FIN3080 Report 123090559

Problem 1

Data collection

As required, the data was downloaded ended up with 1 cvx files: the closing index over 2005/12/31 to 2023/12/31 (the data is 2005 is for calculating the return in 2006/01).

Data manipulation

The downloaded data are daily returns, whereas the inquiry calls for monthly returns. I use the daily returns on the last day of each month to represent monthly returns.

Convert the date column from object to date time format and then changing it to a monthly period. Then select specifically the data with index code "000300" (for CSI 300). The monthly return is derived by .pct change() method.

Then the 2006-01 data is dropped as this is the first index data, with no return, as needed in skewness and kurtosis otherwise it cannot be calculated.

Results

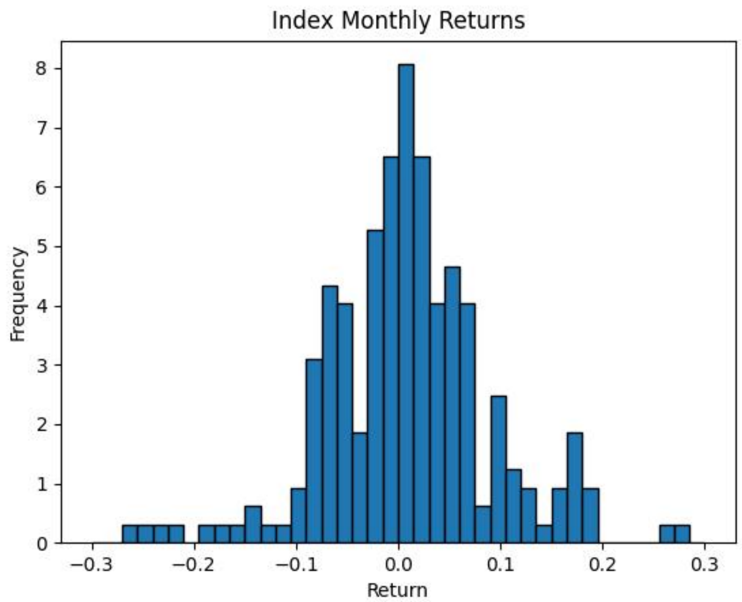
(a)

Summary the statistics, and the results are as followed.

Mean	Standard deviation	Skewness	Kurtosis
0.009042	0.081745	0.018308	1.393663

(b)

Graph as followed



(C)

Simply seeing graph from (b), The skewness value in figure (b) is nearly zero, indicating a nearly symmetrical distribution similar to the normal distribution. The kurtosis value is significantly lower than that of a normal distribution, indicating that the data follows a heavy-tailed distribution.

However simply seeing from the figure cannot assure whether it is normal distribution, so i used shapiro-wilk test to check. In the test H0 is normal distribution, and when p is significant(<0.05), the distribution is not normal. The statistics are as followed.

Shapiro-wilk stats	P-value
0.97128422	0.000228017

So the distribution is NOT normal.

Problem 2

Data Collection

The data are downloaded from CSMAR as required, Returns Without Cash Dividend Reinvested and Market Type from 2017-01 to 2022-53.

Data manipulation

As necessary, the time period should be separated into three parts. In this question, two years will represent one regression period: 2017-2018 (first regression), 2019-2020 (second regression), and 2021-2022 (third regression).

Firstly use .concat() to join the 2 csv files(as downloaded week and week1 files) together. Then select the "1,4,64" market types out as A share mainboards are required. Then groupby by week and calculate market return as required.

Replicate Table 2 from the paper

Firstly, we regress each stock return on the market return in order to calculate the beta in single component models for the first time period. Then, after sorting and rating beta, we divide the stocks into ten categories. We build ten portfolios by merging all of the equities in each group. Then we calculate the difference between the return and the risk-free rate. Finally, we may extract regression results in Table 2 using the second period data.

Regression results of 2nd period data as followed.

Group	Alpha	Alpha t	Alpha p	Beta	Beta t	Beta p	R^2

1	-0.00127	-1.98316	0.05009	0.79934	39.08258	2.18266e-62	0.93855
2	-4.9e-05	-0.09286	0.92619	0.88132	52.52708	1.25029e-74	0.96502
3	0.00021	0.49203	0.62378	0.89614	65.38391	7.18781e-84	0.97714
4	0.00043	0.99721	0.32107	0.94753	69.04839	3.46075e-86	0.97946
5	-1.07e-05	-0.03337	0.97345	1.00091	98.40713	2.38398e-101	0.98978
6	-6.60e-05	-0.16987	0.86545	1.04398	84.57895	7.51855e-95	0.98621
7	-0.00021	-0.57567	0.56613	1.04015	92.02092	1.81869e-98	0.98833
8	-0.00013	-0.27314	0.78531	1.10491	73.84576	4.75962e-89	0.98199
9	0.00030	0.674254	0.50171	1.08707	75.82531	3.53728e-90	0.98290
10	-0.00095	-1.66571	0.09890	1.15054	63.23312	1.88473e-82	0.97560

Discussion:

Compared with the paper, the results are slightly different:

Beta for all portfolios are quite similar and consistent, around 1. and all the p value are significantly small, which means that the effect of market return on stock return is significant. Aligns with the reference paper.

For R squired values, they are all high and no tendency. Indicating that the returns explains most of the regression model. No increase with higher beta, showing that factors of systematic risk influence stock returns. Aligns with the paper

The alpha values are small, around zero. And considering t and p value, there is no evidence to reject the H0 of alpha = 0, which **does not align** with paper.

Replicate table 3

For the 3rd period of 2021-2022, take out the data and calculate the average excess return, then do the regression as required.

The regression results are as followed.

Gamma0	Gamma 1	Gamma0 t	Gamma1 t	R^2	F stats	P
-0.00108	0.00267	-1.05731	2.6088	0.45967	6.80583	0.03119

Discussion:

The value of gamma1 is 0.00267, and t-statistics is greater than 2, p-value is 0.03119 (smaller than 0.05) as well, showing significant positive correlation between returns and systematic risk. This suggests that returns tend to rise with increasing systematic risk, consistent with the principles of the CAPM model, and aligns with the conclusions drawn in the reference paper.

The R-squared value stands at only 0.45967, which is small, similar to the explaining power in the paper.

The alpha values obtained are notably small, approaching zero. The t-value is rather close to 0 (greater than -2), p-value is 0.32, so we fail to find sufficient evidence to reject the H0 that alpha equals zero. This **contradicts** the statements in the reference paper.

Possible reason:

The Chinese stock market has grown rapidly in the years since the report was published, becoming significantly more efficient over time. The growth of a more complete market system, along with updates to the stock issuance process, has helped the overall market mature.

As a result, the balance between risk and return in investments is now more noticeable, with non-systematic risks having less of an impact on pricing. As a result, alpha does not vary significantly from zero.