Lecture 13
Factor Construction and Fama-MacBeth Regressions

Lars A. Lochstoer

UCLA Anderson School of Management

Winter 2025

Overview of Lecture 13

Constructing your own factors / trading strategies

1. Fama-MacBeth regressions and factor-mimicking portfolios
   - Trading on a stock characteristic

Applications of Fama-MacBeth Regressions

Cross-sectional, expected return relation

Let $X_{i,t}$ denote a $1 \times K$ vector of stock characteristics known at time t.

- An example is a stock's market beta estimated using historical data from time $t - j$ to time t.
- Another example is a stock's log book-to-market ratio, $bm_{i,t}$
- $X_{i,t}$ can be a vector and can include a 1 if an intercept term is desired

The Fama-MacBeth procedure we will consider seeks to estimate the $K \times 1$ vector λ in the below relation:

$$E_t \left[R^e_{i,t+1} \right] = X_{i,t} \lambda$$

Note the conditional nature of this statement. We are asking how, if at all, the conditional expected excess returns on stock i are related to the characteristics, $X_{i,t}$

Fama-MacBeth Regression

The Fama-MacBeth procedure we will consider seeks to estimate λ in the below relation:

$$E_t\left[R^e_{i,t+1}\right] = X_{i,t}\lambda$$

Recall that $R^e_{i,t+1} = E_t\left[R^e_{i,t+1}\right] + \eta_{i,t+1}$ where $E\left[E_t\left[R^e_{i,t+1}\right]\eta_{i,t+1}\right] = 0$. Thus:

$$R^e_{i,t+1} = X_{i,t}\lambda - \eta_{i,t+1}.$$

We don't care about the sign of $\eta_{i,t+1}$; it's just an error term, so we write the regression

$$R^e_{i,t+1} = X_{i,t}\lambda + \varepsilon_{i,t+1}.$$

Fama-MacBeth Implementation

For each time t, run the regression (with N observations):

$$R^e_{i,t+1} = X_{i,t}\lambda_{t+1} + \varepsilon_{i,t+1} \text{ for } i = 1, ..., N$$

- Collect $\hat{\lambda}_{t+1}$

The final estimate is

$$\hat{\lambda} = \frac{1}{T}\sum_{t=1}^{T}\hat{\lambda}_{t+1}$$

The squared standard error of the estimate, assuming λ_{t+1} are uncorrelated across time, is:

$$var\left(\hat{\lambda}\right) = \frac{var\left(\hat{\lambda}_{t+1}\right)}{T}$$

Fama-MacBeth Interpretation

Consider the estimate $\hat{\lambda}_{t+1}$:

$$\hat{\lambda}_{t+1} = \left(X_t' X_t\right)^{-1} X_t' R_{t+1}^e$$

where $X_t = [X_{1,t}^{'} \ X_{2,t}^{'} \ \ldots \ X_{N,t}^{'}]'$ is an $N \times K$ matrix.
$R_{t+1}^e = \left[R_{1,t+1}^e \ R_{2,t+1}^e \ \ldots \ R_{N,t+1}^e\right]$ is an $N \times 1$ vector.

- Note that we could easily make the number of stocks time-dependent, N_t.

Now, note that we can write this in terms of portfolio weights (on excess returns):

$$\hat{\lambda}_{t+1} = \mathbf{w}_t' R_{t+1}^e$$

where $\mathbf{w}_t' = \left(X_t' X_t\right)^{-1} X_t'$ is a $K \times N$ matrix of portfolio weights.

- The k'th row of $\mathbf{w}_t'$ are the portfolio loadings for the k'th row of $\hat{\lambda}_{t+1}$
- As is usual, the portfolio loadings do not need to sum to one since we are operating with excess returns

FM Interpretation: Example

Consider the following Fama-MacBeth regression:

$$R^e_{i,t+1} = \lambda^{intercept}_{t+1} + \lambda^{bm}_{t+1} bm_{i,t} + \varepsilon_{i,t+1} \text{ for } i = 1, ..., N$$

In this case, OLS implies that:

$$\begin{aligned}
\lambda^{bm}_{t+1} &= \frac{Cov_{cs}\left(bm_{i,t}, R^e_{i,t+1}\right)}{Var_{cs}\left(bm_{i,t}\right)} \\
&= \sum_{i=1}^{N} \underbrace{\frac{1}{N} \frac{bm_{i,t} - E_{cs}\left[bm_{i,t}\right]}{Var_{cs}\left(bm_{i,t}\right)}}_{=w_{i,t}} \times R^e_{i,t+1}
\end{aligned}$$

where $E_{cs}[\cdot]$, $Var_{cs}(\cdot)$, and $Cov_{cs}(\cdot)$ denote the
cross-sectional mean, variance (and covariance) of $bm_{i,t}$ (with R^e_{t+1}) at time t.

Note that λ^{bm}_{t+1} is a long-short portfolio return

- The (excess) return to a portfolio that is long high book-to-market stocks and short low book-to-market stocks.

FM Interpretation: Example (cont'd)

The final estimate of the Fama-MacBeth regression is

$$\hat{\lambda} = \frac{1}{T}\sum_{t=1}^{T}\hat{\lambda}_{t+1}$$

Thus, the estimated coefficient $\hat{\lambda}^{bm}$ is the mean return to the portfolio that goes long high bm stocks and short low bm stocks.

- The expected excess return to a factor-mimicking portfolio

Note that the intercept is capturing the average return to all stocks as well as a term related to the mean of the explanatory variable, $bm_{i,t}$:

$$\lambda_{t+1}^{intercept} = \frac{\overline{bm_{i,t}^2}}{\overline{bm_{i,t}^2} - \overline{bm_{i,t}}^2}\frac{1}{N}\sum_{i=1}^{N}R_{i,t+1}^{e} - \frac{\overline{bm_{i,t}}}{\overline{bm_{i,t}^2} - \overline{bm_{i,t}}^2}\frac{1}{N}\sum_{i=1}^{N}bm_{i,t}R_{i,t+1}^{e}$$

The math behind these results are on the next two slides.

Regression Math

Note: A *bar* $\bar{\cdot}$ denotes a cross-sectional average

$$X_t' X_t = N \begin{bmatrix} 1 & \overline{bm_{i,t}} \\ \overline{bm_{i,t}} & \overline{bm_{i,t}^2} \end{bmatrix}$$

So

$$\left(X_t' X_t\right)^{-1} = \frac{1}{N} \frac{1}{\overline{bm_{i,t}^2} - \overline{bm_{i,t}}^2} \begin{bmatrix} \overline{bm_{i,t}^2} & -\overline{bm_{i,t}} \\ -\overline{bm_{i,t}} & 1 \end{bmatrix}$$

And

$$\left(X_t' X_t\right)^{-1} X_t' =$$

$$\frac{1}{N} \frac{1}{\overline{bm_{i,t}^2} - \overline{bm_{i,t}}^2} \begin{bmatrix} \overline{bm_{i,t}^2} & -\overline{bm_{i,t}} \\ -\overline{bm_{i,t}} & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & \dots & 1 \\ bm_{1,t} & bm_{2,t} & \dots & bm_{N,t} \end{bmatrix}$$

Regression Math (cont'd)

Then:

$$\left(X_t' X_t\right)^{-1} X_t' =$$

$$\frac{1}{N}\frac{1}{\overline{bm^2_{i,t}} - \overline{bm_{i,t}}^2}\begin{bmatrix} \overline{bm^2_{i,t}} - \overline{bm_{i,t}}bm_{1,t} & \overline{bm^2_{i,t}} - \overline{bm_{i,t}}bm_{2,t} & \dots & \overline{bm^2_{i,t}} - \overline{bm_{i,t}}bm_{N,t} \\ bm_{1,t} - \overline{bm_{i,t}} & bm_{2,t} - \overline{bm_{i,t}} & \dots & bm_{N,t} - \overline{bm_{i,t}} \end{bmatrix}$$

So, finally:

$$\begin{aligned} \lambda^{intercept} &= \frac{1}{N}\frac{\overline{bm^2_{i,t}}\sum_{i=1}^{N} R^e_{i,t+1} - \overline{bm_{i,t}}\sum_{i=1}^{N} bm_{i,t}R^e_{i,t+1}}{\overline{bm^2_{i,t}} - \overline{bm_{i,t}}^2} \\ &= \frac{\overline{bm^2_{i,t}}}{\overline{bm^2_{i,t}} - \overline{bm_{i,t}}^2}\frac{1}{N}\sum_{i=1}^{N} R^e_{i,t+1} - \frac{\overline{bm_{i,t}}}{\overline{bm^2_{i,t}} - \overline{bm_{i,t}}^2}\frac{1}{N}\sum_{i=1}^{N} bm_{i,t}R^e_{i,t+1} \end{aligned}$$

If $\overline{bm_{i,t}} = 0$, $\lambda^{intercept}$ is the equal-weighted excess stock return at time $t+1$.

Take-away

The long-short portfolio sorts we typically see in finance are motivated by *cross-sectional forecasting regressions*

- Find a characteristic that you think forecasts next period stock return in the cross-section
- It is cross-sectional as the intercept captures the average return
  - Thus, we are looking for differences in return *across assets*, while not attempting to forecast the average movement

Let's look at an example of such regressions

- Novy-Marx (2013) and the Profitability Anomaly
- Anomaly, by the way, means alpha relative to existing model(s)

The Profitability Factor

Novy-Marx argues:

- The typical measure of future profitability, current earnings, is inferior to gross profitability
- For instance, investments that are treated as expenses (advertising, R&D, human capital development) are subtracted of profits to get the current earnings, but these investments actually signal higher future profitability
- Earnings are often manipulated, which obscures their relation to future earnings

Of these reasons, Novy-Marx suggests that a better proxy for future profitability is Gross Profitability divided by Book Value of Assets

$$Gross\ Profitability = \frac{\text{Total Revenue (RevT) - Cost of Goods Sold (COGS)}}{\text{Book Assets (AT)}}.$$

The Profitability Factor (cont'd)

Fama-MacBeth regressions showing the the GP characteristics is indeed related to average returns

Table 1.
Fama-MacBeth regressions of returns on measures of profitability.

Panel A reports results from Fama-MacBeth regressions of returns on gross profits (revenues minus cost of goods sold, REVT - COGS) scaled by assets (AT), and income before extraordinary items (IB) and free cash flow (net income plus amortization and depreciation minus changes in working capital and capital expenditures, NI + DP - WCAPCH - CAPX) each scaled by book equity. Panel B repeats the tests employing profitability measures demeaned by industry (Fama-French 49). Regressions include controls for book-to-market (log(B/M)), size (log(ME)), and past performance measured at horizons of one month ($r_{1,0}$) and twelve to two months ($r_{12,2}$). Independent variables are trimmed at the one and 99% levels. The sample excludes financial firms (those with one-digit SIC codes of six), and covers July 1963 to December 2010.

	slope coefficients ($\times 10^2$) and [test-statistics] from regressions of the form $r_{tj} = \boldsymbol{\beta}'\mathbf{x}_{tj} + \epsilon_{tj}$						
independent variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Panel A: straight profitability variables						
gross profitability	0.75 [5.49]			0.69 [5.22]	0.62 [4.63]		0.61 [4.59]
earnings		0.22 [0.84]		0.08 [0.31]		-0.02 [-0.06]	-0.07 [-0.27]
free cash flow			0.27 [2.28]		0.20 [1.64]	0.39 [3.17]	0.33 [2.67]
log(B/M)	0.35 [5.98]	0.30 [4.97]	0.26 [4.59]	0.34 [5.54]	0.30 [5.17]	0.27 [4.48]	0.31 [5.05]
log(ME)	-0.09 [-2.29]	-0.12 [-3.24]	-0.13 [-3.20]	-0.11 [-2.78]	-0.11 [-2.80]	-0.13 [-3.34]	-0.11 [-2.92]
$r_{1,0}$	-5.57 [-13.8]	-5.49 [-13.7]	-5.52 [-13.7]	-5.64 [-14.1]	-5.66 [-14.1]	-5.56 [-13.9]	-5.70 [-14.3]
$r_{12,2}$	0.76 [3.87]	0.78 [4.02]	0.78 [4.02]	0.74 [3.80]	0.74 [3.80]	0.76 [3.93]	0.73 [3.74]

Why do these factors 'work'?

First, simplest model possible (constant growth and discount rates):

$$\frac{P}{D} = \frac{1}{r-g}$$

Thus, high valuations mean low discount rates (expected returns) and/or high growth.

- Thus, with a good estimate of future growth, value investing will 'work' in terms of giving a spread in returns
- Of course, b/m may not continue to work; perhaps we will need better estimates of g
- CAPM alpha may not continue, perhaps in the future variation in r is all explained by the CAPM

Why do these factors 'work'? (cont'd)

Using "clean surplus" accounting (see Ohlson (1995)), we can write the market to book ratio as:

$$\frac{M_t}{B_t} = \frac{1}{B_t} \sum_{\tau=1}^{\infty} \frac{Y_{t+\tau} - \Delta B_{t+\tau}}{(1+r)^{\tau}}.$$

where Y is earnings, B is book value of equity and r is a long-run discount rate

The above equation makes a couple of statements about future expected returns:

1. Holding earnings Y and investment ΔB constant, a higher market value (size or m/b ratio) means lower expected returns.
2. Holding market value and investment constant, higher earnings Y (profitability) means higher expected returns.
3. Holding profitability and market value constant, higher investment means lower expected returns.

In sum: whether of behavioral or risk-based reasons, there are reasons to think portfolio sorts based on value, profitability, and investment may continue to give spreads in average returns

The Fama-French 5-factor Model

FF sorts into factors in several different ways in order to show robustness of the results. The new factors RMV and CMA (Robust minus Weak profitability and Conservative minus Aggressive investment, respectively) are formed in a similar fashion as the HML sort.

The model is:

$$R_{i,t} - R_{f,t} = \alpha_i + b_i \left(R_{M,t} - R_{F,t} \right) + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + \varepsilon_{i,t}$$

The Fama-French 5-factor Model (cont'd)

All factors have significant positive risk premiums (Panel A)

All three last factors (HML, RMW, CMA) have high returns coming mainly from small firms (Panel B)

Panel A: Averages, standard deviations, and *t*-statistics for monthly returns

	2x3 factors					2x2 factors					2x2x2x2 factors				
	R_M-R_F	*SMB*	*HML*	*RMW*	*CMA*	R_M-R_F	*SMB*	*HML*	*RMW*	*CMA*	R_M-R_F	*SMB*	*HML*	*RMW*	*CMA*
Mean	0.50	0.29	0.37	0.25	0.33	0.50	0.30	0.28	0.17	0.22	0.50	0.30	0.30	0.25	0.14
Std Dev	4.49	3.07	2.88	2.14	2.01	4.49	3.13	2.16	1.52	1.48	4.49	2.87	2.13	1.49	1.29
t-statistic	2.74	2.31	3.20	2.92	4.07	2.74	2.33	3.22	2.79	3.72	2.74	2.60	3.43	4.09	2.71

	HML_S	HML_B	HML_{S-B}	RMW_S	RMW_B	RMW_{S-B}	CMA_S	CMA_B	CMA_{S-B}
2x3 factors									
Mean	0.53	0.21	0.32	0.33	0.17	0.16	0.45	0.22	0.23
Std Dev	3.24	3.11	2.69	2.69	2.35	2.68	2.00	2.66	2.47
t-statistic	4.05	1.69	2.94	3.06	1.81	1.48	5.49	2.00	2.29

The paper has many tables and shows the model can account for expected returns on large set of 'anomaly' sorted portfolios, though the low return on small stocks that invest a lot despite low profitability remains a puzzle. Also, the authors show that the HML factor in fact becomes redundant, so really this may be re-couched as a 4-factor model.

One more thing...

Note that a regression beta can be a characteristic.

- For instance, consider regressing firm i's returns on inflation.

Get the 'inflation beta' of stock i $\left(\beta_{i,t}^{\iota nfl}\right)$.

- Ask: if we go long high inflation beta stocks, short low inflation beta stocks do we earn a premium?
- For instance, if high inflation is bad, the premium on such a strategy should be negative (an inflation hedge portfolio)

Fama-MacBeth regressions estimate this premium as the average return to the inflation beta long-short portfolio

- $\lambda_{\text{inflation}}$

References

- Eugene F. Fama, Kenneth R. French, A five-factor asset pricing model, Journal of Financial Economics, Volume 116, Issue 1, April 2015, Pages 1-22
- Robert Novy-Marx, The other side of value: The gross profitability premium, Journal of Financial Economics, Volume 108, Issue 1, April 2013, Pages 1-28
- Ohlson, Contemporary accounting research, 1995, Earnings, Book Values, and Dividends in Equity Valuation.