Dissecting Time-varying Anomalies

Mingjun Sun*

November 12, 2026

Abstract

This paper presents a new methodology to study time-varying risk premia. Averaging Fama–MacBeth coefficient estimates over long windows can obscure changes in characteristic payoffs. Additionally, OLS becomes less reliable when dealing with high-dimensional characteristics. To address this, I propose the time-varying LASSO method that penalizes the L_1-norm of both the coefficients and their temporal differences to encourage both the smoothness and sparsity, while controlling over the length of the estimation window. An empirical analysis of U.S. equities demonstrates additional economic value of my proposed method, as evidenced by achieving higher out-of-sample R^2 and portfolio Sharpe ratios compared to both the Fama–MacBeth and LASSO regressions. The path of risk premia estimates exhibits higher degree of time-variation. The results indicate the method's ability to capture the time-varying dynamics of anomalies.

JEL Codes: C52, C55, C58, G17

Keywords: Asset pricing, Fama–MacBeth regression, Return prediction, Time-variation

*Tepper School of Business, Carnegie Mellon University, Pittsburgh, PA 15213, USA. Email: mingjuns@andrew.cmu.edu. I am indebted to my advisors, Burton Hollifield, Bryan Routledge, Patrick Blonien and Miguel Oliveira, for their invaluable guidance and support. For helpful comments, suggestions, and discussions, I thank Chester Spatt and seminar participants at the Carnegie Mellon University Finance Brownbag.

1 Introduction

The "factor zoo" described by Cochrane (2011) includes many factors and firm characteristics that appear to have significant explanatory power for the cross-section of expected returns. Identifying which of these anomalies explain returns remains a central question in empirical asset pricing. Examples include the book-to-market ratio, momentum, profitability, and firm investment.

Factor risk premia can change over time. For example, Ferson and Harvey (1991) and Ehsani and Linnainmaa (2022) document substantial variation in the premia associated with market capitalization and the market-to-book ratio. Changes in macroeconomic conditions can affect investors' risk aversion, while new information about future cash flows, interest rates, or economic conditions can alter expectations and risk premia. These changes motivate studying how anomaly payoffs evolve over time.

Estimating anomaly payoffs remains challenging despite the extensive literature. The two-pass regressions of Fama and MacBeth (1973) are widely used to analyze the cross-section of returns. In the first step, excess returns are regressed on anomaly variables by ordinary least squares in each period. In the second, the coefficient estimates are averaged over T periods and interpreted as characteristic payoffs or risk premia.

As Raponi and Zaffaroni (2023) emphasize, averaging cross-sectional estimates over T periods can obscure time variation in risk premia. Studies such as Lewellen (2015) and Green et al. (2017) set the smoothing window T to 10 years, producing estimates with little time variation. Averaging over a shorter window preserves more of this variation but reduces statistical power.

A high-dimensional set of characteristics also makes Fama–MacBeth estimates unreliable. Penalized regression addresses this problem by encouraging sparse coefficient estimates. LASSO, introduced by Tibshirani (1996), adds an L_1-norm penalty to the coefficients. Yet the expanding estimation windows used in studies such as Gu et al. (2020) and Freyberger et al. (2020) also produce estimates with little time variation. Figure 1 illustrates the coefficient paths for Fama–MacBeth and LASSO regressions under these conventional choices of T.

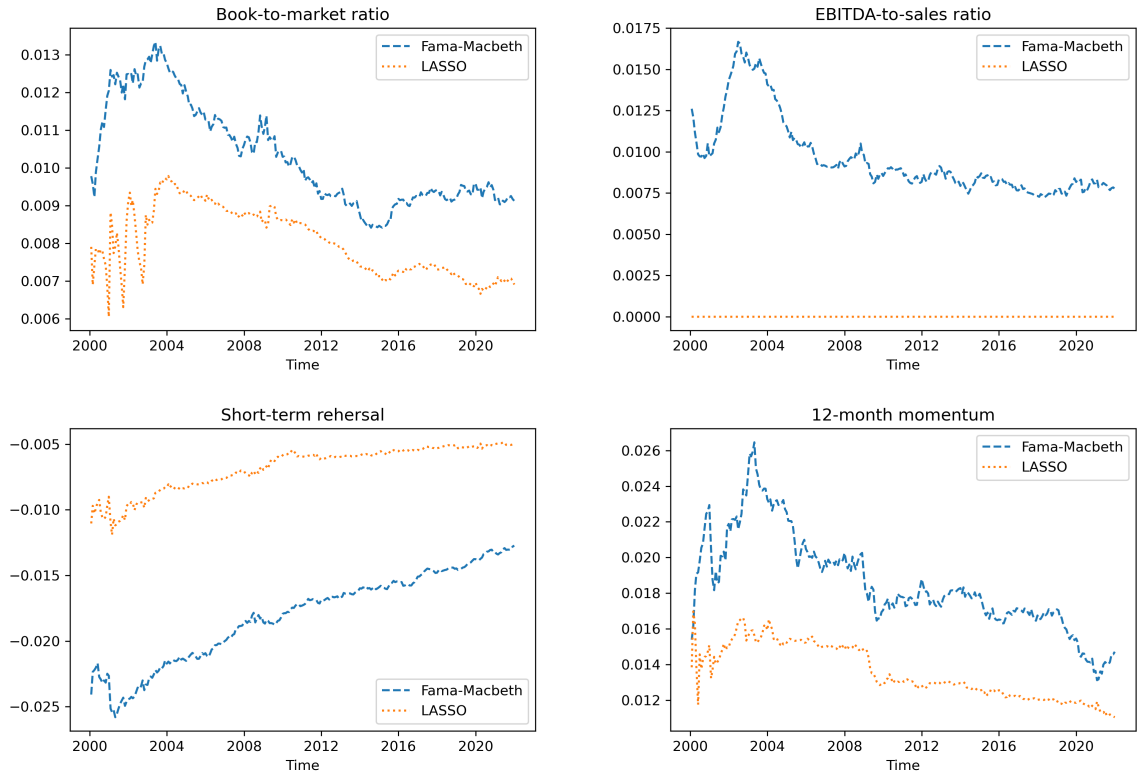


Figure 1: The figure shows the estimated risk premia from Fama–MacBeth and LASSO regressions in each training sample of the out-of-sample analysis. Fama–MacBeth uses a 10-year smoothing window; LASSO uses an expanding window. The sample period is January 1990 to December 2021, and the out-of-sample prediction period is January 2000 to December 2021.

I develop time-varying LASSO to estimate evolving risk premia. The method extends conventional LASSO by penalizing consecutive coefficient differences to control both smoothness and sparsity. The aim is to improve out-of-sample performance while using a shorter estimation window. The method has one cross-sectional parameter (λ_1) and two time-series parameters (λ_2 and T). λ_1 shrinks coefficient estimates toward zero. λ_2 penalizes consecutive coefficient differences, and T sets the estimation-window length.

The penalty λ_2 reflects the persistence of predictive relationships across adjacent estimation periods even as characteristic payoffs evolve. A shorter estimation window T emphasizes recent data but can yield noisy coefficient estimates. Penalizing changes from previous estimates incorporates information over time while allowing coefficients to respond to new observations. The window length controls data recency, and the temporal penalty controls coefficient stability.

I find higher out-of-sample R^2 and portfolio Sharpe ratios for time-varying LASSO than for the Fama–MacBeth and conventional LASSO benchmarks. The coefficient paths and selected predictors also exhibit greater time variation. A simulation study further shows that the predictive gains from temporal regularization depend on the coefficient dynamics, with improvements over conventional LASSO when characteristic payoffs are persistent but evolve over time.

Section 2 reviews the related literature. Section 3 presents the methodology, and Section 4 describes the data. Section 5 reports the empirical analysis of U.S. equities, Section 6 presents the simulations, and Section 7 concludes.

2 Literature Review

My estimator extends characteristic-based return forecasting by combining predictor selection with a penalty on changes in estimated characteristic payoffs. This connects Fama–MacBeth forecasting and penalized regressions with methods for estimating time-varying risk premia.

Fama–MacBeth regressions are widely used to analyze cross-sectional returns. Haugen and Baker (1996) were the first to use them to generate out-of-sample return forecasts. Green et al. (2017) test 94 firm characteristics and find that 8 to 12 are economically significant independent determinants of average returns. Han et al. (2024) extend Fama–MacBeth return forecasting by replacing OLS with machine learning methods in each cross-sectional regression.

The paper also relates to the literature on time-series return predictability. Early studies include Shiller (1981) and Fama and French (1988) for equities and Fama and Bliss (1987) for bonds. More recently, Gagliardini et al. (2016) estimate time-varying risk premia from a large unbalanced panel of individual stock returns. Bessembinder et al. (2023) document a strong correlation between the number of significant factors and both fluctuations in economic conditions and diversity in firm characteristics. Raponi and Zaffaroni (2023) propose a method for estimating time-varying anomalies in conditional asset pricing models. Haddad et al. (2020) argue that time variation in the top principal components can be identified through several of the factors themselves and develop an optimal factor-timing portfolio that combines cross-sectional and time-series return predictability.

Machine learning methods address the high dimensionality of characteristics in cross-sectional return models. Freyberger et al. (2020) propose a group LASSO procedure to select firm characteristics and nonparametrically estimate their effects on expected stock returns. Gu et al. (2020) document substantial economic gains from machine learning in return prediction. Kozak et al. (2020) use model selection to approximate the stochastic discount factor and the mean-variance efficient portfolio. Chen et al. (2024) use deep neural networks to estimate an asset pricing model for stock returns. Bryzgalova et al. (2025) use decision trees to group similar stocks through optimal portfolio splits and characterize the SDF. Kelly et al. (2019) apply PCA to stock returns projected on characteristics to derive a conditional multi-factor model. Lettau and Pelger (2020a) and Lettau and Pelger (2020b) use PCA to estimate and evaluate risk factors. Bryzgalova et al. (2023) estimate latent asset-pricing factors by incorporating portfolio targets in both the cross-sectional and time-series dimensions.

3 Methodology

Section 3.1 reviews pooled OLS and Fama–MacBeth regressions, Section 3.2 introduces LASSO, and Section 3.3 presents my time-varying estimator. Sections 3.4 and 3.5 describe parameter selection and performance evaluation.

Consider N firms and K observed factors. Let the excess return of firm i in month t be $r_{i,t}$. The characteristics[1] of firm i in month t are denoted by $f_{i,t}$, where $f_{i,t} \in \mathbb{R}^K$. In each month t, I model next-period excess return $r_{i,t+1}$ as

$$r_{i,t+1} = \alpha_t + \beta_t^\top f_{i,t} + \epsilon_{i,t}, \tag{1}$$

a linear function of firm characteristics with time-varying coefficients.

3.1 OLS and Fama–MacBeth

Fama–MacBeth estimation first regresses next-period excess returns $r_{i,t+1}$ on firm characteristics $f_{i,t}$ in each period t. It then averages the cross-sectional slope estimates over time to estimate characteristic risk premia.

The first-step regression includes all firm characteristics, as in Equation 2. The error term $\epsilon_{i,t}$ has mean zero, and the slope vector β_t represents characteristic payoffs.

$$r_{i,t+1} = \alpha_t + \beta_t^\top f_{i,t} + \epsilon_{i,t} \qquad \text{for } i = 1, \ldots, N_t \tag{2}$$

[1] They are also anomalies or factors.

To predict returns $\hat{r}_{i,t+1}$, I average the coefficient estimates from months $t-T$ through $t-1$, denoted by $\hat{\beta}_{t-T}$, ... $\hat{\beta}_{t-1}$, and apply them to characteristics observed at the end of month t, denoted by $f_{i,t}$, as in Equation 3:

$$\hat{r}_{i,t+1} = \overline{\alpha}_t + \overline{\beta}_t^\top f_{i,t} \qquad \text{for } i = 1, \ldots, N_t \tag{3}$$

where

$$\overline{\alpha}_t = \frac{1}{T} \sum_{j=t-T}^{j=t-1} \hat{\alpha}_j$$

$$\overline{\beta}_t = \frac{1}{T} \sum_{j=t-T}^{j=t-1} \hat{\beta}_j$$

T is the length of the smoothing window. Equation 3 is the Fama–MacBeth out-of-sample return prediction used in Haugen and Baker (1996), Lewellen (2015) and Green et al. (2017).

As a separate benchmark, I estimate pooled OLS using observations from the preceding T predictor months. For an end-of-month-t forecast of $r_{i,t+1}$, the estimation sample contains the pairs $(f_{i,s}, r_{i,s+1})$ for $s = t-T, \ldots, t-1$. The latest return used for estimation is therefore $r_{i,t}$. Let N_s be the number of available stock observations in predictor month s, and let $M_{t,T} = \sum_{s=t-T}^{t-1} N_s$ be the total number of pooled stock-month observations. Define the pooled mean squared error as

$$\mathcal{L}_{t,T}(\alpha, \beta) = \frac{1}{M_{t,T}} \sum_{s=t-T}^{t-1} \sum_{i=1}^{N_s} \left(r_{i,s+1} - \alpha - \beta^\top f_{i,s}\right)^2.$$

Pooled OLS estimates a common intercept and coefficient vector for this window, whereas the Fama–MacBeth forecast averages estimates from separate monthly regressions. The pooled OLS estimates solve

$$(\hat{\alpha}_t, \hat{\beta}_t) = \arg\min_{\alpha,\beta} \mathcal{L}_{t,T}(\alpha, \beta). \tag{4}$$

The pooled OLS return prediction $\hat{r}_{i,t+1}$ combines the estimated intercept $\hat{\alpha}_t$ and coefficient vector $\hat{\beta}_t$ with characteristics observed at the end of month t, denoted by $f_{i,t}$, as in Equation 5:

$$\hat{r}_{i,t+1} = \hat{\alpha}_t + \hat{\beta}_t^\top f_{i,t} \tag{5}$$

Studies such as Lewellen (2015) and Green et al. (2017)[2] set T to 10 years. Figure 1 shows that the resulting risk-premium estimates vary little over time.

[2] Green et al. (2017) estimate regressions using a 120-month window and apply the resulting mean coefficient estimates to the 94 firm characteristics observed at the end of the month when forecasts are formed.

3.2 LASSO

OLS minimizes the pooled mean squared error in Equation 4. With many characteristics, Fama–MacBeth and OLS estimates become unreliable. LASSO addresses this problem through an L_1-norm penalty that encourages sparsity and performs variable selection. Equivalently, it imposes an L_1-norm restriction on the regression coefficients:

$$\sum_{k=1}^{K} |\beta_k| \leq C \tag{6}$$

where C is a constant. Smaller values of C produce sparser coefficient estimates.

Using the same pooled mean squared error, LASSO estimates the intercept and coefficient vector by solving

$$(\hat{\alpha}_t, \hat{\beta}_t) = \arg\min_{\alpha,\beta} \left\{ \mathcal{L}_{t,T}(\alpha, \beta) + \lambda_1 \|\beta\|_1 \right\}. \tag{7}$$

The L_1 penalty encourages sparse coefficient estimates; the intercept is not penalized. I select λ_1 by validation as described in Section 3.4. Return forecasts use the fitted intercept and coefficient vector as in Equation 5.

The usual approach pools data from $t-T$ to $t-1$ to estimate risk premia for period t. In Gu et al. (2020)[3] and Freyberger et al. (2020), T expands to include all available data. Han et al. (2024) show that the out-of-sample performance of their smoothed LASSO approach consistently improves as T increases. As with Fama–MacBeth regressions, however, long estimation windows produce risk-premium estimates with little time variation.

3.3 Time-varying LASSO

Previous applications of Fama–MacBeth and LASSO regressions have focused mainly on cross-sectional prediction. The conventional choices of $T = 120$ for Fama–MacBeth and an expanding window for LASSO yield risk-premium estimates with little time variation. Yet several studies (Gagliardini et al. (2016), Bessembinder et al. (2023) and Raponi and Zaffaroni (2023)) document substantial variation in the factor zoo over time. Long estimation windows can pool observations from different predictive regimes

[3] In Gu et al. (2020), the initial training period is 1957–1974, the validation period is 1975–1986, and the test period is 1987–2016. The training sample expands by one year at each annual refit, while the 12-year validation window advances by 12 months.

and obscure changes in characteristic payoffs. My objective is to retain responsiveness to recent data while controlling estimation noise.

Characteristic payoffs may vary over time, but allowing coefficients to change independently in every period can lead to imprecise estimates. I therefore penalize changes in coefficients across adjacent estimation periods. This regularization allows characteristic payoffs to evolve while borrowing information from previous estimates. The underlying assumption is persistence in predictive relationships across adjacent periods.

I implement this idea by combining an L_1 penalty on coefficient levels with an L_1 penalty on changes from the preceding estimate. The first encourages sparsity across characteristics; the second shrinks individual coefficient changes toward zero. The temporal penalty can leave some coefficients unchanged while allowing others to adjust when the improvement in fit justifies the penalty.

The resulting estimator solves[4]

$$(\hat{\alpha}_t, \hat{\beta}_t) = \arg\min_{\alpha,\beta} \big\{ \mathcal{L}_{t,T}(\alpha, \beta) + \lambda_1 \|\beta\|_1 + \lambda_2 \|\beta - \hat{\beta}_{t-1}\|_1 \big\}. \tag{8}$$

Here $\hat{\beta}_{t-1}$ is the coefficient vector estimated at the preceding forecast date and is held fixed in the current optimization. As in the LASSO objective, the loss is averaged over the pooled stock-month observations and the intercept is not penalized. I obtain the next-period return forecast from

$$\hat{r}_{i,t+1} = \hat{\alpha}_t + \hat{\beta}_t^\top f_{i,t} \tag{9}$$

The estimator has three parameters: λ_1, λ_2, and T. λ_1 controls cross-sectional shrinkage. Increasing λ_1 shrinks coefficient estimates toward zero.

The parameters T and λ_2 play distinct roles in the time-series dimension. The window length T determines how much historical data enter the current fit, whereas λ_2 determines how strongly the current coefficients are pulled toward the previous estimates. A shorter window emphasizes recent observations but can yield noisier estimates; temporal regularization can stabilize these estimates when predictive relationships persist. This creates a tradeoff between stability and responsiveness: a large λ_2 may delay

[4] Tibshirani et al. (2005) introduce fused LASSO, which penalizes coefficient levels and their successive differences. Yogatama et al. (2011) use time-series regularization to capture first-order temporal effects.

adaptation to abrupt changes. I therefore tune λ_2 by validation for each candidate T, as described below.

When $\lambda_2 = 0$, my approach reduces to LASSO. When $\lambda_1 = \lambda_2 = 0$, it reduces to the pooled OLS benchmark in Section 3.1. This unpenalized estimator fits a common coefficient vector over the estimation window and generally differs from the Fama–MacBeth forecast, which averages coefficients from separate monthly regressions.

3.4 Sample splitting and tuning via validation

I evaluate Fama–MacBeth, pooled OLS, and conventional LASSO over the window grid $T \in \{1, 3, 6, 12, 36, 60, 84, 96, 102, 108, 120\}$. Table 1 reports the Fama–MacBeth and pooled OLS comparisons for these windows except $T = 36$, while Tables 2 and 3 report the full conventional LASSO grid. Time-varying LASSO uses the separate grid $T \in \{1, 6, 12, 24, 60, 72, 96, 120\}$ reported in Section 5.3.

For the two LASSO estimators, I randomly assign 80% of the pooled stock-month observations to a training sample and 20% to a validation sample at each forecast month t and fixed T. These assignments are redrawn each month as the estimation window advances by one month. All returns in these samples are observed by the forecast date. The loss used to fit each candidate model is the mean squared error over its training observations.

For conventional LASSO, I select λ_1 to maximize validation-sample R^2. For time-varying LASSO, I select the pair (λ_1, λ_2) to maximize the same criterion. I then refit the model on all observations in the T-month window and combine the fitted coefficients with characteristics at the end of month t to forecast returns in month $t + 1$. This selection and refitting procedure is repeated each month.

Window lengths are held fixed within each forecasting exercise. Statements about the highest test R^2 or Sharpe ratio across T are ex post comparisons of these exercises; validation selects the penalty parameters conditional on T.

3.5 Performance Evaluation

I evaluate predictive performance using out-of-sample R^2, portfolio returns, and Sharpe ratios.

3.5.1 Individual stock-level analysis

Let $\hat{r}_{i,t+1}$ denote the predicted return of firm i in period $t+1$. I measure excess-return predictive performance using out-of-sample R^2:

$$R^2_{oos} = 1 - \frac{\sum_{(i,t)}(r_{i,t+1} - \hat{r}_{i,t+1})^2}{\sum_{(i,t)} r^2_{i,t+1}} \tag{10}$$

This measure of R^2 pools observations across firms and time. Following Gu et al. (2020), I do not demean excess returns in the denominator. Historical mean returns are noisy and typically forecast future excess returns less accurately than a forecast of zero, making the historical mean an easier benchmark to outperform. I therefore measure R^2 relative to a zero-return forecast.

3.5.2 Portfolio Forecast

Long-short portfolio performance measures the economic value of cross-sectional forecasts. At the end of each month, I sort stocks into 10 deciles by their 1-month-ahead out-of-sample predicted excess returns. Let R_{top} denote the portfolio of stocks in the highest predicted-return decile and R_{bottom} the portfolio in the lowest. The long-short portfolio return R_{ls} is

$$R_{ls} = R_{top} - R_{bottom}$$

I construct value-weighted long-short and long-only portfolios. The long-short portfolio buys stocks in decile 10 and sells stocks in decile 1; the long-only portfolio holds the highest predicted-return stocks, weighted by market capitalization. I also construct equal-weighted versions of these portfolios. For each portfolio, I report average excess returns and Sharpe ratios.

3.5.3 Time-varying risk premia

I examine the paths of estimated risk premia over time.

3.5.4 Time-varying model complexity

Model complexity is the number of factors with nonzero coefficients. I compare its variation over time across models.

4 Data

Jensen et al. (2023) provide a global data set of 406 characteristics and their associated factor portfolios. I use U.S. equities and a subset of 121 characteristics from Jensen et al. (2023).

4.1 Data Preparation

I apply the following filters and transformations.

First, I exclude observations with missing market equity in month t and missing returns in month $t+1$. Second, I exclude stocks in the bottom 1% of market equity to limit the influence of very small, illiquid stocks. Following Green et al. (2017), I replace missing characteristic values with the cross-sectional median of the available observations for that characteristic and month.

Finally, following Freyberger et al. (2020) and Kozak et al. (2020), I rank-transform each characteristic. For characteristic i of stock s at time t, denoted by $f^i_{s,t}$, I rank stocks cross-sectionally by $f^i_{s,t}$ from 1 to n_t, where n_t is the number of stocks at time t with an available value. I then divide each rank by $n_t + 1$:

$$Rank\ Transform^i_{s,t} = \frac{rank(f^i_{s,t})}{n_t + 1}$$

4.2 Descriptive Statistics

Section 8.1 in the Appendix lists and describes the included anomalies.

The sample covers January 1990 to December 2021. The first ten years form the initial estimation sample, and the out-of-sample prediction period runs from January 2000 to December 2021. This period includes the Great Financial Crisis recession, from December 2007 to June 2009, and the COVID-19 recession, from February 2020 to April 2020. The data contain 1,540,526 firm-level observations.

Following Hou et al. (2020), I group the 121 characteristics into six categories: momentum, including momentum over 1-12 months (ret_1_12) and short-term reversal (r_1_0); investment, including 1-year inventory change (inv_gr1a); profitability, including return on operating assets ($ebit_dev$); intangibles, including total accruals ($taccruals_at$) and tangibility ($tangibility$); value, including the book-to-market ratio (be_me) and market equity ($market_equity$); and trading frictions, including 126-day share turnover ($turnover_126d$).

5 An Empirical Study of U.S. Equities

I compare the out-of-sample return predictions and portfolio performance of time-varying LASSO with the Fama–MacBeth and conventional LASSO benchmarks using the U.S. equity sample described in Section 4. I then examine how selected predictors and estimated characteristic payoffs evolve over time.

5.1 OLS and Fama–MacBeth

I compare Fama–MacBeth and pooled OLS forecasts across a grid of values for T. For each T, I apply both methods. For Fama–MacBeth, I first estimate β_t in each month from January 2000 to December 2021. Each month's cross-sectional regression uses only data from month t to estimate β_t. For the chosen T, I then average estimates from months $t-T$ to $t-1$ to predict excess returns in month $t+1$.

For pooled OLS, I estimate β_t in each month from January 2000 to December 2021. In month t, I pool data from months $t-T$ to $t-1$ to estimate β_t and use the fitted coefficients $\hat{\beta}_t$ to predict excess returns in month $t+1$.

Table 1: Out-of-sample monthly stock-level prediction performance (percentage R^2_{OOS})

Methods \ T	1	3	6	12	60	84	96	102	108	120
Fama–MacBeth	-21.51	-8.74	-4.53	-2.32	-0.15	0.03	0.20	0.25	0.16	0.17
OLS	-86.84	-31.32	-15.85	-7.80	-0.94	-0.56	-0.30	-0.21	-0.29	-0.24

The table reports out-of-sample percentage R^2 for Fama–MacBeth and pooled OLS forecasts at different window lengths T. The sample period is January 2000 to December 2021. Pooled OLS estimates β_t in each month from January 2000 to December 2021. For each month t, data from months $t-T$ to $t-1$ are pooled to estimate β_t. The fitted coefficients $\hat{\beta}_t$ predict excess returns in month $t+1$. Fama–MacBeth estimates β_t in each month from January 2000 to December 2021, using only data from month t to estimate β_t. For each T, the estimates from months $t-T$ to $t-1$ are averaged to predict excess returns in month $t+1$.

Table 1 reports out-of-sample R^2 for pooled OLS and Fama–MacBeth forecasts across estimation-window lengths. Both benchmarks attain their highest reported R^2 at $T = 102$. This window also yields higher R^2 than the 120-month window used in earlier applications.[5]

Fama–MacBeth forecasts achieve higher reported out-of-sample R^2 than pooled OLS across the windows with $T > 1$ in Table 1. The two procedures aggregate information

[5] See Green et al. (2017) and Lewellen (2015).

across months differently.

5.2 LASSO

I evaluate LASSO forecasts over a grid of λ_1 and T. λ_1 controls shrinkage, and T is the number of months used for estimation. Larger values of T reduce temporal variation in the coefficients. I apply the following procedure to each pair of λ_1 and T.

I estimate β_t using Equation 7 in each month from January 2000 to December 2021. In month t, I pool data from months $t-T$ to $t-1$ to estimate β_t. The fitted coefficients $\hat{\beta}_t$ predict excess returns in month $t+1$ through Equation 5. The prediction exercise runs from January 2000 to December 2021. It begins in January 2000 because the maximum T is 120 months. When $T = 120$, the first coefficient estimates use pooled data from January 1990 to December 1999 and are combined with January 2000 characteristics to predict February 2000 excess returns.

Table 2: Out-of-sample monthly stock-level prediction performance (percentage R^2_{OOS})

λ_1 \ T	1	3	6	12	36	60	84	96	102	108	120
0	-86.84	-31.32	-15.85	-7.80	-2.04	-0.94	-0.56	-0.30	-0.21	-0.29	-0.24
0.0001	-60.26	-22.60	-11.19	-5.20	-1.19	-0.40	-0.13	0.09	0.17	0.09	0.10
0.0002	-49.05	-18.47	-9.03	-4.10	-0.89	-0.21	-0.02	0.19	0.26	0.16	0.18
0.0003	-42.52	-15.95	-7.75	-3.50	-0.74	-0.14	0.01	0.20	0.27	0.16	0.19
0.0004	-38.11	-14.22	-6.89	-3.12	-0.67	-0.11	0.01	0.19	0.26	0.15	0.18
0.0005	-34.86	-12.97	-6.25	-2.82	-0.63	-0.10	0.01	0.19	0.26	0.15	0.18
0.0006	-32.30	-11.98	-5.77	-2.62	-0.60	-0.08	0.01	0.19	0.26	0.15	0.18
0.0007	-30.27	-11.19	-5.40	-2.46	-0.57	-0.07	0.01	0.18	0.25	0.14	0.17
0.0008	-28.60	-10.54	-5.11	-2.34	-0.55	-0.07	0.01	0.18	0.25	0.13	0.16
0.0009	-27.21	-10.00	-4.86	-2.23	-0.54	-0.06	0.00	0.18	0.24	0.12	0.16
0.001	-26.02	-9.55	-4.66	-2.13	-0.52	-0.07	0.01	0.17	0.23	0.11	0.15
0.0011	-24.98	-9.18	-4.47	-2.04	-0.51	-0.07	0.01	0.17	0.22	0.11	0.14
0.0012	-24.06	-8.85	-4.31	-1.95	-0.49	-0.07	0.01	0.17	0.22	0.10	0.13
0.0013	-23.21	-8.57	-4.17	-1.88	-0.48	-0.07	0.01	0.16	0.21	0.10	0.12
0.0014	-22.46	-8.30	-4.04	-1.82	-0.47	-0.07	0.00	0.15	0.21	0.10	0.11
0.0015	-21.77	-8.07	-3.93	-1.77	-0.46	-0.08	0.00	0.15	0.20	0.10	0.11
0.01	-9.95	-4.52	-2.34	-1.04	-0.31	-0.10	-0.01	0.12	0.17	0.08	0.09
1	-8.06	-3.90	-2.06	-1.03	-0.31	-0.10	-0.01	0.12	0.17	0.08	0.09

The table reports the out-of-sample percentage R^2 of LASSO regressions with different values of λ_1 and T. λ_1 shrinks characteristic coefficients toward zero. T is the window length: in month t, I pool data from months $t-T$ to $t-1$ to estimate β_t. The sample period is January 1990 to December 2021, and the out-of-sample prediction period is January 2000 to December 2021.

Table 2 shows that $T = 102$ yields the highest reported out-of-sample R^2 for each fixed value of λ_1. The largest value is 0.27%, attained at $T = 102$ and $\lambda_1 = 0.0003$.

Table 3: Out-of-sample value-weighted portfolio performance (Sharpe ratio)

λ_1 \ T	1	3	6	12	36	60	84	96	102	108	120
0	0.12	0.11	0.11	0.10	0.14	0.18	0.18	0.21	0.21	0.21	0.20
0.0001	0.13	0.12	0.11	0.10	0.16	0.19	0.21	0.22	0.22	0.23	0.22
0.0002	0.14	0.12	0.12	0.11	0.16	0.19	0.20	0.22	0.22	0.22	0.22
0.0003	0.14	0.13	0.12	0.11	0.16	0.18	0.20	0.21	0.22	0.21	0.21
0.0004	0.14	0.13	0.13	0.11	0.15	0.17	0.18	0.20	0.21	0.20	0.19
0.0005	0.14	0.13	0.13	0.11	0.15	0.17	0.17	0.20	0.21	0.19	0.19
0.0006	0.14	0.13	0.13	0.11	0.14	0.17	0.16	0.19	0.20	0.19	0.17
0.0007	0.14	0.13	0.13	0.11	0.14	0.16	0.16	0.18	0.20	0.19	0.16
0.0008	0.14	0.13	0.13	0.12	0.14	0.16	0.15	0.18	0.19	0.17	0.15
0.0009	0.14	0.14	0.13	0.12	0.14	0.16	0.14	0.16	0.18	0.15	0.17
0.001	0.14	0.14	0.13	0.12	0.14	0.16	0.13	0.17	0.17	0.15	0.14
0.0011	0.14	0.14	0.13	0.12	0.14	0.15	0.13	0.17	0.15	0.14	0.15
0.0012	0.14	0.14	0.13	0.12	0.13	0.15	0.14	0.17	0.16	0.14	0.15
0.0013	0.14	0.14	0.13	0.12	0.12	0.15	0.13	0.16	0.15	0.15	0.14
0.0014	0.14	0.13	0.13	0.12	0.11	0.14	0.13	0.15	0.15	0.13	0.13
0.0015	0.14	0.14	0.13	0.12	0.10	0.13	0.13	0.13	0.15	0.14	0.15
0.01	0.12	0.08	0.09	0.11	0.14	0.13	0.13	0.14	0.13	0.13	0.13
1	0.14	0.14	0.14	0.14	0.14	0.14	0.13	0.12	0.13	0.13	0.12

The table reports out-of-sample monthly Sharpe ratios for portfolios formed from LASSO forecasts at different values of λ_1 and T. λ_1 shrinks characteristic coefficients toward zero. T is the window length: in month t, I pool data from months $t-T$ to $t-1$ to estimate β_t. The sample period is January 1990 to December 2021, and the out-of-sample prediction period is January 2000 to December 2021.

At the end of each month, I sort stocks into deciles by predicted returns $\hat{r}_{i,t+1}$. The value-weighted portfolio buys the highest predicted-return stocks (decile 10) and sells the lowest (decile 1). Table 3 reports out-of-sample monthly Sharpe ratios across combinations of λ_1 and T. For a fixed λ_1, larger values of T correspond to higher out-of-sample Sharpe ratios.

5.3 Time-varying LASSO

I evaluate time-varying LASSO over a grid of λ_1, λ_2, and T. λ_1 shrinks coefficients toward zero, λ_2 penalizes changes between consecutive coefficient estimates, and T sets the window length. To limit computation time, I use $T \in \{1, 6, 12, 24, 60, 72, 96, 120\}$. I apply the following procedure to each combination of λ_1, λ_2, and T.

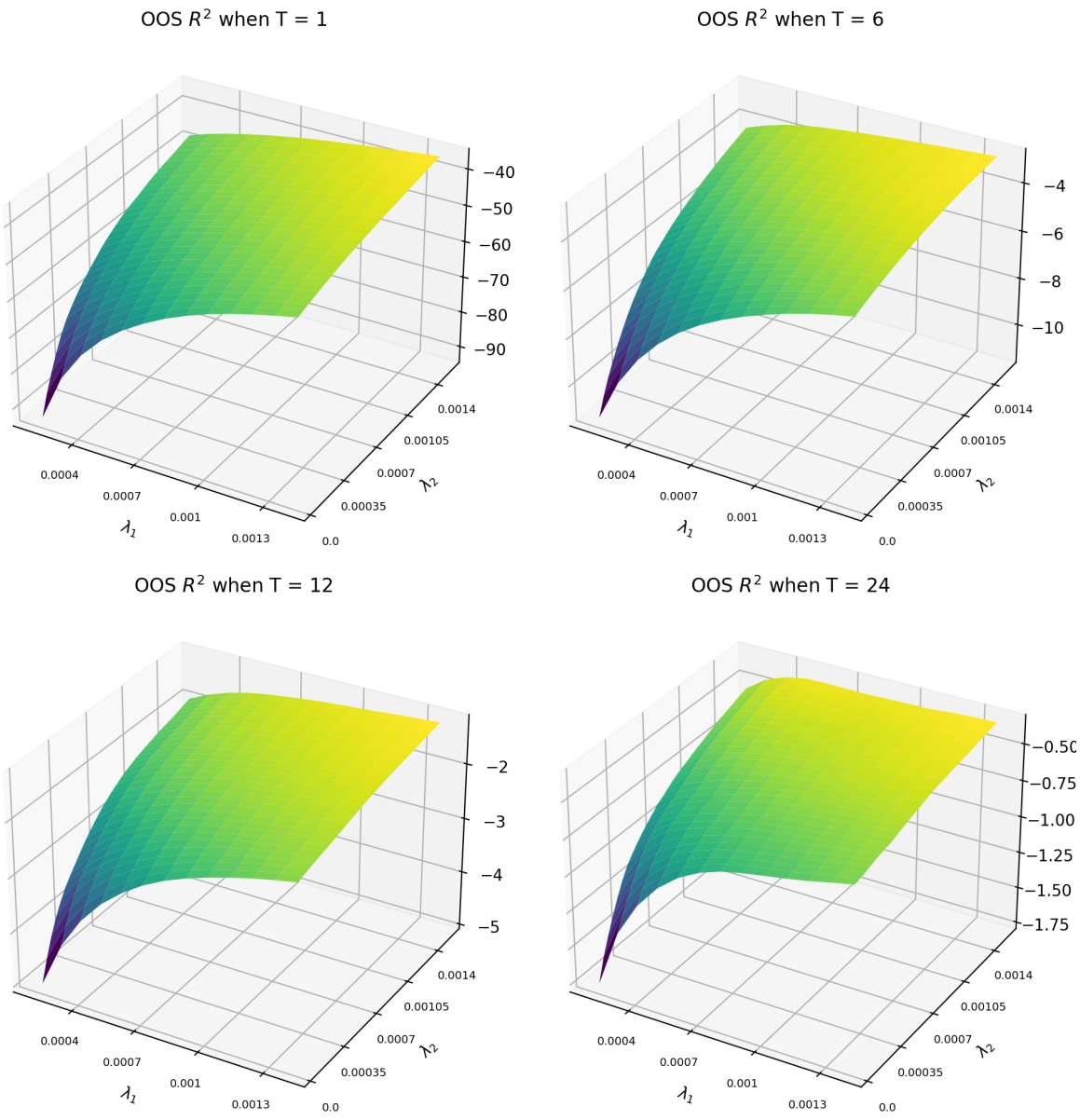


Figure 2: The figure plots the out-of-sample percentage R^2 as a function of λ_1 and λ_2 when $T = 1$, 6, 12 and 24 for time-varying LASSO. λ_1 shrinks characteristic coefficients toward zero, while λ_2 penalizes changes between consecutive coefficient estimates. T is the window length: in month t, I pool data from months $t-T$ to $t-1$ to estimate β_t. The sample period is January 1990 to December 2021, and the out-of-sample prediction period is January 2000 to December 2021.

For the initial month, January 2000, I estimate coefficients using pooled data from January 1990 to December 1999. For each month t from February 2000 onward, I estimate coefficients using pooled data from months $t-T$ to $t-1$, as in Equation 8.

I use the fitted coefficients $\hat{\beta}_t$ to predict excess returns in month $t+1$ through

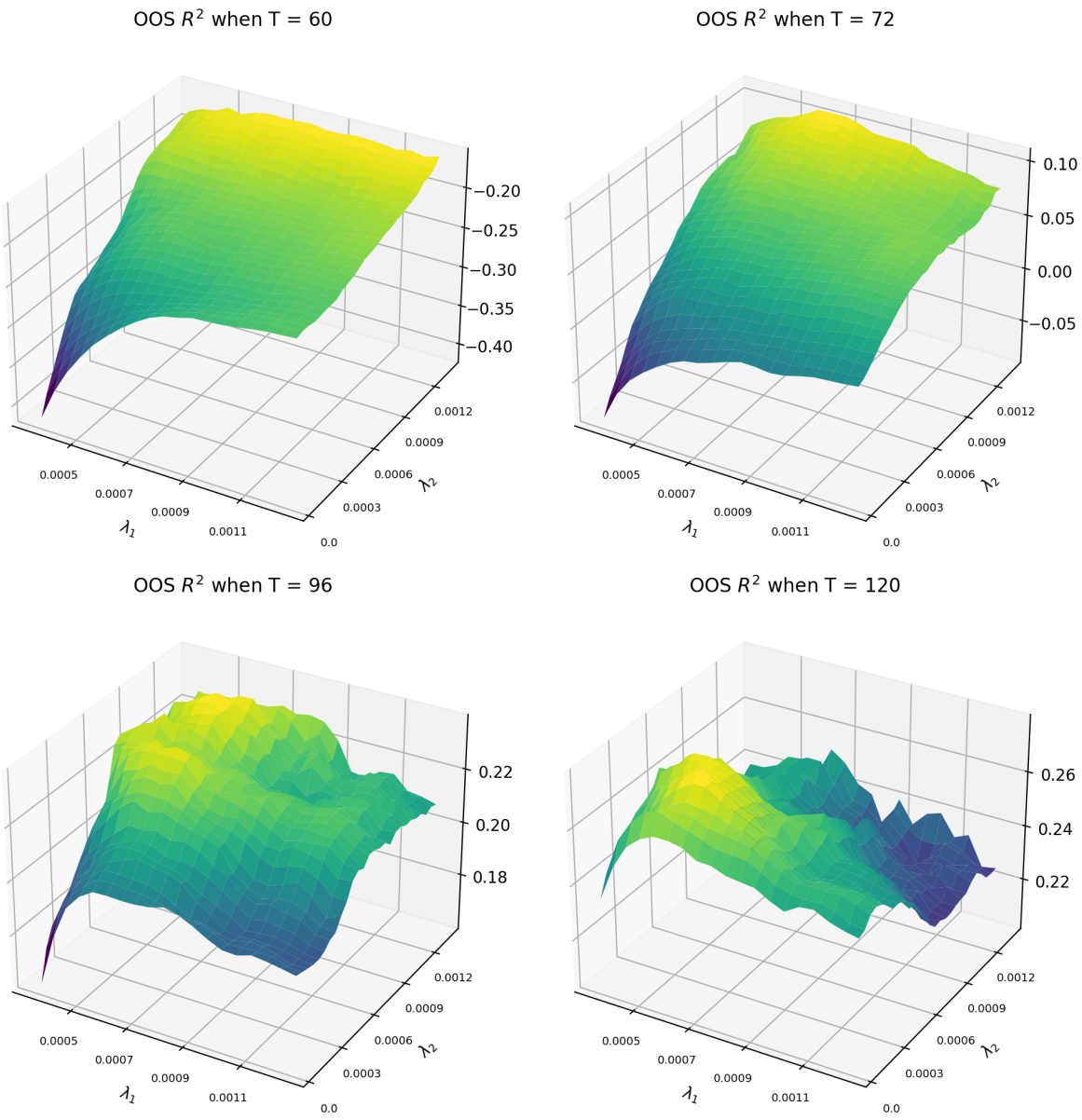


Figure 3: The figure plots the out-of-sample percentage R^2 as a function of λ_1 and λ_2 when $T = 60$, 72, 96 and 120 for time-varying LASSO. λ_1 shrinks characteristic coefficients toward zero, while λ_2 penalizes changes between consecutive coefficient estimates. T is the window length: in month t, I pool data from months $t-T$ to $t-1$ to estimate β_t. The sample period is January 1990 to December 2021, and the out-of-sample prediction period is January 2000 to December 2021.

Equation 9. The prediction exercise runs from January 2000 to December 2021. At the end of each month, I sort stocks into deciles by predicted returns $\hat{r}_{i,t+1}$ and form a value-weighted portfolio that buys the highest predicted-return stocks (decile 10) and sells the lowest (decile 1).

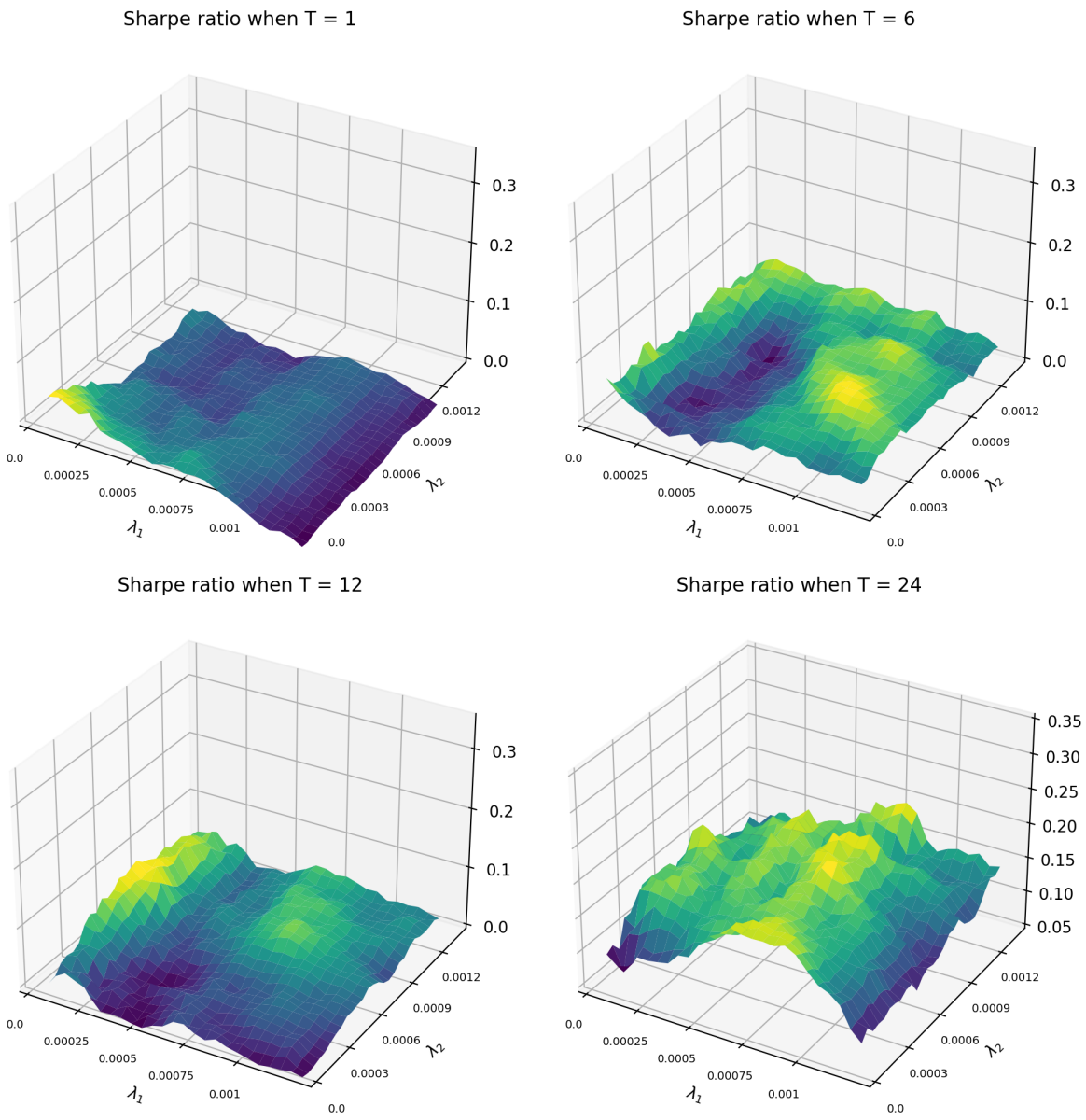


Figure 4: The figure plots value-weighted portfolio Sharpe ratios for time-varying LASSO as a function of λ_1 and λ_2 when $T = 1$, 6, 12, and 24. λ_1 shrinks characteristic coefficients toward zero, while λ_2 penalizes changes between consecutive coefficient estimates. T is the window length: in month t, I pool data from months $t-T$ to $t-1$ to estimate β_t. The sample period is January 1990 to December 2021, and the out-of-sample prediction period is January 2000 to December 2021.

The out-of-sample R^2 comparisons vary λ_1, λ_2, and T. Figure 2 reports R^2 across λ_1 and λ_2 for $T = 1, 6, 12$ and 24. Figure 3 reports R^2 across λ_1 and λ_2 for $T = 60, 72, 96$ and 120.

For a fixed λ_1, varying λ_2 substantially affects out-of-sample R^2. Setting λ_2 to 0 re-

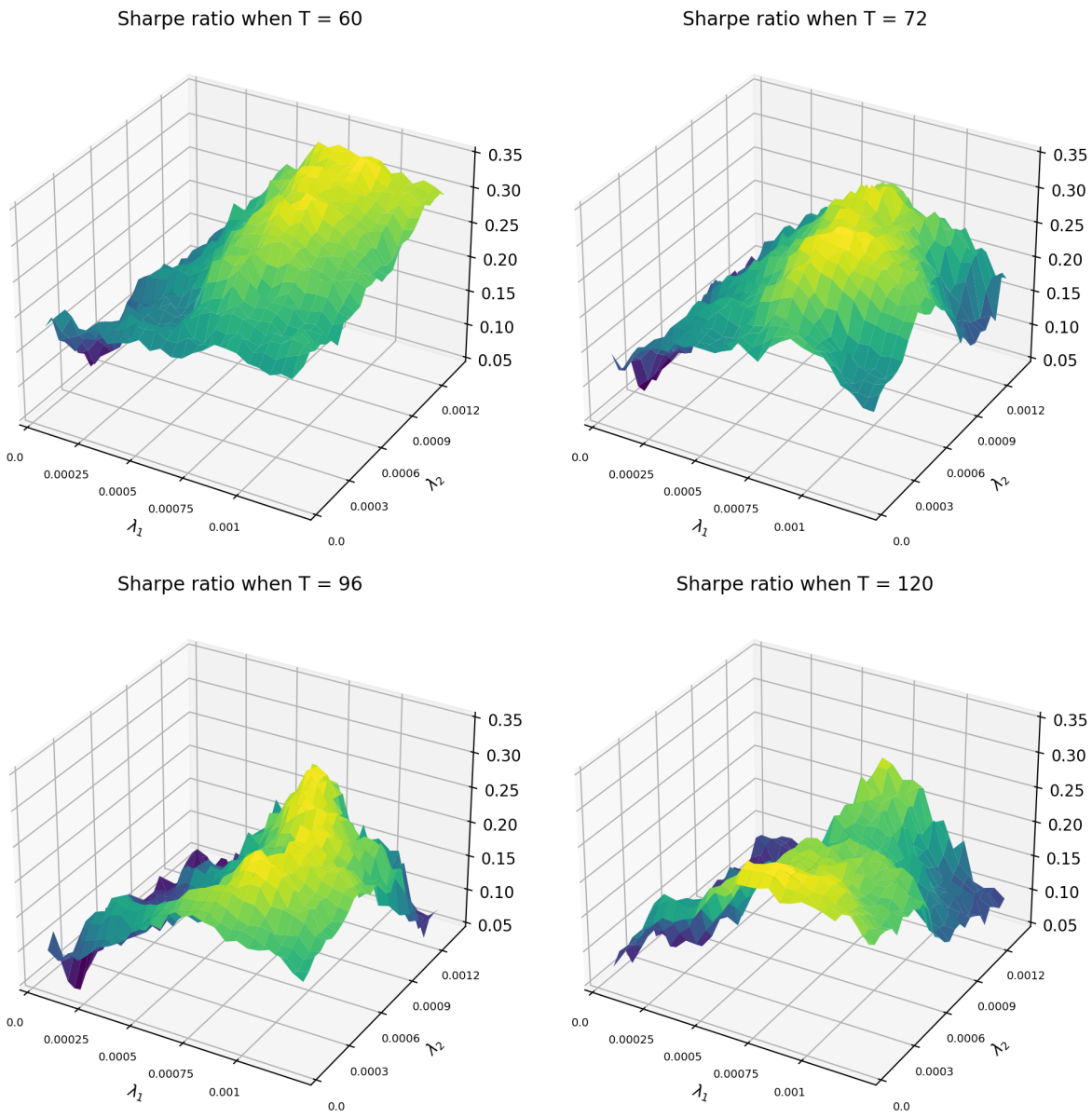


Figure 5: The figure plots value-weighted portfolio Sharpe ratios for time-varying LASSO as a function of λ_1 and λ_2 when $T = 60$, 72, 96 and 120. λ_1 shrinks characteristic coefficients toward zero, while λ_2 penalizes changes between consecutive coefficient estimates. T is the window length: in month t, I pool data from months $t-T$ to $t-1$ to estimate β_t. The sample period is January 1990 to December 2021, and the out-of-sample prediction period is January 2000 to December 2021.

covers conventional LASSO. Across estimation-window lengths T, time-varying LASSO attains higher out-of-sample R^2 than the Fama–MacBeth and conventional LASSO benchmarks. The penalty λ_2 shrinks consecutive coefficient differences. These differences tend to decrease as T increases, so the corresponding optimal λ_2 also decreases.

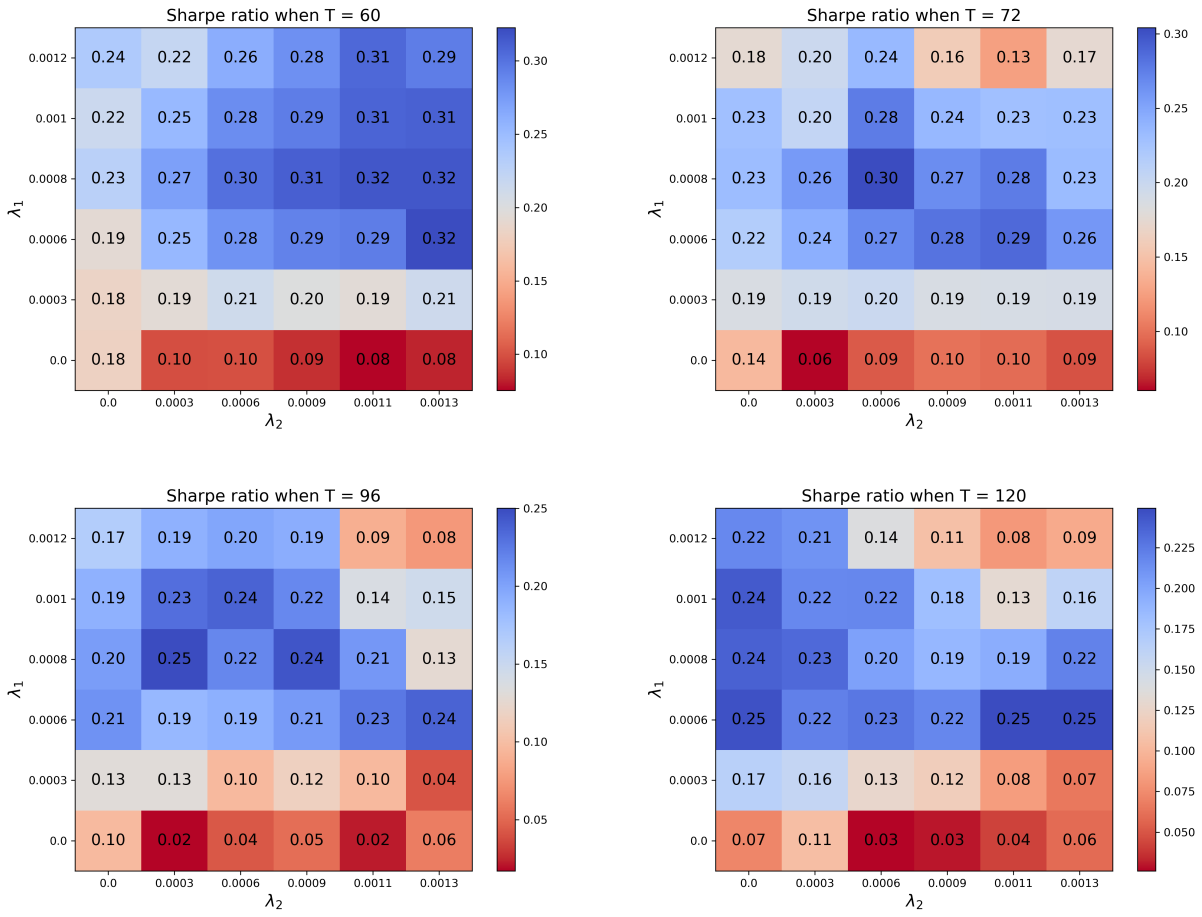


Figure 6: The figure shows Sharpe ratios of value-weighted portfolios formed from time-varying LASSO forecasts across penalty parameters λ_1 and λ_2 for different values of T. λ_1 shrinks characteristic coefficients toward zero, while λ_2 penalizes changes between consecutive coefficient estimates. T is the window length: in month t, I pool data from months $t-T$ to $t-1$ to estimate β_t. The sample period is January 1990 to December 2021, and the out-of-sample prediction period is January 2000 to December 2021.

I report out-of-sample Sharpe ratios and mean excess returns for value-weighted and equal-weighted portfolios across combinations of λ_1, λ_2, and T. Figure 4 plots value-weighted Sharpe ratios across λ_1 and λ_2 for $T = 1, 6, 12$ and 24. Figure 5 plots the same measure across λ_1 and λ_2 for $T = 60, 72, 96$ and 120. Figure 13 reports equal-weighted Sharpe ratios across λ_1 and λ_2 for $T = 60, 72, 96$ and 120. Figure 6 presents Figure 5 as a heatmap.

Figure 7 plots out-of-sample mean excess returns for value-weighted portfolios across λ_1 and λ_2 for $T = 1, 6, 12$ and 24. Figure 8 reports the same measure across λ_1 and λ_2 for $T = 60, 72, 96$ and 120. Figure 9 presents Figure 8 as a heatmap.

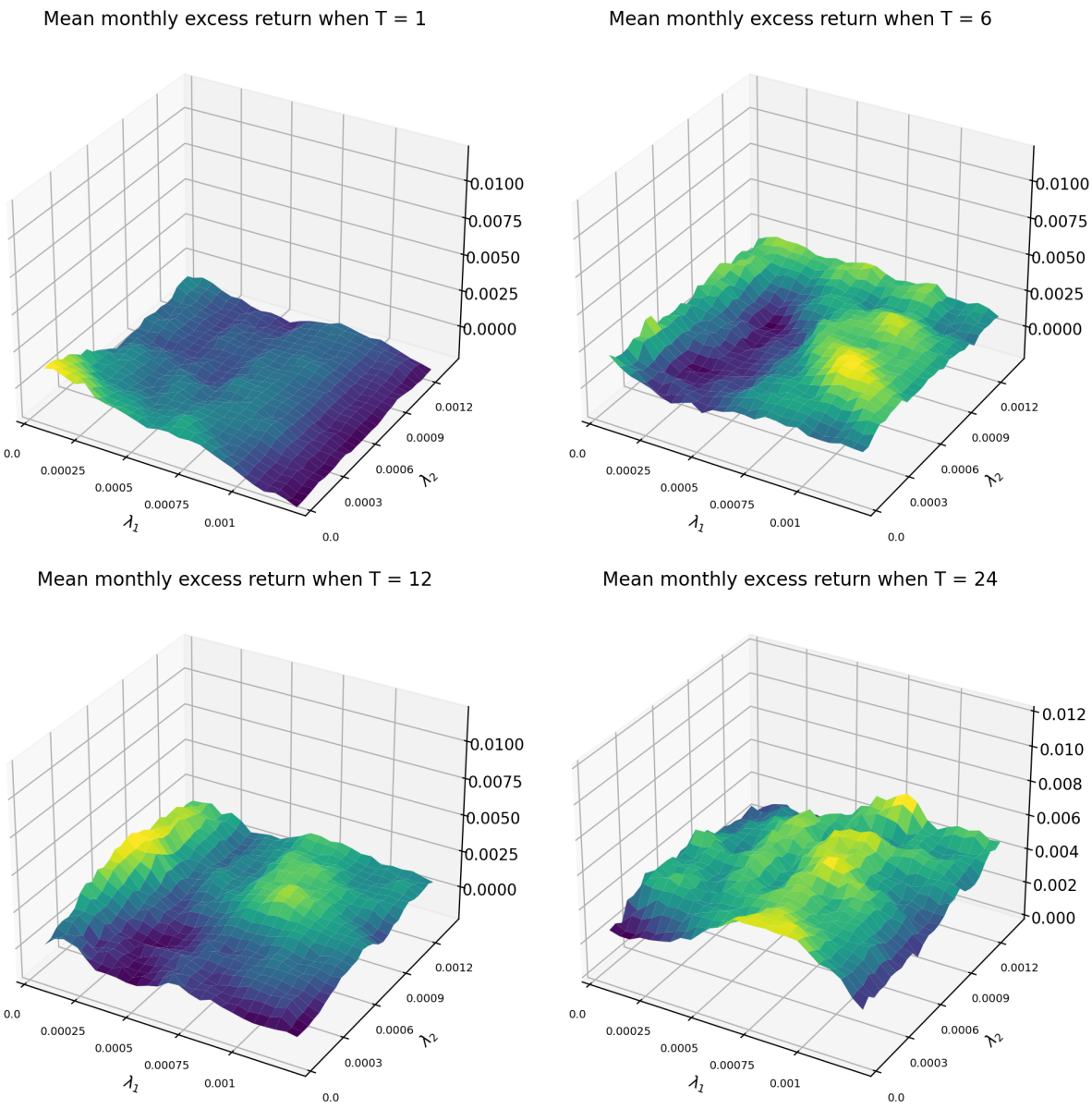


Figure 7: The figure plots mean monthly portfolio excess returns as a function of λ_1 and λ_2 when $T = 1$, 6, 12, and 24 for time-varying LASSO. λ_1 shrinks characteristic coefficients toward zero, while λ_2 penalizes changes between consecutive coefficient estimates. T is the window length: in month t, I pool data from months $t-T$ to $t-1$ to estimate β_t. The sample period is January 1990 to December 2021, and the out-of-sample prediction period is January 2000 to December 2021.

Among the windows $T = 60$, 72, 96, and 120, the highest reported Sharpe ratios and mean excess returns occur at $T = 60$ or 72.

Models with positive penalties on both coefficient levels and coefficient changes achieve the highest reported out-of-sample R^2 and portfolio Sharpe ratios in these

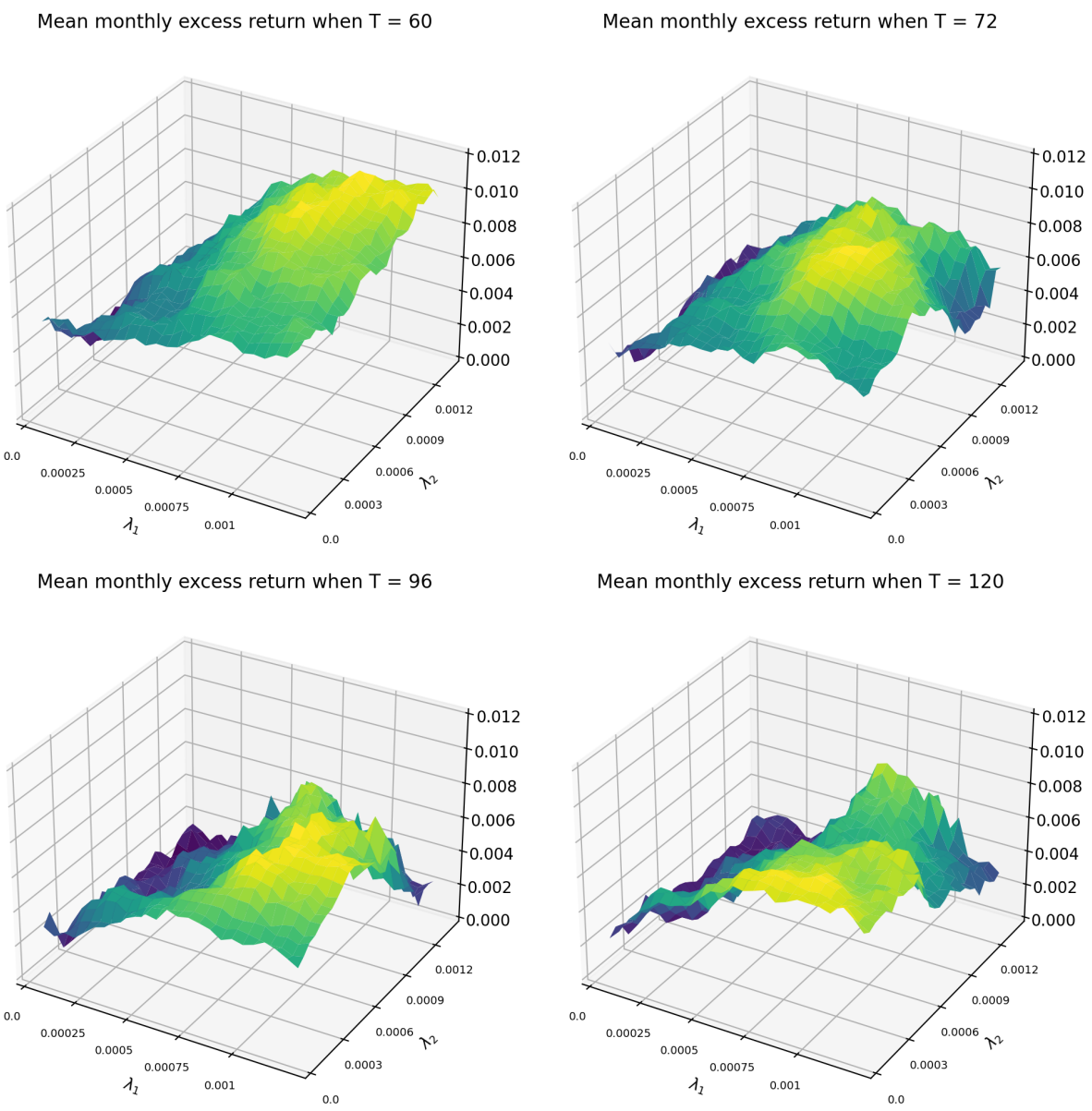


Figure 8: The figure plots mean monthly portfolio excess returns as a function of λ_1 and λ_2 when $T = 60$, 72, 96 and 120 for time-varying LASSO. λ_1 shrinks characteristic coefficients toward zero, while λ_2 penalizes changes between consecutive coefficient estimates. T is the window length: in month t, I pool data from months $t-T$ to $t-1$ to estimate β_t. The sample period is January 1990 to December 2021, and the out-of-sample prediction period is January 2000 to December 2021.

comparisons. These results support the predictive value of combining sparsity across characteristics with temporal regularization.

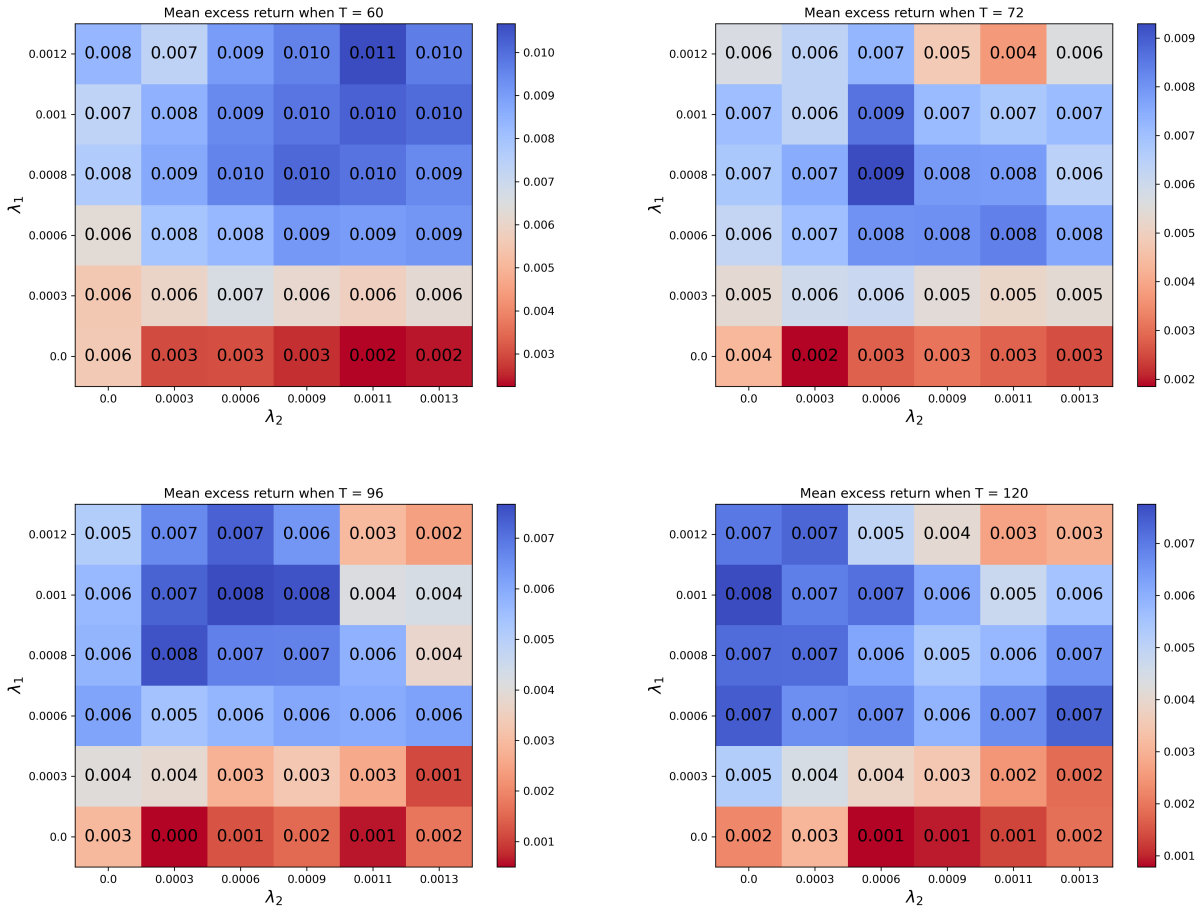


Figure 9: The figure shows mean excess returns of portfolios formed from time-varying LASSO forecasts across penalty parameters λ_1 and λ_2 for different values of T. λ_1 shrinks characteristic coefficients toward zero, while λ_2 penalizes changes between consecutive coefficient estimates. T is the window length: in month t, I pool data from months $t-T$ to $t-1$ to estimate β_t. The sample period is January 1990 to December 2021, and the out-of-sample prediction period is January 2000 to December 2021.

5.4 Time-varying Model Complexity

I next examine how the number of selected characteristics and their estimated coefficients evolve over time. Model complexity is the number of characteristics with nonzero coefficients. Figure 10 plots this count for conventional LASSO and time-varying LASSO with $T = 60$ and 120.

The Fama–MacBeth specification includes all 121 characteristics. In Figure 10, the number selected by conventional LASSO varies little over time. Time-varying LASSO with $T = 60$ shows the largest fluctuations among the plotted models.

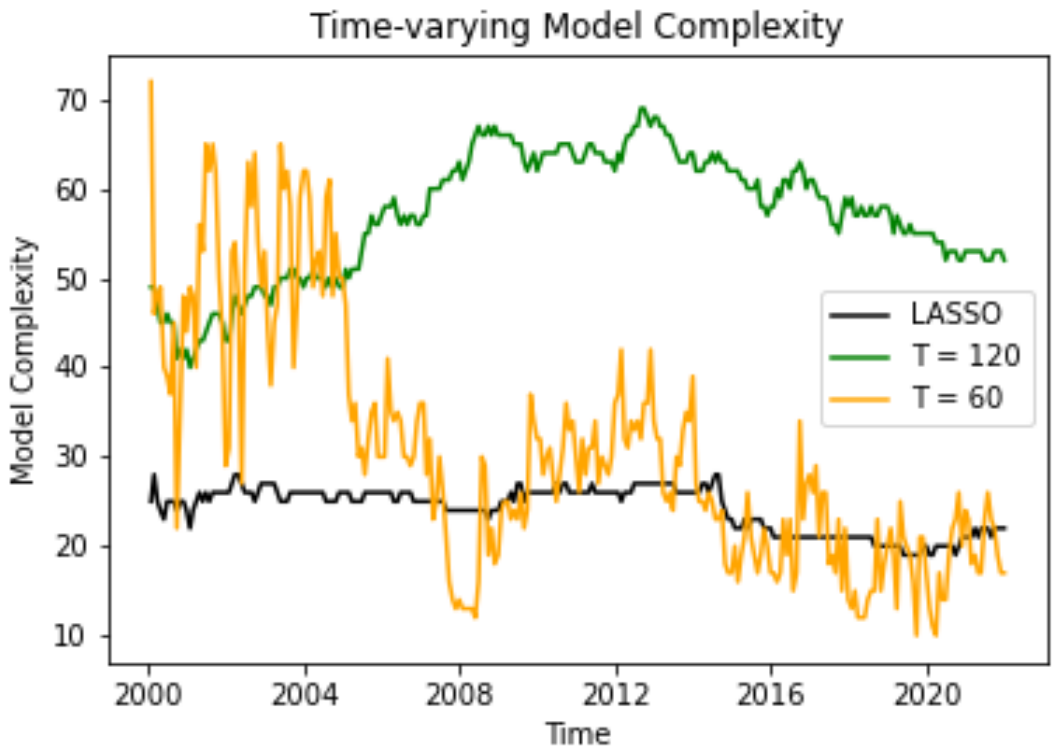


Figure 10: The figure shows model complexity for conventional LASSO and time-varying LASSO with $T = 60$ and 120 in each training sample of the out-of-sample analysis. Model complexity is the number of characteristics with nonzero coefficients. T is the window length: in month t, I pool data from months $t - T$ to $t - 1$ to estimate β_t. The sample period is January 1990 to December 2021, and the out-of-sample prediction period is January 2000 to December 2021.

5.5 Time-varying Risk Premia

Figure 11 compares estimated characteristic payoffs for six anomalies—book-to-market ratio, short-term reversal, 12-month momentum, EBITDA-to-sales ratio, tangibility, and 126-day share turnover—across Fama–MacBeth, conventional LASSO, and time-varying LASSO with $T = 60$ and 120.

Fama–MacBeth regressions retain all characteristics in the specification. In Figure 11, conventional LASSO produces more stable coefficient paths and generally smaller coefficient magnitudes than Fama–MacBeth regressions.

Under time-varying LASSO, coefficients for book-to-market ratio, short-term reversal, and tangibility are zero in some periods and nonzero in others. These paths reflect changes in which characteristics enter the fitted forecasting model.

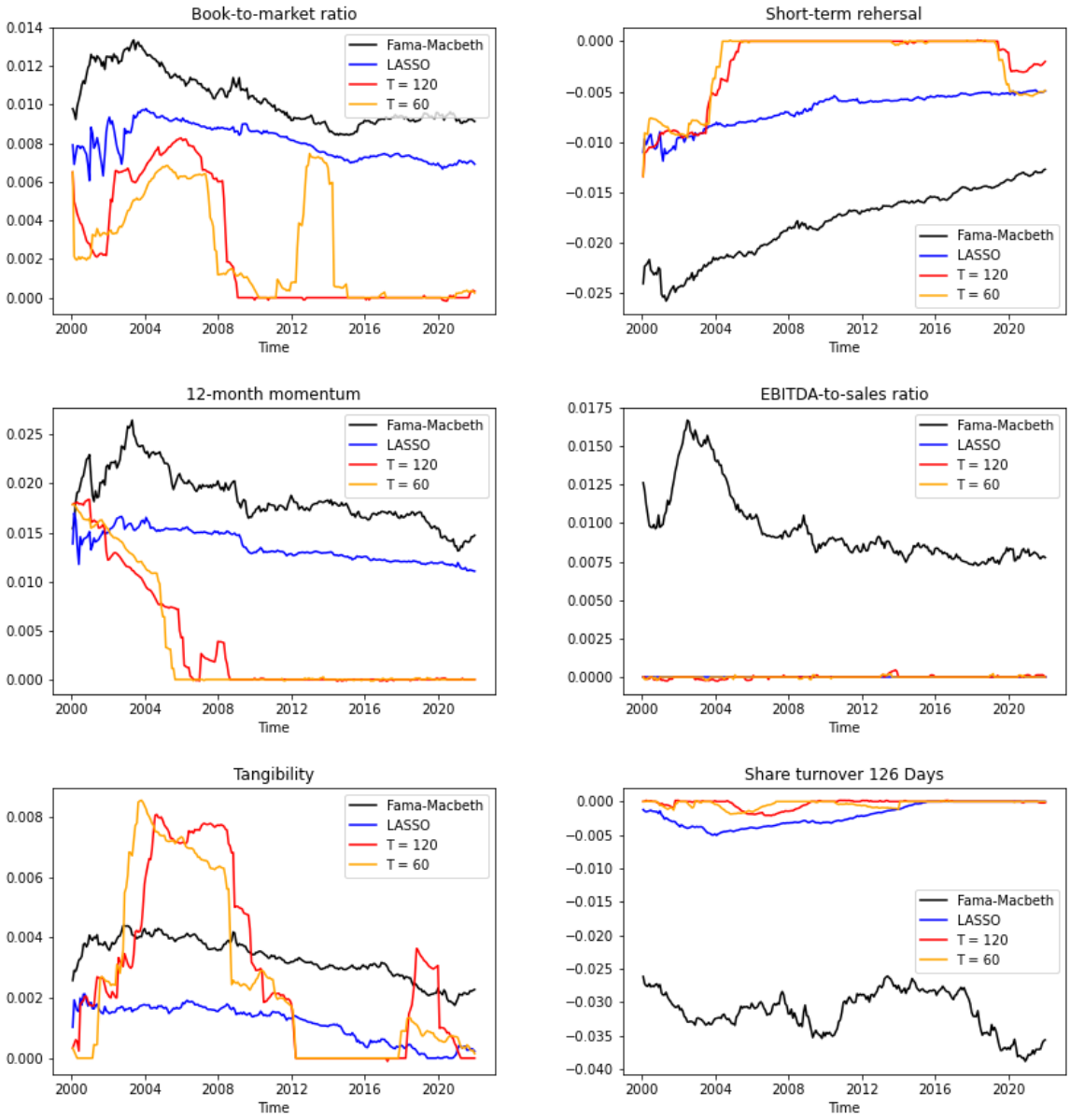


Figure 11: The figure shows estimated risk premia for Fama–MacBeth, conventional LASSO, and time-varying LASSO with $T = 60$ and 120 in each training sample of the out-of-sample analysis. T is the window length: in month t, I pool data from months $t - T$ to $t - 1$ to estimate β_t. The sample period is January 1990 to December 2021, and the out-of-sample prediction period is January 2000 to December 2021.

5.6 Cumulative Performance

Figure 12 plots cumulative wealth for value-weighted and equal-weighted portfolios sorted on out-of-sample forecasts from time-varying LASSO with T = 60 and 120. Cumulative wealth is the value over time of $1 invested in January 2000.

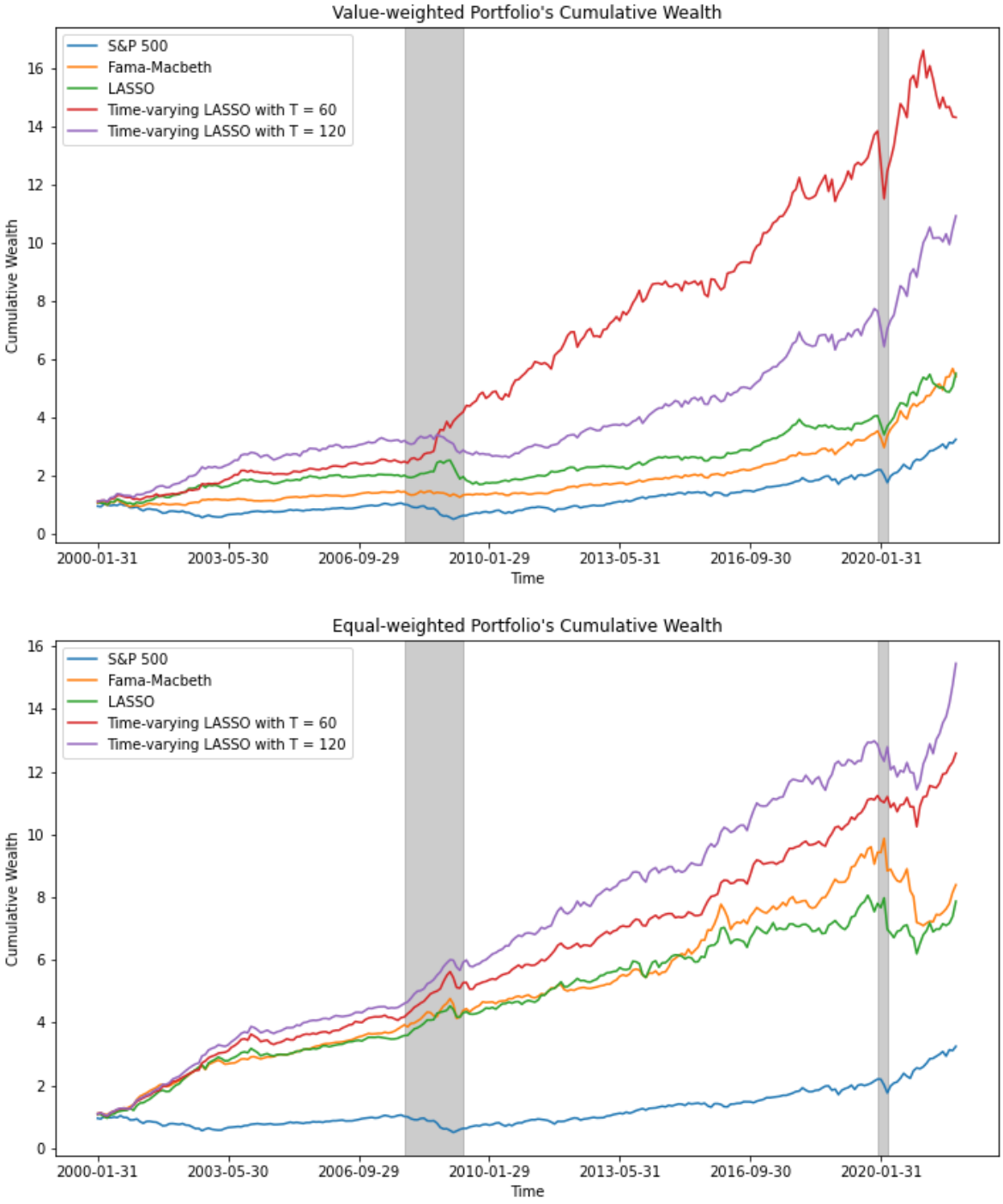


Figure 12: The figure shows cumulative wealth for value-weighted and equal-weighted portfolios sorted on out-of-sample forecasts from time-varying LASSO with $T = 60$ and 120. T is the window length: in month t, I pool data from months $t - T$ to $t - 1$ to estimate β_t. The shaded periods are December 2007 to June 2009 and February 2020 to April 2020. The sample period is January 1990 to December 2021, and the out-of-sample prediction period is January 2000 to December 2021.

Table 4: Out-of-sample predictability (percentage R^2) in simulation

	Assumption 1	Assumption 2 $\rho = 0.9$	Assumption 2 $\rho = 0.5$	Assumption 2 $\rho = 0.1$
Fama–MacBeth	0.37	0.05	0.02	0.01
Pooled OLS	48.9	1.03	0.88	0.46
LASSO	64.9	3.41	1.45	0.80
Time-varying LASSO with T = 60	61.2	3.67	1.62	0.71
Time-varying LASSO with T = 120	63.8	3.79	1.54	0.77

The table reports the out-of-sample percentage R^2 for different data-generating processes and estimation methods. Under Assumption 1, characteristic payoffs are constant over time. Under Assumption 2, payoffs follow an autoregressive process with persistence parameter ρ. Smaller values of ρ imply less persistent coefficient paths.

The time-varying LASSO portfolios attain higher cumulative returns than the Fama–MacBeth and conventional LASSO portfolios under both weighting schemes. During the out-of-sample period, the S&P 500 had a cumulative return of 229%, rising from an index level of 1,455 at the start of 2000 to 4,796 at the end of 2021.

The value-weighted portfolio based on time-varying LASSO with $T = 60$ achieves a cumulative return of 1,329%. The Fama–MacBeth and conventional LASSO portfolios return 439% and 450%, respectively.

For the equal-weighted portfolio, time-varying LASSO with $T = 120$ achieves a cumulative return of 1,445%, compared with 739% for Fama–MacBeth and 686% for conventional LASSO.

6 Simulation

Section 5 documents the empirical predictive performance of time-varying LASSO. I now compare estimators under constant characteristic payoffs and time-varying payoffs with different degrees of persistence. These simulations assess when information from previous estimates improves prediction.

I simulate returns using 12 characteristics, setting six of their risk premia to zero. I use the full data set for these 12 characteristics and compare the methods described in Section 3: Fama–MacBeth, pooled OLS, LASSO, and time-varying LASSO. The simulated 6-factor model for excess returns r_{t+1}, for $t = 1, 2, ..., T$, is

$$r_{i,t+1} = \sum_{s=1}^{6} \beta_{s,t} f_{s,i,t} + \epsilon_{i,t+1}$$

where $\beta_{s,t}$ is the payoff to characteristic s at time t, $f_{s,i,t}$ is characteristic s for firm i at time t, and $\epsilon_{i,t+1}$ is an idiosyncratic error. I resample $\epsilon_{i,t+1}$ with replacement from the empirical residuals. The sample covers 2000 to 2021. I simulate $\beta_{s,t}$ under two assumptions:

1. $\beta_{s,t} = \beta_{s,t-1}$ for $t = 1, 2, ..., T$.

2. $\beta_{s,t} = \rho\beta_{s,t-1} + u_t$ where $u_t \sim N(0, 1 - \rho^2)$, and I choose $\rho = 0.1, 0.5$ and 0.9.

Under the first assumption, characteristic payoffs are constant over time. Under the second, ρ controls persistence across adjacent periods, with larger values implying more persistent payoffs. The penalty parameters for LASSO and time-varying LASSO are tuned by validation as described in Section 3.4.

Table 4 shows that temporal regularization improves out-of-sample predictive accuracy when characteristic payoffs are persistent but evolve over time. Both time-varying LASSO specifications outperform conventional LASSO when $\rho = 0.5$ or 0.9. In contrast, conventional LASSO performs better when payoffs are constant or weakly persistent ($\rho = 0.1$). These results indicate that temporal regularization is most useful when predictive relationships evolve gradually.

7 Conclusion

Temporal regularization incorporates persistence in characteristic payoffs into return prediction. The estimator combines sparsity across predictors with a penalty on coefficient changes across adjacent estimation periods. The estimation window controls data recency. The simulations show that predictive gains depend on payoff persistence.

In U.S. equities, time-varying LASSO improves out-of-sample return prediction and portfolio performance relative to the conventional Fama–MacBeth and LASSO benchmarks. Estimated characteristic payoffs and the set of selected predictors also vary more over time. These findings support forecasting the cross-section of stock returns with evolving characteristic payoffs while regularizing changes across estimation periods.

References

Bessembinder, H. H., Burt, A. P., and Hrdlicka, C. M. (2023). Time series variation in the factor zoo. Unpublished working paper.

Bryzgalova, S., DeMiguel, V., Li, S., and Pelger, M. (2023). Asset-pricing factors with economic targets. Available at SSRN: https://ssrn.com/abstract=4344837.

Bryzgalova, S., Pelger, M., and Zhu, J. (2025). Forest through the trees: Building cross-sections of stock returns. *The Journal of Finance*, 80(5):2447–2506.

Chen, L., Pelger, M., and Zhu, J. (2024). Deep learning in asset pricing. *Management Science*, 70(2):714–750.

Cochrane, J. H. (2011). Presidential address: Discount rates. *The Journal of Finance*, 66(4):1047–1108.

Ehsani, S. and Linnainmaa, J. T. (2022). Factor momentum and the momentum factor. *The Journal of Finance*, 77(3):1877–1919.

Fama, E. and Bliss, R. R. (1987). The information in long-maturity forward rates. *American Economic Review*, 77(4):680–92.

Fama, E. F. and French, K. R. (1988). Dividend yields and expected stock returns. *JournalofFinancialEconomics*, 22(1):3–25.

Fama, E. F. and MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, 81(3):607–636.

Ferson, W. E. and Harvey, C. R. (1991). The variation of economic risk premiums. *Journal of Political Economy*, 99(2):385–415.

Freyberger, J., Neuhierl, A., and Weber, M. (2020). Dissecting Characteristics Nonparametrically. *The Review of Financial Studies*, 33(5):2326–2377.

Gagliardini, P., Ossola, E., and Scaillet, O. (2016). Time-varying risk premium in large cross-sectional equity data sets. *Econometrica*, 84(3):985–1046.

Green, J., Hand, J. R. M., and Zhang, X. F. (2017). The Characteristics that Provide Independent Information about Average U.S. Monthly Stock Returns. *The Review of Financial Studies*, 30(12):4389–4436.

Gu, S., Kelly, B., and Xiu, D. (2020). Empirical Asset Pricing via Machine Learning. *The Review of Financial Studies*, 33(5):2223–2273.

Haddad, V., Kozak, S., and Santosh, S. (2020). Factor Timing. *The Review of Financial Studies*, 33(5):1980–2018.

Han, Y., He, A., Rapach, D. E., and Zhou, G. (2024). Cross-sectional expected returns: New Fama–MacBeth regressions in the era of machine learning. *Review of Finance*, 28(6):1807–1831.

Haugen, R. A. and Baker, N. L. (1996). Commonality in the determinants of expected stock returns. *Journal of Financial Economics*, 41(3):401–439.

Hou, K., Xue, C., and Zhang, L. (2020). Replicating Anomalies. *The Review of Financial Studies*, 33(5):2019–2133.

Jensen, T. I., Kelly, B., and Pedersen, L. H. (2023). Is there a replication crisis in finance? *The Journal of Finance*, 78(5):2465–2518.

Kelly, B. T., Pruitt, S., and Su, Y. (2019). Characteristics are covariances: A unified model of risk and return. *Journal of Financial Economics*, 134(3):501–524.

Kozak, S., Nagel, S., and Santosh, S. (2020). Shrinking the cross-section. *Journal of Financial Economics*, 135(2):271–292.

Lettau, M. and Pelger, M. (2020a). Estimating latent asset-pricing factors. *Journal of Econometrics*, 218(1):1–31.

Lettau, M. and Pelger, M. (2020b). Factors That Fit the Time Series and Cross-Section of Stock Returns. *The Review of Financial Studies*, 33(5):2274–2325.

Lewellen, J. (2015). The cross-section of expected stock returns. *Critical Finance Review*, 4(1):1–44.

Raponi, V. and Zaffaroni, P. (2023). Dissecting anomalies in conditional asset pricing. Available at SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4485448.

Shiller, R. J. (1981). Do stock prices move too much to be justified by subsequent changes in dividends. *AmericanEconomicReview*, 71:421–436.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(1):267–288.

Tibshirani, R., Saunders, M., Rosset, S., Zhu, J., and Knight, K. (2005). Sparsity and smoothness via the fused lasso. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(1):91–108.

Yogatama, D., Heilman, M., O'Connor, B., Dyer, C., Routledge, B. R., and Smith, N. A. (2011). Predicting a scientific community's response to an article. In Barzilay, R. and Johnson, M., editors, *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*, pages 594–604, Edinburgh, Scotland, UK. Association for Computational Linguistics.

8 Appendix

8.1 Data

be_me	Book Equity scaled by Market Equity
ret_12_1	Momentum 1-12 Months
market_equity	Market Equity
ret_1_0	Short Term Reversal
rvol_252d	Return Volatility 252 Days
beta_252d	CAPM Beta 252 Days
qmj_safety	Quality Minus Junk - Safety
rmax1_21d	Maximum Return 21 Days
chcsho_12m	Change in Shares - 12 Month
ni_me	Net Income scaled by ME
eq_dur	Equity Duration
ret_60_12	Momentum 12-60 Months
ope_be	Operating Profit to Equity scaled by BE
gp_at	Gross Profit scaled by Assets
ebit_sale	Operating Profit Margin after Depreciation
at_gr1	Asset Growth 1yr
sale_gr1	Sales Growth 1yr
at_be	Book Leverage
cash_at	Cash and Short Term Investments scaled by Assets
age	Age of the firms in months
z_score	Altman Z-Score
rvol_21d	Return volatility
tangibility	Asset tangibility
dolvol_var_126d	Coefficient of variation for dollar trading volume
turnover_var_126d	Coefficient of variation for share turnover 126 Days
sale_bev	Assets turnover
ret_12_7	Momentum 7-12 Months
ret_6_1	Momentum 1-6 Months
ret_36_12	Momentum 12-36 Months
ret_2_0	Momentum 0-2 Months

ret_3_0	Momentum 0-3 Months
ret_3_1	Momentum 1-3 Months
ret_6_0	Momentum 0-6 Months
ret_12_0	Momentum 0-12 Months
ni_be	Return on equity
niq_at	Quarterly return on assets
prc_highprc_252d	Current price to high price over last year
cop_atl1	Cash-based operating profits-to- lagged book assets
opex_at	Operating leverage
oaccruals_at	Operating Accruals
eqnpo_me	Net payout yield
noa_at	Net operating assets
turnover_126d	Share turnover 126 Days
ppeinv_gr1a	Change PPE and Inventory
dgp_dsale	Change gross margin minus change sales
at_turnover	Capital turnover
chcsho_1m	Change in Shares - 1 Month
inv_gr1a	Inventory Change 1yr
f_score	Piotroski F-Score
zero_trades_21d	Zero Trades
betabab_1260d	Betting Against Beta
bidaskhl_21d	21 Day Bid-Ask High-Low
book_equity	Book Equity
net_income	Net Income
gp_gr1a	Gross Profit Change 1yr
gp_sale	Gross Profit Margin
emp_gr1	Employee Growth 1 yr
div1m_me	Dividend to Price - 1 Month
eqnpo_1m	Net Equity Payout - 1 Month
resff3_6_1	Residual Momentum - 6 Month
rskew_21d	Return Skewness
ivol_ff3_21d	Fama and French Idiosyncratic Vol
aliq_at	Liquidity scaled by lagged Assets

cash_gr1a	Cash and Short-Term Investments Change 1yr
beta_21d	CAPM Beta 21 day
ivol_capm_21d	CAPM Idiosyncratic Vol 21 day
iskew_capm_21d	CAPM Skewness 21 Day
qmj_prof	Quality Minus Junk - Profit
div12m_me	Dividend yield
ebitda_mev	Ebitda-to-market enterprise value
debt_me	Debt-to-market
sale_me	Sales-to-market
prc	Price per share
aliq_mat	Liquidity of market assets
mispricing_mgmt	Mispricing factor: Management
ni_ar1	Earnings persistence
ebitda_gr1a	Operating Profit after Depreciation Change 1yr
ebitda_gr3a	Operating Profit after Depreciation Change 3yr
zero_trades_252d	Number of days with zero trades over 252 days
earnings_variability	Earnings Variability
div3m_me	Dividend to Price - 3 Months
seas_1_1an	1 Year Annual Seasonality
at_gr3	Asset Growth 3yr
ca_gr1	Asset Growth 1yr
ca_gr3	Current Asset Growth 3yr
be_gr1	Book Equity Growth 1yr
cash_gr3a	Cash Growth 3yr
rmax5_21d	Highest 5 days of return
rmax5_rvol_21d	Max Return to Volatility
zero_trades_126d	Number of zero trades (6 months)
mispricing_perf	Mispricing factor: Performance
assets	Assets
sales	Sales
div6m_me	Dividend to Price - 6 Months
chcsho_3m	Change in Shares - 3 Months
chcsho_6m	Change in Shares - 6 Months

ebitda_sale	Operating Profit Margin before Depreciation
ebit_at	Operating Profit after Depreciation scaled by Assets
debt_at	Debt-to-Assets
intrinsic_value	Intrinsic Value
o_score	Ohlson O - Score
beta_60m	60 Month CAPM Beta
cash_lt	Cash Balance scaled by Total Liabilities
ni_emp	Net Income scaled by Employees
sale_emp	Sales scaled by Employees
taccruals_at	Total Accruals
rvolhl_21d	The high-low return volatility
lt_gr1	Total Liabilities Growth 1yr
debt_gr1	Total Debt Growth 1yr
ocf_gr1a	Operating Cash Flow Change 1yr
lti_gr1a	Investment and Advances Change 1yr
capx_gr1	Capital Expenditures Change 1yr
capx_at	Capital Expenditures scaled by Assets
ni_ivol	Net Income Idiosyncratic Volatility
seas_2_5an	Years 2-5 lagged returns, annual
coskew_21d	Coskewness
be_gr1a	Change in common equity
rd_sale	R&D-to-sales
dsale_dinv	Change sales minus change Inventory
dsale_drec	Change sales minus change receivables
ebit_bev	Return on net operating assets

8.2 Additional Plots

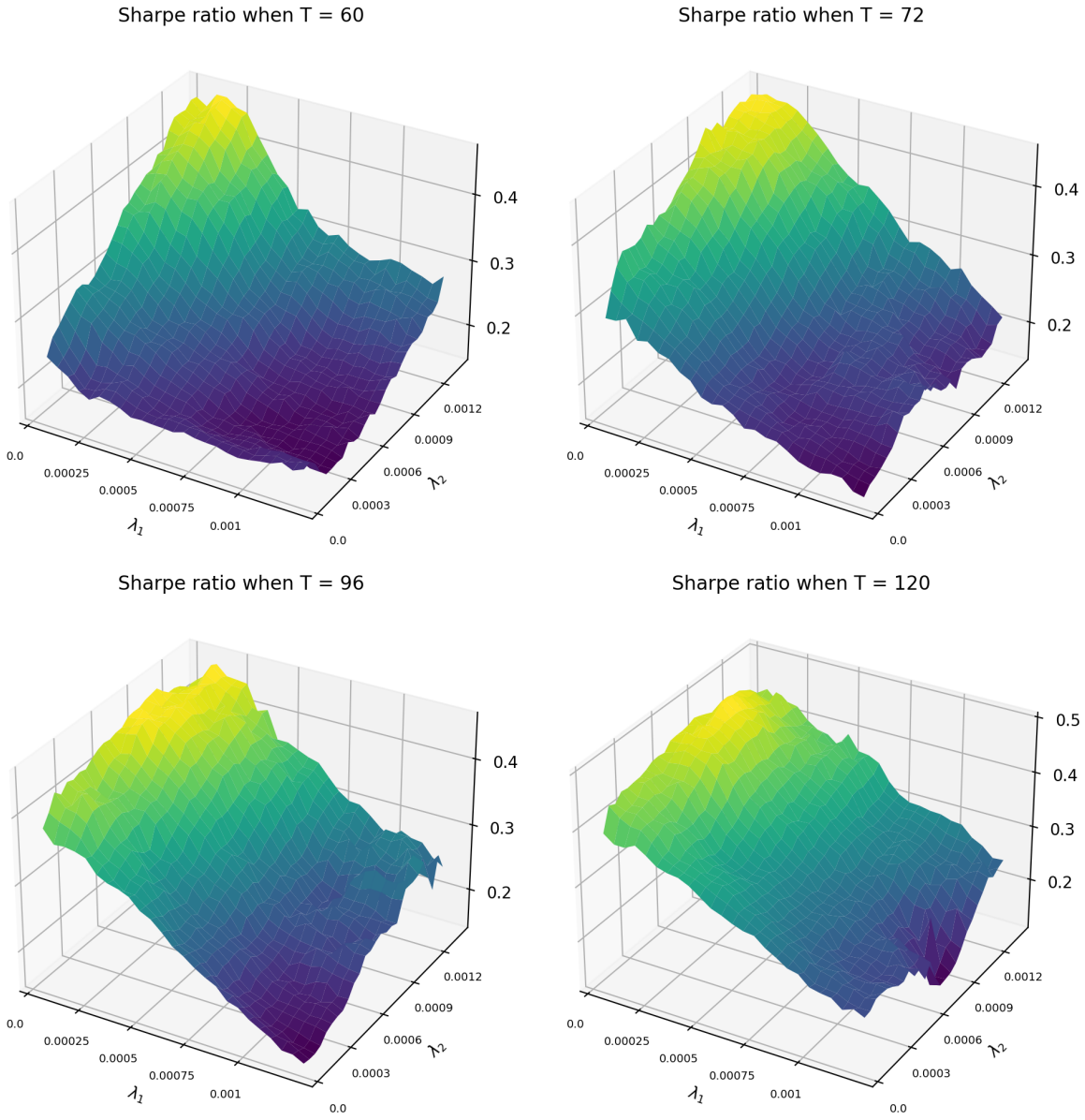


Figure 13: The figure plots equal-weighted portfolio Sharpe ratios for time-varying LASSO as a function of λ_1 and λ_2 when $T = 60$, 72, 96 and 120. λ_1 shrinks characteristic coefficients toward zero, while λ_2 penalizes changes between consecutive coefficient estimates. T is the window length: in month t, I pool data from months $t-T$ to $t-1$ to estimate β_t. The sample period is January 1990 to December 2021, and the out-of-sample prediction period is January 2000 to December 2021.