Econometrics IV's Assignment

Caio Garzeri

Guilherme Luz

Guilherme Masuko

August 2023

1 Question 1

The first question consists of a factor analysis of a large dataset. We consider monthly close-to-close excess returns from a cross-section of 9,456 firms traded in the New York Stock Exchange. The data starts on November 1991 and runs until December 2018. There are 326 monthly observations in total.

In addition to the returns we also consider 16 monthly factors:

- Market (MKT)

- Small-minus-Big (SMB)

- High-minus-Low (HML)

- Conservative-minus-Aggressive (CMA)

- Robust-minus-Weak (RMW)

- earning/price ratio (EP)

- cash-flow/price ratio (CFP)

- dividend/price ratio

- accruals (ACC)

- market beta (BETA)

- net share issues

- daily variance (RETVOL)

- daily idiosyncratic variance (IDIOVOL)

- 1-month momentum (MOM1)

- 36-month momentum (MOM36)

The dataset is organized as an excel file named `returns.xlsx`.

1.1 (a) (30 points)

Compute the principal components of the returns and determine the optimal number of principal factors by one the methods described in Lecture 2. How much of the variance will the factors be able to explain?

1.2 (b) (30 points)

Regress the selected factors on the 16 observed "anomaly" factors described above. How do the "principal component factors" relate to the "anomaly factors"?

1.3 (c) (30 points)

Now, run a principal component analysis on the 16 "anomaly factors" and select the optimal number of principal components using the same criterion adopted in the first item of the exercise. By inspecting the principal eigenvectors can you identify a dominating "anomaly"?

1.4 (d) (30 points)

How do the "anomaly-based principal factors" related to the "return-based principal factors"?

2 Question 2

3 Question 3

References

A Appendix