Predictability of Stock Returns and Asset Allocation under Structural Breaks

D. Pettenuzzo¹  A. Timmermann²

¹Bocconi University
²Dept. of Economics and Rady School of Management, UC San Diego

March 09 2007 / Duke
Current Return predictability debate

- Pesaran and Timmermann (1995)
- Bossaerts and Hillion (1999)
- Lettau and Ludvigsson (2001)
- Goyal and Welch (2003)
- Cooper, Gutierrez and Marcum (2005)
- Campbell and Thompson (2005)
- Cochrane (2006)

Is the weak out-of-sample predictability due to model instability ("breaks")?
Evidence of structural breaks

Weakness of out-of-sample return predictability has been linked to model instability

- Schwert (2003)
- Perez-Quiros and Timmermann (2000)
- Paye and Timmermann (2005)
- Lettau and Van Nieuwerburgh (2005)
- Goyal and Welch (2003, 2007)
Pastor and Stambaugh (2001, p. 1207) “Finance practitioners and academics often elect to rely on more recent data ... motivated in part by concerns that the probability distribution of excess returns changes over time, experiencing shifts known as “structural breaks””

- When did the break occur? (Bai, Bai-Perron)
- How much data to use?
- inefficient only to use post-break data (Pesaran and Timmermann (2006, 2007))
- Can breaks happen in the future?
Methodology

Changepoint model driven by unobserved discrete state variable
After a break, the new parameters of the return forecasting model are drawn from a meta distribution

→ we can forecast out-of-sample even in the presence of future breaks

Approach accounts for structural breaks in return forecasting models

- Chib (1998)
- Pastor and Stambaugh (2001)
- Pesaran, Pettenuzzo and Timmermann (2006)
Return Prediction Model

- Restricted VAR:
  \[ z_t = B' \tilde{x}_{t-1} + u_t \]

- \[ z_t = (r_t, x_t)' \], \[ \tilde{x}_{t-1} = (1, x_{t-1})' \]
  - \( r_t \): excess return at time \( t \)
  - \( x_{t-1} \): return predictor(s)
  - \( u_t \sim N(0, \Sigma) \)

- \( \mu_r, \mu_x \): intercepts in the return and predictor equation
- \( \beta_r, \beta_x \) coefficients on lagged predictor:

  \[ r_t = \mu_r + \beta_r x_{t-1} + \varepsilon_{rt} \]
  \[ x_t = \mu_x + \beta_x x_{t-1} + \varepsilon_{xt} \]
Break process tracked by integer-valued state variable, $S_t$. Conditional on $K$ breaks

\[
\begin{align*}
  z_t &= B'_1 \tilde{x}_{t-1} + u_t, \quad E[u_t u'_t] = \Sigma_1 \quad 1 \leq t \leq \tau_1 \quad s_t = 1 \\
  : \quad & \quad : \quad : \\
  z_t &= B'_{K+1} \tilde{x}_{t-1} + u_t, \quad E[u_t u'_t] = \Sigma_{K+1} \quad \tau_K + 1 \leq t \leq T \quad s_t = K + 1
\end{align*}
\]

$\Upsilon_K = \{\tau_0, ..., \tau_K\}$ : collection of break points

Covariance matrix, $\Sigma_j$, decomposed as follows:

\[
\Sigma_j = \text{diag}(\psi_j) \times \Lambda_j \times \text{diag}(\psi_j)
\]

Both volatilities and correlations can vary across regimes.
Break dynamics is modeled through the transition probability matrix $\tilde{P}$:

$$
\tilde{P} = \begin{pmatrix}
p_{11} & p_{12} & 0 & \cdots & 0 \\
0 & p_{22} & p_{23} & \cdots & 0 \\
0 & \cdots & 0 & p_{Kk} & p_{k,k+1} \\
0 & 0 & \cdots & 0 & p_{k+1,k+1} \\
0 & 0 & \cdots & 0 & 0 \\
0 & 0 & \cdots & 0 & p_{k+2,k+2}
\end{pmatrix}
$$

$p_{j,j}$ is assumed to be independent of $p_{i,i}$, for $j \neq i$, and is drawn from a beta distribution: $p_{j,j} \sim Beta(a, b)$
Apply hierarchical prior setup to forecast returns out-of-sample

Location and scale parameters within each regime, \((B_j, \Sigma_j)\), are drawn from common “meta” distributions

Special cases:
- pooled scenario (parameters are identical across regimes)
- regime-specific scenario (parameters unrelated across regimes)
To characterize the parameters of the meta distribution, we assume

\[ \text{vec}(B)_j \sim N(b_0, V_0), j = 1, \ldots, K + 1 \]
\[ \psi_{j,i}^{-2} \sim \text{Gamma}(v_{0,i}, d_{0,i}) \]
\[ \lambda_{j,ic} \sim N(\mu_{\rho,ic}, \sigma_{\rho,ic}^2) \]
\[ b_0 \sim N(\mu_{\beta}, \Sigma_{\beta}) \]
\[ V_0^{-1} \sim W(B_v, V_0^{-1}) \]

\( W(\cdot) \): Wishart distribution
\( \mu_{\beta}, \Sigma_{\beta}, v_{\beta}, V_{\beta}^{-1} \): prior hyperparameters
\[ v_{0,i} \sim \text{Exp} \left( \rho_{0,i} \right) \]
\[ d_{0,i} \sim \text{Gamma} \left( c_{0,i}, d_{0,i} \right) \]

\( \rho_{0,i}, c_{0,i} \) and \( d_{0,i} \) : prior hyperparameters

Hyperparameters of correlation matrix (truncated to lie on \((-1, 1)):

\[ \mu_{\rho,ic} \sim \mathcal{N} \left( \mu_{\mu,ic}, \tau_{ic}^2 \right) \]
\[ \sigma_{\rho,ic}^{-2} \sim \text{Gamma} \left( a_{\rho,ic}, b_{\rho,ic} \right) \]

\( \mu_{\mu,ic}, \tau_{ic}^2, a_{\rho,ic}, b_{\rho,ic} \) : prior hyperparameters

\[ a \sim \text{Gamma} \left( a_0, b_0 \right) \]
\[ b \sim \text{Gamma} \left( a_0, b_0 \right) \]
Empirical Results

Data

- Monthly data on a portfolio of US stocks comprising firms listed on the NYSE, AMEX and NASDAQ
- Sample: 1926:12-2003:12
- Data source: CRSP
Evidence of Breaks

## I. Excess returns - Dividend Yield

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1368.18</td>
<td>1319.34</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1433.09</td>
<td>1369.33</td>
<td>0.00</td>
<td>Apr-52</td>
</tr>
<tr>
<td>2</td>
<td>1445.83</td>
<td>1383.28</td>
<td>0.00</td>
<td>Feb-52</td>
</tr>
<tr>
<td>3</td>
<td>1454.38</td>
<td>1387.49</td>
<td>0.00</td>
<td>May-40 Oct-52</td>
</tr>
<tr>
<td>4</td>
<td>1459.03</td>
<td>1390.04</td>
<td>0.00</td>
<td>May-40 Feb-58</td>
</tr>
<tr>
<td>5</td>
<td>1462.11</td>
<td>1391.51</td>
<td>0.00</td>
<td>May-40 Apr-52</td>
</tr>
<tr>
<td>6</td>
<td>1476.60</td>
<td>1404.75</td>
<td>0.01</td>
<td>Aug-32 May-40</td>
</tr>
<tr>
<td>7</td>
<td>1482.55</td>
<td>1408.74</td>
<td>0.69</td>
<td>Aug-32 May-40</td>
</tr>
<tr>
<td>8</td>
<td>1482.64</td>
<td>1407.89</td>
<td>0.30</td>
<td>Jun-32 May-40</td>
</tr>
</tbody>
</table>

Table 1: Model comparison and selection of the number of breaks in the return forecasting models. The table shows estimates of the log-likelihood for stock returns and the predictor variable (either the dividend yield or the T-Bill rate), marginal log-likelihood estimates for returns and posterior probabilities for models with different numbers of breaks along with the time of the break points for the different models. The top and bottom panels display results when the predictor for the excess return is the lagged dividend yield (panel I) and the lagged T-Bill rate (panel II), respectively. The data sample is 1926:12 - 2003:12.
Figure 1: Posterior probabilities of breakpoint locations for the return prediction model with seven breaks based on the dividend yield. The estimation sample is 1926:12 - 2003:12.
Evidence of seven Breaks

Break dates reasonably precisely determined

Break locations are associated with major events

- Great Depression (1932)
- Beginning of World War II (1940)
- Major oil price shocks and growth slowdown (1974)
- End of the change in the Fed’s operating procedures (1982)
- Beginning of the bull market of the nineties (1992)

Remaining break dates (1952 and 1958) harder to interpret
Table 2: Parameter estimates for the return ($\mu_r$) forecasting model with seven break points, based on the lagged dividend yield ($\beta_x$) as a predictor variable:

$$r_t = \mu_r + \beta_r x_{t-1} + \epsilon_t,$$

$$\beta_x,$$ 

where $\beta_x$, $\mu_x$, $\sigma_x$, $\mu_r$, $\beta_r$, $\sigma_r$, $\mu_x$, $\beta_x$, $\sigma_x$, $\mu_r$, $\beta_r$, $\sigma_r$, $\mu_x$, and $\sigma_x$ are the estimated parameters. The sample period is 1926:12-2003:12.
### Table 2: Parameter estimates for the return ($r_t$) forecasting model with seven break points, based on the lagged dividend yield ($x_{t-1}$) as a predictor variable:

\[
    r_t = r_j + r_j x_{t-1} + r_x t, \quad r_x N(0; 2) \quad x_{t-1} = x_j + x_j x_{t-1} + x_t, \quad x_t N(0; 2). \]

The sample period is 1926:12-2003:12.
Out-of-sample Break Estimates

Real time estimates of breaks

- Break1
- Break2
- Break3
- Break4
- Break5
- Break6
Out-of-sample Forecasts (1 month)

Pettenuzzo, Timmermann (UCSD) Predictability and Breaks March 9 2007 19 / 31
Out-of-sample Forecasts (12-month)
Out-of-sample Forecasts (12-month)
## Out-of-sample RMSE values

<table>
<thead>
<tr>
<th>horizon</th>
<th>break</th>
<th>no break</th>
<th>prevailing</th>
<th>mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0453</td>
<td>0.0449</td>
<td></td>
<td>0.0451</td>
</tr>
<tr>
<td>2</td>
<td>0.0669</td>
<td>0.0657</td>
<td></td>
<td>0.0662</td>
</tr>
<tr>
<td>3</td>
<td>0.0820</td>
<td>0.0802</td>
<td></td>
<td>0.0812</td>
</tr>
<tr>
<td>4</td>
<td>0.0947</td>
<td>0.0925</td>
<td></td>
<td>0.0939</td>
</tr>
<tr>
<td>5</td>
<td>0.1052</td>
<td>0.1030</td>
<td></td>
<td>0.1052</td>
</tr>
<tr>
<td>6</td>
<td>0.1164</td>
<td>0.1139</td>
<td></td>
<td>0.1167</td>
</tr>
<tr>
<td>7</td>
<td>0.1264</td>
<td>0.1234</td>
<td></td>
<td>0.1271</td>
</tr>
<tr>
<td>8</td>
<td>0.1359</td>
<td>0.1322</td>
<td></td>
<td>0.1364</td>
</tr>
<tr>
<td>9</td>
<td>0.1438</td>
<td>0.1393</td>
<td></td>
<td>0.1441</td>
</tr>
<tr>
<td>10</td>
<td>0.1511</td>
<td>0.1459</td>
<td></td>
<td>0.1516</td>
</tr>
<tr>
<td>11</td>
<td>0.1586</td>
<td>0.1527</td>
<td></td>
<td>0.1592</td>
</tr>
<tr>
<td>12</td>
<td>0.1652</td>
<td>0.1590</td>
<td></td>
<td>0.1665</td>
</tr>
</tbody>
</table>
Assume power utility over terminal wealth:

\[ u(W_{T+h}) = \frac{W_T^{1-\gamma}}{1-\gamma}, \quad \gamma > 0. \]

\( T \rightarrow T + h \): Holding period
\( \gamma \): coefficient of relative risk aversion
\( W_{T+h} \): Terminal wealth (\( W_T = 1 \))

\[ W_{T+h} = (1 - \omega) \exp(r_f h) + \omega \exp(r_f h + r_{T+1} + \ldots + r_{T+h}) \]
Buy-and-hold investor solves the following program

\[
\max_{\omega} E_T \left( \frac{\left( (1 - \omega) \exp(r_f h) + \omega \exp(r_f h + R_{T+h}) \right)^{1-\gamma}}{1 - \gamma} \right)
\]

subject to the no short-sales constraints \( 0 \leq \omega < 1 \)
No Breaks, no Estimation Uncertainty

\[ p(R_{T+h}|\hat{\Theta}, S_{T+h} = 1, Z_T) : \text{predictive return distribution ignoring parameter estimation uncertainty and breaks} \]

\[ \Theta = (\mu, \beta_0, \Sigma) : \text{VAR parameters} \]

\[ \hat{\Theta} : \text{Parameter estimate} \]

\[ Z_T : \text{Information set} \]

Investor maximizes

\[ \max_{\omega} \int u(W_{T+h})p(R_{T+h}|\hat{\Theta}, S_{T+h} = 1, Z_T)dR_{T+h}. \]
Integrating over the posterior distribution, \( \pi(\Theta|S_{T+h} = 1, Z_T) \), leads to the predictive distribution of returns conditioned only on the observed sample and the assumption of no breaks:

\[
p(R_{T+h}|S_{T+h} = 1, Z_T) = \int p(R_{T+h}|\Theta, S_{T+h} = 1, Z_T) \times \pi(\Theta|S_{T+h} = 1, Z_T) d\Theta.
\]

Investor solves the asset allocation problem

\[
\max\int \omega u(W_{T+h}) p(R_{T+h}|S_{T+h} = 1, Z_T) dR_{T+h}.
\]
Integrating out uncertainty about the future number of breaks:

\[
p(R_{T+h} \mid S_T = K + 1, Z_T) = \sum_{j=1}^{n_b+1} p(R_{T+h} \mid S_{T+h} = K + j, S_T = K + 1, Z_T) \times p(S_{T+h} = K + j \mid S_T = K + 1, Z_T).
\]

An investor who considers the uncertainty about the number of out of sample breaks but conditions on \( K \) historical breaks solves

\[
\max_\omega \int u(W_{T+h}) p(R_{T+h} \mid S_T = K + 1, Z_T) \, dR_{T+h}.
\]
Empirical Asset Allocation Results

Interaction between parameter estimation uncertainty and structural breaks:

- Breaks mean that bad draws of the parameters of the return model will eventually cease to affect returns following future breaks.
- Breaks lower the precision of current parameter estimates and increase the importance of parameter estimation uncertainty.
Figure 4: Optimal Asset Allocation as a function of the investment horizon for a buy-and-hold investor with power utility over terminal wealth, $U(W_T + h) = \frac{1}{1 + \gamma W_T + h}$, where $\gamma$ is the coefficient of relative risk aversion. The panels show allocations to stocks under the assumption that the current regime parameters and the dividend yield are set at the values from the regime prevailing during 1958-1974. The dotted line shows allocations starting from the end of the 1952-1974 regime. The dashed line shows the full-sample allocation ignoring breaks and parameter estimation uncertainty. The dashed/dotted line shows allocations based on full-sample parameter values (no breaks) but accounting for parameter uncertainty. Finally, the solid line shows allocations under the model which accounts for both past and future breaks.
Results Based on the Dividend Yield

- Initial value of (persistent) predictor variable matters a lot
- Allocation to stocks increases in the horizon if the initial value of the dividend yield is very low and breaks are ignored
- If past breaks are accounted for but future breaks are ignored, the asset allocation can be flat or increasing in the horizon
- If both past and future breaks are modeled, we see a non-monotonic or sometimes strongly declining allocation to stocks, the longer the investment horizon
- Parameter instability has a larger effect on a buy-and-hold investor’s optimal asset allocation than parameter estimation uncertainty
Conclusion

Our analysis accounts for

1. model uncertainty
2. parameter uncertainty
3. uncertainty about the number and size of historical breaks
4. uncertainty about future (out-of-sample) breaks

Empirical results

- Parameters of standard forecasting models appear to be highly unstable and subject to multiple shifts
- Many of the breaks coincide with important historical events
- Once such breaks are accounted for, the possibility of future breaks has a large impact on the optimal asset allocation