

doi:10.19920/j.cnki.jmsc.2026.01.002

基于信噪比改进预测的金融建模方法与实证研究^①

林艳艳, 朱顺伟, 吴冲锋*

(上海交通大学安泰经济与管理学院, 上海 200030)

摘要: 在经典的线性回归模型基础上, 考虑拟合模型参数估计的随机性因素后, 应该对经典的线性回归模型给予一定的修正才能更好地预测未来. 因此, 本文提出基于信息噪声比修正的L乘子模型, 它架起拟合最优模型与预测最优模型的关系. 文中给出了最优L乘子的显式解, 并在考虑预测变量存在自相关性的情况下, 给出了L的改进表达式. 然后, 本文探讨了L乘子的两种估计方法. 实证上, 通过将L乘子模型应用于股票收益率的预测问题验证了本文理论有效性. 结果表明: 1) 相较于基准的线性回归模型, L乘子模型具有更高的预测精度; 2) 噪声较高、历史样本量较小或预测变量的信息含量较低时, L乘子的修正力度较大, 带来的预测改进效果越明显; 3) 对于均值-方差效用的投资者而言, 这种预测性能的改进可以为其带来投资绩效提升. 这些结论也满足一系列的稳健性检验.

关键词: 线性回归; 噪声; 信息噪声比; 股票收益率; 预测

中图分类号: F83 **文献标识码:** A **文章编号:** 1007-9807(2026)01-0018-21

0 引言

金融经济以及工业过程建模虽然有许多目的, 但是一个非常重要的目的就是希望通过建立合理的模型来改进对未来的预测, 这就需要拟合的模型对未来(真正的样本外数据)具有较好的稳健性. 然而, 在实际许多建模中都非常强调对已有数据的拟合效果, 但是一般都缺乏对未来无限可能性的考虑. 样本数据中一般包括有价值的信息和噪声, 噪声的存在是导致历史拟合最优与未来预测最优之间不一致的主要原因. 如果因变量本身的噪声成分很大, 那么通过最小化均方误差得到的拟合最优模型, 很可能“过度拟合”了因变量中的噪声部分. 历史样本数据中总会存在噪声, 这些样本数据的拟合最优模型常常会“过度利用”数据而很难是预测最优模型^[1-4].

有效而又不过度利用历史数据, 实现去伪

(噪声)存真, 对改善预测效果至关重要. 在传统的线性回归建模中, 经常通过判断系数是否显著来取舍变量, 这种0或1的方法是否合理? 是否存在更好的方法? 如何在充分利用历史数据的同时, 又尽可能避免过度利用数据, 是线性回归预测建模中非常重要的问题.

本研究认为未来发生的实际数据与历史样本数据并不会完全一致, 故需要对历史拟合最优模型进行合理调整. 在充分利用过去数据但又不过度利用的情况下形成期望最优, 这样尽可能保持拟合与预测模型之间的不变性, 把拟合与预测直接联系起来, 实现更好地预测未来. 因此, 本文提出了一种基于数据中的信息噪声比(简称信噪比)来调整样本内估计参数的L乘子模型. L乘子是一个连接拟合模型与预测模型的重要参数. 本研究主要包括以下几个方面:

首先, 在经典多元线性回归模型的框架下, 根

① 收稿日期: 2023-10-07; 修订日期: 2024-10-13.

基金项目: 国家自然科学基金资助项目(72342023; 72371158).

通讯作者: 吴冲锋(1962—), 男, 浙江温州人, 博士, 教授, 博士生导师. Email: cfwu@sjtu.edu.cn

据未来期望平方误差最小化情况下推导出最优 L 乘子的理论表达式和显示解. 当假设多个变量之间相互独立时, L 乘子可进一步简化为具有明确经济含义的形式, 它们取决于信噪比. 接着, 在单变量回归模型的框架下, 考虑变量自相关性的影响, 利用 Amihud 和 Hurvich^[5]的方法, 对回归系数的估计偏差进行纠正, 推导出 L 乘子改进公式. 然后, 给出 L 乘子的两种估计方法并进行比较. 最后, 把 L 乘子模型应用于股票收益率预测的实证研究. 在实证中, 本文利用一系列常用经济变量对股票指数的月度收益率进行预测. 为了全面评估性能, 本文使用统计和经济两方面的指标. 此外, 考虑多变量的情形, 与 LASSO、岭回归这两种现有的收缩方法进行简单比较, 并考虑 L 乘子在 PCA 和 PLS 回归模型中的应用. 实证检验表明相较于基准的线性回归模型, 经 L 乘子修正后的模型具有更高的预测精度, 并且非常显著地改进预测效果.

本文的研究贡献体现在以下几个方面: 首先, 建立了 L 乘子模型, 并从理论上推导出最优 L 乘子的期望表达式以及最优改进性能水平; 其次, 给出两种估计最优 L 乘子的方法, 为实际应用提供可行办法. 最后, 开展广泛且深入的实证研究, 发现理论与实证结果之间存在高度的一致性.

1 文献综述

1.1 Stein 类型收缩技术

在文献中, 减缓由数据噪声引起的过拟合问题一直广受关注. 一种有效的方法是使用收缩 (shrinkage) 技术, 这个方法有着悠久的历史, 最早由 James 和 Stein^[6]进行了初步研究. 该研究表明, 当同时估计多个 (至少 3 个) 随机变量的均值时, 通常的估计量 (如极大似然估计量) 是不可接受的 (inadmissible)^②. 这项早期工作认识到, 包含多个估计量的联合损失函数在各个部分之间建立了联系, 允许在不同估计误差之间进行一定的平衡和补偿. 基于这个思想的收缩方法在不同领域都存在广泛应用, 包括统计学^[1, 7]; 股票收益率预

测^[2-4]; 以及投资组合管理^[8-13].

本研究与以上文献在几个方面存在差异: 第一, 尽管本文的 L 乘子模型具备收缩功能, 但它不同于 Stein 类型的收缩思想. Stein 类型的收缩方法主要侧重于通过同时估计多个参数来提高效率. 相比之下, 本文提出 L 乘子模型的动机在于: 样本数据中噪声的存在导致历史拟合最优与未来预测期望最优之间不一致, 故需要基于数据中的信息噪声比来调整样本内估计参数, 以实现更准确的未来预测. 因此, 即使对于单一参数的估计问题, 估计风险仍然存在, 需要进行合理的修正; 第二, 在多元回归模型的框架下, 本文对每个回归系数 (包括截距项) 都应用不同的 L 乘子进行线性修正. 然而, 需要指出的是, 无论是 Kan 和 Zhou^[10]的投资组合分析还是 Copas^[11]的多元回归分析, 它们都是对所有的估计变量 (均值或回归系数) 进行相同强度的收缩; 第三, 本研究的方法与 Connor^[2]的方法在计算收缩强度方面存在差异. 本文没有采用贝叶斯方法来解决参数的估计风险问题, 而是运用最小化均方预测误差 (MSFE) 准则来确定收缩强度. 这一替代方法在投资组合分析中得到广泛应用.

1.2 岭回归和 LASSO

除了 Stein 类型的收缩方法, 还有另一类收缩技术, 包括岭回归^[14] (Ridge) 和最小绝对收缩及选择算子^[15] (LASSO) 等. 这些方法在最小化均方误差时分别施加 1 范数和 2 范数的惩罚项来估计模型参数, 以实现参数的收缩, 而不是直接对 OLS 估计量进行收缩. 在机器学习领域, 岭回归和 LASSO 作为常用的正则化技术, 近年来已受到国内外众多学者的关注, 并被广泛应用于金融领域的多样化问题研究中. 在国际学术界, Rapach 等^[16]、Gu 等^[17]、Chinco 等^[18]、Freyberger 等^[19]、Kozak 等^[20]、Dong 等^[21]、Mazumder 等^[22]和 Kelly 等^[23]对这些方法在金融领域的应用做出了贡献. 在国内市场, 陈艳和王宣承^[24]、蒋翠侠等^[25]、王小燕等^[26]、李斌等^[27]、姜富伟等^[28]、梁超等^[29]以及郑挺国等^[30]也通过他们的研究, 展示了这些收

② 根据 James 和 Stein^[6]的定义, 一个估计量被称为“可接受的” (admissible), 是指如果不存在其他估计量, 其风险在所有参数点上都更小, 且在某些点上严格更小. 相反, 如果存在这样的估计量, 该估计量被认为是“不可接受的”.

缩技术在金融问题解决中的有效性和适用性。

本研究也可以与岭回归和 LASSO 方法建立一定的联系,并为这些方法中的超参数选择问题提供有益的参考。

2 理论模型

2.1 噪声修正的多元线性回归模型

2.1.1 经典多元线性回归模型

令 $\mathbf{X}_{t-1} = (1, X_{1,t-1}, X_{2,t-1}, \dots, X_{k,t-1})'$ 是 $(k + 1) \times 1$ 维向量的第 $t - 1$ 个观测. 用 $K = k + 1$ 表示回归变量的个数, 包括一个常数项和 k 个解释变量. 假设因变量的第 t 个观测 Y_t 的数据生成过程由下列多元线性回归模型给出

$$Y_t = \mathbf{X}'_{t-1} \boldsymbol{\beta} + \varepsilon_t, t = 1, \dots, T \quad (1)$$

其中 $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_k)'$ 是一个 $K \times 1$ 维的未知参数向量. ε_t 是不可观测的扰动项, 满足均值为 0, 方差为 σ_ε^2 的正态分布. T 为样本观测总数. 如果定义 $\mathbf{Y} = (Y_1, \dots, Y_T)'$, $\mathbf{X} = (\mathbf{X}_0, \dots, \mathbf{X}_{T-1})'$ 以及 $\boldsymbol{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_T)'$, 式(1)具有下列矩阵表示

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (2)$$

其中 \mathbf{Y} 的维度为 $T \times 1$, \mathbf{X} 的维度为 $T \times K$, $\boldsymbol{\beta}$ 的维度为 $K \times 1$, $\boldsymbol{\varepsilon}$ 的维度为 $T \times 1$. 为简化分析, 对每个解释变量进行去均值处理, 以确保它们的样本均值为 0. 此外, 假设 $\mathbf{V} = \frac{1}{T} \mathbf{X}'\mathbf{X}$ 为 $K \times K$ 维的满秩矩阵. 因此, 对于每个观测 $t = 1, \dots, T$, \mathbf{X}_{t-1} (除常数项之外) 可以看作是从均值为 0、方差 - 协方差矩阵为 \mathbf{V} 的多元概率分布中生成的. 假设, \mathbf{X}_{t-1} 的总体均值和方差 - 协方差矩阵等于样本均值和方差 - 协方差矩阵. 因此, 在后续分析中, 样本内和样本外的解释变量来自相同的概率分布.

在经典线性回归模型的假设下, 系数 $\boldsymbol{\beta}$ 所满足的最小二乘(OLS) 估计量为

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{Y} = \left(\sum_{t=1}^T \mathbf{X}_{t-1} \mathbf{X}'_{t-1} \right)^{-1} \times \sum_{t=1}^T \mathbf{X}_{t-1} Y_t \quad (3)$$

于是, 给定预测变量在样本外的观测 \mathbf{X}_T 时, 因变量的第 $T + 1$ 个观测 Y_{T+1} 可预测为

$$\hat{Y}_{T+1} = \mathbf{X}'_T \hat{\boldsymbol{\beta}} \quad (4)$$

在现实世界中, 如果因变量 Y 的噪声较高, 具有较低的信噪比时, 就很难预测. 为此, 本文提出了一种基于信息噪声比修正的 L 乘子模型, 以考虑这种有限样本和数据噪声带来的预测偏差问题, 从而提高模型在样本外的预测精度.

2.1.2 L 乘子模型

在式(4) 中 \hat{Y}_{T+1} 的预测偏差主要源于回归系数 $\hat{\boldsymbol{\beta}}$ 的估计误差及其随机噪声影响. 假设可以采用噪声修正技术来对 OLS 估计量 $\hat{\boldsymbol{\beta}}$ 中的估计误差进行修正, 从而改善 \hat{Y}_{T+1} 的预测性能. 考虑一种最简单的线性去噪技术 $\mathbf{L} = (L_0, L_1, \dots, L_k)'$, 并定义经 L 修正后的预测值为 $\tilde{Y}_{T+1}(\mathbf{L}) = \sum_{i=0}^k L_i \hat{\beta}_i X_{i,T}$, 这个经 L 乘子修正的模型称为 L 乘子模型. 于是, 相对于实际值 Y_{T+1} 的均方预测误差 (MSFE) 可以计算为

$$\begin{aligned} MSFE(\mathbf{L}) &= \mathbb{E} [Y_{T+1} - \tilde{Y}_{T+1}(\mathbf{L})]^2 \\ &= \mathbb{E} \left[Y_{T+1} - \sum_{i=0}^k L_i \hat{\beta}_i X_{i,T} \right]^2 \end{aligned}$$

其中 $X_{0,T} = 1$. 如果令 $\mathbf{Z}_T = (Z_{0,T}, \dots, Z_{k,T})' = (\hat{\beta}_0 X_{0,T}, \dots, \hat{\beta}_k X_{k,T})'$, 则上式可以重新表示为

$$\begin{aligned} MSFE(\mathbf{L}) &= \mathbb{E} [Y_{T+1} - \mathbf{L}'\mathbf{Z}_T]^2 \\ &\text{可以通过最小化 } MSFE(\mathbf{L}) \text{ 来寻找最优的 } \mathbf{L} \text{ 乘子, 即} \\ \mathbf{L}^* &= \underset{\mathbf{L} \in \mathbb{R}^{k+1}}{\operatorname{argmin}} MSFE(\mathbf{L}) \\ &= \underset{\mathbf{L} \in \mathbb{R}^{k+1}}{\operatorname{argmin}} \mathbb{E} [Y_{T+1} - \mathbf{L}'\mathbf{Z}_T]^2 \end{aligned} \quad (5)$$

其中 \mathbb{R} 是所有实数空间. 简单推导可得, 最优 \mathbf{L} 乘子的解析表达式为

$$\mathbf{L}^* = \mathbb{E} [\mathbf{Z}_T \mathbf{Z}'_T]^{-1} \mathbb{E} [\mathbf{Z}_T Y_{T+1}] \quad (6)$$

用 \mathbf{L}^* 来预测的模型称为最优 L 乘子模型, 也可以统称 L 乘子模型.

如果式(5) 关于 \mathbf{L} 的二阶条件 $\mathbb{E} [\mathbf{Z}_T \mathbf{Z}'_T]$ 为正定矩阵, 那么 \mathbf{L}^* 为式(5) 的最小值. \mathbf{L}^* 的两个部分 $\mathbb{E} [\mathbf{Z}_T \mathbf{Z}'_T]$ 和 $\mathbb{E} [\mathbf{Z}_T Y_{T+1}]$ 可以分别计算为

$$\begin{aligned} \mathbb{E} [\mathbf{Z}_T \mathbf{Z}'_T] &= \begin{pmatrix} \mathbb{E} [\hat{\beta}_0^2] & \mathbf{0} \\ \mathbf{0} & \mathbf{W} \end{pmatrix} \\ \mathbb{E} [\mathbf{Z}_T Y_{T+1}] &= (\beta_0^2, \mathbf{B})' \end{aligned}$$

其中

$$\mathbf{W} = \begin{pmatrix} \mathbb{E} [\hat{\beta}_1^2] \sigma_{1,X}^2 & \dots & \mathbb{E} [\hat{\beta}_1 \hat{\beta}_k] V_{1,k} \\ \vdots & \ddots & \vdots \\ \mathbb{E} [\hat{\beta}_k \hat{\beta}_1] V_{k,1} & \dots & \mathbb{E} [\hat{\beta}_k^2] \sigma_{k,X}^2 \end{pmatrix}$$

$$B = \left(\sum_{i=1}^k \beta_i \beta_i V_{1,i}, \dots, \sum_{i=1}^k \beta_i \beta_i V_{k,i} \right)'$$

$V_{i,j} = E[X_{i,T} X_{j,T}]$ 是预测变量方差 - 协方差矩阵 V 中第 i 行第 j 列的元素, $\sigma_{i,x}^2 = V_{i,i}$ 为预测变量 $X_{i,T}$ 的方差. $E[X_T X_T'] = \frac{1}{T} \sum_{t=1}^T X_{t-1} X_{t-1}' = V$.

于是,最优 L 乘子可以表示为

$$L^* = \begin{pmatrix} \mathbb{E}[\hat{\beta}_0^2] & \mathbf{0} \\ \mathbf{0} & W \end{pmatrix}^{-1} \begin{pmatrix} \beta_0^2 \\ B \end{pmatrix}$$

$$= \begin{pmatrix} 1 & \mathbf{0} \\ \mathbb{E}[\hat{\beta}_0^2] & W^{-1} \end{pmatrix} \begin{pmatrix} \beta_0^2 \\ B \end{pmatrix} = \begin{pmatrix} \beta_0^2 \\ \mathbb{E}[\hat{\beta}_0^2] B \end{pmatrix} \begin{pmatrix} 1 \\ W^{-1} \end{pmatrix} \quad (7)$$

可以看出,对于截距项回归系数而言,其 L 乘子具有下列简约形式

$$L_0^* = \frac{\beta_0^2}{\mathbb{E}[\hat{\beta}_0^2]} = \frac{\beta_0^2}{\beta_0^2 + \text{Var}[\hat{\beta}_0]} = \frac{R_{\beta_0}^2}{1 + R_{\beta_0}^2} \quad (8)$$

其中 $R_{\beta_0}^2 = \frac{\beta_0^2}{\text{Var}[\hat{\beta}_0]}$, 称 R_{β_0} 为信噪比.

特殊情形: 当变量之间相互独立.

$$L_i^* = \frac{\beta_i^2}{\mathbb{E}[\hat{\beta}_i^2]} = \frac{\beta_i^2}{\beta_i^2 + \text{Var}[\hat{\beta}_i]} \quad (9)$$

$$= \frac{R_{\beta_i}^2}{1 + R_{\beta_i}^2} \quad (\text{其中 } i = 1, \dots, k)$$

类似于 R_{β_0} , R_{β_i} 同样称为信噪比, $L_i^* \leq 1$. 在经典线性回归模型的假设下, $\mathbb{E}(\hat{\beta}) = \beta$, $\text{Var}(\hat{\beta}) = \sigma_\varepsilon^2 (X'X)^{-1}$. 将式(8)和式(9)结合,可得对于每个 $i = 1, \dots, k$, 最优 L 乘子满足

$$L_i^* = \frac{\beta_i^2}{\mathbb{E}[\hat{\beta}_i^2]} = \frac{\beta_i^2 \sigma_{i,x}^2}{\beta_i^2 \sigma_{i,x}^2 + \frac{\sigma_\varepsilon^2}{T}} \quad (10)$$

因为 $(L_0^*, L_1^*, \dots, L_k^*)'$ 使得式(5)中的均方预测误差 $MSFE(L)$ 达到最小,所以

$$MSFE(L_0^*, L_1^*, \dots, L_k^*) \leq MSFE(1, 1, \dots, 1) \quad (11)$$

成立.

以单变量情形为例,可以计算出 L 乘子模型相对于原始模型的 MSFE 理论改进值如下

$$MSFE(1,1) - MSFE(L_0^*, L_1^*) = \left(\beta_0^2 + \frac{\sigma_\varepsilon^2}{T} \right) \times$$

$$(1 - L_0^*)^2 + \left(\beta_1^2 \sigma_{1,x}^2 + \frac{\sigma_\varepsilon^2}{T} \right) (1 - L_1^*)^2 \geq 0$$

2.2 单变量情形以及变量自相关性调整

本文的实证分析中,将重点关注单一变量情形下的 L 乘子模型对于预测性能的改进效果. 选择关注单变量情况的原因有几点: 首先,多变量非独立情形下的 L 乘子表达式相对复杂,缺乏明确的经济解释. 其次,为了综合多个变量中的有价值信息,可以先进行单变量预测,再考虑如何组合这些单变量预测的结果.

假设一元线性回归模型为

$$Y_t = \alpha + \beta X_{t-1} + \varepsilon_t = 1, \dots, T \quad (12)$$

由上一节的分析可知, α 和 β 的 OLS 估计量 $\hat{\alpha}$ 和 $\hat{\beta}$ 的最优 L 乘子分别为

$$L_\alpha^* = \frac{\alpha^2}{\mathbb{E}[\hat{\alpha}^2]} = \frac{\alpha^2}{\alpha^2 + \frac{\sigma_\varepsilon^2}{T}} \quad (13)$$

$$L_\beta^* = \frac{\beta^2}{\mathbb{E}[\hat{\beta}^2]} = \frac{\beta^2}{\beta^2 + \frac{\sigma_\varepsilon^2}{T\sigma_x^2}} \quad (14)$$

当预测变量存在高度自相关性时,经典线性回归模型的假设就不再成立,导致 OLS 估计量有偏,同时标准误也存在偏差. 鉴于此,Stambaugh^[31] 以及 Amihud 和 Hurvich^[5] 在进行收益率预测的研究中,充分考虑了自变量存在自相关性的情况,并提出了相应的方法来纠正回归系数的估计偏差.

在一元线性模型式(12)的基础上,进一步假设预测变量满足下列一阶自回归过程

$$X_t = \theta + \rho X_{t-1} + u_t, \quad t = 1, \dots, T \quad (15)$$

其中

$$\text{Cov} \left(\begin{bmatrix} \varepsilon_t \\ u_t \end{bmatrix}, \begin{bmatrix} \varepsilon_t & u_t \end{bmatrix} \right) = \begin{pmatrix} \sigma_\varepsilon^2 & \sigma_{\varepsilon u} \\ \sigma_{\varepsilon u} & \sigma_u^2 \end{pmatrix}$$

由 Stambaugh^[31] 的研究可知,预测变量的自相关性问题不会影响截距项估计的无偏性,即 $\mathbb{E}[\hat{\alpha}] = \alpha$. 但 β 的 OLS 估计量 $\hat{\beta}$ 是有偏的,满足

$$m_1' = \mathbb{E}[\hat{\beta} - \beta]$$

$$m_2' = \mathbb{E}[(\hat{\beta} - \beta)^2]$$

其中 m_1' 和 m_2' 如 Stambaugh^[31] 中性质 3 (Proposition 3) 所定义. 此时,OLS 估计量 $\hat{\beta}$ 的方差为

$$\text{Var}[\hat{\beta}] = m_2' - (m_1')^2$$

因此,式(14)中的 L 乘子调整为

$$L_\beta^c = \frac{\beta^2 + m_1' \beta}{\beta^2 + 2m_1' \beta + m_2'} \quad (16)$$

2.3 L 乘子的实证估计

单变量情形下的最优 L 乘子表达式 (13)、式 (14) 和式 (16) 都是理论值. 尽管理论分析表明, L 乘子通过修正“噪声”可以提高经典线性回归模型在样本外的预测性能, 但在实际应用中, L 乘子本身需要使用有限样本进行估计, 因此存在估计误差的问题. 在极端情形下, L 乘子修正模型的实证表现可能会不如基准模型. 理想情况下, 式 (13)、式 (14) 和式 (16) 的整体无偏估计量才是 L 乘子的最佳估计. 然而, 即使对于式 (13) 和式 (14), 获得它们的无偏估计量也是相当具有挑战性的. 因此, 考虑以下两种简单的近似估计 L 乘子的方法.

第一种方法是直接将单变量回归模型参数的 OLS 估计量代入式 (13) 和式 (14), 则有

$$\hat{L}_1^\alpha = \frac{\hat{\alpha}^2}{\hat{\alpha}^2 + \widehat{Var}[\hat{\alpha}]} \quad (17)$$

$$\hat{L}_1^\beta = \frac{\hat{\beta}^2}{\hat{\beta}^2 + \widehat{Var}[\hat{\beta}]} \quad (18)$$

其中 $\hat{\alpha}$ 和 $\hat{\beta}$ 分别为截距项和预测变量系数的 OLS 估计量, $\widehat{Var}[\hat{\alpha}]$ 和 $\widehat{Var}[\hat{\beta}]$ 分别为系数估计值方差 $Var[\hat{\alpha}]$ 和 $Var[\hat{\beta}]$ 的无偏估计值.

由于预测变量进行了去均值处理, 所以 $\hat{\alpha} = \bar{Y}$. 给定 X , $\hat{\alpha}$ 的条件方差可估计为, $\widehat{Var}[\hat{\alpha} | X] = \widehat{Var}[\bar{Y} | X] = \frac{\hat{\sigma}_\varepsilon^2}{T}$. 当然, 如果估计 $\hat{\alpha}$ 的无条件方差, 则有 $\widehat{Var}[\hat{\alpha}] = \widehat{Var}[\bar{Y}] = \widehat{Var}\left[\frac{1}{T} \sum_{i=1}^T Y_i\right] = \frac{\widehat{Var}[Y_i]}{T} = \frac{\hat{\beta}^2 \sigma_X^2 + \hat{\sigma}_\varepsilon^2}{T} \neq \frac{\hat{\sigma}_\varepsilon^2}{T}$. 在之后的实证分析中, 研究发现在计算 L 乘子时, 无论是用条件方差和还是无条件方差来估计 $Var[\hat{\alpha}]$, 都不影响本文的研究结论, 特别是样本窗口足够长时, 两者几乎没有差别.

第二种方法是将分子和分母的整体无偏估计量代入. 因为 $E[\hat{\alpha}^2] = \alpha^2 + Var[\hat{\alpha}]$, $E[\hat{\beta}^2] = \beta^2 + Var[\hat{\beta}]$, 所以 α^2 和 β^2 的无偏估计量分别为 $\hat{\alpha}^2 - \widehat{Var}[\hat{\alpha}]$ 和 $\hat{\beta}^2 - \widehat{Var}[\hat{\beta}]$. 代入式 (13) 和式 (14) 可得

$$\hat{L}_2^\alpha = \frac{\hat{\alpha}^2 - \widehat{Var}[\hat{\alpha}]}{\hat{\alpha}^2} \quad (19)$$

$$\hat{L}_2^\beta = \frac{\hat{\beta}^2 - \widehat{Var}[\hat{\beta}]}{\hat{\beta}^2} \quad (20)$$

比较两种 L 乘子估计值, 可以发现, 由于 $\widehat{Var}[\hat{\alpha}] > 0$ 始终成立, 所以满足

$$\hat{L}_1^\alpha = \frac{\hat{\alpha}^2}{\hat{\alpha}^2 + \widehat{Var}[\hat{\alpha}]} > 1 - \frac{\widehat{Var}[\hat{\alpha}]}{\hat{\alpha}^2} = \hat{L}_2^\alpha$$

同理可得, $\hat{L}_1^\beta > \hat{L}_2^\beta$. 这说明, 尽管两种方法都可以近似估计 L 乘子, 但第二种方法得到的 L 乘子估计值更小, 修正力度更强.

从式 (19) 和式 (20) 可以看出, \hat{L}_2^α 和 \hat{L}_2^β 可能会出现负值的情形. 然而, L 乘子的理论值始终介于 0 到 1 之间. 为此, 需要对第二种方法下的 L 乘子估计值在 0 处进行截断处理^③. 则式 (19) 和式 (20) 可以修改为

$$\hat{L}_2^\alpha = \max\left\{\frac{\hat{\alpha}^2 - \widehat{Var}[\hat{\alpha}]}{\hat{\alpha}^2}, 0\right\} \quad (21)$$

$$\hat{L}_2^\beta = \max\left\{\frac{\hat{\beta}^2 - \widehat{Var}[\hat{\beta}]}{\hat{\beta}^2}, 0\right\} \quad (22)$$

当考虑预测变量的自相关性时, 依照 Stambaugh^[31]的方法推导自相关性调整后的 β 估计值, 记为 $\hat{\beta}^c$, 以及调整后的方差估计值 $\widehat{Var}^c(\hat{\beta}^c)$. 可用 Amihud 和 Hurvich^[5]的方法来计算 $\hat{\beta}^c$ 和 $\widehat{Var}^c(\hat{\beta}^c)$.

3 实证分析

第 2 章的理论推导已经表明, 通过应用 L 乘子对 OLS 估计量进行修正, 可以显著提高 OLS 模型在样本外的预测性能. 下面以股票收益率的时间序列预测对该理论进行实证研究. 股票收益率预测一直以来都备受关注, 已在众多文献中进行了研究^[28, 32-43]. 然而, 股票收益率序列本身的信噪比很低, 很难找到能够持续预测收益率的有效预测.

3.1 数据样本及描述性统计

在实证分析中, 本文使用了一系列经济变量来预测上证指数的月度收益率. 除了采用 Welch

③ 这种截断处理与文献中对 Stein 类型收缩估计量的修正方法相似, 即所谓的“positive part”估计, 例如 Copas^[1]对此有论述.

和 Goyal^[44]研究中提到的 12 个变量之外,还引入了额外的 5 个变量,分别是现金流价格比(c/p)、销售收入价格比(s/p)、营业收入($op\ rev$)、换手率($turnover$)和百度搜索指数($baidu$)^④.共使用了

17 个变量来进行上证指数超额收益率的预测.这些变量的详细说明参见表 1.本文用于构建这些预测变量的数据主要来源于数据库 CSMAR 和数据库 Wind.

表 1 变量说明

Table 1 Variable descriptions

变量名	变量说明
$premium$	上证指数超额收益率,股指对数收益率减去定期存款利率
d/p	股息率,股息除以价格,取对数
d/y	股息收益率,股息除以滞后价格,取对数
e/p	盈利价格比,盈利除以价格,取对数
d/e	股息支付率,股息除以盈利,取对数
c/p	现金流价格比,现金流除以价格,取对数
s/p	销售收入价格比,销售收入除以价格,取对数
b/m	账面市值比,净资产除以价格,取对数
$op\ rev$	营业收入,取对数
$svar$	股票方差,日收益率的平方,在一个月內求和
$turnover$	换手率
tbl	短期国债利率,3 个月期国债到期收益率
lty	长期国债利率,10 年期国债到期收益率
tms	期限利差,10 年期国债与 3 个月期国债到期收益率之差
dfy	违约利差,1 年期 AA 级企业债与 1 年期 AAA 级企业债到期收益率之差
dfr	信用利差,10 年期 AAA 级企业债与 10 年期国债到期收益率之差
$infl$	通货膨胀率
$baidu$	百度搜索指数:将个股名称在百度搜索引擎上出现的次数按照市值进行加权得到市场的搜索指数,取对数

表 2 报告了上证指数超额收益率和预测变量的月度样本时间跨度,以及它们各自的描述性统计结果.可以看出,从 1990 年 12 月 ~ 2023 年 8 月,上证指数的平均超额收益率为 0.5%,标准差为 11.6%,夏普比率为 0.04,这意味着上证指数收益率具有特别低的信噪比.另外,股票超额收益率的一阶自相关系数为 -0.031 ,这表明股票收益率很难被其过去的收益率解释,也难以预测.

相比之下,大多数预测变量具有较高的持续性,17 个预测变量中有 13 个的一阶自相关系数超过了 0.90.根据 Stambaugh^[31]的观点,预测变量的高度持续性会对线性回归估计产生偏差.因此,在后续的样本内分析中,采用 Amihud 和 Hurvich^[5]的方法来校正 OLS 估计量.在样本外分析中,本文主要报告未考虑预测变量高度自相关性的结果,考虑变量自相关性并不影响本文的主要研究结论.

表 2 变量描述性统计

Table 2 Descriptive statistics of variables

变量名	样本数量	样本开始时间	样本结束时间	均值	标准差	中位数	偏度	峰度	一阶自相关系数
$premium$	393	1990-12	2023-08	0.005	0.116	0.003	2.531	25.187	-0.031
d/p	378	1992-03	2023-08	-4.534	0.999	-4.185	-1.730	6.039	0.983
d/y	378	1992-03	2023-08	-4.532	0.998	-4.184	-1.733	6.065	0.983
e/p	386	1991-07	2023-08	-3.254	0.763	-3.074	-1.124	4.507	0.984
d/e	378	1992-03	2023-08	-1.332	0.529	-1.244	-2.678	11.350	0.968
c/p	296	1999-01	2023-08	-2.307	0.683	-2.166	-1.292	4.353	0.978

④ 随着互联网信息技术的快速进步,像百度等网络搜索引擎的应用与发展为投资者提供了一种高效快捷的获取信息的途径.通过这些网络搜索引擎,投资者可以实时获取与股票相关的搜索词条数据,这些数据有望更加及时地反映投资者情绪和关注度的变化,因此可能对股票收益率产生一定的预测作用^[45-47].

续表 2
Table 2 Continues

变量名	样本数量	样本开始时间	样本结束时间	均值	标准差	中位数	偏度	峰度	一阶自相关系数
<i>s/p</i>	386	1991-07	2023-08	-0.733	0.689	-0.569	-0.899	3.439	0.980
<i>b/m</i>	386	1991-07	2023-08	-0.975	0.589	-0.890	-0.882	3.663	0.974
<i>op rev</i>	381	1991-12	2023-08	28.503	2.849	29.272	-1.272	4.441	0.998
<i>svar</i>	393	1990-12	2023-08	0.011	0.061	0.003	16.928	314.297	0.042
<i>turnover</i>	386	1991-08	2023-08	0.436	0.792	0.206	5.210	34.183	0.803
<i>tbl</i>	260	2002-01	2023-08	0.023	0.007	0.022	0.268	2.814	0.913
<i>lty</i>	260	2002-01	2023-08	0.034	0.006	0.033	0.779	3.272	0.944
<i>tms</i>	260	2002-01	2023-08	0.011	0.006	0.010	0.550	3.029	0.836
<i>dfy</i>	196	2007-05	2023-08	0.006	0.003	0.005	1.209	5.704	0.902
<i>dfr</i>	210	2006-03	2023-08	0.012	0.003	0.012	0.019	2.538	0.921
<i>infl_lag</i>	392	1990-12	2023-07	0.040	0.057	0.021	2.311	8.419	0.989
<i>baidu</i>	120	2011-01	2020-12	5.962	0.333	5.932	0.897	4.369	0.885

3.2 样本内回归

本文主要考虑单变量的预测回归,这与许多已有的研究一致,这些研究侧重于包含单一预测变量的回归模型,例如 Campbell 和 Thompson^[48]、Pettenuzzo 等^[49]以及 Welch 和 Goyal^[44]。

表 3 报告了单变量预测回归的全样本(以预测变量的样本长度为准)估计结果。本文给出了回归系数的 OLS 估计量($\hat{\alpha}$ 、 $\hat{\beta}$)、*t* 统计量($t_{\hat{\alpha}}$ 、 $t_{\hat{\beta}}$)、 R^2 统计量(R^2)以及利用 Amihud 和 Hurvich^[5]的方法进行修正后的 OLS 估计量($\hat{\beta}^c$)和相应的 *t* 统计量($t_{\hat{\beta}^c}$)。所有预测变量都进行了标准化处理(均值为 0,标准差为 1),以确保不同变量的回归系数具有可比性。

根据预测变量回归系数的 OLS 估计量的 *t* 值

($t_{\hat{\beta}}$),在本文研究的全样本区间内,大部分变量在预测股票超额收益率方面都不具有统计显著性。在样本内表现中,最好的是长期国债利率(*lty*),其在 1% 的水平上呈现出显著性。根据 Campbell 和 Thompson^[48]的观点,月度回归的 R^2 统计量达到 0.5% 就足以表示变量的预测能力具有经济意义。本文有 7 个预测变量的样本内 R^2 统计量超过了这个水平。

在对变量的自相关性进行修正后,很多变量回归系数和显著性水平都发生了显著变化。这强调在进行统计推断时,对预测变量的高度自相关性进行修正是必要的。另外,对于自相关性最低的变量,即股票方差(*svar*),其自相关系数为 0.042(见表 2),修正自相关性后的回归系数以及 *t* 统计量变化不大。

表 3 样本内预测回归

Table 3 In-sample forecast regressions

变量名	样本数量	$\hat{\alpha}$	$t_{\hat{\alpha}}$	$\hat{\beta}$	$t_{\hat{\beta}}$	$\hat{\beta}^c$	$t_{\hat{\beta}^c}$	R^2 (%)
<i>d/p</i>	377	0.003	0.444	0.006	0.993	0.001	0.290	0.263
<i>d/y</i>	377	0.003	0.444	0.006	1.017	0.001	0.337	0.275
<i>e/p</i>	385	0.005	0.839	-0.009	-1.486	-0.014***	-4.318	0.573
<i>d/e</i>	377	0.003	0.447	0.013**	2.249	0.013**	2.161	1.331
<i>c/p</i>	295	0.002	0.370	0.002	0.430	0.000	-0.002	0.063
<i>s/p</i>	385	0.005	0.840	-0.001	-0.170	-0.006**	-2.077	0.008
<i>b/m</i>	385	0.005	0.840	0.000	-0.053	-0.005*	-1.930	0.001
<i>op rev</i>	380	0.003	0.546	-0.008	-1.406	-0.010*	-1.732	0.521
<i>svar</i>	392	0.005	0.869	-0.003	-0.508	-0.003	-0.551	0.066
<i>turnover</i>	384	0.004	0.750	0.008	1.299	0.008	1.584	0.439
<i>tbl</i>	259	0.001	0.234	-0.010**	-2.192	-0.010**	-2.251	1.835
<i>lty</i>	259	0.001	0.242	-0.012**	-2.672	-0.012**	-2.580	2.702
<i>tms</i>	259	0.001	0.223	0.000	0.103	0.001	0.176	0.004
<i>dfy</i>	195	-0.003	-0.617	0.001	0.211	0.001	0.195	0.023
<i>dfr</i>	209	0.002	0.467	-0.013**	-2.379	-0.013**	-2.418	2.660
<i>infl</i>	391	0.005	0.895	-0.005	-0.889	-0.004	-0.673	0.203
<i>baidu</i>	119	0.000	0.016	-0.009	-1.597	-0.007	-1.413	2.134

注: *, ** 和 *** 分别表示在 10%, 5% 和 1% 的水平下显著。

3.3 样本外预测的统计表现

大量研究^[37, 44]表明,良好的样本内表现并不能保证模型在样本外表现出色.实际上,市场参与者更关心模型在样本外的预测表现.因此,在这一部分,本文将研究这些预测变量是否能够用于在样本外预测收益率,并探讨通过 L 乘子修正噪声是否能够改善原始模型的预测性能.主要通过统计表现来评估模型的预测能力,而在 3.4 节将进一步分析均值 - 方差效用的投资者所获得的经济效益,以评估模型的样本外预测能力.

3.3.1 样本外预测模型

首先,遵循现有文献的常规做法^[50-52],使用以下单变量预测回归模型

$$Y_t = \alpha_i + \beta_i X_{i,t-1} + \varepsilon_t$$

其中 Y_t 表示 t 月上证指数的超额收益率, $X_{i,t-1}$ 为第 i 个预测变量在 $t-1$ 月的观察值, ε_t 是扰动项.与 Welch 和 Goyal^[42]不同,本研究采用了滚动窗口而非扩展窗口来进行股票收益率的样本外预测.这种滚动预测的方法在许多研究中被用来捕捉系数的时变性,以考虑参数的不稳定性,如 Dangi 和 Halling^[53]以及 Wang 等^[54].此外,本文构建的 L 乘子模型旨在修正模型估计中的噪声影响,而估计窗口的长度则直接决定了可用的有价值信息的量.因此,为了观察 L 乘子修正效果随着估计窗口长度的变化,在不同的估计窗口(24 个月到 60 个月)下进行滚动预测.

表 4 L 乘子的描述性统计

Table 4 Descriptive statistics of L-multipliers

	截距项系数的 L 乘子估计值						预测变量系数的 L 乘子估计值					
	\widehat{L}_1^α		\widehat{L}_2^α		$(\widehat{L}_1^\alpha + \widehat{L}_2^\alpha)/2$		\widehat{L}_1^β		\widehat{L}_2^β		$(\widehat{L}_1^\beta + \widehat{L}_2^\beta)/2$	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
Panel A: 估计窗口 = 24												
<i>d/p</i>	0.35	0.24	0.12	0.24	0.24	0.23	0.47	0.29	0.29	0.33	0.38	0.30
<i>d/y</i>	0.35	0.24	0.12	0.24	0.24	0.23	0.48	0.28	0.29	0.33	0.39	0.30
<i>e/p</i>	0.35	0.24	0.12	0.24	0.24	0.23	0.46	0.28	0.27	0.33	0.37	0.30
<i>d/e</i>	0.35	0.24	0.12	0.24	0.24	0.23	0.38	0.29	0.20	0.30	0.29	0.28
<i>c/p</i>	0.37	0.25	0.14	0.26	0.26	0.24	0.46	0.30	0.30	0.33	0.38	0.30
<i>s/p</i>	0.35	0.24	0.12	0.24	0.24	0.23	0.55	0.28	0.40	0.35	0.48	0.31
<i>b/m</i>	0.35	0.24	0.12	0.24	0.24	0.23	0.56	0.28	0.41	0.33	0.48	0.30
<i>op rev</i>	0.35	0.24	0.12	0.24	0.24	0.23	0.38	0.29	0.21	0.31	0.29	0.29
<i>svar</i>	0.37	0.25	0.14	0.26	0.26	0.24	0.43	0.28	0.24	0.32	0.34	0.29
<i>turnover</i>	0.35	0.25	0.13	0.25	0.24	0.23	0.42	0.26	0.21	0.27	0.31	0.26
<i>tbl</i>	0.36	0.24	0.13	0.24	0.24	0.23	0.30	0.26	0.11	0.23	0.20	0.23
<i>lty</i>	0.35	0.24	0.12	0.24	0.24	0.23	0.38	0.27	0.18	0.28	0.28	0.26
<i>tms</i>	0.36	0.25	0.14	0.26	0.25	0.25	0.40	0.34	0.28	0.37	0.34	0.35
<i>dfy</i>	0.36	0.25	0.14	0.26	0.25	0.25	0.47	0.28	0.27	0.34	0.37	0.30
<i>dftr</i>	0.36	0.25	0.14	0.26	0.25	0.25	0.40	0.33	0.26	0.36	0.33	0.34
<i>infl</i>	0.32	0.22	0.08	0.19	0.20	0.19	0.43	0.29	0.24	0.34	0.33	0.31
<i>baidu</i>	0.32	0.22	0.09	0.19	0.20	0.19	0.45	0.27	0.23	0.31	0.34	0.28
Panel B: 估计窗口 = 60												
<i>d/p</i>	0.26	0.24	0.08	0.19	0.17	0.21	0.54	0.24	0.32	0.32	0.43	0.27
<i>d/y</i>	0.26	0.24	0.08	0.19	0.17	0.21	0.54	0.24	0.32	0.32	0.43	0.27
<i>e/p</i>	0.25	0.24	0.08	0.19	0.17	0.20	0.43	0.30	0.27	0.31	0.35	0.30
<i>d/e</i>	0.26	0.24	0.08	0.19	0.17	0.21	0.48	0.29	0.31	0.32	0.40	0.30
<i>c/p</i>	0.27	0.23	0.08	0.20	0.18	0.20	0.58	0.28	0.42	0.36	0.50	0.32
<i>s/p</i>	0.25	0.24	0.08	0.19	0.17	0.20	0.58	0.30	0.47	0.33	0.53	0.31
<i>b/m</i>	0.25	0.24	0.08	0.19	0.17	0.20	0.61	0.27	0.47	0.33	0.54	0.29
<i>op rev</i>	0.26	0.24	0.08	0.19	0.17	0.20	0.38	0.27	0.18	0.28	0.28	0.27
<i>svar</i>	0.25	0.24	0.08	0.19	0.17	0.20	0.44	0.31	0.28	0.34	0.36	0.32
<i>turnover</i>	0.25	0.24	0.08	0.19	0.17	0.20	0.50	0.29	0.32	0.34	0.41	0.31
<i>tbl</i>	0.25	0.22	0.06	0.17	0.16	0.18	0.67	0.27	0.55	0.36	0.61	0.31
<i>lty</i>	0.25	0.22	0.06	0.17	0.16	0.18	0.57	0.35	0.46	0.41	0.51	0.37
<i>tms</i>	0.25	0.22	0.06	0.17	0.16	0.18	0.60	0.27	0.46	0.35	0.53	0.31
<i>dfy</i>	0.19	0.19	0.03	0.12	0.11	0.14	0.45	0.29	0.29	0.32	0.37	0.30
<i>dftr</i>	0.19	0.19	0.03	0.11	0.11	0.14	0.30	0.27	0.13	0.28	0.21	0.27
<i>infl</i>	0.26	0.24	0.08	0.19	0.17	0.20	0.47	0.32	0.31	0.38	0.39	0.34
<i>baidu</i>	0.15	0.14	0.01	0.05	0.08	0.09	0.65	0.14	0.44	0.28	0.55	0.21

注: α 差异不大,因此 L^α 差异也不大.

本文将样本分为样本内(前 T 个观测)和样本外(后 q 个观测)两部分. 首先,用 $\{Y_i\}_{i=1}^T$ 对一个常数项和 $\{X_{i,t-1}\}_{i=1}^T$ 进行回归,从而得到 α_i 和 β_i 的 OLS 估计量,分别为 $\hat{\alpha}_{i,T}$ 和 $\hat{\beta}_{i,T}$. 接着,在给定了第 T 个月预测变量 i 的观察值 $X_{i,T}$ 后,将 $T+1$ 月的股票收益率预测为

$$\hat{Y}_{i,T+1} = \hat{\alpha}_{i,T} + \hat{\beta}_{i,T} X_{i,T}$$

然后,使用 L 乘子对 OLS 估计量进行噪声修正,从而获得修正后的收益率预测值

$$\tilde{Y}_{i,T+1} = \hat{L}_T^\alpha \hat{\alpha}_{i,T} + \hat{L}_T^\beta \hat{\beta}_{i,T} X_{i,T}$$

其中 \hat{L}_T^α 包括两种估计,即 $\hat{L}_{1,T}^\alpha$ 和 $\hat{L}_{2,T}^\alpha$, 它们的计算公式分别为式 (17) 和 式(21); 同样地, \hat{L}_T^β 也包括两种估计,即 $\hat{L}_{1,T}^\beta$ 和 $\hat{L}_{2,T}^\beta$, 计算公式分别为式 (18) 和式 (22). 以这种方式滚动向前,始终保持 T 个训练样本,直到样本外期末,最终获得了 q 个样本外股票收益率预测,记为 $\{\hat{Y}_{i,t+1}\}_{i=T}^{T+q-1}$ 和 $\{\tilde{Y}_{i,t+1}\}_{i=T}^{T+q-1}$. 在本文的实证分析中, T 的取值分别为 24 个月、36 个月、48 个月和 60 个月,而 q 的选择取决于各个预测变量的样本长度.

3.3.2 L 乘子性质

表 4 报告了两种 L 乘子估计值序列的均值和标准差,本文仅展示了滚动窗口为 24 个月和 60

个月的结果,以简化报告. 与理论推导一致,无论是对于截距项的 L 乘子还是预测变量系数的 L 乘子,通常情况下第一种方法的 L 乘子估计值 (\hat{L}_1^α 和 \hat{L}_1^β) 大于第二种方法的估计值 (\hat{L}_2^α 和 \hat{L}_2^β). 而两种 L 乘子的平均值 ($(\hat{L}_1 + \hat{L}_2)/2$) 则介于两者之间.

此外,随着估计窗口的递增(从 24 个月到 60 个月),大多数变量系数的 L 乘子估计值呈递增趋势. 这是因为随着训练样本的扩大,预测变量中的噪声成分减少,导致 L 乘子的修正力度会下降. 值得注意的是,截距项的 L 乘子 (\hat{L}^α) 平均小于预测变量系数的 L 乘子 (\hat{L}^β). 截距项代表股票超额收益率的平均值,这表明收益率本身的噪声较大,需要更强的噪声修正.

为了更清晰地观察 L 乘子序列的演变过程,本文以估计窗口长度等于 36 个月为例,在图 1 中展示了股息率 (d/p) 变量的 L 乘子序列. 图 1(a) 为截距项 (α) 的情况,图 1(b) 为回归系数 (β) 的情况. 为了显示 L 乘子如何随着 OLS 估计量的显著性水平而变化,本文在 L 乘子图的上方绘制了它们的 t 统计量 (t_α 和 t_β) 的柱状图,用水平虚线标记 10% 显著性水平的边界. 虚线以外的部分表示回归系数在 10% 的水平下是显著的.

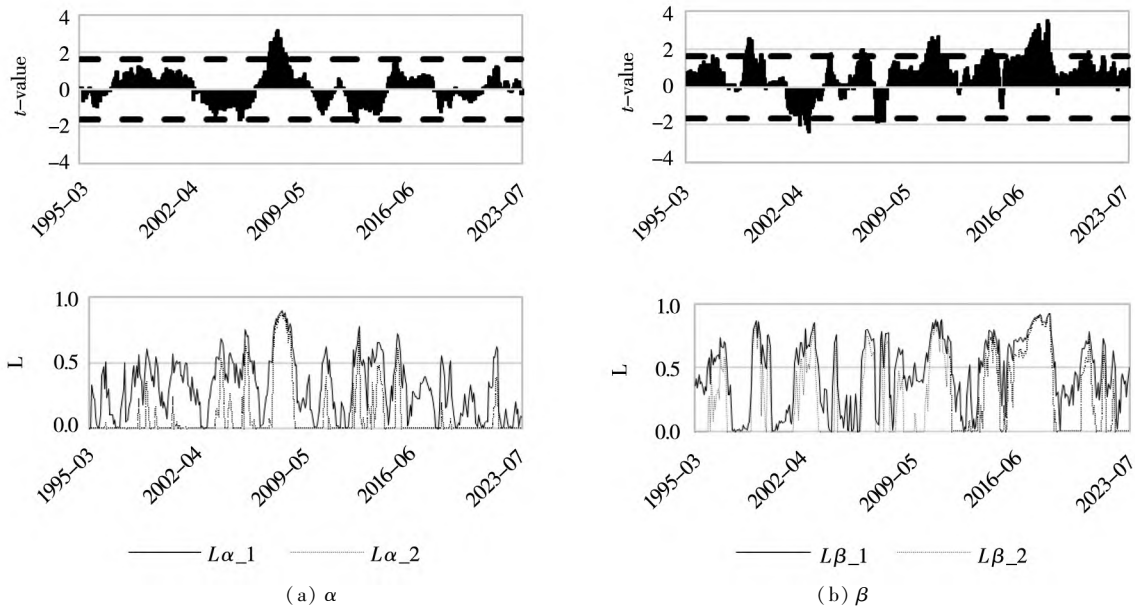


图 1 回归系数的 t 统计量值与 L 乘子的时间序列变化 (d/p , 估计窗口 = 36)

Fig. 1 Time series of regression coefficient t -statistics and L -Multipliers (d/p , estimation window = 36)

从图 1 中可以看出, L 乘子呈现较大的波动,而且不同估计方法下 L 乘子的变化趋势相似. 第二种

估计方法得到的 L 乘子 ($L\alpha_2$ 和 $L\beta_2$) 始终低于第一种估计方法的 L 乘子 ($L\alpha_1$ 和 $L\beta_1$), 而两种估计

方法的平均值($L\alpha_{ave}$ 和 $L\beta_{ave}$)则位于两者之间. 此外, L 乘子的变化与回归系数的 t 值密切相关. 当回归系数的 t 值较小时, L 乘子也较小, 表明对噪声修正的力度更大. 当 t 值较高时, 表示模型在样本内表现较好, 回归系数的噪声也较小, 因此 L 乘子的值相对较高, 对 OLS 估计量的收缩程度也较小.

此外, 从 L 乘子的时间序列图中, 本文发现 L 乘子在某些区间内持续取值较高, 存在一定的聚

集现象. 如图 1(b) 所示, 预测变量 d/p 的回归系数 L 乘子序列在 2018 年附近持续保持较高水平, 说明在这段时间无需对回归系数进行太多修正. 这可能存在两方面的原因: 一方面, 本文认为由于 L 乘子是通过滚动或者扩展的方式进行计算而得, 样本每次替换或者增加一个会使 L 乘子一定程度上表现出集聚性; 另一方面, 股票收益率本身以及预测变量与经济周期有关^[32, 55].

表 5 样本外预测的统计表现 ($\Delta MSFE$)

Table 5 Statistical performance of out-of-sample forecasts ($\Delta MSFE$)

变量	Panel A: 估计窗口 = 24			Panel B: 估计窗口 = 36		
	OLS + L_1	OLS + L_2	OLS + L_ave	OLS + L_1	OLS + L_2	OLS + L_ave
d/p	12.59*	18.77*	15.85*	2.50***	3.31***	2.97***
d/y	12.68*	18.57*	15.79*	2.47***	3.27***	2.94***
e/p	8.93**	11.88**	10.53**	5.52***	7.28***	6.46***
d/e	8.78*	12.93*	11.01*	2.57***	3.48***	3.08***
c/p	3.61**	4.70**	4.24**	0.86*	1.05*	1.01*
s/p	9.85**	12.67**	11.39**	5.51*	8.72*	7.18*
b/m	11.01**	14.19**	12.74**	6.84*	10.72*	8.85*
$op\ rev$	7.36***	9.69***	8.77***	2.41***	2.91***	2.76***
$svar$	8.39*	10.44	9.78*	3.55***	5.25***	4.47***
$turnover$	7.40***	9.71***	8.66***	3.58***	4.77***	4.24***
tbl	1.75**	1.62*	1.75**	1.28	1.51	1.42
lty	1.99**	2.02**	2.09**	1.08**	1.50**	1.32**
tms	0.41	-0.20	0.15	2.20**	2.60**	2.44**
dfy	1.97**	2.15**	2.12**	2.39***	2.91***	2.68***
dfr	4.33***	6.21***	5.35***	1.35	1.64	1.53
$infl$	6.14**	7.54**	7.02**	3.28***	4.32***	3.89***
$baidu$	3.45***	3.44***	3.51***	3.89***	4.58**	4.29***
变量	Panel C: 估计窗口 = 48			Panel D: 估计窗口 = 60		
	OLS + L_1	OLS + L_2	OLS + L_ave	OLS + L_1	OLS + L_2	OLS + L_ave
d/p	1.22**	1.46**	1.38**	1.30***	1.89***	1.62***
d/y	1.23**	1.43**	1.37**	1.29***	1.87***	1.61***
e/p	2.08***	2.59***	2.36***	0.88**	1.32**	1.12**
d/e	1.60***	1.80***	1.74***	1.17***	1.58**	1.41**
c/p	1.71**	2.30**	2.04**	1.33**	1.97**	1.67**
s/p	1.38***	1.41**	1.42**	0.80*	1.11*	0.97*
b/m	1.46***	1.43**	1.47**	0.89**	1.27**	1.10**
$op\ rev$	3.11***	3.92***	3.58***	2.11***	3.01***	2.61***
$svar$	2.21***	2.59***	2.44***	0.99*	1.33*	1.19*
$turnover$	2.25***	2.55***	2.44***	0.79*	1.11**	0.98**
tbl	1.9***	2.52**	2.23**	1.67***	2.20***	1.96***
lty	1.44**	1.92**	1.71**	1.78***	2.37***	2.09***
tms	3.34***	4.39***	3.90***	2.08***	2.80**	2.47***
dfy	1.75***	1.87***	1.83***	1.36***	1.65***	1.52***
dfr	2.61***	3.27***	3.02***	2.05***	2.83***	2.46***
$infl$	2.08***	2.64***	2.42***	1.45***	1.53***	1.55***
$baidu$	3.02**	4.17*	3.64*	1.51*	1.69	1.63

注: *, ** 和 *** 分别表示在 10%, 5% 和 1% 的水平下显著.

3.3.3 统计评价指标

为了评估 L 乘子模型相对于原始模型的样本外预测效果,本文使用均方预测误差 (MSFE) 的减少来进行度量. 定义如下

$$\Delta MSFE(L) = [MSFE(1) - MSFE(L)] \times 10\ 000$$

其中 $MSFE(1)$ 和 $MSFE(L)$ 分别代表在未进行和进行 L 乘子修正的模型下的收益率预测误差, 满足

$$MSFE(1) = \frac{1}{q} \sum_{k=1}^q (Y_{T+k} - \hat{Y}_{T+k})^2$$

$$MSFE(L) = \frac{1}{q} \sum_{k=1}^q (Y_{T+k} - \tilde{Y}_{T+k}(L))^2$$

其中 T 是样本内估计窗口的长度, q 是样本外预测评估期的长度. Y_{T+k} 表示 $T+k$ 月的股票收益率实际值, \hat{Y}_{T+k} 是原始模型下的收益率预测值, $\tilde{Y}_{T+k}(L)$ 是 L 乘子模型的收益率预测值. $\Delta MSFE$ 衡量了 L 乘子模型与原始模型相比在 MSFE 准则下的改进程度. 因此, 当 $\Delta MSFE > 0$ 时, 表明 L 乘子模型提供比原始模型更准确的预测.

为了确定 $\Delta MSFE > 0$ 是否具有统计显著性, 本文按照相关文献的做法^[51, 56, 57], 使用 Clark 和 West^[58] 提出的修正统计量, 以比较嵌套模型之间的预测性能差异^⑤. 具体而言, 为了检验零假设 $\Delta MSFE \leq 0$; 备择假设 $\Delta MSFE > 0$

首先, 需要计算

$$f_{T+k} = (Y_{T+k} - \hat{Y}_{T+k})^2 - (Y_{T+k} - \tilde{Y}_{T+k}(L))^2 + (\hat{Y}_{T+k} - \tilde{Y}_{T+k}(L))^2$$

接着将 $\{f_{T+k}\}_{k=1}^q$ 对一个常数项进行回归, 计算截距项回归系数的 t 统计量, 从而得到 Clark 和 West^[58] 统计量. 最后, 使用标准正态分布来计算相应的单边检验 p 值.

表 5 报告了 L 乘子模型 (OLS + L_1, OLS + L_2 和 OLS + L_ave) 相对于原始模型 (OLS) 的 MSFE 改进情况 ($\Delta MSFE$). 此外, 本文利用 Clark 和 West^[58] 的统计量, 在表中用“*”来表示这种 MSFE 改进是否具有统计显著性. Panel A ~ Panel D 分别记录了不同估计窗口长度下 (从 24 个月 ~ 60 个月) 的结果. 例如, 当估计窗口为 24 个月时, 表示使用最近 24 个月的数据滚动估计模型参数, 然后进行未来一个月的超额收益率预测.

表 5 的结果概括如下: 第一, 对于几乎所有的预测变量, 在不同长度的估计窗口内, L 乘子模型的预测误差都要明显小于原始模型的预测误差, 即 $\Delta MSFE > 0$; 第二, L 乘子修正带来的预测性能提升会随估计窗口长度的增加而逐渐减小. 例如, 对于大多数变量而言, 60 个月估计窗口下的 $\Delta MSFE$ 要小于 24 个月估计窗口下的相应值. 这是因为随着估计窗口的增加, 超额收益率和预测变量的信噪比提升. 此时, 使用 L 乘子进行噪声修正, 带来的改进效果逐渐减小; 第三, 使用 L 乘子修正噪声对于包含较少有价值信息的变量 (即预测能力较弱的变量) 来说更为有效; 第四, 关于数据噪声的修正, 第二种方法下的 L 乘子估计值在表现上通常优于第一种方法.

表 6 样本外预测的统计表现 (R_{OS}^2)

Table 6 Statistical performance of out-of-sample forecasts (R_{OS}^2)

变量	Panel A: 估计窗口 = 24				Panel B: 估计窗口 = 36			
	OLS	OLS + L_1	OLS + L_2	OLS + L_ave	OLS	OLS + L_1	OLS + L_2	OLS + L_ave
<i>d/e</i>	-12.93	-3.02	1.66 **	-0.50 *	-4.15	0.08 *	1.56 **	0.92 **
<i>c/p</i>	-1.51	4.80 **	6.70 **	5.90 **	0.34 *	1.88 *	2.21	2.13 *
<i>svar</i>	-5.73	2.07 **	3.98 ***	3.37 ***	-0.50	3.45 **	5.34 ***	4.47 ***
<i>turnover</i>	-6.6	1.51 **	4.04 **	2.90 **	-3.48	0.68 **	2.07 ***	1.46 **
<i>tbl</i>	1.77 *	4.68 **	4.46 **	4.67 **	1.59 **	3.72 *	4.11 *	3.96 *
<i>lty</i>	-0.76	2.53 *	2.59 *	2.70 *	7.26 **	9.06 **	9.77 ***	9.47 ***
<i>infl</i>	-4.26	1.44 *	2.75 **	2.26 **	-2.16	1.48 *	2.64 *	2.16 *

⑤ 如果使用 Bootstrap 方法直接检验 $\Delta MSFE$ 是否显著大于 0, 本文的研究结果依然成立.

续表 6

Table 6 Continues

变量	Panel C: 估计窗口 = 48				Panel D: 估计窗口 = 60			
	OLS	OLS + L_1	OLS + L_2	OLS + L_ave	OLS	OLS + L_1	OLS + L_2	OLS + L_ave
<i>d/e</i>	-2.75	0.06	0.41 *	0.31 *	-1.43	0.7 **	1.45 **	1.13 **
<i>c/p</i>	-1.66	1.46	2.53	2.06	1.09 *	3.39 *	4.5 **	3.98 **
<i>svar</i>	-1.81	1.90 **	2.54 **	2.29 **	-1.27	0.47	1.07	0.82
<i>turnover</i>	-1.92 *	2.04 **	2.55 ***	2.37 **	1.34 **	2.78 **	3.37 ***	3.13 **
<i>tbl</i>	2.67 *	5.86 *	6.90 *	6.42 *	3.41 *	6.29 *	7.21 *	6.78 *
<i>lty</i>	3.48 *	5.91 **	6.71 **	6.35 **	1.46 *	4.52 **	5.54 **	5.07 **
<i>infl</i>	-0.11 *	3.38 **	4.31 **	3.96 **	0.66 *	3.2 **	3.34 **	3.38 **

注：*，** 和 *** 分别表示在 10%，5% 和 1% 的水平下显著。

除了使用 $\Delta MSFE$ 来比较不同模型的样本外预测性能之外,现有关于股票收益率预测的文献常采用 R_{os}^2 来评估预测模型相对于历史平均基准模型的样本外预测表现^[37, 41-43, 48]. 为了探究 L 乘子模型能否打败历史平均基准模型, 本文从 17 个预测变量中挑选了 7 个具有较高样本外预测表现的变量, 并计算了这些预测变量在不同模型下, 相对于(滚动窗口下)历史平均收益率预测的表现. 表 6 报告了原始模型(OLS)和 L 乘子模型(OLS + L_1, OLS + L_2 和 OLS + L_ave)的 R_{os}^2 . 可以看出, 在许多情况下, 原始模型很难超越简单历史平均预测, 但一旦经过 L 乘子修正, 这些模型在很多情况下都可以显著地打败历史平均预测. 通过对变量中噪声以及截距项的修正, 能够显著提高预测能力, 获得正的样本外 R_{os}^2 .

3.3.4 累计均方预测误差

为了更好地了解 L 乘子模型的样本外预测表现在不同时间段和特定时期的变化情况, 本文按照 Welch 和 Goyal^[44] 以及 Rapach 等^[37] 的方法, 计算了原始模型的累积平方预测误差(CSFE)与 L 乘子模型的 CSFE 之间的差异. 计算方法如下

$$\Delta CSFE_t = \sum_{k=1}^t (Y_{T+k} - \hat{Y}_{T+k})^2 - \sum_{k=1}^t (Y_{T+k} - \tilde{Y}_{T+k}(L))^2$$

以 36 个月的估计窗口为例, 图 2 展示了预测变量 (d/p) 下 $\Delta CSFE_t$ 的时间序列变化趋势. 通过

重新绘制水平零线, 这些图形就能展示在特定的样本外期间内, L 乘子模型是否比原始模型具有更低的 MSFE. 如果曲线在样本外期间结束时比开始时更高, 则意味着 L 乘子模型在该样本外期间内 MSFE 更低. 从图 2 中可以看出, 所有模型的曲线总体呈现出正的斜率, 这表明 L 乘子模型在大多数时期相对于原始模型具有更高的样本外预测精度. 此外, 点线基本上位于点划线的上方, 这表明第二种估计方法下的 L 乘子模型在不同时段内表现持续地优于第一种估计方法下的 L 乘子模型.

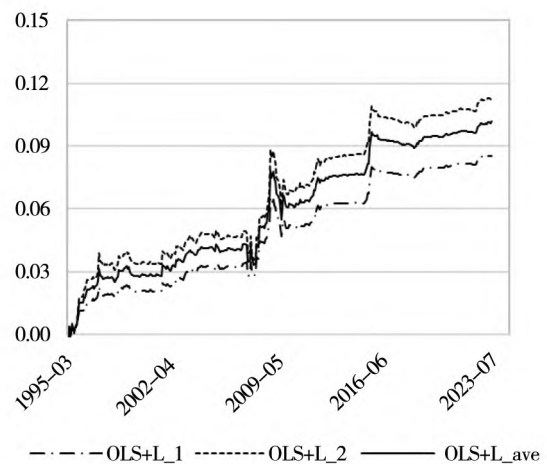


图 2 原始模型的 CSFE 与 L 乘子模型的 CSFE 之差 (d/p , 估计窗口 = 36)

Fig. 2 Difference between CSFE of the original model and CSFE of the L-Multiplier corrected model (d/p , estimation window = 36)

3.4 样本外预测的经济表现

样本外预测的表现表明, 本文提出的 L 乘子

模型明显优于原始模型. 下一个问题是, L 乘子修正方法是否能够为实现现实世界的投资者实现显著的经济收益. 另外, $\Delta MSFE$ 指标的局限性在于它没有明确考虑投资者在样本外期间所承担的风险. 为了解决这些问题, 本文从资产配置的角度, 按照 Campbell 和 Thompson^[48]、Rapach 等^[37]、Neely 等^[50] 和 Jiang 等^[51] 的方法, 继续通过计算确定性等价收益率 (CER) 来衡量 L 乘子修正带来的经济价值.

具体而言, 首先考虑一个相对风险厌恶参数为 γ 的均值 - 方差效用投资者, 该投资者基于原始模型的预测按月构建股票和无风险资产的投资组合. 在 T 月末, 投资者需要决定将其资金的以下比例分配给 $T + 1$ 月的股票资产

$$\hat{w}_{T+1} = \left(\frac{1}{\gamma}\right) \left(\frac{\hat{Y}_{T+1}}{\hat{\sigma}_{T+1}^2}\right)$$

其中 \hat{Y}_{T+1} 是原始模型下的超额收益率预测值. 需要注意的是, 这个过程要求投资者对股票收益率的方差 $\hat{\sigma}_{T+1}^2$ 进行预测. 为了简化分析, 假设投资者使用估计窗口内 (如历史 24 个月 ~ 60 个月) 的月度收益率序列的样本方差来估计^⑥. 因此, 投资组合权重的计算仅因超额收益预测而不同.

在样本外期间, 基于原始模型的预测所构建的投资组合可以获得的 CER 如下

$$\hat{v} = \hat{\mu} - \left(\frac{1}{2}\right)\gamma \hat{\sigma}^2$$

其中 $\hat{\mu}$ 和 $\hat{\sigma}^2$ 分别为基于原始模型构建的投资组

合的样本外均值和方差.

对于投资者使用 L 乘子模型进行股票超额收益率预测并构建投资组合, 相应的权重和 CER 分别为

$$\begin{aligned} \tilde{w}_{T+1}(L) &= \left(\frac{1}{\gamma}\right) \left(\frac{\tilde{Y}_{T+1}(L)}{\tilde{\sigma}_{T+1}^2}\right) \\ \tilde{v}(L) &= \tilde{\mu} - \left(\frac{1}{2}\right)\gamma \tilde{\sigma}^2 \end{aligned}$$

其中 $\tilde{Y}_{T+1}(L)$ 是 L 乘子模型的收益率预测值. $\tilde{\mu}$ 和 $\tilde{\sigma}^2$ 分别为基于 L 乘子模型构建的投资组合的样本外均值和方差.

按照 Rapach 等^[37] 和 Wang 等^[54] 的做法, $\gamma = 3$. 然后, 将 CER 增益定义为投资者基于 L 乘子模型构建投资组合所获得的 CER 与使用原始模型构建投资组合所获得的 CER 之间的差异, 将式子乘以 1 200 以反映年化百分比的 CER 增益, 即

$$\Delta CER = (\tilde{v}(L) - \hat{v}) \times 1\ 200$$

表 7 报告了 L 乘子模型相对于原始模型 (OLS) 的 CER 改进情况 (ΔCER). Panel A ~ Panel D 分别记录了不同估计窗口长度下 (从 24 个月 ~ 60 个月) 的结果. 从表 7 的结果中可以看出, 对于绝大多数变量 ΔCER 为正值, 这些结果基本符合表 5 中的研究结论, 进一步表明 L 乘子模型的预测性能提升, 可以为均值 - 方差效用的投资者带来显著的投资绩效提升.

表 7 样本外预测的经济表现 (ΔCER)

Table 7 Economic performance of out-of-sample forecasts (ΔCER)

变量	Panel A: 估计窗口 = 24			Panel B: 估计窗口 = 36		
	OLS + L_1	OLS + L_2	OLS + L_ave	OLS + L_1	OLS + L_2	OLS + L_ave
d/p	23.44	30.56	27.44	10.06	12.51	11.45
d/y	22.58	29.40	26.38	9.91	12.21	11.22
e/p	19.90	26.29	23.50	10.81	13.91	12.47
d/e	18.69	24.39	21.92	6.44	8.41	7.57
c/p	12.05	14.65	13.76	4.33	4.10	4.41
s/p	19.23	24.24	22.05	9.99	13.42	11.87
b/m	18.22	22.99	20.88	9.40	11.98	10.86

⑥ 可以按照 Campbell 和 Thompson^[44] 和 Neely 等^[50] 的方法, 采用过去五年的月度股票收益率滚动窗口来估计方差预测. 本文的主要研究结论不受这种估计方法的影响.

续表 7
Table 7 Continues

变量	Panel A: 估计窗口 = 24			Panel B: 估计窗口 = 36		
	OLS + L_1	OLS + L_2	OLS + L_ave	OLS + L_1	OLS + L_2	OLS + L_ave
<i>op rev</i>	22.41	30.23	26.70	11.71	12.89	12.51
<i>svar</i>	5.84	5.81	6.15	4.75	5.56	5.26
<i>turnover</i>	9.86	12.34	11.35	6.23	6.77	6.75
<i>tbl</i>	15.56	19.28	17.69	-1.53	-2.93	-2.13
<i>lty</i>	8.99	10.05	9.89	4.19	5.10	4.78
<i>tms</i>	5.12	5.42	5.48	5.03	4.83	5.07
<i>dfy</i>	5.68	7.03	6.63	4.47	4.93	4.81
<i>dfr</i>	15.55	20.68	18.35	11.31	14.10	12.93
<i>infl</i>	9.50	10.63	10.34	6.31	8.24	7.40
<i>baidu</i>	19.91	22.51	21.87	19.87	25.74	23.36
变量	Panel C: 估计窗口 = 48			Panel D: 估计窗口 = 60		
	OLS + L_1	OLS + L_2	OLS + L_ave	OLS + L_1	OLS + L_2	OLS + L_ave
<i>d/p</i>	5.40	6.56	6.10	5.56	7.23	6.50
<i>d/y</i>	5.44	6.38	6.03	5.54	7.19	6.47
<i>e/p</i>	6.58	8.65	7.72	3.82	5.35	4.65
<i>d/e</i>	5.93	7.02	6.57	7.63	10.06	8.99
<i>c/p</i>	5.25	7.13	6.28	2.92	4.02	3.55
<i>s/p</i>	6.14	7.73	7.03	3.78	5.29	4.59
<i>b/m</i>	5.15	6.21	5.77	4.58	6.17	5.43
<i>op rev</i>	11.99	14.70	13.48	8.35	10.80	9.68
<i>svar</i>	5.55	6.71	6.22	6.52	8.40	7.55
<i>turnover</i>	8.81	10.57	9.88	8.51	11.58	10.16
<i>tbl</i>	2.51	2.91	2.77	3.10	3.82	3.52
<i>lty</i>	5.47	7.38	6.50	6.28	7.93	7.16
<i>tms</i>	6.66	8.19	7.54	4.20	5.68	5.01
<i>dfy</i>	3.72	4.22	4.00	4.32	5.50	4.95
<i>dfr</i>	9.93	12.98	11.58	8.68	12.46	10.66
<i>infl</i>	3.71	4.68	4.28	4.36	5.51	4.99
<i>baidu</i>	10.90	14.61	13.12	2.95	3.31	3.21

3.5 多变量预测的应用

已有文献对多变量的降维^[59]进行探讨,岭回归^[14]、LASSO^[15]等收缩方法也能实现多变量情形的预测.本文主要探讨 L 乘子对单变量 OLS 预测模型的改进效果,在此简单探讨多变量预测中 L 乘子模型的适用性,并与现有文献中应用较多的 LASSO、岭回归、主成分分析和偏最小二乘回归方法进行简单比较.考虑到文章篇幅的限制,本文在此处仅展示样本外预测的统计结果,结论也同样适用于相关的经济表现.

3.5.1 与 LASSO、岭回归的简单比较

LASSO、岭回归等是现有应用较多的收缩技

术,多用于多变量预测的情形.为了进行对比,本文考虑单变量预测后进行更进一步的组合的结果,故在此将本文提出的 L 乘子修正方法与这两种技术进行简单的比较.基于此,本文对比多变量 OLS 回归、LASSO 回归、岭回归、单变量 OLS 模型的组合预测以及基于单变量 L 乘子模型(OLS + L_1、OLS + L_2 和 OLS + L_ave)的组合预测这 7 种情形下各模型的 MSFE,具体见表 8.这里涉及的组合预测均为等权组合预测的结果,Panel A ~ Panel D 分别记录了不同滚动窗口长度下(24 个月、36 个月、48 个月和 60 个月)的结果.

在 OLS 模型的基本设定下,考虑 LASSO 和岭回归方法下多元线性回归模型下的 MSFE,此时线性回归模型中对 β 系数进行一定约束,在两种方法下最优解分别为

$$\hat{\beta}^{Lasso} = \arg \min_{\beta} \left(\frac{1}{2(T-1)} \sum_{t=1}^{T-1} \left(Y_{t+1} - \beta_0 - \sum_{k=1}^K \beta_k X_{k,t} \right)^2 + \lambda \sum_{k=1}^K |\beta_k| \right)$$

$$\hat{\beta}^{Ridge} = \arg \min_{\beta} \left(\frac{1}{2(T-1)} \sum_{t=1}^{T-1} \left(Y_{t+1} - \beta_0 - \sum_{k=1}^K \beta_k X_{k,t} \right)^2 + \lambda \sum_{k=1}^K \beta_k^2 \right)$$

以上分别为 LASSO 和岭回归模型中 β 系数的求解式. 本文利用交叉验证(cross validation)方法求解最优 λ , 其中 λ 是基于样本内的数据,通过在 $[0.001, 0.01, 0.1, 1, 10]$ 内搜索来确定.

表 8 与 LASSO 和岭回归的简单比较

Table 8 Simple comparison with LASSO and Ridge regression

	Panel A: 滚动窗口 = 24	Panel B: 滚动窗口 = 36	Panel C: 滚动窗口 = 48	Panel D: 滚动窗口 = 60
多变量 OLS 回归	1 044. 22	96. 06	58. 31	58. 15
多变量 LASSO	45. 76	37. 89	35. 51	36. 16
多变量 Ridge	52. 33	37. 13	36. 17	37. 12
OLS 单变量组合预测	41. 89	35. 96	34. 94	35. 49
(单变量 OLS + L1) 组合预测	41. 50	35. 44	34. 61	35. 30
(单变量 OLS + L2) 组合预测	41. 29	35. 04	34. 45	35. 33
(单变量 OLS + L_ave) 组合预测	41. 37	35. 23	34. 52	35. 31

注: 用 MSFE 衡量, 表中所涉及的组合预测均为等权组合预测的结果.

表 8 报告了多变量 OLS 回归、LASSO 回归、岭回归、单变量 OLS 模型的组合预测以及基于单变量 L 乘子模型(OLS + L_1、OLS + L_2 和 OLS + L_ave)的组合预测这 7 种情形下各模型的 MSFE 值的大小. 结果表明: 第一, 多变量 LASSO 或者岭回归模型优于简单的多变量 OLS 回归, 尤其在短窗口的表现差异更大. 第二, 单变量 OLS 模型组合预测优于 LASSO 或者岭回归. 第三, 基于单变量 OLS 模型的组合预测优于单变量 OLS 模型的组合预测. 实际上基于单变量 L 乘子模型的组合预测本身也是一种非常大的缩小变量系数的方法, 因此基于单变量 L 乘子模型的组合预测相比于单变量 OLS 模型的组合预测总体上改进不是很大, 但也都有所改进, 其它窗口结果也类似.

3. 5. 2 在主成分分析(PCA)和偏最小二乘(PLS)回归中的应用

主成分分析(Principal Component Analysis, PCA)是一种统计技术, 通过降维处理来保留数据中的主要变异性^[60]. 在收益率预测中, PCA 可以有效地从大量相关的金融变量中提取关键信息, 有助于简化数据结构并增强预测模型的

性能.

此外, 由于 PCA 方法是以尽可能多地捕捉预测变量方差为目标, 并不涉及目标变量, 因此所提取的因子未必具有较好的预测能力. 本文以 Huang 等^[55]提出的偏最小二乘(Partial Least Squares Regression, PLS)回归方法为例, 观察 L 乘子模型在有监督的因子模型中的应用效果. 分别考虑对所有预测变量进行 PCA 和 PLS 分析, 获取前 3 个主成分进行多元线性预测回归, 并对每个主成分进行 L 乘子修正. Panel A ~ Panel D 分别记录了不同滚动窗口长度下(24 个月、36 个月、48 个月以及 60 个月), L 乘子模型在 PCA 和 PLS 模型中的应用结果.

表 9 的 Panel A 报告了 PCA 和 PCA 经 L 乘子(L_1、L_2 和 L_ave)修正后这 4 种情形下各模型 MSFE 值大小. 结果表明, 经 L 乘子修正后的 PCA 模型的 MSFE 均小于不修正的 PCA 模型的结果, 说明 L 乘子也能较好运用于 PCA 模型中. 表 9 的 Panel B 报告了 PLS 和 PLS 经 L 乘子(L_1、L_2 和 L_ave)修正后这 4 种情形下各模型 MSFE 值大小, 与 PCA 的结果近似.

表 9 在 PCA 和 PLS 模型中的应用

Table 9 Application in PCA and PLS models

	滚动窗口 = 24	滚动窗口 = 36	滚动窗口 = 48	滚动窗口 = 60
Panel A: PCA 模型				
PCA	60.31	44.59	38.60	38.74
PCA + L1	51.95	41.23	36.98	36.46
PCA + L2	50.75	39.96	36.33	35.63
PCA + L_ave	51.13	40.54	36.59	36.01
Panel B: PLS 模型				
PLS	72.73	47.52	39.07	41.21
PLS + L1	63.36	42.20	36.13	39.63
PLS + L2	61.44	40.71	35.38	40.01
PLS + L_ave	62.28	41.42	35.73	39.79

注：用 MSFE 衡量。

4 稳健性检验

为增强预测的可信度，在此考虑不同的市场指数（如沪深 300 指数、标普 500 指数）^⑦，从统

计表现对不同估计窗口的结论进行稳健性检验，经济表现（ ΔCER ）的结果也类似。出于篇幅的考虑，表 10 ~ 表 11 中仅展示 24 个月、60 个月窗口的结果，发现 L 乘子模型能一致提高预测能力。

表 10 沪深 300 指数收益率预测的统计表现（ $\Delta MSFE$ ）

Table 10 Statistical performance of CSI 300 Index forecasts（ $\Delta MSFE$ ）

变量	Panel A: 估计窗口 = 24			Panel B: 估计窗口 = 60		
	OLS + L_1	OLS + L_2	OLS + L_ave	OLS + L_1	OLS + L_2	OLS + L_ave
<i>d/p</i>	8.22 **	10.87 **	9.63 **	1.35 ***	1.64 **	1.52 ***
<i>d/y</i>	8.13 **	10.82 **	9.56 **	1.33 ***	1.62 ***	1.5 ***
<i>e/p</i>	7.05 ***	9.43 ***	8.34 ***	0.94 ***	1.3 ***	1.13 ***
<i>d/e</i>	5.19 ***	6.37 ***	5.85 ***	1.6 ***	2.02 ***	1.84 ***
<i>c/p</i>	3.07 ***	4.21 ***	3.71 ***	2.31 ***	3.13 ***	2.75 ***
<i>s/p</i>	5.45 ***	7.28 ***	6.43 ***	1.79 **	2.54 **	2.2 **
<i>b/m</i>	5.6 ***	7.53 **	6.64 ***	1.66 **	2.3 **	2.01 **
<i>op rev</i>	6.41 ***	7.95 ***	7.26 ***	1.61 ***	1.82 ***	1.76 ***
<i>svar</i>	3.41 ***	4.18 ***	3.87 ***	2.24 ***	3.09 ***	2.7 ***
<i>turnover</i>	6.04 ***	7.28 ***	6.76 ***	2.33 **	2.98 *	2.7 **
<i>tbl</i>	3.13 **	3.71 ***	3.5 ***	2.07 ***	2.43 ***	2.28 ***
<i>lty</i>	3.35 ***	3.87 ***	3.72 ***	2.31 ***	3.11 ***	2.73 ***
<i>tms</i>	1.19 **	0.99 *	1.15 **	2.03 **	2.52 **	2.3 **
<i>dfy</i>	2.66 ***	3.09 ***	2.96 ***	1.42 ***	1.84 ***	1.65 ***
<i>dfr</i>	5.04 ***	7.4 ***	6.31 ***	1.88 ***	2.57 ***	2.25 ***
<i>infl</i>	2.93 ***	3.16 ***	3.14 ***	1.24 **	1.48 **	1.38 **
<i>baidu</i>	4.55 ***	5.17 ***	4.94 ***	1.16	1.04	1.13

注：*，** 和 *** 分别表示在 10%，5% 和 1% 的水平下显著。

⑦ 考虑预测变量自相关性调整后的结果也近似，总体有所提高。

表 11 标普 500 指数收益率统计表现($\Delta MSFE$)Table 11 Statistical performance of S&P 500 Index returns($\Delta MSFE$)

变量	Panel A: 估计窗口 = 24			Panel B: 估计窗口 = 60		
	OLS + L_1	OLS + L_2	OLS + L_ave	OLS + L_1	OLS + L_2	OLS + L_ave
d/p	0.73 ***	0.75 ***	0.76 ***	0.26 ***	0.23 **	0.25 ***
d/y	0.79 ***	0.84 ***	0.84 ***	0.29 ***	0.28 ***	0.3 ***
e/p	1.19 ***	1.32 ***	1.29 ***	0.4 ***	0.42 ***	0.42 ***
d/e	1.53 ***	1.82 ***	1.71 ***	0.53 ***	0.58 ***	0.56 ***
$svar$	1.69 **	1.94 ***	1.84 **	0.53 **	0.56 **	0.55 **
b/m	0.85 ***	0.85 ***	0.87 ***	0.29 ***	0.32 **	0.31 **
tbl	1.19 ***	1.3 ***	1.28 ***	0.44 ***	0.49 ***	0.47 ***
lty	0.96 ***	1.03 ***	1.02 ***	0.39 ***	0.43 ***	0.42 ***
tms	1.19 ***	1.45 ***	1.35 ***	0.41 ***	0.45 ***	0.44 ***
dfy	0.94 ***	1.05 ***	1.02 ***	0.31 ***	0.34 ***	0.34 ***
dfi	0.94 ***	1.28 ***	1.13 ***	0.22 ***	0.22 **	0.23 ***
$infl$	0.85 ***	1.09 ***	1.0 ***	0.25 ***	0.26 ***	0.26 ***

注: *, ** 和 *** 分别表示在 10%, 5% 和 1% 的水平下显著。

5 结束语

本文在经典的线性回归模型基础上, 提出一个直接、简单且有经济含义的 L 乘子模型. 在考虑参数估计的随机噪声因素后, 证明了基于有限样本的拟合最优模型通过修正能够更好地预测未来, 并从理论上给出最优 L 乘子的显式解以及两种参数估计方法. 本文通过实证检验了 L 乘子模型在样本外预测性能方面具有显著优势. 具体而言, 本文使用经典文献中常用的一系列变量对股票月度超额收益率进行预测, 无论时间长短和样本多少, L 乘子模型始终优于原始模型. 此外, 如果预测变量包含有比较丰富的, 或有价值的信息时, L 乘子模型能够打败历史平均预测. 最后, 本文还考虑多变量的情形,

与 LASSO、岭回归这两种现有的收缩方法进行简单比较, 并考虑 L 乘子在 PCA 和 PLS 回归模型中的应用.

本研究通过理论和实证探讨了如何通过减少噪声的影响来提高预测准确性, 从而提高投资组合的绩效. 从实证结果来看, 不论是国内市场还是国际市场, 不论是统计表现还是经济表现, L 乘子模型都能显著优于原始模型. 这些结果对其他市场预测研究^[61]都具有重要的借鉴意义. 本研究为如何通过减少噪声影响提高收益率预测性能提供了一个新的视角, L 乘子模型具有广泛的适用性. 在未来的研究中, 当信噪比较高的时候, 考虑先进行信息识别再进行噪声修正, 降低 L 乘子修正对信息利用的影响, 同时也可以将模型推广到包括机器学习、金融科技领域大数据建模中噪声处理^[62]在内的多种应用场景.

参 考 文 献:

- [1] Copas J B. Regression, prediction and shrinkage[J]. Journal of the Royal Statistical Society Series B: Statistical Methodology, 1983, 45(3): 311–335.
- [2] Connor G. Sensible return forecasting for portfolio management[J]. Financial Analysts Journal, 1997, 53(5): 44–51.
- [3] Ferreira M A, Santa-Clara P. Forecasting stock market returns: The sum of the parts is more than the whole[J]. Journal of Financial Economics, 2011, 100(3): 514–537.
- [4] Mantilla-García D, Vaidyanathan V. Predicting stock returns in the presence of uncertain structural changes and sample noise[J]. Financial Markets and Portfolio Management, 2017, 31(3): 357–391.

- [5] Amihud Y, Hurvich C M. Predictive regressions: A reduced-bias estimation method[J]. *Journal of Financial and Quantitative Analysis*, 2004, 39(4): 813 – 841.
- [6] James W, Stein C M. Estimation with quadratic loss[J]. *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, 1961, (1): 361 – 380.
- [7] Ashley R. Shrinkage estimation with general loss functions: An application of stochastic dominance theory[J]. *International Economic Review*, 1990, 31(2): 301 – 313.
- [8] Jorion P. Bayes-Stein estimation for portfolio analysis[J]. *Journal of Financial and Quantitative Analysis*, 1986, 21(3): 279 – 292.
- [9] Jorion P. Bayesian and CAPM estimators of the means: Implications for portfolio selection[J]. *Journal of Banking & Finance*, 1991, 15(3): 717 – 727.
- [10] Kan R, Zhou G. Optimal portfolio choice with parameter uncertainty[J]. *Journal of Financial and Quantitative Analysis*, 2007, 42(3): 621 – 656.
- [11] DeMiguel V, Martin-Utrera A, Nogales F J. Size matters: Optimal calibration of shrinkage estimators for portfolio selection [J]. *Journal of Banking & Finance*, 2013, 37(8): 3018 – 3034.
- [12] Kan R, Wang X, Zhou G. Optimal portfolio choice with estimation risk: No risk-free asset case[J]. *Management Science*, 2022, 68(3): 2047 – 2068.
- [13] Kircher F, Rösch D. A shrinkage approach for Sharpe ratio optimal portfolios with estimation risks[J]. *Journal of Banking & Finance*, 2021, (133): 106281.
- [14] Hoerl A E, Kennard R W. Ridge regression: Biased estimation for nonorthogonal problems[J]. *Technometrics*, 1970, 12(1): 55 – 67.
- [15] Tibshirani R. Regression shrinkage and selection via the lasso[J]. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 1996, 58(1): 267 – 288.
- [16] Rapach D E, Strauss J K, Zhou G. International stock return predictability: What is the role of the United States? [J]. *The Journal of Finance*, 2013, 68(4): 1633 – 1662.
- [17] Gu S, Kelly B, Xiu D. Empirical asset pricing via machine learning[J]. *The Review of Financial Studies*, 2020, 33(5): 2223 – 2273.
- [18] Chincio A, Clark-Joseph A D, Ye M. Sparse signals in the cross-section of returns[J]. *The Journal of Finance*, 2019, 74(1): 449 – 492.
- [19] Freyberger J, Neuhierl A, Weber M. Dissecting characteristics nonparametrically[J]. *The Review of Financial Studies*, 2020, 33(5): 2326 – 2377.
- [20] Kozak S, Nagel S, Santosh S. Shrinking the cross-section [J]. *Journal of Financial Economics*, 2020, 135(2): 271 – 292.
- [21] Dong X, Li Y, Rapach D E, et al. Anomalies and the expected market return[J]. *The Journal of Finance*, 2022, 77(1): 639 – 681.
- [22] Mazumder R, Radchenko P, Dedieu A. Subset selection with shrinkage: Sparse linear modeling when the SNR is low[J]. *Operations Research*, 2023, 71(1): 129 – 147.
- [23] Kelly B, Malamud S, Zhou K. The virtue of complexity in return prediction[J]. *The Journal of Finance*, 2024, 79(1): 459 – 503.
- [24] 陈 艳, 王宣承. 基于变量选择和遗传网络规划的期货高频交易策略研究[J]. *中国管理科学*, 2015, 23(10): 47 – 56.
Chen Yan, Wang Xuancheng. Research on high-frequency trading strategies for futures based on variable selection and genetic network programming[J]. *Chinese Journal of Management Science*, 2015, 23(10): 47 – 56. (in Chinese)
- [25] 蒋翠侠, 刘玉叶, 许启发. 基于 LASSO 分位数回归的对冲基金投资策略研究[J]. *管理科学学报*, 2016, 19(3): 107 – 126.
Jiang Cuixia, Liu Yuye, Xu Qifa. Research on hedge fund investment strategies based on LASSO quantile regression[J].

- Journal of Management Sciences in China, 2016, 19(3): 107 – 126. (in Chinese)
- [26] 王小燕, 袁 腾, 段湘斌. 基于零膨胀分位数两部模型的银行贷款违约预测研究[J]. 中国管理科学, 2022, 30(10): 1 – 13.
- Wang Xiaoyan, Yuan Teng, Duan Xiangbin. Research on bank loan default prediction based on the zero-inflated quantile two-part model[J]. Chinese Journal of Management Science, 2022, 30(10): 1 – 13. (in Chinese)
- [27] 李 斌, 邵新月, 李玥阳. 机器学习驱动的基本面量化投资研究[J]. 中国工业经济, 2019, (8): 61 – 79.
- Li Bin, Shao Xinyue, Li Yueyang. Research on fundamental quantitative investment driven by machine learning[J]. China's Industrial Economics, 2019, (8): 61 – 79. (in Chinese)
- [28] 姜富伟, 薛 浩, 周 明. 大数据提升了多因子模型定价能力吗? ——基于机器学习方法对我国 A 股市场的探究[J]. 系统工程理论与实践, 2022, 42(8): 2037 – 2048.
- Jiang Fuwei, Xue Hao, Zhou Ming. Does big data improve the pricing ability of multi-factor models?: An exploration of China's A-share market based on machine learning methods[J]. Systems Engineering: Theory & Practice, 2022, 42(8): 2037 – 2048. (in Chinese)
- [29] 梁 超, 魏 宇, 马 锋, 等. 我国黄金期货价格波动率预测研究: 来自模型缩减方法的新证据[J]. 中国管理科学, 2022, 30(4): 30 – 41.
- Liang Chao, Wei Yu, Ma Feng, et al. Research on the prediction of gold futures price volatility in China: New evidence from model reduction methods[J]. Chinese Journal of Management Science, 2022, 30(4): 30 – 41. (in Chinese)
- [30] 郑挺国, 范馨月, 靳 炜, 等. 通胀预期形成与信息黏性特征: 基于媒体新闻视角[J]. 世界经济, 2023, 46(4): 60 – 82.
- Zheng Tingguo, Fan Xinyue, Jin Wei, et al. Formation of inflation expectations and characteristics of information stickiness: A media news perspective[J]. The Journal of World Economy, 2023, 46(4): 60 – 82. (in Chinese)
- [31] Stambaugh R F. Predictive regressions[J]. Journal of Financial Economics, 1999, 54(3): 375 – 421.
- [32] 姜富伟, 涂 俊, David E, 等. 中国股票市场可预测性的实证研究[J]. 金融研究, 2011, (9): 107 – 121.
- Jiang Fuwei, Tu Jun, David E, et al. An empirical study on the predictability of China's stock market[J]. Journal of Financial Research, 2011, (9): 107 – 121. (in Chinese)
- [33] 蒋志强, 田婧雯, 周炜星. 中国股票市场收益率的可预测性研究[J]. 管理科学学报, 2019, 22(4): 92 – 109.
- Jiang Zhiqiang, Tian Jingwen, Zhou Weixing. Research on the predictability of stock market returns in China[J]. Journal of Management Sciences in China, 2019, 22(4): 92 – 109. (in Chinese)
- [34] 尹海员, 吴兴颖. 投资者高频情绪对股票日内收益率的预测作用[J]. 中国工业经济, 2019, (8): 80 – 98.
- Yin Haiyuan, Wu Xingying. The predictive effect of investors' high-frequency sentiment on intraday stock returns[J]. China Industrial Economics, 2019, (8): 80 – 98. (in Chinese)
- [35] 杨晓兰, 王伟超, 高 媚. 股市政策对股票市场的影响——基于投资者社会互动的视角[J]. 管理科学学报, 2020, 23(1): 15 – 32.
- Yang Xiaolan, Wang Weichao, Gao Mei. The impact of stock market policies on the stock market: From the perspective of investors' social interaction[J]. Journal of Management Sciences in China, 2020, 23(1): 15 – 32. (in Chinese)
- [36] 姜富伟, 胡逸驰, 黄 楠. 央行货币政策报告文本信息, 宏观经济与股票市场[J]. 金融研究, 2021, 492(6): 95 – 113.
- Jiang Fuwei, Hu Yichi, Huang Nan. Textual information in the people's bank of china's monetary policy reports, macroeconomics, and the stock market[J]. Journal of Financial Research, 2021, 492(6): 95 – 113. (in Chinese)
- [37] Rapach D E, Strauss J K, Zhou G. Out-of-sample equity premium prediction: Combination forecasts and links to the real economy[J]. The Review of Financial Studies, 2010, 23(2): 821 – 862.
- [38] Farmer L E, Lars S, Allan T. Pockets of predictability[J]. The Journal of Finance, 2023, 78(3): 1279 – 1341.
- [39] Goyal A, Ivo W, Alexander Z. A comprehensive 2022 look at the empirical performance of equity premium prediction[J]. The Review of Financial Studies, 2024, 37(11): 3490 – 3557.
- [40] 刘 莉, 郝显峰, 王玉东. 基于时变稳健加权最小二乘法的股市收益率预测[J]. 管理科学学报, 2024, 27(1):

- 141 – 158.
- Liu Li, Hao Xianfeng, Wang Yudong. Prediction of stock market returns based on time-varying robust weighted least squares method[J]. *Journal of Management Sciences in China*, 2024, 27(1): 141 – 158. (in Chinese)
- [41] 杨炳铎, 杨子晖, 陈海强. 带有泡沫与崩盘的可预测模型检验[J]. *管理科学学报*, 2023, 26(9): 110 – 124.
- Yang Bingduo, Yang Zihui, Chen Haiqiang. Testing predictable models with bubbles and crashes[J]. *Journal of Management Sciences in China*, 2023, 26(9): 110 – 124. (in Chinese)
- [42] 张 鹏, 党世力, 黄梅雨, 等. 基于机器学习预测股票收益率的两步骤 M-SV 投资组合优化[J]. *中国管理科学*, 2023, 31(12): 96 – 106.
- Zhang Peng, Dang Shili, Huang Meiyu, et al. Two-step M-SV portfolio optimization based on machine learning for stock return prediction[J]. *Chinese Journal of Management Science*, 2023, 31(12): 96 – 106. (in Chinese)
- [43] 康文津, 章 康. 中国 A 股市场流动性冲击与股票回报率关系研究[J]. *中国管理科学*, 2023, 31(7): 68 – 77.
- Kang Wenjin, Zhang Kang. Research on the relationship between liquidity shocks and stock returns in China's A-share market[J]. *Chinese Journal of Management Science*, 2023, 31(7): 68 – 77. (in Chinese)
- [44] Welch I, Goyal A. A comprehensive look at the empirical performance of equity premium prediction[J]. *The Review of Financial Studies*, 2008, 21(4): 1455 – 1508.
- [45] 俞庆进, 张 兵. 投资者有限关注与股票收益——以百度指数作为关注度的一项实证研究[J]. *金融研究*, 2012, (8): 152 – 165.
- Yu Qingjin, Zhang Bing. Limited investor attention and stock returns: An empirical study using Baidu Index as a measure of attention[J]. *Journal of Financial Research*, 2012, (8): 152 – 165. (in Chinese)
- [46] 赵龙凯, 陆子昱, 王致远. 众里寻“股”千百度——股票收益率与百度搜索量关系的实证探究[J]. *金融研究*, 2013, (4): 183 – 195.
- Zhao Longkai, Lu Ziyu, Wang Zhiyuan. Searching for “stocks” amidst a sea of information: An empirical exploration of the relationship between stock returns and Baidu search volume[J]. *Journal of Financial Research*, 2013, (4): 183 – 195. (in Chinese)
- [47] 罗 琦, 张志达, 吴希梅, 等. 股利情绪, 股利迎合与股价崩盘风险——基于百度指数平台搜索量的经验证据[J]. *管理科学学报*, 2023, 26(2): 87 – 103.
- Luo Qi, Zhang Zhida, Wu Ximei, et al. Dividend sentiment, dividend catering, and stock price crash risk: Empirical evidence based on Baidu Index search volume[J]. *Journal of Management Sciences in China*, 2023, 26(2): 87 – 103. (in Chinese)
- [48] Campbell J Y, Thompson S B. Predicting excess stock returns out of sample: Can anything beat the historical average? [J]. *The Review of Financial Studies*, 2008, 21(4): 1509 – 1531.
- [49] Pettenuzzo D, Timmermann A, Valkanov R. Forecasting stock returns under economic constraints[J]. *Journal of Financial Economics*, 2014, 114(3): 517 – 553.
- [50] Neely C J, Rapach D E, Tu J, et al. Forecasting the equity risk premium: The role of technical indicators[J]. *Management Science*, 2014, 60(7): 1772 – 1791.
- [51] Jiang F, Lee J, Martin X, et al. Manager sentiment and stock returns[J]. *Journal of Financial Economics*, 2019, 132(1): 126 – 149.
- [52] Cao C, Simin T, Xiao H. Predicting the equity premium with the implied volatility spread[J]. *Journal of Financial Markets*, 2020, (51): 100531.
- [53] Dangl T, Halling M. Predictive regressions with time-varying coefficients[J]. *Journal of Financial Economics*, 2012, 106(1): 157 – 181.
- [54] Wang Y, Hao X, Wu C. Forecasting stock returns: A time-dependent weighted least squares approach[J]. *Journal of Financial Markets*, 2021, (53): 100568.
- [55] Huang D, Jiang F, Tu J, et al. Investor sentiment aligned: A powerful predictor of stock returns[J]. *The Review of Financial Studies*, 2015, 28(3): 791 – 837.

- [56] Rapach D E, Ringgenberg M C, Zhou G. Short interest and aggregate stock returns[J]. *Journal of Financial Economics*, 2016, 121(1): 46–65.
- [57] Lin H, Wu C, Zhou G. Forecasting corporate bond returns with a large set of predictors: An iterated combination approach [J]. *Management Science*, 2018, 64(9): 4218–4238.
- [58] Clark T E, West K D. Approximately normal tests for equal predictive accuracy in nested models[J]. *Journal of Econometrics*, 2007, 138(1): 291–311.
- [59] 李 斌, 龙 真. 中国股票市场可预测性研究: 基于机器学习的视角[J]. *管理科学学报*, 2023, 26(10): 138–158.
Li Bin, Long Zhen. Research on the predictability of China's stock market: A machine learning perspective[J]. *Journal of Management Sciences in China*, 2023, 26(10): 138–158. (in Chinese)
- [60] Huang D, Jiang F, Li K, et al. Scaled PCA: A new approach to dimension reduction[J]. *Management Science*, 2022, 68(3): 1678–1695.
- [61] 隋 聪, 柳恩茂, 王文俊. 风险中性高阶矩期限结构及收益率的可预测性: 来自原油市场的证据[J]. *系统工程理论与实践*, 2025, 45(9): 2934–2949.
Sui Cong, Liu Enmao, Wang Wenjun. Risk-neutral high-order moment term structure and predictability of returns: Evidence from the crude oil market[J]. *Systems Engineering: Theory & Practice*, 2025, 45(9): 2934–2949. (in Chinese)
- [62] 吴冲锋, 林艳艳. 金融科技: 研究现状与展望[J]. *管理科学学报*, 2024, 27(6): 1–20.
Wu Chongfeng, Lin Yanyan. Financial technology: Research status and prospects[J]. *Journal of Management Sciences in China*, 2024, 27(6): 1–20. (in Chinese)

Financial modeling and empirical methods based on an improved Signal-to-Noise Ratio for prediction

LIN Yan-yan, ZHU Shun-wei, WU Chong-feng*

Antai College of Economics and Management, Shanghai Jiao Tong University, Shanghai 200030, China

Abstract: Considering the random factors involved in estimating model-fitting parameters, the classical linear regression model should be adjusted to improve its predictive performance. Therefore, this paper proposes an L-multiplier model based on Signal-to-Noise Ratio adjustment, which links the optimally fitted model to the best predictive model. The paper provides an explicit solution for the optimal L-multiplier, and considering the presence of autocorrelation in the predictive variables, presents an improved expression for L. Subsequently, this paper discusses two estimation methods for the L-multiplier and compares their respective advantages and disadvantages. Empirically, the theoretical validity of the L-multiplier model has been confirmed by applying the modified L-multiplier model to the problem of predicting stock returns. The results show that: 1) Compared to the baseline linear regression model, the model modified with the L-multiplier has higher predictive accuracy; 2) When the noise is higher, the historical sample size is smaller, or the information content of the predictive variable is lower, the adjustment intensity of the L-multiplier is greater, resulting in a more pronounced improvement in prediction; 3) For investors with mean-variance utility, this improvement in predictive performance can lead to an enhancement in investment returns. These findings also withstand a series of robust tests.

Key words: linear regression; noise; Signal-to-Noise Ratio; stock returns; prediction