 2006; Chernov, Gallant, Ghysels and Tauchen, 2003; Anderson, Bolleslev and Diebold 2009 and others
+ Duffie, Pan Singleton, 2000; Pan, 2002; Eraker, Johannes and Polson 2003; Todordv and Tauchen, 2011 and others
Two Components of Coskewness

- **Volatility Risk (VR) Component**
  - Replicating portfolio of straddle or strangle
  - Returns co-move with volatility ($B > A > C$)
  - Ang, Hodrick, Xing and Zhang, 2006; Jones 2003; Adrian Rosenberg, 2008; Chang, Christoffersen and Jacobs, 2013; etc

- **State-Transition Risk (TR) Component**
  - Investment in vol-of-vol or VVIX
  - Higher returns when volatility moves abruptly ($B/C > 0 > A$)
  - Jones, 2006; Cvitanic, Polimenis and Zapatero, 2008; Ozoguz, 2009; Todorv, 2010, Todorv and Bollerslev, 2011; etc
Two Components of Coskewness

Daily Series of Market Volatility

<table>
<thead>
<tr>
<th></th>
<th>$t_2$</th>
<th>$t_3$</th>
<th>On Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>VR</td>
<td>Hedge (Negative)</td>
<td>Hedge (Negative)</td>
<td>(Negative)</td>
</tr>
<tr>
<td>TR</td>
<td>Hedge (Negative)</td>
<td>Adds Risk (Positive)</td>
<td>(Indeterminable)</td>
</tr>
<tr>
<td>Coskewness Premium</td>
<td>(Negative)</td>
<td>Positive if TR dominates VR</td>
<td></td>
</tr>
</tbody>
</table>
The Coskewness Premium

What do we know about the coskewness premium?

1. Assets with positive coskewness have lower subsequent returns
   (Kraus and Litzenberger, 1976; Harvey and Siddique, 1999, 2000)

2. The premium is time-varying and depends on historical market moments
   (Friend and Westerfield, 1980; Smith, 2006)

3. The premium is related to the size premium (Adesi et al 2004; Vanden 2006)
Two Volatility States: H and L

- States follow a Markov Process with a transition matrix

\[
\Pi = \begin{bmatrix}
\pi_{LL} > 0.5 & \pi_{LH} \\
\pi_{HL} & \pi_{HH} > 0.5
\end{bmatrix}
\]

- Market Process \((h_H > h_L)\):

\[
\begin{align*}
 r_{m,t+1} &= \mu_s - \xi_s J_{s,s'} + \sqrt{h_{t+1}} \epsilon_{t+1}, \quad \epsilon_{t+1} \sim N(0, 1) \\
 \Delta h_{t+1} &= \xi_v J_{s,s'}
\end{align*}
\]

where \(s_t = s, \ s_{t+1} = s', \ \xi_s > 0, \ \xi_v > 0\) and \(J_{L,H} = 1, \ J_{H,L} = -1, \ 0\) o.w.

-\( r_{m,t+1}|J_{s,s'}\) Normal but \(\sqrt{h_{t+1}} \epsilon_{t+1}|F_t\) and \(r_{m,t+1}|F_t\) Normal-mixtures

- Conditional variance of the market

\[
Var[r_{m,t+1}|s] = E[h_{t+1}|s] + (\xi_s)^2 Var[J_{s,s'}|s]
\]

- Risk-averse investor holds the market portfolio
Why the CAPM fails: Conditional Variance of a Portfolio

- Single market factor may be misleading when there are two sources of variations
  Jump and diffusion (Todorv and Bollerslev, 2011; Bollerslev and Li, 2013)

\[ r_{p,t+1} = \alpha_s + \beta_J \xi_s J_{s,s'} + \beta_\epsilon \sqrt{h_{t+1}} \epsilon_{t+1} + e_{p,t+1} \]

\[ \text{Var}[r_{p,t+1} | s] = \beta_\epsilon^2 E[h_{t+1} | s] + (\beta_J \xi_s)^2 \text{Var}[J_{s,s'} | s] \]

- Either a lower exposure to diffusion \((\beta_\epsilon)\) or jump \((\beta_J)\) implies lower variance
A Decomposition of Coskewness

Recall that
\[ r_{m,t+1} = \mu_s - \xi_s J_{s,s'} + \sqrt{h_{t+1}} \epsilon_{t+1} \]
\[ r_{m,t+1}^2 = \mu_s^2 + \xi_s^2 1_{s \neq s'} + h_{t+1} \epsilon_{t+1}^2 + 2\{-\mu_s \xi_s J_{s,s'} + (\mu_s - \xi_s J_{s,s'}) \sqrt{h_{t+1}} \epsilon_{t+1}\} \]

Decomposition of Coskewness

\[ \text{Cov}_t(r_{i,t+1}, r_{m,t+1}^2) = \text{Cov}_t(r_{i,t+1}, h_{t+1} \epsilon_{t+1}^2) \quad : \quad \text{Volatility Risk (VR)} \]
\[ + \xi_s^2 (1 + 2|\tau_s|) \text{Cov}_t(r_{i,t+1}, 1_{s \neq s'}) \quad : \quad \text{Transition Risk (TR)} \]
\[ + 2(\mu_s + \xi_s \tau_s) \text{Cov}_t(r_{i,t+1}, r_{m,t+1}) \quad : \quad \text{Market Beta} \]

where \( \tau_s = \pi_{LH} \) if \( s = L \) and \( -\pi_{HL} \) if \( s = H \)

**Note that** \( \sqrt{h_{t+1}} \epsilon_{t+1} \perp J_{t+1} \)
• **Volatility risk** (VR) Component \[ Cov_t(r_{i,t+1}, h_{t+1}\epsilon^2_{t+1}) \]
  - Higher returns when variations in diffusion \((h_{t+1})\) are high
  - Provides hedge against
    - Low states: negative price jumps
    - High states: positive price jumps NOT occurring

• **Transition risk** (TR) Component \[ Cov_t(r_{i,t+1}, 1_s\neq s') \]
  Jumps are
  - Low states: **positive** in volatility, **negative** in prices
    Provides hedge against negative price jumps
  - High states: **negative** in volatility, **positive** in prices
    Adds risk when positive price jumps do NOT occur
Empirical Predictions

1. High coskewness implies either high VR, high TR or both

2. The coskewness premium related to the VR or TR premium
   2a. The VR premium (AHXZ, 2006) negative and state independent
       • Higher return when valuations are low
   2b. The sign of the TR premium state-dependent
       • Mean-reversion in volatility → jumps likely to be reversed
       • Negative (Positive) premium following negative (positive) volatility jumps
   2c. State-dependency of the coskewness premium is explained by the premia of the components
       • Sign may be driven by the VR, state-dependency resembles that of the TR premium
Stochastic Volatility Model:  
(DPS, 2000, Eraker, Johannes and Polson, 2003)

\[
\begin{align*}
    r_{m,t} &= r_0 + \xi_t J_t + e^{h_t/2} \epsilon_{1,t} \\
    h_t &= \mu_h + \phi_h (h_{t-1} - \mu_h) + \xi_t^v J_t + \sigma_h \epsilon_{2,t}
\end{align*}
\]

where

- \( J \in (-1, 0, 1) \) with probability \( \lambda_n, 1 - \lambda_n - \lambda_p \) and \( \lambda_p \)
- \( \xi_t^v \sim \text{Exp}(\lambda) \), \( \xi_t \sim \text{N}(\rho_0 + \rho_1 \xi_t^v, \sigma^2_\xi) \) and \( (\epsilon_1, \epsilon_2) \sim \text{MVN}(0, I) \)
- Use MCMC to estimate \( \Theta = \{ r_0, \mu_h, \phi_h, \lambda_p, \lambda_n, \lambda, \sigma^2_h, \rho_0, \rho_1, \sigma^2_\xi \} \) and \( h_t, J_t, \xi_t \)
- MCMC algorithm discussed in paper  
- Leverage effect comes through correlated jumps
<table>
<thead>
<tr>
<th>Mode</th>
<th>95% HPD Credible Interval</th>
<th>Mode</th>
<th>95% HPD Credible Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>r</td>
<td>0.070 0.058 0.081</td>
<td>$\mu_h$</td>
<td>-1.178 -1.576 -0.660</td>
</tr>
<tr>
<td>$\phi_h$</td>
<td>0.990 0.988 0.993</td>
<td>$\sigma^2_h$</td>
<td>0.009 0.007 0.011</td>
</tr>
<tr>
<td>$\lambda_p$</td>
<td>0.022 0.013 0.035</td>
<td>$\sigma^2_\xi$</td>
<td>0.693 0.521 0.948</td>
</tr>
<tr>
<td>$\lambda_m$</td>
<td>0.011 0.004 0.020</td>
<td>$\rho_0$</td>
<td>0.002 -0.136 0.144</td>
</tr>
<tr>
<td>$\lambda_3$</td>
<td>3.102 2.384 3.953</td>
<td>$\rho_1$</td>
<td>-0.302 -0.831 -0.054</td>
</tr>
<tr>
<td>$\xi$</td>
<td>0.015</td>
<td>$\frac{1}{T} \sum \exp \frac{h}{2}$</td>
<td>0.834</td>
</tr>
<tr>
<td>T</td>
<td>14600</td>
<td># of Sim</td>
<td>200,000</td>
</tr>
</tbody>
</table>

Daily Series of Market Volatility

![Daily Series of Market Volatility](image)
Empirical Strategy

Volatility Innovations ($VI_t$) = $[h_t - \mu_h - \phi_h(h_{t-1} - \mu_h)]$

$r_{i,t} = \alpha_i + \beta_i r_{m,t} + \beta_{i,VR} VI_t + \beta_{i,TR} |J_t| + \epsilon_{i,t}$

$r_{i,t} = \alpha_i + \beta_i r_{m,t} + \gamma_i r_{m,t}^2 + \epsilon_{i,t}$

- 100 months estimation / 12 months holding
- Evaluate risk-adjusted returns during month T based on whether it is a high state.
  - Enters a high (low) state if $\xi J_{T-1} > 75\%$ (below 25\%) of past 60 months
  - (in the paper) Classification based on RV
### Coskewness-sorted Quintiles

<table>
<thead>
<tr>
<th></th>
<th>Size (1MM USD)</th>
<th>Skew</th>
<th>Idio. Skew</th>
<th>Coskew.</th>
<th>Post-ranking VR</th>
<th>TR</th>
<th>Beta</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bottom 20%</strong></td>
<td>295.21</td>
<td>0.959</td>
<td>1.045</td>
<td>-0.576</td>
<td>-0.091</td>
<td>-0.271</td>
<td>1.198</td>
</tr>
<tr>
<td><strong>Quint 2</strong></td>
<td>756.86</td>
<td>0.690</td>
<td>0.786</td>
<td>-0.223</td>
<td>-0.016</td>
<td>-0.096</td>
<td>0.996</td>
</tr>
<tr>
<td><strong>Quint 3</strong></td>
<td>1,436.11</td>
<td>0.600</td>
<td>0.669</td>
<td>-0.004</td>
<td>-0.006</td>
<td>-0.022</td>
<td>0.944</td>
</tr>
<tr>
<td><strong>Quint 4</strong></td>
<td>2,109.94</td>
<td>0.604</td>
<td>0.637</td>
<td>0.087</td>
<td>0.007</td>
<td>0.035</td>
<td>0.924</td>
</tr>
<tr>
<td><strong>Top 20%</strong></td>
<td>1,647.52</td>
<td>0.899</td>
<td>0.872</td>
<td>0.227</td>
<td>0.077</td>
<td>0.178</td>
<td>1.081</td>
</tr>
</tbody>
</table>

| 5-1            | 1,352.31       | -0.060 | -0.173     | 0.803   | 0.168           | 0.449  | -0.117 |
| T-stats        | (6.51)         | (-1.12) | (-1.01)    | (2.76)  | (3.24)          | (4.18) | (-4.85) |

**Prediction 1:** High coskewness implies either high VR, high TR or both
Portfolios Sorted by $\hat{\beta}_{VR}$ and $\hat{\beta}_{TR}$ (1962-2012)

**Prediction 2a.** The VR premium *negative* and state independent

**Prediction 2b.** The sign of the TR premium state-dependent

<table>
<thead>
<tr>
<th>Classifications</th>
<th>Portfolios Sorted by $\beta_{i,VR}$</th>
<th>Portfolios Sorted by $\beta_{i,TR}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole</td>
<td>Low State</td>
<td>High State</td>
</tr>
<tr>
<td>Bottom 20%</td>
<td>0.21%</td>
<td>0.06%</td>
</tr>
<tr>
<td>Top 20%</td>
<td>-0.13%</td>
<td>-0.15%</td>
</tr>
<tr>
<td>Top-Bottom</td>
<td>-0.34%</td>
<td>-0.21%</td>
</tr>
<tr>
<td>(-3.37)</td>
<td>(-1.07)</td>
<td>(-1.21)</td>
</tr>
</tbody>
</table>

| High-Low of (Top-Bottom) | | |
|-----------------|----------------------------------|
| | -0.06% | 0.67% |
| | (-0.18) | (3.36) |
Can Coskewness be Explained by the Two Components?

2c. State-dependency of the coskewness premium is explained by the premia of the components

- CSK is top - bottom returns of the coskewness sorted portfolios
- VR/TR is the top - bottom returns of the $\beta_{VR}$ and $\beta_{TR}$ sorted portfolios

$$\text{CSK}_t = a_0 + a_1 r_{m,t} + a_2 V R_t + a_3 T R_t + \ldots + e_t$$

<table>
<thead>
<tr>
<th>Model Dep</th>
<th>(1) CSK</th>
<th>(2) CSK</th>
<th>(3) CSK</th>
<th>(4) CSK</th>
<th>(5) CSK</th>
<th>(6) CSK</th>
<th>(7) CSK</th>
</tr>
</thead>
<tbody>
<tr>
<td>MKT</td>
<td>-0.06</td>
<td>-0.09</td>
<td>-0.06</td>
<td>-0.09</td>
<td>-0.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-1.82)</td>
<td>(-3.39)</td>
<td>(-1.91)</td>
<td>(-3.08)</td>
<td>(-2.67)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VR</td>
<td>0.40</td>
<td>0.39</td>
<td>0.40</td>
<td>0.40</td>
<td>0.40</td>
<td>0.41</td>
<td>0.41</td>
</tr>
<tr>
<td></td>
<td>(7.94)</td>
<td>(7.96)</td>
<td>(7.92)</td>
<td>(7.90)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TR</td>
<td>0.42</td>
<td>0.33</td>
<td>0.43</td>
<td>0.43</td>
<td>0.43</td>
<td>0.43</td>
<td>0.43</td>
</tr>
<tr>
<td></td>
<td>(7.20)</td>
<td>(5.21)</td>
<td>(7.06)</td>
<td>(7.13)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HML</td>
<td>-0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-0.98)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMB</td>
<td>-0.34</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-5.77)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UMD</td>
<td>0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.17)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High Dummy</td>
<td></td>
<td>0.55</td>
<td>0.10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\xi J)_{t-1}</td>
<td></td>
<td>(2.24)</td>
<td>(0.48)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adj $R^2$</td>
<td>17.23%</td>
<td>25.00%</td>
<td>28.56%</td>
<td>0.76%</td>
<td>25.10%</td>
<td>0.97%</td>
<td>25.18%</td>
</tr>
</tbody>
</table>

Dec 16th, 2014 17 / 18
Conclusion

Coskewness and volatility risk are related

1) Assets with high exposure to square of the market returns have either high exposure to
   - volatility innovations in the diffusion process or
   - stochastic jumps

2) The coskewness premium also related to the volatility risk premium
   - *Negative* premium driven by the volatility risk premium
   - *Time or state-dependency* driven by the state-transition risk premium
What comes next?
Relation between Momentum-Reversal and VR-TR

How are the VR and TR premia related to momentum (*) or long-term reversal (+)?

- Daniel and Moskowitz (2013)

  “Momentum is correlated with, but not explained by, volatility risk”

- **Positive** volatility jumps:
  - VR performs well $\rightarrow$ Negative VR premium $\rightarrow$ performs badly (+)
  - TR performs well $\rightarrow$ Positive TR premium $\rightarrow$ performs well (*)

- **Negative** volatility jumps:
  - VR performs badly $\rightarrow$ Negative VR premium $\rightarrow$ performs badly (*)
  - TR performs well $\rightarrow$ Negative TR premium $\rightarrow$ performs badly (+)
The Case of a N-State Economy

- N states s.t. state N is the highest state
- Transition Matrix $\Pi = [\pi_{s,s'}]$
  1) Only one step transition possible $\pi_{s,s'} = 0$ if $|s - s'| > 1$,
  2) Transition to higher (lower) states more likely in the low (high) states
     $\pi_{s,s+1} > \pi_{s+1,s+2}, s = 1, ..., S - 2$ and
     $\pi_{s,s-1} > \pi_{s-1,s-2}, s = 3, ..., S$

Role of the components
- VR: Higher returns in higher states when valuations are low
- TR: Provides hedge to the market portfolio if

$$\pi_{s,s+1} > \pi_{s,s-1}$$
**MCMC simulation (Eraker, Johannes and Polson, 2003)**

Steps for MCMC simulation

**Step 1.** Initialize $\Theta^{(0)}, h^{(0)}, \xi^{(0)}$ and $J^{(0)}$

**Step 2.** Sample $\Theta_j^{(1)}$ from $p(\Theta_j | \Theta_{-j}, h^{(0)}, \xi^{(0)}, J^{(0)}) \forall j$ in sequence

**Step 3.** Sample $J^{(1)}$ from $p(J | \Theta^{(1)}, h^{(0)}, \xi^{(0)})$

**Step 4.** Sample $\xi^{(1)}$ from $p(\xi | \Theta^{(1)}, h^{(0)}, J^{(1)})$

**Step 5.** Sample $h^{(1)}$ from $p(h | \Theta^{(1)}, \xi^{(1)}, J^{(1)})$

**Step 6.** Repeat Step 1 - Step 5

Sampling $h$ (highly correlated) at the same time guarantees faster convergence

- Multi-move algorithm using approximation of log-$\chi^2$ by seven normal-mixtures (Kim, Shephard and Chib 1998)
- Adds additional latent variables $a_t$ which takes values 1-7