Lost in Translation: How Predictability Turns Into Performance[*]

Lukas Salcher[†], Sebastian Stöckl[‡] Michael Hanke[§]

This version: July, 2024

Abstract

This paper investigates the impact of return predictability on portfolio performance using mean-variance optimization and various timing strategies. Our approach captures predictability and performs well in conventional timing exercises, showing how predictability leads to substantial improvements in Sharpe ratios. However, for MVO, the relationship between predictability and performance is more complex. While higher predictability leads to increased mean returns, it also introduces significant volatility and dispersion in predicted values, resulting in extreme weights and poor risk-adjusted performance. These counter-intuitive results suggests that mean-variance optimization, may not always yield superior risk-adjusted returns when predictability is increased hinting at a nuanced relationship between predictability and performance in portfolio optimization.

Keywords: Return prediction, forecasting accuracy, portfolio optimization, performance

JEL classification: G11, G14, G17

[*]We would like to thank the participants of the research exchange between the University of Liechtenstein and the University of St. Gallen, as well as the participants of the IDS 2024 for very helpful comments and inputs.

[†]University of Liechtenstein, Liechtenstein Business School, Fürst-Franz-Josef-Strasse, 9490 Vaduz, Liechtenstein; email: lukas.salcherl@uni.li; tel: +423 265 1348 (Corresponding author).

[‡]University of Liechtenstein, Liechtenstein Business School, Fürst-Franz-Josef-Strasse, 9490 Vaduz, Liechtenstein; email: sebastian.stoeckl@uni.li; tel: +423 265 1153.

[§]University of Liechtenstein, Liechtenstein Business School, Fürst-Franz-Josef-Strasse, 9490 Vaduz, Liechtenstein; email: michael.hanke@uni.li; tel: +423 265 1155.

1 Introduction

According to Modern Portfolio Theory by Markowitz (1952), allocating wealth is a two step procedure. The first step involves forming expectations about the future outcome of the risk and the return of each assets within the investment universe, while the second step revolves around optimally combining the assets in a way that maximizes the return given the level of risk an investor is willing to bear. Undeterred by this straightforward theoretical framework published more than seven decades ago, the question of how to form expectations and how to turn them into investment profits is ambiguous and often puzzling for both academic researchers and practitioners. On top of that, the relationship between traditional statistical measures of forecast accuracy and the actual economic value of forecasts is often very fragile, as for example documented by Leitch and Tanner (1991), Cenesizoglu and Timmermann (2012), Reschenhofer et al. (2020) as well as Cederburg et al. (2023).

Explanations for the opaque link between forecasts and performance can be separated into two camps. One strand of the literature revolves around the question of whether returns in general or the equity premium in particular are predictable at all. Welch and Goyal (2008) as well as Goyal et al. (2021) are well known instances for this debate as their results show that the vast majority of predictors that have been published in top journals "lose" their ability to predict future returns on a statistically significant level or no longer yield economically meaningful returns after they have been published.

Contrary to that, the alternative strand of the literature argues that asset returns are predictable to some extent and thus mainly focuses on improving upon the methods and measures used to predict and evaluate predictions. Thus, some researchers have focused on improving the predictive power of a set of variables by imposing economically meaningful restrictions on the sign of the predictions in order to more consistently translate predictions into performance, such as Campbell and Thompson (2008). Alternatively, others focus more strongly on boosting prediction accuracy by optimally combining various forecasts from as set of forecast variables into a single forecast or on advancing prediction techniques by borrowing methods from the realm of machine learning, such as Ridge and LASSO regressions, elastic nets, neural networks or an ensemble thereof pushing for non-linear methods to account for the complexity of modern financial markets (Rapach et al., 2011, 2019; Gu et al., 2020; Kelly et al., 2024).

Finally, when it comes to measuring predictability, the statistical significance of a predictor is typically assessed using the "out-of-sample" R^2 which compares the mean-squared prediction errors of a predictor and its benchmark (Campbell and Thompson, 2008). Alternatively, the (root-)mean-squared-error or the correlation between predictions and realizations are sometimes used (Cenesizoglu and Timmermann, 2012; Kane et al., 2010; Allen et al., 2019).

While the vast majority of existing studies has focused on whether asset returns can be predicted and how to do so, much less is known about how these predictions can be systematically leveraged to generate superior economic outcomes. The issue is further compounded by the fact that the statistical significance of a predictor or a technique often does not translate into economic significance, as impressively demonstrated by Kelly et al. (2024). Against this backdrop, our study seeks to bridge the divide between return predictability and its economic value, focusing on the conversion of predictive accuracy into tangible investment performance. We posit that a reevaluation of the metrics used to measure predictability, coupled with a nuanced understanding of how these metrics relate to economic value, can unlock new avenues for exploiting predictability in investment decision-making.

In order to answer these questions, we set-up a simple environment in which we can precisely control the level of return predictability and thus by extension the accuracy of our predictions. Given this set-up, we further explore how economically valuable our predictions are given the most common investment strategies and universes investors face. In particular, we investigate two different types of assets and two different types of strategies. In terms of assets, we focus on the US equity premium and US industry portfolios. As far as investment strategies are concerned, we focus on investors that focus on simple timing strategies as well as on investors having mean-variance preferences that want to optimally diversify their portfolio.

Our findings indicate that our approach increases predictability across all considered and established measures of predictability, our results are more nuanced when it comes to the economic value of predictions. For timing exercises related to the equity premium, predictability and profitability as measured by the Sharpe ratio increase alongside each other. In a more sophisticated portfolio optimization setting, this relationship does not hold. The Sharpe ratios initially improve with increasing λ-values but eventually decline or fluctuate. We observe this trend across all portfolio sizes.

These findings underscore the complexity of leveraging predictability to enhance portfolio performance. While higher levels of predictability can lead to higher returns, they also introduce greater volatility, which can offset the benefits in terms of risk-adjusted performance. This counter-intuitive result suggests that mean-variance optimization, despite its theoretical appeal, may not always yield superior risk-adjusted returns when predictability is increased.

This research contributes to the existing literature by elucidating the direct link between return predictability and economic performance, challenging the prevailing paradigm that prioritizes statistical significance over economic value. Furthermore, we highlight the importance of choosing the appropriate measure for predictability and economic value given the prediction and investment task at hand. Finally, we demonstrate that even low levels of predictability, when strategically leveraged, can lead to significant economic gains. Thus, we not only advance our theoretical understanding of the predictability-performance nexus but also provide a valuable framework for practitioners seeking to navigate the complexities of financial markets with greater precision and profitability.

2 Literature Review

Given that this paper aims at evaluating how predictability turns into performance it is only natural that we first investigate how returns are currently predicted. This includes a review of methods, measures and variables that are used to predict various kinds of returns. Subsequently, we delver deeper into how the predictability of returns is related to common measures of economic value. Finally, we review what existing literature has to say about how (the lack of) predictability impacts asset allocation.

2.1 Return predictability: Methods, measures and variables

As the concept of informationally efficient capital markets, formalized and popularized by Fama (1970) with the so-called "Efficient Market Hypothesis" (EMH), is one of the most fundamental ones in academic Finance, the question of whether returns are and, from a theoretical perspective, should be predictable has been discussed vividly. While a discussion of that topic per se is beyond the scope of this paper, it is worth mentioning that return predictability only contradicts the EMH if the predictability does not reflect compensation for taking on risk (Rapach and Zhou, 2022). In other words, if predictability cannot be explained via risk and instead is more likely due to behavioural patterns or certain types of market inefficiencies that arise due to frictions (i.e. mispricing) it contradicts the EMH (Rapach and Zhou, 2022).

Having said that, this section mainly focuses on summarizing the literature related to the out-of-sample predictability of the equity premium, industry portfolios and single stock returns. Consequently, we begin with briefly discussing the state-of-the-art when it comes to predicting the equity premium. According to Rapach and Zhou (2022) the most widespread econometric model for predicting returns is a simple linear regression in which the returns of an asset are regressed onto one (or more) lagged variables that with (presumably) predictive ability. More formally, such a predictive regression could take the following form:

$$r_{t+1} = \alpha + \beta x_t + \epsilon_{t+1} \quad \text{for} \quad t = 1, ..., T-1, \tag{1}$$

where ϵ_{t+1} refers to the zero-mean *iid.* error term. Subsequently, it is possible to obtain an out-of-sample return forecast $r_{t+1|t}$ by using the OLS estimates of α and β through t (i.e. $\hat{\alpha}_t$ and $\hat{\beta}_t$). With respect to the question of whether α and β should be estimated using a rolling or expanding window, Dangl and Halling (2012) argue in favour of time-varying coefficients.

The predictive accuracy of such forecasts is typically evaluated using a summary statistic that measures the distance between realized and forecasted returns such as the out-of-sample mean squared forecast error (MSFE). In addition to that, and in an attempt to judge the accuracy of a predictor variable relative to a benchmark prediction Campbell and Thompson (2008) introduced the so-called "out-of-sample R^2" which compares the MSFE of a predictor variable with the MSFE of a simple benchmark such as the prevailing mean. The idea behind this measure is that if the more sophisticated prediction model that takes advantage of the information provided by predictor x_t generates a lower MSFE than a naive benchmark prediction that has not access to this kind of information, this means that the predictor x_t is capable of predicting out-of-sample. Note that we formally introduce and describe both of these measures as well as alternative measures for prediction accuracy in subsection 3.2.2).

As any discussion of equity premium prediction is incomplete without mentioning the seminal paper of Welch and Goyal (2008) we want to stress the importance of their contribution towards facilitating the shift from in-sample dominated prediction exercises towards a more realistic out-of-sample framework in addition to setting new standards for how to assess the predictive abilities of variables in a comparable, reproducible and realistic environment. Furthermore, Welch and Goyal (2008) also present evidence that cast doubt on the predictability of the equity premium as none of the 17 variables they have tested were able to consistently predict the equity premium out-of-sample.

In an attempt to address some of the issues highlighted by Welch and Goyal (2008), Campbell and Thompson (2008) argue that imposing weak, economically motivated restrictions on the return forecasts and the sign of the regression coefficients can already improve the out-of-sample predictions such that they are more accurate than the historical average. Confirming the initial results of Welch and Goyal (2008), Rapach et al. (2010) underline the importance of diversifying forecasts instead of relying on a single predictor variable and thus propose combining forecasts from n univariate predictive regressions to get an equally-weighted combination mean forecast. Alternatively, the authors suggest to take the predictive performance of the individual investors into account when combining the individual predicitons. Regardless of the weighting-scheme, the results of Rapach et al. (2010) indicate that forecast combinations improve the predictive accuracy on a statistically significant level. Besides that, Neely et al. (2014) use the same predictor variables as Welch and Goyal (2008) and another 14 technical indicators as inputs in a predictive regression framework based on principal components and thereby manage to outperform the naive predictions made by the prevailing mean significantly.

In a nutshell, published research indicates that the US equity premium is at least to some extent predictable out-of-sample and on a statistically significant level. Recent advances in the field highlight the importance of extending models beyond the classical framework based on univariate predictive regressions based on OLS and instead rely on methods that are more capable of dealing with the horrendously low signal-to-noise ratio of financial returns. At their core, these novel methods mainly exploit the benefits of the bias-variance trade-off via shrinkage and/or dimension reduction.

2.2 Linking predictability to economic value

While the vast majority of existing studies has focused on whether asset returns can be predicted and how to do so, much less is known about how these predictions can be systematically leveraged to generate superior economic outcomes. The relationship between traditional statistical measures of forecast accuracy and the actual economic value of forecasts is often very fragile, as for example documented by Leitch and Tanner (1991), Cenesizoglu and Timmermann (2012), Reschenhofer et al. (2020) as well as Cederburg et al. (2023). More recently, Kelly et al. (2024) have impressively demonstrated that the positive and significant out-of-sample R^2 of a predictor is not a necessary condition for economically valuable timing signals of that predictor, again highlighting the mismatch between statistical and economic significance. Unfortunately, R^2_{OOS} only focuses on the statistical significance of a predictor variable and neglects its economic significance. As a consequence, R^2_{OOS} often rejects the hypothesis that a specific predictor delivers forecasts that are significantly more accurate than those of the prevailing mean while the same predictor significantly outperforms the prevailing mean from an economic perspective by generating higher utility levels or Sharpe ratios when used in a market timing setting (i.e. Campbell and Thompson (2008) or Kelly et al. (2021)).

Explanations for the opaque link between forecasts and performance can be separated into two camps. One strand of the literature revolves around the question of whether returns in general or the equity premium in particular are predictable at all. Welch and

Goyal (2008) as well as Goyal et al. (2021) are well known instances for this debate as their results show that the vast majority of predictors that have been published in top journals “lose“ their ability to predict future returns on a statistically significant level or no longer yield economically meaningful returns after they have been published.

Contrary to that, the alternative strand of the literature argues that asset returns are predictable to some extent and thus mainly focuses on improving upon the methods and measures used to predict and evaluate predictions. Thus, some researchers have focused on improving the predictive power of a set of variables by imposing economically meaningful restrictions on the sign of the predictions in order to more consistently translate predictions into performance, such as Campbell and Thompson (2008). Alternatively, others focus more strongly on boosting prediction accuracy by optimally combining various forecasts from as set of forecast variables into a single forecast or on advancing prediction techniques by borrowing methods from the realm of machine learning, such as Ridge and LASSO regressions, elastic nets, neural networks or an ensemble thereof pushing for non-linear methods to account for the complexity of modern financial markets (Rapach et al., 2011, 2019; Gu et al., 2020; Kelly et al., 2024).

3 Data & Methodology

3.1 Data

Currently, we use two types of datasets. First of all, for all timing exercises, the US equity premium is at our centre of attention. The data is freely available on the website of Kenneth French[1]. For the sorting and optimization tasks, we use the returns of the 49 industry portfolios that is also freely available on Ken French's website.

3.2 Methodology

In this section we briefly outline our methodology. This includes the set-up we use to come up with return predictions for which we can exactly control the level of predictability. Furthermore, we also examine the most popular measures of predictability as well as the strategies we investigate and how we define their economic value.

[1]See https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

3.2.1 Controlled Return Predictions

We begin this section with a brief outline of the prediction set-up which we use to exactly control the level of predictability inherent in our forecasts. Our model is takes the following form:

$$\hat{r}^{\lambda}_{\tau,t+1} = (1-\lambda)\bar{r}_{\tau} + \lambda r_{t+1} \tag{2}$$

where $\hat{r}^{\lambda}_{t+1}$ are our controlled return predictions and $\bar{r}_{\tau}$ the rolling average over the previous τ months. $\bar{r}_{\tau}$ serves as a basis forecasts which we want to improve. Note that for our purposes we assume that τ can take any of the following values: $\tau \in \{12, 24, 60, 120, 240\}$. We motivate our choice of rolling averages as our basis forecast because we, in accordance with finance theory, associate only a negligible level of actual predictability with such a forecast. Naturally, $\bar{r}_{\tau}$ can be replaced by any other return prediction $\hat{r}_{t+1}$. r_{t+1} refers to the actual realized return in $t+1$ and thus λ represents the degree of look-ahead bias that we introduce in order to control the accuracy of our predictions. More precisely, we investigate different levels of λ by gradually increasing it from 0 to 0.1 by steps of 0.002 which corresponds to a look-ahead bias of 10%. While we keep λ constant over time in our base case, this set-up generally allows for a more flexible modelling of λ to more realistically capture the time-varying nature of predictability and its diversity across assets. Our benchmark predictor is the prevailing mean of the asset(s).

3.2.2 Measuring Prediction Accuracy

By far the most popular measure for assessing the accuracy of predicted returns is the so-called mean squared forecast error ($MSFE$) which is defined in (3):

$$MSFE_i = \frac{1}{T}\Sigma_{t=1}^{T}(r_{i,t+1} - \hat{r}_{i,t})^2 \tag{3}$$

where $r_{i,t+1}$ is the realized return of asset i in time $t{+}1$ and $\hat{r}_{i,t}$ is the time t prediction of the return of asset i in time $t{+}1$. Similarly, the mean absolute forecast error ($MAFE$) can by calculated by using the absolute value of the error term instead of squaring them, as in (4):

$$MAFE_i = \frac{1}{T}\Sigma_{t=1}^{T}|r_{i,t+1} - \hat{r}_{i,t}| \tag{4}$$

In order to assess the relative prediction accuracy of a prediction model, the so-called "out-of-sample R^2" is used:

$$R^2_{OOS} = 1 - \frac{\sum_{t=1}^{T}(r_{i,t} - \hat{r}_{i,t})^2}{\sum_{t=1}^{T}(r_{i,t} - \bar{r}_{i,t})^2} \tag{5}$$

where $\bar{r}_{i,t}$ denotes the prevailing mean of asset i at time t. Essentially, (5) compares the sum of squared forecast errors of a prediction model with the sum of squared errors of a benchmark prediction model. Note that Gu et al. (2020) argue in favour of using 0 as a benchmark prediction when predicting returns for single stocks which is why we consider both cases in our analysis.

While the predictability measures mentioned above are the most common ones, there are also alternative measures that have been used by some researchers. Kane et al. (2010) as well as Allen et al. (2019), for example, both use the correlation between predictions and realizations to measure the accuracy of predictions despite referring to this measure as "R^2". More formally, the "forecast correlation" can be calculated as in (6):

$$\rho_{r,\hat{r}} = \frac{\sum_{t=1}^{T}(r_t - \bar{r}_t)(\hat{r}_t - \bar{\hat{r}}_t)}{\sqrt{\sum_{t=1}^{T}(r_t - \bar{r}_t)^2 \sum_{t=1}^{T}(\hat{r}_t - \bar{\hat{r}}_t)^2}} \tag{6}$$

Beyond that, another simple measure that can be used to proxy prediction accuracy is the so-called "Hit Ratio". The ratio measures the number of directionally correct predictions relative to the toal number of predictions as depcited in (7):

$$HR = \frac{1}{T}\sum_{t=1}^{T} I(sign(\hat{r}_t) = sign(r_t)) \tag{7}$$

where $I()$ is an indicator function that has a value of 1 if the sign of the prediction matches the sign of the realization and 0 otherwise.

Most of the measures mentioned above result in a single number which often is a summary statistic that can to some extent be influenced by a few outliers. This vulnerability in combination with the observation that out-of-sample return predictability seems to be higher in economic downturns as documented by Dangl and Halling (2012) can result in a distorted notion of relatively high predictability although it is concentrated around a few periods or rare events. In fact, some very popular predictor variables such as the aggregate short-interest measure of Rapach et al. (2016) is known to suffer from this phenomenon. To circumvent this issue, Welch and Goyal (2008) entertain the idea of graphically illustrating how predictive outperformance develops over time using cumulative squared forecast error differences (CSFED). These can be computed using (8):

$$CSFED = \sum_{t=1}^{T}[(r_t - \bar{r}_t)^2 - (r_t - \hat{r}_t)^2] \tag{8}$$

Note that all the measures mentioned above mainly focus on assessing the accuracy of return predictions for one asset. Our paper however is not restricted to the classical equity premium prediction task and also investigates predictability in the cross-section.

Thus, we also want to briefly mention the cross-sectional MSFE of Han et al. (2023). In contrast to the conventional MSFE, the cross-sectional MSFE does not investigate prediction errors across time but within the cross-section. Furthermore, the measures demeans the prediction errors and assigns market-cap based weights to each error to reflect the weight of each asset within the cross-section:

$$MSFE_t^{CS} = \frac{1}{n_t} \sum_{i=1}^{n_t} w_{i,t} [(r_{i,t} - \bar{r}_t) - (\hat{r}_{i,t|t-1} - \bar{\hat{r}}_{t|t-1})]^2 \quad \text{for} \quad t = 1, ..., T, \tag{9}$$

where $\bar{r}_t$ and $\bar{\hat{r}}_{t|t-1}$ are value-weighted cross-sectional means for the predicted and realized returns. According to Han et al. (2023), their cross-sectional MSFE is particularly suited for assessing the predictive accuracy in the cross-section as this task typically revolves around making sure that the predictions are accurate relative to one another. Thus, if the predictions for each asset i is off by a certain amount x (i.e. $\hat{r}_{i,t|t-1} = r_{i,t} + x$ for $i = 1, ..., n_t$ then simply value weighting and averaging the MSFE of each stock results in a MSFE of x^2 whereas the cross-sectional MSFE of Han et al. (2023) is 0.

While the value-weighted cross-sectional MSFE is suitable if sorted long-short portfolios are formed in the spirit of the classical asset pricing literature, we argue that this is not the appropriate measure in the case of portfolio optimization. Although the ranking of predicted returns is also important in this context, the actual values do matter much more as they are directly responsible for deriving portfolio weights. Consequently, we suggest to calculate the MSFE for optimized portfolios by explicitly considering the ex-ante optimal weights as in (10):

$$MSFE_t^{opt} = \frac{1}{n_t} \sum_{i=1}^{n_t} |w_{opt,i,t}| (r_{i,t} - \hat{r}_{i,t|t-1})^2 \quad \text{for} \quad t = 1, ..., T, \tag{10}$$

where $w_{opt,i,t}$ refers to the ex-ante optimal weights for asset i at time t. Note that this measure not only assess the aggregate size of the estimation errors within the portfolio but also judges how well the optimizer is capable of ignoring the noise within the parameter inputs. In other words, the measure only takes into account estimation errors that are also relevant for the portfolio (i.e. all assets that have a non-zero weight) and ignores the error within estimates of asset that do not end up in the portfolio.

3.2.3 Investment Strategies

In total, we consider three types of investment tasks: First of all, we focus on simple timing strategies following Goyal et al. (2021). Thus, the first timing strategy is an unscaled long-short strategy that is either long the equity premium if the prediction is above the prevailing median forecast or short if it is below (11):

$$\omega_{Q50,ls,t+1} = \begin{cases} 1 & \text{if } \hat{r}_{t+1} > Q_{50,t,\hat{r}_t} \\ -1 & \text{if } \hat{r}_{t+1} \leq Q_{50,t,\hat{r}_t} \end{cases} \tag{11}$$

Note that we also investigate the timing strategy that only shorts the market if the prediction is extremely bearish (i.e. below the 25th percentile).

The second timing strategy covers the same signals (i.e. the prevailing median foreacst and the prevailing 25th percentile) but is constrained to be long-only in the case of a negative signal (12):

$$\omega_{Q50,lo,t+1} = \begin{cases} 1 & \text{if } \hat{r}_{t+1} > Q_{50,t,\hat{r}_t} \\ 0 & \text{if } \hat{r}_{t+1} \leq Q_{50,t,\hat{r}_t} \end{cases} \tag{12}$$

The third timing strategy is a scaled long-short strategy where the weight of the investment depends on a z-score which is calculated by subtracting the prevailing median forecast from the forecast for this period and then divided by the prevailing standard deviation (13):

$$\omega_{Z,t+1} = Z_{r_{\hat{t+1}}} = \frac{\hat{r}_{t+1} - Q_{50,t,\hat{r}_t}}{\sigma_{t,\hat{r}_t}} \tag{13}$$

For all timing exercises we consider the performance of two benchmark strategies: The first one is based on the prevailing mean as a timing signal but uses the same strategies as described above to derive weights. The second benchmark is a simple (but hard to beat) buy-and-hold strategy.

The second type of investment task is reminiscent of the standard approach to portfolio formation used in asset pricing and revolves around sorting stocks into sub-portfolios based on the magnitude of certain characteristics and then holding the stocks in the top portfolio while the stocks in the bottom portfolio are sold short. In our setting, we sort the cross-section according to the predicted return for the next period. We then proceed with dividing the entire cross-section into five sub-portfolios based on their rank. Subsequently, we form a long-short portfolio by going long the security with the highest predicted return and short the securities with the lowest predicted returns. The weights within each portfolio either depend on the market-capitalization or on the number of assets. As benchmarks, we use the same sorting strategy based on the prevailing mean as well as an equally-weighted portfolio of all assets in the cross-section.

Finally, the third investment task we consider is mean-variance optimization. In order to do so, we repeatedly draw samples of 4 to 30 assets from the set of 49 industry portfolios. At each time t, the security weights w_{t+1} are determined by maximizing the expected utility as in (14):

$$\max_{\omega_{t+1}} \quad \omega_{t+1}^T \hat{\mu}_{t+1} - \frac{\gamma}{2} \omega_{t+1}^T \hat{\Sigma}_{t+1} \omega_{t+1} \tag{14}$$

The vector of relative portfolio weights invested in the N risky assets at time $t+1$ is then given by (15):

$$\omega_{t+1} = \frac{\hat{\Sigma}_{t+1}^{-1} \hat{\mu}_{t+1}}{\mathbf{1}_N \hat{\Sigma}_{t+1}^{-1} \hat{\mu}_{t+1}} \tag{15}$$

In addition to that, we also consider the case with short-sell constraints to investigate the behaviour of optimized long-only portfolios. For all of the mean-variance portfolios, we use the prevailing mean counterparts as benchmark portfolios. Furthermore, we also use the equally-weighted portfolio as an additional benchmark due to its hard-to-beat nature.

3.2.4 Measuring Economic Value

For the purpose of investigating how predictability translates to performance and economic value we first have to define the most important measures of economic value that are used in research. For the majority of timing exercises, the most relevant way of quantifying the economic value of a predictor involves calculating the Sharpe ratio of the realized returns that were generated by investing according to the timing signal. More formally, the Sharpe ratio is defined as in (16):

$$SR = \frac{\bar{r}}{\sigma_r} \tag{16}$$

where $\bar{r}$ denotes the average realized return of the timing strategy in excess of the risk-free rate and σ_r refers to the realized standard deviation of the strategy.

In addition to the Sharpe ratio, we also consider certainty equivalent returns (CER) as a measure for the economic value of a predictor. An investor‘s certainty equivalent return is given by (17):

$$CER = \bar{r} - \frac{1}{2} \gamma \sigma_r^2 \tag{17}$$

The CER can be interpreted as the risk-free rate that an investor requires to forfeit the return of the risky strategy.

Note, that both, the SR and the CER than can be generated using a certain predictor is always compared to the SR and CER of a benchmark predictor such as the prevailing mean. Thus, the actual measure is not the ratio itself but the gain or loss that occurs through relying on the certain predictor in contrast to using the prevailing mean as a predictor.

4 Results

4.1 Lambda vs. conventional measures

First of all, we need to investigate, whether or not our controlled approach to increasing the predictability of returns actually works. Thus, before we consider measures for economic value, we have a look at how predictability changed due to our approach. Table (1) summarizes our results related to how well our measure tracks predictability relative to conventional measures of predictability. In particular, the mean squared forecast error, the mean absolute forecast error, the R^2 for benchmark predictions of 0 and the prevailing mean, the forecast correlation and the hit ratio are depicted for various predictions based on different estimation windows and different levels of λ. Right away, it is noticeable that regardless of the predictability measure, predictability improves as λ increases meaning that we observe the lowest values for MSFE as well as MAFE and the highest values for R^2, FC and HR when λ is the highest. Furthermore, it is worth mentioning that across different lengths of rolling window estimates (i.e. across panels) the levels of predictability do not differ too much although predictability improves according to the majority of measures as the estimation window gets longer.

[PLACE Table 1 HERE.]

Notable exemptions from this observation is the forecast correlation when no artificial predictability is added. For this measure, the highest predictability can be observed in panel A. In addition to that, adding predictability via λ has the strongest impact on the forecast correlation as can be seen across all panels. In the case of $\lambda = 0$, all values are fairly close to 0. Increasing λ to 0.1 however, yields a forecast correlation of 0.3588 in panel A and a forecast correlation of 0.8664 in panel E. Finally, it is also worth mentioning that with the exemption of the hit ratio, the results for predictions based on estimation windows of more than 60 months are always better compared to predictions based on an estimation window of less than 60 months. Interestingly, the hit ratio for the estimation window of 240 months is more in line with that of 12 and 24 months. Also note that regardless of λ and estimation window, all predictions have a hit ratio well above 50% which indicates that even simple rolling averages tend to correctly predict the sign of the next period's return in more than half of all cases. Most likely, this is due to the fact that the equity premium is upward sloping most of the time.

[PLACE FIGURE 1 HERE.]

As R^2 is most commonly used to assess the predictive accuracy of a variable, figure (1) illustrates how the out-of-sample R^2 of predictions based on rolling means increases as we increase λ in our approach. As mentioned, predictions based on a longer rolling window are more accurate than the predictions using a shorter rolling window. Although there is little difference in predictive accuracy once 60 or more months are considered. As all lines in figure (1) are upward sloping, it is evident that the out-of-sample R^2 increases alongside λ. Note that the upper panel compares the MSFE of our predictions to the MSFE of naive predictions of 0 while the lower panel uses the prevailing mean as a benchmark prediction. What is interesting here is the fact that for all levels of λ, the R^2 based on the naive benchmark predictions of 0 are higher compared to their prevailing mean based counterparts. This indicates, that for tasks related to predicting the equity premium, 0 is the the easier benchmark to beat. To further illustrate how much easier a benchmark of 0 is, we investigate at which level of λ each estimation window starts to generate positive out-of-sample R^2 for each benchmark prediction. In the upper panel (i.e. against a benchmark of 0), all predictions achieve an R^2 of 0 or higher at $\lambda = 0.03$. In the lower panel (i.e. against the prevailing mean prediction) all out-of-sample R^2 are above or at 0 at $\lambda = 0.04$. This implies that the threshold for beating the forecasts of the prevailing mean is 1%-point higher than the threshold for a prediction of 0.

[PLACE FIGURE 2 HERE.]

Finally, (2) illustrates how predictability behaves over time. More precisely, the figure depicts cumulative squared forecast error differences of predictions based on a rolling window of 12 (upper panel) and 24 months(lower panel) vs predictions made by the prevailing mean. For illustrative purposes, we only include λ-values of $0, 0.01, 0.05, 0.08,$ and 0.1. Grey bars indicate NBER recession periods. Typically, periods in which lines are upward sloping indicate periods in which the the predictor performs better compared to the prevailing mean while downward sloping lines indicate the opposite. Ideally, a good predictor generates predictability smoothly across time, avoiding periods in which the line is downward sloping or spiking.

As depicted in figure (2), none of the predictors are without periods in which predictability shortly peaks or increases. The predictions based on λ-values of 0 and 0.01 are however continuously downward sloping, confirming the previously documented poor performance of these predictors. In contrast to that, the predictions with a considerable amount of look-ahead bias (i.e. $\lambda \in \{0.05, 0.08, 0.1\}$ are all (slightly) upward sloping. None of the lines depicting cumulative squared forecast error differences are without spikes. Given however, that these spikes do not have a lasting impact on the predictive accuracy as they quickly fade away, none of the results seem to be influenced by “lucky hits“.

4.2 Predictability and economic value

We begin with investigating the the link between predictability and economic value for our timing exercises. Table 2) below depicts monthly mean return (μ), the standard deviation (σ) and the Sharpe ratio ($\frac{\mu}{\sigma}$) of different (scaled), (un-)constrained timing strategies for the equity premium, based on various rolling estimation windows and λ-values. Details on how we incorporated the timing signals in our strategy can be taken from subsection 3.2.3.

While almost all strategies based on unaltered return predictions start out with a negative mean return and thus also a negative Sharpe ratio, as λ increases, the performance of all strategies turns positive. In fact, all three strategies show improved performance metrics with increasing λ-values. The Sharpe ratio consistently improves with higher λ-values, indicating better risk-adjusted performance across different rolling windows. Longer rolling windows generally result in higher mean returns and Sharpe ratios, suggesting that extended historical data provides more robust timing signals. Among the three strategies, the z-score timing strategy outperforms the long-only and long-short strategies in terms of both mean returns and Sharpe ratios, particularly at higher λ-values and longer rolling windows, although there are a few exemptions such as the Sharpe ratios of the rolling estiamtion window of 240 months at the λ-values of $0.08, 0.09$, and 0.1.

[PLACE Table 2 HERE.]

In general terms, these results suggest that higher levels of predictability in timing strategies can significantly enhance performance, especially when combined with longer rolling windows, providing more reliable and profitable investment decisions.

Moving on to analysing the performance of optimized portfolios, table (3) presents the monthly realized mean return (μ), monthly realized standard deviation (σ), and monthly Sharpe ratio ($\frac{\mu}{\sigma}$) of mean-variance optimized (MVO) portfolios. These portfolios are based on repeated random draws of $10, 20$, and 30 assets from a set of 49 industry portfolios with varying λ-values.

Panel A depicts these statistics for portfolios consisting of 10 assets, the monthly mean return increases consistently as λ increases. Starting from 0.0139 at $\lambda = 0$, the mean return grows to 0.1231 at $\lambda = 0.1$. The standard deviation initially decreases from 0.5856 at $\lambda = 0$ to 0.5123 at $\lambda = 0.02$, but then shows significant fluctuations, reaching a peak of 2.9567 at $\lambda = 0.1$. The Sharpe ratio improves initially, peaking at 0.1383 for $\lambda = 0.05$, but then decreases, suggesting a decline in risk-adjusted returns at higher λ-values.

For the portfolios based on 20 assets, the monthly mean return increases markedly with higher λ values, from 0.0148 at $\lambda = 0$ to 0.2014 at $\lambda = 0.1$. The standard deviation decreases initially from 1.2499 at $\lambda = 0$ to 1.0856 at $\lambda = 0.02$, but then shows significant variation, peaking at 5.1828 at $\lambda = 0.07$ and stabilizing around 4.3688 at $\lambda = 0.1$. The Sharpe ratio shows mixed trends, initially increasing to 0.0938 at $\lambda = 0.02$, but later fluctuating and ending at 0.0636 at$\lambda = 0.1$.

For portfolios consisting of 30 assets, the monthly mean return shows a steady increase with rising λ-values, starting at 0.0145 at $\lambda = 0$ and reaching 0.2521 at $\lambda = 0.1$. The standard deviation varies significantly, starting at 1.4188 at $\lambda = 0$, decreasing to 1.1874 at $\lambda = 0.01$, and then fluctuating with a peak of 8.0547 at $\lambda = 0.1$. The Sharpe ratio initially increases to 0.0865 at $\lambda = 0.02$, but then exhibits variability, with a decrease to 0.0456 at $\lambda = 0.1$.

[PLACE Table 3 HERE.]

Figure (**??**) visualizes the results of table (3) in more detail, highlighting the volatile nature of the Sharpe ratios alongside increasing levels of predictability.

[PLACE Figure 3 HERE.]

Across all panels, the monthly mean return increases consistently with higher λ-values, indicating enhanced portfolio returns as λ rises. However, the standard deviation shows significant fluctuations, particularly at higher λ-values, suggesting increased portfolio risk and volatility. This variability is more pronounced in portfolios with a larger number of assets. The Sharpe ratio, generally improves initially with increasing λ-values, but then declines or fluctuates, indicating that while returns increase, the associated risk also rises, which diminishes the risk-adjusted performance at higher λ-values. These findings suggest that while higher λ-values (or more predictability) can lead to higher returns, they also introduce greater volatility, which impacts the risk-adjusted performance. Counter-intuitively, portfolios with a larger number of assets exhibit more pronounced fluctuations in both standard deviation and Sharpe ratios, indicating that despite increased levels of predictability, these benefits may be offset by an increase in the realized volatility of the portfolio returns.

Given the counter-intuitive nature of the results depicted in table (3) we have a closer look at the most important input variable that is used to come up with optimal portfolio weights: the return predictions. Thus, table (4) presents the mean predicted value (Mean Pred), the mean dispersion of the predicted values (SD Pred), and the mean absolute weights (Mean Weight) from standard mean-variance optimization (MVO) for portfolios based on $10, 20$, and 30 assets across different levels of λ.

[PLACE Table 4 HERE.]

Across all panels, the mean predicted value remains relatively stable, indicating that the central tendency of the predicted values does neither significantly change with varying levels of λ nor across the different numbers of assets within each of the cross-sections. In contrast to that, the standard deviation of the predicted values increases gradually with higher λ-values, suggesting increased dispersion in the predictions and thus, potentially more extreme observations. Similarly, the mean absolute weights also show a trend of increasing with higher λ-vales across all portfolio sizes, indicating greater concentration in asset allocations and extreme weights. This trend is more pronounced in portfolios with a larger number of assets, where the mean absolute weights exhibit significant fluctuations, reflecting a higher degree of concentration in specific assets as λ increases.

Figure (4) once visualizes the results of table (4) in more detail, highlighting the increased dispersion in forecasts alongside an increase in absolute portfolio weights and increasing levels of predictability.

[PLACE Figure 4 HERE.]

Taken together with the results presented in table (3), these findings suggest that while the central tendency of predicted values remains stable as λ increases, the dispersion of these predictions also increases leading to more extreme portfolio weights. These extreme weights further increase the standard deviation of the realized portfolio returns. In fact, the increase in volatility even outweighs the higher returns which ultimately results in lower Sharpe ratios. This is particularly counter-intuitive as mean-variance optimization was particularly designed to optimize the Sharpe ratio and treats parameter inputs with certainty. Thus, we would expect that increasing the level of predictability automatically increases the Sharpe ratio.

5 Conclusion

In this study, we aim at investigating how predictability in asset returns translates into portfolio performance using a controlled approach to predictability and various investment strategies. While our approach increases predictability across all considered and established measures of predictability, our results are more nuanced when it comes to the economic value of predictions. For timing exercises related to the equity premium, predictability and profitability as measured by the Sharpe ratio increase alongside each other. In a more sophisticated portfolio optimization setting, this relationship does not hold.

Our findings indicate that while increasing the level of predictability consistently enhances the mean returns of portfolios, it also introduces significant volatility, leading to mixed outcomes in terms of risk-adjusted performance. The Sharpe ratios initially improve with increasing λ-values but eventually decline or fluctuate. We observe this trend across all portfolio sizes.

Additionally, our analysis of the mean predicted values, the standard deviation of predicted values, and the mean absolute weights in MVO portfolios reveals that while the central tendency of predicted values remains stable, the dispersion of these predictions increases with higher levels of predictability. This increased dispersion leads to more extreme portfolio weights and higher volatility in realized portfolio returns, ultimately resulting in lower Sharpe ratios despite higher returns.

These findings underscore the complexity of leveraging predictability to enhance portfolio performance. While higher levels of predictability can lead to higher returns, they also introduce greater volatility, which can offset the benefits in terms of risk-adjusted performance. This counter-intuitive result suggests that mean-variance optimization, despite its theoretical appeal, may not always yield superior risk-adjusted returns when predictability is increased. In conclusion, our study highlights the nuanced relationship between predictability and performance in portfolio optimization.

As a consequence, we are currently working gaining a deeper understanding of the dynamics at work by extending our approach to robust optimization techniques that were specifically designed to deal with estimation errors and parameter uncertainty. Furthermore, we are interested in the the implications of our findings for long-term portfolio allocation (i.e. inter-temporal or dynamic asset allocation). Finally, we explore how the ability to predict returns changes or alters the way investors have to think about, measure and incorporate risk.

References

Allen, D., Lizieri, C., and Satchell, S. (2019). In Defense of Portfolio Optimization: What If We Can Forecast? *Financial Analysts Journal*, pages 1–19.

Campbell, J. Y. and Thompson, S. B. (2008). Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average? *Review of Financial Studies*, 21(4):1509–1531.

Cederburg, S., Johnson, T. L., and O'Doherty, M. S. (2023). On the Economic Significance of Stock Return Predictability. *Review of Finance*, 27(2):619–657.

Cenesizoglu, T. and Timmermann, A. (2012). Do return prediction models add economic value? *Journal of Banking & Finance*, 36(11):2974–2987.

Dangl, T. and Halling, M. (2012). Predictive regressions with time-varying coefficients. *Journal of Financial Economics*, 106(1):157–181.

Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance*, 25(2):383–417. 13410.

Goyal, A., Welch, I., and Zafirov, A. (2021). A Comprehensive Look at the Empirical Performance of Equity Premium Prediction II.

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

Han, Y., He, A., Rapach, D., and Zhou, G. (2023). Cross-Sectional Expected Returns: New Fama-MacBeth Regressions in the Era of Machine Learning.

Kane, A., Kim, T.-H., and White, H. (2010). Forecast Precision and Portfolio Performance. *Journal of Financial Econometrics*, 8(3):265–304.

Kelly, B., Malamud, S., and Zhou, K. (2024). The Virtue of Complexity in Return Prediction. *The Journal of Finance*, 79(1):459–503. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/jofi.13298.

Kelly, B. T., Malamud, S., and Zhou, K. (2021). The Virtue of Complexity in Return Prediction.

Leitch, G. and Tanner, J. E. (1991). Economic Forecast Evaluation: Profits Versus the Conventional Error Measures. *The American Economic Review*, 81(3):580–590. Publisher: American Economic Association.

Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance*, 7(1):77–91.

Neely, C. J., Rapach, D. E., Tu, J., and Zhou, G. (2014). Forecasting the Equity Risk Premium: The Role of Technical Indicators. *Management Science*, 60(7):1772–1791. Publisher: INFORMS.

Rapach, D. and Zhou, G. (2022). Asset Pricing: Time-Series Predictability.

Rapach, D. E., Ringgenberg, M. C., and Zhou, G. (2016). Short Interest and Aggregate Stock Returns. *Journal of Financial Economics*, 121(1):46–65.

Rapach, D. E., Strauss, J. K., Tu, J., and Zhou, G. (2011). Out-of-Sample Industry Return Predictability: Evidence from a Large Number of Predictors. Technical report.

Rapach, D. E., Strauss, J. K., Tu, J., and Zhou, G. (2019). Industry Return Predictability: A Machine Learning Approach. *The Journal of Financial Data Science*, 1(3):9–28. Publisher: Institutional Investor Journals Umbrella.

Rapach, D. E., Strauss, J. K., and Zhou, G. (2010). Out-of-Sample Equity Premium Prediction: Combination Forecasts and Links to the Real Economy. *The Review of Financial Studies*, 23(2):821–862.

Reschenhofer, E., Mangat, M. K., Zwatz, C., and Guzmics, S. (2020). Evaluation of current research on stock return predictability. *Journal of Forecasting*, 39(2):334–351. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/for.2629.

Welch, I. and Goyal, A. (2008). A Comprehensive Look at The Empirical Performance of Equity Premium Prediction. *The Review of Financial Studies*, 21(4):1455–1508. Publisher: Oxford Academic.

A Appendix

Figures

Figure 1: Depicts the out-of-sample R^2 for predictions based on rolling windows of $12, 14, 60, 120$ and 240 months vs our own measure of predictability called λ. For the purpose of this figure, λ has been increased from 0 to 0.1 in steps of 0.002. In the upper panel, the MSFE of predictions based on rolling windows is compared to the MSFE of a naive prediction of 0 while in the lower panel, the benchmark MSFE is that of predictions based on the prevailing mean.

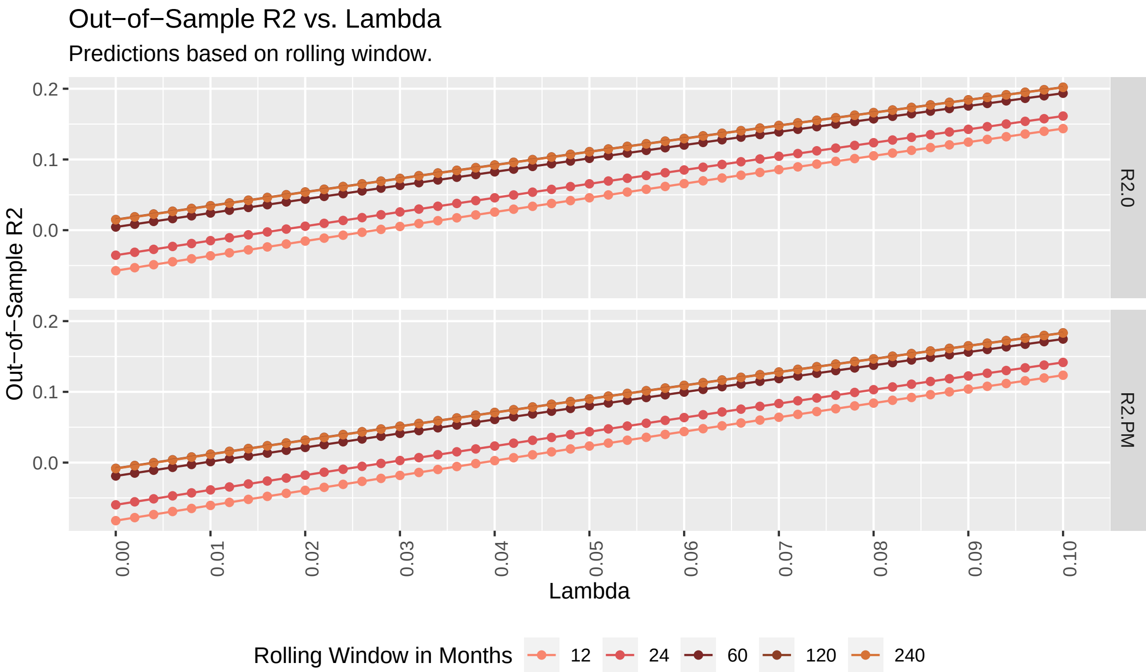


Figure 2: Depicts the cumulative squared forecast error differences of predictions based on a rolling window estimates of 12 (upper panel) and 24 months (lower panel) vs predictions made by the prevailing mean. For illustrative purposes, we only include λ-values of $0, 0.01, 0.05, 0.08$, and 0.1.Grey bars indicate NBER recession periods.

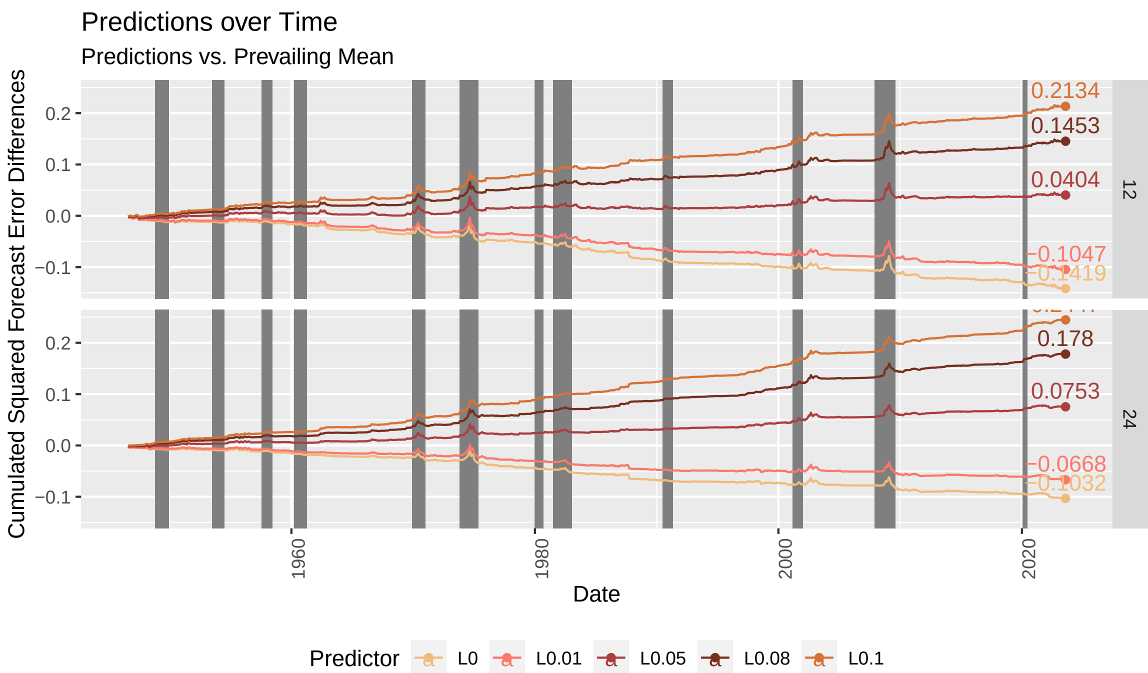


Figure 3: Depicts the monthly realized mean return, the monthly realized standard deviation and the monthly Sharpe ratio of mean-variance optimized portfolios based on repeated random draws of $10, 20$ and 30 assets from the set of 49 industry portfolios alongside increasing values of λ.

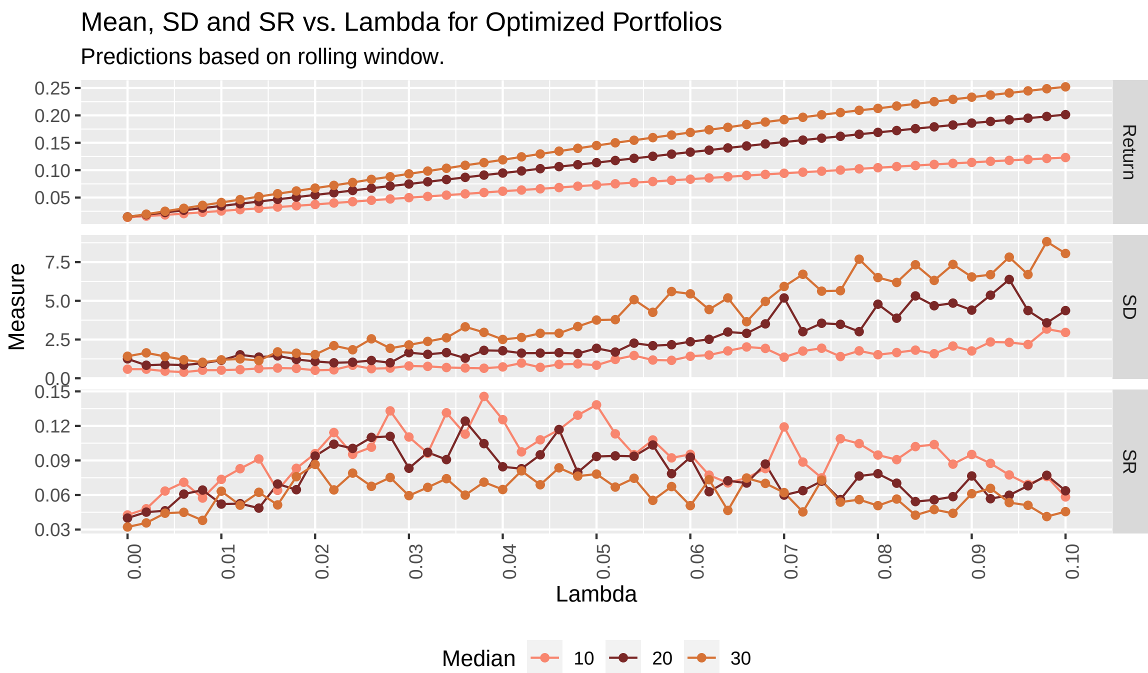
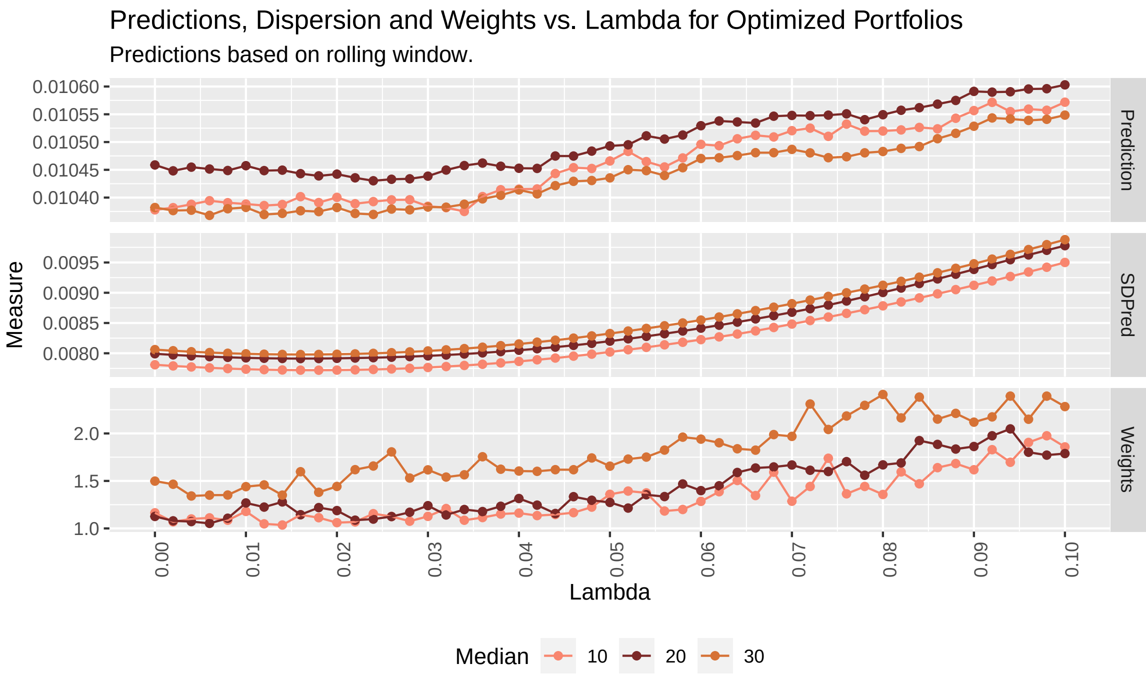


Figure 4: Depicts the mean predicted value, the mean dispersion of the predicted values as well as the mean absolute weights retrieved from standard mean-variance optimization for the portfolios based on $10, 20$, and 30 assets for different levels of λ.



Tables

Table 1: Presents the mean squared forecast error (MSFE), the mean absolute forecast error (MAFE), the out-of-sample R^2 against a prediction of 0 and against the prevailing mean (R2.PM and R2.0), the forecast correlation (FC) as well as the hit ratio (HR) for λ-values between 0 and 0.1. Each panel uses a different rolling windows between 12 and 240 months to create base-forecasts.

Panel A: Rollwing Window = 12 Months											
	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
MSFE	0.0020	0.0020	0.0019	0.0019	0.0019	0.0018	0.0018	0.0017	0.0017	0.0017	**0.0016**
MAFE	0.0341	0.0337	0.0334	0.0330	0.0327	0.0324	0.0320	0.0317	0.0313	0.0310	**0.0307**
R2.PM	-0.0821	-0.0606	-0.0393	-0.0182	0.0027	0.0234	0.0438	0.0641	0.0841	0.1039	**0.1235**
R2.0	-0.0573	-0.0362	-0.0154	0.0052	0.0256	0.0458	0.0658	0.0856	0.1051	0.1245	**0.1436**
FC	0.0209	0.0539	0.0874	0.1212	0.1553	0.1896	0.2239	0.2581	0.2921	0.3257	**0.3588**
HR	0.5892	0.6000	0.6065	0.6108	0.6173	0.6238	0.6346	0.6389	0.6432	0.6508	**0.6595**
Panel B: Rollwing Window = 24 Months											
	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
MSFE	0.0020	0.0019	0.0019	0.0019	0.0018	0.0018	0.0017	0.0017	0.0017	**0.0016**	**0.0016**
MAFE	0.0338	0.0335	0.0332	0.0328	0.0325	0.0322	0.0318	0.0315	0.0311	0.0308	**0.0305**
R2.PM	-0.0597	-0.0386	-0.0178	0.0029	0.0233	0.0436	0.0636	0.0834	0.1030	0.1224	**0.1416**
R2.0	-0.0354	-0.0148	0.0056	0.0258	0.0458	0.0656	0.0851	0.1045	0.1236	0.1426	**0.1613**
FC	-0.0387	0.0104	0.0604	0.1109	0.1617	0.2122	0.2621	0.3109	0.3585	0.4043	**0.4482**
HR	0.5795	0.5881	0.5968	0.6119	0.6184	0.6368	0.6454	0.6519	0.6659	0.6768	**0.6843**
Panel B: Rollwing Window = 60 Months											
	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
MSFE	0.0019	0.0019	0.0018	0.0018	0.0018	0.0017	0.0017	0.0016	0.0016	0.0016	**0.0015**
MAFE	0.0333	0.0329	0.0326	0.0323	0.0319	0.0316	0.0313	0.0309	0.0306	0.0303	**0.0299**
R2.PM	-0.0189	0.0013	0.0214	0.0413	0.0609	0.0804	0.0997	0.1187	0.1376	0.1562	**0.1747**
R2.0	0.0045	0.0243	0.0439	0.0633	0.0825	0.1015	0.1203	0.1390	0.1574	0.1756	**0.1936**
FC	-0.0093	0.0724	0.1542	0.2346	0.3120	0.3851	0.4529	0.5149	0.5709	0.6210	**0.6653**
HR	0.5827	0.6000	0.6097	0.6173	0.6346	0.6541	0.6703	0.6822	0.6908	0.6995	**0.7135**
Panel D: Rollwing Window = 120 Months											
	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
MSFE	0.0019	0.0018	0.0018	0.0018	0.0017	0.0017	0.0017	0.0016	0.0016	0.0016	**0.0015**
MAFE	0.0331	0.0327	0.0324	0.0321	0.0317	0.0314	0.0311	0.0308	0.0304	0.0301	**0.0298**
R2.PM	-0.0082	0.0119	0.0318	0.0514	0.0709	0.0901	0.1092	0.1280	0.1467	0.1651	**0.1834**
R2.0	0.0150	0.0346	0.0540	0.0732	0.0922	0.1110	0.1296	0.1481	0.1663	0.1843	**0.2021**
FC	0.0092	0.1179	0.2246	0.3257	0.4186	0.5016	0.5741	0.6365	0.6895	0.7342	**0.7718**
HR	0.5914	0.6043	0.6195	0.6368	0.6530	0.6638	0.6703	0.6811	0.6984	0.7103	**0.7178**
Panel E: Rollwing Window = 240 Months											
	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
MSFE	0.0019	0.0018	0.0018	0.0018	0.0017	0.0017	0.0017	0.0016	0.0016	0.0016	**0.0015**
MAFE	0.0332	0.0329	0.0325	0.0322	0.0319	0.0315	0.0312	0.0309	0.0305	0.0302	**0.0299**
R2.PM	-0.0085	0.0116	0.0314	0.0511	0.0706	0.0898	0.1089	0.1277	0.1464	0.1649	**0.1831**
R2.0	0.0146	0.0343	0.0537	0.0729	0.0919	0.1107	0.1293	0.1478	0.1660	0.1840	**0.2019**
FC	-0.0252	0.1336	0.2856	0.4214	0.5360	0.6287	0.7021	0.7594	0.8040	0.8389	**0.8664**
HR	0.6054	0.6054	0.6065	0.6108	0.6195	0.6303	0.6411	0.6584	0.6746	0.6800	**0.7005**

Table 2: Presents the monthly mean return, the standard deviation and the Sharpe ratio of different (scaled), (un-)constrained timing strategies for the equity premium based on various rolling estimation windows and λ-values. All investment decision use the predictive signal relative to the prevailing median value to determine the position.

Panel A: Long-Only Timing Strategy												
		0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
μ	12	4.00E-04	8.00E-04	0.0017	0.0020	0.0023	0.0027	0.0030	0.0037	0.0043	0.0049	0.0055
	24	-0.0009	-0.0001	4.00E-04	0.0011	0.0021	0.0030	0.0041	0.0047	0.0056	0.0063	0.0065
	60	-0.0011	-0.0002	0.0012	0.0026	0.0037	0.0050	0.0061	0.0070	0.0074	0.0079	0.0085
	120	-0.0010	0.0013	0.0036	0.0052	0.0064	0.0078	0.0084	0.0090	0.0096	0.0109	0.0111
	240	-0.0016	0.0015	0.0044	0.0063	0.0081	0.0093	0.0106	0.0112	0.0120	0.0123	0.0125
σ	12	0.0274	0.0277	0.0270	0.0270	0.0268	0.0265	0.0266	0.0254	0.0252	0.0251	0.0253
	24	0.0259	0.0264	0.0265	0.0261	0.0254	0.0253	0.0243	0.0244	0.0238	0.0240	0.0241
	60	0.0250	0.0253	0.0257	0.0238	0.0228	0.0225	0.0227	0.0228	0.0228	0.0234	0.0233
	120	0.0222	0.0216	0.0216	0.0213	0.0214	0.0215	0.0213	0.0216	0.0224	0.0236	0.0238
	240	0.0164	0.0141	0.0166	0.0183	0.0203	0.0217	0.0231	0.0234	0.0240	0.0241	0.0242
$\frac{\mu}{\sigma}$	12	0.0140	0.0287	0.0632	0.0738	0.0865	0.1016	0.1144	0.1465	0.1694	0.1969	0.2157
	24	-0.0342	-0.0032	0.0154	0.0441	0.0839	0.1174	0.1680	0.1927	0.2360	0.2609	0.2689
	60	-0.0444	-0.0070	0.0482	0.1108	0.1640	0.2237	0.2698	0.3056	0.3258	0.3373	0.3649
	120	-0.0463	0.0601	0.1667	0.2421	0.2966	0.3616	0.3920	0.4143	0.4294	0.4601	0.4680
	240	-0.0949	0.1090	0.2653	0.3455	0.3978	0.4265	0.4617	0.4782	0.4998	0.5119	0.5180
Panel B: Long-Short Timing Strategy												
		0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
μ	12	8.00E-04	0.0016	0.0034	0.0040	0.0046	0.0054	0.0061	0.0074	0.0086	0.0099	0.0109
	24	-0.0018	-0.0002	8.00E-04	0.0023	0.0043	0.0059	0.0082	0.0094	0.0112	0.0125	0.0130
	60	-0.0022	-0.0004	0.0025	0.0053	0.0075	0.0101	0.0123	0.0139	0.0149	0.0158	0.0170
	120	-0.0021	0.0026	0.0072	0.0103	0.0127	0.0155	0.0167	0.0179	0.0192	0.0217	0.0223
	240	-0.0031	0.0031	0.0088	0.0126	0.0162	0.0185	0.0213	0.0224	0.0239	0.0247	0.0251
σ	12	0.0547	0.0553	0.0540	0.0540	0.0537	0.0529	0.0531	0.0507	0.0505	0.0502	0.0506
	24	0.0518	0.0527	0.0530	0.0522	0.0508	0.0506	0.0485	0.0489	0.0477	0.0480	0.0482
	60	0.0500	0.0506	0.0514	0.0476	0.0455	0.0451	0.0455	0.0456	0.0457	0.0468	0.0466
	120	0.0443	0.0431	0.0432	0.0426	0.0429	0.0429	0.0427	0.0433	0.0448	0.0472	0.0476
	240	0.0328	0.0282	0.0331	0.0365	0.0406	0.0435	0.0461	0.0469	0.0479	0.0482	0.0484
$\frac{\mu}{\sigma}$	12	0.0140	0.0287	0.0632	0.0738	0.0865	0.1016	0.1144	0.1465	0.1694	0.1969	0.2157
	24	-0.0342	-0.0032	0.0154	0.0441	0.0839	0.1174	0.1680	0.1927	0.2360	0.2609	0.2689
	60	-0.0444	-0.0070	0.0482	0.1108	0.1640	0.2237	0.2698	0.3056	0.3258	0.3373	0.3649
	120	-0.0463	0.0601	0.1667	0.2421	0.2966	0.3616	0.3920	0.4143	0.4294	0.4601	0.4680
	240	-0.0949	0.1090	0.2653	0.3455	0.3978	0.4265	0.4617	0.4782	0.4998	0.5119	0.5180
Panel C: z-Score Timing Strategy												
		0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
μ	12	-0.0016	-0.0002	0.0013	0.0028	0.0043	0.0059	0.0074	0.0089	0.0105	0.0120	0.0135
	24	-0.0042	-0.0020	3.00E-04	0.0026	0.0049	0.0072	0.0096	0.0118	0.0139	0.0160	0.0179
	60	-0.0051	-0.0010	0.0033	0.0073	0.0112	0.0146	0.0178	0.0205	0.0229	0.0251	0.0270
	120	-0.0037	0.0018	0.0071	0.0118	0.0158	0.0194	0.0225	0.0251	0.0273	0.0293	0.0309
	240	-0.0052	0.0027	0.0097	0.0155	0.0202	0.0242	0.0274	0.0299	0.0319	0.0334	0.0347
σ	12	0.0583	0.0578	0.0575	0.0572	0.0569	0.0567	0.0566	0.0567	0.0569	0.0571	0.0574
	24	0.0569	0.0559	0.0549	0.0540	0.0535	0.0532	0.0530	0.0531	0.0534	0.0540	0.0548
	60	0.0509	0.0490	0.0476	0.0474	0.0480	0.0493	0.0512	0.0532	0.0555	0.0579	0.0602
	120	0.0423	0.0413	0.0412	0.0422	0.0445	0.0477	0.0510	0.0544	0.0577	0.0607	0.0635
	240	0.0220	0.0205	0.0262	0.0343	0.0424	0.0496	0.0558	0.0608	0.0650	0.0687	0.0717
$\frac{\mu}{\sigma}$	12	-0.0281	-0.0031	0.0231	0.0489	0.0758	0.1034	0.1304	0.1569	0.1838	0.2101	0.2349
	24	-0.0744	-0.0351	0.0057	0.0486	0.0925	0.1363	0.1803	0.2219	0.2604	0.2953	0.3267
	60	-0.0996	-0.0195	0.0687	0.1541	0.2325	0.2972	0.3474	0.3849	0.4132	0.4335	0.4485
	120	-0.0871	0.0434	0.1721	0.2796	0.3556	0.4074	0.4405	0.4615	0.4738	0.4817	0.4856
	240	-0.2355	0.1339	0.3700	0.4517	0.4769	0.4872	0.4912	0.4914	0.4906	0.4871	0.4836

Table 3: Depicts the monthly realized mean return, the monthly realized standard deviation and the monthly Sharpe ratio of mean-variance optimized portfolios based on repeated random draws of $10, 20$ and 30 assets from the set of 49 industry portfolios alongside increasing values of λ.

Panel A: MVO with 10 Assets											
	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
μ	0.0139	0.0255	0.0374	0.0498	0.0616	0.0730	0.0835	0.0943	0.1045	0.1141	0.1231
σ	0.5856	0.5230	0.5123	0.7936	0.7304	0.8406	1.4184	1.3587	1.5188	1.7614	2.9567
$\frac{\mu}{\sigma}$	0.0426	0.0736	0.0960	0.1104	0.1254	0.1383	0.0953	0.1191	0.0947	0.0952	0.0583
Panel B: MVO with 20 Assets											
	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
μ	0.0148	0.0347	0.0548	0.0748	0.0948	0.1138	0.1330	0.1513	0.1690	0.1859	0.2014
σ	1.2499	1.1602	1.0856	1.6529	1.7751	1.9320	2.3577	5.1828	4.7794	4.3982	4.3688
$\frac{\mu}{\sigma}$	0.0400	0.0521	0.0938	0.0832	0.0845	0.0934	0.0928	0.0599	0.0785	0.0765	0.0636
Panel C: MVO with 30 Assets											
	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
μ	0.0145	0.0410	0.0672	0.0933	0.1189	0.1450	0.1689	0.1922	0.2127	0.2332	0.2521
σ	1.4188	1.1874	1.5283	2.1501	2.5033	3.7571	5.4487	5.9230	6.5007	6.5404	8.0547
$\frac{\mu}{\sigma}$	0.0322	0.0633	0.0865	0.0594	0.0648	0.0782	0.0508	0.0621	0.0507	0.0611	0.0456

Table 4: Presents the mean predicted value, the mean dispersion of the predicted values as well as the mean absolute weights retrieved from standard mean-variance optimization for the portfolios based on $10, 20$, and 30 assets for different levels of λ.

Panel A: MVO with 10 Assets											
	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
Mean Pred	0.0104	0.0104	0.0104	0.0104	0.0104	0.0105	0.0105	0.0105	0.0105	0.0106	0.0106
SD Pred	0.0078	0.0077	0.0077	0.0078	0.0079	0.0080	0.0082	0.0085	0.0088	0.0091	0.0095
Mean Weight	1.1646	1.1792	1.0611	1.1270	1.1615	1.3590	1.2846	1.2863	1.3573	1.6177	1.8584
Panel B: MVO with 20 Assets											
	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
Mean Pred	0.0105	0.0105	0.0104	0.0104	0.0105	0.0105	0.0105	0.0105	0.0105	0.0106	0.0106
SD Pred	0.0080	0.0079	0.0079	0.0080	0.0080	0.0082	0.0084	0.0087	0.0090	0.0094	0.0098
Mean Weight	1.1263	1.2683	1.1877	1.2396	1.3149	1.2736	1.3980	1.6680	1.6693	1.8626	1.7882
Panel C: MVO with 30 Assets											
	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.1
Mean Pred	0.0104	0.0104	0.0104	0.0104	0.0104	0.0104	0.0105	0.0105	0.0105	0.0105	0.0105
SD Pred	0.0081	0.0080	0.0080	0.0080	0.0082	0.0083	0.0085	0.0088	0.0091	0.0095	0.0099
Mean Weight	1.4983	1.4399	1.4420	1.6167	1.6041	1.6556	1.9402	1.9692	2.4102	2.1189	2.2828