Bayesian Time Series Analysis of Stock Selection Strategies

Omar Aguilar

ING Investment Management
Quantitative Equity Research
We discuss a general class of Bayesian Dynamic Linear Models for multivariate financial time series. The predictive alpha model attempts to exploit the predictability of the behavior of multiple non-parametric stock selection ranking techniques. The risk component includes the formulation of dynamic factor models with stochastic volatility components for the residual covariance matrix and represents specific varieties of models recently discussed in the growing multivariate stochastic volatility literature. Bayesian inference and computation is developed and explored in a study of the dynamic structure of long/short strategies that shift exposure among various stock selection criteria. We review empirical findings in applying these models to a universe of large capitalization stocks in the US including aspects of model performance in dynamic portfolio allocation. We discuss model assessment, and computational algorithms developed to fit this new class of models and conclude with comments on future potential developments together with model extensions.
Outline

- Investment Management
- Bayesian Forecasting Models for Equity Portfolios
- Stock Selection Strategies
- Summary and Conclusions
Traditional Investment Process:

1. Investment Objectives
   - annualized returns, liabilities, risks, taxes, time horizon
2. Asset Classes
   - equities, bonds, cash, international, alternatives
3. Predictive Return Distributions
4. Portfolio Construction
   - maximize expected utility, constraints, penalties
5. Execution
   - trading, transaction costs, derivatives, mutual funds, other
Stock Analysis

Return distributions

IBM ACF

Lag

IBM (%)
Challenges in Stock Selection

- Information flow is fairly efficient
- Lack of autocorrelation structure
- Ill-behaved, non-gaussian return distributions
- Stock portfolios are traditionally high dimensional
- High levels of signal to noise ratio especially in high frequencies
- Forecast errors are painful...
Dynamic APT Model

Traditional Models are factor driven

\[
\mathbf{r}_t = \mathbf{\alpha}_t + \mathbf{E}_t \mathbf{\beta}_t + \mathbf{e}_t, \quad \mathbf{e}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{S}_t)
\]

where:

- \( \mathbf{r}_t \) is an \( N \) vector of company one-step-ahead returns:
  \[
  r_{jt} = \log(\frac{P_{jt}}{P_{j,t-1}})
  \]
- \( \mathbf{\alpha}_t \) is an \( N \) vector of company specific abnormal returns \( \alpha_{jt} \)
- \( \mathbf{\beta}_t \) is an \( N \) vector of company specific factor sensitivities \( \beta_{jt} \)
- \( \mathbf{E}_t \) is an \( T \times N \) matrix of KNOWN factor payoffs or returns
- \( \mathbf{\beta}_t \) have a random walk evolution
Large Cap Universe Returns: December 2004

S&P stock returns 12/2004

Quantiles of Standard Normal

Rationale
Stock Selection in Large Cap Stocks

Issues:

- Most factor payoffs are calculated from large universes (BARRA)
- Traditionally, predictive power is limited, finding $\alpha_t \geq 0$ is challenging
- Similar stocks trade on similar fundamental and market information

Data:

- Monthly returns from the largest 1000 companies in the US from January 1985 to September 2004

Goals:

- Classification variables could identify extreme positive (negative) performers
- Create clusters of companies with similar characteristics
- Define portfolios with the highest posterior probability of extreme performance (Anomalies)
- Create stock selection strategies and hence profitable equity portfolios
Classification Variables

**Growth/value:** Value stocks (high book/price, E/P, CF/P) outperform growth stocks (low B/P, E/P, CF/P)


**Post-earnings-announcement drift:** Stocks with earnings that beat expectations outperform stocks that miss expectations


**Short-term price reversal:** One-month losers outperform one-month winners

- Jegadeesh (1990), Lo and MacKinlay (1990)

**Intermediate-term price momentum:** Six-month to one-year winners outperform losers


**Earnings quality:** Stocks with cash earnings outperform stocks with noncash earnings


**Analyst Earnings Revisions:** Changes in analyst stock recommendations and earnings estimates predict subsequent stock returns

- Francis and Soffer (1997), Barber, Lehavy, McNichols, and Trueman (2001)
Classification Variables: December 2004

Market Cap

Earnings

Cash Flow

Momentum
Classification Variables: Clusters

Clusters:

- Define $M$ clusters for each classification variable
- Let $r_{it}^{(m)}$ be the vector of returns from cluster $m$ and variable $i$
- For each cluster/variable
  \[ r_{it}^{(m)} | \mu_{it}^{(m)}, s_{it}^{(m)} \sim [\mu_{it}^{(m)}, s_{it}^{(m)}] \]

Applications:

- Find the marginal posterior distribution of $\mu_{it}^{(m)}$
- $\mu_{it}^{(m)}$ represents the return of a portfolio of stocks in cluster $m$ using variable $i$
- In theory the best opportunities will be available in clusters 1 and $M$
- Simple strategy would be to buy stocks in cluster 1 and to sell stocks in cluster $M$
Clusters

Posterior Mean Return

Market Cap
Earnings
Cash Flow
Momentum

Clusters

1 2 3 4 5

1.0 1.5 2.0 2.5 3.0 3.5 4.0 4.5

Posterior Analysis: December 2004
Cluster portfolios imply random effects model for each stock

\[ r_{jt} = \mu_{it} + e_{jt} \]

where

\[ \mu_{it} = \sum_{r=1}^{R} W_{irt}^{(m)} \mu_{irt}^{(m)} I[r_{jit}] \]

Notes:

- for \( r = 1, \ldots, R \) variables
- There are a total of \( R \times M \) means
- The resulting model is a refined version of traditional alpha factor models
- \( W_{irt}^{(m)} \) are the corresponding sensitivities
- The next step is to find the optimal combination of \( W_{irt}^{(m)} \)
Anomalies

Data:

• Monthly values of nine variables from January 1985 to September 2004

• Anomalies

\[ y_{rt} = \mu_{rt}^{(1)} - \mu_{rt}^{(M)} \]

for \( r = 1, \ldots, R \)

Objectives:

• Dynamic regressions to estimate and forecast expected returns of anomalies

• Explore patterns of variability and residual structure over time via modeling \( \Sigma_t \)

• Find potential sources of systematic risk that drive changes in variances and correlations between anomalies

• Develop stock selection strategies by rotating weights on different anomalies via sequential updating
Anomaly Returns

Surprise

Revision

Momentum

Reversal

Quality

Cashtoprice

Relativeval

Growth

Earningsyld
Dynamic Linear Models

Dynamic SUR Model,

\[ y_t = F_t \theta_t + \nu_t, \quad \nu_t \sim N(0, \Sigma_t), \]
\[ \theta_t = \theta_{t-1} + \omega_t, \quad \omega_t \sim N(0, W_t). \]

where:

- \( y_t \) is a \( r \times 1 \) vector of anomaly returns
- \( F_t \) is the regression matrix, a block diagonal matrix with macro-economic and market variables
- \( \theta_t \) is a \( p \times 1 \) vector of regression coefficients
- References include West and Harrison (1997), Zellner, Hong and Min (1995) and Quintana, Chopra and Putnam (1991)
## Explanatory Variables

<table>
<thead>
<tr>
<th>Macro-economic</th>
<th>Market</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inflation</td>
<td>Seasonality</td>
</tr>
<tr>
<td>Short Interest Rates</td>
<td>Profitability</td>
</tr>
<tr>
<td>Yield Curve</td>
<td>Volatility</td>
</tr>
<tr>
<td>Term Structure</td>
<td>Large/Small</td>
</tr>
<tr>
<td>Long Interest Rates</td>
<td>Earnings Growth</td>
</tr>
<tr>
<td>GDP Growth</td>
<td>Value/Growth</td>
</tr>
</tbody>
</table>
Dynamic Factor Models

Special case of Aguilar and West (2000) dynamic factor model for $\Sigma_t$, 

$$\Sigma_t = XH_tX' + \Psi_t$$

- $\nu_t = y_t - F_t\theta_t$
- $X$ is the $r \times k$ factor loadings matrix
- $f_t$ is a $k \times 1$ vector of conditionally independent latent common factors
- $\epsilon_t$ is a $r \times 1$ vector of series-specific quantities
- $\Psi_t = \text{diag}(\psi_{1t}, \ldots, \psi_{rt})$ instantaneous “idiosyncratic” variances and 
- $\epsilon_t$ and $f_s$ are mutually independent for all $t, s$

The observation equation is rewritten as $y_t = F_t\theta_t + Xf_t + \epsilon_t$
Multivariate SV models for the latent factor processes \( f_t \sim N(0, H_t) \)

- For \( z_{it} = \log(f_{it}^2) = \lambda_{it} + \nu_{it} \) and \( \lambda_{it} = \log(h_{it}) \) we assume **latent VAR(1) models**

\[
\begin{align*}
\mathbf{z}_t &= \lambda_t + \nu_t \\
\lambda_t &= \mu_t + \gamma_t \\
\text{where} \quad \mu_t &= \mu_{t-1} + \eta_t \\
\gamma_t &= \Phi \gamma_{t-1} + \omega_t
\end{align*}
\]

with correlated innovations across factors \( \omega_t \sim N(\omega_t|0, \mathbf{U}) \)

- These relationships imply

\[
\lambda_t = \mu_t + \Phi(\lambda_{t-1} - \mu_{t-1}) + \omega_t
\]
Model Fitting and Bayesian Analysis

Inference based on a fixed sample over $t = 1, \ldots, T$

- Bayesian analysis via posterior simulations (GIBBS-Metropolis)
- MCMC samples from the joint posterior for
  - Regression parameters, dynamic factor model parameters and
  - latent processes: factors and volatilities
    
    \[
    \{ f_t, \lambda_t, \psi_{it} : t = 1, \ldots, T \}
    \]

Sequential Analysis over $t = T + 1, T + 2, \ldots$

- **Sequential Particle Filtering** to update posterior samples
  
  \[
  \cdots \rightarrow p(\cdot|D_{t-1}) \rightarrow p(\cdot|D_t) \rightarrow \cdots
  \]

- and compute/revise predictive distributions
  
  \[
  \cdots \rightarrow p(y_t|D_{t-1}) \rightarrow p(y_{t+1}|D_t) \rightarrow \cdots
  \]
Forecast Errors

Surprise

Revision

Momentum

Reversal

Quality

Cashtoprice

Relativeval

Growth

Earningsyld
Stock Selection Strategies

Let $\mathbf{a}_t = (a_{1t}, a_{2t}, \ldots, a_{qt})'$ be the one-step ahead portfolio where $a_{it}$ is weight in the $i$-th anomaly

- Portfolio return at time $t$, $r_t = \mathbf{a}_t' \mathbf{y}_t$

- **Forecast:** predictive distributions $p(\mathbf{y}_t | D_{t-1})$ with $\mathbf{g}_t$ and $\mathbf{G}_t$ denoting the corresponding predictive mean and covariance matrix

**Portfolio Construction:**

- minimize risk: $\mathbf{a}_t' \mathbf{G}_t \mathbf{a}$
  - for a specified risk aversion level
  - anomalies determine portfolios including transaction costs
  - implementation through stock “baskets”
Anomaly Weights

Surprise

Revision

Momentum

Reversal

Quality

Cashtoprice

Relativeval

Growth

Earningsyld
Summary and Conclusions

- Anomaly driven strategies represent great investment opportunities
- Bayesian DLMs adapt to structural changes in capital markets
- Anomaly forecasts are indicators of different faces of the profit cycle in equity markets
- Identification of sources of systematic risk
- Improvements in short-term forecasting for anomalies may suggest industry and sector rotation strategies
- Appealing information ratios for active managers
- Novel and customized MCMC algorithms for fitting and computation of posterior and predictive analysis
- Generalizations include, multiple clustering variables, CART models, number of clusters, stock specific volatilities and variable selection