MS&E 448: Trading Strategies based on Explained Cross-Sectional Returns

Amay Aggarwal, Yu Xin Ang, Varin Satish Nair, Gregory Soh, Chor Seng Tan

Background

Introduction

In contrast to time-series analysis in which the behavior of one or more economic aggregates are traced through time, cross-sectional analysis is a type of observational study that analyzes data at a specific point in time.

Aim is to study why average returns differ across different stocks or portfolios.

E.g. Capital Asset Pricing Model (CAPM), Fama-French 3 factor model etc.

Cross-Sectional Returns

1960s-70s: Sharpe-Lintner-Black (SLB) model: returns are a positive linear function of market beta (CAPM)

1981: Banz found that market equity (ME) helps explain cross-section of average returns provided by beta

1988: Bhandari found that leverage also helps explain cross-section of returns provided by beta and ME

1990s: Others found that factors such as book-to-market value, book-to-market-equity, earning-price ratios etc. were also related to market returns

Cross-Sectional Returns

The asset-pricing literature finds strong cross-sectional predictability in stock returns.

Firm specific characteristics such as size, book-to-market equity, past returns, and investments have been found to be strongly correlated with the firm's subsequent stock returns.

Portfolios constructed on extreme deciles recorded high Sharpe ratios in out--of-sample testing.

Goal: Design our own factor model to construct a portfolio

Project Outline

Our (Data) Universe

We decided to look at constituents of the S&P 500 index, as we aim to build a portfolio based on high market-cap stocks. To avoid survival bias, we considered companies that were on the index as of 01/01/2011.

We collected company specific data/characteristics (e.g. market capitalization, beta, price-to-earnings ratio etc.) of these 500 companies from 2011 to 2018 using WRDS, and also macroeconomic data from the same time period.

Training set: 2011 to 2016, Validation set: 2017, Test set: 2018.

Predictive Variables

In the literature, it seems to be common to start with just a smaller number of predictive variables, before modeling on more variables

Some that are highlighted include: size, book-to-market, past 12 month returns, and beta (Freyberger et al.)

We plan to use ~50 firm specific characteristics as predictive variables in our final model

Portfolio Construction

To do a rolling model where coefficients are constantly updated with time, the time frame to be decided by the results from the validation set

Portfolio construction options:

- Can allocate portfolio based on strength of signal of all stocks
- Can allocate just to the deciles (long the top 10% and short the bottom 10%)

Our aim is to develop a zero-net investment portfolio

Expected Outcomes

- Build models that explain stock returns using stock/industry characteristics and macroeconomic factors
- Create decile sorting of stocks ranked by expected returns
- Long top 10% and short bottom 10% (through value-weighted / equal weighted) to create zero net investment portfolio
- Evaluate returns of the portfolio in relation to average market returns

Preliminary Findings

Fama-French 3 factor model

Designed by Eugene Fama and Kenneth French in 1992 to add value risk (HML) and size risk (SMB) to market risk:

$$r - r_f = \beta_1(r_m - r_f) + \beta_2(HML) + \beta_3(SMB)$$

r: returns (monthly)

r_f: risk free rate (10 year Treasury Bill)

r_m: market return (defined as S&P 500)

HML: high minus low (book-to-market)

SMB: small minus big (market cap i.e. size)

→ Should invest in small and high value stocks

Preliminary Findings/Results

Constructed equal-weighted portfolio based on 2011 - 2016 data

Strategy	Average Daily Returns of SMB	Average Daily Returns of HML
Long top 10%, short bottom 10%	-0.000911777	0.001301399
Long top 30%, short bottom 30%	-0.000310675	0.00053102
Long top 50%, short bottom 50%	-0.000143912	0.0000857543

SMB no longer matters but HML still seems to work!

Next Steps

1. Increasing the number of predictive variables in our own factor model
2. Fitting more complex prediction models, such as neural networks and PCA
   a. First attempt: add all potential firm-level characteristics to the model
   b. Second attempt: look at individual factors and decide their relevance. Filter out factors that do not appear to affect returns
3. From the predictive signals, construct a zero-net investment portfolio
4. Run simulations on our test set to evaluate portfolio performance

References

- Eugene F. Fama, Kenneth R. French. The Cross Section of Expected Stock Returns. The Journal of Finance. 1992
- Jonathan Lewellen. The Cross-Section of Expected Stock Returns. Tuck School of Business Working Paper No. 2511246.2014
- Freyberger, J., Neuhierl, A., Weber, M., 2017. *Dissecting characteristics nonparametrically*. Working Paper 23227, National Bureau of Economic Research.
- Bryan Kelly, Seth Pruitt, and Yinan Su. *Characteristics are covariances: A unified model of risk and return*. Technical report, National Bureau of Economic Research, 2018
- Shihao Gu, Bryan T Kelly, and Dacheng Xiu. *Autoencoder asset pricing models*
- Luyang Chen, Jason Zhu, and Markus Pelger. *Deep learning in asset pricing*. 2018
- Masaya Abe and Hideki Nakayama. *Deep Learning for Forecasting Stock Returns in the Cross-Section*. 2018

QUESTIONS