Empirical Asset Pricing

Master of Research in Finance

Paris Dauphine University

Juan F. Imbet Ph.D.

Grading

- Final Exam 70%
- Project/Homework 30%

Sources and References

- Ralph Koijen and Stijn Van Nieuwerburgh Ph.D. notes on Empirical Asset Pricing.
- John Cochrane's Asset Pricing book.
- John Campbell's Asset Pricing book.

Lesson 1: Return Predictability

Stock Return Predictability

- Average returns on stocks r_m are higher than the returns on short-term nominal bonds r_f.
- The equity premium $\mathbb{E}[r_m - r_f]$ and Sharpe ratio for the U.S. is robust across samples. For a long sample from 1926.7-2021.7, the equity risk premium in the U.S. is 8.3%. Return volatility is 18.5%. The Sharpe Ratio is 0.45.
- The Equity risk premium is similarly large for Europe and Asia Pacific, excluding Japan.
- Japan is a surprising "outlier" with no equity risk premium. (Bonds and stocks have had almost the same expected return).

Stock Return Predictability

- Equity returns are volatile, which makes it challenging to estimate the equity premium precisely. 95 years of data yield a standard error of $18.5/\sqrt{95} \approx 2\%$. A 95%-confidence interval ranges from 4% to 12%. (t-statistic of around 2 times the standard error).
- **Avdis and Wachter (2017)** provide unconditional maximum like-lihood estimators of the equity risk premium.

Time-series predictability and excess volatility.

- Campbell and Shiller (1988) deveop a log-linear approximation of returns that results in a useful accounting identity to understand the link between stock prices, fundamentals (e.g. dividends) and expected returns.

$$r_{t+1} = \log\left(\frac{P_{t+1} + D_{t+1}}{P_t}\right) = \Delta d_{t+1} + pd_t + \log\left(1 + \frac{P_{t+1}}{D_{t+1}}\right)$$

where $pd_t = \log(P_t/D_t)$ is the log-price-dividend ratio and $\Delta d_{t+1} = \log(D_{t+1}/D_t)$ is the log-dividend growth rate.

Campbell-Shiller decomposition

Apply a first-order Taylor approximation to the last term

$$\log\left(1+\frac{P_{t+1}}{D_{t+1}}\right) \approx \kappa_0 + \kappa_1 pd_{t+1}$$

where

$$\kappa_0 = \log\left(1+e^{\bar{pd}}\right), \kappa_1 = \frac{e^{\bar{pd}}}{1+e^{\bar{pd}}} - \kappa_1 \bar{pd}$$
$$\bar{pd} = \mathbb{E}[pd_{t+1}]$$

We can approximate returns as

$r_{t+1} \approx \kappa_0 + \Delta d_{t+1} + \kappa_1 pd_{t+1} - pd_t$

Campbell-Shiller decomposition

Iterate forward

$$pd_t = \frac{\kappa_0}{1-\kappa_1} + \sum_{j=0}^{\infty} \kappa_1^j \Delta d_{t+1-j} - \sum_{j=0}^{\infty} \kappa_1^j r_{t+1-j}$$

And the **transversality** condition

$$\lim_{j \to \infty} \kappa_1^j \mathbb{E}_t pd_{t+j} = 0$$

Note: You can use that transversality condition to test for bubbles. **Giglio, Maggiori and Stroebel (2016)** use this approach to test for bubbles over 700 years of data in the UK.

Present-value relation.

The equation holds ex-post as well as ex-ante

$$pd_t = \frac{\kappa_0}{1-\kappa_1} + \underbrace{\mathbb{E}_t \sum_{j=0}^{\infty} \kappa_1^j \Delta d_{t+1-j}}_{\Delta d_t^H} - \underbrace{\mathbb{E}_t \sum_{j=0}^{\infty} \kappa_1^j r_{t+1-j}}_{r_t^H}$$

- Movements in prices can be attributed to fluctuations in expected growth rates, expected returns or both.

Present-value relation (Variances)

$$Var(pd_t) = Var(\Delta d_t^H) + Var(r_t^H) - 2Cov(\Delta d_t^H, r_t^H)$$

Expected discounted future dividend growth rates or returns have to be volatile or they have to be negatively correlated if prices are to be volatile.

- Shiller (1981) provides the first evidence that prices appear to move more than what is implied by expected dividends, even realized dividends. This is the celebrated excess volatility puzzle.

Present-value relation (Co-variances)

$$
\begin{aligned}
Var(pd_t) &= Cov(\Delta d_t^H - r_t^H, pd_t) \\
&= Cov(\Delta d_t^H, pd_t) - Cov(r_t^H, pd_t) \\
\rightarrow & \\
1 &= \frac{Cov(\Delta d_t^H, pd_t)}{Var(pd_t)} - \frac{Cov(r_t^H, pd_t)}{Var(pd_t)}
\end{aligned}
$$

- First term is the slope of a regression predicting future dividend growth rates with pd_t.
- Second term is the slope of a regression predicting future returns with pd_t.
- The sum of both slopes has to be equal to one. **The dog that did not bark (Lettau and Van Nieuwerburgh, 2008 and Cochrane, 2008)**

Empirical Evidence

- Typical empirical framework ($pd_t = -dp_t$)

$$\begin{aligned} \Delta d_{t+1} &= a_d + \kappa_d dp_t + \epsilon_{d,t+1} \\ r_{t+1} &= a_r + \kappa_r dp_t + \epsilon_{r,t+1} \\ dp_{t+1} &= a_p + \phi dp_t + \epsilon_{p,t+1} \end{aligned}$$

Coefficient Restrictions

We know that (approximately)

$$\begin{aligned} r_{t+1} &= \kappa_0 + \Delta d_{t+1} - \kappa_1 dp_{t+1} + dp_t \\ r_{t+1} &= \kappa_0 + \Delta d_{t+1} - \kappa_1(a_p + \phi dp_t + \epsilon_{p,t+1}) + dp_t \\ r_{t+1} - \Delta d_{t+1} &= (1-\phi\kappa_1)dp_t + \kappa_0 - \kappa_1(a_p + \epsilon_{p,t+1}) \\ Cov(r_{t+1}, (1-\phi\kappa_1)dp_t) - Cov(\Delta d_{t+1}, (1-\phi\kappa_1)dp_t) &= (1-\phi\kappa_1)^2 Var(dp_t) \end{aligned}$$

The present value relation implies a coefficient restriction of the form:

$$\kappa_r - \kappa_d = (1-\phi\kappa_1)$$

Why is it different from one?

Predictive Regressions

If the right-hand side variable is highly persistent $\phi \approx 1$, the OLS estimator of κ_d and κ_r are biased upwards **Stambaugh (1999)**.

Correcting the Bias, Stambaugh 1999.

Consider the general model

$$\begin{aligned} y_{t+1} &= \alpha + \beta x_t + u_{t+1} \\ x_{t+1} &= \theta + \rho x_t + \nu_{t+1} \end{aligned}$$

Plus the assumption on stationarity $|\rho| < 1$. And covariance of the residuals:

$$\Sigma = \begin{bmatrix} \sigma_u^2 & \sigma_{u\nu} \\ \sigma_{u\nu} & \sigma_\nu^2 \end{bmatrix}$$

- **Proposition 4 (Stambaugh, 1999)**

$$\underbrace{\mathbb{E}[\hat{\beta} - \beta]}_{\text{Bias of } \hat{\beta}} = \frac{\sigma_{u\nu}}{\sigma_\nu^2} \underbrace{\mathbb{E}[\hat{\rho} - \rho]}_{\text{Bias of } \hat{\rho}} = -\frac{\sigma_{u\nu}}{\sigma_\nu^2}\left(\frac{1+3\rho}{T}\right) + \underbrace{O(1/T^2)}_{\text{Terms of order } 1/T^2,\, 1/T^3 \ldots}$$

How do we correct the bias?

- We assume that the rhs variables is a random walk (highly persistent). And we estimate the bias as:

$$\hat{\beta}^* = \hat{\beta} - \frac{\hat{\sigma}_{u\nu}}{\hat{\sigma}_\nu^2}\left(\hat{\rho} - 1\right)$$

Where $\hat{\beta}$ and $\hat{\rho}$ are the OLS estimators and the covariances are estimated based on the OLS residuals.

What is the Stock return predictability literature about?

1. Better statistical methods to infer expected returns or expected dividend growth rates given the persistence of the pd ratio.
   - Structural breaks (Lettau and Van Nieuwerburgh, 2008).
   - Filtering methods (Binsbergen and Koijen, 2010).
   - Near-unit root inference (Campbell and Yogo, 2006)
2. Use additional variables to predict returns.
   - Consumption growth (Lettau and Ludvigson, 2001).
   - The cross-section of valuation ratios (Kelly and Pruitt, 2013).
   - The variance risk premium (Bollerslev and Zhou, 2009).
   - Many more predictors, some of which have been called into question by Goyal and Welch (2008).

Econometric issues in return predictability

- Bias and correct test statistics if predictors are persistent (Mankiw and Shapiro (1986), Stambaugh (1999) and Campbell and Yogo (2006)).
- Correct inference in case of long-horizon regressions (Boudoukh, Richardson, and Whitelaw, 2008).
- Poor out-of-sample performance (Goyal and Welch, 2008 and Ferreira and Santa-Clara, 2011).
- In response to Goyal and Welch (2008), it is common practice to include a section on the out-of-sample predictability of a new predictor variable or a new method.

Relation between return predictability and the cross-section of expected returns.

The Stochastic Discount Factor (SDF) approach

Quick derivation - No structure

- A discount factor is just some random variable that generates prices rom payoffs

$$p = \mathbb{E}[mx]$$

- Can we always find such a d.f.? When is it positive? Hint, if there are no arbitrage opportunities.
  The proof comes from the linearity of the expectation operator and the law of one price. See Cochrane (2005) Chapter 4
- The SDF is unique if markets are complete (rarely the case in real life).

But what is m?

We have to come up with a story. One story is that m is related to the marginal utility of consumption.

$$\max_{\eta} u(c_t) + \beta \mathbb{E}_t u(c_{t+1})$$

$$\begin{aligned} c_t &= e_t - p_t \eta \\ c_{t+1} &= e_{t+1} + x_{t+1} \eta \end{aligned}$$

where e_t is endowment and x_{t+1} is the payoff of the asset.

F.O.C

$$\begin{aligned} p_t u'(c_t) &= \beta \mathbb{E}_t u'(c_{t+1}) x_{t+1} \\ p_t &= \mathbb{E}_t \underbrace{\beta \frac{u'(c_{t+1})}{u'(c_t)}}_{m_{t+1}} x_{t+1} \end{aligned}$$

On expected returns (keep this in mind)

Risk free assets

$$1 = R_f \mathbb{E}_t m_{t+1}$$

Risky assets

$$\begin{aligned}
1 &= \mathbb{E}_t m_{t+1} R_{t+1} \\
1 &= Cov(m_{t+1}, R_{t+1}) + \mathbb{E}_t m_{t+1} \mathbb{E}_t(R_{t+1}) \\
\frac{1}{\mathbb{E}_t m_{t+1}} &= \frac{Cov(m_{t+1}, R_{t+1})}{\mathbb{E}_t m_{t+1}} + \mathbb{E}_t(R_{t+1}) \\
\mathbb{E}_t(R_{t+1}) &= R_f - \frac{Cov(m_{t+1}, R_{t+1})}{Var(m_{t+1})} \frac{Var(m_{t+1})}{\mathbb{E}_t m_{t+1}} \\
\mathbb{E}_t(R_{t+1}) &= R_f + \beta\lambda
\end{aligned}$$

The Intertemporal CAPM (ICAPM) - Merton (1973)

- Variables that predict returns in the time-series could also predict returns in the cross-section.
- Setup: A representative agent that derived utility from consumption and a state variable z_t that captures expected return variation.

$$\max_{c_t} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t U(c_t, z_t)$$

Or using the Bellman Equation

$$V(W_t, z_t) = \max_{c_t} U(c_t, z_t) + \beta \mathbb{E}_t V(W_{t+1}, z_{t+1})$$

$$\text{s.t.}$$

$$W_{t+1} = (W_t - c_t)(1 + r_{t+1}^W)$$

where W_t is the agent's wealth, and r_{t+1}^W is the return on wealth.

First-order conditions

Differentiate with respect to c_t and W_t.

$$\frac{\partial V(W_t, z_t)}{\partial c_t} = 0 = U_c(c_t, z_t) - \beta \mathbb{E}_t V_W(W_{t+1}, z_{t+1})(1 + r_{t+1}^W)$$

$$\frac{\partial V(W_t, z_t)}{\partial W_t} = V_W(W_t, z_t) = \beta \mathbb{E}_t V_W(W_{t+1}, z_{t+1})(1 + r_{t+1}^W)$$

therefore

$$V_W(W_t, z_t) = U_c(c_t, z_t)$$

The SDF

$$\mathbb{E}_t R^i_{t+1} - R_f = -\frac{Cov(u'(c_{t+1}), R^i_{t+1})}{\mathbb{E}_t u'(c_{t+1})}$$

$$\mathbb{E}_t R^i_{t+1} - R_f = -\frac{Cov(V_W(W_{t+1}, z_{t+1}), R^i_{t+1})}{\mathbb{E}_t V_W(W_{t+1}, z_{t+1})}$$

First order approximation (plus some Ito's lemma in the original paper)

$$V_W(W_{t+1}, z_{t+1}) \approx V_W(W_t, z_t) + V_{WW}(W_t, z_t)\underbrace{(W_{t+1} - W_t)}_{\Delta W_{t+1}} + V_{Wz}(W_t, z_t)\underbrace{(z_{t+1} - z_t)}_{\Delta z_{t+1}}$$

Replacing

$$\begin{aligned} Cov(V_W(W_{t+1}, z_{t+1}), R^i_{t+1}) &\approx Cov\Big(V_W(W_t, z_t) + V_{WW}(W_t, z_t)\Delta W_{t+1} + V_{Wz}(W_t, z_t)\Delta z_{t+1}, R_{t+1}\Big) \\ &= V_{WW}(W_t, z_t)Cov(\Delta W_{t+1}, R^i_{t+1}) + V_{WZ}(W_t, z_t)Cov(\Delta z_{t+1}, R^i_{t+1}) \end{aligned}$$

Replacing

See Maio and Santa Clara (2012) for a more detailed derivation

$$\mathbb{E}_t R^i_{t+1} - R_f = \gamma Cov\left(R^W_{t+1}, R^i_{t+1}\right) \underbrace{- \frac{V_{Wz}}{V_W}}_{\gamma_z} Cov\left(z_{t+1}, R^i_{t+1}\right)$$

Intuition

- The first component is closely related to the standard CAPM.
- If a state variable hedges against changes in wealth, then

$$V_{Wz} > 0 \rightarrow \gamma_z < 0$$

Which implies that those assets that covary positively with the state variable have a lower expected return (are more expensive).

Fama (1991) critique

- Many of these multifactor models have been justified as empirical applications of the Intertemporal CAPM (ICAPM) (Merton, 1973), leading Fama (1991) to interpret the ICAPM as a "fishing license" to the extent that authors claim it provides a theoretical background for relatively ad hoc risk factors in their models.
- According to Merton, the state variables relate to changes in the investment opportunity set, which implies that they should forecast the distribution of future aggregate stock returns. Moreover, the innovations in these state variables should be priced factors in the cross-section.