Financial Homework

Mattis GILBERT

April 2021

1 Question 1

I started by merging the two datasets. To do so I made sure that the variables on which to merge had the same type as well as names. I then deleted observations that had RET as NA. The reason is threefold: first it facilitates the construction of the function computing the Prospective utilities, second the NAs were mainly concentrated at the first and last dates of PERMNO's observations (there is thus very few "holes", and third it is an assumption that individuals use the returns observed to form their prospective utilities. The only issue is when computing the returns at t+1 where in case of "holes" the returns are shifted. However, the occurrence is seldom and does not impact results.

To compute the Prospective utilities, I built a function looping over PERMNOs. The main function takes a PERMNO, for which the observations were ordered by date. The function selects the slices of 60 observations and for each slice computed the Prospective utility for the 60th observation. It then returned a data-frame with all the observations that had at least 59 preceding observations, containing their Prospective utility as well as all other variables. I could then loop this function over all the PERMNOs of the dataset, row-binding all the resulting data-frames into one final data-frame.

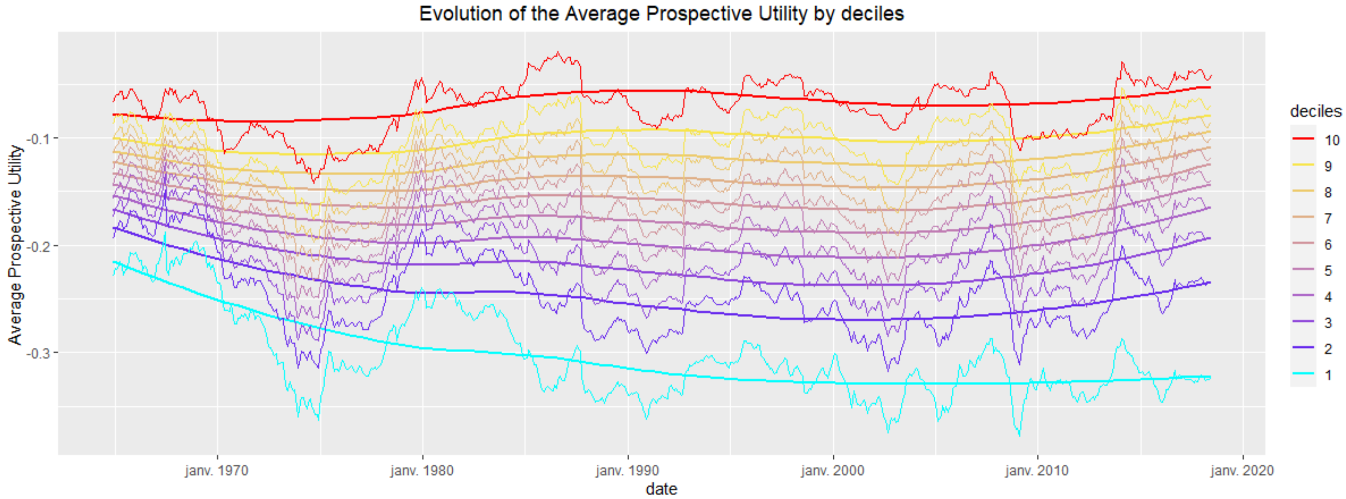
Using parallelization in R the computing time was around 1h40mn. The function I built allowed me to directly choose the values of the different parameters (gamma, delta, alpha) which was efficient to compute the data-frames of question 2 and 3. This resulted in data-frames of 2,421,343 observations over 19,341 PERMNO with an average of 125 months per PERMNO. The oldest month is December 1964, the most recent is June 2018.

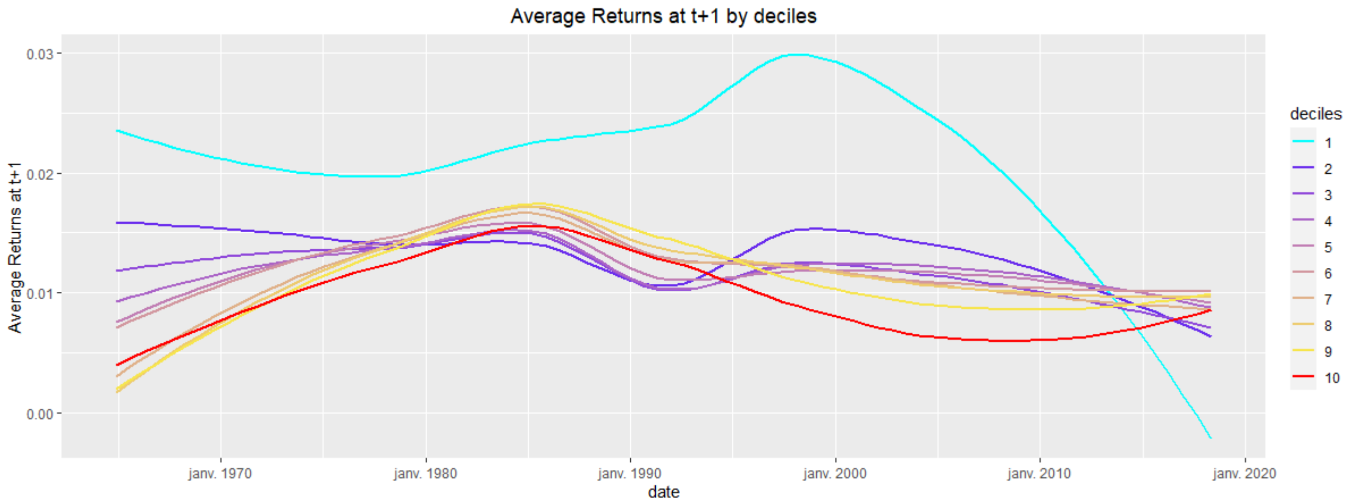
Here are some summary statistics of the data :

deciles	mean_U	mean_RET	mean_RET_t1	mean_ri_rf	mean_ri_rf_t1	mean_PRC
1	-0.3129	-0.0136	0.0207	-0.0171	0.0171	4.0555
2	-0.2526	0.0085	0.0131	0.0050	0.0096	7.4587
3	-0.2217	0.0124	0.0119	0.0088	0.0083	11.6058
4	-0.1976	0.0145	0.0123	0.0110	0.0088	16.4088
5	-0.1768	0.0162	0.0123	0.0127	0.0087	21.7714
6	-0.1577	0.0161	0.0125	0.0126	0.0089	26.6279
7	-0.1394	0.0165	0.0119	0.0130	0.0083	32.5202
8	-0.1210	0.0177	0.0121	0.0141	0.0086	58.6654
9	-0.1006	0.0184	0.0114	0.0148	0.0078	97.8864
10	-0.0674	0.0209	0.0095	0.0174	0.0060	89.0438

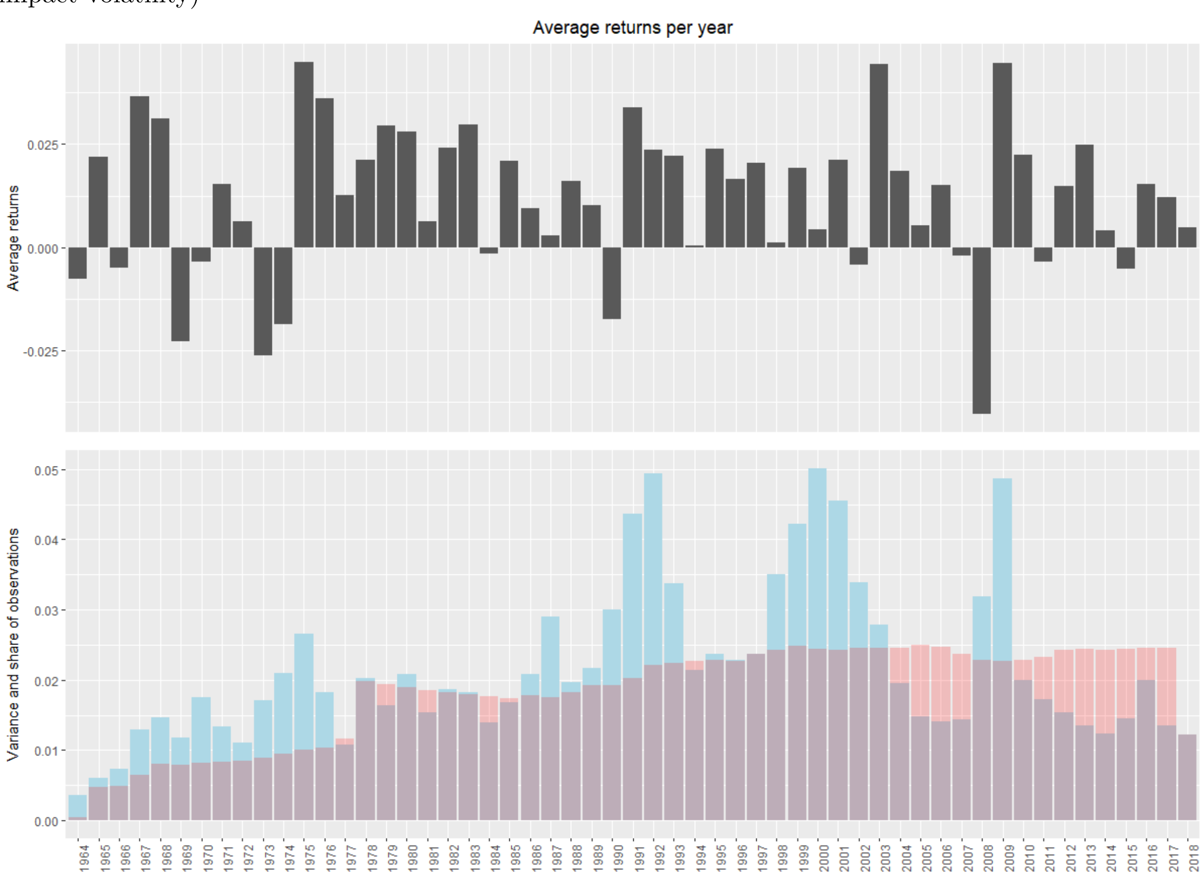
(where ri_rf is the excess return, and "U" stands for Prospective Utility)

I also add some graphs presenting the evolution of the prospective utility by decile over time, as well as the average returns at t+1 of those deciles. From this second graph we clearly see that going long on the first decile and short on the tenth decile seem to make sense.





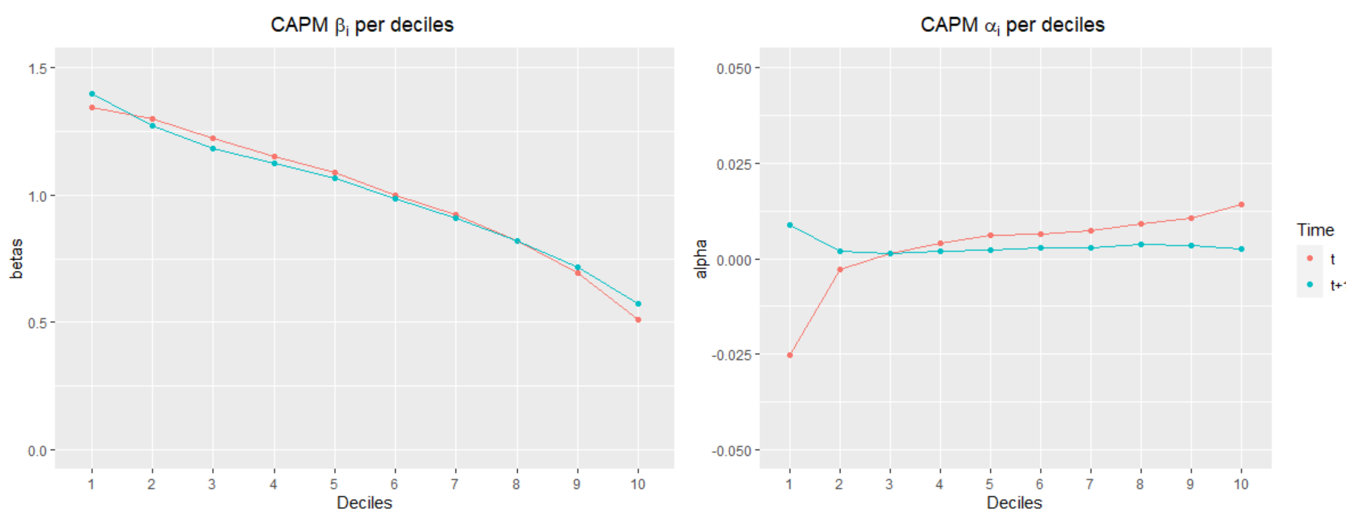
This additional graph shows the average returns per year for the whole sample, as well as the variation of returns in that year. I moreover displayed the share of observations in each year (as it might statistically impact volatility)



In the bottom graph the variance of returns in each year is displayed in blue. The share of observations for a given year is displayed in red (as might explain size of variance)

I computed the deciles based on the prospective utilities. Defined the variables of market-capitalization as SHROUT * PRC. I then took the log of it for the values of market-cap ≥ 1 (NA otherwise). To compute the CAPM I regressed the excess return (ri_rf) on the market excess return (rm_rf) for each decile and extracted the alpha and beta coefficients. This results in the following table, where RET_t1 stands for "returns at t+1" and "alpha" and "betas" are computed at time t. I did append graphs of the evolution of the CAPM alpha and beta with deciles at time t and t+1.

deciles	E_RET_t1	E_ri_rf	E_rm_rf	E_logcap	alpha	betas
1	0.0207	-0.0171	0.0060	9.9209	-0.0252	1.3426
2	0.0131	0.0050	0.0060	10.7770	-0.0028	1.2980
3	0.0119	0.0088	0.0060	11.3270	0.0015	1.2204
4	0.0123	0.0110	0.0060	11.7757	0.0041	1.1490
5	0.0123	0.0127	0.0060	12.1593	0.0061	1.0877
6	0.0125	0.0126	0.0060	12.5069	0.0065	1.0005
7	0.0119	0.0130	0.0060	12.7453	0.0074	0.9206
8	0.0121	0.0141	0.0060	12.9080	0.0092	0.8181
9	0.0114	0.0148	0.0060	12.9945	0.0107	0.6932
10	0.0095	0.0174	0.0060	12.8330	0.0143	0.5119



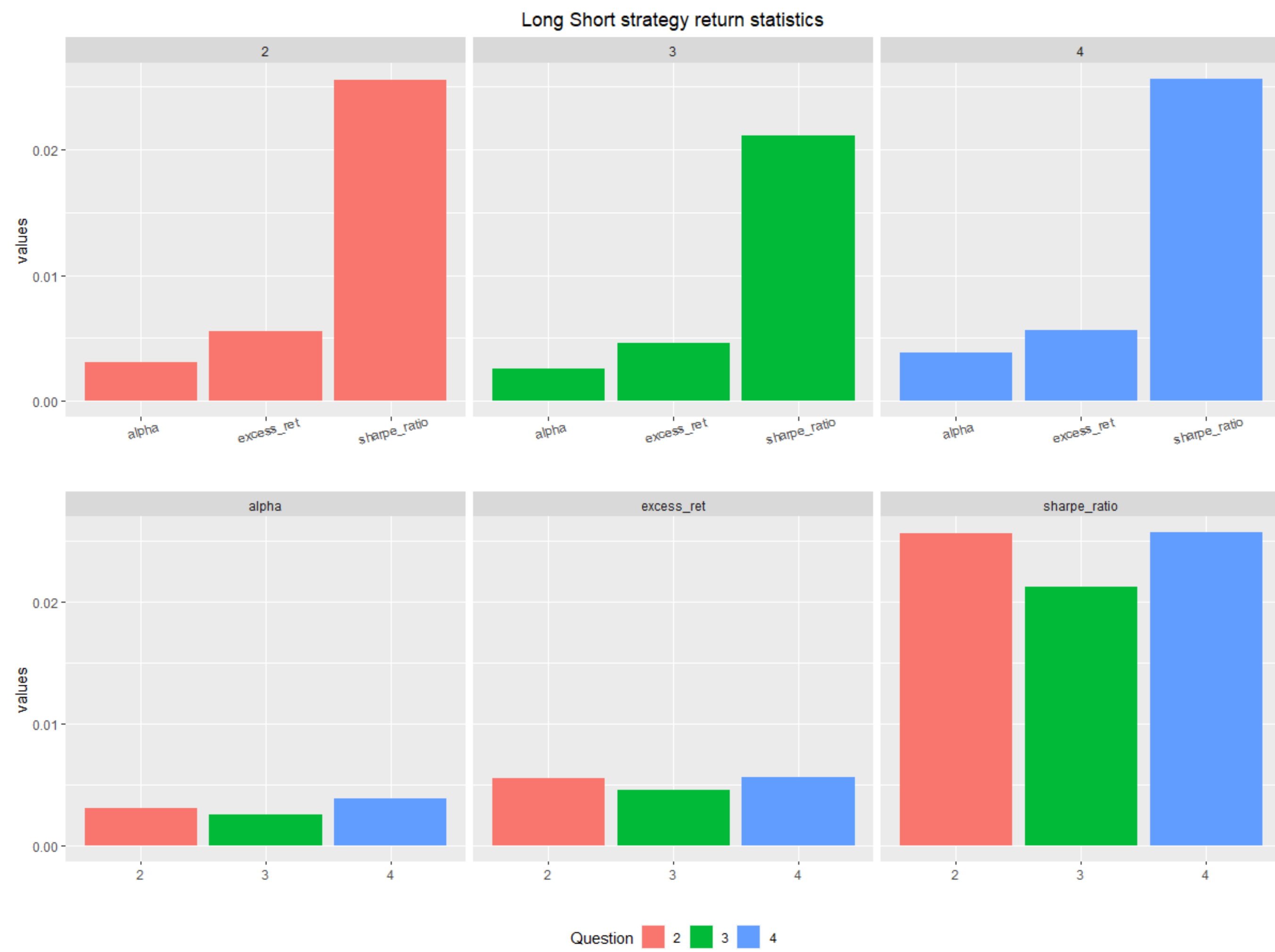
2 Question 2-3-4

For questions 2 and following, I coded a function that takes a data-frame (with Prospective Utilities) and return a table with Sharpe ratio, excess returns and CAPM alpha. The function computes the quantiles at every date, create the variables at t+1 (for returns, risk-free, market excess return) and then filter to keep only the observations in 1^{st} and 10^{th} deciles. The returns of the Long-Short strategy are computed as follows: The 1^{st} decile is short and thus its excess returns at t+1 are RET_T1 - rf_T1 (returns at t+1 – risk free at t+1), the 10^{th} deciles is long and its excess returns at t+1 are rf_T1 - RET_T1. The Sharpe ratio is defined as the average return of the strategy at t+1 divided by its standard deviation.

The results are summarized in this table:

	sharpe_ratio	excess_ret	alpha
Q2	0.02556	0.00552	0.00309
Q3	0.02117	0.00458	0.00255
Q4	0.02565	0.00562	0.00389

I here adjoin charts visually summarizing those results:



We observe that individuals perceiving probabilities correctly tend to perform worse in all measures compared to individuals with distorded perceptions. However, we observe that individuals perform better in every metrics than the when they only exhibit loss aversion.