

Question 1.

Part a.

Here is the list of parameters for FF 5 factor model:

	NoDur	Durbl	Manuf	Enrgy	Chems	BusEq	Telcm	Utils	Shops	HLth	Money	Other
const	-0.133777	-0.055543	-0.124963	-0.102358	-0.218807	0.470277	-0.089466	-0.023175	-0.035256	0.073078	-0.004906	-0.286060
Mkt-RF	0.828660	1.279552	1.151876	0.970547	0.939581	1.117325	0.938007	0.579391	0.966431	0.840885	1.080923	1.070114
SMB	-0.022043	0.349031	0.237822	0.126802	0.015147	0.025131	-0.246775	-0.182297	0.168454	-0.119621	-0.094894	0.194890
HML	-0.094682	0.137872	0.109825	0.537929	-0.046970	-0.432939	-0.018203	0.130547	-0.128320	-0.321886	0.777203	0.107902
RMW	0.557631	0.286298	0.328958	0.095241	0.488445	-0.381526	-0.180245	0.169761	0.482251	0.195381	-0.052916	0.149575
CMA	0.416049	-0.019044	0.142897	0.328097	0.384911	-0.336169	0.291166	0.417191	-0.043433	0.422007	-0.341680	0.047918
Mom	0.008286	-0.338263	-0.071108	0.024384	-0.021783	-0.087325	-0.070522	0.077742	-0.028006	0.063848	-0.046367	-0.042505

Using the 5 factor mode has increase the regression model in the follow in ways:

1. Introducing more important factors(base on t-test),which increase the amount of variance that get explained, results in an overall increase in adjusted R square:

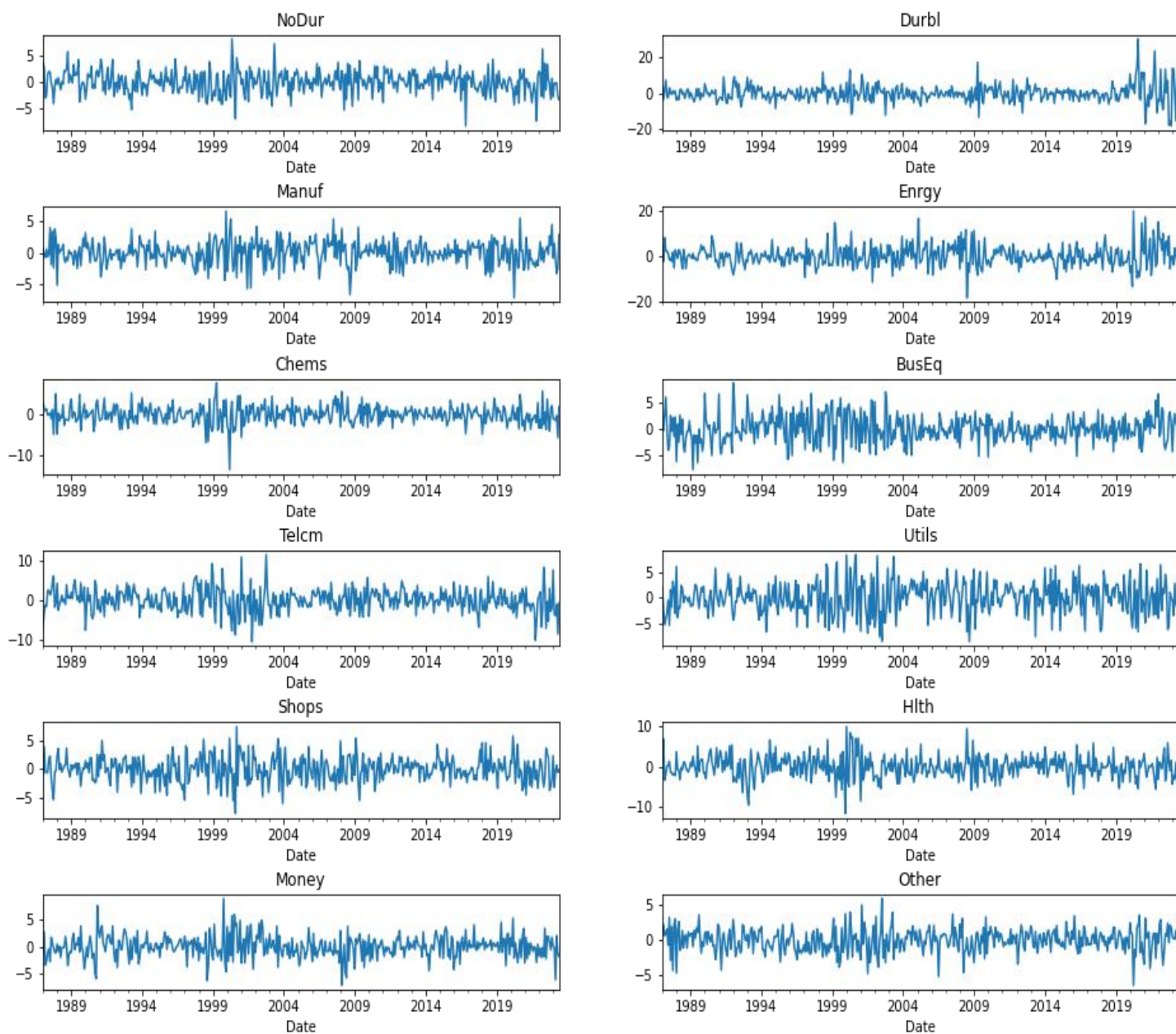
	NoDur	Durbl	Manuf	Enrgy	Chems	BusEq	Telcm	Utils	Shops	HLth	Money	Other
3_factor	0.631364	0.624635	0.864756	0.478463	0.691661	0.841015	0.641637	0.334088	0.751367	0.575700	0.862167	0.891006
5_factor	0.719642	0.627392	0.881618	0.483154	0.747308	0.855796	0.654965	0.356361	0.793695	0.594088	0.868995	0.895606

2. The 5 factor model result in an increase in log-likelihood:

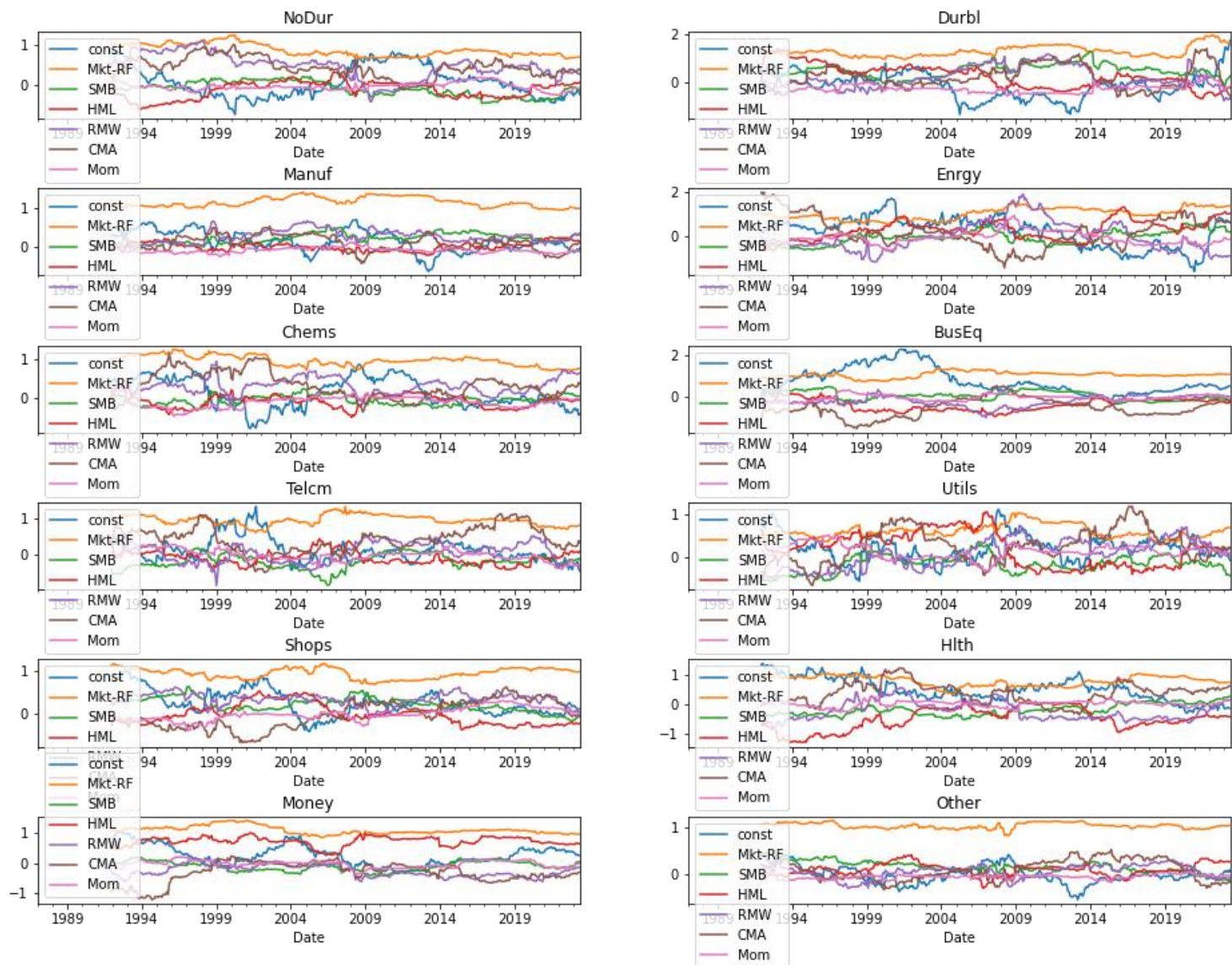
	3_factor	5_factor
NoDur	-1013.600009	-952.636650
Durbl	-1323.809579	-1321.181482
Manuf	-932.151682	-901.974471
Enrgy	-1293.246847	-1290.254439
Chems	-1025.897531	-981.296363
BusEq	-1059.102885	-1036.718269
Telcm	-1112.290282	-1102.976310
Utils	-1137.689303	-1129.225286
Shops	-1014.261559	-972.377026
HLth	-1095.207982	-1084.491678
Money	-943.402136	-931.260574
Other	-854.895311	-844.439334

The Beta estimate is not constant over time. From the residual vs time plot we can see the variance of residual is not constant across time. Also when applying 60 months rolling regression and plot beta vs time, we can clearly see there is a change in the beta estimation.

Plot for residual vs time:



Plots for rolling regression betas:



Part b.

In order to find a more stable estimation of beta, initially I have tried to apply ridge regression and a weighted least square regression. For ridge regression, it increase the beta's stability by shrinking the beta estimation itself. Mostly ridge regression is applied when there exists multicollinearity between factors. However, in our case the estimation is not stable because the residual variance have a time dependency. So, I decide to go with weighted least square where I use the rolling 24 month realized volatility as weights to resolve the time dependent variance cause by different business cycle.

The below photos shows is the variance of rolling beta estimates for ols and wls.(same set up as the rolling regression showed in part a)

	NoDur_ols	NoDur_wls		Telcm_ols	Telcm_wls
const	0.371369	0.394705	const	0.330286	0.282545
Mkt-RF	0.148794	0.148355	Mkt-RF	0.138012	0.130780
SMB	0.176681	0.175525	SMB	0.204349	0.174052
HML	0.205750	0.197253	HML	0.181686	0.188130
RMW	0.303005	0.295474	RMW	0.269498	0.261818
CMA	0.236921	0.216006	CMA	0.352279	0.323143
Mom	0.087991	0.089257	Mom	0.149298	0.149576
	Durbl_ols	Durbl_wls		Utils_ols	Utils_wls
const	0.535954	0.411377	const	0.304926	0.324994
Mkt-RF	0.228271	0.172989	Mkt-RF	0.154634	0.145796
SMB	0.275438	0.260368	SMB	0.205299	0.190891
HML	0.484536	0.414719	HML	0.396881	0.383845
RMW	0.439823	0.364188	RMW	0.325109	0.310144
CMA	0.493418	0.386466	CMA	0.392527	0.381939
Mom	0.163961	0.130393	Mom	0.161262	0.154461
	Manuf_ols	Manuf_wls		Shops_ols	Shops_wls
const	0.272494	0.236947	const	0.301367	0.236117
Mkt-RF	0.127580	0.112405	Mkt-RF	0.123076	0.125186
SMB	0.120082	0.108872	SMB	0.156349	0.154514
HML	0.108372	0.106225	HML	0.236355	0.230652
RMW	0.163163	0.151016	RMW	0.138934	0.128193
CMA	0.199141	0.214842	CMA	0.320804	0.334468
Mom	0.068047	0.063553	Mom	0.146066	0.159250
	Enrgy_ols	Enrgy_wls		Hlth_ols	Hlth_wls
const	0.668009	0.527399	const	0.397790	0.323130
Mkt-RF	0.259637	0.248387	Mkt-RF	0.172383	0.171637
SMB	0.305955	0.316320	SMB	0.195924	0.192666
HML	0.428149	0.408470	HML	0.387189	0.371609
RMW	0.659222	0.616141	RMW	0.318290	0.305662
CMA	0.699540	0.717445	CMA	0.351556	0.320044
Mom	0.274030	0.269755	Mom	0.101025	0.090697
	Chems_ols	Chems_wls		Money_ols	Money_wls
const	0.380727	0.388047	const	0.332574	0.275528
Mkt-RF	0.137389	0.150272	Mkt-RF	0.132135	0.120604
SMB	0.121260	0.117603	SMB	0.155571	0.158232
HML	0.156745	0.153850	HML	0.151601	0.142850
RMW	0.203214	0.236544	RMW	0.190200	0.187714
CMA	0.333887	0.346032	CMA	0.341672	0.355090
Mom	0.136523	0.139343	Mom	0.113298	0.116630
	BusEq_ols	BusEq_wls		Other_ols	Other_wls
const	0.627563	0.512808	const	0.190854	0.207632
Mkt-RF	0.138815	0.135935	Mkt-RF	0.061052	0.053128
SMB	0.220026	0.218984	SMB	0.124267	0.123114
HML	0.279741	0.288347	HML	0.129590	0.118455
RMW	0.246559	0.226584	RMW	0.146045	0.151834
CMA	0.433272	0.423252	CMA	0.215714	0.206052
Mom	0.148679	0.148760	Mom	0.045996	0.044542

From the result using wls result in a decrease in variance for most of the beta estimates. However, the reduction is not constant across all industries and factors. This can be caused by different industry and different factor follow different business characteristic. Using the grouping in Question2, industry in Cyclical tends to have a little more reduction in betas standard deviation than stock in Defensive.

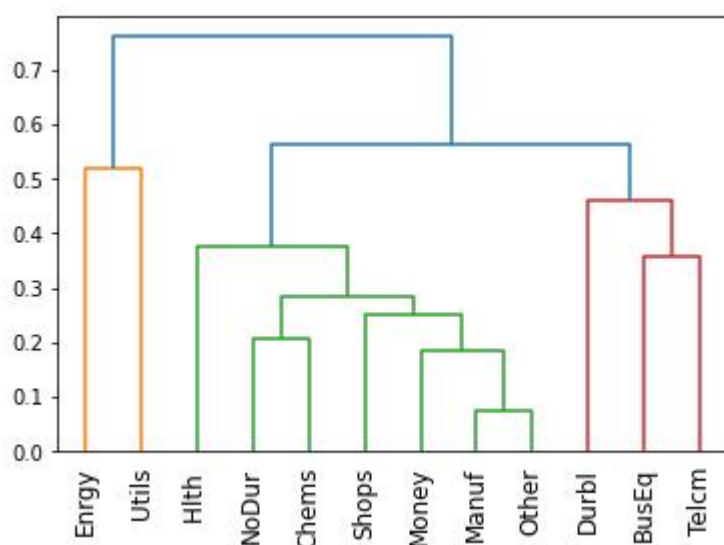
Question 2:

Cyclical stock tend to move follow the health of the over economy. Typically Cyclical stock' beta with market return is greater then 1

Defensive stock tends to provide stable earning regardless of the state of overall economy. Typically Defensive stock with market return is smaller then 1

In order to split the stock into groups. I have tried kmeans on return, hierarchy correlation on pnl and kmeans on Fama-French betas. Kmeans on Fama-French return the best result.

For hierarchy correlation, the result tend to put Energy and Utils in to one subgroup not because they are highly correlated, but because they are not as correlated with other factors. This results does not accurately describe the cluster shown in fundamental analysis



Two groups k means also have the same problem as hierarchy correlation where Enrgy,Telcm,Utils tend to form its own group because they are not correlated with others. Here is a chart of this group

	Enrgy	Utils
Enrgy	1.000000	0.585076
Utils	0.585076	1.000000

The two groups k means on betas ends up returning the best result on classifying Cyclical vs defensive industry where all relative Cyclical industry (market return beta greater than 1) are put into one group

	Durbl	Manuf	BusEq	Shops	Money	Other
Durbl	1.000000	0.831043	0.762587	0.779809	0.766643	0.771442
Manuf	0.831043	1.000000	0.859020	0.836374	0.871396	0.917533
BusEq	0.762587	0.859020	1.000000	0.795535	0.785797	0.814222
Shops	0.779809	0.836374	0.795535	1.000000	0.825638	0.812162
Money	0.766643	0.871396	0.785797	0.825638	1.000000	0.866830
Other	0.771442	0.917533	0.814222	0.812162	0.866830	1.000000

and all relative Defensive industry(market return beta smaller than 1) form another group

	NoDur	Enrgy	Chems	Telcm	Utils	Hlth
NoDur	1.000000	0.610709	0.830573	0.682608	0.706160	0.791365
Enrgy	0.610709	1.000000	0.659765	0.524216	0.585076	0.546846
Chems	0.830573	0.659765	1.000000	0.662353	0.679270	0.756523
Telcm	0.682608	0.524216	0.662353	1.000000	0.621790	0.604098
Utils	0.706160	0.585076	0.679270	0.621790	1.000000	0.615712
Hlth	0.791365	0.546846	0.756523	0.604098	0.615712	1.000000

The component of each group can change over time.

Grouping after 2010:

- 1.Durbl, Manuf, Enrgy,BusEq, Money, Other
- 2.NoDur, Chems, Telcm, Utils, Shops, Hlth

Grouping 2000-2010:

- 1.Durbl, Manuf
- 2.NoDur, Chems, BusEq, Telcm, Utils, Shops, Hlth, Enrgy, Money, Other

Grouping 1990-2000:

- 1.Durbl, Manuf, BusEq, Shops , Money, Other
- 2.NoDur, Chems, Enrgy, Telcm, Utils,, Hlth

Grouping 1980-1990:

- 1.NoDur, Durbl, Manuf, Enrgy, Chems, BusEq, Shops, Hlth, Other
- 2.Telcm, Utils, Money

Also tried a kmeans with 3 groups for betas. However this result in BusEq in a single group, which is not ideal.

Question3

In this question I have tested out 2 different types of portfolio for 2 groups. The first one is the minimum variance portfolio calculated base on historical covariance matrix. The second one is the minimum variance portfolio that calculated base on expanding covariance matrix with a minimum period of 5 years. The result together with equal weighted port folio is shown as below. (Note our portfolio is construct with no target return)

	equal_weight	group1_historical	group2_historical	group1_expanding	group2_expanding
sharpe_ratio	0.612686	0.536287	0.644495	0.516831	0.532755
annulaized_return	0.111715	0.105395	0.096756	0.106521	0.090609
annualized_variance	0.502732	0.533929	0.397911	0.561630	0.441540

For the minimum variance portfolio of group1 we notice it explore a higher annualized variance than equal weight portfolio in both historical weights and expanding weights. The major reason is that the industry in group1 tends to have a high volatility. The equal weighted portfolio for group1 have a variance of 0.62346 which is much higher than the optimized portfolio. The expanding portfolio perform the worst. The reason is because a historical portfolio have a look ahead problem where we construct a portfolio in history with the information from the future, which is not doable in real life. Even with out a look ahead the expanding portfolio perform better than equally weight group1 portfolio.

For minimum variance portfolio of group 2, it has the lowest variance among group1,group2 and equal weighted portfolio. The expanding portfolio perform worse than the historical on for the same reason as group1. Portfolio for group2 also tend to have the lowers annualized return, and this is because the annualized return from a equal weighted group2 portfolio is the lowest which is 0.10087.

From the chart we can tell the equal weight portfolio tend to have a better return than both group1 and group2 portfolio because the optimized target for these portfolio is to minimize variance without considering lost in returns. This can be avoid by adding a target return as in mean-variance portfolio.