Returns Predictability, Cointegration and Factor Models

Carlo A. Favero,[1]

[1]Bocconi University & CEPR

PhD

Cointegration and Predictability of Returns

- The existence of common stochastic trends is at the core of the literature on predictability of returns.
  - the common stochastic trend in dividend and prices in the dynamic dividend growth model (Campbell-Shiller),
  - the common stochastic trend between consumption, income and wealth in cay (Lettau-Ludvigson) ,
  - the common stochastic trend between consumption, income and dividends in (Lettau-Luddvigson),
  - and the common stochastic trend between dividends and consumption in (Bansal et al.)

The Dynamic Dividend Growth Model

$$
\begin{aligned}
r^s_{t+1} &= \kappa + \rho E_t\left(p_{t+1} - d_{t+1}\right) + E_t \Delta d_{t+1} \\
&\quad - \left(p_t - d_t\right) + \rho u^{pd}_{t+1} + u^{\Delta d}_{t+1} \\
\Sigma^m_{j=1} \rho^{j-1} r^s_{t+j} &= \frac{\kappa}{1-\rho} + \Sigma^m_{j=1} \rho^{j-1} E_t\left(\Delta d_{t+j}\right) - \left(p_t - d_t\right) \\
&\quad + \rho^m E_t\left(p_{t+m} - d_{t+m}\right) + \rho^m u^{pd}_{t+m} + \Sigma^m_{j=1} \rho^{j-1} u^{\Delta d}_{t+j}
\end{aligned}
$$

Cointegration in the DDG

- Cointegration between p and d is a necessary prerequisite to validate the model and the empirical evidence.
- (p-d) is a very persistent process. So non-standard distribution should be used to evaluate significance (Valkanov) or a slow moving time-varying mean could be introduced.
- time varying mean could be modelled through rare shifts or slow moving variables (demographics)

CAY

Lettau and Ludvigson concentrate on the cointegrating implications of the intertemporal budget constraint to study the role of fluctuations in aggregate consumption-wealth ratio for predicting stock returns

$$c_t - w_t = E_t \left[\sum_{j=1}^{\infty} \rho^j \left(r_{m,t+j} - \Delta c_{t+j} \right) \right] + \frac{\rho k}{1-\rho}$$

Risk, Returns and Portfolio Allocation with Cointegrated VARs

$$\mathbf{r}_{t+1} = \mu_t + \mathbf{u}_{t+1}$$

where $\mathbf{u}_{t+1}$ is the unexpected log return. Define the k-period cumulative return from period $t+1$ through period $t+k$, as follows:

$$\mathbf{r}_{t,t+k} = \Sigma_{i=1}^{k} \mathbf{r}_{t+i}$$

The term structure of risk is defined as follows:

$$\Sigma_r(k) \equiv \frac{1}{k} Var(\mathbf{r}_{t,t+k} \mid D_t) \tag{1}$$

where $D_t \equiv \sigma\{z_k : k \leq t\}$ consists of the full histories of returns as well as predictors that investors use in forecasting returns.

A cointegrated VAR with predictability

Consider a simple bi-variate first-order VAR for continuously compounded total stock market returns, r_t^s, and the log dividend price,dp_t:

$$\left(z_t - E_z\right) = \Phi_1 \left(z_{t-1} - E_z\right) + \nu_t$$
$$\nu_t \sim \mathcal{N}(0, \Sigma_\nu)$$

where

$$\begin{aligned} z_t &= \begin{bmatrix} r_t^s \\ dp_t \end{bmatrix}, E_z = \begin{bmatrix} E_{r^s} \\ E_{d-p} \end{bmatrix} \\ \Phi_1 &= \begin{bmatrix} 0 & \varphi_{1,2} \\ 0 & \varphi_{2,2} \end{bmatrix} \\ \begin{bmatrix} v_{1,t} \\ v_{2,t} \end{bmatrix} &\sim \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{matrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{matrix} \right] \end{aligned}$$

Predictability and the Term Structure of stock market risk

In our simple bivariate example, the term structure of stock market risk takes the form

$$\sigma_r^2(k) = \sigma_1^2 + 2\varphi_{1,2}\sigma_{1,2}\psi_1(k) + \varphi_{1,2}^2\sigma_{2,2}^2\psi_2(k) \tag{2}$$

where

$$\begin{aligned}
\psi_1(k) &= \frac{1}{k}\Sigma_{l=0}^{k-2}\Sigma_{i=0}^{l}\varphi_{2,2}^{i} \quad k > 1 \\
\psi_2(k) &= \frac{1}{k}\Sigma_{l=0}^{k-2}\left(\Sigma_{i=0}^{l}\varphi_{2,2}^{i}\right)^2 \quad k > 1 \\
\psi_1(1) &= \psi_2(1) = 0
\end{aligned}$$

Predictability and the Term Structure of stock market risk

The total stock market risk can be decomposed in three components: i.i.d uncertainty, σ_1^2 , mean reversion, $2\varphi_{1,2}\sigma_{1,2}\psi_1(k)$, and uncertainty about future predictors, $\varphi_{1,2}^2\sigma_{2,2}^2\psi_2(k)$. Without predictability $(\varphi_{1,2} = 0)$ the entire term structure is flat at the level σ_1^2. This is the classical situation where portfolio choice is independent of the investment horizon. The possible downward slope of the term structure of risk depends on the second term, and it is therefore crucially affected by predictability and a negative correlation between the innovations in dividend price ratio and in stock market returns $(\sigma_{1,2})$, the third term is always positive and increasing with the horizon when the autoregressive coefficient in the dividend yield process is positive.

An Illustration

$$
\begin{aligned}
r^s_{t+1} &= \alpha_r + b_r dp_t + \varepsilon_{1,t+1} \\
dp_{t+1} &= \alpha_{dp} + \varphi dp_t + \varepsilon_{2,t+1} \\
\begin{bmatrix} \varepsilon_{1,t+1} \\ \varepsilon_{2,t+1} \end{bmatrix} &\sim \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{matrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{matrix} \right]
\end{aligned}
$$

$$
\begin{aligned}
E(r^s_{t,t+1} \quad | \quad D_t) &= E(\alpha_r + b_r dp_t + \varepsilon_{1,t+1} \mid D_t) \\
E(r^s_{t+1,t+2} \quad | \quad D_t) &= E(\alpha_r + b_r \alpha_{dp} + b_r \varphi dp_t + b_r \varepsilon_{2,t+1} + \varepsilon_{1,t+2} \mid D_t)
\end{aligned}
$$

$$
\begin{aligned}
\sigma_r^2(1) &= Var(r^s_{t,t+1} \mid D_t) = \sigma_1^2 \\
\sigma_r^2(2) &= \frac{1}{2} Var(r^s_{t,t+2} \mid D_t) \\
&= \frac{1}{2}\sigma_1^2 + \frac{1}{2}\sigma_1^2 + \frac{1}{2} b_r^2 \sigma_2^2 + b_r \sigma_{12}
\end{aligned}
$$

Factor Models

- Factor models have been introduced to reduced dimensionality problem is asset allocation and risk measurement
- The implementation of asset allocation and risk measurement requires the estimation of a very large number of parameters: $\frac{n(n+1)}{2} + n$.
- Factor models allow to simplify the structure of the model and to reduce the number of parameters to be estimated

Factor Models:Time-Series Representation

The statistical distribution of N assets (i=1...n) is conditioned on a vector of K factors **f** (where N is large and K is small)

$$
\begin{aligned}
r^i_{t,t+k} &= \gamma^i_0 + \gamma^{i\prime}_1 \mathbf{f}_{t,t+k} + v^i_{t,t+k} \\
\mathbf{f}_{t,t+k} &= \mu^f + \mathbf{H}^f \epsilon_{t,t+k} \\
\Sigma^f &= \mathbf{H}^f \mathbf{H}^{f\prime}. \\
\mathbf{E}\left(v^i_{t,t+k}, v^j_{t,t+k}\right) &= 0 \\
\mathbf{E}\left(v^i_{t,t+k}, \epsilon^j_{t,t+k}\right) &= 0 \\
\epsilon_{t+k} &\sim \mathcal{D}\left(\mathbf{0}, \mathbf{I}\right)
\end{aligned}
$$

Factor Models: Cross-Sectional representation

The multifactor model has the following cross-sectional representation for the (Nx1) vector of returns at time t

$$
\begin{aligned}
\underset{(Nx1)}{\mathbf{r}_{t,t+k}} &= \underset{(Nx1)}{\alpha} + \underset{(NxK)}{B}\,\underset{(Kx1)}{\mathbf{f}_{t,t+k}} + \underset{(Nx1)}{\mathbf{v}_t} \\
\underset{(Kx1)}{\mathbf{f}_{t,t+k}} &= \underset{(Kx1)}{\mu^f} + \underset{(KxK)}{\mathbf{H}^f}\,\underset{(Kx1)}{\epsilon^f} \\
\Sigma^v &= \begin{bmatrix} \sigma_1 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & 0 \\ .. & .. & .. & .. \\ 0 & 0 & 0 & \sigma_n \end{bmatrix} \\
\Sigma^f &= \mathbf{H}^f \mathbf{H}^{f\prime}.
\end{aligned}
$$

Factor Models as Parsimoniuos Representation

- Using the joint distribution of returns to estimate the variance covariance matrix (the CER model) requires the estimation of n+n(n+1)/2 parameters;
- using a structure of k factors requires the estimation of (2n+nk)+ (k+k(k+1)/2) parameters.
- Think for example of an asset allocation problem with 30 assets and 4 factors. The CER would require the estimation of 505 parameters, the factor model would reduce that number to 194.

A single factor model:The CAPM

The time series representation

$$
\begin{aligned}
\left(r_t^i - r_t^{rf}\right) &= \beta_{0,i} + \beta_{1,i}\left(r_t^m - r_t^{rf}\right) + u_{i,t} \\
\left(r_t^m - r_t^{rf}\right) &= \mu_m + u_{m,t} \\
& \quad u_{i,t} \sim n.i.d.\left(0, \sigma_i^2\right) \\
\begin{pmatrix} u_{i,t} \\ u_{m,t} \end{pmatrix} &\sim n.i.d.\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{ii} & 0 \\ 0 & \sigma_{mm} \end{pmatrix}\right]
\end{aligned}
$$

The CAPM: cross-section representation

Considering the cross-section representation for the returns :

$$
\begin{aligned}
\mathbf{r}_t &= \beta_0 + \beta_1 r_t^m + \mathbf{u}_t \\
r_t^m &= E\left(r^m\right) + \sigma_m \mathbf{u}_{m,t} \\
\Sigma &= \beta_1 \beta_1' \sigma_m^2 + \Sigma_u \\
\mu &= \beta_0 + \beta_1 E\left(r^m\right)
\end{aligned}
$$

and μ, Σ can be obtained with the estimation of 3n+2 parameters only.

Validating Factor Models

- The diagonality of the variance-covariance matrix of the residuals coming from projecting asset returns on factors is a necessary–and testable–requirement for the validity of any factor model.
- Further validation is based on testing restrictions on their coefficients

Validation by testing restrictions

Consider once again the time-series representation of a factor model

$$r_{t+1}^i = \alpha_1 + \beta_i^{f^1} f_{t+1}^1 + \beta_i^{f^2} f_{t+1}^2 + \cdots + \beta_i^{f^k} f_{t+1}^k + v_{t+1} \qquad (3)$$

After having estimated N equations for the N assets you have available the following k vectors of coefficients, each of length N: $\boldsymbol{\beta}^{f1}$, $\boldsymbol{\beta}^{f2}$,..., $\boldsymbol{\beta}^{fk}$. Using the sample of t observations on the returns of the N assets you can compute the vector of length N of average sample returns for the assets: $E(\mathbf{r})$.

Validation by testing restrictions

You can now run the affine expected return-beta cross-sectional regression is:

$$E(\mathbf{r}) = \gamma_0 + \gamma_1 \boldsymbol{\beta}_{f1} + \gamma_2 \boldsymbol{\beta}_{f2} + \cdots + \gamma_k \boldsymbol{\beta}_{fk} + \mathbf{u}$$

A two-step test (FamaMacBeth) for the validity of any factor model can be run by considering the following null hypothesis:

$$\hat{\gamma}_0 = \bar{r}^f, \; \hat{\gamma}_i = E\left(f^i\right)$$

Fama-MacBeth

- care must be exercise in the test as the variance-covariance matrix of the residuals in the cross-sectional regression will not be diagonal and corrections for heteroscedasticity should be implemented.

- note also that, if both test assets and factors are excess returns, the validity of the model can be simply tested by evaluating the null that all intercepts in the time-series model for excess returns are zero.

- this null is inevitably rejected. Two industries have emerged (i) the factors "zoo", thath looks for omitted factors (ii) the performance evaluation industry that classifies fund manager performance according to their

Which Factors ?

Many different set of factors have been considered in the literature :

- Fundamental Factors
  - Fama-French five factors with observable characteristics and estimated betas (MKT, SMB, HML, RMW, CMA and momentum MOM
  - BARRA factors with know time-invariant betas and unobservable factor realizations estimated by cross-sectional regressions (see the program factormodels.R for a practical illustration)
- Macroeconomic Factors (inflation, growth and uncertainty)
- Statistical Factors (for example principal components)

Factor Exposures

Exposure to portfolios to factors can be assessed by computing the share of the total portfolio variance attributable to each factor.

$$r_{t+1}^{i} = \alpha_1 + \beta_i^{f^1} f_{t+1}^{1} + \beta_i^{f^2} f_{t+1}^{2} + \cdots + \beta_i^{f^k} f_{t+1}^{k} + v_{t+1}$$

$$\begin{aligned} Var\left(r_{t+1}^{i}\right) &= Cov\left(r_{t+1}^{i}, r_{t+1}^{i}\right) \\ &= \beta_i^{f^1} Cov\left(f_{t+1}^{1}, r_{t+1}^{i}\right) + \ldots \beta_i^{f^k} Cov\left(f_{t+1}^{k}, r_{t+1}^{i}\right) \\ &+ Cov\left(v_{t+1}, r_{t+1}^{i}\right) \end{aligned}$$

Cointegration and Factor Models

If a model is useful for the long-run, then it should be able to track the *price* of any asset; technically, portfolio and factor prices should share a common stochastic trend. When this long-run dimension of the data is not modeled, two opportunities are missed.

- First, factor specifications could be evaluated and compared based on their capability to capture long-run trends in asset portfolios.
- Second, the presence of a common trend between asset and factor prices implies the presence of a new term in the projection of returns on factors. This term captures temporary deviations of prices from their long-run trends and it is relevant to determine the time-series dynamics of asset returns.

From returns to prices

- Factor models have the general form:

$$r_{i,t+1} = \alpha_i + \beta_i' \mathbf{f}_{t+1} + v_{i,t+1}. \tag{4}$$

  Construct

  - Prices of any test asset as cumulative returns: $\ln P_{i,t} = \ln P_{i,t-1} + \mathbf{r}_{i,t}$.
  - Price-level risk drivers as cumulative returns of the factors: $\ln \mathbf{F}_t = \ln \mathbf{F}_{t-1} + \mathbf{f}_t$.

- Prices and risk-drivers follow stochastic trends that are not related under (4).

The problem(s) with standard factor models

- Standard Factor models that do no relate the stochastic trends in prices and risk drivers generate two potential problems.
    - using wrongly a "good" factor model
    - using correctly a "bad" factor model

A new approach: From prices to returns

- Start with a model that describes the exposure of a given portfolio $P_{i,t}$ to (price-level) risk drivers $\mathbf{F}_t$:

$$\ln P_{i,t} = \alpha_{0,i} + \alpha_{1,i}t + \underbrace{\beta_i' \ln \mathbf{F}_t}_{\text{intrinsic value}} + u_{i,t} \tag{5}$$

- The residuals $u_{i,t}$ are stationary if the risk drivers capture the stochastic trend in the long-run dynamics of prices.

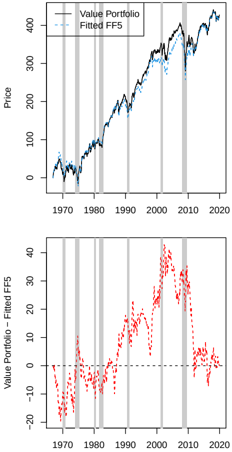
- For ease of exposition assume:

$$u_{i,t+1} = \rho_i u_{i,t} + v_{i,t+1}$$

- Taking first differences of our model in (5) we obtain

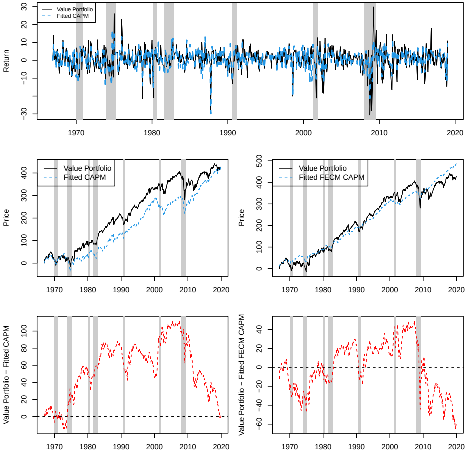
$$r_{i,t+1} = \alpha_{1,i} + \beta_i' \mathbf{f}_{t+1} + \underbrace{(\rho_i - 1)}_{\delta_i} \underbrace{u_{i,t}}_{\equiv ECT_{i,t}} + v_{i,t+1} \tag{6}$$

Price Dynamics in FF5 and its FECM Specification



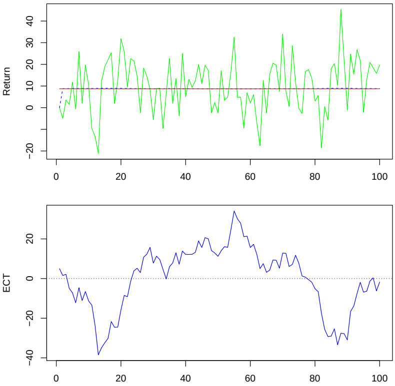
▸ MC simulation ▸ CAPM

Price Dynamics in CAPM and its FECM Specification

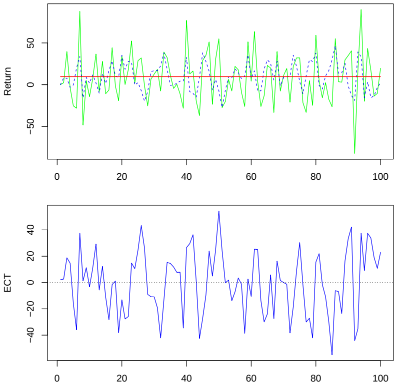


Back

Monte-Carlo Simulation



(a) DGP without cointegration.



(b) DGP with cointegration.

Back

A co-integrated approach to factor modelling

We model the joint distribution of portfolio prices, factors, and risk drivers as follows:

$$
\begin{aligned}
\ln P_{t+1}^{i} &= \alpha_0^i + \alpha_1^i t + \beta_i' \ln \mathbf{F}_{t+1} + u_{t+1}^i \\
u_{t+1}^i &= \rho_i u_t^i + v_{t+1}^i \\
\mathbf{f}_{t+1} &= E\left(\mathbf{f}_{t+1} \mid I_t\right) + \epsilon_{t+1} \\
\ln P_t^i &= \ln P_{t-1}^i + r_t^i \\
\ln \mathbf{F}_t &= \ln \mathbf{F}_{t-1} + \mathbf{f}_t \\
\epsilon_{t+1} &\sim \mathcal{D}\left(\mathbf{0}, \mathbf{\Sigma}\right) \\
Cov\left(v_{t+1}^i, v_{t+1}^j\right) &= 0
\end{aligned}
$$

If u_t^i is stationary, then prices and risk drivers are cointegrated.

What's next?

- ECT, Risk management and Portfolio timing
  - Modeling the relation between risk drivers and asset prices contributes to the description of the dynamics of returns
  - The predictive distribution of returns at time $t+1$ is centered on the ECT observed at time t.
- ECT and time-varying alphas
  - Recall

$$r_{i,t+1} = \alpha_{1,i} + \beta_i' \mathbf{f}_{t+1} + \delta_i ECT_{i,t} + \ v_{i,t+1}.$$

  - The test that the intercept is zero becomes a test that $\widehat{\alpha}_{1,i} + \widehat{\delta}_i ECT_{i,t} = 0$.