Mean reversion strategy

Jaakko Wallenius

Project work, Quantitative Finance (28E35600)
Aalto University

Contents

1 Strategy

The investing strategy developed in this paper was formed coincidentally. Before doing any data analysis, the premise was in fact the opposite of what turned out to work. In the spirit of well known momentum studies (De Bondt & Thaler 1985, Jegadeesh & Titman 1993), the plan was to implement and test a momentum strategy. In short, **buy winners and sell losers**.

In contrast to the well known momentum strategies, I planned to implement a more frequent rebalancing (monthly rebalancing). In the first runs with data, it became clear that **buying losers and selling winners** was superior strategy compared to the converse. In fact, shorting the past winners would not be ideal since the past winners tend to perform well in the short term future (but not as well as the losers).

Because the working name "momentum strategy" didn't live up to its expectations, I had to come up with a new name. In a sense the loser stocks "revert to their mean performance". From here on the strategy of sorting stocks based on their last month's performance and the portfolios formed based on this sorting are referred as **mean-reversion strategy** and **mean-reversion portfolios**.[1] In the full sample that I use for backtesting (restricted CRSP universe from 1999/01 to 2019/12, see more on section Data), the loser portfolio (1st quantile portfolio) outperforms all the other quantiles (see Figure 1).[2]

The phenomenon seems to be persistent (i.e. not just feature of the specific sample and partition). As we increase the number of quantiles (equally increase the number of portfolios) from 5 to 10, the 10-Quantile losers portfolio performs better than the 5-Quantile loser portfolio. The same applies when we increase the number of quantiles from 10 to 20 and from 20 to 50 (see Table 1 and Figure 3). Note that we have:

$$P_{1,Q-50,t} \subset P_{1,Q-20,t} \subset P_{1,Q-10,t} \subset P_{1,Q-5,t}$$

where $P_{1,Q-K,t}$ is the K-Quantile loser portfolio at time t.[3] In the 5-Quantile partition loser and winner portfolios are the most volatile of the portfolios (see Figure 2). This applies also to the higher K-Quantile partitions.

[1] I use words **performance** and **returns** interchangeably. Also to be precise, "quantile" refers to the breakpoint value but I use the word sometimes to refer to the interval between two quantiles (i.e. portfolio).

[2] In the rest of this paper, I may refer the 1st quantile portfolio as the **loser portfolio** or in short **losers**. Similarly the last quantile portfolio may be called the **winner portfolio** or **winners**.

[3] The statement tells that all of the stocks in portfolio $P_{1,Q-50,t}$ are in $P_{1,Q-20,t}$ and so on, i.e. they are subsets.

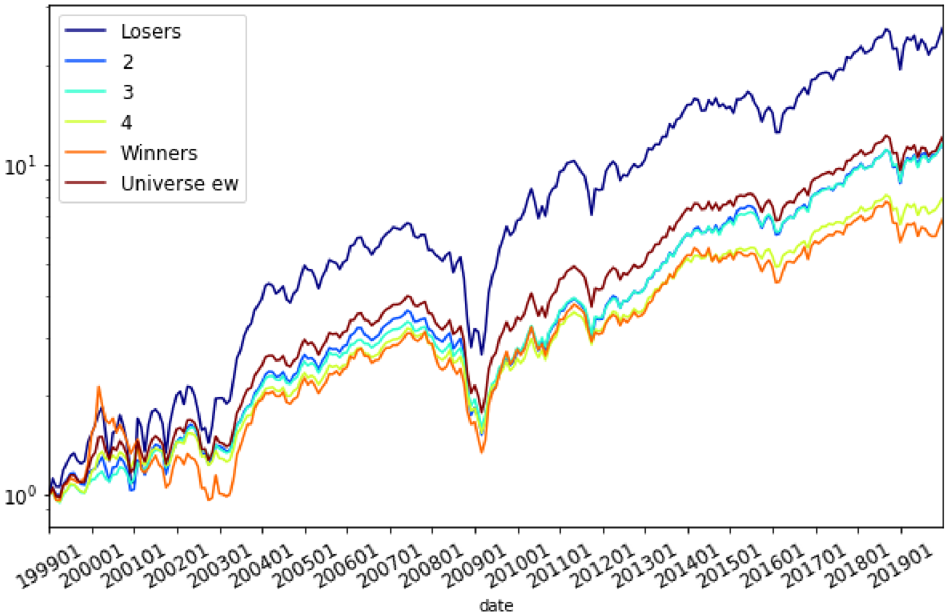


Figure 1: **Cumulative returns (log scale) of 5-Quantile mean reversion portfolios**. 1-Quantile portfolio is named "Losers" (stocks with the worst performance in the last month). Similarly 5-Quantile portfolio is "Winners" (stocks with the best performance in the last month). Portfolios are rebalanced monthly and equally weighted. Portfolios are monitored from 1999/01 to 2019/12 using restricted CRSP data as the investment universe. *Universe ew* is the equally weighted market portfolio of the investment universe. See section Algorithm on how the portfolios are formed and section Data on how the investment universe is restricted.

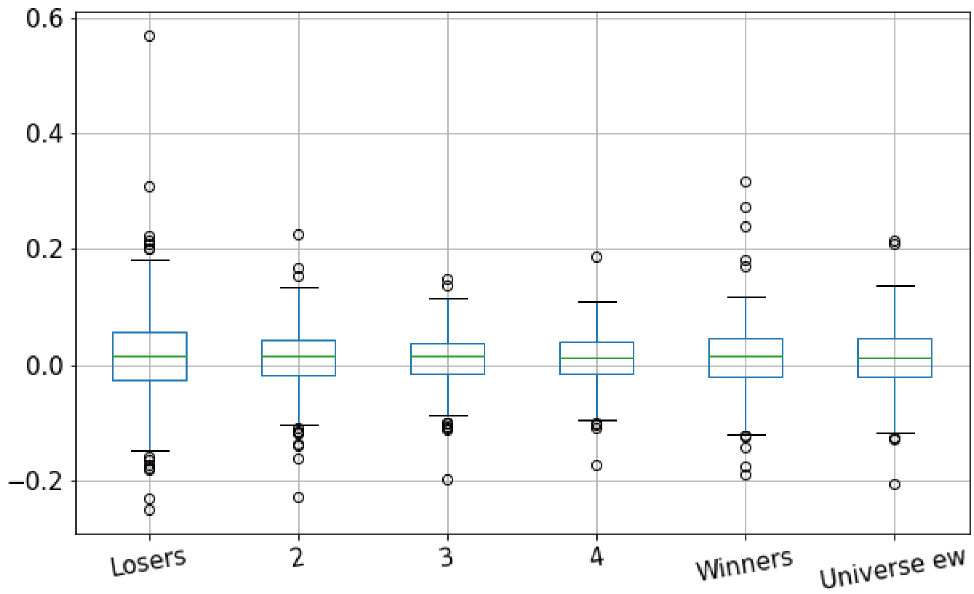


Figure 2: **Box-and-whisker plot of 5-Quantile mean reversion portfolios**. Green lines indicate sample medians. See Figure 1 for definitions.

My investment strategy is to exploit the loser outperformance and further investigate if there exists a feasible way to reduce the portfolio volatility by position sizing[4]

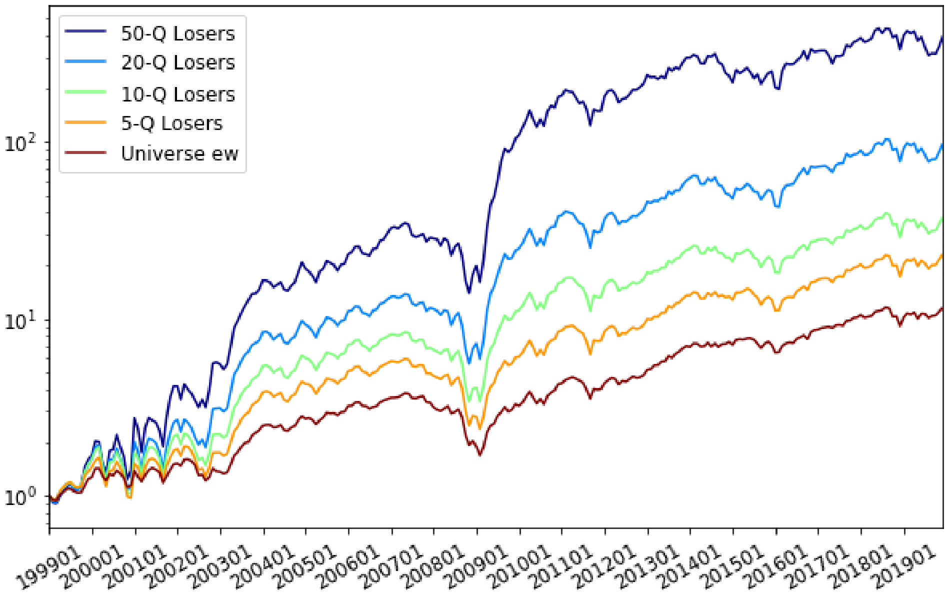


Figure 3: **Cumulative returns (log scale) of K-Quantile partition loser portfolios,** $K = 5, 10, 20, 50$. Portfolios are rebalanced monthly and equally weighted. *Universe ew* is the equally weighted market portfolio of the investment universe. See section Algorithm on how the portfolios are formed and section Data on how the investment universe is restricted.

Table 1: **Summary table of monthly returns of selected** K**-Quantile portfolios (from 1999/01 to 2019/12, 252 months)**. *Avg #* is the average number of positions in the portfolio. *Count* is the number of observations. Returns are **percentages**. See section Algorithm on how the portfolios are formed and section Data on how the investment universe is restricted.

Portfolio	count	mean	std	min	25%	50%	75%	max	Avg #
50-Q Losers	252	3.29	13.58	-27.29	-3.73	1.93	9.25	98.38	78.56
20-Q Losers	252	2.52	11.68	-26.40	-3.27	1.70	7.49	80.42	195.79
10-Q Losers	252	2.00	10.27	-26.45	-2.95	1.81	6.64	70.68	390.94
50-Q Winners	252	1.45	9.70	-33.47	-3.51	1.36	6.00	45.15	78.61
20-Q Winners	252	0.98	5.39	-23.19	-1.90	1.54	4.18	17.84	195.92
10-Q Winners	252	1.36	6.82	-23.53	-2.47	1.75	5.38	35.15	391.27
Universe ew	252	1.16	5.69	-20.56	-2.01	1.19	4.65	21.40	-

[4] It turned out that I was not able find any reasonable way to determine positions sizes such that the risk-reward relation would be enhanced (see section Experimental paths taken).

1.1 Algorithm

The K-Quantile mean-reversion strategy can be compressed into a simple algorithm which filters the investment universe and chooses set of stocks (portfolio) to be bought.[5]

We define step size $t_{k+1} - t_k$ = **a month**, hence holding period is month and holding period return RET_t is calculated from month's first day to its last (where t = month's last day).

1. Sort investment universe into K portfolios by lagged return RET_{t-1} (**in ascending order**)
    - Sorting is based on breakpoints (sample K-quantiles)
    - We have set of portfolios $\{P_1, P_2, ..., P_K\}$
    - P_1 is the portfolio (set of stocks) with the worst performance in the previous month (so called **losers**)
    - P_K is the set of **winner** stocks
2. Buy portfolio P_1 (losers)
    - The positions are **equally-weighted**
3. After month of holding, sell the portfolio and go back to 1. step

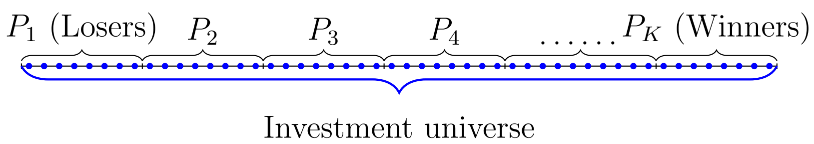


Figure 4: The investment universe is partioned in to K-portfolios after the stocks (blue dots) are ordered ascending by RET_{t-1}.

[5] I programmed a `Python` library (heavily relying on `Pandas`) for this project. See https://github.com/j4bert0/quant-finance-project.

1.2 Long-short portfolios

The K-Quantile mean-reversion strategy has long-short potential.[6] The interesting part is how to choose the shorted portfolio. As noted above, the winner portfolio is not the worst performing. In fact it tends to perform relative well compared to the market portfolio (regardless of the K partition). The worst performing portfolio seems to be the portfolio right behind (or small distance behind) the winners (see Figures 5 for 10-, 20- and 50-Quantile portfolios mean returns).

In general, given the partition K, long-short strategy could be formed by buying long the **Losers** and selling short the **95% portfolio**, which is given by:

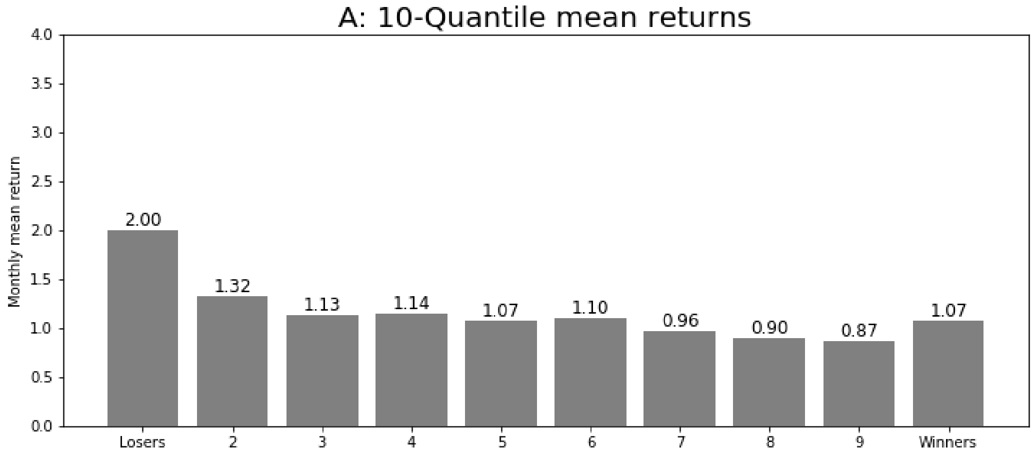
$$\#(P_K) = \lceil K \cdot 0.95 \rceil$$

In the rest of this paper I will further analyze $K = 20$ and $K = 50$ long-short strategies, i.e. shorted portfolio are $\lceil 20 \cdot 0.95 \rceil = P_{19}$ and $\lceil 50 \cdot 0.95 \rceil = P_{47}$, respectively, and the strategies positions are:

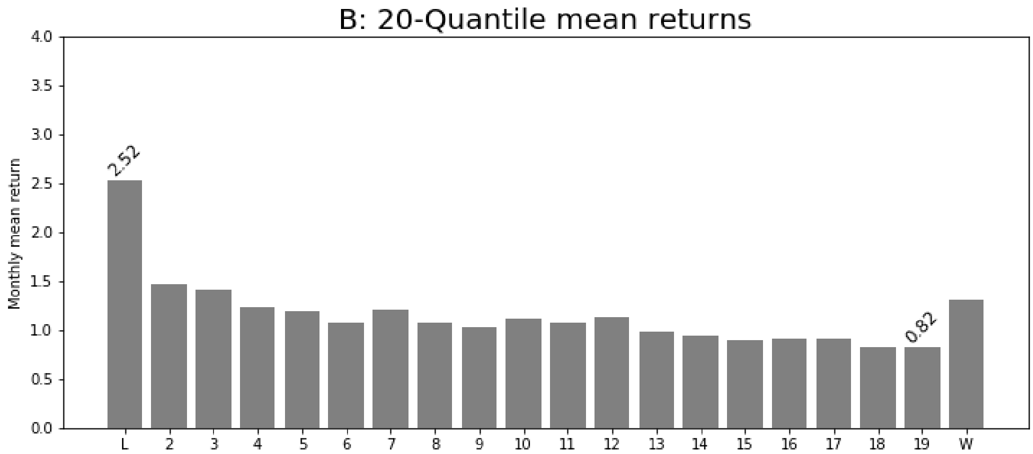
$$\text{20-Quantile long-short:} \quad P_1 - P_{19}$$

$$\text{50-Quantile long-short:} \quad P_1 - P_{47}$$

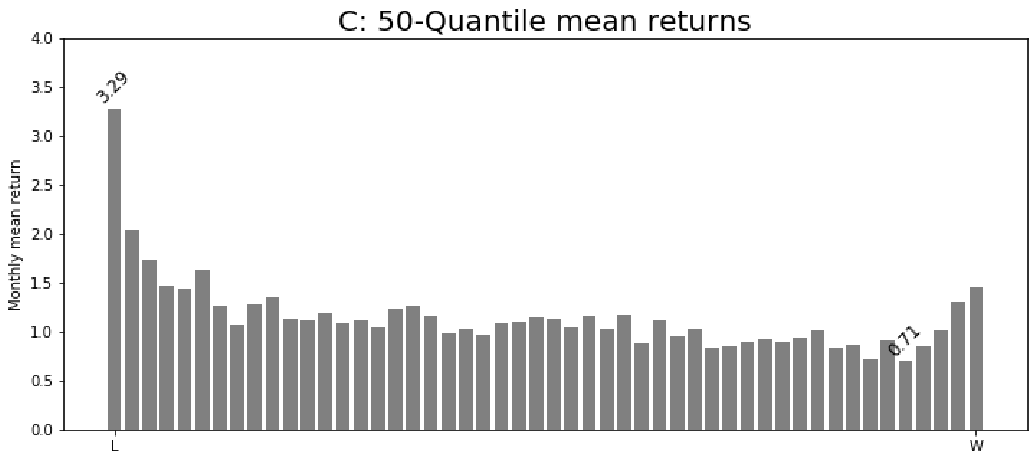
[6]With term long-short portfolio I always refer to **zero-investment** portfolio. Although the zero-investment portfolio may not be applicable in practice it still gives indication how differently set up long-short strategy could work.



(a) **Monthly mean returns of 10-Quantile mean-reversion portfolios**. Best returning portfolio (Losers) has monthly mean return of 2.00%. The worst returning portfolio (9th-quantile) has monthly mean return of 0.87%. Portfolios are monitored from 1999/01 to 2019/12 (see section Data).



(b) **Monthly mean returns of 20-Quantile mean-reversion portfolios**. Best returning portfolio (Losers) has monthly mean return of 2.52%. The worst returning portfolio (19th-quantile) has monthly mean return of 0.82%.



(c) **Monthly mean returns of 50-Quantile mean-reversion portfolios**. Best returning portfolio (Losers) has monthly mean return of 3.29%. The worst returning portfolio (46th-quantile) has monthly mean return of 0.71%.

Figure 5: **Mean return barplots for $K = 10, 20, 50$ Quantile Loser portfolios.**

1.3 Rationale for the strategy

So far the introduced investment strategy is just a pattern of data. It seldom is enough to trust a strategy solely because it has worked in the past. Below is a collection of plausible rationale why the strategy seems to work.

- **Investor overreaction**
  - The phenomenon of investors overreacting has been well documented by the academia for many decades.
  - Because of the overreaction to large price drops, losers may be oversold and thus they rebound quickly to the fair value.
- **Buying losers is uncomfortable**
  - Buying stocks that have lost significant part of their value in the last month is uncomfortable.
- **Popularity of the converse strategy**
  - Since the "buy winners, sell losers" -strategy is so well known, it seems only logical that it will crowd out eventually
  - The good performance of the "buy losers" -strategy might because of this effect.

2 Performance

This section I am going to answers to the most import question of the paper: **how much can we expect profit from the introduced strategies?**[7] The answers is formed by first tabulating strategies' returns for different time periods (Backtesting). Then the returns are critically evaluated by changing some of the assumptions of the investment universe to see if the results are robust (Robustness).

2.1 Backtesting

Table 2 Panel A shows the periodic monthly mean returns for 3-year sub-periods and also for the full backtesting sample. 50-Quantile Losers portfolio has by far the highest monthly mean return (3.29%) for the full sample. Also 50-Quantile Long-short portfolio has very high monthly mean return (2.44%). Table 2 Panel B shows the periodic growth factors of the portfolios. Using the default investment universe (see Data), the 50-Quantile Losers grows by a factor of 481.67 (which is astonishing).

Figures 6 and 7 visualize the backtesting by plotting 3-year runs for 50-Quantile Losers and Winners portfolios. Especially Figure 7 is interesting since it plots all of the 3-year sub-time-serieses for the portfolios. Hence it is kind of a "quasi-simulation" of the portfolio performance.

The periodic performance reveals that strategies have lived their best days in periods 1999/01-2004/12 and 2008/01-2010/12 which both can be considered as "rebound periods" from financial markets crises (dot-com bubble and sub-prime crisis, respectively). Also in the most recent periods (2014/01-2019/12) the strategies have fallen short to the (universe) market portfolio.

2.2 Robustness

I validate the robustness of the investment strategy by altering the data and assumptions. This is done by making two changes to the dataset (investment universe). First, instead of using monthly return data which includes dividends and other distributions, I use data which **does not** include distributions to shareholders. Secondly, I filter out the smallest 30% by lagged market value (ME_{t-1}) from the investment universe instead of the smallest 10% (see section Data on how the default investment universe is structured).

Table 3 shows the periodic performance for the ME filtered investment universe. The mean returns and growth factors are scaled down across the table considerably. Still the relative performance remains: **losers perform better than winners**.

[7] One can argue that "how much risk is involved in the strategy" is the most important question.

Table 2: **Periodic performance of selected K-Quantile mean-reversion portfolios**. Portfolios are rebalanced monthly and equally weighted. *Universe ew* is the equally weighted market portfolio of the investment universe. Portfolios are monitored from 1999/01 to 2019/12 using restricted CRSP data as the investment universe. See section Algorithm on how the portfolios are formed and section Data on how the investment universe is restricted.
The backtesting sample (21 years, 252 months) is partitioned into 7 3-year sub-periods. **Panel A** tabulates the monthly mean return of portfolios for the periods. Entries in Panel A are **percentages**. **Panel B** tabulates the growth factors of portfolios for the periods.

Panel A: Monthly mean returns of 3-year periods

	Periods							FS
Portfolio	'99-'01	'02-'04	'05-'07	'08-'10	'11-'13	'14-'16	'17-'19	'99-'19
50-Q Losers	6.86	5.24	1.11	6.57	1.74	0.62	0.87	**3.29**
20-Q Losers	4.87	4.30	0.63	4.40	1.57	0.77	1.11	**2.52**
10-Q Losers	3.96	3.39	0.41	3.17	1.44	0.52	1.09	**2.00**
5-Q Losers	2.91	2.84	0.40	2.14	1.50	0.65	1.18	**1.66**
50-Q Winners	1.88	2.84	1.02	1.45	1.49	0.55	0.94	**1.45**
20-Q Winners	1.79	2.52	0.93	1.04	1.71	0.26	0.93	**1.31**
10-Q Winners	1.42	2.11	0.68	0.92	1.55	0.20	0.59	**1.07**
5-Q Winners	1.19	1.76	0.73	0.85	1.35	0.38	0.54	**0.97**
50-Q Long-short	5.74	3.42	0.69	5.88	0.52	0.39	0.46	**2.44**
20-Q Long-short	3.83	2.59	0.20	3.61	0.19	0.64	0.86	**1.70**
Universe ew	1.54	1.90	0.55	1.03	1.49	0.67	0.91	**1.16**

Panel B: Growth factors* of 3-year periods

	Periods							FS
Portfolio	'99-'01	'02-'04	'05-'07	'08-'10	'11-'13	'14-'16	'17-'19	'99-'19
50-Q Losers	4.09	5.01	1.38	6.27	1.66	1.08	1.22	**481.67**
20-Q Losers	2.09	3.77	1.18	3.30	1.59	1.19	1.36	**116.30**
10-Q Losers	1.49	2.87	1.10	2.33	1.54	1.12	1.37	**43.26**
5-Q Losers	1.01	2.46	1.11	1.72	1.60	1.20	1.43	**25.92**
50-Q Winners	0.25	2.28	1.38	1.32	1.65	1.11	1.27	**12.18**
20-Q Winners	0.36	2.13	1.35	1.20	1.77	1.03	1.31	**11.28**
10-Q Winners	0.29	1.90	1.24	1.18	1.68	1.02	1.18	**7.31**
5-Q Winners	0.31	1.73	1.27	1.18	1.57	1.10	1.17	**6.83**
50-Q Long-short	3.12	3.07	1.24	6.52	1.15	1.06	1.12	**138.25**
20-Q Long-short	1.49	2.36	1.05	3.20	1.03	1.20	1.32	**32.31**
Universe ew	0.58	1.87	1.19	1.27	1.64	1.23	1.34	**12.12**

***Note**: growth factor is defined:

$$g := \prod_{i=1}^{N} (1 + r_i)$$

where r_i is the return of month i and N is the number of months in the period.

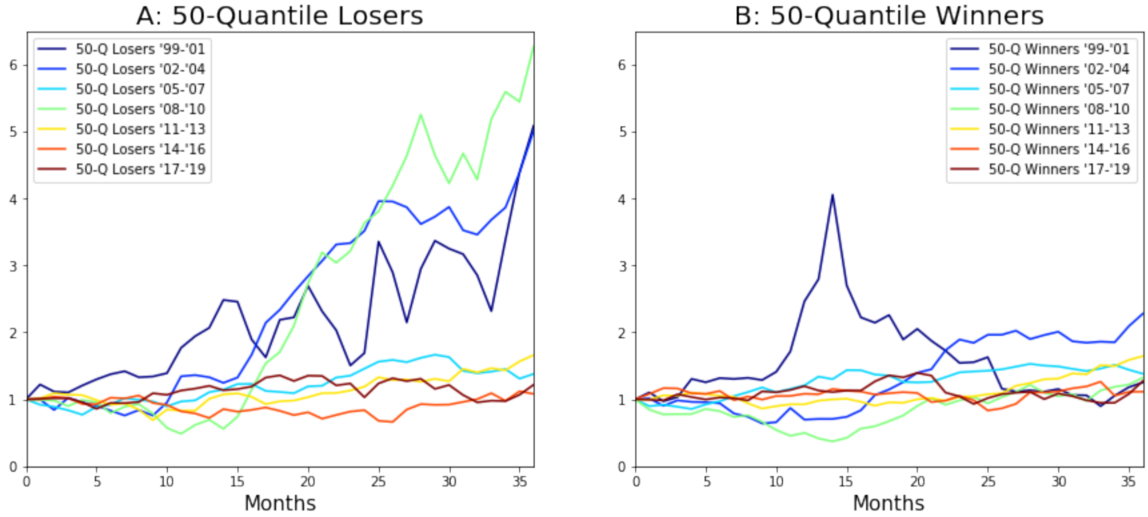


Figure 6: **50-Quantile Losers and Winners, cumulative plot of 3-year runs**. Portfolios are rebalanced monthly and equally weighted. *Universe ew* is the equally weighted market portfolio of the investment universe. See section Algorithm on how the portfolios are formed and section Data on how the investment universe is restricted.

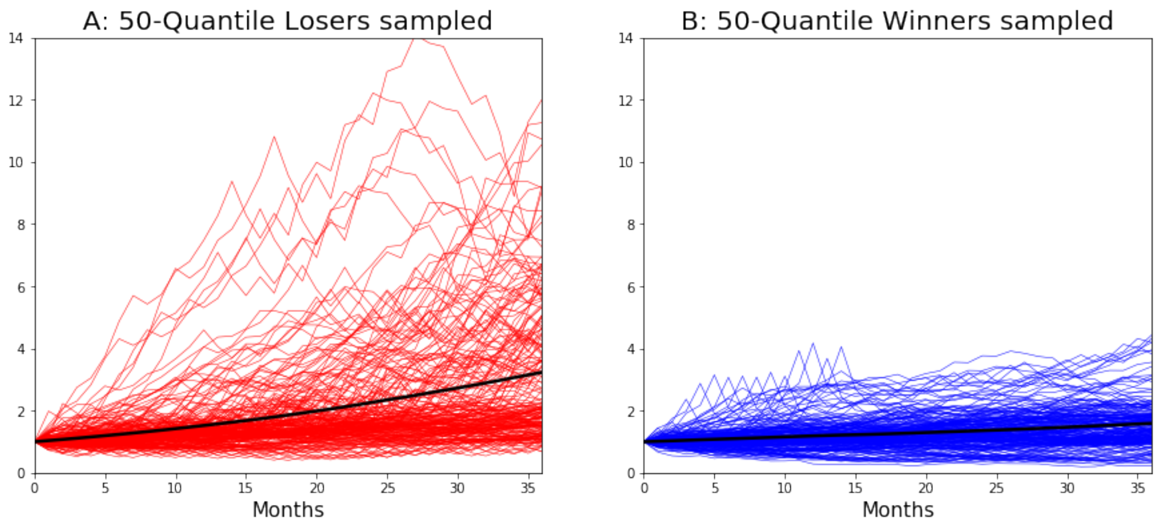


Figure 7: **50-Quantile Losers and Winners, time series sampled**. Portfolios are rebalanced monthly and equally weighted. *Universe ew* is the equally weighted market portfolio of the investment universe. See section Algorithm on how the portfolios are formed and section Data on how the investment universe is restricted.

Table 3: **Periodic performance of selected K-Quantile mean-reversion portfolios (Robustness, ME restricted)**. Portfolios are rebalanced monthly and equally weighted. *Universe ew* is the equally weighted market portfolio of the investment universe. Portfolios are monitored from 1999/01 to 2019/12 using restricted CRSP data as the investment universe. See section Algorithm on how the portfolios are formed and section Data on how the investment universe is restricted. The backtesting sample (21 years, 252 months) is partitioned into 7 3-year sub-periods. **Panel A** tabulates the monthly mean return of portfolios for the periods. Entries in Panel A are **percentages**. **Panel B** tabulates the growth factors of portfolios for the periods.

Panel A: Monthly mean returns of 3-year periods

	Periods							FS
Portfolio	'99-'01	'02-'04	'05-'07	'08-'10	'11-'13	'14-'16	'17-'19	'99-'19
50-Q Losers	4.66	3.04	0.47	4.51	1.09	0.48	0.46	**2.10**
20-Q Losers	2.96	2.73	0.25	2.58	1.18	0.26	1.02	**1.57**
10-Q Losers	2.45	2.32	0.19	1.88	1.30	0.38	1.15	**1.38**
5-Q Losers	1.76	2.05	0.29	1.30	1.37	0.56	1.12	**1.21**
50-Q Winners	2.02	2.42	0.79	0.89	1.44	0.76	0.33	**1.23**
20-Q Winners	1.97	1.80	0.86	0.73	1.61	0.23	0.64	**1.12**
10-Q Winners	1.51	1.60	0.83	0.75	1.45	0.20	0.50	**0.98**
5-Q Winners	1.19	1.29	0.84	0.74	1.26	0.35	0.42	**0.87**
50-Q Long-short	3.85	1.79	-0.32	3.69	-0.30	0.33	0.33	**1.34**
20-Q Long-short	1.91	1.32	-0.55	1.81	-0.11	0.10	0.65	**0.73**

Panel B: Growth factors of 3-year periods

	Periods							FS
Portfolio	'99-'01	'02-'04	'05-'07	'08-'10	'11-'13	'14-'16	'17-'19	'99-'19
50-Q Losers	1.43	2.38	1.10	3.03	1.32	1.02	1.09	**28.04**
20-Q Losers	0.65	2.20	1.03	1.76	1.39	0.99	1.34	**12.13**
10-Q Losers	0.53	1.97	1.03	1.48	1.47	1.06	1.41	**10.04**
5-Q Losers	0.37	1.87	1.07	1.28	1.53	1.15	1.41	**8.75**
50-Q Winners	0.26	1.95	1.27	1.05	1.61	1.20	1.04	**6.56**
20-Q Winners	0.38	1.63	1.32	1.07	1.71	1.01	1.18	**6.47**
10-Q Winners	0.30	1.58	1.31	1.11	1.62	1.02	1.14	**5.60**
5-Q Winners	0.28	1.46	1.32	1.14	1.52	1.09	1.12	**5.20**
50-Q Long-short	1.09	1.75	0.87	3.17	0.85	1.03	1.08	**9.53**
20-Q Long-short	0.25	1.51	0.81	1.73	0.93	0.99	1.24	**2.99**

Note: Here the smallest 30% by lagged market value (ME_{t-1}) are filtered out of the investment universe. In the default investment universe used, smallest 10% are filtered out (see section Data on the investment universe is constructed).

Table 4: **Periodic performance of selected K-Quantile mean-reversion portfolios (Robustness, ME restricted and $RETX$ used).** Portfolios are rebalanced monthly and equally weighted. *Universe ew* is the equally weighted market portfolio of the investment universe. Portfolios are monitored from 1999/01 to 2019/12 using restricted CRSP data as the investment universe. See section Algorithm on how the portfolios are formed and section Data on how the investment universe is restricted.
The backtesting sample (21 years, 252 months) is partitioned into 7 3-year sub-periods. **Panel A** tabulates the monthly mean return of portfolios for the periods. Entries in Panel A are **percentages**. **Panel B** tabulates the growth factors of portfolios for the periods.

Panel A: Monthly mean returns of 3-year periods

	Periods							FS
Portfolio	'99-'01	'02-'04	'05-'07	'08-'10	'11-'13	'14-'16	'17-'19	'99-'19
50-Q Losers	4.65	3.02	0.44	4.42	1.04	0.41	0.40	**2.05**
20-Q Losers	2.94	2.71	0.21	2.50	1.12	0.20	0.95	**1.52**
10-Q Losers	2.42	2.28	0.13	1.79	1.23	0.30	1.06	**1.32**
5-Q Losers	1.71	2.00	0.22	1.20	1.29	0.47	1.02	**1.13**
50-Q Winners	2.00	2.40	0.74	0.83	1.41	0.72	0.31	**1.20**
20-Q Winners	1.94	1.76	0.82	0.66	1.54	0.18	0.59	**1.07**
10-Q Winners	1.48	1.57	0.78	0.67	1.37	0.12	0.43	**0.92**
5-Q Winners	1.14	1.25	0.79	0.65	1.15	0.27	0.34	**0.80**
50-Q Long-short	3.89	1.80	-0.31	3.66	-0.25	0.35	0.35	**1.36**
20-Q Long-short	1.93	1.34	-0.54	1.83	-0.07	0.13	0.67	**0.76**

Panel B: Growth factors of 3-year periods

	Periods							FS
Portfolio	'99-'01	'02-'04	'05-'07	'08-'10	'11-'13	'14-'16	'17-'19	'99-'19
50-Q Losers	1.42	2.36	1.09	2.93	1.29	0.99	1.06	**24.93**
20-Q Losers	0.64	2.18	1.02	1.71	1.37	0.97	1.30	**10.69**
10-Q Losers	0.51	1.94	1.00	1.44	1.44	1.03	1.37	**8.59**
5-Q Losers	0.35	1.84	1.04	1.24	1.48	1.12	1.36	**7.23**
50-Q Winners	0.26	1.93	1.25	1.02	1.60	1.18	1.03	**6.04**
20-Q Winners	0.36	1.61	1.29	1.04	1.67	0.99	1.16	**5.76**
10-Q Winners	0.28	1.56	1.28	1.08	1.58	0.99	1.12	**4.83**
5-Q Winners	0.26	1.43	1.29	1.10	1.47	1.06	1.09	**4.33**
50-Q Long-short	1.11	1.75	0.87	3.13	0.87	1.04	1.09	**9.95**
20-Q Long-short	0.26	1.52	0.81	1.74	0.94	1.00	1.24	**3.16**

Note: Here the smallest 30% by lagged market value (ME_{t-1}) are filtered out of the investment universe. Also returns are computed using CRSP $RETX$ variable (distributions to shareholders not included). In the default investment universe used, smallest 10% are filtered out (see section Data on the investment universe is constructed).

3 Risk analysis

In this section I gauge the strategies' risk with different methods. The methods used can be divided into two classes: **statistical estimates** from portfolios' return time serieses (and risk-return ratios derived from the estimates) and time series regressions with **asset pricing models**.

The first class includes:

- Monthly volatility (standard deviation)
- Monthly semi-deviation (downside volatility)
- Sharpe ratio
- MAR ratio
- Sortino ratio
- Max drawdown (portfolio's maximum drop from the previous peek value)
- min (worst month)
- Skewness
- Kurtosis

See section Definitions on how the measures are defined. Table 5 tabulates these measures for the monitored strategies (portfolios). Table 6 shows the sample correlations of the studied portfolios.[8]

The second class includes:

- Capital asset pricing model (**CAPM**)
- Fama-French 3-factor model (**FF-3**. Includes market, size and value factors)
- Carhart 4-factor model (**CAR-4**. Adds momentum factor to FF-3)
- Fama-French 5-factor model (**FF-5**)

Table 7 tabulates the **risk decomposition** (systematic and unsystematic parts) for the monitored strategies (portfolios). Table 8 tabulates the model **alphas** for the monitored strategies (portfolios).

In addition Figure ref depicts **underwater drawdown plots** for selected strategies. These provide an intuitive way of understanding the downside volatility experienced by the strategies. The underwater drawdown plot (Figures 8, 9) and sampled 3-year runs (Figure 7) are both closely related to the concept of **ruin** (i.e. losing all, or more realistically having to cease operations because of dismal performance).[9]

[8]This is not directly related to risk but still interesting data in my opinion.
[9]See `https://en.wikipedia.org/wiki/Ruin_theory`.

3.1 Risk metrics

Table 5: **Risk metrics (risk-to-reward measures)**. All of the measures all calculated from monthly returns of the full backtesting sample (**default universe** 1999/01-2019/12). See section Definitions on how the measures are defined. Blue indicates the maximum value among the studied portfolios. Red indicates the minimum value among the studied portfolios.

Portfolio	Volatility	Semi-deviation	Sharpe ratio	MAR ratio	Sortino ratio	Max drawdown	min	Skewness	Kurtosis
50-Q Losers	13.58	6.13	0.231	6.17	0.536	-59.79	-27.29	1.88	9.90
20-Q Losers	11.68	5.99	0.203	4.71	0.421	-59.57	-26.40	1.55	8.28
10-Q Losers	10.27	5.95	0.180	3.71	0.336	-59.42	-26.45	1.30	8.20
5-Q Losers	8.67	5.42	0.174	3.25	0.306	-60.02	-25.12	0.94	6.83
50-Q Winners	9.70	6.24	0.135	5.04	0.233	-80.23	-33.47	0.56	3.73
20-Q Winners	8.50	5.38	0.137	3.41	0.244	-71.64	-24.50	0.90	5.86
10-Q Winners	7.54	4.76	0.122	2.41	0.224	-67.13	-22.52	0.82	5.28
5-Q Winners	6.44	4.13	0.128	1.78	0.234	-57.03	-19.00	0.46	3.31
50-Q Long-short	10.49	4.41	0.219	3.77	0.553	-47.60	-24.85	3.05	21.86
20-Q Long-short	8.51	3.93	0.183	2.34	0.433	-40.64	-20.69	3.17	24.64

Table 6: **Correlation table**. Tabulated values are sample **Pearson correlations** of the return serieses of the full sample (**default universe** 1999/01-2019/12). *ewretd* and *vwretd* are equally and valued weighted CRSP market portfolios. *Universe ew* is the equally weighted investment universe market portfolio.

Portfolio	50-Q L	20-Q L	10-Q L	5-Q L	50-Q W	20-Q W	10-Q W	5-Q W	50-Q L-S	20-Q L-S	Univ. ew	ewretd	vwretd
50-Q L	1.00												
20-Q L	0.99	1.00											
10-Q L	0.97	0.99	1.00										
5-Q L	0.95	0.98	0.99	1.00									
50-Q W	0.61	0.64	0.64	0.65	1.00								
20-Q W	0.63	0.66	0.66	0.67	0.97	1.00							
10-Q W	0.65	0.68	0.69	0.70	0.95	0.99	1.00						
5-Q W	0.68	0.71	0.72	0.74	0.92	0.97	0.99	1.00					
50-Q L-S	0.86	0.82	0.79	0.75	0.20	0.18	0.19	0.23	1.00				
20-Q L-S	0.83	0.82	0.80	0.77	0.16	0.14	0.15	0.20	0.96	1.00			
Univ. ew	0.85	0.89	0.92	0.94	0.77	0.82	0.85	0.89	0.52	0.54	1.00		
ewretd	0.87	0.91	0.93	0.95	0.77	0.81	0.84	0.87	0.56	0.57	0.98	1.00	
vwretd	0.70	0.74	0.77	0.80	0.67	0.72	0.75	0.80	0.39	0.41	0.90	0.89	1.00

3.2 Asset pricing models

Below are the tables.

Table 7: **Market risk table (monthly returns)** (CAPM regressions). The backtesting sample (21 years, 252 months **default universe**) is partitioned into 3 cumulative periods ('99-'19, '06-'19 and '13-'19). **Panel A** tabulates regression results for period '99-'19 (full backtesting sample, $N = 252$). **Panel B** tabulates regression results for period '06-'19 ($N = 168$). **Panel C** tabulates regression results for period '13-'19 ($N = 84$). (***), (**) and (*) indicate statistical significance on levels 0.01, 0.05 and 0.10, respectively.

Panel A: Months 1999/01-2019/12 (Full backtesting sample)

Portfolio	α %	α t-stat	β	β t-stat	R squared
50-Q Losers	2.00	3.18 (***)	2.14	14.88 (***)	0.47
20-Q Losers	1.33	2.63 (***)	1.97	17.03 (***)	0.54
10-Q Losers	0.90	2.11 (**)	1.80	18.46 (***)	0.58
5-Q Losers	0.67	2.0 (**)	1.58	20.59 (***)	0.63
50-Q Winners	0.53	1.14	1.45	13.57 (***)	0.42
20-Q Winners	0.43	1.13	1.38	15.72 (***)	0.50
10-Q Winners	0.24	0.74	1.28	17.31 (***)	0.55
5-Q Winners	0.20	0.79	1.17	20.35 (***)	0.62
50-Q Long-short	1.81	2.93 (***)	0.92	6.53 (***)	0.15
20-Q Long-short	1.14	2.3 (**)	0.79	6.95 (***)	0.16

Panel B: Months 2006/01-2019/12

Portfolio	α %	α t-stat	β	β t-stat	R squared
50-Q Losers	0.86	1.52	1.90	14.51 (***)	0.56
20-Q Losers	0.43	0.96	1.77	17.33 (***)	0.64
10-Q Losers	0.11	0.32	1.65	20.9 (***)	0.72
5-Q Losers	0.05	0.19	1.50	25.12 (***)	0.79
50-Q Winners	-0.00	-0.0	1.32	13.91 (***)	0.54
20-Q Winners	-0.06	-0.2	1.29	18.0 (***)	0.66
10-Q Winners	-0.18	-0.71	1.22	20.46 (***)	0.72
5-Q Winners	-0.15	-0.71	1.15	23.51 (***)	0.77
50-Q Long-short	1.09	2.06 (**)	0.71	5.83 (***)	0.17
20-Q Long-short	0.63	1.62	0.62	6.87 (***)	0.22

Panel C: Months 2013/01-2019/12

Portfolio	α %	α t-stat	β	β t-stat	R squared
50-Q Losers	-0.81	-1.15	1.72	8.54 (***)	0.47
20-Q Losers	-0.50	-0.92	1.63	10.36 (***)	0.57
10-Q Losers	-0.54	-1.25	1.55	12.52 (***)	0.66
5-Q Losers	-0.31	-0.93	1.43	14.67 (***)	0.72
50-Q Winners	-0.18	-0.28	1.25	6.8 (***)	0.36
20-Q Winners	-0.32	-0.7	1.28	9.78 (***)	0.54
10-Q Winners	-0.45	-1.24	1.21	11.53 (***)	0.62
5-Q Winners	-0.38	-1.32	1.14	13.57 (***)	0.69
50-Q Long-short	-0.39	-0.57	0.56	2.91 (***)	0.09
20-Q Long-short	0.02	0.05	0.50	3.52 (***)	0.13

Table 8: α **Table**. The backtesting sample (21 years, 252 months **default universe**) is partitioned into 3 cumulative periods ('99-'19, '06-'19 and '13-'19). **Panel A** tabulates regression results for period '99-'19 (full backtesting sample, $N = 252$). **Panel B** tabulates regression results for period '06-'19 ($N = 168$). **Panel C** tabulates regression results for period '13-'19 ($N = 84$). (***), (**) and (*) indicate statistical significance on levels 0.01, 0.05 and 0.10, respectively.

Panel A: Months 1999/01-2019/12 (Full backtesting sample)

Portfolio	CAPM alpha %	FF-3 alpha %	CAR-4 alpha %	FF-5 alpha %
50-Q Losers	2.0 (***)	1.9 (***)	2.44 (***)	2.75 (***)
20-Q Losers	1.33 (***)	1.24 (***)	1.67 (***)	1.92 (***)
10-Q Losers	0.9 (**)	0.8 (**)	1.15 (***)	1.37 (***)
5-Q Losers	0.67 (**)	0.57 (**)	0.83 (***)	0.98 (***)
50-Q Winners	0.53	0.41	0.44	0.62 (*)
20-Q Winners	0.43	0.31	0.31	0.46 (*)
10-Q Winners	0.24	0.12	0.12	0.27
5-Q Winners	0.2	0.08	0.1	0.18
50-Q Long-short	1.81 (***)	1.82 (***)	2.34 (***)	2.51 (***)
20-Q Long-short	1.14 (**)	1.16 (**)	1.59 (***)	1.7 (***)

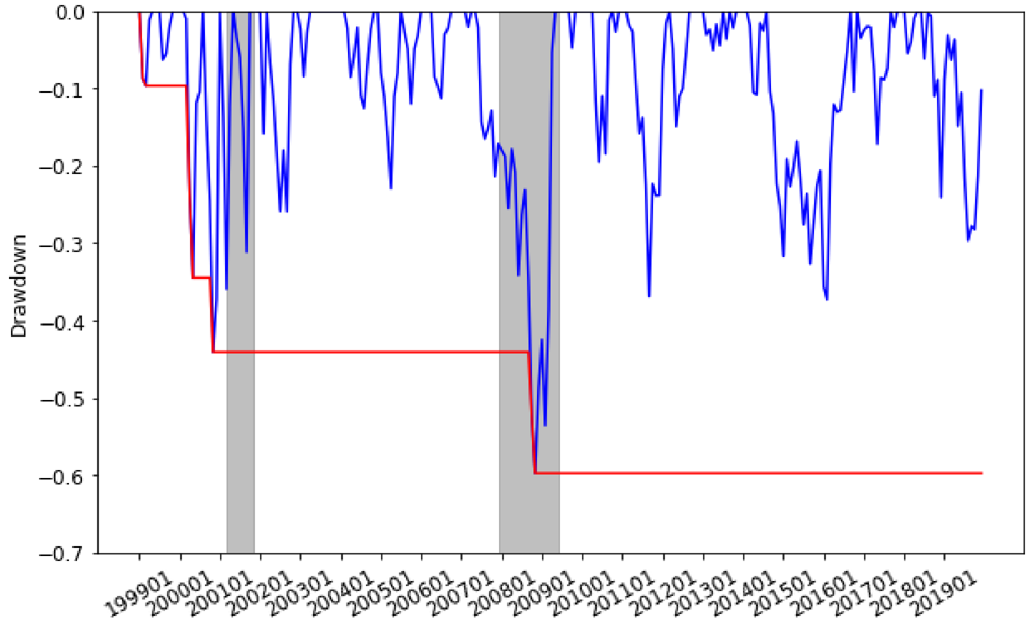
Panel B: Months 2006/01-2019/12

Portfolio	CAPM alpha %	FF-3 alpha %	CAR-4 alpha %	FF-5 alpha %
50-Q Losers	1.16 (**)	1.4 (***)	1.81 (***)	2.06 (***)
20-Q Losers	0.72	0.91 (**)	1.23 (***)	1.53 (***)
10-Q Losers	0.39	0.55 (*)	0.79 (***)	1.04 (***)
5-Q Losers	0.31	0.43 (*)	0.6 (**)	0.82 (***)
50-Q Winners	0.22	0.31	0.38	0.42
20-Q Winners	0.16	0.24	0.32	0.33
10-Q Winners	0.03	0.11	0.19	0.19
5-Q Winners	0.05	0.13	0.2	0.18
50-Q Long-short	1.18 (**)	1.34 (***)	1.67 (***)	1.88 (***)
20-Q Long-short	0.73 (*)	0.83 (**)	1.09 (***)	1.38 (***)

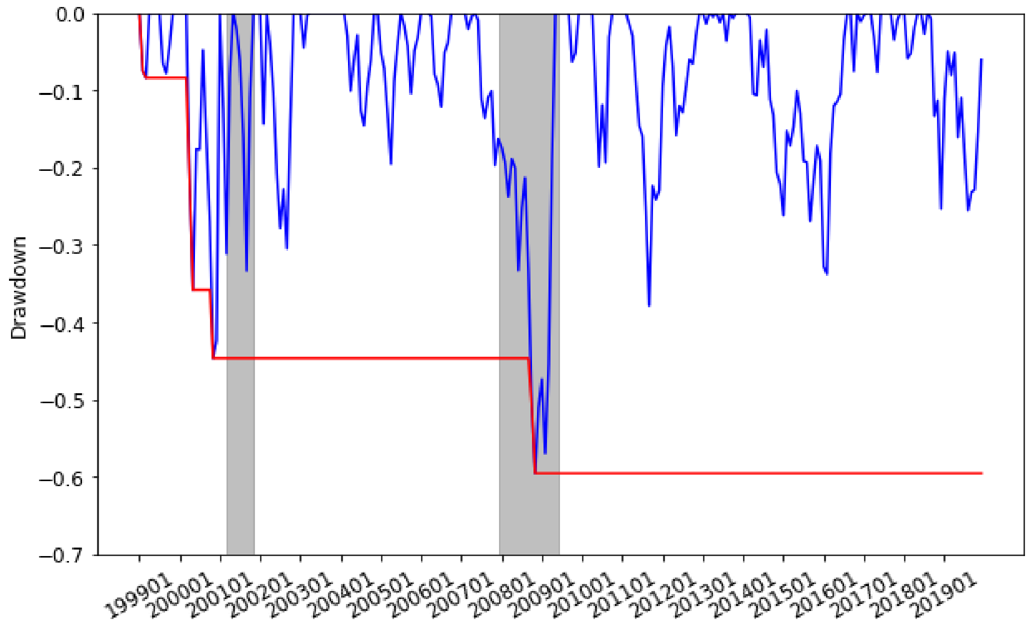
Panel C: Months 2013/01-2019/12

Portfolio	CAPM alpha %	FF-3 alpha %	CAR-4 alpha %	FF-5 alpha %
50-Q Losers	-0.81	0.07	0.13	0.09
20-Q Losers	-0.5	0.24	0.28	0.26
10-Q Losers	-0.54	0.12	0.17	0.14
5-Q Losers	-0.31	0.25	0.29	0.26
50-Q Winners	-0.18	0.5	0.53	0.46
20-Q Winners	-0.32	0.19	0.18	0.17
10-Q Winners	-0.45	0.01	-0.02	-0.01
5-Q Winners	-0.38	0.04	0.0	0.02
50-Q Long-short	-0.39	0.09	0.2	0.12
20-Q Long-short	0.02	0.36	0.44	0.39

3.3 Underwater drawdown plots

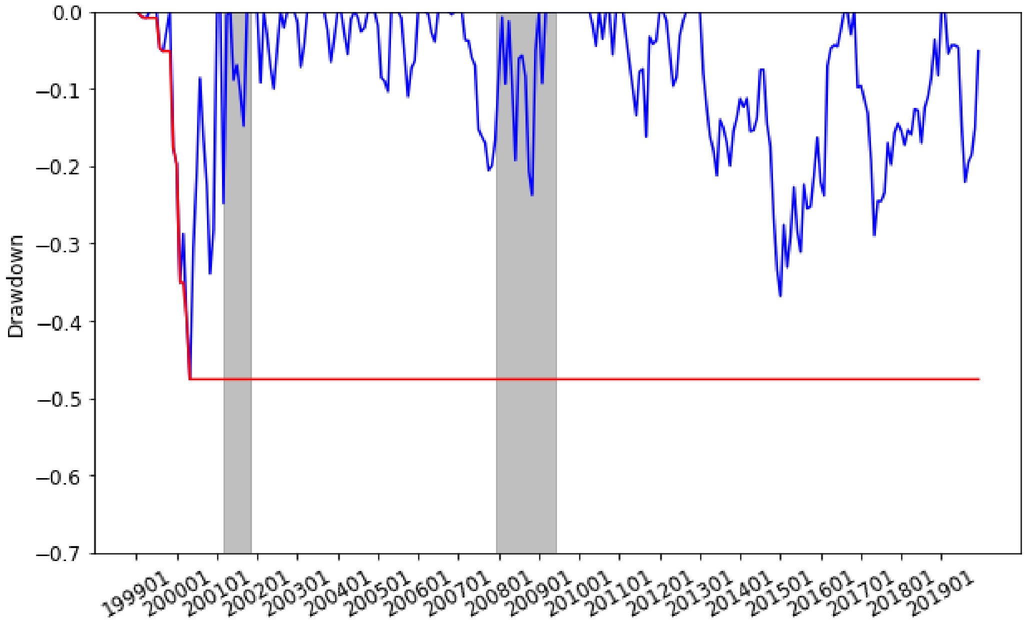


(a) **Drawdown plot of 50-Quantile Losers portfolio**. Red line indicates the rolling-max drawdown. Shaded periods are recessions/economic downturns (Mar 2001- Nov 2001 and Dec 2007- June 2009).

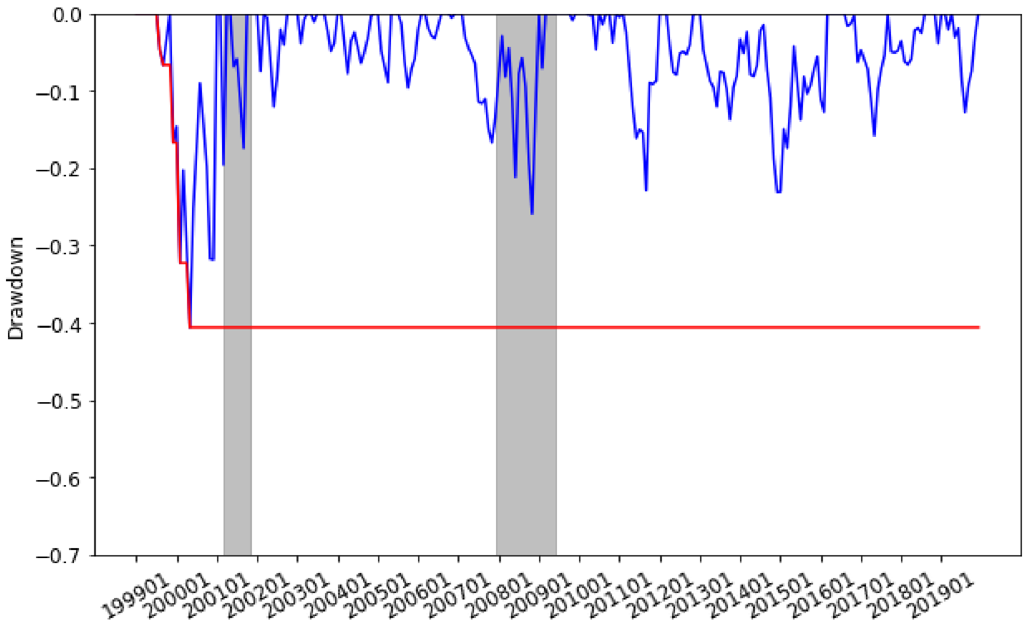


(b) **Drawdown plot of 20-Quantile Losers portfolio**. Red line indicates the rolling-max drawdown. Shaded periods are recessions/economic downturns (Mar 2001- Nov 2001 and Dec 2007- June 2009).

Figure 8: **Long Losers Q-50 and Q-20 drawdown plots.**



(a) **Drawdown plot of 50-Quantile Long-short portfolio**. Red line indicates the rolling-max drawdown. Shaded periods are recessions/economic downturns (Mar 2001- Nov 2001 and Dec 2007-June 2009).



(b) **Drawdown plot of 20-Quantile Long-short portfolio**. Red line indicates the rolling-max drawdown. Shaded periods are recessions/economic downturns (Mar 2001- Nov 2001 and Dec 2007-June 2009).

Figure 9: **Long-short Q-50 and Q-20 drawdown plots.**

4 Summary

- 50-Quantile Losers and 50-Quantile Long-short perform very well in the backtesting sample
  - 50-Quantile Losers grow by factor of **481.67** during 1999/01-2019/12 (28.04 when the universe is ME restricted to 70% largest)
  - 50-Quantile Long-short grow by factor of **138.25** (9.53)
- 50-Quantile Losers and 50-Quantile Long-short produce significant alpha w.r.t. all common asset pricing models in the full backtesting sample (1999/01-2019/12)
- In risk-return terms 50-Quantile Losers and 50-Quantile Long-short are best portfolios of the study

5 Appendix

5.1 Data

I use CRSP Monthly Stock data. The data is retrieved from Wharton Research Data Services (WRDS). Factors are from Kenneth French's webpage (https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html).

I adjust the CRSP data for:

- **Returns** (RET variable)
  - Observations s.t. RET equals C are excluded. These stand for "No valid previous price". There are 11 791 such observations in the dataset.
  - Observations s.t. RET equals B are excluded. These stand for "Off-exchange" (i.e. return calculated from off-exchange price).
    There are 26 446 such observations in the dataset.
  - Observations s.t. RET doesn't exist, RET is set to zero. These stand for "No valid price". There are 24 556 such observations in the dataset.
- **Delisting returns** ($DLRET$ variable)
  - Delisting returns are adjusted according to Shumway (1997) procedure.
  - If $DLRET$ exists for an observation, it is used as the holding period return instead of RET.
    - There are 1 110 such observations in the dataset.
  - If $DLRET$ is not available and the observations $DLSTCD$ is 500, 520, between 551 and 573 inclusive, 574, 580, or 584 then holding period return is set to -1.
- **Prices** (PRC variable)
  - Observation s.t. $|PRC| < 1$ are excluded from the dataset.
- **Sharecodes** ($SHRCD$ variable)
  - Only sharecodes 10 and 11 are of interest (ordinary common shares of US stocks). Hence other sharecodes are excluded from the dataset.

In addition, new variable is defined:

- **Market value of equity** (ME variable)

$$ME = PRC \cdot SHROUT/1000$$

where $SHROUT$ is the share count.

Table 9: **Summary table of the market portfolios**. *Universe ew* is the equally weighted market portfolio of the investment universe. *vwretd* and *ewretd* are CRSP market portfolios, value-weighted and equally weighted, respectively. Portfolios are monitored from 1999/01 to 2019/11 (251 months).

	count	mean	std	min	25%	50%	75%	max
vwretd	251.0	0.006	0.043	-0.185	-0.018	0.012	0.035	0.114
ewretd	251.0	0.009	0.053	-0.205	-0.019	0.009	0.038	0.225
Universe ew	251.0	0.011	0.057	-0.206	-0.020	0.011	0.046	0.214

The data is **filtered by *ME*** to ensure applicability for a significant asset manager. For each month the smallest 10% by lagged *ME* (ME_{t-1}) is filtered out of the data.

As a result of the adjustments and filtering we have restricted investment universe compared to whole CRSP dataset. In the sample period 1999/01-2019/11 this leads somewhat higher return for the "market portfolio" of the universe. Also the restricted universe is more volatile.

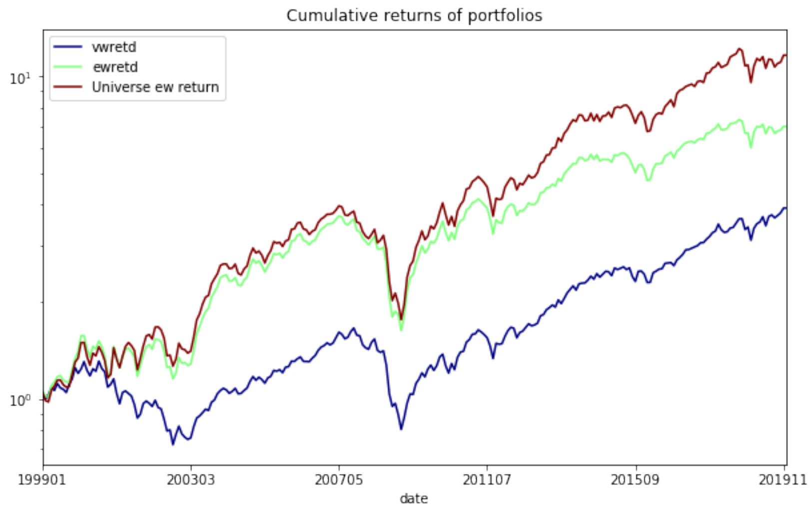


Figure 10: **Cumulative returns of market portfolios (log scale)**. See Table 9 for definitions.

5.2 Code

All of the tables and figures of the paper can be reproduced using jupyter notebooks (running Python). Below all links to sections' notebooks.

- **Strategy** -section figures and tables:
  https://github.com/j4bert0/quant-finance-project/blob/master/research/strategy_intro.ipynb
- **Long-short portfolios** -section figures and tables:
  https://github.com/j4bert0/quant-finance-project/blob/master/research/long_short_portfolios.ipynb
- **Performance (backtesting)** -section figures and tables:
  https://github.com/j4bert0/quant-finance-project/blob/master/research/performance.ipynb
- **Performance (robustness)** -section figures and tables:
  https://github.com/j4bert0/quant-finance-project/blob/master/research/robustness.ipynb
- **Risk analysis** -section figures and tables:
  https://github.com/j4bert0/quant-finance-project/blob/master/research/risk_analysis.ipynb

Most importantly, notebooks on how the data (CRSP and factors) are prepared.

- **Data**:
  https://github.com/j4bert0/quant-finance-project/blob/master/research/full_data.ipynb
- **Factors**:
  https://github.com/j4bert0/quant-finance-project/blob/master/research/factors.ipynb

5.3 Definitions

Let r_i be monthly return observation of portfolio i. Sample of size N of monthly return observations: $S_r^N := (r_1, r_2, r_3, ..., r_N)$. $\bar{r}$ is the arithmetic sample mean. $r_f = (r_{f,1}, r_{f,2}, r_{f,3}, ..., r_{f,N})$ is sample of risk-free interest rate observations. $\bar{r}_f$ is the arithmetic sample mean.

- Monthly volatility (standard deviation)

$$\sigma(S_r^N) = \left(\frac{1}{N-1} \sum_{i=1}^{N} \left(r_i - \bar{r} \right)^2 \right)^{1/2}$$

- Monthly semi-deviation (downside volatility)

$$\sigma_-(S_r^N) = \left(\frac{1}{N-1} \sum_{i=1}^{N} \left(\min \left\{ r_i - \bar{r}, 0 \right\} \right)^2 \right)^{1/2}$$

- Monthly geometric mean

$$G(S_r^N) = \left(\prod_{i=1}^{N} \left(1 + r_i \right) \right)^{1/N} - 1$$

- Max drawdown (portfolio's maximum drop from the previous peek value)
  Porfolios cumulative growth up to k-th month: $C_k = C_k(S_r^N) = \prod_{i=1}^{k}(1 + r_i)$. Maximum cumulative growth up k-th month: $\tilde{C}_k = \max \left(C_j \right)_{1 \leq j \leq k}$

$$MDD = \min \left(C_k / \tilde{C}_k - 1 \right)_{1 \leq k \leq N}$$

- Sharpe ratio

$$SR = \frac{\bar{r} - \bar{r}_f}{\sigma(S_r^N)}$$

- MAR ratio (note: the common MAR definition is scaled by 100)

$$MAR = \frac{G(S_r^N)}{1 - MDD} \cdot 100$$

- Sortino ratio

$$SOR = \frac{\bar{r}}{\sigma_-(S_r^N)}$$

- Skewness: see https://en.wikipedia.org/wiki/Skewness, section "Sample skewness"
- Kurtosis: see https://en.wikipedia.org/wiki/Kurtosis, section "Sample kurtosis"

5.4 Experimental paths charted

In this section I go through some additions to the strategy that I tested but which didn't make the final cut.

- **Position weights**
  - **Weigths by inverse volatility**: scaling the positions by the inverse of past m months' volatility
  - i.e. stock $k \in P_1$ with past m months' volatility $\sigma_{k,m}$ has weight

    $$w_k = \sigma_{k,m}^{-1} \left(\sum_i \sigma_{i,m}^{-1} \right)^{-1}$$

    Where $i \in P_1$
  - Result: scaling weights by inverse volatility reduces portfolio volatility substantially but reduces returns even more
  - Problems: estimating standard deviation from the data
  - **Weigths by volatility**:
  - i.e. stock $k \in P_1$ with past m months' volatility $\sigma_{k,m}$ has weight

    $$w_k = \sigma_{k,m} \left(\sum_i \sigma_{i,m} \right)^{-1}$$

    Where $i \in P_1$
- **Weights by Markowitz optimization**
  - Computationally I made it work but the results were not good

References

[1] De Bondt, W.F.M & Thaler, R. Does the Stock Market Overreact? *Journal of Finance*. 1985.

[2] Jegadeesh, N. & Titman S. Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *Journal of Finance*. 1993.

[3] Shumway, T. The delisting bias in CRSP data. *Journal of Finance*. 1997.