Insights 2.11. Mutual Funds - Basic Historic Analysis: Are you always wrong?
Mutual Funds - Basic Historic Analysis: Are you always wrong?

With reference to the ***US AGG Mutual Funds***, broadly speaking no. For the overall US AGG mutual fund market, assuming the *Bloomberg US Agg Bond TR USD* as the index, there is value in looking at historical price data (and the derived performance measures). As illustrated in Table 1, the overall market (without sampling) generated a 0.52 – 0.58% average annualized excess return over the Index (for each of 1, 2-, 3-, 5- or 10-year periods over the 12/31/2000 – 12/31/2022 evaluation period). With sampling, based on 40+ performance measures as illustrated in Table 2, there can be ***nearly 40% of the funds that are top quartile*** (depending on the performance measure used for selection and evaluation) that generated 0.9-1.7% average annual excess return (ranged 7.9% – (5.1)%) over the 1-2-3 year holding periods (here Excess Return as a selection criteria). However, high performance measure-based selection does not imply stable superior excess returns across all performance measures or periods. Nor does it imply that it's the same funds that are top quartile.

But,

1. Do you want the Index? The lower R^2, high Alpha and low Beta of the Funds implies that the funds in general are not really in line with the Index (Bloomberg US Agg Bond TR USD). This also becomes evident from the peaks and troughs of the Excess Return and Relative Max Drawdowns, where the Index generated a better Sharpe over all of the assessment periods (see Table 1).

2 . Do the incremental 'allocator/distributor' fees erode the generated Excess Return? If the fees you pay are below 0.52% then as a market or via superior fund selection you may be better off (assuming Excess Return is your only objective function).

3. Which performance measures do you value more/less? Some performance measures have promising distributions as selection criteria and can support repeatability. The Explainability Index and Risk of Target (Image 1) gives a control panel for incorporating multiple performance measures.

4. Can you dig deeper for fund level granularity? This is insightful for assessing funds that may be 'better' over time, during certain times/regimes, only in certain markets, against certain benchmarks, at time of entry/exit (given the demonstrated over/under performance), etc. As in Event and Feature Engineering (yes, this Insights piece is basic simpler historical analysis, but we point this out as these questions may be coming up and will be covered in the journey's next pieces (where more data is injected and ML/DL is applied to beyond canned what ifs).

When you are being pitched over 7,000 mutual funds1[1] (in the US alone), how do you know the selection motivations are aligned? Beyond the regulatorily mandated disclosures, the distributors/allocators generally point to the historical performance of the funds and/or forecasted performance under scenarios. This relies on two facets: (a) the benchmark(s) being considered, and (b) the performance measure being evaluated. In this Insights piece, we look at a popular benchmark for the particular Asset Class and the fund performance against that benchmark as measured by 40+ performance measures estimated on a historical basis. Since evaluating so many performance measures can be unwieldy, we also assess the performance via the performance measures unifying framework of Explainability Index (EI) and Risk of Target (RoT)[2].

Data
We filter the US mutual fund data that are categorized as US Aggregate (US AGG), were at least 3 years old (considering 12/31/2000 - 12/31/2022 evaluation period), had over $ 1 billion in AUM and we evaluated the oldest share class. This filtering resulted in 168 funds in 2022 (with a range of 45-269 funds filtered for the analysis over the evaluation period).

Analysis
As a reminder, this Insights piece is the first part of the journey, where both the assessment and evaluation are based on historical price data (and derived performance measures) for both fund and benchmarks. Refer to the Explainability Index paper in footnote 2 for the methodology used for estimating the performance measures.

---

[1] 27,000+ if you assume all share classes. Also, not including SMAs, ETFs, etc.

[2] Hirsa, Ali and Ding, Rui and Malhotra, Satyan, Explainability Index (EI): Unifying Framework of Performance Measures and Risk of Target (RoT): Variability from Target EI (January 23, 2023). Available at SSRN: https://ssrn.com/abstract=4335455
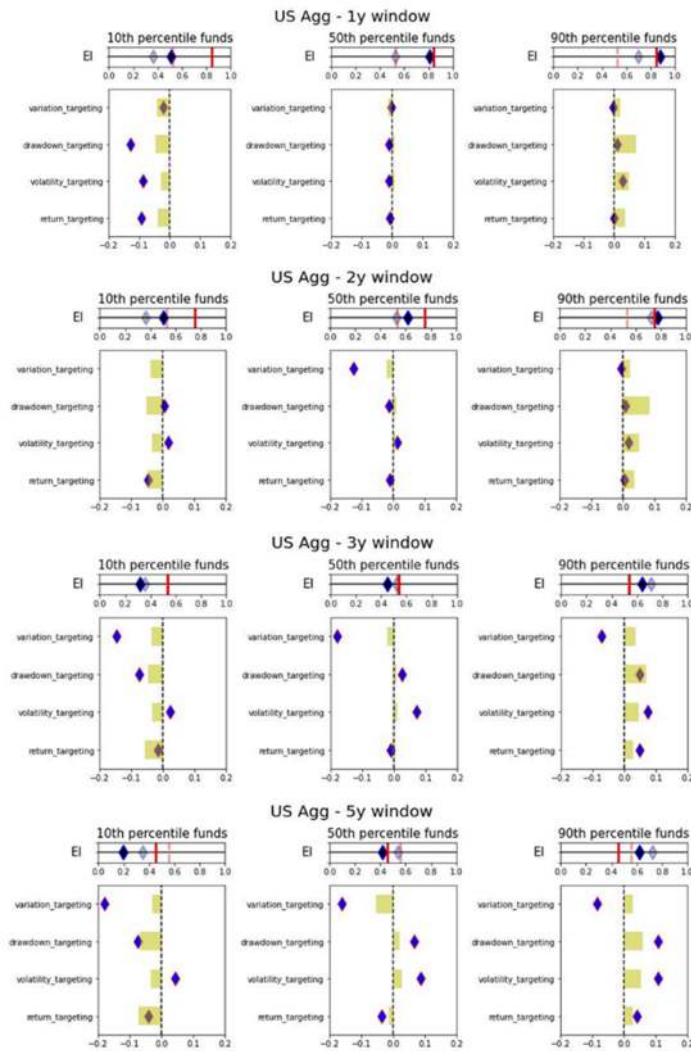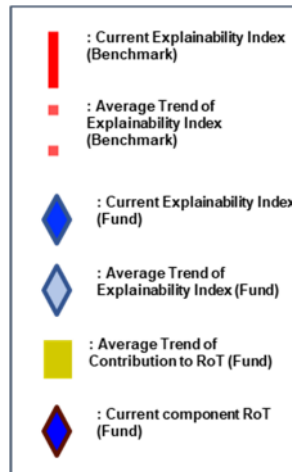
US AGG Market

Although it is difficult (or irrational) to invest in all the funds, it is important to look at the entire market as you never know the performance of the specific fund you have invested in will be (so at a minimum it sets the overall expectations). Therein the point here is to give a datapoint without selection bias for the entire market (as filtered for the Asset Class above), where the alternative is to invest in the Index (directly or via a proxy).
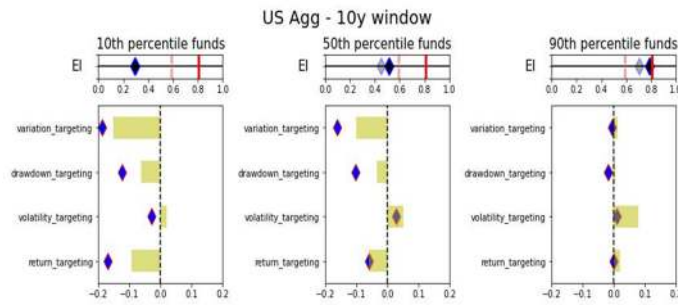
Table 1. Historical Benchmark and Fund Performance.

Funds	All									
Benchmark	BBG US Agg Bond TR USD									
Period	12/31/2000 - 12/31/2022									
		Historical 1 year		Historical 2 years		Historical 3 years		Historical 5 years		Historical 10 years	
Categories	Performance Measures	Benchmark	Funds	Benchmark	Funds	Benchmark	Funds	Benchmark	Funds	Benchmark	Funds
Explainability Index		0.54	0.54	0.54	0.54	0.53	0.52	0.55	0.53	0.61	0.48
Risk of Target			2.10%		2.20%		0.80%		0.90%		-19%
Return Targeting	Alpha		0.65%		0.66%		0.62%		0.61%		0.61%
	Excess Return		0.58%		0.55%		0.52%		0.56%		0.54%
	Return	4.54%	4.85%	4.68%	5.05%	4.84%	5.24%	4.97%	5.45%	5.52%	5.99%
	Average Gain	0.92%	1.12%	0.92%	1.12%	0.92%	1.12%	0.93%	1.12%	0.97%	1.16%
	Batting Average		54.31%		54.55%		54.42%		54.39%		54.08%
	Beta		0.96		0.97		0.97		0.99		0.98
	Bull Beta		0.79		0.81		0.82		0.85		0.85
	M.S. Returns	2.83%	3.15%	2.88%	3.27%	2.93%	3.35%	2.87%	3.37%	2.98%	3.47%
	M.S. Risk Adjusted Returns	2.72%	2.89%	2.77%	3.00%	2.81%	3.08%	2.76%	3.10%	2.86%	3.20%
	Treynor Ratio		36.63%		4.13%		3.84%		4.59%		3.88%
	Up Capture Ratio		1.09		1.10		1.08		1.08		1.07
	Up Period Percent	64.77%	66.07%	65.36%	66.87%	66.01%	67.55%	66.59%	68.08%	67.94%	68.98%
Volatility Targeting	Gain Deviation	3.00%	3.70%	3.07%	3.84%	3.13%	3.91%	3.18%	3.98%	3.37%	4.16%
	Gain Std Dev	2.04%	2.60%	2.10%	2.73%	2.15%	2.80%	2.19%	2.88%	2.28%	3.00%
	Information Ratio		0.31		0.19		0.14		0.12		0.10
	Kappa	3.39	5.90	2.15	2.11	1.84	1.75	1.63	1.45	1.59	1.33
	Loss Deviation	1.65%	2.47%	1.73%	2.65%	1.75%	2.71%	1.78%	2.72%	1.81%	2.73%
	Loss Std Dev	1.89%	2.84%	2.04%	3.22%	2.13%	3.41%	2.19%	3.50%	2.26%	3.57%
	M.S. Risk	0.11%	0.26%	0.11%	0.28%	0.12%	0.28%	0.12%	0.27%	0.13%	0.27%
	Omega	3.25	5.99	2.36	2.54	2.18	2.22	2.01	1.99	1.94	1.90
	Sharpe Ratio	0.94	0.91	0.88	0.81	0.89	0.81	0.87	0.80	0.88	0.80
	Sortino Ratio	3.14	5.70	2.00	2.01	1.74	1.68	1.55	1.42	1.50	1.32
	Tracking Error		2.59%		2.77%		2.86%		2.89%		2.95%
	Volatility	3.28%	4.35%	3.33%	4.52%	3.35%	4.58%	3.38%	4.61%	3.50%	4.70%
Drawdown Targeting	Average Loss	-0.65%	-1.02%	-0.65%	-1.04%	-0.64%	-1.02%	-0.65%	-0.99%	-0.67%	-0.98%
	Bear Beta		0.97		1.06		1.12		1.18		1.15
	Calmar Ratio	5.68	9.31	2.54	2.48	1.85	1.74	1.48	1.27	1.38	0.98
	Down Capture Ratio		1.05		1.04		1.03		1.02		0.97
	Overall Capture Ratio		1.51		1.40		0.98		-3.96		1.06
	Down Period Percent	35.23%	33.51%	34.64%	32.72%	33.99%	32.04%	33.41%	31.54%	32.06%	30.72%
	Max Drawdown	-2.00%	-3.13%	-2.73%	-4.47%	-3.16%	-5.29%	-3.64%	-6.49%	-4.14%	-8.74%
	Var 95%	-5.76%	-5.37%	-5.82%	-5.37%	-5.79%	-5.36%	-5.81%	-5.27%	-5.92%	-5.26%
	CVaR 95%	-8.01%	-8.12%	-8.08%	-8.31%	-8.10%	-8.38%	-8.14%	-8.40%	-8.34%	-8.57%
Variation Targeting	Correlation		79.46%		78.62%		77.84%		77.87%		77.55%
	Kurtosis	-0.03	0.38	0.34	1.07	0.62	1.62	0.94	2.41	1.07	3.23
	Skewness	-0.32	-0.33	-0.70	-0.47	-0.56	-0.54	-0.56	-0.58	-0.61	-0.60
	Jarque Bera	2.31	3.10	3.36	10.32	4.03	23.46	6.57	67.82	10.27	180.09
	R2		70.35%		67.82%		66.16%		65.48%		64.32%
	Residual Std Dev		2.32%		2.55%		2.66%		2.70%		2.81%
	Std Error Alpha		0.22%		0.16%		0.14%		0.10%		0.08%
	Std Error Beta		21.50%		15.83%		13.25%		10.32%		7.39%
	Appraisal Ratio		12.65%		8.55%		6.72%		5.48%		4.45%

A simpler and explainable way to digest all the performance measures in Table 1 is to look at the Explainability Index Frameworks[3] presented below. This highlights the performance measure facets of the Funds that are better or worse than the designated Index.

---

[3] The current is based on annualized over a (x)yr period and the average is based on rolling the (x)yr window every year for 20yrs; we assess the fund/portfolio as of the last evaluation date rolled back. Analysis assumptions: Measures: Absolute. Time Variation: No. Threshold/Scale: Market Index. Categories: Yes. Weights: Equal. Type: Arithmetic. 50th percentile defined as 45-55%.

: Current Explainability Index (Benchmark)

: Average Trend of Explainability Index (Benchmark)

: Current Explainability Index (Fund)

: Average Trend of Explainability Index (Fund)

: Average Trend of Contribution to RoT (Fund)

: Current component RoT (Fund)

US Agg - 1y window

10th percentile funds   50th percentile funds   90th percentile funds

US Agg - 2y window

10th percentile funds   50th percentile funds   90th percentile funds

US Agg - 3y window

10th percentile funds   50th percentile funds   90th percentile funds

US Agg - 5y window

10th percentile funds   50th percentile funds   90th percentile funds

US Agg - 10y window

The Explainabilty Index bridges the final engineering jargon to illustrate and/or manage the performance measures as a control panel per what is important for the selector/allocator.

As illustrated in Figures 1, 2 and 3, the number of Funds that have a higher return than the benchmark is extremely volatile. Averages over the 20-year period for the Asset Class show that 56-62% of the funds beat the Index, which is incredibly high. This itself is a unique maker for the market (as we will see when we look at other Asset Classes), but again this is average and the range is very wide.

Figure 1.

Funds that beat the Index 1-year [higher return]: 56.9% market average

Figure 2.

Funds that beat the Index 2-year [higher return]: 61.4% market average

Figure 3.

Funds that beat the Index 3-year [higher return]: 62.3% market average

As an illustrative example, Figures 4 & 5 look more granularity at the two of the more widely assessed performance measures – Excess Return and Relative Max Drawdown (RMDD). In Figure 4. Excess Return exhibits periods of excessive over or under performance depending on the historic window. Overall, shorter term historic periods show a lot of cyclicality and longer-term periods are more consistent as in general the funds do better than the index from a return perspective.

Figure 4.

Historical Market Excess Return

In Figure 5 the relative MDD shows extended periods of underperformance. This is not surprising as MDD is sticky given the nature of the look back periods. Funds in general have a higher MDD than the index in Table 1.

Figure 5.

Historical Market Relative MDD
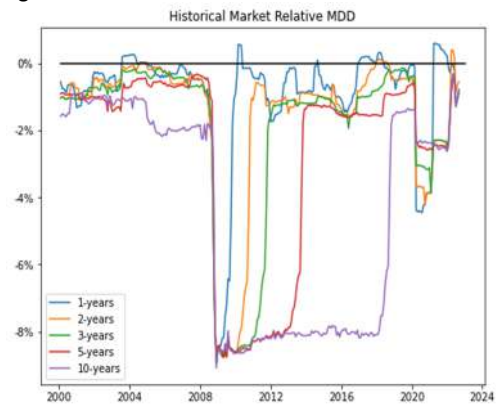
In Figure 6, we see that the funds **have a relatively low R^2 implying that the managers may not really be in line with the Index (especially over longer periods)**. That may also explain the higher Excess Return in Table 1 and the high outperformance averages in Figures 1, 2 and 3.

Figure 6.

Historical Market R2

US AGG Market > Fund Selection (Rolling)

Within the broader market, we try to identify funds that may 'in the future' outperform the market given a particular objective function. For example, as illustrated in Table 1, if the market can generate Excess Return over the Index (albeit with lower Sharpe) the question becomes how good a predictor is that performance measure itself (or others) as a marker for identifying individual funds that have a higher probability of outperformance for that objective function. Further, since we cannot time the entry/exit we conduct the analysis on a rolling basis, where we use every month as a starting point for selection and ending point for the holding period. Final numbers are based on averages across the funds/months. It should be noted that the Tables in this Insights piece have a lot of embedded granularities (some of which we have tried to highlight in the Figures), but all are available upon request. With reference to the overall Insights journey, the selectors/allocators that are somewhat uncomfortable with the more advanced financial engineering methods and/or jargon (or have limited alternative data access) will reside on a spectrum here by using some or a combination of the performance measures covered in Table 1 as their selection and evaluation criteria.

Herein, many studies have been conducted on methods of selecting the 'better' performing funds via leveraging various lenses. In this Insights piece, we remain focused on only using historical fund and benchmark performance data for trying to identify the 'better' performing funds. Further, to keep this practical (as in easily implementable), we assess performance over set holding periods (for 1, 2- or 3-year periods), without rebalancing (or frequent trading) and measured across each of the performance measures as objective/evaluation criteria (versus some x factor (or such) model to assess alpha or other higher order value add).

As a framework, since the allocation can be made at any time, the analysis in Table 2 is based on rolling performance assessment. Where, every month, we take the top decile funds based on the historic performance measure (for each of 1, 2-, 3-, 5- or 10-year periods) and then evaluate the percentage of funds remaining as the top decile selection at the end of the Investment Period (or 1, 2- or 3-years forward). Procedurally for
- Selection, we take the top quartile funds for the historical performance of each measure (for each of 1, 2-, 3-, 5- or 10-year historical periods) and hold them for each of the investment periods (for 1, 2- or 3-year holding periods). This is done on a monthly basis over the entire evaluation period so, depending on the performance measure the fund selection can change.
- Objective/evaluation - for each month, at the end of each investment period (for 1, 2- or 3-year holding periods), we calculate how many of the initially selected funds remain a top fund based on the same performance measure. We also evaluate if the selected funds remain a top fund based on all of the other performance measures. Table 2 illustrates results for Alpha (Higher), Excess Return (Higher) and Return (Higher) as the objective/evaluation criteria (where the results of all other categories and performance measures are available upon request).

Final percentages for the objective/evaluation are based on the averages. Note, over the evaluation period (from 2000-01-31 to 2022-12-31) if the fund had a track record lower than the historical evaluation period, the measure was evaluated from its inception.

Table 2. Selection Performance

Funds	All
Benchmark	BBG US Agg Bond TR USD
Period	12/31/2000 - 12/31/2022

			Objective/Evaluation: Higher Alpha			Objective/Evaluation: Higher Excess Return			Objective/Evaluation: Higher Return		
Category	Measure	Historical Period (years)	Holding 1Y	Holding 2Y	Holding 3Y	Holding 1Y	Holding 2Y	Holding 3Y	Holding 1Y	Holding 2Y	Holding 3Y
Return Targeting	Alpha	1									
		2									
		3									
		5									
		10									

[Remainder of table contains dense numerical data across categories: Return Targeting (Alpha, Excess Return Arith, Return, Average Gain, Batting Average, Beta, Bull Beta, M.S. Returns, M.S. Risk Adjusted Returns, Treynor Ratio, Up Capture Ratio, Up Period Percent), Volatility Targeting (Gain Deviation, Information Ratio, Kappa, Loss Deviation, Loss Std Dev, M.S. Risk, Omega, Sharpe Ratio, Sortino Ratio, Tracking Error, Volatility), Drawdown Targeting (Average Loss, Bear Beta, Calmar Ratio, Down Capture Ratio, Overall Capture Ratio, Down Period Percent, Max Drawdown, VaR 5%, CVaR 5%), Variation Targeting (Correlation, Kurtosis, Skewness, Jarque-Bera, R2, Residual Std Dev, Std Error Alpha, Std Error Beta, Appraisal Ratio), each for Historical Periods 1, 2, 3, 5, 10 years — values not legible at available resolution.]

Overall depending on the performance measure selected there can be an up to 42% chance of being in the top decile (under certain selection criteria, evaluation criteria and investment period). This is a remarkably high and unexpected number, but we again point to the framework setup and its specific application to the US AGG market. Further, it should be noted that it does not imply that it is the same funds that remain top quartile. It should be noted that using a combination of fixed performance measures and weights should generally give results within the performance measure ranges.

For a more granular analysis, let's assess the historic Excess Return as a predictor. As a reminder, the results of all other performance measures are available upon request. From Table 2, assuming we would have invested in the top decile funds as classified by the historic 3-year analysis, where Figure 7, 8 and 9 give the performance of those funds after 1, 2- and 3-year holding periods. Note, average stands for investing equal weights in all of the identified top decile funds, worst stands for picking the worst fund in the top decile every time, and best stands for picking the best fund in the top decile every time.
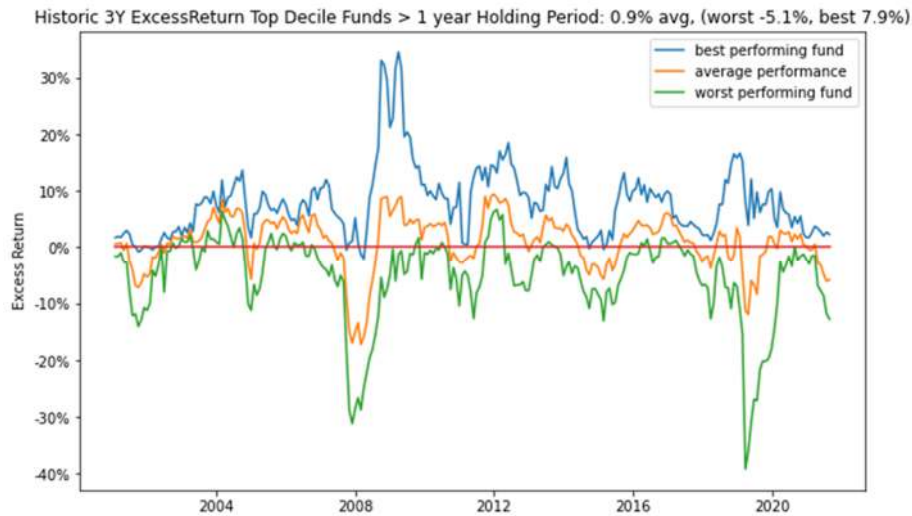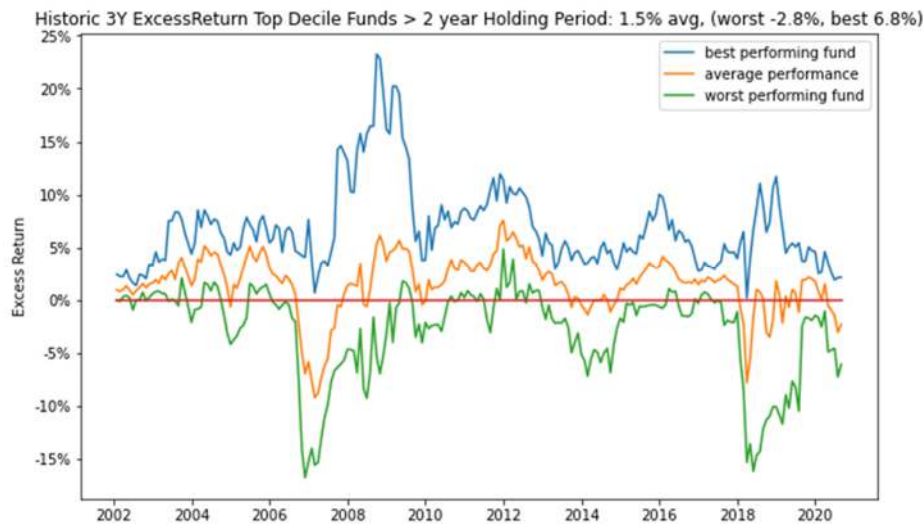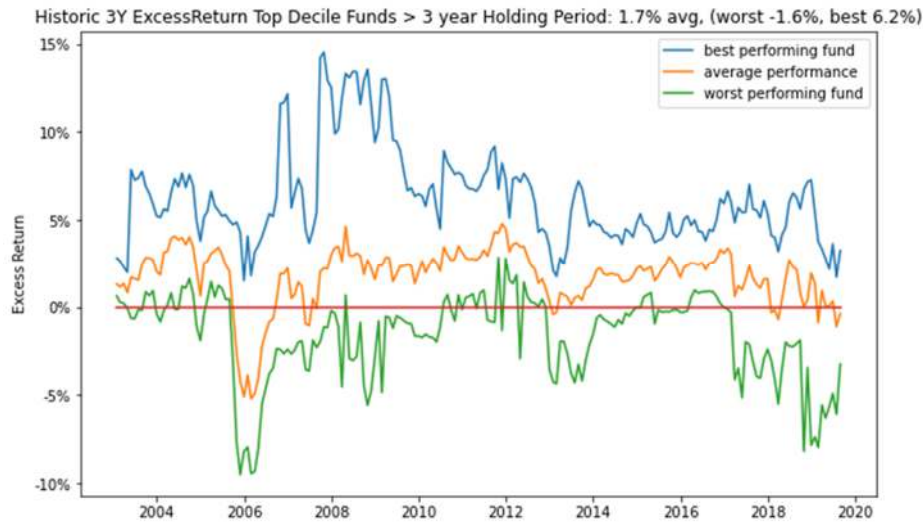
Figure 7.

Figure 8.

Figure 9.

Historic 3Y ExcessReturn Top Decile Funds > 3 year Holding Period: 1.7% avg, (worst -1.6%, best 6.2%)

On average, we find that using just historic (3-year) Excess Return as a selection and evaluation criterion for the increases the selection Excess Return value from 0.52% to 1.7% and similarly for other periods (assuming impact on other performance measures are not considered).

If there is implied value in generally using the performance measures, then the question becomes if any of the performance measures show unique markers that make them better qualified for the selection. As illustrated in Figures 10, 11 and 12, we compare the distributions of funds that generated Excess Return over the index (at the end of the investment periods) with the funds that did not beat the index. Figures 10, 11 and 12 show that Excess Return distributions show it not to be a clear discerning marker.
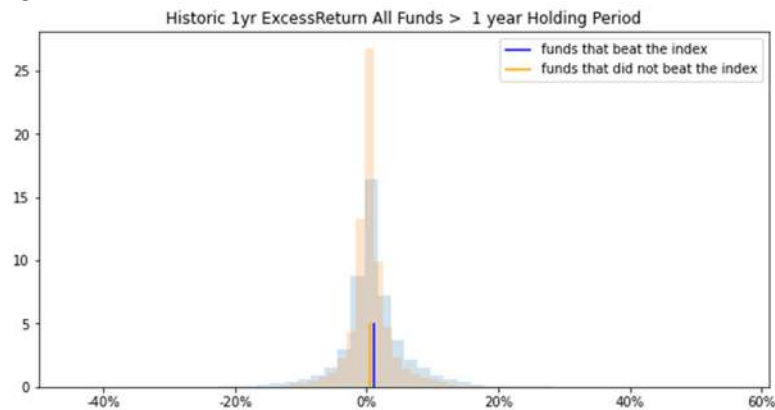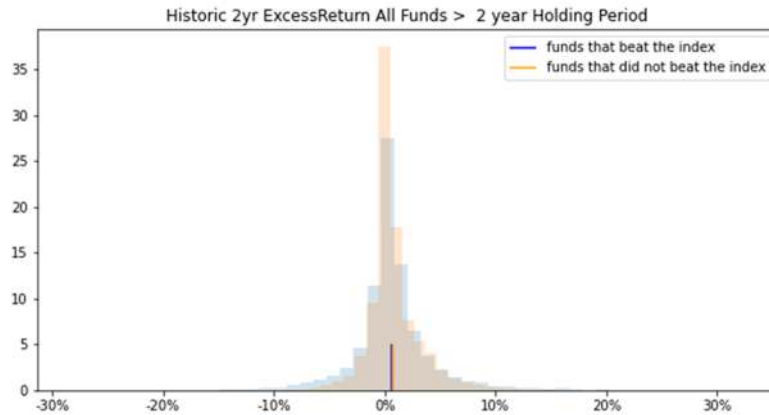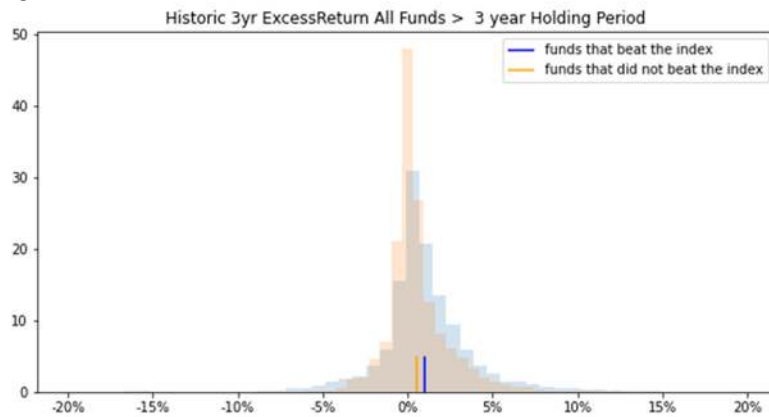
Figure 10.

Historic 1yr ExcessReturn All Funds > 1 year Holding Period

Figure 11.

Historic 2yr ExcessReturn All Funds > 2 year Holding Period

Figure 12.

Historic 3yr ExcessReturn All Funds > 3 year Holding Period

In assessing the distribution profiles of all performance measures listed in Table 1 we isolate the ones that exhibit more pronounced differences (as illustrated in Figures 13, 14, 15 and 16). For example, as illustrated in Figure 13 and 14, funds with a lower R^2 seemingly were more likely to beat the Index. Where as in Figures 15 and 16 show Average Gain as a unique marker also.
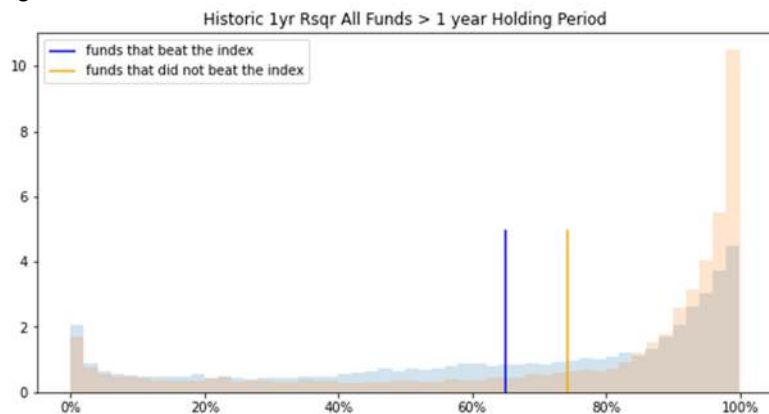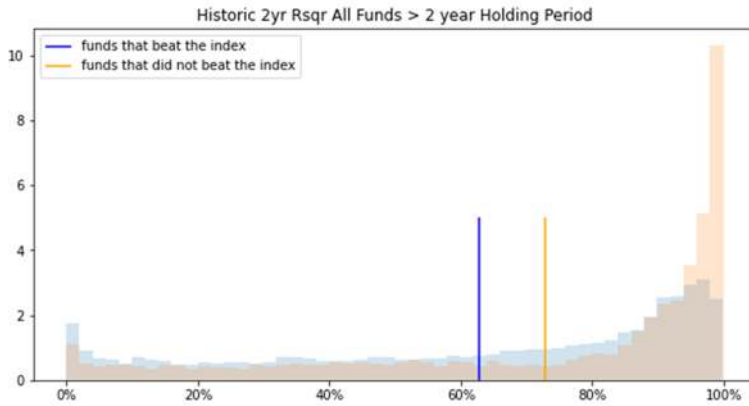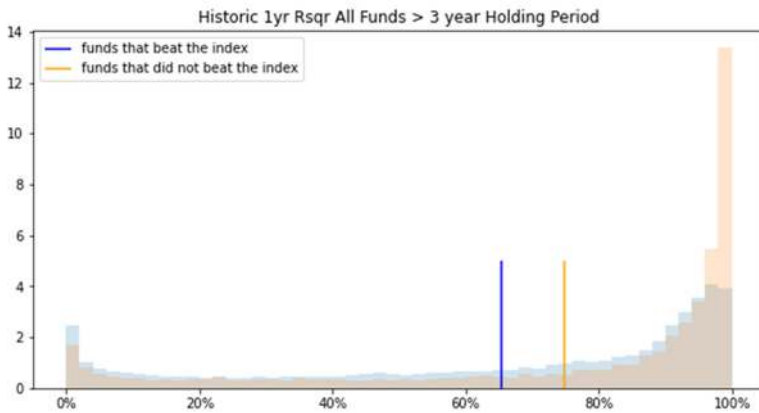
Figure 13.

Historic 1yr Rsqr All Funds > 1 year Holding Period

Figure 14.

Historic 2yr Rsqr All Funds > 2 year Holding Period

Figure 15.

Historic 1yr Rsqr All Funds > 3 year Holding Period

Figure 16.

Historic 1yr Average Gain All Funds > 1 year Holding Period

Figure 17.

Historic 3yr Average Gain All Funds > 2 year Holding Period

Figure 18.

Historic 3yr Average Gain All Funds > 3 year Holding Period
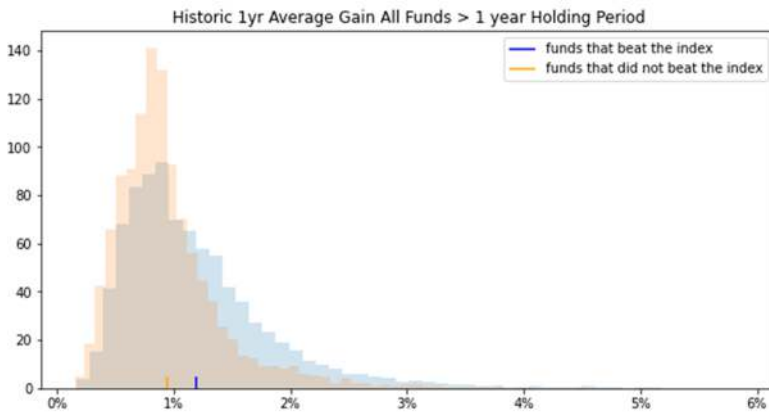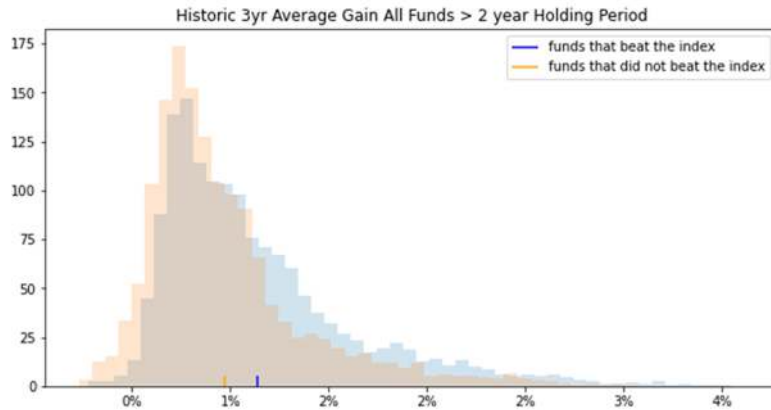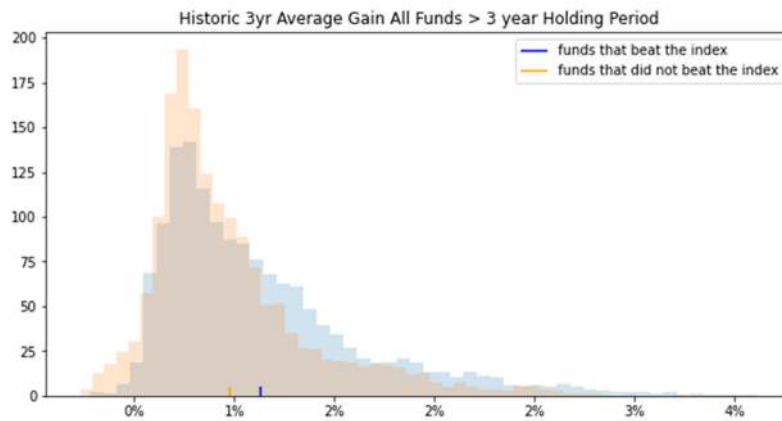
As an illustration, since we find R^2 to have unique marker, we assess the Excess Returns based on using the 3-year historic R2 performance as a selection criteria (Figures 19, 20 and 21). It reflects a tighter spread than using Excess Return and average seems to improve over the short time period.
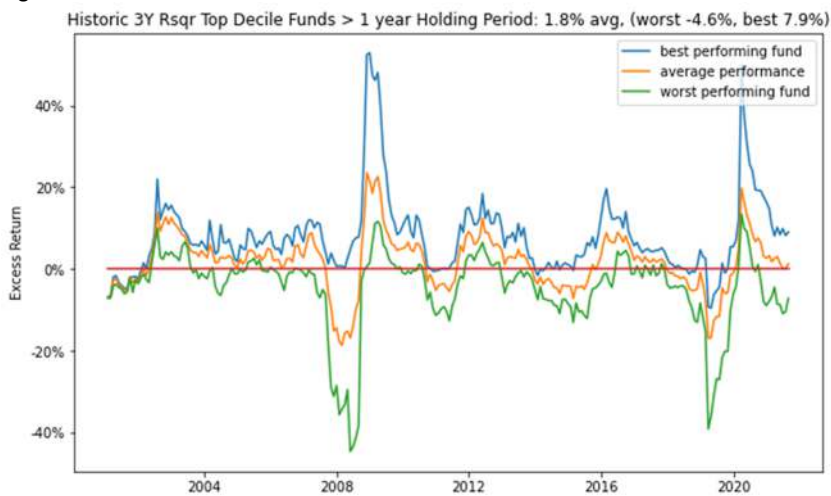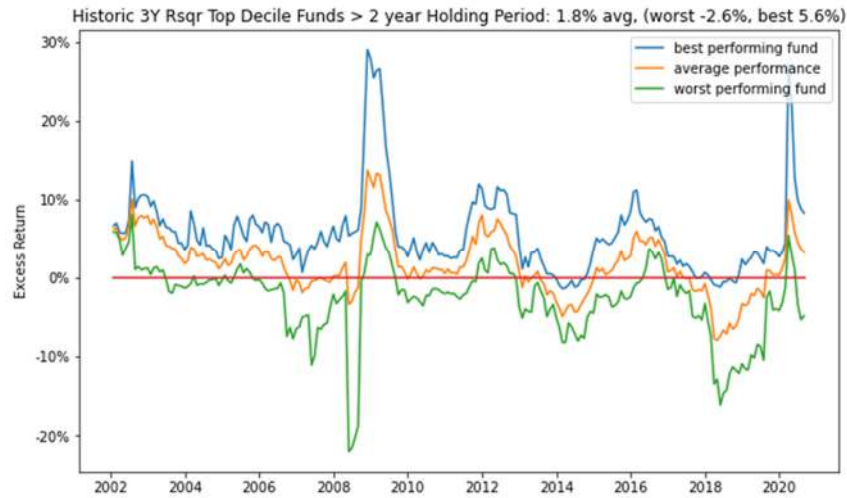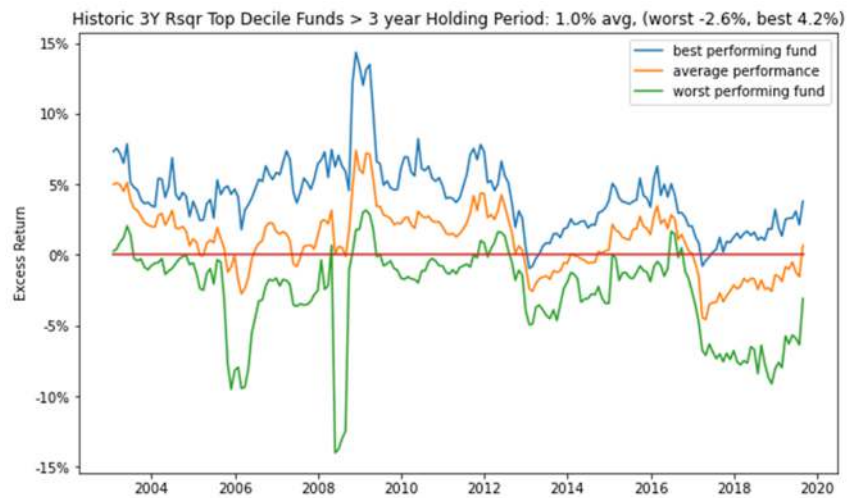
Figure 19.

Historic 3Y Rsqr Top Decile Funds > 1 year Holding Period: 1.8% avg. (worst -4.6%, best 7.9%)

Figure 20.

Historic 3Y Rsqr Top Decile Funds > 2 year Holding Period: 1.8% avg. (worst -2.6%, best 5.6%)

Figure 21.

Historic 3Y Rsqr Top Decile Funds > 3 year Holding Period: 1.0% avg. (worst -2.6%, best 4.2%)

US AGG Market > Fund Selection (point in time)

Looking at monthly rolling performance statistics has lots of embedded nuances, statistics and in general can be overwhelming. In is section, we do a point in time analysis, where we assume that the decisions were made on 2018-12-31 to select the top decile funds based on the 3-year historic Excess Return and we assess the Total Return of the funds in 2019, 2020 and 2021 (as illustrated in Figures 22, 23, 24 and 25).
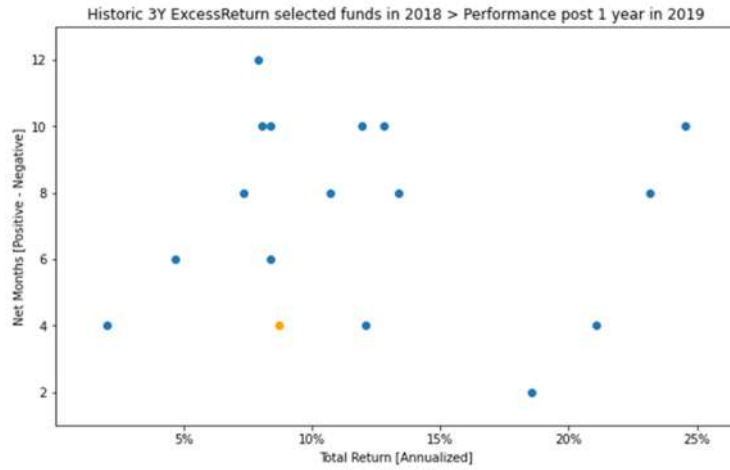
Figure 22.

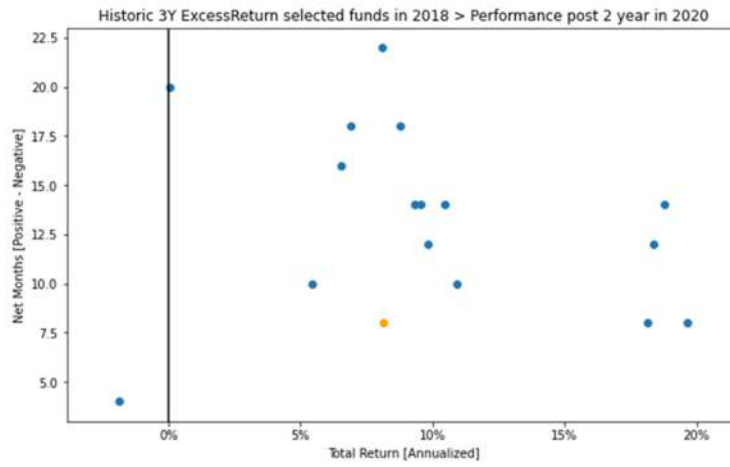Historic 3Y ExcessReturn selected funds in 2018 > Performance post 1 year in 2019

Figure 23.

Historic 3Y ExcessReturn selected funds in 2018 > Performance post 2 year in 2020

Figure 24.

Historic 3Y ExcessReturn selected funds in 2018 > Performance post 3 year in 2021

Figure 25.

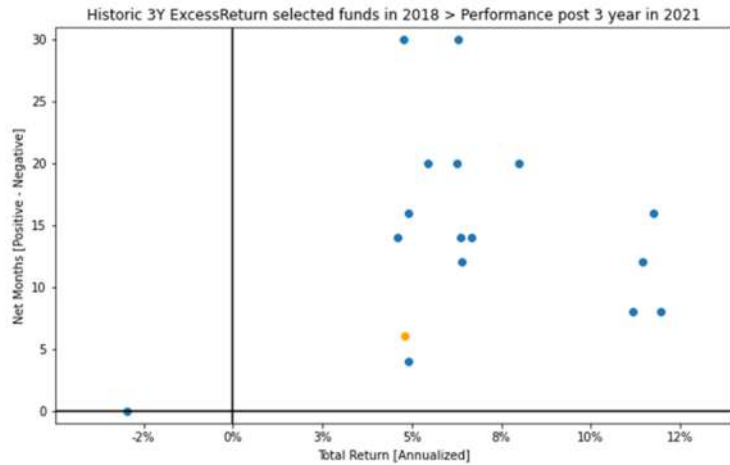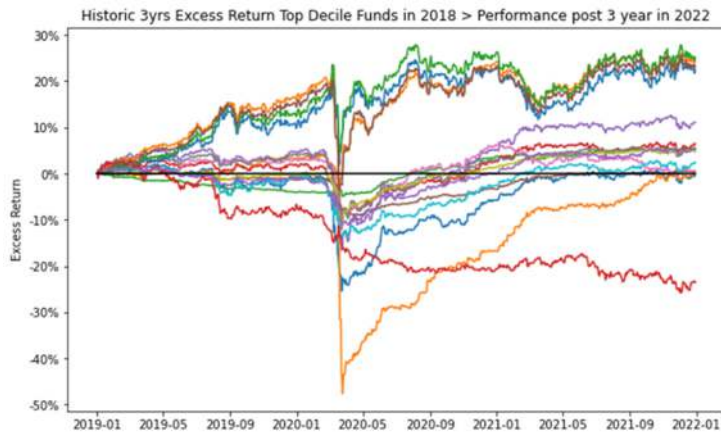Historic 3yrs Excess Return Top Decile Funds in 2018 > Performance post 3 year in 2022

Future performance of the top decile for funds selected on 2018-12-31 according to Excess Return Ranking (Figures 26, 27 and 28). As with Figure 25, over the 3-year period the funds normalize beyond the short-term volatility to trend back to the Excess Return quadrant. Figures 26, 27 and 28 assess the Excess return and net positive/negative returns of the previously selected top quartile funds in the future 1, 2 and 3 year periods.
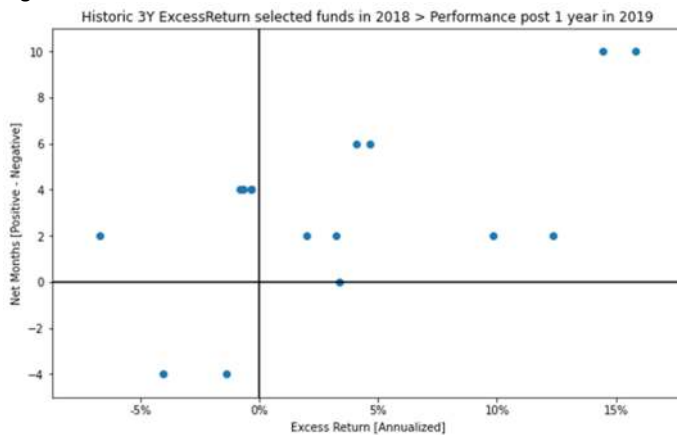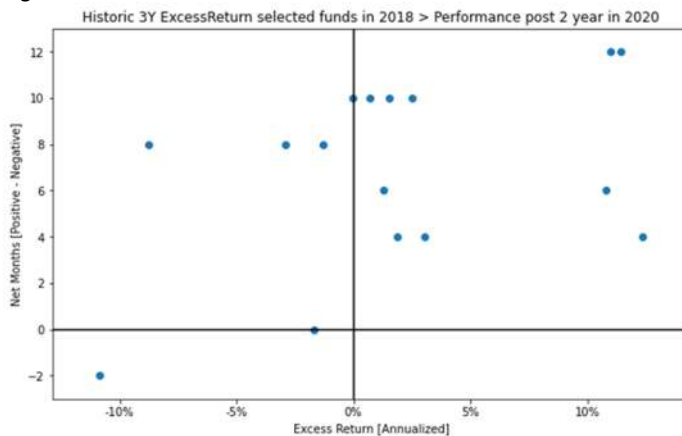
Figure 26.

Historic 3Y ExcessReturn selected funds in 2018 > Performance post 1 year in 2019

Figure 27.

Historic 3Y ExcessReturn selected funds in 2018 > Performance post 2 year in 2020

Figure 28.

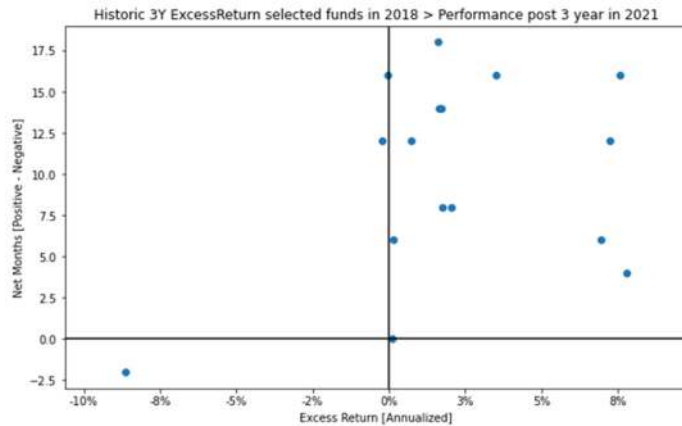Historic 3Y ExcessReturn selected funds in 2018 > Performance post 3 year in 2021

Figure 29 looks at 1 year cumulative return profile funds in the positive quartile of the funds in Figure 26. And, Figure 30 for the funds in Figure 27. Over shorter periods the performance expectations are more volatile.
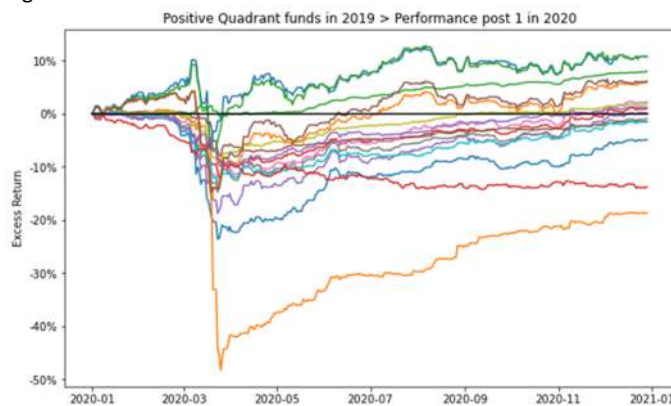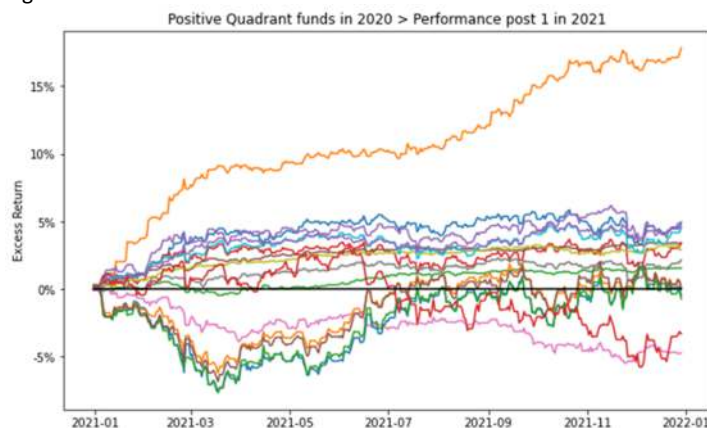
Figure 29.

Positive Quadrant funds in 2019 > Performance post 1 in 2020

Figure 30.

Positive Quadrant funds in 2020 > Performance post 1 in 2021

As highlighted above the historic analysis for the US AGG market seems to show promise when Bloomberg US Agg Bond TR USD is the index. However, deriving true value requires a lot of what ifs for isolating feature and event importance as points of entry/exit can dramatically impact the results due to the volatility shown in Figures 1, 2 and 3. The What ifs can be an anecdotal or iterative process as is expected from a fundamental type analysis.

Contact us for information about a particular fund, performance measure, time period, etc.

Next
Insights 2.10. Mutual Fund Manager Selection - Basic Historic Analysis: Are you always wrong?
Focus: US Large Cap Mutual Funds (US LC)

Background
Insights 2.00. Mutual Fund Manager Selection - Setting up the framework

Insights 2.10. Mutual Fund Manager Selection - Basic Historic Analysis: Are you always wrong?
We begin by holistically looking at the US mutual fund manager landscape from a historical fund price perspective and assess the ability of widely used performance measures for manager selection. This is done both at the market and individual fund level. Then as simple extensions we evaluate regressions for generally fixed weighting schemes of performance measures over fixed time periods and during discrete regimes. We look at simple back testing and predefined simulations. We will give Insights for every Asset Class.

Insights 2.20. Mutual Funds - Is there value in leveraging larger datasets?
We incorporate larger volumes of macro data, market data, performance measures, holding data, alternative data, etc. We introduce forms of feature engineering to generate signals for regimes, factors, indicators and measures using both raw and reduced datasets. We also introduce synthetic data generation to supplement sparse datasets.

Insights 2.30. Mutual Funds - Machine and Deep learning edge?
We incorporate evolving market conditions, performance measures, weights, events, predictions, etc. by leveraging Machine Learning techniques for real time and simulated multivariate analysis. Then we allow the system to do feature and event engineering by assessing various Deep Learning methods.

Extensions can be drawn to other types of managers, assets and markets. Here we will stay at the framework level, but will refer to our other papers that delve into the technical nuances and discoveries. Additionally, we will share similar series of Machine and Deep Learning framework papers for other aspects of the Investment lifecycle - asset allocation, portfolio management, risk management, asset planning, product development, etc.

These are all underpinnings of our Platform, where it is built to support any/all permutation/combination of data/models/visuals.

Email: info@ask2.ai for questions.

NEXT

Insights 2.11. Mutual Fund Manager Selection - Basic Historic Analysis: Are you always wrong?
Focus: US Large Cap