Modeling Risk in Arbitrage Strategies Using Hierarchical Mixture Models

Robert J. Frey, Research Professor and Director
Program in Quantitative Finance
Applied Mathematics and Statistics
Stony Brook University
Abstract

A general class of models is developed that captures the state-dependent character of arbitrage-based hedge fund strategies. Using merger arbitrage as an example, and specific instance is described and its MLE is illustrated.

This modeling approach provides a better description of real hedge fund returns in a framework that is more general and more easily extensible than other alternatives in the open literature.
Benefits

• Economically Rational Foundation
• Improved Signal-To-Noise
• Systematic Hedge $\Rightarrow$ Bet Independence
• Decomposition $\Rightarrow$ Simpler Models
Dangers

• Inefficiencies Small and Transient

• Leverage Required $\Rightarrow$ Hedge Instability

• Model Failure $\Rightarrow$ Catastrophic Returns
Net Result

- Strategies produce stable but modest returns punctuated by intervals of dramatically poor performance.
- Failure occurs when leveraged, ostensibly independent bets behave in ways that are pathological and highly correlated.
HFR Meger Arb vs. S &P 500

Monthly: 1990-2001
HFR Meger Arb vs. S &P 500
Monthly: 1990-2001
Linear Model
Forecast vs. Actual
Mitchel & Pulvino’s Kinked Model
Extended to optimal least squares fit
Generalized r-Squared

\[ r_{\text{generalized}}^2 = 1 - \frac{\sum_{j=1}^{n} (y_j - \hat{y}_j)^2}{\sum_{j=1}^{n} (y_j - \text{E}[y])^2}/(n - k) \]
Parameterized Mixture

\( p[\mathbf{u} | \alpha_1, ..., \alpha_{m-1}] \)

\[ g_1[\mathbf{v} | \beta_1] + \varepsilon_1 \]

\[ g_i[\mathbf{v} | \beta_i] + \varepsilon_i \]

\[ g_m[\mathbf{v} | \beta_m] + \varepsilon_m \]
PFMC Return Model

\[ r = p[u | \alpha_1, \ldots, \alpha_{m-1}]^T(g[v | \beta_1, \ldots, \beta_m] + \phi[\varepsilon | \Omega]) \]

\[ f[r, u, v | \Psi] = \sum_{i=1}^{m} p_i[u | \alpha_i] \phi[r - g_i[v | \beta_i] | \Omega_i] \]
Log Likelihood

\[ \Lambda = \sum_{i=1}^{n} \log \left[ \sum_{j=1}^{m} p_i[u_j | \alpha_i] \phi[r_j - g_i[v_j | \beta_i] | \Omega_i] \right] \]

\[ \Lambda_C = \sum_{i=1}^{m} \sum_{j=1}^{n} z_{i,j} \left( \log \left[ p_i[u_j | \alpha_i] \right] + \log \left[ \phi[r_j - g_i[v_j | \beta_i] | \Omega_i] \right] \right) \]
LMLC Model

\[ p[\alpha^T u] = \begin{pmatrix} p_1 = \frac{e^{\alpha^T u}}{1 + e^{\alpha^T u}} \\ p_2 = \frac{1}{1 + e^{\alpha^T u}} \end{pmatrix} \]

\[ g_i[v | \beta_i] = \beta_i^T v, \quad i = 1, 2 \]
MLE

\[ \Lambda_C \propto \sum_{j=1}^{n} \left( z_{1,j} \log \left[ \frac{e^{\alpha^T u_j}}{1 + e^{\alpha^T u_j}} \right] + z_{2,j} \log \left[ \frac{1}{1 + e^{\alpha^T u_j}} \right] \right) - \]

\[ \frac{1}{2} \sum_{j=1}^{n} \left( z_{1,j} \left( \log[\sigma_1^2] + \frac{(r_j - \beta_1^T v_j)^2}{\sigma_1^2} \right) + z_{2,j} \left( \log[\sigma_2^2] + \frac{(r_j - \beta_2^T v_j)^2}{\sigma_2^2} \right) \right) \]

Estimated using the EM Algorithm
Mixture Function

-0.4 -0.2 0.2 0.4 1
0.8 0.6 0.4 0.2

16
LMLC Single Factor Model

S & P 500
Single Factor Comparison
LMLC Single Factor Errors
Forecast vs. Actual
LMLC Multi-Factor Errors
S & P 500, Lagged S & P 500, Credit Spreads
Linear Multi-Factor Errors

*Fit per Mitchell [1999], Rahl [2000] and Schneeweis [2000]*
## Model Performance Summary

<table>
<thead>
<tr>
<th>Model</th>
<th>$r^2_{\text{generalized}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single Factor Linear</td>
<td>0.188</td>
</tr>
<tr>
<td>Multi-Factor Linear</td>
<td>0.298</td>
</tr>
<tr>
<td>Kinked Single Factor</td>
<td>0.334</td>
</tr>
<tr>
<td>Single Factor LMLC</td>
<td>0.298</td>
</tr>
<tr>
<td>Multi-Factor LMLC</td>
<td>0.428</td>
</tr>
</tbody>
</table>
References


