A State Variable Model of Equity Security Risk

Dan diBartolomeo
Northfield Information
and Brunel University
New York, October 2009
Motivation for the Short Term Risk Forecasts

- Risk models for asset management (as distinct from trading operations) have traditionally focused on estimating portfolio risk from security covariance over time horizons of a year or more
  - Suitable for long term investors such as pension funds
- Investment performance of asset managers is often evaluated over shorter horizons so they are interested in shorter term risk assessment. Hedge funds and other portfolios with high portfolio turnover are even stronger in this preference
- The proliferation of high frequency trading and algorithmic execution methods have created demand for very short horizon risk assessment
A Short Chronology

- A call from Blair Hull in 1996
- diBartolomeo and Warrick (2005) in Linear Factor Models in Finance, edited by Satchell and Knight
- “Short Term Risk from Long Term Models”, Northfield research series, Anish Shah, 2007-2009
- “Equity Portfolio Risk Using Market Information and Sentiment” by diBartolomeo, Mitra and Mitra, 2009
Simple Approach to Short Term Modeling

- The usual answer:
  - Increase the frequency of observations (daily or shorter)
  - Use a shorter sample period
  - Generally need different factors

- There are serious problems with this approach at the individual security level
  - High degree of kurtosis in return distributions (well maybe?)
  - Negative serial correlation due to short term reversal effects
  - Positive serial correlation on illiquid instruments
  - Asynchronous trading across time zones makes correlation estimation very difficult

- Address “shocks” through a GARCH process
What’s the Problem with High Frequency Data?

- Financial markets are driven by the arrival of information in the form of “news” (truly unanticipated) and the form of “announcements” that are anticipated with respect to time but not with respect to content.

- The time intervals it takes markets to absorb and adjust to new information ranges from minutes to days. Generally much smaller than a month, but up to and often larger than a day. That’s why US markets were closed for a week at September 11th.

- GARCH models don’t work well on announcements
  – Market participants anticipate announcements
  – Volume and volatility dry up as investors wait for outcomes
  – Reduce volatility into the announcement and boost it after the announcement, so they are wrong twice
An Example of Anticipation?

- Lets look at a precipitous decline in the implied volatility of options on LUV
  - All days in 2001 prior to September 7, average of .45 with a s.d. of .13
  - September 7, LUV implied = .22
  - September 10, LUV implied = .15
  - All days subsequent to September 17, average of .54 with a s.d. of .18
  - September 10 is in bottom 1% of the universe in implied volatility, September 17 is 91st percentile

- Could this be driven by fundamentals?
Investor Response to Information

- Several papers have examined the relative market response to “news” and “announcements”
  - Ederington and Lee (1996)
  - Kwag Shrieves and Wansley (2000)
  - Abraham and Taylor (1993)
- Jones, Lamont and Lumsdaine (1998) show a remarkable result for the US bond market
  - Total returns for long bonds and Treasury bills are not different if announcement days are removed from the data set
- Brown, Harlow and Tinic (1988) provide a framework for asymmetrical response to “good” and “bad” news
  - Good news increases projected cash flows, bad news decreases
  - All new information is a “surprise”, decreasing investor confidence and increasing discount rates
  - Upward price movements are muted, while downward movements are accentuated
Our Approach is Different

• Continue to use the existing risk models that are estimated from low frequency return observations

• Use new information that is not part of the risk model to adjust various components of the risk forecast to short-term conditions
  – Just ask yourself “How are conditions different now than they were on average during the sample period used for estimation?”

• This approach has multiple benefits
  – We sidestep almost all of the statistical complexities that arise with use of high frequency data
  – We get to keep the existing factor structure of the model so risk reporting remains familiar and intuitive
  – Since our long term and short term forecasts are based on the same factor structure, we can quickly estimate new forecasts for any length time horizon that falls between the two horizons
  – Can be applied to any of our existing models
One Form of Working with “External Information”

- Risk estimates in our short term model of US equities have been conditioned for years based on analysis of stock option implied volatility
  - Every day we look at the implied volatility of options on all US stocks
  - We keep a 30 day moving average of the ratio of implied volatility to historic volatility
  - If the implied volatility/historic ratio jumps because of an information flow to the market (e.g. Bill Gates gets run over by a bus), the specific risk of that stock is adjusted
  - If implied volatility ratio of many related stocks changes, the implied changes in factor variance are also made. This means that the risk forecasts change even for stocks on which no options trade
  - See Chapter 12, *Linear Factor Models in Finance*, Satchell and Knight, editors
“Variety” as External Information

- Solnik and Roulet (2000) examine the dispersion of country returns as a way of estimating correlations between markets.

- Lilo, Mantegna, Bouchard and Potters use the term Variety to describe cross-sectional dispersion of stock returns.
  - They also define the cross-sectional dispersion of CAPM alpha as idiosyncratic variety (noted as $v(t)$).
  - They find that the average correlation between stocks is approximately:

$$ C(t) = \frac{1}{1 + \left( \frac{v^2(t)}{r_m^2(t)} \right) } $$

- diBartolomeo (2000) relates periods of high cross-sectional dispersion to positive serial correlation in stock returns (i.e. momentum strategies working).
Other Conditioning Information

• Estimates of volatility based on high/low/open/close information instead of the dispersion of returns
  – Parkinson, Garman-Klass, Satchell-Wang, etc.

• Yield spreads for different classes of fixed income securities provide an implied default rate and the potential for large negative skew in stock returns

• Implied distribution of asset returns given the implied vols of options on market indices across strike prices

• Number, length and content of text articles on Dow Jones, Reuters and Bloomberg. *Score the content on a lexicon of 2000 keywords and phrases in real time*

• Numbers of internet searches on ticker symbols via Google Trends, Yahoo Finance
Differences Between Long and Short Horizon Risk

- **Negative serial correlation**
  - Daily overreactions & reversals, which cancel out over time, become significant e.g. under leverage

- **Contagion / panic**
  - Liquidity demands can drive up short-term correlations

- **Transient behavior**
  - A long term model intentionally integrates new phenomena slowly: Is the future like the past or are we in and concerned about a present shift?

- **Lots more extreme events in the short-term**
  - 3 std deviations contains less probability mass. 99% VaR is farther away from the mean
Generalizing the Idea

- Take any risk model. e.g. one of our models estimated using monthly observations of returns and security attributes.

- Add flexibility points and fit to information about current conditions.

- Adjust for statistical differences between short and long term returns.

- Many benefits of this approach:
  - Avoids statistical complexities of high frequency data.
  - Keeps with familiar factor structure.
  - Common factor structure for long and short horizons permits interpolating any horizon in between.
  - Works with any factor model.
Statistical Differences: An Extreme Example of Serial Correlation

- Big drops in Aug 2007
  - 8/3: SP500 ↓ 2.7%, R2000 ↓ 3.6%
  - 8/6-8: SP500 ↑ 4.5%, R2000 ↑ 5.3%
  - 8/9: SP500 ↓ 3.0%, R2000 ↓ 1.4%
  - 8/10: SP500 ↑ 0.0%, R2000 ↑ 0.5%
  - 8/13-15: SP500 ↓ 3.2%, R2000 ↓ 4.7%
  - 8/16-17: SP500 ↑ 2.8%, R2000 ↑ 4.5%
  - 8/27-28: SP500 ↓ 3.2%, R2000 ↓ 3.9%
  - 8/29: SP500 ↑ 2.2%, R2000 ↑ 2.5%

- Over the month, SP500 up 1.3%, R2000 up 2.2%
Observing Typical Serial Correlation

- Say 21 trading days in a month
  \[ r_t = \text{return on day } t \]
  \[ r_{\text{month}} = \prod_{t=1\ldots21} (1+r_t) - 1 \]
  \[ r_{\text{21 day sum}} = \sum_{t=1\ldots21} r_t (\text{a month w/o compounding}) \]

- \[ \text{var}(r_{\text{21 day sum}}) = \sum_{t=1\ldots21} \text{var}(r_t) + 2 \sum_{s<t} \text{cov}(r_s,r_t) \]

- If stationary and no covariance between days,
  \[ \text{var}(r_{\text{21 day sum}}) = 21 \times \text{var}(r_{\text{1 day}}) \]

annualized: \[ \frac{250}{21} \times \text{var}(r_{\text{21 day sum}}) = 250 \times \text{var}(r_{\text{1 day}}) \]
(Annualized) 1 day vol > 1 month vol:

Negative Serial Correlation


- daily
- monthly
- sum of 21 daily

Bar chart showing the annualized variance of % return for various indices and currencies from September 2002 to September 2007.
Adjusting for Serial Correlation

- Goal: the variance inflation introduced by negative serial correlation

- Express cumulative return from time 0 → time T in terms of 1 day returns:
  \[ r_{0\rightarrow T} = f (r_1, \ldots, r_T) = \prod_{t=1}^T (1 + r_t) - 1 \]

- Approximate linearly around mean return, \( \mu \)
  \[ r_{0\rightarrow T} \approx f|_{r_1\ldots r_T = \mu} + \nabla f|_{r_1\ldots r_T = \mu} (r - \mu) = \text{constant} + (1 + \mu)^T - 1 \sum_{t=1}^T (r_t - \mu) \]

- Yields cumulative return variance as 1 day return covariances
  \[ \text{var}(r_{0\rightarrow T}) \approx (1 + \mu)^{2T-2} \sum_{1 \leq j,k \leq T} \text{cov}(r_j, r_k) \]
Adjusting for Serial Correlation (2)

- Suppose 1 day returns follow an AR(1)
  \[ r_{t+1} = c + \rho r_t + \varepsilon_{t+1} \]
  \[ \text{same var}(r_t) \text{ every day, } \sigma^2 \]
  \[ \text{n day apart correlation } = \rho^n \]

- Recall previous expression for variance of cumulative return
  as sum of covariances of 1 day returns
  \[ \text{var}(r_{0 \rightarrow T}) \approx (1 + \mu)^{2T-2} \sum_{1 \leq j,k \leq T} \text{cov}(r_j, r_k) \]

- Substitute AR(1) covariances
  Result is in terms of 1 day variance and 1 day apart correlation
  \[ \text{var}(r_{0 \rightarrow T}) \approx (1 + \mu)^{2T-2} \sum_{1 \leq j,k \leq T} \sigma^2 \rho^{|j-k|} \]
  \[ = (1 + \mu)^{2T-2} \sigma^2 \left[ T(1 + \rho)/(1 - \rho) - 2\rho(1 - \rho^T)/(1 - \rho)^2 \right] \]
Adjusting for Serial Correlation (3)

- Now, can project a forecast from one horizon to another by solving out $\sigma^2$ from approximations for $\text{var}(r_{0\rightarrow T_1})$, $\text{var}(r_{0\rightarrow T_2})$
  - e.g. 1 month to 1 week, $T_1=21$, $T_2=5$
    - 1 month to 1 day, $T_1=21$, $T_2=1$

- $\text{var}(r_{0\rightarrow T}) \approx (1 + \mu)^{2T-2} \sigma^2 \left[ T(1 + \rho)/(1 - \rho) - 2\rho(1 - \rho^T)/(1 - \rho)^2 \right]$

- $\sigma^2 \approx \text{var}(r_{0\rightarrow T}) / \{(1 + \mu)^{2T-2} \left[ T(1 + \rho)/(1 - \rho) - 2\rho(1 - \rho^T)/(1 - \rho)^2 \right]\}$

- $T_1/T_2 \text{ var}(r_{0\rightarrow T_2}) \approx \text{var}(r_{0\rightarrow T_1}) (1 + \mu)^{2(T_2 - T_1)}$ 
  $\times \left[ (1 - \rho^2) - 2\rho(1 - \rho^T)/T_2 \right] / \left[ (1 - \rho^2) - 2\rho(1 - \rho^T)/T_1 \right]$
What value for $\rho$?

<table>
<thead>
<tr>
<th>9/02 - 9/07</th>
<th>Variance Ratio: Annualized Daily/Monthly</th>
<th>Measured Serial Correlation of 1 Day Returns</th>
<th>Serial Correlation Implied by Variance Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>S&amp;P 500</td>
<td>2.1</td>
<td>-0.11</td>
<td>-0.36</td>
</tr>
<tr>
<td>NASDAQ</td>
<td>1.4</td>
<td>-0.04</td>
<td>-0.17</td>
</tr>
<tr>
<td>DOW</td>
<td>1.7</td>
<td>-0.10</td>
<td>-0.27</td>
</tr>
<tr>
<td>NIKKEI 225</td>
<td>1.6</td>
<td>0.00</td>
<td>-0.24</td>
</tr>
<tr>
<td>FTSE 100</td>
<td>2.2</td>
<td>-0.13</td>
<td>-0.38</td>
</tr>
<tr>
<td>CAC 40</td>
<td>1.9</td>
<td>-0.04</td>
<td>-0.32</td>
</tr>
<tr>
<td>DAX</td>
<td>1.4</td>
<td>-0.08</td>
<td>-0.19</td>
</tr>
<tr>
<td>$/$€</td>
<td>1.1</td>
<td>-0.01</td>
<td>-0.05</td>
</tr>
<tr>
<td>$/$¥</td>
<td>1.4</td>
<td>-0.02</td>
<td>-0.18</td>
</tr>
<tr>
<td>$/$£</td>
<td>1.2</td>
<td>0.02</td>
<td>-0.09</td>
</tr>
</tbody>
</table>
Bending to Fit Information

- Make model flexible by adding free parameters \( \theta \)

- Example 1: In our global model, let
  \[
  \begin{align*}
  \theta_1 & \ldots \theta_5 \quad \text{scale std dev of the 5 region factors} \\
  \theta_6 & \quad \text{scale std dev of the value/growth factor} \\
  \theta_7 & \quad \text{scale std dev of all the remaining factors} \\
  \theta_8 & \quad \text{scale stock specific std dev for U.S. & E.U. stocks} \\
  \theta_9 & \quad \text{scale stock specific std dev for all other stocks} \\
  \theta_1 & \ldots \theta_9 \geq 0
  \end{align*}
  \]

- Example 2: Correlation tightening/loosening
  correlation matrix = \( \theta_1 \times (11^T) + \theta_2 \times 1 + (1 - \theta_1 - \theta_2) \times \text{original} \)
  \( \theta_1, \theta_2 \geq 0, \quad \theta_1 + \theta_2 \leq 1 \)
Bending to Fit Information (2)

- For each observed statistic (piece of market information)
  e.g. option implied variance of S&P 500
    \( \hat{g} \) = recent value
    \( \hat{g}_{\text{avg}} \) = average value over estimation period

- Find a related statistic that can be predicted by the model
  e.g. variance of S&P 500 portfolio
    \( g_\theta \) = predicted value under model adjusted by \( \theta \)
    \( g \) = predicted value under model

- Choose \( \theta \) to make the increase in the forecast, \( (g_\theta/g) \), match the increase in market information, \( (\hat{g}/\hat{g}_{\text{avg}}) \)
Examples of Global Information

- \( \hat{g} \) = implied variance from options on SP500 index
  \( g \) = variance of SP500
  (a measure of volatility that includes correlation)

- \( \hat{g} \) = average range variance estimates of EAFE constituents
  \( g \) = average variance of EAFE constituents
  (a measure of volatility separate from correlation)

- \( \hat{g} \) = cross-sectional variance of Nikkei 225
  \( g \) = xc var (ignoring difference in means) of Nikkei 225

- \( \hat{g} \) = median squared spread between 2 baskets of securities
  \( g \) = variance of long/short portfolio of same securities
Bayesian Framework

- $p(\theta) = \text{prior distribution on parameters } \theta$
- $p(\hat{g}) = \text{unconditional probability of observing statistics } \hat{g}$

- $p(\hat{g}|\theta) = \text{probability of observations } \hat{g} \text{ given parameters } \theta$
- $p(\theta|\hat{g}) = \text{probability that parameters are } \theta \text{ given observations } \hat{g}$

Maximum a posteriori (MAP) estimate of $\theta$ given $\hat{g}$

- $\theta_{\text{MAP}} = \arg\max_{\theta} p(\theta|\hat{g}) = p(\hat{g}|\theta) \times p(\theta) / p(\hat{g})$
- $\theta_{\text{MAP}} = \arg\max_{\theta} \log p(\theta|\hat{g}) = \log p(\hat{g}|\theta) + \log p(\theta) - \log p(\hat{g})$
- $\theta_{\text{MAP}} = \arg\max_{\theta} \log p(\hat{g}|\theta) + \log p(\theta)$
Making the Intuition Rigorous (2)

• Recall $\theta_{\text{MAP}} = \text{argmax}_\theta \log p(\hat{g}|\theta) + \log p(\theta)$

• Assume observations $\hat{g}$ are noisy observations of predictions $g_\theta$:
  $\hat{g} = g_\theta + \varepsilon$
  $p(\hat{g}|\theta) = p(\hat{g}|g_\theta) = p(\varepsilon = \hat{g} - g_\theta)$

• The distribution of noise determines the fit criterion:
  $\varepsilon_k \sim \text{Gaussian}[0, \sigma_k^2]$ \quad $\rightarrow \log p(\hat{g}|\theta) = \sum_k [\hat{g}_k - g_\theta^k]^2 / 2\sigma_k^2 + \text{const}$
  $\varepsilon_k \sim \text{Laplace}[0, b_k]$ \quad $\rightarrow \log p(\hat{g}|\theta) = \sum_k |\hat{g}_k - g_\theta^k| / b_k + \text{const}$
Summarizing the Framework

- Add parameters to model and tune so increases in predictions match recently observed increases.

- Scale long horizon variance #’s to account for the reversing swings that occur day to day.

- Be aware that short term returns follow a different distributional shape. Extreme events are more frequent.
Tuning A Global Model to Current Conditions

- Factors:
  - Sector, Region, Economic Variables, Pricing Spreads, Blind Factors, Currencies

- Add parameters to make model flexible
  - $\theta_1 \times \text{std dev of E.U. & Scandinavian currencies}$
  - $\theta_2 \times \text{std dev of all other currencies}$
  - $\theta_3 \times \text{std dev of sector factors}$
  - $\theta_4 \times \text{std dev of stock specific risk}$
  - $\theta_5 \times \text{std dev of all remaining factors}$
The Tuned Model

- **Adjusted volatility predictions**

<table>
<thead>
<tr>
<th>Predicted</th>
<th>Data</th>
<th>Observed Relative Level</th>
<th>In Tuned Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>S&amp;P 500</td>
<td>Hi/Lo</td>
<td>2.06</td>
<td>1.93</td>
</tr>
<tr>
<td>Avg of S&amp;P 500 stocks</td>
<td>VIX</td>
<td>1.72</td>
<td>1.72</td>
</tr>
<tr>
<td>Nikkei 225 (¥)</td>
<td>Hi/Lo</td>
<td>1.31</td>
<td>1.32</td>
</tr>
<tr>
<td>S&amp;P Eur 350 ($)</td>
<td>IShare</td>
<td>1.73</td>
<td>1.86</td>
</tr>
<tr>
<td>€ (in $)</td>
<td>Hi/Lo</td>
<td>0.94</td>
<td>0.93</td>
</tr>
<tr>
<td>¥ (in $)</td>
<td>Hi/Lo</td>
<td>1.63</td>
<td>1.63</td>
</tr>
</tbody>
</table>

- **Changes to the model’s factors**

<table>
<thead>
<tr>
<th>Factor</th>
<th>Dilation Applied to Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>E.U. &amp; Scandinavian currencies</td>
<td>0.93</td>
</tr>
<tr>
<td>All other currencies</td>
<td>1.63</td>
</tr>
<tr>
<td>Sector factors</td>
<td>2.14</td>
</tr>
<tr>
<td>Stock specific risk</td>
<td>1.78</td>
</tr>
<tr>
<td>All remaining factors</td>
<td>0.65</td>
</tr>
</tbody>
</table>
Applying the Serial Correlation Adjustment

- Recall \( T_1/T_2 \ var(r_{0 \rightarrow T_2}) \approx var(r_{0 \rightarrow T_1}) \ (1 + \mu)^{2(T_2 - T_1)} \times \left[ (1 - \rho^2) - 2\rho(1 - \rho^{T_2})/T_2 \right] / \left[ (1 - \rho^2) - 2\rho(1 - \rho^{T_1})/T_1 \right] \)

- Suppose \( \rho = -.30, \mu = 0 \)
  - Note: can apply a different serial correlations to each factor and, by security, to stock specific risk

- Monthly \( \rightarrow \) weekly std dev adjustment (\( T_1 = 21, T_2 = 5 \))
  \( = 1.05 \)

- Monthly \( \rightarrow \) daily std dev adjustment (\( T_1 = 21, T_2 = 1 \))
  \( = 1.34 \)
Conclusion

• Two major pieces to this framework
  – Training to fit current market information
  – Adjusting for distributional differences and serial properties

• The principles are general and work with any risk model (and many other types of modeling problems. The approach was motivated by the style of models used in computer vision.)

• Allow us to build a model of security covariance that evolves rapidly as market evolve but does not require estimation with high frequency data that is statistically problematic