國立臺灣大學管理學院財務金融學系

碩士論文

Department of Finance
College of Management
National Taiwan University
Master Thesis

後雙重選擇估計式的有限樣本表現

On the Finite-Sample Performance of the Post-Double Selection Estimator

邱祥鴻

Hsiang-Hung Chiu

指導教授：陳宜廷 博士

Advisor: Yi-Ting Chen, Ph.D.

中華民國112年8月

August, 2023

doi:10.6342/NTU202301747

國立臺灣大學碩士學位論文
口試委員會審定書

後雙重選擇估計式的有限樣本表現
On the Finite-Sample Performance of the Post-Double Selection Estimator

本論文係 邱祥鴻 君(R09723059)在國立臺灣大學財務金融學系暨研究所、所完成之碩士學位論文，於民國111年7月13日承下列考試委員審查通過及口試及格，特此證明

口試委員：

陳安廷 (簽名)	
(指導教授)	
洪志清	羅秉政

系主任、所長 王泓仁 (簽名)

doi:10.6342/NTU202301747

Acknowledgements

本論文能夠順利完成，我真心感謝我的指導教授陳宜廷教授。在過去的兩年時間，陳教授指引我寫作與學習的方向，並且不厭其煩地提供我思考的邏輯，讓我有更全面的思維。他也給予我許多支持和資源，使我初次接觸論文這條路時，能有更深刻的見解，真的非常感謝。此外，也非常感謝秉政教授和志清教授在口試時給予寶貴的建議。

還要特別感謝嵩硯學長的幫助與照顧，為我提供了許多理論的直覺和寶貴建議，讓我對計量方法有更深一層的理解，同時讓我的論文內容更加完整。

此外，我要特別感謝我的朋友們。有了你們的支持，讓我在心情低潮時有所依靠，真的非常謝謝你們。感謝碩班的朋友，智信、新維同學陪我度過枯燥的時光。特別感謝摯友瀞云和彥洵同學，有你們的支持，我才能度過最失落的時光，真的非常感謝。

最後，我要感謝我的父母和家人。他們給予我金錢的支持，也全力支持我的選擇。他們的愛與鼓勵是我前進的動力，無比感激。

邱祥鴻　謹誌

於臺灣大學財務金融學系

一百一十二年

doi:10.6342/NTU202301747

摘要

在各個研究中，測量因果關係是非常重要的。為了維持解釋變數的外生性，研究者可能需要考慮高維度的控制變數。在此架構下，傳統所使用的最小平方估計法並不適用。為了應對這些問題，Belloni 等人（2014）提出了後雙重選擇（Post-Double Selection，PDS）方法，此方法在計量文獻中受到了的重視。雖然 Belloni 等人 (2014) 已經證明了 PDS 方法具有漸近常態性，但在實證使用上，研究者們需要理解 PDS 在有限樣本中的表現。本文探討在 PDS 分析中，不同的統計學習方法進行雙重選擇所得出的有限樣本性質，並且藉此重新檢驗技術指標在實證中的顯著性。

關鍵字：因果推論、統計學習方法、風險溢酬、技術指標、模型選擇

doi:10.6342/NTU202301747

Abstract

Measuring causal effects is crucial in various research. Understanding the impact of an explanatory variable on an outcome variable is essential for evaluating the effectiveness of policy changes or interventions. However, to maintain the exogeneity of explanatory variables, researchers may need to consider high-dimensional control variables. In this framework, the traditional least squares method is inapplicable. To address this problem, Belloni et al. (2014) proposed the post-double selection (PDS) method. This method has received a lot of attention in econometrics. Although Belloni et al. (2014) have proved that the PDS estimator is asymptotically normal under suitable conditions, it is still important to evaluate how the PDS method behaves in finite samples. This study explores the finite-sample performance of the PDS estimator under different choices of statistical learning methods for the double selection. I also apply the PDS method to assess the significance of technical indicators in explaining stock returns.

Keywords: Causal Inference, Statistical Learning, Risk Premium, Technical Indicators, Model Selection

doi:10.6342/NTU202301747

Contents

doi:10.6342/NTU202301747

doi:10.6342/NTU202301747

List of Figures

doi:10.6342/NTU202301747

List of Tables

doi:10.6342/NTU202301747

Denotation

BIC	Bayesian Information Criterion
CV	Cross-Validation
DGP	Data Generating Process
HLMS	High-Dimensional Linear Model Selection
Lasso	Least Absolute Shrinkage and Selection Operator
LS	Least Squares
MA	Moving Average
MCP	Minimax Concave Penalty
MOM	Momentum
OBV	On-Balance Volume
OVB	Omitted Variable Bias
PDS	Post-Double Selection

doi:10.6342/NTU202301747

RMSE	Root Mean Square Error
SCAD	Smoothly Clipped Absolute Deviation
VOL	Volume

doi:10.6342/NTU202301747

1 Introduction

Measuring causal effects is a crucial topic in various research, including medicine, biology, economics, and finance. Understanding the causal effect of an explanatory variable on an outcome variable is essential for evaluating whether a particular policy change, or an intervention, is beneficial or harmful. However, since the explanatory variable may not be automatically exogenous, researchers need to employ a suitable design to avoid the endogeneity problem. A possible solution is to use a suitable set of control variables for applying the conditional independence assumption to causal inference. While the use of control variables has been extensively studied, a critical challenge arises when the number of control variables exceeds the sample size. In this high-dimensional context, the traditional estimation method is infeasible.

To address this high-dimensional problem, it is necessary to impose a suitable assumption of sparsity on the controls. This assumption requires that the number of truly useful control variables is smaller than the sample size in a certain sense. At first sight, one might regard that, under such a sparsity assumption, the choice of control variables might be facilitated using the least absolute shrinkage and selection operator (Lasso) of Tibshirani (1996) (27), or other statistical learning methods, before estimating the causal effect. However, Belloni et al. (2014) (2) demonstrate that this post-selection method might generate an omitted variable bias (OVB), and propose the post-double selection (PDS) method

doi:10.6342/NTU202301747

to avoid the OVB and maintain the exogeneity of the explanatory variable. The PDS method first applies a Lasso-type method to the regression of the explanatory variable on the control variables and the regression of the dependent variable on the control variables for "doubly selecting" suitable control variables. Subsequently, it applies the least squares (LS) method to the regression of the dependent variable on the explanatory variable and the "doubly selected" control variables for estimating the causal effect. The resulting LS estimator for the causal effect is called the PDS estimator. Belloni et al. (2014) (2) prove that, under suitable conditions, the PDS estimator is asymptotically normal.

Many studies use the PDS method to investigate empirical issues. An example of this can be found in the work of Berset et al. (2023) (4), where they utilized the PDS method to examine the impact of fiscal revenue shocks on local fiscal policy. Enke (2020) (9) investigates whether moral values affect voting behavior. Liu-Evans and Mitra (2019) (21) investigate the impact of bank stability on the size of the unregulated, non-tax-paying informal sector, among many others. Moreover, the PDS method has also been applied in financial research. In particular, due to the relatively short history of financial data, researchers often repeatedly use similar data to discover pricing factors of risky assets. This practice might lead to the data-snooping bias, as noted by Lo and MacKinlay (1990) (22) and Harvey et al. (2016) (14), among others. Avoiding this bias is important for financial research. To consider this issue, Feng et al. (2020) (11) apply the PDS method to investigate whether a set of newly released factors are truly useful for explaining the excess return of risky assets by controlling for a large amount of historical factors proposed by different researchers. Their empirical study shows that the PDS method is effective for this empirical investigation.

Nonetheless, it is still important to evaluate the finite-sample performance of the

doi:10.6342/NTU202301747

PDS method because the sample size is indeed limited in empirical studies. Recently, Wüthrich and Zhu (2023) (30) compared the performance of different choices of tuning parameter with Lasso for implementing the PDS method. They found that the PDS estimator may still have OVB in finite samples under certain types of data generating processes (DGP) for some popular choices of tuning parameters. To address this concern, they suggest checking the robustness of the PDS method to OVB by increasing the regularization parameter λ. In addition, Drukker and Liu (2022) (8) evaluate the finite-sample performance of the Neyman-orthogonal estimator, which is related to the PDS method, based on a generalized linear model. In particular, they consider different versions of Lasso with the tuning parameters selected by the Bayesian Information Criterion (BIC), the plug-in method of Belloni et al. (2012) (1), a cross-validation (CV) method and the sure independence screening as well as stepwise methods for variable selection. The simulations of these two studies mainly consider Lasso for variable selection.

Theoretically, the PDS method is based on the "high-dimensional linear model selection (HLMS) condition" of Belloni et al. (2014) (2), which might allow researchers to use not only Lasso but also other statistical learning methods for double selection (under suitable assumptions). This study considers a number of statistical learning methods for double selection and assess the finite-sample performance of the resulting PDS estimators. Specifically, I consider not only the Lasso but also the smoothly clipped absolute deviation (SCAD) method of Fan and Li (2001) (10), the minimax concave penalty (MCP) method of Zhang (2010) (32) and the adaptive Lasso of Zou (2006) (35) for the double selection. These statistical learning methods are all capable of generating sparsity under suitable choices of their tuning parameters. In addition, I consider different selection methods for the tuning parameters, including a CV method and an information-criterion-based selec-

doi:10.6342/NTU202301747

tion method. I adopt a Monte Carlo simulation to evaluate the finite-sample performance of the PDS estimator under different choices of these double-selection methods. In addition, I apply the PDS method with different choices of the double-selection methods to investigate the effectiveness of technical indicators in explaining stock returns.

The remainder of this thesis is organized as follows. In Chapter 2, I present the basic framework of PDS and discuss the statistical learning methods that I consider for double selection. In Chapter 3, I introduce the simulation designs, and present the simulation results. In Chapter 4, I illustrate the empirical application. Finally, I conclude this thesis in Chapter 5.

doi:10.6342/NTU202301747

2 Econometric Methods

In the following, I review the PDS method and the statistical learning methods that I used in this thesis.

2.1 The PDS method

The PDS method is built on the following linear regression:

$$y_i = x_i'\beta_0 + z_i'\gamma_y + \zeta_i, \quad E[\zeta_i|x_i, z_i] = 0, \; i = \{1, ..., N\}, \tag{2.1}$$

$$x_i = z_i'\gamma_x + v_i, \quad E[v_i|z_i] = 0, \tag{2.2}$$

as shown by Equations (2.2) and (2.3) of Belloni et al. (2014) (2), where y_i is the outcome variable, x_i is a vector of explanatory variables with dimension p, β_0 is the parameter vector (that is, the treatment effect when x_i is interpreted as a vector of treatment variables), z_i is a vector of control variables, with dimension k, that contains both irrelevant and relevant controls, γ_y and γ_x are nuisance parameters, and ζ_i and v_i are zero-mean error terms. Suppose that the sample size is N. The high-dimensional context appears when $p \geq N$. This study focuses on the linear model, and hence omits the additional approximation errors considered by Belloni et al. (2014) (2) generated by approximating a non-parametric model.

doi:10.6342/NTU202301747

To obtain a reliable inference for β_0, the PDS method first applies a variable-selection method, such as the Lasso, to the first-stage regression of $x_{i,j}$ on z_i, where $x_{i,j}$ is the jth element of x_i for $j = 1, \ldots, p$. In addition, it applies the same selection method to another first-stage regression of y_i on z_i. Let $\hat{I}_{1,j}$ be the control variables selected for $x_{i,j}$, and $\hat{I}_2$ be the control variables selected for y_i. Define $\hat{I} = (\bigcup_{j=1}^{p} \hat{I}_{1,j}) \cup \hat{I}_2$ as the set of the doubly selected control variables. The PDS estimator for β_0, denoted as $\hat{\beta}$, is the LS estimator based on the second-stage regression of y_i on x_i and the doubly selected control variables; that is,

$$(\hat{\beta}, \hat{\gamma}) = \underset{\beta,\gamma}{\text{argmin}}\{\frac{1}{N}\sum_{i=1}^{N}[(y_i - x_i'\beta - \tilde{z}_i'\gamma)^2]\},$$

where $\tilde{z}_i$ is a subset of z_i defined by $\hat{I}$. An important feature of the PDS method is that the PDS estimator can be shown to be asymptotically normal under suitable conditions. In particular, as shown by Equation (2.10) of Belloni et al. (2014) (2), as $N \to \infty$,

$$\sigma^{-1}\sqrt{N}(\hat{\beta} - \beta_0) \xrightarrow{d} N(0, I_p), \tag{2.3}$$

where $\sigma^2 = (Ev_i^2)^{-1}E(v_i^2\zeta_i^2)(Ev_i^2)^{-1}$, and I_p is the identical matrix with dimension p.

The basic idea underlying the PDS method is to exploit the sparsity (assumption) of the high-dimensional controls for dimension reduction. Obviously, $\hat{\beta}$ is computable only if such a sparsity is assumed and that the selection methods for the first-stage (high-dimensional) regression could suitably capture the sparsity. An important sufficient condition underlying the asymptotic normality is the HLMS condition, which requires the first-stage selection method to satisfy suitable conditions of sparsity and "good estimation quality." Belloni et al. (2014) (2) recommended using a version of Lasso proposed by Belloni et al. (2012) (1), referred to as the plug-in Lasso here, for the double selection. In

doi:10.6342/NTU202301747

this study, I consider more statistical learning methods for the double selection.

2.2 Statistical learning methods

To review these methods, I consider the first-stage reduced-form regression:

$$y_i = z_i'\delta + \epsilon_i,$$

where $\delta := (\delta_1, \ldots, \delta_k)'$ is a k-dimensional parameter vector, and ϵ_i is a zero-mean error. In addition, I consider the statistical learning methods, or said the penalized regressions, that share the following objective function:

$$Q(\delta) + \sum_{j=1}^{k} \mathcal{P}_\lambda(|\delta_j|), \tag{2.4}$$

where $Q(\delta) := \frac{1}{N}\sum_{i=1}^{N}(y_i - z_i'\delta)^2$ is the sample mean squared error of the aforementioned regression, and $\mathcal{P}_\lambda(\cdot)$ is a penalty function that includes λ as a tuning parameter; see, e.g., Equation (2) of Wu and Wang (2020) (29) for this expression of the objective function. By a suitable design of $\mathcal{P}_\lambda(\cdot)$, the minimization of this objective function with respect to δ generates the sparsity of the estimator for δ, which is applicable to the selection of z_i.

2.2.1 Lasso

The Lasso of Tibshirani (1996) (27) sets $\mathcal{P}_\lambda(|\delta_j|) = \lambda|\delta_j|$, and has the solution path:

$$\hat{\delta}^{Lasso}(\lambda) = \underset{\delta}{\text{argmin}}\ \{Q(\delta) + \sum_{j=1}^{k} \lambda|\delta_j|\},$$

doi:10.6342/NTU202301747

which is a function of the tuning parameter λ. A larger λ may make $\hat{\delta}^{Lasso}(\lambda)$ sparser. Therefore, the choice of λ might affect the finite-sample performance of the PDS method via changing the double selection. It is known that Lasso may select the true set of non-zero true coefficients under quite restrictive conditions, as demonstrated by Zhao and Yu (2006) (34). In addition, Leng et al. (2006) (20) demonstrated that Lasso may not have the variable-selection consistency when its objective is designed for prediction. Zou (2006) (35) also shown that Lasso may not serve as an oracle estimator proposed by Fan and Li (2001) (10), which can select correct model when sample size is large enough. The imperfect selection of Lasso may also influence the finite-sample performance of the PDS method. To deal with these problems of Lasso, the penalty function of Lasso needs to be suitably modified.

To refine the original Lasso, Belloni et al. (2012) (1) proposed a modification of the original Lasso, referred to as the plug-in Lasso here. This method generates the following estimator for δ:

$$\hat{\delta}^{\,plugin} = \underset{\delta}{\operatorname{argmin}}\{Q(\delta) + \lambda \sum_{j=1}^{k} |\hat{l}_j \delta_j|\},$$

in which $\hat{l}_j$ is the "penalty loading":

$$\hat{l}_j = \sqrt{\frac{1}{N} \sum_{i=1}^{N} z_{ij}^2 \tilde{\epsilon}_i^2},$$

and λ is a "plug-in" penalty term:

$$\lambda = 2 \cdot c\sqrt{N}\Phi^{-1}(1 - \gamma/2k),$$

where $\Phi^{-1}(\cdot)$ is the quantile function of $N(0, 1)$, and c and γ are, respectively, suggested to be 1.1 and $(1/N)^{0.05}$ in Footnote 10 of Belloni et al. (2014) (2). However, for convenience,

doi:10.6342/NTU202301747

I simply implement the plug-in Lasso using the R package `hdm` which computes the $\tilde{\epsilon}_i^2$ in the penalty loading using an iteration method. Note that, unlike the original Lasso, the λ of the plug-in Lasso is not a free parameter.

2.2.2 SCAD

The SCAD method of Fan and Li (2001) (10) has the solution path:

$$\hat{\delta}^{SCAD}(\lambda) = \underset{\delta}{\text{argmin}} \{Q(\delta) + \sum_{j=1}^{k} \mathcal{P}_\lambda |\delta_j|\},$$

based on a non-convex penalty function:

$$\mathcal{P}_\lambda(|\delta_j|) = \begin{cases} \lambda|\delta_j|, & if\ \delta_j < \lambda, \\ \frac{2a\lambda|\delta_j| - \delta_j^2 - \lambda^2}{2(a-1)}, & if\ \lambda < |\delta_j| < a\lambda, \\ \frac{(a+1)\lambda^2}{2}, & if\ |\delta_j| \geq a\lambda. \end{cases}$$

Following Fan and Li (2001) (10), I set $a = 3.7$. The SCAD penalty matches the Lasso penalty when δ_j is small. However, the SCAD penalty converges to a constant at a quadratic rate when δ_j is sufficiently large. By this design, the SCAD avoids the large bias of the Lasso.

2.2.3 MCP

The MCP method of Zhang (2010) (32) has the solution path:

$$\hat{\delta}^{MCP}(\lambda) = \underset{\delta}{\text{argmin}} \{Q(\delta) + \sum_{j=1}^{k} \mathcal{P}_\lambda |\delta_j|\},$$

doi:10.6342/NTU202301747

based on another non-convex penalty function:

$$\mathcal{P}_\lambda(|\delta_j|) = \begin{cases} \lambda|\delta_j| - \frac{\delta_j^2}{2a}, & if\ |\delta_j| \leq a\lambda, \\ \frac{a\lambda^2}{2}, & if\ |\delta_j| > a\lambda. \end{cases} \tag{2.5}$$

We set $a = 3$. The MCP penalty also matches the Lasso penalty when δ_j is small, but converges to a constant when δ_j is sufficiently large. Like the SCAD, this design also helps correct the large bias of Lasso. Compared with SCAD, MCP deviates from the Lasso penalty and converges to a constant at a faster rate.

2.2.4 Adaptive Lasso

The adaptive Lasso of Zou (2006) (35) has the solution path:

$$\delta^{Ada}(\lambda) = \underset{\delta}{\text{argmin}}\ \{Q(\delta) + \sum_{j=1}^{p} \lambda \hat{w}_j |\delta_j|\},$$

where $\hat{w}_j$ is a weight for δ_j based on an initial estimator for δ_j. Unlike the aforementioned methods, the adaptive Lasso aims to control the bias of Lasso using suitable weights. Following Zou (2006) (35), I set the initial estimator as the jth element of the ridge estimator:

$$\hat{\delta}^{ridge} = \underset{\delta}{\text{argmin}}\{Q(\delta) + \lambda \sum_{j=1}^{p} \delta_j^2\},$$

and set the weight:

$$\hat{w}_j = \begin{cases} 1/|\hat{\delta}_j^{ridge}|, & if\ \hat{\delta}_j^{ridge} \neq 0, \\ 0, & \text{otherwise}, \end{cases}$$

where $\hat{\delta}_j^{ridge}$ is the jth element of $\hat{\delta}^{ridge}$.

doi:10.6342/NTU202301747

2.3 Choices of tuning parameter

The choice of tuning parameter plays a key role in determining the performance of the statistical learning method, with the exception of the plug-in Lasso. In this study, I consider the ten-fold CV method and a BIC method for the choice of the tuning parameter.

2.3.1 The ten-fold CV method

It is standard to choose the tuning parameter of a statistical learning method by minimizing the "k-fold CV," which serves as an estimator for the mean squared prediction errors generated by the statistical learning method. In this study, I consider the ten-fold CV method, and implement this method using the R package functions: `cv.glmnet` and `cv.ncvreg`. Let $CV(\lambda)$ be the ten-fold CV of a statistical learning method evaluated at a specific value of λ. This method chooses λ as:

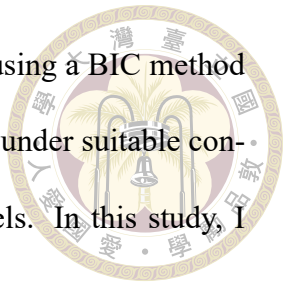
$$\hat{\lambda}^{CV} = \underset{\lambda}{\operatorname{argmin}}\, CV(\lambda).$$

In practice, the minimization is facilitated by examining a set of knots of λ; see, e.g., Hastie et al. (2015) (15).

2.3.2 A BIC method

In the literature, there are also studies that propose using a BIC method for the choice of the tuning parameter. Several studies have found that the tuning parameter selected by the BIC-based selection methods could outperform its counterpart selected by the CV method. See Hui et al. (2015) (19), Xiao and Sun (2019) (31), and Wang et al. (2009)

doi:10.6342/NTU202301747



(28), among others. Zhang et al. (2010) (33) also demonstrated that using a BIC method to select tuning parameters could consistently identify the true model under suitable conditions, while using an AIC method tends to select over-fitted models. In this study, I consider the tuning parameter selected by the following BIC method:

$$\hat{\lambda}^{BIC} = \underset{\lambda}{\text{argmin}}\, BIC(\lambda),$$

where the BIC is of the form:

$$BIC(\lambda) = N \log(N^{-1} \sum_{i=1}^{N} (y_i - z_i' \hat{\delta}(\lambda))^2) + k(\lambda) \log(N),$$

with $k(\lambda)$ denoting the number of non-zero elements of $\hat{\delta}(\lambda)$. This BIC is defined on the solution path of the statistical learning method. This is somewhat different from the conventional BIC for the linear regression. See also Drukker and Liu (2022, Section 2.3.4) (8) for the use of this BIC.

doi:10.6342/NTU202301747

3 Simulation

In this chapter, I present the simulation designs and provide a summary of the simulation results. This simulation considers two designs of DGP. The first design is built on a two-stage model considered by Belloni et al. (2014) (2). The second design is based on a simulation setting which is similar to that of Drukker and Liu (2022) (8).

To investigate how the sample size and the statistical learning method influence the finite-sample performance of the PDS estimator, I consider the following settings of (N, k):

1. $(N, k) = (100, 200)$,
2. $(N, k) = (200, 200)$,
3. $(N, k) = (400, 200)$.

Case 1 and 2 are high-dimensional settings with $N \leq k$. Case 3 is a large-dimensional setting with $N > k$ and a large k. The statistical learning methods being evaluated in this study include Lasso, SCAD, MCP and adaptive Lasso. For each of these methods, I choose the tuning parameter using the 10-fold CV or the BIC method. In addition, I consider the plug-in Lasso method for comparison.

doi:10.6342/NTU202301747

3.1 Simulation designs

In this section, I discuss the two designs of DGP. The first design follows the simulation of Belloni et al.(2014) (2). The main focus of their simulation is to show that the PDS method outperforms the post single-selection method. In comparison, I aim to compare different statistical learning methods for implementing the PDS method. They considered the following DGP:

$$y_i = x_i\beta_0 + z_i'(c_y\gamma_y) + \zeta_i,$$
$$x_i = z_i'(c_x\gamma_x) + v_i,$$

which implies the following two reduced-form regressions:

$$y_i = z_i'(c_x\gamma_x\beta_0 + c_y\gamma_y) + (\zeta_i + \beta_0 v_i),$$
$$x_i = z_i'(c_x\gamma_x) + v_i,$$

that are, respectively, of the R^2's: R_y^2 and R_x^2 which are determined by the coefficients: c_x and c_y; the parameters are set to be $\beta_0 = 0.5$ and

$$\gamma_y = \gamma_x = \left\{j^{-2}\right\}_{j=1}^{k};$$

c_x and c_y are selected to ensure that $R_y^2 = R_x^2 = 0.8$; the error terms: ζ_i and v_i are independent $N(0,1)$ random variables; x_i is one-dimensional, and

$$z_i \sim N(0, \Sigma),$$

where $\Sigma_{j\ell} = (0.9)^{|j-\ell|}$ for $j, \ell = 1, 2, \ldots, k$.

doi:10.6342/NTU202301747

The second design is more intuitive. It is similar to the design considered by Drukker and Liu (2022, Section 3.1) (8). The DGP is of the form:

$$y_i = \beta^b x_i^b + \beta^s x_i^s + \beta^0 x_i^0 + z_i'\gamma_0 + \epsilon_i,$$

where x_i^b, x_i^s and x_i^0 are, respectively, of a big coefficient, a small coefficient and the zero coefficient such that

$$(\beta^b, \beta^s, \beta^0) = (0.5, 0.25, 0),$$

γ_0 is set to be

$$\gamma_0 = \{\underbrace{0.5, ..., 0.5}_{10}, \underbrace{0, ..., 0}_{k-10}\},$$

that is the first ten elements (the remaining elements) of γ_0 are all equal to 0.5 (zero),

$$(x_i^b, x_i^s, x_i^0, z_i')' \sim N(0, \Sigma),$$

where Σ has the (j, ℓ)th element $\Sigma_{j\ell} = (0.9)^{|j-\ell|}$ for $j, \ell = 1, 2, \ldots, k+3$, and $\epsilon_i \sim N(0, 1)$ and independent of the $(x_i^b, x_i^s, x_i^0, z_i')$.

3.2 Simulation results

In this simulation, I set the number of simulation replications to be R =1000.

doi:10.6342/NTU202301747

3.2.1 Performance of the PDS estimator

Following Belloni et al. (2014) (2), I present the simulation results based on a measure of estimation accuracy:

$$RMSE = \sqrt{\frac{1}{R}\sum_{r=1}^{R}(\hat{\beta}_{(r)} - \beta_0)^2},$$

where $\hat{\beta}_{(r)}$ is the PDS estimate of the rth replication, and β_0 is the associated true parameter which is defined in the first simulation design (is set to be β^b, β^s or β^0 in the second simulation design), and a measure of asymptotic validity:

$$Rej = \frac{1}{R}\sum_{r=1}^{R}\mathbb{I}(|\sqrt{N}(\hat{\beta}_{(r)} - \beta)/\hat{\sigma}_{(r)}| \geq Z_{0.975}),$$

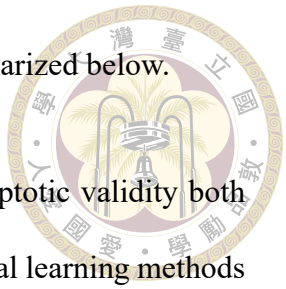
where $\mathbb{I}(\cdot)$ is the indicator function, $Z_{0.975}$ is the 97.5% quantile of $N(0,1)$, and $\hat{\sigma}_{(r)}$ is the standard error of the PDS estimator in the rth replication which is computed following Belloni et al. (2014) (2).

Note that $RMSE$ is the simulated root mean square error (RMSE) of the PDS estimator, which measures the estimation quality of the PDS method. A lower value of $RMSE$ indicates better estimation performance. In addition, Rej is the simulated rejection frequency of a two-sided test at the 5% level. If the PDS method performs well, the asymptotic normality of the PDS method implies that Rej should be close to 5%. A value of Rej that is closer to 5% indicates that the finite-sample performance of the PDS method is closer to the asymptotic normality implied by the theory.

I report $RMSE$ and Rej for the first simulation design in Table 3.1 and for the second simulation design in Table 3.2 (when $\beta_0 = \beta^b$), in Table 3.3 (when $\beta_0 = \beta^s$) and

doi:10.6342/NTU202301747

in Table 3.4 (when $\beta_0 = \beta^0$). The main simulation findings are summarized below.



1. The measure of estimation accuracy and the measure of asymptotic validity both improve as the sample size increases. This holds for all statistical learning methods and to both simulation designs.

2. Among the statistical learning methods, the Lasso and the adaptive Lasso with the tuning parameters, selected by the BIC, tend to outperform other methods in terms of their $RMSE$'s and Rej's.

3. In comparison, the SCAD and the MCP do not perform well in this simulation. This might be related to the fact that these two methods include an additional hyper-parameter "a" which is set to be fixed in this simulation. However, this is a conjecture that needs to be further examined in future studies.

4. For the Lasso and the adaptive Lasso, the BIC tends to outperform the ten-fold CV for the choice of their tuning parameters.

5. Focusing on the second design, the estimation accuracy of the PDS method increases with the value of β_0.

6. Although the plug-in Lasso method and the adaptive Lasso with the tuning parameter selected by the BIC perform have similar finite-sample performance, a closer comparison shows that the adaptive Lasso method tends to be more stable than the plug-in method in terms of their simulation performance.

This simulation shows that, at least for the simulation designs being considered, the adaptive Lasso with the tuning parameter selected by the BIC is a proper statistical learning method for the use of the PDS method.

doi:10.6342/NTU202301747

Table 3.1: $\hat{\beta}$ Performance for Design 1 ($R^2 = 0.8$)

	k	200		200		200	
	N	100		200		400	
Method		$RMSE$	Rej	$RMSE$	Rej	$RMSE$	Rej
Lasso	CV	0.138	0.107	0.079	0.059	0.053	0.053
	BIC	0.111	0.077	0.073	0.055	0.052	0.063
	plug-in	0.107	0.061	0.071	0.048	0.050	0.043
adaptive Lasso	CV	0.119	0.083	0.072	0.041	0.063	0.101
	BIC	0.107	0.06	0.071	0.054	0.053	0.057
SCAD	CV	0.124	0.102	0.075	0.055	0.052	0.053
	BIC	0.124	0.111	0.076	0.071	0.053	0.054
MCP	CV	0.117	0.092	0.074	0.059	0.052	0.058
	BIC	0.142	0.137	0.074	0.058	0.054	0.07

Note. CV represents the tuning parameter chosen through the 10-fold CV.
BIC represents the tuning parameter chosen through the BIC method.

Table 3.2: $\hat{\beta}$ Performance for Design 2 ($\beta_0 = \beta^b$)

	k	200		200		200	
	N	100		200		400	
Method		$RMSE$	$Rej.$	$RMSE$	$Rej.$	$RMSE$	$Rej.$
Lasso	CV	0.262	0.076	0.178	0.072	0.111	0.040
	BIC	0.245	0.070	0.174	0.065	0.110	0.038
	plug-in	0.241	0.061	0.174	0.072	0.110	0.036
adaptive Lasso	CV	0.246	0.067	0.174	0.071	0.120	0.065
	BIC	0.242	0.066	0.174	0.069	0.110	0.040
SCAD	CV	0.267	0.085	0.190	0.065	0.118	0.055
	BIC	0.270	0.069	0.192	0.066	0.118	0.048
MCP	CV	0.276	0.083	0.198	0.081	0.118	0.051
	BIC	0.276	0.075	0.196	0.080	0.118	0.054

Note. CV represents the tuning parameter chosen through the 10-fold CV.
BIC represents the tuning parameter chosen through the BIC method.

doi:10.6342/NTU202301747

Table 3.3: $\hat{\beta}$ Performance for Design 2 ($\beta_0 = \beta^s$)

	k	200		200		200	
	N	100		200		400	
Method		$RMSE$	$Rej.$	$RMSE$	$Rej.$	$RMSE$	$Rej.$
Lasso	CV	0.351	0.072	0.228	0.062	0.153	0.048
	BIC	0.323	0.062	0.225	0.064	0.153	0.056
	plug-in	0.324	0.061	0.226	0.060	0.153	0.054
adaptive Lasso	CV	0.328	0.056	0.226	0.061	0.153	0.053
	BIC	0.325	0.061	0.226	0.060	0.153	0.051
SCAD	CV	0.356	0.064	0.251	0.066	0.157	0.047
	BIC	0.358	0.066	0.254	0.064	0.160	0.056
MCP	CV	0.356	0.074	0.254	0.065	0.158	0.050
	BIC	0.359	0.066	0.254	0.065	0.156	0.054

Note. CV represents the tuning parameter chosen through the 10-fold CV.
BIC represents the tuning parameter chosen through the BIC method.

Table 3.4: $\hat{\beta}$ Performance for Design 2 ($\beta_0 = \beta^0$)

	k	200		200		200	
	N	100		200		400	
Method		$RMSE$	$Rej.$	$RMSE$	$Rej.$	$RMSE$	$Rej.$
Lasso	CV	0.380	0.067	0.228	0.059	0.162	0.052
	BIC	0.334	0.064	0.222	0.056	0.159	0.056
	plug-in	0.339	0.066	0.221	0.057	0.159	0.056
adaptive Lasso	CV	0.348	0.068	0.222	0.056	0.162	0.055
	BIC	0.338	0.066	0.222	0.058	0.159	0.055
SCAD	CV	0.368	0.071	0.246	0.066	0.164	0.058
	BIC	0.369	0.063	0.250	0.059	0.166	0.059
MCP	CV	0.375	0.065	0.241	0.048	0.163	0.054
	BIC	0.387	0.068	0.243	0.054	0.165	0.058

Note. CV represents the tuning parameter chosen through the 10-fold CV.
BIC represents the tuning parameter chosen through the BIC method.

doi:10.6342/NTU202301747

3.2.2 Performance of variable selection

Since the performance of the PDS method heavily depends on the double-selection process, it is important to further evaluate how the statistical learning methods perform for the double selection. For this purpose, I present the double-selection frequencies for the first 20 elements of z_i, generated by different statistical learning methods, in Figure 3.1 for the first simulation design and in Figure 3.2 for the second simulation design in the case where $N = 400$.

Recall that, in the first simulation design, the j control variable has the coefficient j^{-2}, and hence that the relevance of the j control variable rapidly decays to zero as j increases. In particular, $j^{-2} = 0.01$ as $j = 10$. In the second simulation design, the first ten elements of the control variable vector have non-zero coefficients, and the remaining elements have zero coefficients. Thus, I consider the first ten elements of z_i as (rough) relevant controls for the (first) second simulation design.

The main simulation findings of this part are summarized below.

1. In Figures 3.1 and 3.2, both the SCAD and the MCP show an obvious under-selection of the second element of z_i. It is speculated that this might be related to the high correlations among the control variables. In my experience, this problem is refined if I change the simulation setting of $\Sigma_{j\ell}$ to be $\Sigma_{j\ell} = (0.5)^{|j-\ell|}$, while the simulation result is not reported here.

2. The SCAD and the MCP tend to under-select relevant controls in comparison with other statistical learning methods. This might be related to the fact that these two methods include larger penalty terms than others. See Figure 3.3 for a compari-

doi:10.6342/NTU202301747

son between the penalty functions of Lasso, SCAD, and MCP, in which the tuning parameter λ is evaluated at its simulated average from the design 2.

3. The adaptive Lasso tends to have better performance than other statistical learning methods in the double selection. This point is particularly obvious for the second design.

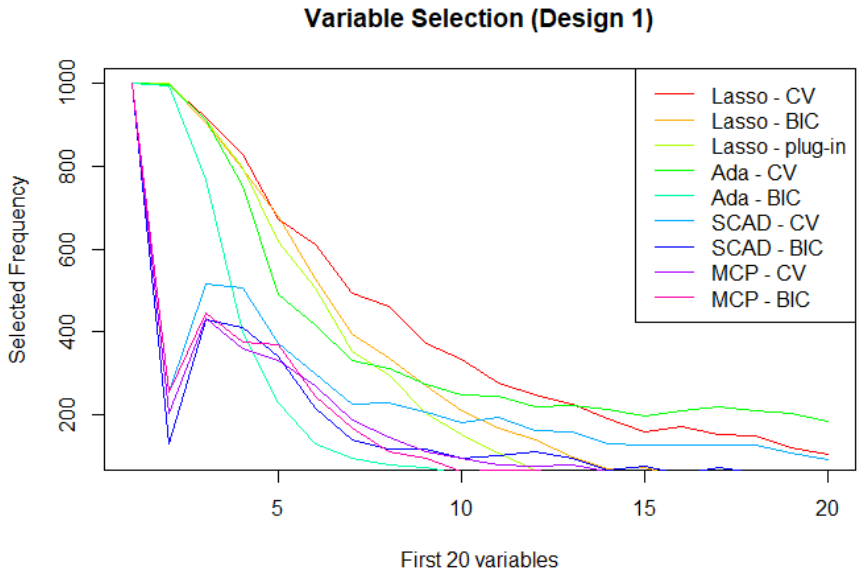


Figure 3.1: Variable Selection (Design 1)

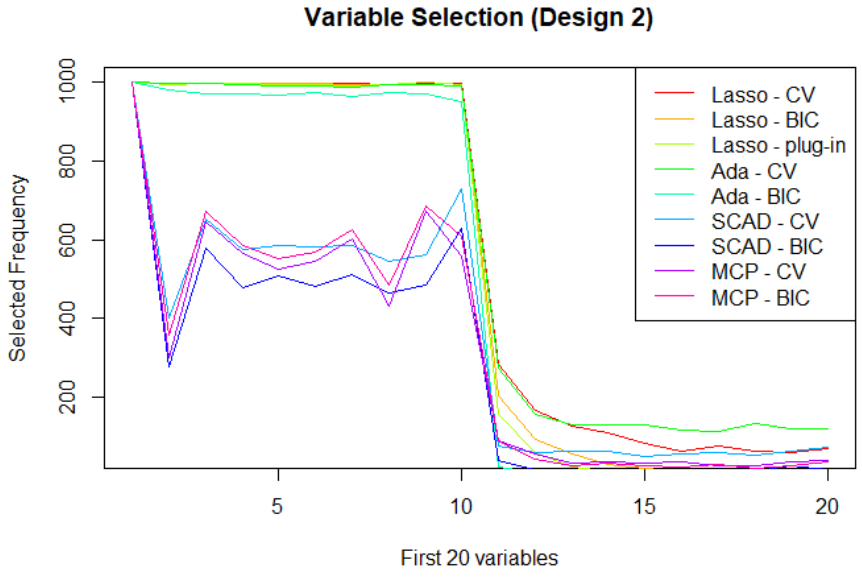


Figure 3.2: Variable Selection (Design 2)

doi:10.6342/NTU202301747

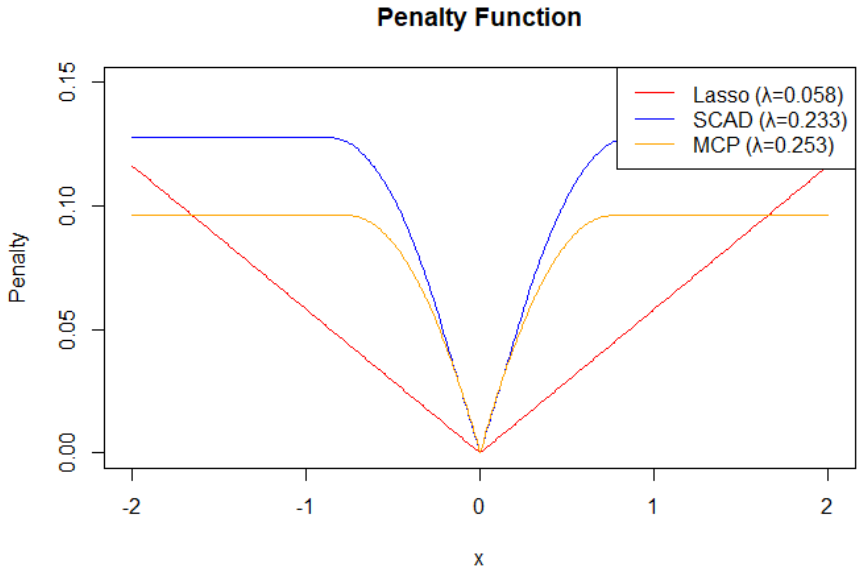


Figure 3.3: Penalty Comparison (Design 2)

doi:10.6342/NTU202301747

4 Empirical Application

In investment, practitioners often utilize the historical information on stock prices and trading volumes to generate "technical indicators" for predicting changes in stock prices. There is also academic research focused on evaluating the profitability of technical-trading strategies; see, e.g., Sullivan et al. (1999) (26), Mitra (2002) (23), Qi and Wu (2006) (25), Hsu et al. (2010) (18), among others. In addition, there are several studies that aim at evaluating whether technical indicators are useful for predicting equity premiums; see, e.g., Dai et al. (2021) (7), Goh et al. (2012) (13) and Henrique et al. (2018) (17), among many others. In particular, Neely et al. (2014) (24) consider the problem of predicting equity premiums using macroeconomic indicators and a set of technical indicators, including the price moving-average rules, the trading-volume-based indicators, and the momentum rules. They considered a linear regression with 14 macroeconomic indicators and 14 technical indicators for assessing whether the technical indicators are useful for predicting equity premiums. By construction, a moving-average variable includes the "window size" as a parameter to be selected. Neely et al. (2014) (24) choose the parameters of the technical indicators in a subjective way. However, as mentioned by Chinco et al. (2019) (6), a subjective choice of parameters for the technical indicators may not be suitable for capturing market information.

To deal with this problem, this study considers a linear regression that includes a

doi:10.6342/NTU202301747

large-dimensional technical indicators with different choices of parameters as predictors without pre-selecting the parameters in a subjective way, and applies the PDS method to assessing the significance of the indicators as a data-driven method for the choice of the parameters (or said for the choice of the technical indicators). For simplicity, this study does not include the macroeconomic indicators, and is fully based on the in-sample analysis.

4.1 Technical indicators

Following Neely et al. (2014) (24), I consider three sets of technical indicators. Let P_t be the stock price or index at time t. The first one is a set of price moving average indicators:

$$S_t^{MA}(s,l) = \begin{cases} 1, & \text{MA}_{s,t} \geq \text{MA}_{l,t}, \\ 0, & \text{otherwise}, \end{cases}$$

where the price moving average:

$$\text{MA}_{j,t} = \frac{1}{j}\sum_{i=0}^{j-1} P_{t-i},\ j = \{s,l\},$$

is dependent on the choice of a short window size s and a long window size $l > s$. This set of indicators reduce short-term fluctuations of price using moving averages and aim to capture the direction of price trend. Given the choice of s and l, $S_t^{MA}(s,l)$ generates a bullish signal, or said a "golden cross," if the short-term moving average curve intersects the long-term moving average curve from below.

The second one is a set of trading-volume moving average indicators. This set of

doi:10.6342/NTU202301747

indicators is built on the concept of "on-balance volume (OBV)." The OBV at time t is defined as:

$$\text{OBV}_t = \sum_{i=1}^{t} \text{VOL}_i D_i,$$

where VOL_i is the volume at time i, and

$$D_i = \begin{cases} 1, & P_i - P_{i-1} \geq 0, \\ -1, & P_i - P_{i-1} < 0. \end{cases}$$

The basic idea of OBV is that an increase in trading volume during price advances (declines) might reflect an anticipation of a bullish (bearish) sentiment. Given OBV, a trading-volume moving average indicator is defined as:

$$S_t^{VOL}(s,l) = \begin{cases} 1, & MA_{s,t}^{OBV} \geq MA_{l,t}^{OBV}, \\ 0, & \text{otherwise}, \end{cases}$$

where

$$\text{MA}_{j,t}^{OBV} = \frac{1}{j} \sum_{i=0}^{j-1} \text{OBV}_{t-i} \; j = \{s,l\}.$$

doi:10.6342/NTU202301747

The third one is a momentum indicator which is defined as:

$$S_t^{MOM}(m) = \begin{cases} 1, & P_t \geq P_{t-m}, \\ 0, & \text{otherwise.} \end{cases}$$

This indicators assesses whether the stock price at time t is higher than the price m months ago. The basic idea of this indicator is that stock price might continue to move in the known direction when the momentum is strong enough. In what follows, I denote these three sets of indicators as: MA, VOL and MOM. The definitions of the three sets of technical indicators follow Equations (2), (4) and (6) of Neely et al. (24).

4.2 Data

In this empirical application, I set P_t to be the S&P500 index at the tth month in the sampling period: January 1950 to December 2022, and define the equity premium at the $t+1$th month as:

$$r_{t+1} = \log(P_{t+1}) - \log(P_t) - \log(1 + r_{f,t+1}),$$

where $r_{f,t+1}$ is the treasury bill rate, which is considered the risk-free rate. I obtain the time-series data for the S&P 500 index and the risk-free rate from Amit Goyal's website, and obtain the data for the S&P 500 volume from the website of "Stooq." The sample size available for the PDS estimation is 863. This calculation is based on the monthly data from 1950 to 2022, which amounts to $73 * 12 = 876$ months. However, we need to subtract the first 12 months because they are used to estimate the signals. Additionally, considering a one-period lag in the predictive regression, the resulting sample size is 863. Since this

doi:10.6342/NTU202301747

empirical application is built on monthly time series, I consider all possible combinations of (s, l, m) within one year for establishing the technical indicators. In particular, there are a total of 66 different choices of (s, l), that is $C_2^{12} = 66$, for the MA and VOL indicators, and 12 different choices of m, that is $m = 1, 2, \ldots, 12$, for the MOM indicators.

4.3 Predictive regression

This empirical analysis is conducted using the following regression model:

$$r_{t+1} = \alpha + S_t'\beta + \epsilon_{t+1}, \tag{4.1}$$

where α is an intercept, S_t is a 144×1 vector of technical indicators, β is a vector of regression coefficients, and ϵ_{t+1} is a zero-mean error. Since the number of parameters is still smaller than the sample size for this regression, the LS method is still computable. However, we might expect that the LS estimator is inefficient because this regression is "large-dimensional." Therefore, this study only considers the LS method a baseline estimation method, and focuses on the PDS method for the estimation of β. To implement the PDS method, I consider the following partition of the regression:

$$r_{t+1} = \alpha + (\beta^c)' F_t^c + \sum_{j \neq c} (\beta^j)' F_t^j + e_{i,t+1}, \quad c \in \{\text{MA, VOL, MOM}\}, \tag{4.2}$$

where F_t^c comprises the 66 MA indicators, the 66 VOL indicators or the 12 MOM indicators, and F_t^j comprises the remaining indicators (that serve as the control variables).

doi:10.6342/NTU202301747

4.4 Empirical findings

I report the calculation results from Equation (4.2) in tables 4.2 to 4.6. In Table 4.1, I report the technical indicators that are significant at the 5% level based on the PDS estimation with different choices of statistical learning methods and their counterparts based on the LS method. This table shows that using the PDS method allows us to identify a greater number of potentially useful technical indicators than the LS method. A possible interpretation is that, compared to the LS method, the PDS method is more efficient for estimation by exploiting the sparsity of controls; see Feng et al. (2020) (11) for a similar fining in a different empirical application. In addition, certain technical indicators are commonly selected by the PDS methods with different statistical learning methods. Examples include MA(10,11), VOL(1,4), MOM(2), and MOM(5). This empirical finding suggests that these indicators are potentially useful for predicting equity premium at least from an in-sample viewpoint. In comparison, it is common to choose (s, l) in a subjective way such as MA(1,9) or VOL(3,12). It is interesting to observe that such a subjective selection is not significant in this empirical application.

Table 4.1: Technical Indicator Selection

		MA(s, l)	VOL(s, l)	MOM(m)
Lasso	CV	(3,5), (1,9), (8,9), (10,11)	(3,6), (6,7)	2, 5
	BIC	(8,9), (10,11)	(1,4), (3,5), (3,6)	2, 5
	plug-in	(3,5), (3,7), (1,9), (8,9), (4,10), (10,11)	(1,4), (3,6)	2, 5
adaptive Lasso	CV	(3,5), (1,9), (8,9), (1,11), (10,11)	(3,5), (9,11)	2, 5
	BIC	(8,9), (10,11)	(1,4), (3,5), (3,6)	2, 5
SCAD	CV	(10,11)	(1,4), (3,6)	2, 5
	BIC	(8,9), (10,11)	(1,4), (3,5), (3,6)	2, 5
MCP	CV	(1,11), (10,11)	(1,4), (6,7)	2, 5
	BIC	(1,11), (10,11)	(1,4), (3,5)	2, 5
OLS		(3,7)	(3,6)	5

Note. CV represents the tuning parameter chosen through the 10-fold CV.
BIC represents the tuning parameter chosen through the BIC method.

doi:10.6342/NTU202301747

Table 4.2: Empirical Results for MA Indicators

	Lasso			adaptive Lasso	
MA(s,l)	CV	BIC	plug-in	CV	BIC
(1,2)	0.003(0.005)	-0.007(0.005)	0(0.006)	0.004(0.006)	-0.006(0.005)
(1,3)	0.002(0.006)	0.001(0.006)	0.005(0.006)	0.003(0.006)	0(0.006)
(2,3)	-0.002(0.005)	0.003(0.005)	-0.002(0.006)	-0.001(0.005)	0.002(0.005)
(1,4)	0.003(0.008)	0.002(0.008)	0.002(0.008)	0.003(0.008)	0.005(0.008)
(2,4)	0.004(0.007)	-0.001(0.007)	0.004(0.007)	0.002(0.007)	-0.001(0.007)
(3,4)	0.003(0.005)	0.004(0.005)	0.004(0.006)	0.003(0.005)	0.005(0.005)
(1,5)	0.013(0.010)	0.017(0.010)	0.012(0.011)	0.014(0.010)	0.018(0.011)
(2,5)	-0.008(0.010)	-0.002(0.010)	-0.007(0.010)	-0.003(0.010)	-0.001(0.010)
(3,5)	-0.018(0.009)*	-0.014(0.009)	-0.021(0.008)**	-0.018(0.008)*	-0.013(0.009)
(4,5)	-0.003(0.006)	0.002(0.006)	-0.002(0.006)	-0.002(0.006)	0.001(0.006)
(1,6)	-0.016(0.012)	-0.018(0.012)	-0.018(0.012)	-0.019(0.012)	-0.017(0.012)
(2,6)	-0.001(0.010)	-0.006(0.010)	-0.002(0.011)	-0.005(0.010)	-0.006(0.010)
(3,6)	-0.003(0.008)	0(0.009)	-0.004(0.008)	-0.001(0.008)	0.002(0.009)
(4,6)	-0.001(0.009)	-0.002(0.009)	0(0.009)	-0.003(0.009)	-0.003(0.009)
(5,6)	0.006(0.006)	0.003(0.006)	0.004(0.006)	0.003(0.006)	0.003(0.006)
(1,7)	-0.009(0.010)	-0.005(0.009)	-0.012(0.010)	-0.009(0.010)	-0.005(0.009)
(2,7)	-0.005(0.012)	0.006(0.013)	-0.006(0.012)	0(0.013)	0.007(0.012)
(3,7)	0.017(0.010)	0.012(0.010)	0.022(0.010)*	0.014(0.010)	0.011(0.009)
(4,7)	-0.007(0.010)	-0.007(0.010)	-0.009(0.010)	-0.005(0.010)	-0.008(0.010)
(5,7)	0.007(0.008)	0.004(0.008)	0.009(0.008)	0.005(0.008)	0.005(0.008)
(6,7)	-0.004(0.006)	-0.005(0.007)	-0.002(0.007)	-0.005(0.006)	-0.005(0.006)
(1,8)	-0.010(0.012)	-0.009(0.011)	-0.011(0.013)	-0.011(0.012)	-0.008(0.012)
(2,8)	0.010(0.013)	0.006(0.014)	0.013(0.012)	0.006(0.013)	0.005(0.014)
(3,8)	-0.008(0.011)	-0.014(0.011)	-0.008(0.011)	-0.012(0.011)	-0.013(0.012)
(4,8)	0.011(0.011)	0.015(0.012)	0.011(0.011)	0.014(0.011)	0.017(0.012)
(5,8)	0.005(0.009)	0.005(0.009)	0.010(0.009)	0.005(0.009)	0.005(0.009)
(6,8)	0.005(0.009)	0.007(0.009)	0.001(0.010)	0.008(0.009)	0.006(0.009)
(7,8)	-0.006(0.007)	-0.006(0.007)	-0.005(0.007)	-0.006(0.007)	-0.005(0.007)
(1,9)	0.039(0.015)*	0.028(0.016)	0.038(0.017)*	0.032(0.015)*	0.027(0.016)
(2,9)	0(0.015)	0.003(0.017)	-0.007(0.017)	0.002(0.016)	0.004(0.017)
(3,9)	-0.010(0.020)	-0.008(0.021)	-0.013(0.021)	-0.007(0.020)	-0.010(0.021)
(4,9)	-0.008(0.015)	-0.017(0.015)	-0.002(0.016)	-0.017(0.015)	-0.019(0.015)
(5,9)	-0.014(0.011)	-0.012(0.011)	-0.018(0.012)	-0.013(0.011)	-0.013(0.011)
(6,9)	0.005(0.012)	0.006(0.012)	0.004(0.012)	0.005(0.012)	0.007(0.011)
(7,9)	-0.007(0.011)	-0.009(0.011)	-0.006(0.011)	-0.006(0.011)	-0.011(0.011)
(8,9)	0.019(0.008)*	0.017(0.008)*	0.018(0.008)*	0.018(0.008)*	0.016(0.008)*
(1,10)	-0.002(0.017)	0.002(0.020)	0.003(0.019)	-0.005(0.018)	0.003(0.020)
(2,10)	0(0.020)	0(0.021)	0.010(0.023)	-0.002(0.020)	-0.001(0.021)
(3,10)	0.007(0.021)	0.010(0.022)	0.006(0.021)	0.009(0.021)	0.013(0.021)
(4,10)	-0.028(0.015)	-0.020(0.016)	-0.036(0.016)*	-0.024(0.015)	-0.018(0.017)
(5,10)	0.007(0.014)	0.005(0.015)	0.008(0.014)	0.005(0.014)	0.003(0.014)
(6,10)	-0.005(0.011)	-0.007(0.012)	-0.005(0.012)	-0.007(0.011)	-0.008(0.012)
(7,10)	0.001(0.014)	0.004(0.014)	0(0.013)	0(0.014)	0.006(0.015)
(8,10)	-0.016(0.012)	-0.018(0.012)	-0.018(0.012)	-0.014(0.012)	-0.017(0.012)
(9,10)	-0.001(0.007)	-0.001(0.007)	-0.001(0.008)	-0.001(0.007)	-0.001(0.007)
(1,11)	0.023(0.016)	0.031(0.017)	0.03(0.016)	0.032(0.016)*	0.03(0.016)
(2,11)	-0.02(0.018)	-0.019(0.018)	-0.021(0.020)	-0.019(0.018)	-0.017(0.018)
(3,11)	0.010(0.016)	0.006(0.017)	0.013(0.017)	0.010(0.017)	0.006(0.018)
(4,11)	0.016(0.016)	0.014(0.017)	0.018(0.016)	0.017(0.016)	0.013(0.018)
(5,11)	0.001(0.015)	0.001(0.015)	-0.001(0.014)	0(0.014)	0.004(0.014)
(6,11)	0.003(0.013)	0.006(0.013)	0.003(0.014)	0(0.012)	0.005(0.013)
(7,11)	0.012(0.016)	0.009(0.016)	0.011(0.016)	0.012(0.016)	0.008(0.016)
(8,11)	-0.002(0.013)	0.004(0.013)	-0.001(0.013)	0(0.013)	0.005(0.013)
(9,11)	0.010(0.012)	0.007(0.011)	0.010(0.012)	0.009(0.012)	0.007(0.012)
(10,11)	-0.022(0.007)**	-0.021(0.007)**	-0.027(0.007)***	-0.022(0.007)**	-0.02(0.007)**
(1,12)	-0.018(0.015)	-0.021(0.014)	-0.021(0.015)	-0.017(0.014)	-0.021(0.014)
(2,12)	-0.001(0.012)	0.002(0.012)	-0.003(0.013)	0.004(0.012)	0(0.012)
(3,12)	-0.007(0.014)	-0.004(0.014)	-0.009(0.014)	-0.010(0.014)	-0.006(0.014)
(4,12)	0.002(0.015)	0.002(0.015)	0.005(0.015)	0.003(0.014)	0.002(0.015)
(5,12)	0.007(0.016)	0.005(0.016)	0.011(0.015)	0.006(0.016)	0.005(0.016)
(6,12)	-0.015(0.011)	-0.015(0.011)	-0.017(0.013)	-0.011(0.011)	-0.014(0.011)
(7,12)	0.013(0.011)	0.011(0.011)	0.014(0.012)	0.010(0.011)	0.011(0.011)
(8,12)	0.010(0.012)	0.008(0.012)	0.015(0.012)	0.008(0.012)	0.007(0.012)
(9,12)	-0.019(0.014)	-0.016(0.014)	-0.019(0.014)	-0.018(0.014)	-0.017(0.014)
(10,12)	0.007(0.009)	0.005(0.009)	0.008(0.010)	0.008(0.009)	0.006(0.009)
(11,12)	0.011(0.007)	0.010(0.007)	0.010(0.007)	0.010(0.007)	0.010(0.007)

Note. This estimator is using Equation (4.2), employing the PDS method, and adjusting covariance with the Newey-West method.
The first column MA(s,l) represents the timing parameters of MA strategy.
The numbers in brackets are standard errors.
* represents the statistical test p-value falls within the range 0.01 to 0.05.
** represents the statistical test p-value falls within the range 0.001 to 0.01.
*** represents the statistical test p-value is less than 0.001.

doi:10.6342/NTU202301747

Table 4.3: Empirical Results for MA Indicators (Cont'd)

	SCAD		MCP		OLS
MA(s,l)	CV	BIC	CV	BIC	
(1,2)	0.004(0.005)	-0.007(0.005)	-0.008(0.004)	-0.006(0.005)	0.001(0.007)
(1,3)	0.001(0.006)	0.001(0.007)	0(0.006)	0.002(0.006)	0.006(0.007)
(2,3)	-0.001(0.005)	0.002(0.005)	0.002(0.005)	0.003(0.005)	-0.002(0.006)
(1,4)	0.004(0.008)	0.004(0.008)	0.002(0.008)	0.003(0.008)	0.003(0.009)
(2,4)	0.002(0.007)	-0.002(0.007)	-0.004(0.007)	-0.003(0.007)	0.005(0.008)
(3,4)	0.001(0.005)	0.005(0.005)	0.002(0.005)	0.003(0.005)	0.002(0.006)
(1,5)	0.013(0.010)	0.017(0.011)	0.017(0.011)	0.017(0.010)	0.013(0.012)
(2,5)	-0.003(0.010)	0(0.010)	0.001(0.009)	0(0.009)	-0.007(0.010)
(3,5)	-0.016(0.009)	-0.013(0.009)	-0.013(0.009)	-0.012(0.009)	-0.022(0.009)
(4,5)	-0.001(0.006)	0.001(0.006)	0.003(0.005)	0.003(0.005)	-0.001(0.006)
(1,6)	-0.018(0.012)	-0.018(0.012)	-0.016(0.012)	-0.017(0.012)	-0.019(0.013)
(2,6)	-0.007(0.010)	-0.005(0.010)	-0.006(0.010)	-0.007(0.010)	0.001(0.012)
(3,6)	0(0.008)	0.002(0.009)	0.004(0.009)	0.002(0.009)	-0.004(0.009)
(4,6)	-0.006(0.009)	-0.001(0.009)	-0.002(0.008)	-0.002(0.008)	0.003(0.009)
(5,6)	0.003(0.006)	0.002(0.006)	0.002(0.006)	0.002(0.006)	0.001(0.007)
(1,7)	-0.010(0.009)	-0.005(0.009)	-0.005(0.010)	-0.006(0.009)	-0.014(0.012)
(2,7)	0.002(0.012)	0.005(0.012)	0.004(0.012)	0.005(0.012)	-0.009(0.013)
(3,7)	0.014(0.009)	0.009(0.009)	0.008(0.009)	0.010(0.009)	0.024(0.011)*
(4,7)	-0.004(0.010)	-0.009(0.010)	-0.008(0.010)	-0.008(0.010)	-0.010(0.011)
(5,7)	0.005(0.008)	0.006(0.008)	0.005(0.008)	0.005(0.008)	0.011(0.008)
(6,7)	-0.003(0.006)	-0.006(0.006)	-0.005(0.006)	-0.006(0.006)	-0.003(0.007)
(1,8)	-0.012(0.012)	-0.007(0.012)	-0.009(0.012)	-0.008(0.012)	-0.005(0.014)
(2,8)	0.006(0.013)	0.006(0.013)	0.008(0.013)	0.007(0.013)	0.013(0.013)
(3,8)	-0.011(0.011)	-0.012(0.011)	-0.012(0.011)	-0.013(0.011)	-0.009(0.012)
(4,8)	0.014(0.011)	0.017(0.012)	0.018(0.012)	0.02(0.012)	0.015(0.012)
(5,8)	0.005(0.009)	0.003(0.008)	0.006(0.008)	0.006(0.009)	0.007(0.010)
(6,8)	0.006(0.009)	0.007(0.009)	0.007(0.009)	0.007(0.009)	0.002(0.011)
(7,8)	-0.005(0.007)	-0.005(0.007)	-0.005(0.007)	-0.004(0.007)	-0.005(0.008)
(1,9)	0.03(0.015)	0.027(0.016)	0.023(0.016)	0.023(0.016)	0.034(0.018)
(2,9)	0.001(0.016)	0.001(0.017)	0.003(0.018)	0.002(0.018)	-0.011(0.017)
(3,9)	-0.009(0.020)	-0.005(0.021)	-0.009(0.021)	-0.009(0.021)	-0.015(0.022)
(4,9)	-0.017(0.014)	-0.019(0.015)	-0.021(0.015)	-0.023(0.015)	-0.005(0.017)
(5,9)	-0.013(0.011)	-0.012(0.011)	-0.013(0.011)	-0.013(0.011)	-0.013(0.013)
(6,9)	0.007(0.011)	0.006(0.012)	0.006(0.011)	0.006(0.011)	0.002(0.013)
(7,9)	-0.009(0.011)	-0.008(0.011)	-0.009(0.011)	-0.009(0.011)	-0.004(0.011)
(8,9)	0.014(0.008)	0.018(0.008)*	0.014(0.008)	0.014(0.008)	0.017(0.009)
(1,10)	0(0.018)	0.001(0.020)	0.004(0.020)	0.004(0.021)	0.009(0.020)
(2,10)	-0.004(0.021)	0(0.021)	-0.004(0.022)	-0.004(0.022)	0.011(0.025)
(3,10)	0.013(0.021)	0.010(0.021)	0.013(0.021)	0.012(0.021)	0.009(0.023)
(4,10)	-0.019(0.016)	-0.021(0.017)	-0.019(0.017)	-0.019(0.017)	-0.041(0.018)
(5,10)	0.004(0.014)	0.003(0.014)	0.001(0.014)	0.003(0.014)	0.011(0.016)
(6,10)	-0.007(0.012)	-0.008(0.012)	-0.007(0.012)	-0.007(0.012)	-0.003(0.012)
(7,10)	0.003(0.015)	0.006(0.015)	0.006(0.015)	0.006(0.015)	-0.003(0.014)
(8,10)	-0.015(0.012)	-0.015(0.012)	-0.016(0.012)	-0.016(0.011)	-0.013(0.013)
(9,10)	0(0.007)	-0.002(0.007)	0(0.007)	-0.001(0.007)	-0.003(0.008)
(1,11)	0.028(0.016)	0.031(0.016)	0.033(0.016)*	0.033(0.016)*	0.023(0.016)
(2,11)	-0.015(0.018)	-0.02(0.018)	-0.015(0.018)	-0.015(0.018)	-0.021(0.023)
(3,11)	0.007(0.017)	0.008(0.017)	0.007(0.017)	0.008(0.017)	0.008(0.018)
(4,11)	0.012(0.017)	0.015(0.017)	0.015(0.018)	0.014(0.018)	0.017(0.017)
(5,11)	0.003(0.014)	0.003(0.014)	0.005(0.013)	0.002(0.014)	-0.002(0.016)
(6,11)	-0.001(0.012)	0.003(0.013)	0.005(0.013)	0.004(0.013)	0.001(0.014)
(7,11)	0.012(0.016)	0.007(0.016)	0.008(0.015)	0.009(0.016)	0.011(0.017)
(8,11)	0.003(0.013)	0.005(0.013)	0.004(0.013)	0.004(0.013)	-0.005(0.014)
(9,11)	0.008(0.012)	0.007(0.012)	0.007(0.012)	0.007(0.012)	0.012(0.013)
(10,11)	-0.022(0.007)**	-0.022(0.007)**	-0.021(0.007)**	-0.022(0.007)**	-0.026(0.007)
(1,12)	-0.018(0.014)	-0.02(0.014)	-0.022(0.013)	-0.021(0.013)	-0.018(0.016)
(2,12)	0(0.012)	0.002(0.013)	-0.002(0.012)	-0.002(0.012)	-0.004(0.014)
(3,12)	-0.009(0.014)	-0.006(0.014)	-0.006(0.014)	-0.006(0.014)	-0.006(0.015)
(4,12)	0.004(0.014)	0.001(0.014)	0.001(0.014)	0.002(0.014)	0.008(0.015)
(5,12)	0.005(0.016)	0.007(0.016)	0.007(0.016)	0.009(0.016)	0.007(0.016)
(6,12)	-0.012(0.011)	-0.015(0.011)	-0.017(0.011)	-0.018(0.011)	-0.016(0.013)
(7,12)	0.009(0.011)	0.010(0.011)	0.011(0.011)	0.010(0.011)	0.016(0.013)
(8,12)	0.007(0.012)	0.005(0.012)	0.009(0.012)	0.008(0.012)	0.017(0.013)
(9,12)	-0.019(0.014)	-0.016(0.014)	-0.016(0.014)	-0.016(0.014)	-0.021(0.015)
(10,12)	0.009(0.009)	0.006(0.009)	0.004(0.009)	0.005(0.009)	0.006(0.011)
(11,12)	0.011(0.007)	0.010(0.007)	0.011(0.007)	0.010(0.007)	0.011(0.008)

Note. This estimator is using Equation (4.2), employing the PDS method, and adjusting covariance with the Newey-West method.
The first column MA(s,l) represents the timing parameters of MA strategy.
The numbers in brackets are standard errors.
* represents the statistical test p-value falls within the range 0.01 to 0.05.
** represents the statistical test p-value falls within the range 0.001 to 0.01.
*** represents the statistical test p-value is less than 0.001.

doi:10.6342/NTU202301747

Table 4.4: Empirical Results for VOL Indicators

	Lasso			adaptive Lasso	
VOL(s,l)	CV	BIC	plug-in	CV	BIC
(1,2)	0.005(0.007)	0.003(0.006)	0.004(0.009)	0.002(0.007)	0.002(0.006)
(1,3)	0.004(0.006)	0.001(0.006)	0.003(0.007)	0.005(0.006)	0.001(0.006)
(2,3)	-0.004(0.005)	-0.001(0.005)	-0.005(0.005)	-0.002(0.005)	-0.001(0.005)
(1,4)	-0.014(0.007)	-0.018(0.008)*	-0.017(0.008)*	-0.015(0.008)	-0.018(0.008)*
(2,4)	-0.002(0.009)	0.001(0.009)	-0.004(0.009)	-0.002(0.009)	0.001(0.009)
(3,4)	0.007(0.006)	0.003(0.006)	0.007(0.007)	0.004(0.006)	0.004(0.006)
(1,5)	0.010(0.008)	0.015(0.009)	0.013(0.009)	0.014(0.009)	0.015(0.009)
(2,5)	0.008(0.009)	0.007(0.009)	0.011(0.010)	0.004(0.010)	0.006(0.009)
(3,5)	-0.011(0.007)	-0.016(0.008)*	-0.013(0.008)	-0.016(0.008)*	-0.016(0.008)*
(4,5)	0.004(0.005)	0.003(0.005)	0.004(0.006)	0(0.005)	0.003(0.005)
(1,6)	-0.006(0.010)	-0.006(0.010)	-0.009(0.010)	-0.005(0.010)	-0.005(0.010)
(2,6)	0(0.010)	-0.003(0.010)	-0.002(0.010)	0(0.009)	-0.003(0.010)
(3,6)	0.019(0.010)*	0.022(0.010)*	0.026(0.010)**	0.019(0.010)	0.021(0.010)*
(4,6)	-0.004(0.008)	-0.004(0.008)	-0.007(0.008)	-0.002(0.008)	-0.002(0.008)
(5,6)	0.001(0.007)	-0.001(0.007)	0.001(0.008)	0.005(0.007)	0.001(0.007)
(1,7)	0.007(0.012)	0.012(0.011)	0.011(0.011)	0.008(0.012)	0.012(0.012)
(2,7)	-0.015(0.012)	-0.012(0.012)	-0.015(0.013)	-0.015(0.011)	-0.012(0.011)
(3,7)	-0.005(0.011)	-0.007(0.011)	-0.006(0.011)	-0.004(0.011)	-0.007(0.011)
(4,7)	0.010(0.010)	0.011(0.011)	0.010(0.011)	0.010(0.010)	0.010(0.011)
(5,7)	-0.003(0.009)	-0.002(0.009)	-0.006(0.010)	-0.006(0.008)	-0.003(0.009)
(6,7)	0.013(0.007)*	0.010(0.007)	0.011(0.007)	0.011(0.007)	0.010(0.007)
(1,8)	-0.008(0.011)	-0.004(0.012)	-0.014(0.011)	-0.007(0.011)	-0.004(0.012)
(2,8)	-0.010(0.012)	-0.008(0.011)	-0.012(0.013)	-0.010(0.011)	-0.007(0.011)
(3,8)	0.009(0.009)	0.007(0.009)	0.009(0.010)	0.007(0.009)	0.007(0.009)
(4,8)	-0.013(0.012)	-0.015(0.012)	-0.008(0.013)	-0.012(0.011)	-0.015(0.012)
(5,8)	-0.009(0.010)	-0.007(0.011)	-0.010(0.011)	-0.006(0.010)	-0.006(0.011)
(6,8)	0(0.008)	0.002(0.009)	0(0.010)	0.001(0.008)	0.002(0.008)
(7,8)	0.009(0.007)	0.009(0.007)	0.007(0.007)	0.006(0.006)	0.007(0.007)
(1,9)	-0.001(0.015)	0(0.015)	0.001(0.015)	0.003(0.015)	0(0.015)
(2,9)	0.004(0.013)	0.002(0.013)	0.013(0.013)	0.001(0.013)	0.002(0.013)
(3,9)	-0.015(0.009)	-0.013(0.010)	-0.014(0.009)	-0.012(0.010)	-0.013(0.009)
(4,9)	-0.004(0.012)	0.002(0.012)	-0.009(0.013)	0.001(0.011)	0.004(0.012)
(5,9)	0.005(0.012)	0.002(0.013)	0.006(0.013)	0.005(0.012)	0.003(0.013)
(6,9)	-0.003(0.012)	-0.001(0.013)	-0.002(0.012)	-0.002(0.012)	-0.002(0.012)
(7,9)	-0.002(0.008)	-0.002(0.009)	0(0.009)	-0.003(0.008)	-0.002(0.009)
(8,9)	0.008(0.006)	0.010(0.006)	0.007(0.006)	0.009(0.005)	0.010(0.006)
(1,10)	-0.02(0.015)	-0.011(0.014)	-0.012(0.016)	-0.019(0.015)	-0.011(0.014)
(2,10)	-0.004(0.016)	-0.009(0.015)	-0.004(0.018)	-0.002(0.016)	-0.009(0.015)
(3,10)	0.02(0.014)	0.018(0.016)	0.021(0.015)	0.018(0.015)	0.02(0.016)
(4,10)	0.014(0.011)	0.012(0.012)	0.012(0.012)	0.009(0.011)	0.012(0.012)
(5,10)	-0.010(0.010)	-0.010(0.011)	-0.011(0.009)	-0.013(0.010)	-0.012(0.011)
(6,10)	-0.018(0.010)	-0.018(0.010)	-0.012(0.010)	-0.018(0.010)	-0.018(0.010)
(7,10)	-0.006(0.009)	-0.008(0.010)	-0.007(0.009)	-0.009(0.009)	-0.007(0.010)
(8,10)	0.007(0.008)	0.006(0.008)	0.008(0.009)	0.008(0.008)	0.006(0.008)
(9,10)	-0.004(0.007)	-0.003(0.007)	-0.005(0.008)	-0.003(0.007)	-0.004(0.007)
(1,11)	0.021(0.011)	0.013(0.012)	0.018(0.013)	0.017(0.011)	0.012(0.012)
(2,11)	0.001(0.016)	0.013(0.017)	-0.006(0.017)	0.007(0.016)	0.014(0.017)
(3,11)	-0.002(0.013)	-0.004(0.015)	-0.006(0.017)	-0.005(0.014)	-0.005(0.015)
(4,11)	0(0.012)	0(0.013)	0.003(0.013)	-0.002(0.013)	-0.002(0.013)
(5,11)	-0.001(0.015)	-0.002(0.015)	0.001(0.014)	0.002(0.016)	-0.002(0.016)
(6,11)	0.011(0.012)	0.012(0.013)	0.010(0.014)	0.011(0.013)	0.011(0.013)
(7,11)	-0.010(0.011)	-0.008(0.012)	-0.015(0.011)	-0.008(0.012)	-0.008(0.012)
(8,11)	-0.005(0.011)	0(0.011)	-0.001(0.011)	-0.003(0.011)	0(0.010)
(9,11)	0.017(0.009)	0.013(0.010)	0.018(0.010)	0.018(0.009)*	0.014(0.009)
(10,11)	0.004(0.007)	0.002(0.008)	0.006(0.007)	0.005(0.007)	0.002(0.008)
(1,12)	0.005(0.011)	0.001(0.012)	0.009(0.013)	0.002(0.012)	0.001(0.012)
(2,12)	0.009(0.014)	0.008(0.016)	0.011(0.016)	0.005(0.015)	0.007(0.016)
(3,12)	-0.001(0.011)	-0.008(0.011)	0.001(0.014)	0.001(0.011)	-0.006(0.011)
(4,12)	0.001(0.012)	-0.001(0.010)	-0.002(0.012)	0.002(0.011)	0(0.010)
(5,12)	0.007(0.012)	0.011(0.013)	0.006(0.013)	0.010(0.012)	0.013(0.013)
(6,12)	-0.007(0.011)	-0.006(0.011)	-0.007(0.012)	-0.005(0.011)	-0.006(0.011)
(7,12)	0.010(0.013)	0.008(0.013)	0.011(0.014)	0.006(0.013)	0.008(0.013)
(8,12)	-0.015(0.010)	-0.015(0.011)	-0.019(0.012)	-0.013(0.010)	-0.014(0.011)
(9,12)	-0.009(0.011)	-0.009(0.012)	-0.009(0.012)	-0.012(0.011)	-0.010(0.012)
(10,12)	-0.006(0.011)	-0.003(0.011)	-0.008(0.011)	-0.001(0.011)	-0.003(0.011)
(11,12)	0.009(0.007)	0.008(0.007)	0.010(0.008)	0.007(0.007)	0.008(0.008)

Note. This estimator is using Equation (4.2), employing the PDS method, and adjusting covariance with the Newey-West method.
The first column VOL(s,l) represents the timing parameters of VOL strategy.
The numbers in brackets are standard errors.
* represents the statistical test p-value falls within the range 0.01 to 0.05.
** represents the statistical test p-value falls within the range 0.001 to 0.01.
*** represents the statistical test p-value is less than 0.001.

doi:10.6342/NTU202301747

Table 4.5: Empirical Results for VOL Indicators (Cont'd)

	SCAD		MCP		OLS
VOL(s,l)	CV	BIC	CV	BIC	
(1,2)	0.005(0.007)	0.002(0.007)	0.005(0.007)	0.002(0.006)	0.004(0.009)
(1,3)	0.003(0.006)	0.001(0.006)	0.005(0.006)	0.004(0.006)	0.003(0.007)
(2,3)	-0.004(0.005)	-0.001(0.005)	-0.004(0.005)	-0.002(0.005)	-0.004(0.006)
(1,4)	-0.015(0.007)*	-0.019(0.008)*	-0.015(0.007)*	-0.015(0.008)*	-0.018(0.008)
(2,4)	-0.002(0.009)	0.001(0.009)	-0.002(0.008)	0.001(0.009)	-0.004(0.010)
(3,4)	0.007(0.006)	0.003(0.007)	0.008(0.006)	0.004(0.006)	0.008(0.008)
(1,5)	0.011(0.008)	0.016(0.009)	0.010(0.008)	0.016(0.009)	0.012(0.009)
(2,5)	0.007(0.009)	0.006(0.010)	0.008(0.009)	0.003(0.010)	0.011(0.010)
(3,5)	-0.012(0.007)	-0.016(0.008)*	-0.012(0.007)	-0.015(0.008)*	-0.015(0.008)
(4,5)	0.002(0.005)	0.003(0.005)	0.003(0.005)	0.002(0.005)	0.004(0.006)
(1,6)	-0.006(0.010)	-0.006(0.010)	-0.005(0.01)	-0.004(0.010)	-0.009(0.011)
(2,6)	0(0.010)	-0.004(0.009)	0.001(0.01)	-0.001(0.010)	-0.004(0.010)
(3,6)	0.02(0.010)*	0.021(0.010)*	0.019(0.01)	0.019(0.010)	0.028(0.010)**
(4,6)	-0.002(0.008)	-0.002(0.008)	-0.003(0.008)	-0.004(0.008)	-0.008(0.009)
(5,6)	0.003(0.007)	0.002(0.007)	0.002(0.007)	0.003(0.007)	0.001(0.008)
(1,7)	0.007(0.012)	0.008(0.012)	0.005(0.012)	0.008(0.011)	0.012(0.012)
(2,7)	-0.015(0.012)	-0.013(0.012)	-0.016(0.012)	-0.015(0.012)	-0.014(0.014)
(3,7)	-0.006(0.011)	-0.005(0.011)	-0.005(0.011)	-0.006(0.011)	-0.004(0.012)
(4,7)	0.009(0.010)	0.010(0.011)	0.009(0.01)	0.014(0.010)	0.011(0.011)
(5,7)	-0.004(0.009)	-0.004(0.009)	-0.003(0.008)	-0.007(0.009)	-0.008(0.010)
(6,7)	0.012(0.006)	0.011(0.007)	0.014(0.007)*	0.010(0.007)	0.012(0.008)
(1,8)	-0.007(0.012)	-0.003(0.012)	-0.007(0.011)	-0.006(0.012)	-0.014(0.012)
(2,8)	-0.009(0.011)	-0.007(0.011)	-0.011(0.011)	-0.008(0.011)	-0.017(0.013)
(3,8)	0.009(0.009)	0.006(0.009)	0.008(0.009)	0.006(0.009)	0.006(0.011)
(4,8)	-0.011(0.012)	-0.015(0.012)	-0.011(0.012)	-0.014(0.011)	-0.011(0.014)
(5,8)	-0.010(0.010)	-0.006(0.011)	-0.008(0.01)	-0.006(0.010)	-0.009(0.011)
(6,8)	0.003(0.008)	0.002(0.008)	0(0.008)	0.003(0.008)	-0.001(0.011)
(7,8)	0.007(0.007)	0.006(0.007)	0.009(0.007)	0.010(0.007)	0.007(0.008)
(1,9)	-0.001(0.014)	0.003(0.014)	0(0.014)	0.002(0.014)	-0.001(0.016)
(2,9)	0.005(0.013)	0.002(0.013)	0.005(0.013)	0.001(0.013)	0.019(0.014)
(3,9)	-0.016(0.009)	-0.012(0.010)	-0.015(0.009)	-0.010(0.009)	-0.012(0.010)
(4,9)	0(0.012)	0.004(0.011)	-0.003(0.011)	0.002(0.012)	-0.008(0.013)
(5,9)	0.004(0.012)	0.002(0.012)	0.003(0.012)	0.004(0.012)	0.005(0.014)
(6,9)	-0.004(0.012)	-0.001(0.012)	-0.003(0.012)	-0.004(0.011)	-0.002(0.013)
(7,9)	0(0.008)	-0.003(0.009)	-0.003(0.008)	-0.003(0.009)	0(0.010)
(8,9)	0.010(0.005)	0.009(0.006)	0.008(0.006)	0.012(0.006)*	0.004(0.007)
(1,10)	-0.017(0.015)	-0.017(0.014)	-0.022(0.015)	-0.019(0.014)	-0.009(0.017)
(2,10)	-0.003(0.015)	-0.010(0.015)	-0.002(0.015)	-0.003(0.015)	-0.003(0.018)
(3,10)	0.018(0.014)	0.019(0.016)	0.021(0.014)	0.016(0.015)	0.019(0.015)
(4,10)	0.012(0.010)	0.009(0.011)	0.013(0.01)	0.008(0.012)	0.015(0.013)
(5,10)	-0.012(0.010)	-0.013(0.011)	-0.011(0.01)	-0.015(0.010)	-0.010(0.011)
(6,10)	-0.017(0.010)	-0.017(0.011)	-0.019(0.01)	-0.017(0.010)	-0.013(0.011)
(7,10)	-0.008(0.009)	-0.008(0.010)	-0.006(0.009)	-0.008(0.010)	-0.006(0.010)
(8,10)	0.007(0.009)	0.007(0.008)	0.007(0.008)	0.007(0.008)	0.009(0.009)
(9,10)	-0.004(0.007)	-0.002(0.007)	-0.005(0.007)	-0.005(0.007)	-0.004(0.008)
(1,11)	0.018(0.012)	0.014(0.012)	0.021(0.011)	0.015(0.012)	0.018(0.013)
(2,11)	0.005(0.015)	0.013(0.016)	0.004(0.015)	0.009(0.016)	-0.011(0.019)
(3,11)	0(0.014)	-0.004(0.015)	-0.002(0.013)	-0.003(0.015)	-0.005(0.017)
(4,11)	-0.001(0.012)	0.001(0.013)	-0.001(0.011)	0(0.013)	0.006(0.015)
(5,11)	-0.002(0.015)	-0.002(0.016)	-0.001(0.014)	0(0.015)	-0.004(0.016)
(6,11)	0.010(0.013)	0.011(0.013)	0.012(0.012)	0.010(0.013)	0.009(0.016)
(7,11)	-0.008(0.012)	-0.007(0.012)	-0.010(0.011)	-0.008(0.012)	-0.016(0.012)
(8,11)	-0.005(0.011)	-0.001(0.010)	-0.005(0.011)	-0.004(0.011)	-0.002(0.011)
(9,11)	0.018(0.009)	0.013(0.009)	0.018(0.009)	0.017(0.009)	0.017(0.011)
(10,11)	0.005(0.007)	0.002(0.008)	0.004(0.007)	0(0.008)	0.007(0.008)
(1,12)	0.004(0.011)	0.002(0.012)	0.004(0.011)	0.005(0.012)	0.008(0.013)
(2,12)	0.005(0.014)	0.008(0.016)	0.006(0.014)	0.004(0.014)	0.016(0.017)
(3,12)	-0.001(0.011)	-0.007(0.011)	0(0.011)	-0.003(0.011)	-0.004(0.015)
(4,12)	0.002(0.011)	0.001(0.011)	0.001(0.011)	0.003(0.011)	-0.001(0.013)
(5,12)	0.009(0.012)	0.013(0.013)	0.009(0.012)	0.012(0.013)	0.008(0.014)
(6,12)	-0.009(0.011)	-0.005(0.011)	-0.007(0.011)	-0.005(0.011)	-0.006(0.013)
(7,12)	0.010(0.013)	0.007(0.013)	0.009(0.013)	0.008(0.013)	0.012(0.014)
(8,12)	-0.014(0.010)	-0.014(0.010)	-0.014(0.01)	-0.016(0.010)	-0.018(0.013)
(9,12)	-0.011(0.011)	-0.010(0.012)	-0.008(0.011)	-0.010(0.012)	-0.012(0.012)
(10,12)	-0.006(0.011)	-0.002(0.011)	-0.006(0.011)	-0.001(0.011)	-0.007(0.012)
(11,12)	0.009(0.007)	0.008(0.008)	0.009(0.007)	0.008(0.007)	0.011(0.008)

Note. This estimator is using Equation (4.2), employing the PDS method, and adjusting covariance with the Newey-West method.
The first column VOL(s,l) represents the timing parameters of VOL strategy.
The numbers in brackets are standard errors.
* represents the statistical test p-value falls within the range 0.01 to 0.05.
** represents the statistical test p-value falls within the range 0.001 to 0.01.
*** represents the statistical test p-value is less than 0.001.

doi:10.6342/NTU202301747

Table 4.6: Empirical Results for MOM Indicators

	Lasso			adaptive Lasso	
MOM(m)	CV	BIC	plug-in	CV	BIC
(1)	0.006(0.005)	0.004(0.004)	0.005(0.005)	0.006(0.005)	0.004(0.004)
(2)	-0.018(0.006)**	-0.016(0.006)**	-0.021(0.006)***	-0.018(0.006)**	-0.017(0.006)**
(3)	0(0.006)	0.002(0.004)	0.001(0.006)	0.001(0.006)	0.002(0.005)
(4)	0.007(0.006)	0.003(0.005)	0.010(0.006)	0.008(0.005)	0.003(0.005)
(5)	0.02(0.007)**	0.017(0.006)**	0.027(0.008)***	0.02(0.006)**	0.017(0.006)**
(6)	-0.011(0.008)	-0.009(0.007)	-0.014(0.008)	-0.011(0.008)	-0.009(0.007)
(7)	-0.007(0.009)	0.003(0.007)	-0.007(0.009)	-0.006(0.009)	0.003(0.007)
(8)	-0.002(0.009)	0.004(0.008)	-0.003(0.009)	-0.003(0.009)	0.006(0.008)
(9)	-0.006(0.009)	-0.006(0.008)	-0.005(0.010)	-0.005(0.009)	-0.007(0.009)
(10)	0.007(0.009)	0.013(0.009)	0.008(0.009)	0.008(0.009)	0.013(0.009)
(11)	0.002(0.008)	-0.007(0.007)	0.003(0.008)	0.001(0.008)	-0.007(0.007)
(12)	-0.001(0.008)	0(0.007)	-0.002(0.007)	0(0.008)	-0.001(0.007)

	SCAD		MCP		OLS
MOM(m)	CV	BIC	CV	BIC	
(1)	0.006(0.004)	0.005(0.004)	0.006(0.004)	0.006(0.004)	0.007(0.006)
(2)	-0.015(0.006)**	-0.018(0.006)**	-0.014(0.006)*	-0.017(0.006)**	-0.023(0.007)
(3)	-0.001(0.005)	0.001(0.005)	0.001(0.005)	0.001(0.005)	-0.001(0.007)
(4)	0.002(0.005)	0.004(0.005)	0(0.005)	0.004(0.005)	0.008(0.007)
(5)	0.015(0.006)*	0.017(0.006)**	0.012(0.006)*	0.017(0.006)**	0.025(0.009)**
(6)	-0.013(0.007)	-0.009(0.008)	-0.013(0.007)	-0.009(0.007)	-0.010(0.009)
(7)	-0.004(0.008)	0.003(0.007)	-0.004(0.008)	0.003(0.007)	-0.003(0.011)
(8)	-0.002(0.008)	0.006(0.008)	0(0.008)	0.005(0.008)	-0.009(0.010)
(9)	-0.008(0.009)	-0.007(0.009)	-0.007(0.008)	-0.005(0.008)	-0.003(0.010)
(10)	0.011(0.009)	0.013(0.009)	0.012(0.008)	0.015(0.009)	0.011(0.011)
(11)	-0.005(0.007)	-0.005(0.007)	-0.004(0.007)	-0.007(0.007)	0.006(0.009)
(12)	0.002(0.008)	0(0.007)	-0.001(0.007)	0(0.007)	-0.006(0.008)

Note. This estimator is using Equation (4.2), employing the PDS method, and adjusting covariance with the Newey-West method.

The first column MOM(m) represents the timing parameter of MOM strategy.

The numbers in brackets are standard errors.

* represents the statistical test p-value falls within the range 0.01 to 0.05.

** represents the statistical test p-value falls within the range 0.001 to 0.01.

*** represents the statistical test p-value is less than 0.001.

doi:10.6342/NTU202301747

5 Conclusions

This study conducted a Monte Carlo simulation and an empirical application to assess the finite-sample performance of the PDS methods with different choices of statistical learning methods for the double selection. The simulation shows that the finite-sample performance of the PDS method may vary with different choices of statistical learning methods in general. In particular, the simulation shows that the plug-in Lasso and the adaptive Lasso, utilizing the tuning parameter selected by the BIC, have proper performance for the double selection. Moreover, the latter tends to choose the control variables in a sparser way than the former. The empirical application also shows that, compared to the least squares method, the PDS method is useful for finding more significant technical indicators signal for predicting the equity premium based on the in-sample analysis. This might be related to the fact that the truly useful technical indicators are indeed highly sparse, it is therefore essential to improve the estimation efficiency by exploiting this sparsity using the PDS method.

However, it is worth noting that this study is not fully comprehensive, and there is still room for further research. An important direction of future research is to assess the finite-sample performance of the PDS method in different modeling frameworks, such as time series models (Hecq et al., 2023)(16), penal models (Belloni et al., 2016)(3), or neural network models (Calvo-Pardoe et al., 2021)(5). Additionally, more statistical learning

doi:10.6342/NTU202301747

methods and other types of selection methods could also be considered. In addition, the PDS method considered in this study relies on the assumption of sparsity. However, it may be worth considering relaxing this assumption and using the principal components method instead (Galbraith and Walsh, 2020)(12). In the empirical scenario, this study focuses on assessing whether technical indicators are useful for explaining the equity premiums. Future research may also consider more control variables, such as the macroeconomic indicators, in evaluating the performance of the technical indicators.

doi:10.6342/NTU202301747

References

[1] Belloni, A., Chen, D., Chernozhukov, V., and Hansen, C. (2012). Sparse models and methods for optimal instruments with an application to eminent domain, *Econometrica*, **80(6)**, 2369–2429.

[2] Belloni, A., Chernozhukov, V., and Hansen, C. (2014). Inference on treatment effects after selection among high-dimensional controls, *The Review of Economic Studies*, **81(2)**, 608–650.

[3] Belloni, A., Chernozhukov, V., Hansen, C., and Kozbur, D. (2016). Inference in high-dimensional panel models with an application to gun control, *Journal of Business and Economic Statistics*, **34(4)**, 590–605.

[4] Berset, S., Huber, M., and Schelker, M. (2023). The fiscal response to revenue shocks, *International Tax and Public Finance*, **30**, 814–848.

[5] Calvo-Pardo, H., Mancini, T., and Olmo, J. (2021). Granger causality detection in high-dimensional systems using feedforward neural networks, *International Journal of Forecasting*, **37(2)**, 920–940.

[6] Chinco, A., Clark-Joseph, A. D., and Ye, M. (2019). Sparse signals in the cross-section of returns, *The Journal of Finance*, **74(1)**, 449–492.

doi:10.6342/NTU202301747

[7] Dai, Z., Zhu, H., and Kang, J. (2021). New technical indicators and stock returns predictability, *International Review of Economics and Finance*, **71**, 127–142.

[8] Drukker, D. M. and Liu, D. (2022). Finite-sample results for Lasso and stepwise Neyman-orthogonal Poisson estimators, *Econometric Reviews*, **41(9)**, 1047–1076.

[9] Enke, B. (2020). Moral values and voting, *Journal of Political Economy*, **128(10)**, 3679–3729.

[10] Fan, J. and Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties, *Journal of the American Statistical Association*, **96(456)**, 1348–1360.

[11] Feng, G., Giglio, S., and Xiu, D. (2020). Taming the factor zoo: a test of new factors, *The Journal of Finance*, **75(3)**, 1327–1370.

[12] Galbraith, J. W. and Zinde-Walsh, V. (2020). Simple and reliable estimators of coefficients of interest in a model with high-dimensional confounding effects, *Journal of Econometrics*, **218(2)**, 609–632.

[13] Goh, J., Jiang, F., Tu, J., and Zhou, G. (2012). Forecasting government bond risk premia using technical indicators, In *25th Australasian Finance and Banking Conference*.

[14] Harvey, C. R., Liu, Y., and Zhu, H. (2016). … and the cross-section of expected returns, *The Review of Financial Studies*, **29(1)**, 5–68.

[15] Hastie, T., Tibshirani, R., and Wainwright, M. (2015). *Statistical learning with sparsity: the Lasso and generalizations*. Boca Raton, FL: CRC Press.

doi:10.6342/NTU202301747

[16] Hecq, A., Margaritella, L., and Smeekes, S. (2023). Granger causality testing in high-dimensional VARs: a post-double-selection procedure, *Journal of Financial Econometrics*, **21(3)**, 915–958.

[17] Henrique, B. M., Sobreiro, V. A., and Kimura, H. (2018). Stock price prediction using support vector regression on daily and up to the minute prices, *The Journal of Finance and Data Science*, **4(3)**, 183–201.

[18] Hsu, P.-H., Hsu, Y.-C., and Kuan, C.-M. (2010). Testing the predictive ability of technical analysis using a new stepwise test without data snooping bias, *Journal of Empirical Finance*, **17(3)**, 471–484.

[19] Hui, F. K., Warton, D. I., and Foster, S. D. (2015). Tuning parameter selection for the adaptive Lasso using ERIC, *Journal of the American Statistical Association*, **110(509)**, 262–269.

[20] Leng, C., Lin, Y., and Wahba, G. (2006). A note on the Lasso and related procedures in model selection, *Statistica Sinica*, **16(4)**, 1273–1284.

[21] Liu-Evans, G. and Mitra, S. (2019). Informality and bank stability, *Economics Letters*, **182**, 122–125.

[22] Lo, A. W. and MacKinlay, A. C. (1990). Data-snooping biases in tests of financial asset pricing models, *The Review of Financial Studies*, **3(3)**, 431–467.

[23] Mitra, S. K. (2002). Profiting from technical analysis in Indian stock market, *Finance India*, **16(1)**, 109–120.

[24] Neely, C. J., Rapach, D. E., Tu, J., and Zhou, G. (2014). Forecasting the equity risk premium: the role of technical indicators, *Management Science*, **60(7)**, 1772–1791.

doi:10.6342/NTU202301747

[25] Qi, M. and Wu, Y. (2006). Technical trading-rule profitability, data snooping, and reality check: evidence from the foreign exchange market. *Journal of Money, Credit and Banking*, **38(8)**, 2135–2158.

[26] Sullivan, R., Timmermann, A., and White, H. (1999). Data-snooping, technical trading rule performance, and the bootstrap, *The Journal of Finance*, **54(5)**, 1647–1691.

[27] Tibshirani, R. (1996). Regression shrinkage and selection via the Lasso, *Journal of the Royal Statistical Society Series B: Statistical Methodology*, **58(1)**, 267–288.

[28] Wang, H., Li, B., and Leng, C. (2009). Shrinkage tuning parameter selection with a diverging number of parameters, *Journal of the Royal Statistical Society Series B: Statistical Methodology*, **71(3)**, 671–683.

[29] Wu, Y. and Wang, L. (2020). A survey of tuning parameter selection for high-dimensional regression, *Annual Review of Statistics and Its Application*, **7**, 209–226.

[30] Wüthrich, K. and Zhu, Y. (2023). Omitted variable bias of Lasso-based inference methods: a finite sample analysis, *The Review of Economics and Statistics*, **105(4)**, 982–997.

[31] Xiao, H. and Sun, Y. (2019). On tuning parameter selection in model selection and model averaging: a Monte Carlo study, *Journal of Risk and Financial Management*, **12(3)**, 109.

[32] Zhang, C.-H. (2010). Nearly unbiased variable selection under minimax concave penalty, *The Annals of Statistics*, **38(2)**, 894–942.

[33] Zhang, Y., Li, R., and Tsai, C.-L. (2010). Regularization parameter selections via

doi:10.6342/NTU202301747

generalized information criterion, *Journal of the American Statistical Association*, **105(489)**, 312–323.

[34] Zhao, P., and Yu, B. (2006). On model selection consistency of Lasso, *The Journal of Machine Learning Research*, **7**, 2541–2563.

[35] Zou, H. (2006). The adaptive Lasso and its oracle properties, *Journal of the American Statistical Association*, **101(476)**, 1418–1429.

doi:10.6342/NTU202301747