Enhanced Index Replication Based on Smart Beta and the Analysis of Distribution Moments

QFRG and DSLab Monthly Meetings

Kamil Korzeń and Robert Ślepaczuk

Quantitative Finance Research Group
Department of Quantitative Finance
Faculty of Economic Sciences,
University of Warsaw

21-06-2021

Agenda

Motivation

- the main goal of this research was to provide a more innovative approach to tracking by:
  - creating a robust replication of the SPXTR Index with a limited number of components, where the tracking error (TE) should be comparable with the most popular ETFs,
  - successfully applying the enhancement (Smart Beta) methods, in order to decrease the level of negative risk associated with the returns of such replication.
- such positive excess return, we aimed to use as the underperformance insurance and the source of dividend for investors
- in the end, we tested the algorithm on out-of-sample data for Q1 2021, to ensure the robustness of our strategy

Main Hypothesis

Hypothesis 1:

With enhancement methods applied, it is possible to create a synthetic index of N components[a] with absolute Tracking Error similar to benchmark ETFs[b]. Such Index will have a return distribution significantly[c] skewed to the right in comparison to the SPXTR Index, therefore positive deviations will be more frequent than negative ones.

---

[a]The number N is smaller than the total number of S&P 500 Index constituents in every holding period.

[b]The TE limit is defined as $[0.5 * TE_L; 2.0 * TE_H]$ where TE_L is the lowest, TE_H is the highest TE among benchmark ETFs.

[c]The statistical significance is verified with a two sample $t - test$ for difference of means with 95% confidence interval.

Research Questions

Research Question 1:

Without the enhancement methods (Kurtosis, Skewness or excess return cushion (ERC), it is possible to create a replicated index using N components that will have a similar return distribution as the SPXTR Index, with statistical significance, applying the statistical tests mentioned in H1.

Research Question 2:

Without the enhancement methods (Kurtosis, Skewness or ERC), it is possible to create a replicated index using N components that will have a similar Tracking Error as benchmark ETFs, within the assumed limits mentioned in H1.

Research Questions cont'd

Research Question 3:

With the enhancement applied (Kurtosis or Skewness), it is possible to create a replicated index using N components that without ERC will have a different return distribution in comparison to the SPXTR Index, with statistical significance, applying the statistical tests mentioned in H1.

Research Question 4:

With the enhancement applied (Kurtosis or Skewness), it is possible to create a replicated index using N components that without ERC will have a higher Positive Tracking Error than benchmark ETFs in absolute terms.

Research Question 5:

While the above (RQ4) is true, the replicated index with ERC will have the Absolute TE within the allowed limits.

Software, Libraries and Hardware

- `macOS 11.0.1`
- `R version 4.0.5`
- data sourcing: `tidyquant 1.0.3` and `getSymbols()` function
- performance metrics: `PerformanceAnalytics 2.0.4`
- additional packages: `ppcor 1.1.0`, `nnls 1.4.0`, `dtw 1.2.3` and `tidyverse 1.3.1`
- 1.6 GHz Dual-Core Intel Core i5; 8 GB 1600 MHz DDR3; Intel HD Graphics 6000 1536 MB

Literature Review

- there are valid empirical results, that physical tracking is better than synthetic in terms of Tracking Error; (Naumenko and Chystiakova, 2015) and (Fassas, 2014)
- most researchers agree, that Smart Beta is not sustainable in the long term, compared to their simple counterparts; (Malkiel, 2014), (Glushkov, 2015) and (Mateus, et al. 2020)
- more sophisticated approaches to tracking and Smart Beta applications based on complex analytical models provide much better returns to investors, than traditional ones; (Fons et al. 2021), (Nadler and Schmidt 2016), (Coelho 2012), (Nicolae et al. 2016) and (Baek et al. 2020)

General Assumptions

Methodology. General Assumptions

- we choose the number of stocks equal to 25 as our baseline approach and 126 trading days as the rebalancing window (for the whole in-sample period)
- for transaction costs we apply a linear approach and set the unit TC to 1%. Therefore changing 1% of a portfolio costs us 0.01% of total assets, regardless of buying or selling
- the process of portfolio allocation during rebalancing is as follows:
  - we calculate the average trading volume in dollar terms. Its rolling (moving) average window is based on the length of the rebalancing period. For the ranking process, we select the moving average value on the rebalancing day and we create a ranking from 1ˆst to 25ˆth stock in terms of average amount traded (the proxy for liquidity). These stocks are given the *available to trade* flag for the selected rebalancing period, while the rest are restricted
  - we proceed to the process of setting the weights, which is significantly different for each of the methods described in the next section

Methodology. General Assumptions cont'd

N *is the number of observations, where each Day is the separate observation*

T *is the number of trading days within a trading year, which is set to* 252

RFR_t *is daily Risk Free Rate, which is set to* 1.5% *per annum*

P_t *is the Daily Adjusted Close Price for day* t

R_t *is the Daily Return for day* t *expressed as* $P_t/P_{t-1} - 1$

R_a, R_b *are vectors of Daily Returns for the investment and benchmark, respectively*

ART absolute rate of return

CRT compounded (annualised) rate of return

ASD annualised standard deviation

ASR annualised Sharpe ratio

MDD maximum drawdown

AIR annualised information ratio

Methodology. General Assumptions cont'd

$$ART = prod(1 + R_a)^{\frac{T}{N}} - 1 \tag{1}$$

$$CRT = prod(1 + R_a) - 1 \tag{2}$$

$$ASD = \sqrt{Var(R_a)} * \sqrt{T} \tag{3}$$

$$ASR = \frac{prod(1 + R_t - RFR_t)^{\frac{T}{N}} - 1}{\sqrt{Var(R_a)} * \sqrt{T}} \tag{4}$$

$$MDD = sup_{x,y \in \{[t_1,\ t_2] : x \leq y\}} \frac{P_x - P_y}{P_x} \tag{5}$$

$$AIR = \frac{ART}{ASD} \tag{6}$$

Methodology. General Assumptions cont'd

$$aTE = \sqrt{\sum \frac{(R_a - R_b)^2}{N\sqrt{T}}} \tag{7}$$

$$pTE = \sqrt{\sum \frac{(R_a - R_b)^2}{N\sqrt{T}}}\ , \textit{ only for observations where } R_a \ > \ R_b \tag{8}$$

$$nTE = \sqrt{\sum \frac{(R_a - R_b)^2}{N\sqrt{T}}}\ , \textit{ only for observations where } R_a \ \leq \ R_b \tag{9}$$

Simple Index Replication

Methodology. Simple Index Replication. Partial Correlation

- we calculate the PCOR Matrix for the whole sample of assets, as the Partial Correlation takes into account not only the correlation between two vectors, but how other variables correlate with them as well
- for each rebalancing period, we calculate the Partial Correlation Matrix, then we replace negative correlation coefficients with zero and standardize the weights, which are based on correlation coefficients. Therefore, we obtain the standardized weights, that sum up to 100%
- in other words, if the asset correlates with the SPXTR Index more, it will receive a higher weight.
- we decided to perform an additional test with Long-Short Portfolio, to test the cost and the impact of skipping the negatively correlated assets in Long-Only Portfolio, where the core of the methodology remains the same

Methodology. Simple Index Replication. Non-Negative Least Squares

- it returns only non-negative coefficients, which allows us to consider Long-Only Portfolio without sacrificing negative coefficients from the output
- the NNLS seems to be an appropriate method for the task that we aim to accomplish due to two reasons:
  - this is a well-known and reliable econometric method, similar to the Ordinary Least Squares with a high degree of explainability in comparison to state-of-the-art machine learning and deep learning methods
  - it sets the coefficient proportionally to the explained variance in the dependent variable
- on the contrary, if the number of stocks available to invest is too low, the method may lose on its explainability power, due to the artificial restriction
- we start by calculating the coefficient vector, then select weights for only the most liquid assets and standardize the weights, such that the sum of weights is equal to 100%

Methodology. Simple Index Replication. Beta Coefficient

- we calculate the Beta Coefficient between each of the constituents and the SPXTR Index
- we invert the coefficient value (divide 1 by the coefficient) to reward assets with Beta coefficient close to one or below, and penalize those which have the Beta coefficient significantly higher than one
- that said, we avoid investing in assets that are much more volatile than the SPXTR Index

Methodology. Simple Index Replication. Dynamic Time Warping

- in this method, the weight output is always higher than zero (we always maintain 25 assets in the portfolio)
- we invert the distance from the algorithm output (divide 1 by the distance), to overweight and reward assets with smaller distance in comparison to those which are more distant to the SPXTR Index
- this method is the most promising, as it takes into account the sequential character of the time series we analyze and – as shown in the section dedicated to the work related to the topic – is especially successful in time series pattern recognition, therefore the effect of it will be interesting to examine in comparison to regular financial and econometric approaches

Enhanced Index Replication

Methodology. Enhanced Index Replication

- by applying the enhancement, we aim to remain as close to the SPX as before, but in absolute terms.
- While the Absolute Tracking Error (aTE) should remain fairly constant, we want to have more positive excess return than the negative.
- In that sense after enhancement, we will be more likely to see a higher Positive Tracking Error (pTE), and lower Negative Tracking Error (nTE).

Methodology. Enhanced Index Replication. Excess Return Cushion

- we would like to keep aside the pTE in a form of accumulated positive excess return, to compensate for future negative excess return (underperformance)
- we are implementing an *excess return cushion* (ERC), that will be more substantial, the higher the accumulated positive excess return (and pTE)
- after each day, when our replicated index is outperforming the SPXTR Index, we match the gains and set the positive excess daily return aside. When our replicated index is underperforming, we use the accumulated return to compensate for the loss. If the *cushion* is not enough, we use it all
- during the rebalancing periods, we test two solutions.
  - we just rebalance the portfolio and pay the TC
  - after rebalancing and paying the TC, we pay out the remaining *cushion* in a form of dividend and reduce the ERC to zero
- to sum up, when we apply the *excess return cushion*, we do not let the synthetic index outperform the SPXTR Index. All alpha is set aside for the future loss compensation and/or dividend payout during rebalancing days.

Methodology. Enhanced Index Replication. Skewness

- we calculate the difference between the asset's skewness and the SPXTR Index's skewness
- we replace negative results with 0 and add 1 to this value, then we multiply the original weights by resulting coefficient
- this method should reward more positively skewed assets, therefore increase the probability of outperforming returns

Methodology. Enhanced Index Replication. Kurtosis

- we create the ratio between the asset's kurtosis and the kurtosis of the SPXTR Index. We use that ratio as the multiplier for original weight
- the assumption is that if the asset has a more centralized distribution of returns (higher kurtosis than the SPXTR Index), it will positively influence the tracking mechanism, as we expect less long-tail returns (either positive or negative)

Data Description

- we aim to track S&P 500 Total Return (SPXTR Index), available on Yahoo Finance under the ^SP500TR ticker
- we source Adjusted Close data for each of the 500 constituents from 2016-01-04 to 2020-12-31
- we additionally source SPY, VOO and IVV Adjusted Close price data, to compare our approach with popular SP 500 ETFs
- we add an additional filter in the form of NA imputation, for the periods where certain equity was listed, but were not considered a part of the S&P 500 Index
- we use trading volumes to build the ranking used to filter Top 25 most liquid constituents during each rebalancing period
- for the whole 5 year period we have 1259 observations (trading days) and 504 securities
- we treat the period from 2016-01-04 to 2020-12-31 as a training sample, and the first quarter of the year 2021 (2021-01-04 - 2021-04-01) as a testing sample

Simple Index Replication

Empirical Results. Simple Index Replication. Partial Correlation

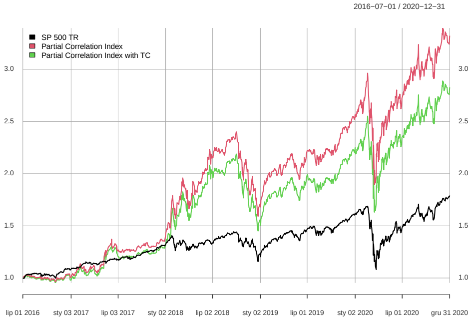
Table 1. Performance Metrics for PCOR Index

Name	PCOR	PCOR_TC	SPX	SPY	VOO	IVV
ART	0.3058	0.2598	0.1342	0.1336	0.1337	0.1323
CRT	2.3221	1.8270	0.8759	0.8708	0.8717	0.8602
ASD	0.2813	0.2825	0.1932	0.1895	0.1930	0.1933
ASR	1.0180	0.8533	0.6070	0.6157	0.6051	0.5973
MDD	0.3608	0.3608	0.3597	0.3575	0.3620	0.3612
AIR	1.0870	0.9197	0.6944	0.7047	0.6925	0.6845
aTE	0.1670	0.1686	0.0000	0.0116	0.0088	0.0097
pTE	0.1348	0.1340	0.0000	0.0108	0.0073	0.0076
nTE	0.1117	0.1151	0.0000	0.0091	0.0069	0.0081

Note: The above table presents the performance metrics for replicated index based on Partial Correlation (Long-Only Portfolio) with and without Transaction Costs included in comparison to the SPXTR Index and three benchmark ETFs. Portfolio allocation is based on the coefficient value, where it rewards higher coefficients. Weights are scaled to sum up to 100

Empirical Results. Simple Index Replication. Partial Correlation cont'd

Figure 1. Cumulative Return of the SPXTR Index and PCOR Index with and without TC



Note: Figure presents the cumulative return for the Index without Transaction Costs, the Index with Transaction Costs and the SPXTR Index. Equity line is calculated by taking the cumulative product of daily returns.

Empirical Results. Simple Index Replication. Partial Correlation L/S

Table 2. Performance Metrics for PCLS Index

Name	PCLS	PCLS_TC	SPX	SPY	VOO	IVV
ART	0.1992	0.1519	0.1342	0.1336	0.1337	0.1323
CRT	1.2649	0.8892	0.8759	0.8708	0.8717	0.8602
ASD	0.2286	0.2316	0.1932	0.1895	0.1930	0.1933
ASR	0.7935	0.5817	0.6070	0.6157	0.6051	0.5973
MDD	0.3916	0.3916	0.3597	0.3575	0.3620	0.3612
AIR	0.8716	0.6557	0.6944	0.7047	0.6925	0.6845
aTE	0.0925	0.0992	0.0000	0.0116	0.0088	0.0097
pTE	0.0709	0.0710	0.0000	0.0108	0.0073	0.0076
nTE	0.0591	0.0740	0.0000	0.0091	0.0069	0.0081

Note: The above table presents the performance metrics for Index based on Partial Correlation (Long-Short Porftolio) with and without Transaction Costs included in comparison to the SPXTR Index and three benchmark ETFs. Portfolio allocation is based on the coefficient value, where it rewards higher coefficients. Weights are scaled to sum up to 100

Empirical Results. Simple Index Replication. Partial Correlation L/S cont'd

Figure 2. Cumulative Return of the SPXTR Index and PCLS Index with and without TC

Note: Figure presents the cumulative return for the Index without Transaction Costs, the Index with Transaction Costs and the SPXTR Index. Equity line is calculated by taking the cumulative product of daily returns.

Empirical Results. Simple Index Replication. Non-Negative Least Squares

Table 3. Performance Metrics for NNLS Index

Name	NNLS	NNLS_TC	SPX	SPY	VOO	IVV
ART	0.2054	0.1761	0.1342	0.1336	0.1337	0.1323
CRT	1.3181	1.0747	0.8759	0.8708	0.8717	0.8602
ASD	0.2235	0.2245	0.1932	0.1895	0.1930	0.1933
ASR	0.8388	0.7063	0.6070	0.6157	0.6051	0.5973
MDD	0.3382	0.3382	0.3597	0.3575	0.3620	0.3612
AIR	0.9190	0.7842	0.6944	0.7047	0.6925	0.6845
aTE	0.0728	0.0752	0.0000	0.0116	0.0088	0.0097
pTE	0.0539	0.0538	0.0000	0.0108	0.0073	0.0076
nTE	0.0481	0.0533	0.0000	0.0091	0.0069	0.0081

Note: The above table presents the performance metrics for Index based on Non-Negative Least Squares with and without Transaction Costs included in comparison to the SPXTR Index and three benchmark ETFs. Portfolio allocation is based on the coefficient value, where it rewards higher coefficients. Weights are scaled to sum up to 100

Empirical Results. Simple Index Replication. Non-Negative Least Squares cont'd

Figure 3. Cumulative Return of the SPXTR Index and NNLS Index with and without TC

Note: Figure presents the cumulative return for the Index without Transaction Costs, the Index with Transaction Costs and the SPXTR Index. Equity line is calculated by taking the cumulative product of daily returns.

Empirical Results. Simple Index Replication. Beta Coefficient

Table 4. Performance Metrics for BETA Index

Name	BETA	BETA_TC	SPX	SPY	VOO	IVV
ART	0.1049	0.0916	0.1342	0.1336	0.1337	0.1323
CRT	0.5663	0.4832	0.8759	0.8708	0.8717	0.8602
ASD	0.2175	0.2177	0.1932	0.1895	0.1930	0.1933
ASR	0.4064	0.3459	0.6070	0.6157	0.6051	0.5973
MDD	0.3829	0.3829	0.3597	0.3575	0.3620	0.3612
AIR	0.4820	0.4205	0.6944	0.7047	0.6925	0.6845
aTE	0.0486	0.0493	0.0000	0.0116	0.0088	0.0097
pTE	0.0344	0.0343	0.0000	0.0108	0.0073	0.0076
nTE	0.0323	0.0337	0.0000	0.0091	0.0069	0.0081

Note: The above table presents the performance metrics for Index based on Beta Coefficient with and without Transaction Costs included in comparison to the SPXTR Index and three benchmark ETFs. Portfolio allocation is based on the coefficient value, where it rewards coefficient close to the value of 1, by setting the weight for each asset as the inverted coefficient value. Weights are scaled to sum up to 100

Empirical Results. Simple Index Replication. Beta Coefficient cont'd

Figure 4. Cumulative Return of the SPXTR Index and BETA Index with and without TC

Note: Figure presents the cumulative return for the Index without Transaction Costs, the Index with Transaction Costs and the SPXTR Index. Equity line is calculated by taking the cumulative product of daily returns.

Empirical Results. Simple Index Replication. Dynamic Time Warping

Table 5. Performance Metrics for DTWA Index

Name	DTWA	DTWA_TC	SPX	SPY	VOO	IVV
ART	0.1314	0.1205	0.1342	0.1336	0.1337	0.1323
CRT	0.7428	0.6688	0.8759	0.8708	0.8717	0.8602
ASD	0.2196	0.2198	0.1932	0.1895	0.1930	0.1933
ASR	0.5216	0.4725	0.6070	0.6157	0.6051	0.5973
MDD	0.3715	0.3715	0.3597	0.3575	0.3620	0.3612
AIR	0.5983	0.5484	0.6944	0.7047	0.6925	0.6845
aTE	0.0472	0.0478	0.0000	0.0116	0.0088	0.0097
pTE	0.0343	0.0342	0.0000	0.0108	0.0073	0.0076
nTE	0.0313	0.0326	0.0000	0.0091	0.0069	0.0081

Note: The above table presents the performance metrics for Index based on Dynamic Time Warping with and without Transaction Costs included in comparison to the SPXTR Index and three benchmark ETFs. Portfolio allocation is based on the distance between the SPXTR Index and particular asset, where it rewards the assets with smaller distance, by setting the weight for each asset as the inverted distance. Weights are scaled to sum up to 100

Empirical Results. Simple Index Replication. Dynamic Time Warping cont'd

Figure 5. Cumulative Return of the SPXTR Index and DTWA Index with and without TC

Note: Figure presents the cumulative return for the Index without Transaction Costs, the Index with Transaction Costs and the SPXTR Index. Equity line is calculated by taking the cumulative product of daily returns.

Enhanced Index Replication

Empirical Results. Enhanced Index Replication. Simple Index w/ ERC

Table 6. Simple Index with ERC Performance for DTWA Index

Name	DTWA_Simple	DTWA_Simple.1	DTWA_Simple.2	SPXTR
ART	0.1205	0.1177	0.1171	0.1382
CRT	0.6688	0.6501	0.6461	0.7907
ASD	0.2198	0.1975	0.2001	0.1966
ASR	0.4725	0.5119	0.5023	0.6169
MDD	0.3715	0.3597	0.3702	0.3597
AIR	0.5484	0.5961	0.5854	0.7030
aTE	0.0478	0.0094	0.0278	0.0000
pTE	0.0342	0.0000	0.0254	0.0000
nTE	0.0326	0.0094	0.0134	0.0000

Note: The above table presents the performance metrics for Simple Index based on Dynamic Time Warping with Transaction Costs included with and without ERC applied for the option with and without dividend payment. The pTE is higher than zero for the options with dividend payment, as the payment is not excluded from the Tracking Error calculation, therefore on each rebalancing day, we observe an abnormal return, if the dividend is paid out. The Absolute Tracking Error (aTE) can be higher as well for the solutions with dividend payment, as we reduce our accumulated ERC to zero along with the event of dividend payout. Additionally, the SPXTR Index is presented for comparison.

Empirical Results. Enhanced Index Replication. Simple Index w/ ERC cont'd

Figure 6. Simple DTWA Index with ERC and the SPXTR Index

Note: Figure presents the cumulative return for the DTWA Index with Transaction Costs, DTWA Index with Transaction Costs and ERC and without Dividends, DTWA Index with Transaction Costs and ERC and with Dividends and the SPXTR Index. Equity line is calculated by taking the cumulative product of daily returns.

Empirical Results. Enhanced Index Replication. Skewness and Kurtosis w/out ERC

Table 7. Kurtosis and Skewness Enhanced Index without ERC Performance for DTWA Index

Name	DTWA_Simple	DTWA_Kurtosis	DTWA_Skewness	SPX
ART	0.1205	0.1439	0.1176	0.1382
CRT	0.6688	0.8309	0.6494	0.7907
ASD	0.2198	0.2127	0.2198	0.1966
ASR	0.4725	0.5963	0.4594	0.6169
MDD	0.3715	0.3526	0.3656	0.3597
AIR	0.5484	0.6764	0.5351	0.7030
aTE	0.0478	0.0533	0.0534	0.0000
pTE	0.0342	0.0368	0.0375	0.0000
nTE	0.0326	0.0367	0.0369	0.0000

Note: The above table presents the performance metrics for Simple Index based on Dynamic Time Warping with Transaction Costs included with Skewness and Kurtosis enhancements without ERC applied. Additionally, the SPXTR Index is presented for comparison.

Empirical Results. Enhanced Index Replication. Skewness and Kurtosis w/out ERC cont'd

Figure 7. Kurtosis and Skewness Enhanced Index without ERC and the SPXTR Index

Note: Figure presents the cumulative return for the DTWA Index with Transaction Costs, DTWA Index with Transaction Costs and Kurtosis Enhancement without ERC, DTWA Index with Transaction Costs Skewness Enhancement without ERC and the SPXTR Index. Equity line is calculated by taking the cumulative product of daily returns.

Empirical Results. Enhanced Index Replication. Skewness and Kurtosis w/ ERC

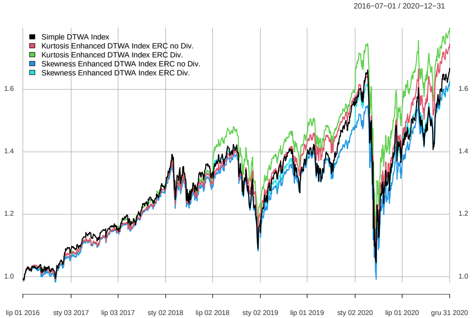
Table 8. Kurtosis and Skewness Enhanced Index with ERC Performance for DTWA Index

Name	DTWA_Simple	DTWA	DTWA_Div	DTWA.1	DTWA_Div.1	SPX
ART	0.1205	0.1318	0.1392	0.1141	0.1135	0.1382
CRT	0.6688	0.7459	0.7973	0.6259	0.6225	0.7907
ASD	0.2198	0.1967	0.1998	0.1976	0.1999	0.1966
ASR	0.4725	0.5847	0.6115	0.4934	0.4851	0.6169
MDD	0.3715	0.3597	0.3597	0.3597	0.3640	0.3597
AIR	0.5484	0.6703	0.6963	0.5773	0.5680	0.7030
aTE	0.0478	0.0050	0.0363	0.0106	0.0327	0.0000
pTE	0.0342	0.0000	0.0357	0.0000	0.0302	0.0000
nTE	0.0326	0.0050	0.0116	0.0107	0.0150	0.0000

Note: The above table presents the performance metrics for Index based on Dynamic Time Warping with Transaction Costs included and for the same method with Transactions Costs and two enhancement methods (Kurtosis and Skewness). Enhancement methods results are presented for the option with and without dividend payment. The pTE is higher than zero for the options with dividend payment, as the payment is not excluded from the Tracking Error calculation, therefore on each rebalancing day, we observe an abnormal return, if the dividend is paid out. The Absolute Tracking Error (aTE) can be higher as well for the solutions with dividend payment, as we reduce our accumulated ERC to zero along with the event of dividend payout. Additionally, the SPXTR Index is presented for comparison.

Empirical Results. Enhanced Index Replication. Skewness and Kurtosis w/ ERC cont'd

Figure 8. Kurtosis and Skewness Enhanced Index with ERC with and without Dividend and the Simple Index



Note: Figure presents the cumulative return for the DTWA Index with Transaction Costs (DTWA TC), DTWA TC with Kurtosis Enhancement with and without Dividends and DTWA TC with Skewness Enhancement with and without Dividends. Equity line is calculated by taking the cumulative product of daily returns.

Out-of-Sample Data

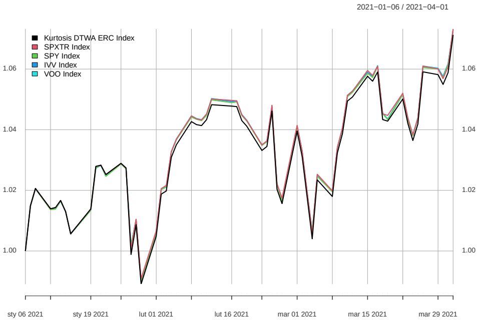
Table 9. Table with Final Test Results for DTWA Index

Name	Simple	Simple.ERC	Kurtosis	Kurtosis.ERC	SPXTR	SPY	IVV	VOO
ART	0.3937	0.3043	0.4371	0.3353	0.3455	0.3436	0.3442	0.3444
CRT	0.0822	0.0653	0.0902	0.0713	0.0732	0.0728	0.0730	0.0730
ASD	0.1932	0.1593	0.1999	0.1587	0.1574	0.1558	0.1554	0.1555
ASR	1.9306	1.7887	2.0801	1.9873	2.0675	2.0765	2.0866	2.0858
MDD	0.0508	0.0422	0.0584	0.0422	0.0422	0.0415	0.0417	0.0416
AIR	2.0378	1.9105	2.1870	2.1124	2.1946	2.2048	2.2152	2.2143
aTE	0.0591	0.0101	0.0759	0.0036	0.0000	0.0079	0.0079	0.0084
pTE	0.0340	0.0000	0.0494	0.0000	0.0000	0.0043	0.0044	0.0047
nTE	0.0346	0.0101	0.0471	0.0036	0.0000	0.0060	0.0062	0.0070

Note: The above table presents the performance metrics on out-of-sample testing period for Index based on Dynamic Time Warping (DTWA Index) with and without ERC, as well as for DTWA Index with Kurtosis Enhancement with and without ERC in comparison to the SPXTR Index and three benchmark ETFs.

Empirical Results. Out-of-Sample Data cont'd

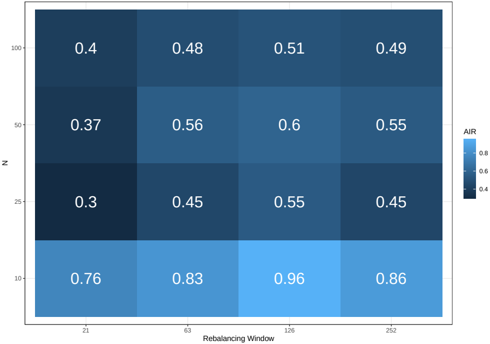
Figure 9. Test Results for DTWA Index with ERC and Kurtosis Enhancement and Benchmark ETFs



Note: Figure presents the cumulative return for the DTWA TC with Kurtosis Enhancement with ERC, the SPXTR Index and three benchmark ETFs. Equity line is calculated by taking the cumulative product of daily returns.

Sensitivity Analysis

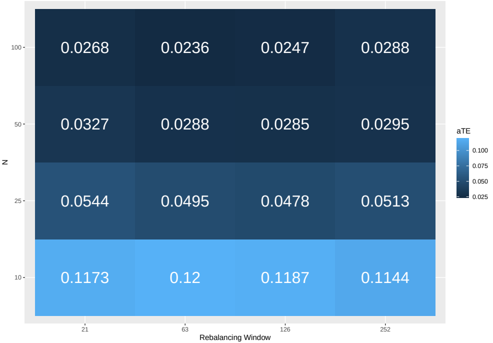
Figure 10. Sensitivity Analysis for Annualized Information Ratio



Note: Figure presents the results of sensitivity analysis on in-sample training data for 16 combinations of N (Number of Assets in Portfolio) and Reb (Length of Rebalancing Period, in trading days).

Sensitivity Analysis cont'd

Figure 11. Sensitivity Analysis for Annualized Absolute Tracking Error



Note: Figure presents the results of sensitivity analysis on in-sample training data for 16 combinations of N (Number of Assets in Portfolio) and Reb (Length of Rebalancing Period, in trading days).

Research Hypotheses Validation

- we confirmed that it is possible to create a replicated index with 25 components that will have a similar return distribution as the SPXTR Index (RQ1)
- with just Simple Index Replication it is not possible to achieve Tracking Error within the limits we assumed as comparable with benchmark ETFs (RQ2)
- with the Enhanced Index Replication, we failed to skew the daily return distribution enough to say that there is a difference in mean return between the SPXTR Index and our replicated index with enhancement methods applied (RQ3)
- we can confirm, that with the enhancement methods in place we can provide a higher Positive Tracking Error that benchmark ETFs (RQ4) and with *excess return cushion* implemented, the Absolute Tracking Error stays within the defined limit, and our Enhanced Index is comparable with benchmark ETFs (RQ5), but not for each applied method
- given all the empirical evidence, we had to reject the main hypothesis of the research, as it assumed - among others - that we would be able to provide return distribution significantly different from the SPXTR Index.

Simple Index Replication

Table 10. Simple Index Replication Research Questions

	Without TC		With TC	
Method	RQ1	RQ2	RQ1.TC	RQ2.TC
PCOR	0.3061	Not Met	0.4256	Not Met
PCLS	0.6445	Not Met	0.8642	Not Met
NNLS	0.6191	Not Met	0.7522	Not Met
BETA	0.8741	Not Met	0.8035	Not Met
DTWA	0.9819	Not Met	0.9609	Not Met

Note: For Research Question 1 (RQ1) we present the p-value of t.test for the difference of the means between two vector of returns. The null hypothesis states, that there is no difference. For Research Question 2 (RQ2) we present the binary outcome of the test, whether a certain method delivers Absolute Tracking Error within the predefined limits. If not, the result of the test is 'Not Met' and 'Met', if otherwise. The TE limit is defined as $[0.5 * TE_L; 2.0 * TE_H]$ where TE_L is the lowest and TE_H is the highest TE among benchmarks.

Enhanced Index Replication

Research Hypotheses Validation. Enhanced Index Replication

Table 11. Enhanced Index Replication Research Questions

	Without Dividend			With Dividend		
Enhancement	RQ3	RQ4	RQ5	RQ3.Div	RQ4.Div	RQ5.Div
Kurtosis	0.9246	Met	Met	0.9246	Met	Not Met
Skewness	0.9455	Met	Met	0.9455	Met	Not Met

Note: For Research Question 3 (RQ3) we present the p-value of t.test for the difference of the means between two vector of returns. The null hypothesis states, that there is no difference. For Research Question 3 (RQ3) and Research Question 4 (RQ4) we present the binary outcome of the test, whether a certain method is compliant with the statement provided in Research Questions. If not, the result of the test is 'Not Met' and 'Met', if otherwise.

Conclusions

- we focused on the tracking behaviour, but with the enhancement method applied, we delivered additional abnormal returns
- we found empirical evidence, that the Dynamic Time Warping is best suited to mimic the SPXTR Index during the analyzed period
- we found that the Kurtosis enhancement performs better and delivers better alpha.
- it is possible to create a replicated Index with only 25 assets in the portfolio, that is on par with the most popular public ETFs
- with ERC implemented, we can beat the benchmark ETFs in terms of performance

Further Research Extensions

This research certainly provides a potential for improvement and further extensions. Based on our learnings:

- there is still a possibility of providing more stable and more robust enhancement methods, than those presented in the research.

We believe that the use of machine learning algorithms or other sequential methods used for pattern recognition can provide a significant uplift on the positive excess return. Such methods, as Dynamic Time Warping have successfully improved the results of the tracking process in our research, therefore we strongly believe that applying them for capturing more alpha, should be a viable application.

References

Amenc, Noël, F Ducolumbier, Felix Goltz, and J Ulhael. 2016. "Ten Misconceptions About Smart Beta." EDHEC Position Paper.

Babić, Slajana, Christophe Ley, Lorenzo Ricci, and David Veredas. 2020. "TailCoR." arXiv Preprint.

Baek, Seungho, Kwan Yong Lee, Merih Uctum, and Seok Hee Oh. 2020. "Robo-Advisors: Machine Learning in Trend-Following Etf Investments." Sustainability 12 (16). Multidisciplinary Digital Publishing Institute: 6399.

Ben-David, Itzhak, Francesco Franzoni, Byungwook Kim, and Rabih Moussawi. 2021. "Competition for Attention in the Etf Space." National Bureau of Economic Research.

Callin, Sabrina, and Olivia Albrecht. 2014. "Making Your Smart Beta Even Smarter." The Journal of Index Investing 5 (1). Institutional Investor Journals Umbrella: 136–45.

Coelho, Mariana Sátiro. 2012. "Patterns in Financial Markets: Dynamic Time Warping." NSBE-UNL.

Dooijeweert, Peter van, and Graham Robertson. n.d. "Managing Beta Exposures in Portable Alpha Strategies."

Dreyer, Anna, and Stefan Hubrich. 2017. "Tail Risk Mitigation with Managed Volatility Strategies." Journal of Investment Strategies 8 (1): 29–56.

Fassas, Athanasios P. 2014. "Tracking Ability of Etfs: Physical Versus Synthetic Replication." The Journal of Index Investing 5 (2). Institutional Investor Journals Umbrella: 9–20.

Fons, Elizabeth, Paula Dawson, Jeffrey Yau, Xiao-jun Zeng, and John Keane. 2021. "A Novel Dynamic Asset Allocation System Using Feature Saliency Hidden Markov Models for Smart Beta Investing." Expert Systems with Applications 163. Elsevier: 113720.

Frino, Alex, and David R Gallagher. 2001. "Tracking SP 500 Index Funds." The Journal of Portfolio Management 28 (1). Institutional Investor Journals Umbrella: 44–55.

Galal, Soheil, Rafael Silveira, and Alison Rapaport. n.d. "Alternative Beta: Redefining Alpha and Beta."

References cont'd

Glushkov, Denys. 2015. "How Smart Are Smart Beta Etfs? Analysis of Relative Performance and Factor Exposure." Analysis of Relative Performance and Factor Exposure.

Hatzakis, Emmanuel D. 2017. "Redefining Indexing Using Smart Beta Strategies." Merrill Lynch Bank of America.

Kahn, Ronald N, and Michael Lemmon. 2016. "The Asset Manager's Dilemma: How Smart Beta Is Disrupting the Investment Management Industry." Financial Analysts Journal 72 (1). Taylor and Francis: 15–20.

Loh, Lixia, and Stoyan Stoyanov. 2014. "Tail Risk of Smart Beta Portfolios: An Extreme Value Theory Approach." EDHEC Risk Institute.

Lyxor, Asset Management. 2019. "Guide to Pragmatic Replication."

Malkiel, Burton G. 2014. "Is Smart Beta Really Smart?" The Journal of Portfolio Management 40 (5). Institutional Investor Journals Umbrella: 127–34.

Mateus, Cesario, Irina B Mateus, and Marco Soggiu. 2020. "Do Smart Beta Etfs Deliver Persistent Performance?" Journal of Asset Management 21 (5). Springer: 413–27.

Nadler, Daniel, and Anatoly B Schmidt. 2016. "Portfolio Theory in Terms of Partial Covariance." SSRN Print:2436478.

Naumenko, Klym, and Olena Chystiakova. 2015. "An Empirical Study on the Differences Between Synthetic and Physical Etfs." International Journal of Economics and Finance 7 (3): 24–35.

Nicolae, Maria-Irina, Éric Gaussier, Amaury Habrard, and Marc Sebban. 2016. "Similarity Learning for Time Series Classification." arXiv Preprint arXiv:1610.04783.

Rosenzweig, Jan. 2020. "Fat Tailed Factors." arXiv Preprint arXiv:2011.13637.

Ryś, Przemysław, and Robert ålepaczuk. 2019. "Machine Learning Methods in Algorithmic Trading Strategy

Thank You for Your Attention!

Kamil Korzeń, k.korzen@student.uw.edu.pl

Quantitative Finance Research Group

and

Robert Ślepaczuk, rslepaczuk@wne.uw.edu.pl

University of Warsaw, Faculty of Economic Sciences, Department of Quantitative Finance

Quantitative Finance Research Group