Singapore Management University

Institutional Knowledge at Singapore Management University

Research Collection School Of Computing and Information Systems | School of Computing and Information Systems

8-2018

# Transaction cost optimization for online portfolio selection

Bin LI
*Wuhan University*

Jialei WANG
*University of Chicago*

Dingjiang HUANG
*East China Normal University*

Steven C. H. HOI
*Singapore Management University*, chhoi@smu.edu.sg

Follow this and additional works at: https://ink.library.smu.edu.sg/sis_research

Part of the Databases and Information Systems Commons, and the Portfolio and Security Analysis Commons

## Citation

LI, Bin; WANG, Jialei; HUANG, Dingjiang; and HOI, Steven C. H.. Transaction cost optimization for online portfolio selection. (2018). *Quantitative Finance*. 18, (8), 1411-1424.
**Available at:** https://ink.library.smu.edu.sg/sis_research/3777

This Journal Article is brought to you for free and open access by the School of Computing and Information Systems at Institutional Knowledge at Singapore Management University. It has been accepted for inclusion in Research Collection School Of Computing and Information Systems by an authorized administrator of Institutional Knowledge at Singapore Management University. For more information, please email cherylds@smu.edu.sg.

**Published in Quantitative Finance, 2017 August, Pages 1-14**
**http://doi.org/10.1080/14697688.2017.1357831**

# Transaction cost optimization for online portfolio selection

BIN LI*†, JIALEI WANG‡, DINGJIANG HUANG§ and STEVEN C. H. HOI*¶

†Department of Finance, Economics and Management School, Wuhan University, Wuhan, P.R. China
‡Department of Computer Science, University of Chicago, Chicago, USA
§School of Data Science and Engineering, East China Normal University, Shanghai, P.R. China
¶School of Information Systems, Singapore Management University, Singapore

To improve existing online portfolio selection strategies in the case of non-zero transaction costs, we propose a novel framework named *Transaction Cost Optimization* (TCO). The TCO framework incorporates the L1 norm of the difference between two consecutive allocations together with the principle of maximizing expected log return. We further solve the formulation via convex optimization, and obtain two closed-form portfolio update formulas, which follow the same principle as Proportional Portfolio Rebalancing (PPR) in industry. We empirically evaluate the proposed framework using four commonly used data-sets. Although these data-sets do not consider delisted firms and are thus subject to survival bias, empirical evaluations show that the proposed TCO framework may effectively handle reasonable transaction costs and improve existing strategies in the case of non-zero transaction costs.

*Keywords*: Portfolio optimization; Transaction costs; Learning in financial models; Investment strategy

*JEL Classification*: C4, C5, C44, C51, G1, G11

## 1. Introduction

We propose a novel online portfolio selection (OLPS) framework, named *Transaction Cost Optimization* (TCO), so as to improve existing strategies with non-zero proportional transaction costs. The framework can be applied to most existing OLPS algorithms. Inspired by our preliminary analysis, the proposed TCO appends a L1 regularization to the traditional objective function of maximizing portfolio's expected log return (Kelly 1956, Li and Hoi 2014). Solving the TCO's optimization problem, we can obtain two closed-form portfolio update formulas and derive two specific algorithms named 'TCO1' and 'TCO2', which follow state of the art mean reversion predictions (Li *et al*. 2012, 2015). Extensive empirical experiments on data-sets‖ show that the derived algorithms are effective in boosting performance in the environment of non-zero proportional transaction costs.

The first key motivation of this study is to improve the out-of-sample performance of existing algorithms when the transaction costs are non-zero. It it widely documented that in frictionless backtesting environments (Huang *et al*. 2013, Li *et al*. 2015), existing OLPS strategies achieved significant success, which seem to beat the best human investors on this planet. On the other hand, as a routine in the backtests, they often simulated their algorithms in an environment of non-zero (proportional) transaction costs, in which the performance degrade exponentially with increasing rates. One crucial problem is that almost no existing algorithm has ever considered the transaction costs issue in their decision-making process, and thus suffers a lot in tests with non-zero costs. The second key motivation is that the market imposed transaction costs are directly related to the L1 norm of the difference between two consecutive allocations. Thus, besides the traditional objective of maximizing a portfolio's expected log return, we add a second term that minimizes the L1 norm of difference between two consecutive allocations, which is equivalent to minimizing the incurred transaction costs.

In summary, our main contributions are fourfold. First, we formulate the problem of online portfolio selection with proportional transaction costs and analyse the cause of transaction costs. Second, we propose a novel framework to handle proportional transaction costs, and derive closed-form solutions.

*Corresponding authors. Email: binli.whu@whu.edu.cn (Bin Li); chhoi@smu.edu.sg (Steven Hoi)

‖To better compare with existing strategies, we employ widely used data-sets, which may be subject to the survival bias.

Based on the solutions, we design two specific algorithms following mean reversion prediction schema. Note that the mathematics behind the solutions is not new, however, we are the first to leverage the mathematical method for online portfolio selection with non-zero transaction costs. Different from other papers on online portfolio selection, our paper uses the new mathematical method and proposes a novel algorithmic framework to deal with transaction costs. Our paper also differs from other papers on their underlying mathematics, including both formulations and algorithms. Finally, empirical evaluations show that the derived algorithms can significantly improve the performance when non-trivial transaction costs exist.

The rest of this paper is organized as follows. Section 2 mathematically formulates the research problem, and section 3 reviews and analyses related work. Section 4 analyses the cause of transaction costs, presents the proposed framework, and derives two specific mean reversion based algorithms. Their effectiveness is validated by extensive empirical studies on real stock markets in section 5. Section 6 summarizes the paper and provides future directions.

## 2. Preliminaries

### *2.1. Problem setting*

In the problem of online portfolio selection, a portfolio manager sequentially (re)distributes his/her capital such that the portfolio's allocation can capture the assets' volatility and maximize his/her terminal capital in the long run. In particular, at the beginning of every period, the manager decides an allocation based on his expectation of the assets' changes in the coming period. He/She then rebalances from current allocation to the decided new allocation. After the rebalance, the portfolio's capital will increase or decrease following the market fluctuations till the end of the period. Note that the allocation on the assets also changes dynamically, as outperforming assets' capital will increase and underperforming assets' capital will decrease. At the beginning of next period, the above procedure repeats. Existing research (Li and Hoi 2014) often assumes that the portfolio rebalance is frictionless and incurs no transaction costs. However, during the rebalance, the market does enforce transaction costs, such as the commission fee paid to brokers, the taxes paid to governments, etc. In the following, we try to mathematically formulate the research problem (Bauer and Kohavi 1999, Albeverio *et al*. 2001, Györfi and Vajda 2008).

We start by introducing some basic terms. Consider a market with $m$ assets to be invested for $n$ periods. Their price changes for period $t$ are represented by a *price relative vector*, i.e. $\mathbf{x}_t \in \mathbb{R}^m_+$. The element $x_{t,i}$ denotes the ratio of change to an investment in asset $i$ for period $t$. For example, $x_{t,i} = 1.1$ means investing in the asset will increase the initial capital by a factor of 1.1, or equivalently an increment of 10%. Note that the element $x_{t,i}$ is defined as $x_{t,i} = \frac{p_{t,i}}{p_{t-1,i}}$, where $p_{t,i}$ denotes the closing price at the end of the $t$th trading day. Investing in an asset for the $t$th trading day means buying the asset at the end of day $t-1$ and hold till the end of day $t$. The allocation over the market (or portfolio) is specified by a *portfolio vector*, denoted as $\mathbf{b} = (b_1, \ldots, b_m)$, where $b_i$ represents the proportion of capital invested in the $i^{\text{th}}$ asset. Typically, we assume the portfolio is self-financed and no margin/shorting is allowed, and then $\mathbf{b} \in \Delta_m$, where $\Delta_m = \left\{\mathbf{b} : \mathbf{b} \in \mathbb{R}^m_+, \sum_{i=1}^m b_i = 1\right\}$. A portfolio selection strategy is a sequence of mappings,

$$\mathbf{b}_t : \mathbb{R}_+^{(t-1)\times m} \rightarrow \Delta_m, \quad t = 2, 3, \ldots,$$

where $\mathbf{b}_t = \mathbf{b}_t(\mathbf{x}_1, \ldots, \mathbf{x}_{t-1})$ denotes the portfolio used for period $t$. Note that as no initial information exists, the portfolio generally starts with uniform portfolio, i.e. $\mathbf{b}_1 = \left(\frac{1}{m}, \ldots, \frac{1}{m}\right)$.

At the beginning of period $t$, a manager (or algorithm here) decides a new portfolio of $\mathbf{b}_t$ and rebalances from current allocation of $\hat{\mathbf{b}}_{t-1}$ to the new allocation $\mathbf{b}_t$. Note that as the current allocation of $\hat{\mathbf{b}}_{t-1}$ is different from $\mathbf{b}_{t-1}$, where $\hat{b}_{t-1,i} = \frac{b_{t-1,i}x_{t-1,i}}{\mathbf{b}_{t-1}\cdot\mathbf{x}_{t-1}}$, $i = 1, \ldots, m$. Let $\gamma_s$ and $\gamma_b$ be the transaction costs rates to be charged during sales and purchases, respectively, and $w_{t-1}$ be the net proportion after transaction costs imposed by the markets. The sale occurs when the proportion before rebalancing is greater than the proportion after rebalancing, i.e. $\hat{b}_{t-1,i} - b_{t,i}w_{t-1} > 0$, while the purchase occurs when $b_{t,i}w_{t-1} - \hat{b}_{t-1,i} > 0$. Clearly, following the common fact that net proportion after transaction costs and transaction costs always sum to 1, we have

$$1 = \underbrace{w_{t-1}}_{\text{net proportion}} + \underbrace{\gamma_s \sum_{i=1}^m \left(\hat{b}_{t-1,i} - b_{t,i}w_{t-1}\right)^+ + \gamma_b \sum_{i=1}^m \left(b_{t,i}w_{t-1} - \hat{b}_{t-1,i}\right)^+}_{\text{transaction costs incurred in sales and purchases}}. \quad (1)$$

Following the conventions (Bauer and Kohavi 1999, Albeverio *et al*. 2001, Györfi and Vajda 2008), we assume $\gamma_s = \gamma_b = \gamma \in [0, 1]$†, and rewrite the equation as,

$$1 = w_{t-1} + \gamma \left\|\hat{\mathbf{b}}_{t-1} - \mathbf{b}_t w_{t-1}\right\|_1. \quad (2)$$

We thus treat the net proportion after transaction costs incurred as a function of two consecutive portfolios and last price relative vector, i.e. $w_{t-1} = w(\mathbf{b}_t, \mathbf{b}_{t-1}, \mathbf{x}_{t-1})$. Note that $w_{t-1}$ can be efficiently solved via any optimization toolbox.

Thus, after rebalancing, the remaining capital becomes $S_{t-1} \times w_{t-1}$, where $S_{t-1}$ denotes the capital at the end of period $t-1$. During the period, the allocation of $\mathbf{b}_t$ changes the capital by a factor of $\mathbf{b}_t^\top \mathbf{x}_t = \sum_{i=1}^m b_{t,i}x_{t,i}$. In sum, for period $t$, the portfolio's capital changes from $S_{t-1}$ to $S_{t-1} \times w_{t-1} \times \left(\mathbf{b}_t^\top \mathbf{x}_t\right)$. As we re-invest, the capital changes multiplicatively, i.e. the portfolio wealth at the end of period $n$ can be expressed as,

$$S_n^{w(\cdot)} = S_0 \prod_{t=1}^n \left[(\mathbf{b}_t \cdot \mathbf{x}_t) \times w(\mathbf{b}_t, \mathbf{b}_{t-1}, \mathbf{x}_{t-1})\right], \quad (3)$$

where $S_0$ is usually set to 1 for convenience. Besides the typical cumulative wealth, traders also interest in the change of portfolio weights, i.e. the *average turnover* over the period,

$$AT_n = \frac{1}{2n}\sum_{t=1}^n \left\|\hat{\mathbf{b}}_{t-1} - \mathbf{b}_t w_{t-1}\right\|_1 = \frac{1}{2n}\sum_{t=1}^n \frac{1 - w_{t-1}}{\gamma}.$$

As a summary, protocol 1 outlines the problem formulation, and also backtests any proposed OLPS algorithm, which is abbreviated to ALG in the protocol.

†Note that this assumption may be released without considering the following derivations.

**Protocol 1:** Online Portfolio Selection with Transaction Costs.

**Input**: ALG: An OLPS algorithm; $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$: Sequence of price relative vectors.

**Output**: $S_n$: Final cumulative wealth.

1 **begin**

2 Initialize variables: $\mathbf{b}_1 = \left(\frac{1}{m}, \ldots, \frac{1}{m}\right)$, $\hat{\mathbf{b}}_0 = (0, \ldots, 0)$, $S_0 = 1$;

3 **for** $t = 1, \ldots, n$ **do**

4 Calculate next portfolio: $\mathbf{b}_t = \text{ALG}(\cdot)$;

5 Rebalance the portfolio from $\hat{\mathbf{b}}_{t-1}$ to $\mathbf{b}_t$;

6 **begin**

7 Solve the net proportion of wealth after transaction costs:

$$1 = w_{t-1} + \gamma \|\hat{\mathbf{b}}_{t-1} - \mathbf{b}_t w_{t-1}\|_1$$

8 Update the wealth after transaction costs: $S' = S_{t-1} \times w_{t-1}$;

9 **end**

10 Receive market price relatives: $\mathbf{x}_t = (x_{t,1}, \ldots, x_{t,m})$;

11 Update the cumulative return: $S_t = S' \times \left(\mathbf{b}_t^\top \mathbf{x}_t\right)$;

12 Record current allocation $\hat{\mathbf{b}}_t$: $\hat{b}_{t,i} = \frac{b_{t,i} \times x_{t,i}}{\mathbf{b}_t \cdot \mathbf{x}_t}$, $i = 1, \ldots, m$;

13 **end**

14 **end**

### 2.2. *Discussions*

In the above model, we follow the conventions and make several general assumptions. Firstly, we assume **proportional** transaction costs on both purchases and sales. Note that in the research of OLPS (Li and Hoi 2014), one common assumption is zero transaction cost. It is widely known that the transaction costs include various components, such as commissions, bid-ask spread and market impact. However, the bid/ask spread and market impact are related to the security's market microstructure (Grinold and Kahn 1999), which is beyond the scope of this study. We therefore consider one typical scenario in practice† and research (Bauer and Kohavi 1999, Albeverio *et al.* 2001, Györfi and Vajda 2008), or the proportional transaction costs. Note that other types of fee structure are also adopted by certain brokers.‡ Moreover, in some markets, transaction costs may be related to stocks, i.e. different stocks may have different rates of transaction costs. For example, the brokers may charge more fees on small and illiquid stocks, and charge less on large and liquid stocks. In this way, our model may incorrectly estimate transaction costs, leading to biased portfolio allocation. Finally, to simplify the research, we also equalize the rates of purchases and sales, which is also widely adopted by related research (Bauer and Kohavi 1999, Albeverio *et al.* 2001, Györfi and Vajda 2008). Note that adopting different rates of purchases and sales are straightforward for our backtest model, i.e. we could directly obtain the net proportion after transaction costs using equation (1) rather than equation (2). Secondly, we assume each asset is arbitrarily divisible, and one can buy and sell required quantities at the last closing price of any given trading period. Thirdly, we assume market behaviour is not affected by any trading strategy.

Note that the above three assumptions are widely used in finance (Cover 1991, Helmbold *et al.* 1998, Borodin *et al.* 2004, DeMiguel *et al.* 2009, 2014, Tu and Zhou 2011, Li *et al.* 2012, 2015, Shen *et al.* 2015, Shen and Wang 2016, 2017) and their implications have been thoroughly discussed (e.g. see Borodin *et al.* 2004, Li *et al.* 2012). In the empirical evaluations, we will follow these assumptions and compare the proposed algorithms with existing algorithms. Bearing in mind that we have made these assumptions throughout this article, even though we can further release them. However, unless we field-test the algorithms in real markets, it is almost impossible to model and implement following the full real market conditions.

### 2.3. *Analysis of transaction costs*

Now we decompose the protocol and analyse the sources related to transaction costs. It is worth special attention that the allocation decision is made by a portfolio manager (or ALG in the protocol 1), but transaction costs are imposed by financial markets. In particular, one iteration of (portfolio) trading process can be decomposed into two separate components. The first component is to decide a portfolio for next period, which is shown in Line 4 of the protocol. The second component is to rebalance the portfolio in the markets, during which market imposes transaction costs, as shown in Line 7 and 8. In the above model, the second component is noncontrollable, while the first is adjustable. Thus, to control (or reduce) the transaction costs imposed by the market,§ the only way is *to choose efficient portfolios that are expected to incur less transaction costs*.

## 3. Related work

Machine learning techniques have been widely used in quantitative finance (Fabozzi *et al.* 2007, Creamer and Freund 2010), including online portfolio selection (Cover 1991, Helmbold *et al.* 1998, Albeverio *et al.* 2001, Agarwal *et al.* 2006, Györfi *et al.* 2006, Tsagaris *et al.* 2012, Li *et al.* 2012, 2015). Although diverse in formulations, these algorithms' main underlying idea is to implicitly or explicitly estimate the distribution of next price relatives (vector), and then maximize the portfolio's expected log return based on the distribution (Kelly 1956, Thorp 1971, Laureti *et al.* 2010, Maclean *et al.* 2010),

$$\mathbf{b}_{t+1} = \arg\max_{\mathbf{b} \in \Delta_m} \mathbb{E}\left[\log\left(\mathbf{b} \cdot \tilde{\mathbf{x}}\right)\right], \qquad (4)$$

where $\tilde{\mathbf{x}}$ denotes the estimated price relative vector. Table 1 summarizes most existing OLPS formulations via the above framework. Some other algorithms also have close connection with the framework but cannot be summarized here, including Universal Portfolios (UP) (Cover 1991), and AntiCorrelation (Anticor) (Borodin *et al.* 2004), etc. For a complete survey of the online portfolio selection algorithms, please refer to Li and Hoi (2014).

On the one hand, these algorithms are successful either theoretically (Cover 1991, Helmbold *et al.* 1998,

†For example, Interactive Broker (www.interactivebrokers.com) charges a fixed set percentage of trade value.

‡For example, Interactive Brokers also charges a fixed commission per share, which is not proportional.

§In imperfect markets, optimized execution (Nevmyvaka *et al.* 2006) can slice large orders into small one so as to limit the impact of large orders, and thus reducing (implicit) transaction costs.

Table 1. A summary of existing algorithms based on the framework of equation (4). R(·) or R(·, ·) denotes the regularization term, and $C_t^j$, $j = 1, 2, 3$ denote the similarity sets obtained with different criteria in their respective studies. $p_i$ denotes the implicit probability of prediction. Note that in some studies (e.g. EG, PAMR, CWMR and OLMAR), their probabilities of predicted price relative can be regarded as 100%. In these cases, they are deterministic, without stochasticity.

| Categories | Methods | Formulations ($\mathbf{b}_{t+1} = \arg\max_{\mathbf{b}\in\Delta_m}$) | Prediction ($\tilde{\mathbf{x}}^i$) | Prob. ($p_i$) |
|---|---|---|---|---|
| In hindsight | BCRP (Cover 1991) | $\sum_{i=1}^n \frac{1}{n} \log \mathbf{b} \cdot \mathbf{x}_i$ | $\mathbf{x}_i, i = 1, \ldots, n$ | $1/n$ |
| Follow the Winner | EG (Helmbold *et al.* 1998) | $\log \mathbf{b} \cdot \mathbf{x}_t - \lambda R(\mathbf{b}, \mathbf{b}_t)$ | $\mathbf{x}_t$ | 100% |
| | SCRP (Granger and Sin 2000) | $\sum_{i=1}^t \frac{1}{t} \log \mathbf{b} \cdot \mathbf{x}_i$ | $\mathbf{x}_i, i = 1, \ldots, t$ | $1/t$ |
| | ONS (Agarwal *et al.* 2006) | $\sum_{i=1}^t \frac{1}{t} \log \mathbf{b} \cdot \mathbf{x}_i - \lambda R(\mathbf{b})$ | $\mathbf{x}_i, i = 1, \ldots, t$ | $1/t$ |
| Follow the Loser | PAMR (Li *et al.* 2012) | $\log \mathbf{b} \cdot \frac{1}{\mathbf{x}_t} - \lambda R(\mathbf{b}, \mathbf{b}_t)$ | $1/\mathbf{x}_t$ | 100% |
| | CWMR (Li *et al.* 2013) | $\text{Prob}(\mathbf{b} \cdot \frac{1}{\mathbf{x}_t}) - \lambda R(\mathbf{b})$ | $1/\mathbf{x}_t$ | 100% |
| | OLMAR (Li *et al.* 2015) | $\log(\mathbf{b} \cdot \tilde{\mathbf{x}}_t) - \lambda R(\mathbf{b})$ | Equation (1) in the paper | 100% |
| Pattern Matching | $\text{B}^{\text{K}}$ (Györfi *et al.* 2006) | $-\sum_{i\in C_t^1} \frac{1}{\lvert C_t^1\rvert} \log \mathbf{b} \cdot \mathbf{x}_i$ | $\mathbf{x}_i, i \in C_t^1$ | $1/\lvert C_t^1\rvert$ |
| | $\text{B}^{\text{NN}}$ (Györfi *et al.* 2008) | $-\sum_{i\in C_t^2} \frac{1}{\lvert C_t^2\rvert} \log \mathbf{b} \cdot \mathbf{x}_i$ | $\mathbf{x}_i, i \in C_t^2$ | $1/\lvert C_t^2\rvert$ |
| | CORN (Li *et al.* 2011) | $-\sum_{i\in C_t^3} \frac{1}{\lvert C_t^3\rvert} \log \mathbf{b} \cdot \mathbf{x}_i$ | $\mathbf{x}_i, i \in C_t^3$ | $1/\lvert C_t^3\rvert$ |

Agarwal *et al.* 2006) or empirically (Györfi *et al.* 2006, Li *et al.* 2012, 2015), but their formulations (or the decision-making step) ignore the transaction costs issue, which all portfolio managers have to pay in the next portfolio rebalance step. Ignoring this unavoidable aspect in the decision-making step will result in sub-optimal performance, as transaction costs in real market are always non-zero. In particular, in case of zero transaction costs, the decision-making component in equation (4) and the portfolio rebalance component (equations (2) and (3)) are consistent, i.e. $\gamma = 0$ results in $w_{t-1} = 1$. However, in real market the transaction cost rate ($\lambda$) is non-zero, thus the two components have discrepancies, resulting in inefficient decisions or lower performance. On the other hand, these algorithms' performance degrades significantly in backtests when the transaction costs are non-zero. Such decision inefficiency and observed performance degradation motivate us to propose a new framework for non-zero transaction costs scenarios.

To the best of our knowledge, only a few existing strategies have been proposed to consider transaction costs for online portfolio selection. The first extension is based on the BCRP benchmark. Bauer and Kohavi (1999) proved that Cover's Universal Portfolios is still universal when the market imposes proportional transaction costs. However, it does not take the transaction costs into decision-making process, thus failing to solve the transaction costs issue. Albeverio *et al.* (2001) proposed a new strategy for online portfolio selection with transaction costs. Its main idea is to maximize the expected return and minimize the distance between consequent portfolios. Similar to Helmbold *et al.* (1998), the authors employed relative entropy as the distance measure (Albeverio *et al.* 2001, equation (3.4)). Closed-form solutions are obtained by solving its optimization problem. The new strategy is equipped with a new prediction method based on 'cross rate'. The algorithms empirically work well on the portfolios of pair stocks. However, the strategy is constrained to work with a pair of two stocks, and has not been extended to more than two stocks. We thus will compare their portfolio generation part with our proposed method in sections 4.2.1 and 4.3, and ignore the comparison of their empirical performance.† Györfi and Vajda (2008) directly incorporated equation (2) to the decision formulas of the Pattern Matching-based approaches. Though straightforward, the decision formulas is hard (or unable for us) to solve by current techniques. Ormos and Urbán (2013) empirically analysed the performance of the Pattern Matching-based algorithms. Das *et al.* (2013) considered the transaction costs issue for the GP algorithm (Helmbold *et al.* 1997), which is an variant of the EG algorithm (Helmbold *et al.* 1998). However, it considers the L1 norm of the difference between two decision portfolios, which is conceptually different from our proposed method in both formulations and algorithms, as shown in sections 4.2.1 and 4.3.

## 4. Transaction costs optimization

### *4.1. Motivation*

Although maximizing the expected log return (or/and risk) is heavily investigated in literatures (Li and Hoi 2014), the unavoidable transaction costs issue is seldomly discussed, especially on the mean reversion based online portfolio selection algorithms (Borodin *et al.* 2004, Li *et al.* 2012, 2015). To motivate our approach, we first analyse the net proportion of wealth after transaction costs imposed by the market, i.e. equation (2). Getting rid of $w_{t-1}$ within the L1 norm, we can bound $w_{t-1}$ as in proposition 4.1, which shows the relationship between $w_{t-1}$ and $\left\|\hat{\mathbf{b}}_{t-1} - \mathbf{b}_t\right\|_1$.

†We also checked the authors' publications and found no clues on how to extend the strategy. We may extend their strategy to more than two stocks, but it is far beyond the scope of this article.

PROPOSITION 4.1 *The net proportion of wealth after transaction costs is bounded as*

$$\frac{1-\gamma}{1-\gamma+\gamma\left\|\hat{\mathbf{b}}_{t-1}-\mathbf{b}_t\right\|_1} \le w_{t-1} \le \frac{1+\gamma}{1+\gamma+\gamma\left\|\hat{\mathbf{b}}_{t-1}-\mathbf{b}_t\right\|_1}.$$

*Proof.* To get the lower and upper bounds for $w_{t-1}$, we have to get rid of $w_{t-1}$ within the L1-norm. Firstly, utilizing the norm inequality, we can get,

$$\begin{aligned}1 &= w_{t-1}+\gamma\|\hat{\mathbf{b}}_{t-1}-\mathbf{b}_t w_{t-1}\|_1\\ &= w_{t-1}+\gamma\|w_{t-1}\hat{\mathbf{b}}_{t-1}+(1-w_{t-1})\,\hat{\mathbf{b}}_{t-1}-\mathbf{b}_t w_{t-1}\|_1\\ &\le w_{t-1}+\gamma w_{t-1}\|\hat{\mathbf{b}}_{t-1}-\mathbf{b}_t\|_1+\gamma\,(1-w_{t-1})\,\|\hat{\mathbf{b}}_{t-1}\|_1\\ &= w_{t-1}+\gamma w_{t-1}\|\hat{\mathbf{b}}_{t-1}-\mathbf{b}_t\|_1+\gamma\,(1-w_{t-1})\,,\end{aligned}$$

and then the lower bound of $w_{t-1}$ is,

$$w_{t-1} \ge \frac{1-\gamma}{1-\gamma+\gamma\left\|\hat{\mathbf{b}}_{t-1}-\mathbf{b}_t\right\|_1}.$$

Similarly, we can derive its upper bound as follows,

$$\begin{aligned}1 &= w_{t-1}+\gamma\|\hat{\mathbf{b}}_{t-1}-\mathbf{b}_t w_{t-1}\|_1\\ &\ge w_{t-1}+\gamma\|\hat{\mathbf{b}}_{t-1}-\mathbf{b}_t w_{t-1}\\ &\quad+\hat{\mathbf{b}}_{t-1}\,(w_{t-1}-1)\,\|_1-\gamma\|\hat{\mathbf{b}}_{t-1}\,(w_{t-1}-1)\,\|_1\\ &= w_{t-1}+\gamma w_{t-1}\left\|\hat{\mathbf{b}}_{t-1}-\mathbf{b}_t\right\|_1-\gamma\,(1-w_{t-1})\,,\end{aligned}$$

and then the upper bound of $w_{t-1}$ is

$$w_{t-1} \le \frac{1+\gamma}{1+\gamma+\gamma\|\hat{\mathbf{b}}_{t-1}-\mathbf{b}_t\|_1}.$$

In summary, we can bound the net proportion of wealth after transaction costs as,

$$\frac{1-\gamma}{1-\gamma+\gamma\|\hat{\mathbf{b}}_{t-1}-\mathbf{b}_t\|_1} \le w_{t-1} \le \frac{1+\gamma}{1+\gamma+\gamma\|\hat{\mathbf{b}}_{t-1}-\mathbf{b}_t\|_1}.$$

□

Obviously, both the upper bound and lower bound are inversely related to $\|\hat{\mathbf{b}}_{t-1}-\mathbf{b}_t\|_1$, i.e. the smaller the L1 norm, the larger the upper/lower bound and thus the value of $w_{t-1}$. Therefore, to obtain a high net proportion of wealth after transaction costs, the proposition motivates us to minimize $\|\hat{\mathbf{b}}_{t-1}-\mathbf{b}_t\|_1$. Note that two cases will lead to $w_{t-1}=1$, i.e. $\gamma=0$ or $\mathbf{b}_t=\hat{\mathbf{b}}_{t-1}$, the former of which denotes zero transaction cost and the latter means no rebalancing. It is worth distinguishing that the first case means that market imposes no transaction costs, while the second is based on the manager's decision. In other words, we have no way to control the transaction costs imposed by the market (the first case), but we can actively decide portfolios considering transaction costs (the second case). As we discussed before, this distinction is seldomly addressed by any previous study. Moreover, note that $\|\hat{\mathbf{b}}_{t-1}-\mathbf{b}_t\|_1$ is conceptually different from $\|\mathbf{b}_{t-1}-\mathbf{b}_t\|_1$ (Das *et al.* 2013). The latter considers the difference of consecutive decision portfolios, which does not precisely reflect the impact of transaction costs.

To handle the transaction costs issue for online portfolio selection, we thus propose a new framework called *Transaction Costs Optimization* (TCO). On the one hand, we follow the line of previous research, and focus on maximizing the portfolio's expected log return. On the other hand, we promptly take transaction costs into account in the decision-making process. That is, we try to minimize the L1 norm of $\|\hat{\mathbf{b}}_{t-1}-\mathbf{b}_t\|_1$, which is connected to the incurred transaction costs as shown in proposition 4.1. Balancing between the two can effectively handle the transaction costs issue and significantly improve the performance when the market imposes non-zero transaction costs.

### *4.2. Formulation*

In this section, we formally formulate the proposed *Transaction Costs Optimization* (TCO) framework. In principle, this framework can be applied to any type of trading principles, including mean reversion, momentum, and pattern matching, etc. However, we restrict our discussion to the domain of mean reversion-based strategy (Li *et al.* 2012, 2013, 2015, as they achieve the state of the art performance.

Straightforwardly, to couple with the transaction costs, two folds of implications exist. One is to *maximize* portfolio's expected log return (Kelly 1956), which is the main principle underlying most existing algorithms. The other is to *minimize* $\|\hat{\mathbf{b}}_{t-1}-\mathbf{b}_t\|_1$, such that the proportion after transaction costs could be maximized according to proposition 4.1. Intuitively, we formulate the following framework.

**Transaction Costs Optimization**: Constrained version

$$\begin{aligned}\mathbf{b}_{t+1} = \arg\min -\overbrace{\mathbb{E}\left\{\log\mathbf{b}\cdot\tilde{\mathbf{x}}_{t+1}\right\}}^{\text{expected } log \text{ return}} +\lambda \overbrace{\left\|\mathbf{b}-\hat{\mathbf{b}}_t\right\|_1}^{\text{term to bound } w_t}\\ \text{s.t.}\quad \mathbf{b}\cdot\mathbf{1}=1,\quad \mathbf{b}\succeq 0,\end{aligned} \tag{5}$$

where $\lambda \ge 0$ is a trade-off parameter to balance the first term and the second term. Minimizing the first term refers to maximize the portfolio's expected log return. Minimizing the second term is equivalent to maximizing the net proportion after transaction costs. Balancing the two terms can effectively control the transaction costs to be incurred in the next portfolio rebalance, while maintaining satisfactory return. If $\lambda$ is small, the whole system concentrates more on obtaining higher expected return; if $\lambda$ is large, the system concentrates more on transaction costs.

Although the above optimization is convex, the non-negativity constraint is an unsolved issue (Helmbold *et al.* 1998). To ease its derivation (Bach *et al.* 2012), we combine the constraint that the portfolio summation equals one into the objective function and discard the non-negativity constraint at the moment. This leads to an unconstrained version, which economically allows shorting. Note that the coefficient of the constraint, $\theta$, will be eliminated with an endogenetic variable forcing the portfolio sums to one.

**Transaction Costs Optimization**: Unconstrained Version

$$\mathbf{b}_{t+1} = \arg\min_{\mathbf{b}\in\mathbb{R}^m} -\mathbb{E}\left\{\log \mathbf{b}\cdot\tilde{\mathbf{x}}_{t+1}\right\} + \theta\left(\mathbf{b}\cdot\mathbf{1}-1\right) + \lambda\left\|\mathbf{b}-\hat{\mathbf{b}}_t\right\|_1. \quad (6)$$

Later we will project the obtained portfolio onto the simplex domain (Duchi *et al.* 2008) such that the solution can be applied to the problem setting in section 2.

#### 4.2.1. Discussion.

The framework in equation (5) is in fact a general framework for the whole Kelly's approaches, including the mean reversion approaches. However, when converting it to the unconstrained version as in equation (6), we adopt the similar process in all the derivations of mean reversion based algorithms (Li *et al.* 2012, 2013, 2015). Economically, it allows increasing leverage at first, and then lowers the leverage by projecting portfolio back to the simplex domain.

We would also like to compare the TCO's formulation with FWGTC's formulations (Albeverio *et al.* 2001, equation (3.1)) and OLU's formulation (Das *et al.* 2013), which are conceptually different. First, FWGTC adopts a similar idea of balancing the expected return and the change of next portfolio from the latest allocation. On the one hand, to obtain an expected return, it predicts the next price relatives' ranking using a proposed 'cross rates' method (Albeverio *et al.* 2001, sections 4.3 and 4.4). One crucial drawback of the prediction method is that it can only handle a pair of two stocks, which is quite limited in real asset management. On the other hand, to constrain the unnecessary changes in the weights, it adopts a relative entropy function to measure the distance of weight change, i.e. $d_{re}\left(\mathbf{b},\hat{\mathbf{b}}_t\right) = \sum_{i=1}^{m}\left(b_{t+1,i}\log\frac{b_{t+1,i}}{\hat{b}_{t,i}}\right)$. Then the final formulation of the FWGTC equals

$$\mathbf{b}_{t+1} = \arg\max_{\mathbf{b}\in\Delta_m} \lambda F_W\left(\mathbf{b},\hat{x}_{t+1}\right) - d_{re}\left(\mathbf{b},\hat{\mathbf{b}}_t\right),$$

where $F_W^{(1)}\left(\mathbf{b},\hat{x}_{t+1}\right) = \mathbf{b}\hat{\mathbf{x}}_{t+1}$ and $F_W^{(2)}\left(\mathbf{b},\hat{x}_{t+1}\right) = \log\left(\hat{\mathbf{b}}_t\cdot\hat{\mathbf{x}}_{t+1}\right) + \frac{\hat{\mathbf{x}}_{t+1}\cdot\left(\mathbf{b}-\hat{\mathbf{b}}_t\right)}{\hat{b}_t\cdot\hat{\mathbf{x}}_{t+1}}$. As we can see from the equations, although its idea is similar to our TCO, they are different from the following two aspects. Conceptually, FWGTC concentrates on prediction for price relatives, while TCO focuses on proposing a framework for transaction costs that applied to all kinds of algorithms in the same category. Technically, TCO's L1 norm is consistent with the return calculation with proportional transaction costs, as in both TCO and FWGTC's models. However, FWGTC's distance function is different from the return calculation, which may diverge from the optimal portfolios.

Secondly, OLU also adopted L1 norm to constrain the change of the portfolio from last portfolio, but not the latest allocation, i.e. $\|\mathbf{b}-\mathbf{b}_t\|$. The TCO, on the other hand, adopts the L1 norm of the difference in the portfolio and the latest allocation, i.e. $\|\mathbf{b}-\hat{\mathbf{b}}_t\|$, where $\hat{b}_{t-1,i} = \frac{b_{t-1,i}x_{t-1,i}}{\mathbf{b}_{t-1}\cdot\mathbf{x}_{t-1}}$, $i = 1,\ldots,m$. The basic trading principle shows that the transaction costs are caused by the deviation from the target portfolio and current allocation, or $\hat{\mathbf{b}}_t$ rather than $\mathbf{b}_t$. While it enjoys theoretical convenience with the latter term (such as proving a bound), it does not follow the general trading principle. To better understand their difference, let us see an example. Let us consider the market with two assets, and stand at the beginning of day 1. Suppose we are running a uniform CRP strategy, that is, every portfolio vector is an equally weighted portfolio ($\mathbf{b}_t = (0.5, 0.5)$, $t = 1,\ldots,n$). The initial allocation is $\mathbf{b}_1 = (0.5, 0.5)$ and day 1's price relative is $\mathbf{x}_1 = (0.8, 1.2)$. At the end of day 1, the allocation becomes $\hat{\mathbf{b}}_1 = (0.4, 0.6)$. At the beginning of day 2, the portfolio will be rebalanced from current allocation ($\hat{\mathbf{b}}_1$, not $\mathbf{b}_1$) to the target allocation ($\mathbf{b}_2$). Here, note that the current allocation, rather than day 1's initial allocation, is meaningful for the rebalancing. Rebalancing to $\mathbf{b}_2 = (0.5, 0.5)$ will incur a transaction cost of $\gamma\|\mathbf{b}_2-\hat{\mathbf{b}}_1\|_1 = 0.2\gamma$. Although both formulations adopt L1 norm of difference between two portfolios, we adopt $\hat{\mathbf{b}}_1$ rather than OLU's $\mathbf{b}_1$. In OLU's cases, as $\mathbf{b}_2 = \mathbf{b}_1$, their model will incur **zero** transaction costs, which is obviously incorrect.

We also want to discuss the role of the L1 norm, which is actually a typical regularization term. However, our intention is not to control the complex of the model (Agarwal *et al.* 2006), or record all historical information (Helmbold *et al.* 1998, Li *et al.* 2012). Motivated to maximize the net proportion of wealth after transaction costs, the L1 term is used to control the effect of transaction costs. In fact, one particular strategy is when $\lambda = +\infty$, then the TCO model degrades to,

$$\mathbf{b}_{t+1} = \arg\min_{\mathbf{b}\in\Delta_m}\left\|\mathbf{b}-\hat{\mathbf{b}}_t\right\|_1 \Rightarrow \mathbf{b}_{t+1} = \hat{\mathbf{b}}_t.$$

Note that the strategy is the *Buy And Hold* (BAH) strategy, which incurs the least transaction costs. Moreover, when $\lambda = 0$, the TCO model is the same as equation (4), which does not consider transaction costs in the decision-making process.

Similar to table 1, equation (5) can be instantiated to different trading schema, in which most existing ideas can be adopted. As shown in table 2, we further adopt some representative predictions in the framework of TCO. In this article, we focus on TCO-MR as mean reversion-based strategies achieved the state of the art performance.

### 4.3. *Algorithm*

In this section, we solve the TCO formulation via Proximal Gradient Descent (Boyd and Vandenberghe 2004, Bach *et al.* 2012) and derive the proposed framework. Proposition 4.2 illustrates the solution of the unconstrained optimization equation (6).

Proposition 4.2 *The solution to the unconstrained TCO formulation (equation (6)) is*

$$\tilde{\mathbf{b}}_{t+\frac{1}{2}} = \eta_t\left(\mathbb{E}\left\{\frac{\tilde{\mathbf{x}}_{t+1}}{\hat{\mathbf{b}}_t\cdot\tilde{\mathbf{x}}_{t+1}}\right\} - \frac{1}{m}\mathbf{1}\cdot\mathbb{E}\left\{\frac{\tilde{\mathbf{x}}_{t+1}}{\hat{\mathbf{b}}_t\cdot\tilde{\mathbf{x}}_{t+1}}\right\}\right),$$

$$\mathbf{b}_{t+1} = \hat{\mathbf{b}}_t + sign\left(\tilde{\mathbf{b}}_{t+\frac{1}{2}}\right)\left[\left|\tilde{\mathbf{b}}_{t+\frac{1}{2}}\right| - \lambda\eta_{t+\frac{1}{2}}\right]_+,$$

*where* $[v]_+ = \max(0, v)$, $sign(v)$ *denotes the sign of* $v$, *and* $\eta_t$, $\eta_{t+\frac{1}{2}}$, $\lambda$ *are parameters to control the learning progress.*

*Proof.* Firstly, we linearize the log function around the current allocation of $\hat{\mathbf{b}}_t$ and eliminate the constant terms, and rewrite equation (6) into the following form,

Table 2. Some representative formulations derived from the TCO framework.

| Categories | Methods | Formulations ($\mathbf{b}_{t+1} = \arg\min_{\mathbf{b}\in\Delta_m}$) | Prediction ($\tilde{\mathbf{x}}^i_{t+1}$) | Prob. ($p_i$) |
|---|---|---|---|---|
| Follow the Winner | TCO-EG | $-\log \mathbf{b}\cdot\mathbf{x}_t + \lambda\|\mathbf{b}-\hat{\mathbf{b}}_t\|_1$ | $\mathbf{x}_t$ | 100% |
| | TCO-SCRP | $-\sum_{i=1}^{t}\frac{1}{t}\log \mathbf{b}\cdot\mathbf{x}_i + \lambda\|\mathbf{b}-\hat{\mathbf{b}}_t\|_1$ | $\mathbf{x}_i, i=1,\ldots,t$ | $1/t$ |
| Follow the Loser | TCO-MR | $-\log \mathbf{b}\cdot\frac{1}{\mathbf{x}_t} + \lambda\|\mathbf{b}-\hat{\mathbf{b}}_t\|_1$ | $1/\mathbf{x}_t$ | 100% |
| Pattern Matching based | TCO-NP | $-\sum_{i\in C_t}\frac{1}{\|C_t\|}\log \mathbf{b}\cdot\mathbf{x}_i + \lambda\|\mathbf{b}-\hat{\mathbf{b}}_t\|_1$ | $\mathbf{x}_i, i\in C_t$ | $1/\|C_t\|$ |

$$\mathbf{b}_{t+1} = \arg\min_{\mathbf{b}\in\mathbb{R}^m} -\mathbb{E}\left\{\frac{\tilde{\mathbf{x}}_{t+1}}{\hat{\mathbf{b}}_t\cdot\tilde{\mathbf{x}}_{t+1}}\cdot\left(\mathbf{b}-\hat{\mathbf{b}}_t\right)\right\} + \frac{1}{2\eta_t}\left\|\mathbf{b}-\hat{\mathbf{b}}_t\right\|_2^2 + \theta\left(\mathbf{b}\cdot\mathbf{1}-1\right) + \lambda\left\|\mathbf{b}-\hat{\mathbf{b}}_t\right\|_1,$$

where $\eta_t$ is a parameter to bound the linear approximation (Bach *et al*. 2012). To simplify the L1 norm, let $\tilde{\mathbf{b}} = \mathbf{b}-\hat{\mathbf{b}}_t$, then $\mathbf{b} = \tilde{\mathbf{b}} + \hat{\mathbf{b}}_t$, and equivalently,

$$\begin{aligned}\tilde{\mathbf{b}}_{t+1} &= \arg\min_{\tilde{\mathbf{b}}\in\mathbb{R}^m} -\mathbb{E}\left\{\frac{\tilde{\mathbf{x}}_{t+1}}{\hat{\mathbf{b}}_t\cdot\tilde{\mathbf{x}}_{t+1}}\cdot\tilde{\mathbf{b}}\right\} + \frac{1}{2\eta_t}\left\|\tilde{\mathbf{b}}\right\|_2^2 \\ &\quad + \theta\left(\left(\tilde{\mathbf{b}}+\hat{\mathbf{b}}_t\right)\cdot\mathbf{1}-1\right) + \lambda\left\|\tilde{\mathbf{b}}\right\|_1 \\ &= \arg\min_{\tilde{\mathbf{b}}} f\left(\tilde{\mathbf{b}}\right) + \lambda\Omega\left(\tilde{\mathbf{b}}\right),\end{aligned}$$

which satisfies the formulation in Duchi and Singer (2009) and is straightforward to solve. In particular, we decompose the above optimization into two steps,

$$\begin{cases}\tilde{\mathbf{b}}_{t+\frac{1}{2}} = \arg\min_{\tilde{\mathbf{b}}} f\left(\tilde{\mathbf{b}}\right) & \text{(7a)}\\ \tilde{\mathbf{b}}_{t+1} = \arg\min_{\tilde{\mathbf{b}}} \frac{1}{2}\left\|\tilde{\mathbf{b}}-\tilde{\mathbf{b}}_{t+\frac{1}{2}}\right\|_2^2 + \eta_{t+\frac{1}{2}}\Omega\left(\tilde{\mathbf{b}}\right) & \text{(7b)}\end{cases}$$

Setting the derivative of equation (7a) to zero,

$$\frac{\partial f}{\partial\tilde{\mathbf{b}}} = -\mathbb{E}\left\{\frac{\tilde{\mathbf{x}}_{t+1}}{\hat{\mathbf{b}}_t\cdot\tilde{\mathbf{x}}_{t+1}}\right\} + \frac{1}{\eta_t}\tilde{\mathbf{b}} + \theta\mathbf{1} = 0,$$

we can get,

$$\tilde{\mathbf{b}} = \eta_t\left(\mathbb{E}\left\{\frac{\tilde{\mathbf{x}}_{t+1}}{\hat{\mathbf{b}}_t\cdot\tilde{\mathbf{x}}_{t+1}}\right\} - \theta\mathbf{1}\right).$$

Multiplying both sides by $\mathbf{1}$ and utilizing the property of $\tilde{\mathbf{b}}\cdot\mathbf{1} = 0$, we can obtain $\theta$,

$$\begin{aligned}0 &= \eta_t\left(\mathbf{1}\cdot\mathbb{E}\left\{\frac{\tilde{\mathbf{x}}_{t+1}}{\hat{\mathbf{b}}_t\cdot\tilde{\mathbf{x}}_{t+1}}\right\} - \theta m\right) \longrightarrow \theta \\ &= \frac{1}{m}\mathbf{1}\cdot\mathbb{E}\left\{\frac{\tilde{\mathbf{x}}_{t+1}}{\hat{\mathbf{b}}_t\cdot\tilde{\mathbf{x}}_{t+1}}\right\}.\end{aligned}$$

Subsequently, we can derive the update formula of $\tilde{\mathbf{b}}_{t+\frac{1}{2}}$,

$$\tilde{\mathbf{b}}_{t+\frac{1}{2}} = \eta_t\left(\mathbb{E}\left\{\frac{\tilde{\mathbf{x}}_{t+1}}{\hat{\mathbf{b}}_t\cdot\tilde{\mathbf{x}}_{t+1}}\right\} - \frac{1}{m}\mathbf{1}\cdot\mathbb{E}\left\{\frac{\tilde{\mathbf{x}}_{t+1}}{\hat{\mathbf{b}}_t\cdot\tilde{\mathbf{x}}_{t+1}}\right\}\mathbf{1}\right).$$

On the other hand, solving equation (7b) results in a closed form update (Duchi and Singer 2009),

$$\mathbf{b}_{t+1} = \hat{\mathbf{b}}_t + sign\left(\tilde{\mathbf{b}}_{t+\frac{1}{2}}\right)\left[\left|\tilde{\mathbf{b}}_{t+\frac{1}{2}}\right| - \lambda\eta_{t+\frac{1}{2}}\right]_+,$$

where $[\mathbf{v}]_+ = \max(0,\mathbf{v})$ and $sign(\mathbf{v})$ returns the sign of each element in $\mathbf{v}$. □

Based on the proposition, we can formulate the TCO framework. However, one problem still unsolved is the prediction schema of $\tilde{\mathbf{x}}_{t+1}$. As shown in section 3, there are several prediction schema employed by existing algorithms. In this article, we derive two specific algorithms following two mean reversion predictions, whose implicit or explicit assumptions are summarized in table 3. In particular, we derive the algorithm based on the prediction of $\tilde{\mathbf{x}}^i_{t+1} = f\left(\mathbf{x}_1^t\right)$ and corresponding probability of 100%. Then the expected value of $\frac{\tilde{\mathbf{x}}_{t+1}}{\hat{\mathbf{b}}_t\cdot\tilde{\mathbf{x}}_{t+1}}$ equals $\frac{f(\mathbf{x}_1^t)}{\hat{\mathbf{b}}_t\cdot f(\mathbf{x}_1^t)}$. To this end, we can obtain proposition 4.3.

PROPOSITION 4.3 *The solution to the unconstrained TCO formulation in equation* (6) *with existing mean reversion prediction, i.e.* $\tilde{\mathbf{x}}_{t+1} = f\left(\mathbf{x}_1^t\right)$ *and a probability of* 100%*, is*

$$\begin{aligned}\tilde{\mathbf{b}}_{t+\frac{1}{2}} &= \eta_t\left(\frac{f\left(\mathbf{x}_1^t\right)}{\hat{\mathbf{b}}_t\cdot f\left(\mathbf{x}_1^t\right)} - \frac{1}{m}\mathbf{1}\left(\mathbf{1}\cdot\frac{f\left(\mathbf{x}_1^t\right)}{\hat{\mathbf{b}}_t\cdot f\left(\mathbf{x}_1^t\right)}\right)\right),\\ \mathbf{b}_{t+1} &= \hat{\mathbf{b}}_t + sign\left(\tilde{\mathbf{b}}_{t+\frac{1}{2}}\right)\left[\left|\tilde{\mathbf{b}}_{t+\frac{1}{2}}\right| - \lambda\eta_{t+\frac{1}{2}}\right]_+,\end{aligned}$$

*where* $[v]_+ = \max(0,v)$ *and* $sign(v)$ *returns the sign of* $v$.

*Proof.* We omit the derivation, which is straightforward. □

To now, we can derive the proposed online portfolio selection algorithms named 'Transaction Costs Optimization' (TCO) in algorithm 2. The whole trading simulation procedure is illustrated in protocol 1 by replacing ALG with TCO. For simplicity, we let $\lambda = \lambda\eta_{t+\frac{1}{2}}$ and adopt a fixed parameter of $\eta_t = \eta$ for all iterations. Ignoring the examples, the algorithm represents a framework for online portfolio selection with (zero or non-zero) transaction costs. The two examples, named TCO1 and TCO2, respectively, show two specifications whose prediction of price relative vector follows the state of the art mean reversion principle. Moreover, to conform to the typical constrained version of online portfolio selection, we follow the existing techniques (Li *et al*. 2012, 2015) and project the final portfolio to the simplex domain.†

†We adopt the *lsqlin* function in Matlab optimization toolbox.

**Algorithm 2:** Transaction Costs Optimization: TCO($\hat{\mathbf{b}}_t, \mathbf{x}_1^t, \lambda, \eta$).

**Input**: $\hat{\mathbf{b}}_t$: last price adjusted portfolio; $\mathbf{x}_1^t$: historical price relatives; $\lambda$: trade-off parameter; $\eta$: smoothing parameter; $w$: window size for TCO-2.

**Output**: $\mathbf{b}_{t+1}$: next portfolio.

**begin**

Estimate next price relative vector: $\tilde{\mathbf{x}}_{t+1} = f\left(\mathbf{x}_1^t\right)$

$$\text{Examples} \begin{cases} \text{TCO-1}: & f_1 = \frac{1}{\mathbf{x}_t} \\ \text{TCO-2}: & f_2 = \frac{1}{w}\left(1 + \frac{1}{\mathbf{x}_t} + \cdots + \frac{1}{\bigodot_{i=0}^{w-2} \mathbf{x}_{t-i}}\right) \end{cases}$$

Calculate variables: $\mathbf{v}_t = \mathbb{E}\left\{\frac{\tilde{\mathbf{x}}_{t+1}}{\hat{\mathbf{b}}_t \cdot \tilde{\mathbf{x}}_{t+1}}\right\}$, $\bar{v}_t = \frac{\mathbf{1} \cdot \mathbf{v}_t}{m}$

$$\text{Examples} \begin{cases} \text{TCO-1}: & \mathbf{v}_t = \frac{f_1}{\hat{\mathbf{b}}_t \cdot f_1} \\ \text{TCO-2}: & \mathbf{v}_t = \frac{f_2}{\hat{\mathbf{b}}_t \cdot f_2} \end{cases}$$

Update portfolio:

$$\tilde{\mathbf{b}}_{t+\frac{1}{2}} = \eta\left(\mathbf{v}_t - \bar{v}_t \mathbf{1}\right)$$

$$\mathbf{b}_{t+1} = \hat{\mathbf{b}}_t + sign\left(\tilde{\mathbf{b}}_{t+\frac{1}{2}}\right)\left[\left|\tilde{\mathbf{b}}_{t+\frac{1}{2}}\right| - \lambda\right]_+$$

Normalize portfolio:

$$\mathbf{b}_{t+1} = \underset{\mathbf{b} \in \Delta_m}{\arg\min} \left\|\mathbf{b} - \mathbf{b}_{t+1}\right\|^2$$

**end**

**4.3.1. Discussions.** Note that there also exist some other types of predictions, such as Pattern Matching-based predictions and Follow the Winner predictions. However, their respective studies solved their optimizations via different techniques from ours in proposition 4.2. It is feasible to incorporate their predictions to the TCO framework, and derive their algorithms accordingly, which however is beyond the scope of this article.

Besides formulations, TCO's derivation technique is also different from that of FWGTC (Albeverio *et al.* 2001, equations (3.9) and (3.10)) and OLU (Das *et al.* 2013). In particular, we derive the TCO's algorithms using Proximal Gradient Descent (Boyd and Vandenberghe 2004, Bach *et al.* 2012), and FWGTC and OLU derived their algorithms using the Lagrange methods and the Alternating Direction Method of Multipliers (ADDM) (Boyd *et al.* 2011), respectively. In addition, TCO's update formulas are additive, while FWGTC's are multiplicative.

### 4.4. *Analysis*

Now we analyse the update formula such that we can better understand its underlying mechanism, and compare it with *Percentage of Portfolio Rebalancing* (PPR) (Institute 2013), which is a rebalance strategy used by various practitioners in case of non-zero transaction costs.

Let us first analyse $\mathbf{v}_t = \mathbb{E}\left\{\frac{\tilde{\mathbf{x}}_{t+1}}{\hat{\mathbf{b}}_t \cdot \tilde{\mathbf{x}}_{t+1}}\right\}$ in proposition 4.2. Without considering the expectation, the denominator is the return without rebalancing, while the numerator is the price relative vector. We thus can view it as a return-adjusted price relative vector. Then, the following step ($\tilde{\mathbf{b}}_{t+\frac{1}{2}} = \eta_t\left(\mathbf{v}_t - \frac{\mathbf{v}_t \cdot \mathbf{1}}{m}\mathbf{1}\right)$) splits the underlying assets into two groups, i.e. outperforming the average and underperforming the average. This is in general reasonable, since investors often transfer weights from the assets that are expected to underperform to the assets that are expected to outperform. Thus, $\tilde{\mathbf{b}}_{t+\frac{1}{2}}$ represents the weights to be transferred, which will transfer from its negative components to positive ones. For the assets that are expected to increase (or decrease) by a small (large) value, the transferred weights ($\tilde{b}_{t+\frac{1}{2},i}$) will also be small (large).

Viewing $\tilde{\mathbf{b}}_{t+\frac{1}{2}}$ as the weights to be theoretically transferred without transaction costs, the second update formula adjusts it in case of non-zero transaction costs. While $\hat{\mathbf{b}}_t$ denotes the current allocation before rebalancing, the remaining terms denote the weights to be transferred in case of non-zero transaction costs. If the absolute value of $\tilde{\mathbf{b}}_{t+\frac{1}{2}}$ is large (small), it deems to affect more (less) to the final performance. Since the transaction costs exist, rebalancing one asset should produce more value than the incurred transaction costs. Therefore, the latter term truncates $\tilde{\mathbf{b}}_{t+\frac{1}{2}}$ if one element's absolute value is below a threshold of $\lambda$, and keeps the original value if its absolute value is above the threshold. If the threshold is zero, then the update degrades to $\mathbf{b}_{t+1} = \hat{\mathbf{b}}_t + \eta_t\left(\mathbf{v}_t - \bar{v}_t \mathbf{1}\right)$, which follows the principle of investments (Li *et al.* 2015). If the threshold is positive, the algorithm will compare the magnitude (or absolute value) of $\mathbf{b}_{t+\frac{1}{2}}$ with the threshold. In other words, such mechanism ensures that the portfolio will only rebalance assets that are expected to deviate a lot and outweigh the transaction costs.

Finally, we want to further investigate the portfolio algorithms. Usually, when transaction costs are included in portfolio selection problems, there is a boundary that the portfolio (or individual stocks) has to cross before manager rebalances the portfolio. In literature, this portfolio rebalancing method is named *Proportion of Portfolio Rebalancing* (PPR) (Institute 2013). The key idea of PPR is to keep a corridor of the weights, which defines the upper bound and lower bound for each asset, and only rebalance the allocation once next weight is out of the corridor. For example, the manager set the corridor for an asset as $W \pm 0.1W$, where $W$ denotes the current weight. The manager thus only rebalances the allocation if the transferred weight is larger than 10% of current weight. Although we do not consider such mechanism in the TCO's formulation, its derived portfolio updates do reflect such thresholds. Obviously, TCO's mechanism is the same as PPR's, which is equivalent to $\lambda = 0.1$. Such coincidence interestingly connects our method to the commonly used rebalancing method in industry.

## 5. Empirical evaluations

### 5.1. *Data-sets*

In the empirical evaluations, we mainly adopt four public data-sets,† i.e. NYSE (O) (Cover 1991), and its following data-set, NYSE (N) (Györfi *et al.* 2012, Li *et al.* 2013), TSE (Borodin *et al.* 2004) and MSCI (Li *et al.* 2012). These publicly available data-sets ensure the reproducibility and make our comparison

†Details of these data-sets, including their compositions, are available at http://olps.stevenhoi.org/, and TCO's Matlab implementation will be available online.

Table 3. Summary of the adopted mean reversion predictions. $\bigodot$ denotes the element-wise product, and $p_i$ denotes the implicit probability of prediction $\tilde{\mathbf{x}}_{t+1}^i$.

| Schema | Prediction ($\tilde{\mathbf{x}}_{t+1}^i = f\left(\mathbf{x}_1^t\right)$) | Prob. ($p_i$) | Reference |
|---|---|---|---|
| 1 | $1/\mathbf{x}_t$ | 100% | PAMR (Li *et al*. 2012)/CWMR (Li *et al*. 2013) |
| 2 | $\frac{1}{w}\left(1+\frac{1}{\mathbf{x}_t}+\cdots+\frac{1}{\bigodot_{i=0}^{w-2}\mathbf{x}_{t-i}}\right)$ | 100% | OLMAR (Li *et al*. 2015) |

fair to existing algorithms. In particular, NYSE (O) contains 36 large cap stocks from the New York Stock Exchange, and 5651 price relatives ranging from 7 March 1962 to 31 December 1984. NYSE (N) has 23 remaining survived stocks till 30 June 2010, totalling 6431 price relatives. TSE contains 88 large cap stocks from the Toronto Stock Exchange (TSE) and 1259 price relatives, ranging from 4 January 1994 to 31 December 1998. The final MSCI data-set is a collection of global equity indices that are the constituents of MSCI World Index.† It contains 24 indices that represent the equity markets of 24 countries around the world, and consists of 1043 trading days ranging from 1 April 2006 to 31 March 2010. While the first three data-sets mainly test the proposed algorithms on stock markets, the fourth data-set test the algorithm's capability on market indices, which may be potentially applicable to 'Fund of Funds' (FOF). Moreover, even though we test the algorithms on the stock related markets, they can be applied to any type of financial markets.

**5.1.1. Discussions on survival bias.** Pioneered by Cover (1991), NYSE (O) is the standard data-set widely used in the online portfolio selection community. The main reason for choosing NYSE (O) is because it allows us to compare the proposed algorithms with all related algorithms in section 3. As the data-set contains 36 large cap NYSE stocks that survived in hindsight for 22 years, it suffers from extreme survival bias. To examine a strategy's profitability over time, Gábor Gelencsér and we‡ created the new NYSE (N) data-set as a continuation of NYSE (O), containing 23 stocks from the 36 NYSE stocks that survived for additional 26 years. Note that the delisted stocks are mainly due to merge and acquisitions, as they are the largest cap firms in the US markets. The NYSE (N) data-set is therefore even worse than the old data-set in this respect. Any serious researcher should be aware of the implicit survival bias in the empirical evaluations.

In fact, there is a trade-off for any fixed (assets) data-set, i.e. length of the data-sets vs. the survival bias. If one data-set is long (such as NYSE (O) or NYSE (N)), it tends to suffer extreme survival bias but is more representative for the whole population. Though the effect of survival bias is weak in data-sets with short durations (such as MSCI and TSE), it may be not large enough to represent the population. Thus, our experiment test bed covers both long-period data-sets ($5000 \sim 6000$ trading days) and short-period data-sets ($\sim$1000 trading days).

Clearly, testing the proposed algorithm on survivor-bias free data-sets can be more realistic. However, to adopt survivor bias free data-sets, we may encounter several challenges. The first challenge is that the delisting return data are missing or hard to collect and calculate. According to Beaver *et al.* (2007), the delisted stocks are mainly due to mergers and acquisitions (51% of their samples) or poor performance (44% of their samples). The returns may be significantly different in the two cases. In cases of mergers and acquisitions, which is hard to collect and process the data, these delisting returns may be positive. In cases of poor performance (e.g. bankruptcy), these returns may be negative. Some studies assign $-100\%$ for these returns (Sloan 1996), while others assign $-30\%$ (Mohanram 2005) or just deleted them (Hribar and Collins 2002). To the best of our knowledge, there is no agreed on method to calculate the returns for delisted stocks.

The second challenge is that even if we can obtain a survivor-bias free data-set, we may encounter challenges in portfolio construction. As survivor-bias free data-sets include delisted stocks, the number of assets will change over time. However, all existing algorithms are designed and implemented with a fixed number of assets (see the problem setting in section 4.2). As a result, all these algorithms have to be modified for the changing number of assets. We note that some accounting and finance studies (such as Beaver *et al.* (2007)) using survivor-bias free data mainly adopt equally weighted portfolio, which simply divides the stocks into several groups and assigns equal weights in each group. Unfortunately, the existing online portfolio selection strategies are usually far more complicated than equally weighted portfolio (c.f. see existing algorithms in section 3), and adapting them to handle changing number of assets is a challenging task. As our main purpose is to propose a framework to tackle the transaction costs issue for online portfolio selection, adapting existing algorithms for the changing number of assets is beyond the scope of this article.

The third reason is that the four data-sets have been widely used in the related studies. To fairly compare with existing studies, we therefore follow existing studies to conduct the empirical experiments on the widely used data-sets. But researchers should keep in mind that the reported results may be biased by the survival bias embedded in the widely used data-sets.

### *5.2. Settings*

The TCO algorithm has two possible parameters, i.e. $\eta$ and $\lambda$. Intuitively, the higher the cost rate is, the less the manager should rebalance. Following the line of OLPS research (Helmbold *et al*. 1998, Li *et al*. 2012), we empirically choose parameter values and later evaluate their sensitivity. It is

†The constituents of MSCI World Index are available on MSCI Barra (http://www.mscibarra.com), accessed on 28 May 2010.
‡Gábor collected till 2006 and we extended the data to 2010.

possible to further alleviate the challenge of parameter selection, such as using expert learning (Borodin *et al.* 2004, Li *et al.* 2013), etc. Note that in some existing studies, optimal performance can be theoretically achieved under certain distributional assumptions, however, TCO's optimal performance cannot be achieved. Naturally, there is no way to choose parameter values so as to obtain optimal out-of-sample performance.

In particular, we set the trade off parameter to $\lambda = 10 \times \gamma$, which is empirically effective, and set the smoothing parameter to $\eta = 10$ in all cases. Further illustrations in section 5.3.4 show that our choices of parameters are not optimal in all cases, and it is easy to blindly choose satisfying parameters.

There are mainly two performance metrics to measure an online portfolio selection strategy with transaction costs. The first metric is *cumulative wealth* (Li and Hoi 2014), which measures the cumulative return starting from one dollar. The higher the cumulative wealth, the less an algorithm suffers from the transaction costs. Another is *average turnover* over the whole period, which measures the changes of portfolios. Lower turnovers indicate more effective portfolio rebalancing strategy in cases of non-zero transaction costs. There also exist various risk adjusted metrics, such as Sharpe Ratio and Calmar ratio. However, they are just complementary to the above two metrics, which will be available in the TCO's package.

Besides, we also conduct statistical tests for TCO's daily return series (Grinold and Kahn 1999). The test separates the return's alpha ($\alpha$) and beta ($\beta$), and outputs the probability that the returns are generated by simple luck. The smaller the probability, the higher the confidence we have on a trading strategy.

**5.2.1. Comparison approaches.** In the experiments, we implement the two proposed TCO strategies, i.e. TCO1 and TCO2. We compare them with a number of benchmarks and existing strategies (Li and Hoi 2014). Below we summarize these algorithms, whose parameters are set according to the recommendations from their respective studies.†

(i) Market: Uniform Buy-And-Hold (BAH) strategy;
(ii) Best: Best stock in hindsight;
(iii) BCRP (Cover 1991): Best Constant Rebalanced Portfolios strategy in hindsight;
(iv) UP (Cover 1991): Cover's Universal Portfolios implemented according to Kalai and Vempala (2002), in which the parameters are set as $\delta_0 = 0.004$, $\delta = 0.005$, $m = 100$, and $S = 500$;
(v) EG (Helmbold *et al.* 1998): Exponential Gradient algorithm with the best parameter $\eta = 0.05$;
(vi) ONS (Agarwal *et al.* 2006): Online Newton Step with the parameters, that is, $\eta = 0$, $\beta = 1$, $\gamma = \frac{1}{8}$;
(vii) OLU (Das *et al.* 2013): Online Lazy Updates with the parameters, that is, $\eta = 20$, $\beta = 0.1$, $\gamma = 0.1$;
(viii) Anticor (Borodin *et al.* 2004): $\text{BAH}_{30}$(Anticor(Anticor)) as a variant of Anticor to smooth the performance, which achieves the best performance among the three solutions;
(ix) CORN (Li *et al.* 2011): Correlation-driven nonparametric learning approach with $W = 5$ and $\rho = 0.1$;
(x) PAMR (Li *et al.* 2012): Passive Aggressive Mean Reversion algorithm with $\epsilon = 0.5$;
(xi) OLMAR (Li *et al.* 2015): Online Moving Average Reversion with $w = 5$ and $\epsilon = 10$;
(xii) RMR (Huang *et al.* 2013): Robust Median Reversion with $w = 5$, $\epsilon = 5$, and $m = 200$.

†We can tune their parameters for better performance. However, it is beyond the scope of this article.

### *5.3. Experimental results*

**5.3.1. Cumulative wealth with fixed transaction costs.** We backtested two reasonable rates of transaction costs (Das *et al.* 2013), i.e. 0.25 and 0.5%. Table 4 illustrates the cumulative wealth achieved by various algorithms under the rates of 0, 0.25 and 0.5%, respectively. For example, 14.50 achieved by BAH on NYSE (O)-0% means that $1 invested using the BAH strategy will grow to $14.50 after 22 years.

From the table, we can draw several observations. Firstly, without transaction costs, BCRP are proved to outperform Best (Cover 1991, proposition 2.1) due to its exploitation of assets' volatility (Luenberger 1998). However, with transaction costs, BCRP may underperform Best, because its constantly rebalancing will cause high transaction costs. For example, although BCRP outperforms Best in most cases, it underperforms Best in columns NYSE (N)-0.5% and MSCI-0.25% & -0.5%. Secondly, most state of the art algorithms (e.g. CORN, PAMR, OLMAR and RMR.) drop significantly when the rates are non-zeros, especially $\gamma = 0.5\%$. The observation shows that this research on tackling non-zero transaction costs is necessary. Thirdly, although the three theoretical guaranteed algorithms (i.e. UP, EG and ONS) achieve a low cumulative wealth with zero transaction costs, their performance degrades much slower than other algorithms. Finally, the proposed TCO algorithms perform much better than the state of the art. In particular, TCO achieves the top two achievements in most non-zero cases. In summary, the research on non-zero transaction costs is necessary and the proposed TCO algorithms can robustly resist reasonable transaction costs.

Moreover, table 5 shows the statistical test of the proposed TCO algorithm in case of three rates, i.e. 0, 0.25 and 0.5%. As rates increase, both TCO's winning ratios over the market and $\alpha$s decrease. The $p$-values show that when $\gamma = 0.25\%$, TCO's performance on most data-sets (except MSCI) is not due to luck. However, when the rate increases to 0.5%, the test's results are consistent with the observations in table 4. In particular, the algorithms cannot beat the market strategy on some data-sets, thus $p$-values become as high as 50%. In summary, the statistical tests verify the observations in table 4, and show the TCO's effectiveness with reasonable transaction costs.

Besides the above observations, we would like to discuss the issue of market efficiency. Note that the discussion on the market efficiency does not prevent traders from exploiting the market for profit. For example, if the market does not satisfy the weak form efficiency, technical trading rules can beat the markets (Bodie *et al.* 2014). The reasons for such high returns on the NYSE (O) data-set may be in twofolds. Firstly, the trading strategies are on a daily basis and once the

Table 4. Cumulative wealth achieved by various strategies on the data-sets with two common transaction cost rates (0.25 and 0.5%), plus zero rate for benchmark. Top two achievements on each column excluding zero rate column and the three benchmarks are highlighted.

| | NYSE(O) | | | NYSE(N) | | | TSE | | | MSCI | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Algorithms | 0% | 0.25% | 0.5% | 0% | 0.25% | 0.5% | 0% | 0.25% | 0.5% | 0% | 0.25% | 0.5% |
| BAH | 14.50 | 14.46 | 14.42 | 18.06 | 18.01 | 17.97 | 1.61 | 1.61 | 1.60 | 0.91 | 0.90 | 0.89 |
| Best | 54.14 | 54.00 | 53.87 | 83.51 | 83.30 | 83.09 | 6.28 | 6.26 | 6.25 | 1.50 | 1.50 | 1.50 |
| BCRP | 250.60 | 182.01 | 132.20 | 120.32 | 98.95 | 81.38 | 6.78 | 6.51 | 6.26 | 1.51 | 1.49 | 1.48 |
| UP | 26.68 | 17.59 | 11.36 | 31.49 | 20.56 | 13.31 | 1.60 | 1.41 | 1.24 | 0.92 | 0.87 | 0.81 |
| EG | 27.09 | 23.08 | 19.66 | 31.00 | 25.74 | 21.37 | 1.59 | 1.52 | 1.46 | 0.93 | 0.91 | **0.88** |
| ONS | 109.91 | 51.47 | 24.26 | 21.59 | 11.24 | 5.85 | 1.62 | 1.21 | 0.90 | 0.86 | 0.72 | 0.60 |
| OLU | 36.04 | 32.81 | 29.88 | 22.58 | 19.97 | 17.66 | 1.71 | 1.64 | **1.58** | 0.46 | 0.45 | 0.44 |
| Anticor | 2.41E+08 | 6.37E+04 | 16.68 | 6.21E+06 | 1.46E+03 | 0.34 | 39.36 | 6.87 | 1.20 | 3.22 | 1.02 | 0.32 |
| CORN | 1.48E+13 | 1.38E+04 | 0.00 | 5.37E+05 | 0.00 | 0.00 | 3.56 | 0.02 | 0.00 | 26.10 | 0.33 | 0.00 |
| PAMR | 5.14E+15 | 2.09E+05 | 0.00 | 1.25E+06 | 0.00 | 0.00 | 264.86 | 2.11 | 0.00 | 15.23 | 0.15 | 0.00 |
| OLMAR | 4.04E+16 | 2.43E+07 | 0.02 | 2.24E+08 | 0.04 | 0.00 | 424.8 | 4.00 | 0.04 | 16.39 | 0.34 | 0.01 |
| RMR | 1.64E+17 | **6.94E+08** | 2.66 | 3.25E+08 | 0.42 | 0.00 | 181.34 | 3.38 | 0.02 | 16.76 | 0.45 | 0.01 |
| TCO1 | 1.35E+14 | **5.53E+09** | **2.31E+06** | 9.15E+06 | **3.80E+03** | **142.00** | 148.99 | **7.73** | 0.92 | 9.68 | **1.52** | **1.13** |
| TCO2 | 1.40E+13 | 3.87E+07 | **1.28E+04** | 2.43E+07 | **2.00E+03** | **55.00** | 153.05 | **31.54** | **4.70** | 5.68 | **1.42** | 0.84 |

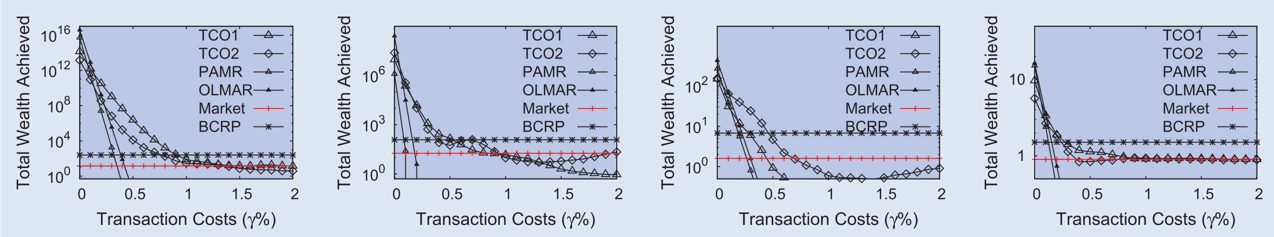
Table 5. Statistical $t$-test of TCO2's performance on the stock data-sets. 'MER' denotes Mean Excess Return.

| | NYSE(O) | | | NYSE(N) | | | TSE | | | MSCI | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Statistics | 0% | 0.25% | 0.5% | 0% | 0.25% | 0.5% | 0% | 0.25% | 0.5% | 0% | 0.25% | 0.5% |
| Length | 5651 | 5651 | 5651 | 6431 | 6431 | 6431 | 1259 | 1259 | 1259 | 1043 | 1043 | 1043 |
| MER (TCO) | 0.0057 | 0.0035 | 0.0021 | 0.0031 | 0.0017 | 0.0011 | 0.0052 | 0.0039 | 0.0025 | 0.0019 | 0.0006 | 0.0000 |
| MER (Market) | 0.0005 | 0.0005 | 0.0005 | 0.0005 | 0.0005 | 0.0005 | 0.0004 | 0.0004 | 0.0004 | 0.0000 | 0.0000 | 0.0000 |
| Winning Ratio | 57.69% | 49.92% | 48.47% | 54.52% | 49.23% | 48.55% | 53.77% | 51.31% | 48.93% | 57.05% | 51.68% | 47.94% |
| $\alpha$ | 0.0051 | 0.0029 | 0.0015 | 0.0026 | 0.0012 | 0.0006 | 0.0046 | 0.0034 | 0.0019 | 0.0019 | 0.0006 | 0.0000 |
| $\beta$ | 1.2824 | 1.3276 | 1.2008 | 1.1449 | 1.0923 | 1.0601 | 1.5311 | 1.5516 | 1.5584 | 1.1526 | 1.1618 | 1.1285 |
| $t$-statistics | 15.464 | 8.0341 | 4.2969 | 7.2494 | 3.0743 | 1.5863 | 3.3166 | 2.4115 | 1.3354 | 5.5155 | 1.572 | 0.0207 |
| $p$-value | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0011 | 0.0564 | 0.0005 | 0.0080 | 0.0910 | 0.0000 | 0.0581 | 0.4917 |

strategies spot a frequent pattern (especially the mean reversion pattern), the power of compound interests will take effect and the strategies will yield an astonishing return without transaction cost. However, frequent daily trading will incur high turnover and thus high transaction costs with non-zero rates. This naturally shows the importance of this study on transaction costs. Secondly, the years in NYSE (O) are relatively old and thus the market is more likely to be (weak-form) inefficient. Thus a strategy may better exploit the market, and gain huge profit. In particular, the NYSE (O) data-set ranges from 1962 to 1984, when the markets are more inefficient, thus a powerful strategy can obtain such a high profit. On contrary, NYSE (N), the subsequent version of NYSE (O), becomes less inefficient, so the same strategy can obtain much less profit.

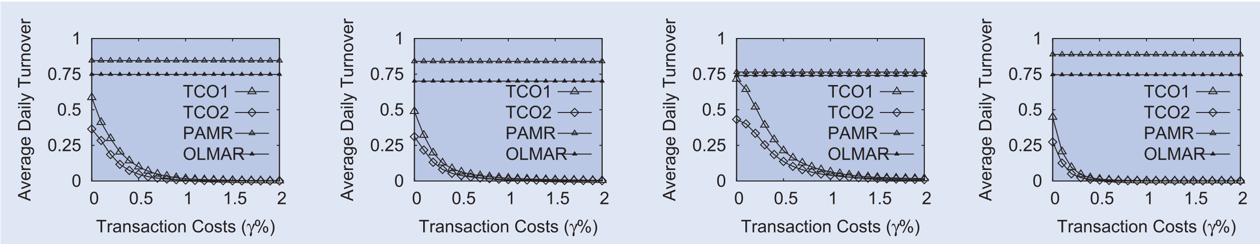
For the different returns between NSYE (O) and MSCI, it may be explained that NYSE (O) and MSCI are different data-sets with different components. While NYSE(O), NYSE (N), TSE are three data-sets on stocks, MSCI is a portfolio of global equity indices. Indicated by table 5—Row 'MER (Market)', MSCI's return is much lower than the other data-sets' return. Moreover, due to the diversification effect, indices are usually less volatile than individual stocks. Thus, strategies that exploit the volatility will make less profit (Li *et al*. 2012, see section 5.4.3 for an examination of the volatility issue in the OLPS community). Secondly, as time goes on, more and more inefficiencies will be exploited for profit, and the financial markets are thus less and less inefficient (or more and more efficient). For example, even on the same market and with similar daily MER, the NYSE (N) data-set yields much less profit than the NYSE (O) data-set. As MSCI ranges from 2006 to 2010, it tends to be more efficient than NYSE, which ranges from 1962 to 2010.

**5.3.2. Cumulative wealth with varying transaction costs.** To better show the effectiveness of the introduced L1 term for non-zero transaction costs, figure 1 compares the cumulative wealth achieved by the proposed TCO strategies and PAMR and OLMAR, which are equipped with the same prediction functions. Two benchmarks, market and BCRP, are also plotted for reference. From the figures, we can draw several observations. Firstly, PAMR and OLMAR decrease exponentially with



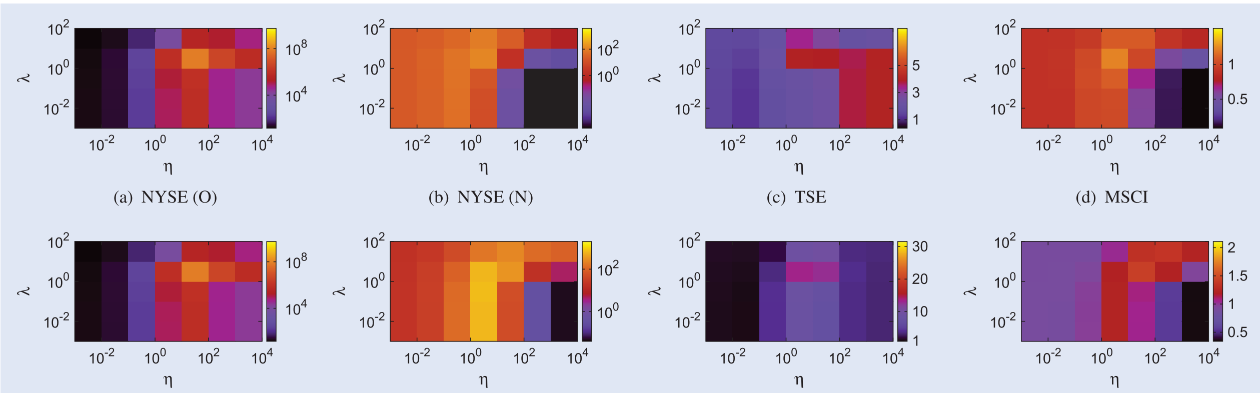

(a) NYSE (O) (b) NYSE (N) (c) TSE (d) MSCI

Figure 1. Cumulative wealth achieved by TCO with respect to varying transaction cost rates.




(a) NYSE (O) (b) NYSE (N) (c) TSE (d) MSCI

Figure 2. Average turnovers achieved by TCO with respect to varying transaction cost rates.




(a) NYSE (O) (b) NYSE (N) (c) TSE (d) MSCI

(e) NYSE (O) (f) NYSE (N) (g) TSE (h) MSCI

Figure 3. Parameter sensitivity of TCO1 and TCO2 with respect to $\eta$ and $\lambda$ at the transaction costs rate of 0.25%. Each point in the heat map corresponds to cumulative wealth achieved with a combination of $\eta$ and $\lambda$. The colour-bars besides each figure show the scale of cumulative wealth. The brighter the colour is, the more wealth it achieves. The horizontal axis denotes varying $\eta$, and the vertical axis indicates varying $\lambda$.

increasing rates. For example, the break-even rates with respect to the market for NYSE (O) are around 0.3 and 0.4%, while for NYSE (N) the break-even rates are 0.1 and 0.2%. Although some institutional investors may have lower rates, their profit will be robbed by non-zero transaction costs. Secondly, TCO1 and TCO2 have much higher break-even rates on all data-sets. In particular, the break even rates are both 1.4% on NYSE (O), and the break-even rates are both 0.9% on NYSE (N). Such observation greatly enhances the algorithms' scalability with respect to non-zero transaction costs. Thirdly, on almost all levels of rates (except some small rates subject to tuning), the proposed TCO can significantly outperform PAMR and OLMAR. In a word, the proposed TCOs significantly improve the performance in case on non-zero transaction costs.

#### 5.3.3. Average turnover with varying transaction costs.

As shown in proposition 4.1, turnover is directly related to the transaction costs. The lower the turnover, the less the imposed transaction costs. Thus, figure 2 compares the TCO's average turnovers with those of PAMR and OLMAR. While PAMR and OLMAR do not consider the transaction costs issue, their average turnovers maintain high levels on all data-sets in order to exploit assets' volatility. The high turnovers unavoidably cause high transaction costs when the proportional transaction costs rate is non-zero. On the other hand, TCO1 and TCO2's turnovers are high when the rates are small and consistently decrease to almost zero as the rates increase. Note that zero turnover is equivalent to no rebalancing, which will not incur any transaction costs. This again verifies that TCO trades off between rates and turnover, to include transaction costs. Finally, connecting figure 2 to figure 1, we can find that in cases of non-zero transaction costs, such trade-off leads to TCO's higher cumulative return than PAMR and OLMAR.

**5.3.4. Parameter sensitivity.** We then evaluate the sensitivity of parameters, i.e. $\eta$ and $\lambda$. The heat maps in figure 3 show the achieved cumulative wealth with various combinations of $\eta$ and $\lambda$ at a reasonable rate of 0.25%. Firstly, we can always observe a peak (the brightest) in the middle region, indicating that these combinations of $\eta$ and $\lambda$ yield relative high cumulative wealth. This observation provides a wide range of feasible parameters that release satisfying performance. Note that the locations of dark regions do not affect the key observation. Secondly, fixing $\lambda$, we can find that as $\eta$ increases, TCO's wealth initially increases and then peaks at a point and finally decreases. Thirdly, fixing $\eta$, we can observe that strategy with different $\lambda$ yields slightly different cumulative wealth. Fixing a small $\eta$, the strategy yields similar wealth around different $\lambda$, as the upper and lower rows have similar colours. Fixing a middle $\eta$, the strategy's wealth initially increases and then peaks at a middle point and finally decreases. Fixing a large $\eta$, the lower rows are darker, which indicates low wealth. Finally, throughout our experiments, we simply set both $\eta$ and $\lambda$ to 10, which release satisfying but obviously not optimal results.

## 6. Conclusions

In this paper, we investigated the problem of online portfolio selection with proportional transaction costs. We firstly formulated the problem as a sequential decision problem, and discovered the relationship between transaction costs and portfolio change. Then we proposed the TCO framework, which trades-off between maximizing expected log return and minimizing transaction costs. An analysis of the closed form update formula shows the connection to a typical rebalancing strategy in industry. Two specific algorithms using existing mean reversion estimation methods have been proposed. Extensive experiments on widely used data-sets show that these derived algorithms are effective in the case of non-zero transaction costs.

Although we focus on the empirical contribution, one important drawback of this article is its lack of theoretical contribution. Thus, in future, we would like to investigate the theoretical aspect of the research problem and try to make some theoretical contributions. One significant drawback in our empirical evaluation is that the widely used data-sets are subject to survival bias. In future, we would like to design online portfolio selection strategies on data-sets free of survivor bias. Moreover, we would like to relax the assumptions of proportional transaction costs. For example, we may assume fixed transaction costs per share, and different rates for purchases and sales. Finally, relaxing other assumptions, such as no market impact or liquidity, will also contribute to the potential deployment of these strategies.


## Acknowledgements

We want to thank the editors and anonymous referees for their helpful comments and suggestions.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## Funding

This work was supported by the National Natural Science Foundation of China [71401128], [91646206], [11501204]; the Academic Team Building Plan for Young Scholars from Wuhan University [WHU2016012]; Natural Scientific Research Program of Wuhan University [2042017kf0225]; the Natural Science Foundation of Shanghai [15ZR1408300]; Shanghai Key Laboratory of Intelligent Information Processing [IIPL-2014-001].


## References


Agarwal, A., Hazan, E., Kale, S. and Schapire, R.E., Algorithms for portfolio management based on the Newton method. In *Proceedings of International Conference on Machine Learning*, Pittsburgh, PA, 2006.

Albeverio, S., Lao, L. and Zhao, X., On-line portfolio selection strategy with prediction in the presence of transaction costs. *Math. Methods Oper. Res.*, 2001, **54**, 133–161.

Bach, F.R., Jenatton, R., Mairal, J. and Obozinski, G., Optimization with sparsity-inducing penalties. *Found. Trends Mach. Learn.*, 2012, **4**, 1–106.

Bauer, E. and Kohavi, R., An empirical comparison of voting classification algorithms: Bagging, boosting, and variants. *Mach. Learn.*, 1999, **36**, 105–139.

Beaver, W., McNichols, M. and Price, R., Delisting returns and their effect on accounting-based market anomalies. *J. Account. Econ.*, 2007, **43**, 341–368.

Bodie, Z., Kane, A. and Marcus, A.J., *Investments, 10e*, 2014 (McGraw-Hill Education: New York).

Borodin, A., El-Yaniv, R. and Gogan, V., Can we learn to beat the best stock. *J. Artif. Intell. Res.*, 2004, **21**, 579–594.

Boyd, S., Parikh, N., Chu, E., Peleato, B. and Eckstein, J., Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found. Trends®Mach. Learn.*, 2011, **3**, 1–122.

Boyd, S. and Vandenberghe, L., *Convex Optimization*, 2004 (Cambridge University Press: New York).

Cover, T.M., Universal portfolios. *Math. Finance*, 1991, **1**, 1–29.

Creamer, G.G. and Freund, Y., Automated trading with boosting and expert weighting. *Quant. Finance*, 2010, **10**, 401–420.

Das, P., Johnson, N. and Banerjee, A., Online lazy updates for portfolio selection with transaction costs. In *Proceedings of the Twenty-Seventh Conference on Artificial Intelligence*, Bellevue, WA, 2013.

DeMiguel, V., Garlappi, L. and Uppal, R., Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *Rev. Financ. Stud.*, 2009, **22**, 1915–1953.

DeMiguel, V., Nogales, F.J. and Uppal, R., Stock return serial dependence and out-of-sample portfolio performance. *Rev. Financ. Stud.*, 2014, **27**, 1031–1073.

Duchi, J., Shalev-Shwartz, S., Singer, Y. and Chandra, T., Efficient projections onto the $l_1$-ball for learning in high dimensions. In *Proceedings of the International Conference on Machine Learning*, Helsinki, 2008.

Duchi, J. and Singer, Y., Efficient online and batch learning using forward backward splitting. *J. Mach. Learn. Res.*, 2009, **10**, 2899–2934.

Fabozzi, F.J., Focardi, S. and Jonas, C., Trends in quantitative equity management: Survey results. *Quant. Finance*, 2007, **7**, 115–122.

Granger, C.W.J. and Sin, C.Y., Modelling the absolute returns of different stock indices: Exploring the forecastability of an alternative measure of risk. *Int. J. Forecast.*, 2000, **19**, 277–298.

Grinold, R. and Kahn, R., *Active Portfolio Management: A Quantitative Approach for Producing Superior Returns and Controlling Risk*, 1999 (McGraw-Hill: New York).

Györfi, L., Lugosi, G. and Udina, F., Nonparametric kernel-based sequential investment strategies. *Math. Finance*, 2006, **16**, 337–357.

Györfi, L., Ottucsák, G. and Walk, H., *Machine Learning for Financial Engineering*, 2012 (World Scientific: Singapore).

Györfi, L., Udina, F. and Walk, H., Nonparametric nearest neighbor based empirical portfolio selection strategies. *Stat. Decis.*, 2008, **26**, 145–157.

Györfi, L. and Vajda, I., Growth optimal investment with transaction costs. In *Proceedings of the International Conference on Algorithmic Learning Theory*, Budapest, 2008.

Helmbold, D.P., Schapire, R.E., Singer, Y. and Warmuth, M.K., A comparison of new and old algorithms for a mixture estimation problem. *Mach. Learning*, 1997, **27**, 97–119.

Helmbold, D.P., Schapire, R.E., Singer, Y. and Warmuth, M.K., On-line portfolio selection using multiplicative updates. *Math. Finance*, 1998, **8**, 325–347.

Hribar, P. and Collins, D.W., Errors in estimating accruals: Implications for empirical research. *J. Account. Res.*, 2002, **40**, 105–134.

Huang, D., Zhou, J., Li, B., Hoi, S.C. and Zhou, S., Robust median reversion strategy for on-line portfolio selection. In *Proceedings of the Twenty-Third international joint conference on Artificial Intelligence*, Beijing, pp. 2006–2012, 2013.

Institute, C., *CFA Program Curriculum: Portfolio: Execution, Evaluation and Attribution, and Global Investment Performance Standards*, CFA Program Curriculum: Level III, 2013 2013 (CFA Institute: Boston, MA).

Kalai, A. and Vempala, S., Efficient algorithms for universal portfolios. *J. Mach. Learn. Res.*, 2002, **3**, 423–440.

Kelly, Jr, J., A new interpretation of information rate. *Bell Syst. Tech. J.*, 1956, **35**, 917–926.

Laureti, P., Medo, M. and Zhang, Y.C., Analysis of kelly-optimal portfolios. *Quant. Finance*, 2010, **10**, 689–697.

Li, B. and Hoi, S.C., Online portfolio selection: A Survey. *ACM Comput. Surv.*, 2014, **36**, 35:1–35:36.

Li, B., Hoi, S.C. and Gopalkrishnan, V., CORN: Correlation-driven nonparametric learning approach for portfolio selection. *ACM Trans. Intell. Syst. Technol.*, 2011, **2**, 21:1–21:29.

Li, B., Hoi, S.C., Sahoo, D. and Liu, Z., Moving average reversion strategy for on-line portfolio selection. *Artif. Intell.*, 2015, **222**, 104–123.

Li, B., Hoi, S.C., Zhao, P. and Gopalkrishnan, V., Confidence weighted mean reversion strategy for online portfolio selection. *ACM Trans. Knowl. Discov. Data*, 2013, **7**, 4:1–4:38.

Li, B., Zhao, P., Hoi, S. and Gopalkrishnan, V., PAMR: Passive aggressive mean reversion strategy for portfolio selection. *Mach. Learn.*, 2012, **87**, 221–258.

Luenberger, D.G., *Investment Science*, 1998 (Oxford University Press: New York).

Maclean, L., Thorp, E. and Ziemba, W., Long-term capital growth: The good and bad properties of the Kelly and fractional Kelly capital growth criteria. *Quant. Finance*, 2010, **10**, 681–687.

Mohanram, P.S., Separating winners from losers among lowbook-to-market stocks using financial statement analysis. *Rev. Account. Stud.*, 2005, **10**, 133–170.

Nevmyvaka, Y., Feng, Y. and Kearns, M.S., Reinforcement learning for optimized trade execution. In *Proceedings of the International Conference on Machine Learning*, Pittsburgh, PA, 2006.

Ormos, M. and Urbán, A., Performance analysis of log-optimal portfolio strategies with transaction costs. *Quant. Finance*, 2013, **13**, 1587–1597.

Shen, W. and Wang, J., Portfolio blending via Thompson sampling. In *Proceedings of the IJCAI International Joint Conference on Artificial Intelligence*, New York, NY, pp. 1983–1989, 2016.

Shen, W. and Wang, J., Portfolio selection via subset resampling. In *Proceedings of the 31th Conference on Artificial Intelligence*, San Francisco, CA, pp. 1517–1523, 2017.

Shen, W., Wang, J., Jiang, Y.G. and Zha, H., Portfolio choices with orthogonal bandit learning. In *Proceedings of the International Joint Conference on Artificial Intelligence*, pp. 974–980, 2015.

Sloan, R.G., Do Stock prices fully reflect information in accruals and cash flows about future earnings? *Account. Rev.*, 1996, **71**, 289–315.

Thorp, E.O., Portfolio choice and the Kelly criterion. In *Proceedings of the Business and Economics Section of the American Statistical Association*, Fort Collins, CO, 1971.

Tsagaris, T., Jasra, A. and Adams, N., Robust and adaptive algorithms for online portfolio selection. *Quant. Finance*, 2012, **12**, 1651–1662.

Tu, J. and Zhou, G., Markowitz meets Talmud: A combination of sophisticated and naive diversification strategies. *J Financ. Econ.*, 2011, **99**, 204–215.