Journal of Mathematics and System Science 10 (2020)
doi: 10.17265/2159-5291/2020.002.005

ST-metric Estimation of Factor Exposures

Nadima El-Hassan[1], Anthony Hall[2], and Ilknur Tulunay[3]

1. Finance Discipline Group, University of Technology, Sydney, Australia

2. Retired

3. School of Mathematics and Statistics, University of NSW, Sydney 2052, Australia

Abstract: Non-parametric methods are treasured in data analysis, particularly infinance. ST -metric is a new concept, introduced by Tulunay (2017). It offers non-parametric methods and a new geometric view to data analysis. In that paper, ST-metric concept has been applied to performance measures of portfolios. In this current paper, we purpose another ST-metric method for finding factor exposures in thefiv e-style-factors model. Here the style factors are value, size, minimum volatility, quality and momentum. The main idea is to find the factor exposures (weights) of the five-factors-model by minimizing the ST-metric between benchmark returns and the constructed factor model returns. We compare ST-metric method with Tracking Error method (TE-method) which is used for factor analysis of major indexes, decomposed into the style factors (tradable via Exchange Traded Funds (ETFs)) by Ang et al. (2018). We show that ST-metric method gives better estimation of the factor exposures (weights) than tracking error method, in general, and further how ST-metric values vary with respect tofluctuations. This explains the reason behind the efficiency of the ST-metric method. We support this idea with empirical evidences.

Key words: ST metric, factor investing modeling, multifactor risk models, exchange traded fund, portfolio management, non-parametric, non-normality, optimisation.

1. Introduction

The idea introducing ST-metric is to analyze return time series by measuring the "distance" between such series. It computes the transformed (cumulative) returns of two assets, viewed as discrete plane curves. In fact, it is a metric of any two discrete plane curves, taking values on the unit interval (0, 1]. In this paper, we purpose a new method tofind factor exposures in the five-style-factors model by using ST-metric. We also show how fluctuations affect the ST-metric value. That provides an understanding of why ST-metric method gives better estimation of the factor exposures (weights). Empirical evidence supports these findings.

ST-metric, defined on discrete plane curves derives from entropy. Some of the advantages of the entropy over standard deviation has been revealed by Philippatos & Wilson (1972) [1]. Mainly, entropy is independent of the mean, contrast to standard deviation. It captures the variation in the data. Entropy is a more general measure than the variance, since it accounts for higher order moments of a probability distribution function [2, 3] expressed that entropy can capture the uncertainty and disorder in a time series without imposing any constraints on the theoretical probability distribution, which constitutes its major advantage.

Entropy is defined as negative expected value of the logarithm of the probabilities (with Boltzmann constant in statistical mechanics). Our point of view is Information Theoretical as in Cover & Thomas (2006) [4]. We note that relative entropy (Kullback-Leibler divergence, Kullback-Leibler Information criterion, KLIC) is viewed as a "distance", but does not satisfy metric conditions (e.g., the triangle inequality fails). The symmetrized relative entropy of the average of two probabilities is called Topsøe distance (not a metric either) [5]. However, it turns out that the square root of Topsøe distance is a metric on probabilities [6]. As well-known, probabilities are required to satisfy

Corresponding author: Ilknur Tulunay, Ph.D., research fields: quantitative Finance, algebra, topology, geometry, statistics. E-mail: ilknur.tulunay@gmail.com.

two conditions: the first condition is being between zero and one, and the second condition is that summing all values up has to be one. Removing the second condition, as generalization, the square root of Topsøe distance on values between zero and one also satisfy metric conditions, called ST-metric[1] [7] for proof). We include the definition in Eq. (1).

Fama & French (1993) [8] identify the five common risk factors in returns on stocks and bonds. Since then, the factor modelling with various number of factors are used in finance discipline. The key idea is that the returns in general can be approximated with a few risk factors. Determining those risk factors is one problem and deciding the optimization method for the coefficient is another problem. Academics' and practitioners' literature are rich with many studies related to these problems. For example, Ang (2014), Fama & French (2015), Clarke et al. (2016), Fama & French (2017), A. et al. (2017), Fama & French (2018) and etc. [9-14]. Particularly, factor models are popular among the fund managers of Exchange Traded Funds (ETFs).

For general information on ETF, we refer to Abner (2010) [15] and Madhavan (2016) [16]. Madhavan & Sobcyzk (2016) [17] emphasizes that Exchange traded funds (ETFs) have grown substantially in diversity, market significance and size in recent years. They found that pricing efficiency varies significantly across funds and is systematically related to cross-sectional measures of liquidity. Piccotti (2018) [18] reported that the liquidity benefits offered by foreign ETFs and fixed income ETFs are revealed to be the most valuable to investors. In our paper, the factor indexes are tracked by investible ETFs, so the exposures are available via actual tradeable portfolios, as stated in Ang et al. (2018) [19].

Ang et al. (2018) [19] focuses on the style risk factors as the systematic factors: value, size, quality, momentum and minimum volatility. They use the tracking error (that is the standard deviation times the absolute tracking differences) as objective function in a factor model to find the factor weights of major indexes. Tracking error smooths the absolute tracking differences and so produces better outcomes for the model optimization than the tracking differences. Our contribution to the subject is to suggest the ST-metric method, in place of tracking error used in Ang et al. (2018) [19]. We examine this idea with some empirical research where the data selection (as being consistent with the data set of Ang et al. (2018) [19] is from Exchange Traded funds (ETF) market in Australia and USA. We found that ST-metric method predominantly estimates better than the tracking error method.

This paper has been organized as follows. Section 2 starts with recalling ST-metric definition and properties. In order to understand the effects of fluctuations on ST-metric values, some sets related to cumulative uniform distribution in the unit interval have been constructed. By using these sets, we establish some ST-metric inequalities, corresponding to different fluctuations, derived from cumulative uniform distribution in the unit interval. We place the proofs in the Appendix A. In Section 3, we explain the methodology of ST-metric and Tracking Error methods. We provide empirical evidence in Section 4. Finally, Section 5 presents our conclusions.

2. ST-Metric

In this section, we give the definition of ST-metric, introduced in Tulunay (2017) [7] and prove some properties which shows how ST-metric accounts for fluctuations. This allows us to understand why ST-metric method offers better tracking the benchmark than tracking error method.

ST-metric can be viewed as a distance between two discrete curves on $(0, 1]\times(0, 1]$. Think of the x-axis as time. Let $I = \{1, 2, . . . , n\}$. Denote time by $t_i \in \mathbb{R}$ for all $i \in I$, where $t_1 < t_2 < \cdot\cdot\cdot < t_n$ are n ordered points on the x-axis (time goes forward). For each $i \in I$,

[1] "ST" presents "square root of Tops_e distance", as ST-metric is an extension of it.

we have the data point $p_i = p(t_i)$. Associating $t_i \to \frac{i}{n}$, we can view the n-data points set, $P = \{p_i \in (0,1] \mid i \in I\}$ as a discrete curve on the square (0,1]×(0,1]. Since there is a unique point for each time t_i, such a n-points set defines a nonsingular discrete curve. Denote by $\mathcal{X}_n$ the set of all such discrete curves with n points. For any $P, Q \in \mathcal{X}_n$, define a function $ST\colon \mathcal{X}_n \times \mathcal{X}_n \mapsto \mathbb{R}$ by

$$ST(P,Q) = \sqrt{\sum_{1\le i\le n} p_i \log\left(\frac{p_i}{\mu(p_i,q_i)}\right) + q_i \log\left(\frac{q_i}{\mu(p_i,q_i)}\right)} \quad (1)$$

where μ is the average, i.e., $\mu(a,b) = \frac{a+b}{2}$ for any real numbers a and b. Define a function φ: $(0,1] \times (0,1] \mapsto \mathbb{R}$ by

$$\varphi(a,b) = a\log\left(\frac{a}{\mu(a,b)}\right) + b\log\left(\frac{b}{\mu(a,b)}\right) \quad (2)$$

for any $a, b \in (0,1]$. By Lemma 1 of Tulunay (2017) [7], we have

$$\varphi \ \geq \ 0 \quad (3)$$

where the equality holds only if a = b. Then we can express ST metric in terms of φ function as

$$ST(P,Q)) = \sqrt{\sum_{i\in I} \varphi(p_i, q_i)} \quad (4)$$

Then the function ST satisfies

(1) For any $P, Q \in \mathcal{X}_n$, $ST(P, Q) \geq 0$, where the equality occurs only if $p_i = q_i$ for all $i \in I$.

(2) For any $P, Q \in \mathcal{X}_n$, $ST(P, Q) = ST(Q, P)$, (symmetry).

(3) For any $P, Q, R \in \mathcal{X}_n$, $ST(P,R) \leq ST(P,Q) + ST(Q,R)$ (triangle inequality). [7]. Hence *ST* is a metric. More precisely, $(\mathcal{X}_n, ST)$ is a metric space. By assuming the convention log 0 = 0, (0, 1] can be replaced by [0, 1]. We assume this convention, whenever we use [0, 1], in place of (0, 1].

If P and Q are probabilities then the ST-metric definition above coincides with the square root of Topsøe distance which is symmetrized Kullback-Leibler divergence (Kullback-Leibler Information Criterion (KLIC), relative entropy) by average of the probabilities [7].

We will now show how ST-metric accounts for fluctuations. To see this, we examine the ST-distance between cumulative uniform distribution and one with a few small up or down movements over the unit interval.

Consider the cumulative uniform distribution, $U_n = \left\{\frac{i}{n} \mid i \in I\right\}$. If we view the n-points, $t_1 < t_2 < \ldots < t_n$ of time in the x-axis as the discretised points on (0, 1], (associating $t_i \to \frac{i}{n}$) then the graph of U_n is the $y = x$ line in the square (0, 1] × (0, 1].

For any $k \in I$ and for each $\varepsilon \in (0,1]$ with $\frac{k}{n} \pm \epsilon \in (0,1]$, define

$$U_n((k,\epsilon)^+) = \left\{\frac{1}{n}, \frac{2}{n}, \ldots, \frac{k-1}{n}, \frac{k}{n} + \epsilon, \frac{k+1}{n}, \ldots, \frac{n-1}{n}, 1\right\} \quad (5)$$

$$U_n((k,\epsilon)^-) = \left\{\frac{1}{n}, \frac{2}{n}, \ldots, \frac{k-1}{n}, \frac{k}{n} - \epsilon, \frac{k+1}{n}, \ldots, \frac{n-1}{n}, 1\right\} \quad (6)$$

for any $k \in I, k \neq 1$ and $k \neq n$. The $U_n((k_2,\varepsilon)^+)$ is the curve obtained from the curve U_n by moving the point $\frac{k}{n}$ to $\frac{k}{n} + \varepsilon$. This creates one up-movement with magnitude ε of the $y = x$ line in the square (0,1]×(0,1] of the real plane. Similar interpretation is valid for $U_n((k,\varepsilon)^-)$ with one down-movements (Fig. 1).

The rest of this section is devoted to show the following properties of *ST*-Metric.

1) *Structure of fluctuations:* The up-movement has a shorter *ST*-distance than the down-movement with the same magnitude at the same point of time.

2) *Timing of fluctuations:* Recent fluctuations have shorter *ST*-distance.

3) *Magnitude of fluctuations:* If the magnitude of a fluctuation increases then *ST*-distance increases.

4) *Quantity of fluctuations:* If the quantity of fluctuations increases then *ST*-distance increases.

The following propositions in this section are mathematical statements of the above facts. Their proofs can be found in Appendix A.

2.1 Structure of Fluctuations

Proposition 1 For any $k \in I$ and $\varepsilon \in (0,1]$, such that $\frac{k}{n} \pm \varepsilon \in (0,1]$

$$ST\left(U_n, U_n\left(\left(k,\varepsilon\right)^+\right)\right) < ST\left(U_n, U_n\left((k,\epsilon)^-\right)\right) \quad (7)$$

Note that ST-metric behaves risk aversive way, as negative movements are counted as the farther than the positive movements. In finance, a measure which accounts for up and down movements is very valuable. Profitable strategies can be constructed if it is used appropriately.

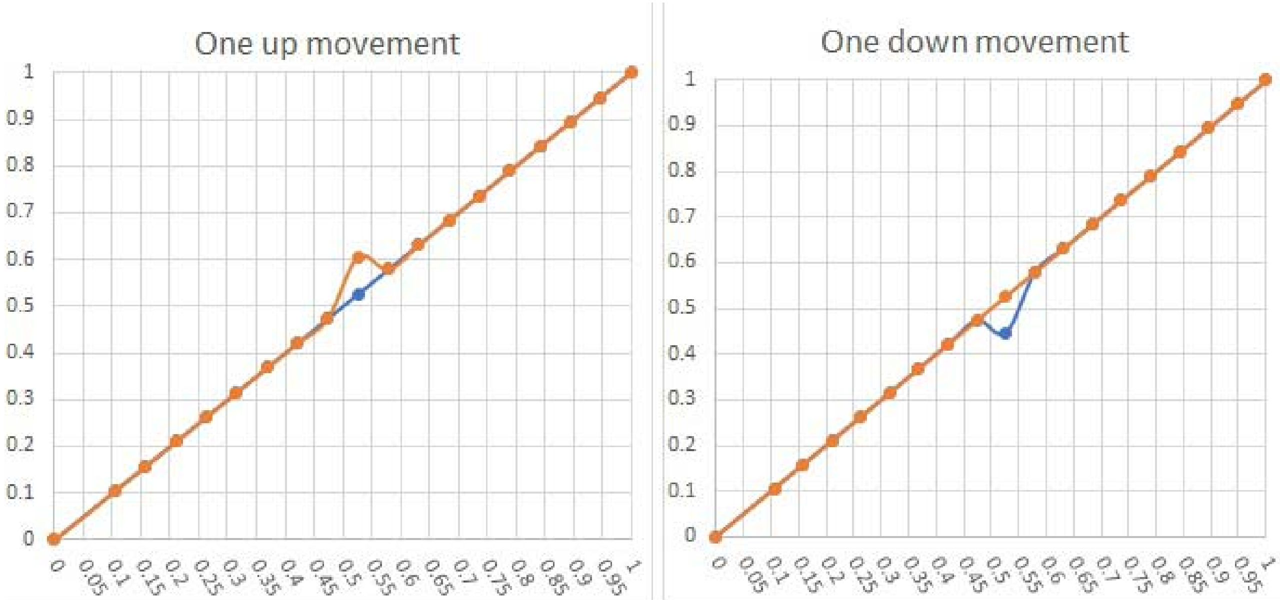

Fig. 1 **The graphs of U_n and $U_n((k,\epsilon)^+)$ for n = 19, k = 11 and ε = 0:08.**

2.2 Timing of Fluctuations

Proposition 2 For $k, l \in I$ and $\varepsilon \in (0,1]$ such that $k < l$ and $\frac{k}{n} \pm \epsilon, \frac{l}{n} \pm \epsilon \in (0,1]$, the followings hold:

$$ST\left(U_n, U_n\left((l,\epsilon)^{\circledast_1}\right)\right) < ST\left(U_n, U_n\left((k,\epsilon)^{\circledast_2}\right)\right) \quad (8)$$

where $\circledast_1, \circledast_2 \in \{+,-\}$.

It is well-known that the recent financial data is more relevant than the farther historical data in calculating the prices or forecasting. The methods, with high weights on the recent data are attractive (e.g., exponentially weighted moving average (EWMA) method).

2.3 Magnitude of Fluctuations

Proposition 3 For $\mathrm{k} \in \mathrm{I}$ and $\epsilon_1, \epsilon_2 \in (0,1]$ such that $\epsilon_1 < \epsilon_2$ and $\frac{k}{n} \pm \epsilon_1, \frac{k}{n} \pm \epsilon_2 \in (0,1]$, we have

$$ST\left(U_n, U_n\left((k,\epsilon_1)^{\circledast_1}\right)\right) < ST\left(U_n, U_n\left((k,\epsilon_2)^{\circledast_2}\right)\right) \quad (9)$$

where $\circledast_1, \circledast_2 \in \{+,-\}$.

Even the magnitude of fluctuations supplies important information for forecasting in finance or more generally in statistics, the usual statistical measures, like standard deviation give either none or very restricted information. Indeed, magnitude does matter in Finance. We note that ST-metric methods do not require to consider the jump events separately.

2.4 Quantity of Fluctuations

Similar to our definition of one fluctuation, we can define two (or more) fluctuations by

$$U_n\left((k_1,\epsilon_1)^{\circledast_1},(k_2,\epsilon_2)^{\circledast_2}\right)=\Big\{\frac{1}{n},\frac{2}{n},\ldots,\frac{k_1-1}{n},\frac{k_1}{n}\circledast_1\epsilon_1,\frac{k_1+1}{n},\ldots$$

$$\frac{\boldsymbol{k_2-1}}{\boldsymbol{n}},\frac{\boldsymbol{k_2}}{\boldsymbol{n}}\circledast_2\epsilon_2,\frac{\boldsymbol{k_2+1}}{\boldsymbol{n}}\cdots,\frac{\boldsymbol{n-1}}{\boldsymbol{n}},\boldsymbol{1}\Big\}$$

where $\circledast_1,\circledast_2\in\{+,-\}$, $k_1,k_2\in I,\varepsilon_1,\epsilon_2\in(0,1]$ and $\frac{k_1}{n}\pm\epsilon_1,\frac{k_2}{n}\pm\epsilon_2\in(0,1]$.

Proposition 4 For $k_1,k_2\in I,\varepsilon_1,\epsilon_2\in(0,1]$ such that $\frac{k_1}{n}\pm\epsilon_1,\frac{k_2}{n}\pm\epsilon_2\in(0,1]$, we have

$$ST\left(U_n,U_n\left((k_i,\epsilon_i)^{\circledast_i}\right)\right)<ST\left(U_n,U_n\left((k_1,\epsilon_1)^{\circledast_1},(k_2,\epsilon_2)^{\circledast_2}\right)\right)\quad(10)$$

where $\circledast_1,\circledast_2\in\{+,-\}$ and i =1, 2.

This is obvious as $\varphi(a,b)>0$ by Eq. (3) and

$$\varphi(a,b)<\varphi(a,b)+\varphi(c,d)$$

for all distinct $a,b,c,d\in(0,1]$. Adding more fluctuations would add positive values to ST-distances. In Finance, more fluctuations imply higher variations of the returns and so less stability and higher risk. In terms of *ST*-metric, the longer the distance, the higher the risk.

Remark 1 Note that the standard deviation of *n* variables can be viewed as the $\mathcal{L}^2$-*Euclidean* distance from a point in an n-dimensional space to the mean point $(\mu,\mu,\ldots,\mu)\in\mathbb{R}^n$ (up to a scalar).

Let's compare the above results of *ST*-metric distance with $\mathcal{L}^2$ -Euclidean distance in an *n*-dimensional space:

$$\mathcal{L}^2(X,Y)=\sqrt{\sum_{1\le i\le n}(x_i-y_i)^2}$$

for $X,Y\in\mathbb{R}^n$. It is straightforward to show that, for any $k\in I$ and for each $\varepsilon\in(0,1]$ with $\frac{k}{n}\pm\varepsilon\in(0,1]$

$$\mathcal{L}^2\left(U_n,U_n\left((k,\epsilon)^{\circledast}\right)\right)=\epsilon$$

where $\circledast\in\{+,-\}$. That is independent of *k*. Hence, $\mathcal{L}^2$-Euclidean distance provides no information about the structure of fluctuation and the timing of fluctuation, but agrees with the *ST*-metric on the properties of magnitude and quantity (Eq. (9) and Eq. (10)): For all $\varepsilon_1,\varepsilon_2\in(0,1]$

$$\epsilon_1=\mathcal{L}^2\left(U_n,U_n\left((k,\varepsilon_1)^{\circledast_1}\right)\right)<\mathcal{L}^2\left(U_n,U_n\left((k,\epsilon_2)^{\circledast_2}\right)\right)=\epsilon_2$$

$$\epsilon_i=\mathcal{L}^2\left(U_n,U_n\left((k_i,\epsilon_i)^{\circledast_i}\right)\right)$$

$$<\mathcal{L}^2\left(U_n,U_n\left((k_1,\varepsilon_1)^{\circledast_1},(k_2,\epsilon_2)^{\circledast_2}\right)\right)=\sqrt{(\epsilon_1^2+\varepsilon_2^2)}$$

Remark 2 Recall that $\mathcal{L}^1$-Euclidean distance in an *n*-dimensional space is defined as

$$\mathcal{L}^1(X,Y)=\sum_{1\le i\le n}|x_i-y_i|$$

for $X,Y\in\mathbb{R}^n$, It is straightforward to show that, for any $k\in I$ and for each $\varepsilon\in(0,1]$ with $\frac{k}{n}\pm\epsilon\in(0,1]$.

$$\mathcal{L}^1\left(U_n,U_n\left((k,\epsilon)^{\circledast}\right)\right)=\epsilon$$

where $\circledast\in\{+,-\}$. That is independent of *k*. Hence, $\mathcal{L}^1$-Euclidean distance provides no information about the structure of fluctuation and the timing of fluctuation, but agrees with the *ST*-metric on the properties of magnitude and quantity (Eq. (9) and Eq. (10)).

3. ST-Metric Method for Solving the Optimization Problem in Five-Factors Model

We consider the five-style-factors model as in the paper of Ang et al. (2018) [19]. The five-style factors and their MSCI indexes are as follows:

1) **Value:** (*MSCI Value Weighted Indexes*) Capture Value factor by weighting according to four fundamental variables (Sales, Earnings, Cash Flow, Book Value).
2) **Size:** (*MSCI Equal Weighted Indexes*) Capture low size effect by equally weighting all stocks in a given parent index.
3) **Quality:** (*MSCI Quality Indexes*) Capture high quality stocks by weighting, based on debt-to-equity, return-on-equity, and earnings variability.
4) **Momentum:** (*MSCI Momentum Indexes*) Reect the performance of high momentum

stocks by weighting based on 6 and 12-month momentum scaled by volatility.

5) **Minimum Volatility:** (*MSCI Minimum Volatility Indexes*) Reect empirical portfolio with lowest forecast volatility using minimum variance optimization.

For more information on factor investing, we refer to Kassam et al. (2013) [20].

Let $S = \{1, 2, \ldots, 5\}$ be the indices for the risk factors and $I = \{1, 2, \ldots, n\}$ the indices of time for the concerned time series. For each $s \in S$ and $i \in I$, denote by $f_{s,i}$ the returns of the risk factor s at time t_i. Let $w_{s,i}$ be the risk exposure (weight) corresponding to each factor $f_{s,i}$, satisfying
(Condition C):

$$\sum_{s\in S} w_{s,i} = 1 \text{ and } w_{s,i} \geq 0 \text{ for all } i \in I \text{ and } s \in S \quad (11)$$

Let b_i be the benchmark returns at time t_i for each $i \in I$. Then the set of all benchmark returns is

$$B = \{ b_i \mid i \in I \}$$

Denote by p_i the constructed portfolio return at time t_i for each $i \in I$ by the factors and the weights. Namely,

$$p_i = \sum_{s \in S} w_{s,i} f_{s,i}$$

Then the set of all constructed portfolio returns is $P = \{ p_i \mid i \in I \}$.

Further, we construct a set, say □n consisting of B and all possible sets P.

3.1 Tracking Error method (TE-method)

Tracking error method has been used in the paper of Ang et al. (2018) [19]. They claimed that “Tracking Error captures the absolute differences in return between the index and the mimicking portfolio and is, thus, a measure of the goodness of fit of the optimization exercise to create a long-only representation”.

Denote by σ the standard deviation of the benchmark returns B. Then Tracking Error is defined as a function $\mathcal{TE} : \mathcal{X}_n \times \mathcal{X}_n \mapsto \mathbb{R}$ by

$$\mathcal{TE}(P, B) = \sigma \sum_{i \in I} \left| p_i - b_i \right|$$

The optimization problem for the Tracking Error method is to find $w_{s,i}$ such that $\mathcal{TE}(P, B)$ is minimal where the minimum is taken over all elements of the set $\mathcal{X}$. Namely,

$$\min_{P \in \mathcal{X}_\infty} \mathcal{TE}(P, B) = \min_{w_{s,i}} \sigma \sum_{i \in I} \left| \sum_{s \in S} w_{s,i} f_{s,i} - b_i \right| \quad (12)$$

subject to Condition C in Equation (11).

3.2 ST-metric Method

For this application, there are some advantages to consider ST-metric defined on values in [0, 1], instead of (0, 1], by assuming the convention log 0 = 0.

In order to be able to apply ST-metric, we first need all values to be in [0, 1]. However, the returns of an asset may not be in [0, 1]. We need to transform all returns onto [0, 1] in a way that timing order of the data and the distributions remain invariant. This problem has been tackled in Tulunay (2017) [7]. Such a transformation is called Maximal Invariant transformation that leaves the parameter space of distributions invariant [21]. In particular, assuming X is a set of random variables, for any fixed $x_0 \in X$, transformations of subtracting x_0 or of dividing by x_0 or a combination of both are all examples of a maximal invariant transformation [21]. Vertically shifting all historical returns above the x-axis (y = 0 line) and vertically shrinking below y = 1 line ensure all returns to be in [0, 1]. Clearly, such transformations leave distributions invariant. Now, we are going to present the maximal invariant transformation that we consider for this application in mathematics language.

Let $D = \{ f_{s,i} \mid s \in S, i \in I \} \cup \{ b_i \mid i \in I \}$ be the set of all available historical returns of all factors and the benchmark. Define a maximal invariant transformation $g: D \mapsto [0,1]$ by

$$g(d) = \frac{d+\alpha}{\beta} \tag{13}$$

where

$$\alpha = \begin{cases} -\min_{d\in D}\{d\} & \text{if } \min_{d\in D}\{d\} < 0 \\ 0 & \text{otherwise} \end{cases}$$

and

$$\beta = \begin{cases} \max_{d\in D}\{d+\alpha\} & \text{if } \max_{d\in D}\{d+\alpha\} > 1 \\ 1 & \text{otherwise} \end{cases}$$

In this case, the constructed portfolio is

$$g(p_i) = \sum_{s\in S} w_{s,i}\, g(f_{s,i})$$

Condition C in Eq. (11) implies that $g(p_i) \geq 0$. Further, since all affine combinations of a nonempty set in an n dimensional real space lie on that set [22], we have $g(p_i) \in [0,1]$. Hence, the range of this transformation, g(D) is a subset of [0; 1].

So ST-metric is well-defined on the set $\{g(P)\} \cup \{g(B)\}$ and

$$\mathcal{ST}(g(P), g(B)) = \sqrt{\sum_{i\in I} \varphi\left(g(p_i), g(b_i)\right)} \tag{14}$$

where the function φ is as in Equation (2). The optimization problem for the ST-metric method becomes finding $w_{s,i}$ such that $ST(g(P), g(B))$ is minimal. Namely,

$$\min_{g(P)\in\mathcal{X}_n} \mathcal{ST}(g(P), g(B)) = \min_{w_{s,i}} \sqrt{\sum_{i\in I} \varphi\left(\sum_{s\in S} w_{s,i}\, g(f_{s,i}), g(b_i)\right)}$$

subject to Condition C in Eq. (11).

Remark 3 In practice, it may be difficult to decide whether the minimal value in Eq. (15) is optimal or not. Tulunay (2017) [7] argue that two discrete curves are indistinguishable if their ST-metric is smaller than and equal to ST-metric of cumulative uniform distribution and Lorenz curve. The ST-metric of cumulative uniform distribution and Lorenz curve only depends on the size n of the time series. It is easy to calculate it. Instead of minimizing the ST-metric in Eq. (15), we can set the problem to approach the ST-metric between the constructed portfolio and the benchmark returns to the ST-metric of cumulative uniform distribution and Lorenz curve. The optimization problem becomes

Find the weights $w_{s,i}$ such that

$$ST(g(P), g(B)) \to \mathcal{ST}\left(U_{n,} L_n\right)$$

where U_n is cumulative uniform distribution as above, and L_n is the Lorenz Curve, that is as a set

$$L_n I = \left\{ \left(\frac{i}{n}\right)^{1.16} \mid \cdot \in \right\}$$

4. Empirical Evidence

Note that we use the data service of Thomson Reuters News Analytics (TRNA) to attain our data sets.

We choose two sets of data: one from Australian market and another one from US market in a similar way of Ang et al. (2018) [19].

4.1 Australian Market Case

Our Australian market data consists of

- The Benchmark: AUMF; iShares Edge MSCI Australia Multifactor ETF (Exchange Traded Funds),
- The indexes to approximate style risk factors are
  - Value: MSCI AUSTRALIA: S VALUE WEIGHTED - PRICE INDEX
  - Size: MSCI AUSTRALIA RISK WEIGHTED - PRICE INDEX
  - Quality: MSCI AUSTRALIA QUALITY - PRICE INDEX
  - Momentum: MSCI AUSTRALIA MOMENTUM - PRICE INDEX
  - Minimum Volatility: MSCI AUSTRALIA MINIMUM VOL(AUD) - PRICE INDEX

The inception date of the benchmark AUMF is 11th of October, 2016. The first-historical-data-available is

2th of November, 2016. The date range is from 2-Nov-2016 to 19-Jun-2018. We take monthly returns for 20 months. The actual weights of the style risk factors, disjoining the benchmark AUMF can be seen in Fig. 2. Apparently, the main driving risk factors for AUMF financial product are size and momentum.

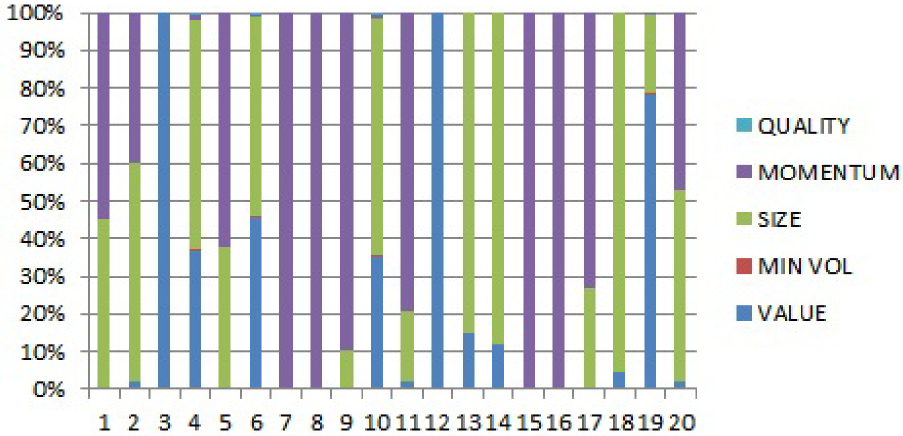

Fig. 2 The Risk Factors Decomposition of AUMF from Oct 2016 to Jun 2018.

We consider two sets of time intervals: 6-months of returns and 3-months of returns. Shifting one month yields 14 scenarios for the 6-monthly experiment and 17 scenarios for the 3-monthly experiment.

For the 6-monthly experiment, the in-sample data consists of monthly returns for 6-months duration. We apply both TE-method and ST-metric methods to calculate monthly weights of the constructed portfolio. We choose the out of sample as the 7^{th} month for forecasts. The forecasted weights for the 7^{th} month is the average of the monthly weights over six months. Then we attain the forecasted return of the constructed portfolio as the weighted averages. We compute actual weights of the factors for the 7^{th} months to compare the results of TE-method and ST-metric methods. The comparison is by looking at two types of errors:

1) Forecast Error: Absolute error of the forecasted value with the actual value for the 7^{th} month,
2) Weights Error: The sum of the square errors of the forecasted weights for the 7^{th} month with the actual weights of the factors.

The 3-monthly experiments are similar to the 6-monthly experiment with in-sample size 3, and the out-of-sample is the 4th month for forecasts. Apart from averaging weights for forecast, we also examine finding one fixed set of weights over 3-months[2].

The success rate of the ST-metric method over TE-method can be computed as

$$\frac{\#\{ST\text{ - } metric\ method\ error\ is\ less\ than\ TE\text{ - }method\ error\ \}}{\#\{\ all\ scenarios\ in\ the\ \exp eriment\ \}}$$

The results are given in Table 1 where Experiment 1 is the 6-monthly experiment, Experiment 2 is the 3-monthly experiment with averaging the weights over 3 months, and Experiment 3 is the 3-monthly experiment with a fixed set of weights over 3-months.

As we mentioned in Remark 3, we find the weights by approaching the ST-metric between the constructed portfolio and the benchmark returns to the ST-metric of cumulative uniform distribution and the Lorenz curve:

$ST(U_6, L_6) = 0.06$ for n = 6 (six monthly experiment)

$ST(U_3, L_3) = 0.09$ for n = 3 (three monthly experiments)

Table 1 The success rate of the ST-metric method over TE-method.

Error type	Experiment 1	Experiment 2	Experiment 3
Forecast error	71%	65%	65%
Weights error	100%	94%	88%

[2] Note that our experiments of finding one set of weights for all 6-months failed. The 6-months fixed weights are practically not acceptable.

4.2 US Market Case

We choose the next set of the data from USA market.

- The Benchmark: Russell 1000 Index,
- The indexes to approximate style risk factors are
  - Value: MSCI USA ENHANCED VALUE - PRICE INDEX
  - Size: MSCI USA RISK WEIGHTED - PRICE INDEX
  - Quality: MSCI USA SECTOR NEUTRAL QLTY - PRICE INDEX
  - Momentum: MSCI USA MOMENTUM - PRICE INDEX
  - Minimum Volatility: MSCI USA MINIMUM VOLATILITY - PRICE INDEX

The actual weights of the style risk factors, disjoining the benchmark are given in Fig. 3.

The success rates of ST-metric method over TE-method for USA market are given in Table 2 where, as in the case of Australian Market above, Experiment 1 is the 6-monthly experiment, Experiment 2 is the 3-monthly experiment with averaging the weights over 3 months, and Experiment 3 is the 3-monthly experiment with a fixed set of weights over 3-months.

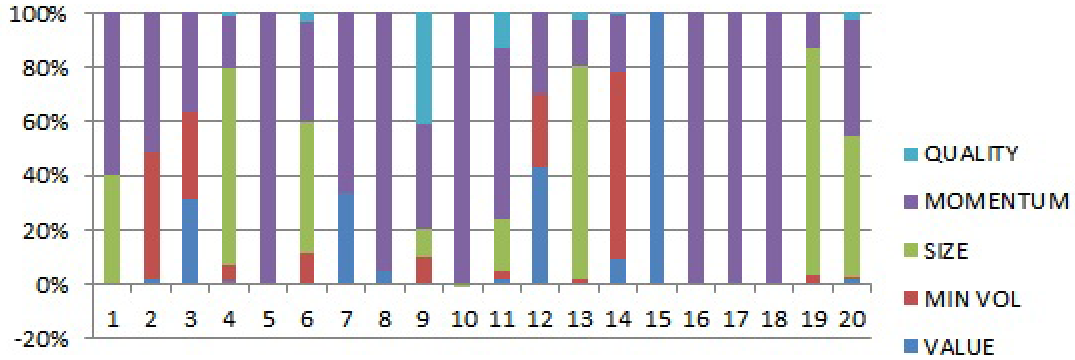

Fig. 3 The Risk Factors Decomposition of Russell 1000 from Oct 2016 to Jun 2018.

Table 2 The success rate of the ST-metric method over TE-method.

Error type	Experiment 1	Experiment 2	Experiment 3
Forecast error	71%	82%	76%
Weights error	93%	100%	100%

5. Conclusion

Time series analysis is generally based on statistical methods. By introducing ST-metric method, we purpose to view time series as transformed discrete curves on the unit interval. By the nature of its definition, the ST-metric accounts for up or down movements, recent or older historical data, magnitude of movements and the quantity of movements. As far as the authors know, there is no statistical method which accounts for all of these features simultaneously. Therefore, ST-metric method is efficient. If it is appropriately applied, ST-metric methods are promising to introduce profitable investment strategies. ST-metric is a new concept, and only two applications have been investigated so far: performance measures of fund managers [7] and finding the factor exposures of factor models. There are many open questions on the subject of ST-metric to discover further. For example, since ST-metric increases with the respect to number of fluctuations, one may wish to investigate how to apply ST-metric as a noise reduction method. If one can understand how to use ST-metric on pricing financial instruments, that would be the start of a new avenue in financial mathematics.

References

[1] Philippatos, G. C. and Wilson, C. J. (1972). “Entropy, Market Risk and the Selection of Efficient Portfolios.” *Applied Economics* 4: 209-220.

[2] Dionisio, A., Menezes, R. and Mendes, D. (2006). “An

Econophysics Approach to Analyse Uncertainty in Financial Markets: An Application to the Portuguese Stock Market." *The European Physical Journal B: Condensed Matter and Complex Systems* 50 (1): 161-164.

[3] Bentes, S. R. & Menezes, R. (2012), "Entropy: A new measure of stock market volatility?" *Journal of Physics: Conference Series* 394 (012033): 1-6.

[4] Cover, T. M. and Thomas, J. A. (2006). *Elements of Information Theory*. Wiley-Interscience.

[5] Topsoe, D. (2000). "Some Inequalities For Information Divergence and Related Measures of Discrimination." *IEEE Transactions on Information Theory* 46 (4): 1602-1609.

[6] Endres, D. M. and Schindelin, J. E. (2003). "A New Metric for Probability Distributions." *IEEE Transactions on Information Theory* 49 (7): 1858-1860.

[7] Tulunay, I. (2017). "A Robust Method for Ordering Performances of Multi-Assets, Based Purely on Their Return Series." *Journal of Mathematics and System Science* 11.

[8] Fama, E. and French, K. (1993). "Common Risk Factors in the Returns on Stocks and Bonds." *Journal of Financial Economics* 33 (1): 3-56.

[9] Ang, A. (2014). *Asset Management: A Systematic Approach to Factor Based Investing*. New York: Oxford University Press.

[10] Fama, E. and French, K. (2015. "A Five-Factor Asset Pricing Model." *Journal of Financial Economics* 116: 1-22.

[11] Clarke, R., De'Silva, H. and Thorley, S. (2016). "Fundamentals of Efficient Factor Investing." *Financial Analysts Journal* 72 (6): 9-26.

[12] Fama, E. and French, K. (2017). "International Tests of a Ve-Factor Asset Pricing Model." *Journal of Financial Economics* 123: 441-463.

[13] Madhavan, A. and Sobczyk, A. (2017). "Estimating Time-Varying Factor Exposures." *Financial Analysts Journal* 73 (4): 41-54.

[14] Fama, E. and French, K. (2018). "Choosing Factors." *Journal of Financial Economics* 128: 234-252.

[15] Abner, D. J. (2010). *The ETF Handbook: How to Value and Trade Exchange-Traded Funds* (3rd ed.). John Wiley & Sons, Inc., Hoboken, New Jersey.

[16] Madhavan, A. N. (2016). *Exchange-Traded Funds and the New Dynamics of Investing*. Oxford Scholarship Online.

[17] Madhavan, A. and Sobcyzk, A. (2016). "Price Dynamics and Liquidity of Exchange-Traded Funds." URL: https://ssrn.com/abstract=2429509

[18] Piccotti, L. R. (2018). "Etf Premiums and Liquidity Segmentation." *The Financial Review* 53: 117-152.

[19] Ang, A., Madhavan, A. and Sobczyk, A. (2018). "What's in Your Benchmark? A Factor Analysis of Major Market Indexes." *The Journal of Portfolio Management* 44 (4): 46-59.

[20] Kassam, A., Gupta, A., Jain, S., Kouzmenko, R. and Briand, R. (2013). "Introducing Msci Index-Metrics: An Analytical Framework for Factor Investing." https://www.msci.com/documents/10199/402635a5-fd5d-498e-985a-1bec88d8b1.

[21] Lehmann, H. L. and Romano, J. P. (2005). *Testing Statistical Hypotheses* (3rd ed.). Springer.

[22] Soltan, V. (2015). *Lectures on Convex Sets*. World Scientific Publishing Co. Pte. Ltd.

Appendix

The Jensen Inequality for any convex function f is given by

$$f\left(\sum_{i=1}^{n} w_i a_i\right) \le \sum_{i=1}^{n} w_i f\left(a_i\right) \quad (16)$$

Where $\sum_{i=1}^{n} w_i = 1$, $w_i \in [0; 1]$, a_i are in the domain of f, for each i = 1, 2, …, n.

Before starting the proofs of the propositions in Section 2, note that the definition of ST-distance between $A = \left\{a_i \in (0,1] \,\|\, i \in I\right\}$ and $B = \left\{b_i \in (0,1] \,\|\, i \in I\right\}$ can be written as

$$\mathcal{ST}(A,B) = \sqrt{\sum_{i \in I} \varphi\left(a_i, b_i\right)}$$

where φ is defined as in Eq. (2).

The proofs of the propositions in Section 2 are as follows.

A.1 Proof of Proposition 1

Proof: Assume $k \in I$, $\epsilon \in (0;1)$ and $\frac{k}{n} \pm \varepsilon \in (0,1]$. Since $\varphi(x, x) = 0$ for all $x \in [0, 1]$, we have

$$\mathcal{ST}\left(U_{n,}U_n\left((k,\varepsilon)^+\right)\right) = \varphi\left(\frac{k}{n^3}, \frac{k}{n} + \varepsilon\right)$$

$$\mathcal{ST}\left(U_{n,}U_n\left(\langle k,\epsilon)^-\right)\right) = \varphi\left(\frac{k}{n}, \frac{k}{n} - \epsilon\right)$$

Hence, it is sufficient to show that

$$\xi = \varphi\left(\frac{k}{n}, \frac{k}{n} - \varepsilon\right) - \varphi\left(\frac{k}{n}, \frac{k}{n} + \varepsilon\right) \geq 0$$

Let $a = \frac{k}{n}$, for simplicity. Then we have

$$\xi = a\log\left(\frac{2a}{2a-\varepsilon}\right) + (a-\varepsilon)\log\left(\frac{2(a-\varepsilon)}{2a-\varepsilon}\right) - a\log\left(\frac{2a}{2a+\varepsilon}\right) - (a+\varepsilon)\log\left(\frac{2(a+\epsilon)}{2a+\varepsilon}\right) \quad (17)$$

$$= a\left(-\log\left(\frac{2a-\varepsilon}{2a}\right)\right) + (a-\varepsilon)\left(-\log\left(\frac{2a-\varepsilon}{2(a-\varepsilon)}\right)\right) \quad (18)$$

$$+a\left(-\log\left(\frac{2a}{2a+\varepsilon}\right)\right) + (a+\varepsilon)\left(-\log\left(\frac{2(a+\varepsilon)}{2a+\varepsilon}\right)\right) \quad (19)$$

Applying Jensen's inequality for the convex function -log, it is straightforward to show that

$$\frac{\xi}{4a} \geq -\log\left(\frac{1}{4a}\left(2a - \varepsilon + \frac{2a^2 + 2(a+\varepsilon)^2}{2a+\varepsilon}\right)\right)$$

$$= -\log\left(\frac{1}{4a}\left(2a - \varepsilon + \frac{(2a+\varepsilon)^2}{2a+\varepsilon}\right)\right)$$

$$= -\log(1) = 0$$

Since $\varphi(a, a\text{-}\epsilon) \neq \varphi(\text{a}, \text{a} + \epsilon)$ for any $\epsilon \in (0; 1)$, $\xi \neq 0$. Hence $\xi > 0$, as required.

A.2 Proof of Proposition 2

Proof: We will only prove the inequality for the case $\circledast_1 = +$ and $\circledast_2 = +$. The others can be proven in a similar way. To prove the first inequality, it is enough to show that

$$\xi = \varphi(a, a+\varepsilon) - \varphi(b, b+\varepsilon) > 0$$

Where $a = \frac{k}{n}$ *and* $b = \frac{l}{n}$ Similar to the calculation in Eq. (17), by Jensen's Inequality for the convex function -log,

$$\frac{\xi}{2(a+b+\varepsilon)} \geq -\log\left(\frac{1}{2(a+b+\varepsilon)}\left(2a + \varepsilon + \frac{2b^2 + 2(b+\varepsilon)^2}{2b+\varepsilon}\right)\right)$$

$$= -\log\left(\frac{1}{2(a+b+\varepsilon)}\left(2a + \varepsilon + \frac{(2b+\varepsilon)^2}{2b+\varepsilon}\right)\right)$$

$$= -\log(1) = 0$$

Since $k < \text{l}$, $\xi > 0$, as required.

A.3 Proof of Proposition 3

Proof: We will only prove the inequality for the case $\circledast_1 = +$ and $\circledast_2 = +$. The others can be proven in a similar way. To prove the first inequality, it is enough to show that

$$\xi = \varphi\left(\frac{k}{n}, \frac{k}{n} + \varepsilon_2\right) - \varphi\left(\frac{k}{n}, \frac{k}{n} + \epsilon_1\right)$$

For simplicity, let $a = \frac{k-1}{n}$ and $b = \frac{k}{n}$. Then we need to show that

$$\frac{\xi}{4a + \epsilon_1 + \epsilon_2} \geq -\log\left(\frac{1}{4a + \varepsilon_1 + \varepsilon_2}\left(2a + \varepsilon_2 + \frac{2a^2 + 2(a + \varepsilon_1)^2}{2a + \varepsilon_1}\right)\right)$$

$$= -\log\left(\frac{1}{4a + \varepsilon_1 + \varepsilon_2}\left(2a + \varepsilon_2 + \frac{(2a + \varepsilon_1)^2}{2a + \varepsilon_1}\right)\right)$$

$$= -\log(1) = 0$$

Since $\epsilon_1 \neq \epsilon_2$, $\xi > 0$, as required.