Return Predictability: Learning from the Cross-Section

Julien Penasse

Luxembourg School of Finance

9th Risk Forum, March 2016
Forecasting stock market returns

- Stock returns appear somewhat predictable:

\[ r_{t+1} = \theta + \beta x_t + u_{t+1} \]

- Excess return
- Predictor
Largely perceived as a stylized fact, i.e. time-varying risk premia

\[ E_t(r_{t+1}) = \theta + \beta x_t \]

Expected excess return

Fama and Schwert (1977); Rozeff (1984); Keim and Stambaugh (1986); Campbell (1987); Campbell and Shiller (1988); Fama and French (1989); Ferson and Harvey (1991); Campbell and Shiller (1991); Hodrick (1992).
Two (related) problems

Pastor and Stambaugh (2012): even after observing more than two centuries of US data, “investors do not know the values of the parameters of the return-generating process, especially the parameters related to the conditional expected return.”

International heterogeneity: strong predictability for some predictors in some countries (see e.g. Ang & Bekaert (2007), Hjalmarsson (2010))
International heterogeneity: illustration

Predictive slopes, predictor: log dividend-price ratio

![Graph showing international heterogeneity with predictive slopes for various countries.](image)
Two (related) problems

Pastor and Stambaugh (2012): even after observing more than two centuries of US data, “investors do not know the values of the parameters of the return-generating process, especially the parameters related to the conditional expected return.”

**International heterogeneity:** strong predictability for some predictors in some countries (see e.g. Ang & Bekaert (2007), Hjalmarsson (2010))

**Main problem:** low signal data.

**This paper:** international heterogeneity may be the work of chance.
Some recent works

Informative priors

\[ r_{t+1} = \theta + \beta x_t + u_{t+1} \]

E.g. Pettenuzzo et al. (2014) restrict \( \theta \) and \( \beta \) such that \( E_t(r_{t+1}) \geq 0 \).

Some recent works

Time-varying parameters

\[ r_{t+1} = \theta_t + \beta_t x_t + u_{t+1} \]

\{\theta_t, \beta_t\} random process


This makes the low signal problem even worse!
Some recent works
Pooled approaches

Cross section of $i = 1, \ldots, N$ countries.

$$r_{i,t+1} = \theta_i + \bar{\beta}x_{i,t} + u_{i,t+1}$$

$\bar{\beta}$ constant across countries.

Ang and Bekaert (2007), Hjalmarsson (2010), Rapach et al. (2013)

But $\beta$s seem to vary significantly across countries...
What I do in this paper

(i) Cross-country equations are *seemingly unrelated*

$$ r_{i,t+1} = \theta_i + \beta_i x_{i,t} + u_{i,t+1} $$

Seemingly unrelated regressions (SUR): each equation could be estimated separately *but efficiency gain* because the innovations are correlated (Zellner, 1962).
What I do in this paper
(ii) Cross-country equations are exchangeable

\[ r_{i,t+1} = \theta_i + \beta_i x_{i,t} + u_{i,t+1} \]

\( \theta_i \) and \( \beta_i \) are random variables: country \( i \)'s return process is a random draw from the cross-section of return processes.
Exchangeability: intuition

A two-country system

\[
\begin{align*}
{r_{us,t+1}} &= \theta_{us} + \beta_{us}x_{us,t} + u_{us,t+1} \\
{r_{uk,t+1}} &= \theta_{uk} + \beta_{uk}x_{uk,t} + u_{uk,t+1}
\end{align*}
\]

Exchangeability = random coefficient structure for the model parameters.
Assume that \( \beta_i \sim N(\bar{\beta}, V_{\bar{\beta}}), \; i = us, uk. \)
Exchangeability: intuition

Noninformative priors

Common assumption in the literature.

Ignorance: prior beliefs on the straight line:

\[ p(\beta_{us}) \propto p(\beta_{uk}) \propto p(\bar{\beta}) \propto 1 \]
Exchangeability: intuition

Informative prior about the common mean

Assume now that we know something about $p(\bar{\beta})$.

$$p(\bar{\beta}) \sim N(m_{\bar{\beta}}, V_{\bar{\beta}})$$

Should we update our priors about individual countries?
Informative prior about the common mean

If we don’t know more about individual countries, it is reasonable to assume

\[ p(\beta_{us}|\bar{\beta}) \sim N(m_{\bar{\beta}}, V_{\bar{\beta}}) \]
\[ p(\beta_{uk}|\bar{\beta}) \sim N(m_{\bar{\beta}}, V_{\bar{\beta}}) \]
Exchangeability: intuition

Informative prior about the common mean

Now assume we do observe UK data $D_{uk}$. Posterior distribution

$$p(\beta_{uk}|D_{uk})$$

Should we overlook the information from the cross section?
Exchangeability: intuition

Informative prior about the common mean

Now assume we do observe UK data $D_{uk}$. Posterior distribution

$$p(\beta_{uk}|D_{uk})$$

We obtain a more precise posterior using the cross section:

$$p(\beta_{uk}|D_{uk}, \bar{\beta})$$

---

**United States**

$p(\beta_{us}|\bar{\beta})$

---

**United Kingdom**

$p(\beta_{uk}|D_{uk})$

$p(\beta_{uk}|\bar{\beta})$

$p(\beta_{uk}|D_{uk}, \bar{\beta})$

---

**Common mean**

$p(\bar{\beta})$
Exchangeability: intuition

**Informative prior about the common mean**

If we also observe US data $D_{us}$, we obtain the posterior distribution (using again the cross section):

$$p(\beta_{us}|D_{us}, \bar{\beta})$$
Exchangeability: intuition

Informative prior about the common mean

In turn, we should update the common mean \( p(\bar{\beta}|D) \) to reflect the information contained in UK and US data.
Exchangeability: intuition

Informative prior about the common mean

If \( p(\bar{\beta}|D) \) changes, we must update

\[
p(\beta_{uk}|\bar{\beta}) \text{ and } p(\beta_{us}|\bar{\beta})
\]
Exchangeability: intuition

Informative prior about the common mean

If $p(\bar{\beta} | D)$ changes, we must update

$$p(\beta_{uk}, \bar{\beta}) \text{ and } p(\beta_{us}|\bar{\beta})$$

and then

$$p(\beta_{uk} | D_{uk}, \bar{\beta}) \text{ and } p(\beta_{us} | D_{us}, \bar{\beta})$$
Exchangeability: intuition

Informative prior about the common mean

... until the distributions converge to steady-state.
Exchangeability: intuition

Informative prior about the common mean

... until the distributions converge to steady-state.
Exchangeability: intuition

Informative prior about the common mean

... until the distributions converge to steady-state.
What I do in this paper

(iii) Equity premium constraints

Restrict $\theta_i$, $\beta_i$ such that

Weak form: \[ E(r_{i,t+1}|D_t) = \theta_i + \beta_i E(x_{i,t}|D_t) \geq 0 \]

Strong form: \[ E_t(r_{i,t+1}|D_t) = \theta_i + \beta_i x_{i,t} \geq 0 \text{ for } t = 1 \ldots T \]
\[ \theta_i + \beta_i \min(x_{i,t}) \geq 0 \]

for all countries $i$ simultaneously.
## Econometric model: overview

### Data-generating process: SUR model

- \( y_{t+1} = X_t \zeta + \epsilon_{t+1} \) where \( \epsilon_{t+1} \sim \mathcal{N}(0, \Sigma) \)
- SUR model for returns + predictors, \( \zeta = (\zeta_1 \ldots \zeta_N) \)
  \( \zeta_i = (\theta_i, \beta_i, \alpha_i, \rho_i) \).
- Covariance \( \Sigma \) parametrized by a simple one-factor structure.

### Prior beliefs

- \( \zeta_i = \bar{\zeta} + \eta_i, \eta_i \sim \mathcal{N}(0, \Delta). \)
- Diffuse prior for \( \bar{\zeta}, \Delta \) and the components of \( \Sigma \).
- Economic restrictions on the equity premium forecasts.

\( \zeta, \bar{\zeta}, \Sigma \) and \( \Delta \) characterize the model and are estimated (Metropolized Gibbs sampler).
Empirical findings

Much less international heterogeneity than previously reported.

- Both for the average equity premium, and for the degree of predictability (i.e. variation of the equity premium).
- The typical allocation to stock is relatively stable across countries.
  - Costly for the “outlier” countries to ignore cross-sectional information.
- The **US = not representative**. The equity premium is larger, predictability is stronger.
Data

- 15 OECD countries
- Quarterly frequency, postwar data (unbalanced panel)
- Equity data: S&P500 value-weighted (United States), MSCI (other OECD countries), from 27 years (Netherlands) to 60 years (US).
- Predictor: log-dividend price ratio.
  - Central result robust to alternative predictors.
Predictive slope
Non informative prior
Predictive slope

Exchangeability and weak equity premium restriction
Predictive slope
Exchangeability and strong equity premium restriction
Long run equity premium (annualized)

Non informative prior
Long run equity premium (annualized)

Exchangeability and weak equity premium restriction
Long run equity premium (annualized)

Exchangeability and strong equity premium restriction
Asset allocation: Illustration
Share of wealth invested in stocks (Weak Equity Premium Restriction vs. 'Naive')

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Optimal (%)</th>
<th>Uninformative prior (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-1.5</td>
<td>0</td>
</tr>
<tr>
<td>Australia</td>
<td>13</td>
<td>29</td>
</tr>
<tr>
<td>Belgium</td>
<td>8</td>
<td>25</td>
</tr>
<tr>
<td>Canada</td>
<td>0</td>
<td>24</td>
</tr>
<tr>
<td>Denmark</td>
<td>2</td>
<td>25</td>
</tr>
<tr>
<td>France</td>
<td>0</td>
<td>21</td>
</tr>
<tr>
<td>Germany</td>
<td>2</td>
<td>20</td>
</tr>
<tr>
<td>Italy</td>
<td>9</td>
<td>16</td>
</tr>
<tr>
<td>Japan</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>Netherlands</td>
<td>5</td>
<td>20</td>
</tr>
<tr>
<td>Norway</td>
<td>3</td>
<td>14</td>
</tr>
<tr>
<td>Spain</td>
<td>0</td>
<td>18</td>
</tr>
<tr>
<td>Sweden</td>
<td>13</td>
<td>30</td>
</tr>
<tr>
<td>Switzerland</td>
<td>18</td>
<td>36</td>
</tr>
<tr>
<td>UK</td>
<td>0</td>
<td>28</td>
</tr>
<tr>
<td>USA</td>
<td>28</td>
<td>84</td>
</tr>
</tbody>
</table>

Bayesian investor with power utility and coeff. of relative risk aversion $A = 5$. 
Previous studies have neglected the cross-sectional nature of return predictability.

By treating the evidence of return predictability jointly, investors
  Make efficient use of the cross-sectional correlation of the data;
  Learn about the common means and variances of the parameters.
Summary

- Previous studies have neglected the cross-sectional nature of return predictability.
- By treating the evidence of return predictability jointly, investors
  - Make efficient use of the cross-sectional correlation of the data;
  - Learn about the common means and variances of the parameters.
- Main empirical results
  - more precise posteriors.
  - much less cross-country dispersion.
  - equity premium constraints weaken the strength of return predictability.
- Not shown today
  - Out-of-sample gains w.r.t. the historical average and uninformative priors.
  - Stocks look much less riskier in the long run.
Econometric model
Modelling the predictor

\[ r_{i,t+1} = \theta_i + \beta_i x_{i,t} + u_{i,t+1} \]
\[ x_{i,t+1} = \alpha_i + \rho_i x_{i,t} + \nu_{i,t+1} \]

- \( u_{i,t+1}, \nu_{i,t+1} \) contemporaneously correlated: Strict exogeneity not respected.

- Cannot treat the regressor as fixed or exogenous in a time series setting (see Wachter and Warusawitharana, 2014)
  \[ \Rightarrow \] must model \( x_{i,t} \).
Econometric model

Modelling the predictor

\[
\begin{align*}
    r_{i,t+1} &= \theta_i + \beta_i x_{i,t} + u_{i,t+1} \\
    x_{i,t+1} &= \alpha_i + \rho_i x_{i,t} + v_{i,t+1}, \quad \rho_i \in (-1, 1)
\end{align*}
\]

- No bubble condition: the dividend-price ratio is stationary, i.e. \( \rho_i \in (-1, 1) \)
- Thus \( E(x_{i,t+1}|D_t) < \infty \) and \( E(r_{i,t+1}|D_t) < \infty \)
- Thus the initial condition \( x_{i,0} \) can be modeled as a draw from the (stationary) data generating process.
  \( \Rightarrow \) Can be used for the purpose of inference (Stambaugh, 1999).
Likelihood function

\[ L(D|\zeta, \Sigma) \propto |\Sigma|^{-T/2} \exp \left[ -\frac{1}{2} \sum_{t=1}^{T} (y_{t+1} - X_t \zeta)' \Sigma^{-1} (y_{t+1} - X_t \zeta) \right] \]

\[ \times \mathcal{N}(\mu_x, V_x). \]

where \( \mathcal{N}(\mu_x, V_x) \) is the unconditional distribution for the vector of predictors \( x_t = \alpha + \rho x_t + \nu_{t+1} \), with

\[ \mu_x = (I_N - \rho)^{-1} \alpha \]

and

\[ \text{vec}(V_x) = (I_{N^2} - (\rho \otimes \rho))^{-1} \text{vec}(\Sigma_x) \]
Predictive volatility

Non informative prior
Predictive volatility
Exchangeability and weak equity premium restriction

![Graph showing predictive volatility over different horizons.](image-url)
Predictive volatility

Exchangeability and strong equity premium restriction
## Summary Statistics

<table>
<thead>
<tr>
<th>Country</th>
<th>Period</th>
<th>Excess return</th>
<th>Log dp ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>s.d.</td>
</tr>
<tr>
<td>Australia</td>
<td>Q1:1971 - Q1:2013</td>
<td>2.54</td>
<td>20.02</td>
</tr>
<tr>
<td>Belgium</td>
<td>Q1:1971 - Q1:2013</td>
<td>3.87</td>
<td>21.58</td>
</tr>
<tr>
<td>Canada</td>
<td>Q1:1971 - Q1:2013</td>
<td>2.95</td>
<td>18.45</td>
</tr>
<tr>
<td>Denmark</td>
<td>Q1:1972 - Q1:2013</td>
<td>5.09</td>
<td>20.17</td>
</tr>
<tr>
<td>France</td>
<td>Q1:1971 - Q1:2013</td>
<td>3.23</td>
<td>21.00</td>
</tr>
<tr>
<td>Germany</td>
<td>Q1:1971 - Q1:2013</td>
<td>3.07</td>
<td>20.04</td>
</tr>
<tr>
<td>Italy</td>
<td>Q1:1971 - Q1:2013</td>
<td>-0.95</td>
<td>24.48</td>
</tr>
<tr>
<td>Japan</td>
<td>Q1:1971 - Q1:2013</td>
<td>3.58</td>
<td>19.93</td>
</tr>
<tr>
<td>Netherlands</td>
<td>Q1:1986 - Q1:2013</td>
<td>4.54</td>
<td>20.50</td>
</tr>
<tr>
<td>Norway</td>
<td>Q1:1979 - Q1:2013</td>
<td>4.79</td>
<td>29.08</td>
</tr>
<tr>
<td>Spain</td>
<td>Q2:1978 - Q1:2013</td>
<td>4.72</td>
<td>23.03</td>
</tr>
<tr>
<td>Sweden</td>
<td>Q1:1971 - Q1:2013</td>
<td>7.70</td>
<td>23.82</td>
</tr>
<tr>
<td>Switzerland</td>
<td>Q1:1974 - Q1:2013</td>
<td>4.94</td>
<td>17.19</td>
</tr>
<tr>
<td>UK</td>
<td>Q1:1971 - Q1:2013</td>
<td>4.34</td>
<td>20.47</td>
</tr>
<tr>
<td>USA</td>
<td>Q1:1953 - Q1:2013</td>
<td>9.60</td>
<td>15.26</td>
</tr>
</tbody>
</table>