MECON6102 Problem Set 1

Xing Mingjie

April 16, 2024

Abstract

This report studies the relationship between monthly stock returns and a set of factors and construct an investment strategy based on various factor models. The report first uses a naive factor regression to estimate the factor loadings of the stock returns. Then, the report uses the Fama-MacBeth regression to estimate the factor risk premia. Finally, the report uses the LASSO regression to select the factors. The report also constructs a mean-variance portfolio based on the factor risk premia.

1 Data

The stock monthly return data is obtained from the CRSP database. The sample creates an unbalanced panel of 573 stocks from 2000-01 to 2022-12. There are in total 82060 observations.

The factors combine Fama-French 5 factors model, q-Factor of Hou et al. (2014), and liquidity factor of Ľuboš Pástor and Stambaugh (2003). Readers are referred to the juoyter notebook for detailed description of the factors.

1.1 Data preparation

To boost calculation, regressions in this exercise is performed on a rolling window with size of 60. So those stocks with less than 60 observations are dropped. The final panel consists of 78233 observations of 390 stocks.

2 Naive Factor Regression

In Figure 1 we can see that the intercept term α is relatively stable over time, oscilliating around -0.01. This is consistent with the theory that the intercept term should be zero in the textbook CAPM model. The small negativity of the intercept could be due to the fact that the stock `10324` loses to the market in the sample period, and having a negative risk premium, especially in the early 2005.

In Figure 2 we can see the risk exposure to market risk free return is arguably consistent acorss time, around 0.005, though there is a secular downward trend, especially since the surge in late 2005. This is consistent with the theory that the market risk premium should be positive and stable over time. The decreasing trend may be due to the fact that the stock `10324` is becoming less risky over time, or the market is becoming more risky over time.

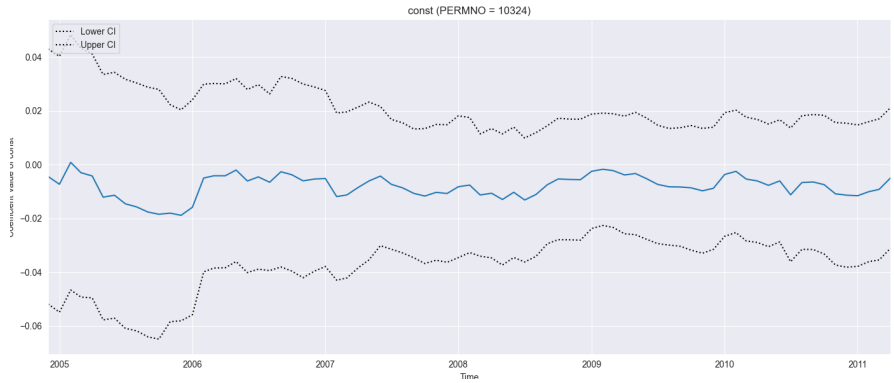

Figure 1: α time series of stock 10324

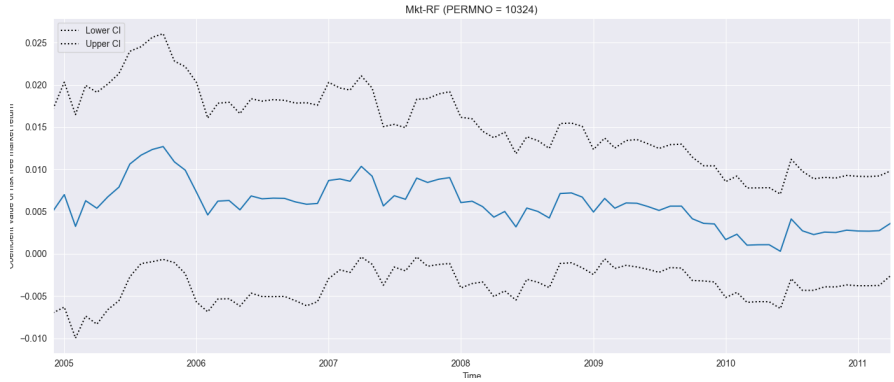

Figure 2: β_{Mkt-RF} time series of stock 10324

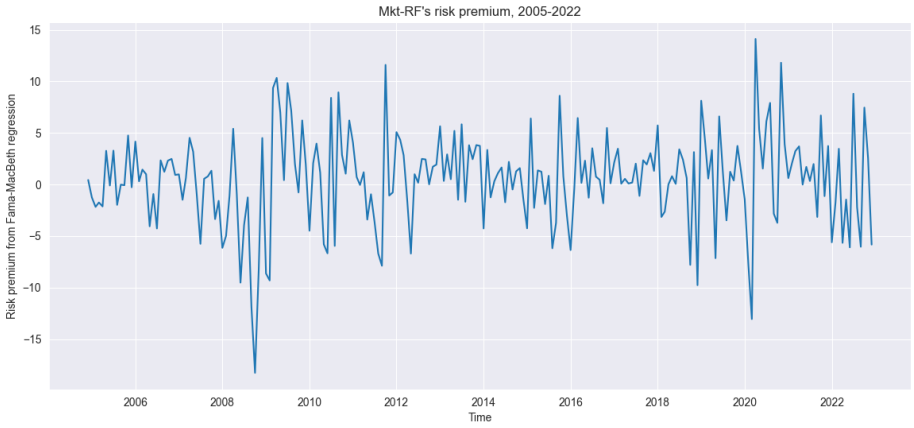

Figure 3: γ_{Mkt-RF} risk premium time series

Table 1: Top 5 factors selected by LASSO each period

date	1st	2nd	3rd	4th	5th
200501	Mkt-RF	HML	Mkt-RF-l5	HML-l3	Mkt-RF-l4
200601	Mkt-RF	Mkt-RF-l2	HML	Mkt-RF-l5	Mkt-RF-l4
200701	Mkt-RF	R_ME	SMB	Mkt-RF-l3	Mkt-RF-l1
200801	Mkt-RF	Mkt-RF-l2	SMB	Mkt-RF-l1	HML
200901	Mkt-RF	Mkt-RF-l3	HML	Mkt-RF-l1	Mkt-RF-l4
201001	Mkt-RF	Mkt-RF-l4	Mkt-RF-l1	Mkt-RF-l2	Mkt-RF-l3
201101	Mkt-RF	Mkt-RF-l1	Mkt-RF-l4	Mkt-RF-l2	HML
201201	Mkt-RF	Mkt-RF-l2	Mkt-RF-l4	Mkt-RF-l1	Mkt-RF-l3
201301	Mkt-RF	Mkt-RF-l4	Mkt-RF-l2	Mkt-RF-l1	Mkt-RF-l3
201401	Mkt-RF	Mkt-RF-l6	Mkt-RF-l5	Mkt-RF-l4	Mkt-RF-l2
201501	Mkt-RF	Mkt-RF-l1	Mkt-RF-l6	Mkt-RF-l2	Mkt-RF-l3
201601	Mkt-RF	Mkt-RF-l3	Mkt-RF-l6	R_ME	Mkt-RF-l1
201701	Mkt-RF	HML	Mkt-RF-l3	Mkt-RF-l5	Mkt-RF-l1
201801	Mkt-RF	HML	SMB	Mkt-RF-l1	R_ME
201901	Mkt-RF	HML	Mkt-RF-l1	Mkt-RF-l5	Mkt-RF-l4
202001	Mkt-RF	HML	Mkt-RF-l1	Mkt-RF-l3	Mkt-RF-l4

3 Fama-MacBeth Regression

In Figure 3 we can see the risk premium of the market risk free return is oscilliating around 0. The negative most value happens in late 2008, possibly due to the financial crisis. The stability of the risk premium is also consistent with the stability of the risk exposure to market risk free return in the naive factor regression.

4 LASSO Regression

Factors are lagged by one to six months to avoid look-ahead bias. First 5 rows of each window are dropped due to missing value in the lagged factors.

LASSO exercise is done without cross validation. The step size is chosen at 0.001. Risk free market return is chosen as the most prevalent factor in the LASSO regression throughout the time, as can be seen in Table 1. This is in line with the L2-Boosting result. Other factors that are frequently chosen include `HML`, `SMB`, `R_ME`, and 1 to 6 month lag of risk free market return. `SMB` is the size factor, `HML` is the value factor, and `R_ME` is the momentum factor. These findings are also consistent with the Gradient Boosting result.

5 Mean-Variance Portfolio

6 Conclusion

Bibliography

Hou, K., Xue, C., and Zhang, L. (2014). Digesting Anomalies: An Investment Approach. *The Review of Financial Studies*, 28(3):650–705.

Ľuboš Pástor and Stambaugh, R. F. (2003). Liquidity risk and expected stock returns. *Journal of Political Economy*, 111(3):642–685.