A Work Project, presented as part of the requirements for the Award of a Master's degree in Finance from the Nova School of Business and Economics.

Trading at the opening bell:

Gap Filling strategy on the E-miniS&P500

Michael Trequattrini

Work project carried out under the supervision of:

Nicholas Hirschey

10-01-2022

Abstract

The objective of the report is to develop a bidirectional gap filling strategy that will be used in the "Quantitative Investment Strategy" field lab's common part, which will create a portfolio of three investment strategies. The dataset here used was obtained from Tradestation and includes historical 1-minute prices of the cash session of the E-miniS&P500 futures from 01/03/2000 to 10/27/2021. For the validation, an innovative methodology is used, which increases the OOS performance stability. All the outcomes and the backtest were carried out by coding an entire engine in Python and accelerating it with Numba.

Keywords: finance, financial markets, trading strategy, hedge funds, modern portfolio theory, portfolio construction

This work used infrastructure and resources funded by Fundação para a Ciência e a Tecnologia (UID/ECO/00124/2013, UID/ECO/00124/2019 and Social Sciences DataLab, Project 22209), POR Lisboa (LISBOA-01-0145-FEDER-007722 and Social Sciences DataLab, Project 22209) and POR Norte (Social Sciences DataLab, Project 22209).

1. Introduction

The most important daily event in the global stock market is the opening gap in the E-miniS&P500. Despite the markets' closure, the economy continues to provide a steady stream of news from around the world. This results in a massive increase in volume and liquidity, which may result in a gap at the beginning of the session. That being said, at the opening bell we will have price movements that can be used to exploit the gap's direction or not. These types of strategies are ideal for illustrating how traders' psychology works. After a gap at the opening of the session, the price can either return to the previous day's close (a phenomenon known as "gap filling") or it can follow the initial direction of the gap (gap continuation).

2. Gap Analysis

2.1 What is a gap?

The classic definition of "gap" is the price difference between the opening of a financial instrument and its closing on the previous day. If the price retraces its steps during the trading day until it reaches the previous day's closing, the gap is defined as "filled" or "closed," and if it exceeds this level until it reaches the maximum/minimum of yesterday, it is defined as "deep filled." Gaps can be given different names depending on the position of the opening price. When the price opens above/below the maximum/minimum of the previous day, we refer to it as a "gap" up/down, whereas a "lap" up/down occurs when the price remains within the range of the previous day. To adhere to editorial standards, when discussing a single case, the terms "gap" and "lap" will be used interchangeably, whereas "cap" will be used when the analysis requires the sum of all possible cases (gap and lap) rather than the single event.

2.2 Advanced gap analysis

Selection is a critical component of implementing a gap filling strategy. Three variables provide useful and reliable information for identifying profitable setups: the size of the gap, the zone in which the gap is realized, and seasonality. As illustrated in Figure 1, there is a 52.58% probability of observing a cap up and a 45.84% probability of observing a cap down. The critical factor is that 62.51% and 65.32% of these caps up/down are filled on the same day[1].

```
ALL CAP UP:

    Number of Cap Up:                   2894
    Percentage of Cap Up:               52.58 % of total bars
    Total Entity of Cap Up:             21013.5 points
    Average Entity of Cap Up:           7.26 points

    Number of Cap Up Refilled:          1809
    Percentage of Cap Up Refilled:      62.51 % of caps up
    Total Entity of Cap Up Refilled:    8554.5 points
    Average Entity of Cap Up Refilled:  4.73 points
```

```
ALL CAP DOWN:

    Number of Cap Down:                   2523
    Percentage of Cap Down:               45.84 % of total bars
    Total Entity of Cap Down:             19042.25 points
    Average Entity of Cap Down:           7.55 points

    Number of Cap Down Refilled:          1648
    Percentage of Cap Down Refilled:      65.32 % of caps down
    Total Entity of Cap Down Refilled:    7917.5 points
    Average Entity of Cap Down Refilled:  4.8 points
```

Figure 1

2.3 Size analysis

The probability of filling a gap decreases as the number of points increases (the size of the gap), both for a gap up and gap down. On the other hand, an increase in the number of points eventually gained results in an increase in the strategy's average trade. Additional information is available in Exhibit 1 of the Appendix.

2.4 Seasonality analysis

In each month, there is an average probability of filling greater than 60%, demonstrating the strength of the bias intended to be exploited. Exhibit 2 in the appendix shows that November (probability of 57.53%) and September (probability of 58.77%) are the most "challenging" months for caps up and down, respectively. Daily subdivisions are shown in Table 1, and it can be seen that the most unfavourable days are Monday and Friday. Since the daily split contains

---

[1] The results presented here are for the cash session of the E-miniS&P500 futures. In addition, interesting results can be seen on instruments such as crude oil, soybeans, and currency futures.

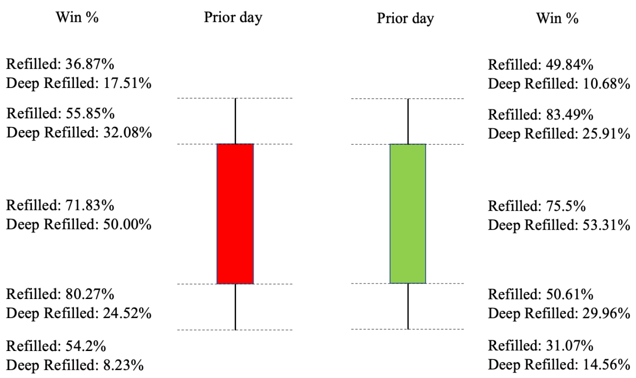
a greater number of observations than the monthly split, this information is more suitable for use as a filter in a trading strategy.

day	cap_up	cap_down	% cap_up refilled	% cap_down refilled	% cap_up deep_refilled	% cap_down deep_refilled
Monday	546	480	57.14	59.17	26.56	29.58
Tuesday	598	510	62.71	65.69	27.76	31.18
Wednesday	590	520	67.63	70.38	32.54	35.00
Thursday	567	525	62.96	68.38	29.45	31.81
Friday	593	488	61.72	62.30	22.93	26.64

Table 1

2.5 Zone analysis

The zone analysis is the most powerful criterion because it takes into account the size of the gap, its location, the daily trend, and traders' psychology. The Infographic 1 illustrates the probability of a gap being filled based on the open position and the previous day's trend. This information is critical because it enables us to select only the most promising configurations.



Infographic 1

3. Long side

3.1 Validation process

The greatest risk when building a trading strategy is overfitting. That is, the risk of over-optimizing the strategy's parameters by producing exceptional results that are unlikely to repeat in the future due to the mutability of markets. A validation system is responsible for determining whether the strategy is in sync with the market and whether the likelihood of remaining in sync in the future is sufficiently high. The classical validation model depicted in Figure 2 (Architecture A) causes a distancing in the past of the strategy's genesis, increasing the likelihood of deterioration due to the mutability of markets. Additionally, the validation outcome is highly dependent on the arbitrary selection of the training and testing periods. Instead of this, the validation model described in Architecture B will be used in this paper. It is based on the theory of periodic signals: if the price series is sampled at a frequency greater than the typical ones that characterize its movements, it is possible to decouple it from particular configurations. As a result, we will create IS periodic intervals of 35000 bars separated by as many OOS periods to obtain 32 blocks for both the learning and testing periods. This "comb" will enable us to model and test the strategy under all possible trend and volatility conditions, significantly increasing the system's robustness. After creating the intervals, the strategy will be tested on the IS aggregate, then on the OOS aggregate, and finally objective methods will be used to determine whether the validation was successful or not.

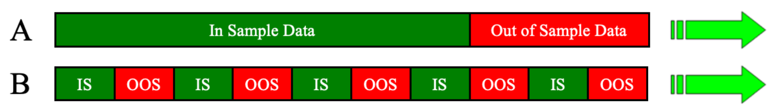


Figure 2

3.1 In Sample period

With the information from the "Zone Analysis" and a selection of areas with a higher probability, we will have two distinct entry setups based on the previous day's positive or negative close. If we come from a positive daily candle, we will consider only gaps whose opening is lower than yesterday's close but greater than yesterday's minimum. On the other hand, for a previous negative daily candle, it is sufficient that the opening is lower than the previous day's close. In both cases, we allow the system to enter regardless of time, but no more than one operation per day is permitted. The only filter proposed in IS (for both setups) is for the price at 1 minute to be greater than the daily 5-period simple moving average. This way, we can profit from the gaps down that occur during a strong short-term bullish trend. If all of the conditions (zone analysis and SMA filter) are met, the setup will be activated, searching for an entry and subsequent exit at the following conditions (the character "#" temporarily replaces the values that will be optimized):

- Enter level (limit order): opening price – #points
- Stop level (limit order): opening price – daily_range_D5 * #
- Target level (limit order): previous day's maximum
- Monetary stop loss e target (stop order): #$
- Exit rule (time exit, stop order): the end of the session

The entity "daily_range_D5" reflects the range of the previous five daily candles, defined as the difference between the maximum of the previous five days' highs and the minimum of the previous five days' lows. We adjust the stop to reflect recent volatility by subtracting a percentage of this range from the opening price[2]. The widespread use of limit orders enables a reduction in slippage, defined as the difference between theoretical and actual execution, which accounts for the majority of fixed costs (commissions + slippage). Obviously, there is still the possibility of meeting an intrinsic slippage because our order could not be executed. For the initial IS version, both the stop loss and monetary target were set to $600 in order to avoid affecting the results due to the ratio of the two measures. Something positive was obtained by using temporary initial parameters for the enter level and stop level, but not much attention is paid to the metrics for the time being, as they are the result of parameters chosen almost randomly. Keeping the stop and target values at $600, a two-variable optimization was

[2] Toby Crabel, a well-known American trader, popularized the technique.

conducted, comparing the system's most critical metrics as the enter level and stop level parameters were varied[3]. The choice must be made among the most stable pairs, where "stable" refers to the ability of the pair not to degrade the strategy's metrics if the market's dynamics change over time. To accomplish this, it is prudent to avoid pairs of parameters that are on the tips or near the cliffs on the 3D optimization surface, preferring instead plateau areas or saddle points where a change in dynamics is desirable in two of the four directions. The objective is to identify a pair of parameters that is stable across multiple metrics. This is because the assonance and unanimity of all metrics' indications is an indicator of robustness and stability.

To ensure objectivity, the Wind Rose algorithm has been implemented in Python, contributing to the creation of an automated process for mapping 3D surfaces to obtain stable values. The pairs identified by the algorithm are those that fall in the green zones. Obviously, this does not mean that the indication should be blindly followed, but it is an extremely useful tool. The surfaces of the major metrics are visible in Figure 3. The average trade was considered as the driving element because, for strategies of this type, it represents the achilles heel. The pair (0.5 for the stop level ; -2 for the enter level) was chosen because its position on a sort of cusp for the average trade almost always involves an improvement in its surroundings. In the other metrics, the pair in question always falls in plateau areas correctly identified by the Wind Rose algorithm, confirming the stability of the pair.

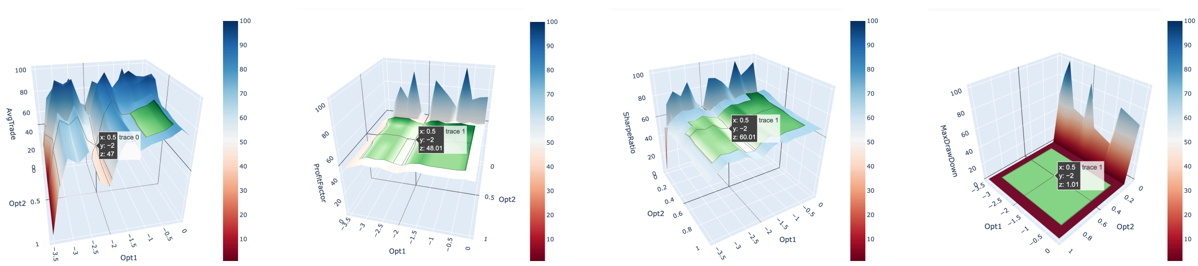


Figure 3

---

[3] The parameters mentioned refer to the "#" shown precedently.

Another combination (0.5 ; -0.75) produces more efficient values, but when we relaunch the same optimization by lowering the stop and monetary target to $300, we can see in Figure 4 how this pair lacks stability across all metrics, in contrast to the original combination (0.5 ; -2), which always falls within the algorithm-identified stable areas. This match bolsters our confidence in our selection and paves the way for further optimization of the stop and monetary target.

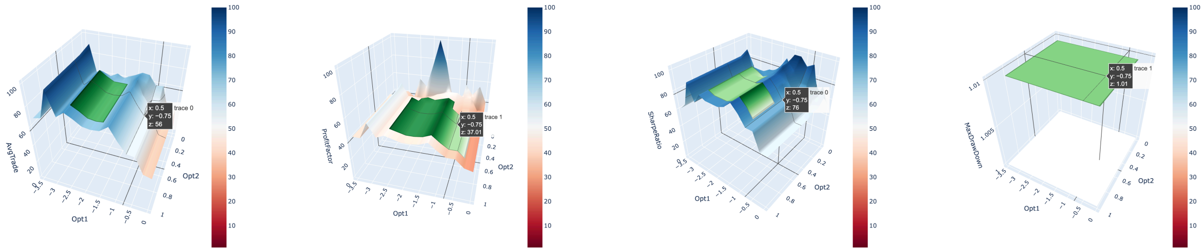


Figure 4

By applying the same optimization logic to the surfaces shown in Exhibit 3 in the Appendix, we obtain the pair ($500 for the monetary stop loss ; $250 for the monetary target). In this case, a stop loss greater than the target was preferred to give the probability of the gap being filled time to mature. With all the parameters in hand, it is possible to evaluate the final version of the IS. While the metrics are intriguing, the average trade value of $28.9 is cause for concern. This is because an average trade of at least $50 is required on this instrument for this type of strategy in this configuration. This does not, however, preclude us from proceeding, as we still have additional filters to apply at the conclusion of validation.

3.2 In Sample vs Out of Sample period

The comparison of IS and OOS for all major metrics is shown in Figure 5. Profit growth of 60% and the consistency of the number of trades are excellent indicators of robustness. The positive changes in the average trade, profit factor, and percent of winning trades, as well as the reduction in drawdown (both average and maximum)[4], are all excellent. As a result, it is clear that the OOS has not only maintained but also improved the metrics. Figure 6 illustrates the superimposition of the distributions of IS and OOS trades. The two distributions almost perfectly overlap, demonstrating how the OOS passes the validation process. To avoid any qualitative approach, it was used a model that compares the different percentiles of the two distributions in order to define a persistence[5] value. The threshold value is 1, and the resulting value is "infinite," which is an extreme case indicative of high persistence.

```
In Sample vs Out of Sample Statistics

Profit:                 [IS] 12343.75 [OOS] 31000.0 -> delta: 60 %
Operations:             [IS] 426 [OOS] 416 -> delta: -2 %

Average Trade:          [IS] 28.98 [OOS] 74.52 -> delta: 61 %
Profit Factor:          [IS] 1.24 [OOS] 1.83 -> delta: 32 %
Percent Winning Trades: [IS] 68.08 [OOS] 72.12 -> delta: 6 %
Reward Risk Ratio:      [IS] 0.58 [OOS] 0.71 -> delta: 18 %
Avg Open Draw Down:     [IS] -1591.34 [OOS] -566.66 -> delta: -181 %
Max Open Draw Down:     [IS] -4737.5 [OOS] -1700.0 -> delta: -179 %

Persistence Distribution Ratio: inf
Persistence Distribution Index (OOS): 100.0 %
Persistence Distribution Index (IS): 0.0 %
```

Figure 5

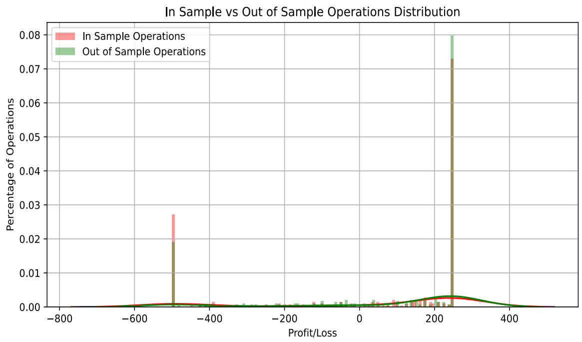


Figure 6

3.3 In Sample + Out of Sample

After passing validation, the system is ready to be tested on the entire dataset, with two additional filters: the first prevents the system from performing a classic Larry Williams style blast-off, requiring that the (body of yesterday's candle / yesterday's range) > 20%, and the second simply prohibits the system from operating on Monday and Friday, as these are the days with the lowest probability of the caps being filled. The filters' purpose is to slightly reduce the number of operations, thereby increasing the system's selectivity, the probability of success, and most importantly, the average trade.

---

[4] The open drawdown is used because it considers the entire life of the trade rather than just its conclusion, which makes it more accurate.
[5] Persistence is defined as a strategy's ability to not deteriorate over time.

The metrics of the strategy calibrated on a single futures contract across the entire dataset are shown in Figure 7. The Sharpe ratio can be calculated in a variety of ways, here we demonstrate two distinct methods: the first yields a value of 1.05 annualized when capital of $50,000 is allocated to the strategy, while the second yields a value of 7.04 when capital is not considered. Given that we are discussing a trading system, greater attention should be paid on the annual Calmar ratio, which, at 3.12, indicates that the strategy earns on average 3.12 times the maximum drawdown it has in a given year. The average trade of $73.06 is sufficient to cover commission costs, and the Kestner ratio of 0.65 indicates that the system is extremely regular. Finally, with a winning rate of 73.55%, the system is excellent from a psychological standpoint. It's possible to see that both the maximum and average open drawdowns are extremely constrained. For additional information on drawdown, see Exhibit 4 in the appendix.

```
Calmar Ratio (yearly): 3.12
Sharpe Ratio (yearly):  1.05
Sharpe Ratio (yearly no capital): 7.04
Kestner Ratio: 0.65

Profit:                     31487.5
Operations:                 431
Average Trade:              73.06

Profit Factor:              1.77
Gross Profit:               72637.5
Gross Loss:                 -41150.0

Percent Winning Trades:     73.55
Percent Losing Trades:      26.45
Reward Risk Ratio:          0.63

Avg Delay Between Peaks:  1993.06
Max Delay Between Peaks:  216705

Avg Time in Trade:  125
Max Time in Trade:  388.0
Min Time in Trade:  1.0

Trades Standard Deviation:   280.02224572057196
Equity Standard Deviation:   9542.38353609134

Avg Open Draw Down:        -625.31
Max Open Draw Down:        -2137.5

Avg Closed Draw Down:      -631.3
Max Closed Draw Down:      -2137.5
```

Figure 7

Figure 8 below illustrates the strategy's profit breakdown on an annual and monthly basis.

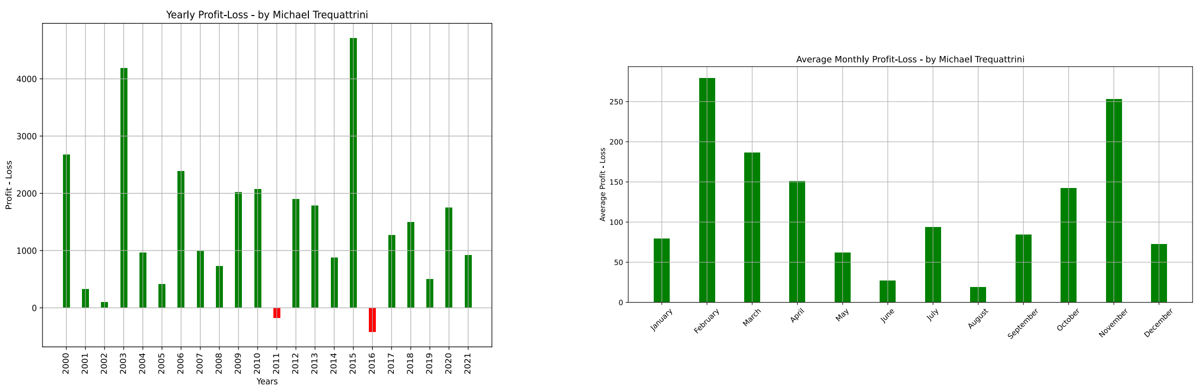


Figure 8

It can be seen that the system has only closed two years in the negative (gross of costs) and that, with the exception of the two outliers of 2003 and 2015, the amount of profit is fairly stable. On a monthly basis, it is worth noting that all months are positive on average. The worst months are June and August. The temptation is to shut down the system during that time, but there are no statistical or theoretical reasons to do so and we don't want to be a victim of any change in market dynamics.

Figure 9 shows the open equity line superimposed on the underlying, along with the IS (yellow) and OOS (light green) validation architectures. On the equity line, the yellow dots indicate new highs.

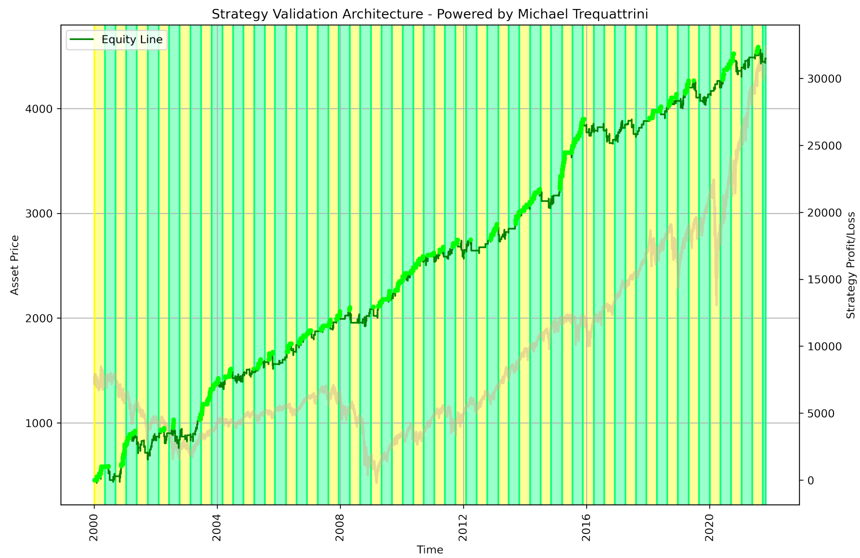


Figure 9

As can be seen, the dense validation architecture (comb-shaped) enabled the strategy to be modelled and tested against all recent and distant market dynamics. This has enabled to increase the robustness while also developing a strategy that is synchronized with the current market. Additionally, when the equity line is compared to the underlying, it is clear that the system responds well to both bullish and bearish trends, as well as flash crashes such as the one caused by Covid-19. This demonstrates the accomplishment of the goal of creating something that can break free from the market trend.

4. Short side

The logic used to develop the short strategy is identical to that used in developing the long strategy. Even in this case, the strategy will have two distinct setups based on the results of the zone analysis. If we come from a positive daily candlestick, we will consider only the caps whose opening is higher than yesterday's close but lower than yesterday's high. Similarly, in the case of a negative daily candlestick, the opening price must be higher than the previous day's close but lower than the previous day's maximum. Everything else, except for the removal of the SMA filter, is the mirror image of the long version. Using the same steps and optimization logic as in the long version, the pair of parameters is again (0.5 for the stop level ; 2 for the enter level)[6]. In this case, the parameters are the same as in the long version. This coincidence should be sought because it reinforces the pair's robustness even more, but it should not be forced. The chosen combination for the stop loss and monetary target is (650$; 650$). Using the same quantitative tools as for the long version (the results of which are shown in the appendix in Exhibit 5), it is possible to confirm that the short version also passes the validation with a persistence value of 4.5. That is, the OOS distribution's percentiles outperform the IS distribution 81.82% of the time.

4.1 In Sample + Out of Sample

Two filters have been added to the final version of the strategy: a prohibition on trading on Mondays and Fridays for the same reasons as on the long side; a gap size < (daily_range_D2 *

[6] The fact that we now have a sell limit order reflects a change in sign for the entry level parameter.

0.2)[7]. With this second filter, the system will narrow the selection even further by trading only small gaps that have the highest probability of being filled. The open equity line with the validation architecture superimposed on the underlying and the profit heatmap are shown in Figure 10. As we can see, the strategy performs admirably during bearish trends and defends itself optimally, even profiting during bullish trends. The heatmap demonstrates how the strategy's profits over the last five years have shown an increasing positive average, indicating the system's affinity with the current market. August has a negative average, but this could be due to outliers, and the system is not turned off in that month without statistical justification.

The strategy has a Calmar ratio of 1.67 on an annual basis, a Kestner ratio of 0.26, a profit factor of 1.5, and an average trade of $74.27. The average open drawdown is -$1040.17, while the maximum open drawdown is -$3550. The annual SR is 1.01 with a $50,000 capital.

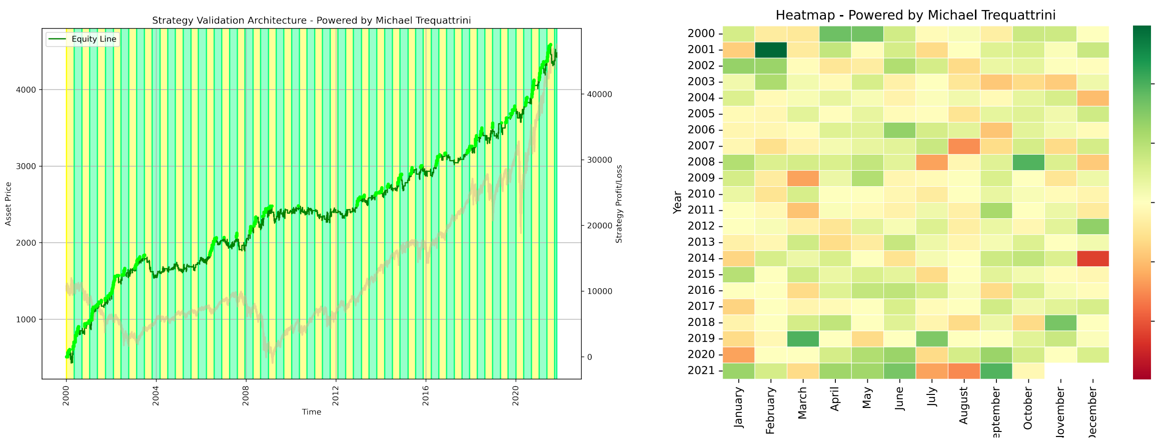


Figure 10

5. Long side + Short side

The only thing left to analyze is the combination of the long and short versions to determine whether their combined effect is beneficial or not. The open equity lines of the two strategies and their aggregation are depicted in Figure 11. At a glance, their sum appears to be profitable

[7] Daily_range_D2 is the difference between the high of the highs and the low of the lows of the previous two days.

and quite regular. This beneficial effect is due to the inverse correlation between the two systems, which can be qualitatively observed during the summer of 2001, the period from June to December 2004, as well as the period from 2009 to 2013. Mutual assistance is quantified by an inverse correlation coefficient of -0.00216.

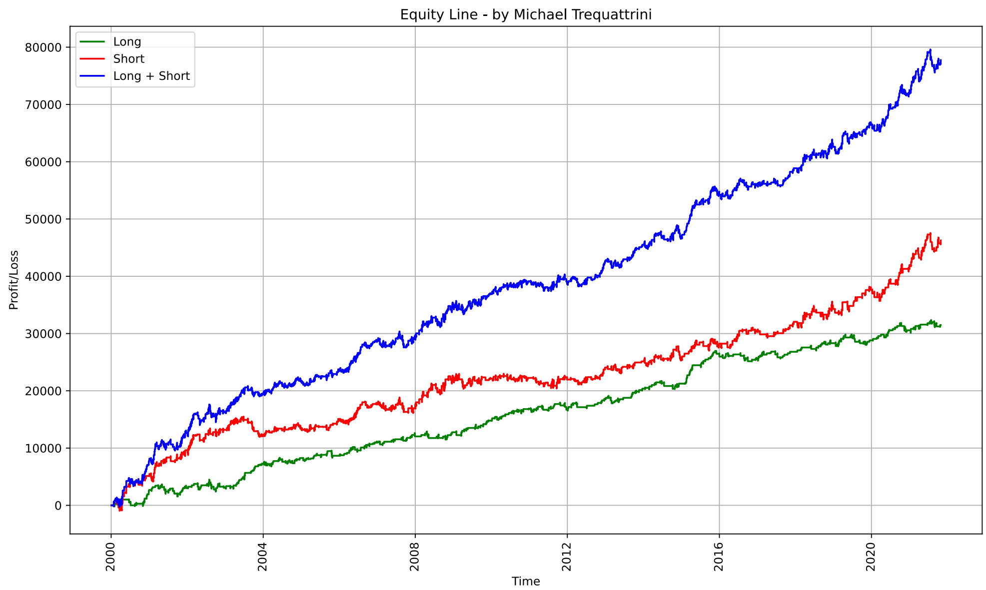


Figure 11

To establish realistic performance metrics for the strategy, it is necessary to first define the optimal capital allocation. There are two factors to consider: the broker's required margin of $14,000 on the E-miniS&P500 and a maximum drawdown of -$3400, which will be multiplied by two to account for any eventuality. This indicates that the minimum capital requirement is $20,800, but to avoid the risk of a margin call and an excessive % drawdown, $50,000 has been chosen. Figure 12 shows the metrics for the strategy. The average trade of $67.54 is sufficient to cover the costs, a Kestner of 0.5 indicates excellent regularity and a maximum drawdown (Exhibit 6 in the appendix) of -$3400 is extremely small. While a Sharpe ratio of 2.0 is excellent, this metric makes more sense for a portfolio of trading strategies whose inverse correlations, positive or negative, affect the portfolio's overall volatility. Most importantly, a

Calmar Ratio of 2.47 indicates that the average annual gain is 2.47 times the annual maximum drawdown.

PERFORMANCE REPORT:

Number of Operations:	1134
Profit:	76593.75
Compound Annual Growth Rate CAGR:	4.31 (capital = 50000)
Annual Return:	6.96 (capital = 50000)
Profit Factor:	1.74
Gross Profit:	179450.0
Gross Loss:	-102856.25
Average Trade:	67.54
Reward Risk Ratio:	1.01
Max Draw Down:	-3400.0 in date 2021-08-29
Draw Down Area:	-575562.5
Avg Draw Down:	-507.55

Max Delay Between Peaks:	62
Calmar Ratio:	22.53
Calmar Ratio Yearly:	2.47
Sharpe Ratio:	4.02
Sharpe Ratio Yearly (Capital = 50000)	2.0
Omega Ratio:	1.74
Kestner Ratio:	0.5
Percent Winning Trades:	42.77
Percent Losing Trades:	57.23
Avg Draw Down NoZero:	-740.75
Largest Winning Trade:	1300.0 in date 2000-04-16
Largest Losing Trade:	-1300.0 in date 2019-05-19

Figure 12

Figure 13 shows the profit divided into annual amounts. The system has shown an average annual increase in profit over the last few years, indicating excellent market synchronization. Additionally, the strategy performed admirably in turbulent years such as 2000, 2001, 2008, and 2020. Only August has a monthly negative average, but for the reasons stated previously, it was decided not to inhibit the system in this month. By examining the sequences of positive and negative months in Figure 14, it is possible to appreciate the profits' regularity.

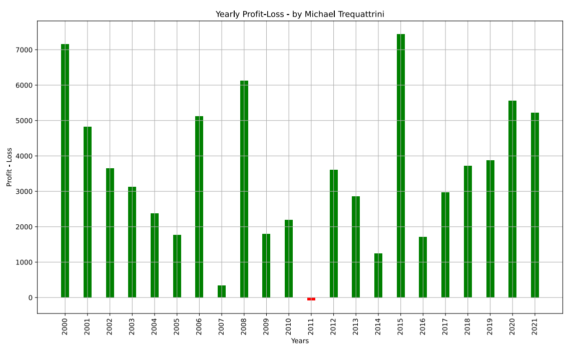


Figure 13

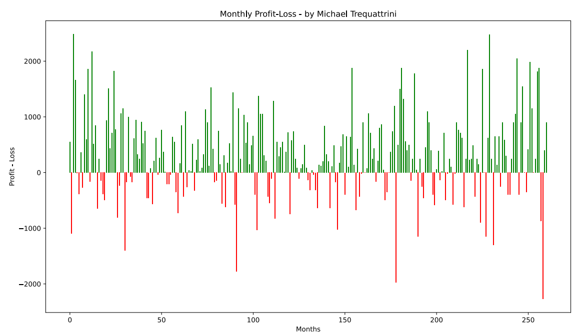


Figure 14

To conclude, we can attest that the system generates consistent profits with minimal drawdowns and is completely immune to market conditions. All of this demonstrates the system's robustness and its ability to be used with real money.

Appendix

Exhibit 1: Size analysis (Cap) on the E-miniS&P500

```
CAP UP 0-2 points:

    Number of Cap Up:                         724
    Percentage of Cap Up:                     13.15 % of total bars

    Number of Cap Up Refilled:                653
    Percentage of Cap Up Refilled:            90.19 % of caps up
    Average Entity of Cap Up Refilled:        1.07 points

    Number of Cap Up Deep Refilled:           292
    Percentage of Cap Up Deep Refilled:       40.33 % of caps up
    Average Entity of Cap Up Deep Refilled:   6.11 points

CAP UP 2-4 points:

    Number of Cap Up:                         604
    Percentage of Cap Up:                     10.97 % of total bars

    Number of Cap Up Refilled:                440
    Percentage of Cap Up Refilled:            72.85 % of caps up
    Average Entity of Cap Up Refilled:        3.03 points

    Number of Cap Up Deep Refilled:           184
    Percentage of Cap Up Deep Refilled:       30.46 % of caps up
    Average Entity of Cap Up Deep Refilled:   8.26 points

CAP UP 4-6 points:

    Number of Cap Up:                         475
    Percentage of Cap Up:                     8.63 % of total bars

    Number of Cap Up Refilled:                282
    Percentage of Cap Up Refilled:            59.37 % of caps up
    Average Entity of Cap Up Refilled:        4.97 points

    Number of Cap Up Deep Refilled:           121
    Percentage of Cap Up Deep Refilled:       25.47 % of caps up
    Average Entity of Cap Up Deep Refilled:   10.18 points

CAP UP 6-8 points:

    Number of Cap Up:                         282
    Percentage of Cap Up:                     5.12 % of total bars

    Number of Cap Up Refilled:                157
    Percentage of Cap Up Refilled:            55.67 % of caps up
    Average Entity of Cap Up Refilled:        7.04 points

    Number of Cap Up Deep Refilled:           79
    Percentage of Cap Up Deep Refilled:       28.01 % of caps up
    Average Entity of Cap Up Deep Refilled:   13.42 points

CAP UP 8plus points:

    Number of Cap Up:                         809
    Percentage of Cap Up:                     14.7 % of total bars

    Number of Cap Up Refilled:                277
    Percentage of Cap Up Refilled:            34.24 % of caps up
    Average Entity of Cap Up Refilled:        14.51 points

    Number of Cap Up Deep Refilled:           130
    Percentage of Cap Up Deep Refilled:       16.07 % of caps up
    Average Entity of Cap Up Deep Refilled:   24.18 points
```

```
CAP DOWN 0-2 points:

    Number of Cap Down:                         702
    Percentage of Cap DOwn:                     12.75 % of total bars

    Number of Cap Down Refilled:                632
    Percentage of Cap Down Refilled:            90.03 % of caps up
    Average Entity of Cap Down Refilled:        1.05 points

    Number of Cap Down Deep Refilled:           345
    Percentage of Cap Down Deep Refilled:       49.15 % of caps up
    Average Entity of Cap Down Deep Refilled:   5.05 points

CAP DOWN 2-4 points:

    Number of Cap Down:                         517
    Percentage of Cap DOwn:                     9.39 % of total bars

    Number of Cap Down Refilled:                395
    Percentage of Cap Down Refilled:            76.4 % of caps up
    Average Entity of Cap Down Refilled:        3.0 points

    Number of Cap Down Deep Refilled:           187
    Percentage of Cap Down Deep Refilled:       36.17 % of caps up
    Average Entity of Cap Down Deep Refilled:   7.18 points

CAP DOWN 4-6 points:

    Number of Cap Down:                         366
    Percentage of Cap DOwn:                     6.65 % of total bars

    Number of Cap Down Refilled:                216
    Percentage of Cap Down Refilled:            59.02 % of caps up
    Average Entity of Cap Down Refilled:        5.11 points

    Number of Cap Down Deep Refilled:           85
    Percentage of Cap Down Deep Refilled:       23.22 % of caps up
    Average Entity of Cap Down Deep Refilled:   8.81 points

CAP DOWN 6-8 points:

    Number of Cap Down:                         222
    Percentage of Cap DOwn:                     4.03 % of total bars

    Number of Cap Down Refilled:                133
    Percentage of Cap Down Refilled:            59.91 % of caps up
    Average Entity of Cap Down Refilled:        6.93 points

    Number of Cap Down Deep Refilled:           57
    Percentage of Cap Down Deep Refilled:       25.68 % of caps up
    Average Entity of Cap Down Deep Refilled:   10.8 points

CAP DOWN 8minus points:

    Number of Cap Down:                         716
    Percentage of Cap DOwn:                     13.01 % of total bars

    Number of Cap Down Refilled:                272
    Percentage of Cap Down Refilled:            37.99 % of caps up
    Average Entity of Cap Down Refilled:        14.87 points

    Number of Cap Down Deep Refilled:           106
    Percentage of Cap Down Deep Refilled:       14.8 % of caps up
    Average Entity of Cap Down Deep Refilled:   20.33 points
```

Exhibit 2: Seasonality analysis on the E-miniS&P500

month	cap_up	cap_down	% cap_up refilled	% cap_down refilled	% cap_up deep_refilled	% cap_down deep_refilled
January	238	204	65.97	65.20	34.03	34.31
February	213	202	63.85	73.76	26.29	36.14
March	254	219	61.02	63.01	28.74	28.31
April	250	197	64.00	65.48	28.40	31.47
May	242	221	61.16	61.54	25.62	25.79
June	259	203	66.02	64.04	32.82	26.60
July	254	203	60.24	63.55	23.23	31.03
August	230	251	62.61	66.14	24.35	31.08
September	231	211	62.34	58.77	25.97	25.59
October	248	230	62.10	68.26	29.44	33.04
November	219	202	57.53	65.84	23.29	33.66
December	256	180	62.89	68.89	30.86	35.00

Exhibit 3: 3D optimization surfaces for the monetary stop and target

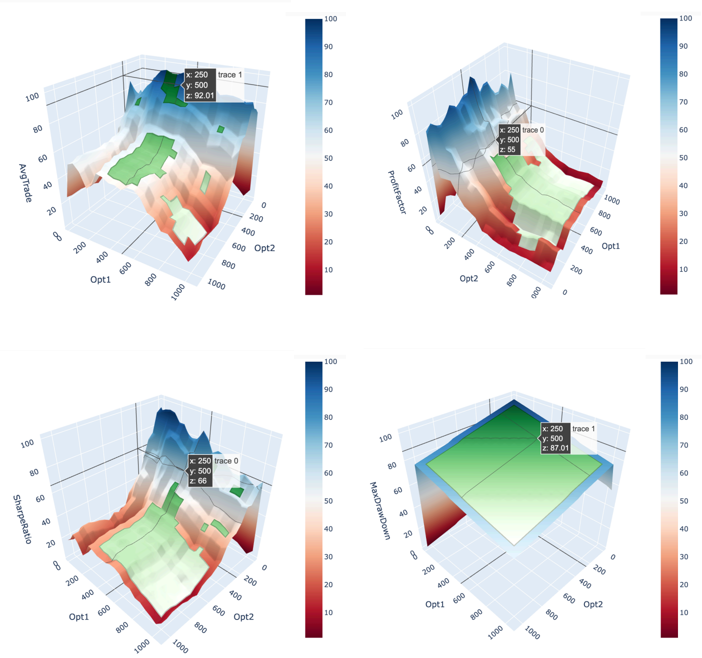


Exhibit 4: Long strategy's drawdown

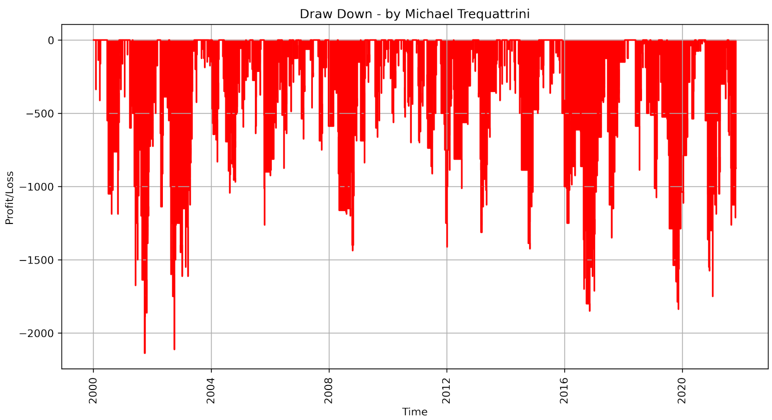


Exhibit 5: Validation analysis for the short side

```
In Sample vs Out of Sample Statistics

Profit:                  [IS] 26052.5 [OOS] 30105.0 -> delta: 13 %
Operations:              [IS] 851 [OOS] 770 -> delta: -11 %

Average Trade:           [IS] 30.61 [OOS] 39.1 -> delta: 22 %
Profit Factor:           [IS] 1.17 [OOS] 1.24 -> delta: 6 %
Percent Winning Trades:  [IS] 57.93 [OOS] 58.31 -> delta: 1 %
Reward Risk Ratio:       [IS] 0.85 [OOS] 0.89 -> delta: 4 %
Avg Open Draw Down:      [IS] -2207.55 [OOS] -1896.05 -> delta: -16 %
Max Open Draw Down:      [IS] -7500.0 [OOS] -5787.5 -> delta: -30 %

Persistence Distribution Ratio: 4.5
Persistence Distribution Index (OOS): 81.82 %
Persistence Distribution Index (IS): 18.18 %
```

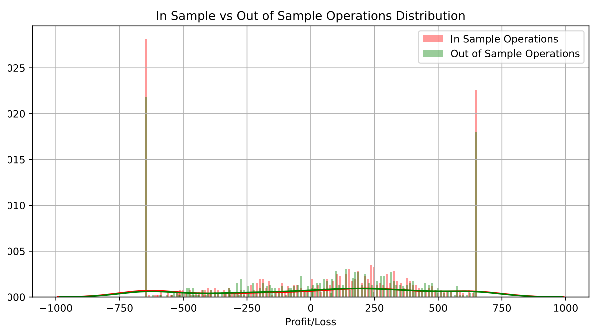
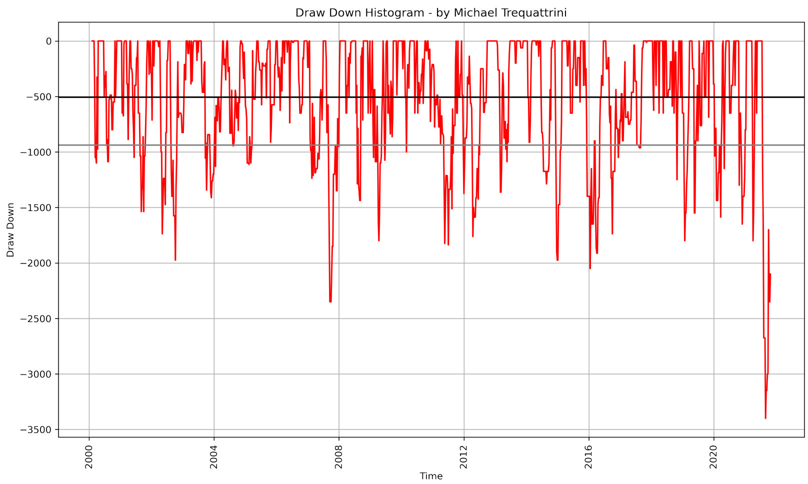


Exhibit 6: Final strategy's drawdown (black line = mean, grey line = 95% percentile)



Bibliography

- Davey, Kevin J. 2014. *Building Winning Algorithmic Trading Systems: A trader's journey from data mining to monte carlo simulation to live trading.* John Wiley & Sons Inc, first edition.
- Fifield, Aaron. *Chat With Traders.* YouTube channel
- *TraderLion*. YouTube channel

A Work Project, presented as part of the requirements for the Award of a Master's degree in Finance from the Nova School of Business and Economics.

TESTING MEAN-VARIANCE ALLOCATION MODELS ON ALTERNATIVE QUANTITATIVE INVESTMENT STRATEGIES

Michael Trequattrini

Gabriel Zancu

Alec Gliga

Work project carried out under the supervision of:

Nicholas Hirschey

10-01-2022

Abstract

This work is the common component of a "Quantitative Investment Strategy" Field Lab project. The objective is to combine the individual strategies created in the individual component of the same project. The strategies are fairly uncorrelated and have high risk-adjusted returns. Portfolio construction relied on mean-variance models namely the maximum Sharpe ratio, minimum volatility, maximum diversification ratio, and equal-risk models. The equal-risk produced the highest returns and lowest risk metrics overall and outperforms common portfolios like the 60/40. The results were further tested against their ex-post optimal counterparts and revealed to be quite similar.

Introduction

This project aims to combine three previously created quantitative investment strategies. Different asset allocation models will be analyzed and the best one, based on risk and return metrics, will be selected. The strategies include a gap-filling trading strategy that uses equity futures, and two cryptocurrency strategies one focusing on cross-sectional momentum and the other focusing on sentiment trends. All three will be explained in more detail below.

Cross-Sectional Momentum in Cryptocurrency

Cryptocurrencies have become popular over the last years due to their extreme returns, and now many want to incorporate them into their portfolios. This quantitative cryptocurrency momentum strategy, crypto momentum from now on, provides a balanced solution that can be adjusted to the risk preference of any investor and may be used individually or as part of a multi-asset portfolio. Historically, it demonstrated high risk-adjusted returns and low correlation with the equity and cryptocurrency markets.

This strategy builds upon cross-sectional momentum, in other words, it goes long on the cryptocurrencies that had the best past performance and shorts the ones with the worst performance. As simple as it seems, it exploits behavioral biases of investors such as the overconfidence of informed traders and the overreaction of investors to price and news. Some argue that the high return is compensation for carrying crash risk, meaning that after big market drawdowns the price of past losers contains a high premium and when the market rebounds, past losers experience large gains and past winners underperform. This phenomenon is again explained by behavioral biases namely the fact that investors become even more risk-averse during extreme situations and ignore statistical probabilities. But this weakness is attenuated by applying a risk scaling mechanism based on volatility forecasts that dynamically changes the weights of the long and short leg so that when the market volatility is high (low) it invests less (more). Another advantage of this risk scaling mechanism is the fact that one can target an upper limit for the strategy's volatility to fit the needs of the portfolio or the investor.

A big distinction from other assets' momentum is that cryptocurrency momentum is only observable in those with the highest market capitalization. This occurs because they are the most liquid ones and liquidity is necessary to withstand demand and supply shocks in order to form momentum. As such, crypto momentum filters for size before filtering for momentum, and at the same time, it avoids the lesser-known and fraud-prone coins. Another strength of this strategy is the fact that transaction costs are low due to the high liquidity and in some countries, the capital gains on cryptocurrencies are not taxed under specific conditions. Additionally, it is a self-financed strategy due to the long-short component which is complementary to the risk scaling method that increases or diminishes leverage depending on the risk target.

The sample on which the strategy has been tested is not very extensive since cryptocurrencies have not existed for a long. Despite that, the results seem robust enough both in- and out-of-sample. The out-of-sample and after-fee crypto momentum returns show a Sharpe ratio of 1.83, and when targeting an annual volatility upper limit of 20%, it achieves an annual volatility of 16.97%, an annual return of 31.05%, and a maximum drawdown of 12.19%. During the same period, the cryptocurrency market achieved a Sharpe ratio of only 1.33 and its drawdown is multiple times higher, even when equalizing the annual volatility. Finally, the strategy demonstrates a high and statistically significant alpha against common equity and crypto factors. The only significant factors were the crypto reversal factor and the equity size factor, in the full model, and both were negative.

Gap filling on the e-mini S&P500

Despite the fact that the majority of markets are now open nearly 24 hours a day, we still have times when markets are closed, in contrast to the real economy, which operates without interruption. At these times of closure, any relevant news can create market imbalances, pushing operators to conduct after-hours trading. This can drag the opening price above or below the previous day's closing price, thereby creating a gap up or down. If, after a gap, the price retraces to at least the previous day's closing price, the gap is defined as "filled," and the phenomenon is called "gap filling". The strategy presented here is based on the empirical fact that 62.51% of gaps up and 65.32% of gaps down on the E-mini S&P500 are filled during the same day. To build the strategy, a statistical study dubbed "Zone analysis" was conducted, which by combining the opening, closing, high, and low prices identifies five zones in which a gap up/down can occur the following day. Thus, with or without additional conditions, we have the probability of a gap filling for each zone. With this information, entry setups for long and short sides are created. To avoid overfitting, both versions use the same entry and exit logic as well as the optimization process. Additionally, limit orders have been used to minimize slippage. Both the long and short versions are profitable, and their combination results in an extremely consistent strategy in terms of profits, owing to an inverse correlation between the two. Only 2011 closed in the negative from 2000 to 2021, and the average annual profit has been increasing in recent years, indicating a strong synchrony of the strategy with the current market. Except for August, every month has a positive average. The system was not turned off in that month due to a lack of statistical justification. With a $50,000 starting capital, the strategy generates an annual return of 6.96%, a profit factor of 1.74, a maximum drawdown of only -6.8% and an average drawdown of -1.48%. Sharpe and Calmar ratios are 2.0 and 2.47 on an annual basis, respectively, indicating very good profitability. In conclusion, a Kestner ratio of 0.5 reflects the strategy's incredible regularity and an average trade of $67.54 is enough to cover the costs. An innovative approach derived from signal theory was used for the validation process: by sampling the price series at a frequency greater than the typical ones that characterize its movements, it is possible to break free from particular configurations. A validation architecture based on this concept enabled the strategy to be modelled and tested in all market and volatility conditions, significantly increasing the strategy's robustness.

Cryptocurrency Google trends

Quantitatively driven investing methods rely on the transformation of data of various types into a signal that reflects the portfolio's position in the assets being traded. The data must have some causal link with the asset's returns for the strategy to be successful and outperform the market. The degree of interest in a specific asset throughout the world is should be a driver of its price, such as when inflation worries increase, investors have historically showed interest in gold as a hedge. In cryptocurrency markets most information about individual assets is retrieved from the internet, and generally investors can be assumed to procure this information using search engines. Due to the fact that Google has a more than 90% market share when it comes to search engines used worldwide, the number of google searches for the name of a particular cryptocurrency should be a good proxy for the general interest in a given currency.

Google search query volumes data is available via the pytrends API commonly used for research projects in this area. For a sample of around 80 cryptocurrencies, a signal has been produced from Google search query volumes, which captures the momentum of the trends data. To put it another way, when search query numbers rise, the strategy takes a long position because the asset is expected to rise, and conversely when query levels fall, we stay out of the market. Short positions were not used since they proved to diminish the strategies performance. To capture the change in search query, simple moving averages, as well as exponential moving averages were used, which resulted in a decently performing strategy that outperformed its benchmark pretty consistently through time and showed a close correlation to the market with a beta hovering around 1.

Data and Methodology

The data used in this project comprises the time series of the three base strategies and the time series for the ETFs that cover the SP500 (SPY), gold market (GLD), and the 20-year US treasury bonds (TLT). The source of the ETF data is Tradestation. The equity futures strategy extends from 2000 to 2021, the cryptocurrency momentum strategy spans from 2015 to 2021, and the cryptocurrency sentiment trend strategy ranges from 2010 to 2021. To have all three strategies present at the same time, the considered time frame will start from 2015, plus the returns are resampled to weekly.

The short time frame is caused by the short lifetime of the cryptocurrency market where, even though cryptocurrencies like Bitcoin have existed since 2010, the majority only appeared after 2015 and onward. This might be a limitation to the validity of our results since this period does not capture all different types of market environments. To that point, we would argue that, based on data from its individual research, the equity strategy performs in a very consistent way and maintains similar return and risk metrics in these 7 years as it did for the last 20 years, making the market environment a less crucial factor. Furthermore, cryptocurrencies only recently have become more sensitive to the rest of the markets and economic information. In these 7 years, two crypto market peaks and troughs were captured which can provide some insight into how the strategies will perform in extreme environments. Although not a perfect situation, we expect the analysis will provide at least minimally valid results.

Five different asset allocation models will be tested in this project, namely equal-weight, minimum volatility, maximum Sharpe ratio, maximum diversification, and equal-risk. Instead of splitting the data between an in- and out-of-sample, the optimization will be computed on a 26-week rolling window. Meaning that at the end of each week, the weight to be attributed to each strategy in the next week is calculated based on the last half-year worth of data. This prevents look-ahead bias while the data remains relevant for that period. There will be a gap in the first 26-weeks that is filled by simply giving equal weights to each asset. The constraints followed by all the models are that the sum of the weights should always be equal to 1 and that the weight of each asset should never be lower than 0 or higher than 1. The models follow the mean-variance framework first introduced by Markowitz (1952). His paper also includes the concept of the tangency and global minimum variance portfolios that in this project will be denominated as maximum Sharpe ratio and minimum volatility. To lay the foundation for all of the models used in this project, the most basic concepts are defined under Equation 1. The equations are constructed in matrix formats to compute the results more efficiently.

The simplest allocation method is the equal-weight portfolio that consists of assigning the same weight to each asset. DeMiguel, Garlappi and Uppal (2009) test the naïve $1/n$ allocation against 14 other models, and they conclude the equal-weighted portfolio on average outperforms the rest in terms of Sharpe ratio. So, this method should provide a good benchmark for the rest of the models. Its strongest points include not requiring the expected returns, which are hard to estimate, nor does it require any other type of forecast for that matter. It usually has a very low turnover which is a positive since more trades mean more costs and most importantly, it is very easy to calculate. Despite these strengths, the model fails to properly distribute the risk between assets.

The minimum volatility portfolio consists of creating a portfolio that has the lowest variance. It uses the sample-based variance and covariance estimates to calculate the total portfolio variance as described by Equation 2. The fact that it does not need expected returns as input is positive, but it still relies on sample estimates exposing it to estimation error. To arrive at the weights for each asset in each period, Equation 3 is followed. Although not optimal by risk-adjusted returns, the minimum variance portfolio has shown positive results both in DeMiguel, Garlappi and Uppal (2009) and Maillard, Roncalli and Teïletche (2010). Some might even relate it to the betting-against-beta factor since market risk is one of the components of variance.

The Maximum Sharpe Ratio model consists of creating a portfolio that has the highest Sharpe ratio (Equation 4) possible, based on the estimated mean return and standard deviation. This should theoretically produce the optimal portfolio, the one with the highest return per unit of risk. But it is not the case in practice, due to estimation error and because assets do not sustain the same return and risk characteristics in all periods. The fact that this model needs estimates for both return and volatility, makes it weaker than the minimum-variance model. The weights to be attributed to each asset, in each period, are produced by Equation 5.

The maximum diversification model follows the research in Choueifaty and Coignard (2008) that built a model based on the idea that diversification is the only "free lunch" in finance. They completely put aside the idea of incorporating returns in the model due to forecast difficulty. The focus is solely on the variance and covariance of the assets creating this way a diversification ratio that is nothing less than the weighted average risk of the assets as a proportion of the total portfolio risk, represented in Equation 6. Assuming that higher risk usually leads to higher returns then, in theory, the maximum diversification ratio should produce the optimal portfolio. Equation 7 was used to arrive at the weights of each asset, in each period.

Last but not least, we have the equally-weighted risk model that budgets risk such that each asset contributes equally to the total portfolio risk, accounting for their levels of correlation. In other words, it consists of a portfolio where each asset has the same risk contribution, which is described in Equation 8. This risk budgeting model is based on the research of Maillard, Roncalli and Teïletche (2010), although they admit the idea is not entirely new. In their study, the equal-risk portfolio reports a higher Sharpe ratio than the equal-weight and minimum variance portfolios, and a maximum drawdown similar to the former. Just like the other models, it has the advantage of not requiring return estimation, but it remains open to variance estimation risk. The minimization of the function described in Equation 9 provides the weights necessary to produce the equal-risk portfolio.

Additionally, two benchmark portfolios were created to serve as a benchmark for the final results. One of the most popular portfolios is the 60/40, which invests 60% in equities and 40% in long maturity bonds. It has historically performed well due to the negative correlation between equities and government bonds, particularly during periods of high volatility when investors seek to mitigate risk. The second portfolio invests equally in stocks, bonds, and gold. Gold provides additional diversification benefits and acts as an inflation hedge. Equities are represented by the SPY ETF, which tracks the S&P 500 index, bonds by the TLT ETF, which tracks 20+ treasury bonds, and gold by the GLD ETF, which tracks the price of gold bullion. Both benchmark portfolios are rebalanced weekly and no transaction costs other than the already implied ETF fees are considered. Their cumulative returns and drawdowns are depicted in Figure 11 and Figure 12.

Results and Analysis

Prior to constructing portfolios using the models just described, it is necessary to analyze the unique characteristics of each strategy in order to gain a better understanding of portfolio dynamics. Table 2 shows the metrics for the three distinct strategies, all of which are net of costs so that we can perform a realistic analysis. What immediately stands out is the volatility difference between the gap filling strategy (Equity Futures in the table) and the two cryptocurrency strategies. The crypto momentum strategy been rescaled to have an annual volatility of no more than 20%, which brings it in line with the most diversified equity portfolios. Furthermore, the crypto trends strategy has been deleveraged, making its volatility essentially equal to that of the crypto momentum strategy. The difference in volatility is due to the strategies' underlyings being different: E-miniS&P500 for the gap filling strategy and a basket of cryptocurrencies for the remaining two. This structural difference is reflected not only in terms of volatility but also in terms of average annual returns. Not surprisingly, crypto strategies generate higher returns than the one that attempts to profit from a small bias at market opening. A Sharpe ratio of 1.56 demonstrates the gap filling strategy's efficiency in managing the relationship between return and volatility. However, for crypto strategies that are by definition more volatile, greater emphasis should be placed on the Sortino Ratio, which is a more efficient measure in high volatility environments. The obtained values are excellent and, combined with the

Sharpe ratio of the Equity Futures strategy, it indicates that the systems presented here are highly efficient. It's worth noting that no strategy's maximum drawdown occurs concurrently with another's. This lays the groundwork for empirical evidence of correlations, which, being predominantly negative, will result in portfolio diversification benefits. A preliminary qualitative analysis can be conducted by examining Figure 1, which summarizes the cumulative returns of all three strategies. Notably, the single largest drawdown occurs in the crypto trends strategy, which suffers a period of weakness between 2018 and early 2019. Figure 2 illustrates how, over the same time period, the correlation between the Equity Futures and Crypto Trends strategies has decreased, causing the gap filling strategy's profitability to inflate and compensate for the loss. Similarly, observe how, from 2017 to 2018, the aforementioned crypto strategy produced excellent results, more than compensating for the lower profitability of the other two systems. Returning to the metrics, each strategy has a very high annual Calmar ratio, indicating that it earns a profit that is more than double the year's maximum drawdown. The skewness reveals how the momenetum strategy on crypto has a nearly normal distribution of returns, the crypto trend is right-skewed, and the gap filling is left-skewed. The Jargue-Bera test values corroborate what has been stated about the normality of the distributions. Simultaneously, it can be seen that the Equity Futures and Cryto Momentum strategies, have a close to mesokurtic distribution. Crypto trends strategy exhibits a leptokurtic distribution (excess kurtosis = 7), indicating a greater likelihood of seeing returns that deviate from the distribution's average value. In other words, abnormally high or low returns are not uncommon (fat tails).

Next, we are going to take a closer look at the portfolio weights assigned to each strategy by the different allocation models. The weights that each model attributes to each strategy, in each period, will vary based on the estimates derived from the previous 26-week data. These estimates include returns, volatility, and correlation between the strategies, and may not sustain in the next period.

First, regarding the minimum volatility portfolio in Figure 3 we can see how it tends to overweight the Equity Futures strategy, which is because it has the lowest volatility of all the strategies. During the second half of 2017, the Equity Futures strategy reaches a weight of almost 1, meaning that almost all the capital of the portfolio was invested in it. We attribute this to the fact that the volatility of both cryptocurrency strategies rose during that period and the correlation between them also spiked thus having a higher marginal contribution to risk. At the end of 2021, the equity futures strategy experienced some turmoil peaking to its lowest attributed weight in the sample.

Figure 4 displays the weights of the maximum Sharpe Ratio allocation method. Here we encounter even more extreme weights than in the previous allocation method since the strategies have certain periods where their risk-adjusted returns outperform the others. During 2018, the Crypto Momentum strategy stagnated and the Crypto Trends strategy retrieved below its peak. This occurred due to the decline of the cryptocurrency market during that period and resulted in lower Sharpe ratios for both. Overall this allocation method has the highest turnover and the least overweight towards equity futures over the whole sample.

The maximum diversification ratio portfolio weights in Figure 5 and equal-risk portfolio in Figure 6 show relatively similar allocations throughout the whole sample. This is most likely due to the fact that both are based on the diversification of risk, in other words attributing the same contribution of risk to each strategy. In both of these methods, the turnover is much smoother compared to the maximum Sharpe ratio model and somewhat similar to the minimum volatility model. This again demonstrates the superior performance of volatility-based models compared to the ones that include returns. Overall, the same general explanations given to the minimum volatility model can be applied to the maximum Sharpe and diversification models, in terms of spikes in the attributed weights.

Given the weights assigned to each strategy in each portfolio, it is possible to begin examining Figure 7, which depicts the log returns on the five portfolios described previously. As can be seen, a preliminary qualitative analysis provides a few intriguing hints. The returns on the equal-risk, maximum diversification, and minimum volatility portfolios all follow a similar pattern. For a more detailed analysis, see the Figure 8 for the three models, which show areas of colour that are similar, indicating that these portfolio construction models fail to produce meaningfully different results when based on only three strategies. We have the Equal-weight and the maximum Sharpe in the first and second positions for cumulative log-returns, respectively. In this case, we can see that there are indeed differences, which can be observed via the respective heatmaps in Figure 9. From mid-2016 to the beginning of 2017, the max Sharpe portfolio shows a majority of zones that lean toward red, in contrast to the equal-weight heatmap, which has greener colours (higher positive returns). From March 2019 to February 2020, the same pattern occurs, with the returns on the two portfolios largely in the opposite direction. In 2018, a turbulent year for both the stock and cryptocurrency markets, the maximum Sharpe allocation outperforms the equal weight, which suffers its largest drawdown. It is self-evident that allocating weights evenly during turbulent market years precludes one from benefiting from the diversification provided by the use of strategies across multiple markets. Concerning risk, Figure 10 illustrates the drawdowns of all portfolios. This comparison is extremely beneficial because it enables us to quickly identify failures in the systems or times in which diversification has not resulted in significant benefits. It is immediately apparent that the portfolios with the highest drawdown are the equal-weight and max Sharpe portfolios, remembering that the other side of the coin for having the highest returns is assuming a greater risk. The other systems experience drawdowns that are more limited in scope and, more importantly, more regular in occurrence, indicating a greater synchrony with market dynamics. Obviously, the best portfolio should not be chosen solely on the basis of return, which is why in Table 3, all of the key metrics for each portfolio are reported in order to enable a quantitative and objective comparison for the purpose of selecting the best. As previously illustrated graphically, equal-weight and maximum Sharpe ratio portfolios have the highest annual returns, but also the highest volatility and maximum drawdown. The remaining three allocations (Min volatility, Maximum diversification, and Equal-risk) are less profitable but also less volatile. The drawdowns are all less than 6%, and given that two of the three strategies operate in the crypto market, these are excellent results. This is especially impressive when compared to traditional benchmark portfolios containing SPY, GLD, and TLT, which have maximum drawdowns that are double, if not triple, and returns that are 2 to 3 times lower. Choosing the optimal portfolio is challenging because, if every metric is considered, the risk is that one will never reach a conclusion. Because the chosen approach is objective, one metric will be used as a selection criterion. The indicator chosen for consideration must be consistent with the portfolio's objective. Given that this is a long-term capital growth strategy, the optimal portfolio must have the highest Sharpe ratio and thus the best one is the equal-risk. This portfolio earns an impressive annual rate of 17.18% and has a maximum drawdown of just -5.37%. The annual log returns of the portfolio, net of costs, are depicted in Figure 13. Interestingly, there has never been a year that ended in the negative in the seven years of history. The portfolio does not make a profit in 2018, a turbulent year for the stock market and the crypto market, but it does not lose money either, ending the year at break-even. It's worth noting that annual profits have been increasing since 2019, indicating excellent synchronicity between the portfolio and the market and providing hope for the years ahead. More interestingly, Figure 14 shows a slightly more detailed analysis, including the average log returns for each individual month. This analysis is beneficial for establishing the presence or absence of bias that could account for systematic variations in portfolio performance. Immediately, it can be noticed how the first six months of the year account for the majority of the annual return. July's average is slightly negative, but the portfolio quickly recovers, with a positive average for every month until the end of the year. In conclusion, given that the average loss in July is negligible in comparison to the positive averages in the other months, and given that each strategy is constructed efficiently and that any potential temporal biases have already been factored in, it has been decided to keep the portfolio active in July as well, in order to avoid being victims of market dynamics changes.

When compared to the two proposed benchmarks, it is clear that the chosen portfolio outperforms them on every metric (return and risk). The log returns of the two benchmarks and the selected portfolio are depicted in Figure 15. As can be seen, the Equal-risk portfolio with its three strategies is capable of mitigating periods of distress. The two benchmarks run into difficulty at four points: in 2015, mid-to-late 2016, 2018, and early 2020 (Covid-19). During the same time periods, the chosen portfolio exhibits no distress, except for 2018, when it closes the year at break-even. This outstanding outcome was made possible by the unique characteristics of the individual strategies. The gap filling strategy is completely indifferent to market conditions and can profit from both bullish and bearish trends in the underlying market (S&P500). Even the cryptocurrency momentum strategy works both long and short, allowing it to be completely independent of market bias. On the other hand, the crypto trends strategy is monodirectional (long only), but its ability to focus exclusively on what is ”hot” at any given time, distances it from general market conditions. Combining these factors results in the creation of a portfolio that, even in its most basic form (equal-weight), produces interesting results while effectively mitigating risk. The portfolio is further improved by applying an equal-risk logic, resulting in the regularity depicted in the figure. The comparison to the two benchmarks shown here is obviously purely hypothetical, as two of the three strategies lack those underlyings. Due to the lack of a specific benchmark for our needs, a classic portfolio followed by an average investor has been chosen, to demonstrate how, by eliminating certain dogmas, it is possible to earn more while risking less.

Mean-variance optimization is based on estimates, as such, the results obtained may vary from the expected. To understand what would theoretically be the highest achievable Sharpe ratio we rely on 20,000 portfolio simulations derived from our three individual strategies. Figure 16 depicts the mean-variance cartesian where, identified by a red and a white star, we have the ex-post maximum Sharpe ratio and minimum volatility portfolios, respectively. These are the points achievable if we could perfectly predict future return and volatility. Represented by circles we have the 5 models we estimated and represented by squares we our strategies. The previously mentioned higher risk and return of the equal-weight and maximum Sharpe models are now graphically visible and located up and to the right of the rest. The ex-post optimal portfolio is, as expected, situated in the upper-most-left corner of the investable universe and the max. diversification and max. Sharpe model portfolios reveal themselves very close. Unfortunately, the minimum volatility model portfolio was not able to achieve the lowest volatility possible as evidenced by the difference to the ex-post example. The lowest possible volatility was 4.97% versus the 5.61 achieved and the highest Sharpe ratio possible was 2.96 versus the 2.77 achieved. Overall, we consider the used rolling estimation method very effective. Different lookback periods could have maybe produced better results, but optimizing this parameter could easily lead us to overfit, and so we chose to use a common window (26-weeks or half a year).

Conclusion

Our objective with this project was to individually create quantitative investment strategies and later combine them to hopefully achieve even superior results. The individual strategies are quite complex and based on non-traditional assets, namely futures and cryptocurrencies. Very succinctly, we have a trading strategy that exploits gaps at the opening bell in equity futures contracts, a cryptocurrency cross-sectional momentum strategy that scales its risk based on volatility forecasts, and a strategy that leverages online query data to extrapolate future returns of cryptocurrencies. The differences in volatilities and sample periods required some standardization to the original version of each strategy.

All three revealed to be uncorrelated between themselves providing us straight away with diversification opportunities. So we chose to proceed to utilize allocation methods based on the mean-variance framework, and instead of using the in- and out-of-sample split for training and testing the data, we relied on a rolling window method. The optimization models used are based on minimum volatility, maximum Sharpe ratio, maximum diversification ratio, equal-risk distribution, and the classic equally-weighted portfolio. Based on empirical evidence of academic research all these models have their merits. Some rely only on volatility which is must easier to estimate than returns, for example, but all have the downside of relying on historical estimates.

Upon assessing the results from each allocation method, and considering that our objective is to produce the highest risk-adjusted returns, we conclude that the equal-risk attribution method was the best since it achieve a Sharpe ratio of 2.88. The equal-weight and diversification ratio method also produced good results, both with a Sharpe ratio of 2.77 however, the former had much higher volatility and returns. While the minimum volatility model produces the lowest volatility of all the strategies and its Sharpe was ahead of only the maximum Sharpe ratio portfolio, which performed the worst. This performance can be attributed to the variability of returns from one period to another. Furthermore, the models have been tested against the exp-post optimal portfolio and the minimum variance portfolio and the results are very close. The lowest possible volatility was 4.97% versus the 5.61 achieved and the highest Sharpe ratio possible was 2.96 versus the 2.77 achieved. This, of course, considers the three individual strategies as the entire investment universe.

The equal-risk method achieved an annual return of 17.19%, annual volatility of 5.97%, and a maximum drawdown of only 5.37%. In the same period, popular portfolios like the 60/40 equity and bond portfolio had a return of only 7.83%, volatility of 10.60%, and a maximum drawdown of 19.64%. Despite containing futures and cryptocurrencies our star portfolio demonstrated to be a good alternative for those who want to grow their capital in a steady but not so slow way. The outcome is not only good but excellent.

We're also aware that our sample length is not very extensive and it does not capture all type of market environments. Despite that, we took precautions to guarantee that our results are as realistic as possible by accounting for transaction cost, and minimizing backtesting pitfalls such as overfitting and look-ahead bias.

.1 Appendix

.1.1 Tables

	Equity Futures	Crypto Momentum	Crypto Trends
Equity Futures	1.000000	0.033858	-0.105228
Crypto Momentum	0.033858	1.000000	-0.031579
Crypto Trends	-0.105228	-0.031579	1.000000

Table 1: Pearson correlation coefficient between the individual strategies.

	Equity Futures	Crypto Momentum	Crypto Trends
Annual Mean	8.81%	31.05%	26.52%
Annual Volatility	5.65%	16.97%	16.98%
Sharpe Ratio	1.56	1.83	1.56
Sortino Ratio	2.20	3.43	2.80
Max. Drawdown	-4.29%	-12.19%	-28.66%
Max. DD Date	2016-01-17	2019-12-01	2018-12-16
Calmar Ratio	2.55	3.93	3.41
Parametric VaR	0.09%	0.28%	0.28%
Skewness	-0.3006	0.0067	1.4297
Excess Kurtosis	0.8693	-0.2370	7.0632
J-B P-Value	0.0004	0.6700	0.0000
Start Date	2015-01-04	2015-01-04	2015-01-04
End Date	2021-07-11	2021-07-11	2021-07-11

Table 2: Performance statistics of the individual strategies. Both cryptocurrency strategies have been volatility scaled. Sortino Ratio is the proportion of the Annual Mean to the annual downside deviation, the Maximum Drawdown is the biggest peak to trough decline observed, Calmar Ratio is the Annual Mean in proportion to the Maximum Drawdown, the Parametric Value-at-Risk is the annual standard deviation multiplied by the 95 percent confidence level z-score, and the J-B P-Value is the p-value for the Jarque-Bera normality test.

	Equal-Weight	Min. Volatility	Max. Sharpe	Max. Diversif.	Equal-Risk	SPY&GLD&TLT	60%SPY&40%TLT
Annual Mean	22.13%	14.38%	20.47%	16.76%	17.18%	6.49%	7.83%
Annual Volatility	8.00%	5.61%	8.77%	6.04%	5.97%	9.60%	10.60%
Sharpe Ratio	2.77	2.57	2.33	2.77	2.88	0.68	0.74
Sortino Ratio	5.50	4.88	3.87	5.36	5.65	0.83	0.84
Max. Drawdown	-8.23%	-3.66%	-10.97%	-5.91%	-5.37%	-14.16%	-19.64%
Max. DD Date	2019-01-27	2019-01-27	2020-04-12	2019-01-27	2019-01-27	2020-03-22	2020-03-22
Calmar Ratio	7.77	6.18	5.98	7.18	7.46	0.68	1.05
Parametric VaR	0.13%	0.09%	0.14%	0.10%	0.10%	0.16%	0.17%
Skewness	0.4976	0.6868	0.5994	0.6501	0.6335	-0.6174	-0.9600
Excess Kurtosis	0.9412	2.0878	2.7774	2.0435	1.9108	10.3659	10.7571
J-B P-Value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Start Date	2015-01-04	2015-01-04	2015-01-04	2015-01-04	2015-01-04	2015-01-04	2015-01-04
End Date	2021-07-11	2021-07-11	2021-07-11	2021-07-11	2021-07-11	2021-07-11	2021-07-11

Table 3: Performance statistics of the results of all the allocation models and the two benchmark portfolios (SPY&GLD&TLT and 60%SPY&40%TLT). Sortino Ratio is the proportion of the Annual Mean to the annual downside deviation, the Maximum Drawdown is the biggest peak to trough decline observed, Calmar Ratio is the Annual Mean in proportion to the Maximum Drawdown, the Parametric Value-at-Risk is the annual standard deviation multiplied by the 95 percent confidence level z-score, and the J-B P-Value is the p-value for the Jarque-Bera normality test.

.1.2 Graphs

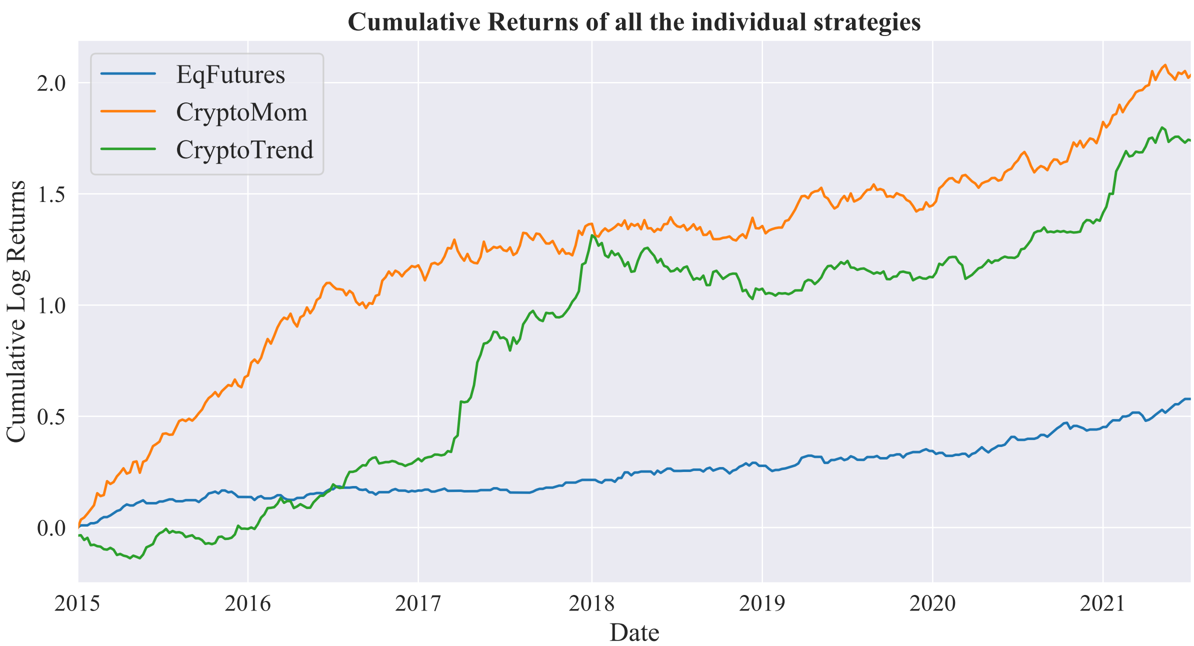


Figure 1: Log-scale cumulative returns of each individual strategy from the beginning of 2015 to mid-2021. EqFutures is the Equity Futures strategy, CryptoMom is the cryptocurrency momentum strategy and CrypotTrend is the cryptocurrency trend strategy.

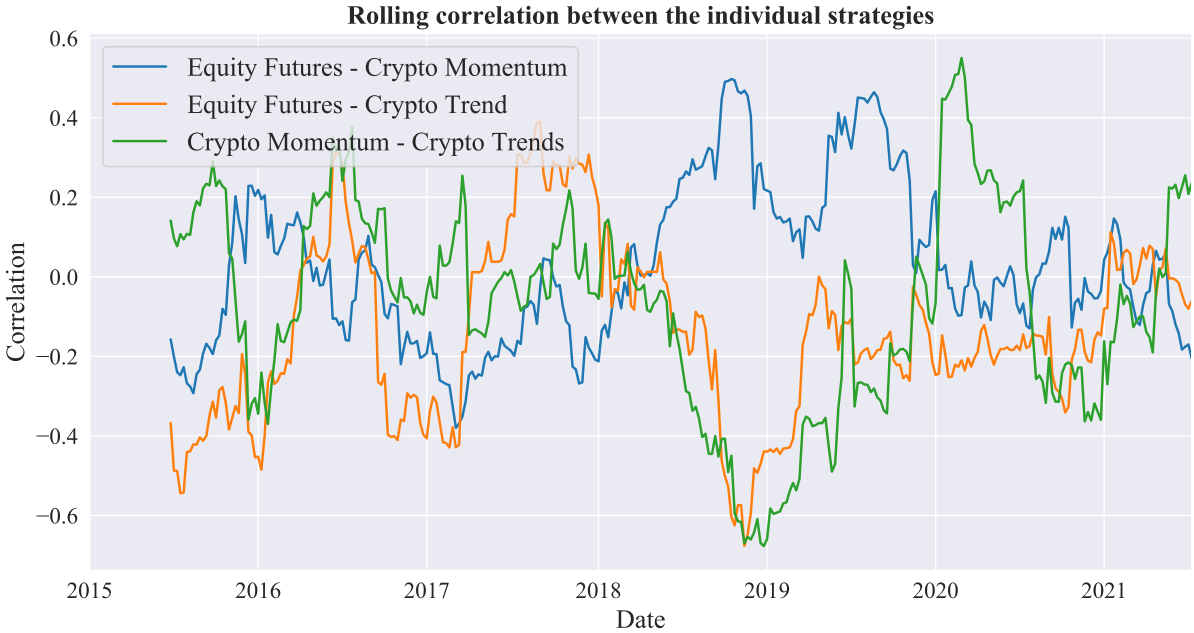


Figure 2: The 26-week rolling Pearson correlation between for the pairs between strategies, from the beginning of 2015 to mid-2021.

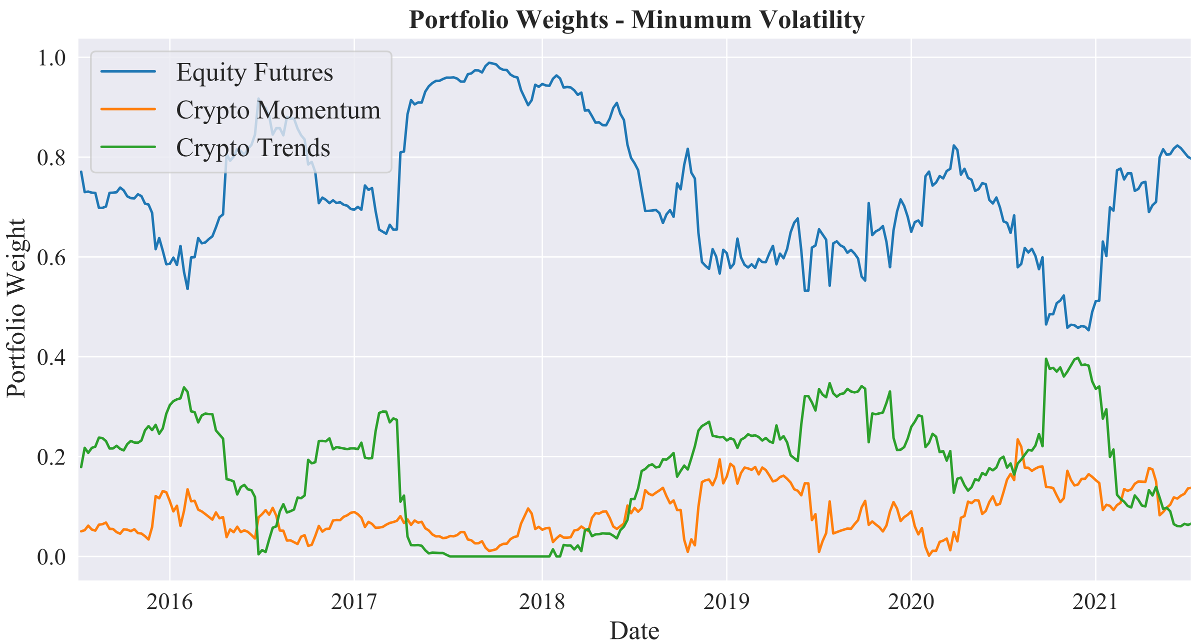


Figure 3: Portfolio weights attributed to each strategy by using a minimum volatility allocation model that rebalances each week and bases its estimates on the previous 26-weeks data.

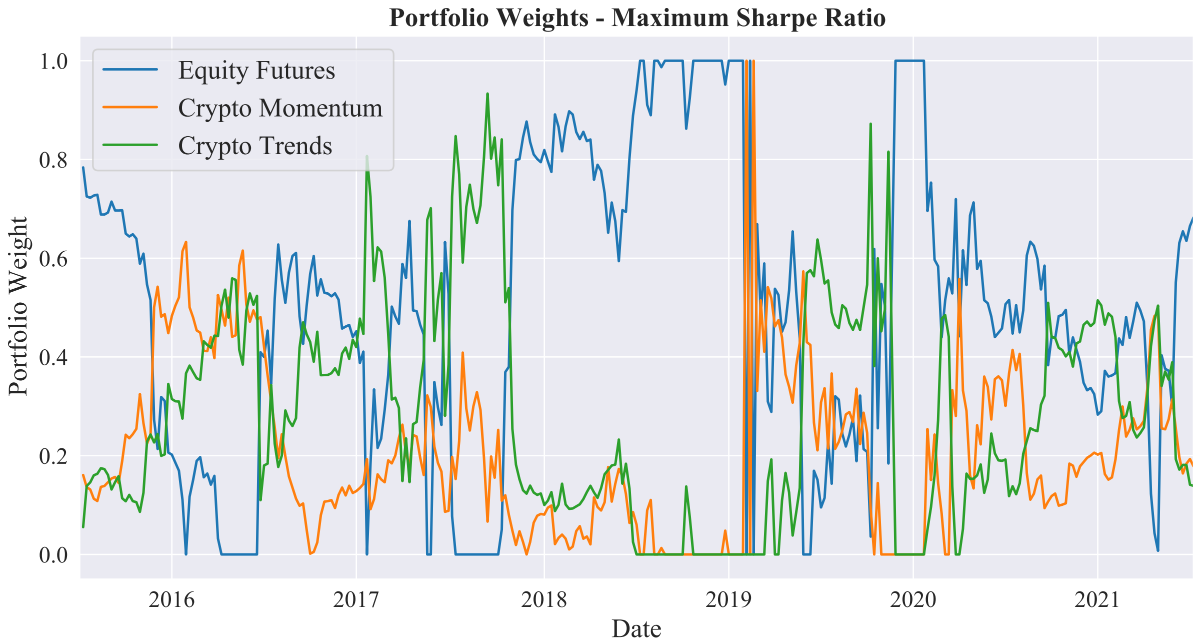


Figure 4: Portfolio weights attributed to each strategy by using a maximum Sharpe Ratio allocation model that rebalances each week and bases its estimates on the previous 26-weeks data.

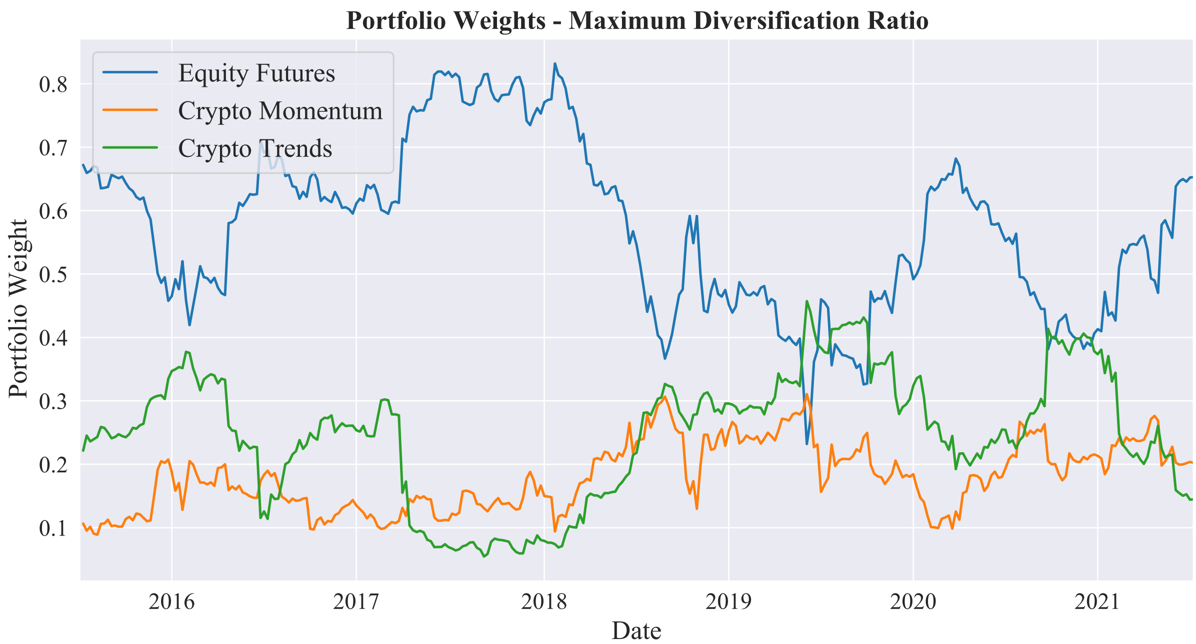


Figure 5: Portfolio weights attributed to each strategy by using a maximum diversification allocation model that rebalances each week and bases its estimates on the previous 26-weeks data.

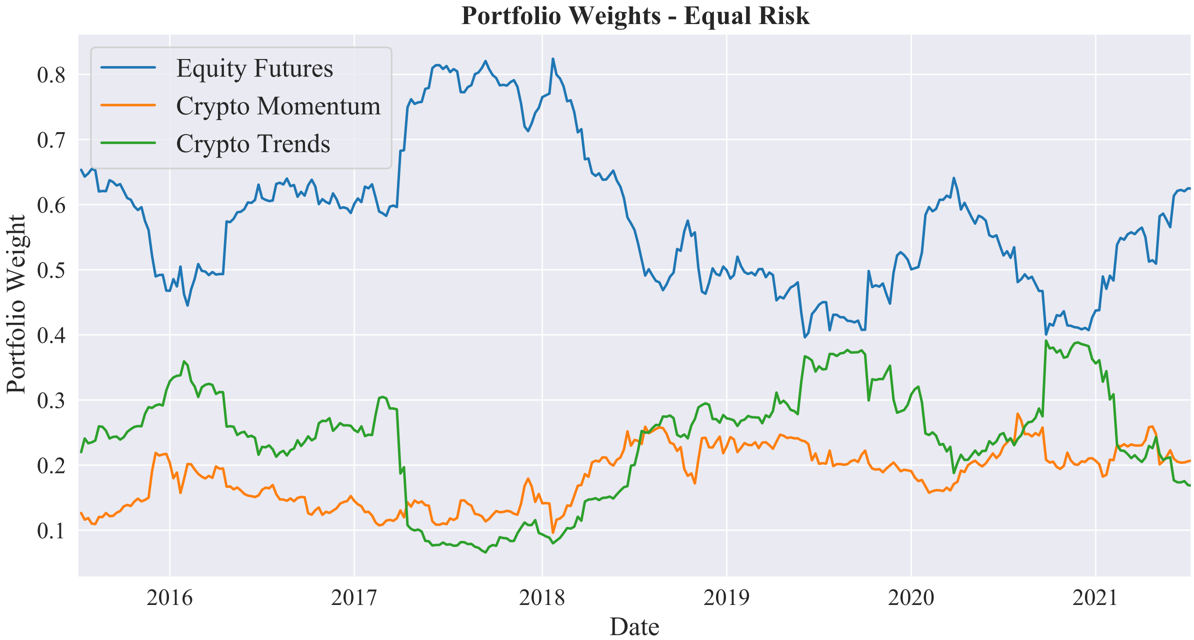


Figure 6: Portfolio weights attributed to each strategy by using an equal-risk distribution allocation model that rebalances each week and bases its estimates on the previous 26-weeks data.

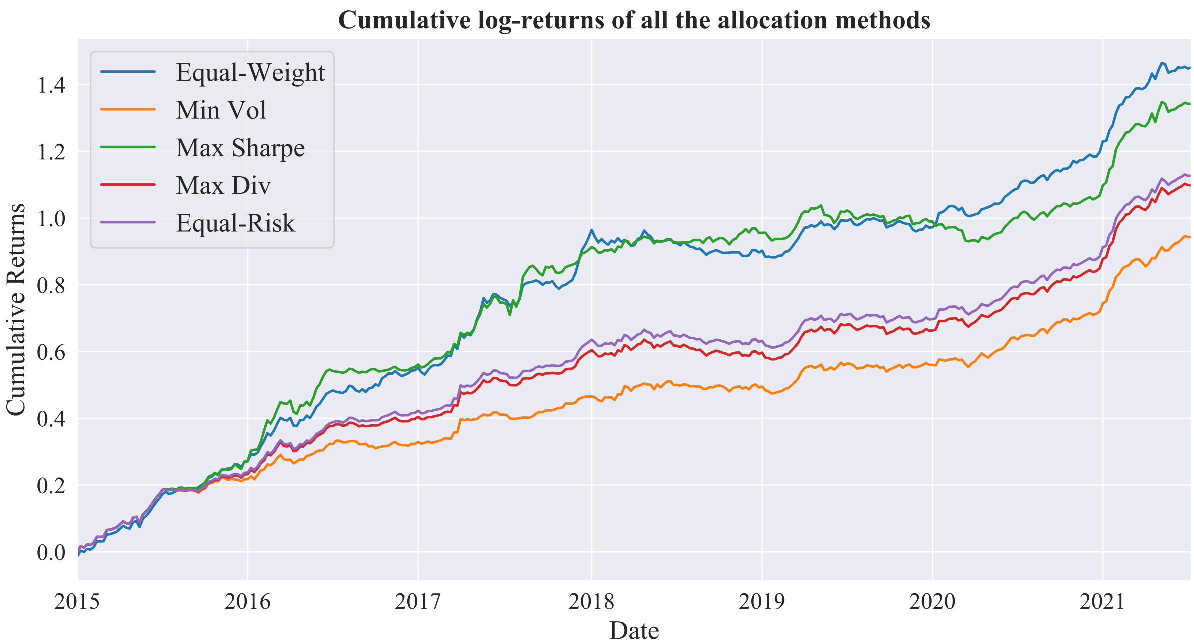


Figure 7: Log-scale cumulative returns of all allocation methods from the beginning of 2015 to mid-2021.

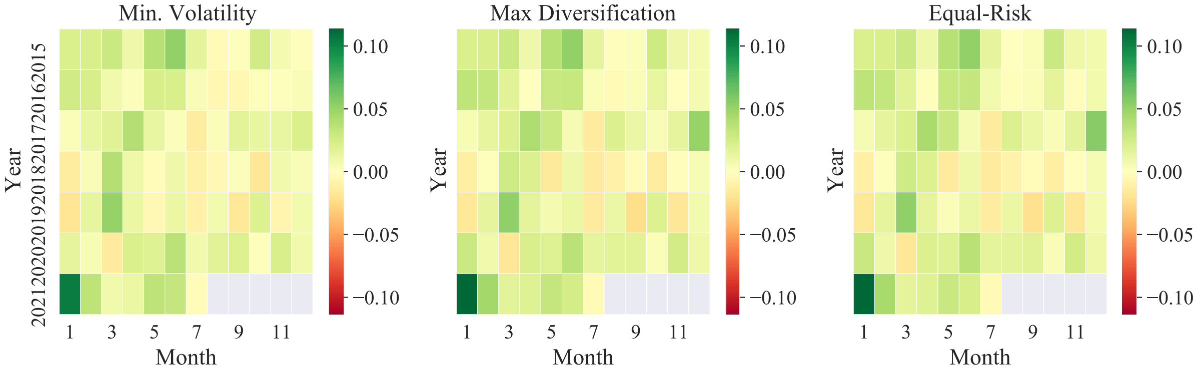


Figure 8: Heat map of monthly returns of the minimum volatility allocation, maximum diversification and equal-risk models from the beginning of 2015 to mid-2021. The color scale indicates the returns, in decimals, attributed to each corresponding color.

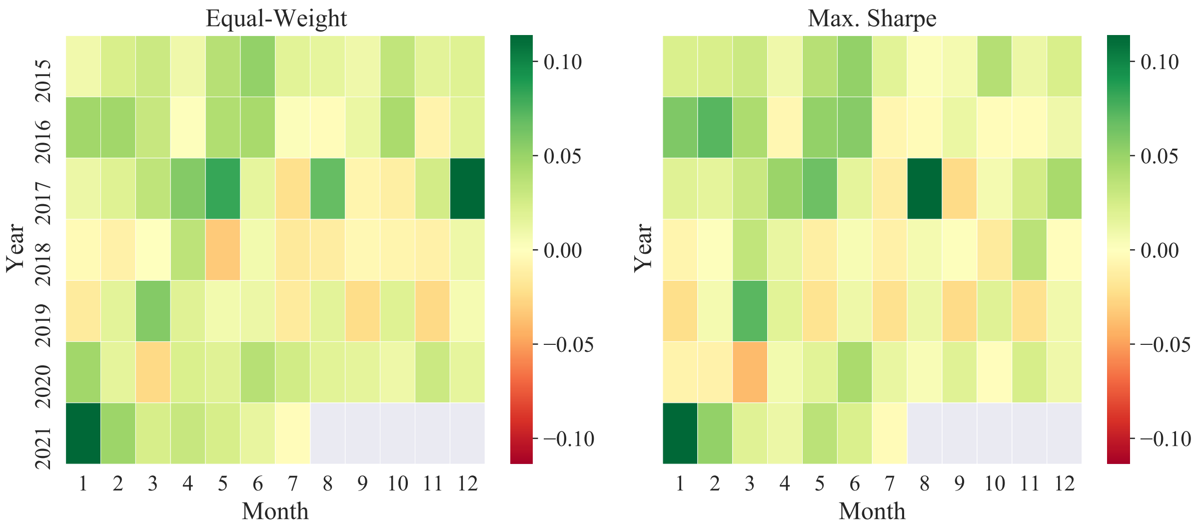


Figure 9: Heat map of monthly returns of the equal weight and maximum Sharpe allocation models from the beginning of 2015 to mid-2021. The color scale indicates the returns, in decimals, attributed to each corresponding color.

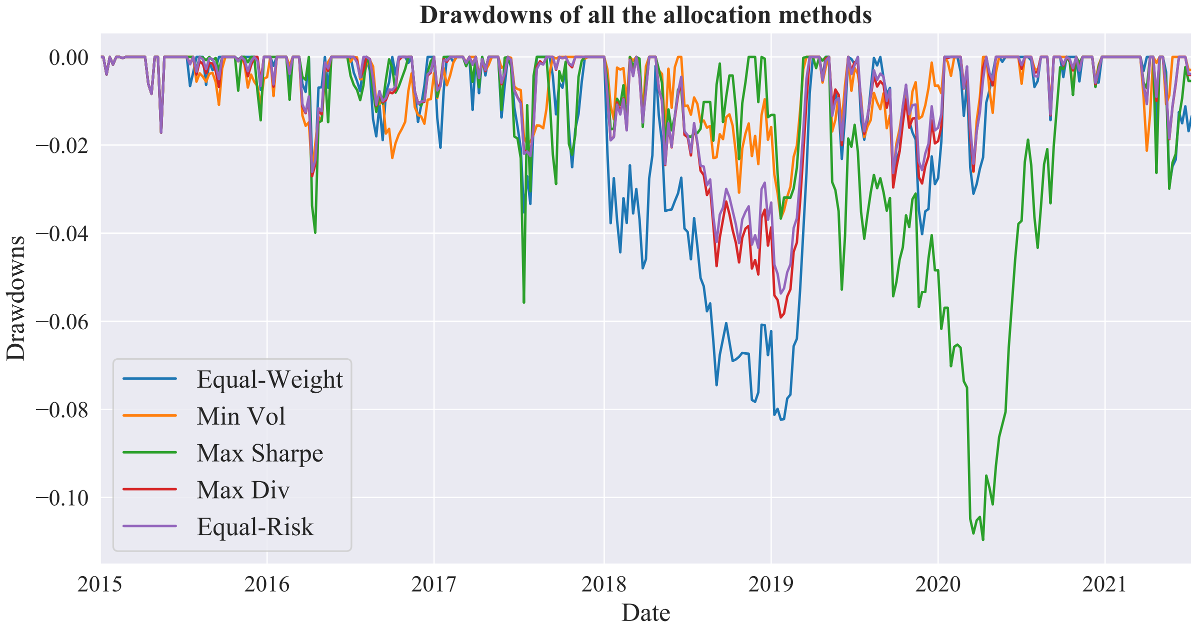


Figure 10: Cumulative drawdown of the returns for all the allocation methods from the beginning of 2015 to mid-2021

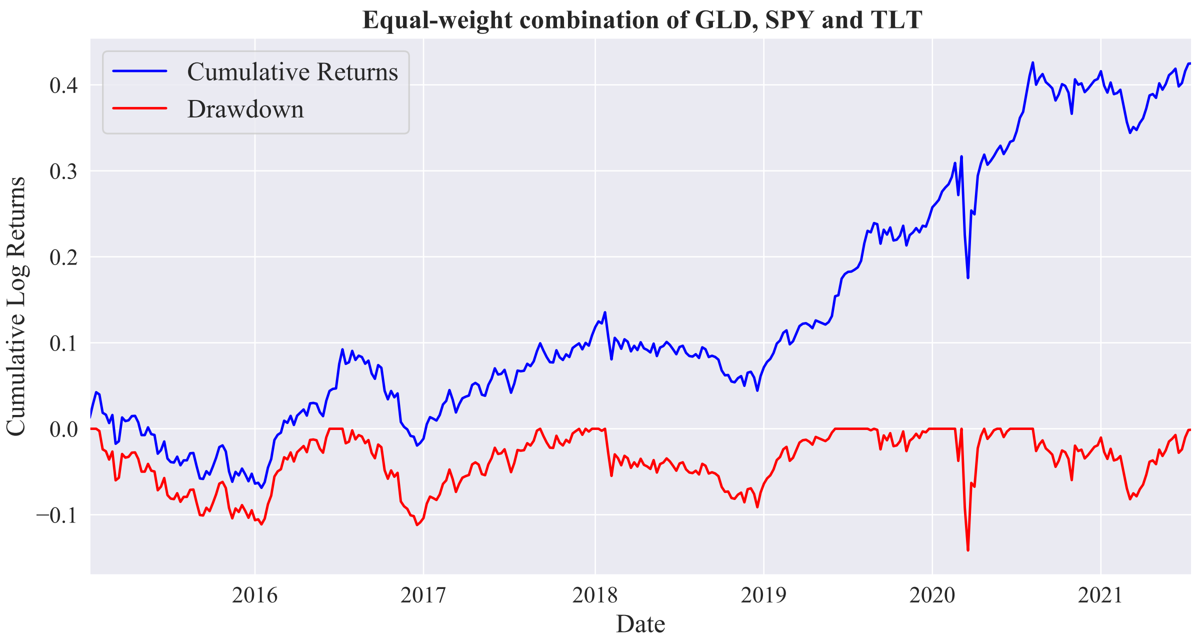


Figure 11: Cumulative returns and drawdown of the of benchmark portfolio with equal weights on the GLD, SPY and TLT ETFs. From the beginning of 2015 to mid-2021.

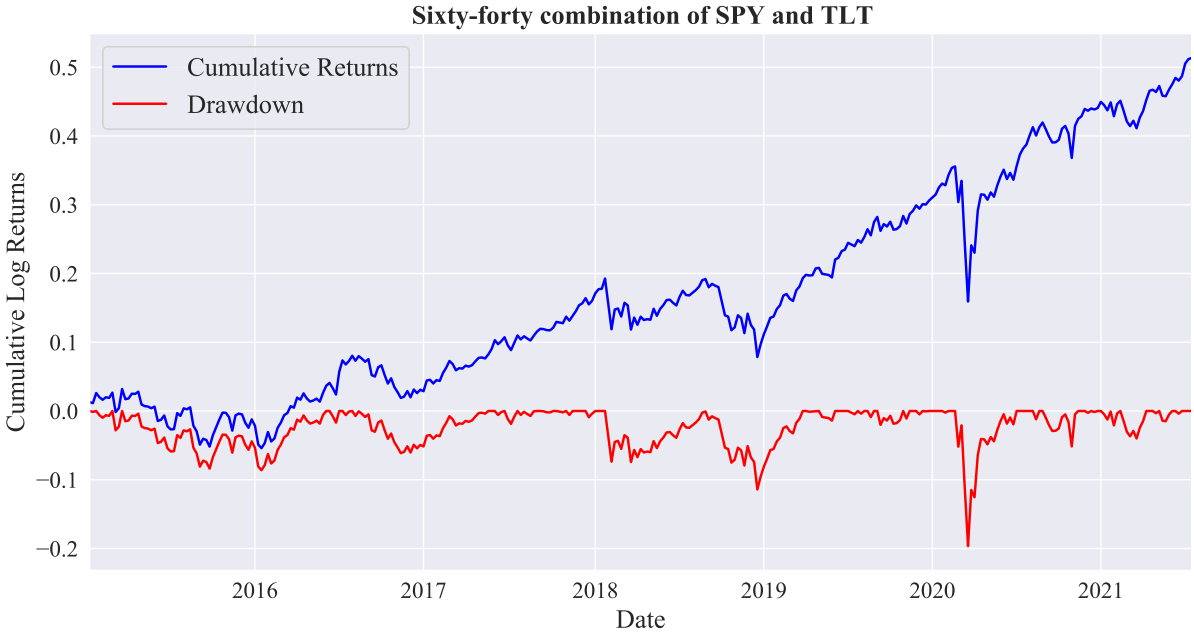


Figure 12: Cumulative returns and drawdown of the of benchmark portfolio with a sixty-forty allocation on the SPY and TLT ETFs, respectively. From the beginning of 2015 to mid-2021.

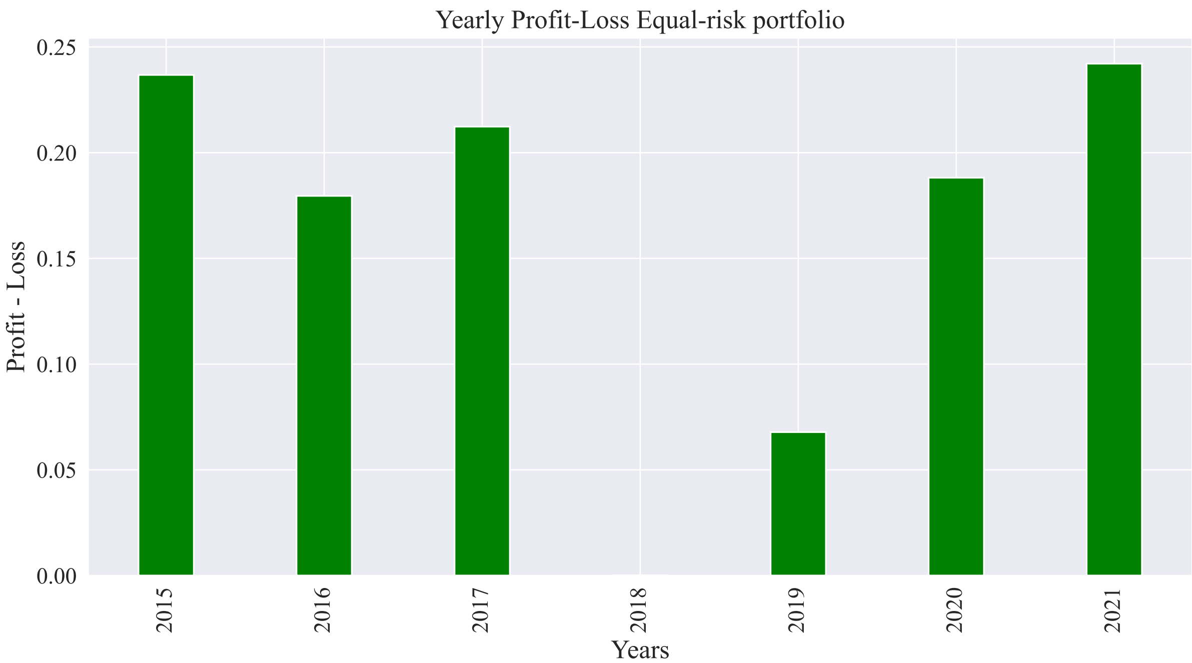


Figure 13: Equal-risk portfolio profit split annually, from 2015 to 2021.

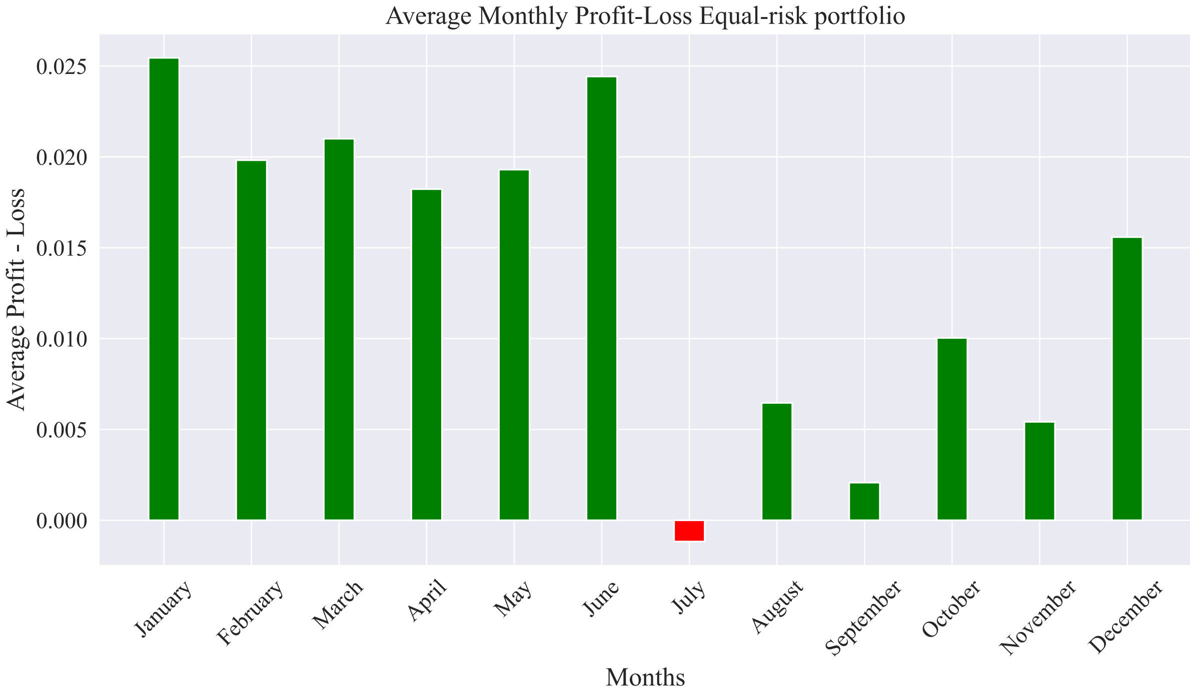


Figure 14: Monthly profit average for the Equal-risk portfolio.

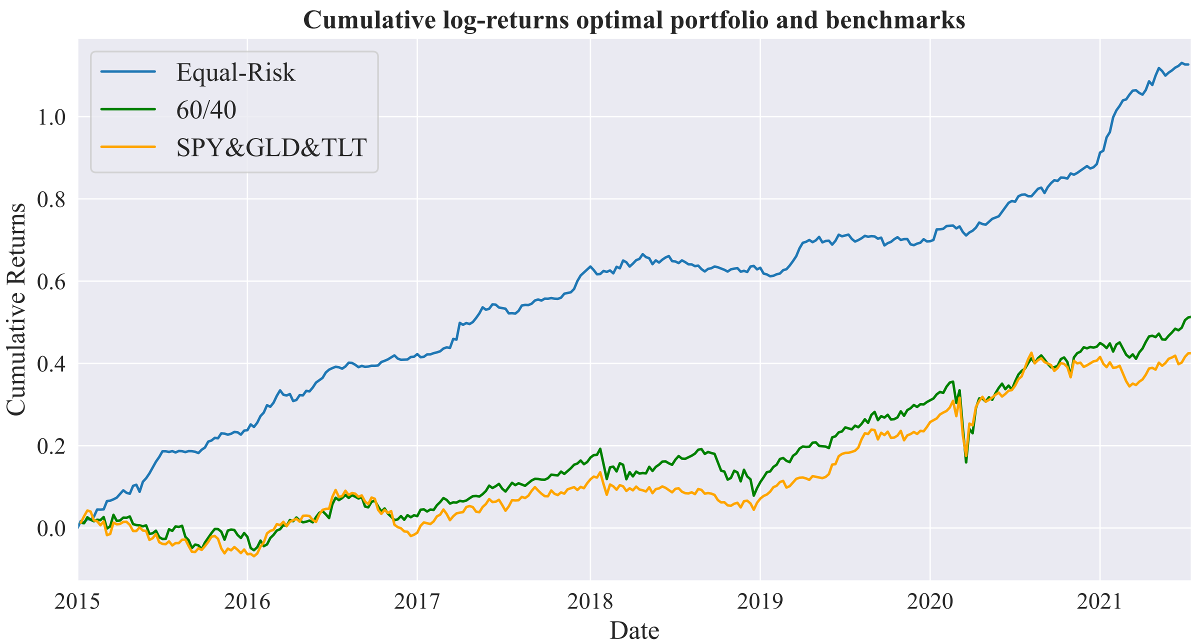


Figure 15: Cumulative returns of the two benchmark portfolios and the Equal-Risk method from the beginning of 2015 to mid-2021

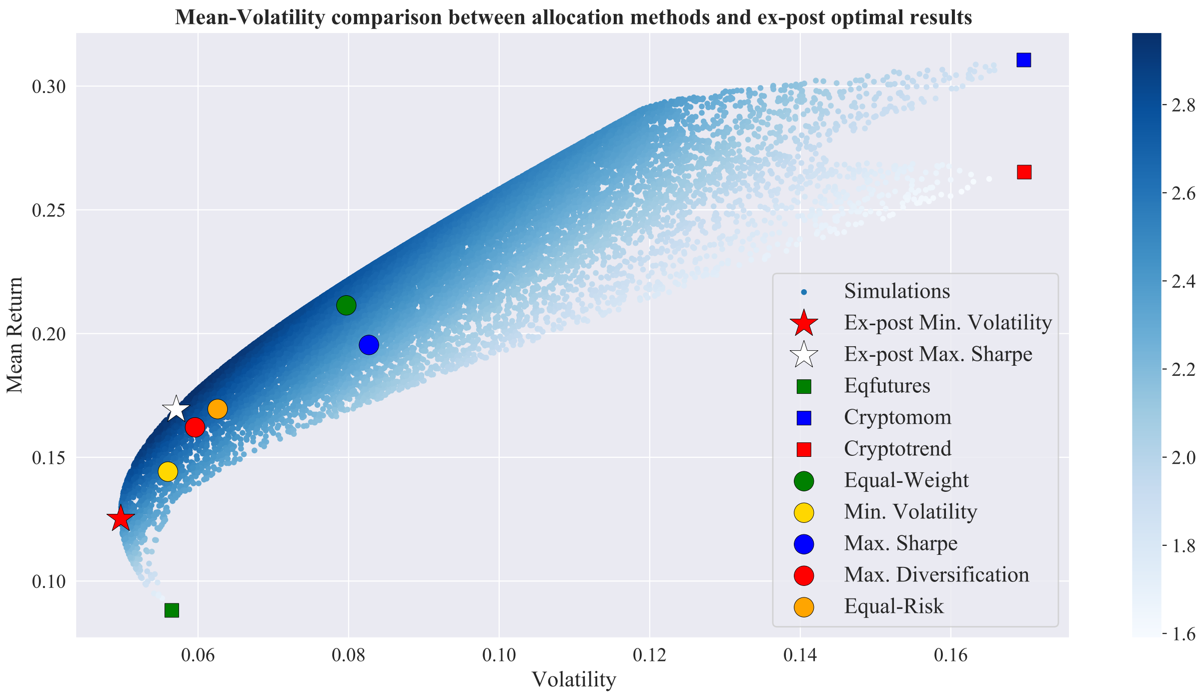


Figure 16: Mean-Volatility environment based on 20 thousand portfolio simulations. Indicating the three initial strategies, the theoretical maximum Sharpe and the minimum volatility portfolios, and the results obtained using historic estimates for each allocation model. The color scale indicates the Sharpe Ratio of each portfolio to the corresponding color.

.1.3 Equations

$$R = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} \quad w = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} \quad V = \begin{bmatrix} \mathrm{var}(1) & \mathrm{cov}(1,2) & \mathrm{cov}(1,3) \\ \mathrm{cov}(1,2) & \mathrm{var}(2) & \mathrm{cov}(2,3) \\ \mathrm{cov}(1,3) & \mathrm{cov}(2,3) & \mathrm{var}(3) \end{bmatrix} \quad \Sigma = \begin{bmatrix} \sigma_{x_1} \\ \sigma_{x_2} \\ \sigma_{x_3} \end{bmatrix}$$

Equation 1: R represents the array of the returns of each asset. w represents the array of the attributed weights of each asset. V is the variance-covariance matrix of each asset. Σ represents the array that contains the variance of each asset. The estimates for each parameter are based on historical samples.

$$\sigma_p^2 = w^T V w$$

Equation 2: σ_p^2 represents the variance of the the portfolio. Where w^T is the transposed array of the weights, and V is the variance-covariance matrix.

$$w^{min.vol} \quad = \quad arg\ min\ (w^T V w)$$

$$\text{subject to } \sum w_i = 1\ and\ 1 \geq w_i \geq 0$$

Equation 3: The minimization of the portfolio variance equation, subject to these constraints, provides the weights necessary to produce the minimum volatility portfolio.

$$sharpe\, ratio = \left(\frac{R_p - R_f}{\sigma_p}\right)$$

Equation 4: The Sharpe ratio provides a metric for returns per unit of risk, where R_p and σ_p are the portfolio's annualized return and standard-deviation. The R_f is considered zero.

$$w^{max.SR} \quad = \quad arg\ max\ \left(\frac{R_p - R_f}{\sigma_p}\right)$$

$$\text{subject to } \sum w_i = 1\ and\ 1 \geq w_i \geq 0$$

Equation 5: The minimization of the Sharpe ratio equation, subject to these constraints, provides the weights necessary to produce the maximum Sharpe portfolio.

$$diversification\ ratio = \frac{w^T \Sigma}{\sqrt{w^T V w}}$$

Equation 6: The diversification ratio provides the weights that produce the most diversified portfolio. V is the variance-covariance matrix, Σ represents the array of variances and W^T is the transposed array of weights.

$$w^{max.DR} \quad = \quad arg\ max\ \left(\frac{w^T \Sigma}{\sqrt{w^T V w}}\right)$$

$$\text{subject to } \sum w_i = 1\ and\ 1 \geq w_i \geq 0$$

Equation 7: The minimization of the Diversification ratio equation, subject to these constraints, provides the weights necessary to produce the maximum diversification portfolio.

$$risk\, contribution = \frac{w_i (V w)_i}{\sigma_p}$$

Equation 8: RC is the risk contribution of asset i to the total risk of the portfolio. V is the variance-covariance matrix, w is the weight of asset i and σ_p is the portfolio standard-deviation.

$$w* \quad = \quad arg\ min \sum_{i=1}^{n} \sum_{j=1}^{n} \left(w_i (Vw)_i - w_j (Vw)_j \right)^2$$

$$\text{subject to } \sum w_i = 1 \text{ } and \text{ } 1 \geq w_i \geq 0$$

Equation 9: The minimization of function f, subject to these constraints, provides the weights necessary to produce the equal-risk portfolio. Function f is the squared error between the observed risk contribution of asset and the target risk contribution. The target weight, in this case, it is a third of the total risk for each asset.

References

Choueifaty, Yves, and Yves Coignard. 2008. "Toward maximum diversification." The Journal of Portfolio Management, 35(1): 40–51.

DeMiguel, Victor, Lorenzo Garlappi, and Raman Uppal. 2009. "Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy?" The review of Financial studies, 22(5): 1915–1953.

Maillard, Sébastien, Thierry Roncalli, and Jérôme Teïletche. 2010. "The properties of equally weighted risk contribution portfolios." The Journal of Portfolio Management, 36(4): 60–70.

Markowitz, Harry. 1952. "Portfolio Selection." The Journal of Finance, 7(1): 77–91.