國立臺灣大學管理學院財務金融研究所

碩士論文

Department of Finance
College of Management
National Taiwan University

Master Thesis

深度強化式學習在指數追蹤的應用

Deep Reinforcement Learning on Index-Tracking

鍾文豪

Wen-Hao Chung

指導教授：韓傳祥 博士

葉小蓁 博士

Advisor: Chuan-Hsiang Han, Ph.D.

Hsiaw-Chan Yeh, Ph.D.

中華民國 107 年 6 月

June, 2018

doi:10.6342/NTU201801196

中文摘要

指數追蹤是一種投資組合管理，藉由建構投資組合來追蹤特定指數的績效，同時極小化追蹤偏離度及追蹤誤差。如果我們知道指數的成分股及成分股的權重或者指數的編制規則，指數追蹤的問題就變得相當容易。如果上述的資訊全部是私有訊息呢？本文提出深度強化式學習方法，在不知道指數成分股、成分股的權重及指數編製規則的情況下，建立指數追蹤投資組合追蹤該指數。本文使用深度強化式學習中的策略梯度，來建構指數追蹤投資組合。策略梯度能夠將狀態訊息轉換成連續的動作，相較於深度 Q-學習，更適合用來做投資組合管理。美國股票市場的所有普通股將作為強化式學習模型的輸入，用來追蹤股票指數（S&P500、NASDAQ Composite）及主動式基金（FSCSX、FBSOX、NASDX）。追蹤偏離度的均方是我們主要衡量指數追蹤的依據。實驗結果顯示，我們提出的深度強化式學習方法所建構的指數追蹤投資組合可以良好的追蹤標的。在樣本外測試期間，追蹤偏離度的均方至少可以達到 2.71E-05 的水準。

關鍵字：指數追蹤、投資組合管理、強化式學習、策略梯度、追蹤偏離度

doi:10.6342/NTU201801196

ABSTRACT

An index-tracking problem is a kind of portfolio management, building a portfolio that tracks the performance of a certain index while minimizing the tracking difference and tracking error. If constituents of the index and portfolio weights of constituents are known or rules to build the index are public information, tracking index is trivial. What if the above information is private? In this paper, we propose a deep reinforcement learning method to build an index-tracking portfolio to track the index while not knowing the constituents of the index, portfolio weights of constituents or rules to build the index. Deep reinforcement learning (RL) with Policy Gradient is deployed to build the index-tracking portfolio. Policy Gradient transform state information to continuous actions which is more suitable for portfolio management than the deep Q-learning. The whole U.S. equity will be put into the deep RL model to track the indexes (S&P 500 and NASDAQ Composite) or the active funds (NASDX, FSCSX and FBSOX). Mean square difference is used as our main measurement for index-tracking. The experiment result shows that the index-tracking portfolio build by the proposed RL method could excellently track the target. The mean square of tracking difference could at least achieve 2.71E-05 in the whole testing period.

Keywords: Index-tracking, Portfolio management, Reinforcement learning, Policy gradient, Tracking difference

doi:10.6342/NTU201801196

CONTENTS

doi:10.6342/NTU201801196

LIST OF FIGURES

doi:10.6342/NTU201801196

LIST OF TABLES

doi:10.6342/NTU201801196

Chapter 1 Introduction

In recent days, more and more people become distrustful of the fund managers and turn out to invest their money on ETFs to track the performance of the index (S&P500). According to researchers at Credit Suisse, in 1989, assets in index-tracking mutual funds totaled only $3 billion, around 1% of total US mutual fund assets. However, by the end of 2016, assets in index-tracking mutual funds and index ETFs had grown to a collective $5.1 trillion, 26% of the US mutual fund market. For those don't have information about constituents of index, portfolio weights of constituents or rules to build the index, index-tracking is not an easy job to deal with. They have to decide which assets should be included in portfolio and how much weight should be allocated to each selected assets while minimizing the tracking difference and the tracking difference. In this paper, we proposed a deep reinforcement learning method to build an index-tracking portfolio to track the index while not knowing the above information.

Machine Learning has achieved a great development on financial market. Support Vector Machine (SVM), Artificial Neural Network (ANN) and Long Short Term Memory (LSTM) have been used to deal with financial problems. One of its popular application is to predict the market movement. Many researchers believe that technical analysis could beat the market. In other words, we could predict the market movement by using the historical data. [1] has used ANN to forecast futures trading volume for six commodities traded on the Winnipeg Commodity Exchange. [2] investigates whether twitter mood could predict the stock market movement by using ANN-based model. [3], [4] proposed a hybrid SVM-based machine learning system for stock market forecasting. And [5] also predicts market by deploying LSTM networks which are considered a state-of-the-art

doi:10.6342/NTU201801196

technique for sequence learning. Though machine learning does a great job on forecasting, it's hard to use it on index-tracking problem. In general, supervised learning cannot deal with index-tracking problem, since we don't even know constituents of the index not to mention the target weight of corresponding assets.

Index-tracking is a kind of portfolio management, building a portfolio that tracks the performance of a certain index while minimizing the tracking difference and tracking error. In this paper, we use deep reinforcement learning (RL) model to build the index-tracking portfolio. Deep RL is drawing much attention due to remarkable achievement in playing complicated video games [6], [7]. [6], [7] deploy the deep Q-learning where actions space is discrete, and cannot be directly applied to portfolio management problems where actions are continuous. In financial management, deep Q-learning has been used for optimal asset allocation [9], [10]. However, due to discrete actions space of Q-learning, the proposed model can only choose which assets to invest. As a result, we turn our model to a general-purpose continuous deep RL framework, the Policy Gradient Algorithms. [10] deploys Policy Gradient with continuous actions space to build a cryptocurrency portfolio. The goal of [10] is to maximize the portfolio returns in the testing period which is not the same as ours, tracking the performance of the index.

In this paper, we proposed a deep RL method with Policy Gradient to build an index-tracking portfolio. We use mean square of tracking difference between the index-tracking portfolio and the index as performance measure. The proposed deep RL method will select the assets from whole financial market and allocate portfolio weights to corresponding assets. In chapter 2, the index-tracking problem will be described in detail. In chapter 3, we introduce the reinforcement learning with Policy Gradients and make it to fit with the index-tracking problem. The experimental results will be presented in chapter 4. Finally, chapter 5 gives the conclusions.

doi:10.6342/NTU201801196

Chapter 2 Index-Tracking Problem

In this chapter, we formulate the index-tracking problem. Suppose the portfolio consists of n assets. The return of the i^{th} asset $r_{i,t}$ and the return of index r_t^I at time t are defined as

$$r_{i,t} = \frac{v_{i,t}}{v_{i,t-1}} - 1, \qquad i = 1 \dots n \tag{2.1}$$

$$r_t^I = \frac{v_t^I}{v_{t-1}^I} - 1 \tag{2.2}$$

where $v_{i,t}$ and v_t^I are the close price of i^{th} asset and the index at time t respectively. The portfolio return at time t is:

$$r_t^p = \sum_{i=1}^{n} w_{i,t} r_{i,t} \tag{2.3}$$

where $w_{i,t}$ is the weight of i^{th} asset in portfolio at time t. $w_{i,t}$ follows

$$\sum_{i=1}^{n} w_{i,t} = 1 \ and \ w_{i,t} \geq 0, \qquad \forall i = 1 \dots n \tag{2.4}$$

doi:10.6342/NTU201801196

Tracking difference δ_t, which can be positive or negative, is the discrepancy between portfolio performance and index performance.

$$\delta_{\mathrm{t}} = r_t^P - r_t^I \tag{2.5}$$

The goal of index-tracking problem is to track the performance of target index. We define mean square of tracking difference δ_t as the performance measure.[1]

$$\frac{1}{T}\sum_{t=1}^{T}\delta_t^2 \tag{2.6}$$

where T is the holding period.

Our goal is not to outperform or to underperform the index. We want to track the performance of index as close as possible. As a consequence, we are going to minimize mean square of tracking difference.

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^{T}\delta_t^2 &= \frac{1}{T}\sum_{t=1}^{T}(r_t^P - r_t^I)^2 \\ &= \frac{1}{T}\sum_{t=1}^{T}\left(\sum_{i=1}^{n} w_{i,t} r_{i,t} - r_t^I\right)^2\end{aligned} \tag{2.7}$$

1 Mean square of tracking difference is actually the same as mean square error of portfolio returns.

doi:10.6342/NTU201801196

Chapter 3 Reinforcement Learning

Reinforcement learning (RL) is an interactive form of learning between an agent (investor) and its environment (market). The agent is provided with information about its environment. The agent then learns to act, without explicitly being told what to do. It discovers by itself the desirable actions from reward obtained for trying those actions. The only goal for the agent is to maximize the reward it gets. Figure 3.1 shows the entire learning process.

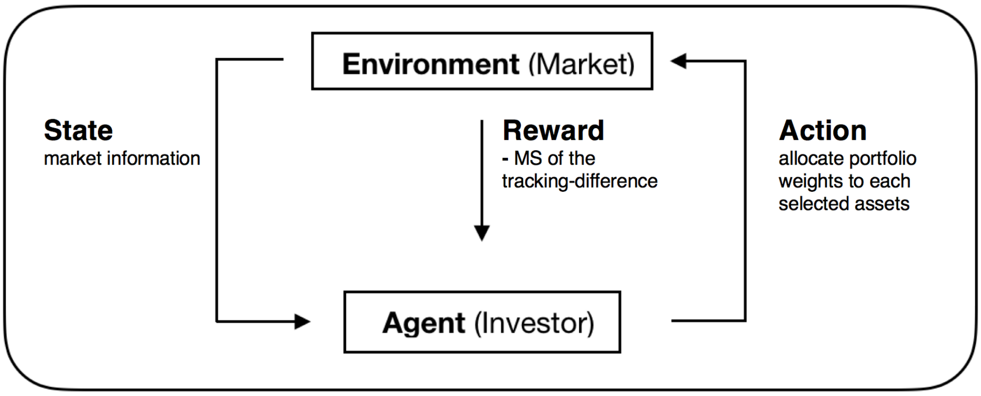

Fig. 3.1 Reinforcement Learning

3.1 Reward Function and Actions

We define the state information provided by market to be the historical returns of assets $S_t \in \mathcal{S}^{n\times k}$ where n is number of assets in the portfolio and k is the number of past period returns, in other words, $S_t = \left(r_{1,t}, r_{2,t}, \dots, r_{n,t}\right)^T and\ r_{i,t} \in \Re^k$. After knowing the state information, the investor chooses portfolio weights $w_t \in \mathcal{W}^n$ of selected assets

doi:10.6342/NTU201801196

and then execute the transaction. The market will return the next state information and a reward ρ_t, the minus of mean square of tracking difference.

$$\rho_t(S_t, w_t) = -\frac{1}{h}\sum_{l=t+1}^{t+h} \delta_l^2, \quad t = 0, 1, \dots, T-h \tag{3.1}$$

where h is the holding period. Note that the market will return a reward that comes from the tracking difference for the next h periods.

For more details, we illustrate inner change of ρ_t. The investor only chooses portfolio weights once in each state. However, w_t will change period by period due to market price change. We add a superscript on portfolio weight to show this.

$$w_t^l = \frac{(1+r_l) \otimes w_t^{l-1}}{(1+r_l) \cdot w_t^{l-1}}, \qquad l = t+1, t+2, \dots, t+h \tag{3.2}$$

where $\otimes$ is the element-wise multiplication operator.

The reward becomes

$$\begin{aligned}\rho_t(S_t, w_t) &= -\frac{1}{h}\sum_{l=t+1}^{t+h} \delta_l^2 \\ &= -\frac{1}{h}\sum_{t+1}^{t+h} (r_l^P - r_l^I)^2 \\ &= -\frac{1}{h}\sum_{l=t+1}^{t+h} \left(w_t^l \cdot r_l - r_l^I\right)^2 \end{aligned} \tag{3.3}$$

doi:10.6342/NTU201801196

3.2 Policy Network

A policy π represents a stochastic mapping from the current state of the system to actions and is specified by a set of policy parameters $\theta \in \Theta$. An investor will allocate the portfolio weights according to his policy, $\pi_\theta : \mathcal{S}^{n \times k} \to \mathcal{W}^n$ where $\mathcal{S}$ and $\mathcal{W}$ are the historical information space and portfolio weights space respectively.

3.2.1 Multilayer Perceptron

A multilayer perceptron (MLP) is a deep, artificial neural network. It is composed of more than one perceptron. They are composed of an input layer to receive the signal, an output layer that makes a decision or prediction about the input, and in between those two, an arbitrary number of hidden layers that are the true computational engine of the MLP. Figure 3.2 shows the structure of MLP.

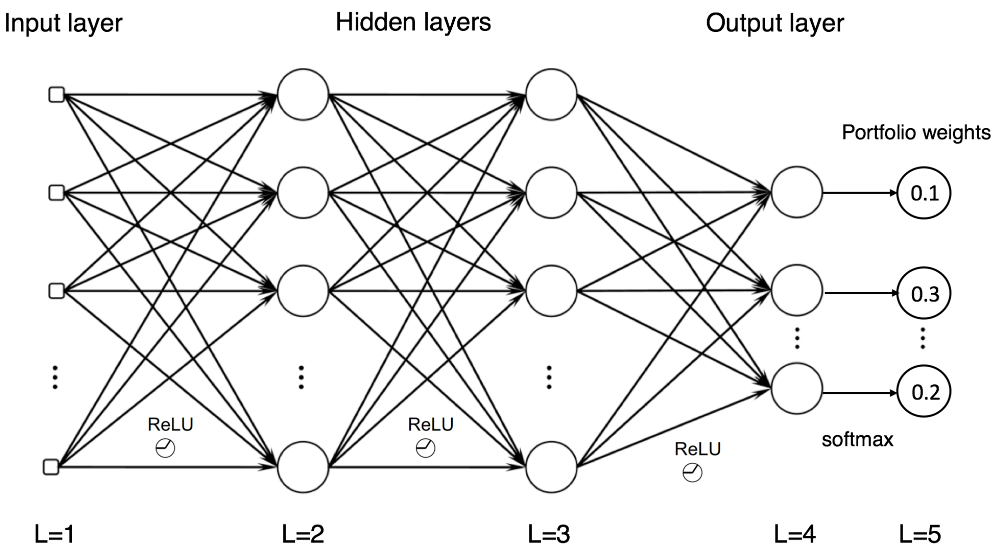

Fig. 3.2 Multilayer Perceptron

Suppose we only use a period of historical returns, $k = 1$. The input of the learning model is historical returns of assets $S_t \in \Re^n$. The activation function between hidden

doi:10.6342/NTU201801196

layers is Rectified Linear Unit (ReLU), $f(x) = \max(x, 0)$. The output activation function is softmax function $g(x)_i = \frac{e^{x_i}}{\sum_{i=1}^{n} e^{x^i}}$ which returns a positive portfolio weight and makes the sum of portfolio weights equal to one. Suppose there are L layers in the MLP and there are s_l units in layer l. Network parameters between layer l and layer $l+1$ are $\theta^{(l)} \in \Re^{s_{l+1} \times (s_l + 1)}$. Now we can formulate the feedforward neural network.

$$x_1 = (1, S_t^T)^T = \left(1, r_{1,t}, r_{2,t}, \dots, r_{n,t}\right)^T$$

$$z_l = f\left(\theta^{(l-1)} x_{l-1}\right)$$

$$x_l = (1, z_l^T)^T, \qquad \forall\, l = 2, 3, \dots, L$$

$$x_L = g(x_{L-1}) = w_t$$

Put it together, we obtain our policy π_θ which maps the historical returns to the portfolio weights.

3.3 Policy Gradient

Policy gradient methods directly store and iteratively improve a parametric approximation of the optimal policy. The goal of policy optimization is to optimize the policy parameters $\theta \in \Theta$ so as to maximize the objective function $J : \Theta \to \Re$

$$\theta^* = \underset{\theta \in \Theta}{argmax} J(\theta;\, \pi) \tag{3.5}$$

We define the objective function by using the reward function in (3.1).

doi:10.6342/NTU201801196

$$J(\theta;\,\pi) = \frac{\rho\big(S_1, \pi_\theta(S_1), S_2, \pi_\theta(S_2), \dots, S_{T-h}, \pi_\theta(S_{T-h})\big)}{\mathrm{T-h}}$$

$$= \frac{1}{T-h}\sum_{t=1}^{T-h} \rho_t(S_t, w_t)$$

$$= -\frac{1}{h(T-h)}\sum_{t=1}^{T-h}\sum_{l=t+1}^{t+h} \delta_l^2 \tag{3.6}$$

where h is the holding period in each state and T is the training period in specific date interval.

To obtain the optimal policy π^*, gradient ascent algorithm is used. After random initialization, the parameters are continuously updated along the gradient direction with a learning rate λ,

$$\theta_{k+1} = \theta_k + \lambda \nabla_\theta J(\theta_k) \tag{3.7}$$

3.4 Portfolio Formation

In this section, we illustrate the process of portfolio formation. Figure 3.4 shows the process of portfolio formation. The reinforcement learning agent will output the portfolio weights of input assets. However, we don't know the constituents of index, we have to define which assets should be selected as input assets. First, we screen out assets from whole financial market. The reinforcement learning agent will allocate less portfolio weights to assets that are insignificant to portfolio. As a consequence, we could screen out the assets whose portfolio weights are too little.

doi:10.6342/NTU201801196

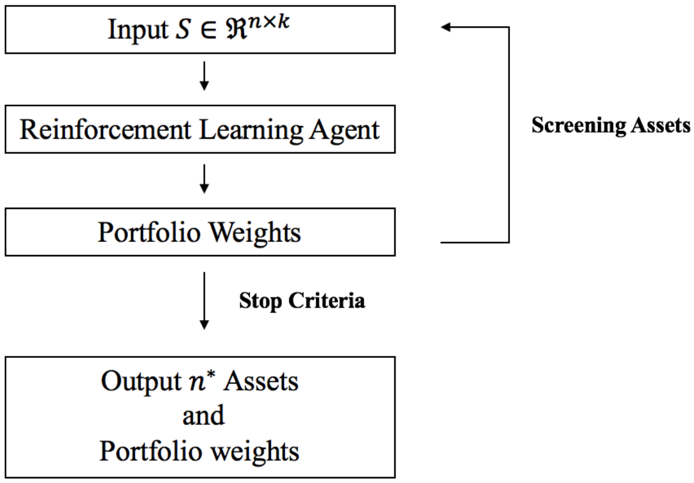

Fig. 3.3 Portfolio Formation

In each assets selection, we reduce by half the number of assets whose portfolio weights are smaller until the number of assets in the portfolio less than optimal number of assets n^*. And then we choose the first n* assets whose portfolio weights are greater as our final output. For a specific asset, its significance level to portfolio may change in each training process, so we don't directly choose the first n* assets whose portfolio weights are greater in the first assets selection. This allows us to avoid omitting the important assets.

doi:10.6342/NTU201801196

Chapter 4 Experiment Results

4.1 Experiment Data

We constrain experiment data in equity class though the proposed reinforcement learning model is not limit to equity class. The U.S. stock universe consists of all common equity in CRSP (sharecodes 10, 11 and 12) from January 2002 to December 2017. We exclude ADRs, REITs and closed-end funds to avoid including index-tracking ETF. Daily and adjust close price in CRPS will be used in experiment.

Source of index data is *Yahoo Finance.* We aim to track performance of the index, however, tracking index such as S&P 500 is not a difficult problem since S&P 500 is built by market capitalization and the constituents seldom change. In order to utilize our deep reinforcement learning method, we include several active funds as the tracked index. We obtain the active funds according to following procedures. First, select all equity funds that survive from January 2002 to December 2017. Secondly, select funds whose 1-year, 3-year and 5-year Sharp ratio are greater than 1 and 10-year Sharp ratio is greater than 0.8. Finally, select funds whose cumulative return from January 2002 to December 2017 are greater than S&P 500 and NASDAQ Composite.

Table 4.1 Description of Indexes and Active Funds

Index/Fund Name	Symbol	Category	Inception Date
S&P 500	GSPC	-	-
NASDAQ Composite	IXIC	-	-
Fidelity Select Software & IT Services Port	FSCSX	Technology	29-Jul-85
Fidelity Select IT Services	FBSOX	Technology	4-Feb-98
Shelton Nasdaq-100 Index Direct	NASDX	Large Growth	18-Jan-00

doi:10.6342/NTU201801196

The table 4.1 gives the description of tracked indexes and active funds. The figure 4.1 shows their cumulative returns from January 2002 to December 2017. Obviously, the selected active funds have much higher cumulative return.

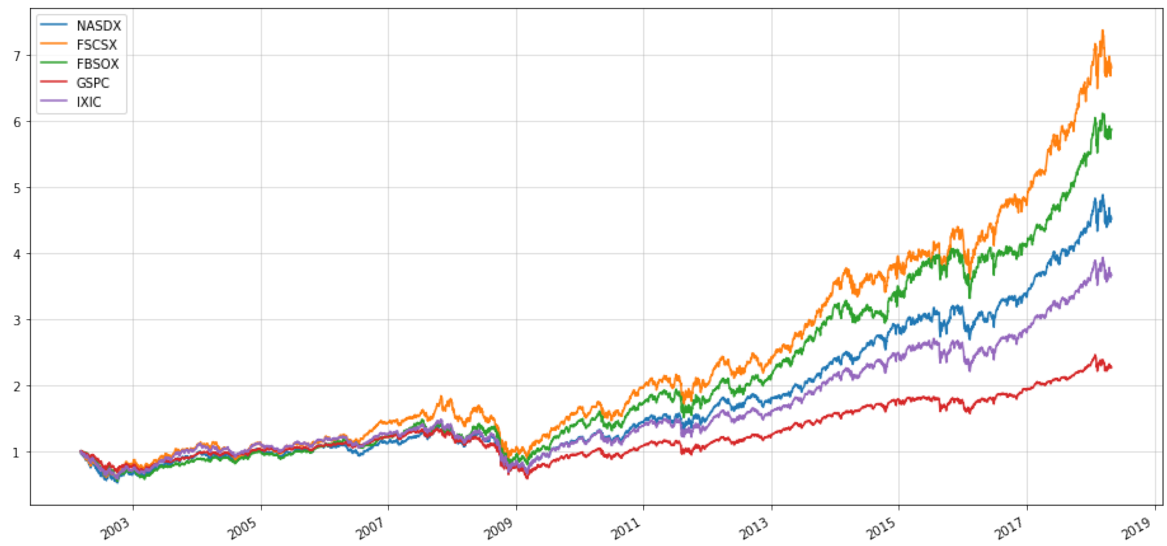

Fig. 4.1 Cumulative Returns of Tracked Indexes and Active Fund

4.2 Results

In the experiment, we roll the window to test whether index-tracking portfolio could track the index or the active funds. We hold the portfolio for 22/66 trading days (about a month/a quarter) in the testing period and then rebalance the portfolio in the end of testing period. As shown in figure 4.2, the rolling window will keep moving forward. In the last rolling window, if testing period is less than 22/66 trading days, the previous testing period will extend to the end of data date. Finally, we concatenate testing period in each window together and compare results with the tracked index.

doi:10.6342/NTU201801196

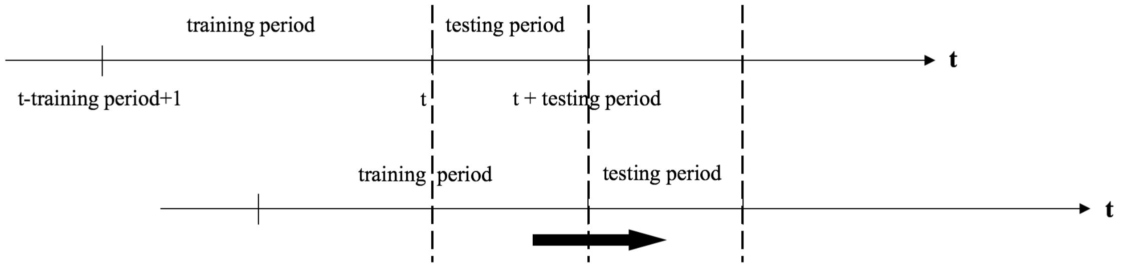

Fig. 4.2 Rolling Window

The optimal numbers of stocks n^* in portfolio are different among indexes and active funds. The optimal number in S&P 500, NASDAQ Composite, FSCSX, FBSOX and NASDX index-tracking portfolios is 500, 100, 100, 60 and 50 respectively. The figure 4.3 and the figure 4.4 show cumulative returns of index-tracking portfolio with different testing period. We can see that testing period affect tracking ability a lot. GSPC (S&P 500), IXIC (NASDAX Composite) and NASDX have better tracking ability with 22 testing period. On the other hand, FSCSX and FBSOX have better tracking ability with 66 testing period. However, no matter which testing period we choose, all of the index-tracking portfolio excellently track the indexes and active funds. That is a surprising result since we select stocks from the whole U.S. equity market which consists of 3000~5000 common stocks (exclude ADRs, REITs and closed-end funds) in each period.

doi:10.6342/NTU201801196

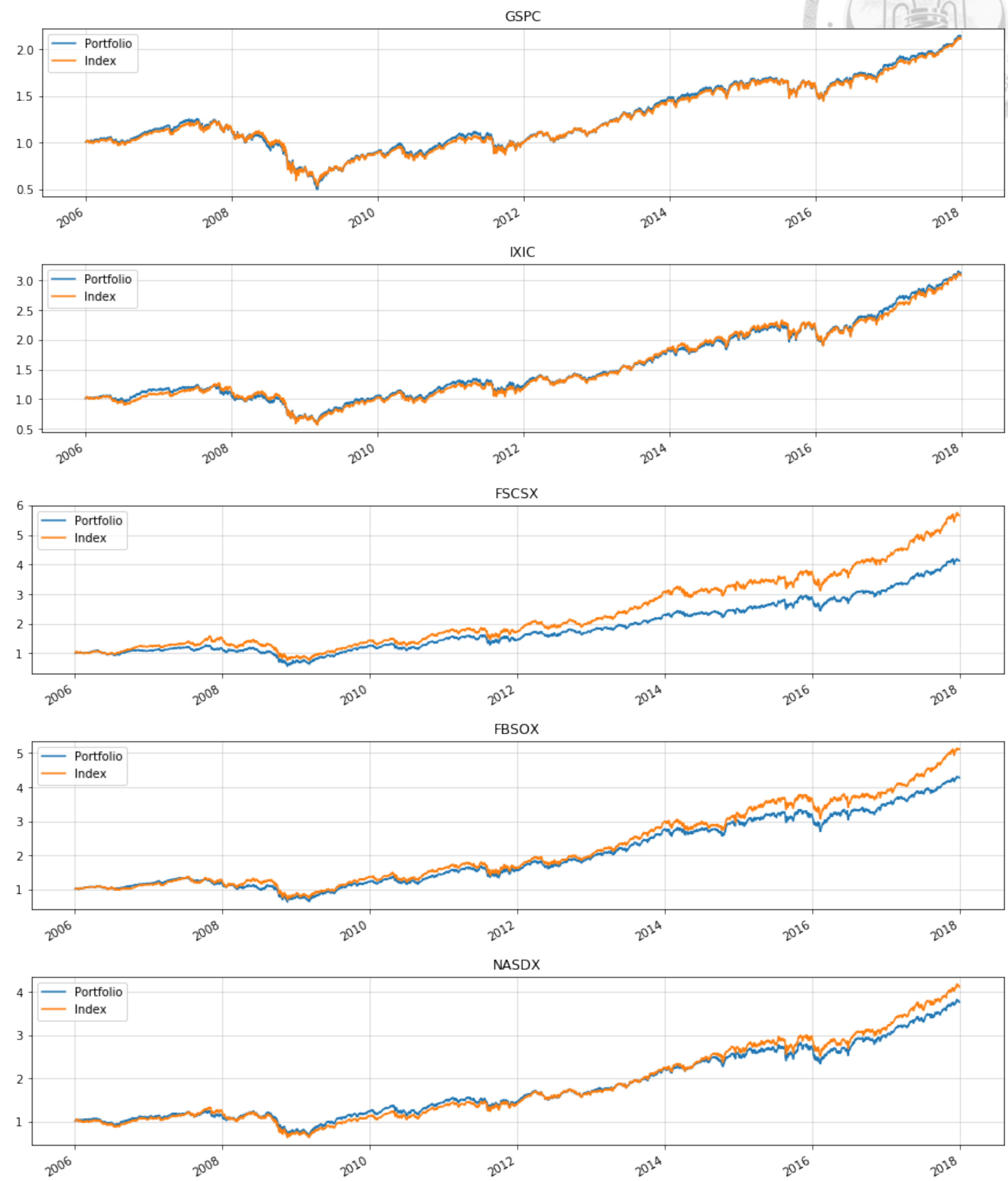

Fig. 4.3 Cumulative Returns of Index-Tracking Portfolio and Index (Testing Period: 22 Trading Days)

doi:10.6342/NTU201801196

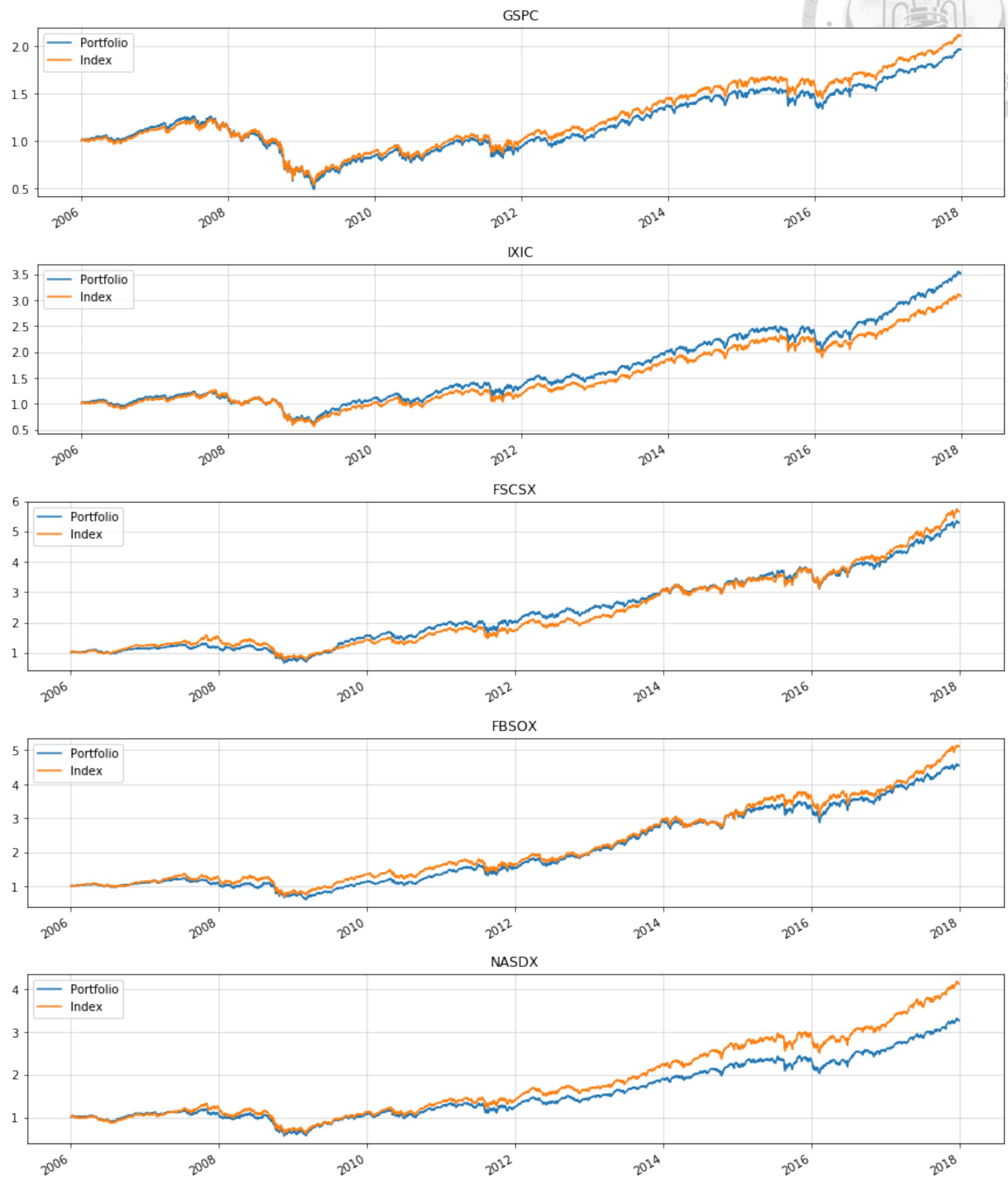

Fig. 4.4 Cumulative Returns of Index-Tracking Portfolio and Index (Testing Period: 66 Trading Days)

doi:10.6342/NTU201801196

We further compare GSPC and IXIC index-tracking portfolios with ETFs.[2] SPDR S&P 500 ETF Trust (SPY) and Fidelity NASDAQ Comp. Index Trk Stk (ONEQ) are the index-tracking ETFs of S&P500 and NASDAX Composite respectively. SPY and ONEQ are the most actively traded index-tracking ETFs and they perfectly track the indexes. The figure 4.5 presents the distribution of tracking difference of index-tracking portfolios and ETFs. All distributions of tracking difference of index-tracking portfolios are bell-shape and centered around zero. The distribution of the index-tracking portfolio with 22 testing period is almost the same as the portfolio with 66 testing period. For GSPC, ETF fits

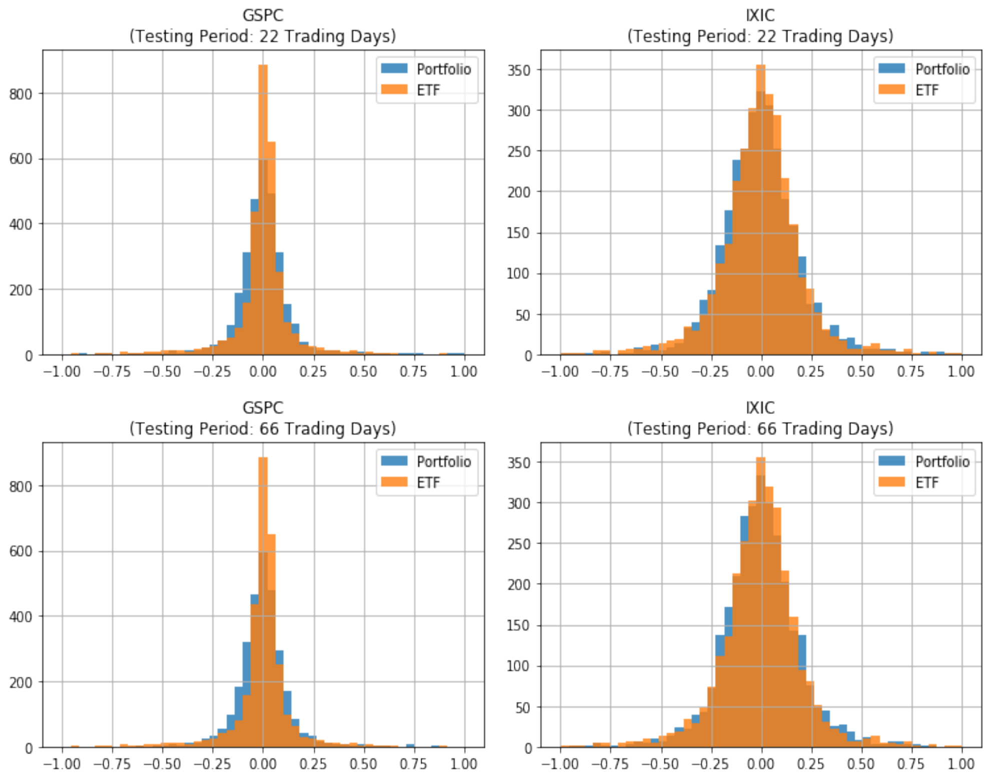

Fig. 4.5 Distribution of Tracking Difference

[2] There is no ETFs for active funds such as FSCSX, FBSOX and NASDX.

doi:10.6342/NTU201801196

better with the index since its distribution of tracking difference has a higher peak and is more centered around zero. For IXIC, the distribution of ETF has only a little difference with the distribution of index-tracking portfolio. The tracking ability may be indifferent according to their distribution.

Our main measurement of index-tracking problem is mean square of tracking difference. Table 4.2 and table 4.3 give a brief summary of tracking difference year by year. In all panels of both tables, the standard deviation of tracking difference in 2008 and 2009 is relative larger than the tracking difference in other years. Besides, the minimum is smaller and the maximum is larger. We can infer that the index-tracing portfolio built by the deep reinforcement learning model do a bad job in 2008 and 2009. One of the possible reason is that the variability of the whole market is larger due to financial crisis which makes RL agents hard to learn. The other reason is that the original tracking difference expanded in 2008 and 2009 because of larger variability.

From the both tables, we can observe that the standard deviation of tracking difference of portfolios of actives funds, NASDX, FSCSX and FBSOX, are larger than index-tracking portfolios of indexes, S&P500 and NASDAQ Composite, in almost all years. Minimum and maximum also exhibit more deviated from zeros. The major reason is that active funds pay dividends while the index-tracking portfolio don’t. The active funds will have a price gap on ex-dividend date which makes tracking difference become larger.

doi:10.6342/NTU201801196

Table 4.2 Tracking Difference Description (Testing period: 22 Trading Days)
Table 4.2 gives the description of tracking difference. Std is the standard deviation of tracking difference. 25%, 50% and 75% are the first, the second and the third quartile respectively.

Year	Mean	Std	Min	25%	50%	75%	Max
Panel A: GSPC							
2006	0.0087%	0.0739%	-0.1867%	-0.0400%	0.0091%	0.0538%	0.2492%
2007	-0.0115%	0.1123%	-0.3808%	-0.0839%	-0.0088%	0.0505%	0.5773%
2008	0.0076%	0.3891%	-1.6064%	-0.1938%	-0.0263%	0.1736%	1.3924%
2009	0.0066%	0.4164%	-1.8917%	-0.1028%	0.0083%	0.1103%	2.0672%
2010	0.0079%	0.0920%	-0.2668%	-0.0518%	0.0113%	0.0669%	0.3488%
2011	-0.0133%	0.0970%	-0.2809%	-0.0713%	-0.0102%	0.0483%	0.2948%
2012	-0.0018%	0.0883%	-0.3560%	-0.0590%	0.0031%	0.0586%	0.1940%
2013	0.0096%	0.0720%	-0.2330%	-0.0326%	0.0104%	0.0458%	0.3760%
2014	-0.0047%	0.0707%	-0.2433%	-0.0465%	-0.0039%	0.0396%	0.2542%
2015	-0.0052%	0.0954%	-0.2778%	-0.0705%	-0.0056%	0.0512%	0.3044%
2016	0.0107%	0.1108%	-0.2989%	-0.0402%	0.0016%	0.0528%	0.6980%
2017	-0.0043%	0.0804%	-0.3223%	-0.0515%	-0.0086%	0.0360%	0.3642%
Panel B: IXIC							
2006	0.0264%	0.1927%	-0.5110%	-0.0864%	0.0145%	0.1236%	1.2570%
2007	-0.0555%	0.1972%	-0.8659%	-0.1522%	-0.0543%	0.0557%	1.0994%
2008	0.0312%	0.4776%	-3.4841%	-0.1664%	0.0233%	0.2060%	1.8949%
2009	0.0130%	0.2800%	-1.1900%	-0.1313%	-0.0197%	0.1297%	1.8554%
2010	0.0063%	0.1707%	-0.3919%	-0.1214%	-0.0028%	0.1202%	0.6383%
2011	-0.0049%	0.1921%	-0.5852%	-0.1428%	-0.0096%	0.1263%	0.7248%
2012	-0.0086%	0.1830%	-0.6896%	-0.1299%	-0.0045%	0.1113%	0.7448%
2013	-0.0127%	0.1388%	-0.4569%	-0.1004%	-0.0037%	0.0604%	0.4378%
2014	0.0000%	0.1657%	-0.6055%	-0.0917%	-0.0009%	0.0988%	0.4571%
2015	0.0017%	0.1547%	-0.3832%	-0.0943%	-0.0093%	0.0909%	0.6851%
2016	0.0192%	0.1673%	-0.5413%	-0.0632%	0.0150%	0.0930%	0.9039%
2017	-0.0120%	0.1513%	-0.8982%	-0.0940%	-0.0107%	0.0749%	0.5633%
Panel C: NASDX							
2006	0.0145%	0.2478%	-0.9348%	-0.1394%	0.0243%	0.1566%	0.8375%
2007	-0.0499%	0.2881%	-1.0866%	-0.2040%	-0.0738%	0.0994%	1.4969%
2008	0.0594%	0.3994%	-1.4107%	-0.1794%	0.0343%	0.2798%	1.6178%
2009	0.0128%	0.2689%	-1.4542%	-0.1443%	0.0268%	0.1725%	0.6930%
2010	-0.0118%	0.2067%	-0.7869%	-0.1382%	-0.0064%	0.1306%	0.6449%
2011	-0.0119%	0.1986%	-0.5790%	-0.1452%	-0.0123%	0.1175%	0.6072%
2012	-0.0117%	0.1918%	-0.5219%	-0.1299%	-0.0158%	0.1040%	0.9138%
2013	-0.0074%	0.2046%	-0.6074%	-0.1285%	-0.0050%	0.1181%	0.9304%
2014	-0.0163%	0.1941%	-0.9655%	-0.1169%	-0.0048%	0.0923%	0.5580%
2015	-0.0121%	0.2818%	-2.3682%	-0.1235%	-0.0032%	0.0995%	1.8900%
2016	0.0084%	0.2369%	-1.3267%	-0.0938%	0.0052%	0.0961%	0.9490%

doi:10.6342/NTU201801196

2017	-0.0137%	0.1424%	-0.7133%	-0.1005%	-0.0014%	0.0592%	0.6758%

Year	Mean	Std	Min	25%	50%	75%	Max
Panel D: FSCSX							
2006	-0.0497%	0.4769%	-2.5825%	-0.3090%	-0.0259%	0.2778%	1.4537%
2007	-0.0588%	0.5413%	-1.8360%	-0.3502%	-0.0473%	0.2573%	2.9550%
2008	0.0276%	1.1621%	-4.7451%	-0.4443%	0.0326%	0.4605%	9.7248%
2009	0.0358%	0.5584%	-1.6563%	-0.2884%	-0.0003%	0.2752%	2.8018%
2010	-0.0120%	0.2484%	-0.7257%	-0.1749%	-0.0136%	0.1434%	0.7875%
2011	-0.0075%	0.2511%	-0.8673%	-0.1504%	0.0013%	0.1575%	0.8099%
2012	0.0038%	0.2612%	-0.8519%	-0.1395%	-0.0114%	0.1404%	1.1724%
2013	-0.0490%	0.2601%	-0.8400%	-0.2269%	-0.0379%	0.1307%	0.6352%
2014	-0.0108%	0.7341%	-7.5772%	-0.2054%	-0.0058%	0.1985%	6.9501%
2015	0.0171%	0.2595%	-1.1505%	-0.1338%	0.0117%	0.1585%	1.1167%
2016	-0.0085%	0.2946%	-2.1559%	-0.1268%	-0.0009%	0.1380%	1.1810%
2017	-0.0138%	0.1832%	-0.6060%	-0.1227%	-0.0164%	0.0825%	0.6319%
Panel E: FBSOX							
2006	0.0105%	0.3651%	-1.8047%	-0.1563%	-0.0075%	0.2183%	1.5595%
2007	-0.0416%	0.5054%	-2.2813%	-0.3954%	-0.0218%	0.2783%	1.5126%
2008	-0.0180%	0.8553%	-2.5227%	-0.4873%	-0.0260%	0.4531%	3.2507%
2009	0.0003%	0.6171%	-1.4386%	-0.3538%	-0.0333%	0.3389%	1.9297%
2010	0.0022%	0.3569%	-0.8574%	-0.2437%	-0.0177%	0.2272%	1.0291%
2011	0.0197%	0.3385%	-0.9588%	-0.1861%	0.0064%	0.2069%	1.7020%
2012	0.0040%	0.2812%	-0.8972%	-0.1884%	-0.0032%	0.1467%	1.3100%
2013	-0.0164%	0.1871%	-0.5553%	-0.1381%	-0.0019%	0.1002%	0.8036%
2014	0.0003%	0.7181%	-8.1514%	-0.1283%	-0.0081%	0.1446%	6.8202%
2015	-0.0289%	0.1871%	-0.6607%	-0.1499%	-0.0265%	0.0774%	0.4859%
2016	0.0154%	0.2318%	-0.8751%	-0.1176%	0.0112%	0.1377%	0.8111%
2017	-0.0317%	0.1922%	-0.6112%	-0.1450%	-0.0348%	0.0557%	0.7165%

doi:10.6342/NTU201801196

Table 4.3 Tracking Difference Description (Testing period: 66 Trading Days)

Table 4.3 gives the description of tracking difference. Std is the standard deviation of tracking difference. 25%, 50% and 75% are the first, the second and the third quartile respectively.

Year	Mean	Std	Min	25%	50%	75%	Max
Panel A: GSPC							
2006	0.0096%	0.0659%	-0.1875%	-0.0335%	0.0094%	0.0543%	0.1969%
2007	-0.0097%	0.1111%	-0.3771%	-0.0695%	-0.0086%	0.0456%	0.5080%
2008	-0.0080%	0.3201%	-1.2547%	-0.1750%	-0.0230%	0.1359%	1.5887%
2009	-0.0097%	0.3359%	-1.6907%	-0.1231%	-0.0122%	0.0877%	1.6052%
2010	0.0046%	0.1025%	-0.3190%	-0.0540%	0.0016%	0.0705%	0.3034%
2011	-0.0098%	0.0906%	-0.2728%	-0.0743%	-0.0062%	0.0481%	0.2493%
2012	-0.0008%	0.0867%	-0.3752%	-0.0555%	0.0043%	0.0609%	0.2106%
2013	0.0049%	0.1074%	-0.4172%	-0.0502%	0.0068%	0.0630%	0.4100%
2014	-0.0064%	0.0822%	-0.5013%	-0.0500%	-0.0066%	0.0395%	0.2270%
2015	-0.0070%	0.0849%	-0.4491%	-0.0626%	-0.0061%	0.0504%	0.3084%
2016	0.0087%	0.1109%	-0.2983%	-0.0541%	0.0058%	0.0557%	0.5248%
2017	-0.0023%	0.0738%	-0.3354%	-0.0456%	-0.0021%	0.0326%	0.2815%
Panel B: IXIC							
2006	0.0172%	0.1825%	-0.5428%	-0.0970%	0.0196%	0.1358%	0.5958%
2007	-0.0383%	0.2071%	-0.8083%	-0.1466%	-0.0359%	0.0642%	1.2699%
2008	0.0397%	0.3208%	-1.0797%	-0.1353%	0.0290%	0.2037%	1.4001%
2009	0.0170%	0.3137%	-1.2500%	-0.1115%	0.0082%	0.1534%	1.3972%
2010	0.0012%	0.1775%	-0.5267%	-0.1201%	-0.0033%	0.0946%	0.6837%
2011	0.0137%	0.1881%	-0.5303%	-0.1066%	0.0019%	0.1314%	0.6992%
2012	-0.0018%	0.1538%	-0.4689%	-0.0915%	0.0003%	0.0930%	0.5161%
2013	-0.0135%	0.1590%	-0.5127%	-0.1085%	-0.0147%	0.0843%	0.5387%
2014	0.0046%	0.1459%	-0.5149%	-0.0892%	-0.0074%	0.0910%	0.5007%
2015	-0.0124%	0.1530%	-0.5955%	-0.1115%	-0.0055%	0.0873%	0.4333%
2016	0.0215%	0.1702%	-0.6137%	-0.0845%	0.0189%	0.1203%	0.5745%
2017	0.0028%	0.1373%	-0.7663%	-0.0834%	0.0027%	0.0745%	0.7190%
Panel C: NASDX							
2006	0.0101%	0.2258%	-0.9240%	-0.1259%	-0.0018%	0.1700%	0.6517%
2007	-0.0628%	0.2808%	-1.0655%	-0.2245%	-0.0722%	0.1025%	1.4036%
2008	0.0017%	0.4743%	-1.9094%	-0.2134%	0.0043%	0.2465%	1.8909%
2009	0.0327%	0.3290%	-1.1436%	-0.1329%	0.0099%	0.1788%	1.5377%
2010	-0.0165%	0.2046%	-0.7646%	-0.1540%	-0.0214%	0.1190%	0.6536%
2011	-0.0152%	0.1897%	-0.5730%	-0.1445%	-0.0202%	0.1140%	0.5494%
2012	-0.0091%	0.1935%	-0.4855%	-0.1238%	-0.0060%	0.1069%	0.8331%
2013	-0.0082%	0.1852%	-0.5772%	-0.1258%	-0.0058%	0.0982%	0.6501%
2014	-0.0052%	0.1731%	-0.4813%	-0.1061%	-0.0028%	0.0964%	0.5321%
2015	-0.0189%	0.2696%	-2.2396%	-0.1237%	-0.0120%	0.0951%	1.9855%
2016	0.0067%	0.2135%	-0.7727%	-0.1037%	-0.0080%	0.1100%	1.0578%
2017	-0.0123%	0.1664%	-0.6800%	-0.1114%	-0.0194%	0.0729%	0.7399%

doi:10.6342/NTU201801196

Year	Mean	Std	Min	25%	50%	75%	Max
Panel D: FSCSX							
2006	-0.0307%	0.4254%	-1.2664%	-0.3163%	-0.0211%	0.2828%	1.5778%
2007	-0.0741%	0.5480%	-2.0672%	-0.3950%	-0.0338%	0.2641%	2.1557%
2008	0.0503%	0.9170%	-2.6856%	-0.4582%	0.0383%	0.4723%	5.9839%
2009	0.0937%	0.9720%	-3.6476%	-0.3617%	0.0417%	0.4938%	4.3411%
2010	0.0099%	0.2425%	-0.8085%	-0.1261%	0.0058%	0.1636%	0.7657%
2011	0.0143%	0.2751%	-0.8907%	-0.1405%	0.0202%	0.1619%	0.9577%
2012	-0.0005%	0.2482%	-1.2700%	-0.1508%	-0.0001%	0.1301%	1.0549%
2013	-0.0612%	0.2387%	-0.7455%	-0.2162%	-0.0648%	0.1017%	0.5496%
2014	-0.0008%	0.7317%	-8.0238%	-0.1787%	-0.0018%	0.1832%	6.8052%
2015	-0.0068%	0.2589%	-1.2809%	-0.1426%	-0.0045%	0.1305%	1.2161%
2016	-0.0091%	0.2820%	-1.5187%	-0.1312%	-0.0054%	0.1305%	0.9481%
2017	-0.0170%	0.1947%	-0.5707%	-0.1430%	-0.0045%	0.0901%	0.5597%
Panel E: FBSOX							
2006	-0.0111%	0.3532%	-1.4851%	-0.2229%	-0.0078%	0.2093%	1.1156%
2007	-0.0496%	0.5316%	-2.7348%	-0.3839%	-0.0352%	0.2751%	1.7698%
2008	0.0050%	0.8817%	-3.0185%	-0.4752%	0.0197%	0.4844%	2.9975%
2009	-0.0269%	0.7382%	-2.7327%	-0.4311%	-0.0645%	0.4436%	2.0437%
2010	0.0013%	0.4089%	-1.0734%	-0.2958%	0.0067%	0.2275%	1.4687%
2011	0.0438%	0.3385%	-0.6805%	-0.1761%	0.0151%	0.2295%	1.3909%
2012	0.0199%	0.2633%	-0.6445%	-0.1477%	0.0190%	0.1568%	1.0349%
2013	-0.0049%	0.2014%	-0.5774%	-0.1255%	-0.0127%	0.1270%	0.8489%
2014	-0.0019%	0.7254%	-8.0445%	-0.1446%	-0.0058%	0.1019%	7.0972%
2015	-0.0267%	0.2096%	-0.7079%	-0.1509%	-0.0347%	0.0986%	0.6263%
2016	0.0202%	0.2329%	-0.6356%	-0.1268%	-0.0014%	0.1397%	0.7745%
2017	-0.0325%	0.2141%	-0.6383%	-0.1680%	-0.0214%	0.0891%	0.6450%

doi:10.6342/NTU201801196

The deep reinforcement learning model will minimize the mean square of tracking difference (MSTD). We expect low MSTD in the end. However, the question is how low is enough for an index-tracking problem. In order to have a concrete understanding about MSTD, we introduce several measurements, Correlation, R Square, Beta and Tracking Error[3], to compare with MSTD. Returns of the index-tracking portfolio and returns of the index are used to calculated the above measurements. Beta is the regression beta coefficient without interception. If Beta is closer to 1, returns of index-tracking portfolio and returns of the index will have more consistent movement, in other words, the tracking ability will be better. Tracking error is the standard deviation of tracking difference. While tracking difference measures the extent to which an index return differs from that of its benchmark index, tracking error indicates how much variability exists among the individual data points that make up the index-tracking portfolio's average tracking difference. Table 4.4 and table 4.5 present the results of tracking measurements. Note that no matter which indexes or active funds we track, the MSTD in 2008 and 2009 is higher. Besides, MSTD of indexes is lower than MSTD of active funds. The result coincides with the results we observe from cumulative returns in figure 4.3 and figure 4.4 and deviation of maximum or minimum of tracking difference in table 4.2 and table 4.3. However, we do not get consistent results from other tracking measurements. For indexes which considered to have a better fit, when MSTD and tracking error reach to their maximum in 2008 and 2009, correlation and R square do not seem to be the lowest. Beta doesn't have a much deviation from one either. Active funds also present the same situation. One

[3] Tracking Error $= std(r^P - r^I)$ which is standard deviation of tracking difference.

doi:10.6342/NTU201801196

of the possible reason is that though MSTD and tracking error reach to their maximum in 2008 and 2009, the direction of movement of index-tracking portfolios and the indexes are similar in the period of financial crisis. As a result, correlation and R square do not decline much. Beta will not change much either. For an investor, we not only consider the scale of deviation from the tracked indexes but also the direction of movement, which give us room for improvement.

For indexes in panel A and panel B and NASDX in panel C, we observe that MSTD and Tracking Error are lower, Correlation and R square are higher, and beta is closer to one in all years with 66 testing period. It indicates that when the testing period is 66 trading days, the tracking ability of index-tracking portfolio is better. This is totally different from the cumulative return we've observed. For active funds in panel D and panel E, MSTD and Tracking Error are lower, Correlation and R square are higher, and beta is closer to one in all years with 22 testing period. It indicates that when the testing period is 22 trading days, the tracking ability of index-tracking portfolio is better. This is totally different from the cumulative return either. The similar situation occurs in active funds. One of the possible reason is that the positive and negative part of tracking difference offset each other, so the discrepancy of cumulative returns between index-tracking portfolio and target becomes smaller. However, the tracking difference itself may be large, which leads to bad results in all tracking measurements. The tracking measurements indeed give us more insight which we cannot observe in cumulative return.

doi:10.6342/NTU201801196

Table 4.4 Tracking Measurement (Testing Period: 22 Trading Days)
Table 4.4 gives the tracking measurement of the index-tracking portfolio with 22 testing period. Correlation means the correlation between returns of index-tracking portfolio and return of the index. We regress the returns of the index-tracking portfolio on returns of index without interception to get R square and Beta. Tracking error is the standard deviation of tracking difference. MSTD is mean square of tracking difference. In Year column, 'all' stands for years from 2006 to 2017.

Year	Correlation	R square	Beta	Tracking Error	MSTD
Panel A: GSPC					
2006	0.9932	0.9863	0.9677	7.39E-04	5.51E-07
2007	0.9939	0.9877	0.9723	1.12E-03	1.27E-06
2008	0.9900	0.9801	1.0315	3.89E-03	1.51E-05
2009	0.9793	0.9591	1.0892	4.16E-03	1.73E-05
2010	0.9969	0.9938	0.9735	9.20E-04	8.49E-07
2011	0.9978	0.9955	0.9936	9.70E-04	9.54E-07
2012	0.9940	0.9880	0.9931	8.83E-04	7.78E-07
2013	0.9948	0.9897	0.9735	7.20E-04	5.25E-07
2014	0.9958	0.9917	0.9538	7.07E-04	4.99E-07
2015	0.9953	0.9905	0.9810	9.54E-04	9.08E-07
2016	0.9910	0.9820	0.9963	1.11E-03	1.23E-06
2017	0.9818	0.9649	0.9459	8.04E-04	6.46E-07
all	0.9894	0.9789	1.0175	1.84E-03	3.39E-06
Panel B: IXIC					
2006	0.9762	0.9522	0.9594	1.93E-03	3.77E-06
2007	0.9837	0.9652	0.9601	1.97E-03	4.18E-06
2008	0.9835	0.9672	1.0022	4.78E-03	2.28E-05
2009	0.9894	0.9792	1.0392	2.80E-03	7.83E-06
2010	0.9909	0.9820	0.9539	1.71E-03	2.90E-06
2011	0.9928	0.9856	0.9990	1.92E-03	3.68E-06
2012	0.9812	0.9628	0.9746	1.83E-03	3.34E-06
2013	0.9837	0.9684	0.9750	1.39E-03	1.94E-06
2014	0.9843	0.9689	0.9121	1.66E-03	2.73E-06
2015	0.9894	0.9790	0.9657	1.55E-03	2.39E-06
2016	0.9860	0.9718	0.9689	1.67E-03	2.83E-06
2017	0.9691	0.9407	0.8967	1.51E-03	2.30E-06
all	0.9854	0.9711	0.9890	2.25E-03	5.07E-06
Panel C: NASDX					
2006	0.9710	0.9426	0.8800	2.48E-03	6.14E-06
2007	0.9714	0.9421	0.8803	2.88E-03	8.52E-06
2008	0.9884	0.9765	0.9865	3.99E-03	1.62E-05
2009	0.9893	0.9790	1.0491	2.69E-03	7.22E-06
2010	0.9859	0.9720	0.9522	2.07E-03	4.27E-06
2011	0.9910	0.9820	0.9882	1.99E-03	3.94E-06

doi:10.6342/NTU201801196

2012	0.9800	0.9605	0.9597	1.92E-03	3.68E-06
2013	0.9641	0.9314	0.9270	2.05E-03	4.17E-06
2014	0.9776	0.9559	0.8952	1.94E-03	3.78E-06
2015	0.9688	0.9385	0.9125	2.82E-03	7.92E-06
2016	0.9720	0.9447	0.9447	2.37E-03	5.60E-06
2017	0.9769	0.9557	0.8839	1.42E-03	2.04E-06
all	0.9821	0.9646	0.9655	2.48E-03	6.13E-06
Panel D: FSCSX					
2006	0.8454	0.7134	0.8392	4.77E-03	2.29E-05
2007	0.8852	0.7824	0.7935	5.41E-03	2.95E-05
2008	0.9101	0.8288	0.9577	1.16E-02	1.35E-04
2009	0.9549	0.9133	1.1022	5.58E-03	3.12E-05
2010	0.9810	0.9624	0.9121	2.48E-03	6.16E-06
2011	0.9890	0.9782	0.9725	2.51E-03	6.29E-06
2012	0.9675	0.9364	0.9907	2.61E-03	6.80E-06
2013	0.9512	0.9050	0.9130	2.60E-03	6.98E-06
2014	0.8219	0.6758	0.5900	7.34E-03	5.37E-05
2015	0.9716	0.9438	0.9541	2.59E-03	6.73E-06
2016	0.9639	0.9291	0.9224	2.95E-03	8.65E-06
2017	0.9660	0.9354	0.8967	1.83E-03	3.36E-06
all	0.9288	0.8629	0.9288	5.14E-03	2.65E-05
Panel E: FBSOX					
2006	0.9069	0.8228	0.7369	3.65E-03	1.33E-05
2007	0.8939	0.7980	0.7146	5.05E-03	2.56E-05
2008	0.9417	0.8871	0.8823	8.55E-03	7.29E-05
2009	0.9318	0.8699	0.9855	6.17E-03	3.79E-05
2010	0.9589	0.9198	0.8840	3.57E-03	1.27E-05
2011	0.9832	0.9664	0.9146	3.39E-03	1.15E-05
2012	0.9586	0.9195	0.9400	2.81E-03	7.88E-06
2013	0.9735	0.9493	0.9789	1.87E-03	3.51E-06
2014	0.8128	0.6610	0.6084	7.18E-03	5.14E-05
2015	0.9840	0.9675	0.9623	1.87E-03	3.57E-06
2016	0.9777	0.9556	0.9291	2.32E-03	5.38E-06
2017	0.9474	0.8986	0.9015	1.92E-03	3.78E-06
all	0.9402	0.8841	0.8790	4.56E-03	2.08E-05

doi:10.6342/NTU201801196

Table 4.5 Tracking Measurement (Testing Period: 66 Trading Days)
Table 4.5 gives the tracking measurement of the index-tracking portfolio with 66 testing period. Correlation means the correlation between returns of index-tracking portfolio and return of the index. We regress the returns of the index-tracking portfolio on returns of index without interception to get R square and Beta. Tracking error is the standard deviation of tracking difference. MSTD is mean square of tracking difference. In Year column, 'all' stands for years from 2006 to 2017.

Year	Correlation	R square	Beta	Tracking Error	MSTD
Panel A: GSPC					
2006	0.9949	0.9895	0.9617	6.59E-04	4.41E-07
2007	0.9939	0.9878	0.9795	1.11E-03	1.24E-06
2008	0.9927	0.9855	1.0142	3.20E-03	1.02E-05
2009	0.9861	0.9724	1.0732	3.36E-03	1.12E-05
2010	0.9962	0.9924	0.9691	1.03E-03	1.05E-06
2011	0.9981	0.9963	0.9855	9.06E-04	8.27E-07
2012	0.9942	0.9885	0.9828	8.67E-04	7.49E-07
2013	0.9884	0.9773	0.9546	1.07E-03	1.15E-06
2014	0.9940	0.9881	0.9526	8.22E-04	6.77E-07
2015	0.9967	0.9933	0.9625	8.49E-04	7.23E-07
2016	0.9910	0.9820	0.9933	1.11E-03	1.23E-06
2017	0.9849	0.9709	0.9435	7.38E-04	5.43E-07
all	0.9918	0.9837	1.0054	1.59E-03	2.51E-06
Panel B: IXIC					
2006	0.9788	0.9577	0.9442	1.83E-03	3.35E-06
2007	0.9824	0.9640	0.9359	2.07E-03	4.42E-06
2008	0.9924	0.9847	1.0012	3.21E-03	1.04E-05
2009	0.9878	0.9760	1.0574	3.14E-03	9.83E-06
2010	0.9898	0.9799	0.9656	1.78E-03	3.14E-06
2011	0.9931	0.9862	1.0039	1.88E-03	3.54E-06
2012	0.9867	0.9737	0.9727	1.54E-03	2.36E-06
2013	0.9787	0.9588	0.9407	1.59E-03	2.54E-06
2014	0.9877	0.9755	0.9289	1.46E-03	2.12E-06
2015	0.9900	0.9799	0.9525	1.53E-03	2.35E-06
2016	0.9855	0.9708	0.9610	1.70E-03	2.93E-06
2017	0.9760	0.9536	0.8937	1.37E-03	1.88E-06
all	0.9882	0.9766	0.9894	2.02E-03	4.08E-06
Panel C: NASDX					
2006	0.9770	0.9544	0.8824	2.26E-03	5.09E-06
2007	0.9722	0.9425	0.8940	2.81E-03	8.25E-06
2008	0.9837	0.9678	0.9493	4.74E-03	2.24E-05
2009	0.9853	0.9711	1.0715	3.29E-03	1.09E-05
2010	0.9861	0.9722	0.9613	2.05E-03	4.20E-06
2011	0.9920	0.9839	1.0037	1.90E-03	3.61E-06

doi:10.6342/NTU201801196

2012	0.9797	0.9599	0.9620	1.94E-03	3.74E-06
2013	0.9710	0.9444	0.9197	1.85E-03	3.42E-06
2014	0.9821	0.9648	0.9141	1.73E-03	2.99E-06
2015	0.9713	0.9432	0.9257	2.70E-03	7.27E-06
2016	0.9773	0.9551	0.9587	2.14E-03	4.55E-06
2017	0.9652	0.9336	0.9004	1.66E-03	2.77E-06
all	0.9807	0.9619	0.9618	2.57E-03	6.61E-06
Panel D: FSCSX					
2006	0.8701	0.7580	0.7997	4.25E-03	1.81E-05
2007	0.8821	0.7755	0.7841	5.48E-03	3.05E-05
2008	0.9400	0.8837	0.9357	9.17E-03	8.40E-05
2009	0.8617	0.7468	1.0582	9.72E-03	9.50E-05
2010	0.9818	0.9640	0.9165	2.43E-03	5.87E-06
2011	0.9878	0.9756	0.9320	2.75E-03	7.56E-06
2012	0.9693	0.9399	0.9501	2.48E-03	6.14E-06
2013	0.9590	0.9176	0.9117	2.39E-03	6.05E-06
2014	0.8200	0.6727	0.6120	7.32E-03	5.33E-05
2015	0.9718	0.9444	0.9321	2.59E-03	6.68E-06
2016	0.9670	0.9351	0.9251	2.82E-03	7.93E-06
2017	0.9611	0.9261	0.8956	1.95E-03	3.81E-06
all	0.9258	0.8574	0.9088	5.21E-03	2.71E-05
Panel E: FBSOX					
2006	0.9107	0.8301	0.7704	3.53E-03	1.24E-05
2007	0.8866	0.7845	0.6653	5.32E-03	2.84E-05
2008	0.9390	0.8821	0.8393	8.82E-03	7.74E-05
2009	0.8927	0.7993	0.8776	7.38E-03	5.43E-05
2010	0.9465	0.8962	0.8423	4.09E-03	1.67E-05
2011	0.9823	0.9642	0.9325	3.39E-03	1.16E-05
2012	0.9639	0.9292	0.9532	2.63E-03	6.94E-06
2013	0.9691	0.9417	0.9738	2.01E-03	4.04E-06
2014	0.8083	0.6536	0.6034	7.25E-03	5.24E-05
2015	0.9801	0.9602	0.9341	2.10E-03	4.45E-06
2016	0.9771	0.9543	0.9521	2.33E-03	5.44E-06
2017	0.9357	0.8767	0.9179	2.14E-03	4.67E-06
all	0.9330	0.8708	0.8511	4.82E-03	2.33E-05

doi:10.6342/NTU201801196

In 2014, index-tracking portfolios of FSCSX and FBSOX have the lowest correlation and R square. Beta deviates from one much as well. Tracking error and MSTD are also relatively high. From the table 4.6, if we drop the return on ex-dividend date and the date before ex-dividend date, all measurements will dramatically improve. Hence one can see that paying dividends significantly affect the tracking measurements.

Table 4.6 Tracking Measurement (Drop Ex-Dividend Date)

Table 4.6 gives the tracking measurement of the index-tracking portfolio. The column, Drop Ex-Dividend Date, means dropping ex-dividend date the date before ex-dividend date in 2014. Correlation means the correlation between returns of index-tracking portfolio and return of the index. We regress the returns of the index-tracking portfolio on returns of index without interception to get R square and Beta. Tracking error is the standard deviation of tracking difference. MSTD is mean square of tracking difference.

Ticker	Drop Ex-Dividend Date	Year	Correlation	R square	Beta	Tracking Error	MSTD
Panel A: 22 Testing Period							
FSCSX	FALSE	2014	0.8219	0.6758	0.5900	7.34E-03	5.37E-05
	TRUE	2014	0.9523	0.9070	0.8089	3.45E-03	1.18E-05
FBSOX	FALSE	2014	0.8128	0.6610	0.6084	7.18E-03	5.14E-05
	TRUE	2014	0.9672	0.9355	0.8921	2.57E-03	6.58E-06
Panel B: 66 Testing Period							
FSCSX	FALSE	2014	0.8200	0.6727	0.6120	7.32E-03	5.33E-05
	TRUE	2014	0.9606	0.9227	0.8507	3.08E-03	9.48E-06
FBSOX	FALSE	2014	0.8083	0.6536	0.6034	7.25E-03	5.24E-05
	TRUE	2014	0.9662	0.9335	0.8883	2.61E-03	6.79E-06

Finally, we compare index-tracking portfolio of GSPC and IXIC to ETFs. Table 4.7 shows the ratio of tracking difference of index-tracking portfolio to tracking difference of the ETF. If the ratio is greater/less than one, the absolute tracking difference of index-

doi:10.6342/NTU201801196

tracking portfolio is larger/smaller.[4] For GSPC in panel A, index-tracking portfolio have less discrepancy in minimum in all years except 2009. And it tends to have large discrepancy in maximum. The similar situation occurs in IXIC index-tracking portfolio. Table 4.8 and table 4.9 present the ratio of the specific tracking measurement of the index-tracking portfolio to the specific tracking measurement of the ETF. If the ratio is greater/less than one, the specific tracking measurement of index-tracking portfolio is larger/smaller. We convert Beta to absolute value of 1-Beta since we care about how Beta close to one. For GSPC in panel A, one thing that we would like to mention are that the ratio of Correlation and R square is less than one, and the ratio of Abs(1-Beta), Tracking Error and MSTD are less than one in all years with 66 testing period. It indicates that index-tracking portfolio tracks the index better when we consider all of the years. However, we cannot make conclusion that the tracking ability of index-tracking portfolio is better or worse since the results fluctuate by year. For IXIC in panel B, no matter which testing period, the ratio of Correlation and R square is less than one, and the ratio of Abs(1-Beta), Tracking Error and MSTD are less than one in almost all years. It strongly indicates that the tracking ability of the index-tracking portfolio is better than the ETF. Taking into account all these factors, we may reasonably come to the conclusion that the deep reinforcement learning model do a great job on index-tracking problem.

[4] The tracking difference of the index-tracking portfolio and the tracking difference of the ETF have the same sign when they reach to maximum or minimum.

doi:10.6342/NTU201801196

Table 4.7 Tracking Difference Ratio of Index-Tracking Portfolio and ETF

Table 4.7 gives the ratio of tracking difference of the index-tracking portfolio to tracking difference of the ETF.

Year	Min	Max	Min	Max
Panel A: GSPC	22 Testing Period		66 Testing Period	
2006	0.23	0.39	0.23	0.31
2007	0.50	1.07	0.50	0.94
2008	0.99	0.47	0.77	0.54
2009	2.23	5.00	1.99	3.88
2010	0.46	0.79	0.55	0.69
2011	0.35	1.69	0.34	1.43
2012	0.53	0.96	0.56	1.04
2013	0.49	1.52	0.88	1.66
2014	0.42	1.91	0.86	1.70
2015	0.48	2.04	0.77	2.07
2016	0.49	5.29	0.49	3.97
2017	0.56	2.58	0.58	1.99
Panel B: IXIC	22 Testing Period		66 Testing Period	
2006	1.78	4.57	1.89	2.16
2007	0.62	0.97	0.58	1.12
2008	0.64	0.43	0.20	0.32
2009	1.24	1.40	1.30	1.06
2010	0.64	0.93	0.86	0.99
2011	0.66	1.00	0.60	0.97
2012	1.00	1.27	0.68	0.88
2013	0.37	0.38	0.41	0.46
2014	0.96	1.14	0.82	1.25
2015	0.47	1.12	0.73	0.71
2016	0.36	0.61	0.41	0.39
2017	1.85	2.07	1.57	2.64

doi:10.6342/NTU201801196

Table 4.8 Tracking Measurement Ratio of Index-Tracking Portfolio to ETF (Testing Period: 22 Trading Days)

Table 4.8 gives the ratio of a specific tracking measurement of the index-tracking portfolio with 22 testing period to a specific tracking measurement of the ETF. Abs(1-Beta) is the absolute value of (1-Beta) which measures the beta how close to one.

Year	Correlation	R square	Abs(1-Beta)	Tracking Error	MSTD
Panel A: GSPC					
2006	1.0318	1.0640	0.9780	0.4312	0.1885
2007	1.0160	1.0321	0.9470	0.5361	0.2904
2008	1.0078	1.0157	2.8334	0.7968	0.6352
2009	0.9832	0.9668	3.5063	2.6699	7.1301
2010	1.0005	1.0010	3.2115	0.9580	0.9246
2011	0.9997	0.9994	0.4894	1.0520	1.1274
2012	0.9998	0.9995	1.7613	1.0185	1.0377
2013	1.0005	1.0008	8016.4262	0.9639	0.9457
2014	1.0017	1.0033	3.4665	0.9164	0.8434
2015	0.9982	0.9964	12.3848	1.2648	1.6044
2016	0.9958	0.9914	3.2139	1.3742	1.9058
2017	0.9974	0.9950	17.5428	1.0681	1.1441
all	1.0003	1.0005	1.3087	1.0212	1.0429
Panel B: IXIC					
2006	0.9825	0.9644	10.7056	1.9242	3.7717
2007	0.9996	0.9965	0.8909	1.0106	1.1023
2008	1.0126	1.0251	0.0203	0.7558	0.5736
2009	1.0073	1.0145	0.7102	0.8329	0.6951
2010	1.0018	1.0036	17.4545	0.9213	0.8500
2011	1.0025	1.0050	0.0204	0.8525	0.7273
2012	1.0070	1.0140	0.9410	0.8544	0.7314
2013	1.0286	1.0559	0.5250	0.6105	0.3758
2014	0.9999	0.9997	8.3350	1.0497	1.1018
2015	1.0047	1.0094	1.5546	0.8380	0.7023
2016	1.0097	1.0190	0.9719	0.7737	0.6064
2017	0.9821	0.9653	3.3072	1.5464	2.4063
all	1.0063	1.0125	0.1887	0.8421	0.7091

doi:10.6342/NTU201801196

Table 4.9 Tracking Measurement Ratio of Index-Tracking Portfolio to ETF (Testing Period: 66 Trading Days)

Table 4.9 gives the ratio of a specific tracking measurement of the index-tracking portfolio with 66 testing period to a specific tracking measurement of the ETF. Abs(1-Beta) is the absolute value of (1-Beta) which measures the beta how close to one.

Year	Correlation	R square	Abs(1-Beta)	Tracking Error	MSTD
Panel A: GSPC					
2006	1.0335	1.0675	1.1592	0.3845	0.1510
2007	1.0160	1.0322	0.7026	0.5304	0.2835
2008	1.0106	1.0213	1.2750	0.6556	0.4300
2009	0.9900	0.9801	2.8762	2.1536	4.6420
2010	0.9998	0.9995	3.7373	1.0674	1.1415
2011	1.0001	1.0001	1.0973	0.9830	0.9774
2012	1.0000	1.0000	4.3939	0.9996	0.9992
2013	0.9940	0.9882	13760.6160	1.4389	2.0746
2014	0.9998	0.9996	3.5544	1.0663	1.1438
2015	0.9996	0.9992	24.3821	1.1261	1.2769
2016	0.9957	0.9914	5.7230	1.3752	1.9029
2017	1.0006	1.0012	18.3197	0.9807	0.9626
all	1.0027	1.0055	0.4016	0.8793	0.7733
Panel B: IXIC					
2006	0.9850	0.9699	14.7078	1.8227	3.3511
2007	0.9983	0.9953	1.4326	1.0615	1.1653
2008	1.0218	1.0437	0.0114	0.5075	0.2615
2009	1.0056	1.0112	1.0383	0.9330	0.8731
2010	1.0007	1.0014	13.0298	0.9585	0.9187
2011	1.0029	1.0057	0.0840	0.8350	0.7010
2012	1.0127	1.0255	1.0135	0.7181	0.5155
2013	1.0233	1.0455	1.2455	0.6994	0.4927
2014	1.0033	1.0065	6.7428	0.9243	0.8552
2015	1.0052	1.0104	2.1534	0.8283	0.6907
2016	1.0092	1.0180	1.2200	0.7871	0.6295
2017	0.9890	0.9785	3.4030	1.4025	1.9677
all	1.0091	1.0183	0.1824	0.7550	0.5703

doi:10.6342/NTU201801196

Chapter 5 Conclusion

With the growth of ETFs, index-tracking problem is becoming more popular. Most of index-tracking ETFs are built by some rules. We propose a deep reinforcement learning method to track a certain index or an active fund without knowing the constituents, portfolio weights and the rules to build the portfolio. We first construct a framework of reinforcement learning to deal with index-tracking problem. The core to solve index-tracking problem is Policy gradient. Policy gradient allows action space to be continuous, which is important to build a portfolio. The experiment shows that deep RL agent tracks better in S&P 500 and NASDAQ Composite index-tracking portfolios. Comparing to SPDR S&P 500 ETF Trust (SPY) and Fidelity NASDAQ Comp. Index Trk Stk (ONEQ), the tracking ability of the index-tracking portfolios slightly behind. However, we can still keep faith with index-tracking portfolio since the average tracking difference from 2006/01/01 to 2017/12/31 is only 0.0009% (S&P 500, 22 testing period), 0.0004% (NASDAQ Composite, with 22 testing period), -0.0022% (S&P 500, 66 testing period), 0.0043% (NASDAQ Composite, with 66 testing period). As for active funds, we're amazing at these results since building portfolio of an active fund is much complicated than building portfolio of an index. Funds managers frequently rebalance the portfolio in order to have a great performance, so the turnover ratio of the portfolio is high. If we exclude the dividend paying factor, we could expect the tracking ability will further boost. In conclusion, index-tracking problem could be solved via deep reinforcement learning with policy gradient. Constituents, portfolio weights and rules to build the portfolios will be no more secret. [10] [10]

doi:10.6342/NTU201801196

Chapter 6 Reference

[1] I. Kaastra and M. S. Boyd, "Forecasting Futures Trading Volume Using Neural Network," *The Journal of Futures Markets,* vol. 15, no. 8, pp. 953-970, December 1995.

[2] J. Bollen, H. Mao and X.-J. Zeng, "Twitter mood predicts the stock market," *Journal of Computational Science,* vol. 2, no. 1, pp. 1-8, March 2011.

[3] R. Choudhry and K. Garg, "A Hybrid Machine Learning System for Stock Market Forecasting," *International Scholarly and Scientific Research & Innovation,* vol. 2, no. 3, pp. 689-692, 2008.

[4] Y. Lin, H. Guo and J. Hu, "An SVM-based Approach for Stock Market Trend Prediction," in *Neural Networks (IJCNN), The 2013 International Joint Conference on*, Dallas, TX, USA, 2014.

[5] T. Fischer, T. Fischera and C. Kraussb, "Deep learning with long short-term memory networks for financial market predictions," 2017.

[6] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra and M. Riedmiller, "Playing Atari with Deep Reinforcement Learning," in *DeepMind Technologies*, 2013.

[7] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg and D. Hassabis, "Human-level control through deep reinforcement learning," *Nature,* vol. 518, p. 529–533, Feburary 2015.

doi:10.6342/NTU201801196

[8] R. Neuneier, "Enhancing Q-Learning for Optimal Asset Allocation," in *NIPS*, 1997.

[9] J. Li, K. Zhang and L. Chan, "Independent Factor Reinforcement Learning for Portfolio Management," in *Intelligent Data Engineering and Automated Learning*, vol. 4881, Springer, Berlin, Heidelberg, 2007.

[10] Z. Jiang, D. Xu and J. Liang, "A Deep Reinforcement Learning Framework for the Financial Portfolio Management Problem," 2017.

doi:10.6342/NTU201801196

doi:10.6342/NTU201801196